Dipartimento di Matematica e Fisica

Sezione di Matematica

Dottorato di Ricerca in Matematica

XXVII Ciclo

New Methods for Degenerate Stochastic Volatility Models

Advisor:

Prof. Marco Papi

Candidate:

Cristina Donatucci

Coordinator of the Ph.D. Program:

Prof. Luigi Chierchia

ACADEMIC YEAR 2015/2016

Contents

Basic Notations

$a.e.$	almost everywhere (with respect to the Lebesgue measure)
$a.s.$	almost surely (with respect to the Lebesgue measure)
$\mathbb{R}^d$	euclidean d-dimensional space
$x \cdot y$	scalar product $\sum_{i=0}^{n} x_i y_i$ of two vectors $x, y \in \mathbb{R}^d$
$\|x\|$	euclidean norm of $x \in \mathbb{R}^k$
$\|\|\sigma\|\|$	matrix norm of $\sigma \in \mathbb{R}^{k \times d}$
$B(x_0, r)$	open ball of center x_0 and radius $r \in \mathbb{R}^d$, $\{x \in \mathbb{R}^d : \|x - x_0\| < r\}$
$\overline{B}(x_0, r)$	closed ball of center x_0 and radius $r \in \mathbb{R}^d$, $\{x \in \mathbb{R}^d : \|x - x_0\| \leq r\}$
∂E	boundary of the set E
int E	interior of the set E
$\overline{E}$	closure of the set E
$a \vee b$	$\min\{a, b\}$, for$a, b \in \mathbb{R}$
$a \wedge b$	$\max\{a, b\}$, for$a, b \in \mathbb{R}$
$\otimes$	Direct Product
$\mathfrak{I} \otimes \mathfrak{J}$	The smallest sigma-algebra which contains $\{U \times T : U \in \mathfrak{I}, T \in \mathfrak{J}\}$
$d(x, C)$	distance from the point x to the closed set C
$\Delta t, \Delta x$	discretization steps, in compact form$\Delta \equiv (\Delta x, \Delta t)$
i	imaginary unit
$\mathbb{1}_\Omega$	characteristic function of the set Ω
δ_{ij}	Kronecker's symbol
$C^k(\Omega)$	Space of functions $u : \Omega \to \mathbb{R}$with continuous k-th derivative on a domain Ω
$C_0^k(\Omega)$	Space of functions with compact support belonging to $C^k(\Omega)$
$L^p(\Omega)$	Space of functions $u :\ \Omega \to \mathbb{R}$ such that $\int_\Omega u^p < +\infty$, $1 \leq p \leq +\infty$
$L^\infty(\Omega)$	Space of functions $u :\ \Omega \to \mathbb{R}$ such that $\sup \|u^p\| < +\infty$, $1 \leq p \leq +\infty$
$L^p_{loc}(\Omega)$	Space of functions u in $L^p(\Omega)$ for any compact set $K \subset \Omega, 1 \leq p \leq +\infty$
L^1_{loc}	Space of locally integrable functions

$\mathcal{M}^{p,d}$	Space of $\mathbb{R}^d$-valued $\mathcal{F}_t-$ adapted progressive measurable processes $(X(t))_t$ such that $\forall q,\ \int_0^t \mathbb{E}[\lvert X(s)\rvert^p]ds < \infty,\ \ p \geq 1$
$\mathcal{M}^{2,n\times d}$	Space of $\mathbb{R}^{n\times d}$-valued $\mathcal{F}_t-$ adapted progressive measurable process $(X(t))_t$ such that $\int_0^t \mathbb{E}[\lvert\lvert X(s)\rvert\rvert^p ds] < \infty]$
$\lvert\lvert u\rvert\rvert_p$	Norm of u in $L^p(\Omega)$
l^p	Space of p-th power summable vectors or sequences
l^∞	Space of bounded vectors or sequences
V, v_j	Numerical solution, as a vector and as a value at the nodex_j
V^n, v_j^n	Numerical solution, as a vector and as a value at the node(x_j, t^n)
$M_{k,d}(\mathbb{R})$	Matrix of dimension $k\times d$ with real values
$M^\top$	Transpose of a matrix B
$Ker(M)$	Kernel of matrix M: all of those vector v for which $Mv = 0$
$Ker^\perp(M)$	The ortogonal space of Kernel of M
USC	The space of real-valued upper semi-continuous functions
LSC	The space of real-valued lower semi-continuous functions

Introduction

One of the central problem in modern mathematical finance is derivative pricing: that is to define a fair price. A derivative is a financial contract which value depends on an underlying asset which can be an equity stock, an interest rate or any different financial asset. The well known Black-Scholes model, after 43 years of its first publishing [8], represents, with its closed form , an universal accepted framework. In this work, we develop a qualitative and quantitative analysis on stochastic volatility models. These models represents a wide known class of models among financial mathematics for the evaluation of options and complex derivatives, starting from the fundamental paper of S.L. Heston (1993, *The Review of Financial Studies*). Moreover, this thesis proposes an interesting researches on both theoretical studies on the solution of Dirichlet problem associated and numerical studies, for the approximation of solution and the model calibration by real data taken from real market.

In Chapter 1 we revise the original Black-Scholes model, which assumes the existence of a risk-free asset B_t and of an underlying asset S_t, following respectively a deterministic and a geometric Brownian motion:

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t dW_t \\ dB_t = rB_t dt \end{cases} \tag{1}$$

where the deterministic constant coefficients μ, σ and r represent respectively the local mean rate of return of the asset, the volatility of the asset and the short-term rate interest. $(W_t)_t$ is a standard one dimensional Brownian process ([34], [41]).

However, as well known from many works in literature ([53], [38], [22]), the original model is not consistent with market prices. In particular, it is unable to correctly reproduce all the vanilla option prices mainly because contracts with different strikes and maturities exhibit different volatilities. In fact, given all the model parameters and the observed price of an European type option it is possible to invert the Black-Scholes formula in order to find the so-called implied-volatility. Thus, the implied volatility is the value to be used such that the Black-Scholes price of a plain vanilla option is equal to the actual price quoted on the market. As Rebonato wittily said ([63]): "Implied volatility is the wrong number to put in the wrong formula to obtain the right price".

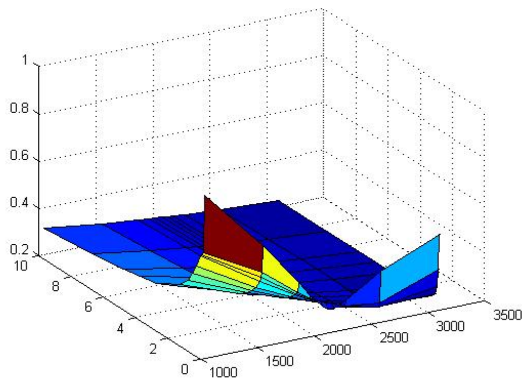


Figure 1: Volatility from Market Data (European call options quoted on the SX5E Index at 1st June 2012).

After an overview of the main mathematical results on the theory of stochastic equations and of financial theory, we have revised results of option pricing. The main strands of research that have been proposed to improve the Black and Schole Model are:

- Local Volatility, [38].

- Stochastic Volatility, [41], [25].
- Levy Processes, [3], [6], [67].

All of these approaches release the Black-Scholes hypothesis of a constant volatility. The class of affine processes, [27], [20], [21], which are a class of time homogeneous Markov processes, includes the above mentioned examples. In an affine model the risk-neutral price of a European-type derivative can be found by solving a system of ordinary differential Riccati equations and then inverting a Fourier transform.

Nevertheless not all the financial models are affine and for the other ones there is both a problem to establish if it is well-posed problem and a problem for the existence of a formula for the evaluation of the solution and of its calibration.

In Chapter 2 we analyze the other models proposed to overcome the Black and Scholes model. So, among the most popular model proposed by literature, in this thesis have been examined the following models:

- Local Volatility Models (LVM). Introduced for the first time in 1994 by Dupire [29] and Derman and Kani [22], these models assume that the diffusion coefficient of the underlying asset is no longer a constant value but instead a deterministic function of time and of the underlying asset itself: $\sigma = \sigma_{LV}(s;t)$

  $$ds_t = rs_t dt + \sigma_{LV}(s_t;t)s_t dW_t \tag{2}$$

  As concerned the Dupire's model, it has been proposed a new implementation, numerically efficient, by which, starting form the option data really quoted on the market, it is obtained an evaluation of all the characteristic parameters of the model. In this way it is obtained a local volatility comparable with the implicit volatility observed. From this approach, it has been highlighted that the local volatility, although it calibrates the no-uniformity of volatility, it presents a not right dynamics of the well known volatilty smile (implicit volatility inferred by market).

- Stochastic Volatility Models (SVM).
  In this class of models the volatility itself is considered to be a stochastic process with its own dynamics. Thus, this is a two-factor model, driven by two correlated Brownian motions, see equation 2.131 below.
- Jump Diffusion Models (JDM) Introduced by Merton these models considers the underlying asset to follow a Levy process with a drift, a diffusion and a jump term;

$$ds_t = rs_t dt + \sigma s_t dW_t + s_t dJ_t \tag{3}$$

Of course, all these three kinds of models have some advantages and disadvantages (which will be analyzed later on). In particular in the last ten years the first two models have been widely studied in academic literature as well as used at the equity trading desks of investment banks. Local Volatility models assume volatility to be a deterministic function of the underlying asset and time, whereas Stochastic Volatility models consider volatility as a random process itself. While the former models are able to be well calibrated to traded vanilla options, the latter can reproduce a more realistic dynamics of implied volatility. The Levy processes can accurately describe heavy-tailed and skewed distributions typical of asset returns and a semi-closed form valuation formulas are available for simple contracts, such as plain vanilla options.

Focus on stochastic volatility models and following the results obtained, among others, also in Pascucci (*Fin. Stoch.*)(2008) and Costantini et P. (*Fin. Stoch.*)(2012) we could know that when the problem at hand does not fit in the class of affine models, the risk-neutral price of a European-type derivative can be computed only by solving a valuation equation. The general form of which (in the time-homogeneous case) is:

$$\begin{cases} \partial_t u(t,x) + Lu(t,x) - c(x)u(t,x) = f(t,x) & (t,x) \in (0,T) \times D \\ u(T,x) = \phi(x) & x \in D \end{cases} \tag{4}$$

where, for every smooth $g : D \subset \mathbb{R}^d \to \mathbb{R}$, the operator L is defined as

$$Lg(x) = \nabla g(x)b(x) + \frac{1}{2}\mathrm{tr}(\nabla^2 g(x)a(x)) + \int_D (g(z) - g(x))m(x,dz), \tag{5}$$

for any $x \in D$. Matrix $a = \sigma\sigma^\top$ corresponds to the diffusion matrix of a stochastic process $(X_t)_t$ with values in the domain $D \subset \mathbb{R}^d$, b represents the drift of the process under a suitable (risk-neutral) probability measure, c is a discount rate function, ϕ is the final payoff function, f is the cost of execution. If presence of jumps, there is also a measure m which summarizes the jump intensity and the probability distribution of $(X_t)_t$.

The Dirichlet problem (4) could have some difficulties, such as:

- The diffusion matrix a is singular on the boundary of the domain D, or is even identically zero in some direction;
- The drift b and the matrix σ are not Lipschitz-continuous up to the boundary of D;
- The coefficients b and a are fast growing near the boundary or at infinity;
- The jump intensity is not bounded;
- The state space D has a boundary, but no boundary conditions are specified.

A result is precisely that under the assumptions below, the pricing problem has a unique viscosity solution. This type of problem, when the operator in the equation is hypoelliptic, is deeply studied in Pascucci (*Fin. Stoch.*)(2008) and is also examined more recently in Costantini et P. ([16]) and we report here the main assumptions for existence and uniqueness of viscosity solutions (which, among the others, allows to deal with singular diffusion matrices). The assumptions needed for such an existence (see chapter 2, 2.2.1

and following) will ensure that any boundary condition will become redundant from an analytical point of view. But, over the hypothesis made on b, σ, f, ϕ, c in the interior of D, the fundamental assumption is the existence of a Lyapunov-type function, that is a no-negative function $V \in C^2(D)$ such that

$$\begin{cases} \int_D V(z)m(x,dz) < +\infty \\ LV(x) \leq C(1+V(x)),\ \forall x \in D, \\ \lim_{x\in D, x\to x_0} V(x) = +\infty\ \forall x_0 \in \partial D, \\ \lim_{x\in D, |x|\to+\infty} V(x) = +\infty. \end{cases} \tag{6}$$

Nevertheless, from a numerical point of view we need a condition even along the boundary which will no force the numerical value to a value which will produce a consistent error. In Ekstrom, E. Tysk (2011) is analyzed for the case of stochastic volatility models the condition enforceable along the boundary and moreover, their assumptions will guarantee the existence of a Lyapunov function. Let X be the stock price process and Y the variance process associated to the underlying X which will determine the payoff ϕ, defined by

$$\begin{cases} dX_t = \sqrt{Y_t} X_t dW_1 \\ X_0 = x_0 \end{cases} \begin{cases} dY_t = \beta(Y_t)dt + \sigma(Y_t)dW_2 \\ Y_0 = y_0 \end{cases}$$

The option price function $u : [0,\infty) \times [0,T] \to \mathbb{R}$ corresponding to a payoff function $\phi : [0,\infty) \to \mathbb{R}$ is given by

$$u(x,y,t) = \mathbb{E}^{x,y,t}[e^{-r(T-t)}g(x(T))], \tag{7}$$

where the indexes indicate that $X(t) = x$ and r is the free-risk rate.

Hypotesis 0.0.1. The drift $\beta \in C([0,\infty))$ is continuously differentiable in x with bounded derivative, and $\beta(0) \geq 0$, $\sigma \in C([0,\infty))$ is such that $\alpha(x) := \frac{1}{2}\sigma^2(x)$ is continuously differentiable with Holder continuous derivative, and $\sigma(x) = 0$ if and only if $x = 0$. The functions β, σ, α_x are all of, at most, linear growth:

$$|\beta(x)| + |\sigma(x)| + |\alpha_x(x)| \leq C(1+x), \tag{8}$$

for all $x \geq 0$. The payoff function $\phi : [0,\infty) \to [0,\infty)$ is continuously differentiable with both ϕ and $\phi^{'}$ bounded.

Theorem 0.0.1. Under Hypothesis 2.3.1 and Assumption 2.3.1, the term structure equation admits a unique solution $u \in C([0,T]\times[0,\infty)) \cap C^1([0,T)\times[0,\infty)) \cap C^{1,2}((0,T)\times(0,\infty))$ which satisfies (2.90) and the terminal condition $u(x,T) = g(x)$, for any $x > 0$, and (2.92).

Then, in Chapter 2 (see Theorem 2.3.3), we state the following result, where some of Tysk's assumptions are relaxed, without loosing those ones needed to have the existence of the Lyapunov function.

Theorem 0.0.2. Let β and $\alpha = \frac{1}{2}\sigma^2$ be locally Lipschitz continuous on $(0,\infty)$. Let $\alpha \in C([0,\infty)) \cap C^1((0,\varepsilon))$, with bounded derivative on $(0,\varepsilon)$, for some $\varepsilon > 0$, $\alpha(y) = 0$ if and only if $y = 0$. Moreover, there exists a positive constant $C > 0$ such that

$$|\beta(y)| + \alpha(y) \leq C(1+y), \tag{9}$$

for all $y > 0$. Assume the following condition:

$$\lim_{y\to 0^+} \left\{\alpha^{'}(y) - \beta(y)\right\} < 0. \tag{10}$$

Then, for every $\phi \in C([0,\infty))$, satisfying $|\phi(S)| \leq C(1+S)$, for all $S \geq 0$, the pricing problem associated to (2.96), with payoff function ϕ, has a unique viscosity solution $u \in C([0,T]\times(0,\infty)^2)$ such that

$$|u(t,S,y)| \leq C\left(1+S\right), \tag{11}$$

for all $S > 0$, $y > 0$, $t \in [0,T]$.

One of the most important stochastic volatility model is the Heston model, introduced in 1993, and nowadays it is probably the most popular stochastic volatility model. Several other models have been derived from Heston model, including also extension with jumps:

$$\begin{cases} dS_t = \mu S_t d_t + \sqrt{V_t} S_t dW_t, \\ dV_t = \kappa(\theta - V_t)d_t + \sigma\sqrt{V_t} dZ_t, \\ dW_t dZ_t = \rho dt. \end{cases} \tag{12}$$

The Heston model is characterized by five constant parameters, namely κ, θ, σ, ρ and the initial value of the variance v_0. We analyze deeply Heston Model: the condition enforceable along the boundary (following the Tysk's result) is verified by the closed-form solution. The reason that makes this model so popular and used is probably the fact that it has a semi-closed form solution for plain vanilla options.

The call t-time price of the European call with strike K and maturity T is the expected discounted value under the risk-neutral measure $\mathbb{Q}$, namely:

$$\begin{aligned} C_t &= e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}[(S_T-K)^+] \\ &= e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}[S_T \mathbb{1}_{S_T>K}] - e^{-r(T-t)}K\mathbb{E}_t^{\mathbb{Q}}[\mathbb{1}_{S_T>K}], \end{aligned} \tag{13}$$

where $\mathbb{E}_t^{\mathbb{Q}}[\cdot]$ denotes the conditional expected value at time $t \in (0,T)$. By analogy with the Black-Scholes formula, the guessed solution of this European option is of the form $C_t = C(t, x_t, V_t)$, where $x_t = \log(S_t)$ and the deterministic function C takes the form (for specified parameters):

$$\begin{cases} C(t,x,v) = e^x P_1(T-t,x,v) - e^{-r\tau} K P_2(T-t,x,v), \\ P_j = \frac{1}{2} + I_j, \qquad j = 1,\, 2, \\ I_j = \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi \log(K)} f_j(\phi;x,v)}{i\phi}\right] \mathrm{d}\phi. \end{cases} \tag{14}$$

This enables a fast valuation of European-style options which becomes of critical importance when calibrating the model to known option prices. The drawback of stochastic volatility models is that the more realistic dynamics comes at the cost of an additional theoretical complexity and a greater difficulty in the numerical solution of the pricing problem and model calibration.

In Chapter 3 we refer closely to the paper presented in Donatucci *et al.* (2015) for a new approach in the calibration of stochastic volatility models, in order to provide an efficient approximation of observed plain vanilla options. The motivation for the introduction of such a viewpoint is mainly based on the fact that the price of a call option obtained in the framework of a stochastic volatility model, depends on the value v_0, the initial volatility, that unfortunately acts like an hidden stochastic variable.

The most simple approach adopted to resolve the estimation of this hidden variable, is considering v_0 as an additional parameter in the calibration procedure.

There exists even a well-known methods of filtering (which is essentially the problem of inferring the current volatility from current and past asset prices) to solve the calibration problem which rely on a linear filter which uses a polynomial state-space formulation of the discrete version of the continuous-time model, see Carravetta F. *at al.* (2000). In contrast, we follow here a different approach inspired on the empirical results proposed by Dragulescu and Yakovenko (2002) in order to calibrate option prices averaging over volatility. This approach permits an extension to a wide class of models where could be one or more hidden values which influences market prices, defining a new notion of arbitrage where the price of the contingent claim could be validated as mean value of option prices. In Chapter 3, we shall derive a closed-form formula for the averaged call price (3.108), given a probability density function Π for the initial volatility. We also generalize this approach by introducing a different notion of abritrage with respect to those given in Duffie (2001). In particular we apply this notion to a market represented by stocks driven by a Brownian motion with coefficients depending by random parameters and we state a general pricing relation for a contingent claim in such a market. In the case of the Heston model, given the functon

$$C_\Pi(S_0, T, K; \Theta) = \int_0^{+\infty} C^H(S_0, v, T, K; \Theta)\Pi(v)\, dv. \tag{15}$$

for a probability density function Π such that $\Pi(v) = 0$ for any $v \leq 0$, we deduce the analogous of equation (2.25) in the following result.

Theorem 0.0.3. (**Average Call Price**) If the function $v \mapsto v\Pi(v)$ belongs to $L^1(\mathbb{R})$, then the following relation holds true:

$$C_\Pi(S_0, T, K; \Theta) = S_0 Q_1(S_0, T, K; \Theta) - e^{-rT} K Q_2(S_0, T, K; \Theta), \tag{16}$$

where

$$Q_j = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\Big[\frac{e^{C_j(T,\phi)+i\phi\log(\frac{S_0}{K})}M_\Pi\big(D_j(T,\phi)\big)}{i\phi}\Big]\,d\phi, \qquad (17)$$

for $j = 1,\ 2,\ S_0,\ T,\ K > 0,\ r > 0,\ \Theta \in \mathcal{H}$, with $\rho \in (-1, 1)$, M_Π being the moment generating function related to Π (see (3.96) in Chapter 3) .

For the Heston model, we also prove that for a large class of probability distributions assumed for the initial volatility parameter, the estimation error in the calibration procedure of option prices is less than the case of the simple pricing formula. Our results are validated with numerical comparisons, on observed call prices, between the proposed calibration method and the classical approach. We have compared the standard methods which considers v_0 as an additional parameter and our approach under three distributions: the Gamma (GAM), the Inverse Gaussian (IG) and the Generalized Inverse Gaussian (GIG), for which the integrands appearing are explicitly known. From a numerical point of view, the calculation of the option price is made somewhat complicated by the fact that the integrands have oscillatory nature. However, the integration can be done in a reasonably simple fashion by the aid of Gauss-Lobatto quadrature. This integration method is capable of handling a wide range of functional forms.

In Chapter 4, we use the principle of finite difference methods to approximate the differential operator defined for the pricing problem studied (using here again the Heston model, as case study). The finite difference methods (FD) for derivatives are one of the simplest and oldest methods to solve differential equations. It was already known by L. Euler (1707-1783) ca. 1768, in one dimension of space and was probably extended to dimension two by C. Runge (1856-1927) ca. 1908.

The principle of finite difference methods consists in approximating the differential operator by replacing the derivatives in the equation using differential quotients. The aim of these methods(which could be different depending on the approximation used) consists into to evaluate the values of a continuous function $f(t, S)$ on a set of discrete points in (t, S) plane (or in a high dimensional ones). The discretization will produce a systems of ODEs as

$$C'(t) = F(t, C(t))(0 \leq t \leq T), C(0) = C_0. \tag{18}$$

This method defines approximation C^n to the exact solutions value $C(t_n)$ subsequently for $n = 1, 2, 3, ...N$. In case of (18) we have

$$F(t, w) = Aw + b(t) \quad 0 \leq t \leq T, w \in \mathbb{R}^m. \tag{19}$$

Thus, each step requires the solution of a system of linear equations involving the matrix $(I - \frac{1}{2}\Delta t A)$ where I denotes the m $\times$ m identity matrix. Generally speaking, finite-difference schemes can be divided into two classes: implicit FD schemes and explicit FD schemes. The θ-schemes refer to those scheme in which are balanced both explicit and implicit scheme. The most famous of these last schemes is the Crank -Nicolson scheme, obtained by taking average of these two schemes. That is, the approximation is obtained as:

$$C^n = C^{n-1} + \frac{1}{2}\Delta t F(t^{n-1}, C^{n-1}) + \frac{1}{2}\Delta t F(t^n, C^n).$$

In our application to the two-dimensional Heston PDE, the dimension usually gets very large and the Crank-Nicholson scheme becomes ineffective. The reason for this is that $(I - \frac{1}{2}\Delta t A)$, and hence the matrices in its LU factorization, possess a bandwidth. For the numerical solution of the semi-discrete Heston PDE we shall study a splitting schemes of the Alternating Direction Implicit (ADI) type. In the past decades, ADI schemes have been successful already in many application areas.

As mentioned above, even if a boundary condition has been proved to not be necessary from an analytical point of view, we still need such a condition for the implementation of any numerical scheme. We decide to implement the PDE defined in [30] (proved that this is consistent with the theoretical solution) and used the Lyapunov function in this implementation: balancing the nodes in the boundary with the function found out. So, established such a function, we will use this one for the implementation of our numerical scheme: For the Heston model it is:

$$L(s,v) = -\log(v) - \log(s) + s\log(s+3) + s(v+1) + v.$$

To discretize the domain, we introduce an equi-distributed grid points corresponding to a spatial step size $\Delta s = 1/(N_s+1)$, $\Delta v = 1/(N_v+1)$ and to a time step $\Delta t = 1/(M+1)$, where M, N_s, N_v are positive integers, number of time steps, of nodes in S and V direction respectively. We define the nodes of the regular grid as:

$$(t_n, s_i, v_j) = (n\Delta t, i\Delta s, j\Delta v)$$

with $n \in 0,..,M+1, i \in 0,...,N_s+1, i \in 0,...,N_v+1$. And we denote as $C^n_{i,j}$ the value of an approximate solution at point (t_n, s_i, v_j) and $C(t,s,v)$ the exact solution of problem. The initial data must also be discretized as:

$$C^0_{i,j} = C^0(s_i, v_j) \quad \forall\, i \in \{1,\ldots,N_s+1\} \quad \forall\, j \in \{1,\ldots,N_v+1\}. \tag{20}$$

The problem is then to find, at each time step, a vector $C_{i,j} \in \mathbb{R}^2$, such that its components are the values $(C^n_{i,j})_{1\le i\le N_s\ 1\le i\le N_v}$. We decompose the matrix A into three submatrices,

$$A = A_0 + A_1 + A_2. \tag{21}$$

We choose the matrix A_0 as the part of A that stems from the FD discretization of the mixed derivative term in. Next, in line with the classical ADI idea, we choose A_1 and A_2 as the two parts of A that correspond to all spatial derivatives in the s- and v-direction, respectively. The FD discretization described implies that A_1, A_2 are essentially tridiagonal and pentadiagonal, respectively. The ADI scheme considered, the Douglas scheme ([43], [24]), develops as

$$\begin{cases} Y_0 = & C_{n-1} + \frac{1}{2}\Delta t F(t_{n-1}, C_{n-1}), \\ Y_j = & Y_{j-1} + \frac{1}{2}\Delta t[F_j(t_n, Y_j) - F_j(t_{n-1}, C_{n-1})], \qquad j = 1,2, \\ C_n = & Y_2. \end{cases}$$

The splitting schemes treats the mixed derivative part F_0 in a fully explicit way. F_1 and F_2 parts are treated implicitly in the schemes. The results obtained are figured as follows below:

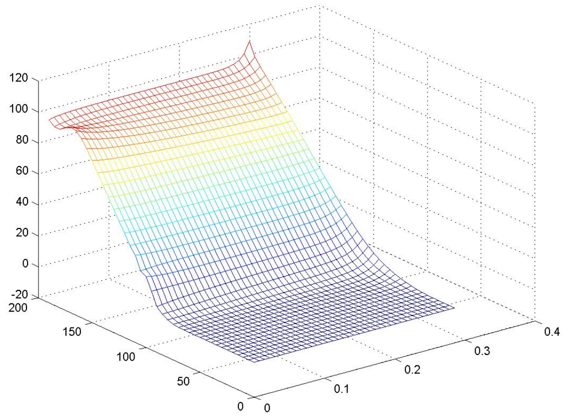


Figure 2: Solution by FD and Lyap

We highlight that our modified scheme converges to the exact solution quicker that the scheme which does not use the Lyapunov function along the boundary. Moreover this scheme gets all the stability requirements of the original scheme. In fact, we derive new linear stability results for this ADI schemes that have previously been studied in the literature ([44], [56], [13]). These results are subsequently used to show that the ADI scheme under consideration are unconditionally stable when applied to finite difference discretizations of general parabolic two-dimensional convection diffusion equations. In the end, the focal point of the last part of analysis is that: even if the numerical scheme without the use of the Lyapunov function is converging to the exact value, the use of the function speeds up this convergence even close to the nodes along the boundary where $v = 0$. The use of the Lyapunov function even in the numerical scheme will allow us to obtain the convergence in a shorter time, decreasing the numerical error.

Chapter 1

Option Pricing

In order to introduce the mathematical framework necessary in the following, we have made an extensive use in this chapter of many of the most known books in financial mathematics. For the more general concept we have employed [3], [15], [48], [66] and [49], while for what concerns the stochastic differential equations we have consulted [60] and [62].

1.1 Stochastic Processes

A stochastic process is a family $(X_t)_{t\in[0,T]}$ of real-valued random variables defined on the same probability space $(\Omega, \mathcal{F}, P)$ indexed by time. The time parameter t can be either discrete or continuous, but for our purposes, we will consider continuous-time stochastic processes. For each realization of the randomness ω, the trajectory $X(\omega) : t \to X_t(\omega)$ defines a function of time, called the sample path of the process. Thus, a stochastic processes can also be viewed as random functions: random variables taking values in function spaces. In a dynamic context, as time goes on, more information is progressively revealed to the observer. It is necessary to add some time dependent ingredient to the structure of our probability space to accommodate this additional feature, that is done by introducing the concept of filtration. Moreover, in the sequel, we consider $X(t)$ and X_t as equivalent notations for the random variable representing the value of the process at time t.

Definition 1.1.1 (Filtration)**.** A filtration on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ is an increasing family of $\sigma - algebras$ $(\mathcal{F}_t)_{t\in[0,T]}$ such that:

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F} \qquad \forall\ 0 \leq s \leq t \leq T. \tag{1.1}$$

A probability space equipped with a filtration is called a filtered space.

The information flow is described by the filtration $\mathcal{F}_t$ and we can now distinguish quantities which are known given the current information from those which are still viewed as random at time t. An $\mathcal{F}_t$-measurable random variable is nothing else but a random variable whose value will be revealed at time t; similarly a process whose value at time t is revealed by the information $\mathcal{F}_t$ is said to be non-anticipating or adapted. We recall here that, if $(E, \mathcal{E})$ and $(F, \mathcal{F})$ are measure spaces, a function $f : F \to E$ is said to be $\mathcal{F}$*-measurable* if for every subset $A \subset \mathcal{E}$, $f^{-1}(A) = \{x \in F : f(x) \in A\} \in \mathcal{F}$.

Definition 1.1.2. Given a filtration $(\mathcal{F}_t)_{t\in[0,T]}$, a stochastic process $(X_t)_{t\in[0,T]}$ is said to be $\mathcal{F}_t$-adapted if, for each $t \in [0, T]$, the random variable X_t is $\mathcal{F}_t$-measurable.

If we consider a $\mathcal{F}_t$-adapted stochastic process $(X_t)_t$ as a function from the measure space $([0, T]\times\Omega, \mathcal{B}([0, T])\times\mathcal{F})$ to $(E, \mathcal{E})$, the following definition is straightforward.

Definition 1.1.3. A stochastic process $\left(X_t\right)_{t\in[0,T]}$ is said to be measurable if the function $(t, \omega) \mapsto X(t, \omega)$ is measurable from $([0, T] \times \Omega, \mathcal{B}([0, T]) \times \mathcal{F})$ to $(E, \mathcal{E})$.

A slightly different notion is introduced in the following definition.

Definition 1.1.4. An $\mathcal{F}_t$-adapted stochastic process $\left(X_t\right)_{t\in[0,T]}$ is said to be progressively measurable if, for every $t \in [0, T]$, the function $(t, \omega) \mapsto X(t, \omega)$ is measurable from $([0, t] \times \Omega, \mathcal{B}([0, t]) \times \mathcal{F}_t)$ to $(E, \mathcal{E})$.

Whenever any other filtration is specified we will assume that the probability space is endowed with the natural filtration generated by the stochastic process under study.

Definition 1.1.5. Given a filtration $(\mathcal{F}_t)_t$ and a random time $\tau : \Omega \to [0, +\infty)$, we say that τ is a $\mathcal{F}_t$-stopping time if

$$\{\tau \leq t\} \in \mathcal{F}_t \quad \forall t \geq 0. \tag{1.2}$$

Moreover, we define the σ-algebra

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_T : A \cap \{\tau \leq t\} \in \mathcal{F}_t \quad , \forall t \in [0, T]\}. \tag{1.3}$$

In other words, a positive random variable that represents the time at which some event is going to take place is an $\mathcal{F}_t$-stopping time if given the information in $\mathcal{F}_t$ one can determine whether the event has happened or not. So, a positive random variable that represents the time at which some event is going to take place is an $\mathcal{F}_t$-stopping time if given the information in $\mathcal{F}_t$ one can determine whether the event has happened or not.

Definition 1.1.6. $\Lambda^p[0, T]$, $p \geq 1$, is the class of progressively measurable stochastic processes $(X_t)_{t\in[0,T]}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ such that

$$\int_0^T |X_s|^p ds < \infty, \tag{1.4}$$

$\mathbb{P}$-almost surely.

Definition 1.1.7. $\mathcal{M}^p[0, T]$, $p \geq 1$, is the class of stochastic processes $(X_t)_{t\in[0,T]} \in \Lambda^p[0, T]$ such that

$$\mathbb{E}\left[\int_0^T |X_s|^p ds\right] < \infty. \tag{1.5}$$

Definition 1.1.8. Consider a probability space $(\Omega, \mathcal{F}, \mathcal{F}, \mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)_{t\in[0,T]}$. The $\mathcal{F}$-adapted process $(W_t)_t$ is a standard d-dimensional Brownian motion if only if its components $(W_t)_t = (W_t^1, W_t^2\ , \ldots,\ W_t^d)$ are one-dimensional independent Brownian motions.

Suppose that $(\mu(t))_t \in \Lambda^1[0,T]$, and that $(\sigma(t))_t \in \Lambda^2[0,T]$, then, for each $x_0 \in \mathbb{R}$

$$X(t) = x_0 + \int_0^t \mu(u)du + \int_0^t \sigma(u)dW(u) \tag{1.6}$$

defines a stochastic process called an *Ito process* [64]. It is customary, and convenient, to express such an equation in differential form, in terms of its stochastic differential:

$$\begin{cases} dX(t) = \mu(t)dt + \sigma(t)dW(t), \\ X(0) = x_0. \end{cases} \tag{1.7}$$

For a given d-dimensional Brownian motion $(W_t)_t = (W_t^1, W_t^2, \dots, W_t^d)$, $d \geq 1$, it is possible to consider an n-dimensional Ito process $X : \Omega \times [0,T] \to \mathbb{R}^n$ on the same filtered probability space, such that each component $X_i(t)$, for $i = 1, \dots, n$, has the stochastic differential

$$\begin{cases} dX_i(t) = \mu_i(t)dt + \sum_{j=1}^d \sigma_i^j(t)dW^j(t), \\ X_i(0) = x_{i,0} \in \mathbb{R}, \end{cases} \tag{1.8}$$

where $(\mu_i(t))_t \in \Lambda^1[0,T]$, $(\sigma_i^j(t)) \in \Lambda^2[0,T]$ for all $i = 1, \dots, n$, $j = 1, \dots, d$. Now suppose that $f : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ is continuous with all the derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x_i}$, $\frac{\partial^2 f}{\partial x_i x_j}$, $i,j = 1, \dots, n$, on $[0,T] \times \mathbb{R}^n$, then the question arises of giving a meaning to the stochastic differential $df(t, X(t))$ of the process $f(t, X(t))$. The well known Ito's lemma gives the analogue of the chain rule for stochastic calculus, assuming the hypothesis above.

Theorem 1.1.1. (**Ito's Lemma**) If $(X(t))_t$ is n-dimensional satisfying (1.8), then the stochastic process $(f(t, X(t)))_t$ has the following stochastic differential:

$$\begin{aligned} df(t, X(t)) \quad &= \quad \left[\frac{\partial f}{\partial t} + \sum_{i=1}^n \mu_i(t) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n (\sigma\sigma^T)_{ij} \frac{\partial^2 f}{\partial x_i x_j} \right] dt \\ &\quad + \sum_{i=1}^n \sum_{j=1}^d \frac{\partial f}{\partial x_i} \sigma_i^j(t) dW^j(t), \end{aligned} \tag{1.9}$$

where, for every $t \in [0,T]$, $\sigma(t)$ stands for the $\mathbb{R}^{n \times d}$ matrix with entries $(\sigma_i^j(t))_{i,j}$. Here all the partial derivatives of f are evaluated at $(t, X(t))$.

Example 1.1.1. An over-abused example of stochastic process was introduced by Black-Scholes model (1973) to describe the evolution for the underlyng stock of an option, is the *Geometric Brownian Motion*:

$$S(t) = s_0 \exp\left\{\sigma W(t) + \left(\mu - \frac{1}{2}\sigma^2\right)t\right\} \tag{1.10}$$

where $(W(t))_t$ is a one-dimensional Brownian motion, $s_0 > 0$, μ and $\sigma > 0$ are constant coefficients. Thus, given the function

$$f(t,x) = s_0 \exp\left\{\sigma x + \left(\mu - \frac{1}{2}\sigma^2\right)t\right\}, \tag{1.11}$$

clearly, it hods $S(t) = f(t, W(t))$. Then, applying Ito's formula, we can derive the stochastic differential of $S(t)$

$$\begin{aligned} dS(t) &= df(t, W(t)) \\ &= (\mu - \frac{1}{2}\sigma^2) f dt + \frac{1}{2}\sigma^2 f dt \\ &= \mu S(t)dt + \sigma S(t)dW(t). \end{aligned} \tag{1.12}$$

By the results reported in next section allow us to establish that the Geometric Brownian motion is the unique solution to the stochastic differential equation (1.12). We recall some useful results related to the concept of martingale.

Definition 1.1.9. Let $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space. An $(\mathcal{F}_t)_{t\geq 0}$-adapted stochastic process $(M_t)_{t\geq 0}$ is a *martingale* if $\mathbb{E}\left[|M_t|\right] < \infty$, for all $t \geq 0$, and

$$\mathbb{E}\left[M_s | \mathcal{F}_t\right] = M_t, \tag{1.13}$$

for all $s \geq t$.

Definition 1.1.10. Let $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space. An $(\mathcal{F}_t)_{t\geq 0}$-adapted stochastic process $(M_t)_{t\geq 0}$ is called *local martingale* if there existis a sequence of stopping times $\{\tau_n\}_{n\geq 1}$ on $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ such that

i) τ_n are almost surely increasing: $\mathbb{P}(\tau_n < \tau_{n+1}) = 1$, for all $n \geq 1$;

ii) τ_n diverge almost surely: $\mathbb{P}(\tau_n \to \infty, \quad \text{as } n \to \infty) = 1$;

iii) the stopped process $(M_{t\wedge\tau_n})_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-martingale, for any $n \geq 1$.

There are two popular generalizations of a martingale that also include cases when the current observation of the process is not necessarily equal to the future conditional expectation. In particular we recall that $(M_t)_{t\geq 0}$ is called sub (*respectively* super) -martingale if the inequality $\geq$ (*respectively* $\leq$) holds in (1.13). A well known result shows that any local martingale bounded from below is a super-martingale (see, for example, [49]).

Theorem 1.1.2. (**Martingale Representation Theorem**) Let $(W_t)_{t\geq 0}$ be n-dimensional Brownian motion on $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Suppose that $(M_t)_{t\geq 0}$ is a martingale on such a space and that $M_t \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, for all $t \geq 0$. Then there exists a unique stochastic process $(g_s)_{s\in[0,t]} \in \mathcal{M}^2[0,t]$, for all $t \geq 0$ such that

$$M_t(\omega) = \mathbb{E}\left[M_0\right] + \int_0^t g_s(\omega) dW_s(\omega),$$

for $\omega \in \Omega$ $\mathbb{P}$-a.s., for all $t \geq 0$.

Theorem 1.1.3. (**Girsanov Theorem**) Let $(W_t)_{t\geq[0,T]}$, $T > 0$, be n-dimensional Brownian motion on $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$. Let $\lambda \in \mathcal{M}^{2,n}[0,T]$ be such that

$$\xi_t^\lambda := \exp\left(-\int_0^t \lambda_s dW_s - \frac{1}{2}\int_0^t |\lambda(s)|^2 ds\right), \qquad t \in [0,T],$$

is a martingale on $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$. Then the process

$$\tilde{W}_t := W_t + \int_0^t \lambda_s ds,$$

is a Brownian motion on the space $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]})$ under the probability measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, defined through by its Radon-Nikodym derivative:

$$\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega) := \xi_T^\lambda(\omega).$$

A sufficient condition under which ξ^λ is a strictly positive martingale is given by the so called *Novikov* condition:

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T |\lambda(s)|^2 ds\right)\right] < \infty.$$

We have also the following representation result for Ito processes.

Theorem 1.1.4. (**Diffusion Invariance Principle**) Let $(W_t)_{t\geq 0}$ be n-dimensional Brownian motion on $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. If $(X_t)_{t\geq 0}$ is a $\mathbb{R}^n$-valued Ito process satisfying:

$$dX_t = \mu_t dt + \sigma_t dW_t,$$

with $(\mu_s)_{s\in[0,t]} \in \mathcal{M}^{1,n}[0,t]$, $(\sigma_s)_{s\in[0,t]} \in \mathcal{M}^{2,n\times n}[0,t]$, for any $t > 0$ and it is a martingale under the probability measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, on $(\Omega, \mathcal{F}_T)$, then there exists a n-dimensional Brownian motion $(\tilde{W}_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ such that $dX_t = \sigma_t d\tilde{W}_t$.

1.2 Stochastic Differential Equations

In this section, we recall some well known results related to the theory of stochastic differential equations (SDEs). Let $\mu : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$ be Borel-measurable functions. Let $(\Omega, \mathcal{F}, \mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ the a filtered probability space and, let $(W_t)_{t\in[0,T]}$ be an adapted d-dimensional standard Brownian Motion of such a space.

Definition 1.2.1. Let Z an $\mathcal{F}_0$-measurable $\mathbb{R}^n$-valued random variable. We will say that the process $(X_t)_{t\in[0,T]}$ is a *solution of the the stochastic differential equation*

$$\begin{cases} dX_t &= \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \\ X_0 &= Z, \end{cases} \tag{1.14}$$

if, for every $t \in [0,T]$, it holds

$$X_t = Z + \int_0^t \mu(u, X_u)d(u) + \int_0^t \sigma(u, X_u)dW_u \tag{1.15}$$

Definition 1.2.2. We say that the stochastic differential equation (1.14) has a *strong solution* on the given probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ with respect to the fixed d-dimensional Brownian motion $(W_t)_t$ if exists a stochastic process $(X_t)_t$ satisfying:

i) (X_t) is adapted to the filtration $\mathcal{F}_{\sqcup t}$;

ii) $\mathbb{P}(X_0 = Z) = 1$;

iii) $(\mu(t, X_t))_t$ and $(\sigma(t, X_t))_t$ belong to $\Lambda^1[0, T]$ and $\Lambda^2[0, T]$, respectively;

iv) With probability one, the process satisfies (1.15).

With this definition at hand the notion of the existence of a strong solution is clear. We will say that strong uniqueness of a solution holds, only if the construction of a strong solution is unique on any probability space carrying the random elements $(W_t)_t$ and Z, where Z is an arbitrary initial condition.

Definition 1.2.3. Suppose that, whenever $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ is a filtered probability space equipped with a d-dimensional Brownian motion $(W_t)_t$ and a $\mathcal{F}_0$-measurable, $\mathbb{R}^n$-valued random variable Z, any two strong solutions $(X_t)_t$, $(X'_t)_t$ of (1.14) with initial condition Z satisfy $\mathbb{P}(\forall\ t \in [0, T]\ : X_t = X'_t) = 1$. Then we say that *strong uniqueness* holds for equation (1.14).

We remark that since solution processes are, by definition, continuous, and $[0, T]$ is separable, it suffices to have the weaker condition $\mathbb{P}(X_t = X'_t) = 1$ for all $t \in [0, T]$ in the above definition.

The concept of strong solution is opposed to the one of weak solution. A strong solution is a weak solution, and if σ is Lipschitz, then any weak solution is a strong solution. In particular two weak solutions on the same space involving the same Brownian Motion are identical, where neither the probability space nor the BM are previously assigned. Both require the existence of a process X_t that solves the integral equation version of the SDE. The difference between the two lies in the underlying probability space

$(\Omega, \mathcal{F}, \mathbb{P})$. A weak solution consists of a probability space and a process that satisfies the integral equation, while a strong solution is a process that satisfies the equation and is defined on a given probability space.

Definition 1.2.4. A weak solution of the stochastic differential equation (1.14) with initial condition Z is a stochastic process $(X_t)_t$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that for some admissible filtration $(\mathcal{F}_t)_t$ and for some Wiener process $(W_t)_t$ adapted to such a filtration, the process $(X_t)_t$ is $(\mathcal{F}_t)_t$-adapted and satisfies the stochastic integral equation (1.15).

In our thesis, we are interested in strong solutions, therefore unless specific cases, in the following we shall assume that the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ and the d-dimensional Brownian motion $(W_t)_t$ are fixed. We recall the following existence and uniqueness result to strong solutions for SDEs. For the sake of simplicity, we shall adopt the notation $||\cdot||$ both for the Euclidean norm a vector in $\mathbb{R}^n$ and for the norm of a matrix, meaning that

$$||\sigma|| = \left[\sum_{i=1}^{n}\sum_{j=1}^{d}|\sigma_i^j|^2\right]^2, \qquad \forall\, \sigma \in \mathbb{R}^{n\times d}.$$

Theorem 1.2.1. (Existence and uniqueness for SDEs)
Consider the SDE (1.14) where $\mu : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0,T]\times\mathbb{R} \to \mathbb{R}^{n\times d}$ are continuous functions satisfying:

***i*)** (Local Lipshitz continuity) for all $\nu \in \mathbb{N}$ there exists $K_\nu \in \mathbb{R}$ so that

$$||\mu(t,x)-\mu(t,y)||^2 + ||\sigma(t,x)-\sigma(t,y)||^2 \le K_\nu||x-y||^2, \tag{1.16}$$

for $||x||, ||y|| \le \nu$, $t \in [0,T]$;

ii) (Growth condition) there exists a constant $K \in \mathbb{R}$ so that

$$||\mu(t,x)||^2 + ||\sigma(t,x)||^2 \le K(1+||x||^2), \tag{1.17}$$

for $x \in \mathbb{R}^n$, $t \in [0,T]$.

iii) The initial condition Z is $\mathcal{F}_0$-measurable and square integrable (i.e. $\mathbb{E}[|Z|^2] < \infty$).

Then (1.14) has a unique strong solution $(X_t)_t \in \mathcal{M}^2[0,T]$ if $\tilde{X}_t$.

The proof of the existence and uniqueness theorem stated above follows the lines of the classical proofs for existence and uniqueness of solutions of ordinary differential equations, with appropriate modifications for the random terms. See, for instance, Karatzas and Shreve (1991) as a reference. It is useful to give here the definition of equivalent measure.

Definition 1.2.5. Given two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on the probability space $(\Omega, \mathcal{F})$, we say that $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ if for any $A \in \mathcal{F}$, with $\mathbb{P}(A) = 0$, it holds $\mathbb{Q}(A) = 0.$. In this case, we can define the Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$ as the $\mathcal{F}$-measurable, $\mathbb{P}$-integrable function $h : \Omega \to \mathbb{R}_+$ satisfying:

$$\mathbb{Q}(A) = \int_A h(\omega) d\mathbb{P}(\omega), \qquad\qquad A \in \mathcal{F}. \tag{1.18}$$

We will use the notation $d\mathbb{Q} = hd\mathbb{P}$ to denote that $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ with Radon-Nikodym derivative h. The measues $\mathbb{P}$ and $\mathbb{Q}$ are said to be equivalent i f $\mathbb{P}$ is absolutely continuous with respect to $\mathbb{Q}$ and $\mathbb{Q}$ is also absolutely continuous with respect to $\mathbb{P}$.

1.3 The Feynman-Kac Theorem

The Faynman-Kac Theorem is extensively used in many applications included in pricing European-type financial derivatives.

Theorem 1.3.1 (Feynman-Kac Theorem). Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space and suppose that $(X_t)_t$ is the solutions of the SDE (1.14), where the coefficients μ and σ satisfy the assumptions of Theorem 1.2.1. Consider the parabolic partial differential equation:

$$\frac{\partial V}{\partial t}(t,x) + \mu^T \nabla V(t,x) + \frac{1}{2} tr(\sigma\sigma^T \nabla^2 V) - rV(t,x) = 0, \tag{1.19}$$

defined for all $x \in \mathbb{R}^n$ and $t \in (0,T)$, subject to the terminal condition $V(T,x) = \Phi(x)$, where $\Phi : \mathbb{R}^n \to \mathbb{R}$ is a continuous function and $r \in \mathbb{R}$ is a positive constant (usually defined as the short-term interest rate in pricing problems). The theorem asserts that if $V \in C([0,T]\times\mathbb{R}^n) \cap C((0,T)\times\mathbb{R}^n) \cap$ is a solution to equation (1.19) satisfying

$$\mathbb{E}\big[\int_0^T \big(e^{-rs}||\nabla V^T \sigma(s,X_s)||^2\big)ds\big] < \infty, \tag{1.20}$$

then, the following holds true

$$V(t,X_t) = \mathbb{E}\big[e^{-r(T-t)}\Phi(X_T)|\mathcal{F}_t\big], \tag{1.21}$$

for any $t \in [0,T]$, $\mathbb{P}$ almost surely.

The Feynman-Kac theorem illustrates the close connection between stochastic differential equations and partial differential equations and it can be used in both directions. One of our objectives is to study enough general conditions to guarantees the existence and the uniqueness of the solution to the Dirichlet problem (1.19) associated with stochastic volatility models used in option pricing. Often in such models, the domain of the problem is not the whole space $\mathbb{R}^n$ and the differential operator is degenerate at the boundary.

1.4 Option Pricing

The stochastic calculus has the appearance of having been expressly designed as a tool for financial analysis, so naturally does it fit the application. Stochastic calculus is now the language of pricing models and risk management. All continuous-time models are based on Brownian motion, despite the fact that most of the results extend easily to the case of a general abstract information filtration.

A financial derivative, for example an option, is an instrument whose value depends on the values of some underlying variables, where the underlying can be a commodity, an interest rate, a stock, a stock index, a currency, etc. This payment is called payoff. An example widely discussed in what follows is the European call option. The buyer of the call option has the right, but not the obligation, to buy an agreed quantity of the underlying from the seller of the option at a certain time (the expiration date) for a certain price (the strike price). The seller incurs a corresponding obligation to fulfill the transaction, that is to sell if the option holder elects to "exercise" the option at expiration. The buyer pays a premium to the seller for this right. Once a financial derivative is defined the first question is the following: "'what is the fair price that the seller of the derivative should charge to the buyer"'. In order to answer this question we introduce a few concepts. First af all we will suppose the market to be frictionless. This means

1. No transaction costs: no cost incurred in making an economic exchange;
2. Perfect liquid markets: the assets traded in the market can be sold without causing a significant movement in the price and with minimum loss of value;
3. No taxes;
4. No restrictions on short sales: the practice of selling securities or other financial instruments that are not currently owned is allowed for any quantity of the financial instrument;
5. no transaction delays.

Finally, we suppose that investors are allowed to trade continuously up to some fixed finite planning horizon T, where all economic activity stops. An essential feature of market is based on the absence of arbitrage opportunities. This assumption can be interpreted as a market equilibrium condition.

1.4.1 The Arbitrage Principle

Let $(W_t)_{t\in[0,T]}$ be a standard d-dimensional Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$. We consider a d-dimensional market with a set of risky non-dividend-paying assets, with price at time t

given by $S_i(t)$, $i = 1, \dots n$, and a riskless asset, a bond, with price at time t given by $S_0(t)$. The price process vector $(S_t)_t = (S_0(t), S_1(t), \dots, S_n(t))_t$ is driven by the following stochastic system of equations:

$$\begin{cases} dS_i(t) = \mu_i(t)dt + \sum_{j=1}^d \sigma_i^j(t) dW_t^j, \\ S_i(0) = s_i > 0 \quad i = 1, \dots, n, \end{cases} \tag{1.22}$$

and

$$\begin{cases} dS_0(t) = r_t S_0(t)dt, \\ S_0(0) = s_0 > 0. \end{cases} \tag{1.23}$$

where μ and σ must meet the conditions of 1.2.1. In particular the short-term interest rate $(r_t)_t$ belongs to $\mathcal{M}^1[0,T]$. Without loss of generality, we can assume that $s_0 = 1$ so that $S_0(t) := \exp\left(\int_0^t r_u du\right)$. We also assume that initial prices s_i, $i = 1, \dots, n$ are known constants.

Definition 1.4.1. An admissible *portfolio* (strategy) for the market represented by (1.22)-(1.23) is a $\mathbb{R}^{n+1}$-valued progressively measurable process $(\theta_t)_t = \{(\theta_0(t), \dots \theta_n(t))\}_t$ with respect to the filtration $(\mathcal{F}_t)_t$ such that $\langle \theta(t), \mu(t) \rangle \in \mathcal{M}^1[0,T]$ and $\langle \theta(t), \sigma^j(t) \rangle \in \mathcal{M}^2[0,T]$ for any $j = 1, \dots, d$. The *value* at time t of the portfolio associated with $(\theta_t)_t$ is given by $\langle \theta_t, S_t \rangle$.

Definition 1.4.2. We say that a portfolio strategy $(\theta_t)_t \in \Lambda(S)$ is *self-financing* if the following holds true:

$$\begin{aligned} \langle \theta_t, S_t \rangle &= \langle \theta_s, S_s \rangle + \int_s^t \langle \theta_u, dS_u \rangle := \\ &= \langle \theta_s, S_s \rangle + \int_s^t \langle \theta_u, \mu_u \rangle du + \sum_{j=1}^d \int_s^t \langle \theta_u, \sigma_u^j \rangle \, dW_u^j, \end{aligned} \tag{1.24}$$

for any $0 \leq s \leq t \leq T$.

The set of all self-financing strategies for the market S is denoted as $\Theta(S)$. We also recall the standard notion of arbitrage, as in Duffie (2001).

Definition 1.4.3. A portfolio strategy $\theta \in \Theta(S)$ is an *arbitrage* if

$$\langle \theta(0), S(0) \rangle \leq 0 \leq \langle \theta(T), S(T) \rangle \qquad \mathbb{P}\text{-a.s.} \tag{1.25}$$

and

$$\mathbb{P}(\langle \theta(T), S(T) \rangle > 0) > 0. \tag{1.26}$$

In financial modelling is a common practice to consider a state price deflator to discount the price process S. In our case (non dividend stocks) we choose $Y(t) := [S_0(t)]^{-1} = \exp(-\int_0^t r_u du)$ as state price deflator.
Therefore, we consider the discounted price process $V(t) = Y(t)S(t)$. In the following with $\mathbb{E}[\cdot]$ we denote the expected value with respect to measure $\mathbb{P}$; if the expected value is calculated with respect to a measure $\mathbb{Q}$ different from $\mathbb{P}$, it will be denoted with $\mathbb{E}^{\mathbb{Q}}[\cdot]$. We say that a probability measure $\mathbb{Q}$ defined on $(\Omega; \mathcal{F}_T)$ is an equivalent martingale measure if $\mathbb{Q}$ is equivalent to $\mathbb{P}$ according to Definition (1.2.5) and the discounted stock-price process $(V_t)_t$ is a $\{\mathcal{F}_t\}_{t\in[0,T]}$-martingale under $\mathbb{Q}$. The existence of an equivalent martingale measure is related to the absence of arbitrage, while the uniqueness of the equivalent martingale measure is related to market completeness, see ([4], [46]). We also recall that Battig (1999), making use of functional analytic methods, gives a very general definition of completeness in a large financial market, which is invariant with respect to a change in probability and independent of No Arbitrage.

Proposition 1.4.1. Completeness of the market is equivalent to uniqueness of the equivalent martingale measure when one of the following conditions is fulfilled:

- the market contains a finite number of assets;
- every asset price process has continuous trajectories;
- the filtration $\mathcal{F}$ (generated by the asset price processes) is strictly left continuous, that is, for all stopping times τ , we have $\mathcal{F}_\tau = \mathcal{F}_{\tau_-}$

1.5 Black and Scholes Model

The Black and Scholes model was first published by Fischer Black and Myron Scholes in the 1973 on the paper entitled "The Pricing of Options and Corporate Liabilities", published in the Journal of Political Economy. After 42 years of its first publication, the problem of derivative pricing which

could overcome the Black and Scholes model has not been completely solved. The difficult concerning derivative pricing is to define a fair price. The just mentioned Black-Scholes model represents a universal accepted framework for derivative pricing. In his now famous doctoral thesis, Bachelier (1900) proposed to model asset prices by an arithmetic Brownian motion, i.e. he suggested the model $dS_t = \mu dt + \sigma dW_t$, for constants $\mu, \sigma > 0$ and for a 1-dimensional Brownian motion. While this was a good first approximation to the dynamics of stock prices, arithmetic Brownian motion has one serious drawback: as S_t is normally distributed with mean $S_0 + \mu t$ and variance $\sigma^2 t$, the asset price can become negative with positive probability, which is at odds with the fact that real-world stock-prices are always nonnegative because of limited liability of the shareholders.
Samuelson (1965) therefore suggested replacing arithmetic Brownian motion by geometric Brownian motion. This model assumes the existence of a risk-free asset B_t and of an underlying asset S_t, following respectively a deterministic and a geometric Brownian motion dynamics:

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t dW_t, \\ dB_t = r B_t dt, \end{cases} \tag{1.27}$$

where the deterministic parameters μ, σ and r represent respectively the local mean rate of return of the asset, the volatility of the asset and the short-term interest rate. Under the Black and Scholes setting, the no arbitrage value driving the price evolution of a European-type option with maturity $T > 0$, is the unique solution to the following partial differential equation (PDE) :

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rSS(t)\frac{\partial C}{\partial S} - rC = 0, \tag{1.28}$$

for all $S > 0$, $t \in (0, T)$, where C is the price of the option as a function of stock price S and time t, with terminal condition, at maturity T, given by the payoff of the option, satisfying suitable boundness conditions. In the case of a European call option, there exists an equivalent martingale measure $\mathbb{Q}$ for the discounted price process of the underlying such that

$$C(t, S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(S_T - K)^+|\mathcal{F}_t]. \tag{1.29}$$

Under such a measure, the price process $(S_t)_t$ is a geometric Brownian motion under $\mathbb{Q}$:

$$S_T = S_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}})}, \tag{1.30}$$

so we can write

$$C(t, S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(S_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}})} - K)^+|\mathcal{F}_t]. \tag{1.31}$$

This relation can explicitly evaluated in closed form. The Black-Scholes model contains some rather strong simplifications of the real world: the model implies that the log-returns are normally distributed but historical returns often displays non-normal features; the model assumes zero transaction costs (without considering that the replicating strategy yields infinitely large losses); the volatility is assumed to be known and constant over time. This last assertion implies that all options on a specific stock should have the same implied volatility. Actually this is not the case. Instead, implied volatility empirically vary with strikes (K) and maturities (T). The implied volatility graphed against strike for a fixed maturity is known as the volatility smile or the volatility skew. Plotting the implied volatility against both strike and time to maturity gives the so called volatility surface.

The names refers to the forms these curves often take, which may vary in different markets. Stock options for example often shows a skew effect, while in the currency world we are more likely to see a smile curve. Implied volatility also seems to vary with time to maturity, a feature commonly referred to as the term structure of volatility. The more we approach the maturity, the more the skew or smile behaviour accentuates. The existence of the volatility smile/surface casts a doubt over the Black-Scholes model and it is one of the why several other smile-consistent model have been developed. The non-constant volatility models can be parted in two categories:

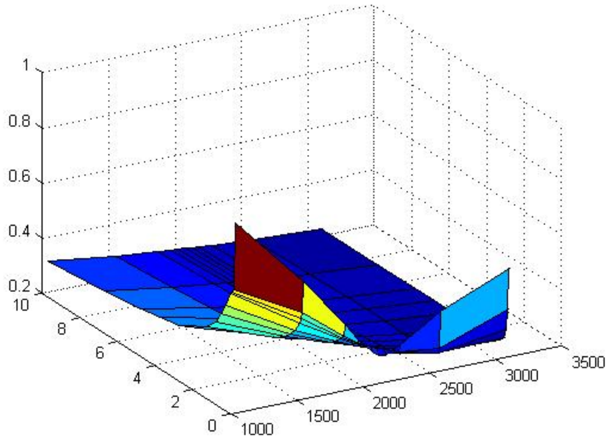


Figure 1.1: Volatility from Market Data (European call options quoted on the SX5E Index at 1st June 2012).

- The model with endogenous volatility, where the volatility is described by a stochastic process that depends on the same risky factor as the underlying. In this case the market is in general complete.

- The exogenous volatility models, where the volatility is described by a stochastic process driven by one or more than one additional risky factors. In this setting the market often is incomplete.

The most popular endogenous volatility models are the local volatility models, where the volatility is assumed to be a function of t and S_t, while in the second class, we can highlighted the stochastic volatility models.

1.6 Affine Models

In order to take into account the empirical evidence of a non constant volatility, several models have been proposed during the last twenty years ([20], [21], [27]). We will give a detailed description of these models in the next chapter. Here we would like to mention only an overview. Among the

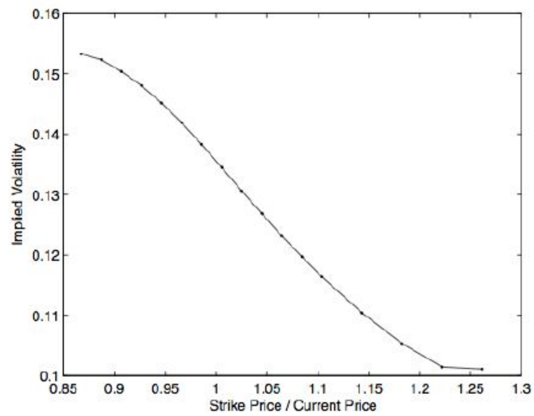


Figure 1.2: Implied volatility of SP500 options at 27-Nov 2005.

most popular models proposed, we recall the exponential Lévy model (which generalize the Black- Scholes model by introducing jumps) and stochastic volatility models.

Both these approaches relaxed the Black-Scholes hypothesis of a constant volatility. The class of affine processes includes the above mentioned examples. We introduce the affine processes as a class of time-homogeneous Markov processes.

Definition 1.6.1. A time-homogeneous Markov process with state space $(D, \mathcal{B}(D))$ is a family

$$(\Omega, (X_t)_{t\geq 0}, (\mathcal{F}_t)_{t\geq 0}, (p_t)_{t\geq 0}, (\mathbb{P})_{x\in D}), \tag{1.32}$$

where

- Ω is a probability space;
- $(X_t)_{t\geq 0}$ is a stochastic process taking values in D;
- $(\mathcal{F}_t)_{t\geq 0} = \sigma(X_s, s \leq t)$;
- $(p_t)_{t\geq 0}$ is a semigroup of transition functions on $(D, \mathcal{B}(D))$;
- $(\mathbb{P})_{x\in D}$ is a probability measures on $(\Omega, \vee_{t\geq 0}\mathcal{F}_t)$.

satisfying

$$\mathbb{E}[f(X_{t+s})|(\mathcal{F}_t)] = \mathbb{E}[f(X_s)], \tag{1.33}$$

$\mathbb{P}$-a.s., for all misurable functions f in D.

In fact, the affine processes are a class of time homogeneous Markov processes characterized by two additional properties. The first one being stochastic continuity, the second one is a condition which characterizes the Fourier-Laplace transform of the one-time marginal distributions. This introduction of affine processes in taken from [27], [21] and [51]. For $u = (v, w) \in \mathbb{C}^m \times \mathbb{C}^n$ we define the function $f_u \in C(D)$ by

$$f_u(x) := e^{\langle u,x\rangle} = e^{\langle v,y\rangle+\langle w,z\rangle}, \quad x = (y, z) \in D. \tag{1.34}$$

Notice that $f_u \in C_b(D)$ if and only if u lies in $U := \mathbb{C}^m_- \times i\mathbb{R}^n$,

Definition 1.6.2. Let

$$(W, (X_t)_{t\geq 0}, (\mathcal{F}_t)_{t\geq 0}, (p_t)_{t\geq 0}, (\mathbb{P})_{x\in D}) \tag{1.35}$$

be a time homogeneous Markov process. The process is said to be an affine process if it satisfies the following properties:

- for every $t \geq 0$ and $x \in D$, $\lim_{s\to t} p_s(x, \cdot) = p_t(x, \cdot)$, weakly,
- there exist functions $\varphi : \mathbb{R}_{\geq 0} \times U \to \mathbb{C}$ and $\Psi : \mathbb{R}_{\geq 0} \times U \to \mathbb{C}^d$ such that

$$\mathbb{E}[e^{\langle u,X_t\rangle}] = \int_D e^{\langle u,\epsilon\rangle} p_t(x, d\epsilon) = e^{\varphi(t,u)+\langle x,\Psi(t,u)\rangle} \tag{1.36}$$

for all $x \in D$ and $(t, u) \in \mathbb{R}_{\geq 0} \times U$.

The affine structure and Markov property give an appealing property for the functions φ and Ψ.

Proposition 1.6.1. The functions φ and Ψ satisfy the following semiflow property: for every $u \in U$ and $t, s \geq 0$

$$\begin{aligned} \Psi(t+s, u) &= \Psi(t, u) + \Psi(s, \varphi(t, u)) \quad \Psi(0, u) = 0, && (1.37)\\ \varphi(t+s, u) &= \varphi(s, \varphi(t, u)) \quad \varphi(0, u) = u. && (1.38) \end{aligned}$$

The functions φ and Ψ can be uniquely chosen so that they are jointly continuous on $\mathbb{R}_{\geq 0} \times U$. See Section 3 in [27] for the proof of the previous proposition. Regularity is also a key feature for an affine process. It gives differentiability of the Fourier-Laplace transform with respect to the time.

Definition 1.6.3. An affine process X is called regular if, for every $u \in U$, the partial derivatives

$$F(u) := \partial_t \Psi(t,u)|_{t=0} \quad R(u) := \partial_t \varphi(t,u)|_{t=0} \tag{1.39}$$

exist for all $u \in U$ and are continuous in

$$U^m = \{u \in \mathbb{C}^d | \sup_{x \in D} \operatorname{Re}(\langle u, x \rangle) \leq m\} \tag{1.40}$$

for all $m \geq 1$.

Theorem 1.6.1. Every affine process is regular. Moreover, on the set $R_{\geq 0} \times U$, the functions Ψ and φ satisfy the following system of generalized Riccati equations:

$$\partial_t \Psi(t,u) = F(\varphi(t,u)) \quad \Psi(0,u) = 0, \tag{1.41}$$

$$\partial_t \varphi(t,u) = R(\varphi(t,u)) \quad \varphi(0,u) = u. \tag{1.42}$$

with

$$\begin{aligned} F(u) &= \langle b, u \rangle + \frac{1}{2} \langle u, au \rangle - c \\ &\quad + \int_{D \setminus \{0\}} \left(e^{\langle u, \epsilon \rangle} - 1 - \langle \pi_j u, \pi_j h(\epsilon) \rangle \right) m(d\epsilon) \end{aligned} \tag{1.43}$$

$$\begin{aligned} R_k(u) &= \langle \beta_k, u \rangle + \frac{1}{2} \langle u, \alpha_k u \rangle - \gamma_k \\ &\quad + \int_{D \setminus \{0\}} \left(e^{\langle u, \epsilon \rangle} - 1 - \left\langle \pi_{j \cup \{k\}} u, \pi_{j \cup \{k\}} h(\epsilon) \right\rangle \right) M_k(d\epsilon) \end{aligned} \tag{1.44}$$

for $k = 1, \cdots, \cdots, d$. The set of parameters

$$(b, \beta, a, \alpha, c, \gamma, m, M) \tag{1.45}$$

is specified by

$$b, \beta_i \in \mathbb{R}^d \qquad i = 1, \cdots, d, \tag{1.46}$$

$$a, \alpha_i \in S^d_+ \qquad i = 1, \cdots, d, \tag{1.47}$$

$$c, \gamma_i \in \mathbb{R}_{\geq 0} \qquad i = 1, \cdots, d, \tag{1.48}$$

$$m, M_i \quad \text{are Levy measures for} i = 1, \cdots, d, \tag{1.49}$$

This set of parameters is called admissible for $D = \mathbb{R}^m_{\geq 0} \times \mathbb{R}^n$, if the conditions in the table below are satisfied with I and J defined as

$$I = \{1, \cdots, m\} \text{ and } J = \{m+1, \cdots, d\}. \tag{1.50}$$

The set of admissible parameters fully characterizes an affine process in D. The theorem just mentioned is referred to Theorem 6.4 in [21]

Example 1.6.1. [Cox-Ingersoll-Ross model (1985)] Consider the one dimensional diffusion $(X_t)_{t\geq 0}$ given by the solution of the following SDE:

$$\begin{cases} dX_t = (b - \beta X_t)dt + \sqrt{\alpha X_t} dW_t, \\ X_0 = x > 0, \end{cases} \tag{1.51}$$

where b, $\beta, \alpha > 0$ are positive coefficients satisfying the well known Feller's condition [34], $2b \geq \alpha$. Actually such a condition avoids the possibility of negative values for the process. The model is often used to describe the evolution of interest rates. It is a type of one-factor model (short-rate model) as it describes interest rate movements as driven by only one source of market risk. The model can be used in the valuation of interest rate derivatives. It was introduced in 1985 by Cox, Ingersoll and Ross [17] as an extension of the Vasicek model (1977). More generally, when the process is at a low level (close to zero), the standard deviation also becomes very small, which dampens the effect of the random shock on the rate. Consequently,

Diffusion
$a_{kl} = 0 \quad \text{for } k \in I or l \in J$
$\alpha_j = 0 \quad \forall j \in J$
$\{\alpha_i\}_{kl} = 0 \quad \text{if } k \in I \setminus \{i\} \text{ or } l \in I \setminus \{i\},$
Drift
$b \in D$
$(\beta_i)_k \geq 0 \quad \forall i \in I \text{ and } k \in I\{i\}$
$(\beta_j)_k = 0 \quad \forall j \in J, \quad k \in I$
Killing
$\gamma_j = 0 \quad \forall j \in J$
Jumps
$supp(m) \subseteq D$ and $\int_{D\setminus\{0\}} \big((\vert\pi_I \epsilon\vert + \vert\pi_J \epsilon\vert^2) \wedge 1\big) m(d\epsilon) < \infty$
$M_j = 0 \quad \forall j \in J$
$supp(M_i) \subseteq D \quad \forall i \in I$ and
$\int_{D\setminus\{0\}} \big((\vert\pi_{I\setminus\{i\}} \epsilon\vert + \vert\pi_{J\cup\{i\}} \epsilon\vert^2) \wedge 1\big) M_i(d\epsilon) < \infty$

Table 1.1: Set of conditions for admissible parameters.

when the process gets close to zero, its evolution becomes dominated by the drift factor, which pushes the process upwards (towards equilibrium). The CIR is an ergodic process, and possesses a stationary distribution. The same process is also used in the Heston model (1993) to model stochastic volatility. The process $(X_t)_t$ is also an example of affine process in $\mathbb{R}_{\geq 0}$. Its Fourier-Laplace transform can be explicitly computed as the solution of the Riccati ordinary differential equation:

$$\begin{cases} \partial_t \Psi(t,u) = -\beta \Psi(t,u) + \frac{1}{2}\alpha \Psi^2(t,u) \\ \Psi(0,u) = u \end{cases} \tag{1.52}$$

and

$$\varphi(t,u) = b \int_0^t \Psi(s,u) ds.$$

Chapter 2

Volatility Models

As mentioned in the previous chapter, several models have been proposed during the last twenty years, developing and generalizing the Black-Scholes framework. As previously exposed in the first chapter the most popular models proposed to overcome such evidences are the exponential Levy model (which generalize the Black- Scholes model by introducing jumps), Local Volatility Models, the Stochastic Volatility model and Jump Diffusion Models. In particular we recall here the last three main approaches:

- Local Volatility Models (LVM).
  Introduced for the first time in 1994 by Dupire [29] and Derman and Kani [22]these models assume that the diffusion coefficient of the underlying asset is no longer a constant value but instead a deterministic function of time and of the underlying asset itself: $\sigma = \sigma_{LV}(s;t)$

$$dS_t = rS_t dt + \sigma_{LV}(S_t t) S_t dW_t \tag{2.1}$$

- Stochastic Volatility Models (SVM).
  In this class of models the volatility itself is considered to be a stochastic process with its own dynamics. Thus, this is a two-factor model, driven by two correlated one-dimensional Brownian motions, see equation 2.131

- Jump Diffusion Models (JDM).

Introduced by Merton these models considers the underlying asset to follow a Levy process with a drift, a diffusion and a jump term;

$$dS_t = rS_t dt + \sigma S_t dW_t + S_t dJ_t, \tag{2.2}$$

Of course, all these three kinds of models have some advantages and disadvantages. In particular in the last ten years the first two models have been widely studied in academic literature as well as used at the equity trading desks of investment banks.

2.0.1 Local Volatility Model

Local Volatility models, introduced for the first time in 1994 by Dupire and Derman and Kani [29], [22], assume that the diffusion coefficient of underlying is no longer a constant value but instead a deterministic function of time and of the underlying asset itself: $\sigma = \sigma_{LV}(s,t)$.

$$dS_t = rS_t dt + \sigma_{LV}(S_t,t) S_t dW_t \tag{2.3}$$

The Local Volatility Models are not so far from the original Black and Scholes model, mainly for the reason that they are a straightforward generalization of the original Black-Scholes framework. The principal advantage is the possibility of a (nearly) perfect fit to the quoted market price.(if we had a continuum of traded vanilla prices for each strike and maturity). Infact, in this case a clear representattion could be possible.

Dupire Equation

The Black-Scholes backward parabolic equation in the variables (s,t) is the Feynman-Kac representation of the discounted expected value of the final option value. It is possible to find the same option price solving a "'dual problem"', namely a forward parabolic equation in the variables (K,T) known as dual Black-Scholes equation or Dupire's equation ([29], [22]).

Proposition 2.0.2. [Dupire's equation] The value of a call option as a function of the strike price K and the time to maturity T given the present value of the stock S, with a continuous dividend rate $d \geq 0$, and short-term rate $r \geq 0$, is given by the following forward parabolic equation known as Dupire's equation:

$$\begin{cases} \frac{\partial C}{\partial T} - \frac{1}{2}\sigma_{LV}(K,T)K^2\frac{\partial^2 C}{\partial K^2} + (r-d)K\frac{\partial C}{\partial K} - dC = 0, \quad \text{on } (0,\infty)\times(0,\infty), \\ C(0;T) = S \quad \forall\, T \in (0,\infty), \\ \lim_{K\to\infty} C(K;T) = 0 \quad \forall\, T \in (0,\infty), \\ C(K;0) = (S-K)_+ \quad \forall\, K \in (0,\infty). \end{cases}$$

Proof. Let's consider the transition probability density function $p(S,T;S_0,0)$ of S_T, given the initial value $S_0 > 0$, at time $t = 0$, related to the following risk-neutral dynamics:

$$dS_t = (r-d)S_t dt + \sigma_{LV}(S_t,t)S_t dW_t \tag{2.4}$$

with the initial value S_0. The diffusion coefficient of the underlying is assumed to be a deterministic function of time and of the underlying asset itself, $\sigma = \sigma_{LV}(s;t)$, and W_t is a standard 1-dimensional Brownian motion. The value of an option is the discounted value of the expected pay-off, for a Call option it means that:

$$\begin{aligned} C(K,T) &= e^{-rT}\int_0^{+\infty}(S-K)_+ p(S,T,S_0,0)dS \\ &= e^{-rT}\int_K^{+\infty}(S-K)p(S,T,S_0,0)dS, \end{aligned}$$

for any $T,\ K > 0$. Deriving the above expression twice with respect to the strike price, we get

$$\frac{\partial C}{\partial K} = e^{-rT}\int_K^{+\infty} p(S,T;S_0 0)dS, \tag{2.5}$$

$$\frac{\partial^2 C}{\partial K^2} = e^{-rT}p(K,T;S_0,0). \tag{2.6}$$

The transition density function of the underlying stock process corresponds to the second order derivative of the call price with respect to the strike (originally due to Breeden and Litzenberg [9]):

$$p(K,T;S_0,0) = e^{rT}\frac{\partial^2 C}{\partial K^2} \tag{2.7}$$

The probability density function solves the Fokker-Planck equation:

$$\frac{\partial p}{\partial T} = \frac{1}{2}\frac{\partial^2}{\partial S^2}[\sigma_{LV}^2(S,T)S^2 p(S,T)] - (r-d)\frac{\partial}{\partial S}[Sp(S,T)]. \tag{2.8}$$

for all $S > 0$, $T > 0$. Here, for the sake of simplicity, we have omitted the explicit dependence of p with respect to S_0 and the initial time. Evaluating the time derivative of relation (2.5), we get

$$\frac{\partial C}{\partial T} = -rC + e^{-rT}\int_K^{+\infty}(S-K)\frac{\partial p}{\partial T}dS, \tag{2.9}$$

and by substituting the time derivative of the probability density function inside the integral as in equation (2.8), the above equation can can be re-written as:

$$\begin{aligned}\frac{\partial C}{\partial T} + rC \quad &= \quad e^{-rT}\int_K^{+\infty}(S-K)\big(\frac{1}{2}\frac{\partial^2}{\partial S^2}[\sigma_{LV}^2(S,T)^2 S^2 p] + \\ &\qquad -(r-d)\frac{\partial}{\partial S}[Sp]\big)dS. \end{aligned} \tag{2.10}$$

By integrating the righthand side by parts twice and leveraging the fact that $p(S,T)$ decays exponentially fast once $S \to \infty$, we deduce:

$$\begin{aligned}\frac{\partial C}{\partial T} \quad &+ \quad rC = e^{-rT}\frac{1}{2}\big([(S-K)\frac{\partial}{\partial S}[\sigma_{LV}^2(S,T)S^2 p(S,T)]]_{S=K}^{S=+\infty} \qquad (2.11)\\ &- \quad \int_K^{+\infty}\frac{\partial}{\partial S}[\sigma_{LV}^2(S,T)^2 S^2 p(S,T)]\big)\\ &- \quad e^{-rT}(r-d)\left([(S-K)Sp(S,T)]]_{S=K}^{S=+\infty} - \int_K^{+\infty} Sp(S,T)dS\right)\\ &= \quad \frac{1}{2}e^{-rT}\sigma_{LV}^2(K,T)^2K^2p(K,T) + (r-d)\left(C + e^{-rt}K\int_K^{+\infty} p(S,T)dS\right).\end{aligned}$$

Using the expression (2.5) and (2.6) in the first and second term of the right hand side respectively we obtain Dupire's equation:

$$\frac{\partial C}{\partial T} = \frac{1}{2}\sigma_{LV}^2(K,T)K^2\frac{\partial^2 C}{\partial K^2} - (r-d)K\frac{\partial C}{\partial K} - dC, \tag{2.12}$$

for all $K > 0$, $T > 0$. □

From Dupire's equation, it is possible to derive a formula to evaluate the local volatility function $\sigma_{LV}(\cdot;\cdot)$ from option prices. This is known as Dupire's formula:

$$\sigma^2_{LV}(K,T) = 2\frac{\partial_T C + (r-d)K\partial_K C + dC}{K^2\partial^2_K C}. \tag{2.13}$$

This seems to be straightforward, if we could observe on the market a continuum of plain vanilla prices. Unfortunately, only some options could be derived from direct observation of market data and the residual ones have to be interpolated and extrapolated starting from these. As it could be easily inferred this entails a computational error (which becomes bigger in the t the denominator) and let the Dupire's formula to be not so useful to put in practice. It is possible to reformulate it in a more suitable form: using the Black-Scholes implied volatility $\sigma(K,t)$ instead of option prices. The idea [35] and [38] is to insert the Black-Scholes formula and its derivative into Dupire's formula. Starting from

$$C = C_{BS}(S_t, t, K, t, \sigma_I(K,t)), \tag{2.14}$$

and applying the rule of differentiation, it can be obtained that

$$\sigma^2_{LV}(K,t) = \frac{\sigma^2_I + 2t\sigma_I(\frac{\partial\sigma_I}{\partial t} + (r-d)K\frac{\partial\sigma_I}{\partial K})}{(1 + d_1 K\sqrt{t}\frac{\partial\sigma_I}{\partial K})^2 + K^2\sigma_I t(\frac{\partial^2\sigma_I}{\partial K^2} - d_1\sqrt{t}(\frac{\partial\sigma_I}{\partial K})^2)} \tag{2.15}$$

with

$$d_1 = \frac{\log\frac{S_0}{K} + (r - d + \frac{1}{2}\sigma^2_I)t}{\sigma_I\sqrt{t}} \tag{2.16}$$

Nevertheless, there is still the problem that implied volatility is not known as a continuous function of strike prices and maturities and in 2004 Gatheral proposed an interesting parametrization of the implied volatility surface called Stochastic Volatility Inspired (SVI) given by the following function:

$$\sigma^{SVI}_I(x,t_n) = \sqrt{C^n_1 + C^n_2[C^n_3(x - C^n_4) + \sqrt{(x-C^n_4)^2 + C^n_5}}, \tag{2.17}$$

where the coefficients mentioned above are calibrated form the market data. Thus, denoting by T_n (for $n = 1\cdots N$), the vector of the N available market maturities, it is possible to evaluate N different SVI functions $\sigma_I(x,t_n)$, for $n = 1 \cdots N$, and then we define for $t \in (t_n, t_{n+1})$:

$$\begin{aligned} \sigma_I^{SVI}(x,t) &= \sigma_I^{SVI}(x,t_n) + \frac{t-t_n}{t_{n+1}-t_n}[\sigma_I^{SVI}(x,t_{n+1}) - \sigma_I^{SVI}(x,t_n)] \\ &= \frac{t_{n+1}-t}{t_{n+1}-t_n}\sigma_I^{SVI}(x,t_n) + \frac{t-t_n}{t_{n+1}-t_n}\sigma_I^{SVI}(x,t_{n+1}). \end{aligned} \quad (2.18)$$

The model assumes the variable x to be the logarithmic forward moneyness $x = \log(K/F_t)$ The five parameters for each maturity:

- $C_1^n \geq 0$ provides the constant component of volatility.Increasing. C_1^n implies a vertical translation of the smile;
- $C_2^n \geq 0$ influences the slope of the volatility in its wings. Increasing C_2^n increases the slopes of both the left and right wings, tightening the smile;
- $C_3^n \in (-1,1)$ provides a counter-clockwise rotation of the smile. Increasing C_3^n decreases (increases) the slope of the left(right) wing;
- C_4^n provides an horizontal translation. Increasing C_4^n translates the smile to the right;
- $C_5^n > 0$ influences the smile curvature. Increasing C_5^n reduces the at-the-money curvature of the smile.

The calibration of coefficients $C_1^n, C_2^n, C_3^n, C_4^n, C_5^n$ which are obviously time-dependent is implemented by solving a least-squares constrained optmization problem, by finding those coefficients such that the implied volatilities of the model are as closest as possible to the ones actually quoted on the market. The idea is to calibrate the parameters starting from the closest maturities till the farthest ones. Let $x^n = [C_1^n, C_2^n, C_3^n, C_4^n, C_5^n]$, then we have to solve a sequence of minimization problems:

$$(M)\begin{cases} \text{for every } n = 1, \cdots, N \text{ find } \ x^n \in W \subset \mathbb{R}^5 \ \text{ such that :} \\ J_n(x_n) = \min_{y \in W} \ J_n(y) \end{cases}$$

where the cost functional is

$$J_n(y) = \sum_{j}^{M} \omega_j \left(\frac{\sigma_I^{\mathrm{SVI}}(K_j, t_n; y) - \sigma_I^{\mathrm{Market}}(K_j, t_n)}{\sigma_I^{\mathrm{Market}}(K_j, t_n)} \right)^2, \tag{2.19}$$

where K_j, $j = 1, \cdots, M$ is the vector of market strike prices, t_n $n = 1, \cdot, N$ is the vector of maturities, $\sigma_I^{\mathrm{Market}}(K_j, t_n)$ is the implied volatility observed on the market for a contract with strike price K_j and maturity t_n; $\sigma_I^{\mathrm{SV}}(K_j, t_n; y)$ is the implied volatility provided by the SVI model for a contract with strike price K_j and maturity t_n, given the vector of coefficients y; ω_j are weights chosen to fit better the central region of the smile. The input data for the minimization problem (M) are shown in table below. Starting from the first maturity, and using the initial guess conditions, the cost functional (2.19) is minimized with a Lsq algorithm. The method minimizes the vector-valued function, which will be the functional cost, using the vector of initial parameter values, x_0, where the lower and upper bounds of the parameters will be specified. We will use the Matlab code function *lsqnonlin*, which uses an interior-reflective Newton method for large scale problems. Matlab defines a large scale problem as one containing bounded/unbounded parameters, where the system is not under-determined, i.e., where the number of equations to solve is more than the required parameters. Given that the underlying of the model is liquidly traded, there should exist a rich set of market prices for calibration. So, hopefully, our system will never be under-determined. The result produced by *lsqnonlin* is dependent on the choice of x_0, the initial estimate. This is, hence, not a global optimizer, but rather, a local one. We have no way of knowing whether the solution is a global/local minimum.

Once the calibration of the coefficients is accomplished for the first maturity, we use these values as new first guess conditions for the next calibration of the coefficients of the second maturity and so on till the last expiration. This procedure is quite natural since consecutive smiles are rather similar and therefore the fitted coefficients would be quite close as well.

Dupire Coefficients	Initial guess	Lower bound	Upper bound
C_1	0.001	0.0001	10
C_2	0.6	0.0001	10
C_3	-0.3	-1	1
C_4	0.05	-10	10
C_5	0.05	0.0001	10

Once calibrating all the coefficients (for each maturity) we can calculate the surface of Implied Volatility. Figure below shows the whole surface of calibrated SVI Implied Volatility $\sigma_I^{SV}(K;T)$ as a function of strike prices and maturities.

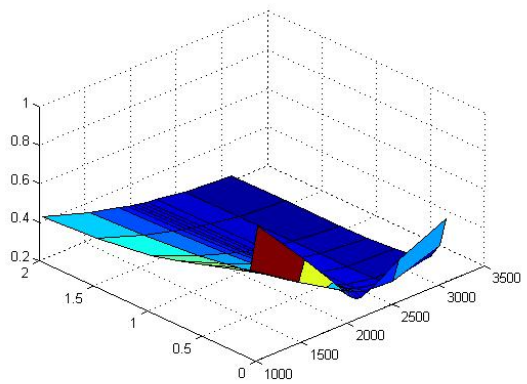


Figure 2.1: Stochastic Volatility Inspired (SVI)

The use of the Dupire's formula combined with the SVI parametrization is one of the most direct, fast, stable and reliable method to reconstruct the Local Volatility surface from marked data.

2.1 Stochastic Volatility Model

The stochastic volatility model has been proposed as a description of data from financial markets by Clark (1973), Tauchen and Pitts (1983), Taylor (1986, 1994), and others. The appeal of the model is that it provides a simple specification for speculative price movements that accounts, in qualitative terms, for broad general features of data from financial markets such as leptokurtosis and persistent volatility. Also, it is related to diffusion processes used in derivatives pricing theory in finance; see Mathieu and Schotman (1994) and references therein. The standard form as set forth, for instance, in Harvey, Ruiz, and Shephard (1994), Jacquier, Polson, and Rossi (1994), and Danielsson (1994), takes the form of an autoregression whose innovations are scaled by an unobservable volatility process, usually distributed as a lognormal autoregression. Unlike the previous section, these kind of models consider the volatility itself as a random process with its own dynamics. The Stochastic Volatility Models, have specular properties to the Local ones: in fact they provide a good smile dynamics over the time, but a bad fit of the present market prices. So, the former models are able to be well calibrated to traded vanilla options, the latter can reproduce a more realistic dynamics of implied volatility. The main feature of a stochastic volatility model (SV) is to consider the volatility itself as a stochastic process. Then, while the standard Black-Scholes model assumes a constant volatility term σ a Stochastic Volatility Model considers volatility as a function of a process S_t. Thus, the model derived is a two-factor model, driven by two correlated Brownian motions on the same filtered probability space, namely W_t and Z_t, as

$$\begin{cases} dS_t = \mu S_t d_t + b(V_t) S_t dW_t, \\ dV_t = a(V_t) d_t + c(V_t) dZ_t, \\ dW_t dZ_t = \rho dt, \end{cases} \tag{2.20}$$

where μ is a constant coefficients which represents the instantaneous rate of return of the stock, $\rho \in [-1, 1]$ is a constant correlation parameter,

a, b, c : $[0, \infty) \to [0, \infty)$ are deterministic and continuous functions, satisfying suitable regularity assumptions such that system (2.20) admits a unique strong solution, for any initial datum $S_0 \geq 0$, $V_0 \geq 0$, satisfying $V_t \geq 0$, for all $t \geq 0$, with probability 1. A SV model assumes two sources of randomness W_t, Z_t and only one traded asset S_t depending on both these sources. In this case, we cannot hedge the risk and the model is said to be incomplete. The concept of completeness of the model is strictly related to the Girsanov theorem and the existence of an equivalent martingale measure (as mentioned in the previous chapter). Actually, if the model is complete then it exists only one equivalent measure and the price of every derivative is uniquely determined. On the other hand, if the model used is incomplete, there exist many different martingale measures. Estimation of the stochastic volatility model presents intriguing challenges, and a variety of procedures have been proposed for fitting the model. Extant methods include method of moments (Dupire and Singleton, 1993; Andersen and Sorensen, 1996); Bayesian methods (Jacquier, Polson, and Rossi, 1994; Geweke, 1994), simulated likelihood (Danielsson, 1994), and Kalman Filtering methods (Harvey, Ruiz, and Shephard, 1994; Kim and Shephard, 1994).

2.1.1 The Heston model

One of the most important stochastic volatility model is the Heston model, introduced in 1993 and nowadays it is probably the most popular stochastic volatility model. Several other models have been derived from Heston model, including also extension with jumps. In the Heston model we have, in line with the notation used above, $b(v) = \sqrt{v}$, (the square of the volatility, the variance V_t, is a Cox-Ingersoll-Ross (CIR) process), $a(v) = \kappa(v - \theta)$ and $c(v) = \sigma\sqrt{v}$. The dynamics is given by the following system:

$$\begin{cases} dS_t = \mu S_t dt + \sqrt{V_t} S_t dW_t, \\ dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dZ_t, \\ dW_t dZ_t = \rho dt, \end{cases} \tag{2.21}$$

The Heston model is characterized by five constant parameters, namely κ, θ, σ, ρ and the initial value of the variance v_0.

- $k > 0$ is the rate of mean reversion. The mean reversion rate can be interpreted as representing the degree of volatility clustering: it means that large moves are followed by large moves, while small moves are more likely to be followed by small moves. The mean reversion parameter controls the curvature of the curve. Increasing the mean reversion parameter flattens the implied volatility smile, see Figure 2.1.1. Decreasing the mean reversion has a similar effect as increasing the volatility of variance in terms of curvature;
- $\theta > 0$ the long term mean;
- $\sigma > 0$ the volatility of volatility;
- $\rho \in [-1, 1]$ si the correlation, which can be interpreted as the correlation between the log-returns and volatility of the asset, affects the heaviness of the tails. Intuitively, if $\rho > 0$, then volatility will increase as the asset price/return increases. This will spread the right tail and squeeze the left tail of the distribution creating a fat right-tailed distribution. Conversely, if $\rho < 0$, then volatility will increase when the asset price/return decreases, thus spreading the left tail and squeezing the right tail of the distribution creating a fat left-tailed distribution (emphasising the fact that equity returns and its related volatility are negative correlated). The correlation, therefore, affects the skewness of the distribution.

In order to guarantee that the volatility process is strictly positive, the parameters $k > 0$, $\theta > 0$ and $\sigma > 0$, must also verify a fundamental con-

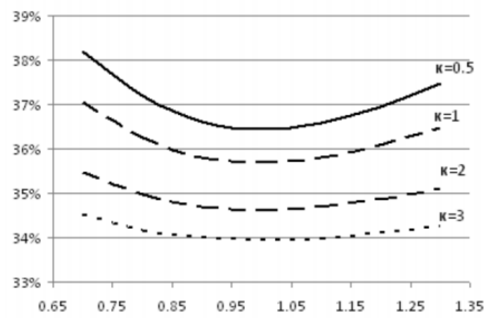


Figure 2.2: The effect of changing the mean reversion κ. Source [39]

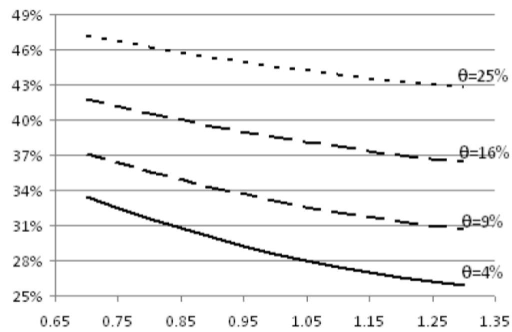
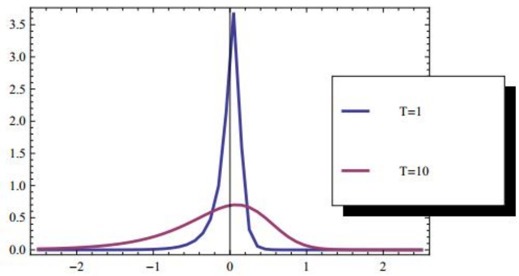


Figure 2.3: The effect of changing the long run variance θ. Source [39]

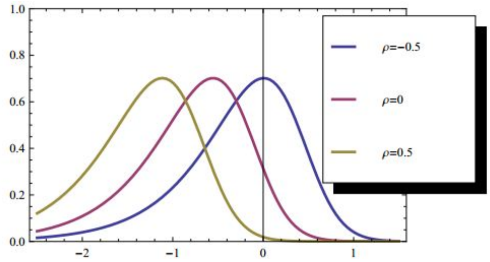
straint, the so-called Feller condition [34]:

$$Fe = \kappa\theta - \frac{1}{2}\sigma^2 \geq 0. \tag{2.22}$$

If this relation is satisfied then, the process V_t is strictly positive: it cannot reach the zero because the drift term pushes it away when it becomes too small. However the reason that makes this model so popular and used is probably the fact that it has a semi-closed form solution for plain vanilla options. This enables a fast and computational efficient valuation of European options which becomes of critical importance when calibrating the model to known option prices. One of the main advantages of the Heston is that the characteristic function for (x_T, V_T) is known in an explicit form, where x_T is the logarithm of the terminal stock price, $x_T = \log(S_T)$. Exploiting the properties of the characteristic function, as it is a the Fourier transform, one can get an approximation of the density function of x_T. Here below the plot of the density function is given for two different maturities. The parameters used are $\kappa = 1.5768$, $\sigma = 0.5751$, $\theta = 0.0398$, $\mu = 0.0175$, $\rho = -0.5711$.



Here below the plot for the density function for different value of correlation. The values used are $r = 0, \kappa = 1.5768, \sigma = 0.5751, \eta = 0.0398, u_0 = 0.0175, T = 10$



2.1.2 A closed-form solution

The main reason for the popularity of the Heston model among stochastich volatility models is that it provides a closed-form solution for pricing vanilla options. This is of great benefit in particular when calibrating against market prices. We also remark that, under the change of measure to a risk-neutral measure $\mathbb{Q}$, the equation in the Heston model keep a shape similar to those of system (2.21), where μ is raplaced by the risk-free interest rate

$r > 0$. Precisely:

$$\begin{cases} dS_t = rS_t dt + \sqrt{V_t} S_t dW'_t, \\ dV_t = \kappa'(\theta' - V_t)dt + \sigma\sqrt{V_t} dZ'_t, \end{cases} \tag{2.23}$$

where $(W'_t)_t$, $(Z'_t)_t$ are Brownian motions under $\mathbb{Q}$. The new parameters are defined as follows:

$$\kappa' = \kappa - \sigma\lambda, \qquad \theta' = \frac{\kappa\theta}{\kappa - \sigma\lambda}, \tag{2.24}$$

where the constant parameter $\lambda \in (0, \kappa/\sigma)$ is used to define the so called market price for the volality risk, $\lambda\sqrt{V_t}$. In particular $(\mu - r)/\sqrt{V_t}$ represents the market price premium for the stock S. Since our focus is mainly on the pricing problem, where the risk-neutral dynamics is considered, with an abuse of notation, in the following we will continue to denote with κ and θ the risk-neutral parameter and with W_t, Z_t the Brownian motions representing the source of randomness. We also observe that Feller's condition (2.22) is still satisfied by considering coefficients κ', θ'.

Then, the no arbitrage price, at time t, of the European call with strike K and maturity T is the expected discounted value under the risk-neutral measure $\mathbb{Q}$, namely:

$$\begin{aligned} C_t &= e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}[(S_T - K)^+] \\ &= e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}[S_T \mathbb{1}_{S_T > K}] - e^{-r(T-t)} K \mathbb{E}_t^{\mathbb{Q}}[\mathbb{1}_{S_T > K}], \end{aligned} \tag{2.25}$$

where $\mathbb{E}_t^{\mathbb{Q}}[\cdot]$ denotes the conditional expected value, given $\mathcal{F}_t$. By analogy with the Black-Scholes formula, the guessed solution of this European option is of the form $C_t = C(t, x_t, v_t)$, where the deterministic function C takes the form

$$\begin{cases} C(t,x,v) = e^x P_1(\tau,x,v) - e^{-r\tau} K P_2(\tau,x,v), \\ P_j = \frac{1}{2} + I_j, \\ I_j = \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi \log(K)} f_j(\phi;x,v)}{i\phi}\right] \mathrm{d}\phi, \end{cases} \tag{2.26}$$

where $\tau = T - t$ is the time to expiration and

$$f_j(\phi; x, v) = \exp\left(C_j(\tau,\phi) + D_j(\tau,\phi)v + i\phi x\right), \tag{2.27}$$

$$D_j = \tfrac{b_j - \rho\sigma i\phi + d_j}{\sigma^2}\left(\tfrac{1-e^{d_j\tau}}{1-g_j e^{d_j\tau}}\right), \tag{2.28}$$

$$C_j = ri\phi\tau + \tfrac{a}{\sigma^2}\left[(b_j - \rho\sigma i\phi + d_j)\tau - 2\log\left(\tfrac{1-g_j e^{d_j\tau}}{1-g_j},\right)\right], \tag{2.29}$$

$$d_j = \sqrt{(b_j - \rho\sigma i\phi)^2 - \sigma^2(2u_j i\phi - \phi^2)}, \tag{2.30}$$

$$g_j = \tfrac{b_j - \rho\sigma i\phi + d_j}{b_j - \rho\sigma i\phi - d_j},$$

where $u_1 = \frac{1}{2}$, $u_2 = -\frac{1}{2}$, $a = k\eta$, $b_1 = k - \rho\sigma$, $b_2 = k$. Despite of such formula, in the Heston model the diffusion coefficient σ is not Lipschitz-continuous up to the boundary. Results obtained in Costantini et al. (2012) for a general class of princing problems is based on reflecting diffusion processes with jumps (including also the Heston model), shows that, in presence of the above features, existence, uniqueness or regularity of solutions of the pricing problem may very well not hold. Even though there is extensive literature on equations with degenerating coefficients, compare the classical reference ([59]), the C^1-regularity of the solution at the boundary is not available in the generality that is needed here. We observe that, as a function of $\tau = T - t$, the function (2.25) satisfies the pricing equation:

$$\begin{aligned}\frac{\partial C}{\partial \tau} &- \frac{1}{2}v\frac{\partial^2 C}{\partial x^2} - (r - \frac{1}{2}v)\frac{\partial C}{\partial x} - \rho\sigma v\frac{\partial^2 C}{\partial x \partial v} + \\ &- \frac{1}{2}\sigma^2 v\frac{\partial^2 C}{\partial v^2} + rC - \kappa(\theta - v)\frac{\partial C}{\partial v} = 0,\end{aligned} \tag{2.31}$$

for any $\tau \in (0, T)$, $x \in \mathbb{R}$. Thus, following [30] and [31], we prove that the solution (2.26) satisfies the boundary condition at $v = 0$ in the classical sense, see Proposition 2.1.1 below.

Proposition 2.1.1. The solution function 2.26 to the pricing problem for a Call option under the Heston model satisfies the boundary condition

$$\frac{\partial C}{\partial \tau} - r\frac{\partial C}{\partial x} - k(\theta - v)\frac{\partial C}{\partial v} + rC = 0. \tag{2.32}$$

along the boundary $v = 0$.

In order to prove the above proposition, we make use of the following results concerning with the asymptotic behavior of the coefficients appearing in the Heston formula (2.26)-(2.27). Actually, for our knowledge the proof of these technical results are not presented in the available literature.

Lemma 2.1.1. Assuming that $\kappa, \theta, \sigma, \tau > 0$ and $\rho \in (-1, 1)$, we obtain the following asymptotics:

$$\lim_{\phi\to\infty} \frac{d_j(\phi)}{\phi} = \sigma\sqrt{1-\rho^2}, \tag{2.33}$$

$$\lim_{\phi\to\infty} g_j(\phi) = -1 + 2\rho^2 + 2i\rho\sqrt{1-\rho^2} = (i\sqrt{1-\rho^2} + \rho)^2, \tag{2.34}$$

$$\lim_{\phi\to\infty} \frac{D_j(\phi)}{\phi} = -\frac{\sqrt{1-\rho^2} + i\rho}{\sigma}, \tag{2.35}$$

$$\lim_{\phi\to\infty} \frac{C_j(\phi)}{\phi} = -i\frac{\kappa\theta^2}{\sigma^2}\rho\tau - \frac{\kappa\theta}{\sigma^2}\sigma\sqrt{1-\rho^2}, \tag{2.36}$$

for $j = 1, 2$.

Lemma 2.1.2. For every $x \in \mathbb{R}$, $v \geq 0, \tau > 0$, κ, θ, r, $K > 0$ and $\rho \in (-1, 1)$, the integrand in I_j (2.26),

$$\mathrm{Re}\left[\frac{e^{-i\phi \log(K)} f_j(\phi; x, v)}{i\phi}\right], \tag{2.37}$$

is bounded as $\phi \to 0^+$ and of exponential decay as $\phi \to \infty$, for $j = 1, 2$.

Proof. (*Proposition 2.1.1*) We evaluate the partial derivatives of the price function. Thanks to Lemma 2.1.2 the passage of the derivative into the integral is admissible:

$$\begin{aligned}\frac{\partial C}{\partial \tau} &= e^x \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)} f_1(\phi; x, v)}{i\phi}\left(\frac{\partial C_1}{\partial \tau} + v\frac{\partial D_1}{\partial \tau}\right)\right] \mathrm{d}\phi \\ &\quad + re^{-r\tau}KP_2 + e^{-r\tau}K\frac{1}{\pi} \times \\ &\quad \times \int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)} f_2(\phi; x, v)}{i\phi}\left(\frac{\partial C_2}{\partial \tau} + v\frac{\partial D_2}{\partial \tau}\right)\right] \mathrm{d}\phi,\end{aligned} \tag{2.38}$$

being the derivative for C_j and D_j the following:

$$\frac{\partial C_j}{\partial \tau} = ri\phi + aD_j, \tag{2.39}$$

$$\frac{\partial D_J}{\partial \tau} = \rho\sigma i\phi D_j - \frac{1}{2}\phi^2 + \frac{1}{2}\sigma^2 D_j^2 + u_j i\phi - b_j D_j, \tag{2.40}$$

for $j = 1$, 2. Collecting all, we have

$$\begin{aligned}
\frac{\partial C}{\partial \tau} &= e^x \frac{1}{\pi} \int_0^\infty \mathrm{Re} \left(\frac{e^{-i\phi \log(K)} f_1(\phi; x, v)}{i\phi} (ri\phi + aD_1 + \right. \\
& \left. + v\rho\sigma i\phi D_1 - \frac{v}{2}\phi^2 + \frac{v}{2}\sigma^2 D_1^2 + vu_1 i\phi - vb_1 D_1) \right) \mathrm{d}\phi + \\
&+ Kre^{-r\tau} P_2 + Kre^{-r\tau} \frac{1}{\pi} \int_0^\infty \mathrm{Re} \left(\frac{e^{-i\phi \log(K)} f_2(\phi; x, v)}{i\phi} (ri\phi + aD_2 + \right. \\
& \left. v\rho\sigma i\phi D_2 - \frac{v}{2}\phi^2 + \frac{v}{2}\sigma^2 D_2^2 + vu_2 i\phi - vb_2 D_2) \right) \mathrm{d}\phi. \qquad (2.41)
\end{aligned}$$

For other partial derivatives, it holds:

$$\begin{aligned}
\frac{\partial C}{\partial x} &= e^x P_1 + e^x \frac{1}{\pi} \int_0^\infty \mathrm{Re} \left[e^{-i\phi \log(K)} f_1(\phi; x, v) \right] \mathrm{d}\phi + \\
& -e^{-r\tau} K \frac{1}{\pi} \int_0^\infty \mathrm{Re} \left[e^{-i\phi \log(K)} f_2(\phi; x, v) \right] \mathrm{d}\phi, \qquad (2.42)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial C}{\partial v} &= e^x \frac{1}{\pi} \int_0^\infty \mathrm{Re} \left[\frac{e^{-i\phi \log(K)}}{i\phi} f_1(\phi; x, v) D_1 \right] \mathrm{d}\phi + \\
& -e^{-r\tau} K \frac{1}{\pi} \int_0^\infty \mathrm{Re} \left[\frac{e^{-i\phi \log(K)}}{i\phi} f_2(\phi; x, v) D_2 \right] \mathrm{d}\phi. \qquad (2.43)
\end{aligned}$$

By using all the previous expressions for the partial derivatives of C, we get:

$$\frac{\partial C}{\partial \tau} - r\frac{\partial C}{\partial x} - \kappa(\theta - v)\frac{\partial C}{\partial v} + rC = \tag{2.44}$$

$$\frac{e^x}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)}f_1(\phi;x,v)}{i\phi}(ri\phi + aD_1 + v\rho\sigma i\phi D_1 - \frac{v}{2}\phi^2 + \right.$$
$$\frac{v}{2}\sigma^2 D_1^2 + vu_1 i\phi - vb_1 D_1)]\mathrm{d}\phi + Kre^{-r\tau}P_2 + \frac{Kre^{-r\tau}}{\pi}\times$$
$$\times\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)}f_2(\phi;x,v)}{i\phi}(ri\phi + aD_2 + v\rho\sigma i\phi D_2 - \frac{v}{2}\phi^2 + \frac{v}{2}\sigma^2 D_2^2 + \right.$$
$$+vu_2 i\phi - vb_2 D_2)]\,\mathrm{d}\phi - \frac{re^x}{2} - re^x\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)}}{i\phi}f_1(\phi;x,v)\right]\mathrm{d}\phi$$
$$-re^x\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[e^{-i\phi\log(K)}f_1(\phi;x,v)D_2\right]\mathrm{d}\phi$$
$$+rKe^{-r\tau}\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)}}{i\phi}f_2(\phi;x,v)\right]\mathrm{d}\phi$$
$$-k(\eta - v)e^x\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)}}{i\phi}f_1(\phi;x,v)D_1\right]\mathrm{d}\phi$$
$$+k(\eta - v)Ke^{-r\tau}\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)}}{i\phi}f_2(\phi;x,v)D_2\right]\mathrm{d}\phi$$
$$+\frac{re^x}{2} + re^x\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)}}{i\phi}f_1(\phi;x,v)\right]\mathrm{d}\phi - \frac{re^{-r\tau}K}{2}$$
$$-re^{-r\tau}K\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)}}{i\phi}f_2(\phi;x,v)\right]\mathrm{d}\phi, \tag{2.45}$$

and deleting similar terms, (2.44) reduces to

$$\frac{e^x}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)}f_1(\phi;x,v)}{i\phi}(v\rho\sigma D_1 - \frac{v\phi}{2i} + \frac{v\sigma^2 D_1^2}{2i\phi} + \right.$$
$$+\frac{v}{2} + \frac{v\rho\sigma D_1}{i\phi})]\mathrm{d}\phi +$$
$$-\frac{Kre^{-r\tau}}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi\log(K)}f_2(\phi;x,v)}{i\phi}(v\rho\sigma D_2 - \frac{v\phi}{2i} + \right.$$
$$\left.\frac{v\sigma^2 D_2^2}{2i\phi} - \frac{v}{2})\right]\mathrm{d}\phi \longrightarrow 0, \qquad v \to 0^+. \tag{2.46}$$

The convergence of the expression above is leveraged by using Lemma 2.1.2. Coming from this lemma, we can state that for $\phi \to 0^+$ the second part of integrand is bounded, as a function of ϕ, as well as the first one and so the integral is convergent. Similarly, as $\phi \to \infty$, the asymptotic behavior of the coefficients $D_j(\phi)$ and $C_j(\phi)$ as functions of ϕ, imply the absolute convergence of (2.44), as $v \to 0^+$ is justified. □

Proposition 2.1.1 suggests that, from the numerical point of view, in building a finite difference method, condition (2.32) can be used as a boundary condition that is coherent with the pricing equation (2.31). Under suitable assumptions on the coefficients. Similar conditions can be stated for other stochastic volatility models, see [30] and [31].

Proof. **Lemma 2.1.2** We evaluate, at first, the behavior of the integrand in I_j as $\phi \to 0^+$.. For the sake of simplicity, we omit the indication of the index j in all the coefficients. Thus, we can write

$$\begin{aligned} d(\phi) &= \sqrt{(b-\rho\sigma i\phi)^2 - \sigma^2(2ui\phi - \phi^2)} \\ &= \sqrt{b^2 - \rho^2\sigma^2\phi^2 - 2ib\rho\sigma\phi - 2iu\sigma^2\phi - \sigma^2\phi^2} \\ &= b\sqrt{1 + \frac{-\rho^2\sigma^2\phi^2 - 2ib\rho\sigma\phi - 2iu\sigma^2\phi - \sigma^2\phi^2}{b}}, \end{aligned} \tag{2.47}$$

and by Taylor expansion, around $phi \approx 0$, we get

$$\begin{aligned} d(\phi) &= b\left[1 + \frac{-\rho^2\sigma^2\phi^2 - 2ib\rho\sigma\phi - 2iu\sigma^2\phi - \sigma^2\phi^2}{2b} + o(\phi)\right] \\ &= b + \frac{1}{2}(-2ib\rho\sigma\phi - 2iu\sigma^2\phi) + o(\phi) \\ &= b - ib\rho\sigma\phi - iu\sigma^2\phi + o(\phi) \\ &= b - hi\phi + o(\phi). \end{aligned} \tag{2.48}$$

where $h = b\rho\sigma - u\sigma^2$. In particular $d(\phi) \to b$ when $\phi \to 0^+$. Now, we consider coefficient $g(\phi)$:

$$\begin{aligned} g(\phi) &= \frac{b-\rho\sigma i\phi + d(\phi)}{b-\rho\sigma i\phi - d(\phi)} \\ &\approx \frac{2b - \rho\sigma i\phi - hi\phi}{hi\phi - \rho\sigma i\phi} \\ &= \frac{g^{(1)}}{i\phi} + g^{(2)}, \end{aligned} \tag{2.49}$$

with $g^{(1)} = 2b/(h - \rho\sigma)$, $g^{(2)} = (\rho\sigma + h)(\rho\sigma - h)$. Therefore, we get the following asymptotic behavior, as $\phi \to 0^+$:

$$\begin{aligned} D(\phi) &= \frac{b - \rho\sigma i\phi + d(\phi)}{\sigma^2}\Big(\frac{1 - e^{d(\phi)\tau}}{1 - g(\phi)e^{d(\phi)\tau}}\Big) \\ &\approx \frac{2b - \rho\sigma i\phi - hi\phi}{\sigma^2}\Big(\frac{1 - e^{d(\phi)\tau}}{1 - g(\phi)e^{d(\phi)\tau}}\Big) \\ &\approx \frac{2b - \rho\sigma i\phi - hi\phi}{\sigma^2}\frac{(1 - e^{d(\phi)\tau})i\phi}{i\phi - g^{(1)}e^{b\tau} - g^{(2)}e^{b\tau}i\phi}. \end{aligned} \tag{2.50}$$

Multiplying for the conjugate of the term in the denominator, we obtain:

$$\begin{aligned} D(\phi) &\approx \frac{2b - \rho\sigma i\phi - hi\phi}{\sigma^2}\frac{(1 - e^{b\tau})i\phi[g^{(1)}e^{b\tau} + (g^{(2)}e^{b\tau} - 1)i\phi]}{[g^{(1)})^2e^{2b\tau} + (1 - (g^{(2)})e^{b\tau}]^2\phi^2} \\ &= -\frac{2b(1 - e^{b\tau})}{\sigma^2 g^{(1)}e^{b\tau}}i\phi + O(\phi^2) \\ &= mi\phi + O(\phi^2), \end{aligned} \tag{2.51}$$

with $m = -2b(1 - e^{b\tau})/[\sigma^2 g^{(1)}e^{b\tau}]$. We use previous estimates to analyze the integrand in I_j (2.27):

$$\begin{aligned} &\mathrm{Re}\left[\frac{e^{-i\phi\log(K)}f_j(\phi; x, v)}{i\phi}\right] = \\ &= -\frac{1}{\phi}e^{\mathrm{Re}(C_j(\phi)) + \mathrm{Re}(D_j(\phi))v}\sin\left[\mathrm{Im}(C_j(\phi)) + \mathrm{Im}(D_j(\phi))v + \phi\log\left(\frac{e^x}{K}\right)\right] \\ &\approx -\frac{1}{\phi}\sin\left\{\phi\left[r\tau + (a + v)m + x - \log K\right]\right\}. \end{aligned} \tag{2.52}$$

as $\phi \to 0^+$. So, being the last term in (2.52) bounded, we argue that the integrand in I_j is also bounded as $\phi \to 0^+$. Now, we consider the behavior as $\phi \to \infty$. By Lemma 2.1.1, we easily deduce that, there exist a constant $M > 0$, such that for ϕ large enough, it holds:

$$\begin{aligned} \left|\mathrm{Re}\frac{e^{-i\phi\log(K)}f(\phi; x, v)}{i\phi}\right| &\leq \frac{1}{\phi}e^{\mathrm{Re}(C_j(\phi)) + \mathrm{Re}(D_j(\phi))v} \\ &\leq Me^{-\phi\left(\frac{\kappa\theta}{\sigma}\tau\sqrt{1-\rho^2} + \frac{1}{\sigma}\sqrt{1-\rho^2}\right)}. \end{aligned} \tag{2.53}$$

Hence the integrand has an exponential decay as $\phi \to \infty$. $\square$

Proof. (*Lemma 2.1.1*) It suffices to prove the result for the case of $j = 1$.

$$\begin{aligned}\lim_{\phi\to\infty}\frac{d_1(\phi)}{\phi} &= \lim_{\phi\to\infty}\sqrt{\sigma^2(1-\rho^2)-\frac{2i\sigma(\rho\kappa-\frac{1}{2}\sigma)}{\phi}+\sigma^2/\phi^2}\\ &= \sigma\sqrt{1-\rho^2}. \end{aligned} \tag{2.54}$$

We use this equation to prove the second limit:

$$\begin{aligned}\lim_{\phi\to\infty} g_1(\phi) &= \frac{-\rho\sigma i+\sigma\sqrt{1-\rho^2}}{-\rho\sigma i-\sigma\sqrt{1-\rho^2}}\\ &= -\frac{(\sigma\sqrt{1-\rho^2}-\rho\sigma i)^2}{\sigma^2(1-\rho^2)+\rho^2\sigma^2}\\ &= -\frac{\sigma^2(1-\rho^2)-2i\sigma^2\rho\sqrt{1-\rho^2}-\rho^2\sigma^2}{\sigma^2}\\ &= -1+2\rho^2+2i\rho\sqrt{1-\rho^2}=(i\sqrt{1-\rho^2}+\rho)^2. \end{aligned} \tag{2.55}$$

Note that $\lim_{\phi\to\infty}|g_1(\phi)| = 1$. Now we need the following estimation to prove the next limit:

$$\begin{aligned}\lim_{\phi\to\infty}\frac{e^{d_1(\phi)\tau}-1}{c_1(\phi)e^{d_1(\phi)\tau}-1} &= \lim_{\phi\to\infty}\frac{1}{c_1(\phi)}\Big(1+\frac{1-c_1(\phi)}{c_1(\phi)e^{d_1(\phi)\tau}-1}\Big)\\ &= \lim_{\phi\to\infty}\frac{1}{c_1(\phi)}=\lim_{\phi\to\infty}\frac{\overline{c_1(\phi)}}{|c_1(\phi)|}\\ &= -1+2\rho^2-2i\rho\sqrt{1-\rho^2}. \end{aligned} \tag{2.56}$$

Moreover

$$\begin{aligned}\lim_{\phi\to\infty}\frac{D_1(\phi)}{\phi} &= \frac{1}{\sigma^2}(-1+2\rho^2-2i\rho\sqrt{1-\rho^2})(-\rho\sigma i+\sigma\sqrt{1-\rho^2})\\ &= \frac{-i\rho-\sqrt{1-\rho^2}}{\sigma}. \end{aligned} \tag{2.57}$$

Finally, consider

$$\begin{aligned}\lim_{\phi\to\infty}\frac{\ln\Big(\frac{c_1(\phi)e^{d_1(\phi)\tau}-1}{c_1(\phi)-1}\Big)}{\phi} &= \lim_{\phi\to\infty}\frac{\ln\Big(e^{d_1(\phi)\tau}\frac{c_1(\phi)}{c_1(\phi)-1}\Big)}{\phi}\\ &= \lim_{\phi\to\infty}\left(\frac{\ln(e^{d_1(\phi)\tau})+\ln(\frac{c_1(\phi)}{c_1(\phi)-1})}{u}\right)\\ &= \lim_{\phi\to\infty}\frac{d_1(\phi)\tau}{\phi}=\sigma\sqrt{1-\rho^2}. \end{aligned} \tag{2.58}$$

Hence, we can obtain the last relation of Lemma 2.1.1:

$$\lim_{\phi\to\infty} \frac{C_1(\phi)}{\phi} = -\frac{\kappa\theta}{\sigma^2}\tau\left(\sigma\sqrt{1-\rho^2} + i\sigma\rho\right). \tag{2.59}$$

□

2.2 Pricing Equation and Viscosity Solutions

Flexibility and availability of closed-form solutions to the pricing equation, represent the success of affine models in financial modelling. As we discussed in the introduction, in presence of an affine model, the computation of the risk-neutral price of (European-type) derivative is reduced to solving a system of ordinary differential equations and then inverting the Fourier transform. Unfortunately, not all financial derivatives could be localized among affine models with affine or exponential-affine final payoff. From a general point of view, if the problem at hand does not fit in the class of affine models, the risk-neutral price can be computed by solving the pricing equation. Here we recall a result proposed in [16] that includes, in the time-homogeneous case, a general enough formulation of the pricing problem:

$$\begin{cases} \partial_t u(t,x) + Lu(t,x) - c(x)u(t,x) = f(t,x), & (t,x) \in (0,T)\times D, \\ u(T,x) = \phi(x), \quad x \in D, \end{cases} \tag{2.60}$$

where, for every smooth $g : D \subset \mathbb{R}^d \to \mathbb{R}$, the operator L is defined as

$$Lg(x) = \nabla g(x) b(x) + \frac{1}{2}\mathrm{tr}(\nabla^2 g(x) a(x)) + \int_D (g(z) - g(x)) m(x, dz), \tag{2.61}$$

for any $x \in D$. Matrix $a = \sigma\sigma^\top$ corresponds to the diffusion matrix of a stochastic process $(X_t)_t$ with values in the domain $D \subset \mathbb{R}^d$, b represents the drift of the process under a suitable (risk-neutral) probability measure, c is a discount rate function, ϕ is the final payoff function, f is the cost of execution. If presence of jumps, there is also a measure m which summarizes the jump intensity and the probability distribution of $(X_t)_t$. Often, in financial modelling, the state space D is not the whole space $\mathbb{R}^d$ and the equation just mentioned above could have, often concurrently, some difficulties, such as:

- The diffusion matrix a is singular on the boundary of the domain D or is even identically zero in some direction. The former singularity arises in some stochastic volatility models, like in the Heston model (1993).
- The drift b and the matrix σ are not Lipschitz-continuous up to the boundary of D. This occurs, for instance, whenever some components are square root-diffusions, like in the CIR (1985) or in the Heston model (1993).
- The coefficients b and a are fast growing near the boundary or at infinity.
- The jump intensity is not bounded.
- The state space D has a boundary, but no boundary conditions are specified in the model. This is the case in most models where D has a boundary.

In presence of the above features, existence, uniqueness and regularity of solutions to system (2.60) may very well not hold. Still, it is common practice in mathematical finance to develop numerical methods taking for granted existence, uniqueness and regularity of the solution to the valuation equation. In contrast, the theory of viscosity solutions allows to deal with singular diffusion matrices (see Appendix A). Viscosity solutions, although a priori not differentiable, are continuous functions if the data are continuous, and are especially well suited to be computed numerically. In a paper of Costantini *et al.* [16], the existence and uniqueness of viscosity solutions for equation (2.60), are proved assuming only that coefficients b and σ are locally Lipschitz-continuous in the interior of D and that exists a Lyapunov-type function.

2.2.1 Existence and Uniqueness results

In this section, we recall the assumptions and the main results proved in [16]. Consider a security derived from the value of a state vector X, with infinitesimal generator (2.61), and with final payoff ϕ at time T . If the valuation equation (2.60) has one and only one viscosity solution u, then the arbitrage-free price of the derivative contract at time $t < T$, for the case $f \equiv 0$, is given by

$$u(t, X_t) = E[e^{-\int_t^T c(X_s)ds}\phi(X_T)|\mathcal{F}_t^X]. \tag{2.62}$$

where $\{\mathcal{F}_t^X\}_t$ stands for the filtration generated by X. The main result proved in [16] concerns with the existence of a unique viscosity solution for (2.60). Let D be a (possibly unbounded) star-shaped open subset of $\mathbb{R}^d$. When the operator (2.61) is a pure differential operator ($m(x, \cdot) = \delta_x$), the following assumption on the coefficients is considered.

Assumption 2.2.1. $a = \sigma\sigma^T$, where $\sigma : D \to \mathbb{R}^{d\times d}$ and $b : D \to \mathbb{R}^d$ are Lipschitz-continuous on compact subsets of D.

Otherwise, the following is given.

Assumption 2.2.2. $a = (a_{i,j})_{i,j=1,..,d}$, $a_{i,j} \in C^2(D)$, $m : D \to M(D)$ (where $M(D)$ denotes the space of finite Borel measures on D) is continuous and

$$\sup_{x\in D}\Big|\int_D g(z)m(x, dz)\Big| < \infty, \quad \forall\, g \in C_c(D). \tag{2.63}$$

Notice that the assumption $a_{i,j} \in C^2(D)$, for all $i, j = 1, ..., d$, implies Assumption 2.2.1.

Assumption 2.2.3. There exists a no-negative function $V \in C^2(D)$ such that:

$$\int_D V(z)m(x, dz) < +\infty, \quad LV(x) \le C(1 + V(x)),\ \forall x \in D, \tag{2.64}$$

$$\lim_{x\in D, x\to x_0} V(x) = +\infty,\ \forall\, x_0 \in \partial D, \qquad \lim_{x\in D,\, |x|\to+\infty} V(x) = +\infty. \tag{2.65}$$

This assumption ensures that the stochastic process corresponding to our system does not blow up in finite time and does not reach the boundary of D, as is the case for most models in finance. In our specific case, σ could, at the same time, be zero in D or on boundary of D and be degenerate (of Hölder-type) on the boundary. The function V is called a Lyapunov-type function for system (2.60). From the probabilistic point of view, the existence of a Lyapunov function ensures that the state process $(X_t)_t$ does not reach the boundary and does not blow up in finite time. The function V also determines the growth rate that we can allow for the data f and ϕ of (2.61). In fact, in [16], the folllowing assumptons are also considered.

Assumption 2.2.4. Let $f \in C([0,T] \times D)$, $c, \phi \in C(D)$ and c is bounded from below. There exists a strictly increasing function $\varphi : [0, +\infty) \to [0, \infty)$ such that

$$s \to s\varphi(s) \text{ is convex, } \lim_{s \to \infty} \varphi(s) = +\infty, \tag{2.66}$$

$$(s_1 + s_2)\varphi(s_1 + s_2) \le C(s_1\varphi(s_1) + s_2\varphi(s_2)), \forall s_1, s_2 \ge 0, \tag{2.67}$$

and

$$\begin{aligned} |f(t,x)|\varphi(f(t,x)) &\le C_T(1 + V(x)), \\ |\phi(x)|\varphi(|\phi(x)|) &\le C_T(1 + V(x)), \end{aligned} \tag{2.68}$$

for all $(t,x) \in [0,T] \times D$.

All the results still hold under the following localized version of Assumptions 2.2.3 and 2.2.4: there exists a sequence of nonegative $C^2(D)$ functions $\{V_n\}_{n \in N}$ such that

$$\sup_{x \in K} \sup_{n} V_n(x) < +\infty, \ \forall\ K \subseteq D \text{ compact}, \tag{2.69}$$

$$\int_D V_n(z) m(x, dz) < +\infty, \forall x \in D. \tag{2.70}$$

$$\lim_{x \to x_0} V_n(x) = +\infty, \forall x_0 \in \partial D, \quad \lim_{x \in D, |x| \to +\infty n \in N} \inf V_n(x) = +\infty, \tag{2.71}$$

$$LV_n(x) \le C(1 + V_n(x)), \ \forall x \in D, |x| \le n, \tag{2.72}$$

$$|f(t,x)|\varphi(|f(t,x)|) \leq C_T(1+\mathrm{inf}_n V_n(x)) \ \forall (t,x) \in [0,T]\times D, \tag{2.73}$$

$$|\phi(x)|\varphi(|\phi(x)|) \leq C(1+\inf_n V_n(x)) \ \forall x \in D. \tag{2.74}$$

For the proof the the following theorems, we refer the reader to [16].

Theorem 2.2.1. Let L be the operator defined by (2.61). Then, for every probability distribution $\mathbb{P}_0$ on D, there exists one and only one stochastic process X which is a solution of the martingale problem for $(L,\mathbb{P}_0)$ with $D(L) = C_c^2(D)$. X is a homogeneous strong Markov process with paths in $D_D[0,\infty)$. Denoting by X^x the process with $\mathbb{P}_0 = \delta_x$, $x \in D$, we have, for every $T > 0$ and $F_t^{X^x}$-stopping time τ such that

$$\sup_{0\leq t\leq T} \mathbb{E}[V(X^x_{t\wedge\tau})] \leq C_T(1+V(x)). \tag{2.75}$$

The following results state the existence and the uniqueness for the viscosity solutions to the valuation equation (2.60).

Theorem 2.2.2. For every $x \in D$, let X^x be the process of Theorem 2.2.1 with $\mathbb{P}_0 = \delta_x$. Then, for every $t \in [0,T]$,

$$\mathbb{E}\left[\left|\phi(X^x_{T-t})e^{-\int_0^{T-t} c(X^x_r)dr} - \int_0^{T-t} f(t+s,X^x_s)e^{-\int_0^s c(X^x_r)dr}ds\right|\right] \leq \infty. \tag{2.76}$$

The function

$$u(t,x) = \mathbb{E}\left[\phi(X^x_{T-t})e^{-\int_0^{T-t} c(X^x_r)dr} - \int_0^{T-t} f(t+s,X^x_s)e^{-\int_0^s c(X^x_r)dr}ds\right] \tag{2.77}$$

is continuous on $[0,T]\times D$ and is a viscosity solution to (2.60) satisfying

$$|u(t,x)|\varphi(|u(t,x)|) \leq C_T(1+V(x)), \ \forall (t,x) \in [0,T]\times D. \tag{2.78}$$

Theorem 2.2.3. There exists only one viscosity solution to (2.60) satisfying (2.78).

We discuss the application of previous results to some pricing problems in finance.

Example 2.2.1. [Heston model]. The Heston model (which will be fully developed in next chapter from the point of view of new calibration techniques), is specified by

$$\begin{cases} dS_t = rS_t d_t + \sqrt{v_t} S_t \, dW_t, \\ dv_t = \kappa(\theta - v_t) d_t + \sigma \sqrt{v_t} \, dZ_t, \end{cases} \tag{2.79}$$

where $(Z_t)_t$ and $(W_t)_t$ are instantaneously correlated Brownian motions: $dW_t dZ_t = \rho \, dt$, with $\rho \in [-1, 1]$. Therefore (2.60) takes the form of

$$Lg(S, v) = rS\partial_S g + \kappa(\theta - v)\partial_v g + \frac{vS^2}{2}\partial_S^2 g + \frac{\sigma^2 v}{2}\partial_v^2 g + \sigma v S \rho \partial_{Sv}^2 g,$$

for all smooth functions $g : (0, \infty)^2 \to \mathbb{R}$. Under Feller's condition (1951) for the volatility process $(v_t)_t$, $2\kappa\theta \geq \sigma^2$, the valuation problem admits, for instance, the following Lyapunov function:

$$V(S, v) = -\log S - \log v + s \log(S + 3) + S(v + 1) + v. \tag{2.80}$$

The valuation problem for an option with linearly increasing payoff function ϕ, with $f = 0$ and $c = r$ (constant), satisfies Assumptions 2.2.3-2.2.4, taking

$$\varphi(s) = \log(s + 3).$$

Therefore there exists a unique viscosity solution to the valuation problem associated to model (2.79) . Indeed, we remark that the existence of this function, expecially for stochastic volatility models, allow to manage the absence of a boundary condition along the critical direction of $v = 0$, where usually the diffusion coefficient loses the Lipschitz-continuity. This point will addressed in Chapter 4.

Example 2.2.2. [Arithmetic Asian Option]. Let r be a constant short-term interest rate and assume that the underlying asset follows

$$dS_t = rS_t dt + \ S_t dW_t, \tag{2.81}$$

under the 1-dimensional Brownian motion $(W_t)_t$, then the valuation problem for the price of the Arithmetic Asian oating-strike put option is

$$\partial_t u + \frac{\sigma^2 S^2}{2}\partial_S^2 u + rS\partial_S u + S\partial_A u - ru = 0, \quad u(T, S, A) = \left(S - \frac{A}{T}\right)_+ . \tag{2.82}$$

where T is the maturity of the option. Here the variable A stands for the value of the average price process $A_t = \int_0^t S_s ds$ and the state space domain is $D = (0, +\infty)^2$. The differential operator has the following coefficients:

$$a(S, A) = \begin{bmatrix} \sigma^2 S^2 & 0 \\ 0 & 0 \end{bmatrix}, \qquad b(S, A) = \begin{bmatrix} rS \\ S \end{bmatrix}.$$

Therefore L is strongly degenerate, in the sense that $a(S, A)$ has everywhere rank 1 on the domain $(0, \infty)^2$ and it is null at the boundary $S = 0$. The valid Lyapunov function is given by

$$V(S, A) = S^{-1}A^{-1} + A^2 + S^2. \tag{2.83}$$

Indeed, we have

$$\begin{aligned} LV(S, A) &= \frac{\sigma^2}{SA} - \frac{r}{SA} - \frac{1}{A^2} + \sigma^2 S^2 + 2rS^2 + 2SA \\ &\leq (\sigma^2 - r)S^{-1}A^{-1} + (2r + \sigma^2 + 1)S^2 + A^2 \\ &= (\sigma^2 - r)V(S, A) + (3r + 1)S^2 + (1 - \sigma^2 + r)A^2 \\ &\leq CV(S, A), \end{aligned}$$

for any $S,\ A > 0$, where $C = 2 + 4r + \sigma^2$. The valuation problem for the arithmetic Asian option is well posed from the point of view of viscosity solutions, taking $\varphi(s) = \log(s + 3)$.

Example 2.2.3. [Geometric Asian Option under Heston model]. Consider model (2.79) where, for the sake of simplicity, we assume that $\rho = 0$; let $x_t = \log(S_t)$ and $G_t = \int_0^t x_s ds$, the geometric Asian put option with maturity T and strike price K has the payoff function $\phi(x, v, G) = (K - e^{G/T})_+$. The valuation equation is

$$\partial_t u + \frac{v}{2}\partial_x^2 u + \frac{\sigma^2 v}{2}\partial_v^2 u + \left(r - \frac{v}{2}\right)\partial_x u + k(\theta - v)\partial_v u + x\partial_G u - ru = 0.$$

As in the previous example, also in this case the operator L is strongly degenerate and degenerates at the boundary $v = 0$. The domain of the problem is $D = \mathbb{R} \times (0, +\infty) \times \mathbb{R}$, with coefficients

$$a(x, v, G) = \begin{bmatrix} v & 0 & 0 \\ 0 & \sigma^2 v & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad b(x, v, G) = \begin{bmatrix} r - \frac{v}{2} \\ k(\theta - v) \\ x \end{bmatrix}.$$

I the condition for the volatility process $2\kappa\theta > \sigma^2$ is satisfied, a suitable Lyapunov function for this case is $V(x, v, G) = v^{-a} + v^2 + x^2 + G^2$, with $a = \frac{2\kappa\theta}{\sigma^2} - 1$. In fact,

$$\begin{aligned} LV(x, v, G) &= av^{-a-1}\left[\frac{\sigma^2}{2}(a+1) - \kappa\theta\right] + \kappa\theta v^{-a} + 2xr - 2xv \\ &\quad + (1 + \sigma^2 + 2\kappa)v - 2\kappa v^2 + 2xG \\ &\leq \kappa\theta v^{-a} + (1 + \sigma^2 + 2\kappa)^2 + r^2 + 2v^2 + 3x^2 + G^2 \\ &\leq C(1 + V(x, v, G)), \qquad\qquad \forall\ x, G \in \mathbb{R},\ \ v > 0, \end{aligned}$$

where $C = \max(r^2 + (1 + \sigma^2 + 2\kappa)^2, \kappa\theta + 3)$. Thus, the associated Dirichlet problem (with $f = 0$ and $c = r$ costant) is well posed in the sense of viscosity solutions, taking $\varphi(s) = s$, given that the payoff function is bounded.

Example 2.2.4. [Local Volatility Model]. Let X_t be the 1-dimensional underlying state process of a contingent claim with payoff $\phi(X_T)$ at maturity $T > 0$, where $\phi : (\alpha, \beta) \to \mathbb{R}$ is a continuous function, and suppose that follows

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \tag{2.84}$$

where $(W_t)_t$ is a 1-dimensional Brownian motion. Let b, $\sigma : (\alpha, \beta) \to \mathbb{R}$ be locally Lipschitz continuous functions on (α, β). The natural domain of the process (from a financial point of view) is the (possibly unbounded) interval (α, β). Given the short-term rate $r > 0$, the valuation equation becomes

$$\partial_t u + \frac{1}{2}\sigma^2(x)\partial_x^2 u + b(x)\partial_x u - ru = 0, \ \ u(T, x) = \phi(x). \tag{2.85}$$

This kind of model includes the so called local volatility model, where $(X_t)_t$ represents the stock price process. The operator L can be degenerate at the boundaries α, β with the coefficient σ that is not Lipschitz continuous up α

and/or β and, at the same time, can be strongly degenerate in the interior of the interval. We consider the following conditions which cover both these degeneracies: There exist $\alpha < p < q < \beta$ such that $\sigma^2 \in C^1(E)$, $\sigma^2 > 0$ in E, $E = (\alpha, p] \cap [q, \beta)$, and

$$\lim_{x \to \alpha^+} \int_x^p \sigma^{-2}(\xi) d\xi = +\infty, \qquad \lim_{x \to \alpha^+} \left[b(x) - \frac{1}{2} \frac{d}{dx} \sigma^2(x) \right] > 0, \tag{2.86}$$

$$\lim_{x \to \beta^-} \int_q^x \sigma^{-2}(\xi) d\xi = +\infty, \qquad \lim_{x \to \beta^-} \left[b(x) - \frac{1}{2} \frac{d}{dx} \sigma^2(x) \right] < 0. \tag{2.87}$$

There exist $\alpha < p' \leq p$, $q \leq q' < \beta$ such that, by (2.86)-(2.87), the function

$$V(x) = \begin{cases} \int_x^{p'} \sigma^{-2}(\xi) d\xi & \text{if } x \in (\alpha, p'), \\ \int_{q'}^x \sigma^{-2}(\xi) d\xi & \text{if } x \in (q', \beta), \end{cases}$$

satisfies $V(x) > 0$ and $LV(x) < 0$, for every $x \in (\alpha, p') \cup (q', \beta)$. Then, we consider an extension of V to the interval $[p', q']$ so that V is strictly positive, with $V \in C^2((\alpha, \beta))$. Therefore, by the continuity of the coefficients b and σ, it is clear that $LV(x)$ is bounded over the closed interval $[p', q']$ and V becomes a Lyapunov function for the Dirichlet problem (2.85) on the domain $D = (\alpha, \beta)$. Then, for any payoff function ϕ satisfying Assumption 2.2.4, the pricing problem (2.85) admits a unique viscosity solution as established in Theorems 2.2.2-2.2.3.

2.3 Boundary Conditions in SV models

By the previous results and examples, we deduce that the existence of the Lyapunov-type function ensures that the underlying state process will never approach the boundary. Nevertheless, from a numerical point of view, a boundary condition along the boundary is needed: a condition coherent with the princing equation possibly reducing the numerical error. In this section we focus our attention on stochastic volatility models whose generale formulation is given in (2.20). The associated pricing equation for a

European-type option is the following:

$$\begin{aligned}\frac{\partial u}{\partial t}+\frac{1}{2}s^2b^2(v)\frac{\partial^2 u}{\partial s^2} &+ \frac{1}{2}c^2(v,t)\frac{\partial^2 u}{\partial v^2}+\rho b(v)c(v,t)\frac{\partial^2 u}{\partial v\partial s}+\\ &+ rs\frac{\partial u}{\partial s}+a(v,t)\frac{\partial u}{\partial v}-ru=0,\end{aligned} \tag{2.88}$$

where r stands for the risk-free rate (assumed to be constant). In [31] the condition enforceable along the boundary is analyzed for the simpler case of a single-factor model. Here, the state process $(X_t)_t$ is described as a nonnegative stochastic process solution to

$$dX_t=\beta(t,X_t)dt+\sigma(t,X_t)dW_t, \tag{2.89}$$

where $(W_t)_t$ is a one dimensional Brownian motion on a given filtered probability space, $\sigma(t,0)=0$ and $\beta(t,0)\geq 0$, for all $t\geq 0$. Under suitable regularity assumptions, the no arbitrage price of an option corresponding to a payoff function $g:[0,\infty)\to[0,\infty)$, with discount function $c:[0,\infty)\to[0,\infty)$, is the unique solution of the following problem:

$$\partial_t u_t(t,x)+\frac{1}{2}\sigma^2(t,x)\partial_x^2 u(t,x)+\beta(t,x)\partial_x u(t,x)=c(x)u(t,x), \tag{2.90}$$

for $(t,x)\in(0,T)\times(0,\infty)$, with terminal condition $u(x,T)=g(x)$, for any $x>0$. We will report here the assumtpions made in ([31]) for time-independent coefficients and for $c(x)\equiv x$ (term-structure equation for an option where X_t stands for the short-term interest rate).

Hypotesis 2.3.1. $\beta\in C([0,\infty))$ is continuously differentiable in x with bounded derivative, $\beta(0)\geq 0$; $\sigma\in C([0,\infty))$ is such that $\alpha(x):=\frac{1}{2}\sigma^2(x)$ is continuously differentiable with Hölder continuous derivative, $\sigma(x)=0$ only if $x=0$. The functions β,σ and α' are all of, at most, linear growth:

$$|\beta(x)|+|\sigma(x)|+|\alpha_x(x)|\leq C(1+x), \tag{2.91}$$

for all $x\geq 0$. The payoff function $g:[0,\infty)\to[0,\infty)$ is continuously differentiable with both g and $g^{'}$ bounded.

Moreover, by formally inserting $x = 0$ in the pricing equation, an additional boundary condition is considered:

$$\partial_t u(t, 0) + \beta(0)\partial_x u(t, 0) = 0, \tag{2.92}$$

for all $t \in (0, T)$. An additional assumption is also considered. Let $(W_t)_t$ be a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

Assumption 2.3.1. The coefficients σ and β are such that, path-wise, uniqueness holds for the equation

$$dY(t) = (\alpha(x) + \beta)(Y(t), t)dt + \sigma(Y(t), t)dW_t \tag{2.93}$$

Theorem 2.3.1. Under Hypothesis 2.3.1 and Assumption 2.3.1, the term structure equation admits a unique solution $u \in C([0, T]\times[0, \infty))\cap C^1([0, T)\times [0, \infty)) \cap C^{1,2}((0, T) \times (0, \infty))$ which satisfies (2.90) and the terminal condition $u(x, T) = g(x)$, for any $x > 0$, and (2.92).

In particular, condition (2.92) becomes a necessary condition for the well posedeness of the pricing problem when the underlying state process can reach the boundary, at $x = 0$. This feature is only partially addressed by authors of [31].

For the case of stochastic volatility models, we recall the result proved in [30] , where the Black-Scholes equation in stochastic volatility models is studied. In particular, in such a work it is proved that there exists a unique classical solution to the parabolic differential equation with a certain boundary behaviour for vanishing values of the volatility. If the boundary is attainable, then this boundary behaviour serves as a boundary condition and guarantees uniqueness in appropriate function spaces. On the other hand, if the boundary is non-attainable, then the boundary behaviour is not needed to guarantee uniqueness, but is nevertheless very useful for instance from a numerical perspective. More precisely, the underlying model is the following. Let S be the stock price process and y the volatility process, defined by

$$\begin{cases} dS_t = S_t\sqrt{y_t}dW_t \\ S_0 = s_0 > 0, \end{cases} \tag{2.94}$$

and

$$\begin{cases} dy_t = \beta(y_t)dt + \sigma(y_t)dZ_t \\ y_0 = y_0 \geq 0. \end{cases} \tag{2.95}$$

respectively, where $(W_t)_t$ and $(Z_t)_t$ are two standard Brownian motions with constant correlation $\rho \in (-1, 1)$. The corresponding pricing operator becomes

$$Lu = \beta(y)\frac{\partial u}{\partial y} + \frac{1}{2}\sigma^2(y)\frac{\partial^2 u}{\partial y^2} + \frac{1}{2}S^2 y\frac{\partial^2 u}{\partial S^2} + \rho\sigma(y)S\sqrt{y}\frac{\partial^2 u}{\partial S \partial y}. \tag{2.96}$$

A classical solution to the valuation equation considered in [30] is a function $u \in C^{1,2}([0,T) \times (0,\infty)^2) \cap C^{1,1,0}([0,T) \times (0,\infty) \times [0,\infty))$, satisfying the system

$$\begin{cases} \partial_t u + Lu = 0 & \text{on } [0,T) \times (0,\infty)^2, \\ u(t,0,y) = g(0) & y \in [0,T] \times [0,\infty), \\ \partial_t u(t,S,0) + \beta(0)\partial_y u(t,S,0) = 0 & (t,S) \in [0,T) \times (0,\infty), \\ u(T,S,y) = \phi(S) & (S,y) \in (0,\infty)^2. \end{cases} \tag{2.97}$$

The model is specified with a zero interest rate and with time-homogeneous coefficients. However generalizations to a deterministic interest rate and time-dependent coefficients are straightforward. Here the main assumptions on the coefficients given in [30] are summarized in the following hypoteses.

Hypotesis 2.3.2. $\beta \in C^1([0,\infty))$ with a Hölder continuous derivative, $\beta(0) \geq 0$. $\sigma(y) > 0$, for all $y > 0$, $\sigma(0) = 0$, and the function $\sigma^2 \in C^1([0,\infty))$ with a Hölder continuous derivative. The growth condition

$$|\beta(y)| + \sigma(y) \leq C(1+y), \tag{2.98}$$

holds for all $y \geq 0$, where C is a constant. The payoff function ϕ is bounded and is twice continuously differentiable on $[0,\infty)$. Moreover, $S \in (0,\infty) \mapsto S\phi'(S)$ and $S \in (0,\infty) \mapsto S^2\phi''(S)$ are bounded.

Thus, the main result proved in [30] is the following Theorem.

Theorem 2.3.2. Assume that Hypothesis 2.3.2 holds. Then, there is at most one classical solution to the pricing problem for (2.95)-(2.94), which is of strictly sublinear growth in S and polynomial in y.

Boundary conditions are included in the notion of a classical solution regardless if the boundary can be hit or not. On the other hand, it is clear that standard payoff functions related to call and put options do not satisfy the assumptions specified in 2.3.2. An extension to the case of the put option and in the case of Heston's model, has been proposed in Ould (2012) (Theorem 3.1), ensuring the following properties of the solution:

1) $u \in C(([0,T) \times [0,\infty)^2) \cap C^{1,0,1}([0,T) \times (0,\infty) \times [0,\infty))$.

2) $u \in C^{1,2,2}([0,T) \times (0,\infty)^2)$.

3) For every $t \in [0,T)$, $v > 0$, the function $S \mapsto u(t,S,v)$ is increasing and strictly convex on $(0,\infty)$;

4) For every $t \in [0,T)$, $S > 0$, the function $v \mapsto u(t,S,v)$ is strictly increasing.

Precisely, under the usual condition $2\kappa\theta > \sigma^2$, no boundary condition is required to state the existence and uniqueness of the solution. We aim to prove here that, for the stochastic volatility model (2.95)-(2.94), if the underlying process does not hit the boundary at $y = 0$, the assumptions given in [31] (see Hypothesis 2.3.2) can be weakened in order to state the existence of uniqueness of a viscosity solution.

Theorem 2.3.3. Let β and $\alpha = \frac{1}{2}\sigma^2$ be locally Lipschitz continuous on $(0,\infty)$. Let $\alpha \in C([0,\infty)) \cap C^1((0,\varepsilon))$, with bounded derivative on $(0,\varepsilon)$, for some $\varepsilon > 0$, $\alpha(y) = 0$ if and only if $y = 0$. Moreover, there exists a positive constant $C > 0$ such that

$$|\beta(y)| + \alpha(y) \leq C(1+y), \tag{2.99}$$

for all $y > 0$. Assume the following condition:

$$\lim_{y \to 0^+} \left\{ \alpha^{'}(y) - \beta(y) \right\} < 0. \tag{2.100}$$

Then, for every $\phi \in C([0,\infty))$, satisfying $|\phi(S)| \leq C(1+S)$, for all $S \geq 0$, the pricing problem associated to (2.96), with payoff function ϕ, has a unique viscosity solution $u \in C([0,T] \times (0,\infty)^2)$ such that

$$|u(t,S,y)| \leq C\,(1+S)\,, \tag{2.101}$$

for all $S > 0$, $y > 0$, $t \in [0,T]$.

Remark 2.3.4. Under the assumptions of the previous Theorem, the pricing problem associated to (2.96) is well posed without imposing a boundary condition at $y = 0$ and at $S = 0$.

Remark 2.3.5. If both β and $\alpha^{'}$ are continuous up to $y = 0$, then condition (2.100) is equivalent into requiring that $\alpha'(0) < \beta(0)$.

Note that when some of the assumptions of Theorem (2.3.3 do not apply, there are explicit examples for which the Dirichlet problem for (2.96) is not well posed. In particular, uniqueness, is not ensured. We consider here two examples.

Example 2.3.1. Let us consider here the Cox-Ingersoll-Ross (CIR) [17] model used for pricing zero coupon bonds, where the underlying described under the empirical measure $\mathbb{P}$ is the short term interest rate, that is

$$dr_t = \kappa(\theta - r)dt + \sigma\sqrt{r_t}dZ_t, \tag{2.102}$$

where $\kappa,\ \theta > 0$ and Feller's condition is satisfied $2\kappa\theta \geq \sigma^2$. In this setting, a well known fact is that the market in not complete (see Proposition 1.4.1) and a market risk-premium function has to be defined. A usual form for such a premium is given in [28] and it takes the form of $\lambda(r) = \psi_0 r^{-1/2} + \psi_1 r^{1/2}$, for some constant parameters $\psi_0,\ \psi_1$. Under the CIR model bond's value $p(t,r)$ satisfies the valuation problem:

$$\begin{cases} \partial_t p + \frac{\sigma^2}{2} r \partial^2_{rr} p + \overline{\kappa}(\overline{\theta} - r)\partial_r p - rp = 0 & \text{on } (0,T) \times (0,\infty), \\ p(T,r) = 1. \end{cases} \tag{2.103}$$

where $\overline{\kappa} = \kappa +\ \psi_1$ and $\overline{\theta} = \frac{\kappa\theta - \sigma\psi_0}{\kappa +\ \psi_1}$. A solution of the pricing problem is:

$$p_1(t,r) = A(t)e^{-B(t)r}, \tag{2.104}$$

where

$$A(t) = \left[\frac{2\gamma e^{(\overline{\kappa}-\gamma)(T-t)/2}}{2\gamma + (\overline{\kappa}-\gamma)[1-e^{-\gamma(T-t)}]} \right]^{2\overline{\kappa}\overline{\theta}/\sigma^2}, \tag{2.105}$$

$$B(t) = \frac{2(1-e^{-\gamma(T-t)})}{2\gamma + (\overline{\kappa}-\gamma)[1-e^{-\gamma(T-t)}]}, \tag{2.106}$$

$$\gamma = \sqrt{\overline{\kappa}^2 + 2\sigma^2}. \tag{2.107}$$

However, when the values of ψ_0, ψ_1 are such that $2\overline{\kappa}\overline{\theta} < \sigma^2$, there is also another solution, made as:

$$p_2(t,r) = \begin{cases} A(t)e^{-B(t)r}\left[1 - \frac{\Gamma(\nu, rz(t))}{\Gamma(\nu,0)}\right], & 0 < t < T \\ 1 & ; \ t = T. \end{cases} \tag{2.108}$$

where

$$\Gamma(\nu, m) = \int_m^{\infty} e^{-\tau}\tau^{\nu-1} d\tau \tag{2.109}$$

is the incomplete gamma function with $\nu = 1 - 2\overline{\kappa}\overline{\theta}/\sigma^2$, and

$$z(t) = \frac{2}{\sigma^2} \frac{e^{-\overline{\kappa}(T-t)}A(t)^{-\frac{1}{\overline{\kappa}\overline{\theta}}}}{\int_t^T e^{-\overline{\kappa}(T-s)}A(s)^{-\frac{1}{\overline{\kappa}\overline{\theta}}}ds}. \tag{2.110}$$

For every $t < T$, $r > 0$, it holds:

$$p_2(t,r) < p_1(t,r), \tag{2.111}$$

with $p_2(T,\cdot) = p_1(T,\cdot)$. Under the inequality

$$2\overline{\kappa}\overline{\theta} < \sigma^2 \tag{2.112}$$

there is not an equivalent martingale measure $\mathbb{Q}$: r can hit zero under $\mathbb{Q}$ even though it cannot under $\mathbb{P}$. In fact measures $\mathbb{P}$ and $\mathbb{Q}$ are not equivalent: they are different for those event which almost surely will never happen (their zero probability events are not the same). The problem is not well posed from the financial point of view and from the analytic point of view, a boundary condition is required. In the CIR model the value of coefficients examined in Theorem 2.3.3 are $\sigma^2(y) = \overline{\sigma}^2 y$ and $\beta(y) = \overline{\kappa}(\overline{\theta} - y)$.. These coefficients verify (2.112) and so do not satisfy condition (i).

Example 2.3.2. This example is to justify the upper bound for the solution (2.101). If it is removed, we can have multiple solutions for the pricing even if all other assumptions of Theorem 2.3.3 hold true. Let us consider a modified version of Heston's model studied in Grünbichler and Longstaff (1996):

$$\begin{cases} dS_t = rS_t\,dt + \sqrt{y_t}S_t\,dW_t, \\ dy_t = \eta^2 dt + \eta\sqrt{y_t}\,dW_t. \end{cases}$$

where $(W_t)_t$ is a standard Brownian motion and $\eta > 0$ is a constant. For this model, the valuation equation is

$$\partial_t u + \frac{1}{2}S^2 y\partial_S^2 u + \frac{1}{2}\eta^2 y\partial_y^2 u + \eta S y\partial_{yS}^2 u + rS\partial_S u + \eta^2\partial_y u - ru = 0,$$

on $(0,T)\times(0,\infty)^2$. Clearly the coefficients $\beta(y) = \eta^2$ and $\sigma(y) = \eta\sqrt{y}$ satisfy (i) and the volatility process $(y_t)_t$ cannot hit zero, if $y_0 > 0$. A regular solution $u_1 = u_1(t,S,y)$ is available for payoff functions $\phi = \phi(S,y)$ corresponding to European equity options and volatility derivatives, see Grünbichler and Longstaff (1996) and also Heston (1993). Let us consider the function $u_2(t,S,y) = u_1(t,S,y) + \Pi(t,y)$, with

$$\Pi(t,y) = \frac{1}{y}\exp\left(-r(T-t) - \frac{2y}{\eta^2(T-t)}\right).$$

Function u_2 is another solution to the pricing problem with the same payoff function as u_1. However u_2 does not satisfy inequality (2.101). In this case there exists an arbitrage involving a short position in the replicating strategy for u_2 and a long position in that for u_1. From the financial point of view, the role of u_2 is related to the presence of bubbles in the market and is studied in [42].

Proof of Theorem 2.3.3. Due to the assumptions made on the coefficients σ and β, Assumptions 2.2.1 and 2.2.4, proposed in [16], are readily verified on $(0,\infty)$. It has to be proved now Assumption 2.2.3. By assumption (i), there exists $c \in (0,\varepsilon)$ such that $\alpha'(y) < \beta(y)$, for all $y \in (0,c)$. Since α' is bounded on $(0,\varepsilon)$, we set H to be the supremum of $|\alpha'|$ on $(0,c)$. Thus, we consider the scale measure for the volatility process:

$$p_c(y) = \int_c^y \exp\left(-2\int_c^\eta \frac{\beta(\tau)}{\sigma^2(\tau)}d\tau\right)d\eta, \tag{2.113}$$

and the speed measure, defined as

$$V_c(y) = \int_c^y p_c'(\xi) \int_c^\xi \frac{2}{p_c'(z)\sigma^2(z)} dz d\xi, \tag{2.114}$$

Under the assumptions made on β and σ, V_c belongs to $C^2((0,\infty))$ and it is strictly positive for all $y > 0$, $y \neq c$. Moreover V_c satisfies

$$\left(\beta(y)\partial_y + \frac{1}{2}\sigma^2(y)\partial_y^2\right) V_c(y) = 1, \tag{2.115}$$

for any $y \in (0,\infty)$. We consider the behavior of $V_c(y)$, around $y = 0$. For every $y \in (0,c)$, it holds:

$$2\int_y^c \frac{\beta(\tau)}{\sigma^2(\tau)} d\tau > \int_y^c \frac{\alpha'(\tau)}{\alpha(\tau)} d\tau = \log \frac{\alpha(c)}{\alpha(y)}.$$

Hence $p_c'(y) > \alpha(c)/\alpha(y)$, for any $0 < y < c$, and (2.99) implies $\beta(y) \leq C(1+y) < C(1+c) =: a$, on the interval $(0,c)$. We can write the following:

$$\begin{aligned}
V_c(y) &= \int_y^c p_c'(\xi) \int_\xi^c \frac{2}{p_c'(z)\sigma^2(z)} dz d\xi \\
&> \int_y^c \frac{\alpha(c)}{\alpha(\xi)} \int_\xi^c \frac{1}{\alpha(z)} \exp\left(-\int_z^c \frac{\beta(\tau)}{\alpha(\tau)} d\tau\right) dz d\xi \\
&> \int_y^c \frac{\alpha(c)}{\alpha(\xi)} \int_\xi^c \frac{1}{\alpha(z)} \exp\left(-a\int_z^c \frac{1}{\alpha(\tau)} d\tau\right) dz d\xi \\
&= \frac{\alpha(c)}{a} \int_y^c \frac{1}{\alpha(\xi)} \left[1 - \exp\left(-a\int_\xi^c \frac{1}{\alpha(\tau)} d\tau\right)\right] d\xi \\
&= \frac{\alpha(c)}{a} \int_y^c \frac{1}{\alpha(\xi)} d\xi - \frac{\alpha(c)}{a^2} \left[1 - \exp\left(-a\int_y^c \frac{1}{\alpha(\tau)} d\tau\right)\right]
\end{aligned} \tag{2.116}$$

From Lagrange's mean value theorem, we deduce the inequality $0 < \alpha(y) \leq Hy$, for all $y \in (0,c)$, implying the divergence of the integrals in (2.116), as $y \to 0^+$. Thus we have proved that that $V_c(y) \to +\infty$, as $y \to 0^+$. Thus, we consider the function

$$V(S,y) = S(1+y) - \log(S) + S\log(S+1) + y + V_c(y). \tag{2.117}$$

Clearly $V \in C^2((0,\infty)^2)$, and it is easy to see that V is strictly positive on

$(0,\infty)^2$. Furthermore, by (2.99), we get

$$
\begin{aligned}
&LV(S,y) = \beta(y)(S+1) + \frac{1}{2}S^2 y\left[\frac{1}{S^2} + \frac{1}{S+1} + \frac{1}{(S+1)^2}\right] \\
&+2\rho\sigma(y)\sqrt{y} + LV_c(y) \\
&\leq C(1+y)(S+1) + \frac{y}{2}(2+S) + \sqrt{2C(1+y)y} + 1 \\
&\leq C(1+y)(S+1) + y + (1+y)S + \sqrt{2C}(1+y) + 1 \\
&\leq (C+1)S(1+y) + (C+\sqrt{2C}+1)(1+y) \\
&\leq A[1+S(1+y)+y] \leq A[1+y+MS(1+y) - M\log(S) + M] \\
&\leq A(M+1)[1+S(1+y) - \log(S) + y] < A(M+1)[1+V(S,y)],
\end{aligned}
$$

where $A := 1+C+\sqrt{2C}$, whenever $M \geq 1/(1-e^{-2})$. In light of Assumption 2.2.3, V represents a Lyapunov function for the operator L in (2.96). Given a function ϕ, such that $|\phi(S)| \leq C(1+S)$, for all $S > 0$ and a positive constant C, it suffices to consider the function $\varphi(S) = \log(S+C)$ in order to satisfy Assumption 2.2.4. for the final datum ϕ. Hence the associated Dirichlet problem satisfies the hypotheses of Theorem 2.2.2. It admits a unique continuous viscosity solution without imposing any boundary condition at $y=0$ and $S=0$. In order to prove the inequality (2.101), we observe that functions $\underline{u}(t,S,y) := -C(1+S)$ and $\overline{u}(t,S,y) := C(1+S)$ represent respectively a viscosity sub/super-solution to the Dirichlet problem associated to (2.96), with final datum ϕ, since $L\overline{u}(t,S,y) = L\underline{u}(t,S,y) = 0$, for all $S>0$, $y>0$, $t\in(0,T)$, using also (2.99). Then (2.101) easily follows by the comparison principle between viscosity sub/super-solutions proved in Theorem 4.2 [16].

□

Remark 2.3.6. In the proof of Theorem 2.3.3 we have used the scale measure for the volatility process, namely the function V_c. It can be proved that that if

$$\lim_{y\to 0^+} V_c(y) = \lim_{y\to+\infty} V_c(y) = +\infty,$$

then the process $(y_t)_t$ cannot hit the origin or explode in finite time.

Remark 2.3.7. It is clear from the proof of Theorem 2.3.3, that the assumption requiring that $\alpha(y) = 0$ if and only if $y = 0$, can be weakened allowing that $\alpha(y) = 0$ for some $y \geq \varepsilon$, keeping the Lipschitz continuity of σ around such points. Indeed, we cannot define the function V_c for all $y > 0$, however it suffices to replace such a function in (2.117) with $W_c \in C^2((0,\infty))$, defined as $W_c(y) := V_c(y)$, for $y \in (0,c)$ and $W_c(y)$ coinciding with a suitable, strictly positive, second order polynomial on $[c,\infty)$. In light of assumption (2.99), such a definition allows to obtain a new Lyapunov-type function for the operator (2.96).

2.4 Regularity of the Viscosity Solution

Here we investigate the regularity of the solution to the problem

$$\begin{cases} \partial_t u(t,S,y) + Lu(t,S,y) = 0 & \text{on } (0,T) \times (0,\infty)^2, \\ u(S,y,T) = \phi(S), & \text{on } (0,\infty)^2, \end{cases} \tag{2.118}$$

where L is the pricing operator associated to the stochastic volatility model (2.96). To this end we recall a well known result related to the initial-boundary value problem (written for t replaced by $T - t$).

$$\begin{cases} \partial_t v + Lv = 0 & \text{on } Q = (0,T) \times B, \\ v(T,x) = \phi(x) & \text{on } B, \\ v(t,x) = \psi(t,x) & \text{on } K, \end{cases} \tag{2.119}$$

where B is a bounded domain in $\mathbb{R}^d$, with C^2 boundary ∂B, $K = [0,T) \times \partial B$, and the differential operator L is defined in (2.60), the integral component being omitted. The coefficients σ and b are continuous functions in $\overline{B}$. Given a set A, $\mathbb{I}_A$ denotes the indicator function of A. In the following $(W_t)_{t\in[0,T]}$ is a d-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with filtration $(\mathcal{F}_t)_{t\in[0,T]}$.

Theorem 2.4.1. Let $a = \sigma\sigma^\top$ be uniformly positive definite (i.e. $\exists\ \lambda > 0$, such that $a(x) \geq \lambda I_d$, for all $x \in B$). Let a, b be uniformly Lipschitz continuous in $\overline{B}$, with ψ continuous in $\overline{K}$, ϕ continuous in $\overline{B}$ and $\psi(T,x) =$

$\phi(x)$, if $x \in \partial B$. Then there exists a unique solution $v \in C([0,T] \times \overline{B}) \cap C^{2,1}((0,T) \times B)$ to the initial-boundary value problem (2.119), given by

$$v(t,x) = \mathbb{E}\left[\psi(\tau, \xi_\tau^{t,x})\mathbb{I}_{\tau<T} + \phi(\xi_\tau^{t,x})\mathbb{I}_{\tau=T}\right], \tag{2.120}$$

where, for every $x \in \mathbb{R}^d$, $t \in [0,T]$, $(\xi_s^{t,x})_{s\in[0,T]}$ is the unique strong solution to the stochastic differential equation

$$d\xi_s^{t,x} = \overline{b}(\xi_s^{t,x})dt + \overline{\sigma}(\xi_s^{t,x})dW_t, \qquad \xi_t^{t,x} = x. \tag{2.121}$$

Here, the coefficients $\overline{b}$, $\overline{\sigma}$ are uniformly Lipschitz continuous extensions of b and σ to $\mathbb{R}^d$, and τ represents the first time in $[t,T)$ that $\xi^{x,t}$ leaves B, if such a time exists and $\tau = T$ otherwise.

The proof can be found in [37] (Theorem 5.2).

Theorem 2.4.2. Under the assumptions of Theorem 2.3.3, the unique viscosity solution to (2.118) satisfying (2.101) belongs to $C^{1,2}((0,T)\times(0,\infty)^2)$.

Proof of Theorem 2.4.2. In light of Theorem 2.3.3, we consider the unique viscosity solution u to problem (2.118). We consider a disk of radius $r > 0$, B_r, with $\overline{B}_r \subset (0,\infty)^2$. Let us consider the initial-boundary value problem (2.119), where $B = B_r$, with $v(T,\cdot,\cdot) = \phi(\cdot)$ (the payoff function) on B_r, $\psi \equiv u$ on $K = [0,T) \times \partial B_r$. From the assumption that $\sigma(y) = 0$ only if $y = 0$ and by $\rho \in (-1,1)$, we deduce that the diffusion matrix of (2.96)

$$a(S,y) = \begin{bmatrix} S^2 y & \rho\sigma(y)\sqrt{y}S \\ \rho\sigma(y)\sqrt{y}S & \sigma^2(y) \end{bmatrix}$$

is uniformly positive definite on any compact subset of $(0,\infty)^2$. Therefore the regularity assumptions on the coefficients β and σ allow to apply Theorem 2.4.1, arguing that $v \in C^{1,2}((0,T) \times B_r)$. Let us fix a point $(\bar{t}, \overline{S}, \overline{y}) \in (0,T) \times B_r$ and let $X^{\overline{x}} := (S^{\overline{S}}, y^{\overline{y}})$ be the unique stochastic process, starting at $\overline{x} := (\overline{S}, \overline{y})$ which is the solution to the martingale problem for $(L, \delta_{\overline{x}})$ given in Theorem 2.2.1. Thus, thanks to Theorem 2.2.2, the process

$$\{u(\bar{t} + t \wedge \tau, X_{t\wedge\tau}^{\overline{x}})\}_{0\le t\le T-\bar{t}},$$

is an $\{\mathcal{F}_t^{X^{\overline{x}}}\}$-martingale, for every $\{\mathcal{F}_t^{X^{\overline{x}}}\}$-stopping time τ. On the other hand, given the stopping time τ_r defined as the first exit time of $X^{\overline{x}}$ from B_r, if $\tau_r < T - \bar{t}$, and $\tau_r = T - \bar{t}$ otherwise, since $X^{\overline{x}}$ is the solution to the martingale problem for L, also the process

$$v(\bar{t} + t \wedge \tau_r, X^{\overline{x}}_{t\wedge\tau_r}) - \int_0^{t\wedge\tau_r} (\partial_s v + Lv)(\bar{t} + s, X^{\overline{x}}_s) ds = v(\bar{t} + t \wedge \tau_r, X^{\overline{x}}_{t\wedge\tau_r})$$

is an $\{\mathcal{F}_t^{X^{\overline{x}}}\}_t$-martingale. Therefore, it is easy to see that

$$u(\bar{t}, \overline{S}, \overline{y}) \quad = \quad \mathbb{E}\left[u(\tau_r, X^{\overline{x}}_\tau)\right] = \mathbb{E}\left[v(\tau_r, X^{\overline{x}}_\tau)\right] = v(\bar{t}, \overline{S}, \overline{y}).$$

Hence the viscosity solution u coincides with v in B_r, implying that $u \in C^{1,2}((0,T) \times (0,\infty)^2)$. □

Remark 2.4.3. By the reasoning adopted in the proof of Theorem 2.4.2, under the assumptions described in Remark 2.3.6, the regularity of the viscosity solution to problem (2.118) can be proved in the neighborhood of any point $(t, S, y) \in (0,T) \times (0,\infty)^2$ where $\sigma(y) > 0$.

2.5 Jump Diffusion Model

An empirical motivation [67] for using jump-diffusion models comes from the fact that asset return distributions tend to have heavier tails than those of normal distribution. However, it is not clear how heavy the tail distributions are, as some people favor power-type distributions, others exponential-type distributions. The two basic building blocks of every jump-diffusion model are the Brownian motion (the diffusion part)and the Poisson process (the jump part). The Brownian motion is a familiar object to every option trader since the appearance of the Black-Scholes model, but a few words about the Poisson process are in order. Take a sequence $\{\tau_i\}_{i\leq 1}$ of independent exponential random variables with parameter λ, that is, with cumulative distribution function $\mathbb{P}(\tau_i \leq y) = e^{-\lambda y}$ and let $T_n = \sum_{i=1}^n \tau_i$. The process

$$N_t = \sum_{n\geq 1} 1_{t\leq T_n} \tag{2.122}$$

is called the Poisson process with parameter λ. For example, if the waiting times between buses at a bus stop are exponentially distributed, the total number of buses arrived up to time t is a Poisson process. The trajectories of a Poisson process are piecewise constant, with jumps of size 1 only. The jumps occur at times T_i and the intervals between jumps (the waiting times) are exponentially distributed. The Poisson process shares with the Brownian motion the very important property of independence and stationarity of increments, that is, for every $t > s$ the increment $N_t - N_s$ is independent from the history of the process up to time s and has the same law as N_{t-s}. The processes with independent and stationary increments are called Levy processes after the French mathematician Paul Levy. For financial applications, it is of little interest to have a process with a single possible jump size. The compound Poisson process is a generalization where the waiting times between jumps are exponential but the jump sizes can have an arbitrary distribution. More precisely, let N be a Poisson process with parameter λ and $\{Y_i\}_{i\geq 1}$ be a sequence of independent random variables with law f. The process

$$X_t = \sum_{i=1}^{N_t} Y_i \tag{2.123}$$

is called compound Poisson process. Combining a Brownian motion with drift and a compound Poisson process, we obtain the simplest case of a jump diffusion process which sometimes jumps and has a continuous but random evolution between the jump times

$$X_t = \mu_t + \sigma B_t + \sum_{i=1}^{N_t} Y_i, \tag{2.124}$$

The Jump Diffusion Model (JDM) was introduced first by Merton in [57], where the stock price is $S_t = S_0 e^{X_t}$ with X_t as above and the jumps Y_i have Gaussian distribution. Several models combining jumps and stochastic volatility appeared in the literature. In the Bates [6] model, one of the most popular examples of the class, an independent jump component is added to the Heston stochastic volatility model:

$$\begin{cases} dX_t = \mu dt + \sqrt{V_t} dW_t^1 + dZ_t, & S_t = S_0 e^{X_t} \\ dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^2 & d\left\langle W^1, W^2 \right\rangle_t = \rho dt. \end{cases} \tag{2.125}$$

Here Z is a compound Poisson process with Gaussian jumps. If J denotes the jump size then $\ln(1+J) \sim N(\ln(1+\bar{k}) - \frac{1}{2}\delta^2, \delta^2)$ for some $\bar{k}$. Under the risk-neutral probability one obtains the equation for the logarithm of the asset price:

$$dX_t = (r - \lambda\bar{k} - \frac{1}{2}v_t)dt + \sqrt{v_t} dW_t + \tilde{Z}_t, \tag{2.126}$$

where $\tilde{Z}_t$ is a compound Poisson process with normal distribution of jump magnitudes. Since the jumps are independent of the diffusion part, the characteristic function for the log-price process can be obtained as:

$$\phi_{X_t}(z) = \phi_{X_t}^D(z)\phi_{X_t}^J(z), \tag{2.127}$$

where:

$$\phi_{X_t}^D(z) = \frac{\exp\left\{\frac{\kappa\theta t(\kappa - i\rho\sigma z)}{\sigma^2 z} + izt(r - \lambda\bar{k}) + izx_0\right\}}{(\cosh\frac{\gamma t}{2} + \frac{(\kappa - i\rho\sigma z)}{\gamma}\sinh\frac{\gamma t}{2})^{\frac{2\kappa\theta}{\sigma^2}}} \exp\left\{-\frac{(z^2 + iz)v_0}{\gamma\coth\frac{\gamma t}{2} + \kappa - i\rho\sigma z}\right\} \tag{2.128}$$

is the diffusion part and

$$\phi_{X_t}^J(z) = \exp\{t\lambda(e^{-\delta^2 z^2/2 + i(\ln(1+\bar{k}) - \frac{1}{2}\delta^2)z} - 1)\} \tag{2.129}$$

is the jump part.

2.6 Local Stochastic Volatility Model

Recently a new model, generalization of the two previous ones, has been proposed: the "'Local-Stochastic Volatility Model"' (LSV). The need for better models became apparent eventually, given the observed drawbacks for both Local Volatility Models and Stochastic Volatility Models in conjunction with various asset classes and/or various financial instruments. A purely stochastic volatility model generates the same smile irrespective of

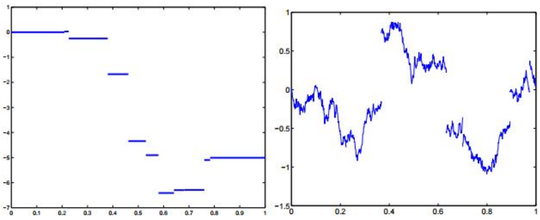


Figure 2.4: Left: Sample path of a compound Poisson process with Gaussian distribution of jump sizes. Right: sample path of a jump-diffusion process (Brownian motion + compound Poisson).

the initial level of spot, and therefore is a ”‘sticky-delta”’ method- the smile stays anchored at points corresponding to the specified deltas, while a local volatility model parametrized by a local function clearly depends on the spot level (and its initial level), and it is therefore ”‘sticky-strike”’. Consequently local stochastic volatility model were introduced in the literature to combine the best characteristic of both models, while minimizing downside. To understand why there was a need for LSV model, we list here below the advantages and disadvantages of both local volatility and stochastic volatility models.

1. Local Volatility Models
   - Advantages
     - Consistent with today’s market price by construction
     - Calibration may not require numerical optimization
   - Disadvantages
     - tend to replicate rather poorly the characteristics of market dynamics for spot and volatility (implied volatility tends to move too much given a change in the spot, no mean-reversion effect)
     - Impossible to tune-up the volatility of implied volatility, as there is simply no parameter for that

  - Forward volatility implied by Local Volatility Models is not realistic, since it flattens out
  - Changes in the underlying imply a general parallel shift of the smile(approximately) under Local Volatility Models, while market experience indicates that often smiles are "'sticky"' and remain invariant under many types of changes

2. Stochastic Volatility Models

- Advantages
  - Tend to be more in line with market dynamics
  - Equipped to model the term-structure (through mean-reversion parameters) and the volatility of the variance (through vol-of-vol parameters)
  - Forward volatility implied by Stochastic Volatility Model has a much more realistic behavior
- Disadvantages
- Calibration is done by Least Squares Optimization, and requires special attention to ensure stability of parameters
- any change in either mean-reversion or vol-of-vol requires re-calibration of other parameters
- Usually does not fit well short term market skew/smile

Local Stochastic Volatility Models aim to incorporate the advantages (and eliminate disadvantages) of these models.

These models consider volatility as the product between a deterministic and a stochastic term. In this way, using an hybrid local-stochastic volatility, it is possible to take the advantages of both the two basic models, which, in fact, can be considered as special cases of this generalized model. It is assumed, therefore, the following stochastic dynamics for the evolution of the underlying asset S_t:

$$\begin{cases} S_t = rS_t dt + b(V_t)\sigma_{LSV}(S_t)S_t dW_t \\ dV_t = a(V_t)dt + c(V_t)dZ_t \\ dW_t dZt = \rho dt \end{cases} \tag{2.130}$$

Example 2.6.1. Described in [54] and [55], this model is representative for a class of Local Stochastic Volatility Model which combines a mean reverting process for the volatility or variance (like in Heston Model), with a general local volatility function acting as multiplication factor for the stochastic volatility. It also incorporates jumps and is formulated as follows:

$$\begin{cases} dS_t = rS_t dt + \sigma_l(S_t)S_t\sqrt{V_t}dW_t + (\exp(j) - 1)dN_t \\ dV_t = \kappa(\theta - V_t)dt + \gamma\sqrt{V_t}dZ_t \\ dW_t dZt = \rho dt \end{cases} \tag{2.131}$$

with N_T is a Poisson processs and W_t, Z_t are Brownian Processes correlated by ρ; V_t represents the stochastic variance and σ_l denotes a local volatility function. We note that the model is the same as in the Heston model plus jumps when $\sigma_l \equiv 1$, and the same as in Dupire's model, when $\gamma = 0$.

Chapter 3

Weighted Average Price

The motivation for the introduction of this new point of view in financial modeling is mainly based on the fact that the price of a call option obtained in the framework of a stochastic volatility model dependes on the value v_0, the initial volatility, that unfortunately acts like an hidden stochastic variable. The most simple approach adopted to resolve the estimation of this hidden variable, is to consider v_0 as an additional parameter in the calibration procedure. In fact, in valuing financial derivatives, the no-arbitrage price of a European-type derivative can be found by a representation formula, where the price is given as a conditional expectation under a *risk-neutral* probability measure. In our approach, we extend that framework, by proving a new version of the fundamental theorem of asset pricing (Harrison Kreps (1979)) for processes depending on random parameters.

This allows to state a no arbitrage pricing formula similar to the classical one, without conflict with classical theory.

Although the application of this new arbitrage context is related to a specific stochastic volatility model, the theoretical results presented, exhibit a much more general significance in financial modeling.

In this chapter we refer closely to the paper presented by Papi, Pontecorvi and Donatucci (2014) for a new approach in the calibration of stochastic volatility models to vanilla options.

3.1 Financial Market Modeling with Random Parameters

We propose a new mathematical framework to price financial instruments derivatives, where the underlying stochastic model depends on some random parameters governed by a pre-defined probability law. For a fixed parameter value, we consider a stochastic Ito process driven by a Brownian motion on the same probability space for the assets dynamics in the market. We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, where we assume that a k-dimensional Brownian motion $(W(t))_t$ is assigned, in order to represent market shocks.
Let $q_0 : (\Omega_0, \mathcal{F}_0) \to (\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$, where $(\Omega_0, \mathcal{F}_0, f_0)$ is a complete probability space; its probability law m_0 is defined as follows:

$$m_0(B) = f_0(\{\omega_0 \in \Omega_0 : q_0(\omega_0) \in B\}) \quad \forall\, B \in \mathcal{B}(\mathbb{R}^p). \tag{3.1}$$

Finally, consider $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), m_0)$ like a complete probability space for the values of random model parameters here described by the outcomes of the random variable q_0. Let $(\tilde{\Omega}_0, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t\in[0,T]}, \tilde{\mathbb{P}}_0)$ a product probability space defined as:

1. $\tilde{\Omega}_0 := \Omega \times \mathbb{R}^p$;
2. $\tilde{\mathcal{F}}_t := \mathcal{F}_t \otimes \mathcal{B}(\mathbb{R}^p)$;
3. $\tilde{\mathbb{P}}_0 := \mathbb{P} \otimes m_0$.

An element of the product space will be denoted as $\tilde{\Omega}_0 \ni \tilde{\omega} = (\omega, q)$. Moreover, we shall use $\mathbb{E}[\cdot]$ to denote the expected value referred to the usual measure $\mathbb{P}$ in the space Ω, meanwhile we shall adopt the notation $\mathbb{E}^{\tilde{\mathbb{P}}_0}[\cdot]$ for the expectation according to the product measure $\tilde{\mathbb{P}}_0$. We consider a k-dimensional market with a set of risky no-dividend-paying assets, with price at time t given by $S_i(t, q)$, $i = 1, \dots k$, and a riskless asset, usually a zero coupon bond, with price at time t given by $S_0(t)$. The price of these asset verify the following q-depending system of stochastic differential equations, for every fixed $q \in \mathbb{R}^p$ (we will omit here and throughout all the chapter the dependence on ω, pointing out only the dependence on q):

$$\begin{cases} dS_i(t,q) = \mu_i(t,q)dt + \sum_{j=1}^{k} \sigma_{i,j}(t,q)dW_j(t), \\ S_i(0,q) > 0, \text{ for } t \in [0,T], \end{cases} \tag{3.2}$$

and

$$\begin{cases} dS_0(t) = r_t S_0(t)dt, \\ S_0(0) = s_0 > 0 \quad \text{for } t \in [0,T]. \end{cases} \tag{3.3}$$

The initial price vector $S(0,q) \in \mathbb{R}^k$ may depend on q, but it is assumed to be independent of $\omega \in \Omega$. The parameter q allows to describe latent factors that cannot be directly observed on the market even if they affect prices. Moreover, it may reflect the uncertainty in the specification of the adopted to describe the evolution of prices.

In such a formulation we can easily find out stochastic volatility models, where the initial volatility is not available, see Example 3.2.2 below. This kind of representation allows also to include the uncertainty in the specification of model parameters. Let $\tilde{\mathcal{M}}^{p,k}[0,T]$, $\tilde{\mathcal{M}}^{p,k\times h}[0,T]$, be the analogous of the spaces of $\mathcal{M}^{p,k}[0,T]$, $\mathcal{M}^{p,n\times m}[0,T]$ in the product space $\tilde{\Omega}_0$, endowed with the filtration $(\tilde{\mathcal{F}}_t)_t$ and under the product measure $\tilde{\mathbb{P}}_0$, for any $p \geq 1$, $k,\ h \in \mathbb{N},\ k,\ h \geq 1$. A similar notation will be used for the space $\Lambda^{p,k}[0,T]$. Thus, we consider also the following standing assumption.

Assumption 3.1.1. $\mu : [0,T] \times \Omega \times \mathbb{R}^p \to \mathbb{R}^k$, $\mu \in \tilde{\mathcal{M}}^{1,k}[0,T]$, $\sigma : [0,T] \times \Omega \times \mathbb{R}^p \to \mathbb{R}^{k\times k}$, $\sigma \in \tilde{\mathcal{M}}^{2,k\times k}[0,T]$ and $r \in \mathcal{M}^1[0,T]$.

Without loss of generality, we can assume that $s_0 = 1$ so that $S_0(t) = \exp(\int_0^t r_u du)$. We also add $r_t S_0(t)$ as the first component in the vector $\mu(t,q)$ and a row of zeros as the first row in the matrix $\sigma(t,q)$, without changing the notation. Thefore μ and σ become $\mathbb{R}^{k+1}$-valued and $\mathbb{R}^{k+1\times k}$-valued processes, respectively.

Definition 3.1.1. A process $\theta : [0,T] \times \Omega \times \mathbb{R}^p \to \mathbb{R}^{k+1}$ is a self-financing portfolio if $(\theta_t)_{t\in[0,T]}$ is $\{\tilde{F}_t\}_{t\in[0,T]}$-progressively measurable, $\langle \theta(t,q), \mu(t,q)\rangle \in$

$\mathcal{M}^1[0,T]$, $\langle\theta(t,q),\sigma_{\cdot,j}(t,q)\rangle \in \mathcal{M}^2[0,T]$, for $j=1,\dots,k$ and $q\in\mathbb{R}^p$, m_0 a.s. and

$$\langle\theta(t,q),S(t,q)\rangle = \langle\theta(0,q),S(0,q)\rangle + \int_0^t \theta(q,s)dS(s,q), \tag{3.4}$$

with $\langle\theta(t,q),S(t,q)\rangle \geq \kappa_\theta$, for all $0\leq t\leq T$, $\tilde{\mathbb{P}}_0$ a.s, for a constant κ_θ depending only by θ.

A portfolio is self-financing if there is no exogenous infusion or withdrawal of money, all trades are financed by selling or buying assets in the portfolio. The boundedness from below required in the definition of self-financing portfolio should be interpreted as a reasonable condition in order to avoid excessive losses in the portfolio value. We will indicate as $\Theta(S)$ the set of self-financing portfolios for the market defined by S. Moreover, a fundamental notion in mathematical finance is the following.

Definition 3.1.2. A portfolio $\theta\in\Theta(S)$ is an *arbitrage* if

$$\langle\theta(0,q),S(0,q)\rangle \leq 0 \leq \langle\theta(T,q),S(T,q)\rangle,\ \ \tilde{\mathbb{P}}_0 \text{ a.s.} \tag{3.5}$$

and

$$\tilde{\mathbb{P}}_0(\langle\theta(T,q),S(T,q)\rangle > 0) > 0. \tag{3.6}$$

Let's define the discounted price process $\tilde{S}(t,q) := Y(t)S(t,q)$, where

$$Y(t) := [S_0(t)]^{-1} = \exp(-\int_0^t r_u du) \tag{3.7}$$

so as to have $\tilde{S}_0(t,q)=1$. Moreover, let $\overline{S} := (\tilde{S}_1,\dots\tilde{S}_k)$ and

$$\overline{\mu}(t,q) = Y(t)S_i(t,q)[\mu_i(t,q)-r_t], \qquad (\overline{\sigma}(t,q))_{i,j} = Y(t)S_i(t,q)\sigma_{i,j}(t,q), \tag{3.8}$$

for all $i,j=1,\dots,k$. It is easy to see that there is no arbitrage in the market defined by S if and only if there is no arbitrage in the market represented by the discounted price process $\overline{S}$. Thus, in absence of arbitrage opportunities, a proportional relationship between the average rate of securities prices variation and the risk related to stocks volatility holds true. In particular, we derive an extension to our market depending on the random parameter

q, of a well known result obtained in the literature by Harrison and Pliska (*Stochastic Proc. and Applications*, (1981)-(1983)) if the market is free of arbitrage, see also [**?**].

Remark 3.1.1. We observe that, in absence of the random parameter q, our definition of arbitrage portfolio coincides with the original given in Definition 1.4.3. Furthermore, if θ is an arbitrage in the sense of Definition 3.1.2, then there exists a Borel set $B_0 \subset \mathbb{R}^p$, with $m_0(B_0) > 0$, such that for every $q_0 \in B_0$, $(\theta(t, q_0))_t$ represents a trading strategy for the market given by (3.2), for $q = q_0$. In other words, there are multiple trading choises that lead to a potential profit for arbitrage, in the classical sense, as in [26].

Theorem 3.1.2. Let the financial market represented by S be arbitrage free. Then, there exists a $\mathbb{R}^d$-valued progressively meadurable process λ in the product space, such that for $\tilde{\omega} \in \tilde{\Omega}_0$, $\tilde{\mathbb{P}}_0$ almost surely, it satisfies

$$\overline{\sigma}(t, \tilde{\omega})\lambda(t, \tilde{\omega}) = \overline{\mu}(t, \tilde{\omega}), \tag{3.9}$$

for $t \in [0, T]$ almost everywhere (a.e.), in the sense of Lebesgue measure on $\mathbb{R}$.

The process λ usually identifies the so called market price of risk, see [**?**]. If $\overline{\sigma}(t, \tilde{\omega})$ is invertible, then λ is uniquely defined by $\lambda(t, \tilde{\omega}) = \overline{\sigma}(t, \tilde{\omega})^{-1}\overline{\mu}(t, \tilde{\omega})$. In order to prove Theorem 3.1.2, we need the following technical results. For the sake of completeness we present their proof in this section, even if the details can be found in [**?**].

Lemma 3.1.1. The following maps:

$$(x, \sigma) \mapsto \mathrm{proj}_{\mathrm{Ker}(\sigma)}(x), \tag{3.10}$$

$$(x, \sigma) \mapsto \mathrm{proj}_{\mathrm{Ker}(\sigma)^{\perp}}(x), \tag{3.11}$$

defined on $\mathbb{R}^d \times M_{k,d}(\mathbb{R})$, and

$$(y, \sigma) \mapsto \mathrm{proj}_{\mathrm{Ker}(\sigma^{\top})}(y), \tag{3.12}$$

$$(y,\sigma) \mapsto \mathrm{proj}_{\mathrm{Ker}(\sigma\top)^{\perp}}(y), \tag{3.13}$$

defined on $\mathbb{R}^k \times M_{k,d}(\mathbb{R})$, are Borel measurable functions.

Here proj_G, $G^{\perp}$ denote the projection on the linear subspace G and the orthogonal complement of the subspace G, respectively.

Corollary 3.1.1. The process $\mathrm{proj}_{\mathrm{Ker}(\overline{\sigma}^{\top}(t,\cdot))}[\overline{\mu}(t,\cdot)]$, $0 \le t \le T$, is progressively measurable in the product space $\tilde{\Omega}_0$, under the extended filtration $\{\tilde{\mathcal{F}}_t\}_t$.

Lemma 3.1.2. In absence of arbitrage in the market given by $\overline{S}$, for $\tilde{\omega} \in \tilde{\Omega}_0$, $\tilde{\mathbb{P}}_0$ almost surely, it holds $\overline{\mu}(t,\tilde{\omega}) \in \mathrm{Im}(\overline{\sigma}(t,\tilde{\omega})$, for $t \in [0,T]$ almost everywhere.

Lemma 3.1.3. Let

$$\Psi_1 : \{(y,\sigma) \in \mathbb{R}^k \times M_{k,d}(\mathbb{R}),\, y \in \mathrm{Im}(\sigma)\} \to \mathbb{R}^d \tag{3.14}$$

where $\Psi_1(y,\sigma)$ is the unique $\xi \in \mathrm{Ker}(\sigma)^{\perp}$ such that $\sigma\xi = y$, and

$$\Psi_2 : \{(x,\sigma) \in \mathbb{R}^d \times M_{k,d}(\mathbb{R}),\, y \in \mathrm{Im}(\sigma^{\top})\} \to \mathbb{R}^d \tag{3.15}$$

where $\Psi_2(x,\sigma)$ is the unique $\eta \in \mathrm{Ker}(\sigma^{\top})^{\perp}$ such that $\sigma^{\top}\eta = y$. Both of such maps are Borel measurable functions.

Proof of Theorem 3.1.2. In accordance with Lemma 3.1.2 and Lemma 3.1.3, the following progressively measurable process:

$$\lambda(t,\tilde{\omega}) := \Psi_1(\overline{\mu}(t,\tilde{\omega}), \overline{\sigma}(t,\tilde{\omega})), \tag{3.16}$$

is well-defined and for $\tilde{\omega} \in \tilde{\Omega}_0$, $\tilde{\mathbb{P}}_0$ almost surely, it satisfies satisfies (3.9) for $t \in [0,T]$ almost everywhere. □

Proof of Lemma 3.1.1. It suffices to prove the result for the first map. The space $M_{k,d}(\mathbb{R})$ is naturally equipped with the Borel σ-algebra induced by the topology associated to the operator norm. $\mathbb{R}^d$ and $\mathbb{R}^k$ are naturally equipped with Borel σ-algebras as well and the product space is equipped with the product σ-algebra. Finally, let $\mathbb{Q}^k$ be the countable dense subset of rational vectors into $\mathbb{R}^k$. Define the following Borel measurable function $F : \mathbb{R}^d \times M_{k,d}(\mathbb{R}) \to \mathbb{R}$ where

$$F(z,\sigma) := \inf_{q\in\mathbb{Q}^k} ||z - \sigma^\top q|| \qquad \forall z \in \mathbb{R}^d,\ \sigma \in M_{k,d}(\mathbb{R}), \tag{3.17}$$

thus

$$\{(z,\sigma) : z \in \mathrm{Im}(\sigma^\top)\} \subset \{(z,\sigma) : F(z,\sigma) = 0\}.$$

On the other hand, if $F(z,\sigma) = 0$, then there exists a sequence $\{q_n\}_n \subset \mathbb{Q}^k$ such that $\lim_{n\to\infty} ||z - \sigma^\top q_n|| = 0$. Since $\mathbb{R}^k = \mathrm{Ker}(\sigma^\top) \oplus \mathrm{Ker}(\sigma^\top)^\perp$, we can decompose every q_n as $q_n = q_{1,n} + q_{2,n}$, where $q_{1,n} \in \mathrm{Ker}(\sigma^\top)$ and $q_{2,n} \in \mathrm{Ker}(\sigma^\top)^\perp$. Obviously, restricted on $\mathrm{Ker}(\sigma^\top)^\perp$, the linear function defined by $\sigma^\top$ is invertible.
Since $\sigma^\top q_{2,n} \to z$, as $n \to \infty$, the sequence $\{q_{2,n}\}_n$ converges to some $q_2 \in \mathrm{Ker}(\sigma^\top)^\perp$ that satisfies $\sigma^\top q_2 = z$. Therefore we deduce that $z \in \mathrm{Im}(\sigma^\top)$. Hence we have proved

$$\{(z,\sigma) : z \in \mathrm{Im}(\sigma^\top)\} = \{(z,\sigma) : F(z,\sigma) = 0\}. \tag{3.18}$$

In particular $\{(z,\sigma) : z \in \mathrm{Im}(\sigma^\top)\}$ is a Borel set. Consequently

$$\begin{aligned} &\{(x,\sigma,\xi) \in \mathbb{R}^d \times M_{k,d}(\mathbb{R}) \times \mathbb{R}^d : \xi = \mathrm{proj}_{\mathrm{Ker}(\sigma)}(x)\} = \\ &= \{(x,\sigma,\xi) : \xi \in \mathrm{Ker}(\sigma),\ x - \xi \in \mathrm{Ker}(\sigma)^\perp\} = \\ &= \{(x,\sigma,\xi) : \xi^\top\sigma^\top = 0,\ x - \xi \in \mathrm{Im}(\sigma^\top)\} \end{aligned} \tag{3.19}$$

is a Borel set too. Define $Q : \mathbb{R}^d \times M_{k,d}(\mathbb{R}) \to \mathbb{R}^d$ such that

$$Q(x,\sigma) := \mathrm{proj}_{\mathrm{Ker}(\sigma)}(x), \qquad \forall\, x \in \mathbb{R}^d,\ \sigma \in M_{k,d}(\mathbb{R}). \tag{3.20}$$

The set in (3.19) is the graph of Q, that is

$$\mathrm{Gr}(Q) := \{(x,\sigma,\xi) : (x,\sigma) \in \mathbb{R}^d \times M_{k,d}(\mathbb{R}),\ \xi = Q(x,\sigma)\}. \tag{3.21}$$

Since $\mathrm{Gr}(Q)$ is a Borel set, Q must be a measurable Borel function. Indeed, for all $B \in \mathcal{B}(\mathbb{R}^d)$, we have

$$\{(x,\sigma) : Q(x,\sigma) \in B\} = \mathrm{proj}_{\mathbb{R}^d \times M_{k,d}(\mathbb{R})} \mathrm{Gr}(Q) \cap (\mathbb{R}^d \times M_{k,d}(\mathbb{R}) \times B) \quad (3.22)$$

and a projection of a Borel set is a Borel set. □

Proof of Lemma 3.1.2. Define, for $0 \le t \le T$, $\tilde{\omega} = (\omega, q) \in \tilde{\Omega}_0$ the following maps:

$$p(t,q) := \mathrm{proj}_{\mathrm{Ker}(\overline{\sigma}(t,q))}[\overline{\mu}(t,q)], \quad (3.23)$$

and

$$\overline{\theta}(t,q) := \begin{cases} \frac{p(t,q)}{|p(t,q)|} & \text{if } p(t,q) \neq 0, \\ 0 & \text{if } p(t,q) = 0. \end{cases} \quad (3.24)$$

We also consider

$$\overline{\theta}_0(t,q) := -\langle \overline{\theta}(t,q), \overline{S}(t,q) \rangle + \int_0^t \overline{\theta}(s,q) d\overline{S}(s,q). \quad (3.25)$$

Then the process $\theta(t,q) := (\overline{\theta}_0(t,q), \overline{\theta}(t,q))$ is self-financing. In fact

$$\langle \theta(t,q), Y(t)S(t,q) \rangle = \overline{\theta}_0 + \langle \overline{\theta}(t,q), \overline{S}(t,q) \rangle = \int_0^t \overline{\theta}(s,q) d\overline{S}(s,q), \quad (3.26)$$

and

$$\langle \theta(0,q), Y(0)S(0,q) \rangle = -\langle \overline{\theta}(0,q), \overline{S}(0,q) \rangle + \langle \overline{\theta}(0,q), \overline{S}(0,q) \rangle = 0. \quad (3.27)$$

Hence

$$\begin{aligned}
\langle \theta(t,q), Y(t)S(t,q) \rangle &= \int_0^t \langle \theta(s,q), \overline{\mu}(s,q) \rangle ds + \int_0^t \langle \theta^\top(s,q) \overline{\sigma}(s,q) \rangle dW(s) \\
&= \int_0^t \langle \overline{\theta}(s,q), \overline{\mu}(s,q) \rangle ds + \int_0^t \langle \overline{\theta}^\top(s,q) \overline{\sigma}(s,q) \rangle dW(s) \\
&= \int_0^t \langle \overline{\theta}(s,q), \overline{\mu}(s,q) \rangle ds \\
&= \int_0^t \Big(\mathbb{1}_{\{p(s,q) \neq 0\}} \langle \frac{p(s,q)}{|p(s,q)|}, p(s,q) + \\
&\quad + \mathrm{proj}_{\mathrm{Ker}(\overline{\sigma}(s,q))^\perp}(\overline{\mu}(s,q)) \rangle \Big) \\
&= \int_0^t \mathbb{1}_{\{p(s,q) \neq 0\}} |p(s,q)| ds \ge, 0 \ \text{ for all } t \in [0,T].
\end{aligned}$$

Since $\langle \theta(T,q), S(T,q)\rangle \geq 0$ and we are assuming a free-arbitrage market, it follows $\langle \overline{\theta}(T,q), \overline{S}(T,q)\rangle = 0$ $\tilde{\mathbb{P}}_0$ a.s., implying for $\tilde{\omega} = (\omega, q) \in \tilde{\Omega}_0$ $\tilde{\mathbb{P}}_0$ a.s, $p(t,q) = 0$ for $t \in [0,T]$ a.e and the result is proved. $\square$

Proof of Lemma 3.1.3. We prove the result only for Ψ_1 (for Ψ_2 the proof is similar). Let:

$$\Delta := \{(y,\sigma,\xi) \in \mathbb{R}^k \times M_{k,d}(\mathbb{R}) \times \mathbb{R}^d : y \in \mathrm{Im}(\sigma), \xi \in \mathrm{Im}(\sigma^\top), \sigma\xi = y\} \quad (3.28)$$

In the proof of Lemma 3.1.1 we have seen that $\{(\sigma,\xi) : \xi \in \mathrm{Im}(\sigma^\top)\}$ is a Borel set. The same argument is valid for $\{(y,\sigma) : y \in \mathrm{Im}(\sigma)\}$.
Thus, Δ is a Borel set, but Δ is the graph of Ψ_1. Then, by Lemma 3.1.1, we deduce that Ψ_1 is a Borel measurable function. $\square$

3.2 q-Depending Risk-Neutral Measure

Definition 3.2.1. A risk-neutral measure (also called equivalent martingale measure) for the market S is a probability measure $\tilde{\mathbb{Q}}_0$ is a probability measure $\tilde{\mathbb{Q}}_0$ on the product space $(\tilde{\Omega_0}, \tilde{\mathcal{F}}_T)$, equivalent to $\tilde{\mathbb{P}}_0$ such that

$$\mathbb{E}^{\tilde{\mathbb{Q}}_0}[\overline{S}(s,q)|\tilde{\mathcal{F}}_t] = \overline{S}(t,q) \quad \forall\, t \leq s \leq T, \;\; t \geq 0, \quad (3.29)$$

$\tilde{\mathbb{P}}_0$ a.s., or equivalently

$$\mathbb{E}^{\tilde{\mathbb{Q}}_0}[S(s,q)|\tilde{\mathcal{F}}_t] = S(t,q)e^{-\int_0^t r_u du}, \quad (3.30)$$

$\tilde{\mathbb{P}}_0$ a.s.

The following results generalize to our setting some well known facts related to free arbritrage markets

Theorem 3.2.1. If S represents an arbitrage free market and there is a market price of risk λ satisfying (3.9) such that

$$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T |\lambda(s,q)|^2\, ds}\right] < \infty \quad \forall\, q \in\; m_0 - \text{a.s.} \quad (3.31)$$

then S admits a risk-neutral measure.

Theorem 3.2.2. If S admits a risk-neutral measure, then it describes an arbitrage free market.

In order to build a risk-neutral measure, we define a weighted version of the Girsanov's exponential [58]. Let us consider a product set $A \times B$, where $A \in \mathcal{F}_T$ and $B \in \mathcal{B}(\mathbb{R}^p)$, a martingale measure $\tilde{\mathbb{Q}}_0$ equivalent to $\tilde{\mathbb{P}}_0$ can be defined through its Radon-Nikodyn derivative, that is:

$$\frac{d\tilde{\mathbb{Q}}_0}{d\tilde{\mathbb{P}}_0}(\omega, q) = \delta_0^\lambda(\omega, q) := e^{-\int_0^T \lambda(s,\omega,q)dW(s) - \frac{1}{2}\int_0^T |\lambda(s,\omega,q)|^2 ds} \tag{3.32}$$

for any $(\omega, q) \in \tilde{\Omega}_0$, $\tilde{\mathbb{P}}_0$ almost surely. In fact, for every $q \in \mathbb{R}^p$, $\delta_0(\cdot, q)$ represents the usual Girsanov's exponential associated with $\lambda(\cdot, q)$. If condition (3.31) (called also Novikov's condition) then, by Girsanov's theorem, the following map is well defined:

$$\begin{array}{rcl} \mathbb{R}^p & \to & \{\text{Probability measures on } (\Omega, \mathcal{F}_T) \text{ equivalent to } \mathbb{P}\ \} \\ q & \mapsto & \mathbb{Q}^q, \qquad \mathbb{Q}^q(A) = \mathbb{E}[\delta_0^\lambda(\cdot, q) \cdot \mathbb{1}_A] \quad \forall A \in \mathcal{F}_T, \end{array}$$

Since for every $A \in \mathcal{F}_T$, $q \mapsto \mathbb{Q}^q(A)$ is also Borel measurable, we can define $\tilde{\mathbb{Q}}_0$ as follows:

$$\tilde{\mathbb{Q}}_0(A \times B) = \int_B \mathbb{Q}^q(A) dm_0(q). \tag{3.33}$$

It is now clear how $\tilde{\mathbb{Q}}_0$ can be extended to all $\tilde{\mathcal{F}}_T$-measurable sets and, of course, $\tilde{\mathbb{Q}}_0$ defines a probability measure on $(\tilde{\Omega}_0, \tilde{\mathcal{F}}_T)$ equivalent to $\tilde{\mathbb{P}}_0$.

For the proof of Theorem 3.2.1, we recall here the concept of π-system and Dinkin-system to exploit the Dynkin's π-λ Theorem [47].

Definition 3.2.2. Let Σ be a non-empty set, and let $\mathcal{D}$ be a collection of subsets of Σ. Then $\mathcal{D}$ is a *Dynkin system* if

1. $\Sigma \in \mathcal{D}$;

2. if $A \in \mathcal{D}$, then $A^c \in \mathcal{D}$;

3. if $A_1, A_2, A_3, \ldots$ is a sequence of subsets in $\mathcal{D}$ such that $A_i \cap A_j = \emptyset$ for all $i \neq j$, then $\displaystyle\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

Definition 3.2.3. Let Σ be a non-empty set, and let $\mathcal{P}$ be a collection of subsets of Σ. Then $\mathcal{P}$ is a *π-system* if

1. $\mathcal{P}$ is non-empty;

2. if A_1 and A_2 are subsets of $\mathcal{P}$ then $A_1 \cap A_2 \in \mathcal{P}$.

Theorem 3.2.3. [Dynkin's π-λ Theorem [47]] If $\mathcal{P}$ is a π-system and $\mathcal{D}$ is a Dynkin system of the same non-empty set Σ and $\mathcal{P} \subseteq \mathcal{D}$, then $\sigma(\mathcal{P}) \subseteq \mathcal{D}$. In othe words, the σ-algebra generated by $\mathcal{P}$ is contained in $\mathcal{D}$.

Proof of Theorem 3.2.1. It suffices to show that, for all $0 \leq t \leq s \leq T$ and for every $G \in \tilde{\mathcal{F}}_t$, it holds

$$\mathbb{E}^{\tilde{\mathbb{Q}}_0}[\overline{S}(s,q)\mathbb{1}_G] = \mathbb{E}^{\tilde{\mathbb{Q}}_0}[\overline{S}(t,q)\mathbb{1}_G]. \tag{3.34}$$

At first we prove that (3.34) holds true on the collection of subsets $\mathcal{I}_t \subset \tilde{\Omega}_0$ where:

$$\mathcal{I}_t := \{A \times B | A \in \mathcal{F}_t \text{ and } B \in \mathcal{B}(\mathbb{R}^p)\}. \tag{3.35}$$

Since the Novikov condition holds for the market price of risk, and by Theorem 3.1.2, we deduce that, for q m_0-a.s., $(\overline{S}(t,q))_{t\in[0,T]}$ is a martingale under the measure $\mathbb{Q}^q$, se also [26] and [48]. Thus, for $A \times B \in \mathcal{I}_t$, we have

$$\begin{aligned} \mathbb{E}^{\tilde{\mathbb{Q}}_0}[\overline{S}(s,q)\mathbb{1}_A(\omega)\mathbb{1}_B(q)] &= \mathbb{E}^{\tilde{\mathbb{P}}_0}[\delta_0(\omega,q)\overline{S}(s,q)\mathbb{1}_A(\omega)\mathbb{1}_B(q)] \\ &= \int_B \mathbb{E}^{\mathbb{Q}^q}[\overline{S}(s,q)\mathbb{1}_A(\omega)]dm_0(q) \\ &= \int_B \mathbb{E}^{\mathbb{Q}^q}[\overline{S}(t,q)\mathbb{1}_A(\omega)]dm_0(q) \\ &= \mathbb{E}^{\tilde{\mathbb{Q}}_0}[\overline{S}(t,q)\mathbb{1}_A(\omega)\mathbb{1}_B(q)]. \end{aligned} \tag{3.36}$$

We observe that closely $\mathcal{I}_t$ is a π-system on $\tilde{\Omega}_0$. Let

$$\tilde{\mathcal{G}}_t := \{G \in \tilde{\mathcal{F}}_t | G \text{ verify } (3.34)\}, \tag{3.37}$$

then it is easy to show that the first two point in the definition of Dynkin system 3.2.2 are verified. We prove here the last point of 3.2.2. Consider a sequence of subsets $A_1, A_2, A_3, \dots$ in $\mathcal{I}_t$ such that $A_i \cap A_j = \emptyset$, for all $i \neq j$, and define

$$G_n = \bigcup_{i=1}^{n} A_i \quad \text{and} \quad G = \bigcup_{i=1}^{\infty} A_i.$$

If $\tilde{\omega} = (\omega, q) \notin G$, then $\mathbb{1}_{G_n}(\tilde{\omega}) = \mathbb{1}_G(\tilde{\omega}) = 0$, for all $n \geq 1$. On the other hand, if $\tilde{\omega} = (\omega, q) \in G$, then:

$$\exists\, \bar{n} \geq 1 \text{ such that } \tilde{\omega} \in G_n, \quad \forall\, n > \bar{n} \text{ and } \mathbb{1}_{G_n}(\tilde{\omega}) = \mathbb{1}_G(\tilde{\omega}) = 1.$$

Hence, $\mathbb{1}_{G_n}$ converges pointwise to $\mathbb{1}_G(\tilde{\omega})$, as $n \to \infty$ and, consequently, we deduce that

$$\overline{S}(\tau, \cdot)\mathbb{1}_{G_n} \to \overline{S}(\tau, \cdot)\mathbb{1}_G, \quad \tilde{\mathbb{Q}}_0\text{-a.s.}, \tag{3.38}$$

as $n \to \infty$, for any $\tau \in [0, T]$. Hence, by (3.36), we have

$$\begin{aligned} \mathbb{E}^{\tilde{\mathbb{Q}}_0}[\overline{S}(s, q)\mathbb{1}_{G_n}] &= \sum_{i=1}^{n} \mathbb{E}^{\tilde{\mathbb{Q}}_0}[\overline{S}(s, q)\mathbb{1}_{A_i}] = \sum_{i=1}^{n} \mathbb{E}^{\tilde{\mathbb{Q}}_0}[\overline{S}(t, q)\mathbb{1}_{A_i}] \\ &= \mathbb{E}^{\tilde{\mathbb{Q}}_0}[\overline{S}(t, q)\mathbb{1}_{G_n}], \end{aligned} \tag{3.39}$$

and using the monotone convergence theorem on both sides (3.39), (3.34) is proved for G. Theorem 3.2.3 implies $\tilde{\mathcal{F}}_t = \sigma(\mathcal{I}_t) \subseteq \tilde{\mathcal{G}}_t \subseteq \tilde{\mathcal{F}}_t$. In particular, from $\tilde{\mathcal{G}}_t = \tilde{\mathcal{F}}_t$ for all $t \in [0, T]$, we deduce that $\overline{S}$ is a $\tilde{\mathbb{Q}}_0$-martingale. □

Proof of Theorem 3.2.2. Let $\tilde{Q}_0$ be an equivalent martingale measure for S. Since for every q, $(S(t, q))_t$ is an Ito process in the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_{t\in[0,T]}), \mathbb{P})$ and since $(W_t)_{t\in[0,T]}$ can be clearly seen as a k-dimensional Brownian motion in the filtered probability space $(\tilde{\Omega}_0, \tilde{\mathcal{F}}_T,$ $(\tilde{\mathcal{F}}_{t\in[0,T]}), \tilde{\mathbb{P}}_0)$ and the assumption on the coefficiente σ, we can apply the diffusion invariance principle (see [49]) to deduce that there exists a k-dimensional standard Brownian motion $(\tilde{W}_t)_{t\in[0,T]}$ under $\tilde{\mathbb{Q}}_0$, such that

$$S(s, \tilde{\omega}) = S(t, \tilde{\omega}) + \int_t^s \sigma(\xi, \tilde{\omega})d\tilde{W}_\xi(\tilde{\omega})$$

for $\tilde{\omega} \in \tilde{\Omega}_0$, $\tilde{\mathbb{P}}_0$ a.s, for all $0 \leq t \leq s \leq T$. Hence for every $\theta \in \Theta(S)$, we have

$$\langle \theta(t, q), S(t, q) \rangle = \langle \theta(0, q), S(0, q) \rangle + \int_0^t \theta(q, \xi)^\top \sigma(\xi, \tilde{\omega})d\tilde{W}_\xi(\tilde{\omega}), \tag{3.40}$$

$\tilde{\mathbb{P}}_0$-a.s., for all $0 \leq t \leq T$. In particular, the portfolio value process $(\langle \theta(t, \cdot), S(t, \cdot) \rangle)_{t \in [0,T]}$ is a local martingale bounded from below on $(\tilde{\Omega}_0, \tilde{\mathcal{F}}_T, \tilde{\mathbb{Q}}_0)$. Hence it is a super-martingale, implying the inequality

$$\mathbb{E}^{\tilde{\mathbb{Q}}_0} \left[\langle \theta(T, \cdot), S(T, \cdot) \rangle \middle| \tilde{\mathcal{F}}_0 \right] \leq \langle \theta(0, \cdot), S(0, \cdot) \rangle,$$

$\tilde{\mathbb{Q}}_0$-a.s. Now, if θ were an arbitrage strategy, it holds

$$\langle \theta(0, \cdot), S(0, \cdot) \rangle \leq 0 \leq \langle \theta(T, \cdot), S(T, \cdot) \rangle,$$

$\mathbb{Q}_0$-a.s. and there would be $A \in \tilde{\mathcal{F}}_T$, with $\tilde{\mathbb{Q}}_0(A) > 0$, such that $\langle \theta(T, \tilde{\omega}), S(T, \tilde{\omega}) \rangle > 0$, for all $\tilde{\omega} \in A$. From the law of iterated expectations, we obtain the following contradiction:

$$\begin{aligned} 0 &< \mathbb{E}^{\tilde{\mathbb{Q}}_0} \left[\langle \theta(T, \cdot), S(T, \cdot) \rangle 1_A \right] = \mathbb{E}^{\tilde{\mathbb{Q}}_0} \left[\mathbb{E}^{\tilde{\mathbb{Q}}_0} \left[\langle \theta(T, \cdot), S(T, \cdot) \rangle 1_A \middle| \tilde{\mathcal{F}}_0 \right] \right] \\ &\leq \mathbb{E}^{\tilde{\mathbb{Q}}_0} \left[\mathbb{E}^{\tilde{\mathbb{Q}}_0} \left[\langle \theta(T, \cdot), S(T, \cdot) \rangle \middle| \tilde{\mathcal{F}}_0 \right] \right] \leq \mathbb{E}^{\tilde{\mathbb{Q}}_0} \left[\langle \theta(0, \cdot), S(0, \cdot) \rangle \right] \leq 0. \end{aligned}$$

Thus the market represented by S is arbitrage free. □

We discuss the application of our modelling framework to two fundamental examples.

Example 3.2.1 (Black and Scholes Model)**.** Now we are going to review the Black and Scholes model with the latter topic. The model, as known, is

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t dW_t, \\ dB_t = r B_t dt \end{cases} \tag{3.41}$$

where $(W_t)_t$ is a one dimensional Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$. The deterministic constant parameters μ, σ and r represent respectively the local mean rate of return of the asset, the volatility of the asset and the short-term rate of interest. As has already been said, one of the mainstay of this model is a closed-form solution for the price of European put and call options while a weakness is based on the constant volatility assu,ption (that makes the model inconsistent with observed market prices). Following our theoretical approach we can consider

σ as a random variable endowed with a given probability distribution with density $f \in L^1(\mathbb{R};\mathbb{R}^+)$, such that $f(\sigma) > 0$, for all $\sigma > 0$, $f(\sigma) = 0$, for all $\sigma \leq 0$; hence $dm_0(\sigma) = f(\sigma)d\sigma$.

In the Black-Scholes model, for a given maturity of the option T, the market price of risk and the Radon-Nikodym derivative of the extended risk-neutral measure on the product space, take the following form:

$$\lambda_\sigma = \frac{\mu - r}{\sigma}, \qquad \frac{d\tilde{\mathbb{Q}}_0}{d\tilde{\mathbb{P}}_0}(\omega,\sigma) = e^{-\lambda_\sigma W_T(\omega) - \frac{1}{2}\lambda_\sigma^2 T}, \tag{3.42}$$

for any $\omega \in \Omega$, $\sigma > 0$. In fact, for every $\sigma > 0$, by Girsanow's theorem, we can define the measure $\mathbb{Q}^\sigma$ as the probability measure on $(\Omega, \mathcal{F}_T)$, equivalent to $\mathbb{P}$, with density $e^{-\lambda_\sigma W(T) - \frac{1}{2}\lambda_\sigma^2 T}$. Thus, it holds

$$\begin{aligned} \mathbb{E}^{\tilde{\mathbb{P}}_0}[e^{-\lambda_\sigma W_T - \frac{1}{2}\lambda_\sigma^2 T}] = \textstyle\int_0^\infty f(\sigma) \int_\Omega e^{-\lambda_\sigma W_T(\omega) - \frac{1}{2}\lambda_\sigma^2 T} d\mathbb{P}(\omega) d\sigma = \\ = \textstyle\int_0^\infty f(\sigma) d\sigma = 1. \end{aligned}$$

In our setting, the no-arbitrage price for the call option, with strike K and maturity T, is based on the well known Black-Scholes formula for the option price:

$$\begin{aligned} C(t,S) &= \int_0^\infty \mathbb{E}^{\mathbb{Q}^\sigma}[(S^\sigma - K)_+] f(\sigma) d\sigma \\ &= \int_0^\infty [SN(d_1^\sigma) - Ke^{-r(T-t)} N(d_2^\sigma)] f(\sigma) d\sigma \\ &=: SN_1^f - Ke^{-rT} N_2^f \end{aligned} \tag{3.43}$$

where

$$d_1^\sigma := \frac{\log(S/K) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \qquad d_2^\sigma := d_1^\sigma - \sigma\sqrt{T-t},$$

represent the usual Black-Scholes coefficients, and

$$N_i^f := \int_0^\infty N(d_i^\sigma) f(\sigma) d\sigma, \qquad i = 1,\, 2. \tag{3.44}$$

These coefficients can be furtherly simplified depending on the choice of the density function f.

Example 3.2.2. Now we are going to review the Heston model in light of our no-arbitrage setting. Let $(W_t^1)_t$, $(W_t^2)_t$ be two independent Brownian motions on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. The Heston model (1993) assumes the following real-world dynamics (i.e. under measure $\mathbb{P}$):

$$\begin{cases} dS_t(v_0) = \mu S_t(v_0)dt + S_t\sqrt{V_t(v_0)}dW_t^1, \\ dV_t(v_0) = \kappa(\theta - V_t(v_0))dt + \sigma\sqrt{V_t(v_0)}\left(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2\right), \end{cases} \tag{3.45}$$

with $S_0(v_0) > 0$, $V_0(v_0) = v_0 > 0$. The parameters $\mu \in \mathbb{R}$, $\kappa, \theta, \sigma > 0$, $\rho \in [-1, 1]$ which are assumed to be constants, satisfy Feller's condition, $2\kappa\theta \geq \sigma^2$, in order to ensure that volatility is strictly positive in finite time (see Revuz and Yor, (1991)). In this case, the uncertainty is on the parameter v_0, the initial volatility.

However, options are usually priced under the risk-neutral measure and they incorporate a volatility risk premium. The risk-neutral model corresponding to equations (3.45), is usually obtained by applying Girsanov's change of measure theorem:

$$\begin{cases} dS_t(v_0) = rS_t(v_0)dt + S_t\sqrt{V_t(v_0)}d\tilde{W}_t^1(v_0), \\ dV_t(v_0) = \kappa^{'}(\theta^{'} - V_t(v_0))dt + \sigma\sqrt{V_t(v_0)}\left(\rho d\tilde{W}_t^1(v_0) + \sqrt{1-\rho^2}d\tilde{W}_t^2(v_0)\right), \end{cases} \tag{3.46}$$

where $(\tilde{W}_t^1(v_0))_{t\in[0,T]}$, $(\tilde{W}_t^2(v_0))_{t\in[0,T]}$ are Brownian motion under the probability measure $\mathbb{Q}^{v_0}$, equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{F}_T)$, with Radon-Nikodym derivative

$$\begin{aligned} \frac{d\mathbb{Q}^{v_0}}{d\mathbb{P}}(\omega) &= \exp\left(-\int_0^T \lambda_1(t,v_0)d\tilde{W}_t^1(\omega,v_0) - \int_0^T \lambda_2(t,v_0)d\tilde{W}_t^2(\omega,v_0)\right. \\ &\quad \left. -\frac{1}{2}\int_0^T |\lambda_1(t,v_0)|^2dt - \frac{1}{2}\int_0^T |\lambda_2(t,v_0)|^2dt\right), \end{aligned}$$

and

$$d\tilde{W}_t^i(v_0) = dW_t^i + \lambda_i(t,v_0)dt, \qquad \text{for } i = 1,2.$$

Here $\{\lambda_1(t,v_0)\}_{t\in[0,T]}$ and $\{\lambda_2(t,v_0)\}_{t\in[0,T]}$ processes in $\mathcal{M}^2[0,T]$, for any $v_0 > 0$. In order to guarantee that the functional form of system (3.45) is preserved under $\mathbb{Q}^{v_0}$, the usual formulation of the market price of risk processes is given by

$$\begin{aligned}\lambda_1(t,v_0)\sqrt{V_t(v_0)} &= \mu - r, &(3.47)\\ \sqrt{1-\rho^2}\lambda_2(t,v_0) &= \lambda\sqrt{V_t(v_0)} - \rho\lambda_1(t,v_0), &(3.48)\end{aligned}$$

for some constant parameter $\lambda \in (-\kappa/\sigma, 0]$. Thus, the following relations hold:

$$\kappa^{'} = \kappa + \sigma\lambda, \qquad \theta^{'} = \frac{\kappa\theta}{\kappa^{'}}. \tag{3.49}$$

We recall that condition (3.47) ensures that discounted stock prices are local martingales under $\mathbb{Q}^{v_0}$, for any fixed $v_0 > 0$. In particular, it holds $\theta^{'} \geq \theta$, that is the risk-neutral measure captures the risk-averse nature under $\mathbb{P}$, since it increases the long-run level value of risk-neutral volatility. Since $(V_t(v_0))_t$ is an affine process, by the results obtained in [27], it admits a continuous dependence on the initial data.

Equation (3.47) represents the equity premium regarded as compensation for accepting the diffusive risk associated with W^1 and W^2. In fact, economic reasoning in Heston (1993) indicates that these compensations take the forms $\lambda_1\sqrt{V_t}$ and $\lambda_2\sqrt{V_t}$. We observe that, whereas the value $\lambda_1\rho+\lambda_2\sqrt{1-\rho^2}$ can be estimated from the asset price alone, the estimation of λ_1, λ_2 needs another class of observations like option prices. 3.4.1 The no-arbitrage price for a European call option, given under the extended market setting, is obtained by integration of the Heston price function C^H (2.26) against the density function f chosen for the initial volatility parameter:

$$C(t,S) = \int_0^\infty C^H(t,\log(S),v_0)f(v_0)dv_0. \tag{3.50}$$

In particular the dependence of the new price function on the volatility is removed. Relation (3.50) can be reduced furtherly to a single integration as we will investigate later in Theorem 3.4.1.

3.3 Averaging over volatility

So, as exposed above, the volatility represents a fundamental factor for the price evaluation of financial primary assets as well as of financial derivatives. We propose a new approach for the estimation parameters in stochastic volatility models. In this section we present an estimation approach which resume the theoretical framework described in previous sections, where the initial volatility is considered as a random variable driven by a given probability density function. This method leads to a new approach in calibrating option prices. We will focus primarily on the Heston model, reserving for future contributions the extension of such a technique to more general stochastic volatility models.

Our technique is mainly based on the work of Dragulescu and Yakovenko (2002) for estimation of the Heston model on historical stock prices. Using the Fourier and Laplace transforms, the authors solve the Fokker-Planck equation for the Heston model exactly and find the joint density function of log-returns and variance as a function of time, conditional on the initial volatility v_0. Thus, they integrate the joint density function over the initial volatility against a particular density function and they obtain a proxy for the marginal density function of log-returns, unconditional with respect to the initial variance. The approximated probability density function, found in Dragulescu and Yakovenko (2002), provides an excellent agreement with observed historical financial data. The intuition consists in supposing that the initial volatility is a random variable distributed according to the stationary distribution of the volatility process.

3.3.1 The stationary distribution of the volatility process

The processes in the Heston model are characterized by the transition density function $P_t(x, v|\ v_0)$, where x stands for the log-return (related to drift μ) and v is the at time t, given the initial conditions $x = 0$ e $v = v_0$ at $t = 0$. The evolution in time of $P_t(x, v|v_0)$ follows the following Fokker-

Planck equation:

$$\frac{\partial}{\partial t}P = \frac{\partial}{\partial v}[\kappa(v-\theta)P] + \frac{1}{2}\frac{\partial}{\partial x}(vP) + \rho\ \frac{\partial^2}{\partial x \partial v}(vP) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(vP) + \frac{{}^2}{2}\frac{\partial^2}{\partial v^2}(vP), \tag{3.51}$$

for $x \in \mathbb{R}$, $v > 0$, $t > 0$. The initial condition is the product of two Dirac delta functions: $P_{t=0}(x, v|v_0) = \delta(x)\delta(v - v_0)$. The marginal density of volatility, conditionated to the initial value v_0, is $\Pi_t(v|v_0) = \int_{-\infty}^{\infty} P_t(x, v|v_0)\, dx$, and it satisfies the Fokker-Planck partial differential equation given by

$$\frac{\partial}{\partial t}\Pi_t(v) = \frac{\partial}{\partial v}[\kappa(v-\theta)\Pi_t(v)] + \frac{{}^2}{2}\frac{\partial^2}{\partial v^2}[v\Pi_t(v)], \tag{3.52}$$

for $t > 0$, $v > 0$, obtained by integration of (3.51) in dx and imposing the initial condition $\Pi_{t=0}(v|v_0) = \delta(v - v_0)$. Due to the Feller's condition, it is proved that the initial value problem associated to (3.52) is well posed for $v \in (0, \infty)$, $t \in (0, \infty)$.

Proposition 3.3.1. The unique stationary solution $\Pi_\star \in C^2((0, \infty))$ of (3.52) which verifies the following conditions:

i) $\Pi_*(v) \geq 0$, for all $v > 0$,

ii) $v\Pi_*(v)$ in integrable on $(0, \infty)$,

iii) $\int_0^\infty \Pi_*(v)dv = 1$,

is the density function related to a Gamma distribution function with parameters $(\alpha, \frac{\theta}{\alpha})$, $\alpha = \frac{2\kappa\theta}{\sigma^2}$, that is:

$$\Pi_*(v) = \frac{v^{\alpha-1}}{\Gamma(\alpha)}\left(\frac{\alpha}{\theta}\right)^{\alpha} e^{-\frac{\alpha v}{\theta}}, \tag{3.53}$$

where $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$ is the Gamma function.

Proof. A stationary solution is the solutions of the ordinary differential equation

$$\kappa\frac{d}{dv}[(v-\theta)\Pi_*(v)] + \frac{\sigma^2}{2}\frac{d^2}{dv^2}[v\Pi_*(v)] = 0, \tag{3.54}$$

which implies that

$$\kappa(v-\theta)\Pi_*(v) + \frac{\sigma^2}{2}\frac{d}{dv}(v\Pi_*(v)) = C, \tag{3.55}$$

for some constant C, that is

$$\kappa(v-\theta)\Pi_*(v)+\frac{v\sigma^2}{2}\frac{d}{dv}\Pi_*(v)+\frac{\sigma^2}{2}\Pi_*(v)=C, \tag{3.56}$$

$$\frac{d}{dv}\Pi_*(v)=\frac{2}{v\sigma^2}\left[C-\left(\kappa v-\kappa\theta+\frac{\sigma^2}{2}\right)\Pi_*(v)\right], \tag{3.57}$$

$$\begin{aligned}&\frac{d}{dv}\Pi_*(v)=\frac{2C}{v\sigma^2}-\frac{2}{v\sigma^2}\left(\kappa v-\kappa\theta+\frac{\sigma^2}{2}\right)\Pi_*(v)=\\&\frac{2C}{v\sigma^2}+\left(-\frac{1}{v}+\frac{2\kappa\theta}{v\sigma^2}-\frac{2\kappa}{\sigma^2}\right)\Pi_*(v)=\frac{2C}{v\sigma^2}+\left(\frac{\alpha-1}{v}-\frac{\alpha}{\theta}\right)\Pi_*(v).\end{aligned} \tag{3.58}$$

From (3.55), by integrating between $v_0>0$ and $\overline{v}>v_0$, we get

$$\kappa\int_{v_0}^{\overline{v}}v\Pi_*(v)dv-\kappa\theta\int_{v_0}^{\overline{v}}\Pi_*(v)dv+\frac{\sigma^2}{2}\big(\overline{v}\Pi_*(\overline{v})-v_0\Pi_*(v_0)\big)=C(\overline{v}-v_0). \tag{3.59}$$

Using conditions $i)$-$iii)$, terms on the left-hand side in (3.59) are bounded for $\overline{v}\to\infty$. Hence, we argue tha $C=0$. Substituting $C=0$ in (3.58), we obtain the following ordinary differential equation:

$$\frac{d}{dv}\Pi_*(v)=\left(\frac{\alpha-1}{v}-\frac{\alpha}{\theta}\right)\Pi_*(v), \tag{3.60}$$

whose generale solution is

$$\Pi_*(v)=Ae^{-\frac{\alpha}{\theta}v+(\alpha-1)\log(v)}=Ae^{-\frac{\alpha}{\theta}v}v^{\alpha-1}, \tag{3.61}$$

for $A>0$, constant. We can find the value of the constant A, thanks to condition $iii)$:

$$\int_0^\infty\Pi_*(v)dv=A\int_0^\infty e^{-\frac{\alpha}{\theta}v}v^{\alpha-1}dv=1, \tag{3.62}$$

implies

$$\begin{aligned}A&=\left(\int_0^\infty e^{-\frac{\alpha}{\theta}v}v^{\alpha-1}dv\right)^{-1}=\left(\int_0^\infty e^{-t}\frac{\theta^\alpha}{\alpha^\alpha}t^{\alpha-1}dt\right)^{-1}\\&=\frac{\alpha^\alpha}{\theta^\alpha}\frac{1}{\int_0^\infty e^{-t}t^{\alpha-1}dt}=\left(\frac{\alpha}{\theta}\right)^\alpha\frac{1}{\Gamma(\alpha)},\end{aligned} \tag{3.63}$$

where the change of variables $v=\frac{\theta}{\alpha}t$ has been used. Replacing the value found for A in (3.61), we finally get the assertion. □

This stationary distribution is a Gamma distribution with coefficients depending on the Heston model parameters. Since x appears in (3.51) only in the derivative operator $\frac{\partial}{\partial x}$, it is convenient to take the Fourier transform:

$$P_t(x, v|v_0) = \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} e^{ip_x x} \bar{P}_{t,p_x}(v|v_0). \tag{3.64}$$

Inserting this expression we find the equation verified by $\bar{P}$:

$$\frac{\partial}{\partial t}\bar{P} = \frac{\partial}{\partial v}[\kappa(v-\theta)\bar{P}] - \big[\frac{p_x^2 - ip_x}{2} v - i\rho\sigma p_x \frac{\partial}{\partial v} v - \frac{\sigma^2}{2}\frac{\partial^2}{\partial v^2} v\big]\bar{P}, \tag{3.65}$$

for $t > 0$, $v > 0$. The latter is simpler than (3.51), because the number of variables has been reduced to two, v and t, whereas p_x only plays the role of a parameter. Since equation (3.65) is linear in v and quadratic in $\frac{\partial}{\partial x}$, it can be simplified by taking the Laplace transform over v:

$$\widetilde{P}_{t,p_x}(p_v|v_0) = \int_0^{+\infty} dv e^{-p_v v} \bar{P}_{t,p_x}(v|v_0). \tag{3.66}$$

The PDE satisfied by $\widetilde{P}_{t,p_x}(p_v|v_0)$ is of the first order

$$\left[\frac{\partial}{\partial t} + \big(Lp_v + \frac{\sigma^2}{2}p_v^2 - \frac{p_x^2 - ip_x}{2}\big)\frac{\partial}{\partial p_v}\right]\widetilde{P} = -\kappa\theta p_v \widetilde{P} \tag{3.67}$$

where

$$L = \kappa + i\rho\sigma p_x \tag{3.68}$$

and the corresponding initial condition is given by

$$\widetilde{P}_{0,p_x}(p_v|v_0) = e^{-p_v v_0}. \tag{3.69}$$

The solution of (3.67) is given by the method of characteristics [33] and is given by:

$$\widetilde{P}_{t,p_x}(p_v|v_0) = \exp\left(-\tilde{p}_v(0)v_0 - \kappa\theta\int_0^t d\tau \tilde{p}_v(\tau)\right) \tag{3.70}$$

where the function $\tilde{p}_v(\tau)$ is the solution of the characteristic (ordinary) differential equation:

$$\frac{d\tilde{p}_v}{d\tau}(\tau) = L\tilde{p}_v(\tau) + \frac{\sigma^2}{2}\tilde{p}_v^2(\tau) - \frac{p_x^2 - ip_x}{2}, \tag{3.71}$$

with the boundary condition $\tilde{p}_v(t) = p_v$ at $\tau = t$. The previous differential equation is of Riccati type, with constant coefficients, and its solution takes the following form:

$$\tilde{p}_v(\tau) = \frac{2M}{\sigma^2}\frac{1}{Ne^{M(t-\tau)}-1} - \frac{L-M}{\sigma^2}, \tag{3.72}$$

where

$$\begin{aligned} M &= \sqrt{L^2+\sigma^2(p_x^2 - ip_x)} \\ N &= 1 + \frac{2M}{\sigma^2 p_v + (L-M)}. \end{aligned} \tag{3.73}$$

Substituting (3.72) in (3.70), we find

$$\widetilde{P}_{t,p_x}(p_v|v_0) = \exp\Big[-\tilde{p}_v(0)v_0 - \frac{\kappa\theta(L-M)t}{\sigma^2} - \frac{2\kappa\theta}{\sigma^2}\log\Big(\frac{N-e^{-Mt}}{N-1}\Big)\Big]. \tag{3.74}$$

Usually, we are interested only in log-returns x and do not care about the volatility v. Moreover, whereas log-returns are directly known from financial data, volatility is a hidden stochastic variable that has to be estimated. Inevitably, such an estimation is done with some degree of uncertainty, which precludes a clear-cut direct comparison between $P_t(x|v_0)$ and financial data. Thus, in [25], the reduced probability distribution is considered:

$$P_t(x|v_0) = \int_0^{+\infty} dv P_t(x,v|v_0) = \int_{-\infty}^{+\infty}\frac{dp_x}{2\pi}e^{ip_x x}\widetilde{P}_{t,p_x}(0|v_0). \tag{3.75}$$

Inserting (3.74) in the last expression, we obtain:

$$\begin{aligned} P_t(x|v_0) &= \int_{-\infty}^{+\infty}\frac{dp_x}{2\pi}\exp\Big(ip_x x - v_0\frac{p_x^2 - ip_x}{L + M\coth(\frac{Mt}{2})}\Big) \\ &\quad\times\exp\Big[-\frac{2\kappa\theta}{\sigma^2}\log\Big(\cosh(\frac{Mt}{2}) + \frac{L}{M}\sinh(\frac{Mt}{2})\Big) + \frac{\kappa L\theta t}{\sigma^2}\Big]. \end{aligned} \tag{3.76}$$

However, equation cannot be directly compared with financial time series data, because it depends on the unknown initial volatility v_0. In order to resolve this problem, Dragulescu and Yakovenko [25] have assumed that v_0 is a random parameter driven by the stationary probability distribution $\Pi_\star$. So, the estimated probability density function for the log-returns, $P_t(x)$, becomes

$$P_t(x) = \mathbb{E}_{\Pi_*}[P_t(x|v_0)] = \int_0^{+\infty}\Pi_*(v_0)P_t(x|v_0)\,dv_0 \tag{3.77}$$

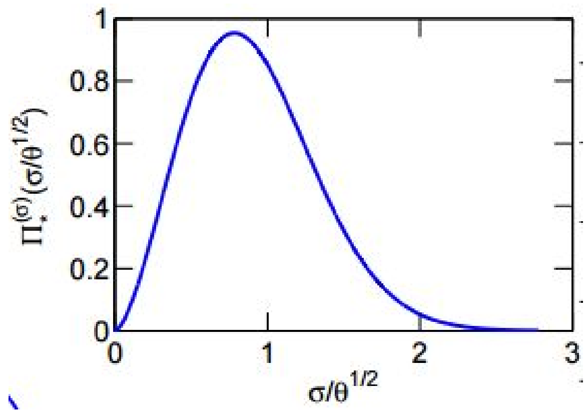


Figure 3.1: See [25]. The stationary probability distribution $\Pi_*(v)$ of variance v, given by Eq. (**??**) and shown for $\alpha = 1.3$, $\gamma = 11.35, \theta = 0.022, \kappa = , 0.618, \mu = 0.143$.

Equation (3.77) can be furtherly simplified with the aim to let an easier implementation from the numerical point of view .

Proposition 3.3.2. For every $t > 0$, $x \in \mathbb{R}$ and for any choice of parameters $\sigma, \kappa > 0$, $\rho \in [-1, 1]$, $\theta \in \mathbb{R}$, with $2\kappa\theta > \sigma^2$, $P_t(x)$ in (3.77) assumes a Fourier type integral representation:

$$P_t(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dp_x e^{ip_x x + F_t(p_x)}, \tag{3.78}$$

where

$$F_t(p_x) = \frac{\kappa\theta L t}{2} - \frac{2\kappa\theta}{2} \log\left[\cosh\left(\frac{Mt}{2}\right) + \frac{M^2 - L^2 + 2\kappa L}{2\kappa M} \sinh\left(\frac{Mt}{2}\right)\right]. \tag{3.79}$$

In order to prove the proposition, we need the following relation for the generalized moment generating function of the Gamma distribution:

$$\int_0^\infty e^{zv} \frac{v^{\alpha-1}}{\Gamma(\alpha)} \left(\frac{\alpha}{\theta}\right)^\alpha e^{-(\alpha v)/\theta} dv = \frac{1}{(1 - \frac{\theta}{\alpha} z)^\alpha}, \tag{3.80}$$

which holds for all $z \in \mathbb{C}$, with $\mathrm{Re}(z) < \alpha/\theta$. In fact it can be easily derived according with the following passages: Let $z = x + iy$ in $\mathbb{C}$, with $x < \alpha/\theta$,, then

$$\int_0^\infty (e^{zv})\frac{v^{(\alpha-1)}}{\Gamma(\alpha)}\left(\frac{\alpha}{\theta}\right)^\alpha e^{-(\alpha v)/\theta}dv = \frac{1}{(1-\frac{\theta}{\alpha}z)^\alpha}.$$

Due to $e^{zv} = e^{xv}(\cos(vy) + i\sin(vy))$, we can write

$$\int_0^\infty (e^{zv})\frac{v^{(\alpha-1)}}{\Gamma(\alpha)}\left(\frac{\alpha}{\theta}\right)^\alpha e^{-(\alpha v)/\theta}dv =$$

$$\frac{\alpha^\alpha}{\Gamma(\alpha)}\int_0^\infty \frac{v^{\alpha-1}}{\theta^\alpha}e^{vx}e^{-(\alpha v)/\theta}\cos(vy)dv + \frac{i\alpha^\alpha}{\Gamma(\alpha)}\int_0^\infty \frac{v^{\alpha-1}}{\theta^\alpha}e^{vx}e^{-(\alpha v)/\theta}\sin(vy)dv.$$

Applying the substitution $v = (-x+\frac{\alpha}{\theta})^{-1}w$, we get

$$\begin{aligned}
&\frac{\alpha^\alpha}{\Gamma(\alpha)\theta^\alpha}\int_0^\infty w^{\alpha-1}\big(\frac{\alpha}{\theta}-x\big)^{-\alpha}e^{-w}\cos\Big(yw\big(\frac{\alpha}{\theta}-x\big)^{-1}\Big)dw+\\
&\frac{i\alpha^\alpha}{\Gamma(\alpha)\theta^\alpha}\int_0^\infty w^{\alpha-1}\big(\frac{\alpha}{\theta}-x\big)^{-\alpha}e^{-w}\sin\Big(yw\big(\frac{\alpha}{\theta}-x\big)^{-1}\Big)dw=\\
&\big(\frac{\alpha}{\theta}\big)^\alpha\big(\frac{\alpha}{\theta}-x\big)^{-\alpha}\int_0^\infty \frac{w^{\alpha-1}}{\Gamma(\alpha)}e^{-w}e^{iAw}dw,
\end{aligned}$$

where we fixed $A := y\big(\frac{\alpha}{\theta}-x\big)^{-1}$. Observe here that the integral in the last line of **??** is the charactheristic function of the Gamma distribution with parameters $(\alpha, 1)$ evaluated in A. Thus, the following equalities allow us to deduce (3.80):

$$\begin{aligned}
&\big(\frac{\alpha}{\theta}\big)^\alpha\big(\frac{\alpha}{\theta}-x\big)^{-\alpha}\int_0^\infty \frac{w^{\alpha-1}}{\Gamma(\alpha)}e^{-w}e^{iAw}dw=\\
&\big(\frac{\alpha}{\theta}\big)^\alpha\big(\frac{\alpha}{\theta}-x\big)^{-\alpha}\frac{1}{(1-iA)^\alpha}=\\
&\big(\frac{\alpha}{\theta}\big)^\alpha\frac{1}{\big(\frac{\alpha}{\theta}-x-iy\big)^\alpha}=\\
&\big(\frac{\alpha}{\theta}\big)^\alpha\frac{1}{\big(\frac{\alpha}{\theta}-z\big)^\alpha}=\frac{1}{(1-\frac{\theta}{\alpha}z)^\alpha}.
\end{aligned}$$

Define the following complex-valued functions:

$$A(p_x,t) = \frac{ip_x - p_x^2}{L + M\coth(\frac{Mt}{2})}, \tag{3.81}$$

$$B(p_x,t) = -\frac{2\kappa\theta}{\sigma^2}\log\Big(\cosh(\frac{Mt}{2}) + \frac{L}{M}\sinh(\frac{Mt}{2})\Big) + \frac{\kappa L\theta t}{\sigma^2}, \tag{3.82}$$

where we omit the dependence on the parameters κ, ρ and σ. In order to prove the central result of this section, we need a technical result, that is not proved in [25].

Lemma 3.3.1. For every $t > 0$, $p_x \in \mathbb{R}$ and for any choice of parameters $\sigma, \kappa > 0$, $\rho \in [-1, 1]$, $\theta \in \mathbb{R}$, with $2\kappa\theta > \sigma^2$, it holds $\mathrm{Re}(A(p_x, t)) \le 0$ and $p_x \in \mathbb{R} \mapsto \exp \mathrm{Re}((B(p_x, t)))$ is integrable on $\mathbb{R}$.

Proof of Proposition 3.3.2. We substitute the full expression of $P_t(x|v_0)$, found in (3.76), into (3.77), to get

$$P_t(x) = \int_0^{+\infty} dv_0 \Pi_*(v_0) \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} e^{ip_x x + v_0 A(p_x,t) + B(p_x,t)}, \tag{3.83}$$

Exchanging the integrals in (3.83), we get

$$P_t(x) = \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} e^{ip_x x + B(p_x,t)} \int_0^{+\infty} \Pi_*(v_0) e^{v_0 A(p_x,t)} dv_0. \tag{3.84}$$

In the second integral we recognize the moment generating function of the Gamma distribution with parameters $(\alpha, \frac{\theta}{\alpha})$ that is evaluated in $A(p_x, t)$, from which results that

$$\int_0^{+\infty} \Pi_*(v_0) e^{v_0 A(p_x,t)} dv_0 = \frac{1}{\left(1 - \frac{\theta}{\alpha} A(p_x, t)\right)^{\alpha}} \tag{3.85}$$

with the condition $\mathrm{Re}(A(p_x, t)) < \alpha/\theta$, which follows from Lemma 3.3.1. Replacing (3.85) in (3.84), we have

$$P_t(x) = \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} \frac{e^{ip_x x + B(p_x,t)}}{\left(1 - \frac{\theta}{\alpha} A(p_x, t)\right)^{\alpha}}. \tag{3.86}$$

And so

$$P_t(x) = \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} e^{ip_x x + B(p_x,t) - \alpha \log\left(1 - \frac{\theta}{\alpha} A(p_x,t)\right)}. \tag{3.87}$$

The proof is easily done by replacing the coefficients A and B in (3.87) with their respective values (3.81)-(3.82). $\square$

Proof of Lemma 3.3.1. We observe that $A(p_x, t) = 0$, if $p_x = 0$, for all $t > 0$. Hence, in the following we assume that $p_x \neq 0$. We change the expression of $A(p_x, t)$ as follows:

$$A(p_x, t) = -\frac{L}{\sigma^2} \frac{(M/L)^2 - 1}{1 + \frac{M}{L} \coth\left(\frac{M}{L} \frac{Lt}{2}\right)}. \tag{3.88}$$

Since $L > 0$ and

$$\mathrm{Re}\left(\frac{M^2}{L^2}\right) = 1 + \frac{\sigma^2 p_x^2}{L^2} > 1,$$

for any $p_x \in \mathbb{R}\backslash\{0\}$, and by $\mathrm{Re}(1/z) = \mathrm{Re}(\overline{z}/|z|^2) = \frac{\mathrm{Re}(z)}{|z|^2}$ for all $z \in \mathbb{C}\backslash\{0\}$, it suffices to show that,

$$\mathrm{Re}\left(\frac{1 + z\coth(\tau\, z)}{z^2 - 1}\right) \geq 0,$$

for any $\tau > 0$, $z = \alpha + i\omega \in \mathbb{C}$, with $\mathrm{Re}(z^2) = \alpha^2 - \omega^2 > 1$. We compute the real part in (3.89). It is easy to see that

$$\frac{1 + z\coth(\xi\, z)}{z^2 - 1} = \frac{\xi}{\eta},$$

where

$$\begin{aligned} \xi = {} & \sinh(\tau\alpha)\cos(\tau\omega) + i\cosh(\tau\alpha)\sin(\tau\omega) + (\alpha + i\omega) \times \\ & [\cosh(\tau\alpha)\cos(\tau\omega) + i\sinh(\tau\alpha)\sin(\tau\omega)], \end{aligned}$$

$$\eta = (\alpha^2 - \omega^2 - 1 + 2i\alpha\omega)[\sinh(\tau\alpha)\cos(\tau\omega) + i\cosh(\tau\alpha)\sin(\tau\omega)].$$

Hence (3.89) is satisfied if and only if $\mathrm{Re}(\xi)\mathrm{Re}(\eta) + \mathrm{Im}(\xi)\mathrm{Im}(\eta) \geq 0$. After some calculation, we get

$$\begin{aligned} & \mathrm{Re}(\xi)\mathrm{Re}(\eta) + \mathrm{Im}(\xi)\mathrm{Im}(\eta) = \\ & (\alpha^2 - \omega^2 - 1)\left[\sinh^2(\tau\alpha)\cos^2(\tau\omega) + \cosh^2(\tau\alpha)\sin^2(\tau\omega)\right] \\ & + \frac{\alpha}{2}(\alpha^2 + \omega^2 - 1)\sinh(2\tau\alpha) - \frac{\omega}{2}(\alpha^2 + \omega^2 + 1)\sin(2\tau\omega). \end{aligned}$$

Since the previous expression is symmetric with respect to ω, it suffices to prove that

$$q := \frac{\alpha}{2}(\alpha^2 + \omega^2 - 1)\sinh(2\tau\alpha) - \frac{\omega}{2}(\alpha^2 + \omega^2 + 1)\sin(2\tau\omega) \geq 0,$$

when $\omega > 0$, with $\sin(2\tau\omega) \geq 0$. We apply the well known inequalities $\sinh(x) \geq x$ and $\sin(x) \leq x$, for any $x \geq 0$, to get

$$\begin{aligned} q & \geq \tau[\alpha^2(\alpha^2 + \omega^2 - 1) - \omega^2(\alpha^2 + \omega^2 + 1)] \\ & = \tau\left(\alpha^4 - \alpha^2 - \omega^4 - \omega^2\right) = \tau(\alpha^2 + \omega^2)\left(\alpha^2 - \omega^2 - 1\right) > 0. \end{aligned}$$

Thus, the result for A is proved. For the complex-valued function $B(p_x, t)$, we have to show that, for every $t > 0$, $\rho \in [-1, 1]$, $\sigma > 0$, $\kappa > 0$, $\theta > 0$, with $2\kappa\theta > \sigma^2$, the function

$$p_x \in \mathbb{R} \mapsto e^{\mathrm{Re}(B(p_x,t))}$$

is integrable on $\mathbb{R}$. To this aim, we estimate the absolute value of

$$w(M, L, t) := \cosh(Mt/2) + (L/M)\sinh(Mt/2). \tag{3.89}$$

Since $w(-M, L, t) = w(M, L, t)$, we can consider only the case in which $M = me^{i\theta/2}$, where $m > 0$, $m^4 = [\kappa^2 + \sigma^2 p_x^2(1-\rho^2)]^2 + {}^2p_x^2(2\rho\kappa - \sigma)^2$ and $\theta \in (-\pi, \pi]$ satisfies

$$\cos\theta = \frac{\kappa^2 + \sigma^2 p_x^2(1-\rho^2)}{m^2}, \qquad \sin\theta = \frac{\sigma p_x(2\kappa\rho - \sigma)}{m^2}.$$

In particular, it holds $\mathrm{Re}(M) > 0$. We study the behavior of $\mathrm{Re}(M)$ and L/M, as $|p_x| \to \infty$.

$$\left|\frac{L}{M}\right|^2 = \frac{|L|^2}{m^2} = \frac{\kappa^2 + \rho^2\sigma^2 p_x^2}{\sqrt{[(1-\rho^2)\sigma^2 p_x^2 + \kappa^2]^2 + \sigma^2 p_x^2(2\rho\kappa - \sigma)^2}},$$

and

$$(\mathrm{Re}(M))^2 = \frac{m^2}{2}(1 + \cos\theta) = \frac{m^2 + \kappa^2 + {}^2p_x^2(1-\rho^2)}{2}.$$

We deduce the following asymptotic regimes as $|p_x| \to \infty$. Let $(\gamma, c) := (\sigma/|2\rho\kappa - \sigma|, 1/2)$ if $\sigma \neq 2\kappa$, whereas $(\gamma, c) := (2, 1)$, if $\sigma = 2\kappa$. Thus

$$\left|\frac{L}{M}\right| \approx \begin{cases} |\rho|/\sqrt{1-\rho^2} & \text{if } \rho \neq 0, \pm 1, \\[2em] \kappa/(\sigma|p_x|) & \text{if } \rho = 0, \\[2em] (\gamma|p_x|)^c & \text{if } \rho = \pm 1, \end{cases}$$

$$\mathrm{Re}M \approx \begin{cases} |p_x|\sqrt{1-\rho^2} & \text{if } \rho \neq \pm 1, \\ \sqrt{|p_x|/(2\gamma)} & \text{if } \rho = \pm 1 \text{ and } 2\kappa \neq \ , \end{cases}$$

whereas, in the case $\rho = 1$ with $2\kappa = \sigma$, we get

$$M = \kappa, \qquad\qquad \frac{L}{M} = 1 + 2ip_x.$$

Let $b = |\rho|/\sqrt{1-\rho^2}$, $a = (b+1)/2$, if $\rho \neq \pm 1$, and let $a = 2$, if $\rho = \pm 1$. By (3.90), there exists $\tilde{p} > 0$ depending on ρ, κ and σ, such that, for all $p_x \in \mathbb{R}$ with $|p_x| > \tilde{p}$, the following inequalities are satisfied:

- If $|\rho| < 1/\sqrt{2}$ $(a < b < 1)$, it holds $|L/M| < a$ and

$$\begin{aligned} |w(M,L,t)| \geq \left|\cosh\left(\frac{Mt}{2}\right)\right| - a\left|\sinh\left(\frac{Mt}{2}\right)\right| \geq \sinh\left(\mathrm{Re}(M)\frac{t}{2}\right) \\ -a\cosh\left(\mathrm{Re}(M)\frac{t}{2}\right) = \frac{(1-a)}{2}e^{\frac{t}{2}\mathrm{Re}(M)} - \frac{1+a}{2}e^{-\frac{t}{2}\mathrm{Re}(M)}. \end{aligned}$$

- If $|\rho| > 1/\sqrt{2}$ $(a > b > 1)$, it holds $|L/M| > a$ and

$$\begin{aligned} |w(M,L,t)| \geq a\left|\sinh\left(\frac{Mt}{2}\right)\right| - \left|\cosh\left(\frac{Mt}{2}\right)\right| \geq a\sinh\left(\mathrm{Re}(M)\frac{t}{2}\right) \\ -\cosh\left(\mathrm{Re}(M)\frac{t}{2}\right) = \frac{(a-1)}{2}e^{\frac{t}{2}\mathrm{Re}(M)} - \frac{a+1}{2}e^{-\frac{t}{2}\mathrm{Re}(M)}. \end{aligned}$$

- If $\rho = \pm 1/\sqrt{2}$ and $2\rho\kappa \neq \sigma$, by replacing ρ^2 with $1/2$ in (3.90) and using the Taylor expansion of $\tau \mapsto 1/\sqrt[4]{1+\tau}$ near $\tau = 0$, we deduce the following equation:

$$\left|\frac{L}{M}\right| = \left(1 + \frac{4\;{}^2p_x^2(\sqrt{2}\kappa - \sigma)^2}{(2\kappa^2 + \sigma^2 p_x^2)^2}\right)^{-1/4} = 1 - \frac{(\sqrt{2}\kappa - \sigma)^2}{\sigma^2 p_x^2} + \mathrm{R}(p_x),$$

where $|p_x^4 \mathrm{Re}(p_x)| \leq \tilde{r}$, for any $|p_x| > \tilde{p}$, with $\tilde{r} > 0$ depending only on σ and κ. Using (3.90), we get the estimate

$$\begin{aligned} |w(M,L,t)| \;\;\geq\;\; & \sinh\left(\mathrm{Re}(M)\frac{t}{2}\right) - \cosh\left(\mathrm{Re}(M)\frac{t}{2}\right) + \left(1 - \left|\frac{L}{M}\right|\right) \times \\ & \times \cosh\left(\mathrm{Re}(M)\frac{t}{2}\right) \\ & \geq e^{-\frac{t}{2}\mathrm{Re}(M)} + \left(\frac{(\sqrt{2}\kappa - \sigma)^2}{\sigma^2 p_x^2} - \frac{\tilde{r}}{p_x^4}\right)\cosh\left(\mathrm{Re}(M)\frac{t}{2}\right) \end{aligned}$$

- If $\rho = 1/\sqrt{2}$ and $2\rho\kappa = \sigma$, it holds $L/M = e^{i\varphi}$, where

$$\cos\varphi = \frac{\kappa}{m}, \qquad \sin\varphi = \frac{p_x}{\sqrt{2}m}, \qquad M = m = \sqrt{\kappa^2 + \;{}^2p_x^2/2}.$$

Hence

$$|w(M,L,t)| \geq \cosh\left(M\frac{t}{2}\right) + \cos(\varphi)\sin\left(M\frac{t}{2}\right) \geq \frac{1}{2}e^{\frac{t}{2}M}.$$

In the specific case $(\rho, 2\kappa) = (1, \sigma)$, we consider the inequality

$$|w(M,L,t)| = \left|\cosh\left(\kappa\frac{t}{2}\right) + (1+2ip_x)\sinh\left(\kappa\frac{t}{2}\right)\right| \geq 2|p_x|\sinh\left(\kappa\frac{t}{2}\right).$$

Let $(\rho, 2\kappa) \neq (1, \sigma)$, then from the above inequalities we deduce that there are constants $c_1 > 0$, $c_2 \geq 0$ depending only on ρ, κ, σ and t such that, for every $|p_x| > \overline{p}$, $B(p_x, t)$ satisfies

$$\mathrm{Re}\left(B(p_x,t)\right) \leq -\frac{2\kappa\theta}{\sigma^2}\log\left[\left(c_1 + c_2/p_x^2\right)e^{\,\zeta\frac{t}{2}|p_x|^c}\right] + \frac{\kappa^2\theta t}{2}, \tag{3.90}$$

where $(c, \zeta) := (1, \sqrt{1-\rho^2})$ if $\rho \neq \pm 1$, whereas $(c, \zeta) := (1/2, 1/\sqrt{2\gamma})$ otherwise. Here $c_2 \neq 0$ only if $\rho = \pm 1/\sqrt{2}$ with $2\rho\kappa \neq \sigma$. Taking the exponential in (3.90), we get the inequality

$$e^{\mathrm{Re}(B(p_x,t))} \leq \frac{\exp(\frac{\kappa^2\theta t}{\sigma^2})}{[c_1 + c_2/p_x^2]^{2\kappa\theta\sigma^2}}\exp\left(-\zeta\frac{\kappa\theta t}{\sigma}|p_x|^c\right), \tag{3.91}$$

for any $|p_x| > \overline{p}$. The function on right-hand is integrable on $\mathbb{R}$, with respect to p_x. Similarly, in the case $(\rho, 2\kappa) = (1, \sigma)$, we have

$$e^{\mathrm{Re}((B(p_x,t)))} \leq \frac{\exp(\frac{\kappa^2\theta t}{\sigma^2})}{[2\sinh(\kappa t/2)]^{2\kappa\theta/\sigma^2}}\frac{1}{|p_x|^{2\kappa\theta/\sigma^2}}. \tag{3.92}$$

Since,$2\kappa\theta > \sigma^2$, also in this case $p_x \mapsto \exp(\mathrm{Re}\left(B(p_x,t)\right))$ is integrable on $\mathbb{R}$. $\square$

As an application of the technique described above, we present the estimation of the Heston model based on closure daily prices of FTSE MIB index (the reference stock market index for the Italian Stock Exchange), from 06/09/2008 to 06/09/2012. The estimation approach is based on the following minimization procedure, on the set of all admissible parameters $\overline{\Omega} = (\mu, \kappa, \theta, \sigma, \rho) \in \mathbb{R}_+^4$, with $2\kappa\theta \geq \sigma^2$, $\rho \in (-1, 1)$.

$$\min_{\overline{\Omega}} S_1(\overline{\Omega}) := \min_{\overline{\Omega}} \sum_x |\log(Emp_\Delta(x)) - \log(P_\Delta(x))|^2. \tag{3.93}$$

where $P_\Delta(x)$ is the proxy for the density of log-returns based on the representation (3.97), Δ is the time frequency of the date (in our case $\Delta \approx 1/250$). Here the comparison is performed using the empirical density $Emp_\Delta(x)$. In the table 3.1 are shown the results of minimization. As it is shown from Figure 3.2, the accuracy of approximation gets down when we get away from the mean value. In accordance with [25], this feature is due to the irregular behavior of the function $F_t(p_x)$ (3.79) in the complex level of p_x: the function $F_t(p_x)$ has, in fact, singularities where the argument of logarithm is null. An improvement of the estimation result can be obtained following an approach which aim to a better approximation of the tails by exponential functions. Let

$$G_\Delta(x) = \begin{cases} P_\Delta(x)e^{(x-a)q^+} & se\ x < a \\ P_\Delta(x) & se\ a \leq x \leq b \\ P_\Delta(x)e^{(-x+b)q^-} & se\ x > b, \end{cases} \tag{3.94}$$

where q^+ e q^- are two additional parameters which will be add to the previous set, which so will become in this case $\overline{\Omega} = (\mu, \kappa, \theta, \sigma, \rho, q^+, q^-)$. The value of a and b are evaluated so that a certain percentage (for example 98%) of desired data will be in the gap $[a, b]$. In this case the minimization procedure is based on the following functional:

$$\min_{\overline{\Omega}} S_2(\overline{\Omega}) := \min_{\overline{\Omega}} \sum_x |\log(Emp_\Delta(x)) - \log(G_\Delta(x))|^2. \tag{3.95}$$

In Table 3.2 are shown the results of the minimization procedure following the last extension while in the Graph 3.3 it could be examine the improvement of approximations on the tails

	Initial values	Results
μ	0.000567	-0.00025
κ	0.045	0.000053
θ	0.0000862	0.00041
σ	0.00245	0.000047
ρ	0	-0.9879
$S_1(\overline{\Omega})$	1664	29.06
Time occured	242.8 secondi	
nr. of iterations	1631	

Table 3.1: Results of the estimation procedure based on the error functional $S_1(\overline{\Omega})$.

	Initial Values	Results
μ	0.000567	-0.0015
κ	0.045	0.1096
θ	0.0000862	0.00040
σ	0.00245	0.00694
ρ	0	-0.9732
q^+	0	55.774
q^-	0	-2.627
$S_2(\overline{\emptyset})$	1832	18.61
time occurred	187.7 seconds	
nr. of iterations	1401	

Table 3.2: Results of the estimation procedure (3.95).

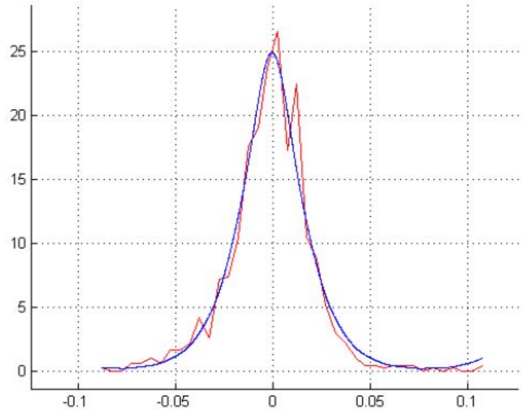


Figure 3.2: The comparison between the empirical density (in red) and the desity $P_\Delta(x)$ (in blue)

We can generalize the above approach by assuming that v_0 has an arbitrary probability distribution with density function $\Pi(v_0)$. Thus, we consider a non-negative function $\Pi \in L^1(\mathbb{R})$, such that $\Pi(v) = 0$, for all $v \leq 0$, and we denote by M_Π its (extended) moment generating function, namely

$$M_\Pi(z) = \int_0^\infty e^{zv}\, \Pi(v)\, dv, \tag{3.96}$$

which is clearly well defined for all $z \in \mathbb{C}$, with $\mathrm{Re}(z) \leq 0$.

According to Dragulescu and Yakovenko [25], we define the probability density function $P_t(x)$ at time t of the log-returns relative to the drift μ, by averaging over v_0 with the weight $\Pi(v_0)$, namely:

$$P_t(x) = \int_0^{+\infty} \Pi(v_0) P_t(x|v_0) dv_0. \tag{3.97}$$

Thus, we can extend the representation of $P_t(x)$ established in Proposition 3.3.2.

Theorem 3.3.1. Under the assumptions on the parameters given in Proposition 3.3.2, if the initial volatility is driven by the density function Π with the above described properties, then the probability density function $P_t(x)$ at time t of log-returns relative to the drift μ, given by (3.97), reduces to

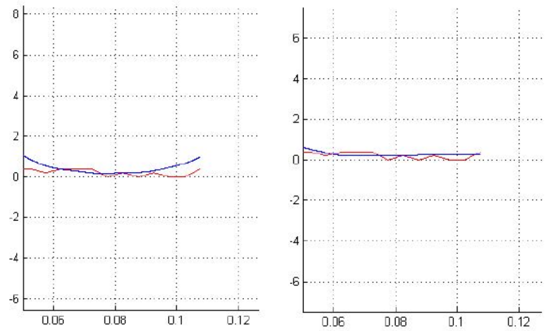


Figure 3.3: Comparison between the approximation obtained using by $P_\Delta(x)$ and $G_\Delta(x)$ (zoom on the tail on the right of dentities).

the following integral:

$$P_t(x) = \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} M_\Pi(A(p_x,t)) e^{ip_x x + B(p_x,t)}. \tag{3.98}$$

Proof of Theorem 3.3.1. By replacing the expression of $P_t(x|v_0)$ in (3.97), we have

$$P_t(x) = \int_0^{+\infty} dv_0 \Pi(v_0) \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} e^{ip_x x + v_0 A(p_x,t) + B(p_x,t)}.$$

In order to change the order of integrals, we show that

$$I \ := \ \int_{-\infty}^{+\infty} dp_x \int_0^{+\infty} dv_0 \frac{\Pi(v_0)}{2\pi} \left| e^{ip_x x + v_0 A(p_x,t) + B(p_x,t)} \right| < \infty. \tag{3.99}$$

By omitting the arguments of the complex-valued functions A and B, we can write

$$\begin{aligned}
I &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dp_x \int_0^{+\infty} dv_0 \Pi(v_0) e^{v_0 \mathrm{Re}(A(p_x,t)) + \mathrm{Re}(B(p_x,t))} \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dp_x e^{\mathrm{Re}(B(p_x,t))} \int_0^{+\infty} dv_0 \Pi(v_0) e^{v_0 \mathrm{Re}(A(p_x,t))} \\
&\le \frac{1}{2\pi} \int_{-\infty}^{+\infty} dp_x e^{\mathrm{Re}(B(p_x,t))} \int_0^{+\infty} dv_0 \Pi(v_0) \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dp_x e^{\mathrm{Re}(B(p_x,t))} \qquad (3.100)
\end{aligned}$$

where we used the fact that $\mathrm{Re}(A(p_x,t)) \le 0$, for any $p_x \in \mathbb{R}$. Thanks to

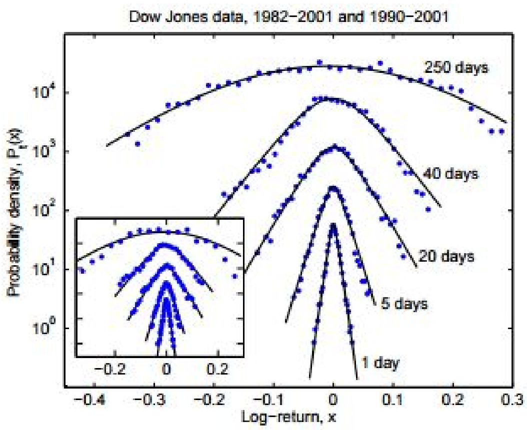


Figure 3.4: See [25]. Probability distribution Pt(x) of log-return x for different time lags t. Points: The 19822001 Dow-Jones data for $t = 1, 5, 20, 40, 250$ trading days. Solid lines: Fit of the data $P_t(x) = \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} e^{\imath p_x x + B(p_x,t)}$.. For clarity, the data points and the curves for successive t are shifted up by the factor of 10 each. Inset: The 19902001 Dow-Jones data points compared with the same theoretical curves.

Lemma 3.3.1, we get (3.99). This implies the following equation:

$$\begin{aligned} P_t(x) &= \int_0^{+\infty} dv_0 \Pi(v_0) \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} e^{ip_x x + v_0 A(p_x,t) + B(p_x,t)} \\ &= \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} e^{ip_x x + B(p_x,t)} \int_0^{+\infty} dv_0 \Pi(v_0) e^{v_0 A(p_x,t)} \\ &= \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} e^{ip_x x + B(p_x,t)} M_\Pi\big(A(p_x,t)\big) \end{aligned}$$

that is the desired expression. $\square$

Example 3.3.1. If $\Pi(v_0)$ is the probability density function of the Gamma distribution, with parameters $\alpha > 0$, $\beta > 0$,

$$\Pi(v_0) = g_{\alpha,\beta}(v_0) := \frac{\beta^\alpha}{\Gamma(\alpha)} v_0^{\alpha-1} e^{-\beta v_0}, \tag{3.101}$$

then the expression (3.98) yields

$$
\begin{aligned}
P_t(x) &= \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} \frac{e^{ip_x x + B(p_x,t)}}{\left(1 - \frac{1}{\beta} A(p_x,t)\right)^{\alpha}} \\
&= \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} e^{ip_x x + B(p_x,t) - \alpha \log\left(1 - \beta^{-1} A(p_x,t)\right)}. \qquad (3.102)
\end{aligned}
$$

The estimation method developed above for stock price data can be extended, with rigorous arguments, technique to the option pricing problem. Furthermore, we prove that, for a large class of probability distributions assumed for the initial volatility parameter, the estimation error in the calibration procedure of option prices is less than the case of the simple pricing formula. Our results are validated with a numerical comparison, on observed call prices, between the proposed calibration method and the classical approach.

The calibration of SV models to synthetic and market option data forms one of the major theme in the literature. Calibrating methods to market data (either option prices or implied volatilities) allows to infer the (risk-neutral) market parameters for the different models and thus to use these models for pricing and hedging purposes.
The cost of using such models, however, is that the calibration and pricing techniques that must be employed are usually quite onerous. The choice of a calibrating routine requires a trade-off between its computational complexity and its accuracy. This leads to a complication that plagues SV models in general. A common solution is to find those parameters which produce the correct market prices of vanilla options. This is called an inverse problem, as we solve for the parameters indirectly through some implied structure.

A well documented and popular method of fitting pricing models to observed data is to find a set of model parameter values that minimizes the square of the differences between the empirical values and the corresponding model values. More specifically, the squared differences between vanilla option market prices and model theoretical prices are minimized over the parameter space:

$$\inf_{\Theta} \sum_{i=1}^{N} w_i \Big(C^{Model}(S_0, K_i, T_i; \Theta) - C_i^{Market}(K_i, T_i) \Big)^2. \tag{3.103}$$

where Θ is the vector of parameter values, $C^{Model}(S_0, T_i, K_i; \Theta)$ and $C_i^{Market}(K_i, T_i)$ denote the i^{th} option price from the model and market dataset, respectively, with strike K_i and maturity T_i, whereas N is the number of options used for calibration. The coefficients w_i, $i = 1, \dots, N$, denote suitable weights. One possible choice for these weights is to set $w_i = 1/N$, for all $i = 1, \dots, N$, making equation (3.103) a measure of mean squared errors Zhu (2010). Alternatively, we could let $w_i = |\text{bid}_i - \text{ask}_i|^{-1}$, where bid_i and ask_i stand for the bid and ask prices of the i^{th} option in the dataset. This would allow us to place more weight on options which are more liquid in the market. A third option that has also been suggested is to use the implied volatilities of the sampled options as weights, a method explored by Cont (2005).

In the Heston model there are essentially four (risk-neutral) parameters that need estimation: $\kappa > 0$, $\theta > 0$, $\sigma > 0$ and $\rho \in [-1, 1]$.

The common approach adopted to overcome this estimation problem, is considering the initial volatility $v_0 > 0$ as an additional parameter in the calibration procedure. An alternative approach can be perfomed with at-the-money (ATM) implied variance, based on the following result from Gatheral (2006).

Theorem 3.3.2. Term structure of the Black-Scholes implied volatility in the Heston Model:

$$\sigma_{ATM}^2 \approx \frac{1}{T} \int_0^T [(v_0 - \theta^{'})e^{-\kappa^{'} t} + \theta] dt = (v_0 - \theta^{'}) \frac{1 - e^{-\kappa^{'} T}}{\kappa^{'} T} + \theta^{'} \tag{3.104}$$

where $\kappa^{'} = \kappa - \frac{1}{2}\rho\sigma$ and $\theta^{'} = \kappa\theta/\kappa^{'}$. The ATM Black-Scholes implied variance in the Heston model converges (in probability) to v_0, as $T \to 0$.

The practical significance of the previous theorem is that, if we assume that the stock process follows the Heston dynamics, then v_0 should be consistent with the short dated at-the-money volatility: there is a linear relationship between the initial variance, v_0, and the Black-Scholes implied variance returned by the Heston model. This estimation method has been considered as a satisfactory estimate for the initial variance v_0, in the Heston model.

Following the method by Dragulescu and Yakovenko (2002), we apply such a method to the calibration of the Heston model and we give mathematical and numerical justification of this approach. In the following, we will consider the set $\mathcal{P}$ of all non-negative Lebesgue-integrable functions $f : \mathbb{R} \to R$ such that $f(v) = 0$, for any $v \leq 0$, almost everywhere (a.e.), and

$$\int_0^{+\infty} f(v)dv = 1. \tag{3.105}$$

Clearly $\mathcal{P}$ is a subset of all probability density functions on $\mathbb{R}$. For every $f \in \mathcal{P}$ and any bounded measurable function $\phi : (0, \infty) \to \mathbb{R}$, let define

$$\mathbb{E}_f[\phi] := \int_0^{+\infty} \phi(v)f(v)dv \tag{3.106}$$

Let the actual price of a call option with maturity T and strike K - in the framework of the Heston model - be denoted as $C^H(S_0, v_0, T, K; \Theta)$ where

$$\Theta \in \mathcal{H} := \left\{(\kappa, \theta, \sigma, \rho) \in \mathbb{R}^4 : \kappa, \theta, \sigma > 0, 2\kappa\theta \geq \sigma^2, \rho \in [-1, 1]\right\}. \tag{3.107}$$

By *averaging over volatility*, we mean the option price functional given by

$$C^H_f(S_0, T, K; \Theta) := \mathbb{E}_f\left[C^H(0, S_0, \cdot, T, K; \Theta)\right], \tag{3.108}$$

for every $S_0 > 0$, $T > 0$, $K > 0$, $\Theta \in \mathcal{H}$, $f \in \mathcal{P}$.
So, if f is replaced with the Dirac delta function $\delta(v - v_0)$ centered at $v_0 > 0$, then (3.108) reduces to $C^H(S_0, v_0, T, K; \Theta)$. The case of a constant initial volatility can be obtained considering v_0 as a random variable distributed according with the density $f \in \mathcal{P}$. Let $C_i^M = C_i^{Market}(K_i, T_i)$, $i = 1, \ldots, N$ be a basket of call prices, all written on the same underlying asset with price S_0, at time $t = 0$. Let $\{w_i \geq 0$, for $i = 1, \ldots, N\}$ be a set of given weights. For the calibration purpose, we consider two objective functionals

$J : \mathcal{H} \times (0, \infty) \to [0, \infty)$ and $J^{'} : \mathcal{H} \times \mathcal{P} \to [0, \infty)$, defined as follows:

$$J(\Theta, v_0) := \sum_{i=1}^{N} w_i |C_i^M - C^H(0, S_0, v_0, T_i, K_i; \Theta)|^2, \tag{3.109}$$

$$J'(\Theta, f) := \sum_{i=1}^{N} w_i |C_i^M - C_f^H(S_0, T_i, K_i; \Theta)|^2. \tag{3.110}$$

$$\tag{3.111}$$

Let $\mathcal{P}^{'}$ be a non-empty subset of $\mathcal{P}$, then define I, I' be the infimum of J over $\mathcal{H} \times (0, \infty)$ and the inf of $J^{'}$ over $\mathcal{H} \times \mathcal{P}^{'}$, respectively.

Theorem 3.3.3. For every $f \in \mathcal{P}$, the integral in (3.108) is bounded. If $\mathcal{P}' \subseteq \mathcal{P}$ is such that for every $v_0 > 0$, there exists a sequence $\{f_n\}_n \subseteq \mathcal{P}'$ satisfying

$$\lim_{n \to \infty} \int_{\mathbb{R}} f_n(v) g(v) \, dv = g(v_0), \tag{3.112}$$

for all bounded, continuous functions $g : \mathbb{R} \to \mathbb{R}$, then $I' \leq I$. Furthermore

$$\inf_{f \in \mathcal{P}'} C_f^H(S_0, T, K; \Theta) < C^H(0, S_0, v_0, T, K; \Theta) < \sup_{f \in \mathcal{P}'} C_f^H(S_0, T, K; \Theta) \tag{3.113}$$

for any $\Theta \in \mathcal{H}$, S_0, $v_0 > 0$ T, $K > 0$.

The chain of inequalities (3.113) shows that the averaged price (3.108) yields a wider range of prices than the standard Heston model, when v_0 varies within a bounded interval, and this holds in practice. Moreover (3.108) represents a no-arbitrage price coherent with the general framework exposed in Setion 3.1.

Before to proceed with the proof of Theorem 3.3.3, we formulate an example for the set $\mathcal{P}^{'}$ which satisfies the assumption of Theorem 3.3.3. Consider the subset $\mathcal{G} \subset \mathcal{P}$ of the probability density functions associated with the Gamma distribution:

$$\mathcal{G} = \left\{ g_{\alpha,\beta} \in \mathcal{P} \, : \, g_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} 1_{x>0}, \quad \alpha, \beta > 0 \right\}. \tag{3.114}$$

We have the following result. Thus, for every $\alpha, \beta > 0$ and $z \in \mathbb{C}$, with $\mathrm{Re}(z) < \beta$, the moment generating function of $g_{\alpha,\beta} \in \mathcal{G}$ is given by

$$M_{\alpha,\beta}(z) = \int_0^\infty e^{zv} \frac{\beta^\alpha}{\Gamma(\alpha)} v^{\alpha-1} e^{-\beta v}\, dv = \left(1 - \frac{1}{\beta} z\right)^{-\alpha}. \tag{3.115}$$

The set $\mathcal{G}$ satisfies condition (3.112). In fact, let $v_0 > 0$ and $g_n \in \mathcal{G}$ be such that $g_n = g_{\alpha_n,\beta_n}$, with $\alpha_n = n$ $\beta_n = n/v_0$, $n > 1$ integer. The characteristic function of g_n is

$$\phi_n(t) = \left(1 - \frac{it}{\beta_n}\right)^{-\alpha_n} = \left(1 - \frac{it\bar{v}}{n}\right)^{-n}, \tag{3.116}$$

for all $t \in \mathbb{R}$. Clearly g_n converges weakly to $\delta(\cdot - v_0)$, since $\phi_n(t) \to e^{itv_0}$, for any $t \in \mathbb{R}$, where

$$\phi(t) = \int_{\mathbb{R}} e^{itx} \delta(x - v_0) dx = e^{itv_0}, \tag{3.117}$$

which is the characteristic function associated to the delta function, centered at v_0.

Remark 3.3.4. We observe that different types of probability distributions satisfy condition (3.112). In particular we mention the Inverse Gaussian distribution (IG) with density function

$$IG_{\alpha,\beta}(x) = \left[\frac{\alpha}{2\pi x^3}\right]^{1/2} \cdot \exp\left(-\frac{\alpha(x-\beta)^2}{2\beta^2 x}\right), \tag{3.118}$$

for $x > 0$, where $\beta > 0$ is the mean and $\alpha > 0$ is the shape parameter. Then it is easy to see that IG_{α_n,v_0} converges weakly to $\delta(\cdot - v_0)$, for any sequence $\alpha_n \to \infty$. In particular, the Inverse Gaussian and Gamma distributions are special cases of the generalized Inverse Gaussian distribution (GIG) having density function

$$GIG_{a,b,p}(x) = \frac{(a/b)^{p/2}}{2K_p(\sqrt{ab})} x^{p-1} \cdot \exp\left[-\frac{1}{2}\left(ax + \frac{b}{x}\right)\right], \tag{3.119}$$

for $x > 0$, with parameters $a,\ b > 0$, $p \in \mathbb{R}$. Here K_p denotes the modified Bessel function of the second kind, where p is not an integer. This is the

Gamma distribution when $a = 2\beta$ and $b \to 0$, $p = \alpha$; it is the Inverse Gaussian if $a = \alpha/\beta^2$, $b = \alpha$ and $p = -1/2$.

Another suitable distribution is a scaled version of noncentral-χ^2 distribution. In fact, as documented in several papers, the distribution of the Heston volatility v_T, conditional on v_t, for $t < T$, is distributed according to such a distribution with parameters derived from the Heston model, see for example Broadie and Kaya (2006).

Proof of Theorem 3.3.3. The function $u(t, S, v) = S$ is a super-solution for the Dirichlet problem associated with the differential operator associated with the Heston model, since $Lu = 0$, for all $S > 0$, $v > 0$, $t \in (0, T)$, and $u(T, S, v) = S > (S - K)^+$ (payoff function for the call price). Thus by the comparison principle proved for in Costantini *et al* (2012), it holds $C^H(0, S_0, v_0, T, K; \Theta) \leq S_0$, for every $S_0 > 0$, $v_0 > 0$, K, $T > 0$, $\Theta \in \mathcal{H}$. This yields

$$0 \leq C_f^H(S_0, T, K; \Theta) = \int_0^\infty C^H(0, S_0, v, T, K; \Theta) f(v)\, dv \leq S_0 < \infty, \tag{3.120}$$

for any $f \in \mathcal{P}$.

The Jensen's inequality implies

$$\begin{aligned} |C_i^M - C_f^H(S_0, T, K; \Theta)|^2 &= |\mathbb{E}_f[C_i^M - C^H(0, S_0, \cdot, T_i, K_i; \Theta)]|^2 \\ &\leq \mathbb{E}_f\left[|C_i^M - C^H(0, S_0, \cdot, T_i, K_i; \Theta)|^2\right] \end{aligned} \tag{3.121}$$

for every $f \in \mathcal{P}$. Summing over $i = 1, \dots, N$, we get

$$J'(\Theta, f) \leq \mathbb{E}_f[J(\Theta, \cdot)] \qquad \forall\, f \in \mathcal{P}'. \tag{3.122}$$

Let $v_0 > 0$ and $\{f_n\}_n \subset \mathcal{P}'$, satisfying (3.112). Writing inequality (3.122) for f_n, leads to

$$I' \leq J'(\Theta, f_n) \leq \mathbb{E}_{f_n}[J(\Theta, \cdot)] = \int_{\mathbb{R}} J(\Theta, v) f_n(v) dv. \tag{3.123}$$

By (3.120), the function $v \mapsto J(\Theta, v)$ is bounded, for any $\Theta \in \mathcal{H}$. Hence, we can take the limit on the right-hand side as $n \to \infty$ to obtain

$$I' \leq J(\Theta, v_0), \tag{3.124}$$

for any arbitrary Θ and $v_0 > 0$, and taking the inf over $(\Theta, v_0) \in \mathcal{H} \times (0, \infty)$. This proves the inequality $I' \leq I$.

We prove the inequality for the sup in (3.113), the other relations can be obtained with similar arguments. Let $\bar{v} > v_0 > 0$, and $\{f_n\}_n \subset \mathcal{P}'$ be a sequence of density functions associated to $\bar{v}$. Let $g \in C^\infty(\mathbb{R})$ be such that $0 \leq g(v) \leq 1$, everywhere, $g(v) = 0$, for $|v| \geq 1$, $g(v) = 1$ for $|v| \leq 1/2$. Define $g_\varepsilon(v) = g\left(\frac{v-\bar{v}}{\varepsilon}\right)$, for every $\varepsilon > 0$. Since $v \mapsto C^H(0, S_0, v, T, K; \Theta)$ is strictly increasing, for every $0 < \varepsilon < \bar{v} - v_0$, we can write

$$\begin{aligned} C^H_{f_n}(S_0, T, K; \Theta) &\geq \int_{\bar{v}-\varepsilon}^{\bar{v}+\varepsilon} C^H(0, S_0, v, T, K; \Theta) f_n(v)\, dv \\ &\geq C^H(0, S_0, \bar{v} - \varepsilon, T, K; \Theta) \int_0^\infty g_\varepsilon(v) f_n(v)\, dv. \end{aligned}$$

Taking the limit as $n \to \infty$, we get

$$\begin{aligned} \sup_{f \in \mathcal{P}^\partial} C^H_f(S_0, T, K; \Theta) &\geq C^H(0, S_0, \bar{v} - \varepsilon, T, K; \Theta) \\ &> C^H(0, S_0, v_0, T, K; \Theta) \end{aligned}$$

for any $S_0 > 0$, $v_0 > 0$, T, $K > 0$, $\Theta \in \mathcal{H}$. □

Following Theorem 3.3.3 the calibration procedure using the average call price (3.108), under the Gamma distribution set $\mathcal{G}$, consists in minimizing the functional

$$I_{\mathcal{G}}(\alpha, \beta, \Theta) := \sum_{i=1}^{N} \omega_i |C_i^M - \mathbb{E}_{g_{\alpha,\beta}}[C^H(0, S_0, \cdot, T_i, K_i; \Theta)]|^2 \tag{3.125}$$

over $(\alpha, \beta, \Theta) \in (0, \infty)^2 \times \mathcal{H}$. Thus, the calibration is achieved by adding to the set of parameters $\mathcal{H}$ two real parameters that describe the distribution of the initial volatility v_0. We also observe that the averaged call price (3.108) is strictly increasing with respect to the scale parameter $\gamma = 1/\beta$. In fact, by the regularity properties of the Heston call price it holds:

$$\begin{aligned}
&\frac{\partial}{\partial\gamma}\mathbb{E}_{g_{\alpha,\beta}}[C^H(0,S_0,\cdot,T,K;\Theta)] \\
&= \frac{\partial}{\partial\gamma}\int_0^\infty C^H\left(S_0,\gamma w,T,K;\Theta\right)\frac{1}{\Gamma(\alpha)}w^{\alpha-1}e^{-w}\,dw \\
&= \int_0^\infty \frac{\partial C^H}{\partial v}\left(S_0,\gamma w,T,K;\Theta\right)\frac{1}{\Gamma(\alpha)}w^{\alpha-1}e^{-w}\,dw > 0. \quad (3.126)
\end{aligned}$$

Hence we can conjecture that the scale parameter represents an estimate of the "true" volatility in the Heston model.

3.4 Average call price formula

In this section we derive a closed-form formula for the averaged call price (3.108), given a probability density function $\Pi \in \mathcal{P}$.

$$C_\Pi(S_0,T,K;\Theta) = \int_0^{+\infty} C^H(S_0,v,T,K;\Theta)\Pi(v)\,dv. \quad (3.127)$$

We will state a result that yields a simplified form, reducing the expression of the price above to a single integration. This will be of great convenience for numerical computation purposes.

Theorem 3.4.1. (**Average Call Price**) If $\Pi \in \mathcal{P}$ satisfies $\mathbb{E}_\Pi[v] < \infty$, then the following relation holds true:

$$C_\Pi(S_0,T,K;\Theta) = S_0Q_1(S_0,T,K;\Theta) - e^{-rT}KQ_2(S_0,T,K;\Theta), \quad (3.128)$$

where

$$Q_j = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\Big[\frac{e^{C_j(T,\phi)+i\phi\log(\frac{S_0}{K})}M_\Pi\big(D_j(T,\phi)\big)}{i\phi}\Big]\,d\phi, \quad (3.129)$$

for $j = 1,\ 2,\ S_0,\ T,\ K > 0,\ r > 0,\ \Theta \in \mathcal{H}$, with $\rho \in (-1,1)$, M_Π being the moment generating function (3.96) related to Π.

Remark 3.4.2. If Π is the pdf associated to the Gamma distribution with parameters (α, β), the integrand in (3.129) reduces to

$$Q_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left[\frac{e^{C_j(T,\phi)+i\phi\log(\frac{S_0}{K})}}{i\phi\big(1-\beta D_j(T,\phi)\big)^\alpha} \right] d\phi. \tag{3.130}$$

In the cases of the Inverse Gaussian distribution (IG) and the Generalized Inverse Guassian distribution (GIG), we can also find an explicit expression for Q_j which are based on the moment generating function of these distributions, respectively given by:

$$M_{IG}(z) = \exp\left[\frac{\alpha}{\beta}\left(1 - \sqrt{1 - \frac{2\beta^2 z}{\alpha}} \right) \right], \tag{3.131}$$

$$M_{GIG}(z) = \left(\frac{a}{a-2z} \right)^{p/2} \frac{K_p(\sqrt{b(a-2iz)})}{K_p(\sqrt{ab})}, \tag{3.132}$$

where K_p is a modified Bessel function of the second kind.

We also need to prove the following Lemma that is not available, for our knowledge, in the actual literature.

Lemma 3.4.1. For any $\phi, \kappa, \theta, \sigma > 0$, $\rho \in (-1, 1)$, $\tau > 0$ we have that

$$\mathrm{Re}(D_j(\tau, \phi)) \leq 0, \tag{3.133}$$

for $j = 1, 2$.

Proof of Theorem 3.4.1. . According to Heston model (2.25),

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty p_j(\phi, \log(S_0), v)\, d\phi, \qquad j = 1,\, 2. \tag{3.134}$$

Therefore it suffices to prove the equation $Q_j = \frac{1}{2} + \frac{1}{\pi} I$, where

$$I = \int_0^\infty \int_0^\infty p_j(\phi, \log(S_0), v)\Pi(v)\, d\phi\, dv. \tag{3.135}$$

Let's first verify the convergence of I. To this end, Lemma 2.1.1-2.1.2 yields

$$\mathrm{Re}(C_j(T, \phi)) < -c\phi, \tag{3.136}$$

as $\phi > \eta$, with $\eta > 0$ chosen large enough, where $c = \kappa\theta T\sqrt{1-\rho^2}/(2\sigma) > 0$. Moreover, it is easy to see that $\mathrm{Im}(D_j(T,\phi)) \to 0$, as $\phi \to 0^+$. Therefore, by Lemma 2.1.1-2.1.2, setting $v = 0$, there exists $\varepsilon > 0$ such that

$$|p_j(\phi, \log(S_0), 0)| \leq 2|\log(S_0/K)| + 2|\mathrm{Im}(\partial_\phi C_j(0,T))|, \quad (3.137)$$

$$|\mathrm{Im}(D_j(T,\phi))| \leq d_j\phi, \quad (3.138)$$

for every $0 < \phi < \varepsilon$, for a constant coefficient d_j depending only on the coefficients in Lemma 2.1.1-2.1.2.

Now it suffices to show that the integrals I_∞ and I_0, defined below, are bounded. They represent respectively the integral on $(\phi, v) \in (\eta,\infty)\times(0,\infty)$ and the integral over $(\phi, v) \in (0,\varepsilon)\times(0,\infty)$. Let us consider them separately:

$$\begin{aligned} I_\infty =& \int_0^\infty \left[\int_\eta^\infty \Pi(v)\frac{1}{\phi}e^{\mathrm{Re}(C_j)+\mathrm{Re}(D_j)v}\left|\sin\left(\mathrm{Im}(C_j)+\mathrm{Im}(D_j)v+\phi\log\frac{S_0}{K}\right)\right| d\phi\right]dv \\ \leq& \int_0^\infty \left[\Pi(v)\int_0^\infty \frac{e^{\mathrm{Re}(C_j)}}{\phi}e^{\mathrm{Re}(D_j)v}\,d\phi\right]dv \leq \int_\eta^\infty \frac{e^{\mathrm{Re}(C_j)}}{\phi}\,d\phi < \infty. \end{aligned}$$

Here we have omitted the dependence of C_j and D_j on (T,ϕ), and we have used Lemma 3.4.1 in the last inequality. Still using Lemma 3.4.1 and inequality $|\sin(x)| \leq |\sin(y)| + |x-y|$, for any pair of real numbers x, y, we get

$$\begin{aligned} I_0 =& \int_0^\infty \Pi(v)\left[\int_0^\varepsilon \frac{1}{\phi}e^{\mathrm{Re}(C_j)+\mathrm{Re}(D_j)v}\left|\sin\left(\mathrm{Im}(C_j)+\mathrm{Im}(D_j)v+\phi\log\frac{S_0}{K}\right)\right| d\phi\right]dv \\ \leq& \int_0^\infty \Pi(v)\left[\int_0^\varepsilon \frac{1}{\phi}e^{\mathrm{Re}(C_j)}\left[\left|\sin\left(\mathrm{Im}(C_j)+\phi\log\frac{S_0}{K}\right)\right| + |\mathrm{Im}(D_j)|v\right]d\phi\right]dv \\ \leq& \int_0^\varepsilon \left|p_j(\phi,\log(S_0),0)\right| d\phi + d_j\varepsilon\,\mathbb{E}_\Pi\left[v\right]. \end{aligned}$$

Hence, I_0 is bounded under the assumption $\mathbb{E}_\Pi[v] < \infty$. So we are allowed to change the order of integration in (3.135):

$$
\begin{aligned}
I &= \int_0^\infty \left[\int_0^\infty \mathrm{Re}\Big[\frac{e^{-i\phi \log(K)} f_j(\phi; \log(S_0), v)}{i\phi} \Big] \Pi(v) dv \right] d\phi \\
&= \int_0^\infty \left[\mathrm{Re}\Big[\int_0^\infty \frac{e^{-i\phi \log(K)} f_j(\phi; \log(S_0), v)}{i\phi} \Pi(v) \Big] dv \right] d\phi \\
&= \int_0^\infty \mathrm{Re}\Big[\frac{e^{i\phi \log \frac{S_0}{K} + C_j(T,\phi)}}{i\phi} \int_0^\infty e^{D_j(T,\phi)v}\, \Pi(v) dv \Big] d\phi \\
&= \int_0^\infty \mathrm{Re} \left[\frac{e^{i\phi \log \frac{S_0}{K} + C_j(T,\phi)}}{i\phi} M_\Pi(D_j(T,\phi)) \right] d\phi. \qquad (3.139)
\end{aligned}
$$

We remark that $\mathrm{Re}(D_j(T,\phi)) \leq 0$ implies that $M_\Pi\big(D_j(T,\phi)\big)$ is well defined for all $T > 0$, $\phi > 0$. Note also that the switch between the integral and the real is allowed since

$$
\begin{aligned}
\left| \frac{e^{-i\phi \log(K)} f_j(\phi; \log(S_0), v)}{i} \Pi(v) \right| &= e^{\mathrm{Re}(C_j(T,\phi)) + \mathrm{Re}(D_j(T,\phi))v} \Pi(v) \\
&\leq e^{\mathrm{Re}(C_j(T,\phi))}\, \Pi(v), \qquad (3.140)
\end{aligned}
$$

and, given that $\mathrm{Re}(C_j(T,\phi))$ does not depend on v, this function is integrable with respect to $v \in (0,\infty)$, for any ϕ. □

Proof of Lemma 3.4.1. As is well known (see Heston (1993)), $\tau \mapsto D_j(\tau,\phi)$ solves, for every $\phi > 0$, a Riccati type equation:

$$
\frac{\partial D_j}{\partial \tau}(\tau,\phi) = A_j(\phi) - B_j(\phi) D_j(\tau,\phi) + R D_j^2(\tau,\phi), \qquad (3.141)
$$

and $D_j(0,\phi) = 0$, where

$$
\begin{aligned}
A_j(\phi) &= i u_j \phi - \frac{1}{2}\phi^2 \\
B_j(\phi) &= b_j - \rho\sigma\phi i \\
R &= \frac{1}{2}\sigma^2.
\end{aligned}
$$

Thus, the function $w(\tau,\phi) = \exp\left(-R \int_0^\tau D_j(t,\phi) dt\right)$ solves the second order differential equation

$$
\partial_\tau^2 w(\tau,\phi) + B_j(\phi)\partial_\tau w(\tau,\phi) + R A_j(\phi) w(\tau,\phi) = 0. \qquad (3.142)
$$

In the sequel we shall use the notation w' to denote the partial derivative $\partial_\tau w(\tau,\phi)$ and $\bar{w}(\tau,\phi)$ for the conjugate of $w(\tau,\phi)$. Computing the real part of D_j leads to

$$\mathrm{Re}(D_j) = -\frac{1}{R}\mathrm{Re}\left(\frac{w'\bar{w}}{|w|^2}\right) = -\frac{1}{R}\frac{w'_R w_R + w'_I w_I}{|w|^2},$$

where $w = w_R + iw_I$ and $w' = w'_R + iw'_I$. Let $\xi(\tau,\phi) = |w(\tau,\phi)|^2$, then of course we can write $\xi' = w'\bar{w} + w\bar{w}'$ and $\mathrm{Re}(D_j) = -\xi'/(2R\xi)$. Moreover, by using equation (3.142), we find

$$\begin{aligned}
\xi'' &= -B_j w'\bar{w} - A_j R w\bar{w} + 2w'\bar{w}' - w(\bar{B}_j\bar{w}' + \bar{A}_j R\bar{w}) \\
&= -\mathrm{Re}(B_j)(w'\bar{w} + w\bar{w}') - i\mathrm{Im}(B_j)(w'\bar{w} - w\bar{w}') - 2R\,\mathrm{Re}(A_j)\xi \\
&\quad + 2w'\bar{w}' = -\mathrm{Re}(B_j)\xi' + 2\mathrm{Re}(iw\bar{w}')\mathrm{Im}(B_j) - 2R\mathrm{Re}(A_j)\xi + \\
&\quad +2w'\bar{w}' = -b_j\xi' - 2\mathrm{Re}(iw\bar{w}')\sigma\rho\phi + \frac{1}{2}\sigma^2\phi^2\xi + 2w'\bar{w}',
\end{aligned}$$

and, by the definition of w, we argue that

$$\begin{aligned}
&\mathrm{Re}(iw\bar{w}') = -R\,\mathrm{Im}(D_j)|w|^2 = -R\,\mathrm{Im}(D_j)\,\xi, \\
&|w'|^2 = R^2\xi|D_j|^2.
\end{aligned}$$

By these relations, we deduce that ξ solves the Cauchy problem:

$$\begin{cases}
\xi''(\tau,\phi) + b_j\xi'(\tau,\phi) - \gamma(\tau,\phi)\xi(\tau,\phi) = 0, \\
\xi'(0,\phi) = 0, \\
\xi(0,\phi) = 1,
\end{cases} \tag{3.143}$$

where $\gamma(\tau,\phi) = (1/2)\sigma^2\phi^2 + 2[R^2|D_j(\tau,\phi)|^2 + R\sigma\rho\phi\,\mathrm{Im}(D_j(\tau,\phi))]$. In order to prove that, for every fixed $\phi > 0$, $\mathrm{Re}(D_j(\tau,\phi)) \leq 0$, for any $\tau > 0$, it suffices to show that $\xi'(\tau,\phi) \geq 0$, for all $\tau > 0$. Assume that this is not true, by way of contradiction. Suppose that there exists $\tau' > 0$ such that $\xi'(\tau',\phi) < 0$. Since $D_j(0,\phi) = 0$, (3.143) yields

$$\xi''(0,\phi) = \frac{1}{2}\sigma^2\phi^2 + 2\gamma(0) = \frac{1}{2}\sigma^2\phi^2 > 0. \tag{3.144}$$

Hence, by the continuity of ξ', ξ'' as functions of τ, the supremum

$$\tau_0 = \sup\{\tau \in (0,\tau') : \xi'(\tau,\phi) \geq 0\}, \tag{3.145}$$

is well defined, since the related set is non-empty, and $0 < \tau_0 < \tau'$, $\xi'(\tau_0, \phi) = 0$. We show that $\xi''(\tau_0, \phi) > 0$. From the differential equation in (3.143) and $\mathrm{Re}(D_j(\tau_0, \phi)) = 0$, this is equivalent to state the inequality

$$\phi^2 + 2\rho\phi A + A^2 > 0, \tag{3.146}$$

where $A = \sigma \mathrm{Im}(D_j(\tau_0, \phi))$. If $A = 0$ the last inequality reduces to $\phi^2 > 0$, that is obviously true. Otherwise, if $A \neq 0$, since $|\rho| < 1$, we get

$$\phi^2 + 2\rho\phi A + A^2 > \phi^2 - 2|A|\phi + A^2 = (\phi - |A|)^2 \geq 0. \tag{3.147}$$

Thus $\xi'(\tau, \phi)$ is positive in a right neighborhood of τ_0. This is in contradiction with the definition of τ_0 and we have proved that $\xi'(\tau, \phi) \geq 0$ for every $\tau,\ \phi > 0$, implying inequality (3.133). □

3.5 Calibration to option prices

The general approach to the calibration of parametric models, such as the Heston model, is to apply a least-square type procedure either in price or implied volatility. Unfortunately, this kind of approach will in general be very sensitive to the choice of the initial point, which will often in practice drive the selection of the local minima the algorithm will converge to. The various explicit formulas come into play to receive a pertaining initial point.

Estimates for the volatility parameter v_0, with the structural parameters $\{\kappa,\ \theta,\ \rho, \sigma\}$ will be needed. The calibration procedure consists in the minimization of the functional in (3.103), where C^{Model} is the Heston call price $C^H(S_0, v_0, T, K; \Theta)$ in the standard case or, otherwise the weighted average call price $C_\Pi(S_0, T, K; \Theta)$, for a given probability distribution density $\Pi \in \mathcal{P}$. In this second case, the density is chosen according to a parameterized family of density functions related to a probability distribution, implying that the set of parameters includes also the parameters of such a distribution.

By the results of Section 3.3, we have compared the standard method which considers v_0 as an additional parameter and our approach under three distributions: the Gamma (GAM), the Inverse Gaussian (IG) and the Generalized Inverse Gaussian (GIG), for which the integrands appearing are explicitly known. From a numerical point of view, the calculation of the option price is made somewhat complicated by the fact that the integrands have oscillatory nature. However, the integration can be done in a reasonably simple fashion by the aid of Gauss-Lobatto quadrature. This integration method is capable of handling a wide range of functional forms. Since the Gauss-Lobatto algorithm is designed to operate on a closed bounded interval, we have used a transformation of the original integral boundaries $(0, \infty)$ to the finite interval $[0, 1]$, as presented in Kahl and Jackel (2005).

In order to evaluate the performance of those methods, we have used three discrepancy measures documented in several works in the literature: the average prediction error (APE), the root mean-square error (RMSE) and the average relative prediction error (ARPE). They are defined as follows:

$$APE = \sum_{n=1}^{N} \frac{|C_i^{Model} - C_i^{Market}|}{\sum_{n=1}^{N} C_i^{Market}} \tag{3.148}$$

$$ARPE = \frac{1}{N} \sum_{i=1}^{N} \frac{|C_i^{Model} - C_i^{Market}|}{C_i^{Market}} \tag{3.149}$$

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{|C_i^{Model} - C_i^{Market}|^2}{N}}, \tag{3.150}$$

where $C_i^{Model} = C^H(S_0, v_0, T_i, K_i; \Theta)$ under the simple Heston model (H) and $C_i^{Model} = C_\Pi(S_0, T_i, K_i; \Theta)$ under the average price method denoted WAPH in the following. T_i and K_i denote respectively the maturity and the strike price of the ith option, all written on a stock with current price S_0. The admissible parameter set for the Heston model has been specified in (3.107).

3.5.1 Estimation results

Our empirical analysis is conducted on a dataset of option prices on the Standard and Poor's 500 Index, which represents the main capitalization-weighted index of 500 stocks in the US market. The index is designed to measure performance of the broad domestic economy through changes in the aggregate market value of 500 stocks representing all major industries. We have considered a first dataset composed by option prices from September 1, 2010 to September 30, 2010 and further dataset of market prices from September 1, 2015 to September 30, 2015. Only call prices that verify standard no-arbitrage bounds have been considered. Moreover, we have tested the model on call options, with the constraint on the moneyness: $0.9 < M < 1.1$, where $M = \frac{K}{S_0}$. Overall, we have considered $8,315$ call prices divided into 21 trading dates and 9 expiry dates for the 2010's set and $1,091$ call prices divided into 21 trading dates and 7 expiry dates for the 2015's set. For each of the considered models for the density function, driving the initial volatility, we have calibrated everyday the corresponding parameters that are 5 in the standard Heston model (denotes as H), where v_0 is a parameters, 6 in both cases of WAPH with GAM and with IG, 7 in the case of WAPH with GIG.

The results of the estimation are summarized in Table 3.3 and in Table 3.5. Precisely, the averages of daily error measures and the parameters for all methods are reported. For what concerns all the error measures, the averaged call price under the GIG distribution seems to perform better in the considered period, while for what concerns the case with the Gamma distribution (GAM) and the Inverse Gaussian distribution (IG) we observe a substantial equality of the performance of these approaches. In fact, the GIG includes the GAM and the IG as special cases. The Heston model does not perform badly, but it is systematically beaten by the weighted average price model, especially for what concerns the RMSE criterion.

In order to better analyze model performance, we have estimate the implied Black-Scholes volatility and for each trading day. Every trading day is associated with the standard deviation of the implied volatility (σ_σ) and the results of the parameter estimation are clustered considering different levels of the standard deviation. In Table 3.7 cluster average RMSE are reported, the results show that the GIG model error does not depend on

	APE	**RMSE**	**ARPE**
Heston	2.1905	0.5182	3.5480
WAPH-GAM	1.2878	0.2145	2.3781
WAPH-IG	1.1230	0.4167	2.5181
WAPH-GIG	1.3400	0.1104	1.0824

Table 3.3: Averages of the daily error measures APE, RMSE and ARPE for the different pricing methods on 2010 database

	APE	**RMSE**	**ARPE**
Heston	4.5448	3.4457	0.0844
WAPH-GAM	3.9909	3.4239	0.1908
WAPH-IG	4.0201	3.1184	0.0923
WAPH-GIG	3.7336	2.9810	0.0700

Table 3.4: Averages of the daily error measures APE, RMSE and ARPE for the different pricing methods on 2015 dataset

the standard deviation of the implied volatility and the conclusion that can be drawn is that clusters with more variability are better described by GIG model.

	k	θ	σ	ρ	v_0		
Heston	0.0252	4.4944	0.4599	-0.6062	0.0024		
	k	θ	σ	ρ	α (shape)	β (rate)	
GAM	0.1178	0.9404	0.4567	-0.6057	0.0016	2.7953	
	k	θ	σ	ρ	α (shape)	β (mean)	
IG	0.1053	0.8843	0.5103	-0.6342	0.0023	0.0082	
	k	θ	σ	ρ	a	b (mean)	p
GIG	0.0145	1.2084	0.4432	-0.6401	5.4502	0.0001	0.0027

Table 3.5: Averages of the daily estimated parameters under the considered pricing methods for 2010 dataset.

	k	θ	σ	ρ	v_0		
Heston	6.5140	0.0919	0.4732	-0.9999	0.0471		
	k	θ	σ	ρ	α (shape)	β (rate)	
GAM	4.3786	0.0282	0.4770	-0.9999	0.0100	4.3888	
	k	θ	σ	ρ	α (shape)	β (mean)	
IG	0.1505	0.2313	0.3483	0.8742	0.0067	0.0124	
	k	θ	σ	ρ	a	b (mean)	p
GIG	4.4033	0.0233	0.2754	-0.8743	2.7055	0.0675	101.0082

Table 3.6: Averages of the daily estimated parameters under the considered pricing methods for the 2015 dataset.

	Heston	**WAPH-GIG**
$1.5\% < \sigma_\sigma \leq 2.0\%$	3.9526	2.5608
$2.0\% < \sigma_\sigma \leq 2.5\%$	4.6239	2.1010
$2.5\% < \sigma_\sigma \leq 3.0\%$	4.1979	1.8574
$3.0\% < \sigma_\sigma \leq 3.5\%$	5.5774	3.1035
$3.5\% < \sigma_\sigma \leq 4.0\%$	6.4538	2.7909
$4.0\% < \sigma_\sigma \leq 4.5\%$	7.8989	3.0723

Table 3.7: Average RMSE for different standard deviation cluster.

A natural extension of the Heston model is to include jumps in the stock price process. In the continuous time setting, Bates (1996), among others, examines the empirical performance of an affine stochastic volatility model (SVJ) using index returns and option data. Bates's model is a Heston process with an added Merton log-normal jump. Other benchmark works for continous-time affine SVJ models has been considered in Pan (2002) and Bates (2006), and variants have been estimated on stock index returns by Andersen *et al*. (2002), Chernov *et al*. (2003),and Eraker *et al*. (2003). These kind of models significantly outperform the Black-Scholes model into fitting the observed implied volatility surface. Intuitively, it makes sense that a jump in the stock price process should trigger a correlated jump in the volatility process in that sudden, large movements in the stock price would cause increased market anxiety around that stock. Therefore, Bates inspired models have been extended by including jumps in the volatility process in addition to those in the stock price process (SVJJ models). In particular, contributions by Broadie *et al*. (2007) and Gatheral (2006) explore the merits and drawbacks of the SVJJ model over Bates-style models. Broadie *et al*. (2007) argue in favour of a stochastic volatility model that incorporates jumps in both the stock price and variance processes, while Gatheral finds that a stochastic volatility model with jumps in the stock price process only produces the best fit to the implied volatility surface. Here, in order to provide a comparison between our technique and the SVJJ model, we have considered the Heston stochastic volatility double jump-diffusions model. However, we remark that such a comparison may not be appropriate since SVJJ models potentially have a superior option market fit while keeping a sound balance between reality and tractability and it allows also a range of jump amplitude distributions. On the other hand, our approach could be extended in order to include the case of models with jumps. In the formulation of the jump stochastic volatility model (SVJJ), we have considered closely [28], with the following risk-neutral dynamics:

$$\begin{cases} dS_t = (r - \lambda\mu_J)S_t dt + S_t\sqrt{v_t}d\tilde{W}_t^1 + J_t S_t dN_t, \\ dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}\left(\rho d\tilde{W}_t^1 + \sqrt{1-\rho^2}d\tilde{W}_t^2\right) + Z_t d\tilde{N}_t. \end{cases} \tag{3.151}$$

Here $(\tilde{W}_t^1)_{t\geq 0}$ and $(\tilde{W}_t^2)_{t\geq 0}$ are independent Brownian motions, $(\tilde{N}_t)_{t\geq 0}$ represents a Poisson process under the risk-neutral measure, independent of W^i, for $i = 1, 2$, with jump intensity $\lambda > 0$. The jump terms in the model are defined through processes $(Z_t, J_t)_{t\geq 0}$ independent of the Brownian motions and the Poisson Process, such that

$$\begin{aligned} Z_t &\sim \text{Exponential}(\mu_V), \\ (1 + J_t)\,|\,Z_t &\sim \text{Log-normal}(\mu_S + \rho_J Z, \sigma_S^2), \end{aligned} \tag{3.152}$$

for each $t > 0$, where $\mu_S \in \mathbb{R}$, $\mu_V > 0$, $\sigma_S > 0$, $\rho_J < 1/\mu_V$ are constant coefficients and

$$\mu_J = \frac{e^{\mu_S + \frac{\sigma_S^2}{2}}}{1 - \rho_J\mu_V} - 1. \tag{3.153}$$

Moreover κ, θ, ρ and σ satisfy the usual constraints for the Heston model, already described in previous sections. For the computation of call prices using model (3.151), we refer the reader to [28]. Table 3.8 includes the estimation of model parameters of the WAPH-GIG model and the SVJJ model (3.151). Even if the SVJJ model implies superior option market fit, there is evidence, from our experimental results, that the WAPH-GIG and SVJJ provide similar performances. This empirical analysis shows that our

	k	θ	σ	ρ	a	b (mean)	p
WAPH-GIG	2.9894	0.0242	0.3796	-0.8499	1.5416	0.0348	15.8841
	k	θ	σ	ρ	v_0		
SVJJ	4.9689	0.0352	0.5913	-0.9999	0.0580		
	λ	μ_S	σ_S	ρ_J	μ_V		
	0.0254	-0.8516	1.4978	-0.0440	39.9852		

Table 3.8: Pricing models estimated parameters, based on September 1, 2015 dataset.

approach is quite promising and represents an improvement over the Heston model, while retaining the same degree of analytical tractability.

	APE	RMSE	ARPE
WAPH-GIG	2.5048	2.9857	0.0491
SVJJ	2.3049	3.0306	0.0452

Table 3.9: Comparison between the error measures for the weighted pricing approach and the SVJJ model, based on September, 2015 dataset.

3.6 A PDE for the Weighted Average Price in the Heston model

In this section we show that the weighted average price function (3.108) corresponds, essentially, to an equivalent one-dimensional model, by developing a partial differential equation for the price function. Then, in Chapter 4, the corresponding Dirichlet problem will be solved by implementing the FD method. In this specific case, we have a one-dimensional problem that reduces the complexity of the weight for the numerical procedure. This is obviously lower in comparison, for instance, with ADI schemes for the Heston pde. Since we are interest the time evolution of the average price (3.108), it is reasonable to assume that, the initial volatility, at time $t \geq 0$ is driven by a probability distribution with density function $\pi = \pi(t, v) \in C^2((0, \infty)^2)$ satisfying the Fokker-Planck equation associated with the variance process in the Heston model:

$$\frac{\partial \pi}{\partial t} = \frac{\partial}{\partial v} \left(\kappa(v - \theta)\pi\right) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial v^2} \left(v\pi\right), \tag{3.154}$$

for any $(t, v) \in (0, \infty)^2$. Thus, for a given density function $\Pi \in L^1((0, \infty))$, for the volatility at time $t = 0$, such that $\mathbb{E}_\Pi[v] < \infty$, the unique solution to (3.154) is obtained by using the transition density function $p(t, v|v_0)$, from v_0, at time $t = 0$, to v, at time t, of the volatility process $p(t, v|v_0)$ that, as proved in [34], is the solution of (3.154) satisfying the initial condition $p(0, v|v0) = \delta(v - v_0)$, provided that $\theta, \sigma, \kappa > 0$ and $q := \frac{2\kappa\theta}{\sigma^2} - 1 \geq 0$.

Actually, we can write

$$\pi(t,v) = \int_0^\infty p(t,v|v_0)\Pi(v_0)dv_0. \tag{3.155}$$

By Karlin and Taylor [50] (page 220), the transition density takes the following explicitly form:

$$p(t,v|v_0) = c\, e^{-\eta - cv}\left(\frac{cv}{\eta}\right)^{q/2} I_q\left[2(cv\eta)^{1/2}\right], \tag{3.156}$$

where $c = 2\kappa e^{\kappa t}/[\sigma^2(e^{\kappa t}-1)]$, $\eta = c\,v_0 e^{-\kappa t}$, and I_q denotes the modified Bessel function of the first kind of order q. It is easy to see that the following properties are also satisfied for all $t > 0$:

$$\begin{cases} \int_0^\infty \pi(t,v)dv = 1, \\ f(t,v) \to 0, \quad \text{as } v \to 0^+, \text{ and as } v \to \infty, \\ \text{for } f(t,v) = v\pi(v,t) \text{ and } f(t,v) = \frac{\partial}{\partial v}[v\pi(t,v)]. \\ \frac{\partial^2}{\partial v^2}[v\pi(t,v)] \text{ is bounded as } v \to 0^+ \end{cases} \tag{3.157}$$

Therefore, we consider the weighted average price for a call option at time $0 \le t < T$, with strike price $K > 0$ and maturity T:

$$C_\pi(t,S;K,T) = \int_0^{+\infty} C^H(t,S,v;K,T)\pi(v,t)dv, \tag{3.158}$$

where C^H is the Heston call price function (2.26). For the sake of simplicity, we omit the dependence on the vector of model parameters $\Theta \in \mathcal{H}$.

Proposition 3.6.1. If $\rho = 0$, $\theta, \sigma, \kappa > 0$ and $q \ge 0$, then there exists a unique, strictly positive, function $V_\pi = V_\pi(t,S;K,T)$, with $V_\pi(\cdot,\cdot;K,T) \in C^{1,2}((0,T)\times(0,\infty))$, such that the weighted average price function C_π satisfies the following final value problem:

$$\frac{\partial C_\pi}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2}[V_\pi C_\pi] + rS\frac{\partial C_\pi}{\partial S} = rC_\pi, \tag{3.159}$$

for any $t \in (0,T)$, $S > 0$, with $C_\pi(T,S) = \max(K-S,0)$.

Proof of Proposition 3.6.1. We recall that C^H is the solution of

$$\begin{aligned}\frac{\partial C^H}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 C^H}{\partial S^2} &+ \frac{1}{2}v\sigma^2\frac{\partial^2 C^H}{\partial v^2} + k(\theta - v)\frac{\partial C^H}{\partial v} + rS\frac{\partial C^H}{\partial S} \\ &- rC^H = 0,\end{aligned}$$

with $C^H(T,S) = \max(S-K,0)$. Multiplying each term in the above equation by π and integrating with respect to v on $(0,\infty)$, we easily obtain

$$\begin{aligned}\int_0^{+\infty} \frac{\partial C^H}{\partial t}\pi dv \quad &+ \int_0^{+\infty} \frac{1}{2} v S^2 \frac{\partial^2 C^H}{\partial S^2}\pi dv + \int_0^{+\infty} \frac{1}{2} v\sigma^2 \frac{\partial^2 C^H}{\partial v^2}\pi dv + \\ &+ \int_0^{+\infty} \kappa(\theta - v)\frac{\partial C^H}{\partial v}\pi dv + \int_0^{+\infty} rS\frac{\partial C^H}{\partial S}\pi dv + \\ &- \int_0^{+\infty} rC^H \pi dv = 0. \end{aligned} \tag{3.160}$$

For every $t \in (0,T)$, $S > 0$, we can apply the mean value theorem to the identity function $(v \mapsto v)$ with respect to the positive and bounded measure $C^H(t,S,v;K,T)\times \pi(t,v)dv$ on $(0,\infty)$, to find a unique $V_\pi = V_\pi(t,S;K,T) > 0$ such that

$$\int_0^{+\infty} vC^H(t,S,v;K,T)\pi(t,v)dv = V_\pi(t,S;K,T)C_\pi(t,S;K,T).$$

By the regularity of C^H and the assumptions on π, we deduce that $V_\pi(\cdot,\cdot;K,T) \in C^{1,2}((0,T)\times(0,\infty))$. Thus, we can write

$$\int_0^{+\infty} \frac{1}{2} vS^2 \frac{\partial^2 C^H}{\partial S^2}\pi(v,t)dv = \frac{1}{2}S^2\frac{\partial^2}{\partial S^2}[V_\pi(t,S;K,T)C_\pi(t,S;K,T)]$$

Let us evaluate now the third term in (3.160), using integration by parts:

$$\begin{aligned}\int_0^{+\infty} \frac{1}{2} v\sigma^2 \frac{\partial^2 C^H}{\partial v^2}\pi dv \quad &= \quad \frac{1}{2}\sigma^2\left(v\pi(t,v)\frac{\partial C^H}{\partial v}\Big|_{v=0}^{v=\infty}\right) + \\ &- \quad \frac{1}{2}\sigma^2\left(\int_0^{+\infty} \frac{\partial C^H}{\partial v}\frac{\partial}{\partial v}[v\pi(t,v)]dv\right),\end{aligned}$$

The first term is null thanks to the properties of π in (3.157) and the properties of $\partial_v C^H$ outlined by Ould in [61]. Thus, we repeat again the integration by parts, to get

$$\begin{aligned}\int_0^{+\infty} \frac{1}{2} v\sigma^2 \frac{\partial^2 C^H}{\partial v^2}\pi dv \quad &= \quad -\frac{1}{2}\sigma^2 C^H \frac{\partial}{\partial v}[v\pi]\Big|_{v=0}^{v=\infty} \\ &+ \quad \frac{1}{2}\sigma^2 \int_0^{+\infty} C^H\frac{\partial^2}{\partial v^2}[v\pi]dv,\end{aligned}$$

where, for the behavior of $\partial_v[v\pi(t,v)]$, as $v \to 0^+$ and for $v \to \infty$, the first term is equal to zero. The first in the equation (3.160) can be rewritten as:

$$\int_0^{+\infty} \frac{\partial C^H}{\partial t}\pi dv = \frac{\partial C_\pi}{\partial t} - \int_0^{+\infty} C^H \frac{\partial \pi}{\partial t}dv. \tag{3.161}$$

Thus, using the partial differential equation satisfied by π (3.154), we can evaluate the second integral in (3.161):

$$\begin{aligned} -\int_0^{+\infty} C^H \frac{\partial \pi}{\partial t} dv &= \int_0^{+\infty} C^H \kappa(\theta - v)\frac{\partial \pi}{\partial v} dv - \int_0^{\infty} \kappa \pi C^H dv + \\ &- \int_0^{+\infty} \frac{\sigma^2}{2} C^H \frac{\partial^2}{\partial v^2}(v\pi) dv, \end{aligned}$$

and integrating by parts the first integral on the right-hand side, we obtain:

$$\begin{aligned} \int_0^{+\infty} \frac{\partial C^H}{\partial t} \pi dv &= \frac{\partial C_\pi}{\partial t} + C^H \kappa(\overline{v} - v)\pi|_{v=0}^{v=\infty} - \int_0^{+\infty} \frac{\partial C^H}{\partial v} \kappa(\theta - v)\pi dv + \\ &- \int_0^{+\infty} \frac{\sigma^2}{2} C^H \frac{\partial^2}{\partial v^2}(v\pi) dv. \end{aligned}$$

Collecting all the previous results in (3.160), we get

$$\begin{aligned} &\frac{\partial C_\pi}{\partial t} - \int_0^{+\infty} \left(\frac{\partial C^H}{\partial v} \kappa(\theta - v)\pi + \frac{\sigma^2}{2} C^H \frac{\partial^2}{\partial v^2}[v\pi] \right) dv + \frac{\sigma^2}{2} S^2 \frac{\partial^2}{\partial S^2}[V_\pi C_\pi] \\ &+ \int_0^{\infty} \frac{\sigma^2}{2} C^H \frac{\partial^2}{\partial v^2}[v\pi] dv + \int_0^{\infty} \kappa(\theta - v) \frac{\partial C^H}{\partial v} \pi dv + rS\frac{\partial C_\pi}{\partial S} - rC_\pi = 0. \end{aligned}$$

Finally, we have:

$$\frac{\partial C_\pi}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2}[V_\pi C_\pi] + rS\frac{\partial C_\pi}{\partial S} = rC_\pi, \qquad (3.162)$$

for any $t \in (0, T)$, $S > 0$, and clearly $C_\pi(T, S) = \max(K - S, 0)$, that proves the result. □

Remark 3.6.1. Equation (3.159) can be easily solved from the numerical point of view thanks to the one dimensional (one space variable S) specification of the problem: the numerical implementation does not require special efforts, it produces a more stable and consistent approximation. Unfortunately, in the equation is still present a term which needs to be evaluated separately, that is the mean value function

$$V_\pi(t, S; K, T) = \frac{\int_0^\infty vC^H(t, S, v; K, T)\pi(v, t)dv}{\int_0^\infty C^H(t, S, v; K, T)\pi(v, t)dv}$$

which, under our weighted average approach, is a equivalent to a local volatility. Moreover it depends on the choice of the initial density function Π. In light of the results of previous sections, if Π is chosen in a class of density functions satisfying the assumptions of Theorem 3.3.3, the calibration problem bases on the average call price allows to reduce the error with respect to observed prices. We expect that Dupire's calibration technique, described in Chapter 1, can be adapted in order to estimate V_π, using option implied volatilities, in agreement with the partial differential equation (3.159). However this is matter for a further study that we do not develop in this thesis. We also recall that, during the years, a different approach for calibrating Local Volatility models was derived independently in several research contributions. In particular, it is shown that the Local Volatility model represents satisfactory and alternative approach to the Stochastic Volatility model. In fact, if we suppose that the underlying asset follows a diffusion process with a stochastic instantaneous variance, then we can think the local volatility as the conditional expectation of the instantaneous volatility.

Chapter 4

Finite Difference Methods

The finite difference methods (FD) for derivatives are one of the simplest and of the oldest methods to solve differential equations. It was already known by Euler, L., (1707-1783) ca. 1768, in one dimension of space and was probably extended to dimension two by Runge, C., (1856-1927) ca. 1908. Theoretical results have been obtained during the last five decades regarding the accuracy, stability and convergence of the finite difference method for partial differential equations. These methods has been used for many application areas such as fluid dynamics, heat transfer, semiconductor simulation and astrophysics, for example. In finance, in particular, acquiring an effective numerical time-discretization method for the spatially discretized problem is a key step. The reason for which finite difference methods are a popular choice for pricing options is that all options satisfy the Black-Scholes PDE or appropriate variants of it. Finite Difference methods can be applied also to American (early exercise) Options and they can also be used for many exotic contracts.

The principle of finite difference methods consists in approximating the differential operator by replacing the derivatives in the equation using differential quotients. The aim of these methods(which could be different depending on the approximation used)is to evaluate the values of a continuous function $f(t, S)$ on a set of discrete points in (t, S) plane (or in a high dimensional ones). The domain is partitioned in space and in time and approximations of the solution are computed at the space or time points: (t, S) plane becomes a mesh grid with mesh points $(i\Delta t, j\Delta S)$. We are interested in the values of $f(t, S)$ at mesh points $(i\Delta t, j\Delta S)$, denoted as

$$f_j^i = f(i\Delta t, j\Delta S) \tag{4.1}$$

Suppose the function f is C^2 in the neighborhood of x. For any $h > 0$ we have:

$$f(x+h) \approx f(x) + hf^{'}(x) + \frac{h^2}{2}f^{''}(x+h_1) \tag{4.2}$$

where h_1 is a number between 0 and h (i.e. $x + h_1$ is point of $]x, x+h[$). For the treatment of problems, it is convenient to retain only the first two terms of the previous expression:

$$f(x+h) = f(x) + hf^{'}(x) + O(h^2) \tag{4.3}$$

where the term $O(h^2)$ indicates that the error of the approximation is proportional to h^2. The approximation (4.2) is known as the forward difference approximant of $f^{'}$. Likewise, we can define the first order backward difference approximation of f' at point x as:

$$f(x-h) = f(x) - hf^{'}(x) + O(h^2) \tag{4.4}$$

Obviously, other approximations can be considered. In order to improve the accuracy of the approximation, we define a consistant approximation, called the central difference approximation, by taking the points $x - h$ and $x + h$ into account

$$f(x+h) = f(x) + hf^{'}(x) + \frac{h^2}{2}f^{''}(x) + \frac{h^3}{6}f^{'''}(\epsilon^+) \tag{4.5}$$

$$f(x-h) = f(x) - hf^{'}(x) + \frac{h^2}{2}f^{''}(x) - \frac{h^3}{6}f^{'''}(\epsilon^-) \tag{4.6}$$

where $\epsilon^+ \in]x, x+h[$ and $\epsilon^- \in]x-h, x[$. By subtracting these two expressions we obtain:

$$\frac{f(x+h) - f(x-h)}{2h} \approx f^{'}(x) + \frac{h^2}{6}(\epsilon) \tag{4.7}$$

where ϵ is a point of $]x-h, x+h[$. We could use:

- Forward difference method for the first derivative

$$f^{'}(x) = \frac{f(x+\Delta x) - f(x)}{\Delta x} + O(\Delta x) \tag{4.8}$$

- Backward difference method for the first derivative

$$f^{'}(x) = \frac{f(x) - f(x-\Delta x)}{\Delta x} + O(\Delta x) \tag{4.9}$$

- Centered difference method for the first derivative

$$f^{'}(x) = \frac{f(x+\Delta x) - f(x-\Delta x)}{2\Delta x} + O(\Delta x)^2. \tag{4.10}$$

And by adding (4.5) and (4.6) we obtain the centered difference method for the second derivative

$$f^{''}(x) = \frac{f(x+\Delta x) - 2u(x) + u(x-\Delta x)}{\Delta x^2} + O(\Delta x)^2 \tag{4.11}$$

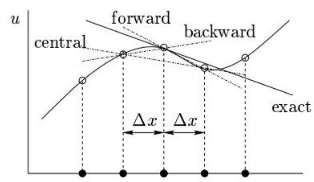


Figure 4.1: Geometric Interpretation

The approximation of f' at point x by forward and backward approximation is said to be consistent at the first order. More generally, we define an approximation at order p of the derivative as

Definition 4.0.1. The approximation of the derivative $f^{'}$ at point x is of order $p(p > 0)$ if there exists a constant $C > 0$, independent of h, such that the error between the derivative and its approximation is bounded by Ch^p (i.e. is exactly $O(h^p)$).

The error between the numerical solution and the exact solution is determined by the error that is committed by going from a differential operator to a difference operator. This error is called the discretization error or truncation error. The term truncation error reflects the fact that a finite part of a Taylor series is used in the approximation. The notion of consistency and accuracy helps to understand how well a numerical scheme approximates an equation. We introduce a formal definition of the consistency that can be used for any partial differential equation defined on a domain Ω and denoted by

$$(\mathrm{L}f)(x) = u(x), \quad \forall x \in \Omega \tag{4.12}$$

where L denotes a differential operator. The notation $(\mathrm{L}f)$ indicates that the equation depends on f and on its derivatives at any point t, x. A numerical scheme can be written, for every index i and j, in a more abstract form as:

$$(\mathrm{L}_h f)(x_j) = u(x_j), \quad \forall j \in 1, ..., M. \tag{4.13}$$

Definition 4.0.2. [Consistency] A finite difference scheme is said to be consistent with the partial differential equation it represents, if for any sufficiently smooth solution f of this equation, the truncation error of the scheme, corresponding to the vector $\epsilon_h \in \mathbb{R}^N$ whose components are defined as

$$(\epsilon_h)_j = (\mathrm{L}_h f)(x_j) - u(x_j), \; \forall j \in 1, ..., M \tag{4.14}$$

tends uniformly towards zero with respect to t and x, when h tends to zero, i.e. if:

$$\lim_{h \to 0} ||\epsilon_h||_\infty = 0. \tag{4.15}$$

Moreover, if there exists a constant $C > 0$, independent of f and of its derivatives, such that, for all $h \in]0, h_0]$ ($h_0 > 0$ given) we have:

$$||\epsilon_h|| \leq Ch^p \tag{4.16}$$

with $p > 0$, then the scheme is said to be accurate at the order p for the norm $|| \, ||$.

Lemma 4.0.1. Suppose f is a C^4 continuous function on an interval $[x - h_0, x + h_0]$, $h_0 > 0$ and exists a constant $C > 0$ such that for every $h \in]0, h_0[$ we have:

$$|\frac{f(x+h) - 2f(x) + f(x-h)}{h^2} - f^{''}(x)| \leq Ch^2 \tag{4.17}$$

The differential quotient $\frac{f(x+h)-2f(x)+f(x-h)}{h^2}$ is a consistent second-order approximation of the second derivative $u^{''}$ of u at point x.

Proof. We use Taylor expansions up to the fourth order to achive the result:

$$f(x+h) = f(x) + hf^{'}(x) + \frac{h^2}{2}f^{''}(x) + \frac{h^3}{6}f^{'''}(x) + \frac{h^4}{24}f^{\mathrm{iv}}(\epsilon^+) \tag{4.18}$$

$$f(x-h) = f(x) - hf^{'}(x) + \frac{h^2}{2}f^{''}(x) - \frac{h^3}{6}f^{'''}(x) + \frac{h^4}{24}f^{\mathrm{iv}}(\epsilon^-) \tag{4.19}$$

where $\epsilon^+ \in]x, x+h[$ and $\epsilon^- \in]x-h, x[$. Like previously:

$$\frac{f(x+h) - 2f(x) + f(x-h)}{h^2} = f^{''}(x) + \frac{h^2}{12}(\epsilon) \tag{4.20}$$

where $\epsilon \in [x-h, x+h]$. Hence, we deduce the relation (4.17) with the constant

$$C = \sup_{x \in [x-h, x+h]} \frac{|f^{iv}|}{12} \tag{4.21}$$

□

Definition 4.0.3. [Convergence] A one-step finite difference scheme approximating a partial differential equation is a convergent scheme if for any solution to the partial differential equation, $f(t, x)$, and solutions to the finite difference scheme, f_i^n, such that f_i^0 converges to $f(0, x)$ as $i\Delta x$ converges to x, then f_i^n converges to $f(t, x)$ as $(n\Delta t, i\Delta x)$ converges to (t, x) as Δt, Δx converge to 0.

Definition 4.0.4. [Stability] A finite difference scheme is said to be stable for the norm $|| \, ||_p$ if there exists two constants $C_1 > 0$ and $C_2 > 0$, independent of h(space step discretization) and δ (time step discretization), such that when h and δ tend towards zero:

$$||f||_p \leq C_1||f_0|| + C_2||u||, \tag{4.22}$$

whatever the initial data f_0 and the source term g.

The most important notion of stability for a PDE is Von Neumann stability. Assume we have a Fourier expansion in space of our function

$$f(t,x) = \sum_k \hat{f}(k) e^{ikx} \tag{4.23}$$

The sum is over k, the Fourier frequencies. Now take for f just one Fourier term

$$f(t,x) = \hat{f}(k) e^{ikx} \tag{4.24}$$

and evaluate it at (x_j, t_n) to get

$$f_j^n = \hat{f}(t_n) e^{ikj\Delta x} \tag{4.25}$$

To simplify notation we can write $\hat{f}^n = \hat{f}(t_n)$. Then

$$f_j^n = \hat{f}^n e^{ikj\Delta x} \tag{4.26}$$

$$f_{j-1}^n = \hat{f}^n e^{(ik(j-1)\Delta x)} \tag{4.27}$$

$$f_{j+1}^n = \hat{f}^n e^{(ik(j+1)\Delta x)} \tag{4.28}$$

$$f_j^{n+1} = \hat{f}^{n+1} e^{(ikj\Delta x)} \tag{4.29}$$

These expressions can be plugged directly into any finite difference scheme to check the stability. The growth rate G is defined as

$$\frac{f_j^{n+1}}{f_j^n}. \tag{4.30}$$

The necessary and sufficient condition for the error to remain bounded is that $|G| < 1$ for all frequencies k. The effect of a single step of the numerical scheme is to multiply the complex exponential by the so-called magnification factor λ:

$$f_j^{n+1} = \lambda e^{ikj\Delta x} \tag{4.31}$$

In other words, e^{ikx} assumes the role of an eigenfunction, with the magnification factor λ being the corresponding eigenvalue of the linear operator governing each step of the numerical scheme. Continuing, we find that the effect of n further iterations of the scheme is to multiply the exponential by the n^{th} power of the magnification factor:

$$f_j^{n+r} = \lambda^r e^{ikj\Delta x} \tag{4.32}$$

Thus, the stability of the scheme will be governed by the size of the magnification factor. If $|\lambda| > 1$, , then λ^r is exponentially growing as $r \to \infty$, and so the numerical solutions become unbounded as $t \to \infty$. This is clearly incompatible with the analytical behavior of solutions to the heat equation, and so a necessary condition for the stability of our numerical scheme is that its magnification factor satisfy

$$|\lambda| \leq 1 \tag{4.33}$$

This method of stability analysis was developed by the mid-twentieh century Hungarian mathematician and father of the electronic computer John von Neumann. Conditional stability means we only have stability on a certain condition. Usually the condition limits Δt in function of Δx.It is important to note that we are checking the stability for a method, not for an equation.

4.1 Heston Model solved by FD

Generally speaking, finite-difference schemes can be divided into two classes: implicit FD schemes and explicit FD schemes. The θ-schemes refer to those scheme in which are balanced both explicit and implicit scheme. The most famous of these last schemes is the Crank -Nicolson scheme, obtained by taking average of these two schemes. In explicit finite difference schemes, the value at time $n+1$ depends explicitly on the value at time n. The major advantage of explicit finite difference methods is that they are
relatively simple and computationally fast but they need a condition (namely

CFL condition) for their stability and if this condition is not satisfied, the solution becomes unstable and starts to wildly oscillate. Instead, for the implicit scheme we have to solve a linear system of equations. The main advantage of implicit finite difference methods is that there are no restrictions on the time step, which is good news if we want to simulate geological processes at high spatial resolution. Taking large time steps, however, may result in an inaccurate solution. Therefore it is always wise to check the results by decreasing the time step until the solution does not change anymore (this is called converge check)

	First Derivative in Space	Second Derivative in Space
Explicit	$\frac{C^n_{i+1,j}-C^n_{i-1,j}}{2ds}$	$\frac{C^n_{i+1,j}-2C^n_{i,j}+C^n_{i-1,j}}{ds^2}$
Implicit	$\frac{C^{n+1}_{i+1,j}-C^{n+1}_{i-1,j}}{2ds}$	$\frac{C^{n+1}_{i+1,j}-2C^{n+1}_{i,j}+C^{n+1}_{i-1,j}}{ds^2}$
θ Methods	$\theta\frac{C^n_{i+1,j}-C^n_{i-1,j}}{2ds}+(1-\theta)\frac{C^{n+1}_{i+1,j}-C^{n+1}_{i-1,j}}{2ds}$	$\theta\frac{C^{n+1}_{i+1,j}-2C^{n+1}_{i,j}+C^{n+1}_{i-1,j}}{ds^2}+(1-\theta)\frac{C^{n+1}_{i+1,j}-2C^{n+1}_{i,j}+C^{n+1}_{i-1,j}}{ds^2}$

Table 4.1: Explicit, Implicit and θ Discretization

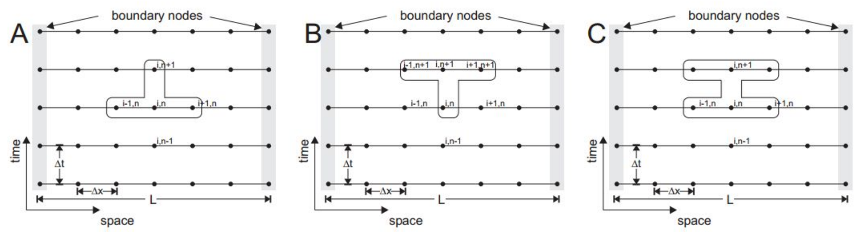


Figure 4.2: A Explicit finite difference discretization. B Implicit finite difference discretization. Cθ-Method (i.d.Crank-Nicolson) discretization.

For 1D modeling, the implicit schemes are superior for two reasons. First they can be made unconditionally stable and so the time step size can be arbitrarily chosen to save computation time and, second, they lead to tridiagonal systems which can be solved efficiently by forward/backward substitution (Trefethen, 1996). However, in higher dimensions, the implicit FD schemes require solving multi-dimensional matrices which greatly increases memory usage and computational cost. As regard our proposal, that is to solve the Heston PDE by FD, a semi-discretization of the Heston PDE, using finite difference schemes on a uniform grids, gives rise to large systems of stiff ordinary differential equations. For the effective numerical solution of these systems, standard implicit time-stepping methods are often not suitable anymore, and tailored time discretization methods are required.

For the numerical solution of the semi-discrete Heston PDE we shall study a splitting schemes of the Alternating Direction Implicit (ADI) type. In the past decades, ADI schemes have been successful already in many application areas. A main and distinctive feature of the Heston PDE, however, is the presence of a mixed spatial-derivative term, stemming from the correlation between the two underlying stochastic processes for the asset price and its variance. It is well known that ADI schemes were not originally developed to deal with such terms. Let $C(s, v, t)$ denote the price of a European option, if at time $T - t$ the underlying asset price is equal to s and its variance is equal to v, and T is the given maturity time of the option. We will study the forward equation in which, with abuse of notation we will use t even if our time variable should be $\tau = T - t$. Hestons stochastic volatility model implies ([41], [53]) that C satisfies

$$(H)\begin{cases} \frac{\partial C}{\partial t} = \frac{1}{2}s^2 v\frac{\partial^2 C}{\partial s^2} + \rho\sigma s v\frac{\partial^2 C}{\partial v\partial s} + \frac{1}{2}\sigma^2 v\frac{\partial^2 C}{\partial v^2} + (r_d - r_f)s\frac{\partial C}{\partial s} + \\ \qquad +k(\eta - v)\frac{\partial C}{\partial v} - r_d u \\ C(0, s, v) = (s - K)^+ \\ C(t, 0, v) = 0 \\ \frac{\partial C}{\partial t} - r\frac{\partial C}{\partial x} - k(\eta - v)\frac{\partial C}{\partial v} + rC = 0 \quad \text{for } v = 0 \end{cases}$$

We will solve the PDE from an initial condition $C(0, s, v) = (s - K)^+$ to some terminal time T, so that the time domain is naturally bound. In the space domain boundary conditions must be provided. These can either be determined by the 'physics' of the problem (a knock-out barrier option has a very clear choice of a boundary) or can be set sufficiently far out as to not

affect the interesting part of the solution. In line with our previous discussion on the boundary condition, the model we are going to evaluate does not guarantee any solution along the boundary of $v = 0$. We first evaluate the PDE (2.32) along the boundary, and second we implemente the Dirichlet conditions along $v = 0$. The use of such a condition will be implemented using the Lypaunov function (2.80)previously studied:the existence of a Lyapunov function ensures that the state vector does not reach the boundary and does not blow up in finite time. As a preliminary step towards the numerical solution of the initial-boundary value problem for the Heston PDE, the spatial domain, is restricted to a bounded set $[0, S_{\text{Max}}] \times [0, V_{\text{Max}}]$ with fixed values $S_{\text{Max}}, V_{\text{Max}}$ chosen sufficiently large. For the initial-boundary value problem we perform a spatial discretization on a cartesian grid by finite difference (FD) schemes.
To discretize the domain $[0, T] \times D$, we introduce an equi-distributed grid points corresponding to a spatial step size $\Delta s = 1/(N_s+1)$, $\Delta v = 1/(N_v+1)$ and to a time step $\Delta t = 1/(M + 1)$, where M, N_s, N_v are positive integers, number of time steps, of nodes in S and V direction respectively. We define the nodes of a regular grid:

$$(t_n, s_i, v_j) = (n\Delta t, i\Delta s, j\Delta v)$$

with $n \in 0, .., M + 1, i \in 0, ..., N_s + 1, i \in 0, ..., N_v + 1$. And we denote as $C^n_{i,j}$ the value of an approximate solution at point (t_n, s_i, v_j) and $C(t, s, v)$ the exact solution of problem. The initial data must also be discretized as:

$$C^0_{i,j} = C^0(s_i, v_j) \quad \forall i \in 1....N_s + 1 \quad \forall j \in 1....N_v + 1 \tag{4.34}$$

The problem is then to find, at each time step, a vector $C_{i,j} \in \mathbb{R}^2$, such that its components are the values $(C^n_{i,j})_{1\leq i\leq N_s \; 1\leq i\leq N_v}$. To develop the numerical scheme we will use the forward approximation for the discretization in the S-direction (it is the analogous one as regard the V-direction):

$$C'(x_{i,j}) \approx \frac{C^n_{i+1,j} - C^n_{i-1,j}}{2\Delta s} \tag{4.35}$$

$$C''(x_{i,j}) \approx \frac{C^n_{i+1,j} - 2C^n_{i,j} + C^n_{i-1,j}}{\Delta s^2} \tag{4.36}$$

For the discretization of the mixed derivative we use the FD scheme:

$$\begin{aligned}\frac{\partial^2 C}{\partial s \partial v}(x_i, v_j) \quad \approx \quad & \frac{C_{i+1,j+1} + C_{i+1,j} + C_{i+1,j-1} + C_{i,j+1}+}{4\Delta s \Delta v} + \\ & + \frac{C_{i-1,j+1} + C_{i,j-1} C_{i-1,j-1} + C_{i-1,j}}{4\Delta s \Delta v}\end{aligned} \tag{4.37}$$

The FD discretization described above of the initial-boundary value problem (H) for the Heston PDE yields an initial value problem for a large system of stiff ordinary differential equations (ODEs),

$$C^{'}(t) = AC(t) + b(t) \;\; (0 \leq t \leq T), \;\; C(0) = C^0. \tag{4.38}$$

Here A is a given $m \times m$ matrix and b(t) $(t \geq 0)$ and C_0 are given mvectors with $m = N_s \times N_v$. The vector C_0 is directly obtained from the initial condition and the vector function b depends on the boundary conditions setted. For each given $t > 0$, the entries of the solution vector $C(t)$ to (H) constitute approximations to the exact solution values $C(t, s, v)$.

4.2 Time discretization and ADI schemes

Acquiring an effective numerical time-discretization method for the spatially discretized Heston problem (H) is a key step in arriving at a full numerical solution scheme for the Heston PDE that is both efficient and robust.

Let $\Delta t > 0$ be a given time step and let temporal grid points be given by $t_n = n\Delta t$ for $n = 0, 1, 2, ..$

A well-known method for the numerical solution of stiff initial value problems for systems of ODEs

$$C'(t) = F(t, C(t)) \quad (0 \leq t \leq T), C(0) = C_0 \tag{4.39}$$

is the Crank-Nicholson scheme or trapezoidal rule. This method defines approximation C^n to the exact solutions value $C(t_n)$ subsequently for $n =$

$1, 2, 3, ...$ as follow:

$$C^n = C^{n-1} + \frac{1}{2}\Delta t F(t^{n-1}, C^{n-1}) + \frac{1}{2}\Delta t F(t^n, C^n). \tag{4.40}$$

In our case of (4.38) we have

$$F(t, w) = Aw + b(t) \quad 0 \leq t \leq T, w \in \mathbb{R}^m \tag{4.41}$$

Thus, each step requires the solution of a system of linear equations involving the matrix $(I - \frac{1}{2}\Delta t A)$ where I denotes the m $\times$ m identity matrix.
Since $(I - \frac{1}{2}\Delta t A)$does not depend on the step index n, one can compute a LU factorization of this matrix once, beforehand, and next apply it in all step to obtain $C^n (n \geq 1)$. The Crank-Nicholson scheme can be practical when the number of spatial grid points $m = N_s \times N_v$ is moderate. In our application to the two-dimensional Heston PDE, however m usually gets very large and the Crank-Nicholson scheme becomes ineffective. The reason for this is that $(I - \frac{1}{2}\Delta t A)$, and hence the matrices in its LU factorization, possess a bandwidth that is directly proportional to $\min(Ns, Nv)$.

4.2.1 A non uniform discretization

For the initial-boundary value problem (H) we perform a spatial discretization on a Cartesian grid by finite difference (FD) schemes. Nerverthless to improve the accuracy and convergence could be even applied a non-uniform meshes (which has recently been considered e.g. by Tavella and Randall [68] and Kluge [52])in both directions such that relatively many mesh points lie in the neighborhood of $s = K$ and $v = 0$, respectively. The application of such non-uniform meshes greatly improves the accuracy of the FD discretization compared to using uniform meshes. This is related to the facts that the initial function (4.34) possesses a discontinuity in its first derivative at $s = K$ and that for $v \approx 0$ the Heston PDE is convection dominated. It is also natural to have many grid points near the point $(s, v) = (K, 0)$as in practice this is the region in the (s, v)-domain where one wishes to obtain option prices. First of all define the grid in the direction of the underlying asset. It is $Ns \geq 1$ and we will take a constant, later fixed, $c > 0$. Define equidistant points $\epsilon_0 < \epsilon_1 < \epsilon_2 < ... < \epsilon_{Ns}$ as

$$\epsilon_i = \sinh^{-1}(-K/c) + i * \Delta\epsilon t \tag{4.42}$$

with

$$\Delta\epsilon = \frac{1}{Ns}[\sinh^{-1}((S-K)/c) - \sinh^{-1}(-K/c)]. \tag{4.43}$$

So a non uniform grid $0 = s_0 < s_1 < .. < s_{Ns} = S$ is defined by the transformation

$$s_i = K + c\sinh(\epsilon_i) \quad (0 \leq i \leq Ns). \tag{4.44}$$

In the v direction we define an integer $m_2 \geq 1$ and a constant $d > 0$. We consider even in this case equidistant points given by $\eta_j = j\Delta\eta$ for $j = 0, 1, .., m_2$ with

$$\Delta\eta = \frac{1}{m_2}\sinh^{-1}(V/d). \tag{4.45}$$

Define so a grid $0 = v_0 < v_1 < .. < v_{Nv} = V$ with

$$Vv_j = d\sinh(\eta_j) \quad (0 \leq j \leq Nv) \tag{4.46}$$

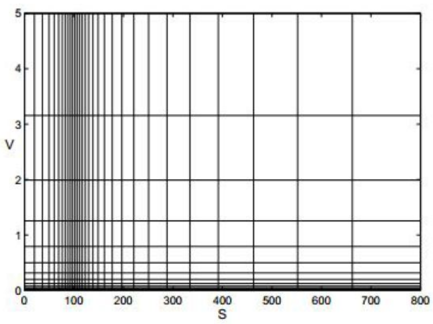


Figure 4.3: Non uniform Grid

So we can formulate even in this case our scheme with finite differences.

Let $f : R \to R$ be a function and let $s_0 < s_1 < s_2 < ... < s_m$ grid points with $\Delta x_i = x_i - x_{i-1}$. For the approximation of our first derivative $f^{'}(s_i)$ and the second derivative $f^{''}(s_i)$ we can consider:

$$f'(s_i) \approx \beta_{i,-1} f(s_{i-1}) + \beta_{i,0} f(s_i) + \beta_{i,1} f(s_{i+1}) \quad (4.47)$$

$$f''(s_i) \approx \gamma_{i,-1} f(s_{i-1}) + \gamma_{i,0} f(s_i) + \gamma_{i,1} f(s_{i+1}) \quad (4.48)$$

$$\frac{\partial^2 f}{\partial v \partial s} \approx \sum_{i,j=-1}^{1} \beta_{i,k} \hat{\beta}_{j,l} f(s_{i+k}, v_{j+l}) \quad (4.49)$$

$$\beta_{i,-1} = \frac{-\Delta s_{i+1}}{\Delta s_i(\Delta s_i + \Delta s_{i+1})} \quad (4.50)$$

$$\beta_{i,0} = \frac{\Delta s_{i+1} - \Delta s_i}{\Delta s_i \Delta s_{i+1}} \quad (4.51)$$

$$\beta_{i,1} = \frac{\Delta s_i}{\Delta s_{i+1}(\Delta s_i + \Delta s_{i+1})} \quad (4.52)$$

$$\gamma_{i,-1} = \frac{2}{\Delta s_i(\Delta s_i + \Delta s_{i+1})} \quad (4.53)$$

$$\gamma_{i,0} = \frac{-2}{\Delta s_i \Delta s_{i+1})} \quad (4.54)$$

$$\gamma_{i,1} = \frac{2}{\Delta s_{i+1}(\Delta s_i + \Delta s_{i+1})} \quad (4.55)$$

Indicate with $\hat{\beta}_{i,k}$ the analogous to $\beta_{i,k}$ but related to the direction of the approximation of y. Here below we can see the relative error which occurs for the underlying value of $S = 100$ in case of uniform and non uniform grid for all the value of volatility valuated:

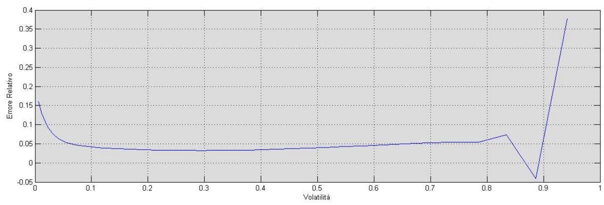


Figure 4.4: Relative error for the Non Uniform Grid

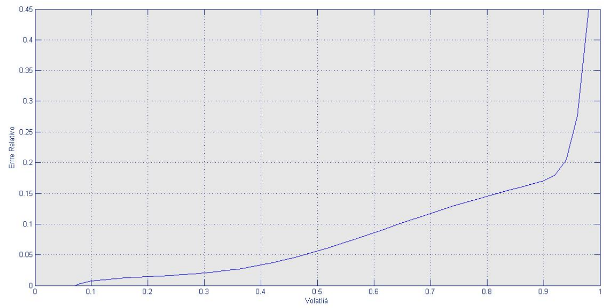


Figure 4.5: Relative error for the Uniform Grid

4.2.2 ADI scheme

For the numerical solution of the semi-discretized Heston problem we shall consider in this paper splitting schemes of the ADI type ([19], [43]).

We decompose the matrix A into three submatrices,

$$A = A_0 + A_1 + A_2. \tag{4.56}$$

We choose the matrix A_0 as the part of A that stems from the FD discretization of the mixed derivative term in. Next, in line with the classical ADI idea, we choose A_1 and A_2 as the two parts of A that correspond to all spatial derivatives in the S- and V-direction, respectively. Taking into account the PDE equation for pricing the r_dC term (that one of degree equal zero, where r_d is the interest rate) is distributed evenly over A_1, A_2. The FD discretization described implies that A_1, A_2 are essentially tridiagonal and pentadiagonal, respectively. Write $b(t)$ as

$$b(t) = b_0(t) + b_1(t) + b_2(t) \tag{4.57}$$

where the decomposition is analogous to that of A. Next, define functions $F_j(j = 0, 1, 2)$ by

$$F_j(t, w) = A_j w + b_j(t) \;\; 0 \leq t \leq T, w \in \mathbb{R}^m. \tag{4.58}$$

The ADI scheme considered, the Douglas scheme ([43], [24]), is developed as

$$\begin{cases} Y_0 = & C_{n-1} + \frac{1}{2}\Delta t F(t_{n-1}, C_{n-1}), \\ Y_j = & Y_{j-1} + \frac{1}{2}\Delta t(F_j(t_n, Y_j) - F_j(t_{n-1}, C_{n-1})) \quad (j = 1, 2) \\ C_n = & Y_2 \end{cases}$$

The splitting schemes treats the mixed derivative part F_0 in a fully explicit way. F_1 and F_2 parts are treated implicitly in the schemes. In the Do scheme, a forward Euler predictor step is followed by two implicit but unidirectional corrector steps, whose purpose is to stabilize the predictor step.

In every step, systems of linear equations need to be solved involving the two matrices $(I - \frac{1}{2}\Delta t A_j)$ for$j = 1, 2$. Like for the CrankNicholson scheme, these matrices do not depend on the step index n, and thus one can determine their LU factorizations once, beforehand, and next apply them in all time steps to compute $C^n (n \geq 1)$. To determine exactly the matrices in our discretization, we fix that are blocks matrices , whose block will have dimension $(N_s - 1) \times (N_s - 1)$. As regard A_1, it will be composed with k blocks only on the main diagonal. Here below we will refer to A_1, A_2, A_0 for the matrix in which is splitted the matrix A; we will refer to A_{i_x} (for $i = 0, 1, 2$ and $x = A, B, C$) to show the blocks distribution of the matrix A_i and with $A_{i_{xy}}$ to indicate the element composing the block of matrix. Every blocks of matrix A_1 (they are equal in each block) made up of:

$$\begin{pmatrix} A1_{B_{1,1,k}} & A1_{B_{1,2,k}} & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & \cdots & A1_{B_{i,i-1,k}} & A1_{B_{i,i,k}} & A1_{B_{i,i+1,k}} & 0 \\ 0 & 0 & \cdots & \cdots & A1_{B_{m_1-1,m_1-2,k}} & A1_{B_{m_1-1,m_1-1,k}} \end{pmatrix}$$

where the index k is referring to the block. The elements of such discretization for $k \neq 1$ and $i \neq 1$ are the following:

$$A_{1_{B_{i,i-1,k}}} = \frac{1}{4}v_{k+1}\frac{s_{i+1}^2}{\Delta s^2} - (r_d - r_f)\frac{s_i}{2\Delta s} \tag{4.59}$$

$$A_{1_{B_{i,i,k}}} = -\frac{1}{2}v_{k+1}\frac{s_{i+1}^2}{\Delta s^2} + \frac{1}{2}r_d \tag{4.60}$$

$$A_{1_{B_{i,i+1,k}}} = \frac{1}{4}v_{k+1}\frac{s_{i+1}^2}{\Delta s^2} + r_d\frac{s_i}{2\Delta s} \tag{4.61}$$

for $i = 1$ we will use the Lyapunov function $\mathrm{L}(s, v)$, introduced in the previous chapter (2.80), replacing the value $S(1)$ in the boundary with $S(1)/V(1)$. Therefore we will have for $k \neq 1$

$$A_{1_{B_{1,1,k}}} = -\frac{1}{2}v_{k+1}\frac{s_2^2}{(\mathrm{L}^2(s_2, v_{k+1}))}\frac{1}{\Delta s^2} + \frac{1}{2}r_d \tag{4.62}$$

$$A_{1_{B_{1,2,k}}} = \frac{1}{4}v_{k+1}\frac{s_2^2}{(\mathrm{L}^2(s_2, v_{k+1}))}\frac{1}{\Delta s^2} + (r_d - r_f)\frac{s_2}{\mathrm{L}(s_2, v_{k+1})}\frac{1}{2\Delta s} \tag{4.63}$$

If $k = 1$ we take into consideration the twin boundary conditions (those coming from the discretization on S-direction and those coming from V-direction). So, for $k = 1$ we have

$$A_{1_{B_{i,i-1,1}}} = \frac{1}{4}v_2\frac{s_{i+1}^2}{\mathrm{L}(s_{i+1}, v_2)}\frac{1}{\Delta s^2} - (r_d - r_f)\frac{s_i}{2\Delta s} \tag{4.64}$$

$$A_{1_{B_{i,i,1}}} = -\frac{1}{2}v_2\frac{s_{i+1}^2}{\mathrm{L}(s_{i+1}, v_2)}\frac{1}{\Delta s^2} + \frac{1}{2}r_d \tag{4.65}$$

$$A_{1_{B_{i,i+1,1}}} = \frac{1}{4}v_2\frac{s_{i+1}^2}{\mathrm{L}(s_{i+1}, v_2)}\frac{1}{\Delta s^2} + (r_d - r_f)s_i\frac{1}{2\Delta s} \tag{4.66}$$

And

$$A_{1_{B_{1,1,1}}} = -\frac{1}{2}v_2\frac{s_2^2}{(\mathrm{L}^3(s_2, v_2))}\frac{1}{\Delta s^2} + \frac{1}{2}r_d \tag{4.67}$$

$$A_{1_{B_{1,2,1}}} = \frac{1}{4}v_2\frac{s_2^2}{(\mathrm{L}^3(s_2, v_2))}\frac{1}{\Delta s^2} + (r_d - r_f)\frac{s_2}{(\mathrm{L}(s_2, v_2))}\frac{1}{2\Delta s} \tag{4.68}$$

The matrix A_2 is a tridiagonal (and not a diagonal one as the previous A_1) matrix made up of blocks of diagonal matrix of dimension $m_1 - 1$, as following mode:

$$\begin{pmatrix} A_{2_B} & A_{2_C} & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & A_{2_A} & A_{2_B} & A_{2_C} & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \cdots & A_{2_A} & A_{2_B} & A_{2_C} \\ 0 & 0 & 0 & \cdots & \cdots & \cdots & A_{2_A} & A_{2_B} \end{pmatrix}$$

The matrix $A_{2_A}, A_{2_B}, A_{2_C}$, will have on diagonal (for $k \neq 1$ e $i \neq 0$, those index which are interested by boundary conditions for $v = 0$ and $S = S_{min}$):

$$A_{2_{A_{i,i,k}}} = \frac{1}{4} v_{k+1} \sigma^2 \frac{1}{\Delta v^2} - \kappa(\theta - v_{k+1}) \frac{1}{2\Delta v} \tag{4.69}$$

$$A_{2_{B_{i,i,k}}} = -\frac{1}{2} v_{k+1} \sigma^2 \frac{1}{\Delta v^2} + \frac{1}{2} r_d \tag{4.70}$$

$$A_{2_{C_{i,i,k}}} = \frac{1}{4} v_{k+2} \sigma^2 \frac{1}{\Delta v^2} + \kappa(\eta - v_{k+2}) \frac{1}{2\Delta v} \tag{4.71}$$

For $k = 1$ the elements on diagonal are:

$$A_{2_{A_{i,1}}} = \frac{1}{4} \frac{v_2}{\mathrm{L}(s_{i+1}, v_2)} \sigma^2 \frac{1}{\Delta v^2} - \kappa(\eta - \frac{v_2}{\mathrm{L}(s_{i+1}, v_2)}) \frac{1}{2\Delta v} \tag{4.72}$$

$$A_{2_{B_{i,1}}} = -\frac{1}{2} \frac{v_2}{\mathrm{L}(s_{i+1}, v_2)} \sigma^2 \frac{1}{\Delta v^2} + \frac{1}{2} r_d \tag{4.73}$$

The matrix A_0 which is referred to the discretizzation for the mixed derivative is a tridiagonal matrix made up of blocks of matrix(in turn tridiagonal of dimension $m_1 - 1$):

$$\begin{pmatrix} A_{0_B} & A_{0_C} & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & A_{0_A} & A_{0_B} & A_{0_C} & 0 & \cdots & \cdots & 0 \\ 0 & 0 & A_{0_A} & A_{0_B} & A_{0_C} & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \cdots & A_{0_C} & A_{0_B} & A_{0_C} \\ 0 & 0 & 0 & \cdots & \cdots & \cdots & A_{0_C} & A_{0_B} \end{pmatrix}$$

The Lyapunov function used is:

$$\mathrm{L}(s, v) = -\log(v) - \log(s) + s\log(s + 3) + s(v + 1) + v$$

. The parameters used for our numerical simulation are:

$$\eta = 0.0022 \quad \sigma = 0.618 \quad r = 0.142$$

$$\kappa = 11.32 \quad T = 1$$

The results obtained are figured as follows below:

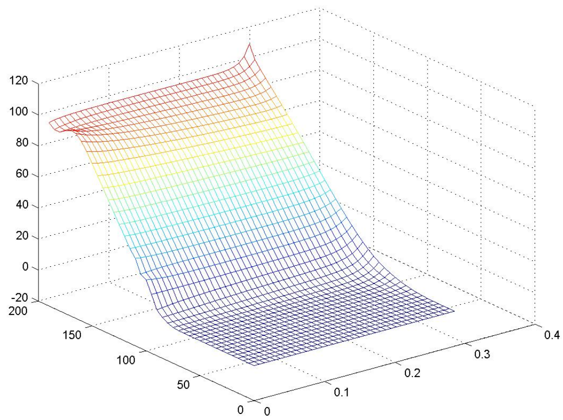


Figure 4.6: Solution by FD and Lyap

While those found out by the exact solution are:

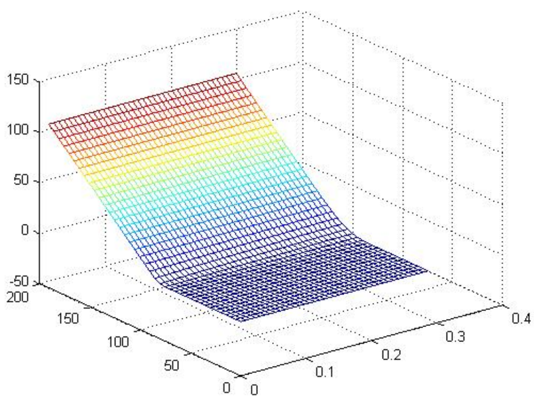


Figure 4.7: Solution by Close Formula

4.2.3 The stability of ADI schemes

Theoretical stability results for all four ADI schemes - relevant to FD discretization of 2D convection-diffusion equations with a mixed derivative term - have been derived in ([44], [56], [13]). These results concern unconditional stability, i.e., without any restriction on the time step Δt. The analysis has been performed following the classical Von Neumann method (Fourier transformation), where the usual assumptions are made that the coefficients are constant, the boundary condition is periodic, the spatial grid is uniform, and stability is considered in the l2-norm.

We derive linear stability results ever for this ADI scheme that use Lyapunov function along the boundary. These results are subsequently used to show that the ADI scheme under consideration is unconditionally stable when applied to finite difference discretizations of general parabolic two-dimensional convectiondiffusion equations. We mention that this scheme above is closely related to so-called Approximate Matrix Factorization methods, cf. e.g. ([45]). When applied to the linear scalar test equation

$$C'(y) = (\lambda_0 + \lambda_1 + \lambda_2)C(t) \tag{4.74}$$

with complex constants λ_j $(0 \leq j \leq 2)$.
Application of ADI scheme gives rise to a linear iteration of the form

$$C^n = R(z_0, z_1, z_2)C^{n-1} \tag{4.75}$$

with $z_j = \Delta t \lambda_j$ with $(j = 0, 1, 2)$ and

$$R(z_0, z_1, z_2) = 1 + \frac{z_0 + z_1 + z_2}{p} \tag{4.76}$$

Here, and throughout this thesis, we adopt the notation

$$z = z_1 + z_2 \quad p = (1 - \frac{1}{2}z_1)(1 - \frac{1}{2}z_2) \quad \theta = \frac{1}{2}. \tag{4.77}$$

The iteration is stable if

$$|R(z_0, z_1, z_2)| \leq 1 \tag{4.78}$$

For our scheme we have that:

$$\begin{cases} Y_0 = & C_{n-1} + \Delta t A C_{n-1} \\ Y_1(1 - \frac{1}{2} dt A_1) = & C_{n-1} + \Delta t A C_{n-1} - \frac{1}{2} \Delta t A_1 C_{n-1} \\ Y_2(1 - \frac{1}{2} dt A_2) = & \frac{C_{n-1} + \Delta t A C_{n-1} - \frac{1}{2} dt A_1 C_{n-1}}{1 - \frac{1}{2} dt A_1} - \frac{1}{2} \Delta t A_2 C_{n-1} \\ C_n = & Y_2 \end{cases}$$

and so, for this scheme, it is readily verified that

$$R(z_0, z_1, z_2) = 1 + \frac{z_0 + z_1 + z_2}{p}. \tag{4.79}$$

Firstly we will consider the diffusion matrix constant and with periodic boundary condition. The value λ_j mentioned above are eigenvalues of the matrices A_j $(0 \leq j \leq k)$. Consider the following condition on $z_0, z_1, z_2 \in C$ (where θ is equal to $\frac{1}{2}$ in our case):

$$p \neq 0 \text{ and } |z_0| \leq |\frac{p}{2\theta}| - |\frac{p}{2\theta} + z| \tag{4.80}$$

Recall that z, p are given by (4.77).

Lemma 4.2.1. Assume (4.80) holds and $\theta \leq \frac{1}{2}$. Then

$$|R(z_0, z_1, ..., z_k)| \leq 1 \tag{4.81}$$

Proof. Define

$$\widetilde{R} = \widetilde{R}(z_0, z_1, z_2, \dot{;} z_k) = \frac{1}{2\theta} + \frac{z + z_0}{p} \tag{4.82}$$

We have

$$|\widetilde{R}| = |\frac{1}{2\theta} + \frac{z}{p} + \frac{z_0}{p}| \leq |\frac{1}{2\theta} + \frac{z}{p}| + |\frac{z_0}{p}| \leq \frac{1}{2\theta}. \tag{4.83}$$

Subsequently,

$$|R(z_0, z_1, z_2, \dot{;} z_k)| = |1 - \frac{1}{2\theta} + \widetilde{R}| \leq 1 - \frac{1}{2\theta} + |\widetilde{R}| \leq 1 - \frac{1}{2\theta} + \frac{1}{2\theta} \tag{4.84}$$

□

The direct verification of condition (4.80) is not straightforward in general, in view of the non-trivial formulas one has for the eigenvalues z_j. We introduce next a condition which is much easier to verify when $k = 2$:

$$\text{Re}(z_1) \leq 0, \ \text{Re}(z_2) \leq 0 \text{ and } |z_0| \leq 2\sqrt{\text{Re}(z_1)\text{Re}(z_2)} \tag{4.85}$$

and note that this condition is independent of the parameter θ.

Lemma 4.2.2. Assume $k = 2$. Then (4.85) $\Rightarrow$ (4.80)

Proof. Define the vectors

$$z_j = \begin{bmatrix} -\sqrt{2\text{Re}(z_j)} \\ |\frac{1+\theta z_j}{\sqrt{z_j}}| \end{bmatrix} j = 1, 2$$

Then

$$||z_j||_2 = \sqrt{-2\text{Re}(z_j) + \frac{|1+\theta z_j|^2}{2\theta}} = \frac{|1-\theta z_j|^2}{\sqrt{2\theta}} \tag{4.86}$$

Condition (4.85) implies $p \neq 0$ (sum of positive amount). Next, we obtain

$$\begin{aligned}
|z_0| + |\frac{p}{2\theta} + z| &= |z_0| + |\frac{(1-\theta z_1)(1-\theta z_2)}{2\theta} + z_1 + z_2| && (4.87)\\
&= |z_0| + |\frac{(1+\theta z_1)(1+\theta z_2)}{2\theta} && (4.88)\\
&\leq 2\sqrt{\text{Re}(z_1)\text{Re}(z_2)} + |\frac{(1+\theta z_1)(1+\theta z_2)}{2\theta}| && (4.89)\\
&= v_1 \cdot v_2 \leq ||v_1||_2||v_2||_2 && (4.90)\\
&= \frac{|1-\theta z_1||1-\theta z_2|}{2\theta} && (4.91)\\
&= |\frac{p}{2\theta}| && (4.92)
\end{aligned}$$

which shows that (4.80) holds. □

In the following theorem (see [44]), we summarize the main results above on the stability requirements 4.78 relevant to $k = 2$. Unless stated otherwise, z_0, z_1, z_2 are assumed to be complex numbers here.

Theorem 4.2.1. Assume $k = 2$ and (4.85) holds. Then the condition (4.78) is fulfilled whenever $\theta \geq \frac{1}{2}$.

4.2.4 Stability for ADI Modified scheme

Following the results obtained in the previous section we are going to validate the results obtained on stability even for the modified scheme proposed for the study about ADI discretization on Heston model. In our specif case we have an equation of the form of

$$\frac{\partial C}{\partial t} = c \cdot \nabla u + \nabla \cdot (D\nabla u) \tag{4.93}$$

where,

$$c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

and

$$D = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$$

to have

$$\frac{\partial C}{\partial t} = (d_{12} + d_{21})C_{sv} + (c_1 C_s + d_{11} C_{ss}) + (c_2 C_v + d_{22} C_{vv}). \tag{4.94}$$

We require that D is positive semi-definite. This is equivalent to

$$d_{11} \geq 0 \; d_{22} \geq 0 \text{ and} (d_{12} + d_{21})^2 \leq 4 d_{11} d_{22}. \tag{4.95}$$

We investigate stability using the model scalar equation above and with λ_j an eigenvalue of A_j for $j = 0, 1, 2$. Taking as example the matrix A_1, whose elements involves the discretization in s-direction. We know that in the Von Neumann analysis we have the solution as

$$C_j^n = \frac{1}{2\pi} \int_{-\frac{\pi}{m_1}}^{\frac{\pi}{m_1}} e^{ijh\xi} \hat{C}(\xi) d\xi \tag{4.96}$$

Applying this replacement to our discretization (referring in this case to the space discretization), we have:

$$c_1 \frac{C_{j+1}^n - C_{j-1}^n}{2\Delta s} + d_{11} \frac{C_{j+1}^n - 2C_j^n - C_{j-1}^n}{\Delta s^2} = \tag{4.97}$$

$$\frac{c_1}{2\pi 2\Delta s} \int_{-\frac{\pi}{j_1}}^{\frac{\pi}{j_1}} (e^{i(j+1)h\xi} + e^{i(j-1)h\xi}) \hat{C}(\xi) d\xi + \tag{4.98}$$

$$+ \frac{d_{11}}{2\pi \Delta s^2} \int_{-\frac{\pi}{j_1}}^{\frac{\pi}{j_1}} (e^{i(j+1)h\xi} - 2e^{ijh\xi} + e^{i(j-1)h\xi}) \hat{C}(\xi) d\xi$$

$$= \frac{1}{2\pi} \int_{-\frac{\pi}{j_1}}^{\frac{\pi}{j_1}} e^{ijh\xi} \hat{C}(\xi) (\frac{c_1}{2\Delta s} (e^{i(j+1)h\xi} + e^{i(j-1)h\xi} + \tag{4.99}$$

$$\frac{d_{11}}{\Delta s^2} e^{i(j+1)h\xi} - 2e^{ijh\xi} + e^{i(j-1)h\xi})) d\xi$$

The scaled eigenvalues $z_r = \lambda_r dt$ are

$$
\begin{aligned}
z_r &= c_r q_r \Big(\frac{1}{2}e^{i\phi_r} - \frac{1}{2}e^{-i\phi_r}\Big) + d_{rr} a_r (e^{i\phi_r} - 2 + e^{-i\phi_r}) && (4.100)\\
&= -2d_{rr} a_r (1 - \cos\phi_r) + i c_r q_r \sin\phi_r && (4.101)
\end{aligned}
$$

and $z_0 = (d_{12} + d_{21}) b[-\sin\phi_1 \sin\phi_2]$ where

$$q_1 = \frac{\Delta t}{\Delta s}\; q_1 = \frac{\Delta t}{\Delta v}\; a_1 = \frac{\Delta t}{\Delta s^2}\; a_2 = \frac{\Delta t}{\Delta v^2}\; b = \frac{\Delta t}{\Delta s \Delta v} \tag{4.102}$$

The angles ϕ_r are integer multiples of $2\pi/m_r$, $(r = 1, 2)$ where as usual m_1, m_2 are the dimension of the grid in the x and y directions, respectively. We have

$$|z_0|^2 \le (d_{12} + d_{21})^2 b^2 \le 4 d_{11} d_{22} b^2 \tag{4.103}$$

and so

$$|z_0|^2 \le 4\mathrm{Re}(z_1)\mathrm{Re}(z_2) \tag{4.104}$$

Thus we have shown that the condition (4.85) is fulfilled, independently of $dt, ds, dv > 0$. and even for those nodes whre the discretization is balanced by the Lyapunov function. By invoking Theorem 4.2.1, we arrive at the following result (see [44])for the Douglas schemes:

Theorem 4.2.2. Consider equation (4.93) for $k = 2$ with 4.85 and periodic boundary condition. Assume that $\theta \le \frac{1}{2}$. Then the scheme are unconditionally stable when applied to equation (H). Moreover, this conclusion remains valid when any other, stable finite difference discretizations for C_s, C_v used.

4.3 Numerical Validation

Consider now value for v_0 close to 0 and with an underlying asset with value $S_0 = 100$, using 30 nodes with both directions and 50 time steps, we obtain the following results:

V_0	Numerical Value with Lyapu	Numerical Value no Lyapu	Theor Value
0.03	13.3193	13.3555	13.4339
0.06	13.6587	13.5071	13.6016
0.12	14.3820	13.9384	14.0529
0.20	14.9670	14.4459	14.5294
0.25	16.0614	15.5247	14.8679

If we implemented the boundary condition which uses the PDE in $v = 0$ we have (for value of $K = 100, V_max = 0.3; T = 1; m_1 = 30; m_2 = 30; N_T = 50; \sigma = 0.618; \eta = 0.0022; \kappa = 11.32; r_d = 0.142; \rho = -0.5$

V_0	Numerical Value with Lyapu	Numerical Value no Lyapu	Exacted Value
0.03	13.0568	9.001	13.4339
0.06	13.3669	9.4182	13.6016
0.12	13.9433	10.2177	14.0529
0.20	14.5173	10.9365	14.5294
0.25	15.6348	12.2431	14.8679

As illustrated in the above table, the use of the Lyapunov function, combined with the right PDE condition evaluated along the boundary for $v = 0$, provides an optimization of the numerical results. As we go far from the boundary (which our process will never approach) the converge of numerical scheme results for both the methods. Nevertheless, the use of the Lyapunov function raises intensely the convergence. Even for those nodes close to the value of $v = 0$ the value of the numerical scheme does not deviate from the exact value for a consistent value. As far as we leave the nodes along the critical boundary, even the scheme without the Lyapunov function will slowly converge to the exact value.

If the grid points, in v direction, are increased (doubled in comparison with the S direction) the results will be as the follow:

	$v_0 = 0.01$			$v_0 = 0.07$			$v_0 = 0.1$		
	Lyap	*NoLyap*	*Theor*	*Lyap*	*NoLyap*	*Theor*	*Lyap*	*NoLyap*	*Theor*
Spot									
93.33	3.8755	6.6806	6.9444	6.3367	7.4265	7.6933	7.1754	7.7940	8.0443
100	11.0391	13.2733	13.3442	13.8818	13.5620	13.6623	14.3542	14.7430	13.8533
106.66	27.8350	19.9175	19.9419	23.0497	20.0379	20.0738	22.4683	20.12441	20.1689

Table 4.2: Price European call No PDE

	$v_0 = 0.15$			$v_0 = 0.2$			$v_0 = 0.25$		
	Lyap	*NoLyap*	*Theor*	*Lyap*	*NoLyap*	*Theor*	*Lyap*	*NoLyap*	*Theor*
Spot									
93.33	8.2347	8.3660	8.5863	9.0592	8.9180	9.0822	10.3598	10.0656	9.5416
100	14.8654	14.0711	14.1884	15.3061	14.4435	14.5294	16.3279	15.4502	14.8679
106.66	22.1653	20.3017	20.3556	22.2125	20.5381	20.5660	22.9365	21.3963	20.7923

Table 4.3: Price European call no PDE

In this case, where it is not used the boundary condition which implemented the PDE along the boundary, the use of the Lyapunov function does not seem to give any improvement for the convergence of numerical solution to the exact solution provided by the close formula. The numerical scheme without the PDE implements in this case Dirichlet condition along both boundary lines. From a theoretical point of view the use of Dirichlet condition is not a fair value: the process will never approach that boundary and the value imposed by the numerical scheme will force the solution to a value never taken.

Again, if we implement the boundary condition which uses the PDE in $v = 0$, taking this time more points in the v direction we have (for value of $K = 100, V_max = 0.3; T = 1; m_1 = 30; m_2 = 60; N_T = 50; \sigma = 0.618; \eta = 0.0022; \kappa = 11.32; r_d = 0.142; \rho = -0.5$):

	$v_0 = 0.01$			$v_0 = 0.07$			$v_0 = 0.1$		
	Lyap	*NoLyap*	*Theor*	*Lyap*	*NoLyap*	*Theor*	*Lyap*	*NoLyap*	*Theor*
Spot									
93.33	5.0768	6.7237	6.9444	6.5971	7.3398	7.6933	7.1638	7.6409	8.0443
100	12.4736	11.6131	13.3442	13.2146	12.2066	13.6623	13.4809	12.4693	13.8533
106.66	20.93656	17.2967	19.9419	20.3387	17.8157	20.0738	20.3204	18.0191	20.1689

Table 4.4: Price European call

	$v_0 = 0.15$			$v_0 = 0.2$			$v_0 = 0.25$		
	Lyap	*NoLyap*	*Theor*	*Lyap*	*NoLyap*	*Theor*	*Lyap*	*NoLyap*	*Theor*
Spot									
93.33	7.9419	8.1164	8.5863	8.6155	8.5899	9.0822	9.8506	9.6714	9.5416
100	13.8944	12.8823	14.1884	14.3246	13.3029	14.5294	15.3846	14.3629	14.8679
106.66	20.4353	18.3416	20.3556	20.6600	18.6871	20.5660	21.5452	19.6844	20.7923

Table 4.5: Price European call

Even in this case, where we increased the number of nodes, the use of the Lyapunov function increases the converge of the numerical solution to the exact solution. The focal point of this part of analysis is that even the numerical scheme without the use of the Lyapunov function is converging to the exact value: the use of the function speeds up this convergence even close to the node for which $v = 0$. The use of function even in the numerical scheme will allow us to obtain the convergence requires in a shorter time, decreasing the numerical error. To underline the error which occurs using this approximation where the boundary conditions were given by the PDE in the boundary of $v = 0$ we report the graphics here below:

To underline the different price obtained using this approximation where the boundary conditions were given by the PDE in the boundary of $v = 0$ we report the graphics here below, where we plot at the same time the theoretical solution and the numerical one:

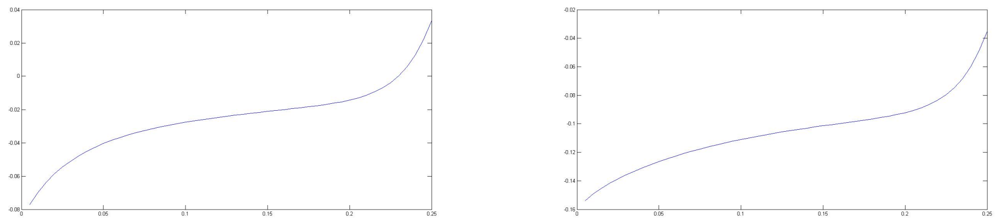


Figure 4.8: Relative error-with and without Lyap

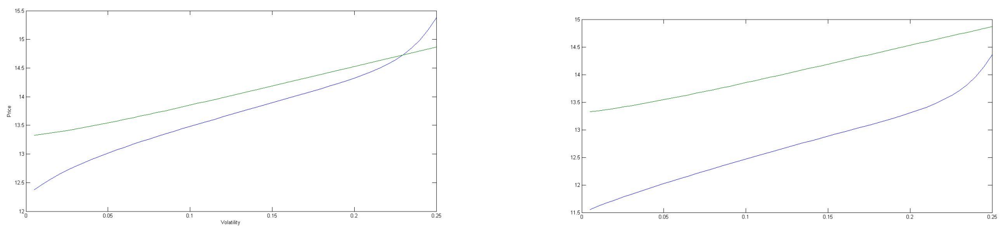


Figure 4.9: Relative error-with and without Lyap

The relative error coming from the approximation which uses the "'Lyapunov" function and PDE on boundary is the following below. We have used three different box, to highlight two volatility macro"'area":

s_0/v_0	0.02	0.05	0.08	0.10	0.11
93.33	-0.3135	-0.2092	-0.1495	-0.1174	-0.1122
100	-0.0588	-0.0405	-0.0314	-0.0276	-0.0261
106.66	0.0379	0.0197	0.0107	0.0075	0.0063
113.33	0.0017	0.0085	0.0110	0.0116	0.0118

The range of errors resulted from the approximation with the numerical scheme implemented and the use of the Lyapunov function is always below

s_0/v_0	0.13	0.14	0.15	0.16	0.17
93.33	-0.0948	-0.0876	-0.0811	-0.0753	-0.0700
100	-0.0034	-0.0223	-0.02512	-0.0201	-0.0189
106.66	0.0048	0.0043	0.0039	0.0037	0.0036
113.33	0.0120	0.0120	0.0121	0.0122	0.0124

10^{-2} (it has to be stressed here that to make a fair comparison between the numerical solution and the exact value, the solution resulted from our approximations should be taken not along or close to the boundary but in the center of grid; for this reason we reported also the other value but we highlighted the center of the grid for a fair evaluation).

Appendix A

Viscosity Solutions

Here we recall the notion of viscosity solutions for our value problem. For a more complete description we refer the reader to [18]. For any starshaped domain $D \in \mathbb{R}^d$, let $USC([0,T]\times D)$ and $LSC([0,T]\times D)$ denote the space of real-valued upper and lower semi-continuous functions on $[0,T]\times D$, respectively. The general form of a valuation equation (in the time-homogeneous case) is the Dirichlet problem:

$$(M)\begin{cases} \partial_t u(t,x) + Lu(t,x) - c(x)u(t,x) = f(t,x) & (t,x) \in (0,T)\times D, \\ u(T,x) = \phi(x), & x \in D \end{cases} \tag{A.1}$$

where, for any smooth function g, the differential operator is

$$Lg(x) = \langle \nabla g(x), b(x)\rangle + \frac{1}{2}\mathrm{tr}(\nabla^2 g(x) a(x)) \tag{A.2}$$

where a, b, c, ϕ, f are continuous functions on $[0,T]\times D$.

Definition A.0.1. Given $u \in USC([0,T]\times D)$, the parabolic super 2-jet of u at the point $(\bar{t},\overline{x}) \in (0,T)\times D$ is

$$\begin{aligned} P^{2,+}u(\bar{t},\overline{x}) \quad := \quad & \{(\partial_t g(\bar{t},\overline{x}), \nabla g(\bar{t},\overline{x}), \nabla^2 g(\bar{t},\overline{x}) : g \in C^{1,2}([0,T]\times D) \text{ and} \\ & (u-g)(t,x) \leq (u-g)(\bar{t},\overline{x}) = 0, \forall (t,x) \in [0,T]\times D\}. \end{aligned}$$

We say that a function g as above is a test function for $P^{2,+}u$ at $(\bar{t}, \overline{x})$. For a function $u \in LSC([0,T] \times D)$, the parabolic lower 2-jet is defined as $P^{2,-}u := -P^{2,+}(-u)$.

In Definition A.0.1, we can assume, without loss of generality, that $u-g$ has a strict global maximum at $(\bar{t}, \overline{x})$, that is,

$$(u-g)(t,x) < (u-g)(\bar{t}, \overline{x})) = 0, \quad \forall (t,x) \in [0,T] \times D, (t,x) \neq (\bar{t}, \overline{x}).$$

A function g satisfying these conditions is called a good *test function* for $P^{2,+}u$ at $(\bar{t}, \overline{x})$.

Definition A.0.2. A function $u \in USC([0,T] \times D)$ (resp. $u \in LSC([0,T] \times D)$) is a viscosity subsolution (resp. supersolution) of (M) if

- for every $(t,x) \in (0,T) \times D$ and any (good) test function g for $P^{2,+}u$ at (t,x) (resp. $P^{2,-u}$), it holds

$$\partial_t g(t,x) + Lg(t,x) - c(x)u(t,x) \geq f(t,x)(resp. \leq),$$

- $u(T,x) \leq \phi(x)$ (resp.,$\geq$) for all $x \in D$.

A function u that is both a viscosity subsolution and a viscosity super solution of (M) is a viscosity solution (clearly, $u \in C([0,T] \times D)$).

Appendix B

Matlab Codes

```
[title={Dupire Volatility}]

function SIG=dupire(x0,T,C,Cp,T1,T2,r)

tau1=(T2-T)/(T2-T1);
tau2=(T-T1)/(T2-T1);
deltaTau=T2-T1;
SigmaI1=sqrt(C(1)+C(2)*(C(3)*(x0 - C(4))+
        sqrt((x0-C(4))^2.0+C(5))));
SigmaI2=sqrt(Cp(1)+Cp(2)*(Cp(3)*(x0-Cp(4))+
        sqrt((x0-Cp(4))^2.0+Cp(5))));

f1=0.5*C(2)/SigmaI1*(C(3)+2.0*(x0-C(4))/sqrt((x0-C(4))^2+C(5)));
f2=0.5*Cp(2)/SigmaI2*(Cp(3)+2.0*(x0-Cp(4))/sqrt((x0-Cp(4))^2
   +Cp(5)));
g1=1.0/SigmaI1*(C(2)*C(5)/(2*((x0-C(4))^2.0+C(5))^1.5)-f1);
g2=1.0/SigmaI2*(Cp(2)*Cp(5)/(2*((x0-Cp(4))^2+Cp(5))^1.5)-f2);
SigmaI=tau1*SigmaI1+tau2*SigmaI2;

SigmaT=(SigmaI2-SigmaI1)/deltaTau;

SigmaX=tau1*f1+tau2*f2;
SigmaXX=tau1*g1+tau2*g2;
```

```
d1=((r+0.5*SigmaI^2.0)*T-x0)/(SigmaI*sqrt(T));
Num=SigmaI^2+2*T*SigmaI*(SigmaT+r*SigmaX);
Den=(1+d1*sqrt(T)*SigmaX)^2+
     SigmaI*T*(SigmaXX-SigmaX-d1*sqrt(T)*SigmaX^2)+0.74;

SIG=sqrt((Num)/(Den));
```

ADI to the Heston Model Calibrated by Lyapunov function along the Boundary

```
function roe=HestonADI(ka,na,sig,rho,v0,r,T,s0,K,m1,m2,NT)

V=zeros(1,m2+1);S=zeros(1,m1+1);
dx=2*s0/m1;dv=2*v0/m2;

S=[0:dx:200];V=[0:dv:0.3];
dt=T/NT;teta=0.5;rd=r;

LY=zeros(m1-1,m2-1);

for i=1:m1-1 for j=1:m2-1
LY(i,j)=-log(S(i+1))-log(V(j+1))+S(i+1)*(log(S(i+1))+3)+
        +S(i+1)*(V(j+1)+1)+V(j+1);
end end LLY=1./LY;

b1=zeros(1,m1-1);b2=zeros(1,m1-1);b3=zeros(1,m1-1);

bb1=zeros(1,m2-1);bb2=zeros(1,m2-1);bb3=zeros(1,m2-1);

g1=zeros(1,m1-1);g2=zeros(1,m1-1);g3=zeros(1,m1-1);

ggg1=zeros(1,m2-1);ggg2=zeros(1,m2-1);ggg3=zeros(1,m2-1);


%%%DERIVATA PRIMA
for i=1:m1-1
```

```
b1(i)=-1/(2*dx);b2(i)=0;b3(i)=1/(2*dx);
end
for i =1:m2-2
bb1(i)=-1/(2*dv);bb2(i)=0;bb3(i)=1/(2*dv);
end

%%%DERIVATA SECONDA
for i=1:m1-1
g1(i)=1/(dx^2);g2(i)=-2/(dx^2);g3(i)=1/(dx^2);
end
for i=1:m2-1
ggg1(i)=1/(dv^2);ggg2(i)=-2/(dv^2);ggg3(i)=1/(dv^2);
end




for k=2:m2-1
for i = 2:(m1-2)
Ax(i,i,k) = 0.5*V(k+1)*S(i+1)^2*g2(i)+
           (rd-rf)*S(i+1)*b2(i)-rd/2;
Ax(i,i-1,k)=0.5*V(k+1)*S(i+1)^2*g1(i)+
            (rd-rf)*S(i+1)*b1(i);
Ax(i,i+1,k)=0.5*V(k+1)*S(i+1)^2*g3(i)+
            (rd-rf)*S(i+1)*b3(i);
end

Ax(1,1,k)=0.5*V(k+1)*(S(2)*LLY(1,k))^2*g2(1)+
          (rd-rf)*(S(2)*LLY(1,k))*b2(1)-rd/2;
Ax(1,2,k)=0.5*V(k+1)*(S(2)*LLY(1,k))^2*g3(1)+
          +(rd-rf)*(S(2)*LLY(1,k))*b3(1);
Ax(m1-1,m1-2,k)=0.5*V(k+1)*S(m1)^2*g1(m1-1)+
              +(rd-rf)*S(m1)*b1(m1-1);
Ax(m1-1,m1-1,k)=0.5*V(k+1)*S(m1)^2*g2(m1-1)+
              +(rd-rf)*S(m1)*b2(m1-1)-rd/2;
```

```
end


for i = 2:(m1-2)
Ax(i,i,1) = 0.5*V(2)*LLY(i,1)*S(i+1)^2*g2(i)+
         +(rd-rf)*S(i+1)*b2(i)-rd/2;
Ax(i,i-1,1)=0.5*V(2)*LLY(i,1)*S(i+1)^2*g1(i)+
          +(rd-rf)*S(i+1)*b1(i);
Ax(i,i+1,1)=0.5*V(2)*LLY(i,1)*S(i+1)^2*g3(i)+
          +(rd-rf)*S(i+1)*b3(i);
end

Ax(1,1,1)=0.5*V(2)*LLY(1,1)*(S(2)*LLY(1,1))^2*g2(1)+
        +(rd-rf)*S(2)*LLY(1,1)*b2(1)-rd/2;
Ax(1,2,1)=0.5*V(2)*LLY(1,1)*(S(2)*LLY(1,1))^2*g3(1)+
        +(rd-rf)*S(2)*LLY(1,1)*b3(1);


Ax(m1-1,m1-2,1)=0.5*V(2)*LLY(m1-1,1)*S(m1)^2*g1(m1-1)+
               +(rd-rf)*S(m1)*b1(m1-1);
Ax(m1-1,m1-1,1)=0.5*V(2)*LLY(m1-1,1)*S(m1)^2*g2(m1-1)+
               +(rd-rf)*S(m1)*b2(m1-1)-rd/2;


for k=1:(m2-1)
A1((k-1)*(m1-1)+1:(k-1)*(m1-1)+m1-1,(k-1)*(m1-1)+1:
        (k-1)*(m1-1)+m1-1)=Ax(1:m1-1,1:m1-1,k);
end


for i=1:m1-1 %% no block
      Ay(i,i,1)=0.5*V(2)*LLY(i,1)*sig^2*ggg2(1)+
      +ka*(na-V(2)*LLY(i,1))*bb2(1)-rd/2;
end

```

```
for i=1:m1-1
Hup(i,i,1)=0.5*V(2)*LLY(i,1)*sig^2*ggg3(1)+
             +ka*(na-V(2)*LLY(i,1))*bb3(1);
end

for k=2:m2-1 for i=1:m1-1
Ay(i,i,k)=0.5*V(k+1)*sig^2*ggg2(k)+
            +ka*(na-V(k+1))*bb2(k)-rd/2;
end end

for k=2:m2-2
for i=1:m1-1
Hup(i,i,k)=0.5*V(k+1)*sig^2*ggg3(k)+
            +ka*(na-V(k+1))*bb3(k);
end end


for k=1:m2-2
for i=1:m1-1
Hdown(i,i,k)=0.5*V(k+2)*sig^2*ggg1(k+1)+
                +ka*(na-V(k+2))*bb1(k+1);
end end


for k=1:(m2-1)
A2((k-1)*(m1-1)+1:(k-1)*(m1-1)+m1-1,(k-1)*(m1-1)+1:
     (k-1)*(m1-1)+m1-1)=Ay(1:m1-1,1:m1-1,k);%%blocks
end
for k=1:(m2-2)
A2((k)*(m1-1)+1:k*(m1-1)+m1-1,(k-1)*(m1-1)+1:
   (k-1)*(m1-1)+m1-1)=Hdown(1:m1-1,1:m1-1,k); %%sttdiagonal
end
for k=1:(m2-2)
A2((k-1)*(m1-1)+1:(k-1)*(m1-1)+m1-1,(k)*(m1-1)+1:
    (k)*(m1-1)+m1-1)=Hup(1:m1-1,1:m1-1,k);%%sprdiagonal
```

```
end


%%%Mixed derivative by Lyapunov
for k=2:m2-1
for i = 2:(m1-2)
M(i,i,k) = rho*sig*S(i+1)*V(k+1)*b2(i)*bb2(k);
M(i,i-1,k)=rho*sig*S(i+1)*V(k+1)*b1(i)*bb2(k);
M(i,i+1,k)=rho*sig*S(i+1)*V(k+1)*b3(i)*bb2(k);
end

M(1,1,k)=rho*sig*S(2)*LLY(1,k)*V(k+1)*b2(1)*bb2(k);
M(1,2,k)=rho*sig*S(2)*LLY(1,k)*V(k+1)*b3(1)*bb2(k);
M(m1-1,m1-1,k)=rho*sig*S(m1)*V(k+1)*b2(m1-1)*bb2(k);
M(m1-1,m1-2,k)=rho*sig*S(m1)*V(k+1)*b1(m1-1)*bb2(k);
end

for i = 2:(m1-2)
M(i,i,1) = rho*sig*S(i+1)*V(2)*LLY(i,1)*b2(i)*bb2(1);
M(i,i-1,1)=rho*sig*S(i+1)*V(2)*LLY(i,1)*b1(i)*bb2(1);
M(i,i+1,1)=rho*sig*S(i+1)*V(2)*LLY(i,1)*b3(i)*bb2(1);

end
M(1,1,1)=rho*sig*V(2)*LLY(1,1)*S(2)*LLY(1,1)*b2(1)*bb2(1);
M(1,2,1)=rho*sig*V(2)*LLY(1,1)*S(2)*LLY(1,1)*b3(1)*bb2(1);
M(m1-1,m1-1,1)=rho*sig*V(2)*LLY(m1-1,1)*S(m1)*b2(m1-1)*bb2(1);
M(m1-1,m1-2,1)=rho*sig*V(2)*LLY(m1-1,1)*S(m1)*b1(m1-1)*bb2(1);


for k=2:m2-2
AxyU(1,1,k)=rho*sig*S(2)*LLY(1,k)*V(k+1)*b2(1)*bb3(k);
AxyU(1,2,k)=rho*sig*S(2)*LLY(1,k)*V(k+1)*b3(1)*bb3(k);
for i=2:(m1-2)
AxyU(i,i,k)=rho*sig*S(i+1)*V(k+1)*b2(i)*bb3(k);
AxyU(i,i-1,k)=rho*sig*S(i+1)*V(k+1)*b1(i)*bb3(k);
```

```
AxyU(i,i+1,k)=rho*sig*S(i+1)*V(k+1)*b3(i)*bb3(k);
end
AxyU(m1-1,m1-1,k)=rho*sig*S(m1)*V(k+1)*b2(m1-1)*bb3(k);
AxyU(m1-1,m1-2,k)=rho*sig*S(m1)*V(k+1)*b1(m1-1)*bb3(k);
end

AxyU(1,1,1)=rho*sig*V(2)*LLY(1,1)*S(2)*LLY(1,1)*b2(1)*bb3(1);
AxyU(1,2,1)=rho*sig*V(2)*LLY(1,1)*S(2)*LLY(1,1)*b3(1)*bb3(1);
for i=2:(m1-2)
AxyU(i,i,1)=rho*sig*V(2)*LLY(i,1)*S(i+1)*b2(i)*bb3(1);
AxyU(i,i-1,1)=rho*sig*V(2)*LLY(i,1)*S(i+1)*b1(i)*bb3(1);
AxyU(i,i+1,1)=rho*sig*V(2)*LLY(i,1)*S(i+1)*b3(i)*bb3(1);
end
AxyU(m1-1,m1-1,1)=rho*sig*V(2)*LLY(m1-1,1)*S(m1)*b2(m1-1)*bb3(1);
AxyU(m1-1,m1-2,1)=rho*sig*V(2)*LLY(m1-1,1)*S(m1)*b1(m1-1)*bb3(1);

for k=1:m2-2
AxyD(1,1,k)=rho*sig*S(2)*LLY(1,k+1)*V(k+2)*b2(1)*bb1(k+1);
AxyD(1,2,k)=rho*sig*S(2)*LLY(1,k+1)*V(k+2)*b3(1)*bb1(k+1);
for i=2:m1-2
AxyD(i,i+1,k)=rho*sig*S(i+1)*V(k+2)*b3(i)*bb1(k+1);
AxyD(i,i-1,k)=rho*sig*S(i+1)*V(k+2)*b1(i)*bb1(k+1);
AxyD(i,i,k)=rho*sig*S(i+1)*V(k+2)*b2(i)*bb1(k+1);
end
AxyD(m1-1,m1-1,k)=rho*sig*S(m1)*V(k+2)*b2(m1-1)*bb1(k+1);
AxyD(m1-1,m1-2,k)=rho*sig*S(m1)*V(k+2)*b1(m1-1)*bb1(k+1);
end

for k=1:(m2-1)
A0((k-1)*(m1-1)+1:(k-1)*(m1-1)+m1-1,(k-1)*(m1-1)+1:
    (k-1)*(m1-1)+m1-1)=M(1:m1-1,1:m1-1,k);%%Block
end
for k=1:(m2-2)
A0((k)*(m1-1)+1:k*(m1-1)+m1-1,(k-1)*(m1-1)+1:
    (k-1)*(m1-1)+m1-1)=AxyD(1:m1-1,1:m1-1,k); %%sttdiagonal
```

```
end
for k=1:(m2-2)
A0((k-1)*(m1-1)+1:(k-1)*(m1-1)+m1-1,(k)*(m1-1)+1:
    (k)*(m1-1)+m1-1)=AxyU(1:m1-1,1:m1-1,k);%%sprdiagonal
end


%%%VETTORE TERMINI NOTI
N0=zeros(NT+1,(m1-1)*(m2-1));
N1=zeros(NT+1,(m1-1)*(m2-1));
N2=zeros(NT+1,(m1-1)*(m2-1));

for k=1:NT+1
for i=1:(m2-1)
N1(k,i*(m1-1))=0.5*V(i+1)*g3(m1-1)*S(m1)^2*max(0,S(m1+1)-K)+
            +(rd-rf)*S(m1)*b3(m1-1)*max(0,S(m1+1)-K);
end
end


for k=2:(NT+1)
t=(k-1)*dt;
for i=1:m1-1
d1(i)=(log(S(i+1)/K)+(rd+0.5*V(1))*t)/(sqrt(V(1)*t));
d2(i)=d1(i)-sqrt(V(1)*t);
FF(i)=S(i+1)*0.5*erfc(-d1(i)/sqrt(2))-
        +K*0.5*erfc(-d2(i)/sqrt(2))*exp(-rd*t);
end
for i=1:m1-1
ddd1(i)=(log(S(i+1)/K)+(rd+0.5*V(m2+1))*t)/(sqrt(V(m2+1)*t));
ddd2(i)=ddd1(i)-sqrt(V(m2+1)*t);

EE(i)=S(i+1)*0.5*erfc(-ddd1(i)/sqrt(2))-
      +K*0.5*erfc(-ddd2(i)/sqrt(2))*exp(-rd*t);
end
```

```
for i=1:m1-1
N2(k,i)=(0.5*sig^2*V(2)*LLY(i,1)*ggg1(1)+
        +ka*(na-V(2)*LLY(i,1))*bb1(1))*FF(i);
N2(k,(m1-1)*(m2-2)+i)=(0.5*sig^2*V(m2)*ggg3(m2-1)+
         +ka*(na-V(m2))*bb3(m2-1))*EE(i);
end

a=(log(S(m1+1)/K)+(rd+0.5*V(1))*t)/(sqrt(V(1)*t));
b=d1(i)-sqrt(V(1)*t);
BB=S(m1+1)*0.5*erfc(-a/sqrt(2))-
   +K*0.5*erfc(-b/sqrt(2))*exp(-rd*t);


%%%first vector
N0(k,1)=rho*sig*V(2)*LLY(1,1)*S(2)*LLY(1,1)*
       *b3(1)*bb1(1)*FF(2)+
       +rho*sig*V(2)*LLY(1,1)*S(2)*LLY(1,1)*
       *b2(1)*bb1(1)*FF(1);
for i=2:m1-2 %%Last and first column
N0(k,i)=rho*sig*V(2)*LLY(i,1)*S(i+1)*b1(i)*bb1(1)*FF(i-1)+
       +rho*sig*V(2)*LLY(1,i)*S(i+1)*b2(i)*bb1(1)*FF(i)+
       +rho*sig*V(2)*LLY(1,i)*S(i+1)*b3(i)*bb1(1)*FF(i+1);
end
N0(k,m1-1)=rho*sig*V(2)*LLY(m1-1,1)*S(m1)*(b3(m1-1)*bb1(1)*BB+
         +b3(m1-1)*bb2(1)*max(0,S(m1+1)-K)+
         +b3(m1-1)*bb3(1)*max(0,S(m1+1)-K)+
         +b2(m1-1)*bb1(1)*FF(m1-1)+b1(m1-1)*bb1(1)*FF(m1-2));
for i=1:(m2-3)
N0(k,(i+1)*(m1-1))=rho*V(i+2)*S(m1)*(b3(m1-1)*bb1(i+1)+
         +b3(m1-1)*bb2(i+1)+
         +b3(m1-1)*bb3(i+1))*max(S(m1+1)-K,0);
end
```

```
%%%Last vector
NO(k,(m1-1)*(m2-2)+1)=rho*sig*V(m2)*S(2)*
            *LLY(1,m2-1)*b3(1)*bb3(m2-1)*EE(2)+
            +rho*sig*V(m2)*S(2)*LLY(1,m2-1)*b2(1)*
            *bb3(m2-1)*EE(1);
for i=2:m1-2 %%First and last column
NO(k,(m1-1)*(m2-2)+i)=rho*sig*V(m2)*S(i+1)*(b3(i)*
            *bb3(m2-1)*EE(i+1)+
            +b1(i)*bb3(m2-1)*EE(i-1)+b2(i)*bb3(m2-1)*EE(i));

NO(k,(m1-1)*(m2-1))=rho*sig*V(m2)*S(m1)*(b3(m1-1)*bb2(m2-1)
            *max(0,S(m1+1)-K)+
            +b3(m1-1)*bb1(m2-1)*max(0,S(m1+1)-K)+
            +b1(m1-1)*bb3(m2-2)*EE(m1-2)+
            +b2(m1-1)*bb3(m2-1)*EE(m1-1)+
            +b3(m1-1)*bb3(m2-1)*max(S(m1+1)-K,0));
end

NO(1,1)=rho*sig*V(2)*LLY(1,1)*S(2)*LLY(1,1)*
        *b3(1)*bb1(1)*max(0,S(m1+1)-K)+
        +rho*sig*V(2)*LLY(1,1)*S(2)*
        *LLY(1,1)*b2(1)*bb1(1)**max(0,S(m1+1)-K);
for i=2:m1-2 %%First and last Column
NO(1,i)=rho*sig*V(2)*LLY(i,1)*S(i+1)*b1(i)*bb1(1)
        *max(0,S(i)-K)+
        +rho*sig*V(2)*LLY(i,1)*S(i+1)*b2(i)*bb1(1)*
.......*max(0,S(i+1)-K)+
        +rho*sig*V(2)*LLY(i,1)*S(i+1)*b3(i)*bb1(1)*
.......*max(0,S(i+2)-K);
end
NO(1,m1-1)=rho*sig*V(2)*LLY(m1-1,1)*S(m1)*(b3(m1-1)
           *bb1(1)*max(0,S(m1+1)-K)+
           +b3(m1-1)*bb3(1)*max(S(m1+1)-K,0)+
         +b3(m1-1)*bb2(1)*max(0,S(m1+1)-K)+
         +b1(m1-1)*bb1(1)*max(0,S(m1-1)-K)+
```

```
        +b2(m1-1)*bb1(1)*max(0,S(m1)-K));

for i=1:(m2-3)%%last and first row of each column
N0(1,(i+1)*(m1-1))=rho*V(i+2)*S(m1)*(b3(m1-1)*bb1(i+1)+
         +b3(m1-1)*bb2(i+1)+
         +b3(m1-1)*bb3(i+1))*max(S(m1+1)-K,0);
end

N0(1,(m1-1)*(m2-2)+1)=rho*sig*V(m2)*S(2)*LLY(1,m2-1)
           *b3(1)*bb3(m2-1)*max(0,S(3)-K)+
           +rho*sig*V(m2)*S(2)*LLY(1,m2-1)*b2(1)*
           *bb3(m2-1)*max(0,S(2)-K);
for i=2:m1-2 %%Last column
N0(1,(m1-1)*(m2-2)+i)=rho*sig*V(m2)*S(i+1)
           *(b3(i)*bb3(m2-1)*max(0,S(i+2)-K)+
           +b1(i)*bb3(m2-1)*max(0,S(i)-K)+
           +b2(i)*bb3(m2-1)*max(0,S(i+1)-K));
end
N0(1,(m1-1)*(m2-1))=rho*sig*V(m2)*S(m1)*(b3(m1-1)*
           *bb2(m2-1)*max(0,S(m1+1)-K)+
           +b3(m1-1)*bb1(m2-1)*max(0,S(m1+1)-K)+
           +b1(m1-1)*bb3(m2-2)*max(0,S(m1-1)-K)+
           +b2(m1-1)*bb3(m2-1)*max(0,S(m1)-K)+b3(m1-1)
           *bb3(m2-1)*max(S(m1+1)-K,0));

N0(k,(m1-1)*(m2-1))=rho*sig*V(m2)*S(m1)*(b3(m1-1)*
         *bb2(m2-1)*max(0,S(m1+1)-K)+
         +b3(m1-1)*bb1(m2-1)*max(0,S(m1+1)-K)+
         +b1(m1-1)*bb3(m2-2)*EE(m1-2)+
.....    +b2(m1-1)*bb3(m2-1)*EE(m1-1)+b3(m1-1)*
         *bb3(m2-1)*max(S(m1+1)-K,0));

for i=1:m1-1
N2(1,i)=(0.5*sig^2*V(2)*LLY(i,1)*ggg1(1)+
       +ka*(na-V(2)*LLY(i,1))*bb1(1))*max(0,S(i+1)-K);
```

```
N2(1,(m1-1)*(m2-2)+i)=(0.5*sig^2*V(m2)*ggg3(m2-1)+
        +ka*(na-V(m2))*bb3(m2-1))*max(0,S(i+1)-K);
                                end


IN=zeros(m1+1,m2+1);
%%%Initial condition
for i=1:m1+1 for j=1:m2+1
IN(i,j)=max(0,S(i)-K);
end end

CC=zeros(m1-1,m2-1); ZZ=zeros(m1-1,m2-1);

D=zeros(m1-1,m2-1);
for i=1:m1-1 for j=1:m2-1
D(i,j)=IN(i+1,j+1);
end end
U0=[];
for i=1:m2-1
U0=[U0;D(:,i)];
end

for iii=2:NT+1
Y0=U0+dt*(A1*U0+A2*U0+A0*U0)+dt*(N0(iii-1,:)+
   +N1(iii-1,:)+N2(iii-1,:))';

QQ=eye((m1-1)*(m2-1))-teta*dt*A1;
SS=Y0-teta*dt*(A1*U0+(N1(iii-1,:))')+
    +teta*dt*(N1(iii,:))';
Y1=linsolve(QQ,SS);

OO=Y1; OO=OO';
for i=1:m2-1
ZZ(:,i)=OO(1+(i-1)*(m1-1):i*(m1-1));
end
```

```
pp=iii;
for e=1:m1-2
GG(e)=ZZ(e,1)+rd*dt*(e+1)*(ZZ(e+1,1)-ZZ(e,1))+
     +dt/dv*ka*(na-V(2)*LLY(e,1))*(ZZ(e,2)-ZZ(e,1))-
     +rd*dt*ZZ(e,1);

N2(pp,e)=(0.5*sig^2*V(2)*LLY(e,1)*ggg1(1)+
       +ka*(na-V(2)*LLY(e,1))*bb1(1))*GG(e);

end
GG(m1-1)=ZZ(m1-1,1)+rd*dt*(m1)*(max(0,S(m1+1)-K)-ZZ(m1-1,1))+
       +dt/dv*ka*(na-V(2)*LLY(m1-1,1))*(ZZ(m1-1,2)-ZZ(m1-1,1))-
       +rd*dt*ZZ(m1-1,1);
N2(pp,m1-1)=(0.5*sig^2*V(2)*LLY(m1-1,1)*ggg1(1)+
       +ka*(na-V(2)*LLY(m1-1,1))*bb1(1))*GG(m1-1);
N0(pp,m1-1)=rho*sig*V(2)*S(m1)*(b3(m1-1)*bb1(1)*BB+
      +b3(m1-1)*bb2(1)*max(0,S(m1+1)-K)+
      +b3(m1-1)*bb3(1)*max(0,S(m1+1)-K)+
      +b2(m1-1)*bb1(1)*GG(m1-1)+
      +b1(m1-1)*bb1(1)*GG(m1-2));

RR=eye((m1-1)*(m2-1))-teta*dt*A2;
WW=Y1-teta*dt*(A2*U0+(N2(iii-1,:))')+teta*dt*(N2(pp,:))';
Y2=linsolve(RR,WW);
U0=Y2;U0=U0';
for i=1:m2-1
CC(:,i)=U0(1+(i-1)*(m1-1):i*(m1-1));
end

for e=1:m1-2
GG(e)=CC(e,1)+rd*dt*(e+1)*(CC(e+1,1)-CC(e,1))+
    +dt/dv*ka*(na-V(2)*LLY(e,1))*(CC(e,2)-CC(e,1))-
    +rd*dt*CC(e,1);
N2(pp,e)=(0.5*sig^2*V(2)*LLY(e,1)*ggg1(1)+
```

```
        +ka*(na-V(2)*LLY(e,1))*bb1(1))*GG(e);
end
GG(m1-1)=CC(m1-1,1)+rd*dt*(m1)*(max(0,S(m1+1)-K)-
         +CC(m1-1,1))+
        +dt/dv*ka*(na-V(2)*LLY(m1-1,1))*(CC(m1-1,2)-
        +CC(m1-1,1))-
        +rd*dt*CC(m1-1,1);
N2(pp,m1-1)=(0.5*sig^2*V(2)*LLY(m1-1,1)*ggg1(1)+
            +ka*(na-V(2)*LLY(m1-1,1))*bb1(1))*GG(m1-1);
U0=U0';end roe=U0(m1/2,m2/2);
```

Acknowledgments

I would like to thank my supervisor Professor Marco Papi, who guided my doctoral studies with the proposal of a stimulating and interesting subject, for his numerous advice and suggestions and patience.

Thanks to Professor Renato Spigler, who supported my studies and followed directly any improvements or lacks in my path. Thanks to Dott. Moreno Concezzi and to Ing. Luca Pontecorvi for their useful and constant collaboration.

The first thank you is to my mum and my sister and Thomas: "cri you are never available." I wasn't never as i wish i were: nevertheless you had the patience and the love to let me go everywhere, even into my silence or into my numbers: without you i would have never been able to do this.

Thanks to Core who believed in me from the beginning :"I trust in you,. go ahead.." you are that person I will never be able to quit..the cliff in a storm, the irrational madness who makes happiness. Andrea: you are the call never missed, the unspoken understood: thanks will never be enough. Thanks to Dado for his hand never left and his heart always open to understand an us and not an I.

Thanks to Lavinia, Eleonora, Chiara and Tommaso and my family, Giulia, Michelina and Cristina for their patience in understanding my own poetry, my absences and my "I can't: a friendship which drives and saves.

I do not like to make a list for every one; nevertheless Id like to thank you my colleagues, Angela, Barbara, Laura, Massimiliano, Marco, Michele e Pino for their kindness and lovely support over a desk, over a business duty.

Finally, I would like to express my deep gratitude to Professor Peter Laurence, who aroused my interest in mathematical finance and supported my studies from the beginning: i would never have be able to say goodbye. This thesis is dedicated also to him.

Bibliography

[1] Ait-Sahalia, Y., Kimmel, R., *Estimating affine multifactor term structure models using closed-form likelihood expansions.* Journal of Financial Economics 98, 113-144, (2010).

[2] Andersen, T. G., Benzoni, L., Lund, J., *An Empirical Investigation of Continuous-Time Equity Return Models.* Journal of Finance 57 (2002).

[3] Appllebaum, D., *Levy Processes and Stochastic Calculus.* Cambridge, Studies in Advanced Mathematics, (2009).

[4] Artzner, P., Heath, D., *Approximate completeness with multiple martingale measures.* Mathematical Finance, 5, 111, (1995).

[5] Bahlali, K., Mezerdi, B., Ouknine, Y., *Pathwise uniqueness and approximation of solutions of stochastic differential equations.* Seminaire de Probabilites, XXXII, 166-187, Lecture Notes in Math., 1686, Springer, Berlin, (1998).

[6] Bates, D.S., *Jumps and Stochastic volatility: Exchange Rate Processes Implicity in Deutsche Mark Opitions.* The Review of Financial Studies 9, 69107, (1996).

[7] Bates, D. S., *Maximum Likelihood Estimation of Latent Affine Processes.* Review of Financial Studies 19, 909965, (2006).

[8] Black F., Scholes M., *The pricing of options and Corporate Liabilities.* Journal of Political Economy 81, 637-59, (1973).

[9] Breeden, D., Litzenberger, R., *Prices of state-contingent claims implicit in option prices.* Journal of Business, 51:621,51, (1978).

[10] Brezis, H., *Analyse fonctionnelle. Theorie et applications.* Masson, Paris, (1983).

[11] Broadie, M., Chernov, M., Johannes, M., *Model Specification and Risk Premia: Evidence from Futures Options.* Journal of Finance 62 (2007).

[12] Buchen, P.W., Kelly, M., *The maximum entropy distribution of an asset inferred from option prices.* Journal of Financial and Quantitative Analysis, 31 (1), (1996).

[13] Tadjeran, C., Meerschaert, M.M., *A second-order accurate numerical method for the two-dimensional fractional diffusion equation.*, Journal of Computational Physics 220, 813823, (2007) .

[14] Chernov, M., Gallant, R., Ghysels, E., Tauchen, J., *Alternative Models for Stock Price Dynamics.* Journal of Econometrics 116, 225257, (2003).

[15] Cont, R., *Financial Modelling with Jump Processes.* Chapman and Hall, CRC Financial Mathematics Series, (2005).

[16] Costantini, C., Papi, M., Ippoliti, F., *Singular risk-neutral valuation equations.* Financ. Stoch. 16(2), 249-274 (2012).

[17] Cox, J.C., Ingersoll, J.E. Jr., Ross, S.A., *A theory of the term structure of interest rates.* Econometrica, 53(2), (1985).

[18] Crandall, M.G., Hitoshi, I., Lions, P.L., *User's Guide to Viscosity Solutions of Second Order Partial Differential Equations.* American Methematical Society 27, 1-67, (1992).

[19] Craig, I.J.D., Sneyd, A.D., *An alternating-direction implicit scheme for parabolic equations with mixed derivatives.* Comp. Math. Appl. 16, 341-350, (1988).

[20] Cuchiero, C., Filipovic, D., Mayerhofer, E., Teichmann, J., *Affine processes on positive semidefinite matrices.* SSRN eLibrary, September, (2009).

[21] Cuchiero, C., Teichmann, J., *Path properties and regularity of affine processes on general state spaces.* ArXiv e-print, July (2011).

[22] Derman, E., Kani, I., *Riding on a smile.* Risk Magazine, 7(2):32-39,(1994).

[23] Donatucci, C., Papi, M., Pontecorvi, L., *Weighted Average Price in the Heston model with Stochastic Volatility.* Accepted for publication on 'Decisions in Economics and Finance' (2015).

[24] Douglas, J., Rachford, H.H., *On the numerical solution of heat conduction problems in two and three space variables.* Trans. Amer. Math. Soc. 82, 421-439, (1956).

[25] Dragulescu, A.A., Yakovenko, V.M., *Probability distribution of returns in the Heston model with stochastic volatility.* Quantitative Finance 2, 443-453, (2002).

[26] Duffie, D., *Dynamic Asset Pricing Theory.* Princeton University Press, New Jersey, (2001).

[27] Duffie, D. , Filipovic, D.,Schachermayer, W., *Affine processes and applications in Finance.* Annals of Applied Probability 13, 984-1053, (2003).

[28] Duffie, D., Pan J., Singleton K, *Transform Analysis and Asset Pricing for Affine Jump Diffusions.* Econometrica 68, 13431376, (2000).

[29] Dupire, B, *Pricing with a smile.* Risk Magazine, 7:18-20, (1994).

[30] Ekstrom, E. Tysk, J., *The Black-Scholes equation in stochastic volatility models.* J. Math. Anal. Appl. 368, 498-507, (2010).

[31] Ekstrom, E. Tysk, *Boundary conditions for the single-factor term structure equation.* J. Math. Anal. Vol.21,No.1 332-350, (2011).

[32] Eraker, B., Johannes, M., Polson, N., *The Impact of Jump in Volatility and Returns.* Journal of Finance, 12691300, (2003).

[33] Evans, C., *Partial Differential Equations.* American Mathematical Society, (1998).

[34] Feller, W., *Two singular diffusion problems.* Ann. of Math. 2, 173-182, (1951).

[35] Fengler, M. R., *Semiparametric modelling of implied volatility.* Lecture Notes. Springer, (2005).

[36] Filipovic, D., *Time-inhomogeneous Affine Processes. Stochastic Processes and their Applications.* 115, 639-659, (2005).

[37] Friedman, A., *Stochastic Differential Equations and Applications.* Vol. 1. Academic Press, New York (1975).

[38] Gatheral, J., *The Volatility Surface: A Practitioners Guide.* John Wiley and Sons, Inc., The United States of America (2006).

[39] Gauthier, P., Rivaille, P.-Y. H., *Fitting the smile: Smart parameters for SABR and Heston.* Available at http://papers.ssrn.com/sol3/papers.cfm?abstract id=1496982, 2009.

[40] Herzberg, F.S., *Stochastic Calculus with Infinitesimals.*, Lecture Notes in Mathematics, Vol. 2067, Springer-Verlag, Berlin Heidelberg, 35-44, (2013).

[41] Heston, S.L., *A closed-form solution for options with stochastic volatility with applications to bond and currency options.* 6, Rev. Financial Studies, 327-343, (1993).

[42] Heston , S. L, Loewenstein, M., Willard, G.A., *The Review of Financial Studies.* 20(2), 359-390, (2007).

[43] Hout, K.J., *ADI schemes in the numerical solution of the Heston PDE.* Numerical Analysis and Applied Mathematics, eds. T. E. Simos et. al., AIP Conf. Proc. 936, 1014, (2007).

[44] Hout, K.J., Welfert, B.D., *Stability of ADI schemes applied to convectiondiffusion equations with mixed derivative terms.* Applied Numerical Mathematics 57, 1935, (2007).

[45] Hundsdorfer, W., Verwer, J.G., *Numerical Solution of Time-Dependent AdvectionDiffusionReaction Equations.* Springer Ser. Comput.Math., vol. 33, Springer, Berlin, (2003).

[46] Kabanov, Y., Kramkov, D., *Large financial markets: asymptotic arbitrage and contiguity.* Probability Theory and its Applications, 39, 222229, (1994).

[47] Kallenberg, O., *Foundations Of Modern Probability.* Springer, (2002).

[48] Karatzas, I., Shreve, S.E., *Methods of Mathematical Finance.* Wiley Finance, (1998).

[49] Karatzas, I., Shreve, S.E., *Brownian Motion and Stochastic Calculus.* Springer (Graduate Texts in Mathematics), (1991).

[50] Karlin, S., Taylor, H.M., *A Second Course in Stochastic Processes.* Academic Press, (1981).

[51] Keller-Ressel, M., W. Schachermayer, W., Teichmann, J., *Regularity of affine processes on general state spaces.* ArXiv e-prints, May (2011).

[52] Kluge, T., *Pricing derivatives in stochastic volatility models using the finite difference method.* Dipl. thesis, TU Chemnitz, (2002).

[53] Lipton, A., *The vol smile problem.* Risk Magazine 15, 6165 (2002).

[54] Lipton, A., *Mathematical Methods for Foreign Exchange.*, World Scientific, Singapore, (2001).

[55] Lipton, A., McGhee, W., *Universal Barriers.* Risk Magazine, (2002).

[56] Meerschaert, M.M., Scheffler , H., Tadjeran, C., *Finite difference methods for two-dimensional fractional dispersion equation.* Journal of Computational Physics 211, 249261, (2006).

[57] Merton, R., *Option pricing when underlying stock returns are discontinuous.* J. Financial Economics, 3, pp. 125144, (1976).

[58] Musiela, M., Rutkowski, M., *Martingale Methods in Financial Modelling.* (2nd ed.). New York: Springer (2004).

[59] Oleinik, O. A., Radkevic, E. V., *Second Order Equations with Nonnegative Characteristic Form.* Plenum Press, New York, (1973).

[60] Oksendal, B., *Stochastic Differential Equations. An Introduction with Applications.* Springer, (2000).

[61] Ould Aly, S.M., *Monotonicity of Prices in Heston Model.* HAL Id: hal-00678437, (2012).

[62] Protter, P.E., *Stochastic Integration and Differential Equations.* Springer, (2003).

[63] Rebonato, R., *Volatility and Correlation: The Perfect Hedger and the Fox.* John Wiley and Sons, (2005).

[64] Revuz, D., Yor ,M., *Continuous Martingales and Brownian Motion.* Fundamental Principles of Mathematical Sciences 293, Springer, (1991).

[65] Rogers, L.C.G., Williams, D., *Diffusions, Markov Processes, and Martingales 2.* Cambridge Mathematical Library, 2nd ed., Cambridge University Press, (2000).

[66] Shreve, S.E., *Stochastic Calculus for Finance II: Continuous-Time Models.* Springer Finance, (2004).

[67] Tankov, P., Voltchkova, E, *Jump-diffusion models: a practitioners guide.*

[68] Tavella, D., Randall, C., *Pricing Financial Instruments.* Wiley, New York, (2000).