POLITECNICO DI MILANO

Scuola di Ingegneria dei Sistemi

Dipartimento di Matematica "F. Brioschi"

Corso di Laurea Magistrale in Ingegneria Matematica

Local Stochastic Volatility Models

Solving the Smile Problem with a Nonlinear Partial Integro-Differential Equation

Tesi di Laurea Magistrale in Ingegneria Matematica

Relatore:
Prof. Emilio Barucci

Correlatore Politecnico di Milano:
Doct. Daniele Marazzina

Correlatore Politecnico di Torino:
Prof. Paolo Brandimarte

Tesi di Laurea di
Daniele Cozzi
Matr. N. 764825

Anno Accademico 2011-2012

Dedicato ai miei nonni, Angelo ed Armando,
senza la cui fatica questo lavoro non sarebbe esistito.

Ringraziamenti

E così sono arrivato alla fine di questo lavoro ma soprattutto alla conclusione di un capitolo ben più importante della mia vita: cinque anni che sembrano cominciati ieri ma che ora finiscono davvero.

Ringrazio innanzitutto il Professor Barucci che al mio ritorno dall'Erasmus ha gentilmente accettato di seguirmi come relatore, assecondando il mio desiderio di svolgere una tesi di finanza in stage. Ringrazio quindi tutto il team di Financial Engineering di Mediobanca Spa, che mi ha accolto e ospitato per sei mesi, durante i quali questo lavoro è nato e si è sviluppato e dove ho imparato molto. Un ringraziamento particolare quindi al Dottor Airoldi, per la preziosa opportunità che mi ha concesso, e soprattutto al Dottor Parruccini e all'Ingegner Biancardi per la paziente e costante supervisione, per tutti i consigli, gli insegnamenti e molto altro. Voglio poi ringraziare il Dottor Marazzina per la cordiale e preziosa collaborazione e la grande disponibilità quale correlatore di questa tesi, specialmente nella parte numerica, ed il Professor Brandimarte per aver accettato di svolgere il ruolo di correlatore ASP per il Politecnico di Torino.

Questa tesi è l'ultimo atto di un lungo percorso, per questo voglio ora ringraziare tutti i compagni di avventura che hanno condiviso con me questi bellissimi cinque anni. Sono stati davvero memorabili. Faticosi, certo, il Poli non fa sconti, però adesso ricordo tutto con grande piacere e già molta nostalgia.

Ringrazio quindi quelli che ci sono sempre stati, sin dall'inizio, quando eravamo il mitico "primo scaglione": Valerio e Fabio innanzitutto. Eravamo davvero un trio fantastico in triennale, ed è per me molto importante vedere che sebbene abbiamo intrapreso strade diverse siamo rimasti comunque amici veri. Ringrazio poi Giuditta per la pazienza, la grandissima amicizia e ospitalità (a Milano come a Lund). E poi Tito, Benny, Diana, Alberto (e Riccardo), Renato e Sam. Abbiamo formato un gruppo certamente eterogeneo ma è davvero bello quando ci ritrovariamo, divertirci ancora insieme.

Un grazie a chi ho trovato man mano lungo il cammino. Certamente il terzetto degli "elettronici" Richi, Marco e Beri, quando però eravamo ancora "fisici" e con Richi e Beri quanti viaggi abbiamo fatto sulla Lecco-Milano. Voglio poi ricordare anche Matteo e Filip, con i quali ho condiviso il fantastico Erasmus ad Uppsala, e gli amici che ho trovato grazie all'esperienza dell'Alta Scuola Politecnica: gli

"informatici" Cipo, Giusto, Mladen e Alain, e i "matematici" Marco e Tommaso, e ancora Richi.

Ci tengo poi a sfruttare questa occasione anche per ringraziare quegli amici davvero importanti, che ci sono sempre stati, indipendentemente dal Poli e, soprattutto, nonostante il Poli. Il mio migliore amico Duccio e poi Eleonora, Silvia e Valeria.

Ringrazio quindi i miei genitori, Giovanni e Marina, che mi hanno permesso di realizzare tutti gli obiettivi che mi ero posto, grazie al loro costante supporto e sostegno, grazie per gli incoraggiamenti, l'educazione e i rimproveri, la comprensione e l'affetto. Ringrazio mia sorella Ilaria per la pazienza, nel sopportarmi come per l'aiutarmi. Grazie anche alle mie nonne Lucia ed Elisa, che mi hanno sempre coccolato e viziato, come é giusto che sia da parte delle nonne. Infine un grazie speciale ad Elena, senza la quale sarebbe stato tutto più triste e difficile. Semplicemente grazie.

Daniele Cozzi
Lecco, Novembre 2012

Sommario

Durante gli ultimi vent'anni vari modelli sono stati proposti per migliorare il classico paradigma di Black-Scholes per la valutazione di contratti derivati su azioni. In particolare il modello originale assumeva che la volatilità del sottostante fosse una costante σ. Al contrario, empiricamente si può osservare come la volatilità implicita σ_I, cioè quel valore che inserito nella formula di Black-Scholes permette di replicare il prezzo di mercato, non sia affatto costante, ma dipenda altresì dal prezzo di esercizio K e dalla scadenza T del contratto. Si osserva dunque sul mercato una superficie di volatilità implicita $\sigma_I(K, T)$.

Tra le varie classi di modelli proposti, due filoni di ricerca, in particolare, sono stati ampiamente sviluppati ed utilizzati: i modelli a Volatilità Locale [17, 21] e i modelli a Volatilità Stocastica [41, 44, 64], nei quali l'ipotesi originale di Black-Scholes di un coefficiente di volatilità costante viene effettivamente rilassata. I modelli a Volatilià Locale considerano la volatilità come una funzione deterministica del titolo sottostante e del tempo $\sigma_{LV}(s, t)$, detta appunto superficie di Volatilità Locale. Si assume dunque una dinamica stocastica per l'evoluzione del titolo sottostante S_t del tipo:

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

I modelli a Volatilità Stocastica, invece, considerano la volatilità stessa come un processo stocastico $b(V_t)$ assumendo dunque che:

$$\begin{aligned} dS_t &= rS_t dt + b(V_t)S_t dW_t \,, \\ dV_t &= a(V_t, t)dt + c(V_t, t)dZ_t \,, \\ dW_t\, dZ_t &= \rho\, dt \,. \end{aligned}$$

Il primo tipo di modelli permette una buona calibrazione rispetto ai prezzi quotati sul mercato delle opzioni europee. Al contrario il secondo tipo di modelli riesce a riprodurre una dinamica più realistica della volatilità implicita σ_I.

Recentemente un nuovo modello è stato proposto, generalizzando i due precedenti: il cosiddetto modello "Local-Stochastic Volatility" [57]. In questo caso la volatilità è data dal prodotto tra una componente deterministica $\sigma_{LSV}(s, t)$, la cosiddetta superficie di volatilità locale-stocastica, ed una componente stocastica $b(V_t)$.

Si assume, dunque, la seguente dinamica stocastica per l'evoluzione del titolo sottostante S_t:

$$\begin{aligned} dS_t &= rS_t dt + b(V_t)\,\sigma_{LSV}(S_t,t)S_t dW_t\,, \\ dV_t &= a(V_t,t)dt + c(V_t,t)dZ_t\,, \\ dW_t\,dZ_t &= \rho\,dt\,. \end{aligned}$$

In questo modo, utilizzando una volatilità ibrida locale-stocastica, è possibile sfruttare i vantaggi di entrambi i modelli base, i quali possono effettivamente essere interpretati come casi particolari di questo nuovo modello generalizzato.

Lo scopo di questo lavoro di tesi è definire e validare una procedura per la calibrazione di un modello Local-Stochastic Volatility per la valutazione di opzioni su azioni. L'idea chiave è quella di calibrare indipendentemente i due modelli base, ossia trovare la superficie di volatilità locale $\sigma_{LV}(s,t)$ ed i parametri del modello a volatilità stocastica a, b, c, ρ. Successivamente è possibile coniugare i parametri dei due modelli così da ottenere una superficie di volatilità locale-stocastica, consistente con i dati di mercato, grazie alla seguente relazione:

$$\sigma_{LSV}(s,t) = \frac{\sigma_{LV}(s,t)}{\sqrt{\mathbb{E}[b^2(V_t)|S_t = s]}}\,.$$

A tal scopo è possibile utilizzare due differenti metodi: risolvere un equazione di Fokker-Planck nonlineare integro-differeziale [57], oppure utilizzare un metodo di proiezioni markoviane [40]. In particolare, in questo lavoro il primo metodo è effettivamente utilizzato per calibrare il modello risolvendo numericamente la seguente equazione per la densità di probabilità congiunta $p(s,v,t)$ per il processo stocastico (S_t, V_t):

$$\begin{aligned} \frac{\partial p}{\partial t} &= \frac{1}{2}\frac{\partial^2}{\partial s^2}\Big[\sigma^2_{LV}(s,t)\,I[p]\,b^2(v)\,s^2\,p(s,v,t)\Big] + \frac{1}{2}\frac{\partial^2}{\partial v^2}\big[c^2(v,t)\,p(s,v,t)\big] \\ &\quad + \frac{\partial^2}{\partial s\,\partial v}\Big[\rho\,\sigma_{LV}(s,t)\sqrt{I[p]}\,b(v)\,c(v,t)\,s\,p(s,v,t)\Big] \\ &\quad - \frac{\partial}{\partial s}[r(t)s\,p(s,v,t)] - \frac{\partial}{\partial v}[a(v,t)\,p(s,v,t)]\,, \end{aligned}$$

dove il termine integrale non lineare $I[p](s,t)$ é dato da:

$$I[p](s,t) = \frac{\int_0^\infty p(s,v,t)\,dv}{\int_0^\infty b^2(v)\,p(s,v,t)dv} = \frac{1}{\mathbb{E}[\,b^2(V_t)|S_t = s]}\,.$$

La ricerca a proposito di tale equazione è piuttosto recente. Introdotta per la prima volta nel 2007 [57], solo nel 2010 alcuni risultati teorici circa l'esistenza di una soluzione sono stati proposti da Abergel e Tachet [1]. Nel 2011 Tachet propone una risoluzione numerica del problema con il metodo delle Differenze Finite [65] mentre Engelmann, Koster and Oeltz [23] usano il metodo dei Volumi Finiti. In questo lavoro ci proponiamo di risolvere numericamente l'equazione utilizzando il metodo degli Elementi Finiti.

Dopo una revisione generale della teoria dei modelli *"Local Stochastic Volatility"*, un particolare tipo di modello è effettivamente implementato. La celebre dinamica di Heston [41] è utilizzata per il comportamento stocastico della volatilità, mentre la formula di Dupire [21] è utilizzata per ricostruire la superficie di Volatilità Locale σ_{LV} a partire dalla superficie di Volatilià Implicita di mercato σ_I. Quest'ultima é interpolata dai dati di mercato attraverso la parametrizzazione SVI suggerita da Gatheral [29]. Il modello a volatilità Locale-Stocastica di Heston-Dupire assume dunque la seguente forma:

$$\begin{aligned} dS_t &= rS_t dt + \sigma_{LSV}(S_t,t)\sqrt{V_t}\, S_t\, dW_t\,, \\ dV_t &= \kappa(\theta - V_t)dt + \eta\sqrt{V_t}\, dZ_t\,, \\ dW_t\, dZ_t &= \rho\, dt\,, \end{aligned}$$

Una volta calibrati i parametri dei due modelli base, $\kappa, \theta, \eta, V_0, \rho$ e σ_{LV} dai dati di mercato, è possibile ricostruire la superficie di volatilità locale-stocastica $\sigma_{LSV}(s,t)$ risolvendo l'equazione integro-differenziale di Fokker-Plack. In questo modo il modello generalizzato è calibrato e pronto ad essere utilizzato per la valutazione di contratti derivati. Andiamo infine a verificare che il modello generalizzato combini effettivamente i pregi di entrambi i modelli di partenza. In particolare si otteniene una buona riproduzione dell'attuale Volatilità Implicita di mercato e al contempo una sua realistica dinamica.

Parole Chiave: Smile della Volatilità Implicita, modello a Volatilità Stocastica di Heston, parametrizzazione SVI per la Volatilità Implicita, formula di Dupire per la Volatilità Locale, calibrazione di modelli Local-Stochastic Volatility, teorema di Gyöngy, metodo delle Proiezioni Markoviane, equazione nonlineare integro-differenziale di Fokker-Planck, metodo degli Elementi Finiti.

Abstract

During the last twenty years several models have been proposed to improve the classic Black-Scholes framework for equity derivatives pricing. In particular, two main strands of research have been widely developed and used: Local Volatility [21, 17] and Stochastic Volatility [44, 64, 41]. Both these approaches relaxed the Black-Scholes hypothesis of a constant volatility. In fact, 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.

Recently a new model, generalization of the two previous ones, has been proposed: the "Local-Stochastic Volatility Model" [57]. This model considers 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.

The aim of this work is to state a clear procedure to calibrate a Local-Stochastic Volatility Model for option pricing and validate it. The key idea is to first calibrate, independently, the Local Volatility surface and the parameters of the Stochastic Volatility model. Afterwards it is possible to combine and merge the parameters of the two basic models to obtain a market consistent local-stochastic volatility surface. Two possible procedures can be used to accomplish this result: solving a nonlinear partial integro-differential Fokker-Planck equation or using a Markovian projection method. In particular, the former method is considered to calibrate the model. Thus the nonlinear Fokker-Planck equation is numerically solved using the Finite Element Method.

The research about this equation is rather recent. Introduced for the first time in 2007 [57], only in 2010 some theoretical results about existence have been proposed by Abergel and Tachet [1]. In 2011 Tachet numerically solved the problem with the Finite Difference Method [65], while Engelmann, Koster and Oeltz [23] used the Finite Volume Method. In this work we solve this equation with the Finite Element Method.

After reviewing the general theory of Local Stochastic Volatility models, a particular instance is actually considered and implemented. The well known Heston dynamics [41] is used for the stochastic behaviour of volatility, while the Dupire formula [21], combined with the SVI parametrization of Implied Volatility [29], is used for the Local Volatility component. First the two basic models are calibrated to real market prices, then the Local Stochastic Volatility surface is evaluated solving the relative nonlinear Fokker-Planck equation. In this way, the generalized model is calibrated and ready to be used for exotic derivatives pricing. Finally, we verify that the Local-Stochastic Volatility model well behaves compared to the two basic ones in terms of Implied Volatility statics and dynamics.

Keywords: Implied Volatility Smile, Heston Stochastic Volatility model, SVI parametrization, Dupire formula for Local Volatility, Local-Stochastic Volatility model calibration, Gyöngy theorem, Markovian projections, nonlinear partial integro-differential Fokker Planck equation, Finite Element Method.

Contents

Introduction

One of the central problem in modern mathematical finance is derivative pricing. 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 difficult concerning derivative pricing is to define a fair price. For this purpose a mathematical theory is needed. The well known Black-Scholes model was first introduced in 1973 and nowadays it represents an universal accepted framework for derivative pricing. In this introduction we briefly recall the main results of the standard theory, following closely [7], for further details see [63, 67]. In particular we discuss both advantages and disadvantages of the standard Black-Scholes model and we revise the main strands of research that have been proposed to improve it. Finally, we outline the structure of this thesis.

The original Black-Scholes 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:

$$dB_t = r\, B_t\, dt, \tag{1}$$
$$dS_t = \mu\, S_t\, dt + \sigma\, S_t\, dW_t, \tag{2}$$

where the deterministic constant μ, σ and r represent respectively the local mean rate of return of the asset, the volatility of the asset and the short rate interest. W_t is a standard Wiener process [51]. Let's consider a *simple contingent claim* of the form:

$$\chi = \phi(S_T), \tag{3}$$

namely a derivative paying at maturity an amount χ depending only on the value S_T of the underlying itself at maturity. The function ϕ is the so called *pay-off* at maturity of the derivative contract. Let's further assume that this contingent claim can be traded on a liquid market and that its price $\pi(t) = \pi(t; \phi)$ has the form

$$\pi(t) = F(S_t, t), \tag{4}$$

for some smooth function F. This means that the price of the derivative at subscription time t depends only on the time itself and on the value of the underlying asset S_t at time t.

Theorem 1. *(Black-Scholes Equation) Assuming that the market is specified by* (1) *and* (2)*, we want to price a contingent claim of the form* (3)*. Then the only pricing function of the form* (4) *which is consistent with the absence of arbitrage is when F is the solution of the following boundary value problem in the domain* $[0,\, T] \times \mathbb{R}^+$.

$$\begin{aligned} \frac{\partial F}{\partial t}(s,t) + rs\frac{\partial F}{\partial s}(s,t) + \frac{1}{2}s^2\sigma^2\frac{\partial^2 F}{\partial s^2}(s,t) - rF(s,t) &= 0, \qquad (5) \\ F(s,T) &= \phi(s). \qquad (6) \end{aligned}$$

This equation is precisely of the form which can be solved using a stochastic representation formula à la Feynman-Kač [51]. The solution is given by

$$F(s,t) = e^{-r(T-t)}E_{s,t}[\phi(X_T)], \qquad (7)$$

where the process X_u is defined by the dynamics:

$$\begin{aligned} dX_u &= rX_u du + \sigma X_u dW_u, \\ X_t &= s. \end{aligned}$$

The process X_t above has precisely the same form of the price process S_T. The only, but important, change is that whereas S_t has the local rate of return μ, the X_t-process has the short rate of interest r as its local rate of return. This is the so called *change of martingale measure* which implies the pricing valuation in a risk-neutral world. It is now possible to state the following central result for derivative pricing [7].

Theorem 2. *(Risk Neutral Valuation) The arbitrage free price of the claim $\phi(S_T)$ is given by $\pi(t;\phi) = F(t,S_t)$, where F is given by the formula*

$$F(s,t) = e^{-r(T-t)}\mathbb{E}^Q_{s,t}[\phi(S_T)], \qquad (8)$$

where the Q-dynamics of S are

$$dS_t = rS_t dt + \sigma S_t dW_t.$$

Let's now consider the problem of pricing a particular financial derivative. The simplest and most common example of derivative is certainly the European call option, also known as *plain vanilla* option. Let's define it formally.

Definition 3. A European call option subscribed at time t with exercise price K and time of maturity T on an underlying asset S is a contract defined by the following features:

- The holder of the option has, at time T, the right to buy one share of the underlying stock at the price K from the underwriter of the option.

- The holder of the option is in no way obliged to buy an underlying stock.
- The right to buy the underlying stock at the price K can only be exercised at time T.

Needless to say that an European call option is a simple contingent claim, for which the contract payoff function is given by

$$\phi(S_T) = \max\left(S_T - K, 0\right).$$

After some calculations it is possible to get the following famous result, which is known as Black-Scholes formula for European options.

Proposition 4 (Black-Scholes formula)**.** *The price of a European call option with strike price K and time of maturity T is given by the formula $\pi(t) = F(S_t, t)$, where*

$$F(s,t) = s\mathcal{N}[d_1(s,t)] - e^{-r(T-t)}K\mathcal{N}[d_2(s,t)]. \tag{9}$$

Here $\mathcal{N}$ denotes the cumulative density function for the normal standard distribution and

$$d_1(s,t) = \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left(\frac{s}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)\right], \tag{10}$$

$$d_2(s,t) = d_1(s,t) - \sigma\sqrt{T-t}\,. \tag{11}$$

In what follows we will indicate the Black-Scholes formula (9) for an European Call option with the notation $C_{BS}(S,t,K,T,r,\sigma)$. Since it allows a closed form formula for several kind of derivatives, the Black-Scholes is a very appealing framework. However, 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 option it is possible to invert the Black-Scholes formula finding the so-called **implied-volatility** σ_I. Thus the implied volatility is the value of σ to use in (9) such that the Black-Scholes price of a plain vanilla is equal to the actual price quoted on the market. As Rebonato wittily said [56]: *"Implied volatility is the wrong number to put in the wrong formula to obtain the right price"*.

For European options under the Black-Scholes model, calculation of the implied volatility seems to be a straightforward exercise since a closed-form presentation exists for the price. However, this closed-form doesn't allow an analytical computation of the implied volatility. Actually, to compute the implied volatility of a contract we have to solve a nonlinear equation. Given at time t the price s of an underlying asset, the interest rate r, an European call option with strike price K and time to maturity T, we want to compute the volatility $\sigma := \sigma_{K,T}$ by solving the non linear equation:

$$C_{BS}(S,t,K,T,r,\sigma) = \nu,$$

where $\nu := C^{Market}$ is the observed market-price of the option. To abbreviate the notation, we set

$$\begin{aligned} f(\sigma) &:= C_{BS}(S,t,K,T,r,\sigma) = s\mathcal{N}(d_1(\sigma)) - e^{-r(T-t)}K\mathcal{N}(d_2(\sigma)), \\ f'(\sigma) &:= \frac{\partial C_{BS}}{\partial \sigma}(S,t,K,T,r,\sigma) = s\sqrt{T-t}\,\mathcal{N}'(d_1(\sigma)), \\ d_i(\sigma) &:= d_i(\sigma,s,K,T,r) \quad \text{for } i = 1,2, \end{aligned}$$

and the non linear equation reads as:

$$f(\sigma) - \nu = 0\,. \tag{12}$$

Needless to say f is a smooth function but it depends on the variable σ in a highly nonlinear way. Therefore there is no closed-form solution. However we can easily solve this nonlinear equation numerically. Since f is differentiable, we can apply each variant of the Newton method [55]. Moreover since f is strictly monotone increasing in σ the equation has a unique solution. In particular the classical Newton-method can be used. Given an initial guess for σ_0, $\forall k > 0$ until $k > k_{max}$:

$$\sigma_{k+1} = \sigma_k - \frac{f(\sigma_k) - \nu}{f'(\sigma_k)}\,. \tag{13}$$

The final value of $\sigma_{k_{max}}$ is a good approximation of the implied volatility $\sigma_I(K,T)$.

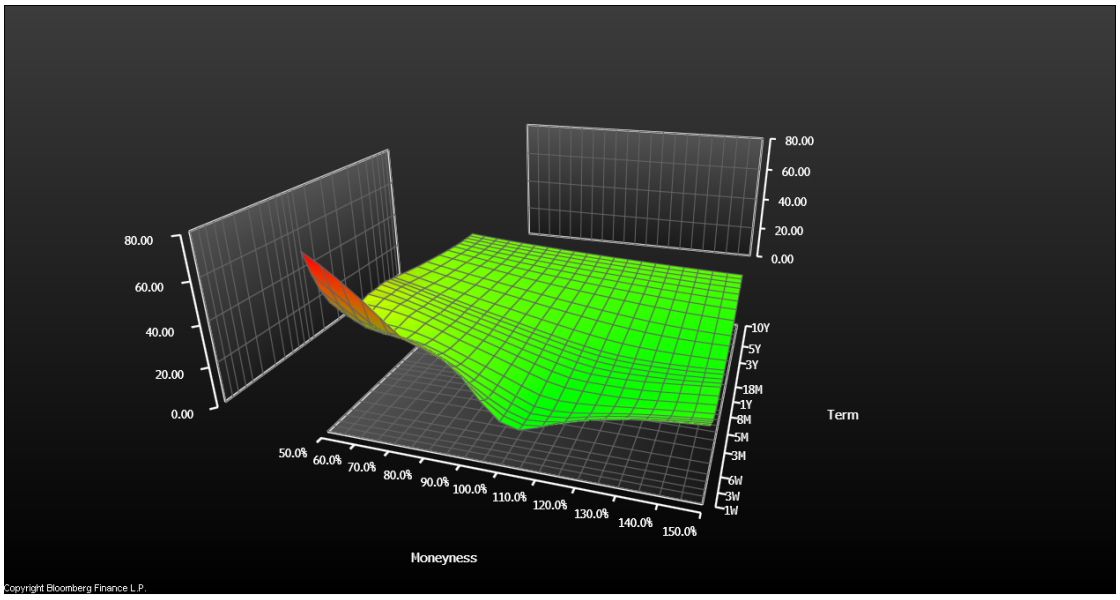


Figure 1: Implied Volatility surface of the SPX500 index at 1^{st} August 2012

The original Black-Scholes model assumes that the volatility is a constant σ across strikes and maturity dates. However it is empirically evident how σ_I depends on the value of the strike and the time to maturity of the option, namely $\sigma_I = \sigma_I(K, T)$. These two effects are, respectively, known as **volatility smile** or **volatility strike structure** and **volatility term structure** of option prices. Before providing some common characteristics of implied volatility surfaces it is useful to introduce the following terminology. Given a call option with strike price K and present value of the underlying asset S_t we say that the option is currently:

- **In-the-money (ITM)** when $S_t > K$;
- **At-the-money (ATM)** when $S_t \simeq K$;
- **Out-the-money (OTM)** when $S_t < K$;

In the first case the option is worth exercising and it is expensive while in the third case the option is worthless and it is cheap. Of course if we are dealing with a put option the terminology is reversed. At any fixed maturity, implied volatility changes with the strike price. In particular almost always in-the-money call options exhibit higher implied volatilities than out-the-money option, while the minimum of implied volatility is usually in the at-the-money region. That's way we talk about the "volatility smile" since the strike structure of implied volatility is usually concave resembling precisely a smile. Concerning the term structure of implied volatility, for any fixed strike, it varies with the maturity. Often options with longer maturity have higher implied volatilities.

In order to take into account the empirical evidence of a non constant volatility several models have been proposed during the last twenty years, developing and generalizing the Black-Scholes framework. In particular we recall the three main approaches:

- **Local Volatility Models (LVM)** Introduced for the first time in 1994 by Dupire [21] and Derman and Kani [17] 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: $\sigma = \sigma_{LV}(s, t)$.
$$dS_t = rS_t dt + \sigma_{LV}(S_t, t)\, S_t\, dW_t\,.$$
- **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 Wiener processes W_t and Z_t.
$$\begin{aligned} dS_t &= rS_t dt + b(V_t)\, S_t\, dW_t\,, \\ dV_t &= a(V_t, t)dt + c(V_t, t)\, dZ_t\,, \\ dW_t dZ_t &= \rho dt\,. \end{aligned}$$

- **Jump Diffusion Models (JDM)** Introduced by Merton [50] these models considers the underlying asset to follow a Levy process with a drift, a diffusion and a jump term;

$$dS_t = rS_t dt + \sigma S_t dW_t + S_t dJ_t \,.$$

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. For this reason we will concentrate on them.

The Local Volatility Model is very popular and rather easy to implement since it is a straightforward generalization of the original Black-Scholes framework. Its principal characteristic and major advantage is the possibility of a (nearly) perfect fit to the quoted market price. In fact as Dupire showed [21] if we had a continuum of traded vanilla prices for each strike and maturity $\Pi_{K,T}$ it would be possible to reconstruct the volatility surface $\sigma(S_t, t)$. In this way the model is consistent with the market, since it is able to reproduce the observed market prices. Unfortunately this model has a wrong implied volatility smile dynamics. Rebonato [56] outlines that the implied volatility smile generated by the LVM tends to become almost flat whereas in the reality the smile persist over time. The Stochastic Volatility Models [41, 64], have specular properties. In fact they provide a good smile dynamics over the time, but a bad fit of the present market prices. For these reasons an interesting solution can be given by the mix between the Local Volatility and a Stochastic Volatility Model, namely a generalized **Local-Stochastic Volatility Model** which combines the realistic smile dynamics of the SVM with the good fit of market price of the LVM. The idea is simply to model the diffusion coefficient of the asset dynamics as the product between a stochastic term $b(V_t)$ and a deterministic function $\sigma_{LSV}(s,t)$:

$$\begin{aligned} dS_t &= rS_t dt + b(V_t)\,\sigma_{LSV}(S_t,t)\,S_t\,dW_t\,, \\ dV_t &= a(V_t,t)dt + c(V_t,t)\,dZ_t\,, \\ dW_t dZ_t &= \rho dt\,. \end{aligned} \tag{14}$$

The academic research about this new kind of model is rather recent. The first contribution was given by Jex, Henderson and Wang in 1999 [45] who first suggested a Local-Stochastic Volatility dynamics and proposed a two-dimensional trinomial tree for the calibration. Developing the idea of mixing the three standard models (LVM, SVM and JDM) Lipton suggested in 2002 a Universal Volatility Model [49] which actually contains as a particular case the LSVM. Some years later other theoretical contributions were given by Alexander and Nogueira [3] and moreover by Ren, Madan and Qian in 2007 [57] who suggested a procedure to calibrate a LSVM. Their work has been further developed in two different strands of research. The first is based on the work of Labordère in 2009 [40] and Guyon

and Labordère in 2012 [35], who exploited the so-called Markovian projections method. The other strand of research has been developed by Abergel and Tachet in 2010 [1] and in 2011 [65] and by Engelmann, Koster and Oeltz in 2012 [23]. It is based on the solution of the Fokker-Planck equation for the probability density function of the model state variables. As we will see this is not a trivial problem since it is a nonlinear partial integral differential equation.

Remark 5 (Lipton's Universal Volatility Model). In 2002 Lipton suggested a universal model for volatility [49] combining all the three main volatility models in a unique Local-Stochastic Volatility Model with Jumps:

$$\begin{aligned} dS_t &= rS_t dt + S_t \sqrt{V_t}\, \sigma_{LSV}(S_t, t) dW_t + S_t dJ_t, \\ dV_t &= \kappa(\theta - V_t)dt + \sqrt{V_t}\, dZ_t. \\ dW_t dZ_t &= \rho\, dt\,. \end{aligned}$$

Although this is a very general and elegant model, providing a lot of flexibility and a realistic dynamics, its complexity has two major drawbacks. The price computation for a simple contingent claim and therefore the parameters calibration to market data is computationally extremely demanding. For this reason it is better to first explore the potentialities of an intermediate generalized model like the Local-Stochastic Volatility which is actually a special case of the Lipton model.

The construction of a model which is able to fit the vanilla prices and the observed volatility smile and, at the same time, that it is able to show a realistic dynamics is of primary importance for the pricing of path-dependent exotic options. This kind of products, like for instance barrier options, are extremely sensitive not only to the asset probability distribution at one time but, moreover, to the time evolution of this probability. For this reason it is critical to find a model which the on one hand provides a realistic evolution dynamics and on the other hand reproduces the observed volatility smile.

The final purpose of this work is to show how to calibrate a particular instance of Local-Stochastic Volatility model, solving with the Finite Element Method a nonlinear parabolic partial integro-differential equation. In order to calibrate this Local-Stochastic volatility to market data we need to accomplish a dual task. First we need to calibrate to market prices a Local Volatility model and a Stochastic Volatility one. Then we need to solve a non linear Fokker-Planck equation in order to merge consistently the two models. In particular this thesis is articulated according to the following structure.

Chapter 1: Understanding Volatility, where the two most common equity models are described, namely Stochastic volatility and Local volatility. A whole understanding of these two models is required since they are the foundation of the Local-Stochastic Volatility model which is built upon them. In particular two specific models, the Heston SVM and Dupire LVM, are, respectively, described with some details since they will be actually used in order to effectively implement a LSV model. The famous Heston model [41] assumes the following two-factor dynamics for the evolution of the underlying asset:

$$\begin{aligned} dS_t &= r\,S_t\,dt + S_t\,\sqrt{V_t}\,dW_t\,, \\ dV_t &= \kappa(\theta - V_t)\,dt + \eta\sqrt{V_t}dZ_t\,, \\ dW_t\,dZ_t &= \rho\,dt\,. \end{aligned} \tag{15}$$

Concerning the Local Volatility, we first interpolate the market Implied Volatility surface σ_I with the SVI parametrization suggested by Gatheral [29] and then we use the Dupire formula [21] to reconstruct the Local Volatility surface σ_{LV} from σ_I:

$$\sigma^2_{LV}(K,T) = \frac{\sigma_I^2 + 2\,T\,\sigma_I\left(\,\frac{\partial\sigma_I}{\partial T} + rK\frac{\partial\sigma_I}{\partial K}\,\right)}{\left(1 + d_1\,K\sqrt{T}\,\frac{\partial\sigma_I}{\partial K}\right)^2 + K^2\,\sigma_I\,T\left(\frac{\partial^2\sigma_I}{\partial K^2} - d_1\sqrt{T}\left(\frac{\partial\sigma_I}{\partial K}\right)^2\right)}\,. \tag{16}$$

The calibration procedure of these two models from market data is described in details. In particular it is formulated as a constrained nonlinear minimization problem of an appropriate cost functional. We also compare the two models explaining why it would be interesting to consider a generalized model. Finally, two important theoretical tools are revised: the Fokker-Planck equation and the Gyöngy theorem. They will be extensively used in the rest of the work

Chapter 2: Local Stochastic Volatility models, where the general theory about Local-Stochastic Volatility models (14) is introduced. The main idea is to first calibrate, independently, from market data, the Stochastic Volatility model parameters and the Local Volatility surface. Given the two basis models, the Stochastic Volatility model parameters can be used for the stochastic dynamics whereas given the Local Volatility surface σ_{LV} it does exists only one consistent Local-Stochastic Volatility surface σ_{LSV}:

$$\sigma_{LSV}(s,t) = \frac{\sigma_{LV}(s,t)}{\sqrt{\mathbb{E}[b^2(V_t)|S_t = s]}} \tag{17}$$

In order to effectively use this equation two different calibration procedures are presented for the model. The first is a stochastic method, based on Markovian

projections and the Gyöngy theorem. The second method is a deterministic one, based on the solution of the Fokker-Planck equation for the probability density function of the model dynamics. In particular this second method is chosen and a review of the few theoretical results nowadays known about this equation is provided.

Chapter 3: The Heston-Dupire model, where a particular instance of Local-Stochastic volatility model is considered. The Heston model is chosen for the stochastic dynamics and the Dupire formula is used for the Local Volatility surface. Thus, we obtain a Heston-Dupire Local Stochastic Volatility Model:

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{V_t}\,\sigma_{LSV}(S_t,t)S_t dW_t\,, \\ dV_t &= \kappa(\theta - V_t)dt + \sqrt{V_t}\,dZ_t\,, \\ dW_t dZ_t &= \rho dt\,. \end{aligned}$$

In order to calibrate from market data the local stochastic volatility surface σ_{LSV} of the model, the relative non linear Fokker-Planck equation is considered:

$$\begin{aligned} &\frac{\partial p}{\partial t} - \frac{1}{2}\frac{\partial^2}{\partial s^2}\Big[v\,s^2\,\sigma_{LV}^2\,I[p]\,p\Big] + \frac{\partial^2}{\partial s\,\partial v}\Big[\rho\,\sigma_{LV}\,I[p]\,\eta\,v\,s\,p\Big] \\ &\quad + \frac{1}{2}\frac{\partial^2}{\partial v^2}\big[\eta^2\,v\,p\big] - \frac{\partial}{\partial s}\Big[r\,s\,p\Big] - \frac{\partial}{\partial v}[\kappa(\theta - v)\,p] \;=\; 0\,, \end{aligned} \tag{18}$$

where the nonlinear integral term $I[p]$ is the inverse of the expected conditional value:

$$I[p](s,t) = \frac{\int_0^\infty p(s,v,t)\,dv}{\int_0^\infty v\,p(s,v,t)dv} \;= \frac{1}{\mathbb{E}[V_t|S_t = s]}\,. \tag{19}$$

Once the equation is solved, we can finally reconstruct the Local-Stochastic volatility surface , $\sigma_{LSV}(x,t)$, thanks to (17):

$$\sigma_{LSV}(s,t) = \sigma_{LV}(s,t)\sqrt{I[p](s,t)}\,. \tag{20}$$

The equation is essentially a time dependent diffusion-advection-reaction equation with a non linear diffusion coefficient. Moreover it is also degenerate to the boundary. For this reason we revise the main theoretical results about this issue. Finally the numerical approximation of the problem with the Finite Element Method is described.

Chapter 4: Numerical Simulations, where the main numerical results are presented. Starting from real market prices we calibrate the Heston SVM and the Dupire LVM, as described in Chapter 1. Afterwards we solve the non linear Fokker-Planck equation and we calibrate in this way the Heston-Dupire model.

Once all the three models are calibrated we compare them with the original market Implied Volatility. Finally we price with the three models Forward-Starting options with a Monte Carlo method in order to compare the Forward Implied Volatility of the models as well.

Chapter 1

Understanding Volatility

Before introducing a Local-Stochastic Volatility model a certain understanding of traditional volatility modelling is required. In fact the generalised model is built upon the two most common equity models, namely Stochastic Volatility and Local Volatility. For this reason in this chapter we provide some details about the two basic models and their relationship. In particular we consider the Heston stochastic volatility model and the formula suggested by Dupire for the local volatility. These models are presented and their calibration to market data described. This chapter is intended to provide some useful tools that will be needed in the rest of this work. Thus we revise also the main results concerning the Fokker-Planck equation and the Gyöngy theorem.

1.1 Stochastic Volatility

A simple observation of equity markets would make natural to model the volatility itself as a stochastic process. This is precisely the main feature of a stochastic volatility model (SVM). While the standard Black-Scholes model assumes a constant volatility term σ a SVM considers volatility as a function $b(\cdot)$ of a stochastic process V_t. Generally speaking a stochastic volatility model assumes the following dynamics.

Definition 1.1 (Stochastic Volatility Model - SVM)**.**

$$\begin{aligned} dS_t &= \mu\, S_t\, dt + b(V_t) S_t dW_t\,, \\ dV_t &= a(V_t, t)dt + c(V_t, t)dZ_t\,, \\ dW_t\, dZ_t &= \rho\, dt\,. \end{aligned} \tag{1.1}$$

As usual S_t denotes the underlying asset, t the time, μ the (deterministic) instantaneous drift and W_t, Z_t are two Wiener processes with correlation ρ. Using non-arbitrage arguments [30], [67] it is possible to show that the price of an

European call option under a SVM satisfies the following equation:

$$\frac{\partial C}{\partial t} + \frac{1}{2}\, s^2\, b^2(v)\, \frac{\partial^2 C}{\partial s^2} + \frac{1}{2}\, c^2(v,t) \frac{\partial^2 C}{\partial v^2} + \rho\, b(v)\, c(v,t)\, s\, \frac{\partial^2 C}{\partial v \partial s} + r\, s\, \frac{\partial C}{\partial s} + \big(a(v,t) - \lambda c(v,t)\big) \frac{\partial C}{\partial v} - r\, C = 0\,.$$

In this equation two new parameters have been introduced, namely r, the usual risk free interest rate, and λ the so called *market price of volatility risk*. While the use of r instead of μ has been already explained previously, describing the Black-Scholes model, some words are needed about λ. The standard BS model assumes only one source of randomness W_t related to one traded asset S_t. In this way it is possible to hedge the risk generated by W_t through S_t. Hence the model is said to be complete, see [7]. On the contrary 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. In fact 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 several different martingale measures and then the same derivative has several possible prices depending on λ. Once the value of λ is chosen it is possible to define a risk neutral drift $\tilde{a} = a - \lambda c$ for the process V_t. In this way it is possible to redefine the dynamics for the SV model in the risk-neutral world as follows:

$$\begin{aligned} dS_t &= r\, S_t\, dt + b(V_t) S_t dW_t\,, \\ dV_t &= \tilde{a}(V_t,t)dt + c(V_t,t)dZ_t\,, \\ dW_t\, dZ_t &= \rho\, dt\,, \end{aligned} \tag{1.2}$$

As observed by Gatheral [30] if we were interested in the connection between the time series of past returns of the underlying and option prices then we would need to investigate the connection between the statistical measure under which the drift of the process V_t is a and the risk-neutral measure under which the drift is given by $\tilde{a}$. However this is not the aim of the present work. Indeed we are interested in fitting models to quoted option prices. From now on, we always assume the dynamics of the various models in risk-neutral terms. In fact we will calibrate the model parameters from market prices imputing in this way the risk-neutral measure. For this reason from now on we will consider only the risk-neutral measure, effectively ignoring the market price of volatility risk.

1.1.1 The Heston model

The Heston model was introduced in 1993 [41] and nowadays it is probably the most popular stochastic volatility model. For this reason it will be used for the stochastic component of the Local-Stochastic Volatility model implemented in chapter 3. In this section we present the main features of this model which will be used in rest of this work. The underlying asset S_t follows the usual log-normal dynamics while the square of the volatility, the variance V_t is a Cox-Ingersoll-Ross (CIR) process [14]:

$$\begin{aligned} dS_t &= r_t\, S_t\, dt + S_t \sqrt{V_t}\, dW_t \\ dV_t &= \kappa(\theta - V_t)\, dt + \eta\sqrt{V_t} dZ_t \\ dW_t\, dZ_t &= \rho\, dt \end{aligned} \tag{1.3}$$

The Heston model is characterized by five constant parameters, namely κ, θ, η, ρ and the initial value of the variance V_0. The parameter θ can be thought as the long term variance, κ as the rate of mean reversion and η as the volatility of volatility. As usual ρ represents the instantaneous correlation between the Brownian motions W_t and Z_t. Since we cannot directly observe V_0 as we do for S_0 we need to calibrate also the initial condition of the variance. Thus we consider V_0 as the fifth parameter. In order to use the model we need to calibrate from the market all these five parameters: κ, θ, η, ρ are strictly positive while $\rho \in (-1, 1)$, being a correlation.

The CIR process is continuous, positive and mean-reverting. These properties make it particularly suited for modelling the square of the volatility V_t. In particular the behaviour of the process near zero is characterized by the so called Feller condition [6, 32]:

$$Fe = \kappa\,\theta - \frac{1}{2}\eta^2 \geq 0\,. \tag{1.4}$$

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. On the contrary, if $Fe < 0$, there is a positive probability that V_t becomes zero for a moment, however, in this case the drift term push the variance to be positive again.

The Heston model has several properties which makes it very suitable for equity option pricing. Stochastic variance is mean-reverting, continuous and positive. The model allows a good fit of market implied volatilities and a realistic smile dynamics. 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 critical when calibrating the model to known option prices. Let's

consider the Heston pricing equation:

$$\begin{aligned}\frac{\partial C}{\partial t} + \frac{1}{2}\, s^2\, v\, \frac{\partial^2 C}{\partial s^2} + \frac{1}{2}\, \eta^2\, v\, \frac{\partial^2 C}{\partial v^2} + \rho\, \eta\, s\, v\, \frac{\partial^2 C}{\partial v \partial s} \\ + r\, s\, \frac{\partial C}{\partial s} + \kappa[\theta - v]\frac{\partial C}{\partial v} - r\, C = 0\,,\end{aligned} \tag{1.5}$$

with the proper initial and boundary conditions. In its original work, Heston looked for a solution similar to the Black-Scholes'one, namely:

$$C(S_t, V_t, t, T) = S_t\, P_1 - K\, e^{-r(T-t)}\, P_2\,. \tag{1.6}$$

He managed to show that this is indeed a solution of the equation defined as follows:

$$\begin{aligned} P_j(S_t, V_t, t, T) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{e^{-i\omega \ln(K)}}{i\omega} f_j(S_t, V_t, t, T, \omega)\right) d\omega\,, \\ f_j(S_t, V_t, t, T, \omega) &= e^{C(T-t,\omega) + D(T-t,\omega)\, V_t + i\,\omega \ln(S_t)}\,, \\ C(\tau, \omega) &= i\omega r + \frac{\kappa\theta}{\eta^2}\left[(b_j - \rho\eta\omega i + d)\,\tau - 2\ln\left(\frac{1 - g e^{d r}}{1 - g}\right)\right]\,, \\ D(\tau, \omega) &= \frac{b_j - \rho\eta\omega i + d}{\eta^2}\left(\frac{1 - e^{d r}}{1 - g e^{d r}}\right)\,, \\ g &= \frac{b_j - \rho\eta\omega i + d}{b_j - \rho\eta\omega i - d}\,, \\ d &= \sqrt{(\rho\eta\omega i - b_j)^2 - \eta^2(2u_j\omega i - \omega^2)}\,, \end{aligned}$$

for $j = 1, 2$, where:

$$u_1 = \frac{1}{2}\,, \quad u_2 = -\frac{1}{2}\,, \quad b_1 = \kappa - \rho\eta\,, \quad b_2 = \kappa\,,$$

Despite this formula looks quite demanding, it is actually rather explicit, easy and fast to evaluate. The only part that requires some computational effort is the evaluation of the integral along a not bounded interval. However such integration can be performed using standard numerical methods.

Although formula (1.6) allows a fast way to compute the Heston prices of plain vanillas a further important improvement has been carried out from Carr and Madan in 1998 [11]. Their idea was to evaluate from the formula (1.6) the probability distribution of the final stock price and from this evaluate the characteristic function, namely the Fourier transform of the probability distribution. Carr and Madan [11] suggested a very efficient and stable method to compute vanilla prices using the Fast Fourier Transform [55] for the valuation of this characteristic function. With this ultimate method the pricing of European options

according to the Heston model is very fast and computationally cheap. Indeed, we will use this FFT method for the calibration of the Heston model, however we do not revise it here since its derivation is rather technical and the procedure is nowadays a common used techniques. For a good description, aside the work of Carr and Madan [11] itself, see [60].

1.1.2 Calibration of Heston parameters

In order to use the Heston model for derivative pricing we first need to find its five parameters: $V_0, \kappa, \theta, \eta, \rho$. A fundamental requirement for any model is to be consistent with the market. Since the most liquid market, concerning equity derivatives, is the market of plain vanillas, a natural choice is to calibrate the model to this market. The calibration of a financial model consists in finding those parameters such that the model prices fit the market ones. This is a well known example of inverse problem and in particular it is known as the inverse problem of mathematical finance. Let's define it more formally.

Let $C^{Market}_{T_i,K_j}$ be the price quoted on the market at the present time $t = 0$ of an European Call option with time maturity T_i and strike price K_j. In particular at any time t there are quoted on the market several call options on the same underlying for different maturities and strikes. Thus, let's define $\{T_i\}_{1,N}$ and $\{K_j\}_{1,M}$ respectively the vector of market maturities and strike prices. Let's now consider the price given by the Heston model concerning the same European call option. We define it as $C^{Heston}_{T_i,K_j}(\mathbf{x})$ where $\mathbf{x} = [V_0, \kappa, \theta, \eta, \rho]$ is the vector of the model parameters. Let $W \subset \mathbb{R}^5$ be the feasible region for these coefficients. The inverse problem we want to solve reads as follows:

$$\text{find } \mathbf{x} \in W \text{ such that } C^{Heston}_{T_i,K_j}(\mathbf{x}) = C^{Market}_{T_i,K_j} \quad \forall i = 1, \cdots, N\,, \forall j = 1, \cdots, M\,.$$

For several reasons this inverse problem turns out to be ill-posed: existence and uniqueness of the solution $\mathbf{x}$ are not guaranteed, as well as a continuous dependence of the solution to market data. In order to calibrate the model it is common practice to transform the inverse problem in an optimization problem. In particular a constrained, non linear least squares minimization problem:

$$\text{find } \mathbf{x} \in W \text{ such that } \mathcal{J}(\mathbf{x}) = \min_{\mathbf{y} \in W} \mathcal{J}(\mathbf{y})\,, \tag{1.7}$$

where the cost functional $\mathcal{J}$ can be defined as:

$$\mathcal{J}(\mathbf{y}) = \frac{1}{2} \sum_{i,j}^{N,M} \omega_{ij} \left(\frac{C^{Heston}_{T_i,K_j}(\mathbf{y}) - C^{Market}_{T_i,K_j}}{C^{Market}_{T_i,K_j}} \right)^2 . \tag{1.8}$$

In this way the cost functional is the sum of the relative errors between model and market prices. It is important to consider the normalised error because

of the great difference between *out-the-money* and *in-the-money* option prices. Moreover since the most sensitive region the model should be able to reproduce is *at-the-money* region it is useful to use the weights ω_{ij}. In this way when minimizing the functional it is possible to leverage the importance assigned to the central part of the smile rather than the external region. Concerning the definition of the feasible set W the choice is not unique. The parameters $\kappa, \theta, \eta, v_0$ must be positive and usually smaller than 1 while the correlation ρ is between 1 and -1, although the extreme cases can produce some misbehaviour. For this reason we can set $W = [0.05, 1.00]^4 \times [-0.99, 0.99]$.

For a numerical solution of the constrained non linear problem (1.7) several iterative methods can be applied. One of the most used is the Levenberg-Marquardt algorithm [34]. However regardless by the particular optimization method used a first guess solution is needed. Moreover, since the optimal solution is generally a local minimum and not a global one, the values of the initial solution are rather critical. Thus it is very important to find a good first guess solution. To do this a very useful tool is the **asymptotic analysis of implied volatility** as suggested in [28, 32]. We do not revise here this topic, widely treated in literature, but we only provide some useful results concerning the Heston model [30]. The idea is to find some asymptotic relations between the implied volatility observed on the market and the parameters of the model. In particular it is possible to prove the following asymptotic behaviour [30].

$$\sigma_I^2(K,T)|_{K=S_0} \simeq v_0 \qquad \text{for } T \to 0 \tag{1.9}$$

$$\sigma_I^2(K,T)|_{K=F_T} \simeq \theta \qquad \text{for } T \to \infty \tag{1.10}$$

$$\frac{\partial \sigma_I^2}{\partial x}\Big|_{K=S_0} \simeq \frac{\rho\eta}{2} \qquad \text{for } T \to 0 \tag{1.11}$$

F_t represents the Forward value of the underlying asset. In the standard Black-Scholes model this is equal to $F_t = S_0 e^{rt}$. The notation x, instead, is used to represent the forward log-moneyness, namely $x = \log(\frac{F}{K})$. Using relations (1.9) and (1.10) it is possible to find an initial guess solution for v_0 and θ. Since generally ρ is approximately equal to -0.8 using (1.11) it is possible to find a value for the volatility of volatility η. Finally to set an initial value for the rate of mean reversion we can choose a value $\kappa \simeq 0.05 + \frac{\eta^2}{2\theta}$. In this way the Feller condition is fulfilled. Summarizing we can use the following first guess condition:

$$\begin{aligned} v_0 &= \sigma_I^2(S_0, 0)\,, \\ \theta &= \sigma_I^2(S_0\, e^{rT_{max}}, T_{max})\,, \\ \rho &= -0.8\,, \\ \eta &= \frac{2\,S_0}{\rho}\,\frac{\partial \sigma_I^2(S_0,0)}{\partial S}\,, \\ \kappa &= 0.05 + \frac{\eta^2}{2\theta}\,. \end{aligned} \tag{1.12}$$

Remark 1.2 (Imposing the Feller condition). A further constrain during the calibration procedure is given by the Feller condition (1.4). Actually the Heston model is well defined even if this condition is violated. However it is a good practice to find a solution which satisfies this condition. To accomplish such result a natural choice is to redefine the feasible set W including this further constrain. However given the non linear nature of this constrain a more efficient way is to impose it with a barrier method [34]. For instance we could minimize the modified cost functional:

$$\tilde{\mathcal{J}} = \mathcal{J} - \omega \log(Fe)\,,$$

where ω is an appropriate tuning parameter and $Fe = 2\kappa\theta - \eta^2$ is the Feller number. With this modified functional it is also important to set a first guess solution such that $Fe > 0$.

Remark 1.3 (Computational considerations). It is important to notice that for every valuation of the cost functional (1.8) we need to compute $N \times M$ vanilla prices with the model. For this reason the time required by the calibration procedure can be quite long. Therefore it is so important to have a semi-closed formula to price plain vanillas as the one found by Heston. Moreover the FFT-method proposed by Carr and Madan further increased the computational efficiency of this semi-closed formula enabling a fast calibration of the Heston model.

1.2 Local Volatility

A not negligible drawback of Stochastic Volatility models is that introducing a new source of randomness we lose the completeness of the model and then the uniqueness of the price. Moreover although they can produce a consistent smile they cannot fit exactly the market prices. For this reason in 1994 Dupire [21] and Derman and Kani [17] introduced a new model generalizing the Black-Scholes'one. They consider a non constant deterministic volatility $\sigma_{LV}(S,t)$, called Local Volatility surface, and they assume the following stochastic differential equation for the evolution of the underlying asset.

Definition 1.4 (Local Volatility Model - LVM).

$$dS_t = (r-d)\,S_t\,dt + \sigma_{LV}(S_t,t)S_t dW_t \tag{1.13}$$

The corresponding pricing equation is a straightforward generalization of the Black-Scholes equation. Thus the price of an European Call option can be computed simply solving the problem below.

Proposition 1.5 (Generalized Black-Scholes equation). *Under a Local Volatility model the price of an European Call option is given by the following generalized Black-Scholes equation:*

$$\begin{cases} \frac{\partial C}{\partial t} + \frac{1}{2}\sigma_{LV}^2(S,t)S^2\frac{\partial^2 C}{\partial^2 S} + (r-d)S\frac{\partial C}{\partial S} - rC &= 0 \quad \text{on } Q = [0,T)\times(0,\infty) \\ C(0,t) &= 0 \quad \forall t \in (0,T) \\ \lim_{S\to\infty} C(S,t) - Se^{-d(T-t)} + Ke^{-r(T-t)} &= 0 \quad \forall t \in (0,T) \\ C(T,S) &= (S-K)^+ \quad \forall S \in (0,\infty) \end{cases}$$

This model seems to be a simple and straightforward generalization of the original Black-Scholes framework since we are simply considering a non constant, deterministic, diffusion coefficient. However it is not straightforward as well to understand how to extract the surface σ_{LV} from the market. Dupire succeed to prove that assuming a continuum of traded plain vanilla on the market it does exist one and only one Local Volatility surface σ_{LV} such that the model is able to perfectly reproduce the market quoted prices. Moreover he provides a closed form formula, known as Dupire'formula, to evaluate σ_{LV}. This formula is actually a corollary of the so called Dupire's equation which is a sort of Fokker-Planck equation where the unknown is the option price instead of the probability density functions. For this reason, before presenting the Dupire's work, we review the essential theory of Fokker Planck equation. Indeed we will need a bit understanding of it also for the next chapters.

1.2.1 The Fokker-Planck Equation

In this section we recall some essential results concerning the Fokker Planck equation (also known as Kolmogorov forward equation). This is the equation ruling the transition probability density function for the solution of a linear stochastic differential equation. The treatment follows closely chapter five of [7]. For further details concerning analytical and numerical solution of Fokker-Planck equation see [58]. Suppose that X_t is a solution of the linear SDE

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t\,. \tag{1.14}$$

Hence the associate infinitesimal generator $\mathcal{A}$ is given by

$$\mathcal{A}f(x,t) = \mu(x,t)\,\frac{\partial f}{\partial x}(x,t) \,+\, \frac{1}{2}\,\sigma^2(x,t)\,\frac{\partial^2 f}{\partial x^2}(x,t)\,.$$

Let's now consider the boundary value problem

$$\begin{cases} \frac{\partial u}{\partial t} + \mathcal{A}u &= 0, \qquad (x,t) \in \mathbb{R} \times (0,\,T) \\ u(x,T) &= I_B(x), \qquad x \in \mathbb{R} \end{cases} \tag{1.15}$$

where I_B is the indicator function of the set $B \subset \mathbb{R}$. It is possible to prove that

$$u(x,t) = \mathbb{E}^{x,t}[I_B(X_T)] = P(X_T \in B | X_t = x),$$

where X_t is the solution of (1.14). This is a very interesting connection between stochastic differential equations and parabolic differential equations. The solution of (1.15) is equal to the probability that the process X_t starting from x at time t belongs to B at time $T > t$. Let's state this result in a formal way.

Proposition 1.6 (Kolmogorov backward equation)**.** *Let X_t be a solution to (1.14). Then the transition probabilities $P(x,t;B,T) = P(X_T \in B | X_t = x)$ are given as the solution to the equation*

$$\begin{cases} \frac{\partial P}{\partial t} + \mathcal{A}P &= 0, \qquad (x,t) \in \mathbb{R} \times (0,\,T) \\ P(x,T) &= I_B(x), \qquad x \in \mathbb{R} \end{cases} \tag{1.16}$$

This equation is known as backward Kolmogorov equation since the differential operator $\mathcal{A}$ works on the "backward variables" (x,t). In fact it is considered the backward probability imposing a terminal condition and the resulting partial differential equation actually evolves backward in time. It is possible to state a dual result known as Kolmogorov forward equation or **Fokker Planck equation** where we look for a forward probability imposing an initial condition.

Proposition 1.7 (Kolmogorov forward equation)**.** *Let X_t be the solution of equation* (1.14) *with the initial condition $X_0 = x_0$ and assume that X_t has a transition density $p(x,t;x_0,0)$. Then p will satisfy the following Kolmogorov forward equation also known as Fokker-Planck equation:*

$$\begin{cases} \frac{\partial p}{\partial t} - \mathcal{A}^* p &= 0, \qquad (x,t) \in \mathbb{R} \times \mathbb{R}^+ \\ p(x,0) &= \delta(x - x_0), \qquad x \in \mathbb{R} \end{cases} \tag{1.17}$$

Where the adjoint operator $\mathcal{A}^*$ is defined by:

$$\mathcal{A}^* f(x,t) = -\frac{\partial}{\partial x}(x,t)\left[\mu(x,t)\, f(x,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\sigma^2(x,t)\, f(x,t)\right] .$$

It is straightforward to extend this result to the multidimensional case which, actually, will be used in the next chapter. Considering the multidimensional process $\mathbf{X}_t$ satisfying the vectorial stochastic differential equation:

$$d\mathbf{X}_t = \boldsymbol{\mu}(\mathbf{X}_t, t)dt + \boldsymbol{\sigma}(\mathbf{X}_t, t)d\mathbf{W}_t \,,$$

where $\boldsymbol{\mu} : \mathbb{R}^N \times \mathbb{R}^+ \to \mathbb{R}^N$ is the drift vector and $\boldsymbol{\sigma} : \mathbb{R}^N \times \mathbb{R}^+ \to \mathbb{R}^{N\times N}$ is the diffusion matrix, both considered as known, deterministic and bounded and Lipschitz continuous functions. The Fokker-Planck equation is:

$$\begin{cases} \frac{\partial p}{\partial t} - \mathcal{A}^* p &= 0, \qquad (\mathbf{x},t) \in \mathbb{R}^N \times \mathbb{R}^+ \\ p(\mathbf{x},0) &= \delta(\mathbf{x} - \mathbf{x_0}), \qquad x \in \mathbb{R}^N \end{cases} \tag{1.18}$$

and the adjoint operator $\mathcal{A}^\star$ is defined by

$$\mathcal{A}^\star f(s,y) = -\sum_{i=1}^{N} \mu_i(s,y)\frac{\partial f}{\partial y_i}(s,y) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial^2}{\partial y_i \partial y_j}[\tilde{\sigma}_{ij}(s,y)f(s,y)]$$

where $\tilde{\sigma}_{ij}^2 = \frac{1}{2}\sum_{k=1}^{N}\sigma_{ik}(\mathbf{x},t)\sigma_{jk}(\mathbf{x},t) = \frac{1}{2}[\sigma(\mathbf{x},t)\sigma^T(\mathbf{x},t)]_{ij}$

1.2.2 The Dupire Equation

The Black-Scholes backward parabolic equation in the variables (S,t) is the *Feynman-Kač* 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**. This equation derived for the first time by Dupire [21] is actually the Fokker-Planck equation for the probability density function of the underlying asset integrated twice. While we need to solve one Black-Scholes equation for each strike and maturity given, solving the Dupire's equation allows us to get the price of an European call option for every strike and maturity, given the present spot value and time.

Proposition 1.8 (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 is given by the following forward parabolic equation known as Dupire's equation.*

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

Proof. Let's consider the transition probability density function $p(S_T,T;S_0,0)$ for the dynamics (2.1) with the initial value S_0. The value of an option is the discounted value of the expected pay-off, for a Call option it means that:

$$\begin{aligned} C(K,T) &= e^{-rT}\int_0^\infty (S_T-K)^+\,p(S_T,T;S_0,0)\,dS_T \\ &= e^{-rT}\int_K^\infty (S_T-K)\,p(S_T,T;S_0,0)\,dS_T\,. \end{aligned} \tag{1.19}$$

Since the final pay-off is the maximum function, it is possible to find a particular form for the probability density function deriving the above expression twice with respect to the strike price, in fact:

$$\frac{\partial C}{\partial K} = -\,e^{-rT}\textstyle\int_K^\infty p(S_T,T;S_0,0)dS_T\,, \tag{1.20}$$

$$\frac{\partial^2 C}{\partial^2 K} = e^{-rT}\,p(K,T;S_0,0)\,. \tag{1.21}$$

We just derived an interesting result originally due to Breeden and Litzenberg [9]. The transition distribution of the asset process is the second derivative of the call price respect to the strike:

$$p(K,T;S_0,0) = e^{rT}\frac{\partial^2 C}{\partial^2 K}\,. \tag{1.22}$$

The time evolution of this probability density function is described by the Fokker-Planck equation (1.17).

$$\frac{\partial p}{\partial T} = \frac{1}{2}\frac{\partial^2}{\partial S_T^2}\left[\sigma_{LV}(S_T,T)^2\,S_T^2\,p(S_T,T)\right] - (r-d)\frac{\partial C}{\partial S_T}\left[S_T\,p(S_T,T)\right]. \tag{1.23}$$

Performing the time derivative of relation (1.19) we get:

$$\frac{\partial C}{\partial T} = -rC + e^{-rT}\int_K^\infty (S_T-K)\frac{\partial p}{\partial T}dS_T \tag{1.24}$$

Substituting the time derivative of the probability inside the integral as in equation (1.23) the above equation can can be written as:

$$\frac{\partial C}{\partial T} + rC = e^{-rT} \int_K^\infty (S_T - K)\Big(\frac{1}{2} \frac{\partial^2}{\partial S_T^2} \big[\sigma_{LV}(S_T,T)^2\, S_T^2\, p\big] - (r-d)\frac{\partial C}{\partial S_T}[S_T\, p]\Big) dS_T\,.$$

Integrating the right hand side by parts twice and assuming that $p(S_T, T)$ decays exponentially fast for $S_T \to \infty$ we obtain:

$$\begin{aligned}
\frac{\partial C}{\partial T} + rC =& e^{-rT}\,\frac{1}{2}\,\Big(\Big[(S_T-K)\frac{\partial}{\partial S_T}[\sigma_{LV}(S_T,T)\,\sigma_{LV}^2\,p(S_T,T)]\Big]_{S_T=K}^{S_T=\infty} \\
&- \int_K^\infty \frac{\partial}{\partial S_T}[\sigma_{LV}(S_T,T)\,\sigma_{LV}^2\,p(S_T,T)]dS_T\Big) \\
&- e^{-rT}(r-d)\Big(\big[(S_T-K)\,S_T\,p(S_T,T)\big]_{S_T=K}^{S_T=\infty} - \int_K^\infty S_T\,p(S_T,T)\,dS_T\Big) \\
=& -\frac{1}{2}e^{-rT}[\sigma_{LV}(S_T,T)^2 S_T^2 p(S_T,T)]_{S_T=K}^{S_T=\infty} + (r-d)e^{-rT}\int_K^\infty S_T p(S_T,T)dS_T \\
=& \frac{1}{2}e^{-rT}\sigma_{LV}(K,T)^2K^2p(K,T) + (r-d)\Big(C + e^{-rT}K\int_K^\infty p(S_T,T)dS_T\Big)\,.
\end{aligned}$$

Substituting in the first and second term of the right hand side respectively (1.21) and (1.20) we finally obtain:

$$\frac{\partial C}{\partial T} + rC = \frac{1}{2}\sigma_{LV}(K,T)^2K^2\frac{\partial^2 C}{\partial K^2} + (r-d)\Big(C + K\frac{\partial C}{\partial K}\Big)\,.$$

This is indeed the Dupire's equation we were looking for:

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

□

1.2.3 The Dupire Formula

Thanks to the Dupire equation we have accomplished a double result. In fact, on the one hand we have now a very useful, dual equation for derivative pricing in the two variable (K, T). On the other hand, we have now a formula to evaluate the local volatility $\sigma_{LV}(s,t)$ from option prices, known as **Dupire formula**:

$$\sigma_{LV}^2(K,T) = 2\,\frac{C_T + (r-d)\,KC_K + d\,C}{K^2C_{KK}} \tag{1.25}$$

Assuming a continuum of option prices quoted on the market for every strikes K and time to maturity T thanks to the above formula it is possible to easily

evaluate the local volatility surface. Moreover this formula ensures existence and uniqueness of a local volatility surface which reproduces exactly the market prices. Unfortunately it is not possible to observe on the market a continuum of plain vanilla prices. In fact only some options with certain strikes and maturities are actually traded. Therefore it is not possible to use directly (1.25) to evaluate σ_{LV} for every K and T. In particular it is needed to interpolate and extrapolate the Call prices from the market and then to numerically approximate the derivatives. This procedure is rather sensitive to numerical instabilities and errors. Particularly critical is the second derivative $\frac{\partial^2 C}{\partial K^2}$ at the denominator which stands alone by itself. This derivative can be very small for options deeply in-the-money or out-the-money and then very sensitive to numerical errors. Furthermore this value is multiplied by K^2 resulting in big errors, sometimes even producing negative values and then resulting in negative variance. Because of these drawbacks the Dupire formula, practically speaking, is not very useful. However it is possible to reformulate it in a more suitable form.

Market prices of options are often quoted in terms of Black-Scholes implied volatility $\sigma_I(K,T)$. Thus it is quite natural to reformulate Dupire formula (1.25) in term of implied volatilities instead of option prices. The idea [27, 30] is to insert the Black-Scholes formula (9) and its derivative into the Dupire formula (1.25) assuming that:

$$C = C_{BS}(S_t, t, K, T, \sigma_I(K,T))\,.$$

Applying the chain rule of differentiation and using the Black-Scholes formula (9) and its T- and K- derivatives it is finally possible to prove [30] that:

$$\sigma_{LV}^2(K,T) = \frac{\sigma_I^2 + 2\,T\,\sigma_I\left(\frac{\partial \sigma_I}{\partial T} + (r-d)K\frac{\partial \sigma_I}{\partial K}\right)}{\left(1 + d_1\,K\sqrt{T}\,\frac{\partial \sigma_I}{\partial K}\right)^2 + K^2\,\sigma_I\,T\left(\frac{\partial^2 \sigma_I}{\partial K^2} - d_1\sqrt{T}\left(\frac{\partial \sigma_I}{\partial K}\right)^2\right)}\,. \qquad (1.26)$$

Where

$$d_1 = \frac{\log\left(\frac{S_0}{K}\right) + (r - d + \frac{1}{2}\sigma_I^2)T}{\sigma_I\sqrt{T}}\,.$$

Although the above formula does not look as nice as (1.25), from a practical point of view it is actually more useful. In particular, now, there is no more a second derivative in the denominator as a lone term. In fact, the second derivative of the implied volatility still appears but now it is one term of a summation. Thus small errors in the approximation of the second derivative will not necessarily lead to large errors in the evaluation of the local volatility surface. For this reason, it seems to be a better idea to use (1.26) instead of (1.25). However, there is still the problem that implied volatility is not known as a continuous function of strike prices and maturities. Instead only few values $\sigma_{I,(K_i,T_j)}$ are known for certain quoted couples (K_i, T_j) of strikes and maturities. Thus some method has to be

used to interpolate and extrapolate the few data points available from the market $\sigma_{I,(K_i,T_j)}$ into a continuous surface $\sigma_I(K,T)$. For this purpose a wide literature has been produced, (see for instance [27, 26, 43]) and several class of methods have been proposed. One of these, the parametric class, assumes that Implied Volatility is a continuous function of K and T depending on some parameters. Once these parameters are correctly calibrated to market prices we actually get $\sigma_I(K,T)$. We now consider a particular parametric model.

1.2.4 SVI parametrization of Implied Volatility

In 2004 Gatheral [29] proposed an interesting parametrization of the implied volatility surface called *Stochastic Volatility Inspired* (SVI). The model assumes that for a given maturity T_n the strike structure of implied volatility is given by the following function:

$$\sigma_I^{SVI}(x,T_n) = \sqrt{C_1^n + C_2^n \left[C_3^n (x - C_4^n) + \sqrt{(x-C_4^n)^2 + C_5^n} \right]}. \qquad (1.27)$$

Where the five coefficients C_1^n, C_2^n, C_3^n, C_4^n and C_5^n have to be suitably calibrated to the market and they will be described below. The idea of SVI is that each time slice of the implied volatility surface is fitted separately and independently and then some interpolation is performed to obtain the time dependence. Thus, denoting by $\{T_n\}_{1,N}$ the vector of the N available market maturities, we calibrate N different SVI functions $\sigma_I(x,T_n)$ for $n = 1$ to N and then we define for $T \in (T_n, T_{n+1})$:

$$\begin{aligned} \sigma_I^{SVI}(x,T) &= \sigma_I^{SVI}(x,T_n) + \frac{T-T_n}{T_{n+1}-T_n}\left[\sigma_I^{SVI}(x,T_{n+1}) - \sigma_I^{SVI}(x,T_n)\right] \\ &= \frac{T_{n+1}-T}{T_{n+1}-T_n}\,\sigma_I^{SVI}(x,T_n) + \frac{T-T_n}{T_{n+1}-T_n}\,\sigma_I^{SVI}(x,T_{n+1}). \qquad (1.28) \end{aligned}$$

The original model assumes the variable x to be the logarithmic forward moneyness $x = \log(\frac{K}{F_T})$. However, coherently with Chapter 3 in this work we choose to use the logarithmic moneyness as variable, thus from now $x = \log(\frac{K}{S_0})$.

This model has several appealing properties. First of all it depends on only five parameters and can be easily and quickly fitted market volatilities for every maturity. Moreover, it is able to correctly reproduce the typical smile structure and, accordingly to what has been proved by Lee [48], it shows a correct asymptotic behaviour: $\sigma_I(x,T) \to \sqrt{x}$ for $x \to \pm\infty$.

The model assumes five parameters [31] for each maturity:

- $C_1^n \geq 0$: this parameter provides the constant component of volatility. Increasing C_1^n implies a vertical translation of the smile;

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

The calibration of the coefficients C_1^n, C_2^n, C_3^n, C_4^n, C_5^n to market data is accomplished in a similar way to the one described for the Heston model 1.1.2. The idea is to find those coefficients such that the implied volatilities of the model (1.27) are as closest as possible to the ones actually quoted on the market. Thus we need to solve a least square constrained, minimization problem very similar to the one which has been introduced for the calibration of the Heston model (1.7). The only, but relevant, difference is that now the parameters we want to calibrate are time-dependent whereas the Heston parameters were not. Thus we need to perform the same calibration procedure for each one of the available maturities T_n.

Let's call $\mathbf{x}^n = [C_1^n, C_2^n, C_3^n), C_4^n, C_5^n]$ the vector of the coefficients to be calibrated for each market maturity T_n and let $W \subset \mathbb{R}^5$ be the feasible region for these coefficients. Thus, we want to find $\mathbf{x}^n$ for every maturity T_n. The idea is to calibrate the parameters $\mathbf{x}^n$, starting from the closest maturities till the farthest ones, solving a sequence of minimization problems:

$$\begin{aligned} &\text{for every } n = 1, \cdots, N \quad \text{find } \mathbf{x}^n \in W \text{ such that:} \\ &\mathcal{J}_n(\mathbf{x}^n) = \min_{\mathbf{y} \in W} \mathcal{J}_n(\mathbf{y})\,. \end{aligned} \tag{1.29}$$

We define the cost functional:

$$\mathcal{J}_n(\mathbf{y}) = \frac{1}{2} \sum_j^M \omega_j \left(\frac{\sigma_I^{SVI}(K_j, T_n; \mathbf{y}) - \sigma_I^{Market}(K_j, T_n)}{\sigma_I^{Market}(K_j, T_n)} \right)^2 , \tag{1.30}$$

where:

$\{K_j\}_{1,M}$ is the vector of market strike prices;

$\{T_n\}_{1,N}$ is the vector of market maturities;

$\sigma_I^{Market}(K_j, T_n)$ is the implied volatility observed on the market for a contract with strike price K_j and maturity T_n;

$\sigma_I^{SVI}(K_j, T_n; \mathbf{y})$ 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 $\mathbf{y}$;

ω_j are weights chosen to fit better the central region of the smile;

1.2.5 Reconstructing the Local Volatility Surface

Once the model is calibrated for each maturity we can finally reconstruct the local volatility surface. For this purpose we notice that (1.26) requires the values of σ_I, $\frac{\partial \sigma_I}{\partial K}$, $\frac{\partial^2 \sigma_I}{\partial K^2}$ and $\frac{\partial \sigma_I}{\partial T}$. Applying the chain rule we have that:

$$\begin{aligned}
\frac{\partial \sigma_I}{\partial K} &= \frac{\partial x}{\partial K}\frac{\partial \sigma_I}{\partial x} = \frac{1}{K}\frac{\partial \sigma_I}{\partial x}\,, \\
\frac{\partial^2 \sigma_I}{\partial K^2} &= \frac{\partial}{\partial K}\left(\frac{1}{K}\frac{\partial \sigma_I}{\partial x}\right) = \frac{1}{K^2}\frac{\partial^2 \sigma_I}{\partial x^2} - \frac{1}{K^2}\frac{\partial \sigma_I}{\partial x}\,.
\end{aligned}$$

Then, we can finally obtain a closed formula for the Local Volatility Surface. Assuming a null dividend yield $d = 0$, for every $x \in \mathbb{R}$ and for $T \in [T_n, T_{n+1}]$ we have that:

$$\sigma_{LV}(x,T) = \sqrt{\frac{\sigma_I^2 + 2\,T\,\sigma_I\left(\frac{\partial \sigma_I}{\partial T} + r\frac{\partial \sigma_I}{\partial x}\right)}{\left(1 + d_1\sqrt{T}\,\frac{\partial \sigma_I}{\partial x}\right)^2 + \sigma_I\,T\left(\frac{\partial^2 \sigma_I}{\partial x^2} - \frac{\partial \sigma_I}{\partial x} - d_1\sqrt{T}\left(\frac{\partial \sigma_I}{\partial x}\right)^2\right)}}\,. \quad (1.31)$$

Where:

$$\begin{aligned}
\sigma_I &= \sigma_I^{SVI}(x,T) = \tau_n\,\sigma_I^{SVI}(x,T_n) + \tau_{n+1}\,\sigma_I^{SVI}(x,T_{n+1})\,, \\
\frac{\partial \sigma_I}{\partial T} &= \frac{1}{T_{n+1}-T_n}\left[\sigma_I^{SVI}(x,T_{n+1}) - \sigma_I^{SVI}(x,T_n)\right], \\
\frac{\partial \sigma_I}{\partial x} &= \tau_n\,f_n(x) + \tau_{n+1}\,f_{n+1}(x)\,, \\
\frac{\partial^2 \sigma_I}{\partial x^2} &= \tau_n\,g_n(x) + \tau_{n+1}\,g_{n+1}(x)\,, \\
d_1 &= \frac{(r + \frac{1}{2}\,\sigma_I^2)T - x}{\sigma_I\sqrt{T}}\,.
\end{aligned}$$

and

$$\begin{aligned}
f_n(x) &= \frac{C_2^n}{2\,\sigma_I^{SVI}(x,T_n)}\left[C_3^n + \frac{(x - C_4^n)}{\sqrt{(x-C_4^n)^2 + C_5^n}}\right], \\
g_n(x) &= \frac{1}{\sigma_I^{SVI}(x,T_n)}\left[\frac{C_2^n\,C_5^n}{2\,\left[(x-C_4^n)^2 + C_5^n\right]^{3/2}} - f_n^2(x)\right], \\
\tau_n &= \frac{T_{n+1} - T}{T_{n+1} - T_n}\,, \\
\tau_{n+1} &= \frac{T - T_n}{T_{n+1} - T_n}\,.
\end{aligned}$$

Although Formula (1.31) is less sensitive than (1.25), it could still happen that the numerator or, most likely, the denominator becomes negative. This

could happen because of numerical instabilities in the computation and because the parametrization (1.27) is not arbitrage free [31]. For this reason, during the numerical implementation of (1.31) it can be useful to add a small positive constant to both numerator and denominator, avoiding in this way to obtain a negative variance and then a complex local volatility. Thus, defining

$$\begin{aligned} N(x,T) &= \sigma_I^2 + 2\,T\,\sigma_I \left(\frac{\partial \sigma_I}{\partial T} + r \frac{\partial \sigma_I}{\partial x} \right), \\ D(x,T) &= \left(1 + d_1 \sqrt{T}\, \frac{\partial \sigma_I}{\partial x} \right)^2 + \sigma_I\, T \left(\frac{\partial^2 \sigma_I}{\partial x^2} - \frac{\partial \sigma_I}{\partial x} - d_1 \sqrt{T} \left(\frac{\partial \sigma_I}{\partial x} \right)^2 \right), \end{aligned}$$

we can consider

$$\sigma_{LV}^2(x,T) = \frac{N(x,T) + \epsilon_1}{D(x,T) + \epsilon_2} \tag{1.32}$$

In order to avoid negative variance, we can also consider a modified version of the cost functional (1.30) introducing a barrier term, like the one suggested for the calibration of the Heston model (1.2). If we set

$$\overline{N}_n = \min_{x \in \Omega} N(x, T_n), \tag{1.33}$$

$$\overline{D}_n = \min_{x \in \Omega} D(x, T_n), \tag{1.34}$$

we can use a modified cost functional:

$$\tilde{\mathcal{J}}_n = \mathcal{J}_n - \omega \big(\log(\overline{N}_n) + \log(\overline{D}_n) \big), \tag{1.35}$$

where ω is an appropriate tuning parameter. In this way we can impose directly in the fitting of the SVI function to market a barrier for negative values of the local variance $\sigma_{LV}^2(x,T)$.

Concluding this section about Local Volatility, it is worth to notice that, in the same way as the Heston model is just one example of Stochastic Volatility model, although the most known, we have here presented in details just a particular instance of Local Volatility model. In fact, we choose to use the SVI parametrization (1.27) to fit market implied volatilities and then the Dupire formula to reconstruct σ_{LV}, however several other possibilities are available in literature. Zühlsdorff [69], for instance, suggested to parametrize directly the Local Volatility surface with an hyperbola, while Coleman [13] suggested the use of splines. Other authors [8, 47] proposed a non parametric approach based on the resolution of an inverse problem for the Dupire equation with a Tykhonov regularization technique. We choose to use the Dupire formula combined with the SVI parametrization since this is a direct, fast, stable and reliable method to reconstruct the Local Volatility surface from marked data.

1.3 Comparing Local and Stochastic Volatility

We conclude this chapter about Local and Stochastic Volatility providing a comparison between these two classes of models. We first revise a different approach to the derivation of Local Volatility, where the Local Volatility surface σ_{LV} turns out to be the conditional expected value of the variance of a Stochastic volatility model. Then we compare the advantages and the disadvantages of the two models justifying the introduction of a generalized class of model.

1.3.1 Local Volatility as conditional expectation

A different approach to Local Volatility was derived independently by Dupire [22] and Derman and Kani [18]. Instead of considering Local Volatility as a different and alternative model to Stochastic Volatility it is possible to show that the first kind of model is actually a particular case of the second one. In fact if we suppose that the underlying asset follows a diffusion process with a stochastic instantaneous variance, then we can think Local Volatility as the conditional expectation of this instantaneous volatility. Let's state this idea more formally following the proof given by Gatheral [30].

Proposition 1.9 (Local Volatility as a Conditional Expectation)**.** *Suppose the asset price, under the risk-neutral equivalent martingale measure, is driven by a stochastic volatility model dynamics:*

$$dS_t = (r-d)S_t dt + \nu_t S_t dW_t \,. \tag{1.36}$$

Where ν_t is itself a stochastic process Then in order to price a contingent claim it is equivalent to use a deterministic local volatility, conditional expectation of the stochastic instantaneous one:

$$\sigma^2_{LV}(K,T) = \mathbb{E}[\nu^2_T | S_T = K] \,. \tag{1.37}$$

Proof. We know by definition that

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

. Let's differentiate once with respect to K. Assuming the usual integrability assumptions in order to interchange the expectation and the derivative operator and assuming also the derivative defined in distributional sense, we get:

$$\frac{\partial C(K,T)}{\partial K} = -e^{-r(T-t)}\mathbb{E}^Q[\mathbf{1}_K(S_T) \,|\, \mathcal{F}_t] \,. \tag{1.38}$$

And differentiating again we obtain:

$$\frac{\partial^2 C(K,T)}{\partial K^2} = e^{-r(T-t)}\mathbb{E}^Q[\delta_K(S_T) \,|\, \mathcal{F}_t] \,. \tag{1.39}$$

Where $\delta_{x_0}(\cdot)$ is the Dirac delta distribution centered in x_0 and $\mathbf{1}_{x_0}(\cdot)$ is the Heaviside function. Note that from the above equation, and recalling (1.22) we have the following meaningful result:

$$p(K,T;S_t,t) = \mathbb{E}^Q[\delta_K(S_T)\,|\,\mathcal{F}_t]\,. \tag{1.40}$$

The probability density function of the underlying asset at maturity being equal to K is the expected value of the Dirac-delta distribution. Now let's take the derivative with respect to T:

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

To evaluate the right-hand side we first interchange the expectation and the derivative operator (assuming this operation admissible) and then we apply the *Tanaka-Meyer formula* a generalization of Itô formula:

$$d(S_T-K)^+ \;=\; \mathbf{1}_K(S_T)dS_T + \frac{1}{2}S_T^2\,\nu^2(S_T,T)\,\delta_K(S_T)dT\,.$$

Using the stochastic dynamics assumed for S_T (1.36), taking expectation and using the fact that $\mathbb{E}[dW_t]=0$ we obtain:

$$d\mathbb{E}^Q[(S_T-K)^+\mid\mathcal{F}_t] = (r-d)\mathbb{E}^Q[S_T\mathbf{1}_K(S_T)\mid\mathcal{F}_t]dT + \mathbb{E}^Q[\frac{1}{2}S_T^2\nu_T^2\,\delta_K(S_T)\mid\mathcal{F}_t]dT\,.$$

Now using (1.38) the first term can be rewritten as:

$$\begin{aligned}\mathbb{E}^Q[S_T\mathbf{1}_K(S_T)\mid\mathcal{F}_t] &= \mathbb{E}^Q[(S_T-K)\mathbf{1}_K(S_T)\mid\mathcal{F}_t] + \mathbb{E}^Q[K\,\mathbf{1}_K(S_T)\mid\mathcal{F}_t]\\ &= e^{r(T-t)}C + e^{r(T-t)}K\frac{\partial C_t}{\partial K}\,.\end{aligned}$$

Concerning the second term, using (1.39), it can be rearranged as:

$$\begin{aligned}\mathbb{E}^Q[\frac{1}{2}S_T^2\nu_T^2\,\delta_K(S_T)\mid\mathcal{F}_t] &= \frac{1}{2}K^2\mathbb{E}^Q[\nu_T^2\mid S_T=K\,]\,\mathbb{E}^Q[\delta_K(S_T)\mid\mathcal{F}_t]\\ &= \frac{1}{2}K^2\mathbb{E}^Q[\nu_T^2\mid S_T=K\,]\,e^{r(T-t)}\frac{\partial^2 C}{\partial K^2}\,.\end{aligned}$$

Combining these expression we find:

$$\begin{aligned}\frac{\partial C}{\partial T} &= -rC + (r-d)C + (r-d)\,K\,\frac{\partial C}{\partial K} + \frac{1}{2}\,K^2\,\mathbb{E}^Q[\nu_T^2\mid S_T=K\,]\,\frac{\partial^2 C}{\partial K^2}\\ &= \frac{1}{2}\,K^2\,\mathbb{E}^Q[\nu_T^2\mid S_T=K\,]\,\frac{\partial^2 C}{\partial K^2} + (r-d)\,K\,\frac{\partial C}{\partial K} - dC\end{aligned}$$

Finally comparing this equation with the Dupire's equation that defines and characterizes the local volatility function it is clear that:

$$\sigma^2(K,T) = \mathbb{E}[\nu_T^2|S_T=K\,|\,\mathcal{F}_t]\,.$$

□

Thanks to the above result, it is interesting to notice how the Local and Stochastic volatility models are strictly related. In fact, Local Volatility can be interpreted more as an average over instantaneous volatilities rather than as a different and independent model. As noted by Gatheral [30] it is likely that Dupire, Derman and Kani introduced Local Volatility thinking that it was not a separate class of models. Rather, the idea was to make a simplifying assumption that would allow to price exotic options in a simpler way and still consistently with the known prices of market vanilla options.

1.3.2 The Gyöngy theorem

The Proposition 1.9 is actually a particular case of the Gyöngy theorem that will be the cornerstone in building a market-consistent Local-Stochastic Volatility model. For this reason we briefly recall here the main results of the work of Gyöngy (1986) [36] about mimicking processes. The idea is that given a SDE with stochastic drift and diffusion coefficients it is possible to construct a mimicking process. This is another SDE with deterministic coefficients such that the solutions of the two equations have the same marginal probability distribution. Before stating the Gyöngy theorem let's first recall the definition of Weak solution for a stochastic differential equation [5].

Definition 1.10 (Weak Solution)**.** The stochastic differential equation

$$\begin{aligned} dX_t &= \mu(X_t,t)dt + \sigma(X_t,t)dW_t\,, \\ X_0 &= x_0\,. \end{aligned}$$

is said to have a weak solution X_t if there exists a a probabilistic space $(\Omega, \mathcal{F}_t, P)$ and a Wiener process $(W_t, \mathcal{F}_t)$ on it such that X_t is an $\mathcal{F}_t$ - adapted stochastic process which satisfies the equation:

$$X_t = x_0 + \int_0^t \mu(X_t,t)dt + \int_0^t \sigma(X_t,t)dW_t\,.$$

Now we can state the main result of Gyöngy's work about mimicking processes.

Theorem 1.11 (Gyöngy 1986)**.** *Let ξ_t be a stochastic process satisfying the following SDE:*

$$\begin{aligned} d\xi_t &= \beta(\omega,t)dt + \delta(\omega,t)dW_t\,, \\ \xi_0 &= x_0\,, \end{aligned}$$

where $(W_t, \mathcal{F}_t)$ is a Wiener process, β and δ are bounded $\mathcal{F}_t$ non anticipative processes such that $\delta\delta^T$ is uniformly positive definite. Then it exists a SDE:

$$dX_t = b(X_t,t)dt + d(X_t,t)dW_t\,,$$

with non-random coefficients given by:

$$\begin{aligned} b(x,t) &= \mathbb{E}[\beta(t)|\xi(t)=x]\,, \\ d^2(x,t) &= \mathbb{E}[\delta\delta^T(t)|\xi(t)=x]\,, \end{aligned}$$

which admits a weak solution X_t having the same one-dimensional probability distribution as ξ_t, $\forall t$.

Equation (1.37) is actually a straightforward application of Theorem 1.11. In fact assuming a stochastic volatility model for the underlying asset:

$$\begin{aligned} dS_t &= (r-d)\,S_t\,dt + b(V_t)S_t dW_t\,, \\ dV_t &= a(V_t,t)dt + c(V_t,t)dZ_t\,, \\ dW_t\,dZ_t &= \rho\,dt\,, \end{aligned}$$

it is possible to obtain the same results for derivative pricing using a local volatility model

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

where the local volatility is defined according to Gyöngy theorem:

$$\sigma_{LV}(s,t) = \mathbb{E}[b^2(V_t)|S_t=s]\,,$$

which is exactly the thesis of Proposition 1.9.

1.3.3 Why do we need a generalized model

We now conclude this chapter, where the two most common classes of equity derivatives models have been introduced comparing their advantages and disadvantages. For this purpose we follow the arguments provided by Fengler [27]

The framework of Local Volatility is very appealing for several reasons. First of all it is a straightforward and natural extension of the classical Black-Scholes model. It assumes only one source of randomness and then it is complete market model. Once we have σ_{LV}, we simply need to solve a standard parabolic equation to price derivatives. Several fast and robust methods are available for this task. Moreover, a Local Volatility model by construction has the very appealing property to fit and reconstruct the market prices. In this way it is possible to completely reproduce the quoted plain vanilla options and consistently price, with the same model, exotic options. This is a very interesting feature since several kinds of exotic derivatives can be replicated with a static hedge [12, 19]. However, in practice it is more common to dynamically hedge. This procedure is based on the computation of the greeks [7, 67] namely the sensitivity of option price and here come the problems with Local Volatility begins since Hagan and

other authors [38] proved that the delta of an option computed with a Local Volatility model is wrong.

Their argumentation is based on the asymptotic relation between Local Volatility $\sigma_{LV}(s)$ and Implied Volatility $\sigma_I(s)$. For simplicity we do not assume a time dependence of the two functions. In particular they showed [37] that given the present stock value S_0 the current Implied Volatility function $\sigma_I(S_0, K)$ is related, at the first order, to the Local Volatility function by the following expression:

$$\sigma_I(S_0, K) \approx \sigma_{LV}\left(\frac{1}{2}(S_0 + K)\right) \tag{1.41}$$

Let's now suppose that the current spot value S_0 changes by ΔS to $S1 = S_0 + \Delta S$. The Local Volatility function σ_{LV} remains the same, but now it is evaluated at the new spot level S_1. Applying the relation (1.41) we should have the following Implied Volatility:

$$\sigma_I(S_1, K) \approx \sigma_{LV}\left(\frac{1}{2}(S_0 + \Delta S + K)\right) \approx \sigma_I(S_0, K + \Delta S)$$

Therefore, according to the model, if the spot increases, the smile shifts on the left, and vice versa. Unfortunately this is exactly the opposite of what is observed on the market. Thus the Implied Volatility dynamics of the LVM is wrong and this negatively affects the Delta of the model. In fact, let's define C^{LVM} and C^{BS} respectively the option price evaluated with the Local Volatility and the Black-Scholes model. We have the following relation:

$$C^{LVM} = C^{BS}(S_0, K, \sigma_I(S_0, K))$$

Differentiating we obtain that the Delta of the LVM is given by the following relation:

$$\Delta^{LVM} = \frac{\partial C^{LVM}}{\partial S} = \frac{\partial C^{BS}}{\partial S} + \frac{\partial C^{BS}}{\partial \sigma_I}\frac{\partial \sigma_I}{\partial S}$$

Since the LVM predicts a wrong dynamics of Implied Volatility the term $\frac{\partial \sigma_I}{\partial S}$ has the opposite sign to the one observed on the market For this reason the Delta of the Local Volatility model turns out to be wrong.

Another issue negatively affecting Local Volatility Models concerns the forward implied volatility. In fact, as observed by Fengler [27] since the market Implied Volatility surface becomes flat for long maturities, so does the Local Volatility surface. For this reason, the LVM implicitly assumes a flat forward skew, namely flat future smiles, which is exactly the opposite of what we expect. Thus, exotic options with a forward starting feature, like Forward-Start options and Cliquets will be priced incorrectly.

For these reasons, if on the one hand Local Volatility models can reprice the market plain vanilla and produce an exact static of Implied Volatility, on the other hand they imply a wrong dynamics of Implied Volatility. This is exactly the opposite of Stochastic Volatility models. In fact, this class of models cannot fit exactly the current prices of vanillas but it exhibits a realistic dynamics of Implied Volatility [46]. Thus, it would be very interesting to consider a model which combines the positive complementary features of LVM and SVM. The next chapter is devoted to the construction of such a model: a generalized Local-Stochastic Volatility model.

Chapter 2

Local Stochastic Volatility models

In this chapter we revise the work of Ren, Madan and Qian [57] about the construction of a consistent Local-Stochastic Volatility Model. Generally speaking a Local-Stochastic Volatility Model (LSVM) couples a Local Volatility Model (LVM) with a Stochastic Volatility Model (SVM). Therefore, depending on the particular choice of the Local Volatility surface and Stochastic Volatility dynamics (e.g. Heston, SABR, Hull-White) different LSV models could be obtained. We now state the model in its whole generality, assuming a generic Stochastic Volatility process and Local Volatility surface σ_{LV}, while in the next chapter we will particularize the model applying the Heston dynamics and the Dupire formula for σ_{LV}. In this chapter we first present how to set a market consistent model and then we present two possible procedures for the calibration. The first is a stochastic method, based on Markovian projections and the Gyöngy theorem 1.11. The second method is a deterministic one and it relies on the solution of the Fokker-Planck equation (1.18) for the probability density function of the model dynamics. Finally we present some further financial applications closely related to our model.

2.1 Model Calibration

We now present the Local Stochastic Volatility model in its whole generality. For this, and assuming the usual notations, let's first recall the dynamics of the two basic models under the risk neutral martingale measure:

Definition 2.1 (Local Volatility Model - LVM).

$$dS_t = (r-d)S_t dt + \sigma_{LV}(S_t,t)S_t dW_t\,, \tag{2.1}$$

Definition 2.2 (Stochastic Volatility Model - SVM).

$$\begin{aligned} dS_t &= (r-d)S_t dt + b(V_t)S_t dW_t\,, \\ dV_t &= a(V_t,t)dt + c(V_t,t)dZ_t\,, \\ dW_t\, dZ_t &= \rho\, dt\,. \end{aligned} \tag{2.2}$$

The main advantage of the LVM is its capability of a (nearly) perfect fit of the market quoted vanillas. If a good calibration of σ_{LV} is performed the model can reproduce the market prices. Unfortunately this model has a major drawback, its implied volatility dynamics is inconsistent with the observed one. On the other hand a Stochastic Volatility model is able to reproduce a consistent dynamics but not to fit exactly the market prices. For this reason we look for a model with a stochastic volatility dynamics which can, at the same time, well replicate the market prices. The intuition is to model the diffusion coefficient of the underlying asset process S_t as the product between a stochastic component $b(V_t)$ and a deterministic one $\sigma_{LSV}(s,t)$. Thus, this generalized Local-Stochastic Volatility Model is described by the following dynamics:

Definition 2.3 (Local Stochastic Volatility Model - LSVM)**.**

$$\begin{aligned} dS_t &= (r-d)S_t dt + \sigma_{LSV}(S_t,t)b(V_t)S_t dW_t\,, \\ dV_t &= a(V_t,t)dt + c(V_t,t)dZ_t\,, \\ dW_t\, dZ_t &= \rho\, dt\,, \end{aligned} \tag{2.3}$$

Remark 2.4 (Stochastic Drift). Recently hybrid local stochastic volatility model have been proposed (see for instance [5] and [35]) where a stochastic drift is considered as well. However we will use deterministic interest rates and dividend yield. In particular, without loss of generality, we will consider $d = 0$.

The key idea of the Local Stochastic Volatility Model is to merge in a consistent way the LVM and the SVM. In particular the two original models can be calibrated independently and simultaneously to market data to obtain their own parameters σ_{LV}, a, b, c. Afterwards it is possible to evaluate the Local Stochastic Volatility surface $\sigma_{LSV}(S,t)$ using these parameters. For this reason, the first thing we need is to find these coefficients with a calibration procedure like the ones described in Section 1.1.2 and 1.2.4. Once the local volatility surface σ_{LV} and the stochastic volatility model parameters a, b, c, ρ have been calibrated it is possible to find the local-stochastic volatility surface σ_{LSV} thanks to the next important result.

Proposition 2.5. *(Local-Stochastic Volatility Surface equation)*
There exists only one LSV surface $\sigma_{LSV}(s,t)$ such that the LSVM reproduces the quoted vanilla options of the LVM. It satisfies the following equation:

$$\sigma^2_{LV}(s,t) = \sigma^2_{LSV}(s,t)\mathbb{E}[\, b^2(V_t)|S_t = s]\,. \tag{2.4}$$

We can give two different proofs of this important result. The first is based on the Gyöngy theorem and it has been provided by Ren, Madan and Qian [57]. The second proof relies on the comparison between the Fokker-Planck equations of the LV and LSV processes and it has been proposed by Tachet [65].

Proof 1 - via Gyöngy theorem . The (2.4) is actually a straightforward application of the Gyöngy theorem 1.11 applied to the LSVM dynamics (2.3). In fact we want the Local Volatility process (2.1) owns the same one-dimensional probability distribution of the Local-Stochastic Volatility process (2.3). The Gyöngy theorem ensures this, provided that:

$$\sigma^2_{LV}(s,t) = \mathbb{E}[\,\sigma^2_{LSV}(s,t)\,b^2(V_t)|S_t = s] = \sigma^2_{LSV}(s,t)\,\mathbb{E}[\,b^2(V_t)|S_t = s]\,.$$

□

Proof 2 - via Fokker-Planck equation . It is possible to prove the result without using the Gyöngy theorem but simply considering the *Fokker-Planck* equation for the joint probability density function $p(s,v,t)$:

$$\begin{aligned}\frac{\partial p}{\partial t} &= \frac{1}{2}\frac{\partial^2}{\partial s^2}\Big[\sigma^2_{LSV}(s,t)\,b^2(v)\,s^2 p(s,v,t)\Big] + \frac{\partial^2}{\partial s\,\partial v}\Big[\rho\,\sigma_{LSV}(s,t)b(v)s\,p(s,v,t)\Big] \\ &+ \frac{1}{2}\frac{\partial^2}{\partial v^2}\big[c^2(v,t)p(s,v,t)\big] - \frac{\partial}{\partial s}[r(t)s\,p(s,v,t)] - \frac{\partial}{\partial v}[a(v,t)p(s,v,t)]\,,\end{aligned} \tag{2.5}$$

with the initial condition $p(s,v,0) = \delta(s-s_0)\delta(v-v_0)$. Taking now the marginal distribution of S_t, $q(s,t) = \int_0^\infty p(s,v,t)\,dv$, we get the mono dimensional Fokker-Planck for q:

$$\frac{\partial q}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial s^2}\big[\sigma^2_{LSV}(s,t)s^2\int_0^\infty \big(b^2(v)\,p(s,v,t)\,dv\big)\big] - \frac{\partial}{\partial s}[r(t)s\,q]\,. \tag{2.6}$$

On the other hand the marginal distribution q of the LVM has to satisfy the following mono dimensional Fokker-Planck equation, with the calibrated Dupire's LV surface:

$$\frac{\partial q}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial s^2}\big[\sigma^2_{LV}(s,t)s^2 q\big] - \frac{\partial}{\partial s}[r(t)s\,q]\,. \tag{2.7}$$

From Dupire's work we know that the solution of equation (2.7) is able to reproduce the market option prices. Since we want the LSVM to reproduce these prices as well, equation (2.6) and (2.7) must be the same one. Comparing the diffusion term is straightforward to identify the two volatility as:

$$\sigma^2_{LV}(s,t)\,q(s,t) = \sigma^2_{LSV}(s,t)\int_0^\infty b^2(v)\,p(s,v,t)\,dv\,.$$

That is:

$$\sigma^2_{LV}(s,t) = \sigma^2_{LSV}(s,t)\frac{\int_0^\infty b^2(v)\,p(s,v,t)\,dv}{\int_0^\infty p(s,v,t)dv} = \sigma^2_{LSV}(s,t)\,\mathbb{E}[b^2(V_t)|S_t = s]\,.$$

□

We have just derived the nonlinear partial integro-differential equation for the joint probability density function $p(s, v, t)$:

$$\begin{aligned}\frac{\partial p}{\partial t} =& \frac{1}{2}\frac{\partial^2}{\partial s^2}\Big[\sigma_{LV}^2(s,t)\,\frac{\int_0^\infty p(s,v,t)\,dv}{\int_0^\infty b^2(v)\,p(s,v,t)dv}\,b^2\,s^2 p\Big] + \frac{1}{2}\frac{\partial^2}{\partial v^2}\big[c^2(v,t)p(s,v,t)\big] \\ &+ \frac{\partial^2}{\partial s\,\partial v}\Big[\rho\,\sigma_{LV}(s,t)\sqrt{\frac{\int_0^\infty p(s,v,t)\,dv}{\int_0^\infty b^2(v)\,p(s,v,t)dv}}\,b\,s\,p\Big] \\ &- \frac{\partial}{\partial s}[r(t)s\,p(s,v,t)] - \frac{\partial}{\partial v}[a(v,t)\,p(s,v,t)]\,.\end{aligned} \tag{2.8}$$

Solving this equation is actually one of the two alternatives we have to calibrate $\sigma_{LSV}(S,t)$, and, it is, the one we will use.

Remark 2.6 (McKean SDE and Non Linear Fokker-Planck). Using the statement of the theorem we can rewrite the dynamics of the LSVM as:

$$\begin{aligned}dS_t &= (r_t - d_t)S_t dt + \frac{\sigma_{LS}(S_t,t)\,b(V_t)}{\sqrt{\mathbb{E}[b^2(V_t)|S_t = s]}}\,S_t\,dW_t\,, \\ dV_t &= a(V_t,t)dt + c(V_t,t)\,dZ_t\,.\end{aligned}$$

As pointed out by Guyon and Labordère [35] the presence of a non-linear Fokker-Planck equation is not surprising considering the particular dynamics followed by the underlying. In fact this is a Mc-Kean equation, namely a SDE where the drift or the diffusion term depend not only on the value of the process but also on its probability distribution. This is actually our case, being the diffusion term depending on the conditional expected value of $b^2(V_t)$.

The very interesting thing about a LSVM is that we can calibrate independently, and therefore simultaneously, the LVM and the SVM and after that we just need to evaluate $\sigma_{LSV}(S,t)$ using the Theorem 2.5. In the same way we were able to prove the theorem both exploiting the Gyöngy theorem and using the Fokker-Planck equation, it is possible to evaluate the LSV surface $\sigma_{LSV}(s,t)$ with two different approaches that have been recently proposed: a stochastic inspired method and a deterministic one. Labordère [40] proposed in 2009 a method exploiting the Gyöngy theorem and he further developed it in [35]. At the same time Engelman, Koster and Oeltz [23] and Abergel and Tachet [1], [65] developed the original work of Ren, Madan and Qian [57] who first proposed to solve the nonlinear partial-integro differential equation (2.8). In the next section we revise the stochastic-inspired approach of Labordère whereas in the rest of this work the second method is actually used.

2.2 The Markovian Projection Method

In this section we revise the recent work of Labordère [40], [35]. He has introduced and developed a method for calibrating LSV surface to market smile exploiting the Gyöngy theorem and the so-called Markovian Projections [52]. This method is rather general and could be applied to any multi factor SVMs without suffering from the curse of dimensionality. However, as noted by Engelmann [23], for a particular two factor model, like the one we would like to develop, the Labordère approach could be less efficient.

Using the notation of Labordère [40] let's consider a LSVM for the Forward F_t in the Forward measure $\mathbb{P}^T$:

$$\begin{aligned} dF_t &= F_t\,\sigma_{LSV}(S_t,t)b(V_t)dW_t\,, \\ dV_t &= a(V_t,t)dt + c(V_t,t)dZ_t\,, \\ dW_t\,dZ_t &= \rho\,dt\,. \end{aligned} \tag{2.9}$$

Let's now define a local martingale X_t that represents the SVM dynamics:

$$\begin{aligned} dX_t &= X_t\,b(V_t)\,dW_t\,, \\ dV_t &= a(V_t,t)dt + c(V_t,t)dZ_t\,, \\ dW_t\,dZ_t &= \rho\,dt\,. \end{aligned} \tag{2.10}$$

The main idea of the method is to use the Gyöngy theorem twice to compute the local volatility surface for the LSVM and the SVM:

$$\begin{aligned} \sigma^2_{LV}(f,t) &= \sigma^2_{LSV}(f,t)\,\mathbb{E}[\,b^2(V_t)\,|F_t=s\,]\,, \\ \sigma^2_{SV}(f,t) &= \mathbb{E}[\,b^2(V_t)\,|X_t=s\,]\,. \end{aligned}$$

Now we proceed using the Markovian projection technique suggested by Piterbarg [52]. Taking the ratio of the above equations we get:

$$\sigma^2_{LSV}(f,t) = \frac{\sigma^2_{LV}(f,t)}{\sigma^2_{SV}(x,t)}\sqrt{\frac{\mathbb{E}[\,b^2(V_t)\,|X_t=x\,]}{\mathbb{E}[\,b^2(V_t)\,|F_t=f\,]}}\,. \tag{2.11}$$

Following [52] we suppose to be able to find a smooth, strictly monotonically increasing (and then invertible) map $H(f,t)$ satisfying the relation $X_t = H(F_t,t)$. In this way we have a bijective correspondence between X_t and F_t, therefore $\mathbb{E}[\,b^2(V_t)\,|X_t=x\,] = \mathbb{E}[\,b^2(V_t)\,|H(F_t,t)=H(f,t)\,] = \mathbb{E}[\,b^2(V_t)\,|F_t=f\,]$ and we get simply:

$$\sigma^2_{LSV}(f,t) = \frac{\sigma^2_{LV}(f,t)}{\sigma^2_{SV}(H(f,t),t)}\,. \tag{2.12}$$

If we could find an explicit expression for the mapping function, we would be able to calibrate the Local-Stochastic Volatility surface given σ^2_{LV} and σ^2_{SV}. This expression is provided by the following result.

Proposition 2.7 (Labordère)**.** *The mapping function $H(f,t)$ is given by the following expression:*

$$H(f,t) = \Phi_t^{-1}\Big(\int_{f_0}^{f} \frac{1}{y\,\sigma_{LV}(y,t)}\,dy\Big)\,, \quad \textit{with} \quad \Phi_t(x) = \int_{\Lambda(t)}^{x} \frac{1}{y\,\sigma_{SV}(y,t)}\,dy\,. \tag{2.13}$$

The integration constant $\Lambda(t)$ satisfies the following equation:

$$\frac{\partial}{\partial t}\Big(\ln\Lambda(t) + \int_{f_0}^{f} \frac{1}{y\,\sigma_{LSV}(y,t)}\,dy\Big) + \frac{1}{2}\,b^2\,\Big(1 - \frac{\partial}{\partial f}\big(f\,\sigma_{LSV}(f,t)\big)\Big) = 0\,. \tag{2.14}$$

Proof. Without loss of generality we can suppose $X_0 = F_0$. Applying Itô's Lemma to the mapping function $X_t = H(F_t,t)$ and using (2.9) we obtain the dynamics of X_t:

$$dX_t = F_t\,\frac{\partial H}{\partial f}\,\sigma_{LSV}(F_t,t)\,b(V_t)\,dW_t + \Big(\frac{\partial H}{\partial t} + \frac{1}{2}F_t^2\,b^2(V_t)\frac{\partial^2 H}{\partial f^2}\,\sigma_{LSV}^2(F_t,t)\Big)dt\,.$$

By assumption this dynamics must be equal to (2.10), and since $X_t = H(F_t,t)$ this implies:

$$dX_t = H(F_t,t)b(V_t)dW_t$$

Equating the last two expression for X_t and identifying the drift and the diffusion terms, we obtain two differential equations:

$$f\,\frac{\partial H}{\partial f}\,\sigma_{LSV}(f,t) = H(f,t)\,, \tag{2.15}$$

$$\frac{\partial H}{\partial t} + \frac{1}{2}f^2\,b^2\,\sigma_{LSV}^2(F_t,t)\,\frac{\partial^2 H}{\partial f^2} = 0\,. \tag{2.16}$$

Let's concentrate on the first equation. Using (2.12) we can write it as:

$$f\,\frac{\partial H}{\partial f}\,\sigma_{LV}(f,t) = H(f,t)\,\sigma_{SV}(H(f,t),t)\,. \tag{2.17}$$

This is a first order, linear ODE for the unknown function H and it can be easily integrated by separation of variables. Using an integration constant $\Lambda(t)$, we obtain:

$$\int_{f_0}^{f} \frac{1}{y\,\sigma_{LV}(y,t)}\,dy = \int_{\Lambda(t)}^{x} \frac{1}{H\,\sigma_{SV}(H,t)}\,dH\,. \tag{2.18}$$

Therefore, inverting the above equation, we can express $H(s,t)$ via harmonic averages:

$$H(f,t) = \Phi_t^{-1}\Big(\int_{f_0}^{f} \frac{1}{y\,\sigma_{LV}(y,t)}\,dy\Big)\,, \quad \text{with} \quad \Phi_t(x) = \int_{\Lambda(t)}^{x} \frac{1}{y\,\sigma_{SV}(y,t)}\,dy\,.$$

In order to evaluate the map $H(f,t)$ we need to define the integration constant $\Lambda(t)$. For this reason we consider now equation (2.16) and we work out an explicit expression for $\Lambda(t)$ from it. For this purpose we shall play a bit with $\frac{\partial H}{\partial t}$ and $\frac{\partial^2 H}{\partial f^2}$. From equation (2.15) we know that

$$\frac{\partial H}{\partial f} = \frac{H(f,t)}{f^2\,\sigma_{LSV}(f,t)}\,,$$

and therefore differentiating again respect to f we gain

$$\begin{aligned}\frac{\partial^2 H}{\partial f^2} &= \frac{1}{f\,\sigma_{LSV}}\frac{\partial H}{\partial f} + \frac{H}{\sigma_{LSV}}\frac{\partial}{\partial f}\Big(\frac{1}{f}\Big) + \frac{H}{f\,\sigma_{LSV}}\frac{\partial \sigma_{LSV}}{\partial f}\\ &= \frac{H}{\sigma_{LSV}^2 f^2}\big(1-\sigma_{LSV} - f\,\frac{\partial \sigma_{LSV}}{\partial f}\big)\\ &= \frac{H}{\sigma_{LSV}^2 f^2}\big(1-\frac{\partial}{\partial f}(f\,\sigma_{LSV})\big)\,.\end{aligned}$$

Substituting this expression in (2.16) we obtain

$$\frac{\partial H}{\partial t} + \frac{1}{2}\,b^2\,H\,\big(1-\frac{\partial}{\partial f}(f\,\sigma_{LSV})\big) = 0\,. \tag{2.19}$$

Now we try to evaluate $\frac{\partial H}{\partial t}$. Considering again equation (2.15) and integrating it directly via separation of variables we obtain:

$$\ln\Big(\frac{H(f,t)}{\Lambda(t)}\Big) = \int_{f_0}^{f}\frac{1}{y\,\sigma_{LSV}(y,t)}\,dy\,,$$

that means

$$H(f,t) = e^{\left(\ln\Lambda(t)+\int_{f_0}^{f}\frac{1}{y\,\sigma_{LSV}(y,t)}\,dy\right)}\,.$$

and therefore the temporal derivative is simply

$$\frac{\partial H}{\partial t} = H\frac{\partial}{\partial t}\Big(\ln\Lambda(t) + \int_{f_0}^{f}\frac{1}{y\,\sigma_{LSV}(y,t)}\,dy\Big)$$

Inserting this expression in equation (2.19) we finally find the equation (2.14) for $\Lambda(t)$ and this concludes the proof.

□

The calibration method here described would be exact if equation (2.14) was exactly satisfied. Since it would be computationally demand to solve this equation we can try to choose the parameter $\Lambda(t)$ such that the residue is smallest as possible. If we replace in equation (2.14) f with f_0 and b^2 with its conditional average $\mathbb{E}[b^2|f=f_0] = \sigma_{SV}^2(f,t)$, we obtain the following approximated equation:

$$\frac{\partial}{\partial t}\Big(\ln\Lambda(t)\Big) + \frac{1}{2}\,\sigma_{SV}^2(f_0,t)\,\Big(1-\sigma_{LSV}(f_0,t) - f_0\,\frac{\partial}{\partial f}\sigma_{LSV}(f_0,t)\Big) = 0\,. \tag{2.20}$$

Therefore, we choose $\Lambda(t)$ such that:

$$\Lambda(t) = e^{\left(\ln f_0 - \frac{1}{2}\int_0^t \sigma_{SV}^2(f_0,\tau)\left(1-\sigma_{LSV}(f_0,\tau)-f_0\frac{\partial}{\partial f}\sigma_{LSV}(f_0,\tau)\right)d\tau\right)}. \tag{2.21}$$

We can finally summarize the calibration procedure of the Local-Stochastic Volatility surface $\sigma_{LSV}(s,t)$ with the Markovian projection method as follows:

Definition 2.8 (Markovian projection Calibration procedure)**.**

1. Calibrate the Local Volatility Surface $\sigma_{LV}(s,t)$
2. Calibrate the Stochastic Volatility Model parameters a, b, c and ρ
3. Evaluate the effective variance of the SVM: $\sigma_{SV}^2(f,t) = \mathbb{E}[b^2(V_t)|X_t = f]$
4. Evaluate iteratively (few steps should be needed) $\sigma_{LSV}(s,t)$ with the following procedure:

   i Setting first $\Lambda(t) = f_0$, evaluate the map $H(f,t)$ using equation (2.13)

   $$H(f,t) = \Phi_t^{-1}\Big(\int_{f_0}^{f} \frac{1}{y\,\sigma_{LV}(y,t)}\,dy\Big), \quad \text{with} \quad \Phi_t(x) = \int_{\Lambda(t)}^{x} \frac{1}{y\,\sigma_{SV}(y,t)}\,dy\,.$$

   ii Compute the LSVM volatility surface $\sigma_{LSV}(s,t)$

   $$\sigma_{LSV}(f,t) = \frac{\sigma_{LV}(f,t)}{\sigma_{SV}(H(f,t),t)}$$

   .

   iii Update the value of $\Lambda(t)$ using equation (2.21) and σ_{LSV} just evaluated.

Although it is very general, this calibration procedure has several critical points. First of all, in the second step we need to evaluate the effective local variance of the SVM, namely $\sigma_{SV}^2(f,t) = \mathbb{E}[b^2(V_t)|X_t = f]$. Computing this volatility surface with a Monte Carlo simulation would be extremely computationally demanding. However a semi-closed formula for σ_{SV} has been evaluated by Ewald [24] for the Heston model using an approach based on the Malliavin Calculus. Labordère [40] has further developed this result for a general SVM, providing the following expression:

$$\sigma_{SV}^2(f,t) = \frac{\mathbb{E}\left[b_T^2\,\frac{e^{-\frac{K^2}{2(1-\rho^2)B_T}}}{B_T}\right]}{\mathbb{E}\left[\frac{e^{-\frac{K^2}{2(1-\rho^2)B_T}}}{B_T}\right]}, \tag{2.22}$$

where:

$$B_T = \sqrt{\int_0^T b_s^2 ds}\,, \quad \text{and} \quad K = \log\frac{f}{f_0} + \frac{1}{2}\int_0^T b_s^2\, ds - \rho\int_0^T b_s^2 dW_s\,.$$

Another issue that should be not underestimated is the evaluation of the map $H(f,t)$ at step $4.ii$ through an (expensive) integral inversion $\forall t \in [0,T]$.

2.2.1 Application to the Heston dynamics

Let's consider the particular case of a Heston SVM, where $b(V_t) = \sqrt{V_t}$, $a(V_t,t) = \kappa(V_t - \theta)$ and $c(V_t,t) = \sqrt{V_t}$. Therefore the dynamics is given by the following SDEs:

$$\begin{aligned} dX_t &= X_t\sqrt{V_t}\,dW_t\,, \\ dV_t &= \kappa(V_t-\theta)dt + \eta\sqrt{V_t}dZ_t\,, \\ dW_t\,dZ_t &= \rho\,dt\,. \end{aligned}$$

Then the formula for the valuation of the effective variance reads as:

$$\sigma_{SV}^2(f,t) = \frac{\mathbb{E}\left[V_T\,\frac{e^{-\frac{K^2}{2(1-\rho^2)B_T}}}{B_T}\right]}{\mathbb{E}\left[\frac{e^{-\frac{K^2}{2(1-\rho^2)B_T}}}{B_T}\right]}\,, \tag{2.23}$$

Where:

$$B_T = \sqrt{\int_0^T V_s ds}\,, \quad \text{and} \quad K = \log\frac{f}{f_0} + \frac{1}{2}\int_0^T V_s\, ds - \rho\int_0^T V_s dW_s\,.$$

To evaluate the expected value we need to simulate $V_T, \int_0^T V_s\,ds, \int_0^T V_s dW_s$ The stochastic process V_t is a CIR process which has a non-central chi-square probability distribution $p(v,t)$ [66]:

$$p(v,t) = \frac{1}{C_0}e^{-\frac{1}{2}(z+\mu)}\sum_{i=0}^{\infty}\frac{\left(\frac{\mu}{2}\right)^i z^{(\frac{d}{2}+i-1)}}{i!\,2^{(\frac{d}{2}+1)}\Gamma(\frac{d}{2}+i)}\,,$$

with

$$C_0 = \frac{\eta^2}{4\lambda}(1-e^{\lambda t}), \qquad d = 4\frac{\lambda\theta}{\eta^2}, \qquad \mu = \frac{4\lambda e^{\lambda t}v_0}{\eta^2(1-e^{\lambda t})}, \qquad z = \frac{v}{C_0}\,.$$

An efficient algorithm to simulate the Heston dynamics, and in particular the CIR process, was proposed by Broadie and Kaya [10]. They first observe that simply integrating the CIR process it holds the following equation:

$$\int_0^t \sqrt{V_s} dW_s = \frac{1}{\eta}\left(V_t - V_0 - \kappa\theta t + \kappa \int_0^t V_s ds\right).$$

Thus all that is required to be simulated is the joint distribution of $\left(V_t, \int_0^\infty V_s ds\right)$. They actually proposed an efficient simulation of this process. Nevertheless this approach is still computationally demanding since we need to implement a proper Monte Carlo simulation to evaluate the Heston effective variance $\sigma^2_{LV}(s,t)$. A different approach to tackle this problem has been proposed by Atlan [5] who first proved a closed-formula for the effective variance of a Bessel process and then showed how to reduce a CIR process to a Bessel one. In this way the effective variance could be computed with simply a Laplace-Inverse transformation.

2.3 The PIDE Method

Ren, Madan and Qian [57] first introduced in 2007 the idea of evaluating σ_{LSV} solving the nonlinear PIDE (2.8). This kind of equation is quite new in literature and the first theoretical results about existence, uniqueness and regularity of the solution of the equation have been recently developed by Abergel and Tachet [1], [65]. They have also proposed a Finite Difference Method algorithm to solve the equation [65], while Engelmann, Koster and Oelz [23] used a Finite Volume Method. In this section we revise the contributions of these authors.

Let's first state the problem formally in its whole generality. We suppose to know the initial values of $s(0) = s_0$ and $v(0) = v_0$ and all the parameters of the model: $\sigma_{LV}(s,t), b(v), a(v,t), c(v,t), r(t)$. We define the nonlinear partial integro-differential operator $\mathcal{L}(p, I[p])$ as:

$$\begin{aligned}\mathcal{L}(p, I[p]) = \frac{1}{2}\frac{\partial^2}{\partial s^2}\Big[\sigma^2_{LV}(s,t)\, I[p]\, b^2(v)\, s^2\, p(s,v,t)\Big] + \frac{1}{2}\frac{\partial^2}{\partial v^2}\big[c^2(v,t)\, p(s,v,t)\big] \\ + \frac{\partial^2}{\partial s\, \partial v}\Big[\rho\, \sigma_{LV}(s,t)\sqrt{I[p]}\, b(v)\, c(v,t)\, s\, p(s,v,t)\Big] \\ - \frac{\partial}{\partial s}[r(t) s\, p(s,v,t)] - \frac{\partial}{\partial v}[a(v,t)\, p(s,v,t)]\,.\end{aligned} \tag{2.24}$$

where the integral term $I[p](s,t)$ is defined as:

$$I[p](s,t) = \frac{\int_0^\infty p(s,v,t)\, dv}{\int_0^\infty b^2(v)\, p(s,v,t) dv} = \frac{1}{\mathbb{E}[\, b^2(V_t) | S_t = s]}\,. \tag{2.25}$$

Thus the non linear partial integro-differential boundary value problem we want to solve is the following one:

$$\begin{cases} \frac{\partial p}{\partial t} - \mathcal{L}(p, I[p]) = 0 & \forall (s,v,t) \in \mathbb{R}^2_+ \times (0,T] \\ p(s,0,t) = 0 & \forall (s,t) \in \mathbb{R}_+ \times (0,T] \\ p(0,v,t) = 0 & \forall (v,t) \in \mathbb{R}_+ \times (0,T] \\ p(s,v,0) = \delta(s-s_0)\,\delta(v-v_0) & \forall (s,v) \in \mathbb{R}_+ \end{cases} \tag{2.26}$$

Remark 2.9 (Conservation equation). Of course, since the unknown solution p is a probability density function it will also satisfy a mass balance equation:

$$\int_{\mathbb{R}^+}\int_{\mathbb{R}^+} p(s,v,t)\,ds\,dv = 1\,, \quad \forall t \in (0,T)\,.$$

this is already implicit in the fact that equation (2.26) is a conservative parabolic differential equation. Still, it is particularly important to remember the model interpretation of the equation unknown solution p.

This nonlinear and nonlocal PIDE is a non trivial problem mainly due to the fact that the integral term in inside the derivative, and its numerical solution will be certainly rather demanding. However this kind of approach has a big advantage. Once the PIDE is solved we know the joint probability density function $p(s,v,t)$ and from this we can easily compute a contingent claim with a simple integration. In fact by Theorem 2 we have that the price F of a simple contingent claim with payoff function at maturity ϕ, is given by:

$$F = e^{-rT}\int_{\mathbb{R}^+}\int_{\mathbb{R}^+} \phi(s)\,p(s,v,T)\,ds\,dv \tag{2.27}$$

2.3.1 The Abergel-Tachet theorem

As observed in [1] if the integral term $I(p)$ would be a constant, the equation becomes a classic linear parabolic PDE that we can easily solve. This is the key observation to build up a fixed point method to solve the problem iteratively. Given a first guess $p_0(s,v,t)$, for every $k \geq 1$ we compute $I[p_{k-1}]$, and then we solve the linearised problem:

$$\frac{\partial p_k}{\partial t} = \mathcal{L}(p_k, I[p_{k-1}])\,,$$

until convergence, namely until $\| p_k - p_{k-1} \|_{H^1_0} < \epsilon$. The critical issue is to understand if such a procedure is stable and which is the initial guess for the solution we should choose. As mentioned above some answers have been provided by Abergel and Tachet in [1]. They have proved that the problem is well posed under some reasonable but strict hypothesis: roughly speaking the volatility of volatility $c(v,t)$ should be small enough and some suitable regularized initial and

boundary condition should be used. In this section we revise their theoretical results presented in [1] and [65].

The real challenge of the problems resides in the fact that the integral term is within the derivatives. Actually a similar problem has been studied by Alvarez and Tourin in [4] who proved existence and uniqueness for the solution of a nonlinear PIDE, where the integral term is outside the derivative. This is indeed the added problem of (2.26):

$$\frac{\partial p}{\partial t} = \frac{1}{2}\sigma_{LV}^2 b^2 s^2 I[p]\frac{\partial^2 p}{\partial s^2} + \rho\,\sigma_{LV}\,b\,c\,s\,\sqrt{I[p]}\frac{\partial^2 p}{\partial s\,\partial v} + \frac{1}{2}\,c^2\,\frac{\partial^2 v}{\partial v^2} - r\,s\,\frac{\partial p}{\partial x} - a\,\frac{\partial p}{\partial v}\,. \quad (2.28)$$

Unfortunately, nowadays, there is not the proof of the well-posedness of the original problem (2.26). In order to prove some result of existence we need to make several assumptions and consider a bit simplified problem. Before stating the theorem proved by Abergel and Tachet [65] let's first introduce some Hölder spaces. Let D be a domain in $\mathbb{R}^{d+1}$, let $h \in (0,1)$. Given a function u, and two points $P = (x,t)$, $Q = (x',t') \in D$ we define the following norms:

$$|u|_{H^{0,h,h/2}(D)} = \sup_D |u| + \sup_{P,Q\in D} \frac{|u(P) - u(Q)|}{(|x-x'|^2 + |t-t'|)^{h/2}}\,, \quad (2.29)$$

$$|u|_{H^{1,h,h/2}(D)} = |u|_{H^{0,h,h/2}(D)} + \sum_{i=1}^{d} \Big|\frac{\partial u}{\partial x_i}\Big|_{H^{0,h,h/2}(D)} + |\partial_t u|_{H^{0,h,h/2}(D)}\,, \quad (2.30)$$

$$|u|_{H^{2,h,h/2}(D)} = |u|_{H^{1,h,h/2}(D)} + \sum_{i,j=1}^{d} \Big|\frac{\partial^2 u}{\partial x_i \partial x_j}\Big|_{H^{0,h,h/2}(D)}\,, \quad (2.31)$$

$$|u|_{H^{1,h/2}(0,T)} = \sup_{t\in(0,T)} |u'(t)| + \sup_{t,t'\in(0,T)} \frac{|u'(t) - u'(t')|}{|t-t'|^{h/2}}\,. \quad (2.32)$$

One says that u is uniformly Hölder continuous with exponent h in D if and only if $|u|_{H^{0,h,h/2}(D)} < \infty$. Then we denote by $H^{0,h,h/2}(D)$ the set of all functions u uniformly Hölder. We denote respectively by $H^{1,h,h/2}(D)$, $H^{2,h,h/2}(D)$ and $H^{2,h,h/2}(D)$ the set of all functions u for which the respective norms above defined are bounded. Those sets are Banach spaces. Now we are ready to state the existence theorem of Abergel-Tachet.

Theorem 2.10 (Abergel-Tachet)**.** *We assume that the following hypothesis hold:*

1. (Domain Localization) *The problem is defined on a bounded domain* $[s_{min}, s_{max}] \times [v_{min}, v_{max}]$ *and let* Ω *denote this rectangle where the corner have been smoothed. We also denote* $Q_T = \Omega \times (0,T]$, $S_T = (0,T] \times \partial\Omega$ *and* $B = \{0\} \times \Omega$

2. (Boundary and Initial Condition) *Given an initial condition p_0 let's define the function Ψ on $\overline{B} \cup S_T$ by:*

$$\begin{aligned} \Psi(s,v,0) &= p_0(s,v) \qquad \text{on } \overline{B}\,, \\ \Psi(s,v,t) &= p_0(s,v) \qquad \text{for } (s,v,t) \in S_T\,, \end{aligned}$$

Let's suppose that this boundary and initial condition Ψ belongs to $H^{2,h,h/2}(S_T)$ and it is strictly positive on $S_T \cup \overline{\Omega}$. Moreover it satisfies the following compatibility conditions on $S_T \cup \overline{\Omega}$:

$$\frac{\partial \Psi}{\partial t} = \mathcal{L}\Big(\Psi, \frac{1}{b^2(v_0)}\Big) \quad \text{and} \quad \frac{\partial \Psi}{\partial t} = \mathcal{L}\big(\Psi, I[\Psi]\,\big)$$

3. (Parameters Regularity) *We suppose a certain regularity of the coefficients: $\sigma_{LV} \in C^3(\mathbb{R}_+ \times \mathbb{R}_+)$, $b \in C^3(\mathbb{R})$, $c \in C^3(\mathbb{R}_+ \times \mathbb{R}_+)$, $r(t) \in H^{1,h/2}(0,T)$ and $a \in H^{1,h,h/2}(\mathbb{R})$*

4. (Ellipticity Condition) *We suppose that $\forall t \geq 0$, $\forall (s,v) \in \tilde{\Omega}$, $\forall (\delta_1, \delta_2) \in \mathbb{R}^2$ the following condition of uniform ellipticity holds:*

$$\frac{s^2}{b^2(v_0)}\,\sigma_{LV}^2(s,t)\,b^2(v)\,\delta_1^2 + 2\rho\,\frac{s}{b(v_0)}\,\sigma_{LV}(s,t)\,b(v)\,c(v,t)\,\delta_1\,\delta_2 + c^2(v,t)\,\delta_2^2 \;\geq K\,(\delta_1^2+\delta_2^2),$$

Given all the previous hypothesis, then for any $T > 0$ there exists a constant b^* such that if $|b^2(v) - b^2(v_0)| \leq b^*$ $\forall v \in (v_{min}, v_{max})$ the problem

$$\begin{cases} \frac{\partial p}{\partial t} = \mathcal{L}(p, I[p]) & \text{on } Q_T\,, \\ p = p_0 & \text{on } S_T \cup \overline{\Omega}\,, \end{cases} \tag{2.33}$$

admits a solution on Q_T, this solution belongs to $H^{2,h,h/2}(Q_T)$.

Since it assumes a bounded and smooth domain and a strictly positive boundary and initial condition, at least from a theoretical point of view the theorem is not completely satisfying. Nevertheless for a numerical approximation of the problem we are forced to consider a bounded domain and an approximation of the Dirac delta as initial condition. Therefore in this case the theorem provides the existence of the solution. Unfortunately as we will see in the next chapter the parameter of the Heston's dynamics do not satisfy the ellipticity condition at the boundary of the domain.

2.4 Further applications

The methodology developed for the construction of the Local-Stochastic Volatility Model and based on the Gyöngy theorem can be applied to other financial models, extensions of the Black-Scholes framework. In this section, following the work of Atlan [5], Ren, Madan, Qian [57] and Tachet [65] we consider the pricing of option on baskets, option on foreign stock and option with a stochastic interest rate.

2.4.1 Foreign Stock

Let's consider a standard two-dimensional Local volatility model for the evolution of the foreign price of a stock S_t and the exchange rate X_t. The risk free dynamics [7] read as follows:

$$\begin{aligned} dS_t &= (r_f - d)\, S_t\, dt \,+\, \sigma_S(S_t, t)\, S_t\, dW_t^S\,, \\ dX_t &= (r_d - r_f)\, X_t\, dt \,+\, \sigma_X(X_t, t)\, X_t\, dW_t^X\,, \\ dW_t^S\, dW_t^X &= \rho(S_t, X_t, t)\, dt\,, \end{aligned}$$

where r_f is the foreign interest rate, r_d is the domestic interest rate and d is the dividend yield. σ_S and σ_X are the two local volatility functions and W_t^S, W_t^X are standard Wiener processes with instantaneous correlation ρ. Let $Y_t = S_t\, X_t$ be the domestic price of the foreign stock. It is possible to write the stochastic dynamics of Y_t as:

$$dY_t \,=\, (r_d - d)\, Y_t\, dt \,+\, \sigma_S\, Y_t\, dW_t^S + \sigma_t^X\, Y_t\, dW_t^X\,.$$

Of course this is a two dimensional process but applying the Gyöngy theorem 1.11 there exists only one monodimensional Markov process with the same marginal distribution:

$$dY_t \,=\, (r_d - d)\, Y_t\, dt \,+\, \sigma_Y\, Y_t\, dW_t^Y\,, \tag{2.34}$$

where the new local volatility σ_Y is defined as the following conditional expectation:

$$\sigma_Y^2(y, t) = \mathbb{E}\left[\left(\sigma_S^2(S_t, t) + 2\rho(S_t, X_t, t)\sigma_X(X_t, t)\sigma_S(S_t, t) + \sigma_X^2(X_t, t)\right) \,|\, Y_t = y\right]\,. \tag{2.35}$$

In this way we have reduced the dimensionality of the problem. The corresponding pricing pde is very easy and fast to be solved, for instance with the Finite Difference Method, once σ_Y is known. Of course we have shifted the computational effort to the evaluation of the conditional expectation. However this form could be particularly useful if the local volatility surfaces σ_S and σ_X are known.

2.4.2 Basket Option

A very similar case is the local volatility model for a weighted basket of stocks S_t^i, where each stock has its own volatility σ_i and the Wiener processes W_t^i are correlated with instantaneous correlation $\rho_{ij}(S_t^i, S_t^j, t)$:

$$dS_t^i = (r - q_i)\, S_t^i\, dt + \sigma_i(S_i, t)\, S_i\, dW_t^i\,.$$

Now let's define a new asset as the weighted sum of the considered stocks:

$$S_t = \sum_{i=1}^{N} w_i\, S_t^i\,. \tag{2.36}$$

Thus the basket S_t follows the SDE:

$$dS_t = \left(rS_t - \sum_{i=1}^{N} w_i q_i S_t^i \right) dt + \sum_{i=1}^{N} w_i \sigma_i(S_t^i, t) S_t^i dW_t^i \tag{2.37}$$

Applying the Gyöngy theorem we can obtain a monodimensional process for the basket:

$$dS_t = (rS_t - q(S_t, t))\, S_t dt + \sigma(S_t, t) S_t dW_t \tag{2.38}$$

where:

$$q(K,t) = \frac{1}{K}\, \mathbb{E}\left[\sum_{i=1}^{N} w_i q_i S_t^i \,\middle|\, S_t = K \right]$$

$$\sigma^2(K,t) = \frac{1}{K^2}\, \mathbb{E}\Big[\sum_{i=1}^{N} w_i^2 \sigma_i^2(S_t^i, t) S_t^{i2} \; + \; 2 \sum_{i,j=1, i\neq j}^{N} w_i w_j \rho(S_t^i, S_t^j, t) \sigma_i(S_t^i, t) \sigma_j(S_t^j, t) S_t^i S_t^j \,\Big|\, S_t = K \Big]$$

The drift q and the diffusion σ can be evaluated using either known probability density function or by Monte Carlo and PDE method, just like we have shown for the Local Stochastic volatility. Once the drift and the diffusion coefficients are known the problem dimension is effectively reduced by N to 1. Dimensional reduction can be crucial when dealing with complex exotic derivatives. In this case we would like to have a fast and efficient way for derivative pricing but pde method suffers the "curse of dimension" [62] and becomes not practical with high dimension problem. Thanks to the Gyöngy theorem we obtain a monodimensional problem easy to solve with standard methods.

Chapter 3

The Heston-Dupire model

In the previous chapter we introduced the general theory of Local Stochastic Volatility models. In this chapter we consider a particular instance of the model obtained combining the stochastic volatility dynamics of the Heston model with the formula suggested by Dupire for the local volatility. Thus, we consider a Heston-Dupire Local-Stochastic Volatility Model. The two base models, and their calibration, have been already described in Chapter 1. Now we consider in details the non linear Fokker-Planck equation (2.26), introduced in the previous chapter, for this Heston-Dupire case. First the boundary value problem is introduced with suitable initial and boundary conditions. Then we provide the numerical approximation of the equation with the Finite Element Method.

3.1 The equation for the Heston dynamics

Let's first introduce the Heston-Dupire model dynamics. Under the risk free martingale measure, it reads as follows:

Definition 3.1 (Heston-Dupire Local Stochastic Volatility Model)**.**

$$\begin{aligned} dS_t &= rS_t dt + \sigma_{LSV}(S_t,t)\sqrt{V_t}\, S_t\, dW_t\,, \\ dV_t &= \kappa(\theta - V_t)dt + \eta\sqrt{V_t}\, dZ_t\,, \\ dW_t\, dZ_t &= \rho\, dt\,, \end{aligned} \tag{3.1}$$

with the initial condition $(S_0, V_0) = (s_0, v_0)$. The parameters κ, θ, η, ρ and v_0 are the same ones of the Heston model presented in Section 1.1.1. Concerning the Local Volatility we use the SVI parametrization (1.27) to reconstruct the Implied Volatility σ_I from the market and then we use the Dupire formula (1.31) to obtain σ_{LV}. Thus, we know from Proposition 2.5 that the local-stochastic volatility surface is:

$$\sigma_{LSV}(s,t) = \frac{\sigma_{LV}(s,t)}{\sqrt{\mathbb{E}[V_t|S_t = s]}}\,.$$

We now assume to know both the Heston parameters $\kappa, \theta, \eta, \rho, V_0$ and the Dupire Local Volatility σ_{LV} and we consider the non linear PIDE for the joint probability density function of the Heston-Dupire Local Stochastic Volatility model. The purpose is to evaluate the Local-Stochastic Volatility surface $\sigma_{LSV}(s,t)$ according to the PIDE method described in Section 2.3

It is useful to reformulate the dynamics (3.1) in terms of the variables $(x,y) = (\log(\frac{s}{s_0}), v)$. The log-moneyness change of variable $x = \log(\frac{s}{s_0})$ is particularly useful for several reason. In fact, dividing the asset variable for its initial value we normalize it. Applying the logarithm, the initial condition x_0 is simply $x_0 = 0$, the domain for the X_t process becomes symmetric and unbounded at the left side as well at the right one and moreover as, we will discuss later, we remove the degeneracy of the differential operator along the x axis. Applying the Ito lemma [51] it is straightforward to obtain the stochastic integral differential equations for the process (X_t, Y_t):

$$\begin{aligned} dX_t &= \big(r_t - \frac{1}{2}\, Y_t\, \sigma^2_{LSV}(X_t,t)\big)\, dt + \sigma_{LSV}(X_t,t)\, \sqrt{Y_t}\, dW_t\,, \\ dY_t &= \kappa(\theta - Y_t)dt + \eta\sqrt{Y_t} dZ_t\,, \\ dX_t dY_t &= \rho dt\,, \end{aligned} \tag{3.2}$$

The corresponding Fokker-Planck PIDE for the evolution of the joint-probability density function $p(x,y,t)$ is:

$$\begin{aligned} \frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial x^2}\Big[y\, \sigma^2_{LSV}(x,t)\; p\Big] + \frac{1}{2}\frac{\partial^2}{\partial y^2}\big[\eta^2\, y\, p\big] + \frac{\partial^2}{\partial x\, \partial y}\Big[\rho\, \sigma_{LSV}(x,t)\, \eta\, y\, p\Big] \\ -\frac{\partial}{\partial x}\Big[\big(r - \frac{1}{2}\, y\, \sigma^2_{LSV}\,\big)\, p\Big] - \frac{\partial}{\partial y}[\kappa(\theta - y)\, p]\,, \end{aligned} \tag{3.3}$$

where the Local-Stochastic volatility $\sigma_{LSV}(x,t)$ is defined as:

$$\sigma_{LSV}(x,t) = \sigma_{LV}(x,t)\sqrt{I[p](x,t)}\,,$$

and the integral term is:

$$I[p](x,t) = \frac{\int_0^\infty p(x,y,t)\, dy}{\int_0^\infty y\, p(x,y,t) dy}\,,$$

Finally, the initial condition is given by:

$$p(x,y,0) = \delta(x)\, \delta(y - y_0)\,, \tag{3.4}$$

where $\delta(\cdot)$ is the usual Dirac Delta distribution and $y_0 = V_0$ is the calibrated initial value of the Heston variance. As already noticed in the previous chapter, the spatial domain of equation (3.3) is the infinite half-plane $\Omega = \mathbb{R} \times \mathbb{R}^+$, while

the time domain is $(0, T]$, being T the longest maturity of the contracts we are interested in. However, for a numerical solution, we need to localize the domain to a finite subset of Ω, imposing appropriate boundary conditions. In the same way we need to use a suitable approximation of the initial condition. In the next sections we consider all these issues, providing the approximated continuous problem that will be actually solved numerically.

3.1.1 Domain localization and boundary conditions

The spatial domain of the solution should be $\Omega = \mathbb{R}^+ \times \mathbb{R}$, however if we want to solve the equation numerically we need to consider a bounded domain $\tilde{\Omega} = [x_{min}, x_{max}] \times [0, y_{max}]$. Assuming that the domain is enough wide, suitable boundary conditions are homogeneous Dirichlet. Thus we need to set the dimensions of the domain: x_{min}, x_{max} and y_{max}. In order to reduce the computational effort of the solution of the problem we would like to set these as smallest as possible. However, we can verify the goodness of the chosen domain only after that the numerical simulation has been performed, observing whether the solution misbehaves or not at the boundary. For this reason, the definition of the domain dimension can be performed with a trial and error approach. However, we can get an a priori idea of the dimension of the domain following the arguments below.

In the standard Black-Scholes model it is easy to show that:

$$\begin{aligned} \mathbb{P}(S_T > U) &< 10^{-8} \quad \text{with} \quad U = S_0 e^{(r-\frac{1}{2}\sigma^2)T+6\sigma\sqrt{T}}, \\ \mathbb{P}(S_T < L) &< 10^{-8} \quad \text{with} \quad L = S_0 e^{(r-\frac{1}{2}\sigma^2)T-6\sigma\sqrt{T}}. \end{aligned}$$

Thus the values U and L can be used as appropriate boundaries of the domain for a numerical solution of the Black-Scholes equation. However, in our model there is not a unique value of σ, instead we have the unknown deterministic function $\sigma_{LSV}(S, t)$. A reasonable approach is then to use the following values:

$$\begin{aligned} x_{min} &= (r - \frac{1}{2}\sigma_{min}^2)T + 6\,\sigma_{max}\sqrt{T}, \\ x_{max} &= (r - \frac{1}{2}\sigma_{max}^2)T - 6\,\sigma_{max}\sqrt{T}, \end{aligned}$$

where

$$\begin{aligned} \sigma_{min} &= \min_{x\in\mathbb{R},\, t\in[0,T]} \sigma_{LV}(x, t), \\ \sigma_{max} &= \max_{x\in\mathbb{R},\, t\in[0,T]} \sigma_{LV}(x, t). \end{aligned}$$

Concerning the variable y since it is defined in $(0, \infty)$ we need to set only an upper boundary y_max. To find an appropriate value let's consider the Fokker-Planck equation for the probability distribution of the stochastic variance Y_t:

$$\frac{\partial}{\partial t}p(y,t) - \frac{\partial}{\partial y}\left(\kappa\left(y-\theta\right)p(y,t)\right) - \frac{1}{2}\eta^2\frac{\partial^2}{\partial y^2}\left(y\,p(y,t)\right) = 0\,. \tag{3.5}$$

This equation is well defined on $[0,\infty)$, $\forall\theta > 0$ [25] and its stationary solution is a gamma distribution [20]:

$$p_\infty(y) = \frac{1}{\Gamma(\alpha)}\Big(\frac{\alpha}{\theta}\Big)^\alpha y^{\alpha-1}\,e^{-\frac{\alpha}{\theta}y}\,, \qquad \text{with } \alpha = \frac{2\kappa\theta}{\eta^2}\,. \tag{3.6}$$

Thus the cumulative density function is:

$$P_\infty(y) = \int_0^y p_\infty(u)du = \frac{\gamma\left(\alpha,\frac{\alpha}{\theta}y\right)}{\Gamma(\alpha)}\,, \tag{3.7}$$

where $\Gamma(x)$ is the gamma function and $\gamma(x,s)$ is the incomplete gamma function. In this way we can find a certain value y_{max} such that $P_\infty(y_{max}) \simeq 1$.

Eventually it is possible to prove that for problem of this kind the truncation error decays exponentially with respect to the size of the domain [42]. For this reason imposing null Dirichlet conditions at the boundaries with a wide enough domain would not affect the solution too much. In the next part of this work we shall use the notation Ω instead of $\tilde{\Omega}$ to indicate the localized domain: $\Omega = [x_{min}, x_{max}] \times [0, y_{max}]$.

3.1.2 Initial condition

For the initial condition of the problem we impose a suitable approximation of the Dirac Delta: $p_0(x,y) = \delta(x)\delta(y-y_0)$. For instance a good choice is given by the product between the fundamental solution of the heat equation on the line and the half-line:

$$p_0(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\,e^{-\frac{x^2}{2\sigma_x^2}}\left(e^{-\frac{(y-y_0)^2}{2\sigma_y^2}} - e^{-\frac{(y+y_0)^2}{2\sigma_y^2}}\right). \tag{3.8}$$

In this way the initial condition vanishes at the boundary $y=0$, it is symmetric respect $x=0$, and, of course, it is a probability density function:

$$p_0(x,0) = 0, \quad p_0(x,y) = p_0(-x,y), \quad \text{and} \quad \int_{\mathbb{R}\times\mathbb{R}^+} p_0(x,y)\,dx\,dy = 1\,.$$

In order to make this initial condition a proper approximation of the real one the parameters σ_x and σ_y should be chosen small enough. However it is interesting to notice that this approximated initial condition has a real model interpretation. In fact, we never know with absolute precision the initial value of s_0 and v_0, but always with a certain indetermination that we could identify with the values σ_x and σ_y. For instance given the bid-ask spread we can define $s_0 = \frac{bid+ask}{2}$ and $\sigma_x = \frac{1}{2}(log(\frac{bid}{s_0}) - log(\frac{ask}{s_0}))$. In the same way, when we calibrate the initial variance V_0 of the Heston model, we never find its exact value, thus we can define σ_y as the uncertainty of the calibration.

3.1.3 Weak formulation

In order to approximate the equation with the Finite Element Method we first need to write the weak formulation of the problem. For this purpose it is useful to rewrite equation (3.3) in the conservative form:

$$\frac{\partial p}{\partial t} - \mathcal{L}(p, I[p]) = 0\,, \tag{3.9}$$

where the non linear elliptic operator $\mathcal{L}(\cdot, I[\cdot])$ is defined as follows:

$$\mathcal{L}(p, I[p]) = \operatorname{div}\big(D[p]\,\nabla p + \mathbf{a}[p]\,p\big)\,, \tag{3.10}$$

being D and $\mathbf{a}$ the diffusion matrix and the advection vector, both depending on the non linear integral term $I[p]$:

$$D[p] = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}\,, \qquad\qquad \mathbf{a}[p] = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}\,.$$

After some calculations, it is possible to prove that the diffusion matrix D and advection vector $\mathbf{a}$ are composed by the following terms:

$$D_{11} = \frac{1}{2}\,y\,\sigma_{LV}^2(x,t)\,I[p]\,, \tag{3.11}$$

$$D_{12} = D_{21} = \frac{1}{2}\,y\,\eta\,\rho\,\sigma_{LV}(x,t)\,\sqrt{I[p]}\,, \tag{3.12}$$

$$D_{22} = \frac{1}{2}\,y\,\eta^2\,, \tag{3.13}$$

$$a_1 = -\,r + \frac{1}{2}\,\eta\,\rho\,\sigma_{LV}\,\sqrt{I[p]} + \frac{1}{2}\,y\,\Big(\sigma_{LV}^2\,I[p] + \frac{\partial}{\partial x}\big(\sigma_{LV}^2\,I[p]\,\big)\Big)\,, \tag{3.14}$$

$$a_2 = -\,\kappa\,\theta + \frac{1}{2}\,\eta^2 + y\,\Big(\kappa + \frac{1}{2}\rho\,\eta\,\frac{\partial}{\partial x}\big(\sigma_{LV}\,\sqrt{I[p]}\,\big)\Big)\,. \tag{3.15}$$

Eventually, assuming zero Dirichlet boundary condition, we can state the continuous boundary value problem that we will solve numerically, as follows:

$$\begin{cases} \frac{\partial p}{\partial t} - \mathcal{L}(p, I[p]) &= 0 & \forall (x,y,t) \in \Omega \times (0,T] \\ p(x,0,t) &= 0 & \forall (x,t) \in [x_{min}, x_{max}] \times (0,T] \\ p(x_{min},y,t) &= 0 & \forall (y,t) \in [0, y_{max}] \times (0,T] \\ p(x_{max},y,t) &= 0 & \forall (y,t) \in [0, y_{max}] \times (0,T] \\ p(x,y,0) &= p_0(x,y) & \forall (x,y) \in \Omega \end{cases} \tag{3.16}$$

The weak formulation [59] of the boundary value problem is readily obtained:

$$\forall t \in (0,T] \quad \text{find } p(t) \in V = H_0^1(\Omega) \quad \text{such that} \quad \int_\Omega \frac{\partial p}{\partial t}\, v\, d\Omega + a_p(p,v) = 0, \quad \forall v \in V\,,$$

where $H_0^1(\Omega)$ is the usual Sobolov space of functions in $H^1(\Omega)$ with null trace on $\partial\Omega$ and $a_p(p,v)$ is defined as:

$$\begin{aligned}
a_{\tilde{p},}(p,v) = & -\int_\Omega v\, \mathcal{L}(p, I[\tilde{p}])\, d\Omega \; = \\
& = \int_\Omega D[\tilde{p}]\, \nabla p \cdot \nabla v\, d\Omega + \int_\Omega p\, \mathbf{a}[\tilde{p}] \cdot \nabla v\, d\Omega = \\
& = \int_\Omega \frac{\partial p}{\partial x}\Big(D_{11}[\tilde{p}]\, \frac{\partial v}{\partial x} + D_{21}[\tilde{p}] \frac{\partial v}{\partial y}\Big) d\Omega \\
& + \int_\Omega \frac{\partial p}{\partial y}\Big(D_{21}[\tilde{p}]\, \frac{\partial v}{\partial x} + D_{22}[\tilde{p}] \frac{\partial v}{\partial y}\Big) d\Omega \\
& + \int_\Omega p\, \Big(a_1[\tilde{p}]\, \frac{\partial v}{\partial x} + a_2[\tilde{p}]\, \frac{\partial v}{\partial y}\Big) d\Omega\,.
\end{aligned}$$

Remark 3.2 (Feller condition). It is interesting to notice that in perfect agreement with the probabilistic theory of the Heston model 1.1.1 the Feller number (1.4) actually rules the advection field of the problem (3.16) along the y direction, namely the variance coordinate. In fact, the advection field for the marginal distribution of the variance is given by the sum of the Feller number and a term depending linearly on y:

$$\begin{aligned}
-a_2 \;\; &= \;\; \kappa\,\theta - \frac{1}{2}\,\eta^2 \; - y\, \Big(\kappa + \frac{1}{2}\rho\,\eta\, \frac{\partial}{\partial x}\big(\sigma_{LV}\, \sqrt{I[p]}\,\big)\, \Big) \\
&= \;\; Fe - y\, \Big(\kappa + \frac{1}{2}\rho\,\eta\, \frac{\partial}{\partial x}\big(\sigma_{LV}\, \sqrt{I[p]}\,\big)\, \Big) \;\approx\; Fe \quad \text{for} \quad y \approx 0
\end{aligned}$$

Thus, in the neighbourhood of $y = 0$ the advection field is ruled by the Feller number. If this is negative, the probability density function $p(x,y,t)$ moves towards the boundary $y = 0$ otherwise if the Feller number is positive.

The idea for the numerical discretization of this problem relies on the linearisation of $a_p(p,v)$ and a finite difference approximation of the time derivative. Thus we will eventually solve a sequence of elliptic problems. For this reason it is important to understand whether the matrix D is positive defined [59]. In fact, in this case, the linearised bilinear form $a_{\tilde{p}}(p,v)$ would be coercive. For this purpose we can use the Sylvester criterion [55]. In our case we must verify that:

$$\begin{aligned}
D_{11} &= \frac{1}{2}\, y\, \sigma_{LSV}^2 > 0\,, \\
D_{11}D_{22} - D_{12}D_{21} &= \frac{1}{4}\, y^2\, \eta^2\, \sigma_{LSV}^2\, \big(1 - \rho^2\big) > 0\,.
\end{aligned}$$

Thus the differential operator is truly elliptic for $y > 0$ and $\rho^2 \neq 1$. The second condition is easily fulfilled since we can impose it during the calibration of the Heston parameters. On the other hand the domain boundary lay precisely on the line $y = 0$. This means that we are dealing with a degenerate elliptic operator at the boundary. For this reason, even for the linearised problem the usual analysis based on the Lax-Milgram theorem cannot be applied. In the next section we revise the main results about degenerate elliptic problem with particular interest about the Heston equation [6]. It is worth to notice that if we had used the original s variable instead of the log-moneyness $x = \log(\frac{s}{s_0})$ the elliptic operator would have been degenerate also along this direction. In fact, the same thing happens in the Black-Scholes equation. However thanks to the well known change of variable $x = \log(\frac{s}{s_0})$ we can easily get rid of this degeneracy.

3.2 Elliptic degenerate boundary value problems

The problem we are dealing with is certainly a non trivial one. Not only it is a non linear partial integral differential equation, but moreover the integro-differential operator is degenerate at the boundary of the domain. In the next section we will show in details how to solve the problem numerically. The non linearity and the time evolution will be essentially resolved with a semi-implicit scheme. In this way at each time step we have to solve just a stationary problem. However it is still a degenerate elliptic problem and the simple analysis based on the Lax-Milgram lemma [59] cannot be applied.

In this section, following closely [15], we revise the main results about existence and uniqueness of elliptic problem with degeneracy at the boundary, with particular interest for the Heston operator $\mathcal{L}_H$, namely the operator $\mathcal{L}$ (3.10) where the Local-Stochastic volatility is set to one: $\sigma^2_{LSV} = \sigma^2_{LV}\, I[p] = 1$. Because the operator degenerates at the boundary $y = 0$ we cannot prove the coercivity of the associated bilinear form $a(u, v) = (\mathcal{L}_H u, v)_{L^2(\Omega)}$ in the usual Sobolev space $H^1_0(\Omega)$. For this reason we need to introduce a broader setting, namely a weighted Sobolev space.

Definition 3.3 (Weighted Sobolev Space)**.** Let $\Omega \subset \mathbb{R} \times \mathbb{R}^+$ be a domain and let's define on Ω the following weight function w:

$$w(x, y) := \frac{2}{\eta^2} y^{\beta-1} e^{-\gamma|x|-\mu y}, \quad \text{with} \quad \beta = \frac{2\kappa\theta}{\eta^2}, \quad \mu = \frac{2\kappa}{\eta^2} \quad \text{and} \quad \gamma > 0. \tag{3.17}$$

Let $L^2(\Omega, w)$ be the space of all the measurable functions $u : \Omega \to \mathcal{R}$ such that

$$\|u\|^2_{L^2(\Omega,w)} := \int_\Omega |u|^2 w(x, y)\, dxdy < \infty, \tag{3.18}$$

Then we define the weighted Sobolev space $H^1(\Omega, w)$ as:

$$H^1(\Omega, w) := \left\{ u \in L^2(\Omega, w) : \|u\|_{H^1(\Omega,w)} < \infty \right\}, \tag{3.19}$$

with $\|u\|^2_{H^1(\Omega,w)} := \int_\Omega y(u_x^2 + u_y^2)\, w(x,y)\, dxdy + \int_\Omega (1+y)u^2\, w(x,y)\, dxdy\,.$

and u_x, u_y are defined in the sense of distributions [59].

Of course the spaces $L^2(\Omega, w)$ and $H^1(\Omega, w)$ are Hilbert spaces provided with the inner products:

$$(u,v)_{L^2(\Omega,w)} := \int_\Omega uv\, w\, dxdy\,; \tag{3.20}$$

$$(u,v)_{H^1(\Omega,w)} := \int_\Omega \big(y\,(u_x v_x + u_y v_y) + (1+y)\,uv\big) w\, dxdy\,. \tag{3.21}$$

Let's now define

$$\Gamma_0 = \bar{\Omega} \cap \mathbb{R} \times \{0\},$$

we can consider two other useful spaces, namely:

- $H_0^1(\Omega, w)$: closure of $C_0^\infty(\Omega)$ in $H^1(\Omega, w)$;
- $H_0^1(\Omega \cup \Gamma_0, w) = V_w$: closure of $C_0^\infty(\Omega \cup \Gamma_0)$ in $H^1(\Omega, w)$.

Using the notation $H_w = L^2(\Omega, w)$ we have that the Heston elliptic operator $\mathcal{L}_H$ defines a bilinear form $a : V_w \times V_w$ such as :

$$a_H(u,v) := (\mathcal{L}_H u, v)_{H_w},$$

where $u \in V_w$ and $v \in C_0^\infty(\Omega)$. The bilinear form that we have introduced is continuous, i.e.,

$$|a_H(u,v)| \le C_1 \|u\|_{V_w} \|v\|_{V_w}, \qquad \forall u, v \in V_w,$$

where C_1 is a positive constant which depends on the model parameters $r, \eta, \kappa, \theta, \rho$. Concerning the coercivity, we have the following important result:

Proposition 3.4 (Gårding inequality)**.** *Let $r, \eta, \kappa, \theta \in \mathbb{R}$ be constants such that*

$$\beta := \frac{2\kappa\theta}{\eta^2} > 1, \quad \eta > 0 \quad \textit{and} \quad -1 < \rho < 1.$$

Then for all $u \in V_w$, there exist two positive constants C_1 and C_2, depending on the model parameters, such that

$$a_H(u,u) \ge \frac{1}{2} C_2 \|u\|_{V_w} - C_3 \|(1+y)^{1/2} u\|_{H_w}$$

Finally thanks to the above results, it is possible to prove a theorem of existence and uniqueness for the following Dirichlet boundary problem [15]

$$\begin{cases} \mathcal{L}_H u &= f \quad \text{on } \Omega \\ u &= g \quad \text{on } \Gamma_1 = \partial\Omega \,/\, \Gamma_0 \end{cases} \tag{3.22}$$

Remark 3.5 (Gårding inequality and the Feller condition). It is interesting to notice that the main hypothesis of Proposition 3.4 is precisely the Feller condition (1.4) already introduced for the Heston dynamics.

Theorem 3.6 (Daskalopoulos, Feehan)**.** *Let $f \in L^2(\Omega, w)$ and $g \in H^2(\Omega, w)$. Suppose there are functions $M, m \in H^2(\Omega, w)$ such that:*

$$m \leq g \leq M \text{ on } \Gamma_1, \quad m \leq M \text{ on } \Omega, \quad \text{and } Am \leq f \leq AM \text{ a.e. on } \Omega,$$

and M, m, f, and g obey

$$(1+y)M,\ (1+y)m,\ (1+y)\frac{1}{2\,f} \in L^2(\Omega, w) \quad \text{and} \quad (1+y)\frac{1}{2\,g} \in H^2(\Omega, w),$$

then there exists a solution $u \in H^2(\Omega, w)$ to the boundary problem, such that $(1+y)u \in L^2(\Omega, w)$, and $m \leq u \leq M$ on Ω and there is a positive constant, C, depending only on the constant coefficients of the operator, A and on the constants such that

$$\|u\|_{H^2(\Omega,w)} \leq C \left(\|(1+y)^{1/2} f\| L^2(\Omega, w) + \|(1+y)^{1/2} g\|_{H^2(\Omega,w)} + \|(1+y)\, u\|_{L^2(\Omega,w)} \right).$$

If there is a function $u \in H^2(\Omega, w)$ satisfying the equation then the solution u is unique.

3.3 Numerical approximation with the Finite Element Method

In this section we consider a possible numerical approximation of problem (3.16). Ren, Madan and Qian [57] first suggested in 2007 to linearise the equation exploiting its time evolution, namely performing a semi-implicit time discretization, applying the same strategy often used to solve the Navier-Stokes equations [53]. In 2011 Tachet [65] proposed a Finite Difference scheme to solve the equation while Engelmann, Koster and Oelz [23] used the Finite Volume method. In this section we propose a Finite Element approximation. For a general introduction about the Finite Element Method see [53].

3.3.1 Time discretization

For the time discretization we introduce a discrete temporal grid t^n for $n = 0 \div N_t$ where $t^0 = 0$ and $t^{N_t} = T$. Since the initial condition is a very peaked bidimensional Gaussian distribution approximating the delta of Dirac, we need to use a very short time step at the beginning of the simulation in order to achieve a stable solution. In fact it is of paramount relevance to correctly simulate the initial evolution of the probability. For this reason we will not use an uniform time grid and we shall denote $\Delta t^n = t^n - t^{n-1}$.

Usually when we deal with non linear evolution equation we can approximate the time derivative with a finite difference scheme that could be explicit, semi-implicit, implicit. Defining $p^n(x, y)$ as the semi-discrete approximation of the solution $p(x, y, t)$, namely $p^n(x, y) \simeq p(x, y, t^n)$ we have three possible choices:

$$\begin{aligned}
\text{Explicit:} \quad & \frac{p^{n+1} - p^n}{\Delta t^n} &= \mathcal{L}\big(\, p^n, I[p^n] \,\big)\,, \\
\text{Semi-Implicit:} \quad & \frac{p^{n+1} - p^n}{\Delta t^n} &= \mathcal{L}\big(\, p^{n+1}, I[p^n] \,\big)\,, \\
\text{Fully Implicit:} \quad & \frac{p^{n+1} - p^n}{\Delta t^n} &= \mathcal{L}\big(\, p^{n+1}, I[p^{n+1}] \,\big)\,.
\end{aligned}$$

It is possible to use these schemes since we both know the initial solution p^0 and the initial value of the non linear integral term $I[p^0]$. In fact given the initial value of the process Y_t, $Y_0 = y_0$, its conditional expectation at time zero is trivially $\mathbb{E}[Y_t|X_t = x]\big|_{t=0} = y_0$, and therefore the initial value of the non linear integral term is known as well:

$$I[p^0](x) = I[p](x, 0) = \frac{\int_0^\infty p(x, y, 0)\, dy}{\int_0^\infty y\, p(x, y, 0) dy} = \frac{1}{\mathbb{E}[Y_t|X_t = x]\big|_{t=0}} = \frac{1}{y_0}\,.$$

We briefly recall the main features of these schemes when the spatial discretization is accomplished with the FEM. The key idea is that moving from the explicit to the fully implicit scheme a greater computational effort is required. At the same time a better numerical stability is achieved.

In particular the **Explicit** treatment of the temporal derivative doesn't require the solution of a linear system but involves a stability condition of the type $\Delta t \leq C\, h^2$, where C is a positive constant and h a characteristic length of the spatial discretization. Such condition could be very restrictive to fulfil because it would need a too large number of time steps. The **Semi-Implicit** scheme on the contrary requires to solve a linear system at each time iteration but it is unconditionally stable. Finally the **Fully-Implicit** scheme lead to the solution of a non linear problem which has to be solved with a fixed point method at each time step. This implies the solution of several linear systems. The main advantages of this approach are the improved stability properties of the solution.

Fixed point method for the implicit discretization

Although it is the most computationally demanding method, it seems that the completely implicit scheme is the most suitable for the considered problem. In particular one of the most critical issues of the non linear equation we are dealing with, is that the unknown solution is a probability density function. Thus its integral over the spatial domain has to be equal to one for every time: $\int_\Omega p^n \, d\Omega = 1$, $\forall n$. We have observed that this conservation condition is much better fulfilled using the implicit method.

The implicit time discretization for the problem in the differential formulation, combined with fixed point iterations, is the following one:

$$\frac{p_{k+1}^{n+1} - p^n}{\Delta t^n} - \mathcal{L}\big(p_{k+1}^{n+1}, I[p_k^{n+1}]\big) = 0\,, \tag{3.23}$$

$$\text{where} \quad I[p_k^{n+1}](x) = \frac{\int_0^\infty p_k^{n+1}(x,y)\,dy}{\int_0^\infty y\, p_k^{n+1}(x,y) dy}\,. \tag{3.24}$$

Considering the next spatial discretization with the Finite Element Method, let's rewrite the implicit scheme in the variational formulation.

Definition 3.7 (Implicit semi-discretized Galerkin problem)**.**

Given $p^0 = p_0$ as in (3.8) and $I[p^0] = \frac{1}{y_0}$, $\quad \forall n = 1 \div N_t \quad$ solve the problem:

Given p^n, $\quad$ let be $p_0^{n+1} = p^n$, $\quad \forall k \geq 1 \quad$ until $\quad \frac{\|p_{k+1}^{n+1} - p_k^{n+1}\|_{H^1}}{\|p_{k+1}^{n+1}\|_{H^1}} < \epsilon$,

find $p_{k+1}^{n+1} \in V$ such that:

$$m(p_{k+1}^{n+1}, v) + a_{p_k^{n+1}}(p_{k+1}^{n+1}, v) + d_{p_k^{n+1}}(p_{k+1}^{n+1}, v) = F(v) \quad \forall v \in V\,, \tag{3.25}$$

where:

$$\begin{aligned}
m(p_{k+1}^{n+1}, v) &= \frac{1}{\Delta t^n} \int_\Omega p_{k+1}^{n+1}\, v \, d\Omega\,, \\
a_{p_k^{n+1}}(p_{k+1}^{n+1}, v) &= \int_\Omega p_{k+1}^{n+1} \Big(a_1[p_k^{n+1}] \frac{\partial v}{\partial x} + a_2[p_k^{n+1}] \frac{\partial v}{\partial y} \Big) d\Omega\,, \\
d_{p_k^{n+1}}(p_{k+1}^{n+1}, v) &= \int_\Omega \frac{\partial p_{k+1}^{n+1}}{\partial x} \Big(D_{11}[p_k^{n+1}] \frac{\partial v}{\partial x} + D_{21}[p_k^{n+1}] \frac{\partial v}{\partial y} \Big) d\Omega\,, \\
&\quad + \int_\Omega \frac{\partial p_{k+1}^{n+1}}{\partial y} \Big(D_{21}[p_k^{n+1}] \frac{\partial v}{\partial x} + D_{22}[p_k^{n+1}] \frac{\partial v}{\partial y} \Big) d\Omega\,, \\
F(v) &= m(p^n, v) = \frac{1}{\Delta t^n} \int_\Omega p^n\, v \, d\Omega\,.
\end{aligned}$$

3.3.2 Spatial discretization

Let's consider the spatial discretization of the problem. For this purpose we introduce a discrete approximation Ω_h using an unstructured triangular mesh of the domain $\Omega = (x_{min}, x_{max}) \times (0, y_{max})$. We define V_h as the discrete approximation on Ω_h of the Sobolev space $V = H_0^1(\Omega)$:

$$V_h = \left\{ v_h \in C^0(\bar{\Omega}) : v_h|_K \in \mathbb{P}_1, \quad \forall K \in \Omega_h, \quad v_h|_{\partial\Omega} = 0 \right\} , \tag{3.26}$$

where K is a generic triangle of the mesh Ω_h and $\mathbb{P}_1$ is the space of the polynomials of degree less or equal to one. We choose this kind of finite elements because the initial solution is very peaked and irregular. Moreover considering the degeneracy at the boundary it would be better use a dense mesh instead of a high order polynomial space.

Mesh Construction

Solving the solution of the considered problem with the Finite Element method has one major advantage compared to Finite Differences [65] and Finite Volumes [23], namely the great flexibility of mesh construction. This is particularly important considering the critical issues of the problem, such as degeneracy at the boundary and a very peaked initial condition. Using a not structured mesh allows to put an adequate number of triangle around these sensitive region, namely around the point $(0, y_0)$ and along the boundary y_0.

Since we are dealing with a Fokker-Planck equation a very important property of the solution is the conservation of the total probability: $\int_\Omega p d\Omega = 1 \, \forall t$. Of course we want the numerical solution p_h to hold this property for every time of the simulation. However this is particularly difficult to accomplish in the very first time steps since the initial condition is very peaked. For this reason it is crucial to use a very thin mesh around $(0, y_0)$ in order to numerically integrate the solution correctly. In the same way it is important to have an adequate number of triangles along the boundary y_0 where the differential operator degenerates.

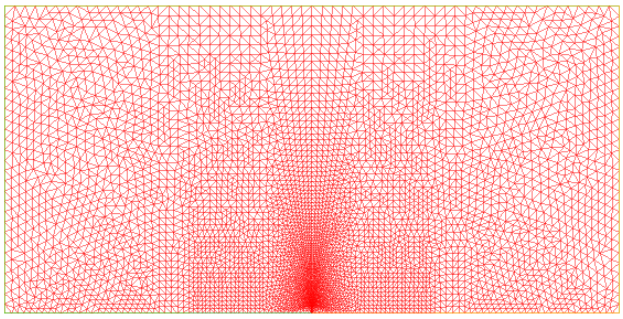


Figure 3.1: Mesh example with 10654 triangles and 5448 vertices

Galerkin problem

The finite element formulation of the Galerkin problem is the following:

Definition 3.8 (Finite element Galerkin problem)**.**

Given $p_h^0 = p_0$ and $I[p_h^0] = \frac{1}{y_0}$, for any $n = 1, \cdots, N_t$ solve the problem:

given p_h^n, and $p_{h,0}^{n+1} = p_h^n$, for any $k \geq 1$, until $\frac{\|p_{h,k+1}^{n+1} - p_{h,k}^{n+1}\|_{H^1}}{\|p_{h,k+1}^{n+1}\|_{H^1}} < \epsilon$,

find $p_{h,k+1}^{n+1} \in V_h$, such that $\forall v_h \in V_h$:

$$m(p_{h,k+1}^{n+1}, v_h) + a_{p_{h,k}^{n+1}}(p_{h,k+1}^{n+1}, v_h) + d_{p_{h,k}^{n+1}}(p_{h,k+1}^{n+1}, v_h) = F(v_h) \quad , \tag{3.27}$$

where

$$\begin{aligned}
\text{(mass)} \quad m(p_{h,k+1}^{n+1}, v_h) &= \frac{1}{\Delta t^n} \int_{\Omega_h} p_{h,k+1}^{n+1}\, v_h \, d\Omega \,, \\
\text{(advection)} \quad a_{p_{h,k}^{n+1}}(p_{h,k+1}^{n+1}, v_h) &= \int_{\Omega_h} p_{h,k+1}^{n+1} \Big(a_1[p_{h,k}^{n+1}] \, \frac{\partial v_h}{\partial x} + a_2[p_{h,k}^{n+1}] \, \frac{\partial v_h}{\partial y} \Big) d\Omega \,, \\
\text{(diffusion)} \quad d_{p_{h,k}^{n+1}}(p_{h,k+1}^{n+1}, v_h) &= \int_{\Omega_h} \frac{\partial p_{h,k+1}^{n+1}}{\partial x} \Big(D_{11}[p_{h,k}^{n+1}] \, \frac{\partial v_h}{\partial x} + D_{21}[p_{h,k}^{n+1}] \frac{\partial v_h}{\partial y} \Big) d\Omega \\
&+ \int_{\Omega_h} \frac{\partial p_{h,k+1}^{n+1}}{\partial y} \Big(D_{21}[p_{h,k}^{n+1}] \, \frac{\partial v_h}{\partial x} + D_{22}[p_{h,k}^{n+1}] \frac{\partial v_h}{\partial y} \Big) d\Omega \,, \\
\text{(right hand side)} \quad F(v_h) &= m(p_h^n, v_h) = \frac{1}{\Delta t^n} \int_{\Omega_h} p_h^n \, v_h \, d\Omega \,.
\end{aligned}$$

Evaluation of the conditional expectation integral

For every fixed point iteration we need to evaluate the local stochastic volatility surface $\sigma_{LSV}(x, t^n) = \sigma_{LV}(x, t^n)\sqrt{I[p_{h,k}^n](x)}$. Thus the non linear integral term $I[p_{h,k}^n](x)$ has to be evaluated. Since we use P1 elements (3.26) the solution $p_h(x_*, y)$ is a linear piecewise continuous function in the variable y. For this reason a simple rectangle-quadrature formula is able to correctly integrate the solution as long as the integration intervals are small enough. Thus we shall evaluate the non linear term with the following expression:

$$I[p_{h,k}^n](x_*) = \frac{\int_0^{y_{max}} p_{h,k}^n(x_*, y) dy}{\int_0^{y_{max}} y \, p_{h,k}^n(x_*, y) dy} \simeq \frac{\sum_i^{N_y} \Delta y_i \, p(x_*, y_i, t) + \epsilon}{\sum_i^{N_y} y_i \, \Delta y_i \, p(x_*, y_i, t) + \gamma\epsilon} \,. \tag{3.28}$$

In order to avoid numerical instabilities when the numerator and the denominator becomes quite small, as suggested by [23] it is useful to add a small stabilization value $\epsilon \simeq 10^{-6}$. The value of γ gives the asymptotic value of the non linear term. Since Y_t follows a C.I.R. process its expected value is given by [63]:

$$\mathbb{E}[Y_t] = y_0 \, e^{-\kappa t} + \theta \left(1 - e^{-\kappa t}\right)$$

Thus $\mathbb{E}[Y_t|X_t = x]\big|_{t=0} = y_0$ while $\mathbb{E}[Y_t]\big|_{t\to\infty} = \theta$. For this reason it is reasonable to use $\gamma = y_0$ for small t and $\gamma = \theta$ for large t.

The nonlinear term is very sensitive. To further stabilize the computation of its first derivative, an averaged and smoothed version of the nonlinear term is used , namely:

$$\tilde{I}[p^n_{h,k}](x_j) = \frac{1}{4}I[p^n_{h,k}](x_{j-1}) + \frac{1}{2}I[p^n_{h,k}](x_j) + \frac{1}{4}I[p^n_{h,k}](x_{j+1})$$

At each iteration using the approximated formula (3.28) it is possible to evaluate the integral term for every nodes of the mesh and then build the finite element local stochastic volatility:

$$\sigma_{h,LSV}(x, y, t^n) = \sum_{i=1}^{N_h} \sigma_{LV}(x_i, t^n)\sqrt{I[p^n_{h,k}](x_i)}\,\varphi_i(x, y)\,, \tag{3.29}$$

where $\{\varphi_i\}_{N_h}$ is the Lagrangian basis of V_h and $\{x_i\}_{N_h}$ are the x-coordinates of the triangulation vertices.

Although the simple mid-point rectangle quadrature formula works fine with P1 finite elements it is important to use an adaptive length of the quadrature intervals. In particular we know that the solution is initially very peaked around the point $(0, y_0$ and almost null elsewhere. In order to correctly integrate this critical point an high number of quadrature nodes are needed. On the other hand few quadrature nodes are required elsewhere. A good automatic algorithm could assign the number and the distribution of the quadrature nodes along the line $(x = x_*, y)$ considering the value of the solution and its first derivative at the previous step.

Stabilization methods

The problem we are dealing with is an advection-diffusion equation with a dominant advection component in some region of the domain. In particular near the boundary $y = 0$ the problem degenerates, and we need a numerical stabilization.

Definition 3.9. (Local Péclet Number) Let's consider the equation (3.9), we call local Péclet number the scalar quantity $\mathbb{P}e_K$ for each $K \in \Omega_h$ [54]:

$$\mathbb{P}e_K(\mathbf{x}) = \frac{h_K|\mathbf{a}(\mathbf{x})|}{2\,\delta_K(\mathbf{x})}\,, \quad \forall \mathbf{x} \in K\,, \tag{3.30}$$

where h_K is a characteristic length of the element K, $\|\mathbf{a}\|$ is the euclidean norm of the advection vector and δ_K is defined as:

$$\delta_K(\mathbf{x}) = \min_{\boldsymbol{\xi}\in\mathbb{R}^2\setminus\mathbf{0}} \frac{\sum_{i,j=1}^{2} D_{ij}(\mathbf{x})\,\xi_i\xi_j}{|\boldsymbol{\xi}|^2}\,, \quad \forall \mathbf{x} \in K\,, \tag{3.31}$$

The Peclet number gives a measure of the dominance of the advection field on the diffusive one. In particular if $\mathbb{P}e > 1$ the numerical solution presents oscillations depending on the domination of the advection term. It is important to remark that the problem we are dealing with degenerates in the limit $y \to 0$. Namely near this boundary we lose the ellipticity and the problem becomes purely advective. For this reason the "stabilization", i.e. the (partial) elimination of the numerical oscillations, assumes a relevant role.

A Galerkin stabilized method is a Galerkin method with an added term which stabilizes the problem enforcing the diffusive term. For this reason we talk about "artificial viscosity" since the idea is to introduce a fictitious diffusive term to supply the real one. We call $s_{p_h^n}(p_{h,k+1}^{n+1}, v_h)$ the additional term to the variational equation (3.27). This depends on the particular stabilization method adopted. In the case of the **Streamline-Upwind Diffusion method** [53] it assumes the following form:

$$s_{p_h^n}(p_{h,k+1}^{n+1}, v_h) = Qh \int_\Omega (\mathbf{a}[p_{h,k}^{n+1}] \cdot \nabla p_{h,k+1}^{n+1}) \, (\mathbf{a}[p_{h,k}^{n+1}] \cdot \nabla v_h) \, d\Omega \qquad \text{with } Q = |\mathbf{a}|^{-1}.$$

This corresponds to add to the initial problem a term which is proportional to the second derivative in the direction of the field $\mathbf{a}$. Thus we are introducing an artificial diffusion only along the advection field. We remark that in this case the coefficient of the artificial viscosity is a tensor. In particular the stabilization term $s(\,,\,)$ can be interpreted as the bilinear form associated to the operator $-\mathrm{div}(\mu_{\mathbf{a}} \nabla p)$ where $[\mu_{\mathbf{a}}]_{ij} = Q\, h\, a_i\, a_j$, being a_i the i-th component of the vector $\mathbf{a}$. The original integro-differential operator (3.10) is thus modified in the following way:

$$\mathcal{L}(p, I[p]) = \mathrm{div}\,(\,[\,D[p] + \mu_{\mathbf{a}}[p]\,]\,\nabla p + \mathbf{a}[p]\, p\,)\,. \tag{3.32}$$

The accuracy of this method is only $\mathcal{O}(h)$. To achieve better results we could use the so called *"strongly consistent methods "*. They consists in stabilization methods with the following property:

$$s_{p_h^n}(p_{h,k+1}^{n+1}, v_h) \quad \text{is such that} \qquad s_{p_h^n}(p, v_h) = 0 \quad \text{for every} \quad v_h \in V_h$$

We observe that the Streamline-Diffusion method is not strongly consistent, being

$$s_{p_h^n}(p, v_h) = Qh \left(\frac{\partial p}{\partial \mathbf{a}}, \frac{\partial v_h}{\partial \mathbf{a}} \right) \quad \text{for every} \quad v_h \in V_h\,.$$

However, we choose to use this stabilization method since strongly consistent method, like the GLS, the SUPG and the DW [53], requires the evaluation of the differential operator $\mathcal{L}$ (2.24) and then the evaluation of the first and second derivatives of the non linear integral term $I[p](x,t)$. Given the numerical evaluation of $I[p](x,t)$ these derivatives could be quite irregular and then these methods could be unstable. For this reason we prefer the simpler Streamline-Upwind diffusion stabilization.

Definition 3.10 (Stabilized Galerkin Finite Element method)**.**

Given $p_h^0 = p_0$ and $I[p_h^0] = \frac{1}{y_0}$, for any $n = 1, \cdots, N_t$:
given p_h^n, and $p_{h,0}^{n+1} = p_h^n$, for any $k \geq 1$, until $\frac{\|p_{h,k+1}^{n+1} - p_{h,k}^{n+1}\|_{H^1}}{\|p_{h,k+1}^{n+1}\|_{H^1}} < \epsilon$,
find $p_{h,k+1}^{n+1} \in V_h$, such that $\forall v_h \in V_h$:

$$m(p_{h,k+1}^{n+1}, v_h) + a_{p_{h,k}^{n+1}}(p_{h,k+1}^{n+1}, v_h) + d_{p_{h,k}^{n+1}}(p_{h,k+1}^{n+1}, v_h) + s_{p_h^n}(p_{h,k+1}^{n+1}, v_h) = F(v_h)\,.$$

Algebraic problem

In order to effectively solve the problem (3.10) we provide its algebraic formulation. For this purpose we introduce a finite element basis $\{\varphi_j\}$ of V_h and we observe that, if (3.10) is satisfied for the basis functions, then it is satisfied by all the functions in V_h. Moreover, being $p_{h,k+1}^{n+1} \in V_h$ for every $n > 0$, then we have

$$p_{h,k+1}^{n+1}(x) = \sum_{j=1}^{N_h} p_{j,k+1}^{n+1} \varphi_j(\mathbf{x}), \qquad \forall \mathbf{x} \in \Omega, \qquad p_{j,k+1}^{n+1} = p_{h,k+1}^{n+1}(\mathbf{x_j})\,,$$

where $\mathbf{x_j}$ is the j mesh node and $p_{j,k+1}^{n+1}$ are the unknowns value of the finite element solution at the node $\mathbf{x_j}$. In this way the (3.10) becomes:

$$\left[\frac{1}{\Delta t^n} M + A[\mathbf{p_k^{n+1}}] + D[\mathbf{p_k^{n+1}}] + S[\mathbf{p_k^{n+1}}]\right] \mathbf{p_{k+1}^{n+1}} = \frac{1}{\Delta t^n}\, \mathbf{p^n}\,, \tag{3.33}$$

where the matrices M, A, D, S are the algebraic equivalent of the mass, advection, diffusion and stabilization bilinear form, defined as:

$$\begin{aligned} M_{i,j} &= m(\varphi_j, \varphi_i)\,, \\ A[\mathbf{p_k^{n+1}}]_{i,j} &= a_{p_{h,k}^{n+1}}(\varphi_j, \varphi_i)\,, \\ D[\mathbf{p_k^{n+1}}]_{i,j} &= d_{p_{h,k}^{n+1}}(\varphi_j, \varphi_i)\,, \\ S[\mathbf{p_k^{n+1}}]_{i,j} &= s_{p_{h,k}^{n+1}}(\varphi_j, \varphi_i)\,. \end{aligned}$$

and we have introduced the unknown vector $\mathbf{p_{k+1}^{n+1}} = [p_{j,k+1}^{n+1}]$.

We observe that we managed to reduce the solution of the nonlinear differential equation into a sequence of non linear systems of algebraic equations:

$$C[\mathbf{p^{n+1}}]\, \mathbf{p^{n+1}} = \frac{1}{\Delta t^n} \mathbf{p^n}\,, \tag{3.34}$$

where:

$$C[p^{n+1}] = \frac{1}{\Delta t^n} M + D[p^{n+1}] + A[p^{n+1}] + S[p^{n+1}]\,.$$

At this point we could rethink the whole problem with an algebraic perspective applying one of the several methods known for the solution of nonlinear systems. Unfortunately since the matrices A, D, S depend in high non linear way on the solution $\mathbf{p^{n+1}}$ it is not possible to get an analytical form of the Jacobian of C and even an approximated evaluation of it would be computationally very demanding. As a matter of fact in order to evaluate C we first need to evaluate the non linear integral term $I[p]$. For this reason the simple fixed point iterations used with a suitable time step Δt behave better than more advanced method like Newton or Quasi-Newton.

Chapter 4

Numerical Simulations

In this chapter the Heston-Dupire model (3.1) is actually implemented. The aim is to first calibrate the model from real market data and then use it to reconstruct the implied volatilities of the model and compare them with the market ones. In particular we consider the market of European call options quoted on the *SX5E* Index at 1^{st} June 2012. The original data can be retrieved in Table A.1 while Figure 4.1 presents the market implied volatility surface $\sigma_I^{Market}(K,T)$ as a function of strike price and time to maturity. The initial asset value S_0 is 2068.67 and we choose to use a flat interest rate $r = 0.01$. We can observe the typical *"Smile"* of implied volatility for short maturities followed by a negative and flatter slope for longer maturities.

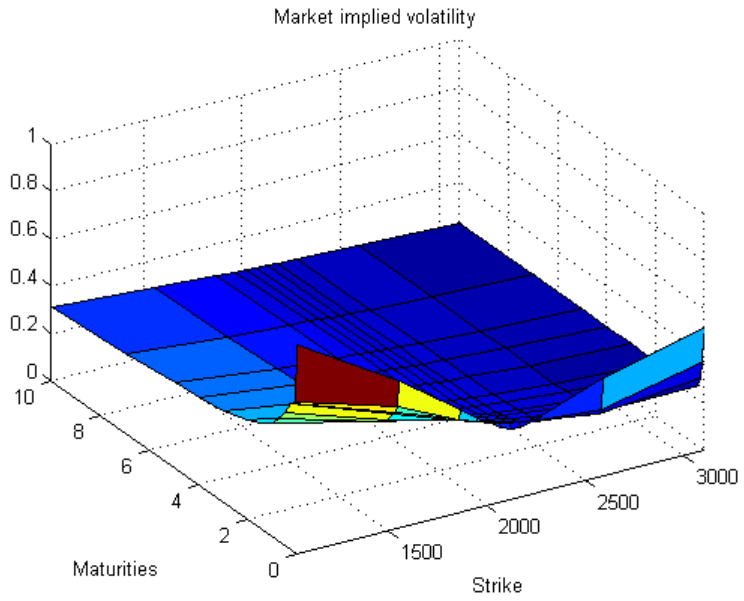


Figure 4.1: Implied volatility surface of market prices $\sigma_I^{Market}(K,T)$

From these data, we first calibrate the Heston parameters and the Dupire Local Volatility. Afterwards we use these values as the parameters of the nonlinear Fokker-Planck equation (3.16) which we solve with the Finite Element method presented in Section 3.3, reconstructing, in this way, the Local-Stochastic volatility surface σ_{LSV}. Then we compare the Implied Volatilities of the three models with the market ones. In order to clearly distinguish the several kind of Implied Volatilities we will deal with, let's introduce the following notation:

σ_I^{Market} represents the implied volatility actually observed on the market A.1;

σ_I^{SVI} is the SVI implied volatility 1.2.4 interpolated and extrapolated from the market one;

$\sigma_I^{LVM}, \sigma_I^{SVM}, \sigma_I^{LSVM}$, represent, respectively, the implied volatility realized by the Dupire Local Volatility model, the Heston Stochastic Volatility model and the Heston-Dupire Local-Stochastic Volatility model;

Finally in Section 4.3 we will present a last concept of volatility, namely the Forward Implied Volatility σ_F.

4.1 Model parameters setting

The first step in order to implement a Local-Stochastic Volatility model is to independently calibrate from market data the two base models. In this section we provide the parameters calibrated for the Dupire LV model and the Heston SV model.

4.1.1 Dupire Local Volatility calibration

Following the procedure described in Section 1.2 the Dupire Local Volatility has been calibrated from market data A.1. We now present the main results of the calibration procedure. The first issue is to make a good fitting of market implied volatilities with the SVI parametrization described in Section 1.2.4. For every maturity T_n we calibrate the five coefficients of the model defining reasonable values for the feasible set W. The input data for the minimization problem (1.29) are shown in Table 4.1.1. Starting from the first maturity, and using the initial guess conditions shown in Table 4.1.1, the cost functional (1.30) is minimized with a Levenberg-Marquardt algorithm [34]. The whole calibration procedure has been implemented with MATLAB® 7.10.0 using the lsqnonlin routine. 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. In this way the calibration procedure is accomplished with a serial method.

Dupire coefficients	Initial guess	Lower bound	Upper bound
C_1	0.001	0.0001	10.0
C_2	0.6	0.0001	10.0
C_3	-0.3	-1.0	1.0
C_4	0.05	-10.0	10.00
C_5	0.05	0.0001	10.00

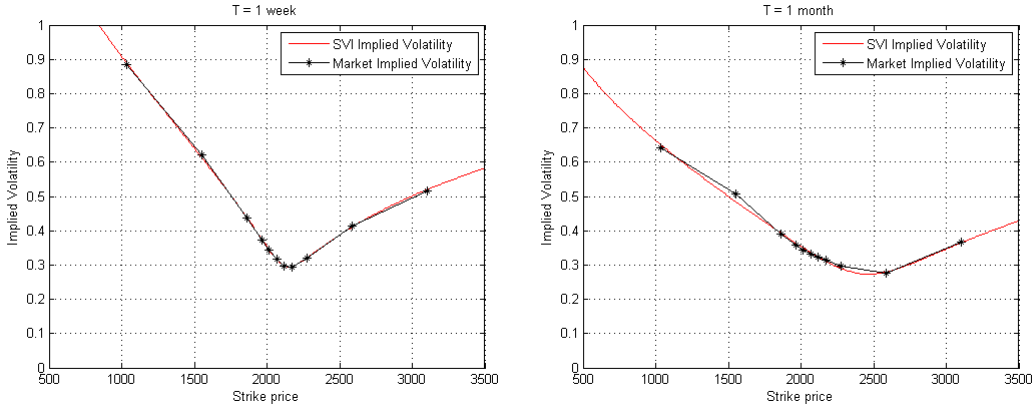
Table 4.1: Input parameter for the calibration of the SVI Implied Volatility

Actually, the SVI parametrization of market data looks to be a robust method and a good fitting can be achieved even if the calibration of the smiles is accomplished in a parallel way. In fact, we can simply calibrate all the smiles independently, starting from the same first guess condition. However, in this case, the time requested for a single calibration is quite longer, up to five times the one needed with the sequential calibration since we do not exploit the information provided by the previous calibrations. Moreover in this way the minimization algorithm can be ended up with quite different coefficients from one maturity to the next one. Thus, the calibrated SVI Implied Volatility surface could be more irregular. For these reasons we choose to use the sequential calibration. Below, Table 4.2 shows the results of the calibration for the Gatheral SVI parametrization of Implied Volatility relative to the first nine maturities.

Calibrated values	T_1 1 week	T_2 1 month	T_3 2 months	T_4 3 months	T_5 6 months
C_1	0.05144	0.00013	0.00010	0.01869	0.00010
C_2	0.7994	0.4792	0.3590	0.2729	0.2285
C_3	-0.2868	-0.0006	-0.2418	-0.6045	-0.5703
C_4	0.0253	0.1756	0.1838	0.1295	0.1611
C_5	0.0022	0.0242	0.0347	0.0401	0.0891

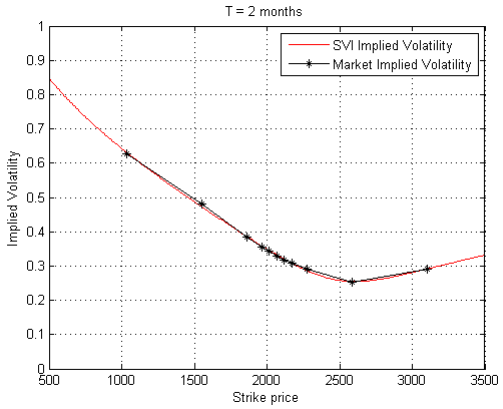
Calibrated values	T_6 9 months	T_7 12 months	T_8 18 months	T_9 24 months
C_1	0.00013	0.00012	0.00014	0.00014
C_2	0.1869	0.1636	0.1289	0.1213
C_3	-0.5657	-0.4734	-0.4541	-0.3763
C_4	0.1960	0.2370	0.2921	0.3122
C_5	0.1172	0.1312	0.1869	0.1972

Table 4.2: Calibrated coefficients of Dupire Local Volatility

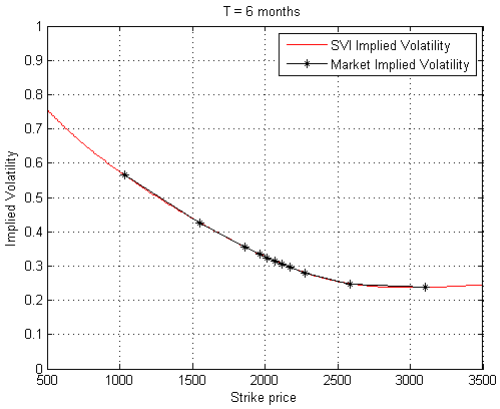


(a)

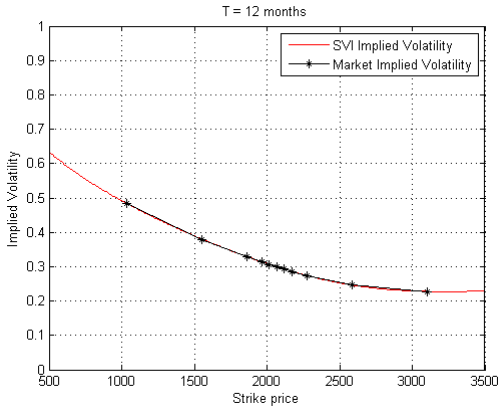
(b)



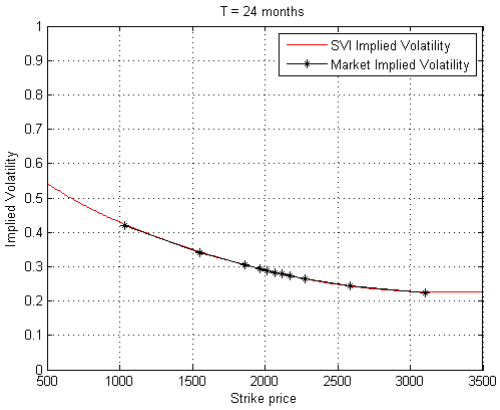
(c)



(d)



(e)



(f)

Figure 4.2: Comparison between the implied volatilities observed on the market and the ones reconstructed with the SVI model for different time to maturity.

Figure 4.2 shows, with a solid red line, the fitted SVI implied volatility compared with the observed market implied volatility. We can see that the model succeed in a very good interpolation of market data for every maturity. Even the marked smile of the very first fixing date 4.2(a) is well reproduced. Figure 4.3 shows the whole surface of calibrated SVI Implied Volatility $\sigma_I^{SVI}(K,T)$ as a function of strike prices and maturities. Comparing Figure 4.3 with 4.1 we can notice the good fitting of the whole SVI - surface with the real one.

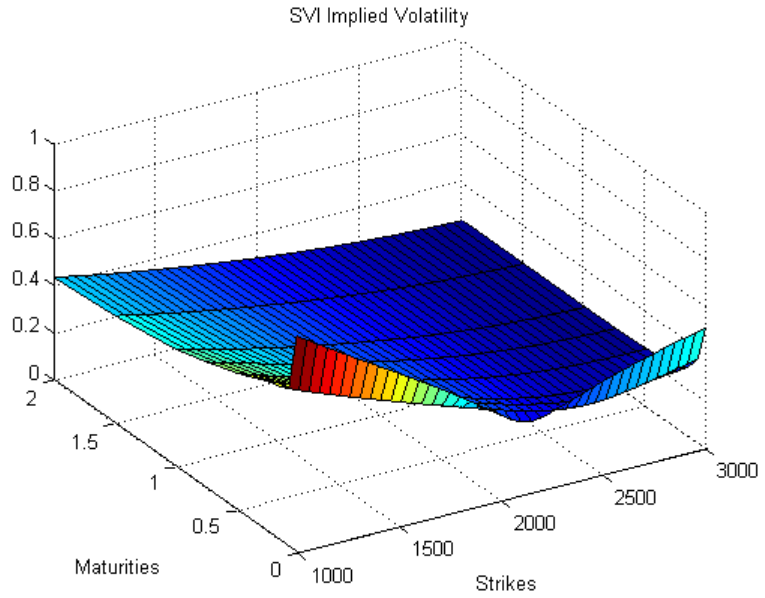


Figure 4.3: Implied Volatility surface of the calibrated SVI parametrization

Once that we have reconstructed the SVI Implied Volatility surface we can use the formula (1.31) to derive the Dupire Local Volatility $\sigma_{LV}(K,T)$, which is shown in Figure 4.4. This procedure is rather delicate since the ratio inside the square root of formula (1.31) may be negative. For this reason, as explained in Section 1.2.5, we add a small positive constant to both the numerator and the denominator. Once $\sigma_{LV}(K,T)$ is reconstructed we can solve the Dupire equation (1.8) with a standard Crank-Nicholson Finite Difference scheme [55, 62] in order to price the plain vanilla contracts quoted on the market. Then, inverting these prices we obtain the model Implied Volatilities $\sigma_I^{LVM}(K,T)$ represented in Figure 4.5. As expected by the theory, we obtain an overall good representation of the SVI Implied Volatility surface 4.3.

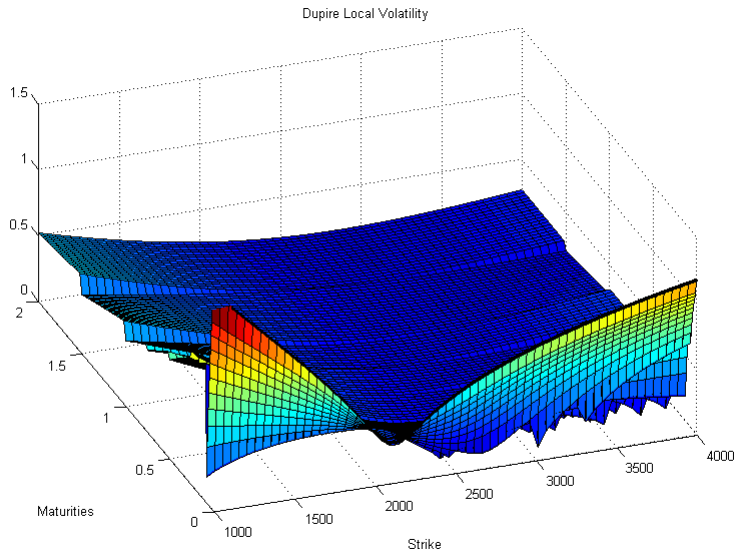


Figure 4.4: Dupire Local Volatility surface $\sigma_{LV}(s,t)$

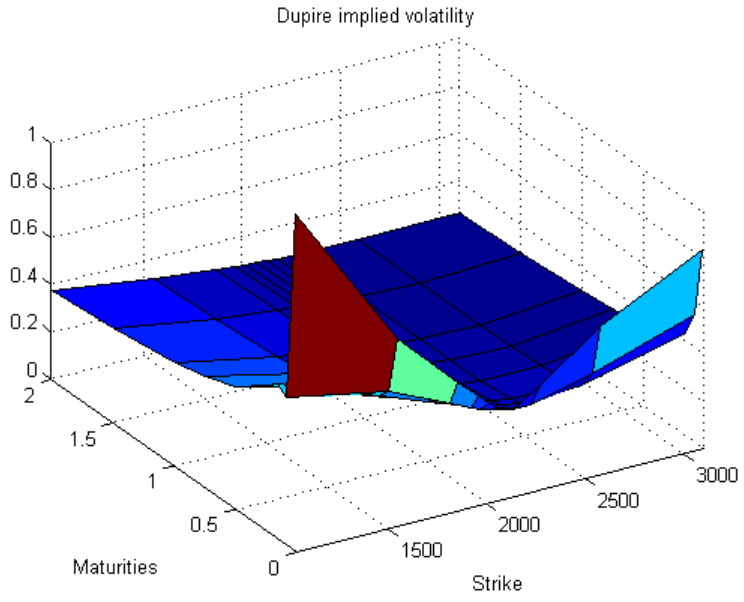


Figure 4.5: Dupire Implied Volatility surface $\sigma_I^{LVM}(s,t)$

4.1.2 Heston Stochastic Volatility calibration

In the same way we did for the Local Volatility, we now present the calibration of the Heston model parameters following the procedure described in Section 1.1.2. Again we need to solve the nonlinear minimization problem (1.7). Thus, we first look for an initial guess solution, according to the recipe (1.12). Then we set

$$\begin{aligned} y_0 &= \left(\sigma_I^{Market}(S_0, T_1)\right)^2 = 0.3181^2 = 0.1012\,, \\ \theta &= \left(\sigma_I^{Market}(S_0\, e^{rT_{max}}, T_1)\right)^2 \simeq 0.2788^2 = 0.0777\,, \\ \rho &= -0.8\,, \\ \eta &= \frac{2S_0}{\rho}\frac{\sigma_I^2(S_0+\Delta S, T_1) - \sigma_I^2(S_0-\Delta S, T_1)}{2\Delta S} = 0.7420\,, \\ \kappa &= 0.05 + \frac{\eta^2}{2\theta} = 3.5912\,. \end{aligned}$$

It is worth to notice that, on the contrary of the calibration problem (1.29) for the SVI fitting of Implied Volatility, the calibration of the Heston parameters(1.7) looks to be quite sensitive to the initial condition. For this reason is important to set an appropriate first guess. Indeed the recipe (1.12) lead, at least in our case, to a very good starting point. Once defined a suitable feasible set W, we use a Levenberg-Marquardt algorithm to solve the minimization problem with MATLAB® 7.10.0. Table 4.1.2 presents the results of the calibration.

Heston parameters	Initial guess	Lower bound	Upper bound	Calibrated value
y_0	0.1012	0.05	5.0	0.1377
θ	0.0777	0.05	5.0	0.2262
κ	3.5912	0.05	5.0	2.4047
η	0.7420	0.05	5.0	0.7802
ρ	-0.8	-0.99	0.99	-0.8189

Table 4.3: Calibrated parameters of Heston Stochastic Volatility

We can see that the Feller condition is verified, being the Feller number:

$$Fe = \kappa\theta - \frac{1}{2}\eta^2 = 0.2396 > 0$$

Once the Heston parameters are calibrated it is possible to use the Heston model to price the plain vanilla contracts quoted on the market and then reconstruct the Implied Volatility surface of the model σ_I^{SVM} as shown in Figure 4.6. We see that the implied volatility of the Heston model σ_I^{SVM} is overall quite realistic with respect to the market implied volatility σ_I^{Market} represented in Figure

4.1. However as we expected the Heston model is not able to completely capture the initial smile on the contrary of the Local Volatility.

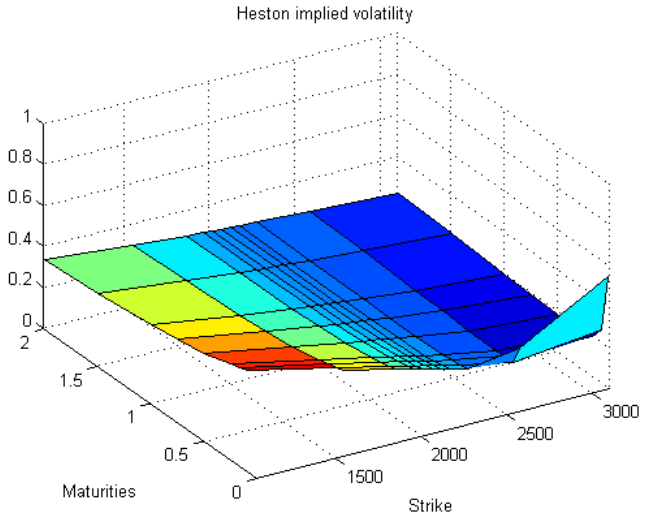


Figure 4.6: Implied volatility surface σ_I^{SVM} of the Heston Stochastic Volatility Model

4.2 Solution of the Fokker-Planck equation

Once that all the coefficients have been finally set, it is possible to solve the non linear Fokker-Planck equation (3.16). In this section we present the results of the numerical simulation performed with the software FreeFem++ 3.12 [39] according to the method already described in Section 3.3. Appendix B presents some excerpts of the code implemented, proving some further details on the algorithm.

The solution of the problem is a probability density function. Thus the most important issue is obtaining a solution such that its integral over the domain is nearly constantly equal to one. In particular we want en error of less than 1%, namely we want that $0.99 \leq \int_{\Omega_h} p_h^n d\Omega \leq 1.01$, $\forall n$. Thanks to several experimental simulations we found that there are five major critical issues to consider in order to obtain such a stable solution. In particular they are related to the intrinsic difficulties of the equation (3.16) namely the combination of the nonlinearity and the degeneracy at the boundary.

- The initial value of the variance y_0 should be not too close to zero, generally speaking $y_0 > 0.05$. Otherwise the initial condition lays in the region where the elliptic operator degenerates and the problem becomes unstable at the very beginning of its evolution.

- The Feller condition should be satisfied. In perfect agreement with the theory, see Remark 3.2, we notice that the Feller number actually rules the advection field of the solution. If $Fe > 0$ then the solution moves away from the boundary $y = 0$. On the contrary if the Feller number is negative the solution moves in the degenerate region.
- The number of mesh triangles around the solution should be quite large. Otherwise there would be not enough element to get an accurate resolution for the evaluation of the total probability. At the same time the solution is quite concentrated, especially at the beginning. For this reason the use of an adapted mesh is a natural choice.
- The application of some fixed point iterations during the same time step increase the stability of the solution. We notice that few iterations should be enough to obtain a good convergence.
- The Dupire Local Volatility surface $\sigma_{LV}(K, T)$ should be smooth and regular. In fact the advection field (3.14), (3.15) depends on the derivatives of σ_{LV}.

Considering all the previous observations, the equation has been solved successfully. Figure 4.7 shows the evolution of the probability density function $p(x, y, t)$ solution of the equation observed at several maturities.

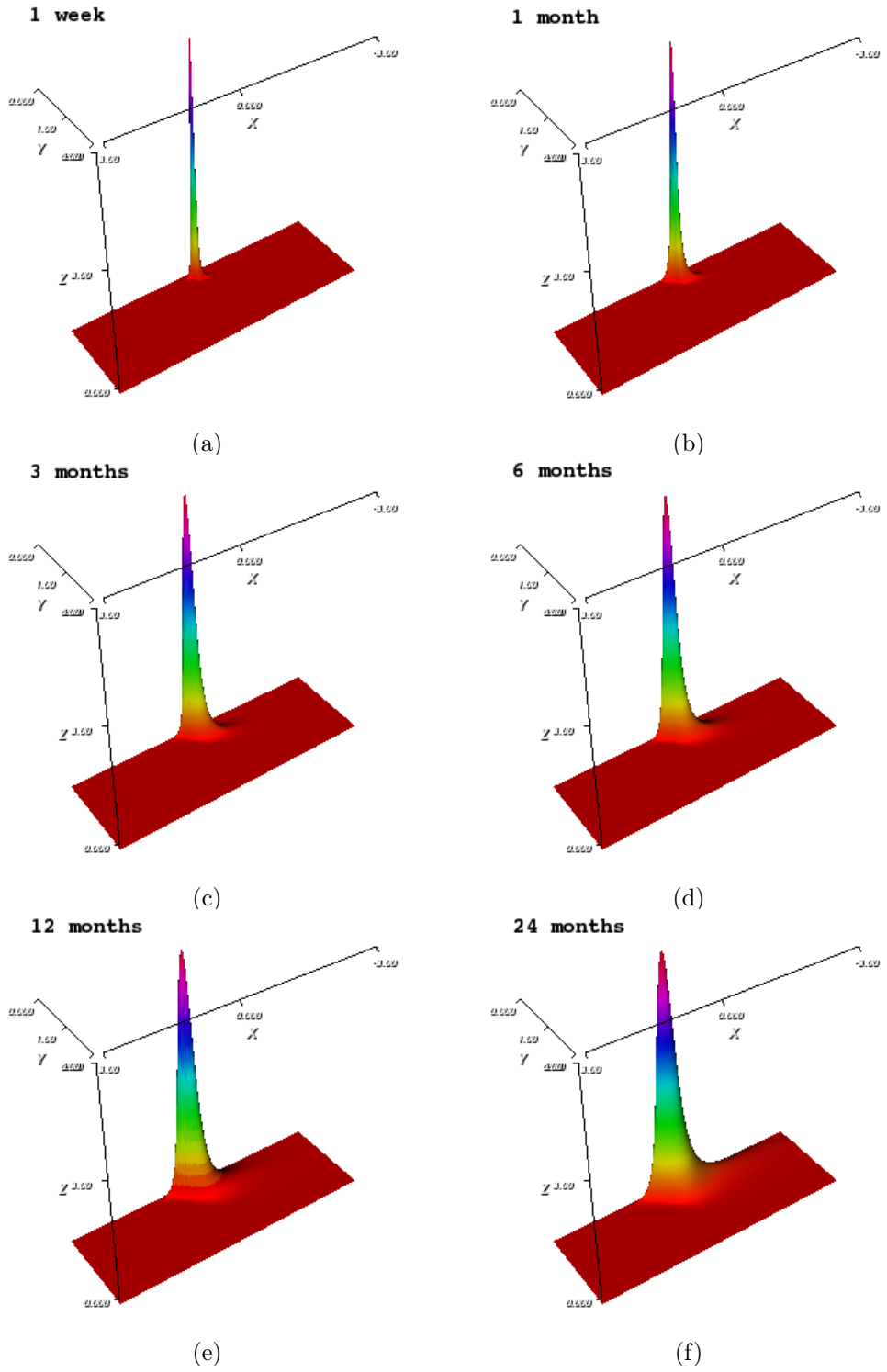
Once the probability density function is retrieved, the price of plain vanilla contracts can be easily evaluated by a simple integration:

$$C^{LSV}(K, T) = e^{-rT} \int_{\Omega} \left(S_0 \, e^x - K \right)^+ p(x, y, T) \, dx \, dy \tag{4.1}$$

Hence we can reconstruct the surface of prices $C^{LSV}(K_j, T_i)$ and then, inverting the Black-Scholes formula (13), obtain the model Implied Volatility surface σ_I^{LSVM}, shown in Figure 4.8. We can notice that the model succeed in reproducing the characteristic smile and it exhibits a non flat implied volatility for longer maturities. Solving the equation we obtain also the conditional expectation function:

$$\mathbb{E}[Y_t | X_t = x] = \frac{\int_0^{y_{max}} y \, p(x, y, t) dy}{\int_0^{y_{max}} p(x, y, t) dy}$$

shown in Figure 4.9(a) and its inverse $I[p](x, t)$, represented in Figure 4.9(b). Then, applying the result 2.4, we can finally evaluate the Local-Stochastic Volatility surface $\sigma_{LSV}(x, t) = \sigma_{LV}(x, t)\sqrt{I[p](x, t)}$, represented in Figure 4.10. In this way the model is completely calibrated to market data and ready to be used for exotic derivatives pricing.



(a) (b) (c) (d) (e) (f)

Figure 4.7: Time evolution of the probability density function $p(x, y, t)$, solution of the Fokker Planck equation, observed at different times: 1 week (a), 1 month (b), 3 months (c), 6 months (d), 1 year (e) and 2 years (f) .

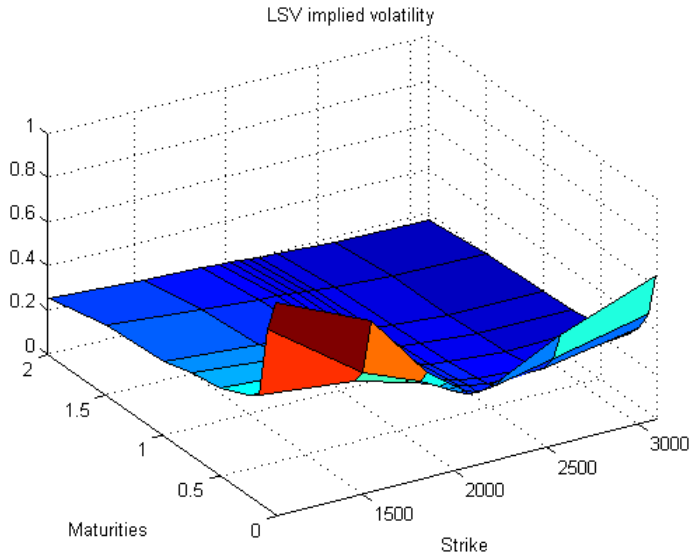
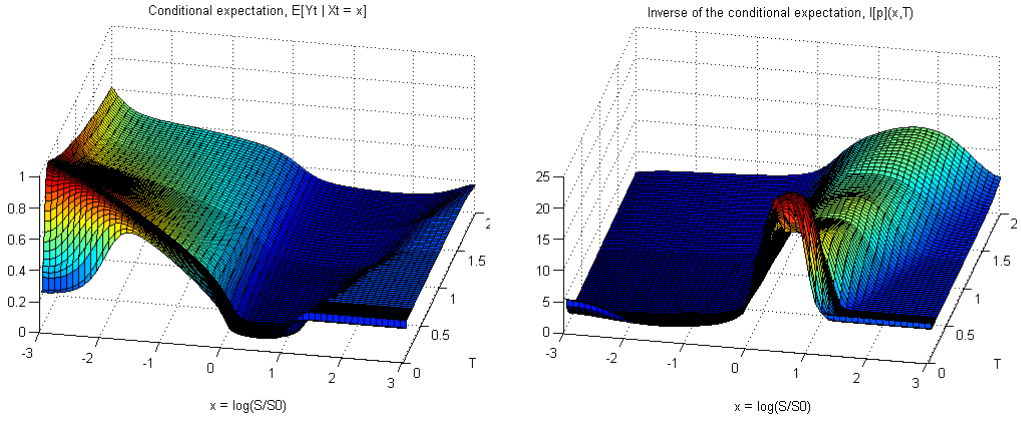


Figure 4.8: Implied volatility surface $\sigma_I^{LSVM}(K,T)$ of the Heston-Dupire model



(a) (b)

Figure 4.9: Conditional expectation function $\mathbb{E}[Y_t|X_t = x] = \frac{\int_0^{y_{max}} y\, p(x,y,t)dy}{\int_0^{y_{max}} p(x,y,t)dy}$ Figure (a) and its inverse $I[p](x,t) = \frac{\int_0^{y_{max}} p(x,y,t)dy}{\int_0^{y_{max}} y\, p(x,y,t)dy}$ Figure (b).

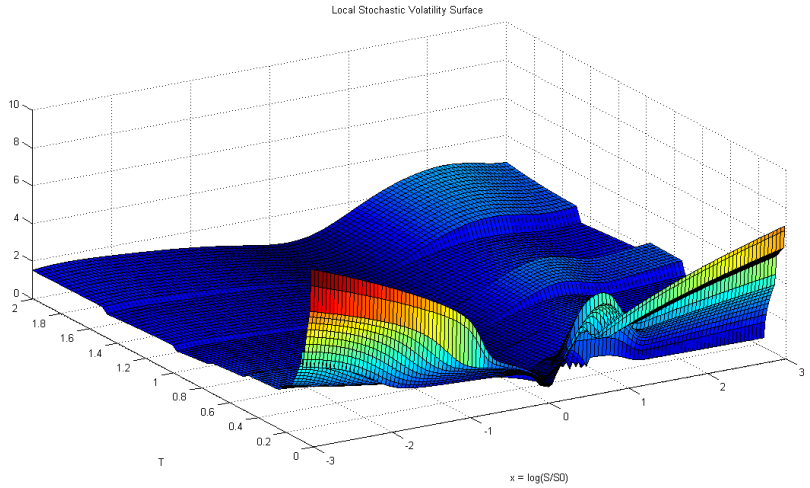


Figure 4.10: Local Stochastic Volatility surface $\sigma_{LSV}(x,t) = \sigma_{LV}(x,t)\sqrt{I[p](x,t)}$

We also show $\sigma_{LSV}(s,t)$ as a function of the S variable in order to make a fair comparison with the original Local Volatility surface $\sigma_{LV}(s,t)$ represented in Figure 4.4. We see that the two surfaces are rather similar although $\sigma_{LSV}(s,t)$ exhibits higher values.

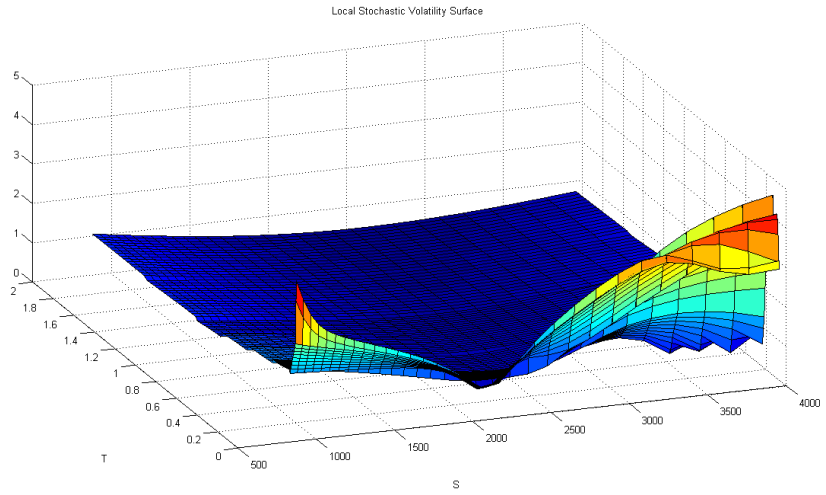


Figure 4.11: Local Stochastic Volatility surface $\sigma_{LSV}(S,t) = \sigma_{LV}(S,t)\sqrt{I[p](S,t)}$

In Figure 4.12 we can finally compare the different implied volatilities of the three model σ_I^{LVM}, σ_I^{SVM} and σ_I^{LSVM} with the real observed ones on the market σ_I^{Market}. Generally we can see how the LSV model (represented with a blue line) succeeds in reproducing the form of market implied volatility (black line) nearly

as well as the LV model (green line) whereas the SV model (red line) exhibits a worst fitting.

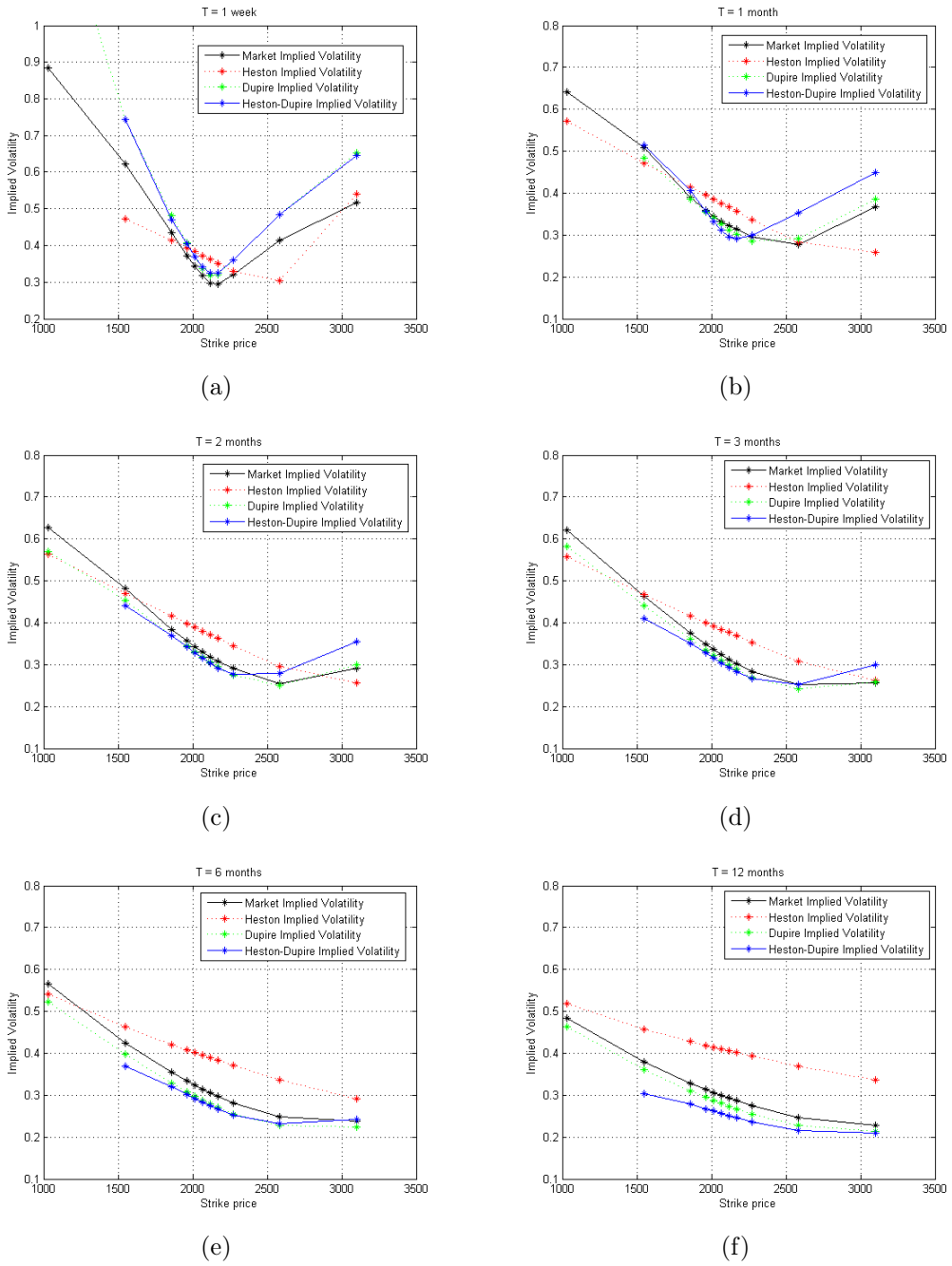


Figure 4.12: Comparison between the Implied Volatilities observed on the market σ_I^{Market} and the ones reconstructed by the three different models σ_I^{LVM} σ_I^{SVM} σ_I^{LSVM}, for different time to maturity.

In Figure 4.13 we plot the Delta $\Delta = \frac{\partial C}{\partial S}$, obtained with the different models. We see that they are comparable and in particular the ones obtained with the LV and LSV model are very close.

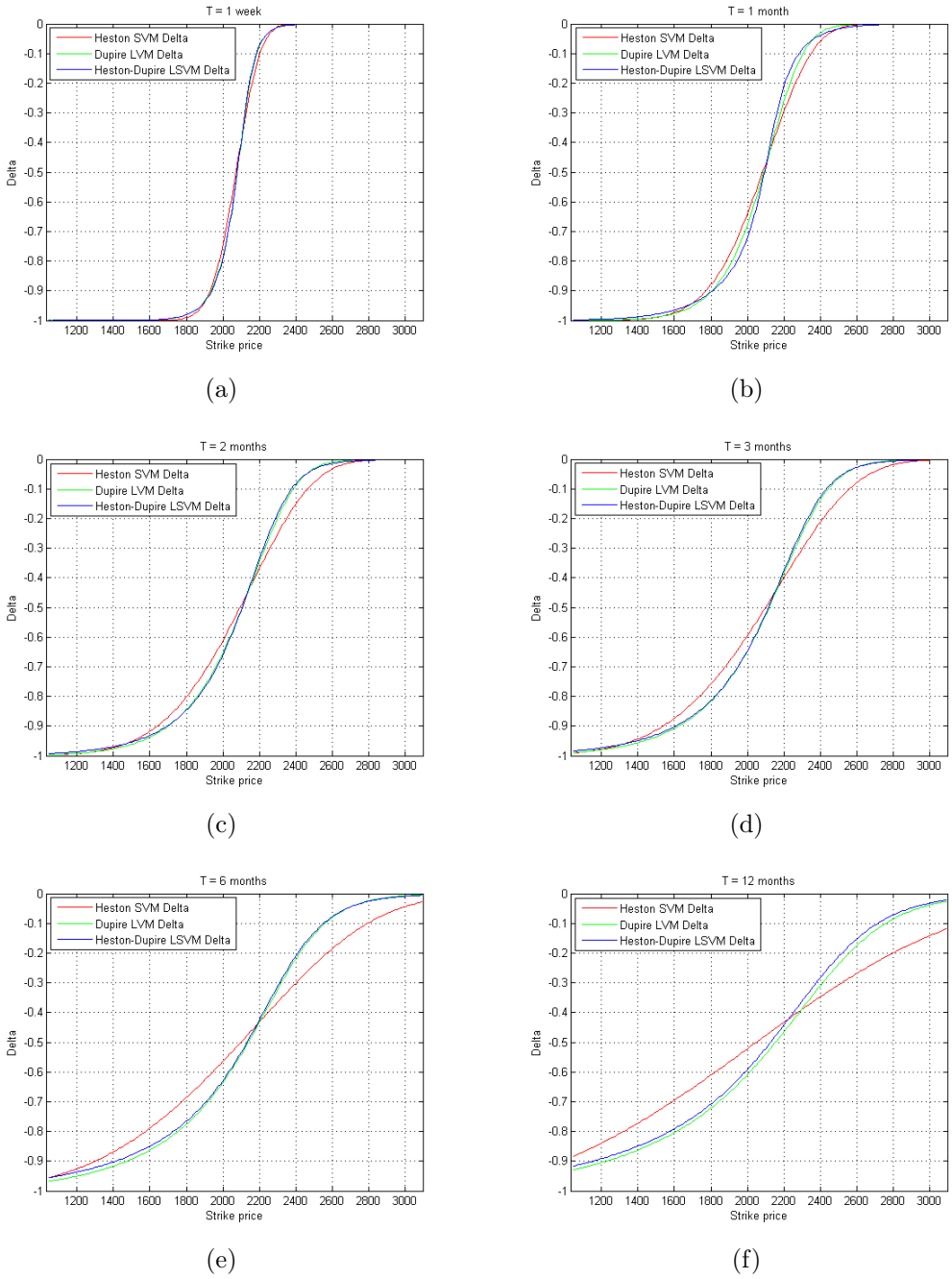


Figure 4.13: Comparison between the greek Delta of the three different models Δ^{LVM}, Δ^{SVM}, Δ^{LSVM} for different time to maturity.

4.3 Forward Volatility

We have seen in Figure 4.12 that the Heston-Dupire model succeeds in reproducing the volatility smile of the Implied Volatility as the Dupire Local Volatility model does. We now want to verify that our LSV model exhibits also a realistic dynamic for the Forward Implied volatility like the one of the Heston model. Before defining this last volatility concept we first need to introduce a particular kind of derivative contracts called Forward-Starting Options.

Let T_1, T_2 be two fixing dates such that $0 < T_1 < T_2$; a **Forward-Starting Option** is an exotic derivative [16, 67], depending on an underlying asset S_t, which contract is defined at time $t = 0$, with maturity T_2 and strike price $K = kS_{T_1}$. Thus, the pay-off at maturity of a forward-starting call option $C_{FS}(T_2)$ is:

$$C_{FS}(T_2) = \max(S_{T_2} - kS_{T_1}, 0)\,, \tag{4.2}$$

where k is the so called "moneyness" of the option. We see that Forward starting options are plain vanilla contracts where the strike price is unknown at the *starting date* $t = 0$ but it is defined at a future *determination date* T_1 proportionally to the asset value S_{T_1}. Therefore the value of this contract at $t = T_1$ is equal to the price of a plain vanilla call option, with strike price $K = kS_{T_1}$ and time to maturity $T_2 - T_1$:

$$C_{FS}(T_1) = C(S_{T_1}, kS_{T_1}, T_2 - T_1)\,.$$

In the classical Black-Scholes framework we have an explicit formula (9) for such contract:

$$\begin{aligned} C_{FS}(T_1) &= C_{BS}(S_{T_1}, 0, kS_{T_1}, T_2 - T_1, r, \sigma)\,, \\ &= S_{T_1}\, C_{BS}(1, 0, k, T_2 - T_1, r, \sigma)\,. \end{aligned}$$

Where we use the fact that the Black-Scholes formula is linear in the spot value if the ratio k between the spot and the strike is constant. Using now the forward value for the asset $S_{T_1} = S_0 e^{rT_1}$ and considering the discounting factor e^{-rT_1} we finally obtain the following formula for the Black-Scholes value of a Forward-Starting Call option:

$$C_{FS}(0) = S_0\, C_{BS}(1, 0, k, T_2 - T_1, r, \sigma)\,. \tag{4.3}$$

At this point, following [2], we can define the **Forward Implied Volatility** $\sigma_F(k, T_1, T_2)$ for a given price $C_{FS}(k, T_1, T_2)$ of Forward-Starting option as the number that used in the formula (4.3) above, makes the price equal to the Black-Scholes'one:

$$C_{FS}(k, T_1, T_2) = S_0\, C_{BS}\big(1, 0, k, T_2 - T_1, r, \sigma_F(k, T_1, T_2)\big)\,. \tag{4.4}$$

We notice that in the same way as Forward Rates become Spot Rates when the time of contracting coincides with the start of the effectiveness time of the interest [7] so does Forward Implied Volatility with Implied Volatility:

$$\sigma_I(kS_0, T) = \sigma_F(k, 0, T)\,.$$

Forward Volatility is a measure of future Implied Volatility [33], which, differently from present Implied Volatility, is not observable on the market and therefore unknown. However it is particularly critical for the pricing of path dependent exotic options like barrier options and cliquet [30] since these kind of instruments are very sensitive to future values of volatility.

The market of Forward-Starting options is not as liquid as the one of Plain Vanillas. Thus, it is not easy to calibrate a model with this kind of instruments. However we expect that σ_F exhibits similar properties than σ_I, such as the presence of smiles, and we want a model able to reproduce this kind of features. For this reason We are now interested in the Forward Volatility implied by the three models we have considered, namely σ_F^{LVM}, σ_F^{SVM}, σ_F^{LSVM} In order to reconstruct these volatilities, we act in the same way we previously did for the Implied Volatility: we first price Forward-Starting option with the different models and then, inverting the Black-Scholes formula we retrieve σ_F.

Given the path dependent nature of Forward-Starting option a natural choice for their pricing is to use a Monte Carlo simulation. In particular, for the Heston-Dupire model we simulate the dynamics (3.2) with a Milstein discretization [30, 62]. Thus, we introduce a suitable time discretization t^n and starting from $X_0 = 0$, $Y_0 = V_0$, $\forall n \geq 0$ we simulate the following paths:

$$X_{n+1} = X_n + \big(r - \frac{1}{2}\,Y_n\,\sigma^2_{LSV}(X_n, t^n)\big)\Delta t^n + \sqrt{Y_n}\sigma_{LSV}(X_n, t^n)\Delta W_1^n\,, \quad (4.5)$$

$$Y_{n+1} = \kappa(\theta - Y_n)\Delta t^n + \Big(\sqrt{V_n} + \frac{\eta}{2}\Delta W_2^n\Big)^2 - \frac{\eta^2}{4}\Delta t^n\,, \quad (4.6)$$

where:

$$\Delta t^n = t^{n+1} - t^n\,, \quad (4.7)$$

$$\Delta W_1^n = \sqrt{\Delta t^n}\,Z_1\,, \quad (4.8)$$

$$\Delta W_2^n = \sqrt{\Delta t^n}\,\Big(\rho\,Z_1 + \sqrt{1-\rho^2}\,Z_2\Big)\,, \quad (4.9)$$

$$\text{and } Z_1, Z_2 \sim \mathcal{N}(0,1)\,. \quad (4.10)$$

Analogous schemes have been used to simulate the Dupire Local Volatility model and the Heston Stochastic Volatility model. In particular we performed 200000 simulations with 20 time steps between each consecutive maturity. Then we priced, with the three models, several Forward Starting contracts with different determination date T_1 and maturity T_2. From these prices, applying formula

(4.4) it is possible to obtain the Forward Volatilities Implied by the three models: σ_F^{LVM}, σ_F^{SVM}, σ_F^{LSVM}. These are represented in Figure 4.14.

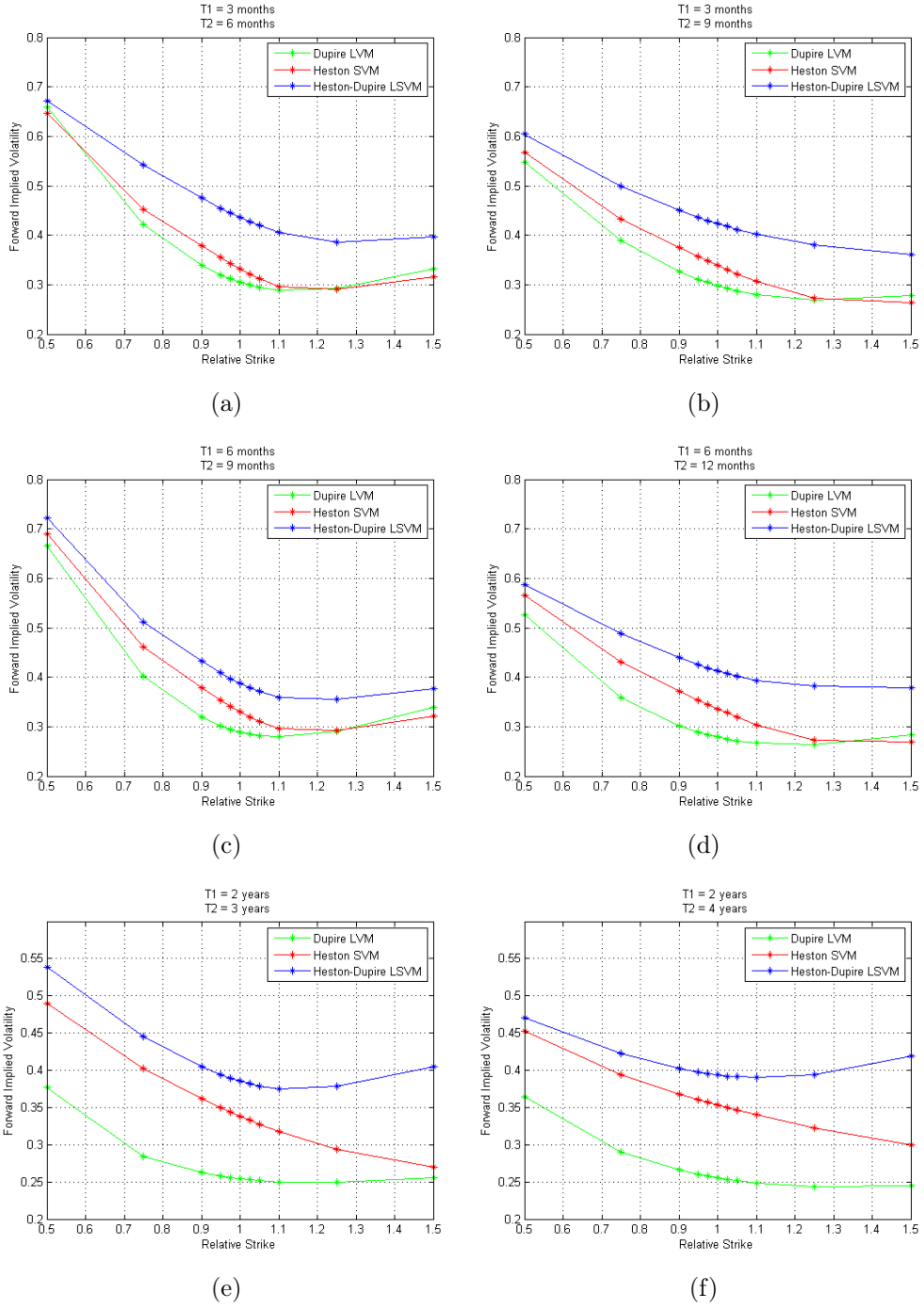


Figure 4.14: Comparison between the Forward Volatility of the different models σ_F^{LVM}, σ_F^{SVM}, σ_F^{LSVM}, for different determination dates T_1 and maturities T_2 .

We see that for close maturities the three models behaves in a similar way. However for longer time to start T_1 of the contracts the Forward Volatility implied by the Local Volatility model σ_F^{LVM} tends to be rather flat while the Stochastic Volatility model and moreover the Local Stochastic Volatility model succeed in reproducing realistic Forward Volatility.

We conclude this section about the numerical simulation of the Heston-Dupire Local-Stochastic Volatility model confirming that this generalized model is able to catch the advantages of both the two basic models. It succeeds in reproducing a realistic statics and dynamics of the Implied Volatility.

Concluding remarks

The aim of this work was to present and validate a new class of Local-Stochastic Volatility models for equity derivatives pricing. Depending on the particular choice of Stochastic dynamics and Local Volatility surface, different LSV models can be obtained. We have considered in details a Heston-Dupire model, where the Heston process is used for the Stochastic dynamics and the Dupire formula, combined with the Gatheral's SVI parametrization, is used for the Local Volatility. In order to calibrate the model, a Finite Element approximation of the nonlinear Fokker-Planck equation has been developed. This LSV model proved to be very interesting since it succeeds in combining the advantages of both the basic models: the good fitting of present market Implied Volatilities of the LVM and the realistic Forward Volatility implied by the SVM.

Unfortunately the improved features of the generalized model are obtained with an additional and not negligible computational cost since the solution of the nonlinear PIDE (3.16) is rather demanding. Moreover we have certain limitations about the input parameters. In fact, it is important to use a smooth Local Volatility surface and suitable Heston parameters which satisfy the Feller condition. The former condition is a natural one also for the Local Volatility model itself but the latter one is quite strict. Indeed, a good calibration of the Heston model often violates the Feller condition.

For these reasons I personally believe that the Local-Stochastic framework for equity volatility modelling is very promising but, for the moment, too much complex and computationally demanding for practical applications at the trading desks of investment banks. Thus an interesting solution would be to look for a simplified Local-Stochastic Volatility model. A first idea, already suggested by [61, 68] is to use some parametric form of the Local Stochastic Volatility surface $\sigma_{LSV}(s,t)$. In this way we do not need any more to solve a nonlinear Fokker-Planck equation to merge consistently the two basic models. On the other hand, with this approach we need to calibrate at the same time $\sigma_{LSV}(s,t)$ and the parameters of the stochastic dynamics. Moreover each price evaluation requires the solution of a bidimensional parabolic PDE. Although this is a linear problem, much easier than the nonlinear equation we have considered so far, we need to solve it several times to effectively calibrate the model with a minimization algorithm. Therefore, this approach could turn out to be even less efficient than

the solution of a nonlinear equation which we need to solve just once.

The very interesting feature of the method we have presented is that we calibrate, independently and possibly in parallel, the two basic models. The calibration of the Gatheral-Dupire LVM and the Heston SVM we considered, are, indeed, quite fast. The bottleneck of the procedure is the solution of the nonlinear PIDE that we need to merge consistently the two models, in order to find $\sigma_{LSV} = \sqrt{I[p]}\,\sigma_{LV}$. For this reason, an interesting strand of research could be to look for an approximate form of the nonlinear term $I[p](x,t)$, instead of solving the Fokker-Planck equation. This is not a trivial task. A first, naive approach would be to set simply $\sigma_{LSV} = \sigma_{LV}$ or $\sigma_{LSV} = \sqrt{I_0}\,\sigma_{LV}$ with $I_0 = \frac{1}{\mathbb{E}[V_t]|_{t\to\infty}} = \frac{1}{\theta}$ but this approach does not work. In fact, we know that the inverse conditional expectation function $I[p](x,t)$ is greatly variable 4.9(b) and approximating it with a constant it is a too strong approximation. However some less trivial possibilities could be explored and for this purpose this thesis work can be a useful benchmark.

Appendix A

Market Data

Here we present the market data used for the calibration of the model. The table below shows the market implied volatilities, at different strike prices and maturity, of European call options quoted on the *EuroStoxx50* Index at 1^{st} June 2012 as retrieved on Bloomberg. In particular, for the numerical simulation of Section 4.2 we have considered maturities up till 2 years.

SX5E Index | 90) Asset | 91) Actions | 92) Settings | Volatility Surface

Euro Stoxx 50 Pr | 2068.66 EUR | Bloomberg | Mid | As of 01-Jun-2012 | 18:35

1) Vol Table | 2) 3D Surface | 3) Term Analysis | 4) Skew Analysis | 5) Dividends | 6) Prices

Moneyness | ○ Listed Exp | ◉ Tenors | Fwd | Dates | Strikes

Exp	50.0%	75.0%	90.0%	95.0%	97.5%	100.0%	102.5%	105.0%	110.0%	125.0%	150.0%
	1034.3	1551.5	1861.8	1965.2	2016.9	2068.7	2120.4	2172.1	2275.5	2585.8	3103.0
1W	88.48	62.15	43.59	37.20	34.37	31.81	29.81	29.40	32.05	41.35	51.72
1M	64.07	50.77	39.00	35.83	34.41	33.19	32.18	31.34	29.60	27.62	36.73
2M	62.72	48.11	38.45	35.60	34.25	32.97	31.81	30.77	29.05	25.39	29.16
3M	62.01	46.38	37.58	34.91	33.63	32.42	31.26	30.17	28.24	25.20	25.73
6M	56.65	42.40	35.44	33.36	32.38	31.44	30.54	29.68	28.07	24.75	23.77
9M	52.35	40.08	34.18	32.41	31.58	30.78	30.02	29.29	27.91	24.71	23.04
1Y	48.43	37.99	32.80	31.30	30.60	29.92	29.27	28.64	27.44	24.59	22.80
18M	44.49	35.66	31.41	30.18	29.61	29.06	28.53	28.02	27.07	24.71	22.57
2Y	41.94	34.19	30.47	29.37	28.86	28.36	27.88	27.42	26.57	24.47	22.45
3Y	38.57	32.27	29.51	28.70	28.32	27.94	27.58	27.23	26.57	24.86	23.03
4Y	37.21	31.60	29.17	28.43	28.08	27.73	27.39	27.05	26.41	24.69	22.71
5Y	35.95	30.99	28.79	28.12	27.80	27.49	27.18	26.87	26.29	24.74	22.95
7Y	33.77	29.72	27.88	27.32	27.05	26.78	26.53	26.27	25.79	24.50	22.93
10Y	31.47	28.20	26.67	26.21	25.98	25.77	25.55	25.35	24.95	23.89	22.55

Table A.1: Market Implied Volatilities on the *EuroStoxx50* Index at 1^{st} June 2012. Source: Bloomberg®

Appendix B

Source Code

In this appendix we provide an excerpt of the code implemented in FreeFem++ 3.12, for the solution of the nonlinear Fokker-Planck equation (3.3). The numerical method has been described in Section 3.3. In this Appendix we present some parts of the basic, not optimized code, in order to show some details of the implemented algorithm.

B.1 Initialization

The first step, of course, is to set all the parameters of the problem:

```
// Domain Dimension
real xlow   = -3.0;     // Left  boundary
real xhigh  =  3.0;     // Right boundary
real yhigh  =  2.0;     // Upper boundary

// Initial condition
real y0     = 0.1377;   // Stochastic volatility value at time zero
real sigmax = 0.02;     // Initial uncertainty on x
real sigmay = 0.02;     // Initial uncertainty on y

// Heston Parameters
real r      =  0.0100;  // Risk free interest rate
real theta  =  0.2262;  // Reference volatility
real kappa  =  2.4047;  // Reversion speed
real eta    =  0.7802;  // Volatility of volatility
real rho    = -0.8189;  // Correlation

// Number of Mesh Elements
int N0      = 50;       // Element around the initial value y0
int N1      = 100;      // Element around the final solution
int N2      = 50;       // Element around the boundaries
```

```
// Temporal Grid Parameters
int Nfixings  = 9;                  // Number of fixing dates
real[int]  fixingdates(Nfixings);   // Fixing dates vector
real[int]  timesteps(Nfixings);     // Time steps between fixing dates

// Fixing dates                     // Timesteps between the fixings
fixingdates(0)  = 7.0/365.0;        timesteps(0) = 10; // 1 Week
fixingdates(1)  = 1.0/12.0 ;        timesteps(1) = 10; // 1 Month
fixingdates(2)  = 2.0/12.0 ;        timesteps(2) = 10; // 2 Month
fixingdates(3)  = 3.0/12.0 ;        timesteps(3) = 10; // 3 Month
fixingdates(4)  = 6.0/12.0 ;        timesteps(4) = 10; // 6 Month
fixingdates(5)  = 9.0/12.0 ;        timesteps(5) = 10; // 9 Month
fixingdates(6)  = 1.0;              timesteps(6) = 10; // 1 Year
fixingdates(7)  = 1.5;              timesteps(7) = 20; // 1.5 Year
fixingdates(8)  = 2.0;              timesteps(8) = 20; // 2 Year

// Fixed Point Iterations
int   Kmax = 50;                    // Maximum number of fixed point iteration
real  toll = 000001;                // Maximum tolerance
```

Then we set the temporal grid.

```
real    dt;                           // Local time step
int     Ntimesteps = 0;               // Total number of time steps

for( k = 0; k < Nfixings; k++)        // Total number of time steps
    Ntimesteps = Ntimesteps + timesteps(k);

real[int]   time(Ntimesteps+2);

dt = fixingdates(0)/timesteps(0);     // Initialization of the time vector
for( n = 0; n < timesteps(0); n++){
    time(n) = n*dt;
    for( k = 0; k < Nfixings-1; k++){
        dt = (fixingdates(k+1) - fixingdates(k) )/timesteps(k+1);
        for( i = 0; i < timesteps(k+1); i++){
            time(n) = fixingdates(k) + i*dt;
            n++;
}}}

time(Ntimesteps)    = fixingdates(Nfixings-1);
time(Ntimesteps+1)  = fixingdates(Nfixings-1) + dt;
```

And we construct the mesh.

```
border l0(t=0,1){ x = 0   ; y = 3.0*y0*t ; };
border l1(t=0,1){ x = alphag*t   ; y = 0 ; };
border l2(t=0,1){ x = alphag + t*(xhigh - alphag)  ; y = 0 ; };
border l3(t=0,1){ x = xhigh ;     ; y = yhigh*t  ; };
border l4(t=0,1){ x = xhigh - t*(xhigh - xlow) ; y = yhigh  ; };
border l5(t=0,1){ x = xlow     ; y = yhigh*(1-t)  ; };
border l6(t=0,1){ x = xlow - t*(alphag + xlow)  ; y = 0 ; };
border l7(t=0,1){ x = - alphag*(1-t)  ; y = 0 ; };

mesh Th = buildmesh(  l0(N0)     + l1(N1) + l2(1.5*N2) + l3(N2)
                    + l6(1.5*N2) + l5(N2) + l4(2*N2)   + l7(N1) );
```

At this point we define the Finite Element spaces and the stiffness matrices.

```
// Finite Element Space
fespace Vh(Th,P1);      int dimVh = Vh.ndof;
fespace Nh(Th,P0);      int dimNh = Nh.ndof;

// Finite Element functions declaration
Vh u, uold, utemp, uend, v;  // Solutions finite element variables
Vh LSV2, LSV, LVol;          // Volatility surfaces
Vh D11, D12, D21, D22;       // Diffusion coefficients
Vh a1, a2;                   // Advection coefficients
Vh err, res;                 // Error and residual

// Bilinear Variational Forms Definition
D11 = 0.5*y*LSV2;
D12 = 0.5*y*rho*eta*LSV;
D21 = 0.5*y*rho*eta*LSV;
D22 = 0.5*y*eta*eta;

a1  = - r + 0.5*eta*rho*LSV + 0.5*y*( LSV2 + dx(LSV2) ) ;
a2  = - kappa*theta + 0.5*eta*eta + y*( kappa + 0.5*rho*eta*dx(LSV));

varf mass(u,v)  = int2d(Th)(u*v);
varf diff(u,v)  = int2d(Th)(   D11*dx(u)*dx(v) + D12*dy(u)*dx(v)
                             + D21*dx(u)*dy(v) + D22*dy(u)*dy(v) );
varf advec(u,v) = int2d(Th)( a1*u*dx(v) + a2*u*dy(v) );
varf rhs(u,v)   = int2d(Th)(uold*v) + on(l1,l2,l3,l4,l5,l6,l7, u = 0);

// FE Matrix Definition
real[int] ff(dimVh);       // RHS vector
real[int] uu(dimVh);       // Solution vector
matrix M = mass(Vh,Vh);    // mass matrix (it is the only time-independent)
matrix D;                  // diffusion matrix
```

```
matrix A;                   // advection matrix
matrix C, UP;               //Auxiliary matrix

real[int] LVOLvec(dimVh);
real[int] LSVvec(dimVh);
real[int] LSV2vec(dimVh);

// Streamline-Upwind Diffusion Numerical Stabilization
Nh   tau         = 0.3*hTriangle/sqrt( a1^2 + a2^2);
varf upwind(u,v) = int2d(Th)( tau*( a1*dx(u) + a2*dy(u))*(a1*dx(v) + a2*dy(v)));
UP               = upwind(Vh,Vh);
```

B.2 Auxiliary Functions

Then we need to initialize the Dupire local volatility surface $\sigma_{LV}(x,t)$, from the fitted SVI Implied Volatility $\sigma_I^{SVI}(x,t)$ given the calibrated coefficients (4.2). We do not show here all the technical details. We just provide the function which evaluates the Local Volatility:

```
func real Lvol( real x0, real T){

   SigmaI1 = sqrt( C1  + C2 *(C3* (x0 - C4 ) + sqrt((x0 - C4 )^2.0 + C5 ) ) );
   SigmaI2 = sqrt( C1p + C2p*(C3p*(x0 - C4p) + sqrt((x0 - C4p)^2.0 + C5p ) ) );
   f1 = 0.5*C2 /SigmaI1*( C3  + 2.0*(x0 - C4 )/sqrt((x0 - C4 )^2.0 + C5 ) );
   f2 = 0.5*C2p/SigmaI2*( C3p + 2.0*(x0 - C4p)/sqrt((x0 - C4p)^2.0 + C5p) );
   g1 = 1.0/SigmaI1*( C2*C5  / ( ( (x0 - C4 )^2.0 + C5  )^1.5 ) - f1 );
   g2 = 1.0/SigmaI2*( C2p*C5p/ ( ( (x0 - C4p)^2.0 + C5p )^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^2.0 + 2.0*T*SigmaI*( SigmaT + r*SigmaX );
   Den = ( 1.0 + d1*sqrt(T)*SigmaX )^2.0
          + SigmaI*T*( SigmaXX - SigmaX - d1*sqrt(T)*SigmaX^2.0 );
   Lvol = sqrt( ( epsNum + Num) / ( epsDen + Den) );
   return Lvol;
}
```

Then we define the function for the evaluation of the inverse conditional expectation. For the sake of clarity this is a simplified, non adapted and not optimized version of the one actually implemented.

```
func real InvCondExp1(real x0){

    num = 0;
    den = 0;

    for(int i=0; i<Nint; i++){          // Rectangle quadrature formula
        ui   = utemp(x0, i*deltay);
        num += deltay*ui;
        den += deltay*(i*deltay)*ui;
    }
    return    ( num + epsInt) /(den + tunInt*epsInt);
};
```

B.3 Problem Solution

At this point we can initialize the solution and solve the problem. The core of the algorithm is essentially constituted by a double loop, the outer for the time evolution and the inner for the solution of the non linearity with fixed point iterations.

```
// Initial condition, approximation of the Dirac Delta
uold  = 1/(2*pi*sigmax*sigmay) * exp( - 0.5*( x / sigmax )^2 )
                               *(  exp( - 0.5*( (y - y0) / sigmay )^2 )
                                 - exp( - 0.5*( (y + y0) / sigmay )^2 ) );

utemp = uold;

// Initial Conditional Expectation
for( i = 0; i < dimVh; i++){
    xv = Th(i).x;
    yv = Th(i).y;
    LVOLvec(i)  = Lvol(xv);
    InvCondExpVec(i) = 1/y0;
    LSV2vec(i)   = (LVOLvec(i))^2*InvCondExpVec(i);
    LSVvec(i)   = sqrt(LSV2vec(i));
}

LVol[] = LVOLvec;
LSV[]  = LSVvec;
LSV2[] = LSV2vec;
```

```
// Time Loop
for( n = 0; n < Ntimesteps+1; n++){

    k  = 0;
    dt = time(n+1) - time(n);
    relativerror = 1.0;

    // ... Updating the Dupire Local Volatility LVol

    // Updating the right hand side
    varf rhs2(u,v)  = int2d(Th)(uold*v/dt) + on(l1,l2,l3,l4,l5,l6,l7, u = 0);
    ff = rhs2(0,Vh);

    // Fixed Point Loop
    while( k < Kmax && relativerror > toll ){

        D11 = 0.5*y*LSV2;
        D12 = 0.5*y*rho*eta*LSV;
        D21 = 0.5*y*rho*eta*LSV;
        D22 = 0.5*y*eta*eta;
        a1  = - r + 0.5*rho*eta*LSV + 0.5*y*( LSV2 + dx(LSV2) ) ;
        a2  = - kappa*theta + 0.5*eta*eta + y*( kappa + 0.5*rho*eta*dx(LSV) );

        D  =  diff(Vh,Vh);
        A  =  advec(Vh,Vh);
        UP =  upwind(Vh,Vh);
        C  = (1/dt)*M + D + A + UP;
        set(C, solver = UMFPACK);

        uu = C^-1*ff;
        uuold = uu;
        u[]   = uu;

        // Computing the Error
        err    = u - utemp;
        L2err  = int2d(Th)( err*err );
        H1err  = int2d(Th)( (dx(err))^2 + (dy(err))^2 ) + L2err;
        L2u    = int2d(Th)( utemp*utemp );
        H1u     = int2d(Th)( (dx(u))^2   + (dy(err))^2 ) + L2u;

        relativerror = H1err/H1u;

        utemp = u;
```

```
        // Computing the Inverse Conditional Expectation
        for( j = 0; j < dimVh; j++){
            xv = Th(j).x;
            InvCondExpVec(j) = InvCondExp2(xv);
            LSV2vec(j)   = ( LVOLvec(j) )^2*InvCondExpVec(j);
            LSVvec(j)   = sqrt(LSV2vec(j));
        }

        LSV[]   = LSVvec;
        LSV2[]   = LSV2vec;

        k++;
    }

    // Update the solution
    uold = u;

}
```

List of Figures

List of Tables

Bibliography

[1] Abergel, F. & Tachet, R. A nonlinear partial integro-differential equation from mathematical finance. *Discrete and Continuos Dynamical Systems, Series A*, 27(3):907–917, 2010.

[2] Albanese, C., H. Lo, & A. Mijatovic. Spectral methods for volatility derivatives. *Quantitative Finance*, 9(6):663–692, 2009.

[3] Alexander, C. & Nogueira, L. Stochastic local volatility. Working Paper, Available Online, 2008.

[4] Alvarez, O. & Tourin, A. Viscosity solutions of nonlinear integro-differential equations. *Ann. Inst. Henri Poincaré*, 13(3), 1996.

[5] Atlan. Localizing volatility. Working Paper, Available Online, 2006.

[6] Bandini A. Problemi ellittico-parabolici degeneri al bordo e buona posizione del problema derivante dal modello di heston. Master's thesis, Politecnico di Milano, 2010.

[7] Björk, T. *Arbitrage theory in continuos time*. Oxfor University Press, 2004.

[8] Bouchouev, I. & Isakov, V. Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. *Inverse Problems*, 15:95–116, 1999.

[9] Breeden, D. & Litzenberger, R. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51:621–651, 1978.

[10] Broadie, M. & Kaya, O. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 54(2):217–231, 2006.

[11] Carr, P. & Madan, D. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2:61–73, 1998.

[12] Carr, P., Ellis, K. & Madan, D. Static hedging of exotic options, journal of finance. *Journal of Finance*, 53(3):1165–1190, 1998.

[13] Coleman, T., Li, Y. & Verma, A. Reconstructing the unknown volatility function. *Journal of Computational Finance*, 3:77–102, 1999.

[14] Cox, J., C., Ingersoll, J., E. & Ross, S. A theory of the term structure of interest rate. *Econometrica*, 54:373–384, 1985.

[15] Daskalopoulos, P. & Feehan, P., M., N. . Existence, uniqueness, and global regularity for variational inequalities and obstacle problems for degenerate elliptic partial differential operators in mathematical finance. 2011.

[16] de Weert, F. *Exotic Options Trading.* The Wiley Finance Series. Wiley, 2011.

[17] Derman, E. & Kani, I. Riding on a smile. *Risk Magazine*, 7(2):32–39, 1994.

[18] Derman, E. & Kani, I. Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility. *International Journal of Theoretical and Applied Finance*, 1(1):61–110, 1998.

[19] Derman, E., Ergener, D. & Kani, I. Static options replication. *Journal of Derivatives*, 2(4):78–95, 1995.

[20] Dragulescu, A., A. & Yakovenko, V., M. Probability distribution of returns in the heston model with stochastic volatility. *Quant. Finance*, 2(6):443–453, 2002.

[21] Dupire, B. Pricing with a smile. *Risk Magazine*, 7:18–20, 1994.

[22] Dupire, B. A unified theory of volatility. In P. Carr, editor, *Derivatives Pricing: The Classic Collection.* Risk Books, 2004.

[23] Engelmann, B., Koster, F., & Oeltz, D. Calibration of the heston stochastic local volatility model: A finite volume scheme. Working Paper, Available Online, 2012.

[24] Ewald, C., O. Local volatility in the heston model: A malliavin calculus approach. *Journal of Applied Mathematics and Stochastic Analysis*, 3:307–322, 2005.

[25] Feller, W. Two singular diffusion problems. *Ann. Math.*, 54(173), 1951.

[26] Fengler, M. R. Arbitrage-free smoothing of the implied volatility surface. Discussion paper, 2005.

[27] Fengler, M. R. *Semiparametric modelling of implied volatility*. Lecture Notes. Springer, 2005.

[28] Forde, M., Jacquier, A. & Mijatović, A. Asymptotic formulae for implied volatility under the heston model. Working paper, 2009.

[29] Gatheral, J. A parsimonious arbitrage-free implied volatility parametrization with application to the valuation of volatility derivatives. Presentation at Global Derivatives, 2004.

[30] Gatheral, J. *The Volatility Surface: a Practitioner's Guide*. Wiley, 2006.

[31] Gatheral, J. & Jacquier, A. Arbitrage-free svi volatility surfaces. Working Paper, 2012.

[32] Gerhold, S. Fast calibration in the heston model. Master's thesis, Technischen Universität Wien, 2012.

[33] Glasserman, A. & Wu, Q. Forward and future implied volatility. *International Journal of Theoretical and Applied Finance*, 14(03):407–432, 2011.

[34] Griva, I., Nash, S., G. & Sofer, A. *Linear and Nonlinear Optimization*. Cambridge University Press, 2009.

[35] Guyon, J. & Henry-Labordère, P. Being particular about calibration. *Risk Magazine*, pages 92–97, 2012.

[36] Gyöngy, I. Mimicking the one-dimensional marginal distributions of processes having an ito differential. *Probability Theory and Related Field*, 71:501–516, 1986.

[37] Hagan, P. & Woodward, D. Equivalent black volatilities. *App. Math. Finance*, 6, 147-157 1999.

[38] Hagan, P., Kumar, D., Lesniewski, A. & Woodward, D. Managing smile risk. *Wilmott magazine*, 1, 84-108 2002.

[39] Hecht, F. & Pironneau, O. Freefem++ manual. www.freefem.org/ff++/, 2011.

[40] Henry-Labordère, P. Calibration of local stochastic volatility models to market smiles. *Risk Magazine*, 20:112–117, 2009.

[41] Heston, S. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.*, 6:327–343, 1993.

[42] Hilber, N., Matache, A., M. & Schwab, C. Sparse wavelet methods for option pricing under stochastic volatility. *J. Comp. Finance*, 8(4):1–42, 2005.

[43] Homescu, C. Implied volatility surface: Construction methodologies and characteristics. Technical report, Wells Fargo Securities, 2011.

[44] Hull, J. & White, A. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2):281–300, 1987.

[45] Jex, M., Henderson, R., & Wang, D. Pricing exotics under the smile. *Risk Magazine*, 12(11):72–75, 1999.

[46] Kruse, S. On the pricing of forward starting options under stochastic volatility. *Berichte des Fraunhofer ITWM*, 53, 2003. Fraunhofer Institut Techno und Wirtschaftsmathematik, Kaiserslautern.

[47] Lagnado, R. & Osher, S. A technique for calibrating derivative security pricing models: numerical solution of an inverse problem. *Journal of Computational Finance*, 1:13–25, 1997.

[48] Lee, R. The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, 14(3):469–480, 2004.

[49] Lipton, A. The vol smile problem. *Risk Magazine*, pages 61–65, 2002.

[50] Merton, R.C. Option pricing when the underlying stock returns are discontinuous. *J. Financ. Econ.*, 3:125–144, 1976.

[51] Oksendal, B. K. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, 6th edition, 2010.

[52] Piterbarg, Y. Markovian projection for volatility calibration. *Risk Magazine*, 20:84–89, 2007.

[53] Quarteroni, A. *Modellistica Numerica per Problemi Differenziali*. Springer-Verlag, 2008.

[54] Quarteroni, A. & Valli, A. *Numerical Approximation of Partial Differential Equations*. Springer, 1994.

[55] Quarteroni, A., Sacco, R. & Saleri, F. *Matematica Numerica*. Springer-Verlag, 2008.

[56] Rebonato, R. *Volatility and Correlation: The Perfect Hedger and the Fox*. John Wiley & Sons, 2005.

[57] Ren, Y., Madan, D. & Qian, M., Q. Calibrating and pricing with embedded local volatility models. *Risk Magazine*, 20:138–143, 2007.

[58] Risken,H. *The Fokker-Planck Equation - Methods of solution and Applications*. Springer, 1996.

[59] Salsa, S. *Equazioni a Derivate Parziali - Metodi, modelli e applicazioni*. Springer-Verlag, 2010.

[60] Schmelzle, M. Option pricing formulae using fourier transforms: theory and application. Working Paper, 2010.

[61] Sepp, A. Stochastic local volatility models: Theory and implementation. Technical report, University of Leicester, 2010.

[62] Seydel, R., U. *Tools for Computational Finance*. Springer-Verlag, 4th edition, 2008.

[63] Shreve, S., E. *Stochastic Calculus for Finance 2: Continuous-time Models*. Springer-Verlag, 2004.

[64] Stein, E., M. & Stein, J., C. Stock price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies*, 4:727–752, 1991.

[65] Tachet, R. *Non parametric model calibration in finance*. PhD thesis, Ecole Centrale Paris, October 2011.

[66] Van Haastrecht, A. & Pelsser, A., J. Efficient, almost exact simulation of the heston stochastic volatility model. *International Journal of Theoretical and Applied Finance*, 31(1), 1-43 2008.

[67] Wilmott, P. *Paul Wilmott on Quantitative Finance*. John Wiley & Sons, 2006.

[68] Wystup, U. Validation of the tremor stochastic-local-volatility model. Technical report, MathFinance AG, 2011.

[69] Zühlsdorff, C. The pricing of derivatives on assets with quadratic volatility. *Applied Mathematical Finance*, 8, 235-262 2001.