## IMPERIAL COLLEGE LONDON

### MENG FINAL YEAR THESIS

# Multiscale Models and Dimensionality reduction in the pricing and hedging of Path Dependent Financial Options

Author: Radu Baltean-Lugojan Supervisor: Dr. Panos Parpas

A thesis submitted in fulfilment of the requirements for the degree of Master of Engineering in the

Department of Computing



June 2013

## Abstract

Multiscale models address the main drawback in a one-factor stochastic volatility framework by providing a better fit to short maturity European-style options. In this project we focus on the complex considerations that a model needs to fulfill to prove practical. We build on the theoretical work of Fouque, Papanicolaou and Sircar, implementing calibrations for both first-order and second order perturbation expansions. This allows us to investigate properties such as goodness of fit, parameter stability and computational tractability of calibrations. We find that the first-order expansion with time-dependency shows decent fits and parameter stability for contracts that are not very close to expiry.

The second order quadratic fits provide more flexibility and improve results, but only at a high computational cost. We develop a hybrid calibration using first order expansion results and a parameter reduction technique which both prove successfull in reducing the dimension of the non-convex optimization problem. A combination of the two techniques allows us to derive a quasi-closed form solution to the second-order calibration, reducing its computational cost.

# A cknowledgements

I would like to thank Dr. Panos Parpas for agreeing to supervise my project and entrusting me with the selected topic, as well as his constant advice. A special thanks goes to my girfriend, Ioana Chivu, for providing me with the much needed support and encouragement in my last years at Imperial. Finally, I would equally like to thank my parents for the financial support and trust they have put in me.

# Contents

A	Abstract i					
A	cknov	wledge	ments	iii		
Li	st of	Figure	25	vii		
Li	ist of	Tables	3	x		
1	Intr	oducti	on	1		
	1.1	Motiva	ation and objectives	2		
	1.2	Contri	butions	2		
2	Bac	kgrour	nd	4		
	2.1	Black-	Scholes Pricing Framework	4		
		2.1.1	Probability Space and Stochastic Processes	4		
		2.1.2	Brownian Motion	5		
		2.1.3	Stochastic Integrals	5		
		2.1.4	Market Model and Risky Asset Price	6		
		2.1.5	Ito's Formula	7		
		2.1.6	Lognormal Risky Asset Price	7		
		2.1.7	European and Path Depedent options	7		
		2.1.8	Black-Scholes Partial Differential Equation and Formula	9		
		2.1.9	The Greeks	12		
		2.1.10	Risk-Neutral Pricing	12		
		2.1.11	Markov Processes and Corresponding Infinitisemal Generators and Mar-	19		
		9 1 19	Complete ve Incomplete Monkrete	15		
	<u></u>	2.1.12 Stocho	complete vs incomplete Markets	10 15		
	2.2		Historia Volatility	10		
		2.2.1	Instone volatility	10		
		2.2.2	Local Volatility	10		
		2.2.0 2.2.0	Local Volatility vs Stochastic Volatility	19 91		
		2.2.4 2 2 5	Stochastic Volatility Models	21 91		
		2.2.0 2.2.6	One-Factor Stochastic Volatility Models	21 99		
		2.2.0 2 2 7	Mean reversion-driven processes	22		
		2.2.1	mean reversion-univen processes	40		

		2.2.8 One-factor models pricing - PDE approach					
		2.2.9 One-factor models pricing - Martingale approach					
		2.2.10 Multi-Factor Stochastic Volatility Models					
	2.3	3 Related Work					
		2.3.1 (K,T,t) Problem and Benchmark For Implied Volatility Models 2					
		2.3.2 Model Selection					
		2.3.3 Summary					
	2.4	Volatility Time Scales					
		2.4.1 Mean Reversion and Markov Process Time Scales					
		2.4.2 Time scales observed in market data					
		2.4.3 Application to stochastic volatility models					
3	Firs	t order perturbation:					
	Out	line and Implementation Results 4					
	3.1	Model Setup					
	3.2	Derivation outline					
		3.2.1 Regular perturbation expansion (around $\delta$ )					
		3.2.2 Singular perturbation expansion (around $\epsilon$ )					
		3.2.3 Price expansion formula					
		3.2.4 Parameter reduction and expansion accuracy					
	3.3	Model Calibration					
	3.4	4 Data Preparation					
		3.4.1 Matlab routines					
		3.4.2 Data preprocessing $\ldots \ldots 5$					
	3.5	Results					
		3.5.1 Goodness of Fit					
		3.5.2 Parameter Time Stability					
	3.6	3.6 Adjusting For Maturity Cycles					
		3.6.1 Time-dependent Fast Scale					
		3.6.2 Time-dependent Function					
		3.6.3 Time-dependent Calibration					
	3.7	Time-dependent Results					
		3.7.1 Maturity cycles correction					
		3.7.2 Goodness of Fit					
		3.7.3 Parameter Time Stability					
	3.8	Optimizing $p$ parameter in $v(t)$					
4	Sec	and order perturbation:					
1	Out	line and Implementation Results 7					
	4.1	Derivation Outline					
	4.2	Calibration Equations					
	4.3	Calibration - First parameter group (10)					
	4.4	Second parameter group (18)					
		4.4.1 Global Optimization in Current Form					
		4.4.2 Integrating constraints in objective function					
		4.4.3 Hybrid Calibration with First-Order Results					
	4.5	Further 3 Parameter Reductions					

		4.5.1	Accuracy of Reduction	97
		4.5.2	Recovering the Original Parameters	98
		4.5.3	Concluding Remarks	98
	4.6	Deriva	tion of Quasi-Closed Form Solution	99
		4.6.1	Concluding remarks	101
5	Cor	nclusio	n	103
	5.1	Future	Work	104

### A Appendix A

### Bibliography

106

105

# List of Figures

2.1	Geometric Brownian motion for $\mu = 0.15, \sigma = 0.20, X_0 = 1$ . The mean path	
	(red), 20 sample paths for $t \in [0; 5]$ (green), and $1\%, 5\%, 10\%, 90\%, 95\%$ and	
	99% percentile paths (red dashed).	8
2.2	Discounted stock price = $xe^{r(T-t)} \approx 102$ = minimum vol	17
2.3	Vertical line shows at-the-money level (= stock price $x$ )	17
2.4	S&P 500 index options implied volatilities (June 1, 2007) 1	18
2.5	S&P 500 index options implied volatilities (September 15, 2005), [1] 1	18
2.6	S&P 500 index level and BlackScholes implied volatility each Wednesday dur-	
	ing the period of June 1988 through December 1993 2	20
2.7	The term structure of at-the-money(ATM) variance skew for various choices	
	of jump diffusion parameters	31
2.8	Simulated paths of $\sigma_t = Y_t$ , with $(Y_t)$ a mean-reverting CIR process (source: [2])	36
2 0	Simulated volatility and corresponding return paths for small and large rates	,0
2.0	of mean-reversion for the OU model (source: [2])	36
2 10	De-meaned Dow Jones Industrial Average (D.IIA) returns over years 1900-2000	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
2.10	[source: svol]	37
2.11	Empirical variogram of the S&P 500 (dashed line) and exponential fit (solid	
	line) from which $\kappa$ is found, for 1999-2000 data.(source: [2])	38
2.12	The first two figures show simulated paths of $D_n$ for $\kappa = 1$ and $\kappa = 250$ ,	
	respectively. The third figure shows S&P 500's normalized fluctuation process	
	from the first 60 trading days of 1994 (source: $[2]$ ) $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$	38
2.13	Intraday variability in 1999-2000 (dashed line is a 4-parameter model fit).	
	Notice that volatility is highest at the beginning of the trading day, with a	
	rebound at the end. [source: short time scale in $S\&P$ ] $\ldots \ldots \ldots 3$	39
91		51
ე.1 ვე	StrP 500 implied veletilities fitted as a function of LMMP on April 10, 2005	)1
3.2	S&F 500 implied volatilities littled as a function of LMMR on April 19, 2005. The circles represent option data for the data, and the line $L \sim b^* + a^{\epsilon} (LMMR)$	
	shows the fit from only a fast factor fit [2]	53
33	SkP 500 implied volatilities fitted as a function of LM $-\tau *$ LMMBon April	,0
0.0	19. 2005. The circles represent option data for the date, and the line $I \approx$	
	$\overline{\sigma} + b^{\delta} \tau + a^{\delta}$ ( <i>LM</i> ) shows the fit from only a slow factor fit. [2]	53
3.4	Term-structure fits on April 19, 2005. The circles are the slope coefficients $\hat{a}_i$	
	of LMMR fitted in the first step of the regression. The solid line is the straight	
	line $(a^{\epsilon} + a^{\delta}\tau)$ fitted in the second step of the regression. The estimates are	
	$(a^{\epsilon}, a^{\delta}) = (-0.0646, -0.1397) \dots \dots$	54

3.5	Term-structure fits on April 19, 2005. The plot shows the $\hat{a}_i$ counterparts, $\hat{b}_i$ slope coefficients for the skew intercepts $\hat{b}_i$ fitted to the straight line $b^* + b^{\delta}\tau$ . The estimates are $(b^{\delta}, b^*) = (0.0164, 0.1417)$	55
3.6	Calibration fit on S&P 500 data (circles) on April 19, 2005, based on fast and slow scale estimates $(a^{\epsilon} b^{\star} a^{\delta} b^{\delta})$ as calculated for figures 3.4 and 3.5	00
37	Average relative fitting error = $3.75\%$	55
0.1	of LMMR fitted in the first step of the regression. The solid line is the straight line $(a^{\epsilon} + a^{\delta}\tau)$ fitted in the second step of the regression.	55
3.8	Term-structure fits on July 26, 2006. The plot shows the $\hat{a}_i$ counterparts, $\hat{b}_i$ slope coefficients for the skew intercepts $\hat{b}_i$ fitted to the straight line $b^* + b^{\delta}\tau$ .	56
3.9	Calibration fit on S&P 500 data (circles) on July 26, 2006, based on fast and slow scale estimates $(a^{\epsilon}, b^{\star}, a^{\delta}, b^{\delta})$ . Average relative fitting error = 6.1%	56
3.10	Time series (2000-2012) of estimated parameters $(a^{\epsilon}, b^{\star}, a^{\delta}, b^{\delta})$ obtained through daily calibration according to a first-order price expansion	58
3.11	Time series (2000-2012) of estimated parameters $(\sigma^{\star}, V_0^{\delta}, V_1^{\delta}, V_3^{\epsilon})$ obtained through daily calibration according to a first-order price expansion	59
3.12	Fitted skew slopes for 2007, using first three maturities only. The ticks on the time axis correspond to option maturity (expiration) dates throughout 2007.	
	$\operatorname{Var}(a^{\epsilon}) = 0.000159$	60
3.13	$v^{-1}$ for $p = 1, 5, 50$ with a cutoff $\Delta t \approx 3$ days $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$	64
3.14	Fitted skew slopes with time-dependent fast scale for 2007, using first three maturities only. The ticks on the time axis correspond to option maturity (expiration) dates throughout 2007. Variance( $a^{\epsilon}$ ) = 0.000108 (-32% reduction compared to time-independent case)	66
3.15	Time dependent term-structure fit on S& P 500 options data from July 26.	00
0.10	2006. The top plot shows the fit of $\hat{a}$ against $\frac{\overline{v_{LT}^{1/2}}}{\tau}$ . The bottom plot shows the	
3.16	fit of $\hat{b}$ by plotting $\hat{b} - b^{\delta}\tau$ against $v_{t,T}^{1/2}$ Time-dependent first-order fits on S& P 500 options data from July 26, 2006.	68
	Improved average relative fitting error $= 3.76\%$ , due to better short maturity fits	68
3.17	Time series (2000-2012) of estimated parameters $(a^{\epsilon}, b^{\star}, a^{\delta}, b^{\delta})$ obtained through daily time-dependent calibration to a first-order price expansion	70
3.18	First-order fits on S& P 500 options data from March 18, 2010. Average relative fitting error = $6.53\%$	71
3.19	Time-dependent $(p = 1)$ first-order fits on S& P 500 options data from March 18, 2010. Average relative fitting error = $5.40\%$	72
3.20	Time-dependent $(p^* = 33)$ first-order fits on S& P 500 options data from March 18, 2010. Average relative fitting error = $3.35\%$	72
3.21		73
4.1		81
4.2	Average relative fit error $= 2.36\%$	82
4.3	Time series (2000-2012) of estimated parameters $(k, l, m, n)$ obtained through daily calibration according to a second-order price expansion	83
4.4	Time series (2000-2012) of estimated parameters $(p, q, s)$ obtained through daily calibration according to a second-order price expansion	84

4.5	Time series (2000-2012) of estimated parameters $(u, v, w)$ obtained through			
	daily calibration according to a second-order price expansion	85		
4.6	Output from running 'rbfSolve' global optimization algorithm	87		
4.7	Time series (2000-2011) of market parameters $(C_2^{\epsilon,\delta}, C_1^{\epsilon,\delta}, C_0^{\epsilon,\delta}, C^{\epsilon,\delta})$ obtained			
	through daily calibration according to a second-order price expansion	90		
4.8	Time series (2000-2011) of market parameters $(A_2^{\epsilon}, A_1^{\epsilon}, A_0^{\epsilon}, A^{\epsilon})$ obtained			
	through daily calibration according to a second-order price expansion	91		
4.9	Time series (2000-2011) of market parameters $(B_2^{\delta}, B_1^{\delta})$ obtained through			
	daily calibration according to a second-order price expansion	92		
4.10	Time series (2000-2011) of market parameters $\left(\frac{V_{1}^{\prime\epsilon}}{\overline{z}}, \frac{V_{1}^{\prime\delta}}{\overline{z}}, \frac{V_{0}^{\prime\delta}}{\overline{z}}, \phi^{\epsilon}\right)$ obtained			
	through daily calibration according to a second-order price expansion	93		

# List of Tables

2.1	from [2]	23
4.1	Properties of estimated parameters (first set)	82

For/Dedicated to/To my...

## Chapter 1

# Introduction

The first pricing model for financial options to provide a closed-form numerical solution, as introduced by Black and Scholes, makes an important assumption by fixing the volatility of the option underlying as a constant. Measuring the volatility (as implied by Black-Scholes) that matches market option prices shows a significant departure of markets reality from the assumptions of the model. Trying to find models consistent with the market implied volatility surface has spurred extensive research both in academia and in industry[Joshi]. However this remains a difficult task as evidenced by existing volatility models failing to fully capture both the characteristics and dynamics of the implied surface [Gatheral]. A good volatility model as detailed in [Related Work] needs to find the right balance between model parameter stability and goodness-of-fit properties. Practitioners most often focus on goodness of fit and resort to a pernicious practice of daily recalibration to adjust for data changes which leads to expensive transaction costs. Meanwhile, the existing literature often focuses on highly complex models that are computationally expensive to calibrate, easily overfit the data [Gatheral] and are not applicable to a wide range of financial derivatives. Furthermore a lack of closed-form numerical solutions impedes the use of computationally tractable methods to calibrate such models to market data in a trivial manner.

More recently, novel techniques were developed in [2] to introduce perturbation expansions to model volatility as a stochastic process on different time scales. Basic stochastic volatility models represent the volatility variation on a single time scale, but fail to capture the skew characteristics of implied volatility. Empirical studies dismiss a deterministic volatility function [Dumas] and show multiple time scales (or dimensions) of volatility variation ([short time scale]). [2] builds a general class of multidimensional stochastic volatility models with multiple extensions and applications, that shows a great improvement in capturing the skew effect. The addition of fast and slow volatility time scales introduces a group of parameters that are easy to calibrate to market and provide greater model flexibility, whilst providing a pricing tool for more exotic options by marking them to liquid option market prices via the implied volatility surface.

### **1.1** Motivation and objectives

The project aims to build on the current multiscale models research in the main reference book [2] at both the first and second order perturbation expansion levels. We seek to build working implementations of the models and investigate the behaviour and limitations of the first-order theory in terms of goodness of fit and parameter stability. Further research scope is given by adapting the first-order expansion to maturity cycles in the presence of a time-dependent fast volatility time scale. For calibration purposes, S&P 500 European options implied volatility data is used, allowing us to compare results directly with our references. Furthermore we aim to contrast results with those obtained through a second order expansion. In the context of the second order theory, a main focus of the project is to understand the calibration complexity and explore dimensionality reduction schemes guided by empirical and theoretical considerations. This is motivated by the explosion of parameter dimensions in the second order theory, which severely impacts the computational aspects of the calibration procedure. Ultimately, we want to determine if multiscale stochastic volatility models are capable of being used by practitioners, and how their calibration can be adjusted to fit that purpose.

### **1.2** Contributions

This project provides the following noteworthy contributions:

- A full Matlab implementation with functions for fast data handling and processing, as well as custom procedures for particular calibration variants considered. The scripts can be either adapted to an industry practitioner environment or be used for direct further experimentation and research with calibration techniques or in pricing applications.
- An extended treatment of maturity cycles in the context of the first-order expansion theory. By optimizing the time-dependency model, severe limitations of the first-order expansion have been found in close proximity to option expiry dates (up to 3 days). For the remainder of the data, empirical evidence showcases the robustness of the model, even compared to the second-order expansion theory.
- A parameter reduction step has been introduced in the second order theory calibration. This reduces the dimensionality of a global optimization problem present in calibration.

The solution is an accurate approximation to the optimal set of parameters, given the regime of the perturbation expansions.

- A hybrid calibration technique for the second order expansion has been tested with moderate success. We take advantage of the first-order calibration results and use them to reduce the dimensionality of the second order calibration.
- A combination of the previous two extensions has lead to a quasi-closed form solution that reduces the dimensionality of the second step in the second order calibration from 18 to 3. The approach break down a high dimensional global optimization problem in a few small sequential optimization steps, with partial analytical solutions.

# Chapter 2

# Background

### 2.1 Black-Scholes Pricing Framework

In order to understand stochastic volatility, perturbation techniques and the models using them to create a platform for pricing path-dependent options, we must initially take a short excursion through the pricing theory put forward by Black and Scholes (). We do this to explain the mathematical concepts inherent to the theory and lay the groundwork for its model independent extension to stochastic volatility (following the work of [2]).

We begin with the simple market model proposed by Samuelson (1973) and used by Black and Scholes to price a risky asset under lognormal assumptions. Eventually we present the extension of the original Black-Scholes framework in the risk-neutral probability space, covering specific derivative instruments in the passing.

#### 2.1.1 Probability Space and Stochastic Processes

As detailed in [3] a probability space is a triple of the form  $(\Omega, \mathscr{F}, P)$ , where:

- $\Omega$  represents the sample space (i.e. the set of all outcomes)
- $\mathscr{F}$  is a  $\sigma$ -field (or  $\sigma$ -algebra), i.e. a nonempty collection of subsets of  $\Omega$  that is closed under the complement and union operations
- $P: \mathscr{F} \to [0,1]$  is a function associating probabilities to events

Following the above definition, a continuous-time stochastic process is defined as a collection of random variables  $(X_t)_{t\geq 0}$ , one for each positive t (time), that are all defined on a given probability space.

#### 2.1.2 Brownian Motion

Brownian Motion (or Wiener Process) is a real-valued stochastic process W with continuous trajectories  $t \to W_t$  that have independent and stationary increments. More formally, a Brownian motion is characterized by:

- 1. W(0) = 0 with probability 1
- 2. For  $0 \le s < t$ , the random increment  $W_t W_s$  is normally distributed, that is  $W_t W_s \sim N(\mu = 0, \sigma^2 = t s)$ .
- 3. For any  $0 \le q < r < s < t$ ,  $(W_r W_q)$  and  $(W_t W_s)$  are independent.

The increasing sequence of  $\sigma$ -fields  $\mathscr{F}_t$  generated by  $(W_t)_{s \leq t}$  is called a *filtration*. The stochastic process  $(X_t)_{t\geq 0}$  is adapted to the filtration  $(\mathscr{F}_t)_{t\geq 0}$  if the random variable  $(X_t)$  is  $\mathscr{F}_t$  measurable for every t, meaning any event  $(X_t \leq x) \in \mathscr{F}_t$ . The independence of the increments of the Brownian motion can be thus rewriten in the form

$$\mathbb{E}\left\{ e^{iu(W_t - W_s)} | \mathscr{F}_s \right\} = \mathbb{E}\left\{ e^{iu(W_t - W_s)} \right\} = e^{-u^2(t-s)/2}$$
(2.1)

which shows the independence of  $W_t - W_s$  from the past  $\mathscr{F}_s$ .

Due to the independence of its increments, the Brownian motion is an ideal model for an entire family of independent infinitesimal increments  $dW_t$ , that are normally distributed, with mean 0 and variance dt. However, the same useful increment independence does not allow us to integrate with respect to  $dW_t$  in the usual "trajectory by trajectory" way (trajectories are of unbounded variation). In order to overcome this problem and use Brownian motion as a model of white Gaussian noise, we define stochastic integrals.

#### 2.1.3 Stochastic Integrals

Let  $(X_t)_{0 \le t \le T}$  be a continuous stochastic process adapted to  $(\mathscr{F}_t)_{0 \le t \le T}$ , the filtration of the Brownian motion up to time T, so that:

$$\mathbb{E}\left\{\int_{0}^{T} X_{t}^{2} dt\right\} < +\infty$$
(2.2)

The stochastic integral of  $(X_t)$  w.r.t the Brownian motion  $(W_t)$  is defined as a limit in the mean-square sense:

$$\int_{0}^{t} X_{s} dW_{s} = \lim_{n \to +\infty} \sum_{i=1}^{n} X_{t_{i-1}} (W_{t_{i}} - W_{t_{i-1}})$$
(2.3)

as the size of the subdivisions  $[t_i, t_{i-1}]$  of the interval  $[0, +\infty]$  so that  $t_0 = 0 < t_1 < t_2 < ... < t_n = +\infty$  goes to zero. The above stochastic integral, as a function of time, defines a continuous square integrable process such that:

$$\mathbb{E}\left\{\left(\int_{0}^{t} X_{s} dW_{s}\right)^{2}\right\} = \mathbb{E}\left\{\int_{0}^{t} X_{s}^{2} ds\right\}$$
(2.4)

and has the **martingale property** (as do all pure diffusion terms of the form below)

$$\mathbb{E}\left\{\int_{0}^{t} X_{u} dW_{u} |\mathscr{F}_{s}\right\} = \mathbb{E}\left\{\int_{0}^{t} X_{s}^{2} ds\right\}$$
(2.5)

and a quadratic variation  $\langle Y \rangle_t$   $(Y_t = \int_0^t X_u dW_u)$  given by:

$$\langle Y \rangle_t = \lim_{n \to +\infty} \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2 = \int_0^t X_s^2 ds$$
 (2.6)

#### 2.1.4 Market Model and Risky Asset Price

The market model underlying the Black-Scholes theory consists of two assets:

1. The riskless asset (bond) with price  $\beta_t$  at time t that obeys the ordinary differential equation

$$d\beta_t = r\beta_t dt \tag{2.7}$$

where r = instantaneous interest rate for lending or borrowing, and where  $\beta_t = e^{(rt)}$ for  $\beta_0 = 1$  and  $t \ge 0$ 

2. The risky asset (stock or stock index) with price  $X_t$  that obeys the stochastic differential equation:

$$dX_t = \mu X_t dt + \sigma X_t dW_t \tag{2.8}$$

where  $\mu = \text{constant}$  mean return rate,  $\sigma > 0 = \text{constant volatility}$ , and  $(W_t)_{t\geq 0}$  is a standard Brownian motion

In integral form, equation (2.8) can be rewriten as:

$$X_{t} = X_{0} + \mu \int_{0}^{t} X_{s} ds + \sigma \int_{0}^{t} X_{s} dW_{s}$$
(2.9)

where the last integral is of the form described in the previous subsection.

#### 2.1.5 Ito's Formula

Whenever a random variable X follows an Ito process (which contains stochastic and nonstochastic terms) of the form:

$$dX_t = \mu(t, X_t)X_t dt + \sigma(t, X_t)X_t dW_t$$
(2.10)

we can compute the differential of a twice differentiable function g of  $X_t$  and time t using the following expression derived by Ito:

$$dg(t, X_t) = \left(\frac{\delta g}{\delta t} + \mu(t, X_t)\frac{\delta g}{\delta x} + \frac{1}{2}\sigma^2(t, X_t)\frac{\delta^2 g}{\delta x^2}\right)dt + \sigma(t, X_t)\frac{\delta g}{\delta x}dW_t$$
(2.11)

#### 2.1.6 Lognormal Risky Asset Price

An immediate application of Ito's formula is switching to a lognormal representation of the risky asset price. First, dividing by price in (2.8) we obtain:

$$\frac{dX_t}{X_t} = \mu dt + \sigma dW_t \tag{2.12}$$

As from ordinary calculus we have  $\int dx/x = \log x$ , we can compute the differential of the logarithm by applying Ito's formula with  $g(t, x) = \log x$ ,  $\mu(t, x) = \mu x$ , and  $\sigma(t, x) = \sigma x$ , and obtain

$$d\log X_t = \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dW_t \tag{2.13}$$

which can be rewritten as:

$$\frac{X_t}{X_0} = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right)$$
(2.14)

The process  $(X_t)$  is called a geometric Brownian motion, as we can see that in the above equation the return is lognormal - the exponential of a non-standard Brownian Motion. Unless the stock is in permanent bankruncpy  $(X_0 = X_t = 0, \forall t)$ , for a positive current price  $X_0 > 0$ ,  $X_t$  always remains positive, as noticeable in figure 2.1.

#### 2.1.7 European and Path Depedent options

A European call option is a contract giving its holder the right, but not the obligation, at a predetermined maturity time T, to exercise the option and buy one unit of the underlying asset for a predetermined strike price. Given  $X_T$  = the price of the underlying asset at



FIGURE 2.1: Geometric Brownian motion for  $\mu = 0.15, \sigma = 0.20, X_0 = 1$ . The mean path (red), 20 sample paths for  $t \in [0; 5]$  (green), and 1%, 5%, 10%, 90%, 95% and 99% percentile paths (red dashed).

maturity, the contract value at maturity (or its payoff) is given by:

$$h(X_T) = (X_T - K)^+ = \left\{ \begin{array}{l} X_T - K & \text{if } X_T > K \\ 0 & \text{if } X_T \le K \end{array} \right\}$$
(2.15)

A European put option is a contract giving its holder the right, but not the obligation, at a predetermined maturity time T, to exercise the option and sell one unit of the underlying asset for the a predetermined strike price. Its payoff is given by:

$$h(X_T) = (K - X_T)^+ = \left\{ \begin{array}{cc} K - X_T & \text{if } X_T < K \\ 0 & \text{if } X_T \ge K \end{array} \right\}$$
(2.16)

European-style derivatives are path-independent as their payoff function is only a function of the stock price value at maturity time T. At a random time t < T, the value of the contract, or derivative price, will vary with t and the current stock price  $X_t = x$ . The problem of pricing the contract becomes equivalent to finding a pricing function P(t, x), that shares the Markov property (memoryless w.r.t. past stock prices) with the stochastic differential equations modelling a stock price process as defined earlier in equation (2.8). American options have the same payoff as European ones with the only difference being that they can be exercised at any time prior to maturity, which includes further complications in their pricing using perturbation theory, and will, hence, not be treated here.

In contrast to the "plain vanilla" European and American options described earlier, any option with a different payoff structure is called "exotic". We will introduce barrier options to illustrate the purpose of the platform for options pricing based on perturbation theory. We will explore perturbation theory - pricing exotic path-dependent options starting with an accurate pricing of vanilla options - in the remainder of this thesis.

Barrier options are path dependent options with a payoff function that depends on whether or not the underlying asset price reaches a target value during the option's lifetime. A downand-out call option on a stock for example, becomes worthless (ceases to exist) if at any time t < T, the stock price  $X_t$  falls below a given barrier level B that is higher than the initial stock price. Its payoff function at expiration is thus determined by:

$$h(X_T) = (X_T - K)^+ \mathbf{1}_{\{inf_t < T X_t > B\}}$$
(2.17)

where the indicator function of the set A is defined by:

$$\mathbf{1}_{A}(x) = \left\{ \begin{array}{cc} 1 & \text{if } x \in A \\ 0 & \text{if } X \notin A \end{array} \right\}$$
(2.18)

Given its payoff structure, we expect its value  $C_{down-out}$  to be less that that of a call C, and by considering its dual knock-in option, the down-and-in call which activates below a threshold B, as mentioned in [4], we have  $C_{down-out} = C - C_{down-in}$ .

#### 2.1.8 Black-Scholes Partial Differential Equation and Formula

There are multiple approaches to arrive at the Black-Scholes formula for pricing European options. First, we can use a discrete approximation of the Brownian motion governing the risky asset's price process and determine the value at each node of the resulting tree by no-arbitrage arguments. Decreasing the time intervals between the levels of the tree, we can obtain a reliable continuous time approximation of the option price. Alternatively, we can take on a more fundamental approach relying on the concept of a martingale measure, as presented in detail by [5], or use a delta-hedging argument (delta is the first derivative of the option price with respect the underlying asset's price, measuring the exposure to the underlier). We will follow the following approach: arriving at the Black-Scholes PDE and its solution by using replication and no-arbitrage arguments and then adapting our stochastic model to risk neutral measures, following the outline in [2].

We therefore introduce the notion of a replicating self-financing portfolio. A trading strategy to replicate the payoff  $h(X_T)$  of a European-style derivative is to replicate it with a pair  $(a_t, b_t)$  of processes that indicate the number of units held at time t of the underlying risky asset and riskless bond respectively. We assume continuus compounding so that the price of the bond at time t is  $\beta_t = e^{rt}$ , given a risk-free interest rate r (for more details refer to [4]). We further assume that the stochastic integrals of  $(a_t)$  and  $b_t$  are well defined.

Therefore the value at time t of this portfolio is  $a_t X_t + b_t e^{rt}$  and the replication condition of matching payoffs at maturity is:

$$a_T X_T + b_T e^{rT} = h(X_T) (2.19)$$

The self-financing condition imposes that all value fluctuations of the portfolio are only due to market fluctuations (changes in the prices of stock and bond portfolio constituents), so that, for example, in order to finance more units of stock we need to sell bond units from our portfolio. The latter can be expressed in differential form as

$$d(a_t X_t + b_t e^{rt}) = a_t dX_t + rb_t e^{rt} dt$$

$$(2.20)$$

or in integral form as

$$a_t X_t + b_t e^{rt} = a_0 X_0 + b_0 + \int_0^t a_s dX_s + \int_0^t r b_s e^{rs} ds, \ 0 \le t \le T$$
(2.21)

We assume that a regular pricing function P(t, x) for a European option exists so that we can apply Ito's formula to its differential. If we construct a self-financing replicating portfolio to match the pricing function during the contract's lifetime, the lack of arbitrage (instant profit with no future payoff) condition requires that:

$$a_t X_t + b_t e^{rt} = P(t, X_t), \text{ for any } 0 \le t \le$$

$$(2.22)$$

Differentiating the above equation, replacing the left-hand side according to (2.20) applying Ito's formula on the right hand side ((2.11)) and expanding the underlying asset price according to (2.8), we obtain:

$$(a_t \mu X_t + b_t r e^{rt})dt + a_t \sigma X_t dW_t = \left(\frac{\partial P}{\partial t} + \mu X_t \frac{\partial P}{\partial x} + \frac{1}{2}\sigma^2 X_t^2 \frac{\partial^2 P}{\partial x^2}\right)dt + \sigma X_t \frac{\partial P}{\partial x}dW_t, \quad (2.23)$$

with all partial derivatives of P evaluated  $(t, X_t)$ . Equating the factors of  $dW_t$  in the above equation yields

$$a_t = \frac{\partial P}{\partial x}(t, X_t) \tag{2.24}$$

and extracting  $b_t$  from the no-arbitrage equation (2.22) yields

$$b_t = (P(t, X_t) - a_t X_t) e^{-rt}.$$
(2.25)

Finally, equating the factors of  $d_t$  in (2.23) we obtain the Black-Scholes partial differential equation

$$\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 X_t^2 \frac{\partial^2 P}{\partial x^2} + r(X_t \frac{\partial P}{\partial x} - P) = 0.$$
(2.26)

One particularity of the above equation is that the rate of return  $\mu$  ("drift" term) does not enter the equation, meaning that market participants will agree on the no-arbitrage price of the contract regardless of their views on the growth rate of the underlying, and thus the risky asset influences the contract price only by the volatility term  $\sigma$  and its current price. This highlights the importance of modelling and estimating volatility.

Introducing the Black-Scholes operator  $\mathscr{L}_{BS}(\sigma)$  given by

$$\mathscr{L}_{BS}(\sigma) = \frac{\partial}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + r(x\frac{\partial}{\partial x} - \cdot)$$
(2.27)

we can rewrite (2.26) as

$$\mathscr{L}_{BS}(\sigma)P = 0. \tag{2.28}$$

The Black-Scholes PDE is solved with the final condition  $h(x) = (x - K)^+$ . Thus, the price of a European call at time t, when the risky asset price is  $X_t = x$ , is denoted by  $C_{BS}(t, x)$ and its closed-form solution is given by the Black-Scholes formula:

$$C_{BS}(t,x) = xN(d_1) - Ke^{-r(T-t)}N(d_2), \qquad (2.29)$$

where

$$d_1 = \frac{\log(x/K) + (r + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}},$$
(2.30)

$$d_2 = d_1 - \sigma \sqrt{T - t},\tag{2.31}$$

and

$$N(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-y^2/2} dy, \qquad (2.32)$$

is the standard Gaussian cumulative distribution function. The input parameters needed for the Black-Scholes formula above for European options are thus:

• r - risk free interest rate

- T t time to maturity
- x underlying risky asset (i.e. stock/stock index) current price
- K option strike price
- $\sigma$  volatility, the only parameter we need an estimate for, as it is not given

The formula for European puts can be easily inferred by noting the put-call parity relation between calls and puts of the same maturity and strike price:

$$C_{BS}(t, X_t) - P_{BS}(t, X_t) = X_t - Ke^{-r(T-t)},$$
(2.33)

#### 2.1.9 The Greeks

The quantities known as the greeks of an options measure the sensitivity of its price in relation to variations of its parameters, and are measured by the relevant partial derivatives of the option price. We will give a succint description of these sensitivities for a call option.

Delta is the first derivative with respect to the stock price and is given by

$$\Delta_{BS} = \frac{\partial C_{BS}}{\partial x} = N(d_1). \tag{2.34}$$

Gamma is the second derivative with respect to the stock price and is given by

$$\Gamma_{BS} = \frac{\partial^2 C_{BS}}{\partial x^2} = \frac{\partial \Delta_{BS}}{\partial x} = \frac{e^{-d_1^2/2}}{x\sigma\sqrt{2\pi(T-t)}}.$$
(2.35)

Vega is the sensitivity to the volatility level and equals

$$\gamma_{BS} = \frac{\partial C_{BS}}{\partial \sigma} = \frac{x e^{-d_1^2/2} \sqrt{T - t}}{\sqrt{2\pi}}.$$
(2.36)

Rho and Theta are the sensitivities with respect to the interest rate r and time to maturity T - t. Finally, it can be proven that there is a relation between Vega and Gamma:

$$\frac{\partial P}{\partial \sigma} = (T-t)\sigma x^2 \frac{\partial^2 P}{\partial x^2}.$$
(2.37)

#### 2.1.10 Risk-Neutral Pricing

An alternative method to price options would start from their discounted expected value, so for instance the option price at time t = 0 of an observed underlying asset price x would equal

$$P(0,x) = \mathbb{E}\left\{ e^{-rT}h(X_T) \right\} = \mathbb{E}\left\{ e^{-rT}h\left(xe^{\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma W_T}\right) \right\},$$
(2.38)

where the last equality was obtained by expanding the asset price process after the lognormal formula. However, as mentioned in [2], since  $W_T$  is normally distributed the expectation reduces to a Gaussian integral which creates an arbitrage opportunity in the above option pricing equation, unless  $\mu = r$ . This is related to the fact that the discounted price  $\tilde{X}_t = e^{-rt}X_t$  is not a martingale, as its differential follows

$$d\tilde{X}_t = (\mu - r)\tilde{X}_t dt + \sigma \tilde{X}_t dW_t, \qquad (2.39)$$

that contains a non-zero drift term if  $\mu \neq r$ .

To overcome this difficulty we introduce a unique risk-neutral probability measure  $\mathbb{P}^*$  equivalent to  $\mathbb{P}$  so that under the new measure the discounted price  $\tilde{X}_t$  is a martingale and the expected discounted payoff of a derivative results in a no-arbitrage price.

The first step is to modify the above equation is to "absorb" the drift term into a martingale term in the form

$$d\tilde{X}_t = \sigma \tilde{X}_t \left[ dW_t + \left(\frac{\mu - r}{\sigma}\right) dt \right].$$
(2.40)

Setting

market price of asset risk = 
$$\theta = \frac{\mu - r}{\sigma}$$
, (2.41)

we define a modified Brownian motion  $W_t^{\star}$  over probability measure  $\mathbb{P}^{\star}$  so that

$$W_t^{\star} = W_t + \int_0^t \theta \, ds = W_t + \theta t \text{ and } d\tilde{X}_t = \sigma \tilde{X}_t dW_t.$$
(2.42)

Using Girsanov's Theorem (described in detail in [5]), it can be proven that the modified process  $W_t^{\star}$  is a Brownian motion process itself, which implies that the discounted price process is a martingale and thus insures that no-arbitrage holds. Starting from a risk-neutral discounted payoff pricing function it can be shown (as described in [5] and [2]) that the price of an European call follows the Black-Scholes formula.

### 2.1.11 Markov Processes and Corresponding Infinitisemal Generators and Martingales

A process  $(X_t)$  for an asset price of the form

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \tag{2.43}$$

is a Markov Process if it posseses the Markov property

$$E\{h(X_s)|\mathscr{F}_s\} = E\{h(X_s^{t,x})\}|_{x=X_t}.$$
(2.44)

where h is a function of the process  $(X_t)$ . In words, the Markov property implies that, to find h from 0 to s, we can start from an intermediate point t in time, where we know the value of the process  $X_t$ .

Motivated by their use in analytic characterizations of both volatility time scales (subsection 3) and Monte Carlo variance reduction techniques (chapter Monte Carlo), we also introduce the notion of infinitisemal generators. We consider a Markov process  $(X_t)$  of the form 2.43, and a twice continuously differentiable function g(x) acting on it. The differential operator  $\mathscr{L}$  acting on g(x) according to

$$\mathscr{L}g(x) = \frac{1}{2}\sigma^2(x)g''(x) + \mu(x)g'(x), \qquad (2.45)$$

is called the infinitisemal generator of the Markov process  $(X_t)$ . For a process of the form 2.43 that is time-inhomogenous ( $\sigma$ ,  $\mu$  and g depend on time t too), its time-dependent infinitesimal generator  $\mathscr{L}_t$  is given by

$$\mathscr{L}_t = \frac{1}{2}\sigma^2(t,x)\frac{\partial^2}{\partial x^2} + \mu(t,x)\frac{\partial}{\partial x}.$$
(2.46)

Applying Ito formula w.r.t x and then (x,t), it can be shown according to 2.5 that the time-homogenous and time-inhomogenous terms

$$M_t = g(X_t) - \int_0^t \mathscr{L}g(X_s)ds, \qquad (2.47)$$

$$M_t = g(t, X_t) - \int_0^t \left(\frac{\partial g}{\partial t} + \mathscr{L}_s g\right)(s, X_s) ds$$
(2.48)

are (corresponding) martingales.

Based on the above, in the time-homegenous case, it can be shown that

$$\frac{d}{dt} \mathbb{E}\{g(X_t)\}|_{t=0} = \lim_{t \downarrow 0} \frac{\mathbb{E}\{g(X_t)\} - g(x)}{t} = \mathscr{L}g(x),$$
(2.49)

which gives the intuitive grounds for infinitisemal generators. They govern the (speed of) evolution over time of the expectation of a Markov process. This property is central to characterising different volatility time scales (see subsection 4).

#### 2.1.12 Complete vs Incomplete Markets

The Black-Scholes lognormal model implies a complete market, meaning a market in which any derivative can be replicated by a self-financing portfolio of stocks and bonds. Another feature of a complete market is the existence of a unique equivalent martingale measure  $\mathbb{P}^*$ so that the discounted prices of traded securities are martingales.

In constrast, an incomplete market does not permit perfect hedging only using bonds and stocks, as it allows a whole family of equivalent martingale measures.

...... (infinitisemal generators?)

### 2.2 Stochastic Volatility Framework

The classical Black-Scholes model considered in the previous section makes several assumptions that severely impact its practical aplicability. First, it makes an incorrect assumption on the lognormal density function of stock prices, as numerous empirical studies and market crashes present evidence of thicker, asymmetric distribution tails. Secondly, there are major discrepancies between the prices Black-Scholes computes and market-quoted prices across maturities and strikes that exhibit an implied volatility skew. The Black-Scholes assumption of constant volatility (historical volatility) has been dismantled by numerous empirical studies, most notably Rubinstein (1985), which brought evidence of random volatility characteristics.

Therefore the emerging practical extensions consisted of local volatility and stochastic volatility models. A more recent trend has been started at Bloomberg and in other academic publications (Stochastic local volatility paper) to bring together the two approaches in so-called "stochastic local volatility models", as previous market events and model misprisings proved many purists on both sides wrong.

Many stochastic volatility models evolved separately to model highly specific financial derivatives; however, a unifying approach using a common parameter base is highly desirable. This is motivated by the need for cross-pricing of multiple derivatives from more liquidly traded options to exotic ones, cross-market pricing, as well as a common base for further research in a hybrid approach with local volatility models. Therefore, in this section I will present the characteristics of both the local and stochastic models and their associated volatility surfaces, and introduce popular stochastic volatility models among practitioners and in academia. Finally, I will differentiate between one-factor and multifactor models.

#### 2.2.1 Historic Volatility

The most natural substitute for the unkonwn volatility parameter in the Black-Scholes formula is historic volatility. The idea is to use historical stock price data in order to estimate  $\sigma$ . Since in reality the volatility is not constant over time, the standard practice is to use historical data for a period of the same length as the time to maturity - in this case six months.

As detailed in [6] we assume to have the standard Black-Scholes model and we sample (observe) the stock price process X at n+1 discrete equidistant points  $t_0, t_1, \ldots, t_n$ , where  $\Delta t$  denotes the length of the sampling interval ( $\Delta t = t_i - t_{i-1}$ ). We thus observe  $X(t_0), \ldots, X(t_n)$ , and, using the fact that X has a log-normal continuous distribution, we define the independent, normally distributed variables  $\xi_1, \ldots, \xi_n$  as

$$\xi_i = \ln \left( \frac{X(t_i)}{X(t_{i-1})} \right), \tag{2.50}$$

with

$$Var[\xi_i] = \sigma^2 \Delta t. \tag{2.51}$$

Thus, an estimate of  $\sigma$  is given by

$$\sigma^* = \frac{X_{\xi}}{\sqrt{\Delta t}},\tag{2.52}$$

where

$$S_{\xi}^{2} = \frac{1}{n-1} \sum_{i=1}^{n} (\xi_{i} - \bar{\xi})^{2}.$$
(2.53)

#### 2.2.2 Implied Volatility Surface

Implied volatility is often used to express discrepancies between the Black-Scholes predicted prices and those quoted in the market, being used by traders in practice to quote derivative prices. As the term "implied" suggests, it quantifies the market expectation of the volatility of an option over its remaining lifetime, without using historic volatility as a predictor of future volatility. As [5] describes, "A more subtle issue is that it is not the past volatility that matters. It is the volatility that occurs during the lifetime of an option which will cause hedging costs and the option should be priced thereby". Given the latter, historical volatility becomes an ill-conditioned predictor of future realized volatility, just as historical stock returns would be bad absolute predictors of future returns.

Therefore, we define implied volatility I = I(t, x; K, T) as the volatility parameter that matches the observed (market) prices and Black-Scholes prices:

$$C_{BS}(t,x;K,T;I) = C^{\text{obs}}$$

$$(2.54)$$



FIGURE 2.2: Discounted stock price =  $xe^{r(T-t)} \approx 102$  = minimum vol.

FIGURE 2.3: Vertical line shows at-themoney level (= stock price x)

and which has the following properties:

- Given the monotonicity of  $\Gamma$  in (2.36)  $C^{\text{obs}} > C_{BS}(t, x; K, T; 0) \rightarrow I > 0$
- Put-call parity (2.33) ensures implied volatilities from put and call option of the same strike price and time-to-maturity are equal  $(I(t; K, T)_{call} = I(t; K, T)_{put})$

If we visualize the implied volatilities across strikes, given all other parameters are fixed, we find it is U-shaped, with a minimum around the discounted stock price pre-1987 "Black Monday" (figure 2.3). Post the 1987 stock market crash, investors became aware of huge negative stock price drops and put a premium on insurance, leading to a negatively-sloped volatility function around the strike price (skew) (figure 2.2).

In addition to the non-linear shape across strikes, if we visualize implied volatilities as a function of moneyness K/X (figure 2.4) or log-moneyness (figure 2.5), we can see that skews are steeper for shorter maturities (as visible by the strands corresponding to each maturity and the 3 - D view respectively). Another qualitative feature is that stock index implied volatilities are higher than historical volatilities ([2]).

The variation of implied volatilities with respect to maturity is called the *term structure*, and the variation with respect to both maturity T and strike K is known as the *implied volatility* surface.

The observed implied volatility skew and empirical studies of historical volatility point to its random nature, contrary to the constant volatility assumption of Black-Scholes. So a natural question is how the smile (or skew) arises. [5] provides an explanation using a real-world example. Suppose a market maker for vanilla options quotes a bid (for buying) and an offer (for selling) in terms of volatility, with a spread between the two. As long as he makes a



FIGURE 2.4: S&P 500 index options implied volatilities (June 1, 2007)



FIGURE 2.5: S&P 500 index options implied volatilities (September 15, 2005), [1]

riskless profit when he is perfectly hedged (the amount he buys from clients = the amount he sells to clients), there will be an equilibrium volatility. However, that volatility will be different across strikes due to different selling and buying behaviour for each strike, as well as the inability to have a perfect hedge (in the model-free sense) between different strikes. One is taking risk on the basis of the imperfections of the model, and thus the skew expresses the market's view of the model's imperfections - in this case Black-Scholes.

However, the traders' need to impose a pricing consistency across a broad range of related financial instruments has led to the use of the "practitioner Black-Scholes", which relies on the short-term stability of the smile in order to input different volatilities for different options, depending on their moneyness ([7]). The practitioner belief that implied volatility is a better predictor of future realized volatility led to extensive research of volatility models that are usually calibrated only from derivative data (traded option prices), without taking into consideration the underlying asset time series. This, so-called, cross-sectional fitting is easier to perform than econometric time series models and provides a pricing platform that can be parameterized through, for example, highly liquid at-the-money European option prices and those same parameters can then be used to price more illiquid options.

#### 2.2.3 Local Volatility

One way to modify the lognormal model of the underlying asset price is to model volatility as a function of time and stock price in the form  $\sigma = \sigma(t, X_t)$ . Thus the asset process becomes

$$dX_t = \mu X_t dt + \sigma(t, X_t) X_t dW_t, \qquad (2.55)$$

and the no-arbitrage price of an European option follows the generalized Black-Scholes PDE of the form (2.26), with only the volatility term changed.

Since the randomness of the volatility follows the randomness of the original lognormal model, the market is still complete, and allows for a unique risk-neutral measure  $\mathbb{P}^*$  under which

$$dX_t = rX_t dt + \sigma(t, X_t) X_t dW_t^{\star}, \qquad (2.56)$$

where r is the risk-free drift rate and  $(W_t^{\star})$  is a  $\mathbb{P}^{\star}$ -Brownian motion

1. Time-dependent only volatility:  $\sigma = \sigma(t)$ 

The risky asset return follows a lognormal distribution analogous to that given by (2.14) and, hence, we use the standard Black-Scholes formula only adjusted for the



FIGURE 2.6: S&P 500 index level and BlackScholes implied volatility each Wednesday during the period of June 1988 through December 1993

new root-mean square volatility parameter  $\sqrt{\sigma^2}$  given by

$$\overline{\sigma^2} = \frac{1}{T-t} \int^T \sigma^2(s) ds.$$
(2.57)

where, using the latter equation, the integral can be backed out between two maturities  $T_1 < T_2$  as

$$\int_{T_1}^{T_2} \sigma^2(s) ds = (T_2 - t)I(t, T_2)^2 - (T_1 - t)I(t, T_1)^2$$
(2.58)

As  $\overline{\sigma^2}$  is different across maturities, there is a term structure of implied volatility; however, this is obtained at the cost of constant recalibration as data changes, creating an inconsistent approach to valuing other options.

#### 2. Underlying price and time dependent volatility: $\sigma = \sigma(t, x)$

For a skew across strike prices to be present,  $\sigma$  needs to depend on the stock price as well. On the downside, however, the model now only allows for perfect correlation (in this case negative) which does not capture the full dynamics of volatility. Empirical studies and economic arguments point that an increase in volatility tends to coincide with a decrease in stock prices. So there is an inverse relation but no perfect correlation between the two, as we can see in figure 2.6, extracted from [8].

The volatility function  $\sigma(t, x)$  can be estimated either in a parametric way (for example fitting  $\sigma(t, x) = \kappa x^{-\gamma}$  to data), or in a nonparametric way as an inverse problem to the generalized Black-Scholes PDE. Under the assumption of observing option prices for a continuum of strikes and maturities, the latter approach has a closed-form solution as detailed in [9]. Thus the Dupire formula defining the local volatility surface is given by

$$\sigma^{2}(T,K) = \frac{\frac{\partial C}{\partial T}(T,K) + rK\frac{\partial C}{\partial K}(T,K)}{\frac{1}{2}K^{2}\frac{\partial^{2}C}{\partial K^{2}}(T,K)}.$$
(2.59)

In practice, to uphold the assumption of continuous strike and maturity dimensions and estimate the partial derivatives in the above formula, interpolation of option prices at the observed discrete values is performed. This creates a high risk of misfit when the grid (K, T) of traded option prices is sparse.

#### 2.2.4 Local Volatility vs Stochastic Volatility

Compared to the stochastic volatility models we will introduce, local volatility models have the advantages of a complete market that allows for a unique risk-neutral measure and a very good fit of the option prices. The weak point of local volatility is its parameter instability is that it produces a highly unstable surface, given the fit's large degree of freedom and constant data changes that can occur even on a weekly basis.

Derman and Kani in [10], Dupire in [9], and Rubinstein in [11] hypothesized that volatility is a deterministic function of asset prices and time, and provided appropriate binomial or trinomial option valuation procedures to account for it. However, a later empirical study by [8] tests the "null hypothesis" of volatility as a deterministic function of asset price and time on S&P 500 index options. The study finds that the model specifications that include the time parameter preponderely lead to overfitting and that the hedge ratios given by the different specifications of the local volatility model are worse than those of the original Black-Scholes model. Thus, the study rejects the hypothesis, advocating stochastic volatility models, unless a generalized local volatility model using past index price changes does not present similar weaknesses.

#### 2.2.5 Stochastic Volatility Models

As [1] points out, stochastic volatility (SV) models are useful as they explain in a selfconsistent way why the "volatility smile occurs". Furthermore, as oppposed to local volatility models that can also fit the smile, SV models assume realistic dynamics for the underlying assets "arising from Brownian motion subordinated to a random clock", rather than ad hoc. [12] suggests that the clock time, or trading time, can be identified with the volume/frequency of trading for each maturity or strike that determines volatility. This just reinforces the idea presented at the end of subsection 2.2.2 that looked at the dynamics of volatility from a practitioner's view. In stochastic volatility models, the asset price process  $(X_t)_{t\geq 0}$  is given by the stochastic differential equation

$$dX_t = \mu_t X_t dt + \sigma_t X_t dW_t^{(0)} \tag{2.60}$$

where  $W^{(0)}$  is a standard Brownian motion,  $(\mu_t)_{t\geq 0}$  is an adapted returns process, and  $(\sigma_t)_{t\geq 0}$ is called the volatility process.  $(\sigma_t)_{t\geq 0}$  must be postitive and satisfy some regularity conditions for the model to be well-defined, but it does not have to be an Ito process (for example it can be a jump process or a Markov chain). Thus a whole family of stochastic volatility processes were developed by practitioners and academia alike. [1] gives a good overview of the various models based on jump-diffusion on log jump-diffusion (Levy) processes as well SV models with jumps (SVJ), concluding that for appropriate parameters, any stochastic volatility with jump model would yield a similar shape of the volatility surface.

A primary characteristic that departs from the local volatility approach is a volatility process that is not perfectly correlated with the Brownian motion  $W^{(0)}$ . This process is modeled to have an independent random component of its own, leading to more flexibility and consistency, but also to an incomplete market.

#### 2.2.6 One-Factor Stochastic Volatility Models

In addition to the Ito process governing the asset price, we want to model volatility as a function of a one-dimensional Ito process, governed by an SDE that is driven by a second brownian motion  $W^{(1)}$ . We introduce a smooth, positive and increasing function f so that  $\sigma_t = f(\mathbf{Y}_t)$  where the process Y is defined by the SDE

$$d\mathbf{Y}_t = \alpha(\mathbf{Y}_t)dt + \beta(\mathbf{Y}_t)dW_t^{(1)}.$$
(2.61)

 $(\alpha : \mathbb{R} \to \mathbb{R}, \ \beta : \mathbb{R} \to \mathbb{R}+)$  Through f we can enforce positivity and extend the framework to include the dependency of volatility on multiple factors. The two Brownian motions are typically correlated with an instantaneous correlation coefficient  $p \in [-1, 1]$ , given by

$$d\langle W^{(0)}, W^{(1)} \rangle_t = pdt,$$
 (2.62)

The parameter p can alternatively be viewed as the correlation between stock price and volatility shocks, as by applying Ito's formula to  $f(\mathbf{Y}_t)$  and extracting the martingale part of  $d\sigma_t$  we obtain

$$cov(d\sigma_t, dX_t) = pf'(\mathbf{Y}_t)\beta(\mathbf{Y}_t)f(\mathbf{Y}_t)X_tdt$$
(2.63)

with a correlation indepedent of time (constant), given by

$$corr(d\sigma_t, dX_t) = p($$
which tends to be negative, as explained in 2.2.3 $).$  (2.64)

#### 2.2.7 Mean reversion-driven processes

Another desirable feature of volatility is mean reversion, as evidenced by empirical studies such as [13], which loosely speaking, increases its probability to drop when it is high and rise when it is low. We can model a mean reversion process by simply incorporating a linear pull-back term in the drift driving an Ito process. To obtain mean-reverting volatility Y that is pulled back around a long-run mean level m at a rate of mean-reversion  $\alpha > 0$ , we define the Y process as

$$d\mathbf{Y}_t = \alpha(\mathbf{Y}_t)dt + \beta(\mathbf{Y}_t)dW_t^{(1)}.$$
(2.65)

Particular volatility driving processes are:

• LN - lognormal

$$d\mathbf{Y}_t = \alpha \mathbf{Y}_t dt + \beta \mathbf{Y}_t dW_t^{(1)} \tag{2.66}$$

• OU - Ornstein-Uhlenbeck

$$d\mathbf{Y}_t = \alpha \mathbf{Y}_t dt + \beta \mathbf{Y}_t dW_t^{(1)} \tag{2.67}$$

• CIR - Cox-Ingersoll-Ross

$$d\mathbf{Y}_t = \alpha \mathbf{Y}_t dt + \beta \mathbf{Y}_t dW_t^{(1)} \tag{2.68}$$

Model Name	Correlation	f(y)	Y Process
Hull-White	p = 0	$f(\mathbf{y}) = \sqrt{\mathbf{y}}$	Lognormal
$\operatorname{Scott}$	p = 0	$f(y) = e^y$	OU
Stein-Stein	p = 0	$f(\mathbf{y}) =  \mathbf{y} $	OU
Ball-Roma	p = 0	$f(\mathbf{y}) = \sqrt{\mathbf{y}}$	CIR
Heston	$p \neq 0$	$f(\mathbf{y}) = \sqrt{\mathbf{y}}$	CIR

TABLE 2.1: from [2]

Popular one-factor models in the literature (table 2.1) rely on the three driving processes (only LN is not mean-reverting) for their properties and analytical tractability rather than any other underlying motivation.

#### 2.2.8 One-factor models pricing - PDE approach

We follow the PDE derivation for a generic one-factor model in order to introduce the (total) risk premium and the volatility risk premium - concepts used further in multi-factor models and their perturbation expansions.
Assume the following Markovian one-factor model:

$$\begin{cases} dX_t = \mu(\mathbf{Y}_t) X_t dt + \sigma_t X_t dW_t^{(0)}, \\ \sigma_t = f(\mathbf{Y}_t), \\ d\mathbf{Y}_t = \alpha(\mathbf{Y}_t) dt + \beta(\mathbf{Y}_t) dW_t^{(1)}, \end{cases}$$

$$(2.69)$$

where we decompose  $W_t^{(1)}$  in terms of  $W_t^{(0)}$  and an independent standard Brownian motion  $W_t^{\perp}$  as in the equation

$$W_t^{(1)} = pW_t^{(0)} + \sqrt{1 - p^2}W_t^{\perp}.$$
(2.70)

To price an European derivative with payoff  $h(X_T)$  we define a pricing function  $P(t, X_t, Y_t)$  depending on time and both aseet and volatility processes. [2] proceeds with the derivation of the PDE for the model defined above by a no-arbitrage argument, using a self-financing replicating portfolio along the lines of the Black-Scholes PDE derivation in subsection 2.1.8. We start by noticing that we cannot have a perfect instantaneous hedge with just the asset that can only balance risk from the diffusion term  $dW_t^{(0)}$ . The term  $dW_t^{\perp}$  must be balanced too and the hedge is only possible with the inclusion of another option with a different expiration date that can also be traded (rebalanced) continuously.

Assume two options with prices  $P^{(1)}(t, x, y)$  and  $P^{(2)}(t, x, y)$  and expiration dates  $T_1$  and  $T_2$  respectively so that  $T_1 < T_2$ . To hedge the first option, we have to find the continuous processes  $\{N_t, A_t, \Sigma_t\}$  defining the quantities allocated for each asset. These processes must ensure that the self-financing portfolio

$$\Pi_t = N_t P^{(1)}(t, X_t, Y_t) - A_t X_t - \Sigma_t P^{(2)}(t, X_t, Y_t)$$
(2.71)

is instantaneously riskless (or hedged) at any time t < T. Differentiating the above equation, expanding acccording to a two-dimensional version of the Ito formula, and hedging all risks, we obtain the PDE:

$$\frac{\partial P}{\partial t} + \frac{1}{2}f^2(\mathbf{y})x^2\frac{\partial^2 P}{\partial x^2} + p\beta(\mathbf{y})xf(\mathbf{y})\frac{\partial^2 P}{\partial x \partial \mathbf{y}} + \frac{1}{2}\beta^2(\mathbf{y})\frac{\partial^2 P}{\partial \mathbf{y}^2}$$
(2.72)

with the terminal condition P(T, x, y) = h(x) and no closed-form solution. In the above equation we introduced the total risk premium function  $\Lambda$  given by

$$\Lambda(t, x, y) = p \frac{(\mu(y) - r)}{f(y)} + \gamma(t, x, y) \sqrt{1 - p^2},$$
(2.73)

where  $\gamma(t, x, y)$  is an arbitrary function. The equation gives the interpretation of the total risk premium  $\gamma$  as a linear combination of the *stochastic Sharpe ratio*  $\frac{(\mu(Y_t) - r)}{f(Y_t)}$  and the volatility risk premium (or market price of volatility risk) $\gamma$ , weighted by the correlation p and its complement  $\sqrt{1-p^2}$ , respectively.

#### 2.2.9 One-factor models pricing - Martingale approach

In this subsection, we follow the risk-neutral theory derivation of no-arbitrage price in order to derive the expression of the model in equation 2.69 in terms of risk-neutral measures. The pricing results obtained extend to general non-Markovian models as well.

We first suppose  $\exists$  an equivalent martingale measure  $\mathbb{P}^*$  under which the discounted stock price  $X_t = e^{-rt}X_t$  is a martingale. Then the no-arbitrage price of the derivative with underlying X and square integrable payoff H is given by

$$V_t = \mathcal{E}^{\star}\{e^{-r(T-t)}H|\mathscr{F}_s\}, \forall t \le T.$$
(2.74)

To construct an equivalent martingale measure, we absorb the drift term of  $\overline{X_t}$  in its martingale term as in equation 2.42, defining the Brownian motions

$$W_t^{(0)\star} = W_t^{(0)} + \int_0^t \frac{(\mu(\mathbf{Y}_s) - r)}{f(\mathbf{Y}_s)} ds,$$
  

$$W_t^{\perp\star} = W_t^{\perp} + \int_0^t \gamma_s ds.$$
(2.75)

By Girsanov's theorem,  $W^{(0)\star}$  and  $W^{\perp\star}$  are independent standard Brownian motions under a measure  $\mathbb{P}^{\star(\gamma)}$  (assumed to be well-defined) given by

$$\frac{d\mathbb{P}^{\star(\gamma)}}{d\mathbb{P}} = \exp\left(-\frac{1}{2}\int_0^T ((\theta_s^{(0)})^2 + (\theta_s^{\perp})^2)ds - \int_0^T \theta_s^{(0)}dW_s^{(0)} - \int_0^T \theta_s^{\perp}dW_s^{\perp}\right)$$
(2.76)

where

$$\theta_t^{(0)} = \frac{\mu(\mathbf{Y}_t) - r}{f(\mathbf{Y}_t)},$$
  

$$\theta_t^{\perp} = \gamma_t.$$
(2.77)

The process  $(\gamma_t)$  (market price of volatility risk) parametrizes the space of equivalent martingale measures  $\{\mathbb{P}^{\star(\gamma)}\}$ . Thus under  $\mathbb{P}^{\star(\gamma)}$  the model (2.69) becomes:

$$\begin{cases} dX_t = rX_t dt + f(\mathbf{Y}_t) X_t dW_t^{(0)\star}, \\ d\mathbf{Y}_t = (\alpha(\mathbf{Y}_t) - \beta(\mathbf{Y}_t) \Lambda_t) dt + \beta(\mathbf{Y}_t) dW_t^{(1)\star}, \\ W_t^{(1)\star} = pW_t^{(0)\star} + \sqrt{1 - p^2} W_t^{\perp\star}, \end{cases}$$
(2.78)

where the total risk premium process is

$$\Lambda_t = p \frac{(\mu(\mathbf{Y}_t) - r)}{f(\mathbf{Y}_t)} + \gamma_t \sqrt{1 - p^2}.$$
(2.79)

A natural question arising from the above risk-neutral model formulation is how to choose a risk-neutral probability and determine the necessary parameters. One method is to estimate the parametrically specified functions  $(f(y), \alpha(y), \beta(y))$ , the constant p, and the present volatility level by using econometric methods in extracting them from historical stock price time-series. Then  $\gamma$  can be estimated by using derivatives data, implying a unique martingale measure under which the market prices contract ([6]).

Alternatively, we can use cross-sectional fitting, which is easier to implement and does not suffer from the drawback of trying to estimate an unobservable volatility process. Instead, if we denote the vector of unknown parameters by  $\vartheta$  for strike price and time to maturity pairs (K, T) in some set  $\chi$ , we solve the problem of least squares errors between model-given and observed call prices:

$$\min_{\vartheta} \sum_{(K,T)\in\chi} \left( C(K,T;\vartheta) - C^{obs}(K,T) \right)^2.$$
(2.80)

#### 2.2.10 Multi-Factor Stochastic Volatility Models

We now introduce models that have at least two stochastic volatility factors (as defined in [2]), as we will develop further perturbation expansions on a two-factor model with different time scales for each of its volatility factors.

If we assume that volatility is driven by a *d*-dimensional diffusion process  $Y \in \mathbb{R}^d$  the Markovian stochastic volatility model (2.69) becomes

$$\begin{cases} dX_t = \mu(\mathbf{Y}_t) X_t dt + \sigma_t X_t dW_t^{(0)}, \\ \sigma_t = f(\mathbf{Y}_t), \\ d\mathbf{Y}_t = \alpha(\mathbf{Y}_t) dt + \beta(\mathbf{Y}_t) d\mathbf{W}_t, \end{cases}$$
(2.81)

where  $\alpha : \mathbb{R}^d \to \mathbb{R}^d, \beta : \mathbb{R}^d \to \mathbb{R}^d_+ \times \mathbb{R}^d_+$  is a diagonal matrix with entries  $\beta_i(\mathbf{Y}_t)$ , and  $\mathbf{W} = (W^{(1)}, \ldots, W^{(d)})$  is a vector of d correlated Brownian motions, which are also correlated with  $W^{(0)}$ . We denote by  $p_{ij}$  the correlations between the Brownian motions driving the volatility factors and by  $p_i$  the correlations between shocks to the factor  $\mathbf{Y}^{(i)}$  and shocks to the stock price:

$$d\langle W^{(i)}, W^{(j)} \rangle_t = p_{ij}dt, \ i, j = 1, \ \dots, \ d,$$
$$d\langle W^{(0)}, W^{(i)} \rangle_t = p_i dt, \ i = 1, \ \dots, \ d.$$

The total instantaneous correlation between stock price and volatility shocks is given by a weighted sum of the the  $p_i$ :

$$\operatorname{corr}(dX_t, d\sigma_t) = \sum_i \left(\frac{\frac{\partial f}{\partial \mathscr{Y}_i}(\mathbf{Y}_t)\beta_i(\mathbf{Y}_t)}{\sqrt{\sum_j \left(\frac{\partial f}{\partial \mathscr{Y}_i}(\mathbf{Y}_t)\beta_i(\mathbf{Y}_t)\right)^2}}\right) p_i$$

We again decompose each of the Brownian motions  $W^{(i)}$  as

$$W^{(i)} = p_i W^{(0)} + \sqrt{1 - p_i^2} W^{(i)\perp}$$

The risk-neutral measures introduced in subsection 2.2.9 can be generalized for the multifactor case. We note that the measures  $\mathbb{P}^{\star(\gamma)}$  are parametrized by a *d*-dimensional market price of volatility risk processes  $\gamma \in \mathbb{R}^d$  adapted to the filtration generated by the d+1 Brownian motions  $(W^{(0)}, W)$ . The Brownian processes  $(W^{(0)\star}, W^{(1)\perp\star}, \ldots, W^{(d)\perp\star})$  are defined under  $\mathbb{P}^{\star(\gamma)}$ , and  $W^{(0)\star}$  and ensure that the discounted stock price is a  $\mathbb{P}^{\star(\gamma)}$ -martingale. The adapted processes are given by

$$W_t^{(0)\star} = W_t^{(0)} + \int_0^t \frac{(\mu(\mathbf{Y}_s) - r)}{f(\mathbf{Y}_s)} ds,$$
  
$$W_t^{(i)\perp\star} = W_t^{(i)\perp} + \int_0^t \gamma_s^{(i)} ds, \ i = 1, \ \dots, \ d.$$

The multi-factor model dynamics under the risk-neutral measure become

$$\begin{cases} dX_t = rX_t dt + f(\mathbf{Y}_t) X_t dW_t^{(0)\star}, \\ d\mathbf{Y}_t = (\alpha(\mathbf{Y}_t) - \beta(\mathbf{Y}_t) \underline{\Lambda}_t) dt + \beta(\mathbf{Y}_t) d\mathbf{W}_t^{\star}, \\ W_t^{(i)\star} = p_i W_t^{(0)\star} + \sqrt{1 - p_i^2} W_t^{(i)\perp\star}, \end{cases}$$
(2.82)

where the vector of risk premium processes given by

$$\underline{\Lambda}_{t} = (\Lambda_{t}^{(1)}, \ldots, (\Lambda_{t}^{(d)}), \\ \Lambda_{t}^{(i)} = p_{i} \frac{(\mu(\mathbf{Y}_{t}) - r)}{f(\mathbf{Y}_{t})} + \gamma_{t}^{(i)} \sqrt{1 - p_{i}^{2}}, \ i = 1, \ d.$$

# 2.3 Related Work

In the previous section we introduced the class of stochastic models that we will be exploring further and provided enough background for an involved discussion and comparison of volatility models. In this section we take a detour from the mathematical and financial background and think about the underlying motivation and problem that has spurred research on volatility in different markets. Then we explore particularities of most notable model classes and provide a comparison based on empirical studies in the current literature. Finally, we conclude why we make the choice of volatility time scales (explained in the next section) and perturbation theory.

#### 2.3.1 (K,T,t) Problem and Benchmark For Implied Volatility Models

To set the criteria that makes a good volatility model, it is optimal to start from the model's use. Often called "market models", volatility models are one approach to pricing and hedging complex derivatives. This is opposed to modelling the evolution of time series of the derivatives' underlier (such as in GARCH-type models), which is inherently harder to extend to derivatives due to the use of complex econometric models. Therefore, volatility models are employed by practioners to help price and hedge more complex instruments in a noarbitrage way given available data on highly liquid traded options. From a financial industry viewpoint, this can be reformulated as: how can I map observed prices to parameters in a self-consistent and relatively easy, computationally tractable way.

As named in [2], the "(K,T,t)-problem" implies: for a given present time t and maturity time T, low-dimensional models can provide a good fit of the skew along the strikes' K axis, but are a lot harder to fit to the volatility term structure along the T-t time-to-maturity axis.

Fitting the term structure therefore requires high-dimensional models that can extract enough random factors and parameters to provide a good fit over both a range of maturities as well as a range of strikes. However, these models very easily suffer from overfitting, which leads to unstable parameters given negligible changes in the underlying data. The obvious example of overfitting are local volatility models that provide an infinite amount of flexibility and features but suffer from great relatively short-term parameter instability.

Given the above, the "holy grail" of volatility models would, among other things, offer a good implied surface fit and, at the same time, good parameter stability. This would mean that the model explains and acounts for the volatility dynamics present in traded options with great accuracy. Thus the (K,T,t) problem of finding balance between high-dimensionality and a good fit can be thought of as a problem of relevant feature extraction. An analogy can be made with the machine-learning technique of Principal Component Analysis (PCA) where the variance in high-dimensional data is explained up to a certain degree by a few major components. Similarly, in volatility analysis, we try to explain the current variance in the implied surface (the skew and term structure). In addition to explaining current variance, which a good model needs to do in a stable manner over time, it needs to identify explicitly or implicitly what exactly shapes and drives the volatility process over time. Multiple practical market factors enter this analysis, as we will explore further the effect of daily, annual, and maturity cycles on the volatility surface. Clearly, modeling directly the evolution of the implied volatility surface (as several model classes do by including stochastic volatility) has its advantages over other approaches (such as local volatility). The drawback to the approach taken is the difficulty of implementing an arbitrage-free model due to the introduced market incompleteness. This constitutes an important benchmark. Equally important is that the model serves as a platform that is easily extendable to other exotic contracts and that has a simple calibration procedure.

Where most practitioners depart, however, from the above criteria, is in their relaxation of the t problem (stability over time) and focus on an extremely accurate daily fit (K,T) problem. This is owed, in part, to traders' mistrust of non-perfect fits, even if this approach is not self-consistent and leads to a practice of daily recalibration and, thus, rehedging of risk positions.

In the course of the next chapters, I will assess model merits based equally on fit and stability, taking the approach most popular in academia, adopted by, among others, Fouque, Papanicolau, Sircar, or Joshi. However, my choice is not also motivated by a consideration that ranks high for practitioners: the transaction cost of daily rehedging or readjusting risk positions associated with pricing models with unstable parameters. To that extent, I quote a relevant and revealing example from [5]:

"The pricing of exotic options is not just about finding prices that are compatible with market dynamics in the sense of being non-arbitrageable; it is equally about realizing those prices via hedging. Thus if a model is to be useful to a trader it must tell him how to hedge and that the hedges must work.

Typically, the way a trader will hedge is to fit the model to the market and then to hedge each of the parameters by using simple options. Typically, an exotic option will be hedged using calls and puts of various maturities. Therefore after fitting the model to the vanilla market, the trader measures the derivative of the price with respect to each parameter of the model, possibly breaking up time into pieces to do this if he is using a variable-parameter model so as to get the exposure of the model to changes in the parameter over the various time slices. The trader then buys a portfolio of vanilla options so as to cancel all these exposures and uses the underlying at the end to remove any residual Delta. The trader returns a day later and repeats the process. The market will have changed a little in the meantime. If the market fit has also changed only a little, his hedge has been successful and the value of his portfolio has only changed a small amount. He need then only make small adjustments to his hedging portfolio to keep himself hedged. However, suppose he runs his fitting routine and it outputs a vastly different parameter-set. He then has a problem: although the market has not changed much his fitter is telling him to totally dissolve his original hedge and set up a new one. In addition, the new fit will probably give a wildly different price for the exotic option. The trader will be very unhappy at this and probably throw the model away (and shoot the quant!).

This means that an important criterion for trading off a model is that it should fit the market stably. That is, if one changes the market slightly, the fit should also change slightly."

#### 2.3.2 Model Selection

In general, model selection is not trivial, although a few books and empirical studies offer some degree of direct comparison especially between model classes. [1] is the authorative text on implied volatility surface dynamics, and [5] offers a good overview of model classes, so we compare intuitively different models and their characteristics based on the two sources, completing our observations with existing empirical evidence.

Among popular classes of models employed both in academic research and by practitioners we can find the following:

- **Derman-Kani or Dupire type models** They are based on local volatility, and calibrated using discrete trees or a closed-form solution using interpolation. Although they provide an excellent fit, their disadvantages compared to stochastic models are explained in subsection 2.2.4.
- Jump-diffusion models They assume a constant volatility and instead focus on modelling random underlying asset jumps, following an equation of the type

$$dX_t = \mu X_t dt + \sigma X_t dW_t + (J-1)X_t dq$$

where dq is a Poisson process independent of the Brownian motion  $W_t$ . To find a closedform solution to this class of models we must make the unrealistic assumption that the jump size J is known in advance. In practice, we would need to specify a distribution of jump sizes that creates an additional hedging asset for each new jump value in the no-arbitrage replication argument. Thus we would need an infinite number of hedging assets making a perfect replicating hedge impossible. This a major drawback compared to the other model classes. Another detrimental feature of the model is that it decays very fast in skewness across maturities. The effects of jumps vanish very quickly as we increase time-to-expiry, as visible in figure 2.7, leaving an unrealistic Black-Scholes-like



FIGURE 2.7: The term structure of at-the-money(ATM) variance skew for various choices of jump diffusion parameters

model for long maturities. Therefore, jump-diffusion models are not a good choice for modelling the implied term structure in its *t*dimension.

• Stochastic Volatility (SV) models Detailed mathematically in section 2.2, simple, one-factor volatility models, as [1] points out, don't fit the observed implied volatility surface for short expirations, but do well for long-term expirations, in contrast to jump-diffusion models. The class includes the Heston model - very popular mainly due to its advantage in having a quasi-closed form solution that is computationally efficient, unlike all other models.

Despite its shortcomings, stochastic volatility offers the posibility of short-term fit improvement by using multiple factors and scales of volatility, which unfortunately leads to a near-impossible closed-form derivation. The work of [2] that this thesis builds on, however, exploits perturbation theory to derive asymptotic approximations for a multi-scale two-factor model. As the empirical work in the next chapters and the multiscale expansion of Heston in [14] shows, the model can be gradually improved in fit and stability and offers further scope to do so. Moreso, it offers a relatively fast calibration and presents you with model-independent parameters for both pricing and hedging exotics.

• Stochastic Volatility with Jumps (SVJ) models The contrasting results of stochstic volatility asset jumps naturally lead to a hybrid approach. We can extend the Heston model for example to include jumps, or even simultaneous jumps in asset and volatility (SVJJ models). [1] notices SVJ clearly ourperforms Heston or SVJJ by producing a better fit. However, the empirical study [15] agrees that SVJ outperforms in terms of

fits, but concludes that the simple SV model offers the best hedging results out of all models. This brings further grounds to try and improve the SV model without jumps, as this thesis proceeds.

• Variance Gamma models This class of models differs considerably from the others in linking the volatility process to new information arrival. As [5] describes - "The idea is that the volatility should be a measure of a stock's sensitivity to information as it arrives, but the amount of information arriving is random also, and needs to be described by a random process itself. One can think of trading volume as a proxy for information arrival, and there is some statistical evidence that stock price returns are more log-normal when rescaled to use trading volume for the time parameter instead of calendar time."

The process lends well to quick Monte Carlo or numeric integral calibration. However the departure from Brownian motion gives the process only jump movements, which determine a fast skew decay for long maturities and also leave the hedging problem unsolved. Variance Gamma models offer a fresh perspective on implied volatility modelling, but are also prone to the downsides of jump models.

• Stochastic Local Volatility (SLV or LSV) models Presented at the beginning of section 2.2, stochastic volatility models were originally introduced as a hybrid that takes advantage of local volatility to fit the implied surface and, in this framework, specific SV dynamics. According to [16] the model calibration is done by first computing the Dupire local volatility on the grid of (K, T-t) and then calibrating the local stochastic volatility on that grid. To price exotic options, backward PDE methods or Monte-Carlo simulations can be used. However, the model is not popular in equity markets due to the convexity of the skew that it creates, with gives it a tendency to underprice volatility products. This shortcoming can be remediated by introducing jumps, though such an approach makes the model cumbersome and its SV parameters hard to interpret. We choose not to pursue this class of models due to their bad scaling to equity markets unless jumps are included, majorly increasing the difficulty of an analytical interpretation. However, SLV models may offer further scope of research as an extension to the SV models and perturbation expansion considered in this thesis.

#### 2.3.3 Summary

Motivated by the (K,T,t) problem, and considering the array of implied volatility model classes, we further pursue a generalized two-factor SV model, each factor being characterized by a different time scale (see next section). Using perturbation techniques we obtain approximations for its parameters and analyse the parameter stability and fit of the model specifications that arise. We start with an SV model characterized by worse fits on shorter maturities (a property shared by all SV models) and improve it in succesive expansions.

### 2.4 Volatility Time Scales

In this section we provide empirical and mathematical background to motivate and describe the use of different mean-reverting volatility time scales for the two volatility factors in our SV model. We follow the self-contained explanations in [2], giving reference to relevant empirical studies.

#### 2.4.1 Mean Reversion and Markov Process Time Scales

Numerous studies among which Merville & Piptea, find implied volatility to be strongly mean-reverting, following a mixed mean-reverting diffusion process with discrete white noise. We build on the intuitive introduction made in 2.2.7 and provide a brief mathematical characterization (for a more involved discussion of the underlying mathematics please refer to [17] or [3]).

In the context of a stochastic volatility framework, we want the volatility process to replicate the (in the stochastic future time-to-expity) integrated square volatility, as given by (2.57):

$$\overline{\sigma^2} = \frac{1}{T-t} \int^T \sigma^2(s) ds.$$
(2.83)

Therefore we need to characterize and parametrize the time scales of fluctuation that govern the evolution of  $\overline{\sigma^2}$  over time.

Mean-reversion is closely related to the concept of ergodicity, which we define as in [2]: A process  $(Y_t)$  is called ergodic if it admits a unique invariant (or stationary) distribution, denoted by  $\Pi$ , and the long-time time average of any measurable bounded function g of the process converges almost surely (a.s.) to the deterministic average with respect to its invariant distribution:

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t g(\mathbf{Y}_s) ds = \int g(\mathbf{y}) \Pi(d\mathbf{y}) \stackrel{def}{=} \langle g \rangle.$$
(2.84)

The invariant (or stationary) distribution  $\Pi$  is defined as an initial distribution for Y<sub>0</sub>so that  $\forall t > 0 \ Y_t$  has the same distribution. It can be interpreted as a generalization of the classical law of large numbers for a sum of i.i.d random variables.

In order to parameterize a time-homogenous Markov Process  $(Y_t)$ , we define its transition semi-group  $P_t$ , acting on the bounded function g introduced above, that satisfies:

$$P_t g(y) = E\{g(Y_t) | Y_0 = y\}.$$
(2.85)

It can be shown that the transition semi-group is time-homogenous too and has the Markov property. For  $\kappa > 0$ , we introduce  $P_t^{(\kappa)} = P_{\kappa t}$  - the transition semigroup of a process  $(Y_t^{(\kappa)})$  that has the same distribution as  $(Y_{\kappa t})$ :

$$E\{g(Y_t^{(\kappa)})|Y_0^{(\kappa)} = y\} = P_t^{(\kappa)}g(y) = P_{\kappa t}g(y) = E\{g(Y_{\kappa t})|Y_0 = y\}.$$
(2.86)

Thus the process  $Y^{(\kappa)}$  evolves exactly like the process Y, but transitions  $\kappa$  times as fast. We can also see the role of  $\kappa$  as a multiplicative factor by defining the infinitisemal generator of  $(Y_t)$  as

$$\mathscr{L}g(\mathbf{y}) = \lim_{t \downarrow 0} \frac{\mathbf{P}_t g(\mathbf{y}) - g(\mathbf{y})}{t},$$
(2.87)

and obtaining

$$\mathscr{L}^{(\kappa)}g(\mathbf{y}) = \lim_{t \downarrow 0} \frac{\mathbf{P}_t^{(\kappa)}g(\mathbf{y}) - g(\mathbf{y})}{t} = \kappa \lim_{t \downarrow 0} \frac{\mathbf{P}_{\kappa t}g(\mathbf{y}) - g(\mathbf{y})}{\kappa t} = \kappa \mathscr{L}g(\mathbf{y})$$
(2.88)

where  $\mathscr{L}^{(\kappa)}$  is the infinitisemal generator of the process  $\mathbf{Y}^{(\kappa)}$ .

It can also be proven that the invariant distributions for Y and  $Y^{(\kappa)}$  are uniquely determined and equal, if they exist. We restrict the Markov process Y to be reversible (i.e. the eigenvalues  $\lambda_i$  of the process generator  $\mathscr{L}$  follow  $0 = \lambda_0 > \lambda_1 > \lambda_2 > \cdots$ ). Using the last condition and eigenfunction expansions (see ...), it can we shown that  $Y_t$  converges exponentially to its invariant distribution at a rate governed by  $\lambda = |\lambda_1|$  (its spectral gap). From equation (2.88), we see that the spectral gap of  $\mathscr{L}^{(\kappa)} = \kappa \mathscr{L}g(y)$  will be  $\kappa\lambda$ , so  $Y^{(\kappa)}$  reverts to its inveriant distribution at an exponential rate  $\kappa\lambda$ .

For an illustration of the effects of varying  $\kappa$  see figures 2.8 and 2.9. As we can observe qualitatively, a higher  $\kappa$ , i.e. a faster rate of mean reversion leads to more frequent high amplitude osciallations in the volatility process, regardless of the choice of the driving factor models.

#### 2.4.2 Time scales observed in market data

In this subsection we explore the empirical grounds for a multi-scale approach to modelling volatility. The multi-factor approach is motivated by an in-depth study of one-factor SV models performance [Gatheral], that points out the limitations of one-factor SV models in fitting the implied surface are similar, regardless of the driving stochastic process. As we can see in figure 2.10, existing empirical work has looked at volatility characteristics for large indexes (S&P 500, DJIA) as a function of demeaned-returns:

$$\frac{dX_t}{X_t} - \mu dt = \sigma_t dW_t, \tag{2.89}$$



FIGURE 2.8: Simulated paths of  $\sigma_t = Y_t$ , with  $(Y_t)$  a mean-reverting CIR process (source: [2])



FIGURE 2.9: Simulated volatility and corresponding return paths for small and large rates of mean-reversion for the OU model (source: [2])



FIGURE 2.10: De-meaned Dow Jones Industrial Average (DJIA) returns over years 1900-2000 [source: svol]

For instance, [18] establishes the split of short-term and long-term traders, giving evidence and insight into a heterogenous market acting on multiple time scales. In addition, an extensive study in [19] focuses on identifying and quantifying the mean rate of fast reversion using S&P data. The study models the normalized fluctuations (or demeaned returns) process as

$$\overline{D}_n = \sigma_n \epsilon_n.$$

Next, the study defines the variagram model as

$$V_j^N = \frac{1}{N} \sum_n (F_{n+j} - F_n)^2$$

where  $F_n = \log |\overline{D}_n| = \log \sigma_n + \log |\epsilon_n|$ , *j* is the lag, and *N* the total number of points. Fitting the variogram by a three-parameter model

$$V_j^N \approx 2\gamma^2 + 2v^2(1 - e^{-\kappa j\Delta t})$$

the [short time scale in S&P] obtains a fit as in figure 2.11, estimating a short mean-reversion time  $1/\kappa = 1.7$  days for the S&P 500 volatility. Analysing the plot of simulated paths using the found mean-reversion rate vs. the S&P 500 realized volatility paths in figure 2.12, we can see the similarities between the two in terms of their normalized fluctuations.

Finally, it is worthy to mention the so-called 'day-effect' of volatility (or intraday volatility variability), illustrated in figure 2.13. However, it is remarkably stable over time and its inclusion in the perturbation expansion does not modify the two-factor model calibration, according to [2].



FIGURE 2.11: Empirical variogram of the S&P 500 (dashed line) and exponential fit (solid line) from which  $\kappa$  is found, for 1999-2000 data.(source: [2])



FIGURE 2.12: The first two figures show simulated paths of  $D_n$  for  $\kappa = 1$  and  $\kappa = 250$ , respectively. The third figure shows S&P 500's normalized fluctuation process from the first 60 trading days of 1994 (source: [2])



FIGURE 2.13: Intraday variability in 1999-2000 (dashed line is a 4-parameter model fit). Notice that volatility is highest at the beginning of the trading day, with a rebound at the end. [source: short time scale in S&P]

#### 2.4.3 Application to stochastic volatility models

Coming back to equation (2.83), by modeling  $\sigma_t = f(\mathbf{Y}_t^{(\kappa)})$ , where Y is mean-reverting and  $f^2$  is  $\Pi$ - integrable we can write

$$\begin{split} \overline{\sigma^2} &= \frac{1}{(T-t)} \int^T f^2(\mathbf{Y}_s^{(\kappa)}) ds \\ &= \frac{1}{(T-t)} \int^T f^2(\mathbf{Y}_{\kappa s}) ds \\ &= \frac{1}{\kappa(T-t)} \int_{\kappa t}^{\kappa T} f^2(\mathbf{Y}_s) ds \\ &= \frac{1}{(T-t)} (T \frac{1}{(\kappa T)} \int_0^{\kappa T} f^2(\mathbf{Y}_s) ds - t \frac{1}{(\kappa t)} \int_0^{\kappa t} f^2(\mathbf{Y}_s) ds). \end{split}$$

The above implies:

1. For fast mean-reversion -  $\kappa$  large

$$\lim_{\kappa \uparrow \infty} \overline{\sigma^2} = \int_{\Omega} f^2(\mathbf{y}) \Pi(dy) \tag{2.90}$$

, where  $\Omega$  is the state space of  $(Y_t) \Rightarrow \overline{\sigma^2}$  converges to a constant in distribution.

2. For slow mean-reversion -  $\kappa$  small

$$\lim_{\kappa \downarrow 0} \frac{1}{(T-t)} \int^T f^2(\mathbf{Y}_{\kappa s}) ds = f^2(\mathbf{Y}_t)$$
(2.91)

 $\Rightarrow \overline{\sigma^2}$  converges to the spot square volatility (assuming f is continuous and bounded).

The two-factor two-scale model we experiment with further is defined by the system of SDE's under real-world measure  $\mathbb{P}$ :

$$\begin{cases} dX_t = \mu(\mathbf{Y}_t, Z_t) X_t dt + (\mathbf{Y}_t, Z_t) X_t dW_t^{(0)}, \\ dY_t = \frac{1}{\epsilon} \alpha(\mathbf{Y}_t) dt + \frac{t_1 +}{\sqrt{\epsilon}} \beta(\mathbf{Y}_t) dW_t^{(1)} \\ dZ_t = \delta c(Z_t) dt + \sqrt{\delta} g(Z_t) dW_t^{(2)}, \end{cases}$$

$$(2.92)$$

where  $Y_t$  and  $Z_t$  model a fast factor with a high factor  $\kappa = 1/\epsilon(\epsilon > 0, \text{ small})$  and a slow factor with a low  $\kappa = \delta(\delta > 0, \text{ small})$  respectively. The correlation structure between the Brownian motions  $(W_t^{(0)}, W_t^{(1)}, W_t^{(2)})$  (vital for the skew effect) is given by

$$\begin{cases} d\langle W^{(0)}, W^{(1)} \rangle_t = p_1 dt, \\ d\langle W^{(0)}, W^{(2)} \rangle_t = p_2 dt, \\ d\langle W^{(1)}, W^{(2)} \rangle_t = p_{12} dt, \end{cases}$$
(2.93)

where  $|p_1| < 1$ ,  $|p_2| < 1$ ,  $|p_{12}| < 1$ , and  $1 + 2p_1p_2p_{12} - p_1^2 - p_2^2 - p_{12}^2 > 0$ . The latter condition ensures covariance matrix of the Brownian motions remains positive definite. Note that the underlying driving processes for the volatility factors are parameterized by  $(\alpha, \beta)$  and (c, g)respectively, and no factor specific process (i.e. OU, CIR) is chosen, creating a factor model - independent framework.

# Chapter 3

# First order perturbation: Outline and Implementation Results

# 3.1 Model Setup

We rewrite 2.92 using the same change of measure described in Chapter 2 to yield the following model form under the risk-neutral measure  $\mathbb{P}^*$ :

$$\begin{cases} dX_t = rX_t dt + f(\mathbf{Y}_t, Z_t) X_t dW_t^{(0)\star}, \\ dY_t = (\frac{1}{\epsilon} \alpha(\mathbf{Y}_t) - \frac{1}{\sqrt{\epsilon}} \beta(\mathbf{Y}_t) \Lambda_1(\mathbf{Y}_t, Z_t)) dt + \frac{1}{\sqrt{\epsilon}} \beta(\mathbf{Y}_t) dW_t^{(1)\star}, \\ dZ_t = (\delta c(Z_t) - \sqrt{\delta} g(Z_t) \Lambda_2(\mathbf{Y}_t, Z_t)) dt + \sqrt{\delta} g(Z_t) dW_t^{(2)\star}, \end{cases}$$
(3.1)

Again the correlation structure between the Brownian motions  $(W_t^{(0)\star}, W_t^{(1)\star}, W_t^{(2)\star})$  is given by

$$\begin{cases}
d\langle W^{(0)\star}, W^{(1)\star} \rangle_t = p_1 dt, \\
d\langle W^{(0)\star}, W^{(2)\star} \rangle_t = p_2 dt, \\
d\langle W^{(1)\star}, W^{(2)\star} \rangle_t = p_{12} dt,
\end{cases}$$
(3.2)

where  $|p_1| < 1$ ,  $|p_2| < 1$ ,  $|p_{12}| < 1$ , and  $1 + 2p_1p_2p_{12} - p_1^2 - p_2^2 - p_{12}^2 > 0$ . Notes and assumptions on model parameters:

- The volatility function f of asset X must be positive, smooth in z and  $f^2(\cdot, z)$  must be integrable with respect to the invariant distribution of Y
- The instantaneous interest rate r is constant

- The path of  $(X_t)$  always describes positive values
- The coefficients α(y) and β(y) describe the dynamics of the mean-reverting diffusion process Y and c(z) and g(z) describe the dynamics of Z respectively, all under the physical measure P. The process Y has a unique invariant distribution denoted by Φ. Finally, the perturbation expansion does not depend on any particular form of these coefficients.
- The functions Λ<sub>1</sub>(y, z) and Λ<sub>2</sub>(y, z) are the combined market prices of volatility risk which determine the risk-neutral pricing measure P<sup>\*</sup>, as introduced in the context of one-factor models in equation 2.73. They are chosen as functions of Markov processes (Y, Z) so that the triple (X,Y,Z) remains a Markov process under P<sup>\*</sup>.

To explicitly show the dependency on the two small parameters  $\epsilon$  and  $\delta$ , we denote the price of an European option as  $P^{\epsilon,\delta}(t, x, y, z)$ , given by

$$P^{\epsilon,\delta}(t, X_t, Y_t, Z_t) = \mathbb{E}^{\star} \{ e^{-r(T-t)} h(X_T) | X_t, Y_t, Z_t \},$$
(3.3)

where h(x) is a smooth an bounded payoff function for an option with maturity T and E<sup>\*</sup> is the expectation under the P<sup>\*</sup> measure. The above price expression depends on the parameter r and functions  $(f, \alpha, \beta, c, g, \Lambda_1, \Lambda_2)$ , which are difficult to estimate and required in order to price a contract using model (3.1).

# 3.2 Derivation outline

Using the multidimensional Feynman-Kac formula (see appendix) we can write  $P^{\epsilon,\delta}$  as the solution of

$$\begin{cases} \frac{\partial P^{\epsilon,\delta}}{\partial t} + \mathscr{L}_{(X,Y,Z)} P^{\epsilon,\delta} - r P^{\epsilon,\delta} = 0, \\ P^{\epsilon,\delta}(T,x,y,z) = h(x), \end{cases}$$
(3.4)

where  $\mathscr{L}_{(X,Y,Z)}$  defines the infinitesimal generator of the Markov process  $(X_t, Y_t, Z_t)$ . Defining the infinitisemal operator

$$\mathscr{L}^{\epsilon,\delta} = \frac{\partial}{\partial t} + \mathscr{L}_{(X,Y,Z)} - r$$

the above system can be rewriten as

$$\begin{cases} \mathscr{L}^{\epsilon,\delta} P^{\epsilon,\delta} = 0, \\ P^{\epsilon,\delta}(T, x, y, z) = h(x), \end{cases}$$

$$(3.5)$$

Now we decompose (expand)  $\mathscr{L}^{\epsilon,\delta}$  into component operators factored by the powers of  $\epsilon$  and  $\delta$  in the form

$$\mathscr{L}^{\epsilon,\delta} = \frac{1}{\epsilon}\mathscr{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathscr{L}_1 + \mathscr{L}_2 + \sqrt{\delta}\mathscr{M}_1 + \delta\mathscr{M}_2 + \sqrt{\frac{\delta}{\epsilon}}\mathscr{M}_3, \tag{3.6}$$

where

$$\mathcal{L}_{0} = \frac{1}{2}\beta^{2}(y)\frac{\partial^{2}}{\partial y^{2}} + \alpha(y)\frac{\partial}{\partial y},$$
  

$$\mathcal{L}_{1} = \beta(y)(p_{1}f(y,z)x\frac{\partial^{2}}{\partial x\partial y} - \Lambda_{1}(y,z)\frac{\partial}{\partial y}),$$
  

$$\mathcal{L}_{2} = \frac{\partial}{\partial t} + \frac{1}{2}f^{2}(y,z)x^{2}\frac{\partial^{2}}{\partial x^{2}} + r(x\frac{\partial}{\partial x} - \cdot),$$
  

$$\mathcal{M}_{1} = g(z) \ (p_{2}f(y,z)x\frac{\partial^{2}}{\partial x\partial z} - \Lambda_{2}(y,z)\frac{\partial}{\partial z}),$$
  

$$\mathcal{M}_{2} = \frac{1}{2}g^{2}(z)\frac{\partial^{2}}{\partial z^{2}} + c(z)\frac{\partial}{\partial z},$$
  

$$\mathcal{M}_{3} = \beta(y)p_{12}g(z)\frac{\partial^{2}}{\partial y\partial z}.$$
  
(3.7)

The expressions of the generators introduced showcase the roles they play in the model dynamics [2]:

- $\epsilon^{-1}\mathscr{L}_0$  is the infinitesimal generator of the process Y under the physical measure  $\mathbb{P}$
- $\mathscr{L}_1$  contains the mixed derivative due to the covariation between X and Y, and the first derivative with respect to y due to the market price of volatility risk  $\Lambda_1$
- $\mathscr{L}_2$  contains the time derivative and is the Black-Scholes operator at the volatility level  $f(\mathbf{y}, z)$ , also denoted by  $\mathscr{L}_{BS}(f(\mathbf{y}, z))$
- $\mathcal{M}_1$  contains the mixed derivative due to the covariation between X and Z, and the first derivative with respect to z due to the market price of volatility risk  $\Lambda_2$
- $\delta \mathscr{M}_2$  is the infinitesimal generator of process Z under the physical measure  $\mathbb{P}$
- $\mathcal{M}_3$  contains the mixed derivative due to the covariation between Y and Z

In the expansions that follow, we assume that the payoff function is smooth and bounded. In reality it is non-differentiable at the strike for options that are close to maturity (due to the hockey stick shape). This case is much more involved (terminal layer analysis), and a detailed presentation can be found in [20].

#### 3.2.1 Regular perturbation expansion (around $\delta$ )

The terms associated with  $\delta$  are small when  $\delta \to 0$  and give rise to a regular perturbation problem. Thus we initially expand  $P^{\epsilon,\delta}$  in powers of  $\sqrt{\delta}$ :

$$P^{\epsilon,\delta} = P_0^{\epsilon} + \sqrt{\delta}P_1^{\epsilon} + \delta P_2^{\epsilon} + \cdots$$
(3.8)

We insert the above expansion in equations (3.6) (both the PDE and terminal condition), grouping generators in terms of powers of  $\delta$  in the form

$$\left(\frac{1}{\epsilon}\mathscr{L}0 + \frac{1}{\sqrt{\epsilon}}\mathscr{L}_1 + \mathscr{L}_2\right)P_0^{\epsilon} + \sqrt{\delta}\left\{\left(\frac{1}{\epsilon}\mathscr{L}0 + \frac{1}{\sqrt{\epsilon}}\mathscr{L}_1 + \mathscr{L}_2\right)P_1^{\epsilon} + \left(\mathscr{M}_1 + \frac{1}{\sqrt{\epsilon}}\mathscr{M}_3\right)P_0^{\epsilon}\right\} + \dots = 0$$
(3.9)

Equating the independent term and the factor of  $\sqrt{\delta}$  with 0, we can define  $P_0^{\epsilon}$ , and  $P_1^{\epsilon}$  respectively, as the unique solutions to the problems:

$$\begin{cases} \left(\frac{1}{\epsilon}\mathscr{L}_{0} + \frac{1}{\sqrt{\epsilon}}\mathscr{L}_{1} + \mathscr{L}_{2}\right)P_{0}^{\epsilon} = 0\\ P_{0}^{\epsilon}(T, x, y, z) = h(x) \end{cases}$$
(3.10)

$$\begin{cases} \left(\frac{1}{\epsilon}\mathscr{L}_{0} + \frac{1}{\sqrt{\epsilon}}\mathscr{L}_{1} + \mathscr{L}_{2}\right)P_{1}^{\epsilon} = -\left(\mathscr{M}_{1} + \frac{1}{\sqrt{\epsilon}}\mathscr{M}_{3}\right)P_{0}^{\epsilon},\\ P_{1}^{\epsilon}(T, x, y, z) = 0\end{cases}$$
(3.11)

where the terminal payoff is assigned to the independent term  $P_0^{\epsilon}$ . Next, we expand  $P_0^{\epsilon}$  and  $P_1^{\epsilon}$  in powers of  $\sqrt{\epsilon}$ .

#### 3.2.2 Singular perturbation expansion (around $\epsilon$ )

The terms associated with  $\epsilon$  are diverging when  $\epsilon \to 0$  and give rise to regular perturbation problem. We first expand  $P_0^{\epsilon}$  in powers of  $\sqrt{\epsilon}$ :

$$P_0^{\epsilon} = P_0 + \sqrt{\epsilon} P_{1,0} + \epsilon P_{2,0} + \epsilon^{3/2} P_{3,0} + \cdots .$$
(3.12)

Inserting the above expansion in the first equation of (3.10) we obtain

$$\frac{1}{\epsilon}\mathscr{L}_{0}P_{0} + \frac{1}{\sqrt{\epsilon}}(\mathscr{L}_{0}P_{1,0} + \mathscr{L}_{1}P_{0}) + (\mathscr{L}_{0}P_{2,0} + \mathscr{L}_{1}P_{1,0} + \mathscr{L}_{2}P_{0}) + \sqrt{\epsilon}(\mathscr{L}_{0}P_{3,0} + \mathscr{L}_{1}P_{2,0} + \mathscr{L}_{2}P_{1,0}) + \dots = 0$$
(3.13)

Equating all factors of powers of  $\sqrt{\epsilon}$  with 0, we obtain multiple equations that are solved using Poisson equations and their centering conditions (see [2] for more details), yielding:

•  $P_0(t, x, z) = P_{BS}(t, x; \overline{\sigma}(z))$  where the volatility  $\overline{\sigma}(z)$  is the averaged effective volatility given by

$$\overline{\sigma}^2(z) = \langle f^2(\cdot, z) \rangle = \int f^2(\mathbf{y}, z) \Phi(dy), \qquad (3.14)$$

and  $\overline{\sigma'} = \partial_z \overline{\sigma}$ .

•  $P_{1,0}^{\epsilon} \stackrel{def}{=} \sqrt{\epsilon} P_{1,0}$  is the term corresponding to the **fast scale correction** and is given by

$$P_{1,0}^{\epsilon}(t,x,z) = -(T-t)\left(-V_3^{\epsilon}(z)D_1D_2 - V_2^{\epsilon}(z)D_2\right)P_{BS}(t,x;\overline{\sigma}(z)),$$
(3.15)

where the fast-scale group parameters  $V_3^{\epsilon}(z)$  and  $V_2^{\epsilon}(z)$  are given by

$$V_{3}^{\epsilon}(z) = -\frac{p_{1}\sqrt{\epsilon}}{2} \langle \beta f(\cdot, z) \frac{\partial \phi}{\partial \mathbf{y}}(\cdot, z) \rangle, \qquad (3.16)$$

$$V_2^{\epsilon}(z) = \frac{\sqrt{\epsilon}}{2} \langle \beta \Lambda_1(\cdot, z) \frac{\partial \phi}{\partial y}(\cdot, z) \rangle, \qquad (3.17)$$

and

$$D_k = x^k \frac{\partial^k}{\partial x^k}.$$
(3.18)

Next, we expand  $P_1^{\epsilon}$  in powers of  $\sqrt{\epsilon}$ :

$$P_1^{\epsilon} = P_{0,1} + \sqrt{\epsilon} P_{1,1} + \epsilon P_{2,1} + \epsilon^{3/2} P_{3,1} + \cdots .$$
(3.19)

Inserting the above expansion in the first equation of (3.11) we obtain

$$\frac{1}{\epsilon} \mathscr{L}_0 P_{0,1} + \frac{1}{\sqrt{\epsilon}} (\mathscr{L}_0 P_{1,1} + \mathscr{L}_1 P_{0,1} + \mathscr{M}_3 P_0) + (\mathscr{L}_0 P_{2,1} + \mathscr{L}_1 P_{1,1} + \mathscr{L}_2 P_{0,1}) + \mathscr{M}_1 P_0 + \mathscr{M}_3 P_{1,0} \\
+ \sqrt{\epsilon} (\mathscr{L}_0 P_{3,1} + \mathscr{L}_1 P_{2,1} + \mathscr{L}_2 P_{1,1} + \mathscr{M}_1 P_{1,0} + \mathscr{M}_3 P_{2,0}) + \dots = 0.$$
(3.20)

Equating all factors of powers of  $\sqrt{\epsilon}$  with 0, we again obtain a system of equations that yields:

•  $P_{0,1}^{\delta} \stackrel{def}{=} \sqrt{\delta} P_{0,1}$  is the term corresponding to the **slow scale correction** and is given by

$$P_{0,1}^{\delta} = (T-t) \left( V_0^{\delta}(z) \frac{\partial}{\partial \sigma} + V_1^{\delta}(z) D_1 \frac{\partial}{\partial \sigma} \right) P_{BS}.$$
(3.21)

where the slow-scale group parameters  $V_0^\delta$  and  $V_1^\delta$  are given by

$$V_0^{\delta}(z) = -\frac{g(z)\sqrt{\delta}}{2} \langle \Lambda_2(\cdot, z) \rangle \overline{\sigma}'(z), \qquad (3.22)$$

$$V_1^{\delta}(z) = \frac{p_2 g(z) \sqrt{\delta}}{2} \langle f(\cdot, z) \rangle \overline{\sigma}'(z).$$
(3.23)

#### 3.2.3 Price expansion formula

Adding together the first order expansion terms  $P(0), P_{1,0}^{\epsilon}, P_{0,1}^{\delta}$ , we get an approximation of an European option price as

$$P^{\epsilon,\delta} \approx \tilde{P}^{\epsilon,\delta} := P_{BS} + P_{1,0}^{\epsilon} + P_{0,1}^{\delta}.$$
(3.24)

Given the price terms formulas found earlier in the previous subsection, the price expansion can be written as a formula dependent on the five market group parameters  $\{\overline{\sigma}(z), V_0^{\delta}(z), V_1^{\delta}(z), V_2^{\epsilon}(z), V_3^{\epsilon}(z)\}$ . The price approximation then takes the form:

$$\tilde{P}^{\epsilon,\delta} = P_{BS} + (T-t) \left[ V_0^{\delta}(z) \frac{\partial}{\partial \sigma} + V_1^{\delta}(z) D_1(\frac{\partial}{\partial \sigma}) + V_2^{\epsilon}(z) D_2 + V_3^{\epsilon}(z) D_1 D_2 \right] P_{BS}.$$
(3.25)

#### 3.2.4 Parameter reduction and expansion accuracy

To obtain a numerical result for the price expansion we need to estimate the five market group parameters  $\{\overline{\sigma}(z), V_0^{\delta}(z), V_1^{\delta}(z), V_2^{\epsilon}(z), V_3^{\epsilon}(z)\}$ . The volatility can theoretically be calculated from historical returns data, while the other four parameters can be estimated from options data.

However, as [2] explains, it is desirable to not rely on historical returns data, mainly due to the complications of estimating historical volatility relative to two volatility time scales. Also,  $V_2^{\epsilon}(z)$ can be interpreted according to its formula in (??) that is dependent on market price of volatility risk, as a volatility level correction. Based on the last two arguments, we "absorb"  $\overline{\sigma}$  and  $V_2^{\epsilon}$  into a corrected volatility level  $\sigma^*$ , given by:

$$\sigma^{\star}(z) = \sqrt{\overline{\sigma}^2(z) + 2V_2^{\epsilon}(z)}.$$
(3.26)

Propagating the changes to the price approximation, (3.24) becomes

$$\tilde{P^{\star}} = P^{\star}_{BS} + (T-t) \left( V^{\delta}_0(z) \frac{\partial P^{\star}_{BS}}{\partial \sigma} + V^{\delta}_1(z) D_1 \frac{\partial P^{\star}_{BS}}{\partial \sigma} + V^{\epsilon}_3(z) D_1 D_2 P^{\star}_{BS} \right),$$
(3.27)

or, expressed in terms of the Greeks (see subsection 2.1.9)

$$\tilde{P^{\star}} = P_{BS}^{\star} + (T - t) \left( V_0^{\delta}(z)7 + V_1^{\delta}(z)x\Delta7 + V_3^{\epsilon}(z)x\Delta(x^2\Gamma) \right).$$
(3.28)

The fast time scale contribution to  $\tilde{P}^{\star}$  is given by parameters  $\sigma^{\star}$  and  $V_3^{\epsilon}$  which account for historical volatility and market price of risk, and the effect of the correlation (or skew)  $p_1$ .

A final point is the accuracy of the first order price approximation, though an outline proof is outside the scope of this thesis. As proved in [20], we have:

$$P^{\epsilon,\delta} = P^* + \theta \left(\epsilon \log |\epsilon| + \delta\right). \tag{3.29}$$

# 3.3 Model Calibration

Denoting  $\tau = T - t$  and using 2.37 we can rewrite equation (3.27) as

$$P^{\star} = P^{\star}_{BS} + \left(\tau V_0^{\delta} + \tau V_1^{\delta} D_1 + \frac{V_3^{\epsilon}}{\sigma^{\star}} D_1\right) \frac{\partial P^{\star}_{BS}}{\partial \sigma}.$$
(3.30)

We then set  $P^{\star} = C_{BS}(I)$  (the Black-Scholes call price at volatility I), trying to match a similar expansion of the implied volatility:

$$I \approx \sigma^* + \sqrt{\epsilon} I_{1,0} + \sqrt{\delta} I_{0,1}. \tag{3.31}$$

Fitting the expansion terms of 3.31 to 3.30, after a few calculations (see [2] for more details) we obtain the simple form:

$$I \approx b^{\star} + \tau b^{\delta} + \left(a^{\epsilon} + \tau a^{\delta}\right) \text{LMMR.}$$
(3.32)

where the parameters  $(b^{\star}, b^{\delta}, a^{\epsilon}, a^{\delta})$  are functions of  $(\sigma^{\star}, V_0^{\delta}, V_1^{\delta}, V_3^{\epsilon})$  given by

$$\begin{cases}
b^{\star} = \sigma^{\star} + \frac{V_{3}^{\epsilon}}{2\sigma^{\star}} \left(1 - \frac{2r}{\sigma^{\star}2}\right) \\
b^{\delta} = V_{0}^{\delta} + \frac{V_{1}^{\delta}}{2} \left(1 - \frac{2r}{\sigma^{\star}2}\right) \\
a^{\epsilon} = \frac{V_{3}^{\epsilon}}{\sigma^{\star}3} \\
a^{\delta} = \frac{V_{1}^{\delta}}{\sigma^{\star}2}
\end{cases}$$
(3.33)

LMMR stands for log-moneyness to maturity ratio and is given by

$$LMMR = \frac{\log(K/x)}{T-t} = \frac{\log(K/x)}{\tau}$$
(3.34)

To recover the group parameters  $(\sigma^{\star}, V_0^{\delta}, V_1^{\delta}, V_3^{\epsilon})$  we just need to invert the  $(b^{\star}, b^{\delta}, a^{\epsilon}, a^{\delta})$  formulas:

$$\sigma^{\star} = b^{\star} + a^{\epsilon} \left( r - \frac{b^{\star 2}}{2} \right)$$

$$V_0^{\delta} = b^{\delta} + a^{\delta} \left( r - \frac{b^{\star 2}}{2} \right)$$

$$V_1^{\delta} = a^{\delta} b^{\star 2}$$

$$V_3^{\epsilon} = a^{\epsilon} b^{\star 3},$$
(3.35)

which means that if we fit  $(b^*, b^{\delta}, a^{\epsilon}, a^{\delta})$  to the implied surface from options data, we have calibrated all needed parameters to data. As the data is sparse on the *T* axis compared to *K*, we first fit the skew maturity by maturity, and then across the term structure.

Thus, the calibration procedure follows two stages:

1. For each maturity *i*, find estimates of  $\hat{a}_i$  and  $\hat{b}_i$  and thus the best linear fit to strikes (j) for that maturity, by solving the least-squares problem

$$\min_{a_i, b_i} \sum_j (I(T_i, K_{ij}) - (a_i (LMMR)_{ij} + b_i))^2,$$
(3.36)

where

$$(\text{LMMR})_{ij} = \frac{\log(K_{ij}/x)}{\tau_i}$$

2. For all strikes, use previously found estimates to find the best linear fit to the term structure, by obtaining intercept  $\hat{a}^{\epsilon}$  and slope  $\hat{a}^{\delta}$  from

$$\min_{a^{\epsilon}, a^{\delta}} \sum_{i} \{ \hat{a}_i - (a^{\epsilon} + a^{\delta} \tau_i) \}^2, \qquad (3.37)$$

and intercept  $\hat{b}^{\star}$  and slope  $\hat{b}^{\delta}$  from

$$\min_{b^{\star}, b^{\delta}} \sum_{i} \{ \hat{b}_{i} - (b^{\star} + b^{\delta} \tau_{i}) \}^{2}.$$
(3.38)

# **3.4** Data Preparation

To enable the parameter stability investigation of this thesis, we perform the model calibration described in 3.3 on options data. To that extent, we employ options index data (S&P 500 options), motivated partly by the fact that index options offer diversification away from specific equity stock events, reflecting the volatility of a broad market under different economic conditions. This choice allows us to compare our results with those obtained in [2] and extend them under time-dependency for the first order theory.

We obtain both index close and European options data on the S&P 500 index from the WRDS (Wharton Research Data Services) database. For each option entry entry we extract:

- date
- expiration date
- call/put flag (0 for puts, 1 for calls)
- call/put flag (0 for puts, 1 for calls)
- strike price
- daily closing bid-ask quotes
- WRDS-computed implied volatility
- index close price (stored separately and used in data preprocessing, see subsection 3.4.2).

#### 3.4.1 Matlab routines

Our programming language of choice is Matlab, as its libraries are best suited for importing, manipulating, visualizing or saving large amounts of data in matrix form in a quick and convenient way. Also, we make explicit use of the statistics, optimization and econometrics toolboxes, which provide useful tools for our implementation. The only weaknesses of Matlab lie in its execution speed (not a compiled language), and as we discovered in our work, its inability to read or write large datasets in files without resorting to expensive caching. For the 2000-2012 data, we have almost 3 million option entries, with between 300 and 1500 entries in any day. To import the > 200MB .csv original data file, we initially considered the 'importData','load', or 'dImread' functions, or even splitting data by year and then recombining it when necessary. Avoiding complications and fitting format and size constraints, we chose 'textscan', as it can handle gracefully both numeric and non-numeric data in a regular expression input (as in our case), and has an explicit way of setting up buffered reading. Experimentation showed that a buffer of  $\approx 300000$  entries read simultaneously works optimally. We made use of the direct array/matrix indexing feature of Matlab, avoiding any costly functions such as 'find'. Another major data bottleneck in Matlab lies in appending values to an existing large .mat file and should be avoided as much as possible. With the latter choices, we can import the 12-year options data file and save it in a .mat file in  $\approx 25.7$ seconds on a 2.8 Ghz Core 2 Duo Windows machine with 4GB RAM memory. Remaining bottlenecks are the Matlab built-in routine 'cell2mat' function that transforms the read cell structure to a matrix structure, and saving the large .mat file on disk. A desirable thing would be the parallelization of the process under which the 'textscan' function runs.

# 3.4.2 Data preprocessing

In the initial step of importing, we filter out option entries with

- bid quotes <\$0.5
- no implied volatility value

and obtain a distribution of the number of option entries for each day as shown in figure 3.1. We can see the exponential increase in option entries in the 'boom' period 2003-2008, which can be speculatively attributed to high levels of hedger/speculator demand for more index option products. The increased numbers of daily active option entries in recent years translates into more option/implied volatility data points across both different strikes and different maturities (see CBOE website for a full product listing and history). Thus, the accelerated 'maturization' of the index options market led to a more fine-grained description of the volatility surface, offering more data to calibrate volatility models to. As a consequence, we have a relatively dense data grid, especially in most recent years, that brings out clear distinctions between the properties of fit offered by different calibration techniques of SV models (as we shall see further).

Next, as the WRDS data differentiates between the implied volatility curves coming from puts and calls, we adopt an adapted blending technique, following the procedure originally described in [Figlewski (2010)] and used in [2] as well. We are interested in each daily set of option data, as we calibrate our model to the implied surface at a daily frequency. Therefore, for each day, for each maturity, we blend call and put implied volatilities, following the procedure:



FIGURE 3.1

#### For each day

- 1. Skip if only data for one maturity is available
- 2. For each maturity
  - 2.1 Skip incorrect maturities (if not Saturday following 3rd Friday of expiration month, see [CBOE])
  - 2.2 Find  $\mathcal{K} =$  set of strikes for which both call and put implied volatilities are available
  - 2.3 Choose cutoff levels

$$L = \max(0.85x, \min(\mathscr{K}))$$
$$H = \min(1.15x, \max(\mathscr{K}))$$

, where x = index close price for the day

- 2.4 If H or L are empty, or they are equal, skip maturity
- 2.5 Discard puts with strike > H and calls with K < L
- 2.6 Select  $\mathscr{U} = \text{set}$  of puts with K < L and calls with K > H that will remain unblended
- 2.7 Select  $\mathscr{B} =$  set of puts and calls with  $K \in (L, H)$
- 2.8 Discard entries in  $\mathscr{B}$  where for a strike only one option type (put or call) is present
- 2.9 For each strike  $K \in \mathscr{B}$  where we have both a put and a call
  - i. Let  $I_p(K)$  and  $I_c(K)$  denote the put and call implied volatilities. We set the implied volatility value I(K) for a strike as

$$I(K) = wI_p(K) + (1-w)I_c(K)$$

where w = w(K) = (H - K)/(H - L).

- 2.10 Discard maturities where  $size(\mathscr{B} \cap \mathscr{U}) = 0$  or 1 (not enough data for the implied volatility surface at one maturity)
- 3. Discard data for a day that has < 2 maturities (not enough data for a day)
- 4. Save pre-processed data for that day

End /For each day



FIGURE 3.2: S&P 500 implied volatilities fitted as a function of LMMR on April 19, 2005. The circles represent option data for the date, and the line  $I \approx b^* + a^\epsilon$  (*LMMR*) shows the fit from only a fast factor fit. [2]



FIGURE 3.3: S&P 500 implied volatilities fitted as a function of  $LM = \tau *LMMRon$  April 19, 2005. The circles represent option data for the date, and the line  $I \approx \overline{\sigma} + b^{\delta}\tau + a^{\delta}$  (*LM*) shows the fit from only a slow factor fit. [2]

# 3.5 Results

Having pre-processed the data (previous section), we will calibrate the first-order option price expansion to market data using the calibration procedure desribed in section 3.3. As we can see in figures 3.2 and 3.3, only fast factor or slow factor fits (corresponding to a single factor model) fail to capture the range of maturities. We can observe these maturities as data strands, with maturities increasing as we go counterclockwise from the leftmost strand.



FIGURE 3.4: Term-structure fits on April 19, 2005. The circles are the slope coefficients  $\hat{a}_i$  of LMMR fitted in the first step of the regression. The solid line is the straight line  $(a^{\epsilon} + a^{\delta}\tau)$  fitted in the second step of the regression. The estimates are  $(a^{\epsilon}, a^{\delta}) = (-0.0646, -0.1397)$ 

#### 3.5.1 Goodness of Fit

Calibrating to the full two-factor model in a two-stage fitting procedure as described earlier, we compute the estimates  $a^{\epsilon}$ ,  $b^{\star}$ ,  $a^{\delta}$ , and  $b^{\delta}$ . In figures 3.4, 3.5 and 3.6 we show the calibration results on the same test date of April 19, 2005 used in [2]. As we can see qualitatively or quantitatively (average relative fit error of 3.75%) the fit is quite good overall, and the entire range of maturities is captured well. To not restrict ourselves to the special case of one test day and have grounds of comparison in terms of goodness of fit futher on, we average the relative fit error for all the data between 2000-2011 considered in our experiments with the measure (expressed in %):

$$GOF_{2000-2011} = \frac{1}{\# \text{of days}} \sum_{\text{daily data}} avg_{\text{data points}}(I - \hat{I})$$
(3.39)

where I the the implied volatility for a data point, and  $\tilde{I}$  is our model estimate. For the case of our first-order calibration we obtain

$$GOF_{2000-2011}^{1st ord} = 5.52\%$$
(3.40)

Figures 3.7, 3.8 and 3.9 for July 26, 2006 though, show that the largest misfitting occurs at the shortest maturities. This leaves room for improvement of the linear fits, and we will examine the cause and correct this behavior in the next section.

#### 3.5.2 Parameter Time Stability

Time-stability of the fitted parameters, as detailed in 2.3.1, is a crucial feature of the model, ensuring its consistency when pricing and hedging other exotic financial options. In order to analyse parameter stability, we look at the time evolution with respect to t of the estimated parameters  $(a^{\epsilon}, b^{\star}, a^{\delta}, b^{\delta})$  (figure 3.10) and the market group parameters  $(\sigma^{\star}, V_0^{\delta}, V_1^{\delta}, V_3^{\delta})$ 



FIGURE 3.5: Term-structure fits on April 19, 2005. The plot shows the  $\hat{a}_i$  counterparts,  $\hat{b}_i$ slope coefficients for the skew intercepts  $\hat{b}_i$  fitted to the straight line  $b^* + b^{\delta}\tau$ . The estimates are  $(b^{\delta}, b^*) = (0.0164, 0.1417)$ 



FIGURE 3.6: Calibration fit on S&P 500 data (circles) on April 19, 2005, based on fast and slow scale estimates  $(a^{\epsilon}, b^{\star}, a^{\delta}, b^{\delta})$ , as calculated for figures 3.4 and 3.5. Average relative fitting error = 3.75%



FIGURE 3.7: Term-structure fits on July 26, 2006. The circles are the slope coefficients  $\hat{a}_i$  of LMMR fitted in the first step of the regression. The solid line is the straight line  $(a^{\epsilon} + a^{\delta}\tau)$  fitted in the second step of the regression.



FIGURE 3.8: Term-structure fits on July 26, 2006. The plot shows the  $\hat{a}_i$  counterparts,  $\hat{b}_i$  slope coefficients for the skew intercepts  $\hat{b}_i$  fitted to the straight line  $b^* + b^{\delta} \tau$ .



FIGURE 3.9: Calibration fit on S&P 500 data (circles) on July 26, 2006, based on fast and slow scale estimates  $(a^{\epsilon}, b^{\star}, a^{\delta}, b^{\delta})$ . Average relative fitting error = 6.1%

(3.11). For the plots in the latter figures, in the blending procedure in 3.4.2, we discard maturities where we don't have both blended and unblended data to eliminate outlier data points. [2] chooses to discard instead the shortest maturity. Qualitatively, the assumptions of the price expansion hold true over the entire period 2000-2012, as  $b^*$  is the leading-order magnitude of volatility compared to small  $a^{\epsilon}, b^{\delta}$  and  $a^{\delta}$  parameters. Also,  $b^*$  and  $b^{\delta}$  are most correlated with large market events such as the financial crisis that started in September 2008, 'indicating a more dramatic jump in implied volatility levels rather than skew slope'[2]. The largest parameter variations are found in periods of economic crises, such as the 'tech bubble' of 2000-2001 and the credit crisis of 2008. The more recent European sovereign-debt crisis did not impact the post-2008 recovery of the S%P 500, as can also be seen through the low variation of fitted parameters throughout 2010-2012. All four estimated parameters are

relatively stable, given their variances over the period 2000-2012:

| |

$$Var_{2000-2012}(a^{\epsilon}) = 0.000328$$
  

$$Var_{2000-2012}(b^{\star}) = 0.0069$$
  

$$Var_{2000-2012}(a^{\delta}) = 0.000696$$
  

$$Var_{2000-2012}(b^{\delta}) = 0.000743$$
  
(3.41)

The means of the estimated parameters are given by:

$$Mean_{2000-2012}(a^{\epsilon}) = -0.0699$$

$$Mean_{2000-2012}(b^{\star}) = 0.2118$$

$$Mean_{2000-2012}(a^{\delta}) = -0.1302$$

$$Mean_{2000-2012}(b^{\delta}) = 0.0011.$$
(3.42)

The above also holds true for the market group parameters, as they are linearly dependent on the estimated parameters through equations 3.35. Their variances are given by:

$$Var_{2000-2012}(\sigma^{\star}) = 0.0072$$

$$Var_{2000-2012}(V_0^{\delta}) = 0.000606$$

$$Var_{2000-2012}(V_1^{\delta}) = 0.000041$$

$$Var_{2000-2012}(V_3^{\epsilon}) = 0.0000034$$
(3.43)

and their means are

$$\sigma^{\star} = 0.2135, V_0^{\delta} = 0.0045, V_1^{\delta} = -0.0067, V_3^{\epsilon} = -0.0010.$$

Overall, on all test days, and over the long time horizon we inspected parameter stability on, our results match those obtained in [2].










FIGURE 3.12: Fitted skew slopes for 2007, using first three maturities only. The ticks on the time axis correspond to option maturity (expiration) dates throughout 2007.  $Var(a^{\epsilon}) = 0.000159$ 

# **3.6 Adjusting For Maturity Cycles**

As we see in figure 3.9 for the test date July 26, 2006, the fit to the shortest maturities suffers a drawback. To bring out the periodic effect that causes a bad short-term fit, we calibrate using only a fast-factor fit  $I \approx b^* + a^{\epsilon}$ LMMR on the first three shortest maturity options (with  $\tau > 3$  days). Performing the calibration during the year 2007 and plotting the slope  $a^{\epsilon}$ as in figure 3.12 we observe a periodicity around the expiration dates of the S&P 500 options (third Friday of each month). The observed periodic behaviour exhibits down 'jumps' in the skew slope  $a^{\epsilon}$  when a 'maturity cycle' ends, i.e. the closest-to-maturity options disappear from the data. The event of an option expiring triggers a chained reaction in the remaining options, with shorter maturities impacted most. We can think of this phenomenon as options that are waiting in a FIFO (First-In-First-Out) queue, arranged by their time-to-maturity. When an option expires, we pop it from the queue, and there is a one-place shift for the remainder of the queue. As an option is closer to expiry itself, the maturity cycle effect (the popping of the front option) on it will be larger. Thus the impact of this one-step shift towards the front of the queue (and expiration) on options is large for short-dated options and negligible for long-dated ones. Looking back at the implied surface of figure 2.5, the effect on short options can be explained by a larger 'responsibility' they have to explain the short-term large skew through a drop to larger negative skew values  $a^{\epsilon}$  on expiry dates.

### 3.6.1 Time-dependent Fast Scale

In order to correct and model the periodic behaviour, we will outline and follow the approach taken in [2]. As we can see in figure 3.12, where the fast-scale fit was studied in isolation, the periodic variation appears in the two-scale model due to the fast scale contribution to volatility. Therefore, the natural and consistent choice is to introduce time variation in the fast scale coefficient  $\epsilon$ . As a consequence, the mean reversion speed increases locally in the period leading up to an expiry date, and then is reset to a low-level until we near the next expiration date. Interestingly, this type of behavior is reminescent of a spiking neuron time series, and could, in theory, be modeled as such. An Izhikevich neuron model could be used to represent a time-dependent function, presenting good computational tractability. However, the flexibility (given by a set of parameters) it brings in neural networks poses difficulty in aligning existing empirical observation with the model, making calibration intrinsically difficult. Instead, we will focus on a simple parametric calendar time model, as introduced in [2].

By introducing a time-dependent function v(t) (which we define in the next subsection), the two-scale model 3.1 becomes:

$$\begin{cases} dX_t = rX_t dt + f(\mathbf{Y}_t, Z_t) X_t dW_t^{(0)\star}, \\ dY_t = \left(\frac{1}{\epsilon v(t)} \alpha(\mathbf{Y}_t) - \frac{1}{\sqrt{\epsilon v(t)}} \beta(\mathbf{Y}_t) \Lambda_1(\mathbf{Y}_t)\right) dt + \frac{1}{\sqrt{\epsilon v(t)}} \beta(\mathbf{Y}_t) dW_t^{(1)\star}, \\ dZ_t = \left(\delta c(Z_t) - \sqrt{\delta}g(Z_t) \Lambda_2(\mathbf{Y}_t, Z_t)\right) dt + \sqrt{\delta}g(Z_t) dW_t^{(2)\star}, \end{cases}$$
(3.44)

with the Brownian correlation structure (3.2) and no arbitrage pricing function (3.3) unchanged. The price PDE (3.5) however, changes to

$$\begin{cases} \mathscr{L}^{\epsilon,\delta}(t)P^{\epsilon} = 0, \\ P^{\epsilon}(T,x,y) = h(x), \end{cases}$$
(3.45)

where the first-order price expansion operator  $\mathscr{L}^{\epsilon,\delta}(t)$  from (3.6) is now time-dependent

$$\mathscr{L}^{\epsilon\delta}(t) = \frac{1}{\epsilon v(t)}\mathscr{L}_0 + \frac{1}{\sqrt{\epsilon v(t)}}\mathscr{L}_1 + \mathscr{L}_2 + \sqrt{\delta}\mathscr{M}_1 + \delta\mathscr{M}_2 + \sqrt{\frac{\delta}{\epsilon v(t)}}\mathscr{M}_3$$
(3.46)

and the operators in the fast-scale terms are substituted by

$$\begin{aligned} \mathscr{L}_{0} &\mapsto \frac{1}{v(t)} \mathscr{L}_{0}, \\ \mathscr{L}_{1} &\mapsto \frac{1}{\sqrt{v(t)}} \mathscr{L}_{1}, \\ \mathscr{M}_{3} &\mapsto \frac{1}{\sqrt{v(t)}} \mathscr{M}_{3}. \end{aligned}$$

Following the derivation from section 3.2 and expanding the price approximation 3.24, we obtain:

$$\tilde{P}^{\epsilon,\delta} = P_{BS} + (T-t) \left( V_0^{\delta}(z) \frac{\partial}{\partial \sigma} + V_1^{\delta}(z) D_1(\frac{\partial}{\partial \sigma}) + \overline{v_{t,T}^{1/2}} (V_2^{\epsilon}(z) D_2 + V_3^{\epsilon}(z) D_1 D_2) \right] P_{BS} \quad (3.47)$$

where we have introduced the time-averaged square of the time-dependent function v(t):

$$\overline{v_{t,T}^{1/2}} = \frac{1}{T-t} \int^T \sqrt{v(s)} ds.$$
(3.48)

From the above we can see that only the fast-scale market group parameters

$$V_2^{\epsilon}(z) \mapsto \overline{v_{t,T}^{1/2}} V_2^{\epsilon}(z),$$
$$V_3^{\epsilon}(z) \mapsto \overline{v_{t,T}^{1/2}} V_3^{\epsilon}(z),$$

are affected by the introduction of time-dependency, with  $V_0$  and  $V_1$  unchanged. Applying the parameter-reduction method from section 3.2.4 we can rewrite (3.47) as

$$\tilde{P}^{\star} = P_{BS}(\sigma_{t,T}^{\star}) + (T-t) \left( V_0^{\delta}(z) \frac{\partial}{\partial \sigma} + V_1^{\delta}(z) D_1(\frac{\partial}{\partial \sigma}) + \overline{v_{t,T}^{1/2}} V_3^{\epsilon}(z) D_1 D_2 \right) P_{BS}(\sigma_{t,T}^{\star}), \quad (3.49)$$

where the time-dependent corrected volatility  $\sigma_{t,T}^{\star}$  (analogue of 3.26) is given by

$$\sigma_{t,T}^{\star} = \sqrt{\overline{\sigma}^2 + 2\overline{v_{t,T}^{1/2}}V_2^{\epsilon}}.$$
(3.50)

*Note*:The accuracy results obtained in the derivation of the first-order perturbation hold true for the time-dependent case as well.

### 3.6.2 Time-dependent Function

As mentioned in the previous subsection, we will use a calendar function v(t) to represent the periodic maturity cycles. Any simple parametric choice is suitable, as long as it represents the upward monthly slope increases in figure 3.12. Another desirable of the calendar function is to allow enough flexibility for a slope increase that is linear w.r.t next maturity date, or is more accelerated as we near the maturity date. As in [2], we choose a normalized polynomial function:

$$v(t) = c_p (n(T-t))^p$$
, with  $p > 0$ , and (3.51)

$$n(T-t) = \inf\{T-t : T-t \ge \Delta t\},$$
(3.52)

where  $\Delta t > 0$  is a small cutoff to prevent a blowup of  $v^{-1}$  to infinite values at maturity dates. An alternative model would be a parameterized exponential

$$v(t) = c_{a,b} \exp(a * n(T - t) - b)$$
, with  $a, b > 0$ ,

where b aligns the function to start from positive time-to-expiry values, and a would play the role of p in the polynomial model, accelerating slope increases. However, we choose not to pursue this model due to its complete dependence of the cutoff value on b, which serves a different purpose. Nevertheless, both models present similar qualitative shape characteristics and suffer from the same limitations, as the slopes in 3.12 appear to be increasing randomly over very short time-frames (order of 1-2 days). Therefore we choose the more flexible and easier to use polynomial function.

Even though expiration dates are not evenly distributed, we must assume that they are approximately  $\Delta T$  apart (in our case,  $\Delta T = 1/12$  - one month) so we can enforce the integrability condition:

$$\frac{\int_{0}^{\Delta T} (v(s))^{1/2} ds}{\Delta T} = 1 \tag{3.53}$$

The latter is necessary to obtain normalized values for v(t), as it implies a normalization constant given by

$$c_p^{1/2} = \frac{1}{(\Delta T)^{p/2}} (1 + \frac{p}{2}).$$
(3.54)

The parameter p influences v(t) and consequently its inverse  $v^{-1}$ . As we can see from 3.44 the calendar function  $v^{-1}$  is directly proportional to the fast scale mean reversion speed. Figure 3.13 shows how various values of p impact the calendar function and implicitly the fast scale mean reversion speed, as we go through succesive monthly expiration dates. Notice how a higher value of p defers speeding the fast scale till near expiration, when it spikes on a scale proportional to p.

To use the calendar function, we need to define its time average square. To this end, we follow [2] and decompose the time to maturity as  $T - t = m_0 \Delta T + \eta$  with  $m_0$  an integer and  $0 \le \eta < \Delta T$  the residual of a monthly cycle. For simplicity, we choose the cutoff  $\Delta t = 0$ . Then, using 3.51, we can rewrite 3.48 as:

$$\overline{v_{t,T}^{1/2}} = \frac{1}{T-t} \int_{0}^{T} (v(s))^{1/2} ds 
= \frac{1}{T-t} \int_{0}^{m_0 \Delta T + \eta} (v(T-s))^{1/2} ds 
= \frac{\sqrt{c_p}}{m_0 \Delta T + \eta} \left( \int_{t}^{t+\eta} (t+\eta-s)^{p/2} ds + m_0 \int_{0}^{\Delta T} (\Delta T - s)^{p/2} ds \right) 
= \frac{1}{(\Delta T)^{p/2}} \frac{(\eta^{1+\frac{p}{2}} + m_0 (\Delta T)^{1+\frac{p}{2}})}{m_0 \Delta T + \eta}.$$
(3.55)

From the above we obtain the approximations:

$$\overline{v_{t,T}^{1/2}} = \begin{cases} \sqrt{\frac{T-t}{\Delta T}} & \text{for } T-t < \Delta T \ (m_0 = 0), \\ 1 & \text{for } T-t \gg \Delta T \ (\text{time-independent}), \end{cases}$$
(3.56)



FIGURE 3.13:  $v^{-1}$  for p = 1, 5, 50 with a cutoff  $\Delta t \approx 3$  days

# 3.6.3 Time-dependent Calibration

Introducing the variables

$$LMMR_v = LMMRv_{t,T}^{1/2},$$
$$\Delta b = b^* - \overline{\sigma}.$$

and given 3.50 and that  $V_2^{\epsilon}(z)$ ,  $V_3^{\epsilon}(z)$  are substituted by their product with  $\overline{v_{t,T}^{1/2}}$  wherever they appear, [2] shows that the time-dependent analogue of (3.32) is:

$$I \approx \overline{\sigma} + \Delta b \overline{v_{t,T}^{1/2}} + a^{\epsilon} \text{LMMR}_v + b^{\delta} \tau + a^{\delta} \text{LM}.$$
(3.57)

To fit the implied volatility surface according to 3.57, we need to obtain estimates of  $(a^{\epsilon}, \Delta b, \overline{\sigma}, b^{\delta}, a^{\delta})$ . As in the time-independent case, we calibrate the model to implied volatility data from a given day, determined by expiration dates  $T_i$  and strike prices  $K_{ij}$ . We again employ the same two-stage procedure described in 3.3, where we first fit on the mode dense dimension K for a fixed T and then fit the sparse term structure in the T dimension.

Thus, the calibration procedure follows two stages:

1. For each maturity *i*, find estimates of  $\hat{a}_i$  and  $\hat{b}_i$  and thus the best linear fit to strikes (j) for that maturity, by solving the least-squares problem

$$\min_{a_i, b_i} \sum_j (I(T_i, K_{ij}) - (a_i(LM)_{ij} + b_i))^2,$$
(3.58)

where

$$(LM)_{ij} = \log(K_{ij}/x)$$

2. For all strikes, use previously found estimates to find the best linear fit to the term structure, by obtaining intercept  $\hat{a}^{\epsilon}$  and slope  $\hat{a}^{\delta}$  from

$$\min_{a^{\delta}, a^{\epsilon}} \sum_{i} \{ \hat{a}_{i} - (a^{\epsilon} \frac{\overline{v_{t, T_{i}}^{1/2}}}{\tau_{i}} + a^{\delta}) \}^{2},$$
(3.59)

and parameters  $\hat{\Delta b}, \hat{\overline{\sigma}}$  and  $\hat{b^{\delta}}$  by performing a multiple linear regression (to not be confused with multiple simple linear regression) on the factors  $\overline{v_{t,T_i}^{1/2}}$  and  $\tau_i$  as

$$\min_{\sigma,\Delta b,b^{\delta}} \sum_{i} \{ \hat{b}_{i} - (\sigma + \Delta b \overline{v_{t,T_{i}}^{1/2}} + b^{\delta} \tau_{i}) \}^{2}.$$
(3.60)

Using  $\Delta b = b^* - \overline{\sigma}$ , from the five estimated parameters we obtain the four group parameters  $(a^{\epsilon}, b^*, b^{\delta}, a^{\delta})$  that we need. As we obtain estimates for each day, we can compare the goodness of fit for a given day and the stability of parameters over time with the time-independent calibration.



FIGURE 3.14: Fitted skew slopes with time-dependent fast scale for 2007, using first three maturities only. The ticks on the time axis correspond to option maturity (expiration) dates throughout 2007.

 $Variance(a^{\epsilon}) = 0.000108$  (-32% reduction compared to time-independent case)

# 3.7 Time-dependent Results

We aim to compare results with the time-independent ones and conclude whether the introduction of time-dependency has improved the two-scale model first-order price expansion. We do this by analysing the maturity cycles correction, goodness of fit on test days, and parameter time-stability. The test data is pre-processed in an identical manner, as described in 3.4.2. As a note, the results in this section are obtained using a constant parameter p = 1in the calendar function v(t) expression.

### 3.7.1 Maturity cycles correction

We keep only the first three shortest maturity options (with  $\tau > 3$  days). We use only a fast factor fit, similar to the one generating 3.12, but now we introduce time-dependency:

$$I \approx b_{t,T}^{\star} + a_{t,T}^{\epsilon} \text{LMMR} \overline{v_{t,T}^{1/2}}$$
(3.61)

By plotting the slope  $a^{\epsilon}$  in 3.14, we can observe the periodic strands have dissapeared, leaving a lower variation.

### 3.7.2 Goodness of Fit

First we examine the goodness of fit properties on the test day of July 26, 2006, which proved to be problematic for the time-independent model as we saw in figure 3.9, and we compare results with [2]. Using the calibration technique described in the previous section, the time-dependent fit can be seen in figure 3.16. Notice the improved fit at the shortest

maturity (leftmost strand), with an overall average relative fit error of 3.76% < 6.1% for the initial time-indepedent model. The calibration was performed assuming p = 1 in the calendar function v(t) expression. The results are consistent with the time-dependent theory outlined at the beginning of this section, as we can see a clustering of points in the bottom plot of 3.15 around 1 on the  $v^{1/2}$  axis (time-indepedent case), with only two points at lower values (two shortest maturities). Further evidence of the fit improvement is brought by averaged relative fit error for 2000-2011, given by

 $\text{GOF}_{2000-2011}^{1\text{st ord TD}} = 3.80\%$ , a 31% reduction of the time-independent model fit error (3.62)

Some numerical values are, however, contradicting those in [2] for the same test day July 26, 2006 and using the same test data obtained from the WRDS database (see section 3.4). In the reference book, smaller values up to a factor of 10 are obtained on the x axes of both of the plots in figure 3.15. The values on the y axes and fitted lines in those plots, as well as figure 3.16, representing fits on the whole term structure, are identical with those in [2]. The different parts of our results, though, seem to be in consistence with equations 3.55 and 3.56 for the calendar function, unlike those in the reference book. We therefore conclude the validity of our experiment which improves the first-order theory, leaving room for either a data error in [2] or a small misinterpretation in our work. Regardless, the empirical validity of the goodness of fit improvement by modeling time-dependency holds true.

### 3.7.3 Parameter Time Stability

We analyse the time-stability of the fitted parameters in the time-dependent case to compare it with the time-independent results in subsection 3.5.2. Therefore we look again at the time evolution with respect to t of the estimated parameters  $(a^{\epsilon}, b^{\star}, a^{\delta}, b^{\delta})$  (figure 3.10) this time without making the conversion to the market group parameters  $(\sigma^{\star}, V_0^{\delta}, V_1^{\delta}, V_3^{\epsilon})$ . For a direct comparison with the time-independent model, market group parameters are not needed as they are obtained linearly from the fitted parameters, whose time stability we put in comparison. Plotting  $(a^{\epsilon}, b^{\star}, a^{\delta}, b^{\delta})$  over the period 2000-2012 in figure 3.17, we obtain similar characteristics with the time-independent model, as expected. The real differences appear when we look at the new parameter(td := time dependent) variances and means over the same period (percent change and time-independent value in paramtheses):

$$Var_{2000-2012}(a_{td}^{\epsilon}) = 0.000254/ - 23\%(0.000328)$$

$$Var_{2000-2012}(b_{td}^{\star}) = 0.0050/ - 28\%(0.0069)$$

$$Var_{2000-2012}(a_{td}^{\delta}) = 0.0013/ + 87\%(0.000696)$$

$$Var_{2000-2012}(b_{td}^{\delta}) = 0.000298/ - 60\%(0.000743)$$
(3.63)



FIGURE 3.15: Time dependent term-structure fit on S& P 500 options data from July 26, 2006. The top plot shows the fit of  $\hat{a}$  against  $\frac{\overline{v_{tT}^{1/2}}}{\tau}$ . The bottom plot shows the fit of  $\hat{b}$  by plotting  $\hat{b} - b^{\delta}\tau$  against  $\overline{v_{tT}^{1/2}}$ 



FIGURE 3.16: Time-dependent first-order fits on S& P 500 options data from July 26, 2006. Improved average relative fitting error = 3.76%, due to better short maturity fits.

$$\begin{aligned} \operatorname{Mean}_{2000-2012}(a_{td}^{\epsilon}) &= -0.0506(-0.06990) \\ \operatorname{Mean}_{2000-2012}(b_{td}^{\star}) &= 0.2084(0.2118) \\ \operatorname{Mean}_{2000-2012}(a_{td}^{\delta}) &= -0.1795(-0.1302) \\ \operatorname{Mean}_{2000-2012}(b_{td}^{\delta}) &= 0.0039(0.0011). \end{aligned}$$

$$(3.64)$$

From the figures above we can observe a decrease in variance in  $(a^{\epsilon}, b^{\star}, b^{\delta})$  and a high variance increase in  $a^{\delta}$ . However, in absolute terms, the variance reduction in the leading magnitude term  $b^{\star}$  offsets any other increase, and thus the time-dependent calibration empirically proves itself superior also in terms of time stability. The only issue that remains to be investigated is the validity of the new means, which present noticeable differences. The means could be used by an industry practitioner to price liquid exotic options for which a closed form price expansion solution has been found (for example barrier options in [? ]). The resulting prices can then be bechmarked against the exchange/inter-dealer network quotes, and thus the time-dependent parameter means could be validated or not. However such practical issues and an in-depth study of the parameters application to exotic options are beyond the scope of this thesis.







FIGURE 3.18: First-order fits on S& P 500 options data from March 18, 2010. Average relative fitting error = 6.53%

# 3.8 Optimizing p parameter in v(t)

We consider a new test day of March 18, 2010, where we study fit results after the application of time dependency with a calendar function parameterized by p = 1. As we can see from figures 3.18 and 3.19, time dependency with p=1 does not improve the fits of shorter maturities. Given our discussion of the effects of varying p in subsection 3.6.2, we allow pto vary between 1 and a limit value of 50. We choose the max limit to be 50 as the slope of the calendar function  $\approx 1$  when p=50 (see figure 3.13) and a higher p would not cause a major change of its slope or shape. Returning to our test day we choose a value of p so we minimize average relative error of fit for that day:

$$p^{\star} = \arg\min_{p} avg_{\text{data points}}(I - \hat{I})$$
(3.65)

The above yields an optimum p = 33 for the test day, and we can see the considerably improved time-dependent fit in figure 3.20 (a 38% reduction of fit error).



FIGURE 3.19: Time-dependent (p = 1) first-order fits on S& P 500 options data from March 18, 2010. Average relative fitting error = 5.40%



FIGURE 3.20: Time-dependent  $(p^* = 33)$  first-order fits on S& P 500 options data from March 18, 2010. Average relative fitting error = 3.35%

It is necessary to mention that the time-dependency model is built with an implicit assumption that if we abstract away the effect of maturity cycles in vanilla markets, we can apply better, more stable parameters to over-the-counter (OTC) markets where the effect is not present. [2] argues that the cycle effect may not be present in different markets such at the OTC options one, where there is an almost continuous range of maturity cycles.



Given the fit improvement when choosing an optimal  $p \neq 1$  for March 18, 2010, we take the last idea to an 'extreme', proposing that p vary freely for all data. Thus we are basically transfering variance from the market parameters we need in pricing exotic options, to the parameter p which we discard as an unwanted pure effect of maturity cycles. The main empirical test of our proposition validity is to see if a p optimized on a daily basis reduces parameter variance and/or improves fits. Calibrating on all available data from 2000-2011, we obtain a better average fit:

$$\text{GOF}_{2000-2011}^{\text{1st ord TD, optimal p}} = 3.26\%$$
, a 14% reduction of the time-dependent model fit error (3.66)

The parameter variances and means (tdp := time dependent with p optimized) over the same period (percent change and p = 1 time-dependent value in parantheses) are:

$$Var_{2000-2012}(a_{tdp}^{\epsilon}) = 0.000427/ + 68\%(0.000254)$$

$$Var_{2000-2012}(b_{tdp}^{\star}) = 0.0049/ - 2\%(0.0050)$$

$$Var_{2000-2012}(a_{tdp}^{\delta}) = 0.0012/ - 8\%(0.0013)$$

$$Var_{2000-2012}(b_{tdp}^{\delta}) = 0.000284/ - 5\%(0.000298)$$
(3.67)

$$\begin{aligned} \operatorname{Mean}_{2000-2012}(a_{tdp}^{\epsilon}) &= -0.0713(-0.0506) \\ \operatorname{Mean}_{2000-2012}(b_{tdp}^{\star}) &= 0.2083(0.2084) \\ \operatorname{Mean}_{2000-2012}(a_{tdp}^{\delta}) &= -0.1449(-0.1795) \\ \operatorname{Mean}_{2000-2012}(b_{tdp}^{\delta}) &= 0.0040(0.0039). \end{aligned}$$

$$(3.68)$$

Though the above values indicate a high increase of  $a^{\epsilon}$  variance, the decrease in one order larger variances of  $b^*$  and  $a^{\delta}$ ) more than offsets it. Thus, empirical results show that a calendar function with an optimized p reduces parameter variance and increases goodness of fit, proving superior to one with a static p. A natural question arises about how the distribution of optimal p values influences fit error. Not surprisingly, we find that keeping only tests days with  $p^* = 50$  produces a GOF = 1.8%, well below the average for all data. Most fit error resides in test dates with  $p^* < 5$ , as we obtain a below average GOF = 2.16% by filtering those dates out. By analysing the data results we discover that low values of p are obtained most often before expiry, with high values right after expiration dates. Since low values of p are responsible for error fit, we reinstate an expiry cutoff (which was previously assumed as 0). Thus we remove from the time series any dates that are closer than  $\Delta t = 3$ days to a maturity date. We discover that this reduces average error of fit for 2000-2011 data considerably:

- GOF= 2.57% for a static p = 1 throughout
- GOF= 2.52% for a daily optimized p

These results indicate that despite the introduction of time depedency, the first order expansion still has problems in fitting options with a time-to-maturity lower than 3 days. A calendar function with an optimized p reduces error fit considerably when short-expiry data is included in the calibration, but is ineffective otherwise. Outside very short expirations, the first-order time-dependent expansion performs well even compared to a second order expansion (as we will see in the next chapter). Furthermore, it has the advantage of a quick and purely deterministic calibration with only 4 market parameters, that we have shown to be relatively stable.

# Chapter 4

# Second order perturbation: Outline and Implementation Results

In this chapter we extend our treatment to the second order perturbation expansion theory. We outline the main results from its derivation in [?], leaving the reader to inspect the given paper for the array of detailed formal proofs underlying the theoretical grounds of the expansion. Of particular importance in that context is the terminal layer analysis performed in the recent paper, which provides a difficult extension of the accuracy results mentioned in (3.29) for the first-order theory. After introducing the implications of the second order expansion, we focus on associated calibration techniques and their computational tractability and results.

# 4.1 Derivation Outline

We start from the same two-scale risk-neutral model (3.1) used in the first-order theory in Chapter 3:

$$\begin{cases} dX_t = rX_t dt + f(\mathbf{Y}_t, Z_t) X_t dW_t^{(0)\star}, \\ dY_t = \left(\frac{1}{\epsilon} \alpha(\mathbf{Y}_t) - \frac{1}{\sqrt{\epsilon}} \beta(\mathbf{Y}_t) \Lambda_1(\mathbf{Y}_t, Z_t)\right) dt + \frac{1}{\sqrt{\epsilon}} \beta(\mathbf{Y}_t) dW_t^{(1)\star}, \\ dZ_t = \left(\delta c(Z_t) - \sqrt{\delta} g(Z_t) \Lambda_2(\mathbf{Y}_t, Z_t)\right) dt + \sqrt{\delta} g(Z_t) dW_t^{(2)\star}, \end{cases}$$
(4.1)

with the correlation structure between the Brownian motions  $(W_t^{(0)\star}, W_t^{(1)\star}, W_t^{(2)\star})$  given by

$$\begin{cases} d\langle W^{(0)\star}, W^{(1)\star} \rangle_t = p_1 dt, \\ d\langle W^{(0)\star}, W^{(2)\star} \rangle_t = p_2 dt, \\ d\langle W^{(1)\star}, W^{(2)\star} \rangle_t = p_{12} dt. \end{cases}$$

$$(4.2)$$

The infinitisemal operator  $\mathscr{L}^{\epsilon,\delta}$  satisfying 3.5 is again expanded as

$$\mathscr{L}^{\epsilon,\delta} = \frac{1}{\epsilon}\mathscr{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathscr{L}_1 + \mathscr{L}_2 + \sqrt{\delta}\mathscr{M}_1 + \delta\mathscr{M}_2 + \sqrt{\frac{\delta}{\epsilon}}\mathscr{M}_3, \tag{4.3}$$

where the  $\mathscr{L}$  and  $\mathscr{M}$  operators follow the same definitions as in equations 3.7, with a change of notation for the market price of risk functions:  $(\Lambda_1, \Lambda_2) \mapsto (\Lambda, \Gamma)$  We avoid the lengthy and involved construction of the expansion from [?], and expose the important results in a simpler top-bottom form. We start from the equation driving all the differences and go to the first-order theory, the price expansion

$$P^{\epsilon,\delta} \approx \overline{P}^{\epsilon,\delta} := P_{0,0} + \sqrt{\epsilon}P_{1,0} + \sqrt{\delta}P_{0,1} + \epsilon P_{2,0} + \delta P_{2,0} + \sqrt{\epsilon\delta}P_{1,1}, \tag{4.4}$$

which is the analogue to the first order equation 3.8. The difference is that we retained three more terms at the end that are of the second order in terms of the  $\sqrt{\epsilon}$  and  $\sqrt{\delta}$  square volatility time scales. The conditions governing the expansion terms are:

$$\begin{array}{ll}
O(1): \langle \mathscr{L}_{2} \rangle P_{0,0} = 0, & P_{0,0}(T, x, z) = h(x), \\
O(\sqrt{\epsilon}): \langle \mathscr{L}_{2} \rangle P_{1,0} = -\mathscr{V}P_{0,0}, & P_{1,0}(T, x, z) = 0, \\
O(\sqrt{\delta}): \langle \mathscr{L}_{2} \rangle P_{0,1} = -\langle \mathscr{M}_{1} \rangle P_{0,0}, & P_{0,1}(T, x, z) = 0, \\
O(\epsilon): P_{2,0} = -\frac{1}{2}\phi D_{2}P_{0,0} + F_{2,0}, & (4.5) \\
\langle L_{2} \rangle F_{2,0} = -\mathscr{A}P_{0,0} - \mathscr{V}P_{1,0}, & F_{2,0}(T, x, z) = 0, \\
O(\delta): \langle L_{2} \rangle P_{0,2} = -\langle \mathscr{M}_{1} \rangle P_{0,1} - \mathscr{M}_{2}P_{0,0}, & P_{0,2}(T, x, z) = 0, \\
O(\sqrt{\epsilon\delta}): \langle L_{2} \rangle P_{1,1} = -\mathscr{V}P_{0,1} - \frac{1}{\sigma'} \mathscr{C} \partial_{z} P_{0,0} - \langle \mathscr{M}_{1} \rangle P_{1,0}, & P_{1,1}(T, x, z) = 0,
\end{array}$$

where the z-dependent operators are given by

$$\begin{cases}
O(1): \langle L_2 \rangle = \partial_t + \frac{1}{2} \overline{\sigma}^2 D_2 + r D_1 - r, \\
O(\sqrt{\epsilon}): \mathscr{V} = V_3 D_1 D_2 + V_2 D_2, \\
O(\sqrt{\delta}): \langle \mathscr{M}_1 \rangle = \frac{2}{\overline{\sigma}} (V_1 D_1 + V_0) \partial_z, \\
O(\epsilon): \mathscr{A} = A_2 D_1^2 D_2 + A_1 D_1 D_2 + A_0 D_2 + A D_2^2, \\
O(\delta): \mathscr{M}_2 = \frac{1}{2} g^2 \partial_{zz}^2 + c \partial_z, \\
O(\sqrt{\epsilon\delta}): \mathscr{C} = C_2 D_1^2 + C_1 D_1 + C_0 + C D_2.
\end{cases}$$
(4.6)

In 4.6 we have introduced multiple group parameters. The A parameters are given by

$$\begin{cases}
A_2(z) = \frac{1}{2} \rho_{xy}^2 \langle \beta(\cdot) f(\cdot, z) \partial_y \psi_1(\cdot, z) \rangle, \\
A_1(z) = -\frac{1}{2} \rho_{xy} (\langle \beta(\cdot) \Lambda(\cdot, z) \partial_y \psi_1(\cdot, z) \rangle + \langle \beta(\cdot) f(\cdot, z) \partial_y \psi_2(\cdot, z) \rangle), \\
A_0(z) = \frac{1}{2} \langle \beta(\cdot) \Lambda(\cdot, z) \partial_y \psi_2(\cdot, z) \rangle, \\
A(z) = -\frac{1}{4} (\langle \phi(\cdot, z) f^2(\cdot, z) \rangle - \langle \phi(\cdot, z) \rangle \langle f^2(\cdot, z) \rangle),
\end{cases}$$
(4.7)

and the C parameters by

$$\begin{cases} C_2(z) = -\rho_{xy}\rho_{xz}\overline{\sigma}'(z)g(z)\langle\beta(\cdot)f(\cdot, z)\partial_y\psi_3(\cdot, z)\rangle, \\ C_1(z) = \rho_{xz}\overline{\sigma}'(z)g(z)\langle\beta(\cdot)\Lambda\partial_y\psi_3\rangle + \rho_{xy}g\langle\beta(\cdot)f(\cdot, z)\partial_y\psi_4(\cdot, z)\rangle, \\ C_0(z) = -\overline{\sigma}'(z)g(z)\langle\beta(\cdot)\Lambda(\cdot, z)\partial_y\psi_4(\cdot, z)\rangle, \\ C(z) = -\frac{1}{2}\rho_{yz}\overline{\sigma}'(z)g(z)\langle\beta(\cdot)\partial_y\phi(\cdot, z)\rangle. \end{cases}$$

$$(4.8)$$

 $(\psi_i, i = \overline{1, 4} \text{ are functions introduced by centering conditions for Poisson functions - we skip the full derivation involving these) We further introduce the$ *B*group parameters

$$\begin{cases} B_2 = \frac{1}{2}g^2(\overline{\sigma}')^2, \\ B_1 = \frac{1}{2}g^2\overline{\sigma}'' + c\overline{\sigma}', \end{cases}$$

$$\tag{4.9}$$

and the notations

$$V'_{j} := \partial_{z} V_{j}, j = \overline{0, 3}, \qquad V_{i}^{\epsilon} := \sqrt{\epsilon} V_{i},$$

$$V_{i}^{\delta} := \sqrt{\delta} V_{i}, \qquad A_{i}^{\epsilon} := \sqrt{\epsilon} A_{i},$$

$$B_{i}^{\delta} := \delta B_{i}, \qquad C_{i}^{\epsilon, \delta} := \sqrt{\epsilon \delta} C_{i}, \qquad (4.10)$$

$$V_{i}^{\prime \epsilon}(z) := \partial_{z} V_{i}^{\epsilon}(z), \qquad V_{i}^{\prime \delta}(z) := \partial_{z} V_{i}^{\delta}(z),$$

$$\phi^{\epsilon}(y, z) := \epsilon \phi(y, z).$$

As in subsection 3.2.4 for the first-order case, we can apply a similar parameter reduction, by absorbing the fast scale  $V_2^{\epsilon}$  parameter in the volatility term

$$\sigma^*(z) := \sqrt{\overline{\sigma}(z)^2 + 2V_2^{\epsilon}(z)},\tag{4.11}$$

and as a consequence remove all  $V_2$  dependent terms. Translating the expansion in price into an expansion in implied volatility we obtain the form

$$I^{\bar{\epsilon},\delta} := I_{0,0} + \sqrt{\epsilon}I_{1,0} + \sqrt{\delta}I_{0,1} + \sqrt{\epsilon\delta}I_{1,1} + \epsilon I_{2,0} + \delta I_{0,2}.$$
(4.12)

# 4.2 Calibration Equations

Matching 4.12 with our model, applying the parameter reduction step, and defining the forward log-moneyness as

$$d := \log(K/xe^{r\tau}), \tag{4.13}$$

[?] proves implied volatility can be rewritten in expanded form as:

$$I^{\bar{\epsilon},\delta} = (\frac{1}{\tau}k + l + \tau m + \tau^2 n) + \frac{d}{\tau}(p + \tau q + \tau^2 s) + \frac{d^2}{\tau^2}(u + \tau v + \tau^2 w),$$
(4.14)

where

$$\begin{cases} O(1/\tau) : k = \frac{3(V_{5}^{\epsilon})^{2}}{2(\sigma^{\epsilon})^{5}} - \frac{A_{2}^{\epsilon}}{(\sigma^{\epsilon})^{3}} - \frac{A^{\epsilon}}{(\sigma^{\epsilon})^{3}} - \frac{\phi^{\epsilon}}{2\sigma^{\epsilon}} \\ O(1) : l = \frac{3V_{1}^{\delta}V_{3}^{\epsilon}}{(\sigma^{\epsilon})^{4}} - \frac{C_{2}^{\epsilon,\delta}}{2(\sigma^{\epsilon})^{2}} - \frac{C^{\epsilon,\delta}}{2(\sigma^{\epsilon})^{2}} + \frac{A_{0}^{\epsilon}}{\sigma^{\epsilon}} + \frac{A_{1}^{\epsilon}}{2\sigma^{\epsilon}} + \frac{A_{2}^{\epsilon}}{4\sigma^{\epsilon}} - \frac{A^{\epsilon}}{4\sigma^{\epsilon}} - \frac{V_{1}^{\delta}V_{3}^{\epsilon}}{(\sigma^{\epsilon})^{3}\sigma^{\epsilon'}} + \sigma^{\epsilon} + \frac{V_{3}^{\epsilon}}{2\sigma^{\epsilon}} \\ O(\tau) : m = \frac{B_{1}^{\delta}}{2} + \frac{C_{0}^{\epsilon,\delta}}{2} + \frac{C_{1}^{\epsilon,\delta}}{4} + \frac{C_{2}^{\epsilon,\delta}}{8} - \frac{C^{\epsilon,\delta}}{8} + \frac{5(V_{1}^{\delta})^{2}}{6(\sigma^{\epsilon})^{3}} - \frac{V_{0}^{\delta}V_{3}^{\epsilon}}{2(\sigma^{\epsilon})^{2}} + \frac{B_{2}^{\delta}}{6\sigma^{\epsilon}} \\ - \frac{2V_{1}^{\delta}V_{1}^{\prime\delta}}{3(\sigma^{\epsilon})^{2}\sigma^{\epsilon'}} + \frac{V_{0}^{\delta}V_{3}^{\prime\epsilon}}{2\sigma^{\epsilon'}} + \frac{V_{1}^{\delta}V_{3}^{\prime\epsilon}}{4\sigma^{\epsilon'}\sigma^{\epsilon'}} + V_{0}^{\delta} + \frac{V_{1}^{\delta}}{2} \\ O(\tau^{2}) : n = \frac{(V_{0}^{\delta})^{2}}{(6\sigma^{\epsilon})} + \frac{V_{0}^{\delta}V_{1}^{\delta}}{6\sigma^{\epsilon}} + \frac{(V_{1}^{\delta})^{2}}{6\sigma^{\epsilon}} - \frac{B_{2}^{\delta}\sigma^{\epsilon}}{12} + \frac{2V_{0}^{\delta}V_{0}^{\prime\delta}}{3\sigma^{\epsilon'}} + \frac{V_{0}^{\delta}V_{1}^{\delta}}{3\sigma^{\epsilon'}} + \frac{V_{1}^{\delta}V_{1}^{\prime\delta}}{6\sigma^{\epsilon'}} \\ O(\frac{d}{r}) : p = -\frac{3(V_{0}^{\delta})^{2}}{2(\sigma^{\epsilon})^{5}} + \frac{A_{1}^{\epsilon}}{(\sigma^{\epsilon})^{3}} + \frac{C_{1}^{\epsilon,\delta}}{2(\sigma^{\epsilon})^{2}} + \frac{C_{2}^{\epsilon,\delta}}{2(\sigma^{\epsilon})^{2}} + \frac{V_{0}^{\delta}V_{1}^{\prime\delta}}{(\sigma^{\epsilon})^{3}\sigma^{\epsilon'}} + \frac{V_{1}^{\delta}V_{1}^{\prime\epsilon}}{(\sigma^{\epsilon})^{3}\sigma^{\epsilon'}} + \frac{V_{1}^{\delta}}{(\sigma^{\epsilon})^{3}\sigma^{\epsilon'}} \\ O(d) : q = -\frac{3V_{0}^{\delta}V_{3}^{\delta}}{3(\sigma^{\epsilon})^{3}} - \frac{5(V_{1}^{\delta})^{2}}{6(\sigma^{\epsilon})^{3}} + \frac{2V_{0}^{\delta}V_{1}^{\delta}}{2(\sigma^{\epsilon})^{2}} + \frac{2V_{0}^{\delta}V_{1}^{\delta}}{3(\sigma^{\epsilon})^{2}\sigma^{\epsilon'}} + \frac{2V_{1}^{\delta}V_{1}^{\epsilon}}{(\sigma^{\epsilon})^{3}} \\ O(d\tau) : s = -\frac{5V_{0}^{\delta}V_{1}^{\delta}}{3(\sigma^{\epsilon})^{3}} + \frac{2V_{0}^{\delta}V_{1}^{\delta}}{(\sigma^{\epsilon})^{5}} + \frac{A_{1}^{\epsilon}}{(\sigma^{\epsilon})^{5}} \\ O(\frac{d}{\tau^{2}}) : u = -\frac{3(V_{3}^{\epsilon})^{2}}{(\sigma^{\epsilon})^{6}} + \frac{C_{2}^{\epsilon,\delta}}{2(\sigma^{\epsilon})^{4}} + \frac{C_{1}^{\epsilon,\delta}}{2(\sigma^{\epsilon})^{4}}} \\ O(d^{2}) : w = -\frac{6V_{1}^{\delta}V_{1}^{\delta}}{(\sigma^{\epsilon})^{5}} + \frac{B_{2}^{\delta}}{2(\sigma^{\epsilon})^{3}} + \frac{2V_{1}^{\delta}V_{1}^{\epsilon}}{3(\sigma^{\epsilon})^{5}\sigma^{\epsilon'}} \\ O(d^{2}) : w = -\frac{7(V_{1}^{\delta})^{2}}{3(\sigma^{\epsilon})^{5}} + \frac{B_{2}^{\delta}}{3(\sigma^{\epsilon})^{3}} + \frac{2V_{1}^{\delta}V_{1}^{\epsilon}}{3(\sigma^{\epsilon})^{5}\sigma^{\epsilon'}} \\ O(d^{2}) : w = -\frac{7(V_{1}^{\delta})^{2}}{3(\sigma^{\epsilon})^{5}} + \frac{B_{2}^{\delta}}{2(\sigma^{\epsilon})^{3}} + \frac{2V_{1}^{\delta}V_{1}^$$

We have two groups of parameters, which we define as:

 $\Theta := \{k, l, m, n, p, q, s, u, v, w\},$ (10 estimated parameters)  $\Phi := \{\sigma^*, V_3^{\epsilon}, V_1^{\delta}, V_0^{\delta}, C_2^{\epsilon,\delta}, C_1^{\epsilon,\delta}, C_0^{\epsilon,\delta}, A_2^{\epsilon}, A_1^{\epsilon}, A_0^{\epsilon}, A^{\epsilon}, B_2^{\delta}, B_1^{\delta}, \frac{V_3'^{\epsilon}}{\overline{\sigma}}, \frac{V_1'^{\delta}}{\overline{\sigma}}, \frac{V_0'^{\delta}}{\overline{\sigma}}, \phi^{\epsilon}\}$ (18 market group parameters) (4.16)

We denote  $I(\tau, d)$  to be the implied volatility observed in the market of a European call option with time-to-maturity  $\tau$  and forward log-moneyness d.  $\hat{I}^{\epsilon,\delta}(\tau, d; \Theta)$  represents the expanded implied volatility of a European call as defined in equation (4.12).

Following [?], the calibration procedure naturally lends itself to two steps:

1. Estimate  $\Theta^*$  (the 10 estimated parameters) around the 10 basis functions

$$\{\frac{1}{\tau}, 1, \tau, \tau^2, \frac{d}{\tau}, d, d\tau, \frac{d^2}{\tau^2}, \frac{d^2}{\tau}, d^2\}$$

so that we minimize the least squares fit error over all maturities i and strikes j in the data:

$$\min_{\Theta} \sum_{i} \sum_{j} (I(\tau_{i}, d_{j}) - I^{\hat{\epsilon}, \delta}(\tau_{i}, d_{j}; \Theta))^{2} = \sum_{i} \sum_{j} (I(\tau_{i}, d_{j}) - I^{\hat{\epsilon}, \delta}(\tau_{i}, d_{j}; \Theta^{*}))^{2}, \quad (4.17)$$

2. Estimate  $\Phi^*$  (the 18 group market parameters) as the minimal least squares set of parameters (L2)

$$\min_{\star} \|\Phi\|^2 = \|\Phi^*\|^2 \tag{4.18}$$

so that the 10 nonlinear constraints in 4.15 hold.

# 4.3 Calibration - First parameter group (10)

The system of polynomial equations 4.15 linking the two sets of parameters make the translation between the two parameter sets highly non-trivial. Therefore we will focus individually on each of the two steps of the calibration procedure and analyse the time stability of each parameter set.

To compute the first set of 10 estimated parameters  $\Theta := \{k, l, m, n, p, q, s, u, v, w\},$ , as emphasized in [?], we must fit the observed volatility surface across all maturities at once, not maturity-by-maturity as in the first order theory. Indeed, the two step maturityby maturity procedure first fitted to a polynomial of d and then three polynomials of  $\tau$  is completely unusable, as we see in figure 4.1, where the fit is extremely biased towards longer maturities.

Therefore we find the parameters that give the least squares fit globally, across all maturities. For this purpose we use a local solver (Matlab 'fmincon') to find the optimal solution, given that the fit is linear in all parameters in (4.14). For goodness-of-fit we use the same test day of July 26, 2006 where the first-order theory struggled to capture short maturities, but improved with the introduction of time-dependency. From figure 4.2 we can observe how the second order expansion captures quadratically the large skew at short maturities reducing the average relative fit error on the test day by 37%. Also, it provides an unbiased fit across maturities as the first step of the calibration was done across all maturities, so there is no need to model a maturity cycles phenomenon.

Time-stability of the fitted parameters can be seen in figures 4.3, 4.4 and 4.5 where they are grouped according to their dominating factor in the implied volatility expansion 4.14. Qualitatively we can observe a high monthly variation in the group (p, q, s), which may be linked to these parameters absorbing the monthly variation related to maturity cycles. This proves there is further research scope in trying to stabilize (p, q, s) by explicitly modelling



FIGURE 4.1

time-dependency. However, time-dependency should impact only these parameters, a task hard to achieve, given our experience with the impact on the slow scale parameters that timedependency had for the first-order theory. Given the almost certain risk that we destabilize the other parameters of leading magnitude and the difficulty of adjusting the second order expansion to time-dependency in a computationally feasible way, we leave this addition for further research.

Quantitatively, over the period 2000-2011, the 10 estimated parameters display statistics in line with first-order parameters, as we can see in table 4.1. Variances are extremely low, except for l, m, q. All parameters have low values and consequently long-term means, except for l, as expected given that it contains the leading volatility magnitude term  $\sigma^*$  at the denominator level.



FIGURE 4.2: Average relative fit error = 2.36%

Parameter	Mean	Variance
k	-1.77-004	3.74-006
1	0.2084	0.0063
m	0.0043	0.0021
n	1.08-005	2.08-004
р	0.0036	7.18-006
q	-0.0884	0.0011
$\mathbf{S}$	-0.1229	7.36-004
u	-1.37-004	1.56-007
v	1.37-004	3.02-005
W	0.0102	8.91-004

TABLE 4.1: Properties of estimated parameters (first set)













### 4.4 Second parameter group (18)

We expect the calibration of the market parameter group  $\Phi$  to be highly non-trivial given the non-linear system of equations 4.15 linking them to the estimated group of parameters. The inversion step to obtain them requires solving a complex global optimization problem. This is evidenced by obtaining multiple local minima and high sensitivity to initial conditions when using a simple random global search function such as the Matlab built-in function 'multiStart'. Therefore, the search space of solutions to equations 4.15 is non-convex and requires either a powerful global solver or approximate techniques. These considerations guide four approaches that we present in detail here. The solution we seek in not necessarily a global optimum, but needs to be obtained computationally fast enough to allow calibration over a long period of time where we seek acceptable time stability. Parameters in  $\Phi$  should also align with our expansion asymptotics expectations ([?]): leading order positive  $\sigma^* < 1$ and small parameters for the remainder of the set bounded by [-1, 1]. These will form our box contraints when searching for a solution.

### 4.4.1 Global Optimization in Current Form

We attempt to find the minimal L2 parameter group  $\Theta$  by using global search in the solution space of 4.15, so our problem is defined as:

$$\min_{\Phi} \|\Phi\|^2 = \|\Phi^*\|^2 \tag{4.19}$$

with the nonlinear constraints given by equations 4.15 and box-constraints  $\sigma^* \in [0, 1], \Phi - \{\sigma^*\} \in [-1, 1],$ 

We initially experimented with the matlab 'globalSearch' function that employs a scatter search for choosing sample points to run a local solver on. Despite a short calibration time (1-2 seconds for a day's data), when tested on 10 days data, the results showed extreme variation from one day to another of a factor of  $10^5$ , thus dismissing their validity.

Given the limitations of Matlab optimization libraries, we obtianed a 25 days trial license of a commercial grade optimization toolbox - TOMLAB (see [website]). The package is written in Matlab and offers a suite of local and global optimization solvers, so its adoption is relatively straightforward. TOMLAB offers a radial basis functions (cubic splines) function for solving global optimization problems called 'rbfSolve' which has the following noteworthy features:

• Constructs a response surface to data collected by evaluating the objective function at a few points

TOMLAB - Tomlab user Demo license 999100. Va	lid to 2013-06-20
Problem: 1: Original Form User given sum( cc f(x_k) + sum( cc	f_k         0.106646458573259580           f(x_*)         0.0040000000000000           onstr          0.00000093585818166           onstr          0.106646552159077740
Solver: rbfSolve. EXIT=0. INFORM=0. Cubic Radial Basis Function Interpolation. Glo Tried 200 f(x), using 200, startup 190 ExD 12	obal solver glcCluster. Local solver snopt : Constrained Latin Hypercube, use 9000 get 190.
FuncEv 200 ConstrEv 9201 Iter 10 CPU time: 545.332696 sec. Elapsed time: 514.92 Optimal vector x: x_k: 0.242888 -0.001565 0.001230 0.072	11000 sec.
-0.003246 0.021/16 0.01346/ -0.013 0.001063 -0.075207 -0.005478 -0.096 Error in 10 nonlinear constraints. cErr -4.840905e-010 -1.418269e-009 -1.491512e -3.004365e-009 -6.481033e-009 -5.271940e	-38 -0.03026/ -0.015342 274 -0.157898 -0.001257 -010 -3.094665e-010 -2.019971e-010 -1.778402e-008 -008 -1.103403e-008
* * *	* * * *

FIGURE 4.6: Output from running 'rbfSolve' global optimization algorithm

- Can be started with either a combination of corner points (proved ineffective), a latin hypercube (an algorithm to split search space in unbiased regions of variation), or with a 'warm start' (resuming computation from an intermediate stage of the optimization). We use the latin hypercube option, as it is effective compared to the few unprecise corner points given by the box contraints. However, we must bear in mind that the number of regions generated from a latin hypercube will grow at a power given by the number of search dimensions (in our case 18 parameters/dimensions).
- Uses multiple strategies to determine target values and search the feasible surface, all of which operate in cycles, alternating local and global searches.
- (Website) mentions the solvers used for global or local search on the response surface are possible to change, recommending the 'glcCluster' and 'glcFast' solvers for global search. Also of great importance: 'glcCluster' will probably be more reliable, however it might spend too much time in the search on the surface at each iteration.'
- The algorithm is optimized to handle costly objective functions and computationally cheap constraints. It does not treat constraints explicitly, but rather uses them as penalties for the objective function.
- Can be parallelized naturally due to employing cycles of computation

Considering the algorithm's features, we run it on S&P 500 data from the test day October 19, 2006, fixing a limit to the maximum function evaluations, to avoid a very lenghty full computation. We obtain the results in figure 4.6, which look remarkably similar to those in [?].  $\sigma^*$  is the leading order term with a realistic value in terms of both the first-order results and the VIX value of 15.5% for the test date ([yahoo finance]). At the same time, remaining parameters are small in comparison, confirming asymptotics expectations. While we have no guarantee for global optimality, we obtained a feasible result with only a partial optimization. However, as we see in figure 4.6, the computation time is prohibitive to conduct a study of the time series stability of the parameters. The two alternative TOMLAB global agorithms 'evo' and 'arbfMIP' have been tested without success, displaying similar results. Therefore, we cannot conclude whether we can reliably use this approach to calibration, unless we experiment on multiple parallel machines with more computing power. We need to either complete the optimization and (possibly) find the global minimum or run the partial one above on the entire 2000-2011 data and inspect stability. Regardless, the main drawback is the lengthy computation time.

Other possible improvements include using a more suitable algorithm start or using a different global solver ('glcCluster' is computationally expensive). Even more importantly, modelling constraints explicitly might speed up the computation as the penalty method that treats them together with the objective functions may start false searches. There may be cases where the objective function is believed to be reduced but in fact the additive constraint divergence is. This improvement is announced in a future release of TOMLAB (see website).

Other state-of-the-art global search algorithms can be used online on the solver server [neos]. The BARON or LINDOGlobal algorithms would be suitable choices, however the problem needs to be rewritten in the GAMS modelling language. This detailed research unfortunately falls outside of the thesis scope, given time constraints.

### 4.4.2 Integrating constraints in objective function

We take this approach to avoid the intermediate set of 10 estimated parameters and due to global solver's properties. As we already mentioned, the TOMLAB global algorithms are optimized for expensive objective functions and cheap constraints. Therefore we choose to incorporate the system of constraints 4.15 directly in the volatility expansion 4.14, bypassing the first calibration stage and fitting directly to the 18 market parameters. As a consequence, we removed all non-linear constraints and now have a computationally expensive function.

#### 4.4.3 Hybrid Calibration with First-Order Results

To fit the 18 market parameters, we either need a more well-adapted solver, or we can try to simplify the parameter problem. To support the latter, we first quote a remark from [? ]: 'Note [...] that the  $V_i^{\epsilon}$  are order  $\sqrt{\epsilon}$ , the  $V_i^{\delta}$  order  $\sqrt{\delta}$  and that they appeared in the first order asymptotic theory in [?]. The new parameters [...] come from the order  $\epsilon$ , order  $\delta$ and order  $\sqrt{\epsilon\delta}$  terms in the the second order expansion respectively.'. We therefore attempt to match the first 4 parameters ( $\sigma^*$ ,  $V_0^{\delta}$ ,  $V_1^{\delta}$ ,  $V_3^{\epsilon}$ ) in the first-order theory with their exact analogues in the second order theory. The theoretical validity of the bridging between the two expansion theories is given by:

- identical operators  $\mathscr{L}$  and  $\mathscr{M}$  given by equations 3.7 for the two expansions
- identical expressions (3.22), (3.23), (3.17), (3.16) for parameters  $V_0^{\delta}, V_1^{\delta}, V_2^{\epsilon}, V_3^{\epsilon}$  in both [2] and [?].
- an identical parameter reduction step 3.26 yielding  $\sigma^*$ , that is dependent on the same parameter  $V_2^{\epsilon}$  and identically defined notion of effective volatility  $\overline{\sigma}$ .

We propose the following calibration procedure for the 18 market parameters:

- 1. Calibrate  $(\sigma^{\star}, V_0^{\delta}, V_1^{\delta}, V_3^{\epsilon})$  using the first-order time-dependent model (time-dependency is included because second order fits don't have any difficulties with short maturities)
- 2. Find optimal  $\Phi' = \Phi \{\sigma^{\star}, V_0^{\delta}, V_1^{\delta}, V_3^{\epsilon}\}$  from

$$\min_{\Phi'} \|\Phi'\|^2 = \|\Phi'^*\|^2 \tag{4.20}$$

so that the 10 nonlinear constraints in 4.15 hold with 4 known parameters replaced by their values found in step 1. Running the above calibration procedure on daily data in the period 2000-2011, we obtain the time-stability results in figures 4.7 4.8, 4.9 4.10















### 4.5 Further 3 Parameter Reductions

Looking at the time stability of the 18 market parameters plotted in the figures from 4.4.3, we can qualitatively observe some very revealing properties that are not immediately apparent in the system of equations 4.15. There seems to be an extremely high correlation in shape between the parameters  $A_2^{\epsilon}$ ,  $A_1^{\epsilon}$  and  $A^{\epsilon}$ . The same correlation can be seen for the evolutions for  $A_2^{\epsilon}$ ,  $A_1^{\epsilon}$  look almost inversely related. The same applies for the evolutions of  $B_1^{\delta}$  and  $C_0^{\epsilon,\delta}$ . The same observation as for parameters A can be noticed for parameters  $(C_2^{\epsilon,\delta}, C_1^{\epsilon,\delta}, C^{\epsilon,\delta})$ . Therefore, we are motivated to revisit equations 4.15 and search for a way to group combinations of these parameters together in some form that allows us to reduce the number of parameters and greatly simplify the 18 market parameters problem. We find that this step is indeed possible and we proceed with it in detail here.

First we look for all equations where the parameters in the pairs  $(A_2^{\epsilon}, A^{\epsilon})$  and  $(A_2^{\epsilon}, A_1^{\epsilon})$  appear. These are k, l, p, u:

$$\begin{cases} k = \frac{3(V_3^{\epsilon})^2}{2(\sigma^*)^5} - \frac{A_2^{\epsilon}}{(\sigma^*)^3} - \frac{A^{\epsilon}}{(\sigma^*)^3} - \frac{\phi^{\epsilon}}{2\sigma^*} \\ l = \frac{3V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} - \frac{C_2^{\epsilon,\delta}}{2(\sigma^*)^2} - \frac{C^{\epsilon,\delta}}{2(\sigma^*)^2} + \frac{A_0^{\epsilon}}{\sigma^*} + \frac{A_1^{\epsilon}}{2\sigma^*} + \frac{A_2^{\epsilon}}{4\sigma^*} - \frac{A^{\epsilon}}{4\sigma^*} - \frac{V_1^{\delta}V_3'^{\epsilon}}{(\sigma^*)^3\sigma^{*'}} + \sigma^* + \frac{V_3^{\epsilon}}{2\sigma^*} \\ p = -\frac{3(V_3^{\epsilon})^2}{2(\sigma^*)^5} + \frac{A_1^{\epsilon}}{(\sigma^*)^3} + \frac{A_2^{\epsilon}}{(\sigma^*)^3} + \frac{V_3^{\epsilon}}{(\sigma^*)^3} \\ u = -\frac{3(V_3^{\epsilon})^2}{(\sigma^*)^7} + \frac{A_2^{\epsilon}}{(\sigma^*)^5} + \frac{A^{\epsilon}}{(\sigma^*)^5} \end{cases}$$

$$(4.21)$$

Simple algebra allows us to regroup the 3 A market group parameters  $(A_2^{\epsilon}, A_1^{\epsilon}, A^{\epsilon})$  in the form:

$$\begin{cases} k = \frac{3(V_3^{\epsilon})^2}{2(\sigma^*)^5} - \frac{A_2^{\epsilon} + A^{\epsilon}}{(\sigma^*)^3} - \frac{\phi^{\epsilon}}{2\sigma^*} \\ l = \frac{3V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} - \frac{C_2^{\epsilon,\delta}}{2(\sigma^*)^2} - \frac{C^{\epsilon,\delta}}{2(\sigma^*)^2} + \frac{A_0^{\epsilon}}{\sigma^*} + \frac{A_1^{\epsilon} + A_2^{\epsilon}}{2\sigma^*} - \frac{A_2^{\epsilon} + A^{\epsilon}}{4\sigma^*} - \frac{V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^3\sigma^{*'}} + \sigma^* + \frac{V_3^{\epsilon}}{2\sigma^*} \\ p = -\frac{3(V_3^{\epsilon})^2}{2(\sigma^*)^5} + \frac{A_1^{\epsilon} + A_2^{\epsilon}}{(\sigma^*)^3} + \frac{V_3^{\epsilon}}{(\sigma^*)^3} \\ u = -\frac{3(V_3^{\epsilon})^2}{(\sigma^*)^7} + \frac{A_2^{\epsilon} + A^{\epsilon}}{(\sigma^*)^5} \end{cases}$$

$$(4.22)$$

We can thus make the following substitutions in (4.22):

$$X_1 = A_2^{\epsilon} + A^{\epsilon}$$

$$X_2 = A_2^{\epsilon} + A_1^{\epsilon}$$
(4.23)

Now considering the 3 C market parameters  $(C_2^{\epsilon,\delta}, C_1^{\epsilon,\delta}, C^{\epsilon,\delta})$ , they appear in the equations l, m, q, v:

$$\begin{cases} l = \frac{3V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} - \frac{C_2^{\epsilon,\delta}}{2(\sigma^*)^2} - \frac{C^{\epsilon,\delta}}{2(\sigma^*)^2} + \frac{A_0^{\epsilon}}{\sigma^*} + \frac{A_1^{\epsilon} + A_2^{\epsilon}}{2\sigma^*} - \frac{A_2^{\epsilon} + A^{\epsilon}}{4\sigma^*} - \frac{V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^3\sigma^{*'}} + \sigma^* + \frac{V_3^{\epsilon}}{2\sigma^*} \\ m = \frac{B_1^{\delta}}{2} + \frac{C_0^{\epsilon,\delta}}{2} + \frac{C_1^{\epsilon,\delta}}{4} + \frac{C_2^{\epsilon,\delta}}{8} - \frac{C^{\epsilon,\delta}}{8} + \frac{5(V_1^{\delta})^2}{6(\sigma^*)^3} - \frac{V_0^{\delta}V_3^{\epsilon}}{2(\sigma^*)^2} + \frac{B_2^{\delta}}{6\sigma^*} \\ - \frac{2V_1^{\delta}V_1^{\prime\delta}}{3(\sigma^*)^2\sigma^{*\prime}} + \frac{V_0^{\delta}V_3^{\prime\epsilon}}{2\sigma^*\sigma^{*\prime}} + \frac{V_1^{\delta}V_3^{\prime\epsilon}}{4\sigma^*\sigma^{*\prime}} + V_0^{\delta} + \frac{V_1^{\delta}}{2} \\ q = -\frac{3V_0^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} - \frac{3V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} + \frac{C_1^{\epsilon,\delta}}{2(\sigma^*)^2} + \frac{C_2^{\epsilon,\delta}}{2(\sigma^*)^2} + \frac{V_0^{\delta}V_3^{\prime\epsilon}}{(\sigma^*)^3\sigma^{*\prime}} + \frac{V_1^{\delta}V_3^{\prime\epsilon}}{(\sigma^*)^3\sigma^{*\prime}} + \frac{V_1^{\delta}}{(\sigma^*)^2} \\ v = -\frac{6V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^6} + \frac{C_2^{\epsilon,\delta}}{2(\sigma^*)^4} + \frac{C_1^{\epsilon,\delta}}{2(\sigma^*)^4} + \frac{V_1^{\delta}V_3^{\prime\epsilon}}{(\sigma^*)^5\sigma^{*\prime}} \end{cases}$$

$$(4.24)$$

We can rewrite the above equations as:

$$\begin{cases} l = \frac{3V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} - \frac{C_2^{\epsilon,\delta} + C^{\epsilon,\delta}}{2(\sigma^*)^2} + \frac{A_0^{\epsilon}}{\sigma^*} + \frac{A_1^{\epsilon} + A_2^{\epsilon}}{2\sigma^*} - \frac{A_2^{\epsilon} + A^{\epsilon}}{4\sigma^*} - \frac{V_1^{\delta}V_3'^{\epsilon}}{(\sigma^*)^3\sigma^{*'}} + \sigma^* + \frac{V_3^{\epsilon}}{2\sigma^*} \\ m = \frac{B_1^{\delta}}{2} + \frac{C_0^{\epsilon,\delta}}{2} + \frac{C_1^{\epsilon,\delta} + C_2^{\epsilon,\delta}}{4} - \frac{C_2^{\epsilon,\delta} + C^{\epsilon,\delta}}{8} + \frac{5(V_1^{\delta})^2}{6(\sigma^*)^3} - \frac{V_0^{\delta}V_3^{\epsilon}}{2(\sigma^*)^2} + \frac{B_2^{\delta}}{6\sigma^*} \\ - \frac{2V_1^{\delta}V_1'^{\delta}}{3(\sigma^*)^2\sigma^{*'}} + \frac{V_0^{\delta}V_3'^{\epsilon}}{2\sigma^*\sigma^{*'}} + \frac{V_1^{\delta}V_3'^{\epsilon}}{4\sigma^*\sigma^{*'}} + V_0^{\delta} + \frac{V_1^{\delta}}{2} \end{cases}$$

$$(4.25)$$

$$q = -\frac{3V_0^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} - \frac{3V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} + \frac{C_1^{\epsilon,\delta} + C_2^{\epsilon,\delta}}{2(\sigma^*)^2} + \frac{V_0^{\delta}V_3'^{\epsilon}}{(\sigma^*)^3\sigma^{*'}} + \frac{V_1^{\delta}V_3'^{\epsilon}}{(\sigma^*)^3\sigma^{*'}} + \frac{V_1^{\delta}V_3'^{\epsilon}}{(\sigma^*)^3\sigma^{*'}} + \frac{V_1^{\delta}}{(\sigma^*)^2} \\ v = -\frac{6V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^6} + \frac{C_2^{\epsilon,\delta} + C^{\epsilon,\delta}}{2(\sigma^*)^4} + \frac{V_1^{\delta}V_3'^{\epsilon}}{(\sigma^*)^5\sigma^{*'}} \end{cases}$$

Again we can make the following substitutions in (4.25):

$$X_3 = C_2^{\epsilon,\delta} + C^{\epsilon,\delta}$$
  

$$X_4 = C_2^{\epsilon,\delta} + C_1^{\epsilon,\delta}$$
(4.26)

Lastly, considering the parameters  $B_1^{\delta}$  and  $C_0^{\epsilon,\delta}$ , they both appear only in the equation m:

$$m = \frac{B_1^{\delta}}{2} + \frac{C_0^{\epsilon,\delta}}{2} + \frac{C_1^{\epsilon,\delta}}{4} + \frac{C_2^{\epsilon,\delta}}{8} - \frac{C^{\epsilon,\delta}}{8} + \frac{5(V_1^{\delta})^2}{6(\sigma^*)^3} - \frac{V_0^{\delta}V_3^{\epsilon}}{2(\sigma^*)^2} + \frac{B_2^{\delta}}{6\sigma^*} - \frac{2V_1^{\delta}V_1^{\prime\delta}}{3(\sigma^*)^2\sigma^{*\prime}} + V_0^{\delta} + \frac{V_1^{\delta}}{2}$$

$$(4.27)$$

where we can obviously make the substitution

$$X_5 = B_1^{\delta} + C_0^{\epsilon,\delta} \tag{4.28}$$
Applying the substitutions (4.23),(4.26) and (4.28) all at once on the equations system 4.15 we obtain

$$\begin{split} & O(1/\tau) : k = \frac{3(V_{5}^{*})^{2}}{2(\sigma^{*})^{5}} - \frac{X_{1}}{(\sigma^{*})^{3}} - \frac{\phi^{\epsilon}}{2\sigma^{*}} \\ & O(1) : l = \frac{3V_{1}^{5}V_{3}^{*}}{(\sigma^{*})^{4}} - \frac{X_{3}}{(2(\sigma^{*})^{2})^{2}} + \frac{A_{0}^{\epsilon}}{\sigma^{*}} + \frac{X_{2}}{2\sigma^{*}} - \frac{X_{1}}{4\sigma^{*}} - \frac{V_{1}^{5}V_{3}^{*}}{(\sigma^{*})^{3}\sigma^{*'}} + \sigma^{*} + \frac{V_{3}^{\epsilon}}{2\sigma^{*}} \\ & O(\tau) : m = \frac{X_{5}}{2} + \frac{X_{4}}{4} - \frac{X_{3}}{8} + \frac{5(V_{1}^{5})^{2}}{6(\sigma^{*})^{3}} - \frac{V_{0}^{5}V_{3}^{\epsilon}}{2(\sigma^{*})^{2}} + \frac{B_{2}^{5}}{6\sigma^{*}} \\ & - \frac{2V_{1}^{5}V_{1}^{'}}{3(\sigma^{*})^{2}\sigma^{*'}} + \frac{V_{0}^{5}V_{3}^{'\epsilon}}{2\sigma^{*}\sigma^{*}} + \frac{V_{1}^{5}V_{3}^{'\epsilon}}{4\sigma^{*}\sigma^{*}} + V_{0}^{5} + \frac{V_{1}^{5}}{2} \\ & O(\tau^{2}) : n = \frac{(V_{0}^{5})^{2}}{6\sigma^{*}} + \frac{V_{0}^{5}V_{1}^{*}}{6\sigma^{*}} + \frac{(V_{1}^{5})^{2}}{(\sigma^{*})^{3}} - \frac{B_{2}^{5}\sigma^{*}}{12} + \frac{2V_{0}^{5}V_{0}^{'\delta}}{3\sigma^{*'}} + \frac{V_{0}^{5}V_{1}^{'\delta}}{3\sigma^{*'}} + \frac{V_{1}^{5}V_{1}^{'\delta}}{3\sigma^{*'}} \\ & O(\frac{d}{r}) : p = -\frac{3(V_{3}^{5})^{2}}{2(\sigma^{*})^{5}} + \frac{X_{2}}{(\sigma^{*})^{3}} + \frac{X_{4}}{2(\sigma^{*})^{2}} + \frac{V_{0}^{5}V_{3}^{'\epsilon}}{(\sigma^{*})^{3}\sigma^{*'}} + \frac{V_{1}^{5}V_{3}^{'\epsilon}}{(\sigma^{*})^{3}\sigma^{*'}} + \frac{V_{1}^{5}}{(\sigma^{*})^{3}} \\ & O(d) : q = -\frac{3V_{0}^{5}V_{1}^{\delta}}{(\sigma^{*})^{4}} - \frac{5(V_{1}^{5})^{2}}{6(\sigma^{*})^{3}} + \frac{2V_{0}^{5}V_{1}^{\delta}}{3(\sigma^{*})^{2}\sigma^{*'}}} + \frac{2V_{0}^{5}V_{1}^{'\delta}}{3(\sigma^{*})^{2}\sigma^{*'}} + \frac{2V_{1}^{5}V_{1}^{'\delta}}{3(\sigma^{*})^{2}\sigma^{*'}} \\ & O(d\tau) : s = -\frac{5V_{0}^{5}V_{1}^{\delta}}{(\sigma^{*})^{3}} - \frac{5(V_{1}^{5})^{2}}{6(\sigma^{*})^{3}} + \frac{2V_{0}^{5}V_{1}^{\delta}}{3(\sigma^{*})^{2}\sigma^{*'}}} + \frac{2V_{0}^{5}V_{1}^{'\delta}}{3(\sigma^{*})^{2}\sigma^{*'}} + \frac{2V_{1}^{5}V_{1}^{'\delta}}{3(\sigma^{*})^{2}\sigma^{*'}} \\ & O(\frac{d^{2}}{\tau^{2}}) : u = -\frac{3(V_{3}^{5}V_{2}^{*}}{(\sigma^{*})^{5}} + \frac{X_{3}}{2(\sigma^{*})^{4}} + \frac{V_{1}^{5}V_{3}^{'\epsilon}}{3(\sigma^{*})^{5}\sigma^{*'}} \\ & O(d^{2}) : w = -\frac{7(V_{1}^{5})^{2}}{(\sigma^{*})^{5}} + \frac{B_{2}^{\delta}}{3(\sigma^{*})^{3}} + \frac{2V_{1}^{5}V_{1}^{'\delta}}{3(\sigma^{*})^{4}\sigma^{*'}} \\ \end{array}$$

Therefore, we have translated the 8 parameters  $(A_2^{\epsilon}, A_1^{\epsilon}, A^{\epsilon}, C_2^{\epsilon,\delta}, C_1^{\epsilon,\delta}, C_0^{\epsilon,\delta}, B_1^{\delta})$  into a new group of 5 parameters  $(X_1, X_2, X_3, X_4, X_5)$ , removing 3 parameters in the process. Next, we formalize the effect the parameter reduction will have on the original L2 minimization problem 4.30, and investigate how we can recover the original parameters.

Given our substitutions

$$\Phi := \{\sigma^*, V_3^{\epsilon}, V_1^{\delta}, V_0^{\delta}, C_2^{\epsilon,\delta}, C_1^{\epsilon,\delta}, C_0^{\epsilon,\delta}, C^{\epsilon,\delta}, A_2^{\epsilon}, A_1^{\epsilon}, A_0^{\epsilon}, A^{\epsilon}, B_2^{\delta}, B_1^{\delta}, \frac{V_3^{\prime\epsilon}}{\overline{\sigma}}, \frac{V_1^{\prime\delta}}{\overline{\sigma}}, \frac{V_0^{\prime\delta}}{\overline{\sigma}}, \phi^{\epsilon}\}$$

gets replaced in the second stage of the calibration by

$$\chi := \{ \sigma^*, \ V_3^{\epsilon}, \ V_1^{\delta}, \ V_0^{\delta}, \ A_0^{\epsilon}, \ B_2^{\delta}, \ \frac{V_3^{\prime \epsilon}}{\overline{\sigma}}, \ \frac{V_1^{\prime \delta}}{\overline{\sigma}}, \ \frac{V_0^{\prime \delta}}{\overline{\sigma}}, \ \phi^{\epsilon}, \ X_1, \ X_2, \ X_3, \ X_4, \ X_5 \}$$

Thus, the second stage of the calibration becomes:

2. Estimate  $\chi^*$  (the 15 group market parameters) as the least squares set of parameters (L2)

$$\min_{\chi} \|\chi\|^2 = \|\chi^*\|^2 \tag{4.30}$$

so that the 10 nonlinear constraints in 4.29 hold.

#### 4.5.1 Accuracy of Reduction

By minimizing  $\chi$  instead of  $\Phi$  we make an implicit approximation to the original optimal parameters. However, we show that the error is negligible in the regime of the second order asymptotics. As we replace 8 original parameters with five sums of pairs of two, our optimization change can be summarized for one pair grouping at a time as:

$$\min_{x,y,\Phi-\{x,y\}} x^2 + y^2 + \|\Phi-\{x,y\}\|^2$$
(4.31)

gets replaced by

$$\min_{a,\Phi-\{x,y\}} a^2 + \|\Phi-\{x,y\}\|^2$$
(4.32)

with

$$x + y = a.$$

Here x and y are interpreted of any of the pairs we substitute and a is one of their corresponding  $(X_i)$  parameters that we introduced. To get to the optimization problem (4.30) for  $\chi$  we just perform the above replacement repeteadly over all relevant parameter pairs. For one pair (x, y) that is substituted, the change in the objective function is thus given by:

$$\Delta = a^2 - x^2 - y^2 = 2xy \tag{4.33}$$

The corresponding aggregate (sum of changes due to additive least squares form) change for the 5 parameter pairs we substitute is then given by:

$$\Delta_{agg} = 2 \left( A_2^{\epsilon} A^{\epsilon} + A_2^{\epsilon} A_1^{\epsilon} + C_2^{\epsilon,\delta} C^{\epsilon,\delta} + C_2^{\epsilon,\delta} C_1^{\epsilon,\delta} + B_1^{\delta} C_0^{\epsilon,\delta} \right)$$
(4.34)

We note that the regime of asymptotic expansion that led to the original calibration procedure, as [?] states, assumes negligibly small values for all other market parameters except for the leading order  $\sigma^*$  parameter. The empirical results over 2000-2011 plotted in the stability figures from subsection 4.4.3 further support this claim. We can observe that the A, Band C parameters forming the  $\Delta_{agg}$  error function are centered around values than tend to zero for  $\epsilon, \delta \to 0$ . Thus the change in the objective function  $\Delta_{agg}$  must be neglijable given the consistency conditions of the second-order expansion asymptotics. We choose to refrain from a more formalized mathematical proof to find the exact order of our accuracy in terms of the volatility time scales parameters  $\epsilon$  and  $\delta$ . This can however be completed by the more mathematically inclined following the accuracy proofs for the first-order expansion in [2].

#### 4.5.2 Recovering the Original Parameters

To recover the original 8 parameters after fitting the 15 market parameters  $\in \chi$ , we use a simple property of the original second stage least squares optimization problem. A simple argument by contradiction shows that in order to minimize the original  $||\Phi||^2$  objective with 18 parameters, given that 16 optimal parameters are known, the residual objective  $x^2 + y^2$  from the remaining 2 parameters x and y must be minimal too. If the residual objective would not be minimal and could be improved, then our entire objective function would not be minimal and could be improved, contradicting the L2 optimality condition on  $\Phi$ . As we replace the pair (x, y) with a (as in 4.31), a will be optimized. Thus, in order to recover optimal parameters x and y from an optimal a, we must enforce the original optimality condition on  $x^2 + y^2$ .

In our case, we recover each of the 5 pairs of parameters from  $\Phi^*$  after finding the optimal L2 set  $\chi^*$ . The same argument from above applies, only now in an additive form, repeated 5 times. Thus we recover an original pair (x, y) as:

$$(x^*, y^*) = \min_{x,y} (x^2 + y^2)$$
, so that  $x + y = a^*$  (4.35)

#### 4.5.3 Concluding Remarks

Given our accuracy results and recovery procedure, we can obtain optimal parameters  $(A_2^{\epsilon}, A_1^{\epsilon}, A^{\epsilon}, C_2^{\epsilon,\delta}, C_1^{\epsilon,\delta}, C_0^{\epsilon,\delta}, B_1^{\delta})$  in  $\Phi^*$  from optimal parameters  $(X_1, X_2, X_3, X_4, X_5)$  in the reduced set  $\chi^*$ .

Thus we bring the total number of 18 market parameters to only 15 and reduce the dimension and difficulty of the global optimization problem posed by converting the 10 estimated paramaters in the 18 market parameters we need in order to use the second order expansion. This reduced problem formulation will improve the computational time and accuracy when solving using any global search algorithm, such as those considered in 4.4.3. With sufficient global search time, we expect results similar to those in section 4.4.3 where the hybrid calibration/model had a market parameters problem dimension of 14 by assuming 4 parameters given by the first-order model calibration. However, we will focus our efforts to employ the results in this section and obtain a quasi-closed form solution for the 18 market parameters in terms of the 10 estimated parameters for the hybrid calibration mentioned earlier.

### 4.6 Derivation of Quasi-Closed Form Solution

In this section we use a combination of hybrid calibration and the parameter reduction described in 4.5 to reduce the troublesome non-linear mapping to market parameters into a quasi-closed form solution in terms of the estimated parameters.

We start from the reduced system of equations

$$\begin{split} k &= \frac{3(V_{3}^{\circ})^{2}}{2(\sigma^{\ast})^{5}} - \frac{X_{1}}{(\sigma^{\ast})^{3}} - \frac{\phi^{\epsilon}}{2\sigma^{\ast}} \\ l &= \frac{3V_{1}^{\delta}V_{3}^{\ast}}{(\sigma^{\ast})^{4}} - \frac{X_{3}}{(2(\sigma^{\ast})^{2})^{2}} + \frac{A_{0}^{\epsilon}}{\sigma^{\ast}} + \frac{X_{2}}{2\sigma^{\ast}} - \frac{X_{1}}{4\sigma^{\ast}} - \frac{V_{1}^{\delta}V_{3}^{\prime\epsilon}}{(\sigma^{\ast})^{3}\sigma^{\ast\prime}} + \sigma^{\ast} + \frac{V_{3}^{\ast}}{2\sigma^{\ast}} \\ m &= \frac{X_{5}}{2} + \frac{X_{4}}{4} - \frac{X_{3}}{8} + \frac{5(V_{1}^{\delta})^{2}}{6(\sigma^{\ast})^{3}} - \frac{V_{0}^{\delta}V_{3}^{\epsilon}}{2(\sigma^{\ast})^{2}} + \frac{B_{2}^{\delta}}{6\sigma^{\ast}} \\ - \frac{2V_{1}^{\delta}V_{1}^{\prime\delta}}{3(\sigma^{\ast})^{2}\sigma^{\ast\prime}} + \frac{V_{0}^{\delta}V_{3}^{\prime\epsilon}}{2\sigma^{\ast}\sigma^{\ast\prime}} + \frac{V_{1}^{\delta}V_{3}^{\prime\epsilon}}{4\sigma^{\ast}\sigma^{\ast\prime}} + V_{0}^{\delta} + \frac{V_{1}^{\delta}}{2} \\ n &= \frac{(V_{0}^{\delta})^{2}}{6\sigma^{\ast}} + \frac{V_{0}^{\delta}V_{1}^{\delta}}{6\sigma^{\ast}} + \frac{(V_{1}^{\delta})^{2}}{6\sigma^{\ast}} - \frac{B_{2}^{\delta}\sigma^{\ast}}{12} + \frac{2V_{0}^{\delta}V_{0}^{\prime\delta}}{3\sigma^{\ast\prime}} + \frac{V_{0}^{\delta}V_{1}^{\delta}}{3\sigma^{\ast\prime}} + \frac{V_{1}^{\delta}V_{1}^{\prime\delta}}{3\sigma^{\ast\prime}} + \frac{V_{1}^{\delta}V_{1}^{\prime}}{3\sigma^{\ast\prime}} \\ p &= -\frac{3(V_{3}^{\delta})^{2}}{(\sigma^{\ast})^{4}} - \frac{3V_{1}^{\delta}V_{3}^{\epsilon}}{(\sigma^{\ast})^{3}} + \frac{X_{4}}{(\sigma^{\ast})^{2}} + \frac{V_{0}^{\delta}V_{3}^{\prime\epsilon}}{(\sigma^{\ast})^{3}\sigma^{\ast\prime}} + \frac{V_{1}^{\delta}V_{3}^{\prime\epsilon}}{(\sigma^{\ast})^{3}\sigma^{\ast\prime}} + \frac{V_{1}^{\delta}}{(\sigma^{\ast})^{2}} \\ q &= -\frac{5V_{0}^{\delta}V_{1}^{\delta}}{(\sigma^{\ast})^{4}} - \frac{5(V_{1}^{\delta})^{2}}{3(\sigma^{\ast})^{2}\sigma^{\ast\prime}} + \frac{2V_{0}^{\delta}V_{1}^{\delta}}{3(\sigma^{\ast})^{2}\sigma^{\ast\prime}} + \frac{2V_{0}^{\delta}V_{1}^{\prime}}{3(\sigma^{\ast})^{2}\sigma^{\ast\prime}} \\ u &= -\frac{3(V_{3}^{\varepsilon})^{2}}{(\sigma^{\ast})^{7}} + \frac{X_{1}}{(\sigma^{\ast})^{5}} \\ w &= -\frac{7(V_{1}^{\delta})^{2}}{(\sigma^{\ast})^{5}} + \frac{B_{2}^{\delta}}{3(\sigma^{\ast})^{3}} + \frac{2V_{1}^{\delta}V_{1}^{\prime\epsilon}}{3(\sigma^{\ast})^{2}\sigma^{\ast\prime}} \\ \end{array}$$

where all left values are known from the first stage of the calibration and we treat the parameters ( $\sigma^{\star}$ ,  $V_0^{\delta}$ ,  $V_1^{\delta}$ ,  $V_3^{\epsilon}$ ) as known values (obtained from the time-dependent calibration to a first-order perturbation expansion).

Given our setup, from the equation for u we easily obtain

$$X_{1} = (\sigma^{*})^{5} \left( u + \frac{3(V_{3}^{\epsilon})^{2}}{(\sigma^{*})^{7}} \right)$$
  
=  $u(\sigma^{*})^{5} + \frac{3(V_{3}^{\epsilon})^{2}}{(\sigma^{*})^{2}}.$  (4.37)

From k we then extract  $\phi^{\epsilon}$  as

$$\phi^{\epsilon} = 2\sigma^{*} \left( \frac{3(V_{3}^{\epsilon})^{2}}{2(\sigma^{*})^{5}} - \frac{X_{1}}{(\sigma^{*})^{3}} - k \right) = \frac{3(V_{3}^{\epsilon})^{2}}{(\sigma^{*})^{4}} - \frac{2X_{1}}{(\sigma^{*})^{2}} - 2k\sigma^{*}$$

$$= \frac{3(V_{3}^{\epsilon})^{2}}{(\sigma^{*})^{4}} - 2u(\sigma^{*})^{3} - \frac{6(V_{3}^{\epsilon})^{2}}{(\sigma^{*})^{4}} - 2k\sigma^{*}$$

$$= -\frac{3(V_{3}^{\epsilon})^{2}}{(\sigma^{*})^{4}} - 2u(\sigma^{*})^{3} - 2k\sigma^{*}.$$
(4.38)

Next, from p we get  $X_2$  as

$$X_2 = p(\sigma^*)^3 + \frac{3(V_3^{\epsilon})^2}{2(\sigma^*)^2} - V_3^{\epsilon}$$
(4.39)

The system of 3 equations (n, s, w)

$$\begin{cases} n = \frac{(V_0^{\delta})^2}{6\sigma^*} + \frac{V_0^{\delta}V_1^{\delta}}{6\sigma^*} + \frac{(V_1^{\delta})^2}{6\sigma^*} - \frac{B_2^{\delta}\sigma^*}{12} + \frac{2V_0^{\delta}V_0'^{\delta}}{3\sigma^{*'}} + \frac{V_0'^{\delta}V_1^{\delta}}{3\sigma^{*'}} + \frac{V_0^{\delta}V_1'^{\delta}}{3\sigma^{*'}} + \frac{V_1^{\delta}V_1'^{\delta}}{3\sigma^{*'}}, \\ s = -\frac{5V_0^{\delta}V_1^{\delta}}{3(\sigma^*)^3} - \frac{5(V_1^{\delta})^2}{6(\sigma^*)^3} + \frac{2V_0'^{\delta}V_1^{\delta}}{3(\sigma^*)^2\sigma^{*'}} + \frac{2V_0^{\delta}V_1'^{\delta}}{3(\sigma^*)^2\sigma^{*'}} + \frac{2V_1^{\delta}V_1'^{\delta}}{3(\sigma^*)^2\sigma^{*'}}, \\ w = -\frac{7(V_1^{\delta})^2}{3(\sigma^*)^5} + \frac{B_2^{\delta}}{3(\sigma^*)^3} + \frac{2V_1^{\delta}V_1'^{\delta}}{3(\sigma^*)^4\sigma^{*'}} \end{cases}$$
(4.40)

has the 3 unknowns  $B_2^{\delta}, \frac{V_1^{\prime\delta}}{\overline{\sigma}}, \frac{V_0^{\prime\delta}}{\overline{\sigma}}$  and therefore can be completely solved. n depends on all 3 unknowns, s on  $\frac{V_1^{\prime\delta}}{\overline{\sigma}}, \frac{V_0^{\prime\delta}}{\overline{\sigma}}$ , and w on  $B_2^{\delta}, \frac{V_1^{\prime\delta}}{\overline{\sigma}}$ . We omit here the explicit analytical solutions for  $B_2^{\delta}, \frac{V_1^{\prime\delta}}{\overline{\sigma}}, \frac{V_0^{\prime\delta}}{\overline{\sigma}}$  as their derivation is rather lengthy. The 3 equation system can also be deterministically solved in Matlab in negligible time, so we will treat the previously mentioned variables as known going forward.

Out of the 10 non-linear equations, we are left with only 4 unconsidered ones with 5 unknowns  $(X_3, X_4, X_5, \frac{V_3^{\prime \delta}}{\overline{\sigma}})$ :

$$\begin{cases} l = \frac{3V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} - \frac{X_3}{2(\sigma^*)^2} + \frac{A_0^{\epsilon}}{\sigma^*} + \frac{X_2}{2\sigma^*} - \frac{X_1}{4\sigma^*} - \frac{V_1^{\delta}V_3^{\prime\epsilon}}{(\sigma^*)^3\sigma^{*\prime}} + \sigma^* + \frac{V_3^{\epsilon}}{2\sigma^*} \\ m = \frac{X_5}{2} + \frac{X_4}{4} - \frac{X_3}{8} + \frac{5(V_1^{\delta})^2}{6(\sigma^*)^3} - \frac{V_0^{\delta}V_3^{\epsilon}}{2(\sigma^*)^2} + \frac{B_2^{\delta}}{6\sigma^*} \\ - \frac{2V_1^{\delta}V_1^{\prime\delta}}{3(\sigma^*)^2\sigma^{*\prime}} + \frac{V_0^{\delta}V_3^{\prime\epsilon}}{2\sigma^*\sigma^{*\prime}} + \frac{V_1^{\delta}V_3^{\prime\epsilon}}{4\sigma^*\sigma^{*\prime}} + V_0^{\delta} + \frac{V_1^{\delta}}{2} \\ q = -\frac{3V_0^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} - \frac{3V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} + \frac{X_4}{2(\sigma^*)^2} + \frac{V_0^{\delta}V_3^{\prime\epsilon}}{(\sigma^*)^3\sigma^{*\prime}} + \frac{V_1^{\delta}V_3^{\prime\epsilon}}{(\sigma^*)^3\sigma^{*\prime}} + \frac{V_1^{\delta}}{(\sigma^*)^3\sigma^{*\prime}} \\ v = -\frac{6V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^6} + \frac{X_3}{2(\sigma^*)^4} + \frac{V_1^{\delta}V_3^{\prime\epsilon}}{(\sigma^*)^5\sigma^{*\prime}} \end{cases}$$
(4.41)

We observe that v depends on unknowns  $(X_3, \frac{V_3^{\prime\delta}}{\overline{\sigma}})$  and l on unknowns  $(X_3, \frac{V_3^{\prime\delta}}{\overline{\sigma}}, A_0^{\epsilon})$ . Taking advantage of the common structure of the unknowns  $(X_3, \frac{V_3^{\prime\delta}}{\overline{\sigma}})$  in both equations we replace

 $X_3$  in *l* from *v* determining  $A_0^{\epsilon}$ :

$$l = \frac{3V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^4} + \frac{A_0^{\epsilon}}{\sigma^*} + \frac{X_2}{2\sigma^*} - \frac{X_1}{4\sigma^*} + \sigma^* + \frac{V_3^{\epsilon}}{2\sigma^*} - \left(v(\sigma^*)^2 + \frac{6V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^4}\right)$$
  

$$\Rightarrow A_0^{\epsilon} = l\sigma^* - \frac{9V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^3} - \frac{X_2}{2} + \frac{X_1}{4} - (\sigma^*)^2 - \frac{V_3^{\epsilon}}{2} + v(\sigma^*)^3$$
  

$$= l\sigma^* - \frac{9V_1^{\delta}V_3^{\epsilon}}{(\sigma^*)^3} - \frac{p(\sigma^*)^3}{2} - \frac{3(V_3^{\epsilon})^2}{4(\sigma^*)^2} + \frac{V_3^{\epsilon}}{2} + \frac{u(\sigma^*)^5}{4} + \frac{3(V_3^{\epsilon})^2}{4(\sigma^*)^2} - (\sigma^*)^2 - \frac{V_3^{\epsilon}}{2} + v(\sigma^*)^3$$
  
(4.42)

Therefore we discard equation l as we used it to find  $A_0^{\epsilon}$  and are left with equations v, q, m that depend on unknowns  $(X_3, \frac{V_3^{\prime\delta}}{\overline{\sigma}}), (X_4, \frac{V_3^{\prime\delta}}{\overline{\sigma}}), (X_3, X_4, X_5, \frac{V_3^{\prime\delta}}{\overline{\sigma}})$  respectively. Using the common algebraic structures around the unknown parameters present in equations v and q in comparison to m, we replace unknown parameters  $(X_3, X_4, \frac{V_3^{\prime\delta}}{\overline{\sigma}})$  in m with known values using v and q. Thus we end up with a deterministic expression for  $X_5$  (we omit cumbersome details here).

At this final stage we are left with only two equations v and q and three unknown parameters  $(X_3, X_4, \frac{V_3^{\prime\delta}}{\overline{\sigma}})$ . Here we cannot proceed further with the reduction and need an L2 minimization on the set of parameters  $\{X_3, X_4, \frac{V_3^{\prime\delta}}{\overline{\sigma}}\}$ .

### 4.6.1 Concluding remarks

In conclusion, using a hybrid calibration and a very accurate approximate parameter reduction we have managed to significantly reduce the complexity of the non-linear inversion problem for market parameters. An implementation of the quasi-closed form solution would therefore proceed as follows:

- 1. Compute market parameters  $(\sigma^*, V_0^{\delta}, V_1^{\delta}, V_3^{\epsilon})$  from a fast time-dependent first order calibration to the volatility surface (as described in Chapter 3).
- 2. Compute the 10 estimated parameters ( $\Theta := \{k, l, m, n, p, q, s, u, v, w\}$ ) from the volatility surface using the first stage of the second order calibration as described in equation (4.17). Notice that first stage calibration is a least squares problem with no constraints, and is therefore a convex problem. This means a solution can be found easily and fast using a simple local solver.
- 3. Compute the 15 parameters

$$\chi := \{\sigma^*, \ V_3^{\epsilon}, \ V_1^{\delta}, \ V_0^{\delta}, \ A_0^{\epsilon}, \ B_2^{\delta}, \ \frac{V_3^{\prime \epsilon}}{\overline{\sigma}}, \ \frac{V_1^{\prime \delta}}{\overline{\sigma}}, \ \frac{V_0^{\prime \delta}}{\overline{\sigma}}, \ \phi^{\epsilon}, \ X_1, \ X_2, \ X_3, \ X_4, \ X_5\}$$

from those found in step 2:

- $\{\sigma^*, V_3^{\epsilon}, V_1^{\delta}, V_0^{\delta} \text{ were found in step 1.} \}$
- $\{A_0^{\epsilon}, B_2^{\delta}, \frac{V_3^{\epsilon}}{\overline{\sigma}}, \frac{V_1^{\delta}}{\overline{\sigma}}, \frac{V_0^{\delta}}{\overline{\sigma}}, \phi^{\epsilon}, X_1, X_2, X_3, X_4, X_5\}$  are found by direct analytical expressions as functions of the 15 estimated parameters and the 4 market parameters already known.
- {V''<sub>3</sub>/σ̄, X<sub>3</sub>, X<sub>4</sub>, } are found by least squares optimization using as constraints the equations for v and q from the equations system 4.36. This is the highest dimension (3) constrained least squares optimization required, reduced from 18 initial dimensions.
- 4. Recover parameters  $(A_2^{\epsilon}, A_1^{\epsilon}, A^{\epsilon}, C_2^{\epsilon,\delta}, C_1^{\epsilon,\delta}, C_0^{\epsilon,\delta}, B_1^{\delta})$  from optimal parameters  $(X_1, X_2, X_3, X_4, X_5)$  found in the previous step. Here we need to perform 5 least squares minimizations of dimension 2 with one linear constraint of the form 4.35.
- 5. We have a complete set  $\Phi^*$  of 18 market parameters for the second-order perturbation theory. The error during the inversion from  $\Theta^*$  is of an order of magnitude comparable to  $\epsilon$  and  $\delta$  (volatility time scales parameters that  $\rightarrow 0$ ).

### Chapter 5

### Conclusion

In the scope of this project, we have performed an extensive empirical investigation, which aided by theoretical considerations, led to several extensions and improvements to the multiscale stochastic volatility theory extensively covered in [2]. We have implemented the implied volatility blending technique described originally in [7] to construct an implied volatility surface fit for calibration. We then demonstrated empirically how time dependency impacts the calibration of the first-order expansion and studied parameter stability and goodness of fit with equal importance. This led to the optimization of the calendar function for maturity cycles that improved on previous results, and also showcased the weaknesses of the first-order expansion.

Our calibration implementations of the second order expansion showed its computational intractability in the form detailed in [2], even when using powerfull global solvers. We took further the problem of finding an optimal set of 18 parameters in the second calibration stage and attempted to reduce the dimensionality of the problem. The hybrid calibration we considered showed moderate success, as a large subset of the parameter set showed very low variation, but a few parameters were highly volatile. We then defined and formalized a parameter reduction technique to eliminate 3 dimensions of the optimization search space for the calibration. This approximate technique can be used with any global solver and is not conditioned on any assumption, which leaves further scope for experimentation. Therefore this is one of the main achievements of the project. Furthermore, conditioning the parameter reduction on a hybrid calibration yielded a quasi-closed form solution for the second stage calibration. Further empirical work is required though to test the properties of this solution. Nevertheless, the project has been largely successfull given the implementations allowing for empirical assessment of model calibration limitations, as well as the ideas developed for dimensionality reduction.

### 5.1 Future Work

A set of extensions could potentially improve the results presented here. However, the work done as part of the project builds a solid platform for both further iterative improvements to the multiscale theory and implementation as well as possible applications:

- The perturbation expansions applicable to European-style derivatives can be further extended to price American-style contracts. This can be done as detailed in [2] by using either importance sampling or control variates to reduce variance of Monte Carlo simulations. However, a strong background in econometrics is preffered, as the simulation may need to be customized for variance reduction.
- We have implicitly assumed a constant interest rate throughout our experiments. A natural extension would be to either employ historical interest rates or model explicitly an interest rate variation dependent on the volatility factors, as proposed in [2]. This has the potential to further improve the time stability of the fitted parameters, particularly those with an evolution that follows economic cycles.
- After applying the parameter reduction step in the second stage of calibration for the second order theory, state-of-the-art global optimization solvers such as BARON or LINDOGlobal could be employed successfully. This approach would remove reliance on the bridging with the first-order theory results and could potentially find globally optimal solutions. Knowledge of the modelling language GAMS is required however.
- A Principal Component Analysis (PCA) method could be devised to provide an empirical framework for comparison between implied volatility models. Thus we could potentially analyse the amount of variance a model variant captures along the maturity or strike axis and how the implied volatility surface distributes variance itself between the two dimensions. This approach can lead to a more fundamental understanding of dynamics of the implied volatility surface and provide a more direct benchmark for comparison amongst models.

## Appendix A

# Appendix A

Appendix Content

### Bibliography

- [1] Jim Gatheral. The Volatility Surface (A practitioner's Guide). Wiley Finance, 2006.
- F. al. Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives. Cambridge University Press. ISBN 9781139502450. URL http://books.google.co. uk/books?id=pBRpVGVqELgC.
- [3] R. Durrett. Probability: Theory and Examples. 4th Edition. Cambridge U Press, 2010.
- [4] J. C. Hull. Options, Futures, and Other Derivatives. Prentice-Hall, Upper Saddle River, N.J., sixth edition, 2006.
- [5] M.S. Joshi. The Concepts and Practice of Mathematical Finance. 2003.
- [6] Tomas (mathmaticien) Bjrk. Arbitrage theory in continuous time. Oxford university press, Oxford, 2004. ISBN 0-19-927126-7. URL http://opac.inria.fr/record=b1123434.
- [7] Stephen Figlewski. The implied risk neutral density for the u.s. market portfolio, 2007.
- [8] B. Dumas, Jeff Fleming, and Robert E. Whaley. Implied volatility functions: Empirical tests, 1995.
- [9] B. Dupire. Pricing with a smile. *Risk*, (7):18–20, 1994.
- [10] E. Derman and I. Kani. The volatility smile and its implied tree, 1994.
- [11] Mark Rubinstein. Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active cboe option classes from august 23, 1976 through august 31, 1978. Research Program in Finance Working Papers 117, University of California at Berkeley, 1981. URL http://EconPapers.repec.org/RePEc: ucb:calbrf:117.
- [12] P. K. Clark. A subordinated stochastic process model with fixed variance for speculative prices, 1973.
- [13] Robert F. Engle and Andrew J. Patton. What good is a volatility model. Quantitative Finance, pages 237–245, 2000.

- [14] Jean-Pierre Fouque and Matthew J. Lorig. A fast mean-reverting correction to heston's stochastic volatility model. SIAM J. Financial Math., 2(1):221–254, 2011.
- [15] Gurdip Bakshi, Charles Cao, and Zhiwu Chen. Empirical performance of alternative option pricing models, 1997.
- [16] Artur Sepp. Stochastic local volatility models: Theory and implementation, 2010.
- [17] G. Yin and Hanqin Zhang. Two-time-scale markov chains and applications to quasibirth-death queues. SIAM journal on applied mathematics, 65, 2005.
- [18] Ulrich A. Muller, Michel M Dacorogna, Rakhal D. Dave, Richard Olsen, Olivier V. Pictet, and Jakob von Weizscker. Volatilities of different time resolutions – analyzing the dynamics of market components. *Journal of Empirical Finance*, 4(2-3):213-239, 1997. URL http://EconPapers.repec.org/RePEc:eee:empfin:v:4:y:1997:i:2-3: p:213-239.
- [19] Jean-Pierre Fouque, George Papanicolaou, Ronnie Sircar, and Knut Solna. Multiscale stochastic volatility asymptotics. SIAM J. MULTISCALE MODELING AND SIMU-LATION, pages 22–42, 2003.
- [20] Jean-Pierre Fouque, George Papanicolaou, Ronnie Sircar, and Knut Solna. Multiscale stochastic volatility asymptotics. SIAM J. MULTISCALE MODELING AND SIMU-LATION, pages 22–42, 2003.
  - [] Jean-Pierre Fouque, Matthew Lorig, and Ronnie Sircar. Second order multiscale stochastic volatility asymptotics: Stochastic terminal layer analysis calibration. Papers, arXiv.org, 2012. URL http://EconPapers.repec.org/RePEc:arx:papers: 1208.5802.