

Pricing and calibration with stochastic local volatility models in a monte carlo setting

van der Stoep, Anton

DOI

[10.4233/uuid:a2314b15-6fee-4be1-bf77-f78b4356d11a](https://doi.org/10.4233/uuid:a2314b15-6fee-4be1-bf77-f78b4356d11a)

Publication date

2019

Document Version

Final published version

Citation (APA)

van der Stoep, A. (2019). *Pricing and calibration with stochastic local volatility models in a monte carlo setting*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:a2314b15-6fee-4be1-bf77-f78b4356d11a>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

**PRICING AND CALIBRATION WITH STOCHASTIC
LOCAL VOLATILITY MODELS IN A MONTE CARLO
SETTING**

PRICING AND CALIBRATION WITH STOCHASTIC LOCAL VOLATILITY MODELS IN A MONTE CARLO SETTING

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 26 maart 2019 om 12:30 uur

door

Anthony Willem VAN DER STOEP

Master of Science in Toegepaste Wiskunde, Universiteit Twente, Nederland
geboren te Barendrecht, Nederland.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. ir. C.W. Oosterlee	Technische Universiteit Delft, promotor
Dr. ir. L.A. Grzelak	Technische Universiteit Delft, copromotor

Onafhankelijke leden:

Prof. dr. C. Vázquez Cendón	Universidade da Coruña, Spanje
Prof. dr. K.J. In 't Hout	Universiteit Antwerpen, België
Prof. dr. P. J.C. Spreij	Universiteit van Amsterdam
Prof. dr. ir. A.W. Heemink	Technische Universiteit Delft
Drs. W.F. van Raaij	Rabobank

Reservelid:

Prof. dr. ir. G. Jongbloed	Technische Universiteit Delft
----------------------------	-------------------------------



Rabobank



Centrum Wiskunde & Informatica



Pricing and Calibration with Stochastic Local Volatility Models in a Monte Carlo Setting
Dissertation at Delft University of Technology
Copyright © 2019 by A.W. van der Stoep
Printed by: Gildeprint

ISBN 978-94-6323-540-2

An electronic version of this dissertation is available at
<http://repository.tudelft.nl/>.

SUMMARY

A general purpose of mathematical models is to accurately mimic some observed phenomena in the real world. In financial engineering, for example, one aim is to reproduce market prices of financial contracts with the help of applied mathematics.

In the *Foreign Exchange* (FX) market, the so-called *implied volatility smile* plays a key role in the pricing and hedging of financial derivative contracts. This volatility smile is a phenomenon that reflects the prices of European-type options for different strike prices; the implied volatility tends to be higher for options that are deeper In The Money and Out of The Money than options that are approximately At The Money. In order for a pricing model to be accepted in the financial industry, it should at least be able to accurately price back the most simple financial derivative contracts, namely European call and put options. In other words, the model should *calibrate* well to the implied volatility smile observed in the financial market. The calibration should not only be accurate, but also reasonably fast.

Another feature we wish the financial asset model to possess, is an accurate pricing of so-called *exotic* financial products. Exotic products are not traded on regular exchanges, but *over-the-counter*, i.e. directly between two parties without the supervision of an exchange. An example is a barrier option, which is a financial contract of which its payoff depends on the possible event that the underlying asset price hits a certain pre-determined level. The model prices of these path-dependent contracts are determined by the transition densities of the relevant underlying asset(s) between future time-points. These transition densities are reflected by the *forward volatility smile* the model implies; in order for the model to accurately price exotic products, it should yield realistic forward volatilities.

The models discussed in this thesis can be considered as enhancements of Dupire's classical and famous *Local Volatility* (LV) model [34, 35], which by its non-parametric local volatility component yields a perfect calibration to any set of arbitrage-free European-type options prices. We will consider the addition to the LV model of stochastic volatility, resulting in the *stochastic local volatility* (SLV) model [75, 80, 82], and we also add stochastic interest rates – for both extensions, a perfect calibration is preserved. As an alternative to the LV model, Hagan et al. [63] introduced the *SABR* model. We further enrich this model by time-dependent parameters and propose an efficient calibration procedure. Also, we introduce a novel asset model dynamics class, the *Collocating Local Volatility* (CLV) model. Similarly to an SLV model, by construction the CLV model is perfectly calibrated to liquid (i.e. heavily traded) market quotes, while maintaining the flexibility in accurately capturing the forward volatility smile.

In some more detail, in Chapter 2 we explain for a general SLV model the local volatility component, which consists of Dupire's local volatility component and a non-trivial conditional expectation. We present a Monte Carlo approach for the efficient evaluation of a general SLV model. The approach is based on an intuitive and easy to imple-

ment non-parametric approximation of the conditional expectation, which consists of assigning asset path realizations to appropriate ‘bins’. This approximation is embedded in a simulation scheme that is strongly based on a so-called *Quadratic Exponential* (QE) scheme [5], which introduces less bias than more common Euler schemes. We show by means of numerical experiments and an error analysis that our approach yields accurate prices for European-type options. Also, we price forward-start options and determine the corresponding forward volatility smile; we observe that the Heston-SLV model preserves the current shape of the implied volatility smile, which is typically more in line with financial market observations.

The time-dependent FX-SABR model is discussed in Chapter 3. In contrast to the constant-parameter SABR model introduced by Hagan et al. [63], the time-dependent model can be calibrated to an *implied volatility surface*; in this way, it captures as much market information as possible. However, the calibration of time-dependent parameters is non-trivial. We propose an efficient calibration approach that is based on *effective parameters*, which can be considered as ‘sophisticated averages’ of the corresponding time-dependent parameters. By considering the qualitative effects of the SABR parameters on the shape of the implied volatility smile, we derive ‘mappings’ between the time-dependent parameters and their ‘constant counterparts’. The mappings allow for an efficient calibration of the time-dependent parameters. Numerical experiments confirm that both the separate and combined performance of the effective parameters are accurate. Also, we numerically show that the effective parameters derived yield highly satisfactory calibration results. In a barrier option pricing experiment, the time-dependent FX-SABR model yields more accurate prices than the traditional LV and SABR models.

We consider two types of *hybrid* local volatility models in Chapter 4, namely the SABR and Heston models enhanced by Dupire’s local volatility component, and the Local Volatility model enriched with stochastic interest rates. For both model classes one needs to determine an, above mentioned, non-trivial (conditional) expectation, which is expensive to compute and cannot be extracted from the market quotes. In this chapter, we propose a second generic efficient Monte Carlo approach to these hybrid models, which consists of two projection steps. The first step is based on *stochastic collocation*, where in general a certain ‘expensive’ variable of interest Y is approximated by a function of a ‘cheaper to evaluate’ random variable X – more concrete, a Lagrange interpolation is established through so-called optimal *collocation points* that are determined based on the distribution of X . The second projection step relies on standard regression techniques. We numerically show that our approach yields a fast Monte Carlo evaluation and highly accurate pricing results for European-type options. Also, we provide an error analysis for a model consisting of two correlated Geometric Brownian Motions, the ‘2D-GBM’ model.

In Chapter 5 we introduce a novel asset price model class, namely the *Collocating Local Volatility* (CLV) model. The CLV model consists of two elements, a kernel process and a local volatility function. The kernel process can be chosen freely and determines the forward volatility smile the model implies. The local volatility function, which is constructed based on stochastic collocation, ensures a perfect calibration to financial market quotes. We compare three different kernel processes within the CLV model framework, namely the Ornstein-Uhlenbeck (OU) and Cox-Ingersoll-Ros (CIR) process

and the Heston model. For each of these kernels, we consider the effect of the kernel parameters on the shape of the forward volatility smile. Subsequently, we calibrate the OU-CLV, CIR-CLV and Heston-CLV models to FX barrier option prices observed in the market by means of a Monte Carlo simulation, where we make use of Brownian bridge techniques.

With the work in this thesis, one is able to price complicated FX derivatives by efficient Monte Carlo methods, while an accurate calibration to liquid market quotes is preserved. The work developed in this PhD thesis is based on journal articles, that have either been published or been submitted during the doctoral research period.

SAMENVATTING

Een doel van wiskundige modellen is om zo nauwkeurig mogelijk bepaalde fenomenen uit de wereld te modelleren. In de financiële sector is bijvoorbeeld een doel om marktprijzen van financiële contracten te reproduceren met behulp van toegepaste wiskunde.

In de *buitenlandse valutamarkt* (de ‘FX markt’) speelt bijvoorbeeld de zogenoemde *impliciete volatiliteit* een belangrijke rol bij het waarderen van financiële derivaten. De geobserveerde impliciete volatiliteitsfunctie is een fenomeen dat de prijzen van Europese opties representeert voor verschillende uitoefenprijzen; de impliciete volatiliteit neigt hoger te zijn voor opties die dieper In The Money en Out The Money zijn, dan voor opties die bij benadering At The Money zijn. Een andere eis aan een waarderingsmodel om geaccepteerd te worden in de financiële industrie, is dat het tenminste de meest simpele financiële derivaten nauwkeurig kan terugwaarderen, namelijk Europese call- en putopties. Met andere woorden, het model moet goed *gecalibreerd* kunnen worden aan de impliciete volatiliteitsfunctie die in de financiële markt wordt waargenomen. De calibratie moet niet alleen accuraat zijn, maar ook snel genoeg.

Een andere gewenste eigenschap van een financieel model, is dat het nauwkeurig zogenoemde *exotische* financiële producten kan waarderen. Exotische producten worden niet verhandeld op reguliere beurzen, maar direct tussen partijen zonder de tussenkomst van een beurs. Een voorbeeld is een barrieroptie, waarvan de uitbetaling afhangt van de mogelijke gebeurtenis dat een onderliggende wisselkoersprijs een vooraf bepaald niveau bereikt. De modelprijzen van deze padsafhankelijke contracten worden bepaald door de ‘transitiekansen’ van de relevante onderliggende prijzen, tussen toekomstige tijdstippen. Deze ‘transitiekansen’ worden gerepresenteerd door de *voorwaartse volatiliteitsfunctie* die het model impliceert; een model moet een realistische voorwaartse volatiliteit impliceren om exotische financiële producten accuraat te kunnen waarderen.

De modellen die in dit proefschrift behandeld worden, kunnen beschouwd worden als verbeteringen van Dupire’s klassieke en beroemde *Locale Volatiliteitsmodel* (LV) [34, 35], dat vanwege een niet-parametrische locale volatiliteitscomponent perfect gecali-breerd kan worden aan een willekeurige set van arbitrage-vrije Europese optieprijzen. We zullen de toevoeging van stochastische volatiliteit aan het LV model beschouwen, hetgeen resulteert in het zogeheten *stochastische locale volatiliteitsmodel* (SLV) [75, 80, 82], en we voegen ook stochastische rente toe — met beide uitbreidingen wordt een perfecte calibratie behouden. Als een alternatief voor het LV model, introduceerden Hagan et al. [63] het *SABR* model. Wij verrijken dit model met tijdsafhankelijke parameters en stellen een efficiënte calibratieprocedure voor. Ook introduceren we in dit proefschrift een nieuwe modelklasse, namelijk het *Collocerende Locale Volatiliteitsmodel* (CLV). Net als het SLV model, kan het CLV model perfect gecali-breerd worden aan de prijzen van intensief verhandelde opties, terwijl het model de flexibiliteit behoudt om accuraat de voorwaartse volatiliteitsfunctie te behouden.

In meer detail, in Hoofdstuk 2 leggen we voor een algemeen SLV model de locale

volatiliteitscomponent uit, dat bestaat uit Dupire's locale volatiliteit en een niet-triviale conditionele verwachting. We presenteren een Monte Carlo methode voor een efficiënte evaluatie van een algemeen SLV model. De methode is gebaseerd op een intuïtieve niet-parametrische benadering van de conditionele verwachting, die gemakkelijk geïmplementeerd kan worden. In deze methode kennen we de realisaties van de onderliggende prijzen toe aan de 'deelintervallen'. De benadering is geïmplementeerd in een simulatieschema dat gebaseerd is op het *Kwadratische Exponentiële* schema [5], dat minder ruis veroorzaakt dan de conventiële Euler schema's. Door middel van numerieke experimenten en een foutenanalyse laten we zien dat onze methode accurate prijzen bepaalt voor Europese opties. Ook waarderen we 'voorwaartsstartende' opties en bepalen de bijbehorende 'voorwaartse volatiliteitsfunctie'; we observeren dat het Heston-SLV model de vorm van de volatiliteitsfunctie behoudt, die in overeenstemming is met observaties in de financiële markt.

Het tijdsafhankelijke FX-SABR model wordt behandeld in Hoofdstuk 3. In tegenstelling tot het constante SABR model geïntroduceerd door Hagan et al. [63], kan het tijdsafhankelijke model gecalibreerd worden aan een *impliciet volatiliteitsvlak*; zodoende kan het zo veel mogelijk marktinformatie meenemen. De calibratie van tijdsafhankelijke parameters is echter niet triviaal. We stellen een efficiënte calibratiemethode voor die gebaseerd is op *effectieve parameters*, die beschouwd kunnen worden als 'geavanceerde gemiddeldes' van de bijbehorende tijdsafhankelijke parameters. Door de kwalitatieve effecten van de SABR parameters op de vorm van de impliciete volatiliteitsfunctie te beschouwen, leiden we 'projecties' af tussen tijdsafhankelijke parameters en de 'constante tegenhangers'. De projecties leiden tot een efficiënte calibratie van de tijdsafhankelijke parameters. Numerieke experimenten tonen aan dat zowel de individuele als de gecombineerde effecten van de effectieve parameters nauwkeurig zijn. Ook tonen we numeriek aan dat het gebruik van de effectieve parameters tot de gewenste calibratieresultaten leidt. In een experiment waarin we een barrieroptie waarderen, geeft het tijdsafhankelijke FX-SABR model accuratere prijzen dan de traditionele LV en SABR modellen.

We beschouwen twee typen *hybride* locale volatiliteitsmodellen in Hoofdstuk 4, namelijk de SABR en Heston modellen die uitgebreid zijn met Dupire's locale volatiliteitscomponent, en het Locale Volatiliteitsmodel waaraan stochastische rente is toegevoegd. Voor beide modelklassen moet men een, ook bovengenoemde, niet-triviale (conditionele) verwachting bepalen, die numeriek gezien 'duur' is om te berekenen en niet uit marktprijzen kan worden afgeleid. In dit hoofdstuk stellen we een tweede, algemene, efficiënte Monte Carlo methode voor die uit twee projectiestappen bestaat. De eerste stap is gebaseerd op stochastische collocatie, waarin in het algemeen een zekere 'dure' variabele Y benaderd wordt door een functie van een 'goedkoper te evalueren' variabele X — concreter, we bepalen een Lagrange interpolatie met *collocatiepunten* die berekend worden op basis van de verdeling van X . De tweede projectiestap steelt op standaard regressietechnieken. Numeriek tonen we aan dat onze methode tot een snelle Monte Carlo evaluatie leidt waarin met een hoge nauwkeurigheid Europese opties gewaardeerd kunnen worden. Ook doen we een foutenanalyse voor een model dat uit twee gecorreleerde geometrische Brownse bewegingen bestaat, het '2D-GBM' model.

In Hoofdstuk 5 introduceren we een nieuwe modelklasse, namelijk het *Collocerende*

Locale Volatiliteitsmodel (CLV). Het CLV model bestaat uit twee elementen, een kernproces en een locale volatiliteitsfunctie. Het kernproces kan vrij gekozen worden en bepaalt de voorwaartse volatiliteitsfunctie. De locale volatiliteitsfunctie is gebaseerd op stochastische collocatie, en zorgt voor een perfecte calibratie aan de financiële marktprijzen. We vergelijken drie verschillende kernprocessen binnen het CLV model, de Ornstein-Uhlenbeck (OU) en Cox-Ingersoll-Ross (CIR) processen en het Heston model. Voor elk kernproces beschouwen we het effect van de kernparameters op de vorm van de volatiliteitsfunctie. Vervolgens calibreren we met Monte Carlo simulatie de OU-CLV, CIR-CLV en Heston-CLV modellen aan prijzen van FX barrieropties in de markt, waarbij we gebruik maken van Brownse brugtechnieken.

Met het werk in dit proefschrift, kunnen niet-triviale FX derivaten gewaardeerd worden met efficiënte Monte Carlo methoden, terwijl een accurate calibratie aan marktprijzen behouden blijft. Het werk in dit proefschrift is gebaseerd op artikelen die gepubliceerd of ingediend zijn tijdens het promotietraject.

CONTENTS

Summary	v
Samenvatting	ix
1 Introduction	1
1.1 Basics of Risk-Neutral Option Pricing	1
1.2 Volatility Modelling in the FX Market	3
1.2.1 The Local Volatility model	4
1.2.2 Stochastic Volatility models	5
1.2.3 Stochastic Local Volatility models	5
1.2.4 The Collocating Local Volatility Model	7
1.3 Outline of the thesis	8
2 The Heston Stochastic-Local Volatility Model: Efficient Monte Carlo Simulation	9
2.1 Introduction	9
2.2 Stochastic-Local Volatility Model	11
2.2.1 Specifying $\sigma(t, S(t))$	11
2.3 Novel Technique for $\mathbb{E}[\psi^2(V(t)) S(t) = K]$	15
2.3.1 Non-parametric method	15
2.3.2 Continuous approximation	20
2.3.3 Efficient simulation scheme	20
2.4 Numerical Results.	23
2.4.1 European call options	24
2.4.2 Forward starting options	25
2.4.3 Calculation time	26
2.5 Error Analysis	27
2.5.1 Bound on pricing error.	28
2.5.2 Performance	31
2.5.3 Numerical experiment: choice of bins	32
2.6 Conclusion	32
2.A Additional Pricing Experiments	35
2.B Proof of Lemma 2.5.1	35
3 The Time-Dependent FX-SABR Model: Efficient Calibration based on Effective Parameters	37
3.1 Introduction	37
3.2 Time-dependent FX-SABR Model with Local Volatility	39
3.2.1 Time-dependent FX-SABR model	40
3.2.2 Local volatility compensator	41

3.3	Calibration Problem	42
3.3.1	Calibration set-up	44
3.4	Effective Parameters	46
3.4.1	Effective vol-vol parameter.	46
3.4.2	Effective term structure	49
3.4.3	Effective correlation	57
3.5	Calibration & Pricing	59
3.5.1	Calibration procedure	60
3.5.2	Calibration results	61
3.5.3	Pricing barrier options	64
3.6	Conclusion	66
3.A	Characteristic Function Recovery	67
3.A.1	Distribution of R_j	67
3.A.2	Recovery procedure	68
3.B	Additional Calibration Experiment	69
4	A Novel Monte Carlo Approach to Hybrid Local Volatility Models	71
4.1	Introduction	71
4.1.1	Stochastic local volatility models.	72
4.1.2	Local volatility model with stochastic interest rates	73
4.1.3	Stochastic collocation basics.	73
4.2	Stochastic Local Volatility Models	75
4.2.1	Establishing $\mathbb{E}[V(t) S(t) = K]$.	77
4.2.2	Enhancements	80
4.2.3	Numerical experiments	81
4.3	Local Volatility Model with Stochastic Rates.	87
4.3.1	Establishing $\mathbb{E}^Q \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t) > K} \right]$	89
4.3.2	Numerical experiments	92
4.4	Conclusion	94
4.A	Error Analysis & Discussion	99
4.A.1	Stochastic collocation error	99
4.A.2	Regression error	100
4.B	Proofs of Lemma 4.2.1 and the result in Section 4.3.1	101
5	Collocating Local Volatility: A Competitive Alternative to Stochastic Local Volatility Models	105
5.1	Introduction	105
5.2	The Collocating Local Volatility Model	106
5.2.1	Martingale considerations	110
5.3	The OU-CLV, CIR-CLV and Heston-CLV models	111
5.3.1	The OU-CLV model	112
5.3.2	The CIR-CLV model	113
5.3.3	The Heston-CLV model	117

5.4	Calibration to FX Barrier Options	121
5.4.1	Monte Carlo simulation framework	122
5.4.2	Pricing barrier options: a Brownian bridge approach	124
5.4.3	Calibration of the OU-CLV, CIR-CLV and Heston-CLV models to FX barrier options	126
5.5	Conclusion	129
5.A	Optimal Collocation Points	131
5.B	Numerical Experiment: Pricing a Discretely Monitored Barrier Option	132
5.C	Effect of CIR Parameters, Omitting the Level Effect	133
5.D	Effect of Heston Parameters, Omitting the Level Effect	134
6	Conclusions and Outlook	137
6.1	Conclusions.	137
6.2	Outlook	138
	References	141
	141
	Curriculum Vitae	151
	List of Publications	153
	List of Attended Conferences with Presentation	155

1

INTRODUCTION

1.1. BASICS OF RISK-NEUTRAL OPTION PRICING

In the financial industry, a financial *derivative* is a contract that is based on the value of some underlying *asset*, like the price of a stock (i.e. part of a company), commodity price, interest rate or a foreign exchange (FX) rate. One of the most basic derivatives is the *European option*, where a distinction is made between a ‘call’ and ‘put’ variant. A European call option gives the holder the right, but not the obligation, to *buy* a certain amount of the underlying asset for a certain price, the strike K , at a certain time T in the future. The put variant gives the holder the right to *sell*. In a mathematical notation, defining $S(t)$ as the underlying price at a certain time-point t , the European call option’s payoff at T can be represented as

$$C(T, K) = \max(S(T) - K, 0). \quad (1.1.1)$$

For a European put option we have

$$P(T, K) = \max(K - S(T), 0).$$

In order to properly determine the current price of a derivative, we model the dynamics of the underlying price $S(\cdot)$ with a certain stochastic process – in financial mathematics we typically model this process by a *stochastic differential equation* (SDE). A well-known SDE in finance is the *Geometric Brownian Motion* (GBM), driven by a Brownian motion $W(\cdot)$ that is described by a Wiener process:

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \quad S_0 := S(0), \quad (1.1.2)$$

where $W(\cdot)$ is defined under a certain probability measure, and μ and σ determine the so-called *drift* and *diffusion* of the process.

In pricing financial derivatives, we typically work with models that are free of *arbitrage* – that is, one cannot make more *risk-free* profit than the *risk-free interest rate*, which is the theoretical rate of return of some investment with no risk of a financial loss in a

certain time-period. Let r denote the risk-free interest rate on a certain *money-savings* account with value $M(\cdot)$, which is determined by $dM(t) = rM(t)dt$, $M(0) = 1$. This implies that 1 euro *today*, put on the money-savings account, is worth $M(T) = e^{rT}$ at time T . We can also invert this relation: a certain value at T , $X(T)$, is worth today $\frac{X(T)}{M(T)}$. Now, the no-arbitrage property of a model is implied by its *martingale* property, which means for a certain stochastic quantity $X(\cdot)$:

$$\mathbb{E} \left[\frac{X(T)}{M(T)} \middle| \mathcal{F}(t) \right] = \frac{X(t)}{M(t)}, \text{ for all } 0 \leq t \leq T, \quad (1.1.3)$$

where the σ -field $\mathcal{F}(t)$ represents the filtration, which is all the information that is available up to and including time t . The probability measure under which (1.1.3) holds is the so-called *risk-neutral measure*, which we will refer to by \mathbb{Q} . In order for the stochastic process in (1.1.2) to be a martingale under the risk-neutral measure, one can show that $\mu = r$ needs to hold.

In the risk-neutral \mathbb{Q} -measure today's price of a financial contract is determined by the future payoff you are expecting to receive, as seen from today. Coming back to the European call option with its payoff specified in (1.1.1), its current value at t is given by

$$C(t, K) = M(t) \mathbb{E}_t^{\mathbb{Q}} \left[\frac{C(T, K)}{M(T)} \right], \quad (1.1.4)$$

where the \mathbb{Q} -superscript indicates that the expectation is taken under the risk-neutral measure, and the t -subscript denotes the conditioning on the filtration $\mathcal{F}(t)$.

In certain specific cases, the expectation in (1.1.4) can be calculated *analytically*. However, in many cases this is not possible, and one may resort to solving a partial differential equation (PDE). The connection between the PDE and an expectation as the one in (1.1.4) is formalized by the *Feynman-Kac theorem*, which loosely states that the solution $u(t, x)$ of a PDE with a certain form and final condition $u(T, x) = \psi(x)$, can be written as a conditional expectation in terms of a stochastic variable $X(\cdot)$ that is governed by some general SDE $dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t)$. One of the most famous PDEs in finance is the *Black-Scholes pricing PDE*, which was derived based on the concept of a *replicating portfolio* [16]. Solving this PDE led to closed-form pricing formulas of European call and put options:

$$C^{\text{BS}}(T, K, \sigma) = F_{\mathcal{N}(0,1)}(d_1)S(t) - F_{\mathcal{N}(0,1)}(d_2)Ke^{-r(T-t)}, \quad (1.1.5)$$

$$P^{\text{BS}}(T, K, \sigma) = F_{\mathcal{N}(0,1)}(-d_2)Ke^{-r(T-t)} - F_{\mathcal{N}(0,1)}(-d_1)S(t), \quad (1.1.6)$$

where $S(t)$ denotes today's price of the underlying asset, $F_{\mathcal{N}(0,1)}(\cdot)$ is the standard normal CDF and d_1 and d_2 are given by

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

$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

Next to a PDE-based solution method, another alternative approach that is widely applied to calculate the value and risks of financial derivatives, is based on *Monte Carlo simulation*, which will be the general solution method in this thesis.

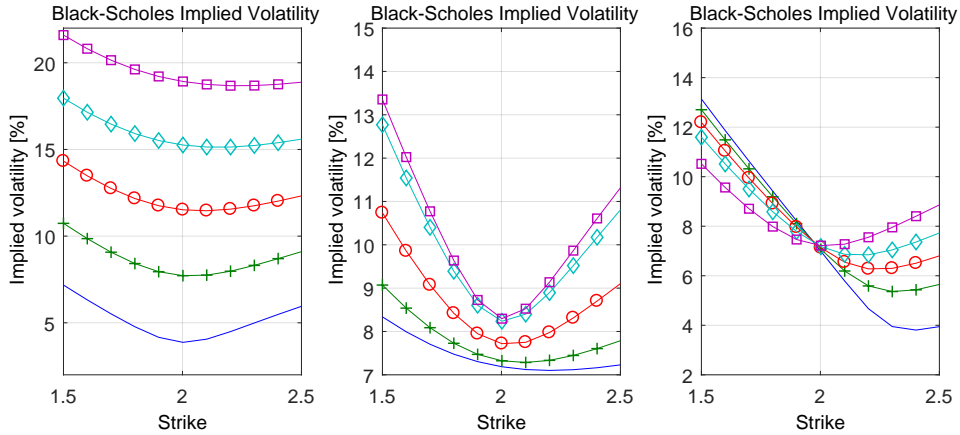


Figure 1.1.1: The implied volatility smile: its level (left), curvature (middle) and skewness (right).

The prices of European-type options are quoted in the market. It is a convention to quote the prices in terms of the *implied volatility* σ_{implied} , which is the value of σ for which the following equality holds:

$$C^{\text{BS}}(T, K, \sigma) = C^{\text{market}}(T, K),$$

where $C^{\text{BS}}(\cdot, \cdot, \cdot)$ denotes the Black-Scholes price of a European call option, see equation (1.1.5), and $C^{\text{market}}(\cdot, \cdot)$ denotes the market price. For every strike K , we generally obtain a different implied volatility value. In particular, in the FX market we typically observe a so-called *implied volatility smile*, as for the lower and higher strikes a higher implied volatility is observed as for the strikes 'in the middle'. Related to this, in this thesis we will refer to the *forward volatility smile*; this is the implied volatility smile corresponding to a *forward-start option*, which is a European call or put option that starts at some time T_1 in the future and has a payoff at $T_2 > T_1$ given by:

$$V_{\text{Forw.St.Call}}(T_2, K) = \max(S(T_2) - KS(T_1), 0).$$

In this thesis, we will refer to the *level*, *curvature* and *skewness* of the implied volatility smile. In Figure 1.1.1 we display these smile properties. Loosely speaking, the level is the 'height' of the smile, the curvature is a measure for the 'convexity', and the skewness can be seen as the 'steepness' of the smile.

The observations of an implied volatility smile and a forward volatility smile in the market, led to a more advanced modelling of volatility, which will be discussed in the next section.

1.2. VOLATILITY MODELLING IN THE FX MARKET

Over the last decades, the foreign exchange (FX) market has rapidly grown to become the world's largest and most liquid OTC market [1, 125]. According to the latest triennial survey by the Bank for International Settlements [2], an estimated \$5.09 trillion changes

hands *every day*. Foreign exchange swaps are the most actively traded instruments, followed by spot trading. Spot trading is characterized by immediate delivery of a foreign currency or commodity ‘on the spot’, as opposed to futures contracts, which typically expire before a physical delivery.

The fast expansion of the FX market, together with the rise of complex products¹, led to an increasing demand for modelling FX rates in a sophisticated way. As a consequence, there is ample literature on modelling FX rates and many stochastic models are available. In the financial industry typically a so-called three-factor pricing model is used [42, 97, 111], where FX dynamics are assumed to be lognormal and the domestic and foreign interest rates follow a Hull-White one factor Gaussian model [67]. However, this model is not capable though of generating a skew effect which we observe in the FX market. In order to overcome this issue, several researchers applied local volatility [97] and stochastic volatility models [56, 122] in an FX context, which will be discussed in Sections 1.2.1 and 1.2.2, respectively. Later, the class of stochastic local volatility models gained popularity – see Section 1.2.3. Also, a novel model class was introduced, the Collocating Local Volatility model [54], that can be calibrated to liquid market quotes and at the same time FX exotic options – see Section 1.2.4.

1.2.1. THE LOCAL VOLATILITY MODEL

For many years the Local Volatility (LV) model, introduced by Dupire [35] and Derman & Kani [34], has been considered a standard model for pricing and managing the risk of structured financial products. Given $S(\cdot)$ as the underlying price, the LV model is represented by the following stochastic differential equation (SDE), under the risk-neutral \mathbb{Q} -measure:

$$dS(t) = rS(t)dt + \sigma_{LV}(t, S(t))S(t)dW(t), \quad S_0 := S(0),$$

where r is a constant interest rate and $\sigma_{LV}(\cdot, S(\cdot))$ is the local volatility term, which is expressed in terms of derivatives of arbitrage-free prices of European call options:

$$\sigma_{LV}^2(t, K) = \frac{\frac{\partial C(t, K)}{\partial t} + rK \frac{\partial C(t, K)}{\partial K}}{\frac{1}{2} K^2 \frac{\partial^2 C(t, K)}{\partial K^2}}. \quad (1.2.1)$$

Here $C(t, K)$ denotes the price of a European call option with maturity t and strike K . By the non-parametric local volatility term in (1.2.1) the LV model satisfies a necessary condition to be accepted in the financial industry; by its construction, the LV model can be perfectly calibrated to any set of arbitrage-free European-type option prices.

Despite for its desirable property of a perfect calibration, the LV model has certain drawbacks. It exhibits a *flattening* forward smile, which may not be in line with market observations. As a result, the LV model is possibly not well-capable of pricing financial contracts that are sensitive to the forward volatility smile, such as cliquets and barrier options [23, 101]. This problem is often addressed by adopting a so-called *sticky-skew technique* which is based on the forward volatilities “as seen today”.

¹Like the so-called Power-Reverse Dual-Currency [111] and the Equity-CMS Chameleon, but also more standard barrier options.

Also, the LV model does not always accurately predict the direction of the smile movement with respect to a change in value of the underlying, which could imply unstable hedges [10, 63, 76].

Further, the model assumes deterministic interest rates, which may imply mispricing of long-term interest rate sensitive hybrids, like the *Power-Reverse Dual-Currency (PRDC)* notes in the FX market, see e.g. [17, 32, 97]. Introducing a short-rate process in the model, like Vasicek, Black & Karasinski [15] and Hull & White [67] may enhance the pricing results.

1.2.2. STOCHASTIC VOLATILITY MODELS

To overcome potential issues with the Local Volatility model, one may resort to the class of stochastic volatility models, which have also been widely employed in the financial industry. Stochastic volatility (SV) models can be characterized by a volatility term which is driven by an additional SDE. For example, a well-established SV model, the Heston model [66], is governed by the following set of SDEs, where the variance $V(\cdot)$ is driven by CIR dynamics [26]:

$$\begin{aligned} dS(t) &= rS(t)dt + \sqrt{V(t)}S(t)dW_x(t), S_0 := S(0), \\ dV(t) &= \kappa(\bar{V} - V(t))dt + \gamma\sqrt{V(t)}dW_v(t), V_0 := V(0), \end{aligned}$$

with $dW_x(t)dW_v(t) = \rho_{x,v}dt$, γ denoting the volatility of the variance – also denoted by ‘vol-vol’ – and κ is the speed of mean reversion to the long-term variance \bar{V} . The SV models can be considered as more appropriate choices [37, 46] for pricing forward volatility sensitive derivatives. Also, in the SV models the volatilities change, to a certain extent, independently of a spot price change – local volatility models do not possess this characteristic [104].

Although the SV models certainly have beneficial features, they typically cannot be very well calibrated to a given set of arbitrage-free European vanilla option prices. In particular, the pricing of short-maturity options in the equity market by the Heston model may often be unsatisfactory [37].

1.2.3. STOCHASTIC LOCAL VOLATILITY MODELS

As an alternative to the ‘pure’ LV model and SV models and with the purpose to combine features of both, the class of stochastic local volatility (SLV) models was developed around the year 2000 [75, 80, 82]. Since then, SLV models² have been given a lot of attention – e.g. [23, 75, 80, 93, 104, 117]. According to Lipton et al., SLV models are the *de facto* standard for pricing FX options in practice [81]. They combine the beneficial characteristics of the standard Local Volatility model – an almost perfect calibration to liquid European-type options – and well-established stochastic volatility models such as the Heston model, which often yield more realistic forward smiles and prices of exotic options.

²Formally, we should make a distinction between SLV models with a parametric local volatility component, such as the SABR model [63], and SLV model with a non-parametric local volatility component (which is expressed in terms of arbitrage-free European-type option prices) – in this section we elaborate on the latter type.

The general SLV model can be represented by

$$dS(t) = rS(t)dt + \sigma_{\text{SLV}}(t, S(t))\psi(S(t), V(t))S(t)dW_x(t), \quad S_0 := S(0), \quad (1.2.2)$$

$$dV(t) = a_v(t, V(t))dt + b_v(t, V(t))dW_v(t), \quad V_0 := V(0), \quad (1.2.3)$$

with $dW_x(t)dW_v(t) = \rho_{x,v}dt$ and $\psi(S(\cdot), V(\cdot))$ controlling the stochastic volatility. The drift and diffusion of the variance process are determined by $a_v(\cdot, V(\cdot))$ and $b_v(\cdot, V(\cdot))$, respectively.

Note that the general SLV model described by the system of the SDEs in (1.2.2) and (1.2.3) can collapse to either the pure SV model or to the LV model. If we set the local volatility component $\sigma_{\text{SLV}}(t, S(t)) = 1$, then the model boils down to a *pure* stochastic volatility model. On the other hand, if the stochastic component of the variance $b_v(t, V(t))$ is equal to 0, the model reduces to a local volatility model. Two popular stochastic volatility models which fit into our framework are the Heston SV model [66], with the variance process driven by the CIR dynamics [26], $\psi(V(t)) = \sqrt{V(t)}$ with $a_v(t, V(t)) = \kappa(\bar{v} - V(t))$ and $b_v(t, V(t)) = \gamma\sqrt{V(t)}$, and the Schöbel-Zhu model [108] with $\psi(V(t)) = V(t)$ and $a_v(t, V(t)) = \kappa(\bar{v} - V(t))$, $b_v(t, V(t)) = \gamma$. Parameter κ controls the speed of mean-reversion, \bar{v} controls a long-term mean and γ determines the volatility of the process $V(t)$.

An essential part of SLV models involves a particular conditional expectation, which is present in the term $\sigma_{\text{SLV}}(\cdot, S(\cdot))$:

$$\sigma_{\text{SLV}}^2(t, K) = \frac{\sigma_{\text{LV}}^2(t, K)}{\mathbb{E}[\psi^2(S(t), V(t)) | S(t) = K]}, \quad (1.2.4)$$

where $\sigma_{\text{LV}}^2(\cdot, \cdot)$ is ‘Dupire’s local volatility’, see (1.2.1). The conditional expectation is non-trivial, as the joint distribution of $S(\cdot)$ and $V(\cdot)$ in its denominator is unknown. Also, there is no direct link with the market quotes. Nonetheless, several approaches have been developed to evaluate SLV models, amongst others solving a Kolmogorov forward PDE forward one step at a time [23, 32, 104], recovering simultaneously the conditional expectation and the complete stochastic local volatility component. In this iterative procedure the joint density of $S(\cdot)$ and $V(\cdot)$ is solved for all time-points. PDE-discretization techniques are common practice in the financial industry in a hybrid local volatility context. As an alternative to the standard ADI methods, Lipton et al. [81], in a Quadratic Local Stochastic Volatility (QLSV) framework, introduce a Galerkin-Ritz inspired method for solving a system of PDEs and demonstrate that it is efficient. Another approach to handle the problem of computational burden is presented in [117], who employ GPUs to accelerate the computations.

The *Markovian projection* technique has also been applied in an SLV context [64, 98]. Although this method is generally applicable, it involves several conditional expectations that typically need to be approximated. Moreover, the technique does not preserve marginal distributions of order higher than one. This may result in a significant mismatch in prices of contracts depending on stock values at multiple times, such as American and barrier options, implied by the original and projected models.

Other attempts for solving the SLV model are presented in [115], where a Levenberg-Marquardt optimization technique for a non-linear Fokker-Planck equation is applied

and in [32], where zero correlation is assumed between the volatility process and the underlying asset, yielding an efficient simulation of the extended Schöbel-Zhu model [108]. In a more general framework, based on the theory of generalized Wiener functionals, see e.g. [124], An and Li [4] provide closed-form expansions for evaluating a general conditional expectation that involves marginal distributions which are generated by stochastic differential equations. In [86] for a general class of stochastic local volatility models a family of asymptotic expansions for European-style option prices and implied volatilities is derived. Further, in [93] the authors derive an asymptotic expansion for forward-starting options in a multi-factor local-stochastic volatility model, which results in explicit approximation formulas for the forward implied volatility. Recent developments on stochastic local volatility models have been made by e.g. [27, 28, 86, 123, 129]

1.2.4. THE COLLOCATING LOCAL VOLATILITY MODEL

The original idea of *stochastic collocation* is to project uncertainty onto a probability space with known properties and conditions [9, 131]. Collocation methods have been addressed in various disciplines for uncertainty quantification, see e.g. [45, 127]. Generally speaking, in collocation methods the purpose is to satisfy governing differential equations at a discrete set of points, in the corresponding probability space. Two of the main approaches of high-order stochastic collocation methods are the Lagrange interpolation approach, see e.g. [131], and the pseudo-spectral generalized polynomial chaos approach from e.g. [130].

Recently, the stochastic collocation method was applied in a financial context [57, 58], where in a sampling setting a particular stochastic variable of interest Y , of which computing samples $y_n = F_Y^{-1}(u_n)$ (with u_n drawn from $U \stackrel{d}{=} \mathcal{U}([0, 1])$) is an expensive exercise, was approximated by a function of a more convenient ‘cheap to evaluate’ random variable X . The approximation relies on the fact that the CDFs of Y and X are equal in distribution, i.e.

$$F_Y(Y) \stackrel{d}{=} U \stackrel{d}{=} F_X(X), \text{ where } U \stackrel{d}{=} \mathcal{U}([0, 1]),$$

which implies for a certain function $g(\cdot)$:

$$F_Y(g(x_n)) = F_X(x_n), \quad y_n = g(x_n),$$

where x_n and y_n denote samples of their corresponding distributions. Once an approximation for $g(\cdot)$ has been determined – we can establish an N th order Lagrange polynomial $g_N(\cdot)$ interpolating through the ‘expensive’ collocation values $y_i = F_Y^{-1}(F_X(x_i))$ – sampling from the ‘expensive’ distribution Y can be performed via sampling from the ‘cheap’ distribution X .

A model class that can serve as an alternative to the more well-established volatility models discussed in Sections 1.2.1-1.2.3, is the novel Collocating Local Volatility (CLV) model, introduced in [54]. The CLV model can be represented as follows³:

$$\begin{aligned} S(t) &= g_N(t, X(t)), \\ dX(t) &= \mu(t, X(t))dt + \sigma(t, X(t))dW(t), \quad X_0 := X(0), \end{aligned}$$

³For notation purposes, the general dynamics of a 1-dimensional kernel process are given – the dynamics of an n -dimensional kernel process are presented in Chapter 4.

where $g_N(\cdot, X(\cdot))$ is a local volatility function and $X(t)$ denotes a kernel process. The kernel process can be any stochastic process and is able to control the forward volatility smile implied by the CLV model. By an appropriate choice of the kernel process and its parameter values, the CLV model can accurately price exotic options. At the same time, a perfect calibration to arbitrage-free European-type option prices is preserved by the local volatility function $g_N(\cdot, X(\cdot))$, which is constructed by means of the stochastic collocation described above.

1.3. OUTLINE OF THE THESIS

This thesis is organized as follows. In Chapter 2 we discuss SLV models in general, and in particular the Heston-SLV model. We develop a ‘non-parametric method’ that admits a rapid Monte Carlo evaluation. Essentially, the non-parametric method consists of assigning every (s, v) -realization in a Monte Carlo simulation to an appropriate ‘bin’, resulting in an accurate approximation of the non-trivial conditional expectation in (1.2.4), discussed in Section 1.2.3.

Subsequently, in Chapter 3 we address the problem of calibrating the time-dependent FX-SABR model by determining ‘mappings’ between time-dependent parameters and their constant parameter ‘equivalents’, the so-called ‘effective parameters’. We use real FX market data to assess the performance of the time-dependent FX-SABR model for pricing FX barrier options, and compare them against the prices implied by the Local Volatility model, the constant parameter FX-SABR model and the time-dependent FX-SABR model enhanced by a non-parametric local volatility component.

We study hybrid local volatility models in Chapter 4, in particular the Local Volatility model enhanced by stochastic interest rates and the SLV models SABR-LV and Heston-SLV, which are the SABR and Heston models to which a local volatility component is added. Based on stochastic collocation, in combination with standard regression techniques, we establish approximations for the relevant non-trivial (conditional) expectations.

Next, in Chapter 5 we discuss the novel Collocating Local Volatility model. We consider three different kernel processes, the Ornstein-Uhlenbeck and Cox-Ingersoll-Ros processes and the Heston model. For these kernel processes we study the effect of the kernel parameters on the forward volatility smile. In addition, we calibrate the kernel processes to FX barrier option prices.

Lastly, Chapter 6 concludes and we give an outlook for future research.

2

THE HESTON STOCHASTIC-LOCAL VOLATILITY MODEL: EFFICIENT MONTE CARLO SIMULATION

In this chapter we propose an efficient Monte Carlo scheme for simulating the stochastic volatility model of Heston [66] enhanced by a non-parametric local volatility component. This hybrid model combines the main advantages of the Heston model and the local volatility model introduced by Dupire [35] and Derman & Kani [34]. In particular, the additional local volatility component acts as a “compensator” that bridges the mismatch between a non-perfectly calibrated Heston model and the market quotes for European-type options. By means of numerical experiments we show that our scheme enables a consistent and fast pricing of products that are sensitive to the forward volatility skew. Detailed error analysis is also provided.

Keywords: Heston Stochastic-Local Volatility, HSLV, Stochastic Volatility, Local Volatility, Heston, Hybrid Models, Calibration, Monte Carlo.

2.1. INTRODUCTION

In this chapter we consider a hybrid model which includes stochastic as well as local volatility. We focus on the Heston stochastic volatility model enhanced by a non-parametric local volatility component. Such a model, by construction, allows a high-quality calibration to plain vanilla options, even for an initial set of Heston parameters which is not very well calibrated to market data.

The evaluation of these stochastic-local volatility (SLV) models is however not trivial. As the stock’s overall volatility consists of two different types of volatilities (the stochastic and the local) it is challenging to account, in the calibration process, for the correlation

This chapter is based on the article ‘The Heston Stochastic-Local Volatility Model: Efficient Monte Carlo Simulation’, published in *International Journal of Theoretical and Applied Finance*, 17(7):1450045, 2014 [118].

between these two.

Although the SLV hybrid models are rather new in the financial industry, a number of attempts for efficient model evaluation have been made already. Ren *et al.* [104] proposed a stochastic volatility model driven by a lognormal volatility process and developed a tailor-made algorithm for solving the corresponding Kolmogorov forward PDE. An extension of this technique to the Heston SLV was presented in [37] where a finite volume scheme for the model evaluation was used. Clark [23] discusses SLV models in an FX context. He mentions that for solving the forward Kolmogorov equation, one can e.g. use explicit finite differencing or ADI timestepping. Although the PDE-discretization techniques are common practice in the financial industry in the context of the local-volatility component, explicit finite differences methods are typically only stable for a very large number of time-grid points requiring significant computational burden. In [129] a Modified Craig-Sneyd (MCS) scheme was used in the context of stochastic local volatility models, which is unconditionally stable for two-dimensional convection-diffusion equations with a mixed derivative term [69, 70]. Tian *et al.* engaged a parallel GPU platform to accelerate these computations [116].

The authors in [99] moved away from the direct solution of the SLV model and derived via the Markovian projection closed-form approximations to prices of European options on various underlyings. Work on Markovian projections in the context of the SLV models has also been presented in [64], where a so-called “effective local volatility” was derived. The Markovian projections can be widely applied but require a number of conditional expectations to be determined. Very often these expectations are not available analytically and brute-force assumptions need to be imposed so that approximations can be defined [98]. Although mathematically appealing the Markovian projection technique preserves only marginal densities and does not keep marginal distributions of orders higher than one intact. Due to this, prices of financial securities depending on stock values at multiple times, such as American options and barriers, may significantly differ between the original model and the projected model.

Another attempt for solving the SLV model was presented in [115], where a Levenberg-Marquardt optimization technique for a non-linear Fokker-Planck equation was applied. Another approach for simulation was proposed in [32] by Deelstra and Rayée. By assuming zero correlation between the volatility process and the underlying asset, it is possible to efficiently simulate the extended Schöbel-Zhu model.

We present a Monte Carlo approach for efficient simulation of the Heston SLV model. In particular, we develop a non-parametric numerical scheme for efficient model evaluation. The scheme is model independent and can be applied to all SLV hybrid models, including those based on the SABR model. The technique introduced does not require any advanced methods which makes it intuitive and easy to implement. A similar idea was presented in [60, 77], based on kernel estimators in an interacting particle system.

The outline of this chapter is as follows. In Section 2.2 we derive the full-scale SLV model and highlight the issues related to efficient model evaluation. Section 2.3 constitutes the core of this chapter. We show there how, for a Monte Carlo simulation scheme, *nontrivial* conditional expectations can be evaluated efficiently. We also discuss the simulation of the full-scale model and present how the “unbiased” Monte Carlo scheme for the Heston model [5] can be adopted to the Heston SLV (HSLV) model. In Section

2.4 some numerical examples are presented. We particularly concentrate on forward-volatilities implied by the Heston SLV model. Section 2.5 focuses on the theoretical assessment of the model error and Section 2.6 concludes.

2.2. STOCHASTIC-LOCAL VOLATILITY MODEL

The stochastic-local volatility (SLV) model under consideration is driven by the following system of Stochastic Differential Equations (SDEs):

$$dS(t)/S(t) = r dt + \sigma(t, S(t))\psi(V(t))dW_x(t), \quad (2.2.1)$$

$$dV(t) = a_v(t, V(t))dt + b_v(t, V(t))dW_v(t), \quad (2.2.2)$$

$$dW_x(t)dW_v(t) = \rho_{x,v}dt, \quad (2.2.3)$$

with correlation $\rho_{x,v}$ between the corresponding Brownian motions, $\sigma(t, S(t))$ is the local volatility component, $\psi(V(t))$ controls the stochastic volatility, parameters $a_v(t, V(t))$ and $b_v(t, V(t))$ determine the drift and diffusion of the variance process, respectively, and r is a constant interest rate.

The SLV model described by equations (2.2.1) and (2.2.2) is not completely determined as $\sigma(t, S(t))$ is left unspecified. This function can take different forms. It can be, for example, given by the constant elasticity of variance model, i.e. $\sigma(t, S(t)) = \hat{\sigma}S^\beta(t)$, which is a well-known parametric form for describing the volatility movements in terms of the underlying asset $S(t)$. Choosing a parametric form for the local volatility, $\sigma(t, S(t))$, although very flexible and well-accepted, has an undesired feature which is the need for model calibration, i.e. one needs to determine the SV parameters and the LV parameters in the calibration procedure. As the calibration may not always guarantee a sufficiently-well fit to market data, we concentrate on non-parametric forms for $\sigma(t, S(t))$ here.

The main concept for deriving a non-parametric LV component $\sigma(t, S(t))$ is as follows: it is well-known that from market data for the European-style options one can determine the market implied density¹, $\hat{f}_S(x)$, of the stock $S(T)$. Furthermore, by deriving the Kolmogorov forward equation for the underlying model, we are able to determine the density, $f_S(x)$, of the stock driven by the SDEs (2.2.1) and (2.2.2). In a general setting these densities differ and only for a *perfectly* calibrated model they are identical. As in the SLV framework we have one free parameter available, namely $\sigma(t, S(t))$, we may choose the local component so that the densities implied from the market and the model are equal. In the following we derive an expression for the local volatility component $\sigma(t, S(t))$ in the stochastic-local volatility model. Although the main result of the derivations (equation (2.2.9)) can be found in the literature (e.g. [104]), we include these to emphasize the role of the local volatility component as a “compensator”, which is explicitly defined in terms of market prices.

2.2.1. SPECIFYING $\sigma(t, S(t))$

Let us start with a European call option whose price is given by:

$$C(t_0, t, S(t_0), K) = \frac{M(t_0)}{M(t)} \mathbb{E}[(S(t) - K)^+ | \mathcal{F}(t_0)],$$

¹This result is shown in Lemma 2.2.1

where the expectation is evaluated under the risk-neutral measure \mathbb{Q} and the money-savings account $M(t)$ is given by $dM(t) = rM(t)dt$ (with constant interest rate r and $M(t_0) = 1$). In the following derivations, we leave filtration $\mathcal{F}(t_0)$ out in the notation and we introduce the short-hand notation $C(t, K) := C(t_0, t, S(t_0), K)$.

In order to obtain the dynamics of the call option price, we apply Itô's lemma:

$$\begin{aligned} dC(t, K) &= \left(d\frac{1}{M(t)} \right) \mathbb{E}[(S(t) - K)^+] + \frac{1}{M(t)} d\mathbb{E}[(S(t) - K)^+] \\ &= -\frac{r}{M(t)} \mathbb{E}[(S(t) - K)^+] dt + \frac{1}{M(t)} \mathbb{E}[d(S(t) - K)^+], \end{aligned} \quad (2.2.4)$$

where Fubini's theorem justifies the equality $d\mathbb{E}[(S(t) - K)^+] = \mathbb{E}[d(S(t) - K)^+]$. Regarding the right-hand side in (2.2.4), we cannot apply Itô's lemma for the evaluation of $d(S(t) - K)^+$, as the convex function $h(x) = (x - a)^+$ is not differentiable at point $x = a$. Therefore, we will make use of a generalized version of Itô's lemma, known as the Tanaka-Meyer formula [78, 100]:

Theorem 2.2.1 (Tanaka-Meyer formula). *Given a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, $t_0 \leq t < \infty$, let $X(t) = X(t_0) + \widetilde{M}(t) + \widetilde{N}(t)$ be a continuous semimartingale, where $\widetilde{M} = \{\widetilde{M}(t), \mathcal{F}(t)\}$ is a continuous local martingale², $\widetilde{N} = \{\widetilde{N}(t), \mathcal{F}(t)\}$ is a càdlàg adapted process³ of locally bounded variation. Then, for $h(x) = (x - a)^+ := \max(x - a, 0)$ with $a \in \mathbb{R}$:*

$$h(X(t)) = h(X(t_0)) + \int_{t_0}^t \mathbb{1}_{X(u) > a} d\widetilde{M}(u) + \int_{t_0}^t \mathbb{1}_{X(u) > a} dV(u) + \frac{1}{2} \int_{t_0}^t h''(X(u)) (d\widetilde{M}(u))^2.$$

Proof. A full proof can be found in Tanaka [114]. □

Applying the Tanaka-Meyer formula⁴, we get

$$(S(t) - K)^+ = (S(t_0) - K)^+ + \int_{t_0}^t \mathbb{1}_{S(u) > K} dS(u) + \frac{1}{2} \int_{t_0}^t \delta(S(u) - K) (dS(u))^2,$$

which in a differential form is given by:

$$d(S(t) - K)^+ = \mathbb{1}_{S(t) > K} dS(t) + \frac{1}{2} \delta(S(t) - K) (dS(t))^2,$$

where $\delta(\cdot)$ is defined as the Dirac delta function, which could be characterized as:

$$\delta(x) = \begin{cases} +\infty & x = 0, \\ 0 & x \neq 0. \end{cases}$$

² \widetilde{M} is a local martingale provided that there is a nondecreasing sequence $\{\tau_k\}$ of stopping times with the property that $\mathbb{P}(\tau_k \rightarrow \infty \text{ as } k \rightarrow \infty) = 1$ and such that for each k the stopped process $\widetilde{M}(t)^{(k)} = \widetilde{M}(t \wedge \tau_k) - \widetilde{M}(t_0)$ is a martingale.

³ $\widetilde{N}(t)$ is defined on the real numbers (or a subset of them) and is everywhere right-continuous and has left limits everywhere.

⁴ By taking $X(t) := S(t)$ we immediately notice that $S(t)$ is a semimartingale, as $S(t) = S(t_0) + r \int_{t_0}^t S(u) du + \widetilde{M}(t)$, where $\int_{t_0}^t S(u) du$ is a càdlàg adapted process of locally bounded variation and $\widetilde{M}(t)$ is an \mathcal{H}^1 martingale and thus a local martingale as well (every martingale is a local martingale).

Substituting the dynamics of $S(t)$, we obtain:

$$\begin{aligned} d(S(t) - K)^+ &= \mathbb{1}_{S(t) > K} (rS(t)dt + \sigma(t, S(t))\psi(V(t))S(t)dW_x(t)) \\ &\quad + \frac{1}{2} \delta(S(t) - K) \sigma^2(t, S(t)) \psi^2(V(t)) S^2(t) dt. \end{aligned}$$

The dynamics of the call price can be written as:

$$\begin{aligned} dC(t, K) &= -\frac{r}{M(t)} \mathbb{E}[(S(t) - K)^+] dt \\ &\quad + \frac{1}{M(t)} \mathbb{E} \left[\mathbb{1}_{S(t) > K} \left(rS(t)dt + \sigma(t, S(t))\psi(V(t))S(t)dW_x(t) \right) \right] \\ &\quad + \frac{1}{2M(t)} \mathbb{E} \left[\delta(S(t) - K) \sigma^2(t, S(t)) \psi^2(V(t)) S^2(t) \right] dt. \end{aligned}$$

We simplify this equation by using the equality

$$\begin{aligned} \mathbb{E}[(S(t) - K)^+] &= \mathbb{E}[\mathbb{1}_{S(t) > K} (S(t) - K)] \\ &= \mathbb{E}[\mathbb{1}_{S(t) > K} S(t)] - K \mathbb{E}[\mathbb{1}_{S(t) > K}]. \end{aligned}$$

This gives us the following preliminary result:

Result 2.2.1. *The dynamics of the European call option price $C(t, K) := C(t_0, t, S(t_0), K)$ with $S(t)$ and $V(t)$ following the dynamics as given in (2.2.1) and (2.2.2), respectively, are given by*

$$dC(t, K) = \frac{rK}{M(t)} \mathbb{E}[\mathbb{1}_{S(t) > K}] dt + \frac{1}{2M(t)} \mathbb{E}[\delta(S(t) - K) \sigma^2(t, S(t)) \psi^2(V(t)) S^2(t)] dt,$$

where each expectation is conditional on $\mathcal{F}(t_0)$.

In the following, we use another result in our derivations:

Lemma 2.2.1. *The European call option price $C(t, K)$ with $S(t)$ and $V(t)$ following dynamics as given in (2.2.1) and (2.2.2), respectively, satisfies*

$$-\frac{\partial C(t, K)}{\partial K} = \frac{1}{M(t)} \mathbb{E}[\mathbb{1}_{S(t) > K} | \mathcal{F}(t_0)] \text{ and } \frac{\partial^2 C(t)}{\partial K^2} = \frac{f_S(K)}{M(t)},$$

where f_S is the marginal probability density function of $S(t)$.

Lemma 2.2.1 states a well-established result, see e.g. [46]. We return to the dynamics of the call price given in Result 2.2.1 where we include the results from Lemma 2.2.1, i.e.

$$dC(t, K) = -rK \frac{\partial C(t, K)}{\partial K} dt + \frac{1}{2M(t)} \mathbb{E} \left[\delta(S(t) - K) \sigma^2(t, S(t)) \psi^2(V(t)) S^2(t) \right] dt,$$

which is equivalent to:

$$2M(t) \left(dC(t, K) + rK \frac{\partial C(t, K)}{\partial K} dt \right) = \mathbb{E} \left[\delta(S(t) - K) \sigma^2(t, S(t)) \psi^2(V(t)) S^2(t) \right] dt =: A(t) dt.$$

We denote by $A(t)$,

$$\begin{aligned} A(t) &= \iint_{\mathbb{R}} \delta(s-K) \sigma^2(t,s) \psi^2(u) s^2 f_{V,S}(u,s) ds du \\ &= \int_{\mathbb{R}} \psi^2(u) \left(\int_{\mathbb{R}} \delta(s-K) s^2 \sigma^2(t,s) f_{V,S}(u,s) ds \right) du. \end{aligned} \quad (2.2.5)$$

Using properties of the Dirac delta function ⁵ the inner integral simplifies to:

$$\int_{\mathbb{R}} \delta(s-K) s^2 \sigma^2(t,s) f_{V,S}(u,s) ds = K^2 \sigma^2(t,K) f_{V,S}(u,K). \quad (2.2.6)$$

Then, the expression for $A(t)$ is given by

$$A(t) = K^2 \sigma^2(t,K) \int_{\mathbb{R}} \psi^2(u) f_{V,S}(u,K) du, \quad (2.2.7)$$

which is equivalent to:

$$A(t) = K^2 \sigma^2(t,K) f_S(K) \mathbb{E}[\psi^2(V(t)) | S(t) = K].$$

The dynamics are given by:

$$dC(t,K) = -rK \frac{\partial C(t,K)}{\partial K} dt + \frac{1}{2M(t)} K^2 \sigma^2(t,K) f_S(K) \mathbb{E}[\psi^2(V(t)) | S(t) = K] dt.$$

Using the second equation in Lemma 2.2.1, we obtain:

$$dC(t,K) = \left(-rK \frac{\partial C(t,K)}{\partial K} - \frac{1}{2} K^2 \sigma^2(t,K) \mathbb{E}[\psi^2(V(t)) | S(t) = K] \frac{\partial^2 C(t,K)}{\partial K^2} \right) dt, \quad (2.2.8)$$

which can be expressed as:

$$\sigma^2(t,K) \mathbb{E}[\psi^2(V(t)) | S(t) = K] = \frac{\frac{\partial C(t,K)}{\partial t} + rK \frac{\partial C(t,K)}{\partial K}}{\frac{1}{2} K^2 \frac{\partial^2 C(t,K)}{\partial K^2}} =: \sigma_{LV}^2(t,K),$$

where $\sigma_{LV}(t,K)$ denotes Dupire's local volatility [35]. We eventually find the following relation:

$$\sigma^2(t,K) = \frac{\sigma_{LV}^2(t,K)}{\mathbb{E}[\psi^2(V(t)) | S(t) = K]}. \quad (2.2.9)$$

The local volatility component $\sigma^2(t,K)$ consists of two ingredients: the deterministic local volatility $\sigma_{LV}(t,K)$ and the conditional expectation $\mathbb{E}[\psi^2(V(t)) | S(t) = K]$. Numerical evaluation of $\sigma_{LV}(t,K)$ is already well-established in the literature, see for example [6, 24, 30]. On the other hand, the efficient computation of the conditional expectation in (2.2.9) is not yet established. The difficulty lies in the fact that the joint distribution of the variance V and the stock S , $f_{V,S}$, is unknown. This is due to the fact that the stock process, $S(t)$, contains a local-volatility component $\sigma_{LV}(t,S)$ which is also not known analytically. The evaluation of the unknown expectation can be either derived by solving a Kolmogorov forward PDE (e.g. [32, 37]) or by applying a Markovian projection approximation [64, 98, 99]. In this chapter we concentrate on the Monte Carlo evaluation of the stochastic-local volatility model. In the next section we present a numerical method which leads to efficient Monte Carlo model evaluation.

⁵ $\int_{\mathbb{R}} \delta(t-T) f(t) dt = f(T)$

2.3. NOVEL TECHNIQUE FOR $\mathbb{E}[\psi^2(V(t))|S(t) = K]$

In this section we present a new efficient evaluation of a general stochastic-local volatility model. In particular, by an Euler discretization we simulate the SLV model (2.2.1)-(2.2.2), as follows:

$$s_{i+1,j} = s_{i,j} + r s_{i,j} \Delta_t + \sigma(t_i, s_{i,j}) s_{i,j} \psi(v_{i,j}) \sqrt{\Delta_t} Z_x, \quad s_{0,j} = S(t_0), \quad (2.3.1)$$

$$v_{i+1,j} = v_{i,j} + a_v(t_i, v_{i,j}) \Delta_t + b_v(t_i, v_{i,j}) \sqrt{\Delta_t} Z_v, \quad v_{0,j} = v(t_0), \quad (2.3.2)$$

for $j = 1, \dots, N$ and $i = 0, \dots, M$ where $Z_x = Z_1$, $Z_v = \rho_{x,v} Z_1 + (1 - \rho_{x,v}^2)^{1/2} Z_2$, with Z_1 and Z_2 two independent standard normal variables. Further, Δ_t is the equidistant time-step, given by $\Delta_t = T/M$, with M indicating the number of time steps and T stands for final time. N corresponds to the total number of Monte Carlo paths.

Using expression (2.2.9) for $\sigma(t, S)$, System (2.3.1)-(2.3.2) becomes:

$$s_{i+1,j} = s_{i,j} + r s_{i,j} \Delta_t + \sqrt{\frac{\sigma_{\text{LV}}^2(t_i, s_{i,j})}{\mathbb{E}[\psi^2(V(t_i))|S(t_i) = s_{i,j}]}} s_{i,j} \psi(v_{i,j}) \sqrt{\Delta_t} Z_x, \quad (2.3.3)$$

$$v_{i+1,j} = v_{i,j} + a_v(t_i, v_{i,j}) \Delta_t + b_v(t_i, v_{i,j}) \sqrt{\Delta_t} Z_v. \quad (2.3.4)$$

To determine the values of the paths for the *next* time-step, t_{i+1} , one needs to establish two main components, $\sigma_{\text{LV}}^2(t_i, s_{i,j})$ and $\mathbb{E}[\psi^2(V(t_i))|S(t_i) = s_{i,j}]$. As indicated, efficient evaluation of $\sigma_{\text{LV}}^2(t_i, s_{i,j})$ is already well-established in the literature [6, 30]. This is not the case for evaluation of the conditional expectation. The main difficulty in its evaluation is that the conditioning has to be performed on each individual stock realization $s_{i,j}$, i.e. as we simulate a discretized system for (S, V) , each realization of $s_{i,j}$ has exactly one corresponding realization of the variance $v_{i,j}$ and this makes the evaluation of the conditional expectation difficult.

In the next subsection we present a non-parametric method for evaluating the conditional expectation.

2.3.1. NON-PARAMETRIC METHOD

Suppose that for a given time t_i , $i = 1, \dots, M$ we have N pairs of Monte Carlo realizations $(s_{i,1}, v_{i,1})$, $(s_{i,2}, v_{i,2})$, \dots , $(s_{i,N}, v_{i,N})$ for which we wish to evaluate the conditional expectation in (2.3.3). As for each $s_{i,j}$ we have exactly one value $v_{i,j}$ the conditional expectation will always be equal to $\psi^2(v_{i,j})$, which is undesired. Such a problem is a natural consequence of discretization of the continuous system (S, V) . Obviously, in order to obtain an accurate estimate we would need to have an infinite set of paths, which is practically unfeasible.

The idea to overcome this problem is to *group* the pairs of realizations into bundles which would provide a more accurate estimate for the desired expectation. Let us divide the range of $S(t_i)$ into ℓ mutually exclusive bins $(b_1, b_2]$, $(b_2, b_3]$, \dots , $(b_\ell, b_{\ell+1}]$, with $b_1 \geq 0$ and $b_{\ell+1} < \infty$.

Now, for any particular stock realization $s_{i,j}$, for which $s_{i,j} \in (b_k, b_{k+1}]$ for some $k \in \{1, 2, \dots, \ell\}$, we introduce the following approximation:

$$\mathbb{E}[\psi^2(V(t_i))|S(t_i) = s_{i,j}] \approx \mathbb{E}[\psi^2(V(t_i))|S(t_i) \in (b_k, b_{k+1})]. \quad (2.3.5)$$

If we define the left and right boundaries of (b_k, b_{k+1}) to be $s_{i,j}-\varepsilon$ and $s_{i,j}+\varepsilon$, respectively, we obtain the following:

$$\begin{aligned} \mathbb{E}[\psi^2(V(t_i))|S(t_i) = s_{i,j}] &= \lim_{\varepsilon \rightarrow 0^+} \mathbb{E}[\psi^2(V(t_i))|S(t_i) \in (s_{i,j} - \varepsilon, s_{i,j} + \varepsilon)] \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{\mathbb{E}[\psi^2(V(t_i))\mathbb{1}_{S(t_i) \in (s_{i,j} - \varepsilon, s_{i,j} + \varepsilon)}]}{\mathbb{Q}[S(t_i) \in (s_{i,j} - \varepsilon, s_{i,j} + \varepsilon)]}. \end{aligned} \quad (2.3.6)$$

In the limiting case where both boundaries of the bin are equal to $s_{i,j}$ the approximation of the conditional expectation boils down to its exact value. This is an indication for the appropriateness of the approximation in (2.3.5). The open question that remains is how to choose *proper* bin boundaries b_k for $k = 1, \dots, \ell + 1$. We consider the two following choices in a Monte Carlo simulation framework.

We first order all the stock paths $s_{i,1}, s_{i,2}, \dots, s_{i,N}$ and obtain the following sequence: $\bar{s}_{i,1} \leq \bar{s}_{i,2} \leq \dots \leq \bar{s}_{i,N}$, where $\bar{s}_{i,1}$ and $\bar{s}_{i,N}$ are the minimal and maximal values at time-step i , respectively. Then, we choose the bin boundaries $b_{i,k}$, $k = 1, \dots, \ell + 1$. A straightforward way is specifying these such that the bins have the same size. We can also choose the boundaries depending on the number of paths per bin. These two choices are established as follows:

1. Define the bins with respect to *an equidistant grid* specified on the domain $\bar{s}_{i,1} = b_{i,1} < b_{i,2} < \dots < b_{i,\ell+1} = \bar{s}_{i,N}$ such that for any $u, v \in \{1, \dots, \ell\}$, $u \neq v$, $b_{i,u+1} - b_{i,u} = b_{i,v+1} - b_{i,v}$. This is established by:

$$b_{i,k} = \bar{s}_{i,1} + \frac{k-1}{\ell}(\bar{s}_{i,N} - \bar{s}_{i,1}), \quad k = 1 \dots \ell + 1. \quad (2.3.7)$$

2. Specify the bins so that each bin contains an *approximately equal number* of Monte Carlo paths:

$$b_{i,1} = \bar{s}_{i,1}, \quad b_{i,\ell+1} = \bar{s}_{i,N}, \quad b_{i,k} = \bar{s}_{i,(k-1)N/\ell}, \quad k = 2 \dots \ell. \quad (2.3.8)$$

After determination of the bins, each pair $(s_{i,j}, v_{i,j})$ is assigned to a bin according to its $s_{i,j}$ value. Let us denote the path numbers corresponding to the k th bin \mathcal{B}_k at time t_i by $\mathcal{J}_{i,k}$, that is $\mathcal{J}_{i,k} := \{j | (s_{i,j}, v_{i,j}) \in \mathcal{B}_k\}$. Further, N_k is defined as the number of paths in the k th bin, so $N_k = |\mathcal{J}_{i,k}|$. We then have:

$$\begin{aligned} \mathbb{E}[\psi^2(V(t_i))|S(t_i) = s_{i,j}] &\approx \frac{\mathbb{E}[\psi^2(V(t_i))\mathbb{1}_{S(t_i) \in (b_{i,k}, b_{i,k+1})}]}{\mathbb{Q}[S(t_i) \in (b_{i,k}, b_{i,k+1})]} \\ &\approx \frac{\frac{1}{N} \sum_{j=1}^N \psi^2(v_{i,j})\mathbb{1}_{s_{i,j} \in (b_{i,k}, b_{i,k+1})}}{\mathbb{Q}[S(t_i) \in (b_{i,k}, b_{i,k+1})]} \\ &= \frac{1}{N\alpha(k)} \sum_{j \in \mathcal{J}_{i,k}} \psi^2(v_{i,j}), \end{aligned} \quad (2.3.9)$$

where $\alpha(k) := \mathbb{Q}[S(t_i) \in (b_{i,k}, b_{i,k+1})]$ represents the probability of the stock being in the k th bin. The second approximation is established by switching between the expectation

and the average, which is based on a finite number of $(s_{i,j}, v_{i,j})$ -pairs. The value of $\alpha(k)$ depends on the way the bins are chosen:

$$\alpha(k) = \begin{cases} N_k/N, & k = 1 \dots \ell & \text{for bins defined as in (2.3.7),} \\ 1/\ell & & \text{for bins defined as in (2.3.8).} \end{cases}$$

Remark 2.3.1. *As it will be shown in Section 2.5.3, the choice of bins affects the convergence of the non-parametric method. If we define the bins according to (2.3.8), bins close to the mean of the joint density are much smaller than bins in the tails. This is desirable as the region close to the mean contains many more observations, requiring a higher accuracy and thus smaller bin sizes. We will choose bins according to (2.3.8) in numerical experiments.*

We summarize the non-parametric method in Algorithm 1.

```

for each time-step  $t_i, i = 1 \dots M$  do
1   Generate  $N$  pairs of observations  $(s_{i,j}, \psi^2(v_{i,j}))$ ,  $j = 1 \dots N$ .
2   Order the elements  $\bar{s}_{i,j}$ :  $\bar{s}_{i,1} \leq \bar{s}_{i,2} \leq \dots \leq \bar{s}_{i,N}$ .
3   Determine the boundaries of  $\ell$  bins  $(b_{i,k}, b_{i,k+1}]$ ,  $k = 1 \dots \ell$  according to either (2.3.7)
   or (2.3.8).
4   For the  $k$ th bin approximate the conditional expectation by
    $\mathbb{E}[\psi^2(V(t_i)) | S(t_i) \in (b_{i,k}, b_{i,k+1}]] \approx \frac{1}{N\alpha(k)} \sum_{j \in \mathcal{J}_{i,k}} \psi^2(v_{i,j})$ , where  $\mathcal{J}_{i,k}$  is the set of
   path numbers  $j$  for which the observations are in the  $k$ th bin at the  $i$ th time-step and
    $\alpha(k)$  represents the probability of the stock being in the  $k$ th bin, which is determined
   by the choice of bins.
end

```

Algorithm 1: Non-parametric method

We present two illustrative examples where this method for calculating the conditional expectation is applied. Error analysis will be discussed in Section 2.5.

Experiment 2.3.1 (Illustrative examples). In order to illustrate how the introduced algorithm works we present two experiments. First, we consider a simple Monte Carlo simulation consisting of 9 paths. In the second experiment we apply the algorithm to calculate the conditional expectation for the Heston model. We start with some assumptions. The initial values for the stock and variance process are $S(t_0) = s_0 = 1$ and $V(t_0) = v_0 = 0.1$, respectively.

In order to obtain path realizations at time t_1 , we need to determine the conditional expectation $\mathbb{E}[V(t_0)|S(t_0) = s_0]$, which trivially gives $\mathbb{E}[V(t_0)|S(t_0) = s_0] = v_0 = 0.1$. This holds for all paths in this experiment. To determine the paths at time t_2 we calculate the expectation $\mathbb{E}[V(t_1)|S(t_1) = s_{1,j}]$, $j = 1 \dots 9$, as follows. First, we choose the number of bins, $\ell = 3$, and sort the pairs $(s_{1,j}, v_{1,j})$ according to their $s_{1,j}$ values. Then, we assign each pair to a bin – also according to the $s_{1,j}$ values. Finally, we calculate for each bin an approximation of the conditional expectation. The procedure is illustrated in the tables

below:

j	t_0		t_1		j	t_1		
	(s_0, ν_0)	$\mathbb{E}[V(t_0)]$	$(s_{1,j}, \nu_{1,j})$	$(s_{1,j}, \nu_{1,j})$		$(s_{1,j}, \nu_{1,j})$	$(s_{1,j}, \nu_{1,j})$	
1	(1, 0.1)	0.1	(1.9, 0.09)	⇒	9	(0.4, 0.25)	} ⇒ $\mathbb{E}[V(t_1) S(t_1) \in (0, 0.9)]$ $\approx \frac{1}{3} \sum_j \nu_{1,j} = 0.2,$	
2	(1, 0.1)	0.1	(0.9, 0.15)		4	(0.5, 0.20)		
3	(1, 0.1)	0.1	(1.2, 0.15)		2	(0.9, 0.15)		
4	(1, 0.1)	0.1	(0.5, 0.20)		} ⇒ $\mathbb{E}[V(t_1) S(t_1) \in (0.9, 1.3)]$ $\approx \frac{1}{3} \sum_j \nu_{1,j} = 0.1,$	6	(1.1, 0.07)	
5	(1, 0.1)	0.1	(1.6, 0.06)			3	(1.2, 0.15)	
6	(1, 0.1)	0.1	(1.1, 0.07)		} ⇒ $\mathbb{E}[V(t_1) S(t_1) \in (1.3, 1.9)]$ $\approx \frac{1}{3} \sum_j \nu_{1,j} = 0.067.$	8	(1.3, 0.08)	
7	(1, 0.1)	0.1	(1.7, 0.05)	5		(1.6, 0.06)		
8	(1, 0.1)	0.1	(1.2, 0.08)	7	(1.7, 0.05)			
9	(1, 0.1)	0.1	(0.4, 0.25)	1	(1.9, 0.09)			

Let us consider a more practical example. We consider the Heston stochastic volatility model. In this model $\psi(x) = \sqrt{x}$ and the conditional expectation reads $\mathbb{E}[V(t_i)|S(t_i) = s_{i,j}]$. The reason for considering the *pure* Heston model is that we are able to determine the conditional expectation by the 2D-COS method [105]. A discussion of this calculation is provided in Section 2.5.2.

In Figure 2.3.1 we compare the results for the conditional expectation from the proposed scheme and the reference obtained by Fourier expansions. Each plot includes a contour plot of the recovered density, the corresponding conditional expectation, and its approximation. In the simulations we considered 10^5 Monte Carlo paths with 2, 5 and 20 bins, respectively. We choose bins that contain equal numbers of realizations. This yields smaller bins close to the mean of the joint density, see also Remark 2.3.1. The approximation obtained by the algorithm introduced converges to the reference.

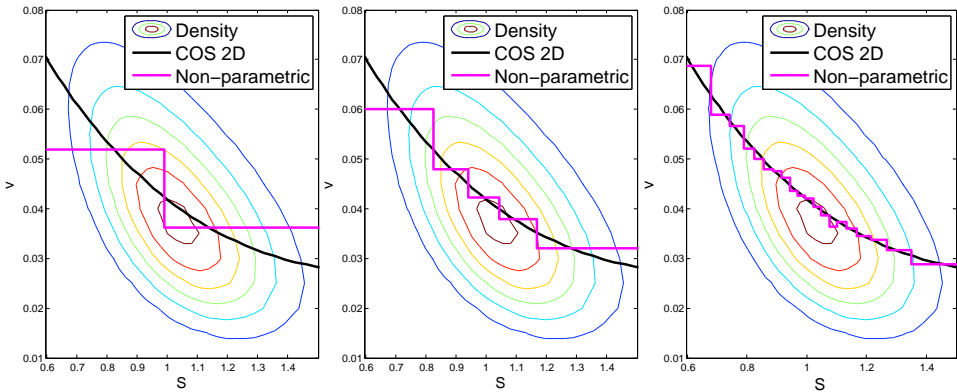


Figure 2.3.1: The approximation obtained by the non-parametric method converges to the conditional expectation recovered by the COS method as the number of bins increases (2, 5 and 20 respectively).

Remark 2.3.2. An alternative method for estimating $\mathbb{E}[\psi^2(V(t_i))|S(t_i) = s_{i,j}]$ is by projecting $\psi^2(V(t_i))$ on a set of orthogonal polynomials $\zeta_k(\cdot)$, $k = 1 \dots n$:

$\psi^2(V(t_i)) = \sum_{k=1}^n \beta_k \zeta_k(S(t_i)) + \epsilon$. By regressing $\psi^2(V(t_i))$ on functions $\zeta_k(\cdot)$ the conditional expectation can be calculated, as

$$\mathbb{E}[\psi^2(V(t_i))|S(t_i) = s_{i,j}] \approx \mathbb{E}\left[\sum_{k=1}^n \beta_k \zeta_k(S(t_i)) \middle| S(t_i) = s_{i,j}\right] = \sum_{k=1}^n \beta_k \zeta_k(s_{i,j}). \quad (2.3.10)$$

The above approximation for $\psi^2(V(t_i))$ is based on the assumption that the conditional expectation is an element of the L^2 -space of square integrable functions. The conditional expectation can be represented as a linear function of elements of a countable orthonormal basis. Applying the approximation in (2.3.10) the discrete scheme described by (2.3.3) and (2.3.4) becomes:

$$\begin{aligned} s_{i+1,j} &= s_{i,j} + r s_{i,j} \Delta_t + \sqrt{\frac{\sigma_{LV}^2(t_i, s_{i,j})}{\sum_{k=1}^n \beta_k \zeta_k(s_{i,j})}} s_{i,j} \psi(v_{i,j}) \sqrt{\Delta_t} Z_x, \\ v_{i+1,j} &= v_{i,j} + \kappa(\bar{v} - v_{i,j}) \Delta_t + \gamma \sqrt{v_{i,j}} \sqrt{\Delta_t} Z_v, \end{aligned} \quad (2.3.11)$$

where κ is the speed of mean-reversion, \bar{v} is the long-run variance and γ is the volatility of the variance process ('vol-vol'). Although intuitive and straightforward, the regression-based alternative possesses the drawback that the Feller condition must be satisfied to guarantee a positive conditional expectation for the whole range of arguments. We show a numerical test in Figure 2.3.2, where we consider a simple quadratic polynomial: $\zeta_1(x) = 1$, $\zeta_2(x) = x$ and $\zeta_3(x) = x^2$.

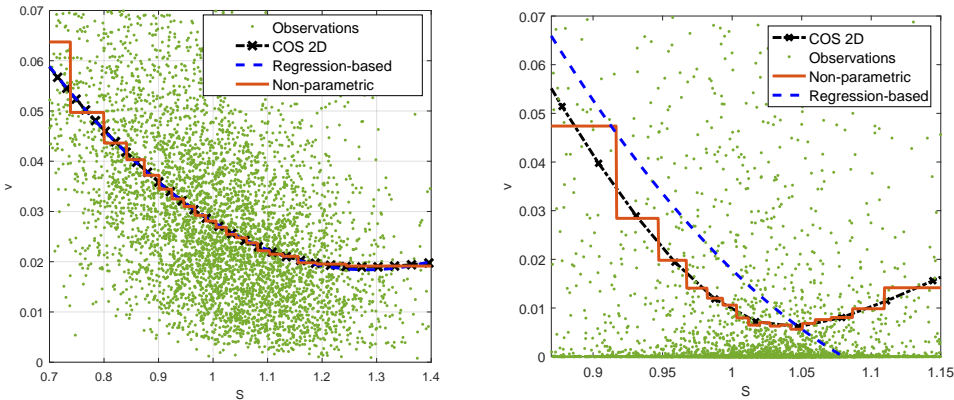


Figure 2.3.2: The regression-based alternative: Feller satisfied (left) and not satisfied (right). Feller must be satisfied in order to guarantee a non-negative approximation of the conditional expectation. The non-parametric method does not suffer from this restriction, as we see on the right.

Since in practice the Feller condition is often violated, regression-based methods require additional tuning like including high-order polynomials or constraining of regression coefficients. As such model improvements need to be done on case-by-case basis we consider the non-parametric approach as preferable and use it throughout this chapter.

2.3.2. CONTINUOUS APPROXIMATION

As the conditional expectation is continuously differentiable, we prefer its approximation to satisfy this property too. Furthermore, at the right-hand side of Figure 2.3.2 we observe that at the left boundary of the strike range the fit of the non-parametric approximation to the reference may be improved. In order to obtain a continuous approximation that establishes this, we consider a linearization of the estimated expectation obtained by the non-parametric method. This can be done by connecting the mid-points of the approximations of the non-parametric method, see Figure 2.3.3. The mid-point approximation is continuous, but not necessarily continuously differentiable.

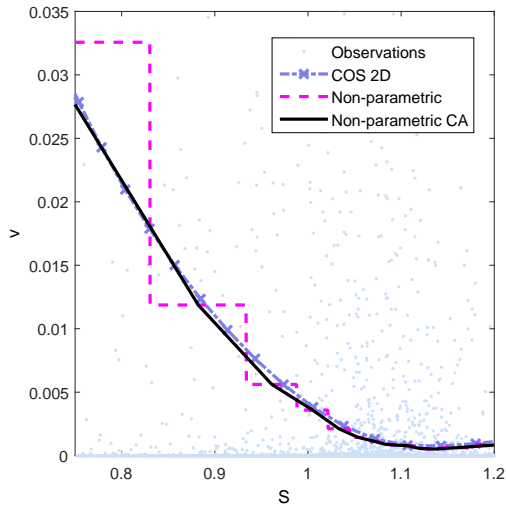


Figure 2.3.3: The continuous approximation ('CA') gives a better fit to the theoretical conditional expectation, which is recovered by the COS method.

In the following we refer to the continuous approximation as 'non-parametric method'.

Remark 2.3.3 (Generalization to early-exercise options). *The approximation of the expectation in the definition of the stochastic-local volatility model by means of the continuous approximation based on the bins does not depend on any specific option type or payoff. In this respect it can be combined with European, forward starting options (as presented in this chapter), but for example also with early-exercise options where it can be combined with the Least Squares Method [83] or Stochastic Grid Bundling Method [73].*

2.3.3. EFFICIENT SIMULATION SCHEME

The CIR-type process used for the dynamics of the variance in the Heston model does not allow for negative realizations. Unfortunately, when applying the basic Euler discretization scheme the variance process can become negative with non-zero probability. Although several fixes like "absorption at zero" for handling negative variance realizations are known in the literature (see [85] for an overview), these improved methods are typically not free of bias.

In this section we adopt the Quadratic Exponential (QE) scheme introduced by Andersen [5] and apply it for simulating the Heston SLV model. The main difference for Monte Carlo simulation between the pure Heston and the Heston SLV models lies in the fact that the variance of the latter is not only driven by the stochastic volatility, but also by the local volatility component, which is state-dependent. This requires an additional “freezing approximation”, which is not present in the derivation of the original QE scheme. Numerical experiments show that the additional approximation still yields an accurate simulation scheme.

We start by recalling the dynamics of the Heston SLV model, expressed in terms of independent Brownian motions:

$$\begin{aligned} dS(t)/S(t) &= rdt + \sigma(t, S(t))\sqrt{V(t)}\left(\rho_{x,v}d\widetilde{W}_v(t) + \sqrt{1 - \rho_{x,v}^2}d\widetilde{W}_x(t)\right), \\ dV(t) &= \kappa(\bar{V} - V(t))dt + \gamma\sqrt{V(t)}d\widetilde{W}_v(t), \end{aligned}$$

where $\rho_{x,v}$ denotes correlation between the $S(t)$ and $V(t)$ processes. The discretization of $X(t) := \log(S(t))$ (“log-stock”), with $\hat{\sigma}(t, X(t)) := \sigma(t, e^{X(t)})$, reads:

$$\begin{aligned} X(t + \Delta_t) &= X(t) + \int_t^{t+\Delta_t} \left(r - \frac{1}{2}\hat{\sigma}^2(s, X(s))V(s) \right) ds \\ &+ \rho_{x,v} \int_t^{t+\Delta_t} \hat{\sigma}(s, X(s))\sqrt{V(s)}d\widetilde{W}_v(s) + \sqrt{1 - \rho_{x,v}^2} \int_t^{t+\Delta_t} \hat{\sigma}(s, X(s))\sqrt{V(s)}d\widetilde{W}_x(s). \end{aligned} \quad (2.3.12)$$

The variance process $V(t + \Delta_t)$ follows a *scaled* non-central chi-squared distribution, i.e.

$$V(t + \Delta_t) \sim c(\Delta_t)\chi^2(d, \lambda(t, V(t))),$$

with

$$c(\Delta_t) = \frac{\gamma^2}{4\kappa}(1 - e^{-\kappa\Delta_t}), \quad d = \frac{4\kappa\bar{V}}{\gamma^2}, \quad \lambda(t, V(t)) = \frac{4\kappa e^{-\kappa\Delta_t}}{\gamma^2(1 - e^{-\kappa\Delta_t})}V(t), \quad (2.3.13)$$

and $\chi^2(d, \lambda(t, V(t)))$ representing a noncentral chi-squared distribution with d degrees of freedom and non-centrality parameter $\lambda(t, V(t))$. Furthermore, by integrating the variance process, we find:

$$\int_t^{t+\Delta_t} \sqrt{V(s)}d\widetilde{W}_v(s) = \frac{1}{\gamma} \left(V(t + \Delta_t) - V(t) - \kappa\bar{V}\Delta_t + \kappa \int_t^{t+\Delta_t} V(s)ds \right). \quad (2.3.14)$$

In the last integral in (2.3.12) the local and stochastic volatilities are coupled. This complicates the simulation as we are not able to directly use the integrated variance from (2.3.14). As any Monte Carlo simulation involving a local-volatility component requires many time-steps, we perform local-freezing of $\hat{\sigma}(s, X(s))$ in (2.3.12), i.e.

$$\int_t^{t+\Delta_t} \hat{\sigma}(s, X(s))\sqrt{V(s)}d\widetilde{W}_v(s) \approx \hat{\sigma}(t, X(t)) \int_t^{t+\Delta_t} \sqrt{V(s)}d\widetilde{W}_v(s). \quad (2.3.15)$$

Due to the approximation in (2.3.15), we can use (2.3.14) in (2.3.12):

$$\begin{aligned}
 X(t + \Delta_t) \approx & X(t) + \int_t^{t+\Delta_t} \left(r - \frac{1}{2} \hat{\sigma}^2(s, X(s)) V(s) \right) ds \\
 & + \frac{\rho_{x,v} \hat{\sigma}(t, X(t))}{\gamma} \left(V(t + \Delta_t) - V(t) - \kappa \bar{V} \Delta_t + \kappa \int_t^{t+\Delta_t} V(s) ds \right) \\
 & + \sqrt{1 - \rho_{x,v}^2} \int_t^{t+\Delta_t} \hat{\sigma}(s, X(s)) \sqrt{V(s)} d\tilde{W}_x(s).
 \end{aligned} \tag{2.3.16}$$

In the Euler discretization all integrals w.r.t. time can be approximated by $\int_a^b f(x) dx \approx (b - a) f(a)$. The discretized process for $X(t)$ then reads:

$$\begin{aligned}
 X(t + \Delta_t) \approx & X(t) + r \Delta_t - \frac{1}{2} \hat{\sigma}^2(t, X(t)) V(t) \Delta_t \\
 & + \frac{1}{\gamma} \rho_{x,v} \hat{\sigma}(t, X(t)) (V(t + \Delta_t) - V(t) - \kappa \bar{V} \Delta_t + \kappa V(t) \Delta_t) \\
 & + \sqrt{1 - \rho_{x,v}^2} \int_t^{t+\Delta_t} \hat{\sigma}(s, X(s)) \sqrt{V(s)} d\tilde{W}_x(s).
 \end{aligned}$$

Furthermore, by the Itô isometry we have

$$\int_t^{t+\Delta_t} \hat{\sigma}(s, X(s)) \sqrt{V(s)} d\tilde{W}_x(s) \sim \tilde{Z}_x \sqrt{\int_t^{t+\Delta_t} \hat{\sigma}^2(s, X(s)) V(s) ds}, \tag{2.3.17}$$

where $\tilde{Z}_x \sim N(0, 1)$. The integral at the right-side of (2.3.17) can also be approximated by the Euler discretization, i.e. $\int_t^{t+\Delta_t} \hat{\sigma}^2(s, X(s)) V(s) ds \approx \hat{\sigma}^2(t, X(t)) V(t) \Delta_t$. With this, the discretization scheme becomes

$$\begin{aligned}
 v_{i+1,j} & \sim c(\Delta_t) \chi^2(d, \lambda(t_i, v_{i,j})), \\
 x_{i+1,j} & = x_{i,j} + r \Delta_t - \frac{1}{2} \hat{\sigma}^2(t_i, x_{i,j}) v_{i,j} \Delta_t + \frac{\rho_{x,v}}{\gamma} \hat{\sigma}(t_i, x_{i,j}) (v_{i+1,j} - \kappa \bar{v} \Delta_t + v_{i,j} c_1) \\
 & \quad + \rho_1 \sqrt{\hat{\sigma}^2(t_i, x_{i,j}) v_{i,j} \Delta_t} \tilde{Z}_x,
 \end{aligned}$$

with $\rho_1 = (1 - \rho_{x,v}^2)^{1/2}$, $c_1 = \kappa \Delta_t - 1$, where $c(\Delta_t)$, d and $\lambda(t, V(t))$ are defined in (2.3.13) and

$$\hat{\sigma}^2(t_i, x_{i,j}) \stackrel{\text{def}}{=} \sigma^2(t_i, e^{x_{i,j}}) = \frac{\sigma_{\text{LV}}^2(t_i, s_{i,j})}{\mathbb{E}[V(t_i) | S(t_i) = s_{i,j}]}. \tag{2.3.18}$$

In (2.3.18) we compute Dupire's local volatility,

$$\sigma_{\text{LV}}^2(t_i, s_{i,j}) = \left. \frac{\frac{\partial C(t, s)}{\partial t} + r s \frac{\partial C(t, s)}{\partial s}}{\frac{1}{2} s^2 \frac{\partial^2 C(t, s)}{\partial s^2}} \right|_{s=s_{i,j}, t=t_i},$$

by using the following finite difference approximations:

$$\left. \frac{\partial C(t, s_{i,j})}{\partial t} \right|_{t=t_i} \approx \frac{C(t_i + h_1, s_{i,j}) - C(t_i, s_{i,j})}{h_1}, \quad \left. \frac{\partial C(t_i, s)}{\partial s} \right|_{s=s_{i,j}} \approx \frac{C(t_i, s_{i,j} + h_2) - C(t_i, s_{i,j})}{h_2},$$

and

$$\left. \frac{\partial^2 C(t_i, s)}{\partial s^2} \right|_{s=s_{i,j}} \approx \frac{C(t_i, s_{i,j} + h_2) - 2C(t_i, s_{i,j}) + C(t_i, s_{i,j} - h_2)}{h_2^2}. \quad (2.3.19)$$

For stability reasons the derivatives are often expressed in terms of the implied volatilities [32]. As in practice a continuum of European call prices in time-to-maturity and strike is not available, some interpolation may be required. Detailed discussion on this is provided in the literature (e.g. [6]).

Numerical comparisons between the Euler and the original QE scheme have been provided in the literature too [5]. We perform an experiment with our version of the Monte Carlo scheme in the follow-up section, because our scheme is slightly different due to the local volatility component which requires an additional approximation (local freezing of the state-dependent local volatility).

Experiment 2.3.2 (Efficient simulation scheme). We compare the scheme we propose with the basic Euler discretization scheme. We consider parameter values based on Case III of Andersen [5], i.e. for $T = 5$ we consider the Heston SLV model with $\kappa = 1.05$, $\gamma = 0.95$, $\bar{v} = 0.0855$, $\nu_0 = 0.0945$, $\rho_{x,v} = -0.315$ and $r = 0$. We perform a Monte Carlo simulation consisting of 20 seeds, $5 \cdot 10^4$ paths per seed. The number of bins is set to 20.

For different time-step sizes Δ_t and strike prices K we calculate the absolute error in the implied volatilities $|\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{SLV}}|$, where $\bar{\sigma}_{\text{market}}$ and $\bar{\sigma}_{\text{SLV}}$ denote volatilities implied by the market and by the HSLV model, respectively. As in [5], we generate synthetic *reference implied volatility* values given a pure Heston model, applying a Fourier-based pricing technique. Analogously, we calculate by means of finite difference approximations the derivatives of European call prices involved in Dupire's local volatility component.

Results are presented in Table 2.1. The efficient ('low-bias') simulation scheme outperforms the Euler scheme: it gives a higher accuracy and we observe faster convergence to the reference for a decaying time-step size.

		Error (%): $ \bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{SLV}} $					
		70%		100%		150%	
K	Δ_t	Euler	Low-bias	Euler	Low-bias	Euler	Low-bias
1		6.05 (0.08)	5.91 (0.11)	6.06 (0.09)	5.65 (0.11)	5.23 (0.17)	4.67 (0.16)
1/2		3.81 (0.12)	1.53 (0.14)	4.12 (0.12)	1.37 (0.17)	3.51 (0.19)	0.86 (0.25)
1/4		2.70 (0.12)	0.64 (0.14)	3.01 (0.13)	0.55 (0.14)	2.66 (0.20)	0.31 (0.24)
1/8		1.71 (0.13)	0.31 (0.19)	1.92 (0.15)	0.25 (0.20)	1.74 (0.23)	0.13 (0.25)
1/16		0.98 (0.16)	0.22 (0.17)	1.08 (0.17)	0.19 (0.19)	1.04 (0.22)	0.13 (0.27)
1/32		0.41 (0.26)	0.15 (0.18)	0.45 (0.30)	0.12 (0.18)	0.37 (0.42)	0.07 (0.23)

Table 2.1: Average error $|\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{SLV}}|$ from Monte Carlo simulations of the HSLV model with the Euler and efficient ('low-bias') schemes using 20 random seeds, for multiple time-step sizes Δ_t and strikes K . Numbers in parentheses are standard deviations over the seeds.

2.4. NUMERICAL RESULTS

Using the simulation scheme introduced in the previous section, we perform some Monte Carlo experiments for the pricing of European call and forward-starting options with the

Heston SLV model. The results are compared against the Heston and the standard local volatility models.

In this experiment we investigate the performance of the Heston SLV model with respect to the quality of a pre-calibrated Heston model. In the case the Heston model is well calibrated, we expect a limited contribution of the local volatility component. On the other hand, if the underlying Heston model is not sufficiently well calibrated, the local volatility contribution should be more pronounced. We then expect the quality of the fit to be more sensitive to the estimation of the conditional expectation discussed in the previous sections. In the simulations we thus distinguish two cases: a case in which the Heston model is well calibrated and one in which it is insufficiently well calibrated. Each of these variants of the Heston model is used in the Heston SLV model.

The simulation of the European-style options is performed for maturities (in years) $T = \{0.5, 2, 5, 8, 10\}$, while the pricing of the forward starting options will be done for the following pairs: $\{T_1, T_2\} = \{2, 4\}$ and $\{T_1, T_2\} = \{6, 8\}$. Our Monte Carlo simulation is performed with $5 \cdot 10^5$ paths and 100 time-steps per year. The number of bins is set to 20. In Section 2.5.3 we show that the accuracy of the non-parametric method is already satisfactory for a smaller number of bins.

2.4.1. EUROPEAN CALL OPTIONS

In this experiment we consider some real-life examples. As described the input for the Heston SLV model is a calibrated Heston model. It is therefore important to check how well the Heston SLV model performs depending on the quality of the pre-calibrated Heston model. In this section we consider two scenarios where the Heston model is well and insufficiently calibrated to market data.

In the first experiment we consider the Heston SLV model with the well-calibrated underlying Heston model. For times to maturity 2 and 8, years we display the results in Figure 2.4.1. As the mismatch between the pure Heston model and the market is small, the contribution of the local-volatility component $\sigma_{LV}(t, S(t))$ is limited in the first test.

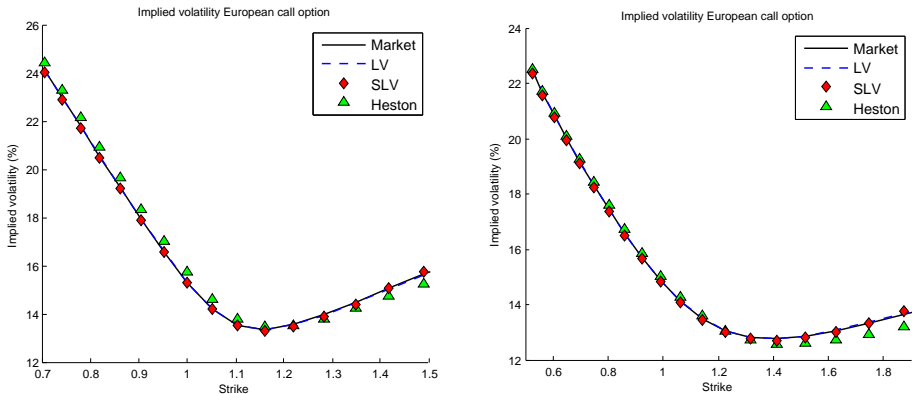


Figure 2.4.1: Implied volatility European call option. $T = 2$ (left), $T = 8$ (right), well calibrated Heston model.

In Figure 2.4.2 we display results of the second experiment, in which the Heston

model is insufficiently calibrated. We observe that in this case the local volatility term can compensate for the large gap between the market and the Heston model.

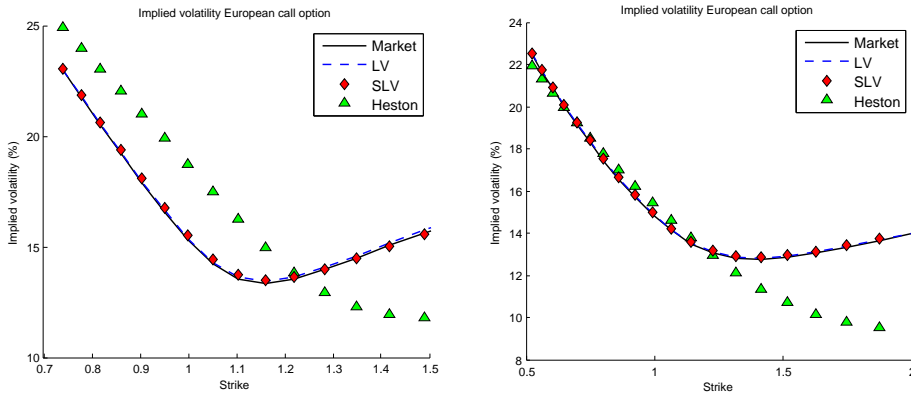


Figure 2.4.2: Implied volatility European call option. $T = 2$ (left), $T = 8$ (right), insufficiently calibrated Heston model.

For times to maturity 0.5, 2, 5, 8 and 10 years, results are given in Table 2.2 where we display the market implied volatility $\bar{\sigma}_{\text{market}}$ and the error in implied volatilities $\epsilon_{\text{model}} := \bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{model}}$, where $\bar{\sigma}_{\text{model}}$ denotes the volatility implied by a particular model. We observe a good fit of the LV model as well as the Heston SLV model to the quotes. The fit of the Heston SLV model can be explained by the mimicking theorem of Gyöngy [62], which states that given a general Itô process, a Markov process containing a local volatility component with the same marginal distributions exists.

For illustration purposes, in Appendix 2.A we show the pricing errors for Case II, $T = 2$ obtained by using 20, 30 and 40 bins. The error for close-to-ATM strikes decreases when 20, 30 and 40 are chosen, respectively.

2.4.2. FORWARD STARTING OPTIONS

With the pricing of European call options we see that the local volatility term in the Heston SLV model acts a compensator that bridges the gap between the market and calibrated Heston SV prices – even in the case of an unsatisfactory calibration. In this experiment we price forward starting options that start at time T_1 years and mature at time T_2 . As prices of forward starting options are not observable in the market, we discriminate between the LV, SLV and the calibrated Heston models. We first consider the case with $T_1 = 2$, $T_2 = 4$ and a well calibrated Heston model, see the plot on the left side in Figure 2.4.3. We observe that the volatility implied by the LV model is much flatter than the volatilities implied by the Heston SV and SLV models. Although it has approximately the same value at the lower and upper bounds of the strike range, this does not hold in the ATM region. As the Heston model is almost perfectly calibrated, it is no surprise that the Heston and SLV implied volatilities are almost identical.

For the case where the Heston model is insufficiently calibrated, we observe that the SLV model provides a forward smile that is located between the ones implied by the He-

European call options – error in implied volatilities [%]							
T	Strike	$\tilde{\sigma}_{\text{market}}$	ϵ_{LV}	Case I		Case II	
				ϵ_{SLV}	ϵ_{H}	ϵ_{SLV}	ϵ_{H}
0.5y	0.81	25.17	0.05	0.06	-0.32	-0.04	-1.87
	0.90	21.75	-0.03	0.01	-0.47	-0.15	-3.03
	1.00	17.70	-0.11	-0.06	-0.63	-0.32	-4.45
	1.11	15.26	-0.12	-0.01	-0.28	-0.24	-3.71
	1.24	16.00	-0.10	-0.08	0.16	0.21	0.18
2y	0.73	23.38	0.00	0.08	-0.33	-0.04	-1.82
	0.81	20.82	-0.02	0.05	-0.39	-0.09	-2.46
	1.00	15.28	-0.01	0.02	-0.45	-0.20	-3.44
	1.24	13.65	0.01	0.10	0.08	-0.06	0.07
	1.53	15.94	0.03	-0.18	0.42	0.06	4.15
5y	0.60	22.70	0.02	0.12	-0.16	-0.04	-0.39
	0.75	19.23	0.01	0.09	-0.22	-0.09	-0.99
	1.00	14.69	0.01	0.07	-0.24	-0.13	-1.45
	1.32	12.79	0.02	0.09	0.14	-0.12	0.97
	1.75	14.41	0.06	-0.09	0.46	-0.06	4.44
8y	0.52	22.46	-0.01	0.08	-0.06	-0.05	0.49
	0.70	19.16	-0.01	0.06	-0.12	-0.08	-0.09
	0.99	14.88	-0.01	0.05	-0.15	-0.10	-0.55
	1.41	12.76	-0.01	0.05	0.17	-0.09	1.38
	1.87	13.66	-0.02	-0.09	0.45	-0.06	4.14
10y	0.48	22.40	0.03	0.08	-0.02	-0.04	0.87
	0.66	19.22	0.02	0.06	-0.08	-0.06	0.30
	0.98	15.09	0.01	0.04	-0.12	-0.07	-0.18
	1.46	12.82	0.01	0.04	0.18	-0.06	1.46
	2.17	13.83	-0.02	-0.20	0.50	0.00	4.62

Table 2.2: Errors in implied volatilities for the local volatility (ϵ_{LV}), Heston SLV (ϵ_{SLV}) and the pure Heston (ϵ_{H}) models for a well (Case I) and insufficiently (Case II) calibrated Heston model for multiple times to maturity and strikes.

ston and the LV models. One may consider the results by the SLV model to represent somehow “advanced interpolation” between the Heston and the local volatility models, i.e. the SLV model can be considered as a non-linear combination of the LV and the SV models.

The results we obtain are in line with those of Engelmann et al. [37], who observe that the forward SLV implied volatilities do not become flat as for the local volatility model and preserve a shape very similar to the Heston model. For $T_1 = 6$ and $T_2 = 8$ similar results are observed, see Figure 2.4.4.

2.4.3. CALCULATION TIME

Considering the speed of the non-parametric method, we calculate the time it takes to evaluate the conditional expectation for a given number of paths or bins. First, we investigate the relation between calculation time and the number of bins. We fix the number of paths at 10^5 . Table 2.3 shows that the calculation time behaves linearly in the number of bins. This also holds for the relation between the calculation time and the number of paths (the number of bins is fixed at 20), see Table 2.4.

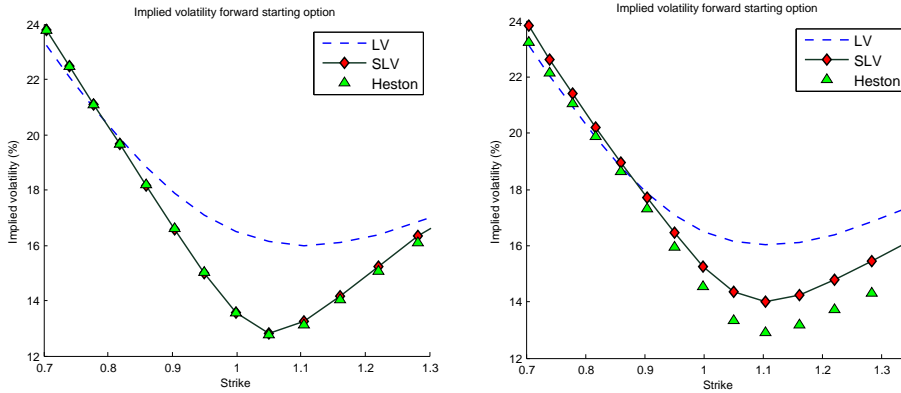


Figure 2.4.3: Implied volatility forward starting option. $T_1 = 2$, $T_2 = 4$, well calibrated (left) and insufficiently well calibrated (right) Heston model.

In our numerical experiments we choose 20 bins and 10^5 paths, which implies that the calculation of the conditional expectation by the non-parametric method takes less than $2.5 \cdot 10^{-2}$ seconds per time step⁶. As an indication regarding the total CPU time, we can solve the problem for $T = 1$ with $5 \cdot 10^4$ paths, 100 time-steps a year, and 5 bins in approximately 14 seconds (in Matlab on a i5-2400 CPU @3.10GHz, 4GB). In Section 2.5.3 we consider the dependence of the accuracy and the number of bins in more detail for the non-parametric method. A result of this analysis is that the accuracy of the method is quite insensitive to the number of bins.

#Bins	10	20	30	40	50	60	70	80	90	100
Time [ms]	15.6	23.7	32.9	41.6	49.9	59.0	67.9	75.9	84.5	93.9

Table 2.3: Timing results for different numbers of **bins** (number of paths fixed at 10^5).

#Paths [10^5]	1	2	3	4	5	6	7	8	9	10
Time [ms]	28.0	51.1	77.9	105.0	131.5	158.9	185.9	214.1	244.3	269.0

Table 2.4: Timing results for different numbers of **paths** (number of bins fixed at 20).

2.5. ERROR ANALYSIS

In Section 2.3 we tested the performance of the non-parametric method (and of the regression-based alternative) for a pure Heston SV model. We considered this model, as

⁶Theoretically, both experiments should result in exactly the same calculation time, i.e., the second time in Table 2.3 should be equal to the first time in Table 2.4. This is not the case due to Monte Carlo noise.

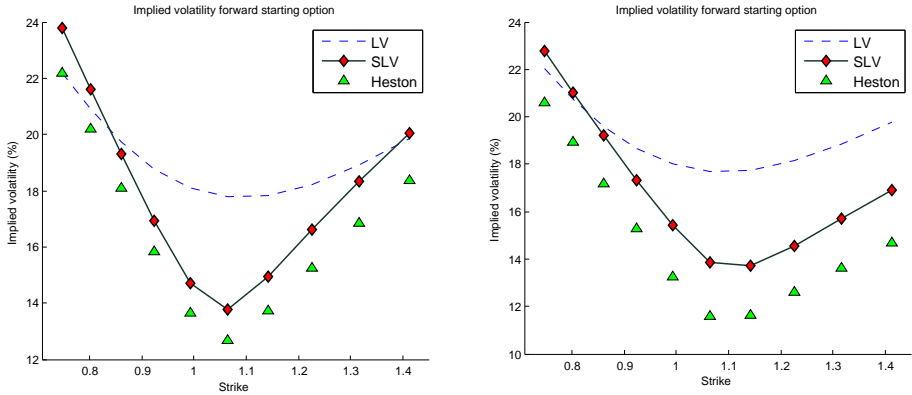


Figure 2.4.4: Implied volatility forward starting option. $T_1 = 6$, $T_2 = 8$, well calibrated (left) and poorly calibrated (right) Heston model.

the 2D-COS method [105] provided us with an accurate approximation of the conditional expectation. In particular, we compared an approximation from the non-parametric method to the recovered conditional expectation. However, our main interest lies in the Heston SLV model. In the Monte Carlo pricing of European call options and forward starting options a bias is introduced, which is due to three error sources. The first error originates from approximations in the calculation of Dupire’s local volatility term (2.3.18). In particular, we use finite differences for the three derivatives in (2.3.19). Further, the discretization of the continuous dynamics to the efficient simulation scheme introduced a discretization error. Last, at each time step we approximate $\mathbb{E}[V(t)|S(t) = s]$ by means of the non-parametric method. These three sources of error generate an error $e = C - \tilde{C}$, where C is the call price from the “original” Heston SLV model and \tilde{C} is the price obtained by the discrete Heston SLV model.

Ignoring the bias originating from the finite differences and the discretization (since these errors are well understood), the price mismatch is driven by the difference in conditional expectations $\|g - \hat{g}\|$, based on the governing PDEs, as in [59]. Here $g(s) := \mathbb{E}[V(t)|S(t) = s]$ denotes the conditional expectation (in the Heston SLV model) and \hat{g} is its piecewise linear continuous approximation we obtain by the non-parametric method. We now provide a pricing error bound that is implied by the mismatch between g and \hat{g} .

2.5.1. BOUND ON PRICING ERROR

In this section we turn to classical PDE error analysis, to make some statements about the approximation errors encountered.

By non-arbitrage assumptions, one can derive the HSLV PDE, which defines the value of a European-style option:

$$\begin{aligned}
0 &= \frac{\partial C}{\partial t} + rs \frac{\partial C}{\partial s} + \kappa(\bar{V} - V) \frac{\partial C}{\partial V} + \frac{1}{2} V s^2 \frac{\sigma_{LV}^2(t, s)}{g(s)} \frac{\partial^2 C}{\partial s^2} + \frac{1}{2} \gamma^2 V \frac{\partial^2 C}{\partial V^2} \\
&\quad + \rho_{x, v} \gamma V s \sqrt{\frac{\sigma_{LV}^2(t, s)}{g(s)}} \frac{\partial^2 C}{\partial s \partial V} - rC,
\end{aligned} \tag{2.5.1}$$

with $t \in [0, T)$, $g(s) = \mathbb{E}[V(t)|S(t) = s]$ and spatial coordinates $\{s, V\} \in [0, +\infty) \times [0, +\infty)$. Of course, we will solve the discrete version of the PDE on a finite domain.

Since the expectation in (2.5.1) is not known analytically we can estimate it by means of Monte Carlo simulation. The resulting, approximating pricing PDE then reads:

$$\begin{aligned}
0 &= \frac{\partial \tilde{C}}{\partial t} + rs \frac{\partial \tilde{C}}{\partial s} + \kappa(\bar{V} - V) \frac{\partial \tilde{C}}{\partial V} + \frac{1}{2} V s^2 \frac{\sigma_{LV}^2(t, s)}{\hat{g}(s)} \frac{\partial^2 \tilde{C}}{\partial s^2} + \frac{1}{2} \gamma^2 V \frac{\partial^2 \tilde{C}}{\partial V^2} \\
&\quad + \rho_{x, v} \gamma V s \sqrt{\frac{\sigma_{LV}^2(t, s)}{\hat{g}(s)}} \frac{\partial^2 \tilde{C}}{\partial s \partial V} - r\tilde{C},
\end{aligned} \tag{2.5.2}$$

with the pre-calibrated function $\hat{g}(s)$, as described in Section 2.3.1. The PDEs in (2.5.1) and (2.5.2) can be written, in shorthand notation, as follows:

$$\frac{\partial C}{\partial t} + L_1 C = 0, \quad \frac{\partial \tilde{C}}{\partial t} + L_2 \tilde{C} = 0, \tag{2.5.3}$$

with the corresponding operators L_1 , as in (2.5.1) and L_2 , as in (2.5.2). Again, C is the solution from the full-scale HSLV PDE, whereas \tilde{C} is the solution from the approximating PDE with the estimated function $\hat{g}(t)$. Both PDEs are accompanied by the same boundary and final conditions. For the error, $e := C - \tilde{C}$, we find:

$$\frac{\partial e}{\partial t} + L_1 C - L_2 \tilde{C} = 0, \tag{2.5.4}$$

which can be re-written as:

$$\frac{\partial e}{\partial t} + L_1 C - (L_1 \tilde{C} + (L_2 - L_1) \tilde{C}) = 0, \tag{2.5.5}$$

and we arrive at the following equation:

$$\frac{\partial e}{\partial t} + L_1 e = (L_2 - L_1) \tilde{C}, \tag{2.5.6}$$

subject to homogeneous boundary and final conditions. Notice that the right-hand side of the equation serves as a source term.

Based on the form in (2.5.6), multiplying both sides by e , and integration over domain Ω , gives us:

$$\int_{\Omega} e \frac{\partial e}{\partial t} d\Omega + \int_{\Omega} e L_1 e d\Omega = \int_{\Omega} e (L_2 - L_1) \tilde{C} d\Omega. \tag{2.5.7}$$

Integration by parts, as follows,

$$\int_{\Omega} e \frac{\partial e}{\partial t} d\Omega = \frac{1}{2} \frac{d}{dt} \int_{\Omega} e^2 d\Omega = \frac{1}{2} \frac{d}{dt} \|e\|_{L^2(\Omega)}^2, \tag{2.5.8}$$

inserted in equation (2.5.7), results in:

$$\frac{1}{2} \frac{d}{dt} \|e\|_{L^2(\Omega)}^2 = - \int_{\Omega} e L_1 e d\Omega + \int_{\Omega} e (L_2 - L_1) \tilde{C} d\Omega. \quad (2.5.9)$$

Applying classical PDE theory, in particular the Lax-Friedrich inequality and Grönwall's lemma (for more details, see [59, 89]), gives:

$$\begin{aligned} \|e\|_{L^2(\Omega)} &\leq \int_0^t \|(L_2 - L_1) \tilde{C}\|_{L^2(\Omega)} e^{\alpha(s-t)} ds \\ &\leq \frac{1}{\alpha} (1 - e^{-\alpha t}) \sup_{s \in (0,t)} \|(L_2 - L_1) \tilde{C}\|_{L^2(\Omega)} \\ &\leq \frac{1}{\alpha} \sup_{s \in (0,t)} \|(L_2 - L_1) \tilde{C}\|_{L^2(\Omega)}, \end{aligned}$$

where α is some positive constant, related to the V -ellipticity of the form $\int_{\Omega} e L_1 e d\Omega$ [89].

We use the notation $U := (L_2 - L_1) \tilde{C}$ and find the following operator:

$$U = \frac{1}{2} V \sigma_{LV}^2(t, s) s^2 \left[\frac{\hat{g}(s) - g(s)}{\hat{g}(s)g(s)} \right] \frac{\partial^2 \tilde{C}}{\partial s^2} + \rho_{x,v} \gamma V s \sigma_{LV}(t, s) \left[\frac{\sqrt{\hat{g}(s)} - \sqrt{g(s)}}{\sqrt{\hat{g}(s)g(s)}} \right] \frac{\partial^2 \tilde{C}}{\partial s \partial V}.$$

Assessing the appropriate norm yields:

$$\begin{aligned} \|U\|_{L^2(\Omega)} &= \left\| \frac{1}{2} V \sigma_{LV}^2(t, s) s^2 \left[\frac{\hat{g}(s) - g(s)}{\hat{g}(s)g(s)} \right] \frac{\partial^2 \tilde{C}}{\partial s^2} \right. \\ &\quad \left. + \rho_{x,v} \gamma V s \sigma_{LV}(t, s) \left[\frac{\sqrt{\hat{g}(s)} - \sqrt{g(s)}}{\sqrt{\hat{g}(s)g(s)}} \right] \frac{\partial^2 \tilde{C}}{\partial s \partial V} \right\|_{L^2(\Omega)}, \end{aligned}$$

which can be bounded by:

$$\begin{aligned} \|U\|_{L^2(\Omega)} &\leq \frac{1}{2} s_{\max}^2 V |\sigma_{LV}^2(t, s)| \left| \frac{\hat{g}(s) - g(s)}{\hat{g}(s)g(s)} \right| \left\| \frac{\partial^2 \tilde{C}}{\partial s^2} \right\|_{L^2(\Omega)} \\ &\quad + |\rho_{x,v} \gamma s_{\max} V | \sigma_{LV}(t, s) \left| \frac{\sqrt{\hat{g}(s)} - \sqrt{g(s)}}{\sqrt{\hat{g}(s)g(s)}} \right| \left\| \frac{\partial^2 \tilde{C}}{\partial s \partial V} \right\|_{L^2(\Omega)}. \end{aligned}$$

Then, we have:

$$\begin{aligned} \|e\|_{L^2(\Omega)} &\leq \frac{1}{\alpha} \sup_{s \in (0,t)} \left(\frac{1}{2} s_{\max}^2 V |\sigma_{LV}^2(t, s)| \left| \frac{\hat{g}(s) - g(s)}{\hat{g}(s)g(s)} \right| \left\| \frac{\partial^2 \tilde{C}}{\partial s^2} \right\|_{L^2(\Omega)} \right. \\ &\quad \left. + |\rho_{x,v} \gamma s_{\max} V | \sigma_{LV}(t, s) \left| \frac{\sqrt{\hat{g}(s)} - \sqrt{g(s)}}{\sqrt{\hat{g}(s)g(s)}} \right| \left\| \frac{\partial^2 \tilde{C}}{\partial s \partial V} \right\|_{L^2(\Omega)} \right). \quad (2.5.10) \end{aligned}$$

This latter inequality bound gives a representation of the parameters and functions that have an impact on the error made when solving for \tilde{C} , as an approximation for C . As both $\frac{\partial^2 \tilde{C}}{\partial s \partial V}$ and $\frac{\partial^2 \tilde{C}}{\partial s^2}$ are small for large V -values, the error is governed by the difference

between the conditional expectation g from the HSLV model and its approximation \hat{g} , which is obtained from the non-parametric method, as discussed in Section 2.3.1.

Regarding the difference in conditional expectations g and \hat{g} , it is clear that the accuracy of the approximation \hat{g} improves if the number of paths and bins simultaneously approach infinity, i.e.

$$\lim_{N \rightarrow \infty, \ell \rightarrow \infty} \|g - \hat{g}\| = 0,$$

where N and ℓ denote the number of paths and bins, respectively. In the following, we quantify the performance of the non-parametric method by considering the pure Heston SV model.

2.5.2. PERFORMANCE

To assess the performance of the non-parametric method, we need to determine a *highly accurate* reference value. As pointed out, it is difficult to find this conditional expectation explicitly. We can analyze the performance of approximating the conditional expectation for the case of the *pure* Heston model, i.e. the case where $\sigma_{LV}(t, s) = 1$.

We make use of the COS method [38, 105]. This introduces a well-understood error between the recovered and the theoretical conditional expectation for the Heston SV model. For a more detailed discussion on this, see [38].

Let $g_H(s)$ be the conditional expectation in the Heston SV model, obtained by the 2D-COS method. We determine an approximation \hat{g}_H by means of the non-parametric method. To measure the performance of the non-parametric method, we are interested in the mismatch in conditional expectations $\|g_H - \hat{g}_H\|$. In the L^2 -norm, the mismatch from the Monte Carlo simulation can be written as:

$$\sum_{i=1}^M \|g_H - \hat{g}_H\|_{L^2(\Omega)}^2 = \sum_{i=1}^M \sum_{k=1}^{\ell} \int_{\mathcal{B}_{i,k}} (g_H(s) - \hat{g}_H(s))^2 ds, \quad (2.5.11)$$

where Ω is the s domain and $\mathcal{B}_{i,k}$ denotes the k th bin at the i th time-step, $k = 1 \dots \ell$, $i = 1 \dots M$. Note that g_H is a smooth function, whereas \hat{g}_H is piecewise linear.

We now specify the error $\int_{\mathcal{B}_{i,k}} (g_H(s) - \hat{g}_H(s))^2 ds$ for one particular bin and time-step, which we state in a lemma.

Lemma 2.5.1. *For an arbitrary bin \mathcal{B} with boundaries $[b_l, b_r]$, the error between g_H and \hat{g}_H has size*

$$\begin{aligned} \|g_H - \hat{g}_H\|_{L^2(\mathcal{B})}^2 &= c_1^2 \Delta s + \frac{1}{12} \left(c_2^2 - 2c_1 g_H^{(2)}(s_m) \right) \Delta s^3 \\ &\quad + \frac{1}{240} \left(2(g_H^{(2)}(s_m))^2 - c_1 g_H^{(4)}(s_m) - c_2 g_H^{(3)}(s_m) \right) \Delta s^5 + \mathcal{O}(\Delta s^7), \end{aligned}$$

where s_m denotes the midpoint of $[b_l, b_r]$ and $\Delta s := b_r - b_l$, $\Delta \hat{g}_H := \hat{g}_H(b_r) - \hat{g}_H(b_l)$, $c_1 := \frac{1}{2} \left((g_H(b_l) - \hat{g}_H(b_l)) + (g_H(b_r) - \hat{g}_H(b_r)) \right)$ and $c_2 := -g_H^{(1)}(s_m) + \frac{\Delta \hat{g}_H}{\Delta s}$.

Proof. For a proof, see Appendix 2.B. □

2.5.3. NUMERICAL EXPERIMENT: CHOICE OF BINS

We now discuss the performance of the non-parametric method with respect to the choice of bins. In particular, we consider the error $|g_H(K) - \hat{g}_H(K)|$ for $K = 40\%$, $K = 100\%$ and $K = 160\%$. The bins are either chosen with respect to an equidistant grid – see (2.3.7) – or are equally weighted as in (2.3.8). Parameter values are $\gamma = 0.2$, $\kappa = 0.2$, $r = 0$, $\rho_{x,v} = -0.6$, $S_0 = 1$, $v_0 = 0.04$, $\bar{v} = 0.04$ and we consider the error at particular time $t = 2$. We choose the number of bins between 1 and 20. Our Monte Carlo simulation is performed with 10^6 paths. Results are displayed in Figure 2.5.1.

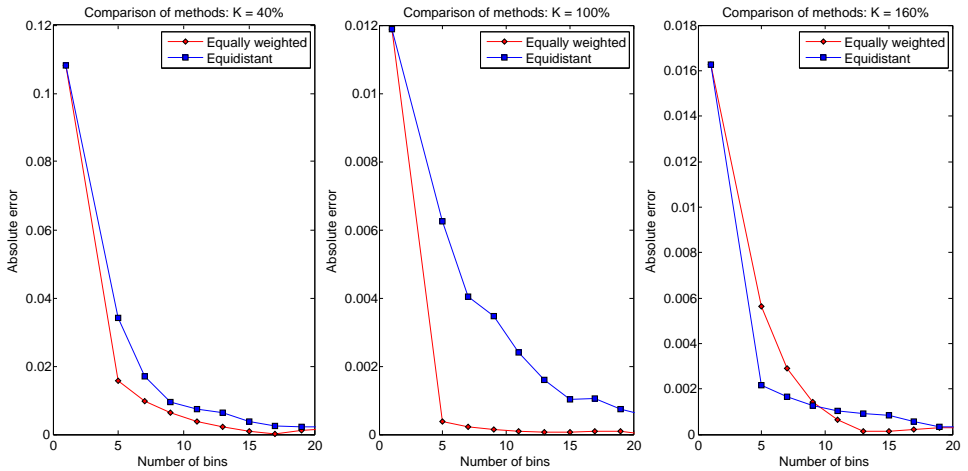


Figure 2.5.1: Convergence of the non-parametric method for two choices of bins.

For deep in-the-money and out-the-money strikes the choice of bins does not affect the performance of the non-parametric method. However, at $K = 100\%$ choosing equally weighted bins yields a faster convergence to the reference $g_H(100)$ than when we choose bins equidistantly. This is due to the natural weighting of the bins defined by (2.3.8), which provides highest accuracy in the ATM region. Further, for all strikes we note that in order to have a high-quality estimate of the conditional expectation, it is not required to use a large number of bins.

In Figure 2.5.1 we observe highly satisfactory convergence up to 15 bins. With a further increase of the number of bins convergence may stagnate since there is an insufficient number of paths in each bin. In such a case the number of Monte Carlo paths may need to be increased and we should improve the interpolation (continuous approximation, see Section 2.3.2).

2.6. CONCLUSION

In this chapter we have presented a new Monte Carlo scheme for the efficient evaluation of a general Stochastic-Local Volatility model. We have considered the Heston Stochastic-Local Volatility model in numerical experiments. For evaluating this model we have approximated a non-trivial conditional expectation in a non-parametric way,

which is intuitive and easy to implement. This approximation is embedded in a simulation scheme that is strongly based on the QE scheme of Andersen [5] and introduces less bias than more common Euler schemes. By means of numerical experiments and an error analysis we have shown that European-style options can accurately be priced by our method. Furthermore, it enables a consistent and fast pricing of products that are sensitive to the forward volatility smile.

APPENDIX

2.A. ADDITIONAL PRICING EXPERIMENTS

In this appendix, we price some more European call options ($T = 2$) with the Heston SLV model using 20, 30 and 40 bins, where the Heston model is insufficiently calibrated (Case II in Table 2.2). Results are given in Table 2.A.1. The error for close-to-ATM strikes decreases when 20, 30 and 40 bins are chosen, respectively.

$\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{SLV}} [\%], \text{ Case II}, T = 2$			
Strike	20 bins	30 bins	40 bins
0.73	-0.04	-0.01	-0.02
0.81	-0.09	-0.03	-0.02
1.00	-0.2	-0.07	-0.05
1.24	-0.06	-0.03	0.00
1.53	0.06	0.02	0.08

Table 2.A.1: SLV pricing errors for Case II, $T = 2$, with 20, 30 and 40 bins.

2.B. PROOF OF LEMMA 2.5.1

For an arbitrary bin \mathcal{B} with boundaries $[b_l, b_r]$, the piecewise linear continuous approximation of $g(s)$ can be specified by

$$\widehat{g}_H(s) = \frac{\Delta \widehat{g}_H}{\Delta s} s + \frac{\widehat{g}_H(b_l) b_r - \widehat{g}_H(b_r) b_l}{\Delta s}, \quad (2.B.1)$$

where $\widehat{g}_H(b_l)$ and $\widehat{g}_H(b_r)$ are approximations of $g_H(b_l)$ and $g_H(b_r)$, respectively, and $\Delta s := b_r - b_l$ and $\Delta \widehat{g}_H := \widehat{g}_H(b_r) - \widehat{g}_H(b_l)$. As $g_H(s)$ is smooth, we can express it as a Taylor series around the midpoint of $[b_l, b_r]$, which we call s_m :

$$g_H(s) = g_H(s_m) + \sum_{n=1}^{\infty} \frac{g_H^{(n)}(s_m)}{n!} (s - s_m)^n. \quad (2.B.2)$$

Assuming Δs to be small, we compute the square of the local L_2 error

$$\|g_H - \widehat{g}_H\|_{L^2(\mathcal{B})}^2 = \int_B (g_H(s) - \widehat{g}_H(s))^2 ds = \int_{b_l}^{b_r} (g_H(s) - \widehat{g}_H(s))^2 ds, \quad (2.B.3)$$

where we use (2.B.1) and (2.B.2) up to some significant order. Combining (2.B.2) and (2.B.3):

$$\|g_H - \widehat{g}_H\|_{L^2(\mathcal{B})}^2 = \int_{b_l}^{b_r} \left(g_H(s_m) - \widehat{g}_H(s) + \sum_{n=1}^{\infty} \frac{g_H^{(n)}(s_m)}{n!} (s - s_m)^n \right)^2 ds. \quad (2.B.4)$$

We now derive an expression for $g_H(s_m) - \widehat{g}_H(s)$. The first step is plugging in (2.B.1). This gives:

$$g_H(s_m) - \widehat{g}_H(s) = g_H(s_m) - \frac{\Delta \widehat{g}_H}{\Delta s} s - \frac{\widehat{g}_H(b_l) b_r - \widehat{g}_H(b_r) b_l}{\Delta s}. \quad (2.B.5)$$

Using the Taylor series expression in (2.B.2), we have for an arbitrary s :

$$g_H(s_m) = g_H(s) - \sum_{n=1}^{\infty} \frac{g_H^{(n)}(s_m)}{n!} (s - s_m)^n.$$

Expanding $g_H(s_m)$ at the boundary points and by plugging this result into (2.B.5) we find

$$\begin{aligned} g_H(s_m) - \widehat{g}_H(s) &= \frac{1}{2} (g_H(b_l) + g_H(b_r)) - \frac{1}{2} \left(\sum_{n=1}^{\infty} \frac{g_H^{(n)}(s_m)}{n!} \left(-\frac{1}{2}\Delta s\right)^n + \sum_{n=1}^{\infty} \frac{g_H^{(n)}(s_m)}{n!} \left(\frac{1}{2}\Delta s\right)^n \right) \\ &\quad - \frac{\Delta \widehat{g}_H}{\Delta s} s - \frac{\widehat{g}_H(b_l) b_r - \widehat{g}_H(b_r) b_l}{\Delta s}, \end{aligned}$$

where we have used the relations $b_l - s_m = -\frac{1}{2}\Delta s$ and $b_r - s_m = \frac{1}{2}\Delta s$. Odd terms in the two Taylor series cancel each other out. Even terms are equal. This results in:

$$g_H(s_m) - \widehat{g}_H(s) = \frac{1}{2} (g_H(b_l) + g_H(b_r)) - \sum_{n=2,4,6} \frac{g_H^{(n)}(s_m)}{n!} \left(\frac{1}{2}\Delta s\right)^n - \frac{\Delta \widehat{g}_H}{\Delta s} s - \frac{\widehat{g}_H(b_l) b_r - \widehat{g}_H(b_r) b_l}{\Delta s}.$$

Now, after some algebraic manipulations we end up with

$$g_H(s_m) - \widehat{g}_H(s) = - \sum_{n=2,4,6} \frac{g_H^{(n)}(s_m)}{2^n \cdot n!} (\Delta s)^n - \frac{(s - s_m) \Delta \widehat{g}_H}{\Delta s} + c_1 + \mathcal{O}(\Delta s^8),$$

where

$$c_1 := \frac{1}{2} ((g_H(b_l) - \widehat{g}_H(b_l)) + (g_H(b_r) - \widehat{g}_H(b_r))).$$

The constant c_1 can be considered as the average error at the boundaries of the interval. Plugging this result into (2.B.4) yields

$$\begin{aligned} \|g_H - \widehat{g}_H\|_{L^2(\mathcal{O})}^2 &= \int_{b_l}^{b_r} \left(\sum_{n=1}^6 \frac{g_H^{(n)}(s_m)}{n!} (s - s_m)^n - \sum_{n=2,4,6} \frac{g_H^{(n)}(s_m)}{2^n \cdot n!} (\Delta s)^n - \frac{(s - s_m) \Delta \widehat{g}_H}{\Delta s} + c_1 + \mathcal{O}(\Delta s^8) \right)^2 ds \end{aligned} \quad (2.B.6)$$

Evaluating (2.B.6) yields the result in Lemma 2.5.1.

3

THE TIME-DEPENDENT FX-SABR MODEL: EFFICIENT CALIBRATION BASED ON EFFECTIVE PARAMETERS

We present a framework for efficient calibration of the time-dependent SABR model [40, 63, 92] in an FX context. In a similar fashion as in [96], we derive effective parameters, which yield an accurate and efficient calibration. On top of the calibrated FX-SABR model, we add a non-parametric local volatility component, which naturally compensates for possible calibration errors. By means of Monte Carlo pricing experiments we show that the time-dependent FX-SABR model enables an accurate and consistent pricing of barrier options and outperforms the constant-parameter SABR model and the traditional local volatility model [34, 35]. We also discuss the role of the local volatility component in the valuation of barrier options.

Keywords: Time-Dependent SABR, FX, Calibration, Effective Parameters, local volatility, Monte Carlo, Path-Dependent.

3.1. INTRODUCTION

The pricing and hedging of complex path-dependent financial products requires an accurate calibration to prices of European-type options with different expiry dates, which contain information about the market behaviour through time. The model should also reflect realistic implied volatility smile dynamics, both with respect to the forward smile [115] and, secondly, to the underlying.

The payoff of a path-dependent product is determined by the evolution of the underlying through time, i.e. its price depends on the transition densities from one future

This chapter is based on the article ‘The Time-Dependent FX-SABR Model: Efficient Calibration based on Effective Parameters’, published in *International Journal of Theoretical and Applied Finance*, 18(6):1550042, 2015 [119].

state to another [10]. The future transition densities implied by a particular model are reflected by the forward implied volatility smiles it produces. Although the local volatility model [34, 35] can be calibrated perfectly to any set of arbitrage-free European-type option prices, it exhibits a flattening of the forward smile. This may lead to a mispricing of financial products that are sensitive to the forward implied volatility skew, like forward starting options, cliquets and path-dependent products. Alternatives for pricing such contracts are stochastic volatility models, which predict that the forward smile has a comparable shape as the smile observed today and typically yield more accurate results [10, 37, 46].

The ‘local dynamics’ of the implied volatility smile, i.e. with respect to the underlying, are also relevant, especially for hedging purposes. One of the main motivations for Hagan et al. [63] to introduce the *SABR model* was the typically inaccurate smile movement as the underlying changes predicted by the local volatility model. In particular, this model predicts that the smile shifts to higher prices as the underlying moves to lower prices, which is more extreme than the market behavior or may be even opposite to it, resulting in unstable hedges [10, 63, 76]. In contrast, the smile implied by the SABR model ‘follows’ the underlying¹.

On the base of the previous discussion, we consider here the *time-dependent SABR model* in an FX context. As this model involves time-dependent parameters, it allows for calibration to European-type option prices with different expiry dates. Furthermore, compared to the local volatility model, it yields more realistic forward implied volatility smiles and thirdly, this model is able of capturing the smile dynamics with respect to the underlying more accurately.

In [63] the time-dependent SABR model was already presented. Osajima [92] derived an asymptotic expression for the implied volatility. Furthermore, he introduced a new ‘FX hybrid SABR model’ and gave an asymptotic expansion formula for implied volatilities. Fernandez et al. [40] apply GPU technology for the Monte Carlo calibration of the static and time-dependent SABR models. They assume a time-dependent vol-vol parameter and correlation under the condition that these parameters decrease over time. In [50] the authors assume piecewise-constant parameters and show how the asymptotic expansion of the bivariate transition density of the underlying and its stochastic volatility presented in [128] are used in the calibration. Further, in [79] a closed-form approximation of the option price for the time-dependent SABR model is derived.

From a theoretical point of view, implied volatility expansion formulas, as in e.g. [79, 92], yield highly efficient calibration. However, in a practical sense these formulas typically only work under certain parameter conditions.

In this chapter, we calibrate the time-dependent SABR model by means of *effective parameters*. Effective parameters can be considered as ‘sophisticated averages’ of the corresponding time-dependent parameters. In [96] the effective parameters approach has been followed with respect to the time-dependent Displaced Diffusion Stochastic Volatility (DDSV) model. In a similar fashion, we derive here effective parameters by considering the qualitative effects of the SABR parameters on the shape of the implied

¹Rebonato [102] supports the conclusions of Hagan et al. regarding the dynamics of the smile with respect to the underlying in an FX context, but he points out that these may not be valid when stochastic interest rates are involved.

volatility smile. By means of numerical experiments we show that our approach yields an accurate and efficient calibration. Moreover, the idea behind the effective parameters is intuitive and the resulting formulas are straightforward and relatively easy to implement.

As the calibration may not be perfect due to e.g. possible inaccuracies in the mapping of the time-dependent to effective parameters, we add a non-parametric local volatility component (see e.g. [104]), which ‘bridges’ the mismatch between the prices in the market and the ones implied by the calibrated model. See also the model in the previous chapter.

The present chapter is organized as follows. In Section 3.2 we present the dynamics of the time-dependent FX-SABR model. We also specify the non-parametric local volatility component. In Section 3.3 we describe the calibration problem for the time-dependent SABR model and in which way effective parameters facilitate efficient calibration. Thereafter, in Sections 3.4.1, 3.4.2 and 3.4.3 we state results regarding the effective vol-vol, term structure and correlation parameters, respectively. Subsequently, we calibrate the time-dependent SABR model in Section 3.5. We show that the local volatility component yields an enhancement in the experimental results. Further, we price standard barrier options and compare results for the constant-parameter FX-SABR model, the local volatility model and the time-dependent FX-SABR model. We also discuss the role of the local volatility component. Section 3.6 concludes.

3.2. TIME-DEPENDENT FX-SABR MODEL WITH LOCAL VOLATILITY

In this section we present the time-dependent FX-SABR model. As the spot dynamics involve time-dependent zero-coupon bonds, we cannot directly apply Hagan’s formulas [63]. We resolve this issue by fixing the expiry of the zero-coupon bonds at the largest time to maturity, which represents the terminal payment date.

We first calibrate the time-dependent FX-SABR model. Subsequently we add a non-parametric local volatility component, which compensates for possible calibration inaccuracies. The local volatility component can compensate for *any* calibration error, which is a consequence of the mimicking theorem of Gyöngy [62]². A perfect fit to a European-type option price can be obtained, as its price is determined by the distribution of the underlying at a *particular point in time*.

Despite this feature of the local volatility component, the stochastic volatility parameters need to be calibrated accurately. European-type option prices with different expiry dates provide insight in the market behavior over time and our target is to ‘capture’ this information in the model by replicating these prices. Further, for the hedging and pricing of path-dependent products, a model should reflect realistic smile dynamics, both with respect to the forward implied volatility smile and regarding changes in the underlying. As the SABR model, compared to the local volatility model, typically captures these features more accurately [10, 46, 63, 76], we will reduce the contribution of the local volatility component by an accurate calibration.

²The theorem states that given a general Itô process, a Markov process containing a local volatility component with the same marginal distributions as the former exists.

3.2.1. TIME-DEPENDENT FX-SABR MODEL

Let $r_d(t)$ and $r_f(t)$ denote the deterministic domestic and foreign interest rates, respectively, and $M_d(t)$ and $M_f(t)$ are corresponding money accounts, determined by

$$dM_d(t) = r_d(t)M_d(t)dt, \quad dM_f(t) = r_f(t)M_f(t)dt.$$

Let $y(t)$ be the spot FX, expressed in units of domestic currency per unit of a foreign currency. Further, define

$$P_d(t, T) := M_d(t)\mathbb{E}^{\mathbb{Q}} \left[\frac{1}{M_d(T)} \middle| \mathcal{F}(t) \right], \quad P_f(t, T) := M_f(t)\mathbb{E}^{\mathbb{Q}^f} \left[\frac{1}{M_f(T)} \middle| \mathcal{F}(t) \right]$$

as the domestic and foreign zero-coupon bonds, respectively, see e.g. [88, 97].

The *time-dependent FX-SABR model* assumes the following dynamics under the domestic risk-neutral \mathbb{Q} -measure:

$$dy(t) = (r_d(t) - r_f(t))y(t)dt + \omega(t)\sigma(t) \left(\frac{P_d(t, T)}{P_f(t, T)} \right)^{1-\beta} y^\beta(t) dW_y^{\mathbb{Q}}(t), \quad (3.2.1)$$

$$d\sigma(t) = \gamma(t)\sigma(t)dW_\sigma^{\mathbb{Q}}(t), \quad (3.2.2)$$

with $y(0) = y_0$, $\sigma(0) = 1$, $dW_y^{\mathbb{Q}}(t)dW_\sigma^{\mathbb{Q}}(t) = \rho_{y,\sigma}(t)dt$ and $\rho_{y,\sigma}(t)$, $\gamma(t)$ and β denoting the correlation, vol-vol parameter and skew parameter, respectively. Further, T denotes the maturity time. The skew parameter is typically set at $\beta = 0.5$ [103]. For calibration purposes the volatility dynamics are scaled, which introduces the term structure parameter $\omega(t)$ in (3.2.1).

As this work aims to apply the SABR model in an FX context, it is convenient to price under the forward measure. The traditional SABR model describes the dynamics of the forward under the corresponding forward measure. Assuming a grid of N expiries T_i , $1, 2, \dots, N$, the FX forward $y^{T_i}(t) := y(t) \frac{P_f(t, T_i)}{P_d(t, T_i)}$ is a martingale under the domestic T_i -forward measure and the SABR model consistently prices an option with expiry date T_i . However, *simultaneous* pricing of options with different expiries raises consistency issues when using a single set of time-dependent parameters. We resolve this by writing the dynamics of $y^{T_i}(\cdot)$ with respect to T_N , which represents the terminal payment date³.

More concretely, let T_1, T_2, \dots, T_N be a set of expiries and suppose that the spot dynamics (3.2.1)-(3.2.2) involve the zero-coupon bonds $P_d(t, T_N)$ and $P_f(t, T_N)$. The dynamics of the forward $y^{T_i}(t)$ corresponding to an arbitrary expiry T_i then read:

$$\begin{aligned} dy^{T_i}(t) &= d \left(y(t) \frac{P_f(t, T_i)}{P_d(t, T_i)} \right) \\ &= \frac{P_f(t, T_i)}{P_d(t, T_i)} dy(t) + (r_f(t) - r_d(t)) y(t) \frac{P_f(t, T_i)}{P_d(t, T_i)} dt. \end{aligned}$$

³A similar approach was followed in [50]. The authors in [21] resolve the issue in a stochastic interest rates framework by projecting the volatility term on a lognormal distribution, which yields forward dynamics that are in the desired SABR form.

Substituting the dynamics of $y(t)$ in (3.2.1) yields⁴:

$$dy^{T_i}(t) = \left(\frac{P_f(t, T_i)}{P_d(t, T_i)} \right)^\beta \left(\frac{P_f(t, T_i)}{P_d(t, T_i)} \right)^{1-\beta} \omega(t) \sigma(t) \left(\frac{P_d(t, T_N)}{P_f(t, T_N)} \right)^{1-\beta} y^\beta(t) dW_y^{T_i}(t).$$

By definition of the zero-coupon bond, we have, for deterministic interest rates $P(T_i, T_N) = \frac{P(t, T_N)}{P(t, T_i)}$.

$$dy^{T_i}(t) = \omega(t) \sigma(t) \left(\frac{P_d(T_i, T_N)}{P_f(T_i, T_N)} \right)^{1-\beta} (y^{T_i}(t))^\beta dW_y^{T_i}(t).$$

Scaling the forward dynamics results in the following model:

$$dy^{T_i}(t) = \omega_1(t) \sigma(t) (y^{T_i}(t))^\beta dW_y^{T_i}(t), \quad y^{T_i}(0) =: \bar{y}_0^{T_i} = 1, \quad (3.2.3)$$

$$d\sigma(t) = \gamma(t) \sigma(t) dW_\sigma^{T_i}, \quad \sigma(0) = 1, \quad (3.2.4)$$

with $dW_y^{T_i}(t) dW_\sigma^{T_i}(t) = \rho_{y, \sigma}(t) dt$ and

$$\omega_1(t) := \omega(t) \left(\frac{P_d(T_i, T_N)}{y_0^{T_i} P_f(T_i, T_N)} \right)^{1-\beta}. \quad (3.2.5)$$

As model (3.2.3)-(3.2.4) is in ‘SABR form’, we can apply Hagan’s formulas [63] under the assumption of constant parameter values. This is a particularly useful property for the calibration framework we propose, which is based on effective parameters (see Section 3.3). The additional constants $P_d(T_i, T_N)$ and $P_f(T_i, T_N)$ in (3.2.5) allow for the calibration of model (3.2.1)-(3.2.2) across multiple expiries. Without these terms the spot dynamics are forward dependent, i.e. from the forward dynamics of $y^{T_i}(\cdot)$ and $y^{T_j}(\cdot)$, $i \neq j$ different dynamics for the spot FX and thus different models can be derived.

For the sake of notation, $y^{T_i}(t)$ and $\sigma(t)$ denote the scaled forward and volatility dynamics, respectively, unless otherwise mentioned. An exception holds for the initial forward: $y_0^{T_i}$ denotes the original initial forward and $\bar{y}_0^{T_i} = 1$ corresponds to the model with scaled forward dynamics.

3.2.2. LOCAL VOLATILITY COMPENSATOR

Calibration of model (3.2.1)-(3.2.2) may not be perfect. For this reason, we add a non-parametric local volatility component that acts as a ‘compensator’ for the mismatch between the market and calibrated model prices.

Adding the local volatility component $\sigma_{\text{SLV}}(t, y(t))$ to model (3.2.1)-(3.2.2) yields the following $y(t)$ dynamics:

$$dy(t) = (r_d(t) - r_f(t)) y(t) dt + \sigma_{\text{SLV}}(t, y(t)) \omega(t) \sqrt{V(t)} \left(\frac{P_d(t, T)}{P_f(t, T)} \right)^{1-\beta} y^\beta(t) dW_y^Q(t),$$

⁴As the Radon-Nikodym derivative [48] is $\Lambda_Q^{T_i} = \frac{P_d(t, T_i) M(0)}{P_d(0, T_i) M(t)} = 1$, we have $dW_y^{T_i}(t) = dW_y^Q(t)$.

where $V(t)$ denotes variance to avoid double use of the letter σ . Defining⁵

$$\psi(y(t), V(t)) := \omega(t) \sqrt{V(t)} \left(\frac{P_d(t, T)}{P_f(t, T)} \right)^{1-\beta} y^{\beta-1}(t), \quad (3.2.6)$$

the spot dynamics are given by:

$$dy(t)/y(t) = (r_d(t) - r_f(t)) dt + \sigma_{\text{SLV}}(t, y(t)) \psi(y(t), V(t)) dW_y^{\mathbb{Q}}(t), \quad (3.2.7)$$

$$dV(t) = \gamma^2(t) V(t) dt + 2\gamma(t) V(t) dW_{\sigma}^{\mathbb{Q}}(t), \quad V(0) = 1, \quad (3.2.8)$$

with $y(0) = y_0$ and $dW_y^{\mathbb{Q}}(t) dW_{\sigma}^{\mathbb{Q}}(t) = \rho_{y,\sigma}(t) dt$.

In the previous chapter and in [118] we applied the Tanaka-Meyer formula [114] and well-known relations between the option price and the underlying's marginal density based on the mimicking theory of Gyöngy [46, 62] to obtain the following result:

$$\sigma_{\text{SLV}}^2(t, K) = \frac{\sigma_{\text{LV}}^2(t, K)}{\mathbb{E}[\psi^2(y(t), V(t)) | y(t) = K]}.$$

Here $\sigma_{\text{LV}}^2(\cdot)$ denotes Dupire's local volatility term [35], which is either expressed in terms of European call prices or in implied volatilities $\bar{\sigma}(\cdot)$. In the numerical experiments we choose the latter:

$$\sigma_{\text{LV}}^2(t, K) = \frac{\bar{\sigma}^2(T, K) + 2\bar{\sigma}(T, K) T \left(\frac{\partial \bar{\sigma}(T, K)}{\partial T} + r K \frac{\partial \bar{\sigma}(T, K)}{\partial K} \right)}{\left(1 - \frac{Ky}{\bar{\sigma}(T, K)} \frac{\partial \bar{\sigma}(T, K)}{\partial K} \right)^2 + K \bar{\sigma}(T, K) T \left(\frac{\partial \bar{\sigma}(T, K)}{\partial K} - \frac{1}{4} K \bar{\sigma}(T, K) T \left(\frac{\partial \bar{\sigma}(T, K)}{\partial K} \right)^2 + K \frac{\partial^2 \bar{\sigma}(T, K)}{\partial K^2} \right)} \Bigg|_{T=t}.$$

Substituting (3.2.6) yields

$$\begin{aligned} \sigma_{\text{SLV}}^2(t, K) &= \frac{\sigma_{\text{LV}}^2(t, K)}{\mathbb{E} \left[\left(\omega(t) \sqrt{V(t)} \left(P_d(t, T) / P_f(t, T) \right)^{1-\beta} y^{\beta-1}(t) \right)^2 \Big| y(t) = K \right]} \\ &= \frac{\sigma_{\text{LV}}^2(t, K)}{\omega^2(t) \left(P_d(t, T) / P_f(t, T) \right)^{2-2\beta} K^{2\beta-2} \mathbb{E} [V(t) | y(t) = K]}. \end{aligned} \quad (3.2.9)$$

The conditional expectation appearing in (3.2.9) can be evaluated efficiently by the non-parametric method presented in the previous chapter.

3.3. CALIBRATION PROBLEM

In this section we discuss an important target of this work, namely the calibration of model (3.2.1)-(3.2.2). For this we need to price basic options on the FX rate with discounted value

$$C(t, T, K) = \mathbb{E}^{\mathbb{Q}} \left[\frac{M_d(t)}{M_d(T)} (y(T) - K)^+ \Big| \mathcal{F}(t) \right],$$

⁵For this model $\psi(\cdot)$ also depends on the underlying, $y(t)$, in contrast to the Heston-SLV model discussed in [118] and the previous chapter. As we condition on $y(t) = K$ though, this issue is resolved in a natural way.

where \mathbb{Q} denotes the domestic risk-neutral measure. As we mentioned at the beginning of Section 3.2, we will calibrate model (3.2.1)-(3.2.2) as accurately as possible in order to enable an accurate and consistent pricing and hedging of path-dependent and other forward volatility sensitive products.

In the calibration we make use of *effective parameters*, which are ‘sophisticated averages’ of their corresponding time-dependent counterparts. In this section we discuss how effective parameters facilitate efficient calibration.

The effective FX-SABR model is given by the following dynamics under the T_i -forward measure, $i = 1, \dots, N$:

$$d\tilde{y}^{T_i}(t) = \tilde{\omega}_1 \tilde{\sigma}(t) (\tilde{y}^{T_i}(t))^\beta dW_y^{T_i}(t), \quad \tilde{y}^{T_i}(0) = \bar{y}_0^{T_i} = 1, \quad (3.3.1)$$

$$d\tilde{\sigma}(t) = \tilde{\gamma} \tilde{\sigma}(t) dW_\sigma^{T_i}(t), \quad \tilde{\sigma}(0) = 1, \quad (3.3.2)$$

with $dW_y^{T_i}(t)dW_\sigma^{T_i}(t) = \tilde{\rho}_{y,\sigma} dt$ and

$$\tilde{\omega}_1 := \tilde{\omega} \left(\frac{P_d(T_i, T_N)}{y_0^{T_i} P_f(T_i, T_N)} \right)^{1-\beta}, \quad \tilde{\omega} := \sigma_0. \quad (3.3.3)$$

This is just the constant-parameter version of model (3.2.3)-(3.2.4), derived in Section 3.2.

Remark 3.3.1. We have defined model (3.2.3)-(3.2.4) such that $\bar{y}_0^{T_i} = 1$. Therefore, the formulas in this section can be simplified. However, for the sake of completeness we include the term $\bar{y}_0^{T_i}$.

In model (3.3.1)-(3.3.2) the vol-vol parameter $\tilde{\gamma}$ mainly accounts for curvature [103]. A second-order effect is that a higher vol-vol value results in a higher smile level. This effect may be negligible, but for the given set of parameter values⁶ it is significant. Further, $\tilde{\gamma}$ slightly affects skew. The effects are displayed in the left-hand graph of Figure 3.3.1.

The initial volatility $\tilde{\omega}$ mainly affects the smile level, as can be observed in the graph in the middle of Figure 3.3.1. On the base of an approximation formula, Hagan et al. [63] discuss that the ATM level of the implied volatility smile traverses along the *backbone* $\tilde{\omega}/(\bar{y}_0^{T_i})^{1-\beta}$. The initial volatility has a secondary, but marginal skew effect [103].

Correlation $\tilde{\rho}_{y,\sigma}$ also has two effects. It primarily affects the skew, as we see in the right-hand graph of Figure 3.3.1. Hagan et al. [63] quantify this effect as $\frac{1}{2} \tilde{\rho}_{y,\sigma} \lambda \log(K/\bar{y}_0^{T_i})$, which they refer to as ‘vanna skew’, with $\lambda = (\tilde{\gamma}/\tilde{\omega})(\bar{y}_0^{T_i})^{1-\beta}$ measuring the ‘strength’ of the vol-vol parameter $\tilde{\gamma}$ compared to the local volatility $\tilde{\omega}/(\bar{y}_0^{T_i})^{1-\beta}$. A second-order effect of correlation is curvature adjustment: a more negative value for $\tilde{\rho}_{y,\sigma}$ yields a decrease in curvature. The effect of the skew parameter β is threefold. We first mention its skew effect. A smaller value for β implies a more negative skew (a ‘steeper downward slope’ of the implied volatility smile). This effect is most clearly visible for the initial forward rate equal to 1, see the graph in the middle of Figure 3.3.2. Hagan et al. [63] determine $-\frac{1}{2}(1-\beta) \log(K/\bar{y}_0^{T_i})$ as being the skew implied by the skew parameter. Secondly, for an initial forward rate smaller than 1, a higher β value implies a downward shifting,

⁶In all figures the parameter values, if not varying, are: $y_0 = 2$, $\beta = 0.5$, $\tilde{\gamma} = 0.5$, $\tilde{\omega} = 0.1$, $\tilde{\rho}_{y,\sigma} = 0$ and $T_i = 1$.

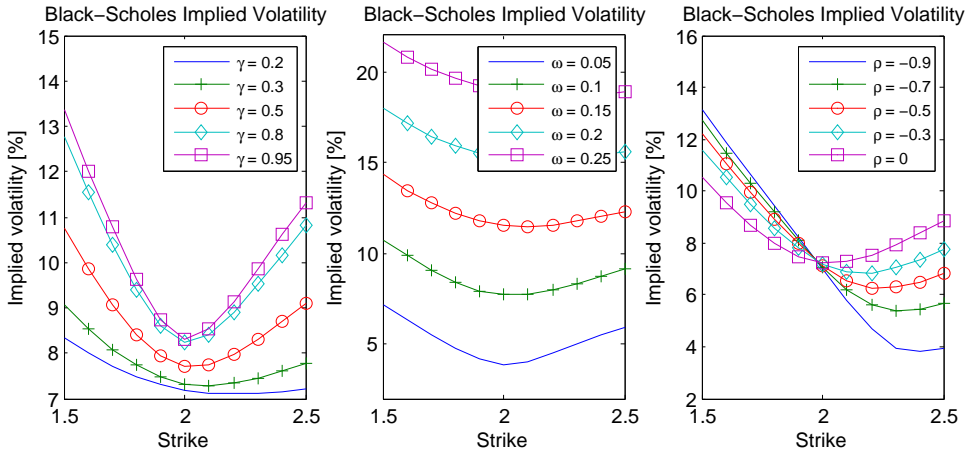


Figure 3.3.1: Effect of vol-vol parameter $\tilde{\gamma}$ (left), initial volatility $\tilde{\omega}$ (middle) and correlation $\tilde{\rho}_{y,\sigma}$ (right) on the shape of the implied volatility smile, $y_0 = 2$, $\beta = 0.5$, $T_i = 1$.

whereas for initial forward rate larger than 1 this effect is opposite (left-hand graph and right-hand graph of Figure 3.3.2, respectively). This effect is quantified by the backbone $\tilde{\omega}/(\bar{y}_0^{T_i})^{1-\beta}$. Last, β has a third-order curvature effect [103]. As we have defined model (3.3.1)-(3.3.2) such that $\bar{y}_0^{T_i} = 1$, the middle graph of Figure 3.3.2 applies.

Based on the smile effects just described, we will derive the effective parameters. We subsequently use the effective model (3.3.1)-(3.3.2) to calibrate the time-dependent parameters $\gamma(t)$, $\omega(t)$ and $\rho_{y,\sigma}(t)$. It will turn out that calibrating $\gamma(t)$ and $\rho_{y,\sigma}(t)$ yields a fit to the market in both curvature and skew. What remains, is a mismatch in level, which is compensated for by calibrating $\omega(t)$. We will add a local volatility component on top of the calibrated time-dependent SABR model to compensate for possible calibration inaccuracies.

3.3.1. CALIBRATION SET-UP

The difficulty of calibrating time-dependent parameters lies in the following. It is market practice to calibrate model (3.3.1)-(3.3.2) for a grid of – say N – expiries. The N parameter values obtained hold from $t = 0$ up to the corresponding expiries. In the case of a time-dependent parameter though, we are interested in calibrating *one* time-dependent function that is consistent with the market prices at *all* (N) expiries.

Let us elaborate on this problem. By means of calibrating the effective model (3.3.1)-(3.3.2) in the common way we find, amongst others, vol-vol parameter values $\tilde{\gamma}_1^{\text{mar}}$, $\tilde{\gamma}_2^{\text{mar}}$, ..., $\tilde{\gamma}_N^{\text{mar}}$ that correspond to the time-intervals $[0, T_1]$, $[0, T_2]$, ..., $[0, T_N]$, respectively. We call these ‘market effective’ parameters. Intuitively, $\tilde{\gamma}_i^{\text{mar}}$ ‘captures’ all information up to T_i . We can also extract market effective term structure values $\tilde{\omega}_1^{\text{mar}}$, $\tilde{\omega}_2^{\text{mar}}$, ..., $\tilde{\omega}_N^{\text{mar}}$ and correlations $\tilde{\rho}_{y,\sigma,1}^{\text{mar}}$, $\tilde{\rho}_{y,\sigma,2}^{\text{mar}}$, ..., $\tilde{\rho}_{y,\sigma,N}^{\text{mar}}$. In the following, we only describe how the time-dependent vol-vol parameter $\gamma(t)$ is obtained. The same procedure applies for finding $\omega(t)$ and $\rho_{y,\sigma}(t)$.

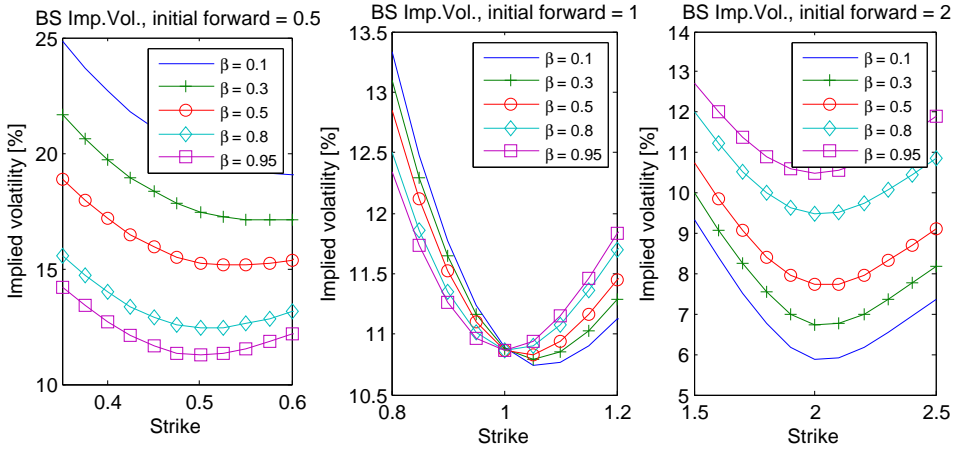


Figure 3.3.2: Effect of β on the shape of the implied volatility smile for initial forward values $\bar{y}_0^{T_i} = 0.5$ (left), $\bar{y}_0^{T_i} = 1$ (middle) and $\bar{y}_0^{T_i} = 2$ (right), $\tilde{\gamma} = 0.5$, $\tilde{\omega} = 0.1$, $\tilde{\rho}_{y,\sigma} = 0$, $T_i = 1$.

The time-dependent vol-vol parameter $\gamma(t)$ has to satisfy two requirements. The first one is trivial: our goal is to determine *one* time-dependent function. Secondly, it should relate to *all* expiries T_i , $i = 1, 2, \dots, N$ given in the market. That is, at the given expiries the values of the time-dependent parameter $\gamma(T_i)$ must yield the same implied volatility as the market effective parameter $\tilde{\gamma}_i^{\text{mar}}$ does.

If only the latter requirement had to be satisfied, we could just choose $\gamma(T_i) = \tilde{\gamma}_i^{\text{mar}}$ for all expiries. In this case we would have N constant functions $\gamma(\cdot)$ that ‘live’ on different time-intervals. The first requirement though complicates the problem of finding a time-dependent parameter: mapping the set of market effective parameters $\{\tilde{\gamma}_1^{\text{mar}}, \tilde{\gamma}_2^{\text{mar}}, \dots, \tilde{\gamma}_N^{\text{mar}}\}$ onto *one* time-dependent parameter $\gamma(t)$ is not straightforward. An easier problem is to transfer from a time-dependent parameter to its ‘effective equivalent’, so to find a mapping

$$\{\gamma(t), 0 \leq t \leq T_N\} \rightarrow \{\tilde{\gamma}_1^{\text{mod}}, \tilde{\gamma}_2^{\text{mod}}, \dots, \tilde{\gamma}_N^{\text{mod}}\}, \quad (3.3.4)$$

where $\tilde{\gamma}_i^{\text{mod}}$, $i = 1, 2, \dots, N$, denote the effective vol-vol parameter values implied by the model, in particular by the time-dependent vol-vol parameter.

Suppose we have established mapping (3.3.4). Subsequently, in the calibration we numerically find $\gamma(t)$ such that $\tilde{\gamma}_i^{\text{mod}} = \tilde{\gamma}_i^{\text{mar}}$, $i = 1, \dots, N$, or on the base of the matching implied volatilities obtained using $\tilde{\gamma}_i^{\text{mod}}$ and $\tilde{\gamma}_i^{\text{mar}}$, respectively. From a computational point of view, finding $\gamma(t)$ in this way is *significantly less expensive* than repeatedly applying a pricing method in the calibration that is suitable for time-dependent parameters [96] (e.g. Monte Carlo simulation or a PDE-based approach). In a similar way we find $\omega(t)$ and $\rho_{y,\sigma}(t)$.

3.4. EFFECTIVE PARAMETERS

Based on their effects on the shape of the implied volatility smile, we derive the effective parameters. In the following subsections we derive mappings for the effective vol-vol parameter, term structure and correlation, respectively. The results obtained are used in the calibration in Section 3.5.

3.4.1. EFFECTIVE VOL-VOL PARAMETER

The vol-vol parameter mainly affects curvature. Curvature is introduced⁷ by adding stochastic volatility to the CEV model [25], which results in the SABR model. An appropriate measure for curvature is the ‘realized volatility’, which is defined as

$$\int_0^{T_i} \omega_1(t) \sigma(t) dW^{T_i}(t),$$

where $\omega_1(t)$ is a deterministic scaling parameter specified by (3.2.5).

We determine the effective vol-vol parameter such that the realized volatilities of the time-dependent and effective models are equal in distribution, that is

$$\int_0^{T_i} \omega_1(t) \sigma(t) dW^{T_i}(t) \stackrel{d}{=} \int_0^{T_i} \tilde{\omega}_1 \tilde{\sigma}(t) dW^{T_i}(t).$$

More concretely, we obtain the effective vol-vol parameter by matching the *moments* of the realized volatilities.

In the following lemma we first state the main result of this section. We subsequently provide a proof for it.

Lemma 3.4.1 (Effective vol-vol parameter). *By matching moments of the realized volatilities of the time-dependent and effective models, the effective vol-vol parameter $\tilde{\gamma}$ corresponding to the expiry T_i is obtained by the following equation:*

$$\begin{aligned} & \int_0^{T_i} \omega_1^2(t) \left(\int_0^t \omega_1^2(s) e^{6 \int_0^s \gamma^2(u) du + \int_s^t \gamma^2(u) du} ds \right) dt \\ &= \frac{1}{5} \left(\frac{\int_0^{T_i} \omega_1^2(t) e^{\int_0^t \gamma^2(u) du} dt}{e^{\tilde{\gamma}^2 T_i} - 1} \right)^2 \left(\frac{1}{6} e^{6\tilde{\gamma}^2 T_i} - e^{\tilde{\gamma}^2 T_i} + \frac{5}{6} \right). \end{aligned}$$

Proof. It is easy to see that first moment matching of the realized volatilities does not give conclusive results: trivially, $\mathbb{E} \left[\int_0^{T_i} \omega_1(t) \sigma(t) dW^{T_i}(t) \right] = \mathbb{E} \left[\int_0^{T_i} \tilde{\omega}_1 \tilde{\sigma}(t) dW^{T_i}(t) \right] = 0$. We therefore proceed by matching the variances of the realized volatilities, i.e.

$$\begin{aligned} \mathbb{E} \left[\left(\int_0^{T_i} \omega_1(t) \sigma(t) dW^{T_i}(t) \right)^2 \right] &= \mathbb{E} \left[\left(\int_0^{T_i} \tilde{\omega}_1 \tilde{\sigma}(t) dW^{T_i}(t) \right)^2 \right] \\ \Leftrightarrow \int_0^{T_i} \omega_1^2(t) \mathbb{E}[\sigma^2(t)] dt &= \tilde{\omega}_1^2 \int_0^{T_i} \mathbb{E}[\tilde{\sigma}^2(t)] dt, \end{aligned}$$

⁷Neglecting the higher-order curvature effects of the correlation and skew parameter.

which gives

$$\tilde{\omega}_1^2 = \frac{\int_0^{T_i} \omega_1^2(t) \mathbb{E}[\sigma^2(t)] dt}{\int_0^{T_i} \mathbb{E}[\tilde{\sigma}^2(t)] dt}.$$

As

$$\mathbb{E}[\sigma^2(t)] = e^{\int_0^t \gamma^2(u) du}, \quad \mathbb{E}[\tilde{\sigma}^2(t)] = e^{\tilde{\gamma}^2 t},$$

we have

$$\tilde{\omega}_1^2 = \frac{\int_0^{T_i} \omega_1^2(t) e^{\int_0^t \gamma^2(u) du} dt}{\int_0^{T_i} e^{\tilde{\gamma}^2 t} dt} = \tilde{\gamma}^2 \frac{\int_0^{T_i} \omega_1^2(t) e^{\int_0^t \gamma^2(u) du} dt}{e^{\tilde{\gamma}^2 T_i} - 1}. \quad (3.4.1)$$

Equation (3.4.1) yields two unknowns: $\tilde{\omega}_1$ and $\tilde{\gamma}$. In order to have a system with two equations and two unknowns that we can solve for $\tilde{\gamma}$, we match another higher moment as follows:

$$\mathbb{E} \left[\left(\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt \right)^2 \right] = \tilde{\omega}_1^4 \mathbb{E} \left[\left(\int_0^{T_i} \tilde{\sigma}^2(t) dt \right)^2 \right]. \quad (3.4.2)$$

Evaluating the left-hand side of (3.4.2) gives:

$$\mathbb{E} \left[\left(\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt \right)^2 \right] = \mathbb{E} \left[\int_0^{T_i} \left(\int_0^{T_i} \omega_1^2(s) \sigma^2(s) ds \right) \omega_1^2(t) \sigma^2(t) dt \right].$$

By symmetry around the point $s = t$, we have

$$\begin{aligned} \mathbb{E} \left[\left(\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt \right)^2 \right] &= 2 \mathbb{E} \left[\int_0^{T_i} \int_0^t \omega_1^2(s) \sigma^2(s) \omega_1^2(t) \sigma^2(t) ds dt \right] \\ &= 2 \int_0^{T_i} \omega_1^2(t) \int_0^t \omega_1^2(s) \mathbb{E}[\sigma^2(s) \sigma^2(t)] ds dt. \end{aligned} \quad (3.4.3)$$

To evaluate $\mathbb{E}[\sigma^2(s) \sigma^2(t)]$, we use the dynamics of the squared volatility $\sigma^2(t)$, which can easily be derived:

$$d\sigma^2(t) = \sigma^2(t) \left(\gamma^2(t) dt + 2\gamma(t) dW_\sigma^{T_i}(t) \right).$$

Applying basic Itô calculus yields

$$\mathbb{E}[\sigma^2(s) \sigma^2(t)] = e^{6 \int_0^s \gamma^2(u) du + \int_s^t \gamma^2(u) du}.$$

Combining this result with (3.4.3) gives for the left-hand side of (3.4.2):

$$\mathbb{E} \left[\left(\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt \right)^2 \right] = 2 \int_0^{T_i} \omega_1^2(t) \left(\int_0^t \omega_1^2(s) e^{6 \int_0^s \gamma^2(u) du + \int_s^t \gamma^2(u) du} ds \right) dt. \quad (3.4.4)$$

In a similar way we obtain for the right-hand side

$$\tilde{\omega}_1^4 \mathbb{E} \left[\left(\int_0^{T_i} \tilde{\sigma}^2(t) dt \right)^2 \right] = \frac{2\tilde{\omega}_1^4}{5\tilde{\gamma}^4} \left(\frac{1}{6} e^{6\tilde{\gamma}^2 T_i} - e^{\tilde{\gamma}^2 T_i} + \frac{5}{6} \right). \quad (3.4.5)$$

We now substitute (3.4.1) in the right-hand side of (3.4.5) and equation (3.4.2) becomes:

$$\begin{aligned} & \int_0^{T_i} \omega_1^2(t) \left(\int_0^t \omega_1^2(s) e^{6 \int_0^s \gamma^2(u) du + \int_s^t \gamma^2(u) du} ds \right) dt \\ &= \frac{1}{5} \left(\frac{\int_0^{T_i} \omega_1^2(t) e^{\int_0^t \gamma^2(u) du} dt}{e^{\tilde{\gamma}^2 T_i} - 1} \right)^2 \left(\frac{1}{6} e^{6\tilde{\gamma}^2 T_i} - e^{\tilde{\gamma}^2 T_i} + \frac{5}{6} \right). \end{aligned}$$

We numerically solve this equation for $\tilde{\gamma}$. The other unknown parameter, $\tilde{\omega}_1$, has vanished. \square

Remark 3.4.1 (Piecewise-constant parameters). *In the case of piecewise-constant parameters we can derive analytical expressions for the integrals in (3.4.1) and (3.4.4), which significantly speeds up the calibration procedure.*

NUMERICAL EXPERIMENT

Let $y_0 = 2$, $\beta = 0.5$, $\omega(t) = 0.15$, $r_d = 0.05$ and $r_f = 0.02$. Further, we assume $\rho_{y,\sigma}(t) = 0$.

Remark 3.4.2 (Interest rates). *In the calibration ‘effective’ and time-dependent domestic and foreign interest rates can be extracted from the initial (non-scaled) forward $y_0^{T_i}$ and spot y_0 via the relation $y_0^{T_i} := y_0 \frac{P_f(0, T_i)}{P_d(0, T_i)}$.*

Remark 3.4.3 (Zero correlation). *In the numerical experiments in this section and Section 3.4.2 we choose the correlation to be zero. Assuming zero correlation yields the advantage that the exact zero-correlation pricing formula of Antonov et al. [7] can serve as a benchmark. In the Monte Carlo simulation we apply a basic first-order Taylor approximation scheme. Typically, Monte Carlo simulation schemes of the SABR model are biased, especially for large vol-vol parameter values and small initial forward rates [22]. In order to make sure that results are not affected by the Monte Carlo bias, we can use the effective parameters both in a Monte Carlo simulation and in the zero-correlation pricing formula [7] to price European call options. In the numerical experiments in this section and Sections 3.4.2 and 3.4.3, we have confirmed that the results are free of Monte Carlo bias. We have also verified that for non-zero correlation values the performance of the effective parameters is similar to the performance as shown in this section and in Section 3.4.2.*

We assume values for $\gamma(t)$ as given in Table 3.4.1. We choose the vol-vol parameter to be decreasing over time, as in the FX market the curvature of the implied volatility smile typically diminishes for longer expiries. Effective vol-vol parameter values are also given in Table 3.4.1. In Figure 3.4.1 we display the Black-Scholes implied volatility smiles obtained by simulating the time-dependent model (3.2.1)-(3.2.2) and the effective model (also in spot measure) for $T = 1$, $T = 2$ and $T = 5$ with the corresponding $\tilde{\gamma}$ values (number of paths is $5 \cdot 10^5$, number of time-steps per year is 200). The curvature fit is highly satisfactory⁸.

⁸We have derived $\tilde{\gamma}$ by only considering the vol-vol’s primary effect (which is on the smile’s curvature). Due to this, a marginal level mismatch occurs between the time-dependent model and the effective model, as the vol-vol parameter also has a level effect. In order to obtain a level fit again, we adjust the term structure value of the effective model slightly by exploiting the techniques of Section 3.4.2.

t	$[0, \frac{1}{2})$	$[\frac{1}{2}, 1)$	$[1, 2)$	$[2, 3)$	$[3, 5]$
$\gamma(t)$	1	0.8	0.5	0.3	0.2
t	$\frac{1}{2}$	1	2	3	5
$\tilde{\gamma}$	1	0.911	0.785	0.692	0.565

Table 3.4.1: Time-dependent and effective vol-vol values

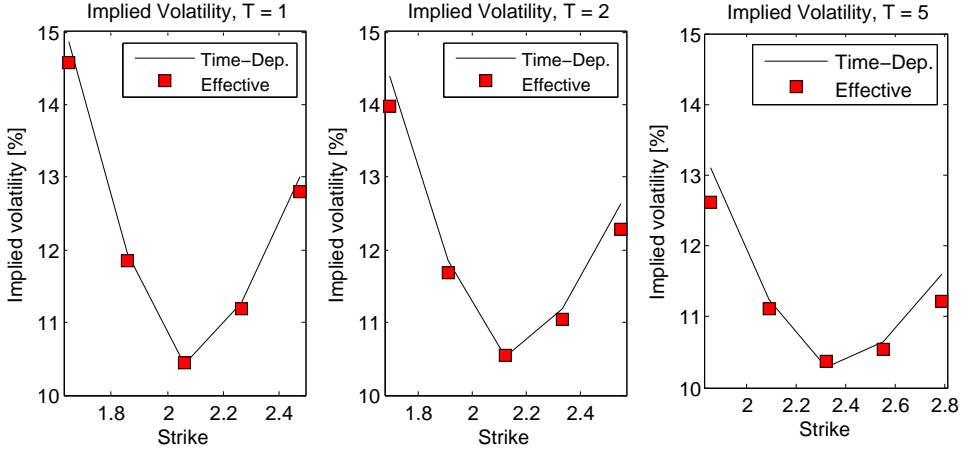


Figure 3.4.1: Performance of the effective vol-vol parameter for $T = 1$ (left), $T = 2$ (middle) and $T = 5$ (right).

3.4.2. EFFECTIVE TERM STRUCTURE

In Section 3.3 we mentioned that the parameter ω mainly affects the level of the implied volatility smile. Therefore, in this section we derive the effective term structure $\tilde{\omega}$ by matching ‘smile levels’ of the time-dependent and effective models. As the level of the smile is completely determined by the ATM implied volatility value, we match ATM prices of the time-dependent and effective models, that is:

$$\mathbb{E} \left[\left(y^{T_i}(T_i) - \bar{y}_0^{T_i} \right)^+ \right] = \mathbb{E} \left[\left(\tilde{y}^{T_i}(T_i) - \bar{y}_0^{T_i} \right)^+ \right], \quad \bar{y}_0^{T_i} = 1. \tag{3.4.6}$$

Lemma 3.4.2. *We assume in SDE system (3.2.3)-(3.2.4) for the forward y^{T_i} lognormal dynamics. The expected ATM payoff at time T_i is then given by*

$$\mathbb{E} \left[\left(y^{T_i}(T_i) - 1 \right)^+ \right] = \mathbb{E} [g(x)],$$

with

$$g(x) := 2\Phi \left(\frac{1}{2} \sqrt{x} \right) - 1, \quad x := \int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt,$$

where Φ denotes the standard normal cumulative distribution function.

Proof. Assuming lognormal dynamics for the forward, we have

$$y^{T_i}(T_i) = e^{-\frac{1}{2} \int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt + \int_0^{T_i} \omega_1(t) \sigma(t) dW_y^{T_i}(t)}.$$

Suppose that the path $\{\omega_1(t)\sigma(t), 0 \leq t \leq T_i\}$ is given. Defining

$$Z(T_i) := \int_0^{T_i} \omega_1^2(t)\sigma^2(t)dt, \quad \tilde{\sigma}^2 := \frac{Z(T_i)}{T_i},$$

we have

$$y^{T_i}(T_i) = e^{-\frac{1}{2}\tilde{\sigma}^2 T_i + \tilde{\sigma} W_y^{T_i}(T_i)},$$

and also

$$\mathbb{E} \left[(y^{T_i}(T_i) - 1)^+ \mid \omega_1(t)\sigma(t), 0 \leq t \leq T_i \right] = \Phi(d_+) - \Phi(d_-),$$

with

$$d_{\pm} = \frac{\log(1) \pm \frac{1}{2}\tilde{\sigma}^2 T_i}{\tilde{\sigma}\sqrt{T_i}} = \pm \frac{1}{2}\tilde{\sigma}\sqrt{T_i} = \pm \frac{1}{2}\sqrt{Z(T_i)}.$$

As a consequence, the conditional expectation can be written as:

$$\begin{aligned} \mathbb{E} \left[(y^{T_i}(T_i) - 1)^+ \mid \omega_1(t)\sigma(t), 0 \leq t \leq T_i \right] &= \Phi\left(\frac{1}{2}\sqrt{Z(T_i)}\right) - \Phi\left(-\frac{1}{2}\sqrt{Z(T_i)}\right) \\ &= 2\Phi\left(\frac{1}{2}\sqrt{Z(T_i)}\right) - 1. \end{aligned}$$

Applying the Tower property yields:

$$\begin{aligned} \mathbb{E} \left[(y^{T_i}(T_i) - 1)^+ \right] &= \mathbb{E} \left[\mathbb{E} \left[(y^{T_i}(T_i) - 1)^+ \mid \omega_1(t)\sigma(t), 0 \leq t \leq T_i \right] \right] \\ &= \mathbb{E} [g(x)], \end{aligned}$$

with

$$g(x) := 2\Phi\left(\frac{1}{2}\sqrt{x}\right) - 1, \quad x := \int_0^{T_i} \omega_1^2(t)\sigma^2(t)dt.$$

□

Analogously, for the effective model we have $\mathbb{E} \left[(\tilde{y}^{T_i}(T_i) - 1)^+ \right] = \mathbb{E} [g(\tilde{x})]$ with $\tilde{x} := \tilde{\omega}_1^2 \int_0^{T_i} \tilde{\sigma}^2(t)dt$. We continue with approximating $g(\cdot)$ by a simpler function, which yields an efficient evaluation of the effective term structure.

Remark 3.4.4. *In this section we use the superscripts (ℓ) and (r) to indicate approximation errors corresponding to the left-hand side and right-hand side of equation (3.4.6), respectively.*

Lemma 3.4.3. *Approximating $g(\cdot)$ by the corresponding Taylor series yields*

$$\mathbb{E} \left[(y^{T_i}(T_i) - 1)^+ \right] = \frac{1}{\sqrt{2\pi}} \mathbb{E} [\sqrt{x}] + \epsilon_1^{(\ell)}, \tag{3.4.7}$$

with

$$\epsilon_1^{(\ell)} := \frac{2}{\sqrt{\pi}} \left(-\frac{1}{3} \mathbb{E} [z^3] + \frac{1}{10} \mathbb{E} [z^5] - \dots \right), \quad z := \frac{1}{2}\sqrt{x/2}, \quad x := \int_0^{T_i} \omega_1^2(t)\sigma^2(t)dt.$$

Proof. By definition, we have

$$\Phi\left(\frac{1}{2}\sqrt{x}\right) = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{1}{2}\sqrt{x/2}\right) \right], \quad (3.4.8)$$

where $\operatorname{erf}(\cdot)$ denotes the (Gauss) error function⁹. Defining

$$x := \int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt, \quad z := \frac{1}{2}\sqrt{x/2},$$

we approximate the integrand of the error function by the first term of the corresponding Taylor series around 0:

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \frac{(-t^2)^0}{0!} dt + \bar{\epsilon}_T^{(\ell)} = \frac{2}{\sqrt{\pi}} z + \bar{\epsilon}_T^{(\ell)}, \quad (3.4.9)$$

with

$$\bar{\epsilon}_T^{(\ell)} := \frac{2}{\sqrt{\pi}} \left(-\frac{1}{3} z^3 + \frac{1}{10} z^5 - \dots \right),$$

denoting the truncation error. Combining (3.4.8) and (3.4.9) we obtain

$$g(x) = \frac{1}{\sqrt{2\pi}} \sqrt{x} + \bar{\epsilon}_T^{(\ell)}$$

and thus

$$\mathbb{E}[(y^{T_i}(T_i) - 1)^+] = \mathbb{E}[g(x)] = \frac{1}{\sqrt{2\pi}} \mathbb{E}[\sqrt{x}] + \epsilon_T^{(\ell)},$$

with

$$\epsilon_T^{(\ell)} := \frac{2}{\sqrt{\pi}} \left(-\frac{1}{3} \mathbb{E}[z^3] + \frac{1}{10} \mathbb{E}[z^5] - \dots \right).$$

□

Remark 3.4.5 (Qualitative analysis of $\epsilon_T^{(\ell)}$). *The truncation error $\epsilon_T^{(\ell)}$ is quantified as follows. We distinguish between two cases, namely $\frac{P_d(T_i, T_N)}{y_0^{T_i} P_f(T_i, T_N)} \leq 1$ and $\frac{P_d(T_i, T_N)}{y_0^{T_i} P_f(T_i, T_N)} > 1$. For the former case, by the definition of $\omega_1(\cdot)$, the truncation error is largest for $\beta = 1$. Typically, $\omega(\cdot) = \mathcal{O}(10^{-1})$ and as $\beta = 1$, we have $\omega_1(\cdot) = \omega(\cdot) = \mathcal{O}(10^{-1})$. Further, as the scaled volatility $\sigma(\cdot)$ has initial value 1, $\mathbb{E}[z] = \mathcal{O}(10^{-1})$. For realistic values for the vol-vol parameter the leading term in the truncation error is $\mathbb{E}[z^3]$ and is assumed to have a lower order of magnitude than $\mathbb{E}[z]$. For $\beta < 1$ the truncation error is smaller. In case $\frac{P_d(T_i, T_N)}{y_0^{T_i} P_f(T_i, T_N)} > 1$, the truncation error is largest for $\beta = 0$, and $\omega(\cdot) = \mathcal{O}(10^{-1})$ results in $\omega_1^2(\cdot) = \mathcal{O}((10y_0^{T_i})^{-2})$. As the scaled volatility $\sigma(\cdot)$ has initial value 1, typically $\mathbb{E}[z] = \mathcal{O}((10y_0^{T_i})^{-1})$. Realistic vol-vol*

⁹The Gauss error function is defined as

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

values imply that $\mathbb{E}[z^3]$ is the leading term in the truncation error. A smaller $y_0^{T_i}$ value implies a larger truncation error and the contribution of the non-parametric local volatility compensator is more significant. For $\beta > 0$ the truncation error is smaller. In the numerical experiments in Section 3.4.2 we show that relatively large $\omega(\cdot)$ and $\gamma(\cdot)$ values imply a slightly less accurate level fit, which is in line with the former qualitative analysis.

Analogous to the time-dependent model, the expected ATM payoff corresponding to the effective model is given by

$$\mathbb{E}[(\tilde{y}^{T_i}(T_i) - 1)^+] = \frac{1}{\sqrt{2\pi}} \mathbb{E}[\sqrt{\tilde{x}}] + \epsilon_T^{(r)}, \quad (3.4.10)$$

with

$$\epsilon_T^{(r)} := \frac{2}{\sqrt{\pi}} \left(-\frac{1}{3} \mathbb{E}[\tilde{z}^3] + \frac{1}{10} \mathbb{E}[\tilde{z}^5] - \dots \right), \quad \tilde{z} := \frac{1}{2} \sqrt{\tilde{x}/2}, \quad \tilde{x} := \tilde{\omega}_1^2 \int_0^{T_i} \tilde{\sigma}^2(t) dt.$$

For the truncation error $\epsilon_T^{(r)}$ a similar analysis holds as in Remark 3.4.5. Substitution of (3.4.7) and (3.4.10) in equation (3.4.6) gives

$$\frac{1}{\sqrt{2\pi}} \mathbb{E} \left[\sqrt{\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt} \right] + \epsilon_T^{(\ell)} = \frac{\tilde{\omega}_1}{\sqrt{2\pi}} \mathbb{E} \left[\sqrt{\int_0^{T_i} \tilde{\sigma}^2(t) dt} \right] + \epsilon_T^{(r)}. \quad (3.4.11)$$

In order to evaluate the expectations, we derive closed-form expressions for the ‘ $\frac{1}{2}$ th moments’ of the integrals in (3.4.11).

Lemma 3.4.4 (Effective term structure). *An approximation of the expected ATM payoff corresponding to the time-dependent model in (3.2.3)-(3.2.4) is given by*

$$\mathbb{E}[(y^{T_i}(T_i) - 1)^+] = \frac{1}{\sqrt{2\pi}} \left(\omega_1(0) \sqrt{\Delta_t} \hat{\phi}_{Y_M}(-\frac{1}{2}i) \right) + \epsilon_T^{(\ell)} + \epsilon_I^{(\ell)} + \epsilon_F^{(\ell)}, \quad (3.4.12)$$

with $\hat{\phi}_{Y_M}(\cdot)$ denoting an approximation of the characteristic function of

$Y_M := \log \left(\sum_{j=1}^M \frac{\omega_1^2(t_j) \sigma^2(t_j)}{\omega_1^2(0)} \right)$, with $t_j = j \Delta_t$, $\Delta_t = T_i / M$, $j = 1, \dots, M$. The error $\epsilon_I^{(\ell)}$ is due to an integral approximation and $\epsilon_F^{(\ell)}$ is introduced in the characteristic function approximation. Further, from (3.4.12) and a similar result for the effective model, the effective term structure corresponding to the expiry T_i is given by

$$\tilde{\omega} = \tilde{\omega}_1 \left(\frac{y_0^{T_i} P_f(T_i, T_N)}{P_d(T_i, T_N)} \right)^{1-\beta}, \quad \text{with} \quad \tilde{\omega}_1 = \frac{\omega_1(0) \hat{\phi}_{Y_M}(-\frac{1}{2}i)}{\hat{\phi}_{\tilde{Y}_M}(-\frac{1}{2}i)} + \varepsilon, \quad (3.4.13)$$

where $\hat{\phi}_{\tilde{Y}_M}(\cdot)$ is an approximation of the characteristic function of $\tilde{Y}_M := \log \left(\sum_{j=1}^M \tilde{\sigma}^2(t_j) \right)$ and ε represents the different error terms introduced, which are marginal and do not significantly affect the performance of the effective term structure parameter.

Proof. We start with approximating the integral

$$\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt = \Delta_t \sum_{j=1}^M \omega_1^2(t_j) \sigma^2(t_j) + \bar{\epsilon}_1^{(\ell)}, \quad (3.4.14)$$

with $\Delta_t = T_i/M$ and $\bar{\epsilon}_1^{(\ell)}$ denoting the error term corresponding to the time-interval $[0, T_i]$. Given the integral approximation we have

$$\mathbb{E} \left[\sqrt{\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt} \right] = \omega_1(0) \sqrt{\Delta_t} \mathbb{E} \left[\left(\sum_{j=1}^M \frac{\omega_1^2(t_j) \sigma^2(t_j)}{\omega_1^2(0)} \right)^{1/2} \right] + \epsilon_1^{(\ell)}, \quad (3.4.15)$$

with $\Delta_t = T_i/M$ and $\epsilon_1^{(\ell)}$ is due to the integral approximation error $\bar{\epsilon}_1^{(\ell)}$ in (3.4.14). We evaluate the expectation as follows. Given the characteristic function $\phi_{Y_M}(\cdot)$ of

$$Y_M := \log \left(\sum_{j=1}^M \frac{\omega_1^2(t_j) \sigma^2(t_j)}{\omega_1^2(0)} \right), \quad (3.4.16)$$

we have

$$\mathbb{E} \left[\left(\sum_{j=1}^M \frac{\omega_1^2(t_j) \sigma^2(t_j)}{\omega_1^2(0)} \right)^{1/2} \right] = \mathbb{E} \left[e^{\frac{1}{2} Y_M} \right] = \phi_{Y_M} \left(-\frac{1}{2} i \right)$$

and thus

$$\mathbb{E} \left[\sqrt{\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt} \right] = \omega_1(0) \sqrt{\Delta_t} \widehat{\phi}_{Y_M} \left(-\frac{1}{2} i \right) + \epsilon_1^{(\ell)} + \epsilon_F^{(\ell)}, \quad (3.4.17)$$

where $\epsilon_F^{(\ell)}$ denotes the error in the approximation of the characteristic function $\widehat{\phi}_{Y_M}(\cdot)$, which is introduced in the procedure of recovering it. In Appendix 3.A we describe the recovery procedure, which was developed in the context of Asian options in [132]. Furthermore, combining (3.4.17) with the result in (3.4.7) gives¹⁰

$$\mathbb{E} [(y^{T_i}(T_i) - 1)^+] = \frac{1}{\sqrt{2\pi}} \left(\omega_1(0) \sqrt{\Delta_t} \widehat{\phi}_{Y_M} \left(-\frac{1}{2} i \right) \right) + \epsilon_T^{(\ell)} + \epsilon_1^{(\ell)} + \epsilon_F^{(\ell)}. \quad (3.4.18)$$

For the effective model we can derive

$$\mathbb{E} [(\tilde{y}^{T_i}(T_i) - 1)^+] = \frac{1}{\sqrt{2\pi}} \left(\tilde{\omega}_1 \sqrt{\Delta_t} \widehat{\phi}_{\tilde{Y}_M} \left(-\frac{1}{2} i \right) \right) + \epsilon_T^{(r)} + \epsilon_1^{(r)} + \epsilon_F^{(r)}, \quad (3.4.19)$$

where $\widehat{\phi}_{\tilde{Y}_M}(\cdot)$ is an approximation of the characteristic function of $\tilde{Y}_M := \log \left(\sum_{j=1}^M \tilde{\sigma}^2(t_j) \right)$. Given the identities (3.4.18) and (3.4.19), equation (3.4.11) can be written as

$$\frac{1}{\sqrt{2\pi}} \left(\omega_1(0) \sqrt{\Delta_t} \widehat{\phi}_{Y_M} \left(-\frac{1}{2} i \right) \right) + \epsilon_1^{(\ell)} + \epsilon_F^{(\ell)} + \epsilon_T^{(\ell)} = \frac{\tilde{\omega}_1}{\sqrt{2\pi}} \left(\sqrt{\Delta_t} \widehat{\phi}_{\tilde{Y}_M} \left(-\frac{1}{2} i \right) \right) + \epsilon_1^{(r)} + \epsilon_F^{(r)} + \epsilon_T^{(r)}$$

¹⁰In this step the error terms $\epsilon_1^{(\ell)}$ and $\epsilon_F^{(\ell)}$ are divided by $\sqrt{2\pi}$ and we keep the same notation for the 'new' error terms.

and we obtain

$$\tilde{\omega}_1 = \frac{\omega_1(0)\hat{\phi}_{Y_M}(-\frac{1}{2}i) + \epsilon_1^{(\ell)} + \epsilon_F^{(\ell)} + \epsilon_T^{(\ell)} - \epsilon_T^{(r)}}{\hat{\phi}_{\tilde{Y}_M}(-\frac{1}{2}i) + \epsilon_1^{(r)} + \epsilon_F^{(r)}} = \frac{\omega_1(0)\hat{\phi}_{Y_M}(-\frac{1}{2}i)}{\hat{\phi}_{\tilde{Y}_M}(-\frac{1}{2}i)} + \epsilon,$$

where ϵ represents the different error terms. By definition (3.3.3) we arrive at the result in equation (3.4.13). \square

Lemma 3.4.5. *For the integral approximation error $\epsilon_1^{(\ell)}$ in (3.4.12) it holds that: $\lim_{M \rightarrow \infty} \|\epsilon_1^{(\ell)}\|_{L^2}^2 = 0$.*

Proof. For the error $\bar{\epsilon}_1^{(\ell)}$ introduced in (3.4.14) we have:

$$\begin{aligned} \mathbb{E} \left| \bar{\epsilon}_1^{(\ell)} \right|^2 &= \mathbb{E} \left[\left| \int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt - \Delta_t \sum_{j=1}^M \omega_1^2(t_j) \sigma^2(t_j) \right|^2 \right] \\ &= \mathbb{E} \left[\left| \sum_{j=1}^M \int_{t_{j-1}}^{t_j} \{ \omega_1^2(t) \sigma^2(t) - \omega_1^2(t_j) \sigma^2(t_j) \} dt \right|^2 \right] \\ &\leq \sum_{j=1}^M \mathbb{E} \left[\int_{t_{j-1}}^{t_j} | \omega_1^2(t) \sigma^2(t) - \omega_1^2(t_j) \sigma^2(t_j) |^2 dt \right]. \end{aligned} \quad (3.4.20)$$

Convergence of the integral approximations is evident. An important result in stochastic calculus (see e.g. [78, 109]) states that for each fixed $T > 0$ and for any bounded, adapted and measurable process $X(\cdot)$, there exists a sequence $\{X^{(M)}(\cdot)\}_{M=1}^\infty$ of simple processes (which are, by definition, bounded¹¹) such that

$$\lim_{M \rightarrow \infty} \mathbb{E} \left[\int_0^T |X^{(M)}(t) - X(t)|^2 dt \right] = 0.$$

In our case, defining $X(t) := \omega_1^2(t) \sigma^2(t)$ and $X^{(M)}(t) := \sum_{j=1}^{M-1} \omega_1^2(t_j) \sigma^2(t_j) \mathbb{1}_{\{t_{j-1} \leq t < t_j\}}(t) + \omega_1^2(t_M) \sigma^2(t_M) \mathbb{1}_{\{t_{M-1} \leq t \leq t_M\}}(t)$ as a simple process approximating $X(t)$ on the interval $[0, t_M]$ with $t_M = T_i$ (both $X(\cdot)$ and $X^{(M)}(\cdot)$ satisfy the regular conditions), we have:

$$\begin{aligned} \lim_{M \rightarrow \infty} \mathbb{E} \left[\int_0^{T_i} \left| \sum_{j=1}^{M-1} \omega_1^2(t_j) \sigma^2(t_j) \mathbb{1}_{\{t_{j-1} \leq t < t_j\}}(t) + \omega_1^2(t_M) \sigma^2(t_M) \mathbb{1}_{\{t_{M-1} \leq t \leq t_M\}}(t) - \omega_1^2(t) \sigma^2(t) \right|^2 dt \right] &= 0 \\ \Leftrightarrow \lim_{M \rightarrow \infty} \mathbb{E} \left[\sum_{j=1}^M \int_{t_{j-1}}^{t_j} | \omega_1^2(t_j) \sigma^2(t_j) - \omega_1^2(t) \sigma^2(t) |^2 dt \right] &= 0 \end{aligned}$$

and thus

$$\lim_{M \rightarrow \infty} \sum_{j=1}^M \mathbb{E} \left[\int_{t_{j-1}}^{t_j} | \omega_1^2(t_j) \sigma^2(t_j) - \omega_1^2(t) \sigma^2(t) |^2 dt \right] = 0. \quad (3.4.21)$$

¹¹Besides the boundedness, the random variable in each piece of the simple process – $\sigma^2(t_j)$ in our case – is $\mathcal{F}(t_{j-1})$ -measurable.

Combining (3.4.20) and (3.4.21) yields

$$\lim_{M \rightarrow \infty} \left\| \bar{\epsilon}_1^{(\ell)} \right\|_{L^2}^2 := \lim_{M \rightarrow \infty} \mathbb{E} \left| \bar{\epsilon}_1^{(\ell)} \right|^2 = 0,$$

where $\|X\|_{L^2} := (\mathbb{E}|X|^2)^{1/2}$ is defined as the norm in L^2 space. As the error $\epsilon_1^{(\ell)}$ in (3.4.15) propagates from $\bar{\epsilon}_1^{(\ell)}$, we have $\lim_{M \rightarrow \infty} \left\| \epsilon_1^{(\ell)} \right\|_{L^2}^2 = 0$. \square

Remark 3.4.6 (Analysis of $\epsilon_F^{(\ell)}$). As stated by [132], three different types of errors are involved in the characteristic function recovery, namely a truncation error ϵ_t , an error of the Fourier cosine expansion ϵ_f and an error term introduced by applying the Clenshaw-Curtis quadrature ϵ_q . The truncation error is defined as

$$\epsilon_t(Y_M) := \int_{\mathbb{R} \setminus [a, b]} f_{Y_M}(y) dy, \quad (3.4.22)$$

with $f_{Y_M}(\cdot)$ denoting the probability density function of Y_M defined in (3.4.16). By definition (3.4.22), the truncation error decreases as the interval $[a, b]$ increases and the error is not dominant for a sufficiently large integration range. Further, from [38] we know that for a probability density function $f(y|x) \in \mathbb{C}^\infty[a, b]$ the error ϵ_f of the Fourier cosine expansion is bounded by

$$|\epsilon_f(N, [a, b])| \leq R(N) e^{-(N-1)\xi}, \quad (3.4.23)$$

where $\xi > 0$ is a constant and the term $R(N)$ is changing less than exponentially with respect to N , the number of Fourier cosine terms. So ϵ_f decays exponentially with respect to N , i.e. $\lim_{N \rightarrow \infty} \epsilon_f(N, [a, b]) = 0$. This error is related to $\epsilon_F^{(\ell)}$, the error in the recovered characteristic function $\hat{\phi}_{Y_M}(\cdot)$, as follows:

$$\left| \epsilon_F^{(\ell)} \right| = \mathcal{O} \left((M-1) (|\epsilon_f| + |\epsilon_q|) \right). \quad (3.4.24)$$

Equations (3.4.23) and (3.4.24) show that if the number of monitoring dates M increases, we need to increase the number of Fourier expansion terms N to compensate for this and to reach a specified level of accuracy, i.e. $\lim_{M, N \rightarrow \infty} \epsilon_F^{(\ell)} = 0$ (neglecting the error ϵ_q).

For the effective model we can also show that for the integral approximation error $\epsilon_1^{(r)}$ in (3.4.19) $\lim_{M \rightarrow \infty} \left\| \epsilon_1^{(r)} \right\|_{L^2}^2 = 0$ and $\lim_{M, N \rightarrow \infty} \epsilon_F^{(r)} = 0$.

Remark 3.4.7 (Level effect). The assumption of lognormal dynamics in (3.2.3)-(3.2.4) does not significantly affect the quality of the effective term structure for $\beta \neq 1$. The reason for this is the marginal ATM level effect of the CEV exponent, as we have defined model (3.2.3)-(3.2.4) such that $\bar{y}_0^{T_i} = 1$. We could already see this when discussing the smile effects of the skew parameter in Section 3.3. For the effective model we can quantify the ATM level effect of β on the base of Hagan's formula for ATM options [63]. Our numerical experiments in Section 3.4.2 confirm that the lognormality assumption still yields accurate results when assuming $\beta = 0.5$. Further, in the calibration procedure we add a non-parametric local volatility component that compensates for possible calibration inaccuracies introduced by the lognormality assumption.

Remark 3.4.8 (Alternative approaches). We have followed two alternative approaches to find the effective term structure. From the characteristic functions $\widehat{\phi}_{Y_M}(\cdot)$ and $\widehat{\phi}_{\widetilde{Y}_M}(\cdot)$ the corresponding probability density functions are derived. Successively, we compute approximations of $\mathbb{E}[g(x)]$ and $\mathbb{E}[g(\widetilde{x})]$ with x and \widetilde{x} as defined earlier, without approximating $g(\cdot)$. As a second approach, we have implemented Curran’s Asian option pricing method [29] and the enhanced “Curran 2M+” method [31, 84]. From a grid of Asian option prices with underlyings $\sum_{j=1}^M \omega_1^2(t_j)\sigma^2(t_j)/M$ and $\sum_{j=1}^M \widetilde{\sigma}^2(t_j)/M$ the corresponding densities are derived by the well-known relation $f(K, T; y_0) = \partial^2 C(\cdot)/\partial K^2$ [35] and we successively calculate approximations of $\mathbb{E}[g(x)]$ and $\mathbb{E}[g(\widetilde{x})]$. In both approaches $\widetilde{\omega}_1$ is determined by an optimization procedure in which $\mathbb{E}[g(x)]$ and $\mathbb{E}[g(\widetilde{x})]$ are repeatedly recalculated. The problem remains how to efficiently determine the effective term structure value. In contrast, the benefit of equation (3.4.11) is the closed-form approximation of $\widetilde{\omega}_1$, which yields an efficient evaluation and highly accurate results (see Section 3.4.2).

NUMERICAL EXPERIMENT

In this section we test the performance of the effective term structure $\widetilde{\omega}$ given by (3.4.13) for four cases. For all experiments $y_0 = 2$, $\beta = 0.5$, domestic and foreign interest rates are $r_d = 0.05$ and $r_f = 0.02$, respectively, and $\rho_{y,\sigma}(t) = 0$.

We first test the stand-alone performance of $\widetilde{\omega}$ in the Cases I, II and III (the Monte Carlo simulation consists of $5 \cdot 10^5$ paths and 200 time-steps per year). Time-dependent parameter values are given in Table 3.4.2. We only consider the last expiry $T = 5$. In Case

	t	$[0, \frac{1}{2}]$	$[\frac{1}{2}, 1]$	$[1, 2]$	$[2, 3]$	$[3, 5]$
Case I	$\gamma(t)$	0.7	0.7	0.7	0.7	0.7
	$\omega(t)$	0.1	0.12	0.14	0.16	0.18
Case II	$\gamma(t)$	0.3	0.3	0.3	0.3	0.3
	$\omega(t)$	0.1	0.12	0.14	0.16	0.18
Case III	$\gamma(t)$	0.3	0.3	0.3	0.3	0.3
	$\omega(t)$	0.3	0.36	0.42	0.48	0.54
Case IV	$\gamma(t)$	1	0.8	0.5	0.3	0.2
	$\omega(t)$	0.1	0.12	0.14	0.16	0.18

Table 3.4.2: Time-dependent parameter values

$\gamma(t)$ has a rather extreme value, especially for the larger times to maturity, whereas the $\omega(t)$ value is chosen to be moderate. For Case III the opposite holds. In Case II both $\gamma(t)$ and $\omega(t)$ have moderate values. In Case IV we test the combined performance of the effective vol-vol and effective term structure parameters across multiple expiries. In this case we assume parameter values which we typically observe in the FX markets.

The effective parameter values are given in Table 3.4.3. Results are displayed in Figures 3.4.2 and 3.4.3. The Cases II and IV yield a highly satisfactory fit. Results are slightly less accurate for a relatively large vol-vol parameter (Case I) or term structure (Case III). The reason for this is that the truncation error corresponding to the Taylor approximation of $\text{erf}(z)$ in (3.4.9) increases for larger z values, see Remark 3.4.5. This is a minor issue, as for typical FX markets $\omega(t)$ and – for relatively large expiries – $\gamma(t)$ assume moderate values, which are comparable with the values in Case IV.

Case	I	II	III	Case IV				
t	5	5	5	t	1	2	3	5
$\tilde{\gamma}$	0.679	0.316	0.316	$\tilde{\gamma}$	0.937	0.803	0.707	0.575
$\tilde{\omega}$	0.157	0.156	0.466	$\tilde{\omega}$	0.110	0.123	0.132	0.146

Table 3.4.3: Effective parameter values

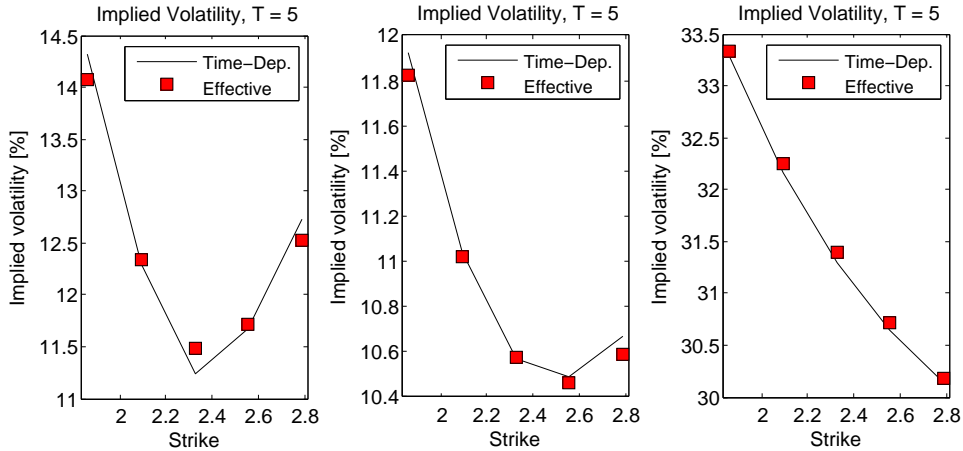


Figure 3.4.2: Stand-alone performance of the effective term structure parameter for Cases I, II and III, $T = 5$ ($\gamma(\cdot)$ is constant over time).

3.4.3. EFFECTIVE CORRELATION

At this point, we have determined the effective vol-vol and term structure parameters. As the numerical experiments in Section 3.4.2 show, the combined use of these effective parameters yields an accurate fit in both curvature and level. In the experiments we assumed zero correlation and set $\beta = 0.5$. By assuming these parameters to be constant, a time-dependent skew effect cannot be accounted for. To resolve this, we derive an effective correlation.

We obtain the effective correlation by considering the *vanna skew* of the SABR model. Based on an approximation of their main pricing formula ('Hagan's formula'), Hagan et al. [63] define the vanna skew (corresponding to the effective model (3.3.1)-(3.3.2) as¹²:

$$\tilde{v}(\tilde{\rho}_{y,\sigma}) := \frac{1}{2} \tilde{\rho}_{y,\sigma} \tilde{\lambda} \log\left(\frac{K}{\bar{y}_0^{T_i}}\right), \quad \tilde{\lambda} = \frac{\tilde{\gamma}}{\tilde{\omega}} (\bar{y}_0^{T_i})^{1-\beta}.$$

The vanna skew is the part of the skew which is caused by $\tilde{\rho}_{y,\sigma}$. The other part of the skew is mainly caused by the skew parameter β (see Remark 3.4.9). The parameter $\tilde{\lambda}$ measures the 'strength' of the vol-vol parameter $\tilde{\gamma}$ compared to the local volatility, $\tilde{\omega}/(\bar{y}_0^{T_i})^{1-\beta}$.

¹²We have defined model (3.2.3)-(3.2.4) such that $\bar{y}_0^{T_i} = 1$ and we can simplify the formulas. However, for the sake of completeness we include the term $\bar{y}_0^{T_i}$.

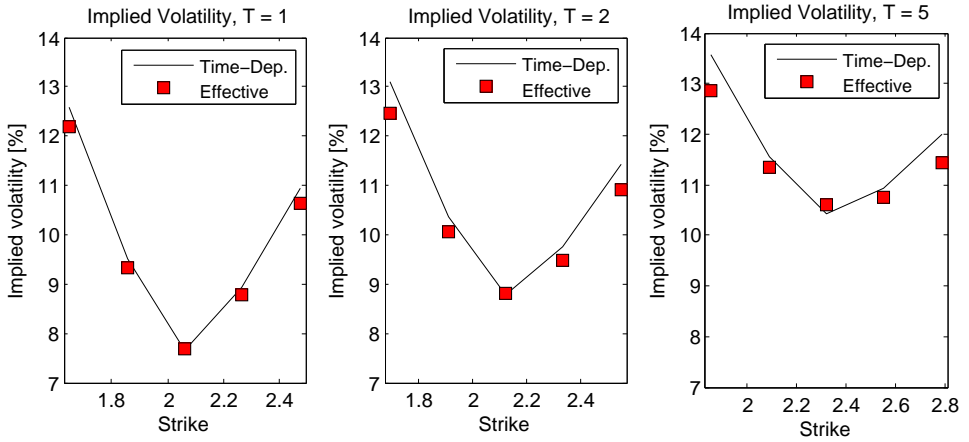


Figure 3.4.3: Combined performance of the effective vol-vol and term structure parameters for Case IV, $T = 1$ (left), $T = 2$ (middle) and $T = 5$ (right).

We define the effective correlation corresponding to the expiry T_i as:

$$\tilde{\rho}_{y,\sigma} := \arg \min_{\rho} \left(\tilde{v}(\rho) - \frac{1}{T_i} \int_0^{T_i} v(\rho_{y,\sigma}(t)) dt \right), \quad (3.4.25)$$

with

$$v(\rho_{y,\sigma}(t)) := \frac{1}{2} \rho_{y,\sigma}(t) \lambda(t) \log \left(\frac{K}{\bar{y}_0^{T_i}} \right), \quad \lambda(t) = \frac{\gamma(t)}{\omega(t)} (\bar{y}_0^{T_i})^{1-\beta}.$$

In other words, the effective correlation $\tilde{\rho}_{y,\sigma}$ is defined as the correlation value for which the *vanna skew* corresponding to the effective model at $t = T_i$ equals the average *vanna skew* corresponding to the time-dependent model over $[0, T_i]$.

Lemma 3.4.6 (Effective correlation). *Suppose that the effective term structure $\tilde{\omega}$ and effective vol-vol parameter $\tilde{\gamma}$ have been established. From definition (3.4.25) it follows that the effective correlation corresponding to the expiry T_i is given by:*

$$\tilde{\rho}_{y,\sigma} = \frac{\tilde{\omega}}{\tilde{\gamma} T_i} \int_0^{T_i} \frac{\rho_{y,\sigma}(t) \gamma(t)}{\omega(t)} dt. \quad (3.4.26)$$

Proof. By its definition (3.4.25), we obtain the effective correlation by solving

$$\begin{aligned} \tilde{v}(\tilde{\rho}_{y,\sigma}) &= \frac{1}{T_i} \int_0^{T_i} v(\rho_{y,\sigma}(t)) dt \\ \Leftrightarrow \frac{1}{2} \tilde{\rho}_{y,\sigma} \tilde{\lambda} \log \left(\frac{K}{\bar{y}_0^{T_i}} \right) &= \frac{1}{2 T_i} \log \left(\frac{K}{\bar{y}_0^{T_i}} \right) \int_0^{T_i} \rho_{y,\sigma}(t) \lambda(t) dt, \end{aligned} \quad (3.4.27)$$

with

$$\lambda(t) = \frac{\gamma(t)}{\omega(t)} (\bar{y}_0^{T_i})^{1-\beta}.$$

Equation (3.4.27) yields

$$\tilde{\rho}_{y,\sigma} = \frac{1}{\tilde{\lambda} T_i} \int_0^{T_i} \rho_{y,\sigma}(t) \lambda(t) dt = \frac{\tilde{\omega}}{\tilde{\gamma} T_i} \int_0^{T_i} \frac{\rho_{y,\sigma}(t) \gamma(t)}{\omega(t)} dt.$$

□

The effective correlation parameter is independent of the initial forward and β .

Remark 3.4.9 (Effective skew parameter). *Hagan et al. [63] determine $-\frac{1}{2}(1-\beta) \log\left(K/\bar{y}_0^{T_i}\right)$ as being the skew implied by the skew parameter. In a similar fashion as for the effective correlation, we may derive an effective skew parameter. We do not include the derivations though, as we do not assume a time-dependent skew parameter.*

NUMERICAL EXPERIMENT

In this section we test the combined performance of the effective vol-vol parameter $\tilde{\gamma}$, effective term structure $\tilde{\omega}$ and effective correlation $\tilde{\rho}_{y,\sigma}$ (the Monte Carlo simulation consists of $5 \cdot 10^5$ paths and 200 time-steps per year).

Let $y_0 = 2$, $\beta = 0.5$, $r_d = 0.05$ and $r_f = 0.02$. Time-dependent and effective parameter values are provided by Tables 3.4.4 and 3.4.5, respectively. Results are highly satisfactory, see Figure 3.4.4.

t	$[0, \frac{1}{2}]$	$[\frac{1}{2}, 1]$	$[1, 2]$	$[2, 3]$	$[3, 5]$
$\gamma(t)$	1	0.8	0.5	0.3	0.2
$\omega(t)$	0.1	0.12	0.14	0.16	0.18
$\rho_{y,\sigma}(t)$	-0.9	-0.8	-0.7	-0.6	-0.5

Table 3.4.4: Time-dependent parameter values

3.5. CALIBRATION & PRICING

In this section we calibrate the time-dependent FX-SABR model to market data. We compare the calibration results for this model and the constant-parameter FX-SABR and local volatility models. We perform three calibration experiments. In each experiment we calibrate the time-dependent model to 2 expiries. Subsequently, in Section 3.5.3, we price barrier options with corresponding times to maturity.

We consider USD/AUD FX market prices quoted on 12 June 2013 from a market data vendor. Domestic currency is USD, foreign currency is AUD. Initial spot is $y_0 = 0.9548$. Implied volatilities are quoted for 5 different strikes. The third strike corresponds to the

t	$\frac{1}{2}$	1	2	3	5
$\tilde{\gamma}$	1.000	0.937	0.803	0.707	0.575
$\tilde{\omega}$	0.100	0.110	0.123	0.132	0.146
$\tilde{\rho}_{y,\sigma}$	-0.900	-0.840	-0.739	-0.673	-0.606

Table 3.4.5: Effective parameter values

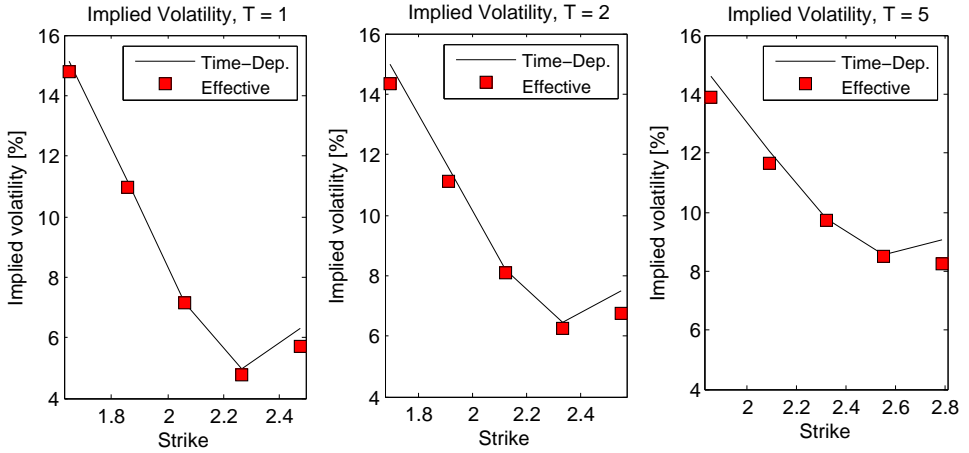


Figure 3.4.4: Combined performance of the effective vol-vol parameter, effective term structure and effective correlation for $T = 1$ (left), $T = 2$ (middle) and $T = 5$ (right).

ATM level. Data is provided for expiries between 1 day and 5 years. The ATM term structure exhibits a downward trend up to 1 year. For longer expiries the ATM level moves in the opposite direction. Also, an increasingly pronounced skew can be observed over time, whereas the amount of curvature slightly declines. Further, the market data consists of mid-prices and bid-ask spreads of up-out barrier put options with expiries of 3 months, 1 year and 2 years.

Remark 3.5.1 (Deterministic interest rates). *In Section 3.2 we introduced the time-dependent FX-SABR model with deterministic interest rates. The deterministic interest rates assumption merely serves the purpose of transparency and readability. Assuming stochastic interest rates would complicate notation. The mappings of the effective parameters are not affected by the deterministic interest rates assumption. As our numerical experiments in this section confirm, the effective parameters facilitate accurate calibration to the FX market, also for longer times to maturity when stochastic interest rate effects become more important.*

3.5.1. CALIBRATION PROCEDURE

The (implicit) mappings derived for effective parameters imply the following functional dependencies:

$$\tilde{\gamma}^{\text{mod}} = f_1(\gamma(t), \omega(t)), \tilde{\omega}^{\text{mod}} = f_2(\gamma(t), \omega(t), \tilde{\gamma}^{\text{mod}}), \tilde{\rho}_{y,\sigma}^{\text{mod}} = f_3(\gamma(t), \omega(t), \rho_{y,\sigma}(t), \tilde{\gamma}^{\text{mod}}, \tilde{\omega}^{\text{mod}}),$$

where $\gamma(t)$, $\omega(t)$ and $\rho_{y,\sigma}(t)$ are time-dependent parameters in the time-dependent model (3.2.1)-(3.2.2) and $\tilde{\gamma}^{\text{mod}}$, $\tilde{\omega}^{\text{mod}}$ and $\tilde{\rho}_{y,\sigma}^{\text{mod}}$ are their ‘effective equivalents’. Similar as in the experiments in Sections 3.4.1, 3.4.2 and 3.4.3, we assume that the time-dependent parameters are piecewise-constant.

The calibration consists of four stages. First, we calibrate the effective SABR model

(3.3.1)-(3.3.2) using Hagan's formulas¹³. For each expiry we obtain a set of *market effective* parameters $\{\tilde{\gamma}^{\text{mar}}, \tilde{\omega}^{\text{mar}}, \tilde{\rho}_{y,\sigma}^{\text{mar}}\}$ – see also Section 3.3. Secondly, we calibrate the time-dependent SABR model (3.2.1)-(3.2.2). As $\gamma(t)$ and $\rho_{y,\sigma}(t)$ both have a curvature (and skew) effect, we calibrate these parameters simultaneously on the base of the implied volatilities. As at this point $\omega(t)$ is not established yet, we apply the approximations

$$f_1(\gamma(t), \omega(t)) \approx f_1(\gamma(t), \tilde{\omega}^{\text{mar}}), \quad (3.5.1)$$

$$f_3(\gamma(t), \omega(t), \rho_{y,\sigma}(t), \tilde{\gamma}^{\text{mod}}, \tilde{\omega}^{\text{mod}}) \approx f_3(\gamma(t), \rho_{y,\sigma}(t), \tilde{\gamma}^{\text{mod}}, \tilde{\omega}^{\text{mar}}). \quad (3.5.2)$$

Successively, we calibrate $\omega(t)$ in order to obtain a level fit. Thirdly, to compensate for the approximation (3.5.2) in the effective correlation mapping $f_3(\cdot)$, we perform another calibration iteration for $\rho_{y,\sigma}(t)$ separately, given the values for $\gamma(t)$ and $\omega(t)$ obtained in the second calibration stage. We observe that an additional calibration iteration for $\gamma(t)$ (to compensate for the $f_1(\cdot)$ mapping approximation (3.5.1)) does not yield a significant improvement in results. In the fourth stage, we add the local volatility component $\sigma_{\text{SLV}}(\cdot)$ specified in Section 3.2.2 on top of the time-dependent SABR model to compensate for calibration inaccuracies. We summarize the calibration routine in Algorithm 2.

Calibration routine:

- 1 Calibrate the effective SABR model (3.3.1)-(3.3.2). This yields market effective parameters $\{\tilde{\gamma}_i^{\text{mar}}, \tilde{\omega}_i^{\text{mar}}, \tilde{\rho}_{y,\sigma,i}^{\text{mar}}\}$, $i = 1, \dots, N$.
- 2 Calibrate $\rho_{y,\sigma}(t)$ and $\gamma(t)$ simultaneously. For this the mapping approximations (3.5.1) and (3.5.2) are required. Successively, calibrate $\omega(t)$ separately in order to obtain a level fit.
- 3 Calibrate $\rho_{y,\sigma}(t)$ separately, given the values for $\omega(t)$ and $\gamma(t)$ just obtained. Now the original mapping on the left-hand side of (3.5.2) can be applied.
- 4 Add the local volatility component $\sigma_{\text{SLV}}(\cdot)$ to compensate for calibration inaccuracies.

Algorithm 2: Calibration procedure

3.5.2. CALIBRATION RESULTS

In Tables 3.5.1-3.5.3 the calibration errors are given, defined as $\sigma_{i,\text{mod}} - \sigma_{i,\text{mar}}$, where $\sigma_{i,\text{mod}}$ and $\sigma_{i,\text{mar}}$ are the corresponding model and market implied volatilities in percentages, respectively (i denotes the strike). The total absolute error $\varepsilon_{\text{tot}} := \sum_{i=1}^5 |\sigma_{i,\text{mod}} - \sigma_{i,\text{mar}}|$ is also provided. The results (also for the constant-parameter SABR model) are obtained by Monte Carlo simulation runs consisting of $5 \cdot 10^5$ paths and 1000 time-steps per year for Experiment I and 200 time-steps per year for Experiments II and III.

For the local volatility model and the time-dependent model with the local volatility component the calibration results are most accurate. Comparing the calibration errors of the time-dependent model with and without the local volatility component, we

¹³It is well-known that Hagan's formulas are biased for extreme strikes and large times to maturity, see e.g. [71, 91, 94, 113]. In [7] it is pointed out that for maturities larger than 10 years the error in implied volatility can be 1% or more, even for ATM values. However, in this chapter no bias as a result of Hagan's formulas is introduced. We have confirmed this for the effective model by assuming zero correlation and applying the analytical pricing method of Antonov et al. [7]. We obtain the same ATM level, which implies that our results are free of bias due to Hagan's formula or the Monte Carlo simulation (correlation does not affect smile level, see Section 3.3).

observe that the local volatility component reduces the calibration error. By its construction, it naturally ‘bridges’ the mismatch between the market and (time-dependent FX-SABR) model prices (see e.g. [118]). Further, except for Experiment I, the constant-parameter and time-dependent SABR models overall yield a similar calibration accuracy for the second expiry date. For the constant-parameter SABR model the calibration error is substantial for the intermediate expiry, as we calibrate this model to the last expiry date. For illustration purposes the implied volatility smiles corresponding to Experiment III are displayed in Figure 3.5.1. The calibration of the time-dependent model is accurate. Inclusion of the non-parametric local volatility component yields a highly satisfactory fit to the market. Similar figures were obtained for Experiments I and II.

Expiry	Strike	Time-dep.+LV	Time-dep.	LV	Constant
2M	0.87	0.031	0.276	0.109	-0.802
	0.91	-0.037	-0.011	-0.011	-0.655
	0.95	0.015	0.042	0.035	-0.431
	0.98	0.016	0.289	0.040	-0.347
	1.01	-0.026	0.616	0.020	-0.404
	ϵ_{tot}	0.125	1.234	0.215	2.638
3M	0.85	-0.050	0.346	0.110	-0.195
	0.91	-0.100	-0.033	-0.017	-0.164
	0.95	-0.025	0.012	0.035	0.004
	0.99	-0.015	0.361	0.044	0.095
	1.02	-0.081	0.821	0.022	0.089
	ϵ_{tot}	0.272	1.572	0.228	0.548

Table 3.5.1: Calibration errors in Experiment I (in %). $\epsilon_{\text{tot}} := \sum_{i=1}^5 |\sigma_{i,\text{mod}} - \sigma_{i,\text{mar}}|$, where i indicates strike.

Expiry	Strike	Time-dep.+LV	Time-dep.	LV	Constant
6M	0.81	-0.029	-0.261	0.166	-1.303
	0.89	-0.084	-0.208	-0.003	-0.959
	0.94	-0.013	0.047	0.037	-0.476
	1.00	0.039	0.301	0.091	-0.213
	1.04	-0.080	0.314	-0.002	-0.424
	ϵ_{tot}	0.246	1.131	0.298	3.374
1Y	0.76	-0.088	0.222	0.160	-0.443
	0.86	-0.167	-0.105	-0.063	-0.325
	0.93	-0.046	-0.038	0.028	-0.032
	1.00	0.035	0.246	0.113	0.171
	1.07	-0.127	0.409	-0.032	0.016
	ϵ_{tot}	0.463	1.019	0.396	0.986

Table 3.5.2: Calibration errors in Experiment II (in %). $\epsilon_{\text{tot}} := \sum_{i=1}^5 |\sigma_{i,\text{mod}} - \sigma_{i,\text{mar}}|$, where i indicates strike.

Calibrated parameter values are provided by Table 3.5.4. The time-dependent vol parameter decreases over time, which implies that curvature in the model implied

Expiry	Strike	Time-dep.+LV	Time-dep.	LV	Constant
1Y	0.76	0.002	-0.769	0.153	-1.638
	0.86	-0.098	-0.406	-0.066	-0.674
	0.93	0.020	0.059	0.025	0.114
	1.00	0.102	0.291	0.102	0.432
	1.07	-0.062	-0.003	-0.047	-0.039
	ε_{tot}	0.285	1.529	0.393	2.898
2Y	0.69	0.014	0.314	0.155	-0.410
	0.82	-0.100	-0.060	-0.068	-0.326
	0.91	0.029	-0.024	0.038	-0.047
	1.01	0.144	0.230	0.140	0.173
	1.11	-0.030	0.362	-0.038	-0.007
	ε_{tot}	0.317	0.990	0.439	0.965

Table 3.5.3: Calibration errors in Experiment III (in %). $\varepsilon_{\text{tot}} := \sum_{i=1}^5 |\sigma_{i,\text{mod}} - \sigma_{i,\text{mar}}|$, where i indicates strike.

volatility surface declines. This effect was also observed in the market data. Further, the term structure parameter decreases in Experiments I and II and increases in Experiment III, which is in line with the observation in the market of a decreasing term structure up to 1 year and an increase for longer expiries. Further, the correlation becomes more negative, which implies a more pronounced skew effect implied by the model. The market surface exhibits the same feature.

Exper.	γ_1	γ_2	ω_1	ω_2	ρ_1	ρ_2
I	1.597	0.468	0.117	0.107	-0.389	-0.996
II	1.008	0.468	0.110	0.100	-0.440	-0.611
III	0.695	0.203	0.106	0.111	-0.456	-0.944

Table 3.5.4: Calibrated piecewise-constant parameter values

Remark 3.5.2 (Multiple expiries). *We only show and discuss the barrier option pricing results for the model that we calibrated to 2 expiries in Experiments I, II and III. The reason for this is the following. For this particular dataset the improvement in the pricing of barrier options (see the follow-up section) turns out to be marginal when we calibrate the model to more than 2 expiries. This is a property of this specific dataset and not a model property: for this dataset we do not need to calibrate to more than 2 expiries, although the model – by means of the effective parameters derived – can be calibrated accurately to multiple expiries if necessary. In Appendix 3.B we calibrate to 4 expiries, namely 1, 2, 3 and 4 years.*

As an indication regarding CPU times, we calibrate the piecewise-constant $\omega(t)$, $\gamma(t)$ and $\rho_{y,\sigma}(t)$ for Experiment III in approximately 14 seconds (in Matlab on an i5-2400 CPU @3.10GHz, 3101Mhz, 4 Cores, 4 Logical Processors).

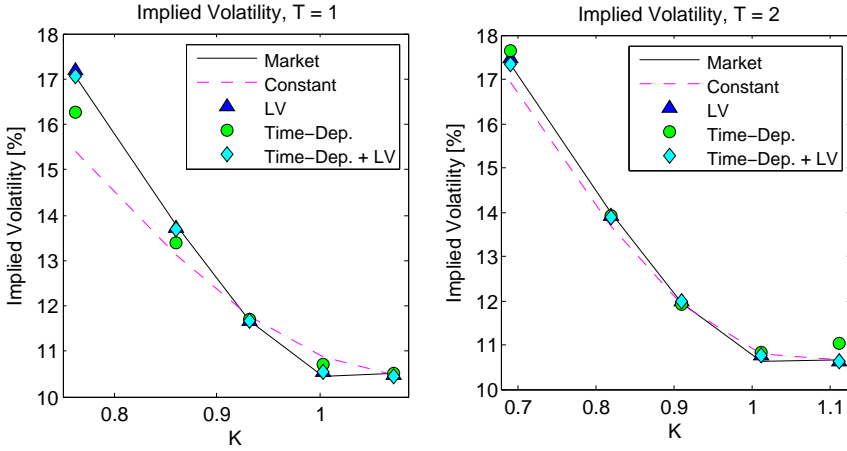


Figure 3.5.1: Black-Scholes implied volatility smiles for Experiment III: 1 year (left) and 2 years (right).

3.5.3. PRICING BARRIER OPTIONS

In this section we price barrier options by the local volatility model, the constant-parameter SABR model and the time-dependent SABR model with and without local volatility component. In contrast to European-type options, the prices of barrier options do not only depend on the distribution of the underlying at the time to maturity, but also on the dynamics of the implied volatility smile through time [115]. The prices are determined by the underlying’s future transition densities, which are reflected by the forward implied volatility smiles a particular model produces [10]. The dynamics of the implied volatility smile with respect to the underlying are also relevant, especially for hedging purposes.

We price up-out barrier put options with different strikes and barriers. The value of an up-out barrier put option is given as follows. Define B as the barrier level and assume for the initial spot FX rate $y_0 < B$. The discounted value of an up-out put option with strike K and expiry T at an arbitrary time $t \in [0, T]$ reads

$$P_{\text{UO}}(t, T, K) := \mathbb{E}^{\mathbb{Q}} \left[\frac{M_d(t)}{M_d(T)} (K - y(T))^+ \mathbb{1}_{(\max_{t \in [0, T]} y(t) < B)} \middle| \mathcal{F}(t) \right],$$

where $M_d(\cdot)$ is the domestic money account determined by $dM_d(t) = r_d(t)M_d(t)dt$, with $r_d(\cdot)$ denoting the domestic interest rate. In the numerical experiments we assume $K < y_0 < B$, so each option starts out-of-the-money. Pricing results are provided in Table 3.5.5.

In each experiment the local volatility model gives rise to significantly higher prices compared to the other stochastic volatility models: 6 of 9 prices are higher than the ask price observed in the market. Each of the models with stochastic volatility only yields one price that is not within the bid-ask spread. Also, the errors corresponding to the stochastic volatility models are significantly smaller. This suggests that the market consensus seems to price barrier options with a ‘SABR-like’ stochastic volatility model, instead of the local volatility model. A plausible reason for this is the fact that the latter

Exper.	Expiry	Barrier	Strike	LV	Constant	Time-dep.	Time-dep.+LV
I	3M	1	0.9	97.53	92.76	91.12	90.39
	3M	1	0.85	37.41	34.51	36.50	33.08
	3M	0.97	0.9	(70.12)	62.91	60.35	59.54
II	1Y	1	0.8	(114.17)	92.71	95.40	93.08
	1Y	1	0.85	(178.39)	(153.26)	(152.67)	(152.11)
	1Y	0.97	0.8	(72.05)	54.19	55.48	54.55
III	2Y	1	0.75	(165.25)	129.96	135.67	134.24
	2Y	1.05	0.7	140.48	117.44	127.46	121.33
	2Y	0.97	0.75	(91.82)	69.14	72.16	72.08
ϵ_{tot}				146.00	51.43	38.97	46.53

Table 3.5.5: Model prices up-out put options multiplied by a factor 10^4 . Results in brackets are *not* within bid-ask spread. The error is defined as $\epsilon_{\text{tot}} := \sum_{i=1}^9 |B_{i,\text{mod}} - B_{i,\text{mar}}|$, where i indicates the particular barrier option and $B_{i,\text{mod}}$ and $B_{i,\text{mar}}$ are the corresponding model and mid-market prices, respectively.

typically does not accurately ‘capture’ the forward implied volatility smile [37, 46]. As we mentioned at the beginning of this section, e.g. in [115] the relevance of the smile dynamics for the pricing of path-dependent derivatives is pointed out. Besides this, the predicted implied volatility smile moves with respect to the underlying are also relevant, especially for the hedging of complex products [63, 76]¹⁴.

To price the up-out put option correctly, the model should not only imply the correct smile dynamics. It should also price plain vanilla options accurately across a grid of multiple expiries, as these provide information about the market behavior over time. In Tables 3.5.1-3.5.3 we observed that the constant-parameter SABR model is only well-calibrated to the last expiry. This is trivial, as the calibrated parameter values only correspond to this maturity. By assuming time-dependent parameters the calibration error at an intermediate expiry date is reduced significantly: the plain vanilla options are priced more accurately across multiple expiries. We observe that the time-dependent model yields prices which are closer to the reference compared to the constant-parameter SABR model.

By incorporating the non-parametric local volatility component the time-dependent model yields less accurate up-out put option prices. This may be due to the fact that the local volatility model predicts inaccurate smile dynamics. However, the calibration error reduces significantly (see Tables 3.5.1-3.5.3). We therefore tend to prefer the model with the local volatility component, as it may reduce the additional costs of hedging the barrier option due to a better calibration performance.

Last, the results indicate that for each model and expiry, the up-out put option prices

¹⁴Barrier options can be priced and hedged in a dynamic or (quasi)-static way. Dynamic hedging models price an exotic option on the base of the costs of dynamically hedging the product with a portfolio of the underlying asset and European-type options (in an analogous way the Black-Scholes model prices vanilla European options based on the costs of a time-dependent hedge with the underlying asset). In this approach the hedging risks and costs may be substantial, as pointed out by e.g. Derman, Ergener and Kani [33]. As an alternative, they adopt a static hedging model, which values the barrier option based on the cost of a replication strategy which requires an unchanging hedge portfolio consisting of European-type options.

are consistent across different barriers and strikes. More specifically, each model yields a higher up-out put option price for a higher barrier and/or higher strike. This is intuitive, as a higher barrier implies a smaller probability that the barrier is reached and that the option becomes worthless. Also, a higher strike means that the underlying can be sold for a larger amount of money.

3.6. CONCLUSION

In this chapter we have presented a framework for accurate and efficient calibration of the time-dependent SABR model in an FX context. By considering the effects of the SABR parameters on the shape of the implied volatility smile, we have derived 'effective equivalents' of the time-dependent vol-vol, term structure and correlation parameters. Numerical experiments show that both the separate and combined performance of the effective parameters are accurate, which results into highly satisfactory calibration results. A non-parametric local volatility component can compensate for the calibration inaccuracies. In our barrier option pricing experiments, the time-dependent FX-SABR model outperforms the traditional local volatility and constant-parameter SABR models. Our results seem to indicate that there is a market consensus of pricing barrier options by a 'SABR-like' stochastic volatility model. Plausible reasons are the facts that the local volatility model typically does not accurately 'capture' the forward implied volatility smile and the implied volatility smile moves with respect to the underlying, which are relevant features for the accurate pricing and hedging of path-dependent derivatives.

APPENDIX

3.A. CHARACTERISTIC FUNCTION RECOVERY

Define

$$R_j := \log \left(\frac{\omega_1^2(t_j)\sigma^2(t_j)}{\omega_1^2(t_{j-1})\sigma^2(t_{j-1})} \right), \quad j = 1, \dots, M, \quad (3.A.1)$$

and stochastic process Y_j , which is given by

$$Y_j := \log \left(\frac{\omega_1^2(t_{M-j+1})\sigma^2(t_{M-j+1})}{\omega_1^2(t_{M-j})\sigma^2(t_{M-j})} + \frac{\omega_1^2(t_{M-j+2})\sigma^2(t_{M-j+2})}{\omega_1^2(t_{M-j})\sigma^2(t_{M-j})} + \dots + \frac{\omega_1^2(t_M)\sigma^2(t_M)}{\omega_1^2(t_{M-j})\sigma^2(t_{M-j})} \right). \quad (3.A.2)$$

It is easy to see that

$$Y_1 = \log \left(\frac{\omega_1^2(t_M)\sigma^2(t_M)}{\omega_1^2(t_{M-1})\sigma^2(t_{M-1})} \right) = R_M$$

and

$$Y_M = \log \left(\sum_{j=1}^M \frac{\omega_1^2(t_j)\sigma^2(t_j)}{\omega_1^2(0)\sigma_0^2} \right) = \log \left(\sum_{j=1}^M \frac{\omega_1^2(t_j)\sigma^2(t_j)}{\omega_1^2(0)} \right) \quad (3.A.3)$$

as $\sigma_0 = 1$. To recover the approximated characteristic function $\hat{\phi}_{Y_M}(\cdot)$ corresponding to Y_M , we need to know the distribution of R_j , $j = 1, \dots, M$ in (3.A.1).

3.A.1. DISTRIBUTION OF R_j

We determine the distribution of R_j , $j = 1, \dots, M$ defined in (3.A.1) by first deriving the dynamics of $Z(t) := \omega_1^2(t)\sigma^2(t)$. As the $\sigma^2(t)$ -dynamics are given by

$$d\sigma^2(t) = \sigma^2(t) (\gamma^2(t)dt + 2\gamma(t)dW_\sigma(t)),$$

we obtain:

$$\begin{aligned} dZ(t) &= [d\omega_1^2(t)]\sigma^2(t) + \omega_1^2(t)d\sigma^2(t) + d\omega_1^2(t)d\sigma^2(t) \\ &= \left(\gamma^2(t) + \frac{d\omega_1^2(t)}{dt} \frac{1}{\omega_1^2(t)} \right) Z(t)dt + 2\gamma(t)Z(t)dW_\sigma(t). \end{aligned}$$

The solution then leads

$$Z(t) = Z_0 \exp \left(\int_0^t -\gamma^2(s)ds + 2 \int_0^t \gamma(s)dW_\sigma(s) \right) \exp \left(\int_0^t \frac{d\omega_1^2(s)}{ds} \frac{1}{\omega_1^2(s)} ds \right).$$

The integral in the second exponential term is simplified to:

$$\int_0^t \frac{d\omega_1^2(s)}{ds} \frac{1}{\omega_1^2(s)} ds = \int_0^t \frac{d\omega_1^2(s)}{\omega_1^2(s)} = [\log(\omega_1^2(s))]_{s=0}^{s=t} = \log \left(\frac{\omega_1^2(t)}{\omega_1^2(0)} \right)$$

and the solution becomes¹⁵:

$$\begin{aligned} Z(t) &= Z_0 \exp\left(-\int_0^t \gamma^2(s) ds + 2 \int_0^t \gamma(s) dW_\sigma(s)\right) \frac{\omega_1^2(t)}{\omega_1^2(0)} \\ &= \omega_1^2(t) \exp\left(-\int_0^t \gamma^2(s) ds + 2 \int_0^t \gamma(s) dW_\sigma(s)\right). \end{aligned} \quad (3.A.4)$$

Thus, we have:

$$\begin{aligned} R_j &= \log \left[\frac{\omega_1^2(t_j) \exp\left(-\int_0^{t_j} \gamma^2(s) ds + 2 \int_0^{t_j} \gamma(s) dW_\sigma(s)\right)}{\omega_1^2(t_{j-1}) \exp\left(-\int_0^{t_{j-1}} \gamma^2(s) ds + 2 \int_0^{t_{j-1}} \gamma(s) dW_\sigma(s)\right)} \right] \\ &= \log\left(\frac{\omega_1^2(t_j)}{\omega_1^2(t_{j-1})}\right) - \int_{t_{j-1}}^{t_j} \gamma^2(s) ds + 2 \int_{t_{j-1}}^{t_j} \gamma(s) dW_\sigma(s). \end{aligned}$$

So R_j is distributed as follows:

$$R_j \sim N(\mu_{R,j}, \sigma_{R,j}^2), \quad \mu_{R,j} = \log\left(\frac{\omega_1^2(t_j)}{\omega_1^2(t_{j-1})}\right) - \int_{t_{j-1}}^{t_j} \gamma^2(s) ds, \quad \sigma_{R,j}^2 = 4 \int_{t_{j-1}}^{t_j} \gamma^2(s) ds.$$

Hence, the characteristic function of R_j is specified by

$$\phi_{R_j}(u) = e^{iu\mu_{R,j} - \frac{1}{2}u^2\sigma_{R,j}^2}, \quad (3.A.5)$$

with $\mu_{R,j}$ and $\sigma_{R,j}^2$ given above.

3.A.2. RECOVERY PROCEDURE

Given the characteristic function (3.A.5), we recover the characteristic function of Y_M as follows. Let $k = 0, 1, \dots, N-1$ and $l = 0, 1, \dots, N-1$ and define $u_k := \frac{k\pi}{b-a}$ and $u_l = \frac{l\pi}{b-a}$. We follow a recursion procedure, *which starts at the end of the time-interval*. By the definition of Y_j in (3.A.2) and R_j in (3.A.1), we have

$$Y_1 = \log\left(\frac{\omega_1^2(t_M)\sigma^2(t_M)}{\omega_1^2(t_{M-1})\sigma^2(t_{M-1})}\right) = R_M.$$

We now proceed according to the following recovery procedure:

1. As $Y_1 = R_M$, we have $\phi_{Y_1}(u_k) = \phi_{R_M}(u_k)$.
2. If we go one step further, we obtain $\phi_{Y_2}(u_k) = \phi_{R_{M-1}}(u_k)\phi_{Z_1}(u_k) \approx \phi_{R_{M-1}}(u_k)\widehat{\phi}_{Z_1}(u_k) =: \widehat{\phi}_{Y_2}(u_k)$, with

$$\widehat{\phi}_{Z_1}(u_k) = \frac{2}{b-a} \sum_{l=0}^{N-1} \operatorname{Re} \left\{ \phi_{Y_1}(u_l) e^{-ia u_l} \right\} \int_a^b (e^x + 1)^{i u_k} \cos((x-a)u_l) dx.$$

¹⁵Equation (3.A.4) makes sense, as $Z(t)/\omega_1^2(t) = \exp\left(\int_0^t -\gamma^2(s) ds + 2 \int_0^t \gamma(s) dW_\sigma(s)\right)$ or equivalently $\sigma^2(t) = \exp\left(-\int_0^t \gamma^2(s) ds + 2 \int_0^t \gamma(s) dW_\sigma(s)\right)$.

3. In the next iteration we obtain $\widehat{\phi}_{Y_3}(u_k) = \phi_{R_{M-2}}(u_k)\widehat{\phi}_{Z_2}(u_k)$, with

$$\widehat{\phi}_{Z_2}(u_k) = \frac{2}{b-a} \sum_{l=0}^{N-1} \operatorname{Re} \left\{ \widehat{\phi}_{Y_2}(u_l) e^{-ia u_l} \right\} \int_a^b (e^x + 1)^{i u_k} \cos((x-a)u_l) dx.$$

4. We continue in this way until we have obtained $\widehat{\phi}_{Y_{M-1}}(u_k)$.

In the last, M th step, we set $u = -\frac{1}{2}i$ and we calculate

$$\widehat{\phi}_{Y_M}(u) = \phi_{R_1}(u)\widehat{\phi}_{Z_{M-1}}(u),$$

with

$$\widehat{\phi}_{Z_{M-1}}(u) = \frac{2}{b-a} \sum_{l=0}^{N-1} \operatorname{Re} \left\{ \widehat{\phi}_{Y_{M-1}}(u_l) e^{-ia u_l} \right\} \int_a^b (e^x + 1)^{i u} \cos((x-a)u_l) dx.$$

In the recursion, the integral $\int_a^b (e^x + 1)^{i u_k} \cos((x-a)u_l) dx$, $k, l = 0, \dots, N-1$ has to be calculated only once (except for the M th step, where $u_k = u$). One can either approximate it numerically by the Clenshaw-Curtis quadrature rule, which is based on an expansion of the integrand in terms of Chebyshev polynomials, or evaluate the integrals numerically. The Clenshaw-Curtis quadrature rule is described in detail by [132].

3.B. ADDITIONAL CALIBRATION EXPERIMENT

For the dataset described in Section 3.5 we perform an additional calibration experiment. We calibrate the time-dependent FX-SABR model to expiries 1, 2, 3 and 4 years. Results are provided in Figure 3.B.1. Except for the first expiry, the calibration of the time-dependent model is accurate. Also for the first expiry the improvement of the time-dependent model compared to the constant-parameter model is significant.

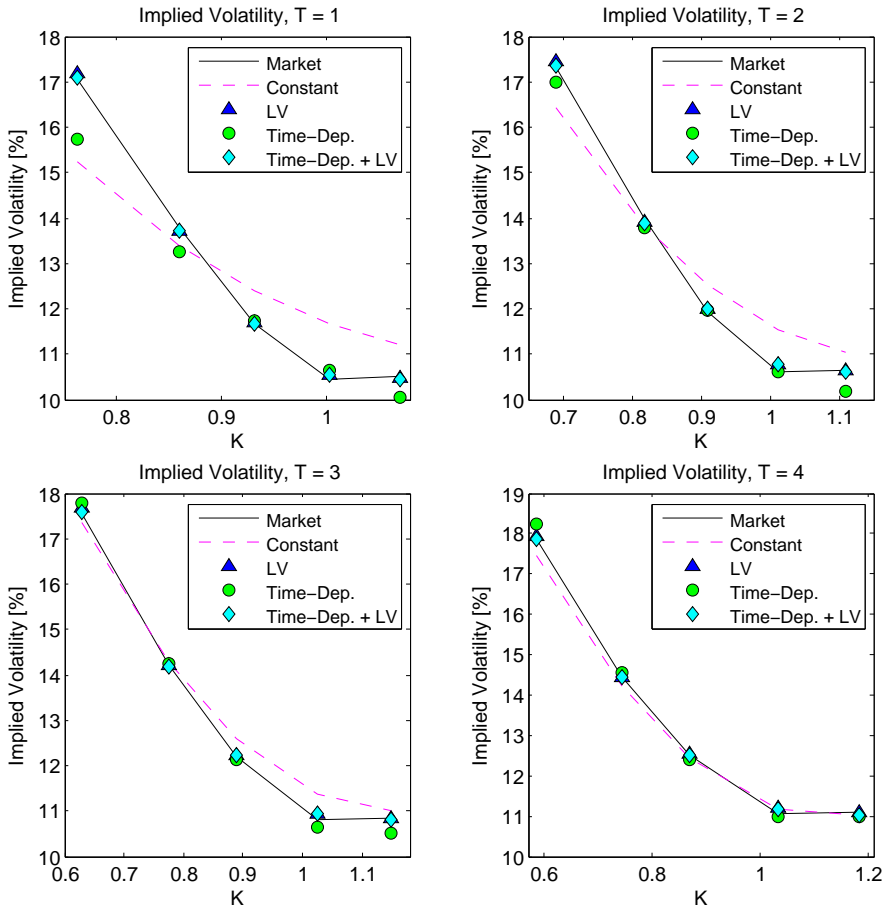


Figure 3.B.1: Black-Scholes implied volatility smiles for 1, 2, 3 and 4 years.

4

A NOVEL MONTE CARLO APPROACH TO HYBRID LOCAL VOLATILITY MODELS

We present in a Monte Carlo simulation framework a novel approach for the evaluation of hybrid local volatility [34, 35] models. In particular, we consider the stochastic local volatility model – see e.g. [80, 81, 98, 115] and Chapter 2 – and the local volatility model incorporating stochastic interest rates – see e.g. [8, 32, 97, 104]. For both model classes, a particular (conditional) expectation needs to be evaluated, which cannot be extracted from the market and is expensive to compute. We establish accurate and ‘cheap to evaluate’ approximations for the expectations by means of the stochastic collocation method [9, 12, 90, 107, 131], which was recently applied in the financial context [57, 58], combined with standard regression techniques. Monte Carlo pricing experiments confirm that our method is highly accurate and fast.

Keywords: Local volatility, Monte Carlo, hybrid, stochastic volatility, stochastic local volatility, stochastic interest rates, stochastic collocation, regression, SABR, Heston, Hull-White.

4.1. INTRODUCTION

In this chapter, we propose for two types of *hybrid local volatility models* a novel, highly efficient Monte Carlo simulation method. We consider stochastic local volatility (SLV) models and the local volatility model incorporating stochastic interest rates. By construction, these hybrid models can be calibrated perfectly to the plain vanilla market.

This chapter is based on the article ‘A Novel Monte Carlo Approach to Hybrid Local Volatility Models’, published in *Quantitative Finance*, 17(9):1347-1366, 2017 [120].

While (partially) inheriting particular desirable features from their ‘pure’ stochastic volatility counterparts or including stochastic interest rates, these hybrid models yield an enhancement in the pricing of long-dated FX and equity-linked structured financial products. Although this makes these models attractive to the financial industry, their evaluation is not trivial. A particular (conditional) expectation needs to be established, which cannot be extracted from the market quotes. The *stochastic collocation method* [9, 12, 90, 107, 131], see Section 1.2.4, which was recently applied in the financial context [57, 58], allows us to determine for both types of hybrid local volatility models the expectation in a way that is highly accurate and fast.

In this chapter we consider the SABR and Heston models, enhanced by a non-parametric local volatility component, which we refer to as the *SABR-Local Volatility*¹ (SABR-LV) and *Heston-Stochastic Local Volatility* (Heston-SLV or H-SLV) models. Also, we study the local volatility model incorporating stochastic interest rates governed by Hull-White dynamics, the so-called *Local Volatility-Hull White* (LV-HW) model. The SABR-LV and Heston-SLV models, compared to the traditional local volatility model, typically yield a more stable hedging performance and a more accurate pricing of forward volatility sensitive products. Enriching the local volatility model with (Hull-White) stochastic interest rates enhances the pricing of long-dated FX and equity-linked structures. We evaluate the SABR-LV, Heston-SLV and LV-HW models based on a method that combines stochastic collocation and standard regression techniques.

The chapter is organized as follows. We briefly touch upon stochastic local volatility models in Section 4.1.1. Subsequently, in Section 4.1.2, we provide an overview of the evaluation approaches for the local volatility model incorporating stochastic interest rates. The basics of stochastic collocation are discussed in Section 4.1.3. In Section 4.2 we present and numerically test the stochastic collocation based approach for the evaluation of stochastic local volatility models. We subsequently apply this method to the local volatility model enhanced by stochastic interest rates in Section 4.3. Section 4.4 concludes this chapter’s work.

4.1.1. STOCHASTIC LOCAL VOLATILITY MODELS

The class of stochastic local volatility models was developed by Jex et al. [75] and Lipton [80, 82]², amongst others. As e.g. pointed out in [23] and [81], for the pricing of FX options SLV models are typically used. Also in the Bloomberg note [115] stochastic local volatility models are presented in an FX context. SLV methods have been described in detail in Section 1.2.3 and also in Chapter 2 of this thesis.

In an SLV framework the conditional expectation of the form $\mathbb{E}[\psi^2(V(t)) | S(t) = K]$ needs to be established, see e.g. [36]. The exact form of $\psi(\cdot)$ depends on the choice of stochastic volatility model, e.g. for the Heston model $\psi(x) = \sqrt{x}$. The conditional expectation cannot be extracted from the market.

In Chapter 2 we introduced in a Monte Carlo setting a non-parametric method for the evaluation of the problematic conditional expectation, which relies on splitting the Monte Carlo realizations in bins. A similar idea was presented in e.g. [61], based on

¹Note that, in fact, the ‘pure’ SABR model is already a stochastic local volatility model with a *parametric* local volatility component.

²In [82] Lipton and McGhee present a more general form of stochastic local volatility models including jumps.

kernel estimators in an interacting particle system. All the above-mentioned numerical techniques are found to be relatively costly, or limited in applicability.

4.1.2. LOCAL VOLATILITY MODEL WITH STOCHASTIC INTEREST RATES

Regarding local volatility in a stochastic interest rates (SIR) framework, the literature is not very rich. In [8] Gyöngy's [62] mimicking techniques are used to incorporate stochastic interest rates in a local volatility framework. More generally, it is shown how Gyöngy's theorem can be used to relate any continuous stochastic volatility model with stochastic interest rates to a local volatility with deterministic interest rates.

In [97] Piterbarg states that the slope of the FX volatility is a major factor affecting the values of Power Reverse Dual Currency swaps. He therefore comes up with a skew-enabled model, namely the local volatility model with domestic and foreign interest rates following Hull-White dynamics, which serves as an extension to the traditional three-factor log-normal model (without skew). For the stability of the calibration, a CEV specification for the local volatility function is chosen. This yields an essentially instantaneous calibration procedure, which is based on a Markovian representation technique³ of the dynamics of the forward FX rate and skew averaging techniques. The calibration basically 'captures' mainly the slope of the implied volatility.

Further, in [104] an expression is derived for the local volatility in a stochastic interest rates framework, consisting of the particular expectation $\mathbb{E}^T [r(T) \mathbb{1}_{S(T) > K}]$, which they compute by iteratively solving the corresponding Kolmogorov forward equation forward in time. Benhamou et al. [14] specify the bias between the local volatility with and without stochastic interest rates. By means of numerical experiments they illustrate the importance of this bias which, in line with intuition, gets larger for longer maturity. In another chapter, based on his work on perturbation methods for local volatility models, Benhamou [13] presents and numerically tests the expansions approximating the prices of European options in a local volatility model with stochastic interest rates. In numerical experiments, similar as in [97], a CEV diffusion for the spot is chosen. In [32] the authors present, in an FX context with stochastic interest rates, four methods to compute the local volatility function for different strike prices and time-points. Although this paper provides a clear overview of the ways the local volatility component can be computed, no concrete calibration or pricing experiments are included. Last, in [61] the authors evaluate the Ho-Lee/Dupire hybrid model by an approach that is based on McKean's particle method.

4.1.3. STOCHASTIC COLLOCATION BASICS

In this section, we briefly discuss the basics of the stochastic collocation method. The original idea of stochastic collocation is to project uncertainty onto a probability space with known properties and conditions [9, 131]. In particular, we approximate a variable of interest, Y , which is expensive to compute, by a function of a more convenient 'cheap to evaluate' random variable X .

Collocation methods have been studied and employed in various disciplines for uncertainty quantification, see e.g. [45, 127, 131]. In collocation methods the target is

³In [98] Piterbarg formalizes this procedure as the *Markovian projection* method.

to satisfy the governing differential equations at a discrete set of points, in the corresponding probability space. Two of the main approaches of high-order stochastic collocation methods are the Lagrange interpolation approach, see e.g. [131], and the pseudo-spectral generalized polynomial chaos approach from e.g. [130].

We explain the stochastic collocation method in a sampling setting. Suppose we wish to sample values y_n from the distribution of Y . This is typically established by first drawing samples u_n from a standard uniform distribution $U \stackrel{d}{=} \mathcal{U}([0, 1])$ and subsequently applying the inversions $y_n = F_Y^{-1}(u_n)$. However, this sampling approach is not preferred in the case that the inversion $F_Y^{-1}(\cdot)$ is *expensive* – the sampling is not performed in an efficient way. By the stochastic collocation method this issue is overcome.

The stochastic collocation technique relies on the fact that $F_Y(Y) \stackrel{d}{=} U \stackrel{d}{=} F_X(X)$, for an arbitrary random variable X , with $U \stackrel{d}{=} \mathcal{U}([0, 1])$, i.e. the CDFs of Y and X (not Y and X themselves) are equal in distribution. In a sampling setting, with y_n and x_n denoting samples from the distributions of Y and X , respectively, the target is to find a function $g(\cdot)$ such that

$$F_Y(g(x_n)) = F_X(x_n), \quad y_n = g(x_n). \quad (4.1.1)$$

When the function $g(\cdot)$ is determined, sampling from Y can be performed by sampling from X , without performing the expensive inversion $F_Y^{-1}(\cdot)$, which is needed when sampling in the traditional way. Trivially, equation (4.1.1) implies $g(\cdot) = F_Y^{-1}(F_X(\cdot))$. However, the task is to find a function which does *not* require many expensive inversions $F_Y^{-1}(\cdot)$. This can be achieved in a polynomial chaos expansion framework, where a sample y_n is approximated in terms of Lagrange basis polynomials $\ell(\cdot)$ evaluated at a sample of X , x_n , as

$$y_n = g(x_n) \approx g_N(x_n) := \sum_{i=1}^N y_i \ell_i(x_n), \quad \ell_i(x_n) := \prod_{j=1, j \neq i}^N \frac{x_n - x_j}{x_i - x_j},$$

where x_i and x_j are so-called *collocation points*, y_i is the exact ‘expensive’ evaluation at the collocation point x_i , i.e. $y_i = F_Y^{-1}(F_X(x_i))$. Choosing the interpolation polynomial in the Lagrange form is well-accepted in the field of Uncertainty Quantification (when the stochastic collocation method is applied), see e.g. [107]. By a change of basis it can be written in terms of monomials, $g_N(x_n) = a_0 + a_1 x_n + a_2 x_n^2 + \dots + a_{N-1} x_n^{N-1}$, where the coefficients a_0, a_1, \dots, a_{N-1} are obtained by solving a linear system $\mathbf{V}\mathbf{a} = \mathbf{y}$, with matrix \mathbf{V} denoting the Vandermonde matrix (see e.g. [58] for more details).

Once the function $g_N(\cdot)$ is established by N expensive inversions $F_Y^{-1}(\cdot)$, we are able to generate any number of samples y_n without significant additional cost. The collocation points x_i can basically be chosen arbitrarily, however we choose them in an *optimal way*, i.e. based on the zeros of an orthogonal polynomial.

In this chapter we choose X to be standard normally distributed. This implies that the *optimal* collocation points are the zeros of the Hermite polynomials (abscissas of the Gauss-Hermite quadrature)[3, 58]. It turns out that choosing X to be normally distributed works highly satisfactory; the method is accurate and efficient, as the inversion of a normal distribution is ‘cheap’. We therefore do not consider other distributions for X , which may yield a method with similar accuracy, but more expensive inversions. What is more, the Cameron-Martin Theorem [19] states that polynomial chaos approximations based on the normal distribution converge to any distribution.

In the following sections we employ the stochastic collocation method for the efficient Monte Carlo evaluation of stochastic local volatility models and the local volatility model incorporating stochastic interest rates.

4.2. STOCHASTIC LOCAL VOLATILITY MODELS

In this section we discuss stochastic local volatility models. In [75, 80, 82], amongst others, the class of stochastic local volatility models was introduced, which combine properties of the traditional local volatility model [33, 35] and stochastic volatility models, like the SABR model [63] and the Heston model [66]. According to Clark [23] neither the ‘sticky-delta’ property of stochastic volatility models nor the ‘sticky-strike’ characteristic corresponding to the local volatility model is in line with the actual smile behaviour in FX markets; the reality is somewhere between the two and therefore typically a stochastic local volatility model is used. In line with this, as pointed out in [81], SLV models are *de facto* standard for pricing FX options.

Assuming constant deterministic interest rate r , no dividends and instantaneous correlation $\rho_{s,v}$, the general SLV model dynamics under the risk-neutral \mathbb{Q} -measure⁴ read

$$dS(t)/S(t) = rdt + \sigma(t, S(t))\psi(S(t), V(t))dW_s^{\mathbb{Q}}(t), \quad S(0) = S_0, \quad (4.2.1)$$

$$dV(t) = a_v(t, V(t))dt + b_v(t, V(t))dW_v^{\mathbb{Q}}(t), \quad V(0) = V_0, \quad (4.2.2)$$

with $dW_s^{\mathbb{Q}}(t)dW_v^{\mathbb{Q}}(t) = \rho_{s,v}dt$ and $\sigma^2(t, K) = \sigma_{LV}^2(t, K)/\mathbb{E}^{\mathbb{Q}}[\psi^2(S(t), V(t)) | S(t) = K]$, where $\sigma_{LV}^2(t, K)$ denotes Dupire’s local volatility component [35]:

$$\sigma_{LV}^2(t, K) = \frac{\frac{\partial C(t, K)}{\partial t} + rK \frac{\partial C(t, K)}{\partial K}}{\frac{1}{2}K^2 \frac{\partial^2 C(t, K)}{\partial K^2}}.$$

For notation purposes, we suppress the \mathbb{Q} -superscript from this point on. By choosing $\psi(S(t), V(t)) = \sqrt{V(t)}S^{\beta-1}(t)$, $a_v(t, V(t)) = \gamma^2 V(t)$ and $b_v(t, V(t)) = 2\gamma V(t)$ we obtain the *SABR-LV model*⁵:

$$dS(t)/S(t) = rdt + \sigma(t, S(t))\sqrt{V(t)}S^{\beta-1}(t)dW_s(t), \quad S(0) = S_0, \quad (4.2.3)$$

$$dV(t) = \gamma^2 V(t)dt + 2\gamma V(t)dW_v(t), \quad V(0) = V_0, \quad (4.2.4)$$

with $dW_s(t)dW_v(t) = \rho_{s,v}dt$. In the SABR-LV model the local volatility component is specified by

$$\sigma^2(t, K) = \frac{\sigma_{LV}^2(t, K)}{K^{2\beta-2}\mathbb{E}[V(t) | S(t) = K]}. \quad (4.2.5)$$

⁴Note that the general SLV model as described by equations (2)-(3) is an incomplete market model, which implies that a unique risk-neutral pricing measure does not exist, see e.g. [41].

⁵To prevent double use of the σ -notation we write the variance dynamics instead of the more common volatility dynamics. The traditional SABR model dynamics are given by the following two SDEs [63, 103]:

$$dF^T(t) = \sigma(t)(F^T(t))^{\beta}dW_s^T(t), \quad d\sigma(t) = \gamma\sigma(t)dW_v^T(t),$$

with $F^T(\cdot)$ denoting the forward corresponding to expiry T and $dW_s^T(t)dW_v^T(t) = \rho_{s,v}dt$.

We use a standard Euler discretization scheme to simulate the SABR-LV model.

The choices $\psi(S(t), V(t)) = \sqrt{V(t)}$, $a_v(t, V(t)) = \kappa(\bar{V} - V(t))$ and $b_v(t, V(t)) = \gamma\sqrt{V(t)}$ provide us with the *Heston-SLV (H-SLV) model*:

$$dS(t)/S(t) = rdt + \sigma(t, S(t))\sqrt{V(t)}dW_s(t), \quad S(0) = S_0, \quad (4.2.6)$$

$$dV(t) = \kappa(\bar{V} - V(t))dt + \gamma\sqrt{V(t)}dW_v(t), \quad V(0) = V_0, \quad (4.2.7)$$

with $dW_s(t)dW_v(t) = \rho_{s,v}dt$ and

$$\sigma^2(t, K) = \frac{\sigma_{LV}^2(t, K)}{\mathbb{E}[V(t)|S(t) = K]}. \quad (4.2.8)$$

Similar dynamics are presented in the chapter of Jex et al. [75]. In order to simulate the Heston-SLV model, we use an *adapted version* of the *Quadratic Exponential (QE) scheme* introduced in [5], which we derive in [118]. The difference between the original and the adapted version lies in the fact that only the latter incorporates the local volatility component. Let $i = 0, 1, \dots, M$ and $j = 1, 2, \dots, N$ indicate the time-step and path, respectively. Defining $\Delta_t := T/M$ as the time-step size, the discretization scheme for $V(\cdot)$ and $X(\cdot) := \log(S(\cdot))$ reads

$$v_{i+1,j} \sim c(\Delta_t)\chi^2(d, \lambda(t_i, v_{i,j})), \quad v_{0,j} = V_0 \quad (4.2.9)$$

$$\begin{aligned} x_{i+1,j} &= x_{i,j} + r\Delta_t - \frac{1}{2}\hat{\sigma}^2(t_i, x_{i,j})v_{i,j}\Delta_t + \frac{\rho_{s,v}}{\gamma}\hat{\sigma}(t_i, x_{i,j})(v_{i+1,j} - \kappa\bar{v}\Delta_t + v_{i,j}c_1) \\ &\quad + \rho_1\sqrt{\hat{\sigma}^2(t_i, x_{i,j})}v_{i,j}\Delta_t Z, \quad x_{0,j} = \log(S_0) \end{aligned} \quad (4.2.10)$$

with $Z \stackrel{d}{=} \mathcal{N}(0, 1)$, $\rho_1 := (1 - \rho_{s,v}^2)^{1/2}$, $c_1 := \kappa\Delta_t - 1$ and

$$c(\Delta_t) := \frac{\gamma^2}{4\kappa}(1 - e^{-\kappa\Delta_t}), \quad d := \frac{4\kappa\bar{V}}{\gamma^2}, \quad \lambda(t, V(t)) := \frac{4\kappa e^{-\kappa\Delta_t}}{\gamma^2(1 - e^{-\kappa\Delta_t})}V(t), \quad (4.2.11)$$

where $\chi^2(d, \lambda(t, V(t)))$ represents a noncentral chi-squared distribution with d degrees of freedom and non-centrality parameter $\lambda(t, V(t))$. Further, the local volatility component reads

$$\hat{\sigma}^2(t_i, x_{i,j}) \stackrel{\text{def}}{=} \sigma^2(t_i, e^{x_{i,j}}) = \frac{\sigma_{LV}^2(t_i, s_{i,j})}{\mathbb{E}[V(t_i)|S(t_i) = s_{i,j}]}. \quad (4.2.12)$$

Numerical comparisons between the Euler and the *original* QE scheme have been provided in the literature [5]. In Chapter 2 we numerically demonstrated that the *adapted* QE scheme outperforms the standard Euler scheme: it yields a higher accuracy and a faster convergence to the reference for a decaying time-step size.

Equation (4.2.12) makes clear that in a Monte Carlo simulation framework, for both the SABR-LV model and the Heston-SLV model, we need to evaluate the conditional expectation for *each path, at each time-step*. A closed-form representation does not exist, as the joint distribution of $S(\cdot)$ and $V(\cdot)$ is unknown. We require the evaluation to be *efficient* and *accurate* – if it is not, the error introduced accumulates in the simulation and the results are biased. The principle of stochastic collocation [9, 131], discussed in Section 4.1.3, allows for an evaluation that satisfies both requirements.

4.2.1. ESTABLISHING $\mathbb{E}[V(t)|S(t) = K]$

In this section we evaluate the conditional expectation of interest $\mathbb{E}[V(t)|S(t) = K]$, which is present in both the SABR-LV model (4.2.3)-(4.2.4) and the Heston-SLV model (4.2.6)-(4.2.7). Our approach essentially consists of two projection steps. We first project $V(\cdot)$ and $S(\cdot)$ on standard normal random variables, where, by means of stochastic collocation, $\mathbb{E}[V(t)|S(t) = K]$ is decomposed into a *series of conditional expectations*. Secondly, similar to e.g [73, 83], each of these conditional expectations, which are expressed in terms of standard normal random variables, is approximated by a projection on a set of basis functions and applying standard regression techniques.

We start by projecting $S(\cdot)$ at a given fixed time t , on a standard normal random variable $X \stackrel{d}{=} \mathcal{N}(0, 1)$ via the function $g(\cdot)$, defined by

$$g(\cdot) := F_{S(t)}^{-1}(F_X(\cdot)), \quad (4.2.13)$$

which ensures

$$S(t) \stackrel{d}{=} g(X)$$

and, moreover, for elements $S(t) = s$ and $X = x$:

$$s = g(x). \quad (4.2.14)$$

In a similar way, we project $V(t)$ on a standard normal random variable $Z \stackrel{d}{=} \mathcal{N}(0, 1)$:

$$V(t) \stackrel{d}{=} h(Z), \quad h(\cdot) := F_{V(t)}^{-1}(F_Z(\cdot)), \quad (4.2.15)$$

which also yields for elements $V(t) = v$ and $Z = z$:

$$v = h(z). \quad (4.2.16)$$

The conditional expectation can be written in terms of X and Z :

$$\mathbb{E}[V(t)|S(t) = K] = \mathbb{E}[h(Z)|g(X) = K]. \quad (4.2.17)$$

The joint distribution of X and Z is not known analytically. Although X and Z are both normally distributed, the joint distribution of X and Z is *not* bivariate normal⁶ – only the reverse holds in general. We therefore cannot evaluate the right-hand side of (4.2.17) analytically and we proceed by determining an approximation for it. This is established by approximating the function $h(\cdot)$ by a polynomial $h_{N_V}(\cdot)$ with degree $N_V - 1$, which is obtained by the stochastic collocation method with N_V collocation points. In particular, given the collocation points z_i , that are *a priori* known, we compute the corresponding *exact* evaluations of $V(t)$:

$$v_i = h(z_i) = F_{V(t)}^{-1}(F_Z(z_i)), \quad i = 1, 2, \dots, N_V. \quad (4.2.18)$$

⁶A well-known test for multi-variate normality is Mardia's, see [87], which is based on multivariate extensions of skewness and kurtosis measures.

Next, we apply Lagrangian interpolation through the v_i -values – choosing the interpolation polynomial in the Lagrange form is well-accepted in the field of uncertainty quantification (when the stochastic collocation method is applied), see e.g. [107]. For an arbitrary value $V(t) = v$, it holds that

$$v = h_{N_V}(z) + \epsilon_1(z) := \sum_{i=1}^{N_V} v_i \ell_i(z) + \epsilon_1(z), \quad \ell_i(z) := \prod_{k=1, k \neq i}^{N_V} \frac{z - z_k}{z_i - z_k}, \quad (4.2.19)$$

where $\epsilon_1(z)$ denotes the interpolation error corresponding to the particular argument z . By a change of basis we can write the Lagrange polynomial in terms of monomials:

$$h_{N_V}(z) = a_0 + a_1 z + \dots + a_{N_V-1} z^{N_V-1}, \quad (4.2.20)$$

where the coefficients $a_0, a_1, \dots, a_{N_V-1}$ are obtained by solving a linear system involving a Vandermonde matrix, see [58] for more details. Given (4.2.20), we approximate the conditional expectation on the right-hand side of (4.2.17) as follows:

$$\begin{aligned} & \mathbb{E} [h(Z) | g(X) = K] \\ &= \mathbb{E} [h_{N_V}(Z) + \epsilon_1(Z) | g(X) = K] \\ &= \mathbb{E} [h_{N_V}(Z) | g(X) = K] + \mathbb{E} [\epsilon_1(Z) | g(X) = K] \\ &= a_0 + a_1 \mathbb{E} [Z | X = g^{-1}(K)] + \dots + a_{N_V-1} \mathbb{E} [Z^{N_V-1} | X = g^{-1}(K)] \\ & \quad + \bar{\epsilon}_1(K), \end{aligned} \quad (4.2.21)$$

with $\bar{\epsilon}_1(K) := \mathbb{E} [\epsilon_1(Z) | g(X) = K]$. The inversions of functions $g(\cdot)$ and $h(\cdot)$, defined in (4.2.13) and (4.2.15), respectively, are cheap, as both merely consist of (1) the inversion of a standard normal random variable and (2) the evaluation of $F_{S(t)}(\cdot)$ or $F_{V(t)}(\cdot)$. As CDFs are strictly monotonic, the inversions of $g(\cdot)$ and $h(\cdot)$ provide a *bijective* mapping between the original probability space and the new space.

As we mentioned, the joint distribution of X and Z is not analytically known. To approximate the conditional expectations in (4.2.21), we assume that we can approximate the conditional expectation $\mathbb{E} [Z^p | X = x]$ in terms of functions of x , the *basis functions* $\psi_{kp}(\cdot)$, $k = 1, 2, \dots, n$, $p = 1, 2, \dots, N_V - 1$:

$$\mathbb{E} [Z^p | X = x] = \sum_{k=1}^n b_{kp} \psi_{kp}(x) + \epsilon_{2p}. \quad (4.2.22)$$

Equation (4.2.22) is motivated rigorously by assuming that the conditional expectations in (4.2.21) are elements of the L^2 -space of square integrable functions. As the L^2 -space is a Hilbert space, it possesses a countable orthonormal basis and the conditional expectations (which are deterministic functions) can be expressed as a linear combination of the elements of this basis. A similar idea is used in [83] in the context of valuing American options by simulation, where the value of continuing with the option is expressed as a conditional expectation. We approximate the conditional expectation in (4.2.22) by using the first n orthogonal polynomials⁷ $\{1, x, x^2, \dots, x^{n-1}\}$ and, similar as in e.g. [73, 83],

⁷Other, more complex types of basis functions we may use are the Laguerre, Hermite, Legendre, Chebyshev, Gegenbauer and Jacobi polynomials, see e.g. Chapter 22 of [3]. In this chapter we do not consider these basis functions, as the set of simple polynomials $\{1, x, \dots, x^{n-1}\}$ already yields highly satisfactory results, see the numerical experiments in Section 4.2.3.

we apply OLS regression to compute the corresponding coefficients, which yields

$$\mathbb{E}[Z^p | X = g^{-1}(K)] = \widehat{\mathbb{E}}[Z^p | X = g^{-1}(K)] + \widehat{\varepsilon}_{2p}, \quad (4.2.23)$$

with $\widehat{\mathbb{E}}[Z^p | X = g^{-1}(K)] = \widehat{\beta}_{0p} + \widehat{\beta}_{1p}g^{-1}(K) + \widehat{\beta}_{2p}(g^{-1}(K))^2 + \dots + \widehat{\beta}_{n-1,p}(g^{-1}(K))^{n-1}$. Combining this result with (4.2.21) yields

$$\mathbb{E}[V(t) | S(t) = K] = a_0 + \sum_{p=1}^{N_V-1} a_p \sum_{k=0}^{n-1} \widehat{\beta}_{kp} (g^{-1}(K))^k + \bar{\varepsilon}_1(K) + \bar{\varepsilon}_2, \quad (4.2.24)$$

with $g(\cdot)$ defined in (4.2.13), $\bar{\varepsilon}_1(K) := \mathbb{E}[e_1(Z) | X = g^{-1}(K)]$ and $\bar{\varepsilon}_2 := \sum_{p=1}^{N_V-1} \widehat{\varepsilon}_{2p}$.

A brief analysis of the errors $\bar{\varepsilon}_1(\cdot)$ and $\bar{\varepsilon}_2$ can be found in Appendix 4.A. In [9], in an elliptic PDE framework, a rigorous convergence analysis of the stochastic collocation method is provided, where exponential convergence with respect to the number of ‘‘Gauss points’’ is proven. Our numerical experiments in Section 4.2.3 for a base case are in line with this.

Computation of the approximation in (4.2.24) is efficient, as it only requires N_V inversions of $F_{V(t)}(\cdot)$, see (4.2.18). Determining x and z and the OLS estimates $\widehat{\beta}_{0p}, \widehat{\beta}_{1p}, \dots, \widehat{\beta}_{n-1,p}$ does not involve significant computational cost. In the following, we refer to the approach presented in this section as the ‘stochastic collocation – regression’ or ‘SC-R’ approach.

In a Monte Carlo simulation framework, we apply the SC-R approach as described in Algorithm 3. In this algorithm $i = 1, 2, \dots, M$ denotes the time-step and $j = 1, 2, \dots, N$ indicates the path. For simulating the SABR-LV model we use a standard Euler discretization scheme, whereas for the Heston-SLV model we apply the adapted version of the QE scheme, see equations (4.2.9)-(4.2.11). After the Monte Carlo simulation described in Algorithm 3 we price European call options based on the obtained values for $S(\cdot)$ at the time of maturity. This results in the model implied volatility values $\bar{\sigma}_{\text{model}}$ displayed in Figures 4.2.4 and 4.2.5 and the errors reported in Tables 4.2.2 and 4.2.3. In the Monte Carlo simulation, it may be necessary to apply one or more of the enhancements we describe in the follow-up section.

for each time-step $t_i, i = 1, 2, \dots, M$ do

- 1 | Generate N pairs $(s_{i,j}, v_{i,j}), j = 1, 2, \dots, N$ by going forward one time-step in the Euler scheme (SABR-LV model) or the adapted QE scheme (Heston-SLV model).
- 2 | Compute $\mathbb{E}[V(t_i) | S(t_i) = s_{i,j}]$ using the SC-R approach, see equation (4.2.24).
- 3 | Establish the local volatility component $\sigma^2(t_i, s_{i,j})$ by equation (4.2.5) for the SABR-LV model or equation (4.2.8) for the Heston-SLV model – use its value in step 1.

end

- 4 | Price European call options based on the obtained values for $S(\cdot)$ at the time to maturity.

Algorithm 3: Pricing European call options by a Monte Carlo simulation of the SABR-LV and Heston-SLV models, incorporating the SC-R approach (Section 4.2.1).

Remark 4.2.1. *In a stochastic local volatility framework, directly applying OLS regression may yield reasonable results as well. However, as we describe in Remark 2.3.2 of Chapter 2, non-negativity of the conditional expectation cannot be guaranteed for cases where*

the Feller condition is violated and improvements must be made. Further, by applying stochastic collocation we can use the analytical expression of the CDF of $V(\cdot)$ in order to obtain values for the coefficients a_0, a_1, \dots . Moreover, in the context of the local volatility model with stochastic interest rates, see Section 4.3, by projecting $S(\cdot)$ on a standard normal random variable we can employ the analytical expression for moments of a truncated standard normal random variable, see Result 4.3.1.

4.2.2. ENHANCEMENTS

In this section, we discuss three adaptations to the stochastic collocation – regression method which may enhance the results.

First, we observe that at the boundaries of the X -domain (recall $X := g^{-1}(K) = F_X^{-1}(F_{S(t)}(K))$) the performance of the regression deteriorates due to the presence of a small number of observations, which may yield a significant increase of $\bar{\epsilon}_2$. We therefore set for $K \leq s_{\min}$: $\mathbb{E}[V(t)|S(t) = K] = \mathbb{E}[V(t)|S(t) \leq s_{\min}]$ and for $K \geq s_{\max}$: $\mathbb{E}[V(t)|S(t) = K] = \mathbb{E}[V(t)|S(t) \geq s_{\max}]$, where s_{\min} and s_{\max} are percentiles of the $S(t)$ -distribution, i.e. $s_{\min} = F_{S(t)}^{-1}(p_{s,\min})$ and $s_{\max} = F_{S(t)}^{-1}(p_{s,\max})$. Here $0 \leq p_{s,\min} < p_{s,\max} \leq 1$ denote fractions of the total number of Monte Carlo realizations. *In all pricing experiments in Section 4.2.3, we apply this adaptation and choose $p_{s,\min} = 0.1$ and $p_{s,\max} = 0.9$.*

The approximation of the expectation in (4.2.24) is not guaranteed to be positive. This may be problematic in the case that a significant part of the variance realizations is close to zero, e.g. if the Feller condition in the Heston-SLV model is strongly violated. In this case we may split the conditional expectation in two parts in the following way:

$$\begin{aligned} \mathbb{E}[V(t)|S(t) = K] &= \mathbb{E}[V(t)|S(t) = K, V(t) \leq v^*] \mathbb{Q}[V(t) \leq v^*] \\ &\quad + \mathbb{E}[V(t)|S(t) = K, V(t) > v^*] (1 - \mathbb{Q}[V(t) \leq v^*]). \end{aligned} \quad (4.2.25)$$

The first conditional expectation we approximate by $\mathbb{E}[V(t)|V(t) \leq v^*]$, the second conditional expectation is approximated by the stochastic collocation – regression approach. We can choose v^* to be a fixed value, or based on a fixed percentile p_v^* , i.e. $v^* = F_{V(t)}^{-1}(p_v^*)$. We prefer the latter, as in this case at each time-step in the Monte Carlo simulation we naturally control the fraction of the total number of observations on which we apply the stochastic collocation – regression approach. So we obtain

$$\begin{aligned} \mathbb{E}[V(t)|S(t) = K] &= (\mathbb{E}[V(t)|V(t) \leq v^*] + \epsilon_3) F_{V(t)}(v^*) \\ &\quad + \left(a_0 + \sum_{p=1}^{N_V-1} a_p \sum_{k=0}^{n-1} \hat{\beta}_{kp} (g^{-1}(K))^k + \bar{\epsilon}_1(K) + \bar{\epsilon}_2 \right) (1 - F_{V(t)}(v^*)) \\ &:= \mathcal{V}(K) + \epsilon, \end{aligned} \quad (4.2.26)$$

where ϵ denotes the approximation error. By means of this adaptation we leave out the smallest variance realizations when applying the stochastic collocation – regression approach, which makes it less likely that the corresponding SC-R approximation yields negativity. To this approximation we moreover add the positive term $(\mathbb{E}[V(t)|V(t) \leq v^*]) \cdot F_{V(t)}(v^*)$.

Although the former adaptation guarantees non-negativity for $\mathcal{V}(K)$ in (4.2.26), in extreme cases the fraction of $V(\cdot)$ -realizations close to zero is substantial, and we would

need to choose a relatively large value for p_v^* to ensure non-negativity of $\mathcal{V}(K)$. This would make the approximation for the conditional expectation inaccurate, as in this case it is for a large part determined by the naive approximation $\mathbb{E}[V(t)|V(t) \leq v^*]F_{V(t)}(v^*)$. Therefore, in the case that the approximation $\mathcal{V}(K)$ still yields negative values for an appropriate value of p_v^* (in our numerical experiments we choose p_v^* in the range 0.01 – 0.1), we apply another correction, namely

$$\mathbb{E}[V(t)|S(t) = K] = \mathcal{V}(K) + \epsilon - \min_K \{0, (1 + \delta)\mathcal{V}(K)\},$$

with $0 < \delta < 1$. This correction is interpreted as follows: in the case that a part of $\mathcal{V}(K)$ is negative, we apply a vertical “shift” such that it becomes positive. If $\mathcal{V}(K)$ is completely non-negative, the vertical shift is zero⁸. This correction *guarantees* non-negativity of the approximation of the conditional expectation.

4.2.3. NUMERICAL EXPERIMENTS

In this section we test the accuracy of the approximation of the conditional expectation in (4.2.24). We first test the method for a base case where an analytical reference value is available. Subsequently, we consider the SABR-LV and Heston-SLV models in a Monte Carlo simulation framework. In particular, given a pre-specified market, we add to an either poorly or satisfactorily calibrated ‘pure’ SABR or Heston model the local volatility component, consisting of Dupire’s local volatility and the conditional expectation approximation (4.2.24),

THE 2D-GBM MODEL: A BASE CASE

We start by testing the approximation of the conditional expectation (4.2.24) for a model which is given by two correlated Geometric Brownian Motions (GBMs):

$$dY_1(t) = \sigma_1 Y_1(t) dW_1(t), \quad dY_2(t) = \sigma_2 Y_2(t) dW_2(t), \quad Y_1(0) = y_{10}, \quad Y_2(0) = y_{20}, \quad (4.2.27)$$

with $dW_1(t)dW_2(t) = \rho dt$. The expectation of $Y_2(t)$ conditional on the event $Y_1(t) = y_1$ is [4]

$$\mathbb{E}[Y_2(t)|Y_1(t) = y_1] = y_{20} \left(\frac{y_1}{y_{10}} \right)^{\rho \frac{\sigma_2}{\sigma_1}} e^{t(\frac{1}{2}\rho\sigma_1\sigma_2 - \frac{1}{2}\sigma_2^2\rho^2)}. \quad (4.2.28)$$

Let $y_{10} = 1$, $y_{20} = 0.05$, $\rho = -0.5$ and $t = 5$. In a stochastic volatility model these parameter values are representative choices for S_0 , V_0 and $\rho_{s,v}$, respectively. As a first experiment, suppose we choose $N_{Y_2} = 6$ collocation points and $n = 7$ basis functions. In Figure 4.2.1 we compare the reference (4.2.28) and the approximation (4.2.24) obtained by the stochastic collocation – regression (SC-R) approach for a moderate case (left, $\sigma_1 = \sigma_2 = 0.3$) and a more extreme case (right, $\sigma_1 = \sigma_2 = 0.9$). An excellent fit is obtained.

The reference (4.2.28) allows for a numerical analysis of the errors $\bar{\epsilon}_1(\cdot)$ and $\bar{\epsilon}_2$, which are introduced by the stochastic collocation method and the regression step, respectively. We choose the parameter values just mentioned and $\sigma_1 = \sigma_2 = 0.3$. We make use of the result in the following lemma.

⁸Numerical experiments demonstrate that merely applying the third correction, i.e. applying a vertical shift, typically yields worse pricing results compared to *combining the second and third corrections* mentioned in Section 4.2.2.

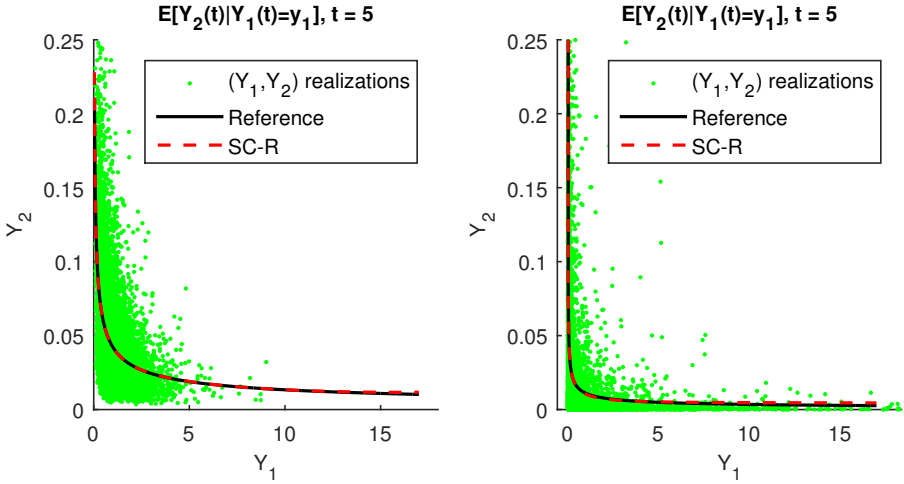


Figure 4.2.1: Comparison of the reference (4.2.28) and the SC-R approximation (4.2.24) for a moderate case (left) and a more extreme case (right).

Lemma 4.2.1. *Given the two-dimensional model (4.2.27). Let X and Z denote standard normal random variables and assume for an arbitrary t that the elements $Y_i(t) = y_i$, $i = 1, 2$, $X = x$, $Z = z$ are related by $y_1 = g(x)$, $y_2 = h(z)$, with $g(\cdot)$ and $h(\cdot)$ defined in (4.2.13) and (4.2.15), respectively. This implies that X and Z are jointly bivariate normally distributed.*

Proof. For a proof of Lemma 4.2.1, see Appendix 4.B. □

Recall the error due to the stochastic collocation method: $\bar{\epsilon}_1(K) := \mathbb{E}[\epsilon_1(Z) | g(X) = K]$. In a Monte Carlo simulation framework, for a fixed t , let $\bar{\epsilon}_1(y_{1j})$ denote the error corresponding to the j th realization of $Y_1(t)$, y_{1j} . From (4.2.21) and (4.2.28) it follows that it is given by

$$\bar{\epsilon}_1(y_{1j}) = y_{20} \left(\frac{y_{1j}}{y_{10}} \right)^{\rho \frac{\sigma_2}{\sigma_1}} e^{t(\frac{1}{2}\rho\sigma_1\sigma_2 - \frac{1}{2}\sigma_2^2\rho^2)} - f(y_{1j}),$$

with

$$f(y_{1j}) := a_0 + a_1 \mathbb{E}[Z | X = g^{-1}(y_{1j})] + \dots + a_{N_{Y_2}-1} \mathbb{E}[Z^{N_{Y_2}-1} | X = g^{-1}(y_{1j})]. \quad (4.2.29)$$

As Z and X are jointly bivariate normally distributed, see Lemma 4.2.1, we are able to evaluate each conditional expectation in (4.2.29) analytically. For arbitrary $p \in \{1, 2, \dots, N_{Y_2}-1\}$, applying the Cholesky decomposition, straightforward calculus yields

$$\mathbb{E}[Z^p | X = g^{-1}(y_{1j})] = \sum_{k=0}^p \binom{p}{k} \rho^{p-k} (1-\rho^2)^{\frac{k}{2}} (g^{-1}(y_{1j}))^{p-k} \mu_k, \quad (4.2.30)$$

with $\mu_k = (k-1)!!$ if k is even and $\mu_k = 0$ if k is odd. The double exclamation marks stand for the “double factorial”. For an even integer $n > 0$ it is defined as $n!! = n \cdot (n-2) \cdot (n-$

4) ...6 · 4 · 2 and for an odd integer $n > 0$ it is $n!! = n \cdot (n-2) \cdot (n-4) \dots 5 \cdot 3 \cdot 1$ and, by an extension, $-1!! = 1$. Further, by definition, $0!! = 1$.

Given (4.2.30), for different N_{Y_2} values we compute $\mathcal{E}_1 := \log\left(\frac{1}{N} \sum_{j=1}^N |\bar{\epsilon}_1(y_{1j})|\right)$, where N denotes the total number of observations. In Figure 4.2.2 on the left-hand side \mathcal{E}_1 is displayed against the number of collocation points. An exponential convergence is observed, which is in line with [9], where in an elliptic PDE framework a rigorous proof for exponential convergence of the stochastic collocation method is provided. The error does not decrease further for $N_{Y_2} > 14$, as machine precision has been reached ($\exp(-36) \approx 2 \cdot 10^{-16}$).

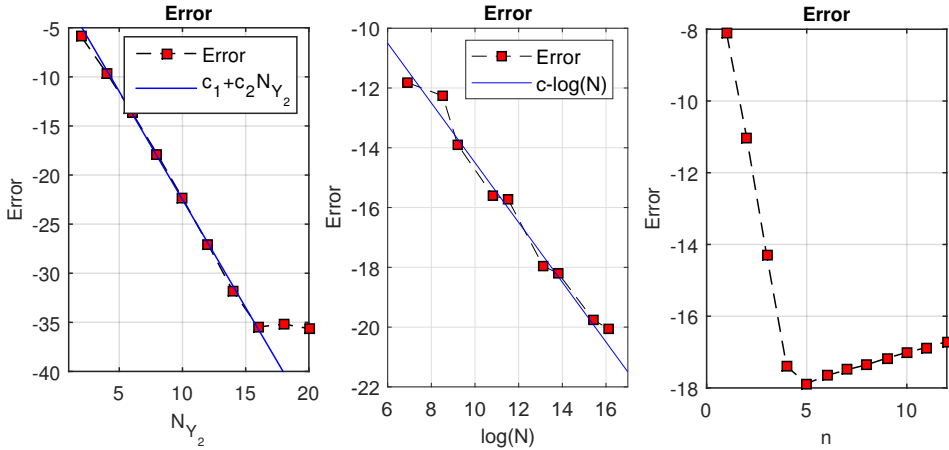


Figure 4.2.2: Left: the relation between the error $\mathcal{E}_1 = \log\left(\frac{1}{N} \sum_{j=1}^N |\bar{\epsilon}_1(y_{1j})|\right)$ and the number of collocation points N_{Y_2} ($N = 10^6$). Middle: the relation between the error $\mathcal{E}_2 = \log\left(\frac{1}{N} \sum_{j=1}^N \bar{\epsilon}_{2j}^2\right)$ and the logarithm of the number of realizations N ($N_{Y_2} = 14$, $n = 5$). Right: the relation between the error \mathcal{E}_2 and the number of basis functions n ($N_{Y_2} = 14$, $N = 10^6$).

We proceed with analyzing $\bar{\epsilon}_2 := \sum_{p=1}^{N_{Y_2}-1} \hat{\epsilon}_{2p}$, the error due to regression. Define $\bar{\epsilon}_{2j} := \sum_{p=1}^{N_{Y_2}-1} \hat{\epsilon}_{2pj} = f(y_{1j}) - \hat{f}(y_{1j})$, $j = 1, 2, \dots, N$, with $f(\cdot)$ given by (4.2.29), where the conditional expectations $\mathbb{E}[Z^p | X = g^{-1}(y_{1j})]$, $p = 1, 2, \dots, N_{Y_2} - 1$ are evaluated by the analytical formula (4.2.30), and $\hat{f}(\cdot)$ denotes

$$\hat{f}(y_{1j}) := a_0 + a_1 \hat{\mathbb{E}}[Z | X = g^{-1}(y_{1j})] + \dots + a_{N_{Y_2}-1} \hat{\mathbb{E}}[Z^{N_{Y_2}-1} | X = g^{-1}(y_{1j})], \quad (4.2.31)$$

where $\hat{\mathbb{E}}[Z^p | X = g^{-1}(y_{1j})]$, $p = 1, 2, \dots, N_{Y_2} - 1$ is obtained by OLS regression. We consider the logarithm of the mean squared error: $\mathcal{E}_2 := \log\left(\frac{1}{N} \sum_{j=1}^N \bar{\epsilon}_{2j}^2\right)$. We observe for \mathcal{E}_2 a convergence of order $\mathcal{O}(-\log(N))$, see the plot in the middle of Figure 4.2.2, where we consider $N = 10^3, 5 \cdot 10^3, 10^4, 5 \cdot 10^4, 10^5, 5 \cdot 10^5, 10^6, 5 \cdot 10^6, 10^7$.

Last, we study the dependence of \mathcal{E}_2 on the number of basis functions, see the right-hand plot of Figure 4.2.2, where we consider $n = 1, 2, \dots, 12$ basis functions. We observe that for $n = 5$ the smallest error is achieved. For $n > 5$ the increase in \mathcal{E}_2 is due to overfitting, where oscillations in the approximation of the conditional expectation may occur.

In practice, for the Heston-SLV and SABR-LV models we typically choose 4 – 6 collocation points; our numerical experiments confirm that with this number of collocation points sufficiently accurate results are obtained. In general, we choose the number of basis functions n in the range 5 – 9, depending on how extreme the parameters of the calibrated ‘pure’ Heston or SABR model are. In the follow-up section we consider the performance of the stochastic collocation – regression approach for the Heston-SLV and SABR-LV models in more detail.

THE SABR-LV AND HESTON-SLV MODELS

In this section we test the performance of the stochastic collocation – regression approach for the SABR-LV model (4.2.3)-(4.2.4) and the Heston-SLV model (4.2.6)-(4.2.7). Stochastic local volatility models, that are considered as the standard for pricing in an FX context, combine desirable features of a stochastic volatility model, e.g. preserving the shape of the forward volatility smile and reflecting more realistic smile dynamics, and the local volatility model, namely a perfect calibration to arbitrage-free European plain vanilla options.

We first consider the SABR-LV model. Let $S_0 = 1$, $V_0 = 0.05$, $\beta = 0.5$, $\gamma = 0.5$, $\rho = -0.5$ and $t = 2$. We generate for this parameter set (S, V) -realizations by simulating the ‘pure’ SABR model. Given the realizations at $t = 2$, for different numbers of basis functions we compare approximations obtained by the stochastic collocation – regression approach (4.2.24) and the non-parametric approach, see Chapter 2, using 10 bins, which serves as a reference. We consider $N_V = 4$ and $n = 3, 5, 7$. Results are displayed in Figure 4.2.3. For $n > 7$ no significant increase in accuracy was observed. For the Heston-SLV model we expect a similar increase in accuracy of the stochastic collocation – regression approach.

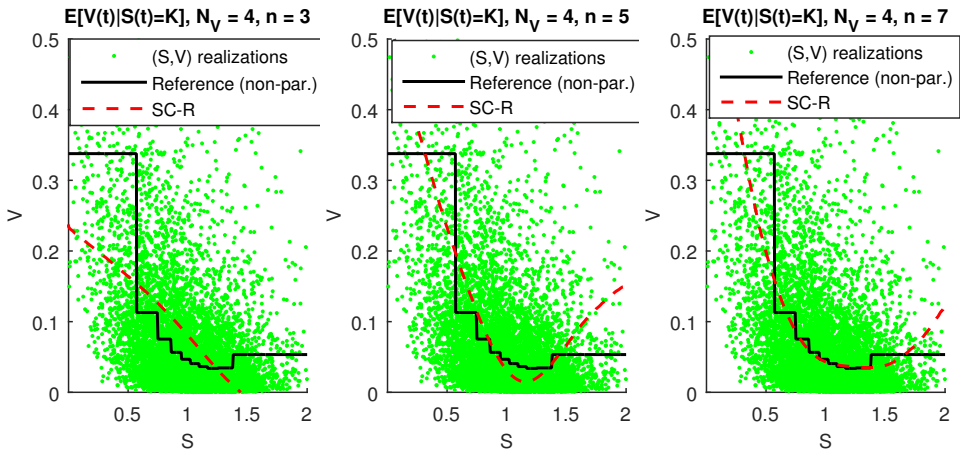


Figure 4.2.3: The conditional expectation approximation (4.2.24) obtained by the stochastic collocation – regression approach for $n = 3$ (left), $n = 5$ (middle) and $n = 7$ (right) compared to the non-parametric method [118]. Number of collocation points is $N_V = 4$.

To assess whether for a given number of collocation points N_V and basis functions n

the stochastic collocation – regression approach performs sufficiently accurate, we perform pricing experiments. In particular, given *a priori* specified market implied volatilities, we price European call options by a ‘pure’ Heston or SABR model and the Heston-SLV (‘H-SLV’) and SABR-LV model, respectively. By definition of stochastic local volatility, the Heston-SLV and SABR-LV models should yield implied volatilities that *perfectly* match the ones corresponding to the market. At each time-step in the Monte Carlo simulation we establish the conditional expectation according to (4.2.24). The Monte Carlo simulation consists of $2 \cdot 10^5$ paths (20 seeds, each seed constitutes 10^4 paths) and 200 time-steps per year, unless otherwise mentioned. We generate synthetic market prices by the Heston model, which we assume to be calibrated perfectly to the market. For this we choose some parameter sets from [23] which may be encountered in typical FX markets, see Table 4.2.1 (market data as of 16 September 2008).

Given the market data, we assume a both satisfactorily and poorly calibrated Heston model (parameter values are 20% and 80% off, respectively). On top of this model we add the local volatility component to compensate for the calibration error. For each set in Table 4.2.1, we consider the implied volatilities corresponding to (1) the market (the perfectly calibrated Heston model with parameters given in Table 4.2.1), (2) the (satisfactorily or poorly) calibrated ‘pure’ Heston model, (3) the Heston-SLV model and (4) the traditional local volatility model. Given the expiry T , similar as in [97] we consider the strike prices $K_i = \exp(0.1\delta_i\sqrt{T})$ (we choose $S_0 = 1$ and zero interest rate), with $\delta_i = -1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5$. We simulate the Heston-SLV model according to the *adapted QE scheme* given by equations (4.2.9)-(4.2.12). Implied volatilities for the ‘pure’ Heston model are obtained by a standard Fourier pricing technique.

For all cases we choose $N_V = 6$. For ‘market’ Sets 1 and 2 we choose $n = 5$ basis functions. For Set 3 $n = 7$ for the poorly calibrated case and $n = 15$ for the satisfactorily calibrated case. For Set 4 these numbers are $n = 5$ and $n = 9$, respectively. As mentioned earlier, in all pricing experiments we apply the first adaptation to the method described in Section 4.2.2. Further, for Sets 3 and 4, the satisfactorily calibrated case, we make use of the second and third adaptation specified in Section 4.2.2; we choose $p_v^* = 0.01$ and $\delta = 0.1$. For Sets 1 and 3 the results are provided by Tables 4.2.2 and 4.2.3, respectively. Given the standard deviations, we observe that for both the local volatility model and the Heston-SLV model for all strike prices the reference is within the 95%-confidence interval⁹. A lower standard deviation can be obtained by increasing the number of Monte Carlo paths¹⁰. For Set 4 we report the implied volatilities in Figure 4.2.4. The results for Set 2 are essentially the same as these for Set 1 and therefore, to save some space, they are not presented.

We proceed with similar pricing experiments for the SABR-LV model; we consider the cases where the ‘pure’ SABR model is satisfactorily and poorly calibrated to the market data (generated by the Heston model with the parameters as specified in Table 4.2.1).

⁹The boundaries of the 95%-confidence interval are $\mu(\bar{\sigma}_{1,\text{model}}, \bar{\sigma}_{2,\text{model}}, \bar{\sigma}_{3,\text{model}}, \dots) \pm 1.96 \cdot \sigma(\bar{\sigma}_{1,\text{model}}, \bar{\sigma}_{2,\text{model}}, \bar{\sigma}_{3,\text{model}}, \dots)$, with $\mu(\cdot)$ and $\sigma(\cdot)$ denoting the mean and standard deviation, respectively, and $\bar{\sigma}_{i,\text{model}}$ stands for the model implied volatility (obtained from Monte Carlo) corresponding to the i th seed.

¹⁰E.g., when repeating the experiment for the Heston-SLV model ($N_V = 6$, $n = 5$), given ‘Heston market’ Set 1, with 20 seeds, $5 \cdot 10^3$ paths per seed, we obtain the errors 0.02, 0.00, 0.01, 0.01, 0.00, 0.03, 0.01 and corresponding standard deviations 0.02, 0.01, 0.01, 0.01, 0.01, 0.01, 0.01, 0.02.

Set	Ccypair	T	V_0	$\rho_{s,v}$	γ	κ	\bar{V}
1	EURGBP	1	0.01	0.23	0.21	1.50	0.01
2	EURUSD	2	0.02	-0.14	0.20	0.75	0.02
3	AUDJPY	3	0.07	-0.54	0.93	0.50	0.07
4	USDJPY	5	0.02	-0.71	0.39	0.30	0.02

Table 4.2.1: Heston parameters in typical FX markets (market data as of 16 September 2008, see [23]).

'Heston market' Set 1						
K	Sat. calibrated			Poorly calibrated		
	ϵ_H	ϵ_{H-SLV}	ϵ_{LV}	ϵ_H	ϵ_{H-SLV}	ϵ_{LV}
0.86	0.32	0.04 (0.09)	0.04 (0.11)	1.48	0.01 (0.11)	0.02 (0.09)
0.90	0.10	0.04 (0.08)	0.05 (0.09)	1.00	0.01 (0.09)	0.01 (0.08)
0.95	0.09	0.05 (0.08)	0.06 (0.09)	0.72	0.01 (0.09)	0.01 (0.09)
1.00	0.15	0.04 (0.09)	0.04 (0.09)	0.77	0.01 (0.11)	0.01 (0.10)
1.05	0.08	0.01 (0.10)	0.03 (0.10)	1.13	0.02 (0.12)	0.02 (0.11)
1.11	0.06	0.03 (0.13)	0.03 (0.12)	1.64	0.01 (0.12)	0.02 (0.15)
1.16	0.22	0.00 (0.17)	0.02 (0.14)	2.20	0.01 (0.15)	0.03 (0.19)

Table 4.2.2: Errors $\epsilon_{\text{model}} := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{model}}|$ in % corresponding to 'Heston market' Set 1. 'Sat.' stands for the satisfactorily calibrated Heston model and $\bar{\sigma}$ denotes the Black-Scholes implied volatility. Numbers in parentheses are standard deviations over the seeds.

Contrary to the Heston-SLV case, we report the results for Set 3 in Figure 4.2.5 and for Sets 2 and 4 in Tables 4.2.4 and 4.2.5, respectively. Given the standard deviations, we observe that for the SABR-LV model for all strikes – except for $K = 0.72$ and $K = 0.80$ for the 'Heston market' Set 4 – the reference is within the 95%-confidence interval. Further, for the standard SABR model, only for 'Heston market' Set 2, the satisfactorily calibrated case, the reference is within the 95%-confidence interval. Here we leave out the highly accurate results corresponding to Set 1, to save some space. All results are obtained with $N_V = 4$, $n = 5$ (Sets 1, 2) and $n = 7$ basis functions (Sets 3, 4) – these numbers correspond to the first experiment in this section. In all pricing experiments we apply the first adaptation to the method described in Section 4.2.2. Both the second and the third correction mentioned in Section 4.2.2 are *not* used. However, for Sets 3 and 4, in the calibration of the 'pure' SABR model we include a constraint on the vol-vol parameter, see Remark 4.2.2.

Remark 4.2.2 (Limitations of the SC-R approach). *Considering expiries up to 6 years, for the Heston-SLV model the stochastic collocation – regression approach yields highly accurate results for the calibrated 'pure' Heston parameters that satisfy $F := \frac{2\kappa\bar{V}}{\gamma^2} - 1 \geq -0.8$, regardless of the 'Heston market' we assume. For extreme cases for which $F \approx -0.8$ we typically choose $N_V = 6$, the number of basis functions n in the range 7 – 9 and make use of all enhancements described in Section 4.2.2, with p_v^* in the range 0.01 – 0.1 and $\delta = 0.1$. Trivially, in the calibration one can control to which extent the Feller condition is violated by imposing constraints on the parameters, such that the stochastic collocation – regression approach works without the enhancements of Section 4.2.2 and for lower numbers of*

‘Heston market’ Set 3						
K	Sat. calibrated			Poorly calibrated		
	ϵ_H	ϵ_{H-SLV}	ϵ_{LV}	ϵ_H	ϵ_{H-SLV}	ϵ_{LV}
0.77	1.14	0.01 (0.20)	0.04 (0.18)	1.09	0.13 (0.18)	0.05 (0.20)
0.84	1.26	0.05 (0.18)	0.04 (0.17)	2.17	0.09 (0.18)	0.05 (0.18)
0.92	1.35	0.07 (0.18)	0.04 (0.18)	3.19	0.06 (0.19)	0.04 (0.16)
1.00	1.33	0.11 (0.19)	0.04 (0.19)	4.01	0.03 (0.19)	0.03 (0.15)
1.09	1.03	0.14 (0.22)	0.01 (0.21)	4.30	0.01 (0.20)	0.02 (0.17)
1.19	0.34	0.07 (0.28)	0.01 (0.25)	3.66	0.05 (0.21)	0.02 (0.22)
1.30	0.44	0.10 (0.35)	0.01 (0.30)	2.18	0.06 (0.24)	0.04 (0.27)

Table 4.2.3: Errors $\epsilon_{\text{model}} := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{model}}|$ in % corresponding to ‘Heston market’ Set 3. ‘Sat.’ stands for the satisfactorily calibrated Heston model and $\bar{\sigma}$ denotes the Black-Scholes implied volatility. Numbers in parentheses are standard deviations over the seeds.

‘Heston market’ Set 2				
K	Sat. calibrated		Poorly calibrated	
	ϵ_{SABR}	$\epsilon_{\text{SABR-LV}}$	ϵ_{SABR}	$\epsilon_{\text{SABR-LV}}$
0.81	0.02 (0.23)	0.01 (0.25)	2.93 (0.21)	0.04 (0.25)
0.87	0.03 (0.18)	0.04 (0.19)	2.87 (0.14)	0.05 (0.19)
0.93	0.04 (0.14)	0.04 (0.16)	2.79 (0.11)	0.04 (0.16)
1.00	0.03 (0.13)	0.02 (0.14)	2.73 (0.10)	0.03 (0.15)
1.07	0.04 (0.14)	0.00 (0.15)	2.72 (0.11)	0.01 (0.16)
1.15	0.08 (0.14)	0.05 (0.16)	2.78 (0.12)	0.03 (0.17)
1.24	0.15 (0.17)	0.16 (0.20)	2.90 (0.14)	0.11 (0.21)

Table 4.2.4: Errors $\epsilon_{\text{model}} := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{model}}|$ in % corresponding to ‘Heston market’ Set 2. ‘Sat.’ stands for the satisfactorily calibrated Heston model and $\bar{\sigma}$ denotes the Black-Scholes implied volatility. Numbers in parentheses are standard deviations over the seeds.

collocation points and basis functions. For our approach to work for the SABR-LV model, for the ‘Heston market’ Sets 3 and 4 we need to impose in the calibration of the ‘pure’ SABR model the constraints $\gamma < 0.55$ and $\gamma < 0.4$, respectively, which seems very reasonable in practice.

4.3. LOCAL VOLATILITY MODEL WITH STOCHASTIC RATES

In this section, we present an evaluation approach for the local volatility model extended with stochastic interest rates. *The method is similar as the one employed in the stochastic local volatility context*; based on the fact that cumulative distribution functions are equally distributed, we first project $S(\cdot)$ on a standard normal random variable. Subsequently, we apply regression to approximate a conditional expectation.

As pointed out in [32], in the long-dated FX options market the effect of interest rate volatility becomes increasingly relevant for a longer expiry and may become as important as that of the FX spot volatility. Further, also in an FX context, Piterbarg [97] considers for the pricing of *Power-Reverse Dual-Currency* (PRDC) swaps, the local volatility model incorporating stochastic interest rates, assuming that the domestic and foreign

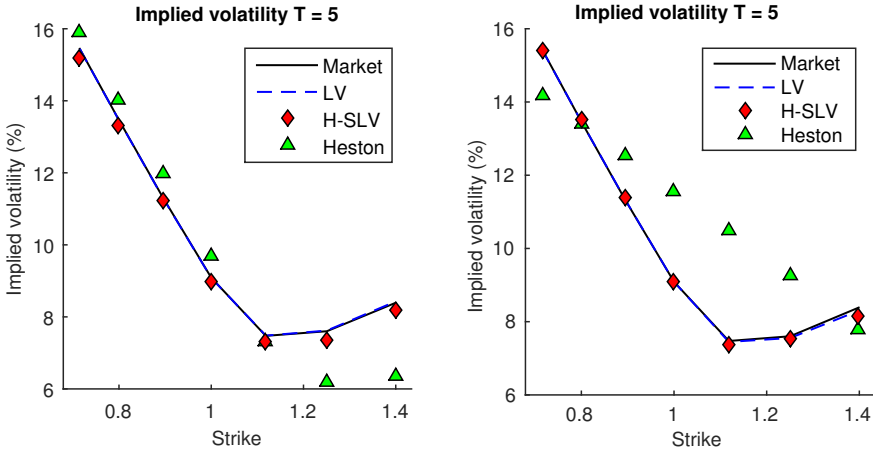


Figure 4.2.4: Black-Scholes implied volatilities corresponding to Set 4 with the satisfactorily (left) and poorly (right) calibrated Heston model. ‘H-SLV’ stands for the Heston-Stochastic Local Volatility model. Results are obtained with $2 \cdot 10^5$ paths (2 seeds, each seed constitutes 10^5 paths) and 200 time-steps per year.

interest rates follow Hull-White dynamics. He states that FX options exhibit a significant volatility skew and, moreover, that PRDC swaps, due to their structure, are highly sensitive to it. Therefore, the assumption of lognormality of the FX rates in the standard three-factor pricing model is not appropriate to price and hedge long-dated FX products. Further, as pointed out in [13], long-term callable path-dependent equity options require an appropriate modelling of the underlying asset process and, moreover, the early-exercise feature – in particular for a large time-span – suggests interest rates risk. Enhancing the local volatility model with stochastic interest rates is also the subject of research in [8, 61, 104], amongst others.

Similar in e.g. [13, 32, 97], let the interest rate be governed by Hull-White dynamics, which is also the case in the traditional three-factor pricing model. Under the risk-neutral \mathbb{Q} -measure, the dynamics of the *Local Volatility-Hull White (LV-HW) model* are given by the following system of equations (see e.g. [8]):

$$dS(t)/S(t) = r(t)dt + \sigma(t, S(t))dW_s^{\mathbb{Q}}(t), \quad (4.3.1)$$

$$dr(t) = \lambda(\theta(t) - r(t))dt + \eta dW_r^{\mathbb{Q}}(t), \quad (4.3.2)$$

with $dW_s^{\mathbb{Q}}(t)dW_r^{\mathbb{Q}}(t) = \rho_{r,s}dt$. In the interest rate process the speed of mean reversion λ and the volatility coefficient η are related with the time-dependent term structure function $\theta(\cdot)$ via $\theta(t) = f(0, t) + \frac{1}{\lambda} \frac{\partial}{\partial t} f(0, t) + \frac{\eta^2}{2\lambda^2} (1 - e^{-2\lambda t})$ (the expression for $\theta(\cdot)$ is obtained by decomposing the Hull-White model, see e.g. [95]), which yields a model fit with the initial yield curve, where $f(0, t)$ denotes the initial instantaneous forward rate corresponding to expiry t , defined by $f(0, t) = -\partial \log(P(0, t)) / \partial t$ and $P(0, t)$ is the current value of the zero-coupon bond, which is given by

$$P(0, t) = \exp\left(-\int_0^t \psi(s)ds + A(t)\right), \quad (4.3.3)$$

‘Heston market’ Set 4				
K	Sat. calibrated		Poorly calibrated	
	ϵ_{SABR}	$\epsilon_{\text{SABR-LV}}$	ϵ_{SABR}	$\epsilon_{\text{SABR-LV}}$
0.72	1.08 (0.21)	0.54 (0.22)	2.44 (0.22)	0.44 (0.21)
0.80	0.78 (0.14)	0.32 (0.13)	1.98 (0.14)	0.25 (0.13)
0.89	0.32 (0.09)	0.11 (0.08)	1.33 (0.08)	0.10 (0.08)
1.00	0.35 (0.06)	0.10 (0.06)	0.46 (0.05)	0.03 (0.06)
1.12	0.73 (0.06)	0.10 (0.08)	0.14 (0.06)	0.04 (0.07)
1.25	0.10 (0.09)	0.01 (0.09)	0.32 (0.08)	0.00 (0.09)
1.40	0.44 (0.13)	0.11 (0.14)	0.82 (0.13)	0.09 (0.14)

Table 4.2.5: Errors $\epsilon_{\text{model}} := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{model}}|$ in % corresponding to ‘Heston market’ Set 4. ‘Sat.’ stands for the satisfactorily calibrated Heston model and $\bar{\sigma}$ denotes the Black-Scholes implied volatility. Numbers in parentheses are standard deviations over the seeds.

with $\psi(t) = r_0 e^{-\lambda t} + \lambda \int_0^t \theta(s) e^{-\lambda(t-s)} ds$ and $A(t) = \frac{\eta^2}{2\lambda^3} (\lambda t - 2(1 - e^{-\lambda t}) + \frac{1}{2}(1 - e^{-2\lambda t}))$.

From the expression for the instantaneous forward rate, the initial interest rate r_0 is implied by the identity $r(t) = f(t, t)$. Further, the local volatility component reads

$$\sigma^2(t, K) = \frac{\frac{\partial C(t, K)}{\partial t} - K \mathbb{E}^{\mathbb{Q}} \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t) > K} \middle| \mathcal{F}(t_0) \right]}{\frac{1}{2} K^2 \frac{\partial^2 C(t, K)}{\partial K^2}}, \quad (4.3.4)$$

where $M(\cdot)$ denotes the value of the money account, determined by $dM(t) = r(t)M(t)dt$. As we always consider $t_0 = 0$, we leave out the filtration for notational purposes from now on.

In the local volatility component the expectation $\mathbb{E}^{\mathbb{Q}} \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t) > K} \right]$ is problematic in a calibration sense, as no direct link with the market quotes can be observed [32]. Also, no analytical expressions for the joint distribution of $r(t)/M(t)$ and $S(t)$ are available. Further, the discretization scheme suggests that *for each time-step* in the simulation, the expectation, which in principle is a deterministic function of $s_{i,j}$, needs to be evaluated *for each path*. This is expensive and undesirable. In the following section we present a novel approach for the evaluation of the expectation, which is both efficient and accurate.

4.3.1. ESTABLISHING $\mathbb{E}^{\mathbb{Q}} \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t) > K} \right]$

In this section, we determine an approximation for the *non-trivial* expectation in the local volatility component (4.3.4). Similar to the approach for evaluating stochastic local volatility models, the method essentially consists of two projection steps. We first apply a projection on a standard normal random variable, employing the equality in distribution of cumulative distribution functions, and subsequently we make use of ordinary least squares regression.

We start by applying a change of measure:

$$\mathbb{E}^{\mathbb{Q}} \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t) > K} \right] = P(0, t) \mathbb{E}^t \left[r(t) \mathbb{1}_{S(t) > K} \right]. \quad (4.3.5)$$

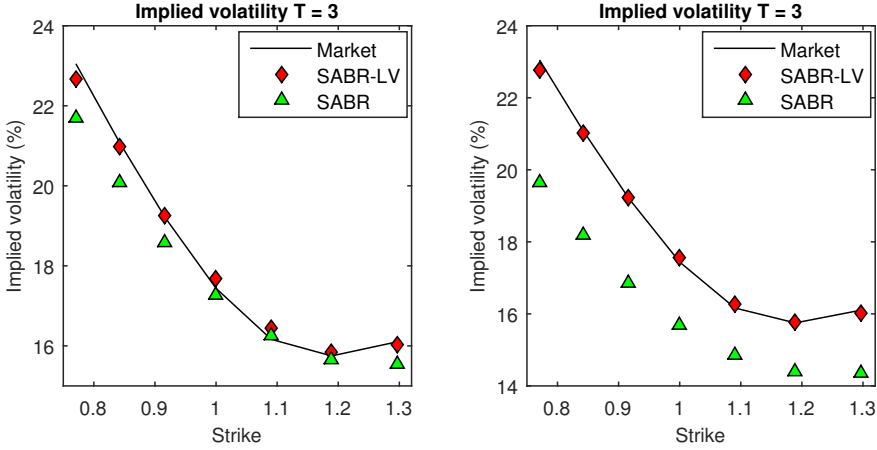


Figure 4.2.5: Black-Scholes implied volatilities corresponding to Set 3 with the satisfactorily (left) and poorly (right) calibrated SABR model. Results are obtained with $2 \cdot 10^5$ paths (2 seeds, each seed constitutes 10^5 paths) and 200 time-steps per year.

At the right-hand side, under the t -forward measure, $r(\cdot)$ is normally distributed with mean

$$\mu_r^t(t) = r(0)e^{-\lambda t} + \int_0^t \tilde{\theta}(u)e^{-\lambda(t-u)} du, \quad \tilde{\theta}(u) := \lambda\theta(u) + \frac{\eta^2}{\lambda} \left(e^{-\lambda(t-u)} - 1 \right),$$

and standard deviation

$$\sigma_r^t(t) = \left(\frac{\eta^2}{2\lambda} \left(1 - e^{-2\lambda t} \right) \right)^{1/2}.$$

Further, the CDF of $S(t)$ under the t -forward measure, denoted by $F_{S(t)}^t(\cdot)$, can be derived from the following well-known relation (see e.g. [46]):

$$\frac{\partial C(t, K)}{\partial K} = -P(0, t) \mathbb{Q}^t [S(t) > K],$$

where $C(t, K)$ is the price at $t = 0$ ('today') of a European call option with maturity t and strike K and $P(0, t)$ denotes the zero-coupon bond with expiry t . This relation directly implies

$$F_{S(t)}^t(K) = 1 - \mathbb{Q}^t [S(t) > K] = 1 + \frac{1}{P(0, t)} \frac{\partial C(t, K)}{\partial K}.$$

To evaluate the expectation at the right-hand side of equation (4.3.5), for a fixed t we project $S(t)$ onto a standard normal distribution $X \stackrel{d}{=} \mathcal{N}(0, 1)$ via the function $g(\cdot)$, defined by¹¹ $g(\cdot) := F_{S(t)}^{-1}(F_X(\cdot))$, which ensures $S(t) \stackrel{d}{=} g(X)$ and, moreover, for elements $S(t) = s$ and $X = x$: $s = g(x)$. This element-wise equality implies $x = g^{-1}(s) = F_X^{-1}(F_{S(t)}^t(s))$,

¹¹For notation purposes we suppress the t -superscript in the inverse of the CDF of $S(t)$.

which yields for the expectation in (4.3.5):

$$\mathbb{E}^{\mathbb{Q}} \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t) > K} \right] = P(0, t) \mathbb{E}^t [r(t) \mathbb{1}_{g(X) > K}] = P(0, t) \mathbb{E}^t [r(t) \mathbb{1}_{X > g^{-1}(K)}]. \quad (4.3.6)$$

We proceed with the second projection step. Trivially, from (4.3.6) we write (as $F_X(g^{-1}(K)) = F_{S(t)}^t(K)$),

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t) > K} \right] &= P(0, t) \mathbb{E}^t [r(t) | X > g^{-1}(K)] (1 - F_X(g^{-1}(K))) \\ &= P(0, t) \left(\mu_r^t(t) + \sigma_r^t(t) \mathbb{E}^t [Z | X > g^{-1}(K)] \right) (1 - F_{S(t)}^t(K)), \end{aligned}$$

where Z is a standard normal random variable. *Similar to the approach presented in the stochastic local volatility setting*, to evaluate the conditional expectation, we apply a projection on a set of basis functions $\psi_k(\cdot)$, $k = 1, 2, \dots, n$ ($n < \infty$) which depend on X . We again choose the simple set of orthogonal polynomials $\{1, x, x^2, \dots, x^{n-1}\}$ and apply OLS regression to compute the corresponding coefficients, which yields

$$\begin{aligned} \mathbb{E} [Z | X > g^{-1}(K)] &= \hat{\beta}_0 + \hat{\beta}_1 \mathbb{E} [X | X > g^{-1}(K)] + \hat{\beta}_2 \mathbb{E} [X^2 | X > g^{-1}(K)] \\ &\quad + \dots + \hat{\beta}_{n-1} \mathbb{E} [X^{n-1} | X > g^{-1}(K)] + \epsilon. \end{aligned}$$

The truncated moments of X allow for an analytic evaluation, which we state in Result 4.3.1.

Result 4.3.1 (Moments of a truncated standard normal random variable). *Given $X \stackrel{d}{=} \mathcal{N}(0, 1)$ and define $m_i := \mathbb{E} [X^i | X > a]$, with $a \in \mathbb{R}$. The truncated moments are given by*

$$m_i = (i-1)m_{i-2} + \frac{a^{i-1} f_{\mathcal{N}(0,1)}(a)}{1 - F_{\mathcal{N}(0,1)}(a)}, \quad i = 1, 2, \dots,$$

with $m_{-1} = 0$, $m_0 = 1$ and $f_{\mathcal{N}(0,1)}(\cdot)$ and $F_{\mathcal{N}(0,1)}$ denoting the standard normal probability density and cumulative distribution functions, respectively.

Combining results yields

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t) > K} \right] &= P(0, t) \left(\mu_r^t(t) + \sigma_r^t(t) \left(\hat{\beta}_0 + \hat{\beta}_1 \mathbb{E} [X | X > g^{-1}(K)] + \hat{\beta}_2 \mathbb{E} [X^2 | X > g^{-1}(K)] \right. \right. \\ &\quad \left. \left. + \dots + \hat{\beta}_{n-1} \mathbb{E} [X^{n-1} | X > g^{-1}(K)] \right) \right) (1 - F_{S(t)}^t(K)) + \epsilon. \end{aligned} \quad (4.3.7)$$

The error ϵ is introduced in the regression step and is discussed in Appendix 4.A. In Section 4.2.3 we demonstrated a decrease of this error with respect to the number of Monte Carlo realizations.

ALTERNATIVE APPROACH

Instead of applying regression, an alternative approach is based on the *assumption* that $r(\cdot)$ and X in (4.3.6) are governed by a joint bivariate normal distribution. For two jointly

normally distributed random variables $X_1 \stackrel{d}{=} \mathcal{N}(\mu_1, \sigma_1)$ and $X_2 \stackrel{d}{=} \mathcal{N}(\mu_2, \sigma_2)$, correlated with correlation parameter ρ , the following result holds:

$$\mathbb{E} \left[X_1 \mathbb{1}_{X_2 > k} \right] = \left(\mu_1 + \rho \sigma_1 \frac{f_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)}{1 - F_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)} \right) (1 - F_{X_2}(k)), \quad (4.3.8)$$

where $f_{\mathcal{N}(0,1)}(\cdot)$ and $F_{\mathcal{N}(0,1)}(\cdot)$ are the standard normal PDF and CDF, respectively, and $F_{X_2}(\cdot)$ is the CDF corresponding to the random variable X_2 . A proof of this result is given in Appendix 4.B.

By the result in (4.3.8) the expectation in (4.3.6) is approximated as follows:

$$\mathbb{E}^t \left[r(t) \mathbb{1}_{X > g^{-1}(K)} \right] = \left(\mu_r^t(t) + \rho_{r,X}^t(t) \sigma_r^t(t) \frac{f_{\mathcal{N}(0,1)}(g^{-1}(K))}{1 - F_{\mathcal{N}(0,1)}(g^{-1}(K))} \right) (1 - F_X(g^{-1}(K))) + \epsilon,$$

with $g^{-1}(K) = F_X^{-1}(F_{S(t)}^t(K))$, where the error term ϵ is introduced by assuming that $r(t)$ and X are jointly bivariate normally distributed under the t -forward measure. Further, as $F_X(g^{-1}(K)) = F_{S(t)}^t(K)$, we have

$$\mathbb{E}^Q \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t) > K} \right] = P(0, t) \left(\mu_r^t(t) + \rho_{r,X}^t(t) \sigma_r^t(t) \frac{f_{\mathcal{N}(0,1)}(g^{-1}(K))}{1 - F_{S(t)}^t(K)} \right) (1 - F_{S(t)}^t(K)) + \epsilon, \quad (4.3.9)$$

where the correlation parameter is numerically (i.e. based on the Monte Carlo paths) established by applying a change of measure:

$$\rho_{r,X}^t(t) \stackrel{\text{def}}{=} \frac{\mathbb{E}^t[r(t)X] - \mathbb{E}^t[r(t)]\mathbb{E}^t[X]}{\sigma_r^t(t)\sigma_X^t(t)} = \frac{\mathbb{E}^t[r(t)X]}{\sigma_r^t(t)} = \frac{1}{P(0, t)\sigma_r^t(t)} \mathbb{E}^Q \left[\frac{r(t)}{M(t)} X \right],$$

with $X = g^{-1}(K)$, $P(0, t)$ is defined in (4.3.3) and $\sigma_r^t(t) = \left(\frac{\eta^2}{2\lambda} (1 - e^{-2\lambda t}) \right)^{1/2}$.

A comparison of (4.3.7) and (4.3.9) makes clear that the latter can be considered as a special case of the more generic expression in (4.3.7). We apply the SC-R approach in (4.3.7) or the alternative in (4.3.9) in a Monte Carlo simulation framework according to Algorithm 4. As for Algorithm 3, the indices $i = 1, 2, \dots, M$ and $j = 1, 2, \dots, N$ denote the time-step and path, respectively.

Our numerical experiments in Section 4.3.2 indicate that the approximations (4.3.7) and (4.3.9) show a similar performance for the shorter expiries, whereas for the longer expiry dates the former outperforms the latter. The reason for this is the fact that the *error* due to the bivariate normality assumption becomes more pronounced for longer expiries, i.e. the joint distribution of $r(\cdot)$ and X in (4.3.6) then resembles less a bivariate normal distribution.

4.3.2. NUMERICAL EXPERIMENTS

In this section we test the accuracy of the approximation in (4.3.7) and the alternative (4.3.9). We price European call options by means of a Monte Carlo simulation of the Local Volatility-Hull White model. The Monte Carlo simulation consists of $2 \cdot 10^5$ paths

```

for each time-step  $t_i, i = 1, 2, \dots, M$  do
1  Generate  $N$  pairs  $(s_{i,j}, r_{i,j}), j = 1, 2, \dots, N$  by going forward one time-step in the
   standard Euler discretization scheme of the Local Volatility-Hull White model.
2  Compute  $\mathbb{E}^{\mathbb{Q}} \left[ \frac{r(t_i)}{M(t_i)} \mathbb{1}_{S(t_i) > s_{i,j}} \right]$  according to either (4.3.7) or the alternative in (4.3.9).
3  Establish the local volatility component  $\sigma^2(t_i, s_{i,j})$  by equation (4.3.4) – use its value in
   step 1.
end
4  Price European call options based on the obtained values for  $S(\cdot)$  at the time to maturity.
    
```

Algorithm 4: Pricing European call options by a Monte Carlo simulation of the LV-HW model, incorporating the SC-R approach (4.3.7) or the alternative (4.3.9).

Set	T	λ	η	$\rho_{r,s}$
A	1, 2, 5	0.01	0.01	0.6
B	5, 10, 15	0.01	0.007	-0.15

Table 4.3.1: Hull-White model parameters as in [55] (Set A) and [56, 97] (Set B).

Set	V_0	$\rho_{s,v}$	γ	κ	\bar{V}
1	0.04	-0.9	1	0.5	0.04
2	0.04	-0.5	0.9	0.3	0.04
3	0.09	-0.3	1	1	0.09

Table 4.3.2: ‘Heston market’ parameters as in [5].

(20 seeds, each seed constitutes 10^4 paths) and 200 time-steps per year, unless otherwise mentioned. At each time-step in the simulation we either use the approximation (4.3.7) or approximate the expectation according to (4.3.9). We consider two sets of Hull-White parameters in the literature, see Table 4.3.1. In line from the literature we choose for both sets $r_0 = 0.02$.

We generate synthetic market data by applying Fourier techniques to the Heston model, which we assume to be calibrated perfectly to the market. We choose three sets of Heston parameters for which the Feller condition is strongly violated, namely the parameter sets presented in [5], see Table 4.3.2. In the regression we choose $n = 5$ basis functions, so we consider the first four moments of the truncated standard normal distribution. Similar to [97], given the expiry T we consider the strikes $K_i = F_0^T \exp(0.1\delta_i\sqrt{T})$, with $F_0^T = S_0/P(0, T) = 1/P(0, T)$ (as $S_0 = 1$) denoting the initial forward and $\delta_i = -1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5$.

For the two shortest expiries per Hull-White set the results are provided by Tables 4.3.3, 4.3.4 (‘Heston market’ Set 1), 4.3.5, 4.3.6 (‘Heston market’ Set 2), 4.3.7 and 4.3.8 (‘Heston market’ Set 3). We report the absolute error $\epsilon := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{LV-HW}}|$ in %, with $\bar{\sigma}$ denoting the Black-Scholes implied volatility. The error $\epsilon_{\text{alternative}}$ corresponds to the alternative approach of Section 4.3.1. Given the standard deviations, we observe that for all LV-HW experiments both the SC-R approach and the alternative yield 95%-confidence intervals¹² that cover the reference implied volatility, except for one case: the alternative approach with ‘Heston market’ Set 1, Hull-White Set B, $T = 10$. For the expiries $T = 5$ and $T = 15$ corresponding to Hull-White Sets A and B, respectively, the results are displayed

¹²The boundaries of the 95%-confidence interval are $\mu(\bar{\sigma}_{1,\text{model}}, \bar{\sigma}_{2,\text{model}}, \bar{\sigma}_{3,\text{model}}, \dots) \pm 1.96 \cdot \sigma(\bar{\sigma}_{1,\text{model}}, \bar{\sigma}_{2,\text{model}}, \bar{\sigma}_{3,\text{model}}, \dots)$, with $\mu(\cdot)$ and $\sigma(\cdot)$ denoting the mean and standard deviation, respectively, and $\bar{\sigma}_{i,\text{model}}$ stands for the model implied volatility (obtained from Monte Carlo) corresponding to the i th seed.

‘Heston market’ Set 1, Hull-White Set A					
T = 1			T = 2		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.88	0.00 (0.27)	0.01 (0.25)	0.84	0.01 (0.22)	0.03 (0.21)
0.92	0.01 (0.23)	0.01 (0.22)	0.90	0.01 (0.20)	0.02 (0.18)
0.97	0.02 (0.21)	0.02 (0.19)	0.97	0.01 (0.18)	0.01 (0.16)
1.02	0.04 (0.17)	0.04 (0.16)	1.04	0.03 (0.15)	0.01 (0.13)
1.07	0.05 (0.13)	0.05 (0.11)	1.12	0.04 (0.11)	0.06 (0.08)
1.13	0.05 (0.10)	0.05 (0.09)	1.20	0.01 (0.12)	0.02 (0.10)
1.19	0.04 (0.11)	0.04 (0.10)	1.29	0.02 (0.15)	0.03 (0.15)

Table 4.3.3: Errors $\epsilon := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{LV-HW}}|$ in % corresponding to the ‘Heston market’ Set 1, Hull-White Set A ($\bar{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

4

‘Heston market’ Set 1, Hull-White Set B					
T = 5			T = 10		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.79	0.04 (0.15)	0.13 (0.14)	0.76	0.01 (0.11)	0.38 (0.10)
0.88	0.03 (0.13)	0.11 (0.12)	0.89	0.00 (0.11)	0.31 (0.08)
0.90	0.02 (0.12)	0.07 (0.10)	1.04	0.01 (0.10)	0.21 (0.08)
1.11	0.00 (0.09)	0.02 (0.08)	1.22	0.03 (0.09)	0.05 (0.07)
1.24	0.01 (0.08)	0.09 (0.07)	1.43	0.06 (0.08)	0.26 (0.06)
1.38	0.01 (0.11)	0.10 (0.10)	1.68	0.01 (0.08)	0.50 (0.05)
1.55	0.02 (0.16)	0.03 (0.14)	1.96	0.01 (0.18)	0.28 (0.14)

Table 4.3.4: Errors $\epsilon := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{LV-HW}}|$ in % corresponding to the ‘Heston market’ Set 1, Hull-White Set B ($\bar{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

in Figures 4.3.1 and 4.3.2. With ‘LV-HW alt.’ we denote the alternative approach.

In general, for both the approximation (4.3.7) and its alternative (4.3.9) the results are highly satisfactory. For the shorter expiries the two methods show a comparable performance, however for $T = 10$ and $T = 15$ we observe that the regression-based approach outperforms the alternative – we clearly observe this in Figure 4.3.2. The reason for this is that the alternative approach relies on the bivariate normality assumption of $r(\cdot)$ and X in (4.3.6). The error introduced by the bivariate normality assumption becomes more pronounced for longer expiries, since the joint distribution of $r(\cdot)$ and X resembles less a bivariate normal distribution when going forward in time. In the left-hand plot of Figure 4.3.2 we observe a slight mismatch at the right-hand side of the strike range. This is *not* due to the performance of the approximation methods, but due to general Monte Carlo bias. We conclude this based on a simulation of the LV-HW model applying an ‘exact’, ‘brute-force’ approach to compute the expectation $\mathbb{E}^{\mathbb{Q}} \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t) > K} \right]$. The same bias was observed.

4.4. CONCLUSION

In this chapter we considered in a Monte Carlo simulation framework two classes of hybrid local volatility models, namely stochastic local volatility models and the local

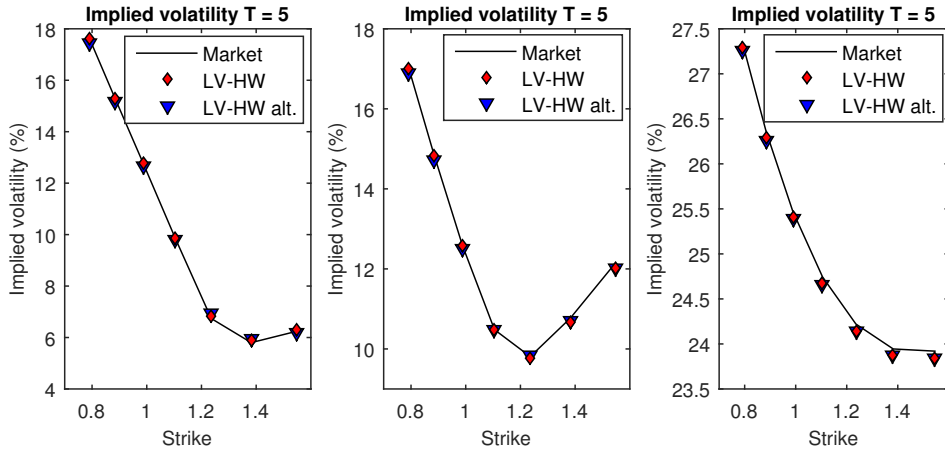


Figure 4.3.1: Black-Scholes implied volatilities corresponding to Hull-White Set A, $T = 5$ and the ‘Heston market’ Sets 1 (left), 2 (middle) and 3 (right), using 200 time-steps per year for Sets 1 and 2. For Set 3 500 time-steps per year are used, as 200 time-steps did not yield highly satisfactory results. ‘LV-HW alt.’ denotes the alternative approach presented in Section 4.3.1. Results are obtained with $2 \cdot 10^5$ paths (2 seeds, each seed constitutes 10^5 paths).

‘Heston market’ Set 2, Hull-White Set A					
$T = 1$			$T = 2$		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.88	0.09 (0.27)	0.09 (0.27)	0.84	0.05 (0.23)	0.07 (0.23)
0.92	0.08 (0.24)	0.07 (0.24)	0.90	0.05 (0.21)	0.06 (0.21)
0.97	0.06 (0.21)	0.06 (0.20)	0.97	0.05 (0.19)	0.05 (0.18)
1.02	0.05 (0.17)	0.04 (0.16)	1.04	0.04 (0.18)	0.03 (0.17)
1.07	0.05 (0.16)	0.04 (0.16)	1.12	0.05 (0.20)	0.04 (0.19)
1.13	0.05 (0.17)	0.05 (0.17)	1.20	0.09 (0.24)	0.08 (0.24)
1.19	0.08 (0.19)	0.07 (0.19)	1.29	0.14 (0.31)	0.13 (0.31)

Table 4.3.5: Errors $\epsilon := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{LV-HW}}|$ in % corresponding to the ‘Heston market’ Set 2, Hull-White Set A ($\bar{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

‘Heston market’ Set 2, Hull-White Set B					
$T = 5$			$T = 10$		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.79	0.08 (0.20)	0.11 (0.20)	0.76	0.05 (0.15)	0.20 (0.15)
0.88	0.07 (0.18)	0.09 (0.18)	0.89	0.04 (0.14)	0.14 (0.15)
0.90	0.05 (0.16)	0.05 (0.16)	1.04	0.01 (0.14)	0.04 (0.14)
1.11	0.03 (0.16)	0.00 (0.16)	1.22	0.02 (0.15)	0.09 (0.16)
1.24	0.04 (0.18)	0.00 (0.19)	1.43	0.00 (0.20)	0.13 (0.19)
1.38	0.06 (0.24)	0.04 (0.24)	1.68	0.05 (0.28)	0.05 (0.27)
1.55	0.10 (0.32)	0.09 (0.32)	1.96	0.06 (0.37)	0.05 (0.35)

Table 4.3.6: Errors $\epsilon := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{LV-HW}}|$ in % corresponding to the ‘Heston market’ Set 2, Hull-White Set B ($\bar{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

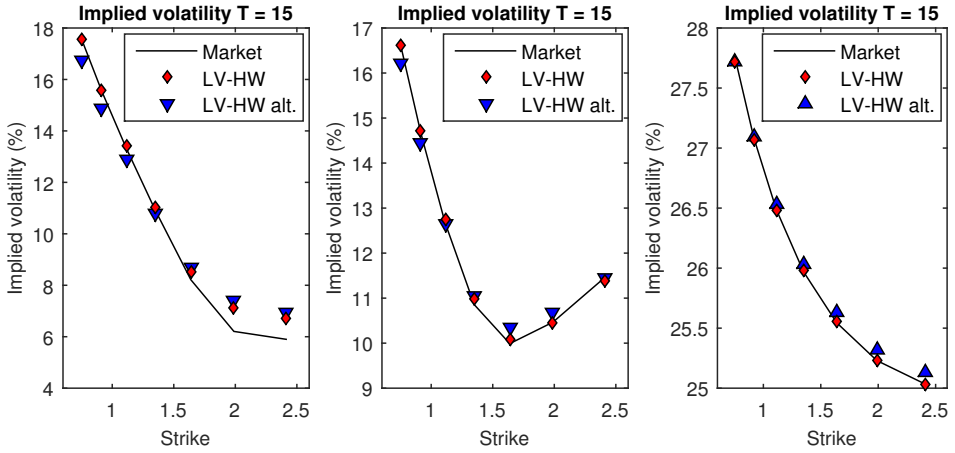


Figure 4.3.2: Black-Scholes implied volatilities corresponding to Hull-White Set B, $T = 15$ and the ‘Heston market’ Sets 1 (left), 2 (middle) and 3 (right). ‘LV-HW alt.’ denotes the alternative approach presented in Section 4.3.1. Results are obtained with $2 \cdot 10^5$ paths (2 seeds, each seed constitutes 10^5 paths) and 200 time-steps per year.

‘Heston market’ Set 3, Hull-White Set A					
$T = 1$			$T = 2$		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.88	0.14 (0.26)	0.14 (0.26)	0.84	0.09 (0.29)	0.10 (0.29)
0.92	0.13 (0.24)	0.13 (0.24)	0.90	0.08 (0.28)	0.09 (0.28)
0.97	0.12 (0.23)	0.11 (0.23)	0.97	0.08 (0.28)	0.08 (0.27)
1.02	0.11 (0.24)	0.11 (0.24)	1.04	0.08 (0.29)	0.08 (0.28)
1.07	0.10 (0.25)	0.10 (0.25)	1.12	0.08 (0.31)	0.08 (0.31)
1.13	0.09 (0.27)	0.09 (0.27)	1.20	0.09 (0.34)	0.09 (0.34)
1.19	0.09 (0.29)	0.09 (0.29)	1.29	0.11 (0.37)	0.11 (0.37)

Table 4.3.7: Errors $\epsilon := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{LV-HW}}|$ in % corresponding to the ‘Heston market’ Set 3, Hull-White Set A ($\bar{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

‘Heston market’ Set 3, Hull-White Set B					
$T = 5$			$T = 10$		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.79	0.04 (0.27)	0.04 (0.28)	0.76	0.01 (0.35)	0.01 (0.38)
0.88	0.03 (0.28)	0.03 (0.29)	0.89	0.00 (0.36)	0.01 (0.38)
0.90	0.03 (0.30)	0.03 (0.30)	1.04	0.00 (0.38)	0.01 (0.40)
1.11	0.03 (0.31)	0.03 (0.32)	1.22	0.02 (0.40)	0.01 (0.43)
1.24	0.04 (0.32)	0.03 (0.33)	1.43	0.02 (0.43)	0.01 (0.45)
1.38	0.05 (0.34)	0.04 (0.35)	1.68	0.03 (0.46)	0.01 (0.48)
1.55	0.05 (0.38)	0.04 (0.39)	1.96	0.04 (0.49)	0.01 (0.52)

Table 4.3.8: Errors $\epsilon := |\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{LV-HW}}|$ in % corresponding to the ‘Heston market’ Set 3, Hull-White Set B ($\bar{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

volatility model extended with stochastic interest rates. For both model classes a *non-trivial* (conditional) expectation needs to be evaluated, which cannot be extracted from the market quotes and is expensive to compute. In this chapter we presented a novel, efficient approach to the evaluation of these expectations. The method essentially consists of two projection steps; the first projection employs the equality in distribution of cumulative distribution functions, which stands at the basis of the stochastic collocation method, the second projection step relies on standard regression techniques. By means of numerical experiments we confirm that our approach facilitates an efficient Monte Carlo evaluation and yields highly accurate pricing results for European-type options.

APPENDIX

4.A. ERROR ANALYSIS & DISCUSSION

In this section we briefly discuss the (asymptotics of the) errors in (4.2.24), which are due to the stochastic collocation method ($\bar{\epsilon}_1(\cdot)$) and the projection on an orthonormal basis and subsequently applying regression ($\bar{\epsilon}_2$).

4.A.1. STOCHASTIC COLLOCATION ERROR

The first error is $\bar{\epsilon}_1(K) := \mathbb{E}[\epsilon_1(Z) | X = g^{-1}(K)]$, with Z denoting a standard normal random variable. It is introduced by projecting, for a given t , the random variable $V(t)$ on Z via a Lagrange polynomial $h_{N_V}(\cdot)$, which interpolates through the collocation values $v_i = F_{V(t)}^{-1}(F_Z(z_i))$, where the *collocation points* z_i are chosen in an *optimal way*, namely based on the zeros of Hermite polynomials.

We start the analysis of $\bar{\epsilon}_1(\cdot)$ by the following (in)equalities:

$$|\mathbb{E}[h(Z) - h_{N_V}(Z)]| \leq \mathbb{E}[|h(Z) - h_{N_V}(Z)|] \quad (\text{by Jensen's inequality}) \quad (4.A.1)$$

$$\leq \left(\mathbb{E}[(h(Z) - h_{N_V}(Z))^2] \right)^{1/2}, \quad (4.A.2)$$

where the latter inequality is a standard relation between L^p -norms, see e.g. [112].

As pointed out in the error analysis in [58], the advantage of using optimal collocation points is that the stochastic collocation method can be connected to the computation of integrals by Gauss quadrature, which for the general function $\Psi(\cdot)$, weight function $f_Z(\cdot)$ and quadrature weights ω_i , $i = 1, 2, \dots, N_V$ reads:

$$\mathbb{E}[\Psi(Z)] = \int_{\mathbb{R}} \Psi(z) f_Z(z) dz = \sum_{i=1}^{N_V} \Psi(z_i) \omega_i + \epsilon_{N_V}. \quad (4.A.3)$$

In this chapter we choose the collocation variable $Z \stackrel{d}{=} \mathcal{N}(0, 1)$, for which a simple relation between the ‘stochastic collocation pairs’ $\{z_i, \omega_i\}_{i=1}^{N_V}$ and the Gauss-Hermite quadrature pairs $\{z_i^H, \omega_i^H\}_{i=1}^{N_V}$ exists. Whereas in the stochastic collocation method the weight function $f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}z^2)$ is used, Gauss-Hermite quadrature is based on $f_Z(z) = \exp(-z^2)$. However, standard calculus yields $\int_{\mathbb{R}} \Psi(z) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \int_{\mathbb{R}} \Psi(z\sqrt{2}) \frac{1}{\sqrt{\pi}} e^{-z^2} dz$. From this one can show the relations $z_i^H = z_i / \sqrt{2}$ and $\omega_i^H = \omega_i \sqrt{\pi}$, which implies that for a standard normal collocation variable the error due to stochastic collocation is given by the error for the Gauss-Hermite quadrature, which is, see e.g. [3], given by:

$$\epsilon_{N_V} = \frac{N_V! \sqrt{\pi}}{2^{N_V}} \frac{\Psi^{(2N_V)}(\hat{\xi})}{(2N_V)!}, \quad (4.A.4)$$

with $\hat{\xi} \in [\min(\mathbf{z}) \max(\mathbf{z})]$, $\mathbf{z} = (z_1, z_2, \dots, z_{N_V})$. By choosing $\Psi(\mathbf{z}) = (h(\mathbf{z}) - h_{N_V}(\mathbf{z}))^2$, with $h(\cdot) = F_{V(t)}^{-1}(F_Z(\cdot))$ and $h_{N_V}(\cdot)$ the corresponding approximating Lagrange polynomial, equation (4.A.3) yields the error:

$$\begin{aligned} \mathbb{E}[(h(Z) - h_{N_V}(Z))^2] &= \int_{\mathbb{R}} (h(\mathbf{z}) - h_{N_V}(\mathbf{z}))^2 f_Z(\mathbf{z}) d\mathbf{z} \\ &= \sum_{i=1}^{N_V} (h(z_i) - h_{N_V}(z_i))^2 \omega_i + \epsilon_{N_V} \\ &= \epsilon_{N_V}, \end{aligned}$$

as $h(z_i) = h_{N_V}(z_i)$. $\Psi(\mathbf{z}) = (h(\mathbf{z}) - h_{N_V}(\mathbf{z}))^2$ can be written more explicitly as the square of the standard Lagrange interpolation error, see e.g. [3]:

$$\Psi(\mathbf{z}) = \left(\frac{1}{N_V!} \frac{d^{N_V} h(\mathbf{z})}{d\mathbf{z}^{N_V}} \Big|_{\mathbf{z}=\hat{\xi}} \prod_{i=1}^{N_V} (z - z_i) \right)^2, \quad (4.A.5)$$

with $\hat{\xi} \in [\min(\mathbf{z}) \max(\mathbf{z})]$, $\mathbf{z} = (z_1, z_2, \dots, z_{N_V})$. This error may be bounded by choosing $\mathbf{z} = \hat{\xi}$ for which $\left| \frac{d^{N_V} h(\mathbf{z})}{d\mathbf{z}^{N_V}} \right|$ attains its maximum.

Substituting (4.A.5) in (4.A.4) yields a complete specification for ϵ_{N_V} , and it can be shown that it converges to zero as $N_V \rightarrow \infty$ (under the condition that $\Psi(\cdot)$ is sufficiently smooth). Then, from the inequalities (4.A.1) and (4.A.2), combined with the identity

$$\mathbb{E}[h(Z) - h_{N_V}(Z)] = \int_{\mathbb{R}} \mathbb{E}[h(Z) - h_{N_V}(Z) | X = x] f_X(x) dx,$$

the error $\bar{\epsilon}_1(x) := \mathbb{E}[h(Z) - h_{N_V}(Z) | X = x]$ converges to zero as $N_V \rightarrow \infty$, for arbitrary $x \in \mathbb{R}$.

4.A.2. REGRESSION ERROR

The second error term $\bar{\epsilon}_2 := \sum_{p=1}^{N_V-1} \hat{\epsilon}_{2p}$ is due to the projection of the unknown conditional expectations $\mathbb{E}[Z^p | X = g^{-1}(K)]$ on a set of basis functions $\{1, x, x^2, \dots, x^{n-1}\}$ and applying OLS regression. This is polynomial regression, which is a special case of multiple linear regression. In a Monte Carlo simulation framework, given N observations for the underlying $S(\cdot)$, s_j , we write, as we evaluate the local volatility component in $K = s_j$:

$$\mathbb{E}[Z^p | X = x_j] = b_{0p} + b_{1p}x_j + \dots + b_{n-1,p}x_j^{n-1} + \epsilon_{2pj}, \quad j = 1, 2, \dots, N, \quad (4.A.6)$$

with $x_j := g^{-1}(s_j)$ and ϵ_{2pj} is the *unobserved* error term corresponding to the j th realization. We apply OLS regression to compute $\hat{\beta}_{kp}$, $k = 0, 1, \dots, n-1$, $p = 1, 2, \dots, N_V-1$, which are estimates for b_{kp} . This yields $\hat{\mathbb{E}}[Z^p | X = x_j] = \hat{\beta}_{0p} + \hat{\beta}_{1p}x_j + \dots + \hat{\beta}_{n-1,p}x_j^{n-1} + \hat{\epsilon}_{2pj}$, $j = 1, 2, \dots, N$.

Let $\mathbf{x} = (1 \ x \ x^2 \ \dots \ x^{n-1})$ be an $N \times n$ matrix and denote its j th row by \mathbf{x}_j . Under some standard assumptions for the regression model (4.A.6), e.g. for $j = 1, 2, \dots, N$ the errors ϵ_{2pj} should have conditional mean zero, i.e. $\mathbb{E}[\epsilon_{2pj} | \mathbf{x}] = 0$ ('strict exogeneity'), amongst others, the *Gauss-Markov theorem* – see e.g. [52] – states that $\hat{\boldsymbol{\beta}}_p := (\hat{\beta}_{0p}, \hat{\beta}_{1p}, \dots, \hat{\beta}_{n-1,p})$

is the Best Linear Unbiased Estimator ('BLUE') of $\mathbf{b}_p := (b_{0p}, b_{1p}, \dots, b_{n-1,p})$ amongst all $\hat{\boldsymbol{\beta}}_p$ candidates. $\hat{\boldsymbol{\beta}}_p$ is 'best' in a least squares sense, i.e. it is the *unique* value for $\hat{\boldsymbol{\beta}}_p$ for which the sum of squared residuals $\sum_{j=1}^N (z_j^p - \mathbf{x}_j \hat{\boldsymbol{\beta}}_p^T)^2$ is minimized, with $z_j = F_Z^{-1}(F_{V(t)}(v_j))$. If we additionally assume that the error terms ϵ_{2pj} , $j = 1, 2, \dots, N$ in (4.A.6) are independent and identically distributed with mean zero and finite variance σ^2 , one can prove that $\hat{\boldsymbol{\beta}}_p$ is a *consistent* estimator of \mathbf{b}_p (convergence in probability)

$$\lim_{N \rightarrow \infty} \mathbb{P}(|\hat{\boldsymbol{\beta}}_p - \mathbf{b}_p| \geq \varepsilon) = 0 \quad \forall \varepsilon > 0,$$

and, moreover, applying the central limit theorem, that $\hat{\boldsymbol{\beta}}_p$ is asymptotically normal (convergence in distribution):

$$\hat{\boldsymbol{\beta}}_p \xrightarrow{d} \mathcal{N}\left(\mathbf{b}_p, \frac{\sigma^2}{N} \mathbf{Q}^{-1}\right) \text{ if } N \rightarrow \infty, \text{ with } \mathbf{Q} \text{ defined by } \lim_{N \rightarrow \infty} \mathbb{P}\left(\left|\frac{\mathbf{x}'\mathbf{x}}{N} - \mathbf{Q}\right| \geq \varepsilon\right) = 0 \quad \forall \varepsilon > 0.$$

\mathbf{Q} is a positive definite matrix.

Remark 4.A.1 (Number of basis functions). *Although $\hat{\boldsymbol{\beta}}_p$ is the 'best' estimator of \mathbf{b}_p in the sense that it has the lowest variance compared to all other unbiased estimators $\hat{\boldsymbol{\beta}}_p$, the absolute error can still be significant, e.g. due to an inappropriate choice of the basis functions or the number of basis functions. Tests for the significance of the polynomial terms, in particular the highest order term, can be conducted, where the null hypothesis states that $\hat{\beta}_{kp} = 0$ for some $k = 0, 1, \dots, n-1$. Related to this, one can employ either a forward selection procedure or a backward elimination procedure, in which the model is successively fit in increasing or decreasing order under statistical testing, respectively. Other criteria one can consider to test whether a multiple linear regression model is well-constructed are e.g. the well-known R^2 -value (coefficient of determination) and the condition number of the matrix involved in the regression, which is a measure for the ill-posedness and multicollinearity of the problem.*

4.B. PROOFS OF LEMMA 4.2.1 AND THE RESULT IN SECTION 4.3.1

In this section we provide proofs for Lemma 4.2.1 and the result in Section 4.3.1, which we state here as Lemmas 4.B.1 and 4.B.2, respectively.

Lemma 4.B.1. *Given the two-dimensional model (4.2.27). Let X and Z denote standard normal random variables and assume for an arbitrary t that the elements $Y_i(t) = y_i$, $i = 1, 2$, $X = x$, $Z = z$ are related by $y_1 = g(x)$, $y_2 = h(z)$, with $g(\cdot)$ and $h(\cdot)$ defined in (4.2.13) and (4.2.15), respectively. This implies that X and Z are jointly bivariate normally distributed.*

Proof. We start by writing (4.2.27) in terms of independent Brownian motions $\widetilde{W}_1(\cdot)$ and $\widetilde{W}_2(\cdot)$:

$$dY_1(t) = \sigma_1 Y_1(t) d\widetilde{W}_1(t), \quad dY_2(t) = \sigma_2 Y_2(t) \left(\rho d\widetilde{W}_1(t) + \sqrt{1 - \rho^2} d\widetilde{W}_2(t) \right), \quad (4.B.1)$$

with $Y_1(0) = y_{10}$ and $Y_2(0) = y_{20}$. The solution to (4.B.1) reads

$$Y_1(t) = y_{10} \exp\left(-\frac{1}{2}\sigma_1^2 t + \sigma_1 \widetilde{W}_1(t)\right), \quad Y_2(t) = y_{20} \exp\left(-\frac{1}{2}\sigma_2^2 t + \sigma_2 \left(\rho \widetilde{W}_1(t) + \sqrt{1-\rho^2} \widetilde{W}_2(t)\right)\right), \quad (4.B.2)$$

respectively. One can easily show $\log(Y_i(t)) \stackrel{d}{=} \mathcal{N}(\bar{\mu}_i, \bar{\sigma}_i)$, with $\bar{\mu}_i := \log(y_{i0}) - \frac{1}{2}\sigma_i^2 t$ and $\bar{\sigma}_i := \sigma_i \sqrt{t}$, $i = 1, 2$. Further, the element-wise equality $y_1 = g(x)$, with $g(\cdot)$ specified in (4.2.13), implies, as $X \stackrel{d}{=} \mathcal{N}(0, 1)$,

$$x = g^{-1}(y_1) = F_X^{-1}(F_{Y_1(t)}(y_1)) = F_{\mathcal{N}(0,1)}^{-1} F_{\mathcal{N}(0,1)}\left(\frac{\log(y_1) - \bar{\mu}_1}{\bar{\sigma}_1}\right) = \frac{\log(y_1) - \bar{\mu}_1}{\bar{\sigma}_1},$$

which yields for y_1 and similarly for y_2 , substituting the values of $\bar{\mu}_i$ and $\bar{\sigma}_i$, $i = 1, 2$:

$$y_1 = y_{10} \exp\left(-\frac{1}{2}\sigma_1^2 t + \sigma_1 \sqrt{t}x\right), \quad y_2 = y_{20} \exp\left(-\frac{1}{2}\sigma_2^2 t + \sigma_2 \sqrt{t}z\right),$$

thus

$$Y_1(t) = y_{10} \exp\left(-\frac{1}{2}\sigma_1^2 t + \sigma_1 \sqrt{t}X\right), \quad Y_2(t) = y_{20} \exp\left(-\frac{1}{2}\sigma_2^2 t + \sigma_2 \sqrt{t}Z\right). \quad (4.B.3)$$

Equations (4.B.2) and (4.B.3) imply $X = \frac{1}{\sqrt{t}} \widetilde{W}_1(t)$ and $Z = \frac{1}{\sqrt{t}} \left(\rho \widetilde{W}_1(t) + \sqrt{1-\rho^2} \widetilde{W}_2(t)\right) = \rho X + \sqrt{1-\rho^2} \widetilde{Z}$, with $\widetilde{Z} := \frac{1}{\sqrt{t}} \widetilde{W}_2(t)$. As we are able to express Z in terms of X via the Cholesky decomposition, Z and X are jointly bivariate normally distributed. \square

Lemma 4.B.2. *For two jointly normally distributed random variables $X_1 \stackrel{d}{=} \mathcal{N}(\mu_1, \sigma_1)$ and $X_2 \stackrel{d}{=} \mathcal{N}(\mu_2, \sigma_2)$, correlated with correlation parameter ρ , the following result holds:*

$$\mathbb{E}[X_1 \mathbb{1}_{X_2 > k}] = \left(\mu_1 + \rho \sigma_1 \frac{f_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)}{1 - F_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)} \right) (1 - F_{X_2}(k)),$$

where $f_{\mathcal{N}(0,1)}(\cdot)$ and $F_{\mathcal{N}(0,1)}(\cdot)$ are the standard normal PDF and CDF, respectively, and $F_{X_2}(\cdot)$ is the CDF corresponding to the random variable X_2 .

Proof. We start by writing

$$\mathbb{E}[X_1 \mathbb{1}_{X_2 > k}] = \mathbb{E}[\mathbb{E}(X_1 \mathbb{1}_{X_2 > k} | X_2 > k)] = \mathbb{E}[\mathbb{1}_{X_2 > k} \mathbb{E}(X_1 | X_2 > k)]. \quad (4.B.4)$$

For the inner expectation we set $X_2 = \sigma_2 Z_2 + \mu_2$ and $X_1 = \sigma_1 \left(\rho Z_2 + \sqrt{1-\rho^2} Z_1\right) + \mu_1$, where Z_2 and Z_1 are independent standard normal random variables. The expression of X_1 in terms of Z_2 is established by assuming that X_1 and X_2 are bivariate normally distributed (the joint distribution of two normal random variables does not need to be bivariate normal. Only the reverse holds in general.). Straightforward calculus yields

$\mathbb{E}(X_1 | X_2 > k) = \rho\sigma_1 \mathbb{E}(Z_2 | Z_2 > (k - \mu_2)/\sigma_2) + \mu_1$. The conditional expectation is given by

$$\mathbb{E}(Z_2 | Z_2 > (k - \mu_2)/\sigma_2) = \frac{f_{\mathcal{N}(0,1)}\left(\frac{k - \mu_2}{\sigma_2}\right)}{1 - F_{\mathcal{N}(0,1)}\left(\frac{k - \mu_2}{\sigma_2}\right)}. \text{ So}$$

$$\mathbb{E}(X_1 | X_2 > k) = \rho\sigma_1 \frac{f_{\mathcal{N}(0,1)}\left(\frac{k - \mu_2}{\sigma_2}\right)}{1 - F_{\mathcal{N}(0,1)}\left(\frac{k - \mu_2}{\sigma_2}\right)} + \mu_1.$$

Substituting this result in (4.B.4) yields the result in the lemma. \square

5

COLLOCATING LOCAL VOLATILITY: A COMPETITIVE ALTERNATIVE TO STOCHASTIC LOCAL VOLATILITY MODELS

We discuss a competitive alternative to the stochastic local volatility models, namely the Collocating Local Volatility (CLV) model, introduced in [54]. The CLV model consists of two elements, a 'kernel process' that can be efficiently evaluated and a local volatility function. The latter, based on stochastic collocation – e.g. [9, 45, 127, 131] – connects the kernel process to the market and allows the CLV model to be perfectly calibrated to European-type options. In this chapter we consider three different kernel process choices: the Ornstein-Uhlenbeck (OU) and Cox-Ingersoll-Ross (CIR) processes and the Heston model. The kernel process controls the forward smile and allows for an accurate and efficient calibration to exotic options, while the perfect calibration to liquid market quotes is preserved. We confirm this by numerical experiments, in which we calibrate the OU-CLV, CIR-CLV and Heston-CLV models to FX barrier options.

Keywords: Collocating Local Volatility, stochastic local volatility, Monte Carlo, stochastic collocation, calibration, forward volatility, barrier options.

5.1. INTRODUCTION

In this chapter we consider an alternative to SLV models, namely the *Collocating Local Volatility* (CLV) model, introduced in [54]. The CLV model is composed of a kernel process and a local volatility function, which is constructed based on stochastic collocation

This chapter is based on the article 'Collocating Local Volatility: A Competitive Alternative to Stochastic Local Volatility Models', submitted for publication, 2018 [121].

[9, 12, 90, 107, 131] and, as a consequence, admits a perfect calibration to arbitrage-free European-type option prices. The kernel process can be *any* stochastic process – in the case however that the moments of the kernel variable exist and are numerically stable, optimal collocation points can be determined by which the local volatility function is defined [58].

The CLV model allows for flexibility regarding the forward smile. The forward smile is governed by the kernel process and reflects the transition densities between future states of the underlying, which determine the price of a path-dependent product [10]. By an appropriate choice of the kernel process and its parameter values, the CLV model is well-capable of pricing exotic options, while maintaining a fit to liquid European-style options. In this chapter we consider three different kernel processes: the Ornstein-Uhlenbeck (OU) and Cox-Ingersoll-Ross (CIR) processes and the Heston model.

Another advantageous property of the CLV model is the fact that the local volatility function only needs to be evaluated at the time-points of interest. In addition, the kernel process typically allows for large time-steps in a simulation. This particularly holds if an analytical solution is available (as for e.g. the OU and CIR kernel processes), however also for other processes efficient simulation schemes exist, e.g. the Heston model can be efficiently simulated by Andersen's QE scheme [5].

The CLV model, by its flexibility in controlling the forward smile and its rapid Monte Carlo evaluation, allows for an efficient Monte Carlo calibration to exotic options, while the fit to European-type options is preserved.

The present chapter is organized as follows. In Section 5.2 we present the CLV model and elaborate on its advantageous properties. We establish the corresponding pricing PDE along the lines of the derivation of the Black-Scholes pricing PDE – we employ the notion of martingales. Also, for application purposes we describe the evaluation steps of the CLV model in a Monte Carlo simulation framework. Subsequently, in Section 5.3 we consider three choices for the kernel process, namely OU and CIR dynamics and the Heston model; we describe the characteristics and consider the effect of the kernel parameters on the shape of the forward smile. Based on this analysis, in Section 5.4 we calibrate the OU, CIR and Heston kernel processes to FX barrier options. Last, Section 5.5 concludes.

5.2. THE COLLOCATING LOCAL VOLATILITY MODEL

In this section we discuss the main characteristics of the Collocating Local Volatility (CLV) model [54]. Also, we write the model in a standard form and derive its pricing PDE.

The CLV model is represented as follows, under the risk-neutral \mathbb{Q} -measure:

$$S(t) = g_N(t, \mathbf{X}(t)), \tag{5.2.1}$$

$$d\mathbf{X}(t) = \bar{\mu}(t, \mathbf{X}(t))dt + \bar{\sigma}(t, \mathbf{X}(t))d\mathbf{W}^{\mathbb{Q}}(t), \mathbf{X}(t_0) = \mathbf{X}_0, \tag{5.2.2}$$

where $g_N(\cdot, \mathbf{X}(\cdot)) : [t_0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$, $\bar{\mu}(\cdot, \mathbf{X}(\cdot)) : [t_0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\bar{\sigma}(\cdot, \mathbf{X}(\cdot)) : [t_0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$ and $\mathbf{W}^{\mathbb{Q}}(\cdot)$ is an n -dimensional Brownian motion. The model consists of two elements that are evaluated separately. The first building block is the *kernel process* $\mathbf{X}(\cdot)$ in (5.2.2). The second building block is the CLV element (5.2.1), which connects the kernel process

to liquid market quotes via the *local volatility function* $g_N(\cdot, \cdot)$, which is based on the *stochastic collocation method* [9, 12, 90, 107, 131].

An advantageous property of the CLV model is that, by construction, function $g_N(\cdot, \cdot)$ guarantees an almost perfect calibration to arbitrage-free European-type option prices, *independently of the kernel parameter values*. Basically this function, given liquid market quotes for expiries T_1, \dots, T_M , yields a highly accurate interpolation through the pairs $(x_{i,j}, s_{i,j})$, $i = 1, \dots, M$, $j = 1, \dots, N$, with $x_{i,1}, \dots, x_{i,N}$ and $s_{i,1}, \dots, s_{i,N}$ representing the *collocation points* and *collocation values* corresponding to T_i , respectively. The collocation points may be established based on the moments of the kernel variable (see Remark 5.2.1). The collocation values are computed by

$$s_{i,j} = F_{S(T_i)}^{-1}(F_{X(T_i)}(x_{i,j})),$$

where the cumulative distribution function of $S(T_i)$ under the risk-neutral measure is specified by equation (5.2.3) in Lemma 5.2.1.

The function $g_N(\cdot, \cdot)$ is an interpolation through the $s_{i,j}$ -values, given particular t and $X(\cdot)$ values. Choosing $g_N(\cdot, \cdot)$ in the Lagrange form is well-accepted in the field of Uncertainty Quantification, see e.g. [106]. However, this choice does not guarantee monotonicity in the strike direction, which is a desirable property. We therefore choose a piecewise cubic Hermite interpolation, which is guaranteed to be monotonic and continuously differentiable, see e.g. [44].

Lemma 5.2.1 (Market-implied CDF of $S(\cdot)$ under the risk-neutral measure). *The market-implied CDF of $S(T)$ under the risk-neutral \mathbb{Q} -measure is given by*

$$F_{S(T)}(x) = 1 + e^{rT} \left. \frac{\partial C^{\text{mkt}}(T, K)}{\partial K} \right|_{K=x}, \quad (5.2.3)$$

with $C^{\text{mkt}}(T, K)$ denoting today's arbitrage-free price of a European call option with strike K and expiry T and r denotes a constant interest rate.

Proof. In general, the discounted value of a standard European option with an expiry T and strike K at time t under the risk-neutral \mathbb{Q} -measure is

$$C(T, K) = M(t) \mathbb{E}^{\mathbb{Q}} \left[\left. \frac{(S(T) - K)^+}{M(T)} \right| \mathcal{F}(t) \right],$$

where $S(\cdot)$ is the underlying and $M(\cdot)$ stands for the money account, determined by $dM(t) = rM(t)dt$ with a constant interest rate r . The discounted value of the option is given by (suppressing the filtration notation):

$$C(T, K) = e^{-rT} \mathbb{E}^{\mathbb{Q}} [(S(T) - K)^+] = e^{-rT} \int_K^{\infty} (x - K) f_{S(T)}(x) dx, \quad (5.2.4)$$

where $f_{S(T)}(\cdot)$ is the market-implied PDF of $S(T)$ under the \mathbb{Q} -measure. Differentiating and applying Leibniz' integration rule gives

$$e^{rT} \frac{\partial C(T, K)}{\partial K} = -1 + F_{S(T)}(K),$$

which we write for an arbitrary argument x as:

$$F_{S(T)}(x) = 1 + e^{rT} \left. \frac{\partial C^{\text{mkt}}(T, K)}{\partial K} \right|_{K=x},$$

where we added the ‘mkt’ superscript to emphasize that we obtain the CDF from the market quotes. This concludes the proof of Lemma 5.2.1. \square

Remark 5.2.1 (Optimal collocation points). *Optimal collocation points $x_{i,1}, \dots, x_{i,N}$ can be calculated based on the first $2N$ moments of the underlying kernel variable at T_i , $X(T_i)$ [58]. In this case the collocation points are zeros of the orthogonal polynomial corresponding to (the probability density function of) the kernel variable and can be computed by an eigenvalue method. By choosing optimal collocation points, the stochastic collocation method can be connected to the computation of integrals by Gauss quadrature.*

Remark 5.2.2 (Relation between $X(t)$ and $S(t)$). *Ideally, for a given t , the relation between the distribution of the kernel variable and the market-implied distribution is approximately linear or, stated differently, the densities of $X(t)$ and $S(t)$ resemble each other. This yields a small approximation error [58] and optimal results.*

5

Besides for its almost perfect calibration, a second beneficial property of the CLV model is the fact that, in e.g. a Monte Carlo simulation framework, we do not need to evaluate $g_N(\cdot, \cdot)$ at each time-step, which is the case for the standard Local Volatility model [34, 35]. For example, for pricing a European-type option we simulate the kernel process (5.2.2) up to the option’s maturity T and subsequently compute $g_N(T, X(T))$ (5.2.1). In the case that the time-points of interest are specified on a coarse grid, we prefer a simulation method for the kernel variable which is low-biased for large time-steps¹. Moreover, in the case that the kernel process has an analytical solution (e.g. the Ornstein-Uhlenbeck and Cox-Ingersoll-Ross processes), it allows for an *exact* simulation method with large time-steps.

A third advantageous characteristic of the CLV model is its flexibility in controlling the forward smile, while maintaining an almost perfect fit to European-type options by construction. For the CLV model, as discussed in [54], the autocorrelation of the kernel process affects the forward smile. As such, the choice of kernel process and the kernel parameter values determine the forward smile generated by the CLV model, without affecting the almost perfect calibration to European-type options. The payoff of a path-dependent product is determined by the evolution of the underlying through time, i.e. its price depends on the transition densities from one future state to another [10].

As the CLV model is flexible in controlling the forward smile and can be efficiently evaluated, it allows for an efficient Monte Carlo calibration to exotic options, while the fit to European-type options is maintained. In Section 5.3 we consider the forward smile

¹For example, in the case of a Heston kernel process we would simulate $X(\cdot)$ by employing the *QE scheme* of Andersen [5], which allows for large time-steps. An alternative would be to make use of the so-called *exact simulation scheme* proposed by Broadie and Kaya [18], which is based on acceptance-rejection sampling of the variance process coupled with certain Fourier inversion computations. As presented in [58], by employing the Stochastic Collocation Monte Carlo sampler the exact simulation can be performed efficiently and accurately.

for three different choices of the kernel process, and in Section 5.4 we calibrate the kernel process to FX barrier options.

Typically, in the field of financial engineering, the dynamics of the underlying $S(\cdot)$ are presented, as opposed to the non-standard model representation in equations (5.2.1)-(5.2.2). We can write this model in a more standard way by applying Itô's lemma. Introducing the short-hand notation $g_N := g_N(t, X(t))$, assuming that $X(\cdot)$ is a one-dimensional kernel process and that the relevant partial derivatives $\partial g_N / \partial X$, $\partial^2 g_N / \partial X^2$ and $\partial g_N / \partial t$ exist, $S(\cdot)$ follows an Itô process which is governed by the same Wiener process as $X(\cdot)$, under the risk-neutral \mathbb{Q} -measure:

$$dS(t) = \left(\frac{\partial g_N}{\partial t} + \mu^{\mathbb{Q}}(t, X(t)) \frac{\partial g_N}{\partial X} + \frac{1}{2} \frac{\partial^2 g_N}{\partial X^2} \sigma^2(t, X(t)) \right) dt + \frac{\partial g_N}{\partial X} \sigma(t, X(t)) dW^{\mathbb{Q}}(t).$$

Analogous to the derivation of the Black-Scholes pricing PDE, we derive the CLV pricing PDE. To express the PDE merely in terms of derivatives to X , we define

$$\tilde{C}(t, X(t)) := C(t, g_N(t, X(t))) = C(t, S(t)),$$

with $C(t, S(t))$ representing the value of a European option on the underlying $S(\cdot)$ ('plain vanilla contingent claim').

Lemma 5.2.2 (CLV pricing PDE). *Given the CLV model under the risk-neutral \mathbb{Q} -measure with a general one-dimensional kernel process $X(\cdot)$:*

$$\begin{aligned} S(t) &= g_N(t, X(t)), \\ dX(t) &= \mu^{\mathbb{Q}}(t, X(t)) dt + \sigma(t, X(t)) dW^{\mathbb{Q}}(t). \end{aligned}$$

Suppose that the partial derivatives of $g_N := g_N(t, X(t))$, $\partial g_N / \partial X$, $\partial^2 g_N / \partial X^2$ and $\partial g_N / \partial t$ exist. Also, assume that the money account $M(\cdot)$ is determined by $dM(t) = rM(t)dt$, with r denoting a constant interest rate. Then $\tilde{C} := \tilde{C}(t, X(t))$ is governed by

$$\frac{\partial \tilde{C}}{\partial t} + \mu^{\mathbb{Q}}(t, X) \frac{\partial \tilde{C}}{\partial X} + \frac{1}{2} \sigma^2(t, X) \frac{\partial^2 \tilde{C}}{\partial X^2} - r\tilde{C} = 0, \quad (5.2.5)$$

with the final condition

$$\tilde{C}(T, X(T)) = \Phi(g_N(T, X(T))),$$

where $\Phi(\cdot)$ is a payoff function depending on the final state of $g_N(\cdot, \cdot)$.

Proof. By Itô's lemma, introducing the short-hand notation $\tilde{C} := \tilde{C}(t, X(t))$, we obtain:

$$d\tilde{C}(t, X(t)) = \left(\frac{\partial \tilde{C}}{\partial t} + \mu^{\mathbb{Q}}(t, X(t)) \frac{\partial \tilde{C}}{\partial X} + \frac{1}{2} \sigma^2(t, X(t)) \frac{\partial^2 \tilde{C}}{\partial X^2} \right) dt + \sigma(t, X(t)) \frac{\partial \tilde{C}}{\partial X} dW^{\mathbb{Q}}(t).$$

Introducing the short-hand notation

$$\Pi(t, X(t)) := \frac{\tilde{C}(t, X(t))}{M(t)},$$

and substituting the dynamics of $M := M(t)$ and $\tilde{C} := \tilde{C}(t, X(t))$, the dynamics of the discounted option value are given by

$$d\Pi(t, X(t)) = \frac{1}{M} \left(\frac{\partial \tilde{C}}{\partial t} + \mu^{\mathbb{Q}}(t, X(t)) \frac{\partial \tilde{C}}{\partial X} + \frac{1}{2} \sigma^2(t, X(t)) \frac{\partial^2 \tilde{C}}{\partial X^2} - r \tilde{C} \right) dt + \frac{\sigma(t, X(t))}{M} \frac{\partial \tilde{C}}{\partial X} dW^{\mathbb{Q}}(t).$$

In an arbitrage-free world we require that the option value discounted by the money account is a *martingale* in the risk-neutral \mathbb{Q} -measure. As such, the drift term is zero, i.e.

$$\frac{\partial \tilde{C}}{\partial t} + \mu^{\mathbb{Q}}(t, X) \frac{\partial \tilde{C}}{\partial X} + \frac{1}{2} \sigma^2(t, X) \frac{\partial^2 \tilde{C}}{\partial X^2} - r \tilde{C} = 0.$$

The final condition on $\tilde{C}(\cdot)$ is given in terms of the payoff function $\Phi(\cdot)$ that depends on the final state of $g_N(\cdot, \cdot)$:

$$\tilde{C}(T, X(T)) = \Phi(g_N(T, X(T))).$$

This concludes the proof of Lemma 5.2.2. □

5

The PDE (5.2.5) is solved backwards in time, given the condition at the time to maturity. It follows from the Feynman-Kac theorem that a solution of (5.2.5) at an arbitrary time $t_0 < T$ is given by

$$\frac{\tilde{C}(t_0, X(t_0))}{M(t_0)} = \mathbb{E}^{\mathbb{Q}} \left[\frac{\Phi(g_N(T, X(T)))}{M(T)} \middle| \mathcal{F}(t_0) \right].$$

5.2.1. MARTINGALE CONSIDERATIONS

The CLV method does not necessarily guarantee risk-neutral drift conditions. In particular, the following condition does not hold for any set of kernel parameter values:

$$\mathbb{E}^{\mathbb{Q}} \left[\frac{g_N(T_2, X(T_2))}{g_N(T_1, X(T_1))} \middle| \mathcal{F}(t_0) \right] = \frac{F(t_0, T_2)}{F(t_0, T_1)}, \tag{5.2.6}$$

with $t_0 < T_1 < T_2$ and $F(t_0, T) = S_0 e^{r(T-t_0)}$ denoting the forward. This implies that $S(\cdot) = g_N(\cdot, X(\cdot))$ is *not* a martingale. However, by a proper choice of the kernel process and its parameter values, we are able of controlling the autocorrelation within the CLV model. As such, the condition in (5.2.6) can be approximated accurately. In Section 5.2.1 we will provide numerical evidence for this. In addition, an idea similar to the CLV model has already been proposed by e.g. Jäckel in [72], where an ‘inverse quantile function’ is employed to map the quantile of an arbitrary distribution to the level of the underlying. He establishes a link with the class of *Markov-functional models* [43, 68], and observes that the method does not necessarily guarantee risk-neutral drift conditions. Last, there is a unique scalar diffusion that is perfectly calibrated to the market for any time *and* is a martingale, namely Dupire’s Local Volatility (LV) model [34, 35]. Although from a theoretical point of view (5.2.6) may not necessarily be satisfied by the CLV model, its construction yields a flexibility that allows for a calibration to exotic options – in a traditional LV setting this is not possible.

KERNEL PARAMETER EFFECT ON MARTINGALITY

In this section we numerically show that the condition in (5.2.6) can be accurately approximated by determining a suitable kernel parameter value. We consider an OU-CLV model with the OU model parameters $X_0 = 1$, $\gamma = 0.3$ and $\theta = 0.5$. We vary the value of the mean reversion parameter κ and calculate the ratio

$$v := \mathbb{E} \left[\frac{g_N(T_2, X(T_2))}{g_N(T_1, X(T_1))} \right] \quad (5.2.7)$$

and compare it against the theoretical ratio of the forwards $v^* := F(t_0, T_2)/F(t_0, T_1)$ (we set $t_0 = 0$). We assume a Heston market parameterization with parameters $\kappa = 0.5$, $\gamma = 0.3$, $\rho_{x,v} = -0.1$ and $V_0 = \bar{V} = 0.04$ (the Feller condition is violated). Further, we set $r = 0.05$ and $S_0 = 1$.

Results² are given by Table 5.3.1 – we report (5.2.7) and the absolute difference in percentage

$$\epsilon := 100 \cdot \left| \frac{v - v^*}{v^*} \right|.$$

For $\kappa = -0.025$ ϵ is less than 5 basis points and the condition in (5.2.6) is accurately approximated.

(T_1, T_2)	(1,2)	(2,3)	(3,4)	(4,5)
v^*	1.0513	1.0513	1.0513	1.0513
v for:				
$\kappa = -0.1$	1.0497 (0.15)	1.0486 (0.26)	1.0473 (0.38)	1.0463 (0.48)
$\kappa = -0.025$	1.0513 (0.00)	1.0516 (0.03)	1.0518 (0.05)	1.0515 (0.02)
$\kappa = -0.01$	1.0516 (0.03)	1.0522 (0.09)	1.0533 (0.19)	1.0526 (0.12)
$\kappa = 0.01$	1.0521 (0.08)	1.0531 (0.18)	1.0539 (0.25)	1.0561 (0.46)
$\kappa = 0.1$	1.0541 (0.27)	1.0574 (0.59)	1.0608 (0.91)	1.0658 (1.39)
$\kappa = 0.5$	1.0634 (1.16)	1.0790 (2.64)	1.0967 (4.32)	1.1161 (6.17)

Table 5.2.1: Effect of the OU kernel parameter κ on the ratio of underlying prices v , as defined in (5.2.7). Between brackets we report the absolute difference in percentage $\epsilon := 100 \cdot \left| \frac{v^* - v}{v^*} \right|$. For $\kappa = -0.025$ results are most accurate (in **bold**).

As was already pointed out in [54], in the OU-CLV model only κ has an effect on the autocorrelation. Even more accurate results can be expected for richer kernel processes, like the CIR process or the Heston model, where more parameters affect the autocorrelation and can enhance the fit to the theoretical ratio of forward rates.

5.3. THE OU-CLV, CIR-CLV AND HESTON-CLV MODELS

As already mentioned in [54], the kernel process (5.2.2) can be specified in principle freely. In the case its moments are analytically available and numerically stable, the

²We use $N = 6$ collocation points and apply a Monte Carlo simulation with 1 time-step per year and 5 seeds, with 10^5 paths per seed.

optimal collocation points can be established (see Remark 5.2.1 and Appendix 5.A) – however, this is not a strict requirement.

The choice of an appropriate kernel process is subtle. On one hand, the process should be evaluated efficiently, e.g. the Monte Carlo simulation may consist of large time-steps. On the other hand, the process should be sufficiently rich to represent realistic dynamics, implying a *realistic forward smile behaviour*.

In this section we discuss three different choices for the kernel process, namely an Ornstein-Uhlenbeck (OU), a Cox-Ingersoll-Ross (CIR) and a Heston process. We discuss the main characteristics of the OU-CLV, CIR-CLV and Heston-CLV models and give special attention to the forward smiles. More specifically, for each of the kernel processes we consider the effect of the various kernel parameters on the shape of the forward smile. Based on this analysis, we calibrate the OU-CLV, CIR-CLV and Heston-CLV models to FX barrier options in Section 5.4.

5.3.1. THE OU-CLV MODEL

The OU-CLV model is given by the following equations, under the risk-neutral \mathbb{Q} -measure:

$$S(t) = g_N(t, X(t)), \tag{5.3.1}$$

$$dX(t) = \kappa(\theta - X(t))dt + \gamma dW^{\mathbb{Q}}(t), \quad X(t_0) = X_0. \tag{5.3.2}$$

Remark 5.3.1 (Specification of $X(\cdot)$). *Given the filtration at $t_0 = 0$, the solution to (5.3.2) reads*

$$X(t) = X_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}) + \frac{\gamma}{\sqrt{2\kappa}} e^{-\kappa t} W^{\mathbb{Q}}(e^{2\kappa t} - 1), \tag{5.3.3}$$

which is normally distributed with the mean and variance

$$\mu_X(t) = X_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}), \quad \sigma_X^2(t) = \frac{\gamma^2}{2\kappa} (1 - e^{-2\kappa t}), \tag{5.3.4}$$

respectively.

As the moments of $X(\cdot)$ are analytically available and numerically stable, we can calculate the optimal collocation points [54, 58] resulting in exponential convergence with respect to the number of collocation points. Also, as the distribution of $X(\cdot)$ is known, for the pricing of a standard European-type option by the OU-CLV model *one* time-step is sufficient³.

For the pricing of exotics though, multiple time-steps are necessary, see e.g. the numerical experiment in Appendix 5.B where we price a discretely monitored barrier option. Whether this price is realistic, depends on the *forward smile* the OU-CLV model implies. We therefore consider the smile corresponding to a *forward-start option*, which provides the holder at a future time $T_1 > t_0$ with a European option with a maturity $T_2 > T_1$ and a strike $K \cdot S(T_1)$. At T_2 the pay-off of this option of the ‘call type’ is [88]:

$$C_{\text{Forw.St.}} = \max(S(T_2) - K \cdot S(T_1), 0). \tag{5.3.5}$$

³In step 3 of Algorithm 5, $X(T)$ is generated by sampling \mathcal{M} realizations x_m , $m = 1, \dots, \mathcal{M}$ from a normal random variable with mean and variance as in equation (5.3.4).

Numerical experiments make clear that only the mean reversion parameter κ has an effect on the forward smile corresponding to the OU-CLV model – see Figure⁴ 5.3.1. On the right-hand side we display the forward volatility smiles omitting the level effect to make the curvature effect visible (more precisely, we shift the smile downwards such that its minimum is at 0). The primary effect of the mean reversion parameter is on the level: an increase in κ yields an increase in level. A secondary effect is on the curvature, see the right-hand plot of Figure 5.3.1. As mentioned in [54], changing κ affects the filtration of the Brownian motion of the solution of $X(\cdot)$ in (5.3.3). As such, an OU process with mean reversion parameter κ_1 cannot be expressed as a linear combination of an OU process with a different mean reversion parameter value κ_2 . Because of this, a change of κ affects the autocorrelation of the paths of $g_N(\cdot, \cdot)$ and the forward smile. In Table 5.3.1 we summarize the kernel parameter effects on the shape of the forward smile.

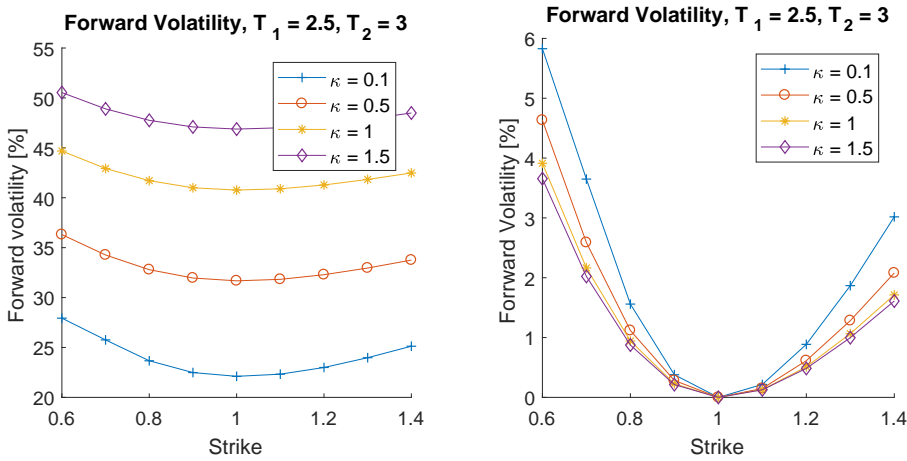


Figure 5.3.1: Effect of κ on the forward volatility smile for the OU-CLV model, with $T_1 = 2.5$ and $T_2 = 3$. In the right-hand plot all smiles are shifted to zero level to make the curvature effect more visible. Other OU parameters are: $X_0 = 1, \gamma = 0.3, \theta = 0.5$.

5.3.2. THE CIR-CLV MODEL

The Cox Ingersoll Ross-Collocating Local Volatility (CIR-CLV) model is represented by the following two equations, under the risk-neutral \mathbb{Q} -measure:

$$S(t) = g_N(t, X(t)), \tag{5.3.6}$$

$$dX(t) = \kappa(\theta - X(t))dt + \gamma\sqrt{X(t)}dW^{\mathbb{Q}}(t), X(t_0) = X_0. \tag{5.3.7}$$

Result 5.3.1 (Specification of $X(\cdot)$). *Given the filtration at $t_0 = 0$, the solution to (5.3.7) is distributed as a scaled non-central chi-square random variable $\chi^2(d, \lambda(t))$ with d degrees*

⁴We omit the pictures corresponding to the other kernel parameters, as no effect was observed. We assume a Heston market parameterization with parameters $\kappa = 0.5, \gamma = 0.3, \rho_{x,v} = -0.1, V_0 = 0.04$ and $\bar{V} = 0.04$. Also, $r = 0$ and $S_0 = 1$. Further, we use $N = 6$ collocation points and 2 time-steps per year and $5 \cdot 10^5$ paths in the Monte Carlo simulation (5 seeds, each seed constitutes 10^5 paths).

	Level	Curvature	Skewness
κ	+	-	0
γ	0	0	0
X_0	0	0	0
θ	0	0	0

Table 5.3.1: Separate effects of the OU kernel parameters on the level, curvature and skewness of the forward smile implied by the OU-CLV model. A '+'/'-' represents a higher/lower volatility smile level, more/less curvature or more/less skewness in the case of increasing a particular kernel parameter. A '0' stands for no effect.

of freedom and non-centrality parameter $\lambda(t)$, i.e.

$$X(t) \stackrel{d}{=} c(t)\Lambda(t), \text{ with } \Lambda(t) \stackrel{d}{=} \chi^2(d, \lambda(t)), \tag{5.3.8}$$

where

$$c(t) = \frac{1}{4\kappa} \gamma^2 (1 - e^{-\kappa t}), \quad d = \frac{4\kappa\theta}{\gamma^2}, \quad \lambda(t) = \frac{4\kappa X_0 e^{-\kappa t}}{\gamma^2 (1 - e^{-\kappa t})}. \tag{5.3.9}$$

The n th moment of $\Lambda(t)$ is given by

$$\mathbb{E}[\Lambda^n(t)] = 2^{n-1} (n-1)! (d + n\lambda(t)) + \sum_{k=1}^{n-1} \frac{(n-1)! 2^{k-1}}{(n-k)!} (d + k\lambda(t)) \mathbb{E}[\Lambda^{n-k}(t)].$$

In computing the first moment, the summation term disappears and $0! = 1$, resulting in $\mathbb{E}[\Lambda(t)] = d + \lambda(t)$, thus

$$\mathbb{E}[X(t)] = c(t) (d + \lambda(t)).$$

For $X(\cdot)$ given by (5.3.8) an explosion of moments may occur. E.g., for the kernel parameters $X_0 = 1, \theta = 1, \kappa = 1, \gamma = 0.1$ we have at $t = 1$ the values $d = 400, \lambda(1) \approx 233$. Given these values, the 4th moment of $\Lambda(\cdot) \stackrel{d}{=} \chi^2(d, \lambda(t))$ has a value with an order of magnitude of 10^{11} . Due to the large moment values, numerical instabilities in computing the collocation points may occur⁵. As an *alternative approach*, we use the collocation points $z_j, j = 1, \dots, N$ corresponding to the *standard normal random variable* $Z \stackrel{d}{=} \mathcal{N}(0, 1)$ – see Remark 5.3.2. Given these, we compute the collocation points corresponding to expiry $T_i, i = 1, \dots, M$, as follows:

$$x_{i,j} := x_j(T_i) = F_{X(T_i)}^{-1}(F_Z(z_j)) = c(T_i) F_{\Lambda(T_i)}^{-1}(F_Z(z_j)), \quad i = 1, \dots, M, \quad j = 1, \dots, N.$$

Note that this implies the collocation values

$$s_{i,j} = F_{S(T_i)}^{-1}(F_{X(T_i)}(x_{i,j})) = F_{S(T_i)}^{-1}(F_{X(T_i)}(c(T_i) F_{\Lambda(T_i)}^{-1}(F_Z(z_j)))) = F_{S(T_i)}^{-1}(F_Z(z_j)), \tag{5.3.10}$$

which shows that we do *not* need the CDF of $X(\cdot)$ for computing the collocation values.

⁵For example, for $N = 8$ we obtain negative collocation points, although the distribution of $X(\cdot)$ does not allow for negative values.

Remark 5.3.2 (Normal distribution). *The reason why we choose a standard normal distribution in the alternative approach is twofold. First, even for a fundamental distribution as the standard normal, results are highly accurate – this is also the case in e.g. [54, 58]. By choosing a different distribution, results may be further enhanced. Secondly, as mentioned in [58], choosing the normal distribution is also motivated by the Cameron-Martin Theorem [19], which states that polynomial chaos approximations based on the normal distribution converge to any distribution.*

Remark 5.3.3 (Loss of optimality of collocation points). *A drawback of the alternative approach from Remark 5.3.2 is that the collocation points are not the zeros of the orthogonal polynomial $p_N(\cdot)$ that corresponds to the weight function $f_X(\cdot)$, with X denoting the non-central chi-square distributed random variable – see also Appendix 5.A. As a consequence, the method can not be connected to Gauss quadrature, see equations (5.A.1) and (5.A.2) in Appendix 5.A, as the x_i -values do not correspond to the quadrature weights ω_i , which are one-to-one connected to the weight function $f_X(\cdot)$. In fact, there would be a mismatch between the x_i and ω_i values and the error is not longer (completely) determined by the quadrature error.*

Remark 5.3.4 (The case $\gamma = 0$). *In the case $\gamma = 0$, the OU and CIR kernel processes (5.3.2) and (5.3.7) are equivalent and deterministic. This is not a relevant case, as the CLV framework relies on the projection of an ‘expensive’ random variable on a ‘cheaper’ random variable, which is the essence of stochastic collocation.*

EFFECT OF ‘LINEARIZATION’

As we stated in Remark 5.2.2, optimal results are established if for a given t the distributions of $X(t)$ and $S(t)$ approximately resemble each other. For an expiry T_i this implies a close-to-linear relation between the collocation points $x_{i,j}$ and collocation values $s_{i,j}$, through which $g_N(\cdot, \cdot)$ interpolates. Given a set of market data and a set of kernel parameters, it may turn out though that $g_N(\cdot, \cdot)$ is highly non-linear, which may affect the performance of the stochastic collocation method and the eventual fit of the CLV model to European-type market prices negatively.

For illustration purposes, suppose that the market data is parameterized by the Heston model, with parameters given by Case II of Andersen [5]: $V_0 = \bar{V} = 0.04$, $\kappa = 0.3$, $\gamma = 0.9$ and $\rho_{xv} = -0.5$. Also, $r = 0$ and $S_0 = 1$. We choose kernel parameters⁶ $\gamma = 1.5$, $\kappa = 0.5$, $\theta = 0.5$, $X_0 = 1$, and price a European call option with expiry $T = 4$; results are displayed in Figure 5.3.2. To judge the performance of the interpolant $g_N(t, x)$, we display the theoretical function

$$g(t, x) := F_{Y(t)}^{-1}(F_{X(t)}(x)) = F_{S(t)}^{-1}\left(F_{\Lambda(t)}\left(\frac{x - a(t)}{c(t)b(t)}\right)\right),$$

with $a(\cdot)$ and $b(\cdot)$ denoting grid-stretching coefficients [58] and $Y(t) = S(t)$ representing the ‘expensive’ random variable which we project on the kernel variable $X(t) \stackrel{d}{=} c(t)\Lambda(t)$ with $\Lambda(t) \stackrel{d}{=} \chi^2(d, \lambda(t))$ ($c(t)$ and $\lambda(t)$ are specified in (5.3.9)).

⁶We use 8 collocation points and make use of grid-stretching with $p_{\min} = 1 \cdot 10^{-3}$ and $p_{\max} = 0.999$ – for more details on grid-stretching, see e.g. [58]. The Monte Carlo simulation constitutes 10^5 paths (20 seeds, each seed constitutes $5 \cdot 10^3$ paths) and 1 time-step per year.

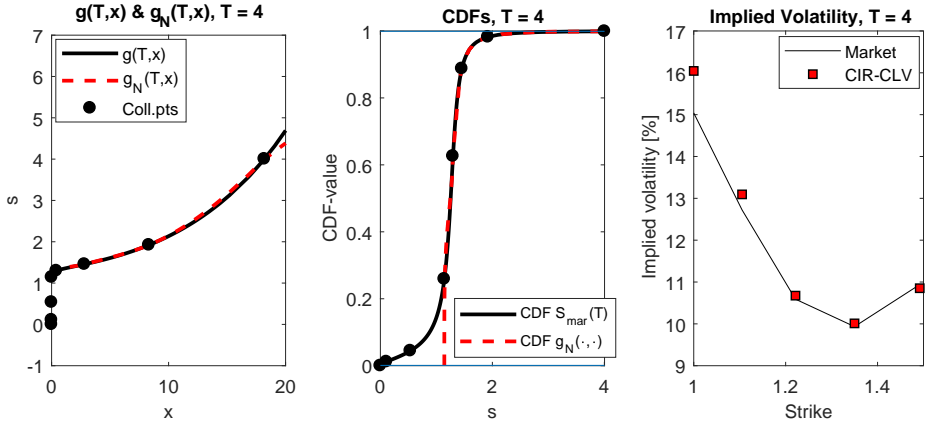


Figure 5.3.2: Results for the case described in Section 5.3.2, with $\gamma = 1.5$. With 'CDF $S_{\text{mar}}(T)$ ' we denote the market-implied CDF specified in equation (5.2.3).

5

As we observe in the middle plot of Figure 5.3.2, the CDF of $g_N(\cdot, \cdot)$ is clearly not in line with the market-implied CDF. As a result, the implied volatility fit for the lower 2 strikes is not accurate. The reason for this lies in the highly non-linear behaviour of $g_N(\cdot, \cdot)$ close to zero. A way to resolve this issue is by adjusting the kernel parameters such that the relation between $S(\cdot)$ and $X(\cdot)$ is closer to linear. Setting $\gamma = 0.75$ results in a better performance of the CLV method close to $x = 0$ and implies a more accurate implied volatility fit, see Figure 5.3.3.

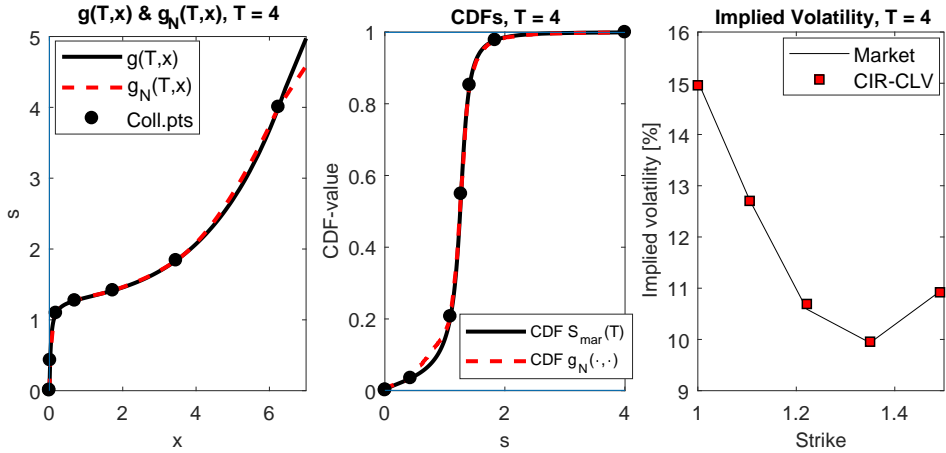


Figure 5.3.3: Results for the case described in Section 5.3.2, with $\gamma = 0.75$. With 'CDF $S_{\text{mar}}(T)$ ' we denote the market-implied CDF specified in equation (5.2.3).

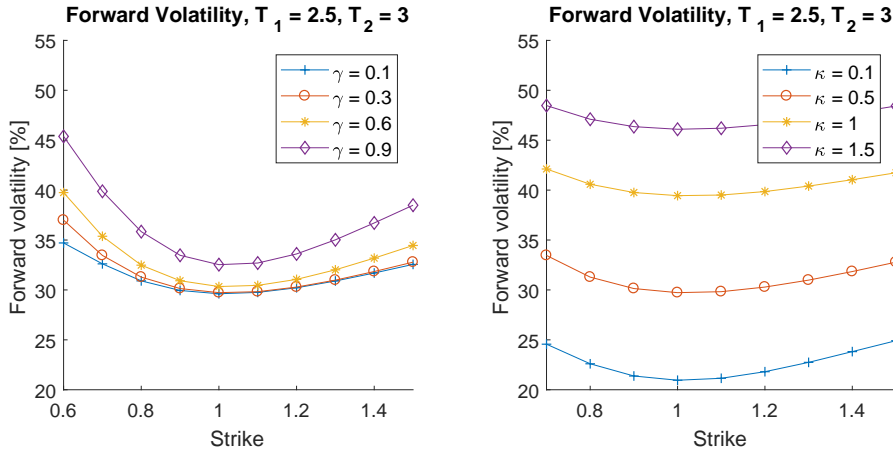


Figure 5.3.4: Effect of γ (left) and κ (right) on the forward volatility smile for the CIR-CLV model, with $T_1 = 2.5$ and $T_2 = 3$. The ‘base case’ CIR parameters (if not varied) are: $\gamma = 0.3$, $\kappa = 0.5$, $\theta = 0.5$ and $X_0 = 1$.

THE FORWARD SMILE

In this section we consider the effect of the CIR kernel parameters on the forward smile corresponding to the CIR-CLV model by pricing a forward-start option, with a pay-off at T_2 given by (5.3.5). With a ‘positive effect’ we mean that increasing a particular kernel parameter results in a higher volatility smile level, more curvature or skewness. Results⁷ are given in Figures 5.3.4 and 5.3.5. We clearly observe that the volatility of variance has a positive effect on both the level and the curvature of the forward smile. The speed of mean reversion has mainly a positive level effect, but also a negative curvature effect, i.e. a larger value of κ implies less curvature. Further, X_0 only has a negative level effect. Last, θ has a positive level effect and a slightly negative curvature effect.

Some effects may be quite difficult to observe due to the level effect. As such, in Appendix 5.C we display Figures 5.3.4 and 5.3.5, but without the level effect. In Table 5.3.2 we summarize the effects of the CIR parameters on the forward smile.

5.3.3. THE HESTON-CLV MODEL

The Heston-CLV model is defined as follows:

$$\begin{aligned}
 S(t) &= g_N(t, X(t)), \\
 dX(t) &= rX(t)dt + \sqrt{V(t)}X(t)dW_x(t), \quad X(0) = X_0, \\
 dV(t) &= \kappa(\bar{V} - V(t))dt + \gamma\sqrt{V(t)}dW_v(t), \quad V(0) = V_0,
 \end{aligned}$$

with $dW_x(t)dW_v(t) = \rho_{x,v}dt$ and where κ , γ , \bar{V} and $\rho_{x,v}$ are the rate of mean reversion, the volatility of variance, the long-term variance and the correlation, respectively.

⁷We assume a Heston market parameterization with parameters $\kappa = 0.5$, $\gamma = 0.3$, $\rho_{x,v} = -0.1$, $V_0 = 0.04$ and $\bar{V} = 0.04$. Also, $r = 0$ and $S_0 = 1$. Further, we use $N = 8$ collocation points, apply grid-stretching with $p_{\min} = 1 \cdot 10^{-3}$ and $p_{\max} = 0.999$ and the Monte Carlo simulation consists of 2 time-steps per year and $5 \cdot 10^5$ paths (5 seeds, each seed constitutes 10^5 paths).

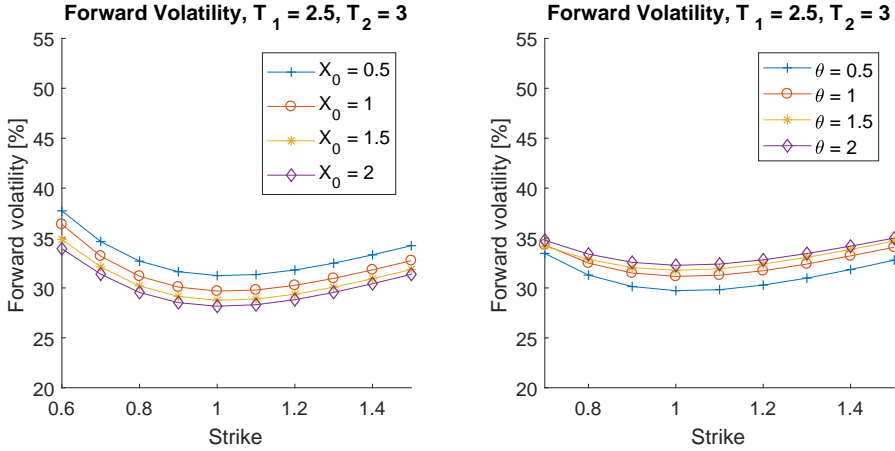


Figure 5.3.5: Effect of X_0 (left) and θ (right) on the forward volatility smile for the CIR-CLV model, with $T_1 = 2.5$ and $T_2 = 3$. The 'base case' CIR parameters (if not varied) are: $\gamma = 0.3$, $\kappa = 0.5$, $\theta = 0.5$ and $X_0 = 1$.

	Level	Curvature	Skewness
γ	+	+	0
κ	+	-	0
X_0	-	0	0
θ	+	-	0

Table 5.3.2: Separate effects of the CIR kernel parameters on the level, curvature and skewness of the forward smile implied by the CIR-CLV model. A '+'/'-' represents a higher/lower volatility smile level, more/less curvature or more/less skewness in the case of increasing a particular kernel parameter. A '0' stands for no effect.

According to [74], the Heston model stands out from the class of stochastic volatility models mainly for two reasons. First, the volatility process is non-negative and mean-reverting, which is typically observed in the markets. Secondly, a fast and easily implemented semi-analytical solution for the pricing of European-type options is available. In particular, efficient numerical Fourier-based techniques exist, which allow for a fast calibration.

ESTABLISHING $x_{i,j}$ AND $s_{i,j}$

From the characteristic function of $\hat{X}(\cdot) := \log(X(\cdot))$ (see e.g. [38]), which is defined as

$$\phi_{\hat{X}(t)}(u) := \mathbb{E} \left[e^{iu\hat{X}(t)} \middle| \mathcal{F}(t_0) \right],$$

we may compute the k th moment of $X(t)$ as follows (we suppress the condition on $\mathcal{F}(t_0)$ for notation purposes):

$$\phi_{\hat{X}(t)}(-ki) = \mathbb{E} \left[e^{k\hat{X}(t)} \right] = \mathbb{E} \left[e^{k\log(X(t))} \right] = \mathbb{E} \left[X^k(t) \right].$$

From the moments of $X(\cdot)$ we obtain the collocation points $x_{i,j}$, $i = 1, 2, \dots, M$, $j = 1, 2, \dots, N$.

There is a drawback to this approach though. One element of the characteristic function of the Heston model is $\sqrt{(\kappa - \gamma\rho_{x,v}iu)^2 + (u^2 + iu)\gamma^2}$. When an *imaginary argument* u is used, it cannot be guaranteed that the expression below the square root is non-negative, which may result in inaccurate numerical moment values.

As an alternative approach, similar to the CIR-CLV model, we use the collocation points z_j , $j = 1, \dots, N$ corresponding to the standard normal random variable $Z \stackrel{d}{=} \mathcal{N}(0, 1)$ – see also Remark 5.3.2. Given these, we compute the collocation points corresponding to expiry T_i , $i = 1, \dots, M$, according to

$$x_{i,j} := x_j(T_i) = F_{X(T_i)}^{-1}(F_Z(z_j)), \quad (5.3.11)$$

and the collocation values as in (5.3.10).

To compute the collocation points according to (5.3.11), we define $\widehat{X}(\cdot) := \log(X(\cdot))$ and note

$$F_{X(T_i)}(x_{i,j}) = F_{\widehat{X}(T_i)}(\widehat{x}_{i,j}).$$

The CDF of $\widehat{X}(\cdot)$ can be obtained efficiently in a ‘COS-like’ fashion, see [38]. Given the approximation for its PDF

$$f_{\widehat{X}(t)}(x) \approx \sum_{k=0}^{N-1} F_k \cos\left(k\pi \frac{x-a}{b-a}\right), \quad F_k = \frac{2}{b-a} \operatorname{Re} \left\{ \phi_{\widehat{X}(t)}\left(\frac{k\pi}{b-a}\right) \cdot \exp\left(-i \frac{k a \pi}{b-a}\right) \right\},$$

with N denoting the number of terms, we obtain the CDF of $\widehat{X}(\cdot)$ as follows:

$$F_{\widehat{X}(t)}(x) = \sum_{k=0}^{N-1} \frac{2}{b-a} \operatorname{Re} \left\{ \phi_{\widehat{X}(t)}\left(\frac{k\pi}{b-a}\right) \cdot \exp\left(-i \frac{k a \pi}{b-a}\right) \right\} \psi_k(a, b, x),$$

with

$$\psi(a, b, x) = \begin{cases} \frac{b-a}{k\pi} \sin\left(k\pi \frac{x-a}{b-a}\right), & \text{if } k = 1, 2, \dots, N-1, \\ x-a, & \text{if } k = 0. \end{cases}$$

Note that $\sum_{k=0}^{N-1}$ represents a summation where the first term ($k = 0$) is multiplied by $1/2$. In our numerical experiments we typically use $N = 2^{12}$, $a = -10$ and $b = 10$. Given the CDF, we obtain $\widehat{x}_{i,j} = F_{\widehat{X}(T_i)}^{-1}(F_Z(z_j))$ and $x_{i,j} = e^{\widehat{x}_{i,j}}$.

CHOICE OF X_0 AND r

An initial calibration of the Heston kernel parameters may enhance the performance of the CLV model, in particular of the stochastic collocation method, as a pre-calibration may ‘linearize’ the relationship between $S(\cdot)$ and the kernel parameter $X(\cdot)$ – see Remark 5.2.2 and Section 5.3.2. Essentially, this means that we prefer that $X(\cdot)$ and $S(\cdot)$ are similar in a distributional sense, which yields a small approximation error [58] and a more optimal performance of the stochastic collocation method.

A first ‘calibration’ step is done by determining values for certain kernel process parameters such that the following condition holds (‘first moment matching’):

$$\mathbb{E}^Q[X(t)] = F(0, t), \quad \text{for all } t \in [0, T], \quad (5.3.12)$$

where the initial forward (in an FX context) is given by

$$F(0, t) = S_0 \frac{P_f(0, t)}{P_d(0, t)} = S_0 e^{(r_d - r_f)t},$$

with $P_f(0, t)$ and $P_d(0, t)$ denoting the foreign and domestic zero-coupon bond prices, respectively, extracted from the market quotes. For the Heston kernel process (assuming constant interest rate r) a standard result is

$$\mathbb{E}^Q \left[\frac{X(t)}{M(t)} \right] = \mathbb{E}^Q \left[\frac{X(t)}{M_0 e^{rt}} \right] = \frac{X_0}{M_0}.$$

Assuming $M_0 = 1$ without loss of generality, the previous equation implies

$$\mathbb{E}^Q [X(t)] = X_0 M(t) = X_0 e^{rt}. \tag{5.3.13}$$

From the result in (5.3.13) we easily see that the condition in (5.3.12) is satisfied by the kernel process parameter choices $X_0 = S_0$, $r = r_d - r_f$, for *arbitrary* time t and *arbitrary* values of the other kernel parameters κ , γ , \bar{V} and V_0 .

5

THE FORWARD SMILE

In this section we consider the effect of the Heston kernel parameters on the forward smile corresponding to the Heston-CLV model by pricing a forward-start option, with a pay-off at T_2 given by (5.3.5). We use the Quadratic Exponential (QE) scheme of Andersen [5]. We assume a Heston market parameterization with parameters $\kappa = 0.5$, $\gamma = 0.3$, $\rho_{x,v} = -0.1$, $V_0 = 0.04$ and $\bar{V} = 0.04$. Also, $r = 0$ and $S_0 = 1$. Further, we use $N = 6$ collocation points and the Monte Carlo simulation consists of 32 time-steps per year and $5 \cdot 10^5$ paths (5 seeds, each seed constitutes 10^5 paths). Results are given in Figures 5.3.6, 5.3.7 and 5.3.8.

Increasing γ yields a more pronounced implied volatility smile; the curvature increases. For the parameter κ this effect is opposite; a higher value of κ implies less curvature. The correlation parameter $\rho_{x,v}$ yields a rotation of the smile; varying the parameter from -1 to $+1$ yields a counterclockwise rotation, i.e. less skewness. Besides for that, $\rho_{x,v}$ also has some positive level effect. Further, the long-run variance \bar{V} and the initial variance V_0 mainly have a level effect – note that a larger value of V_0 yields a lower level. X_0 does not have any effect, which implies that a parallel shift of the kernel variable distribution is ‘ignored’ by the mapping between the distribution of the kernel variable X and the market-implied distribution. We summarize the effect of the kernel parameters on the shape of the forward smile in Table 5.3.3. To observe the curvature and skewness effects more clearly, in Appendix 5.D we display Figures 5.3.6, 5.3.7 and 5.3.8 without the level effect.

The effects of the Heston kernel parameters *in the Heston-CLV model* can be compared to the effects of the parameters in the standard Heston model – see e.g. [126]. In both the Heston-CLV model and the Heston model the volatility of variance γ (mainly) has a positive effect on the curvature of the smile⁸. For both models the initial variance

⁸In this short paragraph on the qualitative effects of the Heston model parameters we make use of [126], which merely mentions the first-order effects. It *may* be possible that the parameters have second-order effects, which is the case for e.g. the Heston-Displaced Diffusion model, see [96].

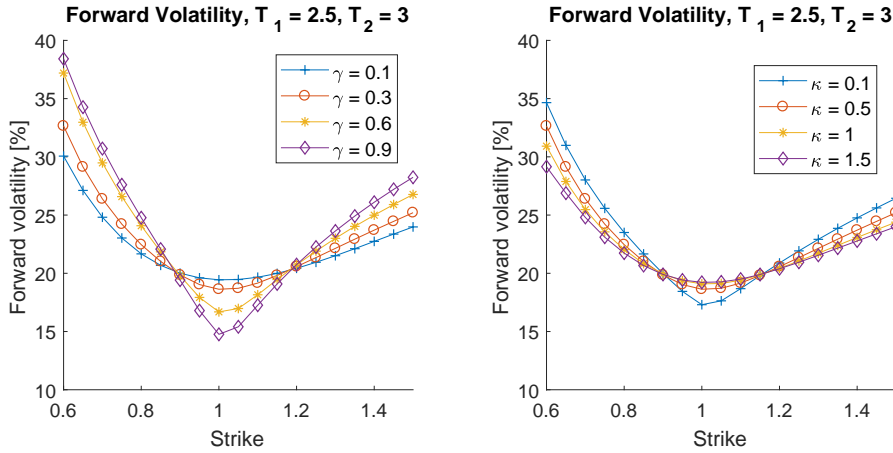


Figure 5.3.6: Effect of γ (left) and κ (right) on the forward volatility smile for the Heston-CLV model, with $T_1 = 2.5$ and $T_2 = 3$. The ‘base case’ Heston parameters (if not varied) are: $\kappa = 0.5, \gamma = 0.3, \rho_{x,v} = 0, \bar{V} = 0.2, V_0 = 0.2$ and $X_0 = 1$.

	Level	Curvature	Skewness
γ	0	+	0
κ	0	-	0
$\rho_{x,v}$	+	0	-
\bar{V}	+	-	+
V_0	-	0	0
X_0	0	0	0

Table 5.3.3: Separate effects of the Heston kernel parameters on the level, curvature and skewness of the forward smile. A ‘+’/‘-’ represents a higher/lower volatility smile level, more/less curvature or more/less skewness in the case of increasing a particular kernel parameter. More skewness represents a more *negative* slope of the forward smile, i.e. a *clockwise* rotation. A ‘0’ stands for no effect.

V_0 and the long-term variance \bar{V} have a level effect, although for the Heston-CLV model V_0 has a negative effect on the level, whereas in the Heston model this is positive. In both the Heston-CLV model and the Heston model the mean reversion parameter κ has a negative curvature effect. In both models the correlation parameter $\rho_{x,v}$ (mainly) accounts for the skewness (‘steepness’) of the smile – in both models a more negative correlation implies a more negative slope, i.e. more skewness. For the Heston-CLV model though $\rho_{x,v}$ also has some level effect.

5.4. CALIBRATION TO FX BARRIER OPTIONS

The CLV model, by its flexibility in controlling the forward smile and its rapid Monte Carlo evaluation, allows for an efficient Monte Carlo calibration to exotic options, while the fit to European-type options is preserved. In this section we calibrate the OU-CLV, CIR-CLV and Heston-CLV models to FX barrier options by Monte Carlo simulation.

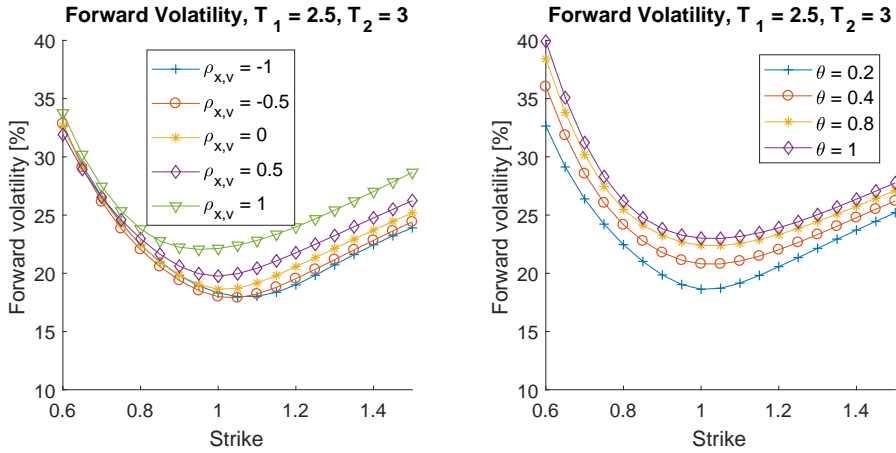


Figure 5.3.7: Effect of $\rho_{x,v}$ (left) and \bar{V} (right, denoted by θ) on the forward volatility smile for the Heston-CLV model, with $T_1 = 2.5$ and $T_2 = 3$. The ‘base case’ Heston parameters (if not varied) are: $\kappa = 0.5$, $\gamma = 0.3$, $\rho_{x,v} = 0$, $\bar{V} = 0.2$, $V_0 = 0.2$ and $X_0 = 1$.

5

The transition densities between future states are reflected by the forward smile a model implies. As a consequence, in the calibration of the CLV model to barrier options we should calibrate the kernel parameters which affect the forward smile. Moreover, ideally, given particular kernel dynamics, to achieve the most accurate calibration we should calibrate kernel parameters which affect *different characteristics* of the shape of the forward smile, namely its level, curvature and skewness.

If the distribution of the kernel variable is analytically known, one time-step is sufficient to price back European-type options – this e.g. holds for the OU-CLV and CIR-CLV models. In the case of a discretely monitored barrier option, for these models the kernel process only needs to be simulated on a time grid consisting of the monitoring dates, see e.g. the numerical example in Appendix 5.B.

In the case of a continuously monitored barrier option, the kernel process needs to be simulated on a dense time-grid – the barrier option is typically monitored on a daily basis, which implies 250 time-steps (business days) per year. To accelerate the Monte Carlo calibration procedure of the kernel parameters to continuously monitored barrier options, we employ *Brownian bridge* techniques, which we describe in Section 5.4.2. First however, in Section 5.4.1 we describe the general steps of applying the CLV model in a Monte Carlo simulation framework.

5.4.1. MONTE CARLO SIMULATION FRAMEWORK

We apply the CLV model in a Monte Carlo simulation framework. Its evaluation basically consists of three parts. First, we compute the collocation points and collocation values. Subsequently, we simulate the kernel variable and compute the local volatility function for the time-points of interest. Last, given the local volatility function values, we establish the price of the relevant financial contract in a standard way. In Algorithm 5 we describe the steps in more detail. Note that in the calibration to exotic options we

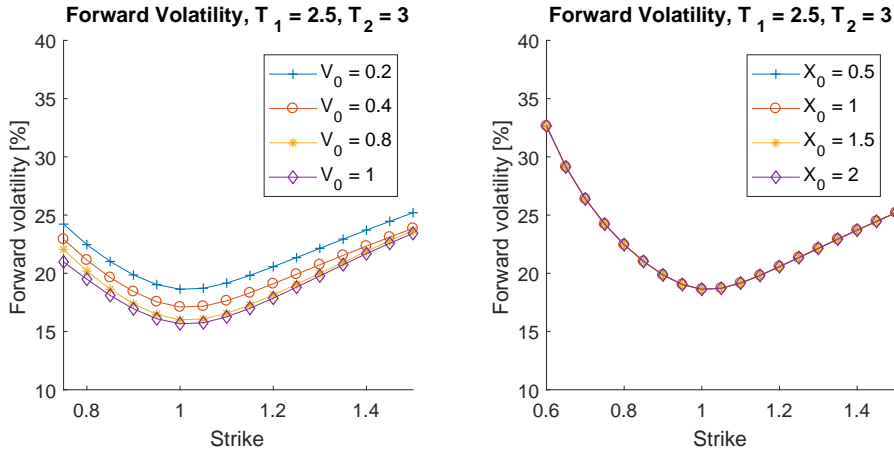


Figure 5.3.8: Effect of V_0 (left) and X_0 (right) on the forward volatility smile for the Heston-CLV model, with $T_1 = 2.5$ and $T_2 = 3$. The ‘base case’ Heston parameters (if not varied) are: $\kappa = 0.5$, $\gamma = 0.3$, $\rho_{x,v} = 0$, $\bar{V} = 0.2$, $V_0 = 0.2$ and $X_0 = 1$.

repeatedly perform the steps 2, 3 and 4 for different kernel parameter value ‘candidates’. In Appendix 5.B we present a basic numerical experiment in which we perform the steps

- 1 Given liquid European-type option prices $C(T_i, K_\ell)$, $i = 1, \dots, M$, $\ell = 1, \dots, L$, establish for each expiry T_i the market-implied CDF $F_{S(T_i)}(x)$ as given in equation (5.2.3), which may be in a parameterized form.
- 2 For each expiry T_i compute N collocation points $x_{i,1}, \dots, x_{i,N}$ and the corresponding collocation values $s_{i,1}, \dots, s_{i,N}$ via $s_{i,j} = F_{S(T_i)}^{-1}(F_{X(T_i)}(x_{i,j}))$, $j = 1, \dots, N$.
- 3 Simulate the kernel variable $X(\cdot)$. At relevant time-points t_k , $k = 1, \dots, K$ in the simulation, compute $g_N(t_k, x_{k,m})$, with $x_{k,m}$ denoting the value of $X(t_k)$ corresponding to the m th path, $m = 1, \dots, \mathcal{M}$.
- 4 Compute the price of the relevant financial contract.

Algorithm 5: Applying the CLV model in a Monte Carlo simulation framework.

in Algorithm 5 to price a discretely monitored barrier option.

Remark 5.4.1 (Evaluation of the market-implied CDF). *Regarding the market-implied CDF in step 1 in Algorithm 5, the derivative in (5.2.3) may be computed by finite differences. However, this approach may not be arbitrage-free as inter- and extrapolation of market volatilities or prices needs to be applied. Another possibility is first calibrating a particular parameterization to the market quotes, enabling us to compute the derivative in (5.2.3) (semi-)analytically, by e.g. an arbitrage-free ‘Hagan implied density’ [57] or Fourier-based pricing techniques, see e.g. [20].*

Figure 5.4.1 provides an illustration of the CLV model in a Monte Carlo simulation

framework. The local volatility function $g_N(\cdot, \cdot)$ transforms the original paths of the kernel variable in such a way that the resulting $S(\cdot)$ -paths yield a perfect calibration to liquid market quotes.

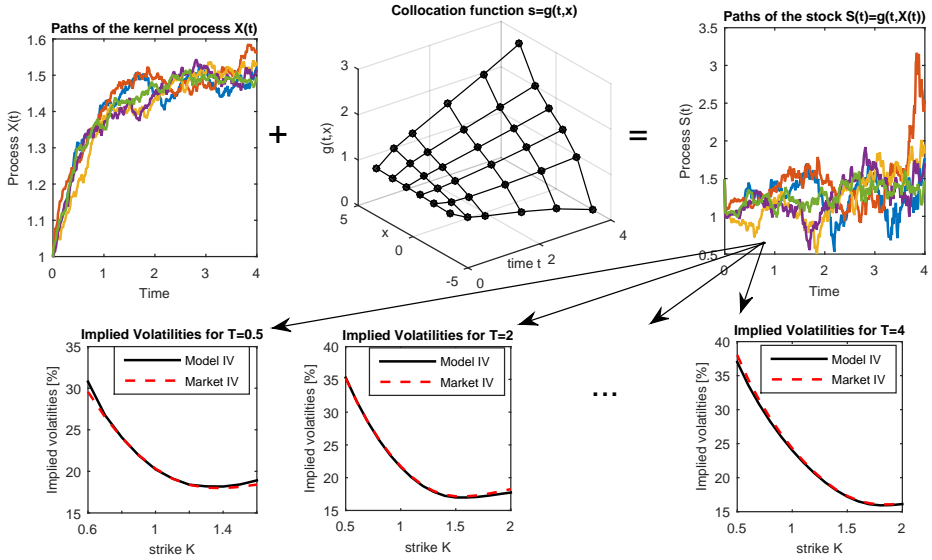


Figure 5.4.1: The CLV model in a Monte Carlo simulation framework.

5.4.2. PRICING BARRIER OPTIONS: A BROWNIAN BRIDGE APPROACH

Pricing continuously monitored barrier options by a Monte Carlo simulation implies an approximation, as we make steps in time and as such monitor on a discrete basis. A way to reduce the error introduced by this approximation is by making use of Brownian bridge techniques [11, 49, 51], where the *conditional hitting probability* is taken into account.

We explain the concept in more detail by considering an up-out barrier put option – for other single barrier products similar results hold. Defining $S(\cdot)$ as the underlying, the price at t_0 of an up-out put option with strike K , barrier B , starting time t_0 and time to maturity T is given by:

$$C_{\text{UO-Put}}(t_0, T, K) := \frac{M_d(t_0)}{M_d(T)} \mathbb{E}^{\mathbb{Q}} \left[(K - S(T))^+ \mathbf{1}_{\left(\max_{t \in [t_0, T]} S(t) < B\right)} \middle| \mathcal{F}(t_0) \right], \quad (5.4.1)$$

with $M_d(\cdot)$ defined as the domestic money account. As for the CLV framework, $S(t) = g_N(t, X(t))$, we have

$$\begin{aligned} V_{\text{UO-Put}}(t_0, T, K) &= \frac{M_d(t_0)}{M_d(T)} \mathbb{E}^{\mathbb{Q}} \left[(K - g_N(T, X(T)))^+ \mathbf{1}_{\left(\max_{t \in [t_0, T]} (g_N(t, X(t)) < B\right)} \middle| \mathcal{F}(t_0) \right] \\ &= \frac{M_d(t_0)}{M_d(T)} \mathbb{E}^{\mathbb{Q}} \left[(K - g_N(T, X(T)))^+ \mathbf{1}_{\left(\max_{t \in [t_0, T]} (X(t) < g_N^{-1}(B))\right)} \middle| \mathcal{F}(t_0) \right], \end{aligned}$$

where we employ the fact that by construction $g_N(\cdot, \cdot)$ is a monotone function.

Defining

$$\bar{B} := g_N^{-1}(B),$$

we write the option value – suppressing the discounting, the \mathbb{Q} -superscript and filtration, for notation purposes – as follows:

$$\begin{aligned} V_{\text{UO-Put}}(t_0, T, K) &= \mathbb{E} \left[\mathbb{E} \left[(K - g_N(T, X(T)))^+ \mathbf{1}(\max_{t \in [t_0, T]} X(t) < \bar{B}) \middle| X(t_0), X(T) \right] \right] \\ &= \mathbb{E} \left[(K - g_N(T, X(T)))^+ \cdot \left(1 - \mathbb{Q} \left(\max_{t \in [t_0, T]} X(t) \geq \bar{B} \middle| X(t_0), X(T) \right) \right) \right]. \end{aligned}$$

We are interested in the probability

$$\mathbb{Q} \left(\max_{t \in [t_0, T]} X(t) \geq \bar{B} \middle| X(t_0), X(T) \right), \quad (5.4.2)$$

i.e. the probability that the maximum of $X(\cdot)$ on $[t_0, T]$ hits or crosses \bar{B} , given $X(t_0)$ and $X(T)$ (trivially, if $X(t_0) \geq \bar{B}$ and/or $X(T) \geq \bar{B}$, this probability is 1). In the following, we explain how this *conditional hitting probability* can be approximated.

Given a general one-dimensional process under the \mathbb{Q} -measure:

$$dX(t) = a(X(t))dt + b(X(t))dW^{\mathbb{Q}}(t), \quad X_0 := X(0).$$

For simulating this process we use the Euler discretization

$$\hat{X}_{k+1, m} = \hat{X}_{k, m} + a(\hat{X}_{k, m})\Delta_t + b(\hat{X}_{k, m})(W_{k+1, m} - W_{k, m}), \quad (5.4.3)$$

with $k = 0, \dots, K-1$ and $m = 1, \dots, \mathcal{M}$ indicating the time-step and path, respectively, and $\Delta_t := T/K$, $\hat{X}_0 := X_0$, $\hat{X}_k := \hat{X}(t_k)$, with $t_k = k\Delta_t$.

Result 5.4.1 (Simulation of the maximum). *Given the values $\hat{X}_{k, m}$ and $\hat{X}_{k+1, m}$, the maximum*

$$\bar{X}_{k, m} := \max_{t \in [t_k, t_{k+1}]} \hat{X}(t)$$

can be simulated by

$$\bar{X}_{k, m} = \frac{1}{2} \left(\hat{X}_{k+1, m} + \hat{X}_{k, m} + \sqrt{(\hat{X}_{k+1, m} - \hat{X}_{k, m})^2 - 2b^2(\hat{X}_{k, m})\Delta_t \log(U_{k, m})} \right), \quad (5.4.4)$$

with $U_{k, m} \stackrel{d}{=} \mathcal{U}[0, 1]$ being independent across the time-steps. A step to arrive at (5.4.4) is that

$$\{X(t), t_k \leq t \leq t_{k+1}\}$$

is approximated by an arithmetic Brownian motion with constant parameters $a(\hat{X}_{k, m})$ and $b(\hat{X}_{k, m})$. Conditional on the endpoint $X(t_{k+1})$, $\{X(t), t_k \leq t \leq t_{k+1}\}$ is a Brownian bridge.

Using (5.4.4) in Result 5.4.1, one can compute the probability that the discretized process hits the barrier \bar{B} in the k th step, conditional on the values of $\hat{X}_{k,m}$ and $\hat{X}_{k+1,m}$. Straightforward calculus yields

$$\begin{aligned} & \mathbb{Q}\left(\bar{X}_{k,m} \geq \bar{B} \mid \hat{X}_{k,m}, \hat{X}_{k+1,m}\right) \\ &= \begin{cases} 1 & \text{if } \hat{X}_{k,m} \geq \bar{B} \text{ and/or } \hat{X}_{k+1,m} \geq \bar{B}, \\ \exp\left(-\frac{2}{b^2(\hat{X}_{k,m})\Delta t}(\hat{X}_{k,m} - \bar{B})(\hat{X}_{k+1,m} - \bar{B})\right) & \text{if } \hat{X}_{k,m} < \bar{B} \text{ and } \hat{X}_{k+1,m} < \bar{B}. \end{cases} \end{aligned} \quad (5.4.5)$$

Applying the discretization scheme in (5.4.3), the conditional hitting probability in (5.4.2) can be approximated by

$$\begin{aligned} \mathbb{Q}\left(\max_{t \in [t_0, T]} X(t) \geq \bar{B} \mid X(t_0), X(T)\right) &= 1 - \mathbb{Q}\left(\max_{t \in [t_0, T]} X(t) < \bar{B} \mid X(t_0), X(T)\right) \\ &\approx 1 - \prod_{k=0}^{K-1} \left\{1 - \mathbb{Q}\left(\bar{X}_{k,m} \geq \bar{B} \mid \hat{X}_{k,m}, \hat{X}_{k+1,m}\right)\right\}, \end{aligned}$$

with $\mathbb{Q}\left(\bar{X}_{k,m} \geq \bar{B} \mid \hat{X}_{k,m}, \hat{X}_{k+1,m}\right)$ given in (5.4.5). Substituting this result in (5.4.2) yields:

$$\begin{aligned} V_{\text{UO-Put}} &\approx \frac{M_d(t_0)}{M_d(T)} \mathbb{E}^{\mathbb{Q}}\left[\left(K - g_N(T, X(T))\right)^+ \cdot \prod_{k=0}^{K-1} \left\{1 - \mathbb{Q}\left(\bar{X}_k \geq \bar{B} \mid \hat{X}_k, \hat{X}_{k+1}\right)\right\}\right] \\ &= \frac{1}{\mathcal{M}} \frac{M_d(t_0)}{M_d(T)} \sum_{m=1}^{\mathcal{M}} \left[\left(K - g_N(T, \hat{X}_{K,m})\right)^+ \prod_{k=0}^{K-1} \left\{1 - \mathbb{Q}\left(\bar{X}_{k,m} \geq \bar{B} \mid \hat{X}_{k,m}, \hat{X}_{k+1,m}\right)\right\}\right], \end{aligned} \quad (5.4.6)$$

with $\mathbb{Q}\left(\bar{X}_{k,m} \geq \bar{B} \mid \hat{X}_{k,m}, \hat{X}_{k+1,m}\right)$ given in (5.4.5). In the calibration of the CLV model to continuously monitored barrier options we make use of the expression in equation (5.4.6).

5.4.3. CALIBRATION OF THE OU-CLV, CIR-CLV AND HESTON-CLV MODELS TO FX BARRIER OPTIONS

In this section we calibrate the OU-CLV, CIR-CLV and Heston-CLV models to continuously monitored FX barrier option prices. In particular, for these 3 models we perform the following 2 steps:

1. **Calibration:** given a particular kernel process, calibrate the relevant kernel parameter(s) to market barrier option prices by Monte Carlo simulation⁹. The Monte Carlo simulation runs consist of 10 time-steps per year and we employ the Brownian bridge technique described in Section 5.4.2 – in particular, we use the result in equation (5.4.6).
2. **Pricing:** given the calibrated kernel parameter values, price the up-out barrier put options by a standard Monte Carlo procedure of the CLV model with 250 time-

⁹The target function value is $\sum_{i=1}^N \left(V_i - V_i^{\text{CLV}}\right)^2$, with V_i and V_i^{CLV} denoting the mid-market price and the CLV price of the i th up-out put option, respectively, and N is the number of barrier options we calibrate to.

steps per year to determine the calibration error, which is defined as

$$\epsilon := \sum_{i=1}^N |V_i - V_i^{\text{CLV}}|, \quad (5.4.7)$$

with V_i and V_i^{CLV} denoting the mid-market price and the CLV price of the i th up-out put option, respectively, and N is the number of barrier options we calibrate to.

We consider USD/AUD FX market prices quoted on 12 June 2013 from a market data vendor. Domestic currency is USD, foreign currency is AUD. Initial spot price is $S_0 = 0.9548$. The dataset consists of 7 expiry dates, namely 0.5, 0.75, 1, 2, 3, 4 and 5 years. For each expiry 5 implied volatility quotes are given, of which the middle (third) one is the ATM volatility. We calibrate to 9 continuously monitored *up-out barrier put options* with different barrier and strike values, of which the price is given by (5.4.1).

For all three kernel process choices (OU, CIR and Heston) we compute $x_{i,j}$ and $s_{i,j}$ values for the expiries 1/365, 2/365, 3/365, 4/365, 1/52, 2/52, 1/12, 1/6, 1/4, 1/2, 3/4, 1 and 2, for which market volatility quotes are available¹⁰. In both the calibration and in the pricing afterwards we use 10^4 paths¹¹. The target function is defined as the sum of squared errors of the 9 barrier options together, where the error is defined as the difference between the model and mid market prices.

THE OU-CLV MODEL

In Section 5.3.1 we observed that only the mean reversion parameter κ of the OU-CLV model affects the shape of the forward smile. Therefore, we calibrate κ and the other parameters are set to $X_0 = 1$, $\theta = 0.1$, $\gamma = 0.25$, which were just chosen values in [54]. Further, we use 6 collocation points. In the calibration we price for 20 ' κ candidates' -1, -0.9, ..., 0.9, 1 (excluding $\kappa = 0$) up-out barrier put options by (5.4.6) in a Monte Carlo simulation.

The calibration results in $\kappa = 0.1$ and takes 8 seconds¹². Results are displayed in Table 5.4.1. Two barrier option prices are outside the bid-ask spread. By construction, the OU-CLV model calibrates perfectly to European-type options – in Figure 5.4.2 we display the implied volatilities corresponding to 3M, 1Y and 2Y corresponding to the market, market parameterization and the OU-CLV model.

THE CIR-CLV MODEL

In the CIR-CLV model, parameter γ has the most pronounced curvature effect, see Section 5.3.2. It also has a level effect. We therefore calibrate γ and leave the other parameters fixed, namely $\kappa = 0.1$, $\theta = 0.1$ and $X_0 = 1$. In the calibration we price for 20 ' γ candidates' on a uniform grid between 0.1 and 1 up-out put options by (5.4.6) in a Monte Carlo simulation.

¹⁰We use a SABR market parameterization.

¹¹In fact, we use 10 seeds with each seed constituting 10^3 paths. For the OU-CLV case we apply antithetic sampling, i.e. per seed we use $5 \cdot 10^2$ paths and $5 \cdot 10^2$ 'antithetic paths'.

¹²Processor: Intel(R) Core(TM) i7-4790 CPU @ 3.60GHz, 3601 Mhz, 4 Core(s), 8 Logical Processor(s). Available Physical Memory is 9.08 GB of in total 16 GB. Simulated with MATLAB.

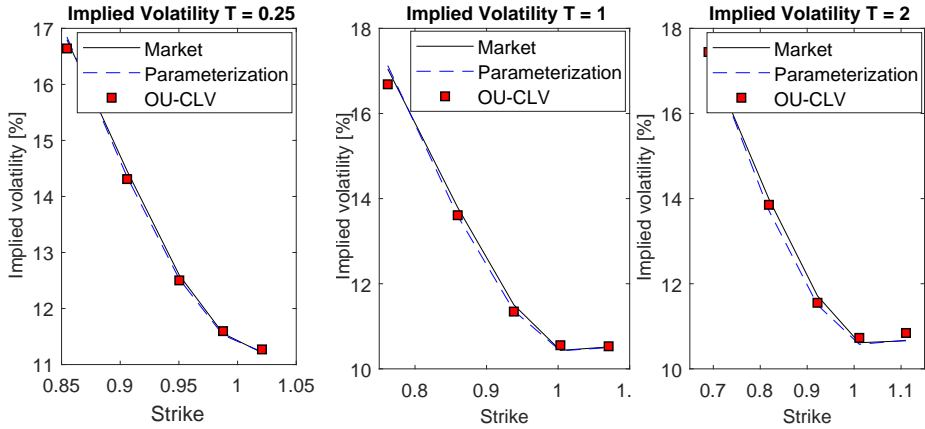


Figure 5.4.2: Implied volatilities corresponding to the OU-CLV model with $\kappa = 0.1$, $X_0 = 1$, $\theta = 0.1$, $\gamma = 0.25$.

5

The calibration results in $\gamma = 0.526$ and takes 39 seconds^{12, 13}. Results are displayed in Table 5.4.1. Two barrier option prices are outside bid-ask spread. The calibration accuracy may be further enhanced by calibrating an additional parameter. *Additionally, we obtain an accurate calibration to the European-type options by construction, similar to the results in Figure 5.4.2 for the OU-CLV model.*

THE HESTON-CLV MODEL

As we observed in Section 5.3.3, the mean reversion κ , volatility of variance γ and long-term variance \bar{V} have a curvature effect – see also Table 5.3.3. As the curvature effect of γ is most pronounced, we calibrate γ and fix κ and \bar{V} . Assuming κ to be constant is justified by observations in [47], where it is argued that the effect on the implied volatility surface of increasing κ is similar to decreasing γ . Besides for \bar{V} , the only parameter having a skewness effect is $\rho_{x,v}$. We therefore also calibrate $\rho_{x,v}$. Last, V_0 has a level effect. We can already achieve an accurate level fit *a priori* by setting $V_0 = \bar{\sigma}_{\text{mkt}}^2(K_{\text{ATM}}, T_{\text{min}})$, with $T_{\text{min}} = 1/365$ and $\bar{\sigma}_{\text{mkt}}(K, t)$ denoting the market implied volatility corresponding to strike K and expiry t . The other fixed parameter values are $\kappa = 0.3$, $X_0 = 1$, $V_0 = \bar{\sigma}_{\text{mkt}}^2(K_{\text{ATM}}, T_{\text{min}}) = 0.0625$, with $T_{\text{min}} = 1/365$ and $\bar{V} = \bar{\sigma}_{\text{mkt}}^2(K_{\text{ATM}}, T_{\text{max}}) = 0.0137$, with $T_{\text{max}} = 2$. We use $N = 6$ collocation points.

In the calibration we choose 5 γ and 5 $\rho_{x,v}$ candidates uniformly between 0.1 and 1 and -1 and 0 , respectively – so in total we consider 25 $(\gamma, \rho_{x,v})$ -pairs. For each of the 25 pairs we price the up-out put options by (5.4.6) in a Monte Carlo simulation.

The calibration results in $\gamma = 0.1$ and $\rho_{x,v} = 0$ and takes 16 seconds¹². Results are displayed in Table 5.4.1. One barrier option price is outside bid-ask spread. The total calibration error (5.4.7) is smaller than for the OU-CLV and CIR-CLV models. The reason is that the Heston kernel process is *richer*; compared to the OU-CLV and CIR-CLV model,

¹³This calibration is relatively slow compared to the calibration of the OU-CLV and Heston-CLV models. The reason is that the MATLAB functionality `ncx2cdf(x, d, lambda)` is relatively slow for a large value of x (in fact we divide by a small $c(T_i)$ value, see (5.3.9)) and $\lambda(T_i)$ values, which is the case for the shortest T_i expiries.

Expiry	Barrier	Strike	OU-CLV	CIR-CLV	Heston-CLV
3M	1	0.9	97.2	88.8*	96.3
3M	1	0.85	36.8	26.0*	35.4
3M	0.97	0.9	67.3*	60.3	64.3*
1Y	1	0.8	99.2	92.9	98.7
1Y	1	0.85	161.2	157.9	162.1
1Y	0.97	0.8	50.8	50.8	52.3
2Y	1	0.75	128.4	132.0	147
2Y	1.05	0.7	109.5*	121.7	131.1
2Y	0.97	0.75	53.6	61.9	66.4
Calibration error ϵ			53.9	38.4	33.5
Calibration time¹²			8s	39s	16s

Table 5.4.1: Pricing up-out put options with the OU-CLV, CIR-CLV and Heston-CLV models after the calibration. The Monte Carlo simulation consists of 10^4 paths and 250 time-steps per year. ϵ is defined as the sum of absolute errors over the 9 barrier options, i.e. $\epsilon := \sum_{i=1}^9 |V_i - V_i^{\text{CLV}}|$, with V_i and V_i^{CLV} denoting the mid-market price and the CLV price of the i th up-out put option, respectively. The red values marked by an asterisk (*) are not within bid-ask spread.

the Heston-CLV model is more flexible in capturing the forward smile, and as such in the calibration to forward volatility sensitive products, like barrier options. The calibration accuracy may be further enhanced by calibrating an additional parameter.

5.5. CONCLUSION

In this chapter we discussed a competitive alternative to stochastic-local volatility models, namely the Collocation Local Volatility (CLV) model, introduced in [54]. In the CLV model the local volatility function, based on stochastic collocation [9, 45, 127, 131, 131], connects a relatively simple and easy to implement kernel process to the market, resulting in a perfect calibration to liquid market quotes. The local volatility function only needs to be evaluated at the time-points of interest, e.g. the monitoring dates of a discrete barrier option. Moreover, efficient simulation schemes for the kernel process exist. Further, a proper choice of the kernel process allows the CLV model to be flexible to ‘capture’ the forward smile and, as such, price path-dependent products.

As the CLV model is sufficiently flexible for controlling the forward smile and can be efficiently evaluated, it allows for a rapid calibration to exotic options, while the fit to European-type options is preserved. In Section 5.4 we employ Brownian bridge techniques to calibrate the kernel process to continuously monitored barrier options in an efficient way; the calibration of the Heston-CLV model to the barrier options costs 16 seconds.

Based on the calibration results in Section 5.4, we prefer the Heston-CLV model for the calibration to continuously monitored path-dependent options. Its calibration is reasonably fast and yields the smallest calibration error; the reason is that among the kernel processes considered, the Heston kernel process allows for the most flexibility in capturing the forward smile. Further, for the Heston model low-bias large time-stepping Monte Carlo schemes exist, e.g. Andersen’s QE scheme [5], which can be employed for

the simulation of the Heston kernel.

In the case that we are interested in an exotic option monitored on a coarse grid, the CIR-CLV model may be preferred. Brownian bridge techniques are not needed as the CIR kernel process allows for large time-steps. For example, for the pricing of forward-start options in Section 5.3.2 the Monte Carlo simulation merely consists of 2 time-steps per year. Additionally, the CIR parameters allow for controlling the level and the curvature of the forward smile, see Section 5.3.2.

In terms of simplicity and calibration speed (to exotics), the OU-CLV model outperforms the CIR-CLV and Heston-CLV models. In the OU-CLV model only the mean-reversion parameter κ has an effect on the forward volatility. As analytical expressions for its moments are available and the moments are numerically stable, we can compute optimal collocation points. Similar to the CIR-CLV model, the OU-CLV model allows for large time-steps.

APPENDIX

5.A. OPTIMAL COLLOCATION POINTS

One of the relevant theorems with respect to computing optimal collocation points is the following [39]:

Theorem 5.A.1 (Recurrence in orthogonal polynomials). *For any given density function $f_X(\cdot)$, a unique sequence of monic orthogonal polynomials $p_i(x)$ exists, with $\deg(p_i(x)) = i$, which can be constructed by*

$$p_{i+1}(x) = (x - \alpha_i)p_i(x) - \beta_i p_{i-1}(x), \quad i = 0, 1, \dots, N-1,$$

with $p_{-1}(x) \equiv 0$, $p_0(x) \equiv 1$ and the recurrence coefficients

$$\alpha_i = \frac{\mathbb{E}[X p_i^2(X)]}{\mathbb{E}[p_i^2(X)]}, \quad i = 0, 1, \dots, N-1, \quad \beta_i = \frac{\mathbb{E}[p_i^2(X)]}{\mathbb{E}[p_{i-1}^2(X)]}, \quad i = 1, 2, \dots, N-1,$$

with $\beta_0 = 0$.

Proof. For a proof, see [39]. □

The recurrence coefficients α_i and β_i can be obtained via the moments of X . In particular, one can express the first N coefficients in terms of the elements of a lower triangular matrix R , which is obtained by the Cholesky decomposition of a matrix $M = R^T R$ that constitutes the first $2N$ moments of X . Given the recurrence coefficients, the *optimal* collocation points x_1, x_2, \dots, x_N are the zeros of the orthogonal polynomial $p_N(x)$ and can be computed by an eigenvalue method. Based on this, the only requirement for X to be an appropriate kernel variable for which we can compute N collocation points, is the existence of the first $2N$ moments.

By choosing the collocation points as zeros of the orthogonal polynomial $p_N(\cdot)$ – see Theorem 5.A.1 – the stochastic collocation method can be connected to the computation of integrals by Gauss quadrature, which for the function $\Psi(\cdot)$ (which is required to be well approximated by a polynomial function), weight function $f_X(\cdot)$ and quadrature weights ω_i , $i = 1, 2, \dots, N$ reads:

$$\mathbb{E}[\Psi(X)] = \int_{\mathbb{R}} \Psi(x) f_X(x) dx = \sum_{i=1}^N \Psi(x_i) \omega_i + \epsilon_N. \quad (5.A.1)$$

By choosing $\Psi(x) = (g(x) - g_N(x))^2$, with $g(\cdot) = F_S^{-1}(F_X(\cdot))$ and $g_N(\cdot)$ the approximating polynomial function, we have

$$\mathbb{E}[(g(X) - g_N(X))^2] = \int_{\mathbb{R}} (g(x) - g_N(x))^2 f_X(x) dx = \sum_{i=1}^N (g(x_i) - g_N(x_i))^2 \omega_i + \epsilon_N = \epsilon_N, \quad (5.A.2)$$

since $g(x_i) = g_N(x_i)$, $i = 1, 2, \dots, N$. So, in L^2 the error is determined by the quadrature error. Further, when choosing e.g. X to be the standard normal distribution, a simple linear relation between the stochastic collocation pairs (x_i, ω_i) and the Gauss-Hermite quadrature pairs (x_i^H, ω_i^H) exists.

5.B. NUMERICAL EXPERIMENT: PRICING A DISCRETELY MONITORED BARRIER OPTION

We present a numerical experiment in which we perform the steps in Algorithm 5 to price a discretely monitored up-out call option:

$$V_{\text{UO-call}}(t_0, T, K) := \mathbb{E}^{\mathbb{Q}} \left[\frac{(S(T) - K)^+}{M(T)} \mathbb{1}_{(\max_{t \in \mathcal{T}} S(t) < B)} \middle| \mathcal{F}(t_0) \right],$$

with $B = 1.5$, $K = 0.5$, $T = 3$ and quarterly monitoring dates, i.e. $\mathcal{T} = \{3M, 6M, \dots, 3Y\}$. We use the Ornstein-Uhlenbeck CLV (OU-CLV) model

$$\begin{aligned} S(t) &= g_N(t, X(t)), \\ dX(t) &= \kappa(\theta - X(t)) dt + \gamma dW^{\mathbb{Q}}(t), \quad X(t_0) = X_0, \end{aligned}$$

with $X_0 = 1$, $\kappa = 1$, $\gamma = 0.5$ and $\theta = 0.5$. Given the filtration at $t_0 = 0$, the kernel variable $X(\cdot)$ is normally distributed with mean and variance as in (5.3.4).

We successively apply the steps of Algorithm 5:

1. We generate synthetic market data by the Heston model with the parameters $\kappa = 0.5$, $\gamma = 1$, $\rho_{x,v} = -0.7$, $V_0 = 0.04$, $\bar{V} = 0.04$, $r = 0$ and $S_0 = 1$. We assume that liquid market quotes are available for the expiries 1D, 2D, 3D, 4D, 1W, 2W, 1M, 2M, 3M, 6M, 9M, 1Y, 2Y, 3Y, 4Y and 5Y. In the left-hand plot of Figure 5.B.1 the market-implied CDF is displayed, obtained by Fourier pricing techniques.
2. We use $N = 6$ collocation points. Given the optimal collocation points z_j , $j = 1, \dots, N$ of a standard normal random variable $Z \stackrel{d}{=} \mathcal{N}(0, 1)$ (for $N = 6$ these are $-3.3243, -1.8892, -0.6167, 0.6167, 1.8892, 3.3243$), the optimal collocation points of $X(t)$ are given by $x_j(t) = \mathbb{E}[X(t)] + \sqrt{\text{Var}[X(t)]} \cdot z_j$ [54] and as such

$$x_{i,j} := x_j(T_i) = \mathbb{E}[X(T_i)] + \sqrt{\text{Var}[X(T_i)]} \cdot z_j, \quad i = 1, \dots, M, \quad j = 1, \dots, N.$$

Given the collocation points, by inversion of the market-implied CDF we obtain the collocation values $s_{i,j} = F_{S(T_i)}^{-1}(F_{X(T_i)}(x_{i,j}))$. In the right-hand plot of Figure 5.B.1 the triplets $(x_{i,j}, T_i, s_{i,j})$ are displayed.

3. We simulate the kernel variable with time-steps of 3 months (4 time-steps per year). The Monte Carlo simulation consists of 10^4 paths¹⁴. In order to evaluate $g_N(\cdot, \cdot)$ at time-points that are not part of the market data – so *not* 0.1, 0.25, 0.5, 1, 2, 3, 4 and 5 years – we interpolate between or extrapolate the $s_{i,j}$ values, see e.g. the interpolated $g_N(2.5, \cdot)$ -values in the right-hand plot of Figure 5.B.1.

¹⁴In fact, we use 10 seeds, each seed constitutes 10^3 paths.

4. Last, we price the up-out call option: $V_{\text{UO-call}}(t_0, T, K) = 0.4739$. The calibration to the European-type options is guaranteed to be almost perfect, see Figure 5.B.2.

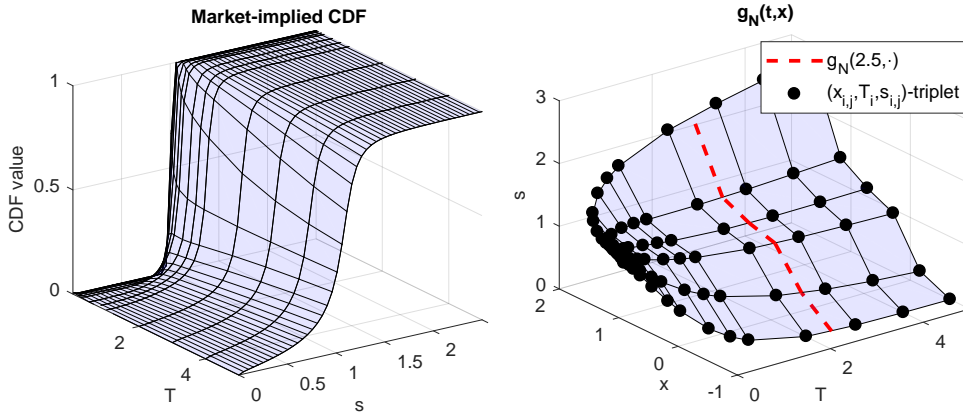


Figure 5.B.1: The market-implied CDF (left) and the $(x_{i,j}, s_{i,j})$ -pairs (right) corresponding to the numerical experiment in Appendix 5.B. The red dashed line indicates the monotone interpolation at $t = 2.5$.

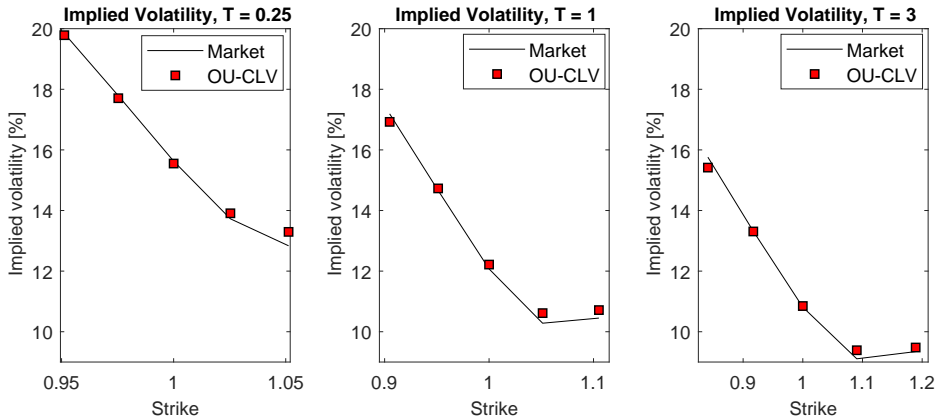


Figure 5.B.2: The implied volatilities for the expiries $T = 0.25$, $T = 1$ and $T = 3$ corresponding to the numerical experiment in Appendix 5.B.

5.C. EFFECT OF CIR PARAMETERS, OMITTING THE LEVEL EFFECT

In Figures 5.C.1 and 5.C.2 we display the effect of the CIR kernel parameters on the shape of the forward smile, see Section 5.3.2, but without the level effect.

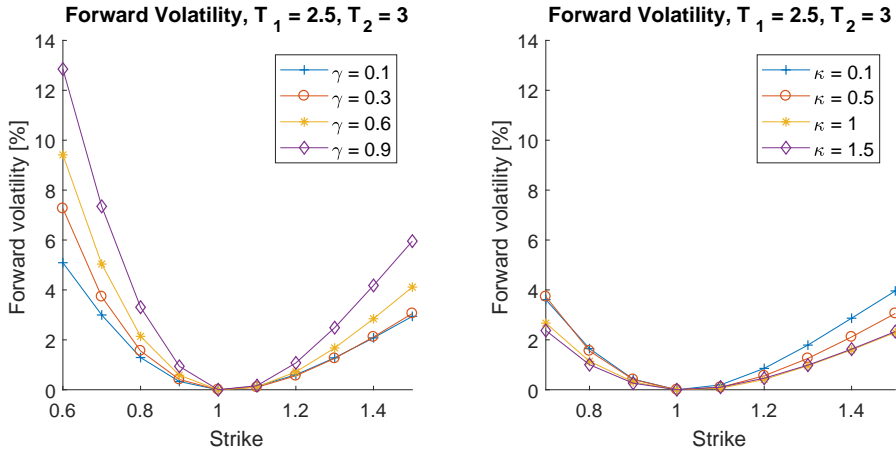


Figure 5.C.1: Effect of γ (left) and κ (right) on the forward volatility smile for the CIR-CLV model, omitting the level effect, with $T_1 = 2.5$ and $T_2 = 3$. The 'base case' CIR parameters (if not varied) are: $\gamma = 0.3$, $\kappa = 0.5$, $\theta = 0.5$ and $X_0 = 1$.

5

5.D. EFFECT OF HESTON PARAMETERS, OMITTING THE LEVEL EFFECT

In Figure 5.D.1 we display the effect of γ and κ on the shape of the forward smile, see Section 5.3.3, but without the level effect. We do not display the effects of all parameters, to save some space.

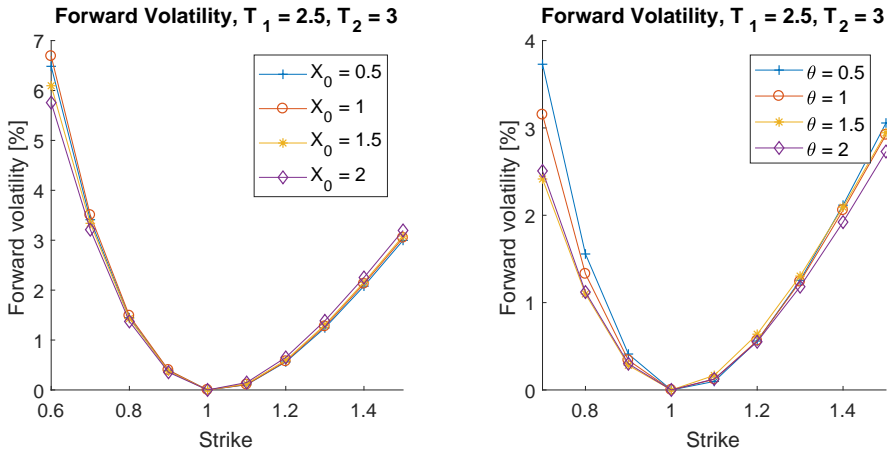


Figure 5.C.2: Effect of X_0 (left) and θ (right) on the forward volatility smile for the CIR-CLV model, omitting the level effect, with $T_1 = 2.5$ and $T_2 = 3$. The 'base case' CIR parameters (if not varied) are: $\gamma = 0.3$, $\kappa = 0.5$, $\theta = 0.5$ and $X_0 = 1$.

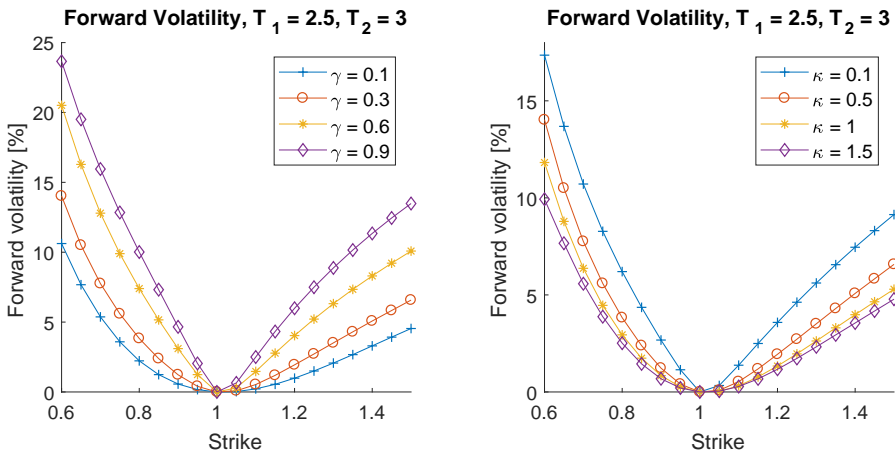


Figure 5.D.1: Effect of γ (left) and κ (right) on the forward volatility smile for the Heston-CLV model, omitting the level effect, with $T_1 = 2.5$ and $T_2 = 3$. The 'base case' Heston parameters (if not varied) are: $\kappa = 0.5$, $\gamma = 0.3$, $\rho_{x,v} = 0$, $\bar{V} = 0.2$, $V_0 = 0.2$ and $X_0 = 1$.

6

CONCLUSIONS AND OUTLOOK

6.1. CONCLUSIONS

In this thesis we solved various problems with respect to the calibration and pricing of enhanced local volatility models in an FX context. All our model evaluation methods are Monte Carlo based.

In Chapter 2 we presented a novel Monte Carlo scheme for the efficient evaluation of a general stochastic local volatility (SLV) model. In numerical experiments we examined the Heston-SLV model. We introduced a non-parametric approximation for the non-trivial conditional expectation, which is intuitive and easy to implement. To enhance the Monte Carlo evaluation of the SLV model, we incorporated the non-parametric approximation in a simulation scheme that is based on the QE scheme of Andersen [5]. Numerical experiments confirmed that European-type options can be priced with high accuracy by our method. Also, we have numerically shown that our method allows for a consistent pricing of products that are sensitive to the forward volatility smile.

Subsequently, in Chapter 3 we established in an FX context a framework that facilitates the accurate and efficient calibration of the time-dependent SABR model. To this purpose, by considering in detail the effects of the SABR parameters on the shape of the implied volatility smile, we derived 'effective parameters', which are the constant 'equivalents' of the time-dependent vol-vol, term structure and correlation parameters. By numerical experiments we confirmed that the derived mappings between the time-dependent and effective parameters result into highly accurate calibration results. In an experiment with real market data, the time-dependent SABR model implies FX barrier option prices that are closest to the market, compared to the Local Volatility (LV) model, the constant-parameter SABR model and the time-dependent SABR model enhanced by a non-parametric local volatility component.

Next, we considered different hybrid local volatility models in Chapter 4, namely the LV model enhanced with stochastic interest rates, and the SABR and Heston models to which a local volatility component is added. For all models, a non-trivial (conditional) expectation needs to be determined. We developed a novel, efficient approach to the

evaluation of this expectation, which consists of two projection steps; the first projection step relies on the equality in distribution of cumulative distribution functions, which is a basic element of stochastic collocation, and the second projection step is based on standard regression techniques. We numerically showed that our approach results into an efficient Monte Carlo evaluation and highly accurate pricing results for European-type options.

Last, in Chapter 5 we discussed the Collocating Local Volatility (CLV) model. We established the corresponding pricing PDE analogous to the derivation of the Black-Scholes pricing PDE. For three different kernel processes, namely the Ornstein Uhlenbeck and CIR processes and the Heston model, we considered the effects of the kernel parameters on the shape of the forward volatility smile. Based on this, we calibrated the corresponding CLV models to market prices of continuously monitored FX barrier option prices. The calibration is Monte Carlo based, where we applied Brownian bridge techniques to reduce the approximation error introduced by the discrete time-stepping. Based on the calibration results, we preferred the Heston-CLV model to the OU-CLV and CIR-CLV models. The Heston-CLV model implied the smallest calibration error and the calibration was reasonably fast. The OU-CLV model stood out from the other two CLV models in terms of simplicity and calibration speed.

6.2. OUTLOOK

As a result of the financial crisis, the market's concern with respect to counterparty risk grew substantially and implied significant changes to regulatory frameworks. The *Basel II* and *Basel III* accords imposed regulatory requirements on the amount of capital that financial institutions need to reserve, in order to compensate for potential losses in the case of a counterparty default. Parallel to the adaption of the financial industry to the post-crisis regulatory standards, academia also shifted their attention to the modelling of various risk measures, such as Expected Exposure (EE), Potential Future Exposure (PFE), and several valuation adjustments, e.g. Credit Valuation Adjustment (CVA), Debt Valuation Adjustment (DVA) and Funding Valuation Adjustment (FVA), among others, commonly referred to as *xVA* [53]. Future research could possibly focus on the impact of different hybrid local volatility model dynamics on the values of *xVA* risk measures – as a starting point, the work in [110] could be considered. As risk measure calculations are portfolio based, one could make a distinction between different portfolio contents. For example, for a portfolio mainly consisting of forward volatility sensitive products, stochastic local volatility (SLV) dynamics are expected to be a more appropriate choice than 'pure' local volatility dynamics. In this thesis we put the focus on an efficient *Monte Carlo approach* of SLV models – as such, this research fits well in the *xVA* framework, where the simulation of market risk drivers plays a key role and, in combination with the pricing of portfolios, brings along computational challenges.

In another possible research direction, one could consider model calibration with *artificial neural networks* (ANNs). In general, ANNs can be considered as an extension of regression and have increasingly been applied in the fields of image and speed recognition. One could start with further improving the work performed in [65], in which a single-factor Hull-White model is calibrated utilizing neural networks. Subsequently, more complicated models can be calibrated, such as the time-dependent FX-SABR model

discussed in Chapter 2 of this thesis. Also, ANNs may yield an alternative to the Monte Carlo calibration of the CLV model to FX barrier option prices, performed in Chapter 5.

With these new challenges, there is plenty of exciting research ahead of us.

REFERENCES

- [1] Special FX – CLS Keeps the Market Safe from Settlement Risk but Needs to Add More Currencies. *The Economist*, September 2013.
- [2] Triennial Central Bank Survey – Foreign exchange turnover in April 2016. Technical report, Bank for International Settlements, September 2016.
- [3] M. Abramowitz and I. A. Stegun. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th edition, 1972.
- [4] Y. An and C. Li. Approximating Local Volatility Functions of Stochastic Volatility Models: A Closed-Form Expansion Approach. *Probability in the Engineering and Informational Sciences*, 29(04):547–563, 2015.
- [5] L. Andersen. Simple and Efficient Simulation of the Heston Stochastic Volatility Model. *Journal of Computational Finance*, 11(3):1–42, 2008.
- [6] J. Andreasen and B. Hуже. Volatility Interpolation. *Risk*, 3:86–89, 2011.
- [7] A. Antonov, M. Konikov, and M. Spector. SABR Spreads its Wings. *Risk*, pages 58–63, August 2013.
- [8] M. Atlan. Localizing volatilities. *ArXiv preprint math/0604316*, 2006.
- [9] I. Babuška, F. Nobile, and R. Tempone. A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.
- [10] G. Baker, R. Beneder, and A. Zilber. FX Barriers with Smile Dynamics. *Available at SSRN 964627*, 2004.
- [11] P. Baldi. Exact asymptotics for the probability of exit from a domain and applications to simulation. *The Annals of Probability*, pages 1644–1670, 1995.
- [12] J. Beck, R. Tempone, F. Nobile, and L. Tamellini. On the Optimal Polynomial Approximation of Stochastic PDEs by Galerkin and Collocation Methods. *Mathematical Models and Methods in Applied Sciences*, 22(09):1–33, 2012.
- [13] E. Benhamou, E. Gobet, and M. Miri. Analytical Formulas for a Local Volatility Model with Stochastic Rates. *Quantitative Finance*, 12(2):185–198, 2012.
- [14] E. Benhamou, A. Rivoira, and A. Gruz. Stochastic Interest Rates for Local Volatility Hybrids Models. *Wilmott Magazine*, March 2008.

- [15] F. Black and P. Karasinski. Bond and Option Pricing when Short Rates are Lognormal. *Financial Analysts Journal*, 47(4):52–59, 1991.
- [16] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
- [17] D. A. Bloch and Y. Nakashima. Multi-Currency Local Volatility Model. *Available at SSRN 1153337*, 2008.
- [18] M. Broadie and Ö. Kaya. Exact Simulation of Stochastic Volatility and Other Affine Jump Diffusion Processes. *Operations Research*, 54(2):217–231, 2006.
- [19] R. H. Cameron and W. T. Martin. The Orthogonal Development of Non-Linear Functionals in Series of Fourier-Hermite Functionals. *Annals of Mathematics*, 48(2):385–392, 1947.
- [20] P. Carr and D. Madan. Option Valuation using the Fast Fourier Transform. *Journal of Computational Finance*, 2(4):61–73, 1999.
- [21] B. Chen, L. A. Grzelak, and C. W. Oosterlee. Calibration and Monte Carlo Pricing of the SABR-Hull-White model for Long-Maturity Equity Derivatives. *Journal of Computational Finance*, 15(4):79–113, 2012.
- [22] B. Chen, C. W. Oosterlee, and H. Van Der Weide. A Low-Bias Simulation Scheme for the SABR Stochastic Volatility Model. *International Journal of Theoretical and Applied Finance*, 15(2), 2012.
- [23] I. J. Clark. *Foreign Exchange Option Pricing: A Practitioners Guide*. John Wiley & Sons, 2011.
- [24] T. Coleman, Y. Li, and A. Verma. Reconstructing the Unknown Local Volatility Function. *Journal of Computational Finance*, 2:77–102, 1998.
- [25] J. Cox. Notes on Option Pricing I: Constant Elasticity of Diffusions. *Unpublished draft*, 1975.
- [26] J. Cox, J. Ingersoll, and S. Ross. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53:385–407, 1985.
- [27] A. Cozma, M. Mariapragassam, and C. Reisinger. Calibration of a Four-Factor Hybrid Local-Stochastic Volatility Model with a New Control Variate Particle Method. *Preprint*, 2017.
- [28] A. Cozma, M. Mariapragassam, and C. Reisinger. Convergence of an Euler Scheme for a Hybrid Stochastic-Local Volatility Model with Stochastic Rates in Foreign Exchange Markets. *SIAM Journal on Financial Mathematics*, 9(1):127–170, 2018.
- [29] M. Curran. Valuing Asian and Portfolio Options by Conditioning on the Geometric Mean Price. *Management Science*, 40(12):1705–1711, 1994.

- [30] S. De Marco, P. Friz, and S. Gerhold. Rational Shapes of Local Volatility. *Risk*, 2:82–87, 2013.
- [31] G. Deelstra, J. Liinev, and M. Vanmaele. Pricing of Arithmetic Basket Options by Conditioning. *Insurance: Mathematics and Economics*, 34(1):55–77, 2004.
- [32] G. Deelstra and G. Rayée. Local Volatility Pricing Models for Long-Dated FX Derivatives. *Applied Mathematical Finance*, pages 1–23, 2012.
- [33] E. Derman, D. Ergener, and I. Kani. Static Options Replication. *The Journal of Derivatives*, 2(4):78–95, 1995.
- [34] E. Derman and I. Kani. Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility. *International Journal of Theoretical and Applied Finance*, 1(1):61–110, 1998.
- [35] B. Dupire. Pricing With a Smile. *Risk Magazine*, 7(1):18–20, 1994.
- [36] B. Dupire. A Unified Theory of Volatility. *Derivatives Pricing: The Classic Collection*, pages 185–196, 1996.
- [37] B. Engelmann, F. Koster, and D. Oeltz. Calibration of the Heston Stochastic Local Volatility Model: A Finite Volume Scheme. *Available at SSRN 1823769*, 2011.
- [38] F. Fang and C. W. Oosterlee. A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions. *SIAM Journal on Scientific Computing*, 31(2):826–848, 2008.
- [39] J. Favard. Sur les polynomes de tchebicheff. *CR Acad. Sci. Paris*, 200:2052–2053, 1935.
- [40] J. Fernández, A. Ferreiro, J. García, A. Leitao, J. López-Salas, and C. Vázquez. Static and Dynamic SABR Stochastic Volatility Models: Calibration and Option Pricing using GPUs. *Mathematics and Computers in Simulation*, 94:55–75, 2013.
- [41] J.-P. Fouque, G. Papanicolaou, and K. Sircar. Stochastic Volatility Correction to Black-Scholes. *Risk*, 13(2):89–92, 2000.
- [42] R. Frey and D. Sommer. A Systematic Approach to Pricing and Hedging International Derivatives with Interest Rate Risk: Analysis of International Derivatives under Stochastic Interest Rates. *Applied Mathematical Finance*, 3(4):295–317, 1996.
- [43] C. P. Fries. Markov Functional Modeling of Equity, Commodity and other Assets. 2006.
- [44] F. N. Fritsch and R. E. Carlson. Monotone Piecewise Cubic Interpolation. *SIAM Journal on Numerical Analysis*, 17(2):238–246, 1980.
- [45] B. Ganapathysubramanian and N. Zabarar. Sparse Grid Collocation Schemes for Stochastic Natural Convection Problems. *Journal of Computational Physics*, 225(1):652–685, 2007.

- [46] J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. John Wiley & Sons, Ltd, 2006.
- [47] P. Gauthier and P.-Y. Rivaille. Fitting the Smile, Smart Parameters for SABR and Heston. *Available at SSRN 1496982*, 2009.
- [48] H. Geman, N. El Karoui, and J.-C. Rochet. Changes of Numéraire, Changes of Probability Measure and Option Pricing. *Journal of Applied Probability*, pages 443–458, 1995.
- [49] P. Glasserman. *Monte Carlo Methods in Financial Engineering*, volume 53. Springer Science & Business Media, 2003.
- [50] P. Glasserman and Q. Wu. Forward and Future Implied Volatility. *International Journal of Theoretical and Applied Finance*, 14(3):407–432, 2011.
- [51] E. Gobet. Advanced Monte Carlo methods for barrier and related exotic options. *Handbook of Numerical Analysis*, 15:497–528, 2009.
- [52] W. H. Greene. *Econometric Analysis*, 5th Edition. 2002.
- [53] J. Gregory. *The XVA Challenge: Counterparty Credit Risk, Funding, Collateral and Capital*. John Wiley & Sons, 2015.
- [54] L. A. Grzelak. The CLV Framework-A Fresh Look at Efficient Pricing with Smile. *International Journal of Computer Mathematics*, forthcoming, 2019.
- [55] L. A. Grzelak and C. W. Oosterlee. On the Heston Model with Stochastic Interest Rates. *SIAM Journal on Financial Mathematics*, 2(1):255–286, 2011.
- [56] L. A. Grzelak and C. W. Oosterlee. On Cross-Currency Models with Stochastic Volatility and Correlated Interest Rates. *Applied Mathematical Finance*, 19(1):1–35, 2012.
- [57] L. A. Grzelak and C. W. Oosterlee. From Arbitrage to Arbitrage-Free Implied Volatilities. *Journal of Computational Finance*, 20(3):31–49, 2016.
- [58] L. A. Grzelak, J. A. S. Witteveen, M. Suarez, and C. W. Oosterlee. The Stochastic Collocation Monte Carlo Sampler: Highly Efficient Sampling from “Expensive” Distributions. *Quantitative Finance*, doi: 10.1080/14697688.2018.1459807, 2018.
- [59] S. Guo, L. A. Grzelak, and C. W. Oosterlee. Analysis of an Affine Version of the Heston–Hull–White Option Pricing Partial Differential Equation. *Applied Numerical Mathematics*, 72:143–159, 2013.
- [60] J. Guyon and P. Henry-Labordère. The Smile Calibration Problem Solved. *Available at SSRN 1885032*, 2011.
- [61] J. Guyon and P. Henry-Labordère. Being Particular about Calibration. *Risk*, 25(1):88, 2012.

- [62] I. Gyöngy. Mimicking the One-Dimensional Marginal Distributions of Processes having an Itô Differential. *Probability Theory and Related Fields*, 71(4):501–516, 1986.
- [63] P. S. Hagan, D. Kumar, A. S. Lesniewski, and D. E. Woodward. Managing Smile Risk. *Wilmott Magazine*, pages 84–108, 2002.
- [64] P. Henry-Labordère. Calibration of Local Stochastic Volatility Models to Market Smiles: a Monte-Carlo Approach. *Risk*, pages 112–117, 2009.
- [65] A. Hernandez. Model Calibration with Neural Networks. *Risk*, June 2017.
- [66] S. L. Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6:327–343, 1993.
- [67] J. Hull and A. White. One-Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities. *Journal of Financial and Quantitative Analysis*, 28(02):235–254, 1993.
- [68] P. Hunt, J. Kennedy, and A. Pelsser. Markov-Functional Interest Rate Models. *Finance and Stochastics*, 4(4):391–408, 2000.
- [69] K. J. In 't Hout and C. Mishra. Stability of the Modified Craig–Sneyd scheme for Two-Dimensional Convection–Diffusion Equations with Mixed Derivative Term. *Mathematics and Computers in Simulation*, 81(11):2540–2548, 2011.
- [70] K. J. In 't Hout and B. Welfert. Unconditional Stability of Second-Order ADI Schemes Applied to Multi-Dimensional Diffusion Equations with Mixed Derivative Terms. *Applied Numerical Mathematics*, 59(3-4):677–692, 2009.
- [71] O. Islah. Solving SABR in Exact Form and Unifying it with LIBOR Market Model. Available at SSRN 1489428, 2009.
- [72] P. Jäckel. A Practical Method for the Valuation of a Variety of Hybrid Products. In *ICBI Global Derivatives Conference slides*, 2005.
- [73] S. Jain and C. W. Oosterlee. The Stochastic Grid Bundling Method: Efficient pricing of Bermudan options and their Greeks. *Applied Mathematics and Computation*, 269:412–431, 2015.
- [74] A. Janek, T. Kluge, R. Weron, and U. Wystup. FX smile in the Heston model. In *Statistical Tools for Finance and Insurance*, pages 133–162. Springer, 2011.
- [75] M. Jex, R. Henderson, and D. Wang. Pricing Exotics under the Smile. *Risk*, 12(11):72–75, 1999.
- [76] S. Johnson and H. Lee. Capturing the Smile. *Risk*, 16(3):89–95, 2003.
- [77] B. Jourdain and M. Sbai. Coupling Index and Stocks. *Quantitative Finance*, 12(5):805–818, 2012.

- [78] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1991.
- [79] K. Larsson. Dynamic Extensions and Probabilistic Expansions of the SABR Model. *Available at SSRN 1536471*, 2010.
- [80] A. Lipton. The Vol Smile Problem. *Risk*, 15(2):61–66, 2002.
- [81] A. Lipton, A. Gal, and A. Lasis. Pricing of Vanilla and First-Generation Exotic Options in the Local Stochastic Volatility Framework: Survey and new Results. *Quantitative Finance*, 14(11):1899–1922, 2014.
- [82] A. Lipton and W. McGhee. Universal Barriers. *Risk*, 15(5):81–85, 2002.
- [83] F. A. Longstaff and E. S. Schwartz. Valuing American Options by Simulation: a Simple Least-Squares Approach. *Review of Financial Studies*, 14(1):113–147, 2001.
- [84] R. Lord. Partially Exact and Bounded Approximations for Arithmetic Asian Options. *Journal of Computational Finance*, 10(2):1, 2006.
- [85] R. Lord, R. Koekkoek, and D. Van Dijk. A Comparison of Biased Simulation Schemes for Stochastic Volatility Models. *Quantitative Finance*, 10:177–194, 2010.
- [86] M. Lorig, S. Pagliarani, and A. Pascucci. Explicit Implied Volatilities for Multifactor Local-Stochastic Volatility Models. *Mathematical Finance*, 27(3):926–960, 2017.
- [87] K. V. Mardia. Applications of Some Measures of Multivariate Skewness and Kurtosis in Testing Normality and Robustness Studies. *Sankhyā: The Indian Journal of Statistics, Series B*, pages 115–128, 1974.
- [88] M. Musiela and M. Rutkowski. *Martingale Methods in Financial Modelling – Second Edition*. Springer, 1997.
- [89] J. Necas. *Direct Methods in the Theory of Elliptic Equations*. Springer, 2012.
- [90] F. Nobile, R. Tempone, and C. G. Webster. A Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.
- [91] J. Oblój. Fine-Tune Your Smile: Correction to Hagan et al. *Wilmott Magazine*, 2008.
- [92] Y. Osajima. The Asymptotic Expansion Formula of Implied Volatility for Dynamic SABR Model and FX Hybrid Model. *Available at SSRN 965265*, 2007.
- [93] A. Pascucci and A. Mazzon. The Forward Smile in Local-Stochastic Volatility Models. *Journal of Computational Finance*, 20(3):1–29, 2016.
- [94] L. Paulot. Asymptotic Implied Volatility at the Second Order with Application to the SABR Model. In *Large Deviations and Asymptotic Methods – Springer Proceedings in Mathematics & Statistics 110*, 2015.

- [95] A. Pelsser. *Efficient Methods for Valuing Interest Rate Derivatives*. Springer Science & Business Media, 2000.
- [96] V. Piterbarg. Time to Smile. *Risk*, 18(5):71–75, 2005.
- [97] V. Piterbarg. Cross-Currency Exotics: Smiling Hybrids. *Risk*, 19(5):66–71, 2006.
- [98] V. Piterbarg. Markovian Projection Method for Volatility Calibration. *Risk*, April 2007.
- [99] V. Piterbarg and B. Capital. Modern Approaches to Stochastic Volatility Calibration. In *Proceedings of the WBS 3rd Fixed Income Conference, Amsterdam*, volume 41, 2006.
- [100] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer, 2005.
- [101] R. Rebonato. *Volatility and Correlation in the Pricing of Equity, FX, and Interest-rate Options*. John Wiley & Sons, Ltd, 1999.
- [102] R. Rebonato. *Volatility and Correlation: The Perfect Hedger and the Fox*. John Wiley & Sons, 2005.
- [103] R. Rebonato, K. McKay, and R. White. *The SABR/LIBOR Market Model: Pricing, Calibration and Hedging for Complex Interest-Rate Derivatives*. Wiley.com, 2011.
- [104] Y. Ren, D. Madan, and M. Q. Qian. Calibrating and Pricing with Embedded Local Volatility Models. *Risk*, 20(9):138–143, 2007.
- [105] M. J. Ruijter and C. W. Oosterlee. Two-dimensional Fourier Cosine Series Expansion Method for Pricing Financial Options. *SIAM Journal on Scientific Computing*, 34(5):B642–B671, 2012.
- [106] M. Sankaran. Approximations to the Non-central Chi-square Distribution. *Biometrika*, 50:199–204, 1963.
- [107] S. Sankaran and A. L. Marsden. A Stochastic Collocation Method for Uncertainty Quantification and Propagation in Cardiovascular Simulations. *Journal of Biomechanical Engineering*, 133(3):031001, 2011.
- [108] R. Schöbel and J. Zhu. Stochastic Volatility with an Ornstein–Uhlenbeck Process: an Extension. *European Finance Review*, 3(1):23–46, 1999.
- [109] S. E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer, 2004.
- [110] S. Simaitis, C. de Graaf, N. Hari, and D. Kandhai. Smile and Default: the Role of Stochastic Volatility and Interest Rates in Counterparty Credit Risk. *Quantitative Finance*, 16(11):1725–1740, 2016.
- [111] J. Sippel and S. Ohkoshi. All Power to PRDC Notes. *Risk*, 15(11):1–3, 2002.

- [112] J. M. Steele. *Stochastic Calculus and Financial Applications*, volume 45. Springer Science & Business Media, 2001.
- [113] A. Takahashi, K. Takehara, and M. Toda. *Computation in an Asymptotic Expansion Method*. John and Wiley & Sons, Ltd, 2007.
- [114] H. Tanaka. Note on Continuous Additive Functionals of the 1-Dimensional Brownian Path. *Z. Wahrscheinlichkeitstheorie*, 1:251–257, 1963.
- [115] G. Tataru and T. Fisher. *Stochastic Local Volatility*. Quantitative Development Group, Bloomberg Version 1, 2010.
- [116] Y. Tian, Z. Zhu, F. Klebaner, and K. Hamza. A Hybrid Stochastic Volatility Model Incorporating Local Volatility. In *Fourth International Conference on Computational and Information Sciences*, pages 333 – 336, 2012.
- [117] Y. Tian, Z. Zhu, G. Lee, F. Klebaner, and K. Hamza. Calibrating and Pricing with a Stochastic-Local Volatility Model. *The Journal of Derivatives*, 22(3):21–39, 2015.
- [118] A. W. van der Stoep, L. A. Grzelak, and C. W. Oosterlee. The Heston Stochastic-Local Volatility Model: Efficient Monte Carlo Simulation. *International Journal of Theoretical and Applied Finance*, 17(7):1450045, 2014.
- [119] A. W. van der Stoep, L. A. Grzelak, and C. W. Oosterlee. The Time-Dependent FX-SABR Model: Efficient Calibration based on Effective Parameters. *International Journal of Theoretical and Applied Finance*, 18(6):1550042, 2015.
- [120] A. W. van der Stoep, L. A. Grzelak, and C. W. Oosterlee. A Novel Monte Carlo Approach to Hybrid Local Volatility Models. *Quantitative Finance*, 17(9):1347–1366, 2017.
- [121] A. W. van der Stoep, L. A. Grzelak, and C. W. Oosterlee. Collocating Local Volatility: a Competitive Alternative to Stochastic Local Volatility Models. Submitted for publication, 2018.
- [122] A. Van Haastrecht and A. Pelsser. Generic Pricing of FX, Inflation and Stock Options under Stochastic Interest Rates and Stochastic Volatility. *Quantitative Finance*, 11(5):665–691, 2011.
- [123] L. von Sydow, S. Milovanović, E. Larsson, K. J. In 't Hout, M. Wiktorsson, C. W. Oosterlee, V. Shcherbakov, M. Wyns, A. Leitao, S. Jain, T. Haentjens, and J. Waldén. BENCHOP-SLV: The BENCHmarking Project in Option Pricing – Stochastic and Local Volatility Problems. *International Journal of Computer Mathematics*, Forthcoming, 2019.
- [124] S. Watanabe. Analysis of Wiener Functionals (Malliavin Calculus) and its Applications to Heat Kernels. *The Annals of Probability*, 15:1–39, 1987.
- [125] T. Weathers. *Foreign Exchange: a Practical Guide to the FX Markets*, volume 309. John Wiley & Sons, 2006.

-
- [126] R. Weron and U. Wystup. Heston's model and the smile. In *Statistical tools for finance and insurance*, pages 161–181. Springer, 2005.
- [127] J. A. S. Witteveen and G. Iaccarino. Simplex Stochastic Collocation with Random Sampling and Extrapolation for Nonhypercube Probability Spaces. *SIAM Journal on Scientific Computing*, 34(2):A814–A838, 2012.
- [128] Q. Wu. Series Expansion of the SABR Joint Density. *Mathematical Finance*, 22(2):310–345, 2012.
- [129] M. Wyns and K. J. In 't Hout. An Adjoint Method for the Exact Calibration of Stochastic Local Volatility Models. *Journal of Computational Science*, 24:182–194, 2018.
- [130] D. Xiu. Efficient Collocational Approach for Parametric Uncertainty Analysis. *Communications in Computational Physics*, 2(2):293–309, 2007.
- [131] D. Xiu and J. S. Hesthaven. High-Order Collocation Methods for Differential Equations with Random Inputs. *SIAM Journal on Scientific Computing*, 27(3):1118–1139, 2005.
- [132] B. Zhang and C. W. Oosterlee. Efficient Pricing of European-Style Asian Options under Exponential Lévy Processes Based on Fourier Cosine Expansions. *SIAM Journal on Financial Mathematics*, 4(1):399–426, 2013.

CURRICULUM VITÆ

Anthonie Willem VAN DER STOEP

05-10-1988 Born in Barendrecht, the Netherlands

EDUCATION

2001–2007 Gymnasium
CSG Calvijn, Rotterdam, the Netherlands

2007–2010 Bachelor Technische Wiskunde (with honours)
University of Twente, Enschede, the Netherlands

2010–2012 Master Applied Mathematics
University of Twente, Enschede, the Netherlands

2012–2019 PhD Researcher
Rabobank, Utrecht, the Netherlands
Centrum Wiskunde & Informatica, Amsterdam, the Netherlands
Delft University of Technology, Delft, the Netherlands

2019 PhD Applied Mathematics
Delft University of Technology, Delft, the Netherlands
Thesis: Pricing and Calibration with Stochastic Local Volatility Models in a Monte Carlo Setting
Promotor: Prof. dr. ir. C.W. Oosterlee
Copromotor: Dr. ir. L.A. Grzelak

LIST OF PUBLICATIONS

4. A.W. van der Stoep, L.A. Grzelak and C.W. Oosterlee, Collocating Local Volatility: a Competitive Alternative to Stochastic Local Volatility Models. Submitted for publication, 2018.
3. A.W. van der Stoep, L.A. Grzelak and C.W. Oosterlee, A Novel Monte Carlo Approach to Hybrid Local Volatility Models, *Quantitative Finance*, 17(9):1347-1366, 2017.
2. A.W. van der Stoep, L.A. Grzelak and C.W. Oosterlee, The Time-Dependent FX-SABR Model: Efficient Calibration based on Effective Parameters, *International Journal of Theoretical and Applied Finance*, 18(6):1550042, 2015.
1. A.W. van der Stoep, L.A. Grzelak and C.W. Oosterlee, The Heston Stochastic-Local Volatility Model: Efficient Monte Carlo Simulation, *International Journal of Theoretical and Applied Finance*, 17(7):1450045, 2014.

LIST OF ATTENDED CONFERENCES WITH PRESENTATION

Presentations:

5. 8th General AMaMeF Conference, Amsterdam, the Netherlands, June 2017.
4. Financial Mathematics Winter School, Lunteren, the Netherlands, January 2017.
3. International Conference of Computational Finance, Greenwich, England, December 2015.
2. SIAM Conference on Financial Mathematics and Engineering, Chicago, USA, November 2014.
1. Multi-ITN STRIKE and WWCSC Mini-Workshop in Stochastic Computing and Optimization, Würzburg, Germany, October 2014.

Posters:

2. Models and Numerics in Financial Mathematics, Leiden, the Netherlands, May 2015.
1. 38th Woudschoten Conference, Woudschoten, the Netherlands, October 2013.