Calibration to Vanilla and Barrier Options with the Gyöngy and Brunick–Shreve Markovian Projections

Matthieu Mariapragassam

Trinity 2017



Green Templeton College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Contents

Abstract xiii

Acknowledgments xv

Statement of Originality xvii

1 Introduction 1

1.1 Background and literature . . . 2
1.2 A motivating example . . . 3
1.3 Outline and contribution . . . 6
1.4 Mimicking an Itô process . . . 9

2 A Forward Equation for Barrier Options under the Brunick–Shreve Markovian Projection 11

2.1 Introduction . . . 11
2.2 Setup and main result . . . 12
2.3 A Dupire-type formula for barrier options . . . 19
2.4 A forward partial-integro differential equation for barrier options . . . 20
2.4.1 Formulation as PIDE . . . 21
2.4.2 Finite difference approximation . . . 23
2.5 Numerical results . . . 27
2.5.1 Pricing under the mimicking Brunick–Shreve model . . . 27
2.5.2 Numerical validation . . . 30
2.6 Conclusions . . . 32

3 Calibration of a Four-Factor Hybrid LSV Model with a New Control Variate Particle Method 35

3.1 Introduction . . . 35
3.2 Model definition . . . 37
3.3 A necessary and sufficient condition for exact calibration . . . 39
3.4 Fast calibration with a new control variate particle method . . . 50
3.4.1 Market data set . . . 50
3.4.2 Variance reduction for standard expectations . . . 51

3.4.3 Variance reduction for the Markovian projection . . . 54
3.4.4 Shifted CIR model . . . 56
3.4.5 Four-factor hybrid stochastic volatility model calibration . . . 59
3.4.6 One-factor local volatility model calibration . . . 62
3.4.7 Two-factor Heston-type LSV model calibration by PDE . . . 64
3.5 Four-factor Heston-type LSV 2CIR++ model calibration . . . 75
3.5.1 Monte Carlo QE-scheme . . . 76
3.5.2 Calibration algorithm with the control variate particle method . . . 78
3.5.3 Calibration results . . . 82
3.5.4 Stress scenario . . . 85
3.6 Pricing of an AutoCallable Barrier Dual-Currency Note . . . 86
3.7 Conclusion . . . 90

4 Calibration of Local-Stochastic and Path-Dependent Volatility Models to Vanilla and No-Touch Options with Markovian Projections 93
4.1 Introduction . . . 93
4.2 Models and market . . . 95
4.2.1 Model definition . . . 95
4.2.2 Market data . . . 97
4.3 Condition for exact calibration to up-and-out call prices . . . 100
4.3.1 Mimicking result and forward equation . . . 100
4.3.2 Heston-type LSV and LMSV models . . . 109
4.3.3 Fokker–Planck PDE for the joint density of (S, M, V) . . . 115
4.4 Improved numerical solution of the forward PIDE . . . 116
4.4.1 Pricing vanilla options . . . 117
4.4.2 Space discretisation . . . 118
4.4.3 BDF2 scheme with variable step size . . . 121
4.4.4 Non-uniform mesh construction and numerical tests . . . 122
4.5 Calibration of the Brunick–Shreve model . . . 126
4.5.1 Gradient operator with respect to the volatility parameters . . . 128
4.5.2 Calibration with regularised best-fit algorithm . . . 129
4.5.3 The importance of regularisation . . . 131
4.5.4 The underlying local volatility . . . 132
4.6 Calibration of the LMSV model by 2D particle method . . . 134
4.7 Calibration of the Heston-type LSV-LVV model . . . 141
4.7.1 Model definition . . . 143
4.7.2 Calibration methodology . . . 143
4.7.3 Full Monte Carlo approach . . . 148

4.8 Calibration results and model comparison . . . 150
4.8.1 Smile dynamics . . . 153
4.9 Adding stochastic rates . . . 156
4.10 The Fokker–Planck equation for a two-dimensional diffusion with path-dependence . . . 160
4.10.1 Proof of Theorem 4.6 . . . 160
4.10.2 Finite element solution . . . 166
4.10.3 Mesh construction . . . 174
4.10.4 Solving the Fokker–Planck PDE for calibrated models . . . 178
4.11 Conclusion . . . 180

5 Conclusion **183**

Bibliography **185**

List of Figures

1.2.1 Volatility impact for up-and-out call options 6
1.3.1 Analogy of results for vanilla and barrier options 7

2.3.1 Analogy of results for European and barrier options. 19
2.4.1 PIDE Domain and Boundaries 22
2.5.1 Mesh construction for backward equation. The spot grid is recomputed for every PDE layer to be $(x_j(y_i), y_i)$, where $x_j = j\Delta_x$, $j < N_x - 2$, $x_{N_x-j} = y_{i-j}$, $j \in \{0, 1, 2\}$, $\Delta_x = y_{i-2}/(N_x - 2)$ and $N_x = \lceil y_i/\Delta_y \rceil$ to keep the spacing approximately constant. . 30
2.5.2 Assumed Brunick–Shreve volatility surface as per (2.5.5), where $a = 0.04, b = 0.2, \sigma = 0.2, \rho = m = 0$, and $T \in \{0.1, 0.5, 1\}$. 31
2.5.3 Forward PIDE vs Backward PDE:up-and-out call option prices with Brunick–Shreve volatility, $S = 100$ and $T = 1$. 32

3.2.1 Full calibration routine flowchart 39
3.4.1 Market volatility surface for the EUR-USD (18/03/2016) 51
3.4.2 The instantaneous forward rates and the continuously-compounded spot rates. 57
3.4.3 The market- and model-implied term structures of cap volatilities. 59
3.4.4 EURUSD Local volatility function calibrated by forward PDE and fixed-point algorithm 65
3.4.5 Market spot marginal density computed from the Dupire forward PDE with calibrated σ_{LV} 66
3.4.6 $\mathbb{P}_2$ element with 6 degrees of freedom 71
3.4.7 Finite element triangular mesh refined around $x = S_0$ and $z = 0$ 71
3.4.8 $p = v^{-\beta}\phi$ using ϕ computed with the standard PDE 72
3.4.9 $p = v^{-\beta}\phi$ computed directly from the change of variables PDE 73
3.4.10 Calibrated leverage function α for the 2-factor LSV model (3.2.2) 75
3.5.1 Error convergence with and without control variates. Error is computed on all quoted deltas and maturities (log-scale) 83
3.5.2 Calibration fit with 4 000 particles for the 4-factor LSV model with and without control variates 84
3.5.3 Variance reduction factor as a function of time for the conditional expectation estimator p_N^* (left, using a log-log scale) and the standard expectations estimators $\gamma_{1,N}^*$ and $\gamma_{2,N}^*$ (right, using a log scale) 85

3.5.4 Calibration fit with 500 000 particles for the 4-factor LSV model vs 4-factor SV model 86
3.5.5 Calibrated leverage function α for the 4-factor LSV model (3.2.1) 88

4.2.1 USD Zero-coupon bond prices. 97
4.2.2 EUR Zero-coupon bond prices. 97
4.2.3 Market volatility surface for EUR-USD (28/03/2013). 98
4.2.4 Foreign no-touch up market prices 99
4.3.1 Steady state of ψ for the uncertain volatility model (4.3.13) in accordance with the LSV-LVV calibrated parameters from Table 4.3.1. 112
4.3.2 Explosion of ψ for the uncertain volatility model with $\bar{\xi} \in [0.9, 1.5]$. 112
4.4.1 Finite difference hyperbolic grid for the forward PIDE (4.4.1) initialised with "blank layers". 123
4.4.2 $\phi(B, B, T)$ as a function of the barrier level for $T = 1$ initialised with "blank layers". 700 strike steps, 100 BDF2 time steps + Rannacher initialisation. 125
4.4.3 $\phi(B, B, T)$ as a function of the barrier level for $T = 1$ initialised without "blank layers". 700 strike steps, 100 BDF2 time steps + Rannacher initialisation. 125
4.4.4 $\phi(B, B, T)$ as a function of the barrier level for $T = 1$ initialised with "blank layers". 700 strike steps, 10 BDF2 time steps + Rannacher initialisation. 125
4.4.5 $\phi(B, B, T)$ as a function of the barrier level for $T = 1$ initialised with "blank layers". 700 strike steps, 10 Crank–Nicolson time steps + Rannacher initialisation. 125
4.4.6 Brunick–Shreve volatility used for numerical test with smooth transition at the boundary of the domain. The construction is such that $\sigma_{BSV}(s, m, t) = \sqrt{\sigma_{LV}(s, t)\sigma_{LV}(m, t)}$. 127
4.4.7 Up-and-out call prices computed with the forward PIDE for different values of strikes and barrier levels and $T = 1$. 127
4.5.1 Brunick–Shreve volatility function $T = 1Y$ with regularisation. 132
4.5.2 Brunick–Shreve volatility function $T = 5Y$ with regularisation. 133
4.5.3 Brunick–Shreve volatility function $T = 1Y$ with no regularisation. 133
4.5.4 Brunick–Shreve volatility function $T = 5Y$ with no regularisation. 134
4.5.5 EURUSD local volatility calibrated by forward Dupire PDE. 135
4.5.6 EURUSD local volatility projected from the calibrated Brunick–Shreve volatility. . . 135
4.6.1 Joint density of the spot and running maximum after calibration at time $T = 1$. . . 142
4.6.2 Calibrated leverage function for the LMSV model at time $T = 1$. 142
4.7.1 Calibrated leverage function for the LSV-LVV model. 147
4.7.2 Calibrated local vol-of-vol $\xi(S, t)$ for the LSV-LVV model. 148
4.7.3 Full Monte Carlo vs forward PIDE calibration comparison for a short- and long-term maturity pillar, N_T is the number of time steps per year and N the number of particles. 150
4.7.4 Local vol-of-vol function calibrated by forward PIDE 200 strike points, 100 000 particles and 100 time steps per year. 151

4.7.5 Local vol-of-vol function calibrated by full Monte Carlo, 500 000 particles and 100 time steps per year. 151
4.8.1 Calibration fit to foreign no-touch options for all models; see Table 4.8.2 for tabulated errors. 154
4.8.2 Implied volatility "model to sticky-strike spread" in % for an upward spot shift of 1.26% for all calibrated models with $T = 2Y$ 155
4.10.1 Possible triangulation of Ω_h. 171
4.10.2 Mesh refinement near $\mathbf{p}_0 = (S_0, S_0, v_0)$. 175
4.10.3 FEM mesh for the joint density of (S_t, M_t) under model (4.2.5). 177
4.10.4 FEM mesh for the joint density of (S_t, M_t, V_t) under model (4.3.25). 178
4.10.5 Forward PIDE vs quadrature price for up-and-out call options with $B = 110\% \times S_0$ on a ladder of strikes. Average relative difference over all plotted points is 0.042%. . . 179
4.10.6 FEM solution of the PDE (4.3.26) with no variance term for $t = 1$. 179
4.10.7 Monte Carlo vs quadrature price for up-and-out call options with $B = 120\% \times S_0$ on a ladder of strikes. Average relative difference over all plotted points is 0.079%. . . . 180
4.10.8 FEM solution of the PDE (4.3.26) for $t = 1$ and $z = v_0$. 180

List of Tables

2.5.1 Difference between forward and backward solutions over strikes between 0 and 120, and up-and-out barriers between 100 and 120, all for maturity 1. 31
2.5.2 Results and relative errors for different strikes; barrier = 120, forward = 105.13. . . . 32

3.4.1 The calibrated CIR parameters 60
3.4.2 The calibrated Heston parameters 61
3.5.1 Calibration error in absolute volatility (% unit) for the 4-factor LSV model vs 4-factor Heston–2CIR++ model 87
3.5.2 Average error in absolute volatility (% unit) for a high-volatility stress scenario on the rate processes. N is the number of particles. 87
3.6.1 Contract parameters for an AutoCallable Barrier Dual-Currency Note 89
3.6.2 Monte Carlo NPVs, in % of N, for the 5Y AutoCallable Barrier Dual-Currency Note . 90
3.6.3 The Monte Carlo estimates of the NPV of the contract specified in Table 3.6.1 for different numbers of time steps (per year) and 5×10^7 sample paths (for a standard deviation of 9.29×10^{-4}), the difference from the previous NPV estimate and the empirical order of convergence. 91

4.3.1 Uncertain volatility bounds and Heston parameters. 111
4.4.1 Price of an up-and-out call option for $K = 80\% \times S_0$, $B = 110\% \times S_0$ and $T = 1$. . . 126
4.4.2 Convergence order in number of strike steps N_K for 60 time steps 126
4.4.3 Convergence order in number of time steps N_T for 1200 strike steps 126
4.6.1 The calibrated Heston parameters. 136
4.7.1 Full Monte Carlo vs PIDE average one-touch price relative error comparison. "Timing factor" is the time needed for the calibration, relative to the PIDE approach computational time. N_T is the number of time steps per year, N the number of particles and N_K the number of strike points in the PIDE. 149
4.8.1 Average absolute error in % for the implied volatility for all models. 152
4.8.2 Average absolute error in % for foreign no-touch quotes for all models; see Fig. 4.8.1 for error plots. 153

Abstract

In this thesis, we present a novel approach to the calibration of diffusion models to vanilla and barrier options with the Gyöngy [84] and Brunick–Shreve [25] Markovian projection results. Firstly, we derive a forward equation for arbitrage-free barrier option prices in continuous semi-martingale models, in terms of Markovian projections of the stochastic volatility process. This leads to a Dupire-type formula for the coefficient derived by Brunick and Shreve for their mimicking diffusion and can be interpreted as the canonical extension of local volatility for barrier options. Secondly, we treat the problem of long-dated foreign-exchange option pricing and propose a novel and generic calibration technique to vanilla options for four-factor foreign-exchange hybrid local-stochastic volatility models with stochastic short rates. We build upon the particle method introduced by Guyon and Labordère [82] and combine it with new variance reduction techniques in order to accelerate convergence. Finally, we derive the necessary and sufficient condition for the exact calibration to up-and-out call options and provide a step-by-step procedure to calibrate a Brunick–Shreve volatility model, a Heston-type local-stochastic volatility model with local vol-of-vol and a path-dependent local-maximum-stochastic volatility model. While the Brunick–Shreve model is calibrated with our forward PIDE for barriers, both stochastic volatility models require an interesting two-dimensional extension of the particle method. We then derive and prove the self-consistency condition for perfect calibration to barrier options for path-dependent models with stochastic domestic and foreign short rates, where techniques from our previous work can be combined, in order to price accurately long-dated derivatives with barrier feature.

Acknowledgments

First and foremost, I would like to warmly thank my main supervisor Prof. Christoph Reisinger who made this work possible, our collaboration fruitful and our discussions interesting and stimulating. His guidance, enthusiasm, friendliness and above all, mathematical expertise have been of invaluable help at every step of the DPhil. It has been an honour and great pleasure to have worked on such an enriching project with him.

I would like to express my gratitude to my co-supervisor Prof. Marek Musiela for the extensive conversations about stochastic modelling and for sharing his immense knowledge of quantitative finance and derivatives pricing.

I would like to express my appreciation to my co-supervisor Prof. Ben Hambly for his precious advice on stochastic analysis and collaboration on Chapter 2.

I would like to thank Dr Alan Bain and Dr Simon McNamara, our industry partners, for their insightful comments, discussions about the foreign-exchange market and above all, for initiating the project at BNP Paribas London.

A special thanks to Prof. Vathana Ly Vath for introducing me to quantitative finance, teaching me stochastic calculus and for all his help during the application process.

Some ideas of this thesis have been inspired by my work experience at Murex, and numerous exciting conversations with colleagues. I would like to particularly thank Dr Laurent Tur, who taught me countless calibration techniques, presented me pricing difficulties of barrier options and was always willing to listen to my ideas.

On the personal level, I could not thank enough my partner Emily, she provided me with undivided support, always tried to understand my research, and made spare time pleasant and fun. She brought a much appreciated stability throughout the DPhil, that helped me handle setbacks and difficulties positively. I am infinitely grateful to my mother Claudine, father Gilbert, brother Luca, uncle Toni, grand-parents Maria and Domenico and the Mariapragassam family who never doubted my choices and always helped me achieve my unconventional projects and goals. My family believed in the DPhil experience from start to finish, and always made sure, that my studies were enjoyable. A special thought goes to my best friend and half-brother Jonathan, who always supported my decisions, despite his very sensitive personality which made it hard for him to see me leave.

It has been a pleasure to get to know Andrei and Kees, now two good friends with whom I have got the chance to share the enriching DPhil experience. Andrei and I had the opportunity to collaborate on two projects including part of Chapter 3. This created a stimulating work environment and very enjoyable coffee breaks with the group.

This thesis has been written at the University of Oxford, and the research project has been funded by the Oxford-Man Institute of Quantitative Finance and BNP Paribas London. This offered me a wonderful environment and opportunity to work on the ideas that Christoph and I decided to develop. I had the invaluable chance to meet extremely talented, smart and interesting people. I appreciated every single discussion with colleagues in the Mathematical Institute and the Oxford-Man Institute of Quantitative Finance.

Statement of Originality

I hereby declare that this thesis contains no material which has been accepted or is currently being submitted for any other degree, diploma, certificate or other qualifications at the University of Oxford or elsewhere. To the best of my knowledge, this thesis contains no material previously published and precise reference is made when a previously published result is used or discussed.

This thesis includes one paper published in the peer-reviewed Quantitative Finance journal [86], which is co-authored with my supervisor Prof. Christoph Reisinger and Prof. Ben Hambly.

The remainder of the thesis contains two additional unpublished papers, one of which is co-authored with my supervisor Prof. Christoph Reisinger and colleague Andrei Cozma [41], the other is co-authored with Prof. Christoph Reisinger and is based on Chapter 4.

The ideas, development and writing up of the thesis, and the three papers included in it, were the principal responsibility of myself, Matthieu Mariapragassam, working under the supervision of Prof. Christoph Reisinger.

All contributions that are not entirely my work are Subsections 3.4.5 and 3.4.4 and were written by the co-author Andrei Cozma for the article [41].

1 Introduction

In 2016, the number of derivative contracts traded on exchanges was 25.22 billion, setting a new record for the derivatives industry. It is a strong indicator that derivative products are re-gaining popularity both on listed and Over-The-Counter (OTC) markets. The risk-management of derivatives requires a model which closely agrees with market quotes for simple, liquid contracts like vanilla options. While the prices themselves of such liquid contracts are given by the market within a small bid-ask spread and therefore the model is not required to predict them, the hedge parameters used to risk manage even these simple products cannot be observed from the market and have to be estimated from a model. It is then of crucial importance that the model has a similar dynamic behaviour as the market.

Moreover, a model is required for the pricing and hedging of more exotic OTC contracts for which no agreed market price is available. In this manner, a model acts as a powerful interpolation-extrapolation tool for different classes of contracts.

The process that ensures a model matches the current market quotes is called calibration. Therefore, a model is said to be calibrated on a financial product when it re-prices the market quotes of the considered derivative. To be able to trust the calibrated model for the pricing of more exotic OTC contracts, it is important that the products used for the calibration have similar features as the products to be priced. For instance, to accurately price a product which is exposed to tail risk, it is crucial that the calibration instruments are similarly sensitive to large price moves.

Nowadays, the calibration process often solely relies on vanilla option quotes, and a widely used measure is the difference between the model implied volatility surface and the market volatility surface. Therefore, a model is said to be calibrated on a financial product when it re-prices, at best, the market quotes of the considered derivative. The case of calibration to vanilla options is widely documented and has been considered extensively in the literature since the work of Dupire [54] and Gyöngy [84] in the context of local volatility models. Nowadays, stochastic volatility models are the industry standard because of their superior dynamics and flexibility over more simplistic, one-dimensional alternatives. The additional parameters they embed allow the practitioner to account for more information, implied from prices of quoted derivatives, and incorporated in the model through a calibration procedure. While the valuation of a contract can be made generic with a Monte Carlo engine coupled with an Euler scheme, the calibration is often highly model dependent. This makes the implementation and validation of new models costly and time-consuming for financial institutions.

The focus of this thesis is the numerically efficient calibration of a general class of volatility models to vanilla and options with barrier features.

1.1 Background and literature

In Jex, Hendersen and Wang [98], one of the first Local-Stochastic Volatility (LSV) models was introduced, combining the dynamic properties of the stochastic volatility with the accuracy of the local volatility to calibrate vanilla options. In Ren, Madan and Qian [129] a fast calibration method is proposed, based on the Fokker–Planck PDE of the two-dimensional LSV diffusion model, to calibrate almost exactly the market volatility surface. It combines a smart utilisation of the Dupire formula and the calibration of the leverage function, step-wise forward in time, of the model. The underlying idea relies on the one-dimensional mimicking process introduced in [84, 54] and the computation of the Markovian projection, i.e. the expectation of the stochastic variance conditional on the value of the spot. In Guyon and Labordère [82], the authors generalise the approach to higher-dimensional stochastic volatility models and discuss the Monte Carlo-based particle method to estimate the Markovian projection. The main contribution of Ren et al. [129] and Guyon and Labordère [82] is the generality of the method that can be applied to a large class of continuous stochastic volatility model with leverage function.

While the accurate repricing of vanilla options is a must-have for all models, there is an open question of which additional derivative contract to include in the calibration routine. For instance, vanilla options only encode information about the marginal density of the spot, such that it becomes harder to understand the behaviour of the model for path-dependent options such as Asian, barrier or forward-start contracts. More particularly, an LV model and an LSV model, both calibrated to vanillas, can value a similar contract very differently, and a 5% absolute difference for a no-touch option price is very common. This brings into question which model to trust and under what setting, when the richer model sometimes does not provide a better price if not calibrated adequately. It then becomes critical to extend the scope of calibration instruments and include dynamics-dependent products, such as forward-start options, a popular addition in equity markets for example [14]. Foreign-exchange markets (FX) quote touch contracts such as no-touch options, a product that pays one if the underlying rate does not cross an up-and-out barrier. Barrier products of these types and, more generally, up-and-out call options, are extensively traded over-the-counter for most asset classes. Since the price discrepancy for barrier options between two models can be significant, the impact is naturally present for all popular exotic contracts with a knock-in or knock-out feature, such as the AutoCallable note and the Accumulator [35, 153].

A few published works address the question of touch options calibration from different angles and under different assumptions. For example, in Crosby and Carr [31] a particular class of models gives a calibration to both vanilla and barrier options. Model-independent bounds on the price of double

no-touch options were inferred from vanilla and digital option quotes in Cox and Obloj [38]. Pironneau [121] proves that the Dupire equation is still valid for a given barrier level, in a local volatility setting. While powerful, these approaches do not allow for a model agnostic calibration technique, similar to what [129, 82] introduced for vanilla options. This is the focus of this thesis.

1.2 A motivating example

Consider an LV pricing model driven by the following stochastic differential equation under the risk-neutral measure $\mathbb{Q}$

$$dS_t = (r(t) - q(t))S_t\,dt + \sigma(S_t, t)S_t\,dB_t \tag{1.2.1}$$

with σ a bounded local volatility and where r and q are deterministic risk-free rate and dividend yield, respectively and where B is a standard Brownian motion. Assume that the real market model, under a physical measure $\mathbb{P}$ is

$$d\bar{S}_t = \mu(t)\bar{S}_t\,dt + \bar{\sigma}_t\bar{S}_t\,dW_t^{\mathbb{P}}$$

with $\bar{\sigma}_t$ an arbitrary Itô process and a drift μ. In the remainder of the section we work under $\mathbb{Q}$, an equivalent martingale measure of $\mathbb{P}$, such that the discounted spot $\bar{S}$ capitalised with dividends is a martingale and therefore

$$d\bar{S}_t = (r(t) - q(t))\bar{S}_t\,dt + \bar{\sigma}_t\bar{S}_t\,dW_t\,,$$

where W is a standard Brownian motion. Note that a hedge that works almost surely under $\mathbb{Q}$ also works almost surely under $\mathbb{P}$ since the two measures agree on which events have probability one (see [135]). Therefore the analysis can be conducted under $\mathbb{Q}$ without loss of generality. The following derivation is similar to [101] and tailors it to barrier contracts. Denote by M_t and $\bar{M}_t$ the running maximum of the processes S_t and $\bar{S}_t$, respectively. Let a portfolio of wealth X_t be composed of a short position in an up-and-out call option C of maturity T, strike K and barrier B and delta-hedged with a long position of $\Delta = \frac{\partial C}{\partial x}$ shares with

$$C(x, y, t) = \mathbb{E}^{\mathbb{Q}}\Big[e^{-\int_t^T r(u)\,du}(S_T - K)^+\mathbf{1}_{M_t<B}\,|\,(S_t, M_t) = (x, y)\Big]\,.$$

Therefore, the portfolio evolution is

$$\begin{aligned} dX_t &= -dC\Big(\bar{S}_t, \bar{M}_t, t\Big) + \Delta\Big(\bar{S}_t, \bar{M}_t, t\Big)\,d\bar{S}_t \\ &+ r(t)\Big(X_t - \Delta\Big(\bar{S}_t, \bar{M}_t, t\Big)\bar{S}_t + C\Big(\bar{S}_t, \bar{M}_t, t\Big)\Big)\,dt \\ &+ q(t)\Delta\Big(\bar{S}_t, \bar{M}_t, t\Big)\bar{S}_t\,dt\,. \end{aligned} \tag{1.2.2}$$

And additionally, by Itô-Doeblin's lemma,

$$dC\Big(\bar{S}_t, \bar{M}_t, t\Big) = \frac{\partial C}{\partial t}\Big(\bar{S}_t, \bar{M}_t, t\Big)\, dt + \Delta\Big(\bar{S}_t, \bar{M}_t, t\Big)\, d\bar{S}_t + \frac{1}{2}\frac{\partial^2 C\Big(\bar{S}_t, \bar{M}_t, t\Big)}{\partial x^2}\, d\Big\langle \bar{S}_t \Big\rangle\,, \tag{1.2.3}$$

since $\frac{\partial C}{\partial y} d\bar{M}_t = 0$, as we discuss in Chapter 2, and $\bar{M}$ is of finite-variation (see [135]). Furthermore, in Chapter 2 we show that the backward local volatility pricing PDE still holds with the addition of a Neumann boundary condition, such that we can set $x = \bar{S}_t$ and $y = \bar{M}_t$, and write

$$\begin{aligned} \frac{\partial C}{\partial t}\Big(\bar{S}_t, \bar{M}_t, t\Big) &= -(r(t) - q(t))\bar{S}_t \frac{\partial C}{\partial x}\Big(\bar{S}_t, \bar{M}_t, t\Big) + r(t) C\Big(\bar{S}_t, \bar{M}_t, t\Big) \\ &- \frac{1}{2}\sigma^2\Big(\bar{S}_t, t\Big)\bar{S}_t^2 \frac{\partial^2 C}{\partial x^2}\Big(\bar{S}_t, \bar{M}_t, t\Big)\,. \end{aligned} \tag{1.2.4}$$

By combining (1.2.2), (1.2.3) and (1.2.4) and integrating we get for the discounted portfolio wealth process

$$e^{-\int_0^t r(u)\,du} X_t = -\frac{1}{2}\int_0^{\tau \wedge t} e^{-\int_0^u r(s)\,ds} \bar{S}_u^2 \frac{\partial^2 C\Big(\bar{S}_u, \bar{M}_u, u\Big)}{\partial x^2}\Big(\bar{\sigma}_u^2 - \sigma^2\Big(\bar{S}_u, u\Big)\Big)\, du, \tag{1.2.5}$$

where τ is the knock-out time, where no re-calibration of the local volatility happens in between $[0, t]$ and with zero initial wealth. Assume σ is a local volatility function perfectly calibrated to call options. We want to define $\bar{\sigma}$ such that we can arbitrarily control the sign of $\bar{\sigma}_u^2 - \sigma^2\Big(\bar{S}_u, u\Big)$ when the deal is live, i.e. when $\bar{M}_t \leq B$, and correct the calibration to vanillas when the deal has knocked-out, i.e. when $\bar{M}_t > B$. Therefore, we write $\bar{\sigma}$ as a function of both $\bar{S}$ and $\bar{M}$ as

$$\bar{\sigma}\Big(\bar{S}_t, \bar{M}_t, t\Big) = \sigma\Big(\bar{S}_t, t\Big)\sqrt{\mathbf{1}_{\bar{M}_t \leq B}\left(1 + \epsilon\,\mathrm{sgn}\left(\frac{\partial^2 C\Big(\bar{S}_t, \bar{M}_t, t\Big)}{\partial x^2}\right)\right) + \mathbf{1}_{\bar{M}_t > B}\beta\Big(\bar{S}_t, t\Big)}\,,$$

with $\mathrm{sgn}(x) = 1$ when $x > 0$, $\mathrm{sgn}(x) = -1$ when $x < 0$ and $\mathrm{sgn}(0) = 0$ and where the calibration condition to vanilla options, as explained in Section 1.4 is

$$\mathbb{E}\Big[\bar{\sigma}^2\Big(\bar{S}_t, \bar{M}_t, t\Big) \,|\, \bar{S}_t = x\Big] = \sigma^2(x, t)\,,$$

which gives for any $\epsilon > 0$,

$$\beta(x, t) = \frac{1 - \mathbb{E}\Big[\mathbf{1}_{\bar{M}_t \leq B}\Big(1 + \epsilon\,\mathrm{sgn}\Big(\frac{\partial^2 C\left(x, \bar{M}_t, t\right)}{\partial x^2}\Big)\Big) \,|\, \bar{S}_t = x\Big]}{\mathbb{E}\Big[\mathbf{1}_{\bar{M}_t > B} \,|\, \bar{S}_t = x\Big]}\,,$$

and ensures that $\bar{\sigma}$ is perfectly calibrated to vanilla options, similarly to σ. The up-and-out call option C, and therefore the portfolio wealth X, are not sensitive to scenarios where $\bar{M}_t > B$ since the deal

would be knocked out. Therefore,

$$\frac{\partial^2 C\left(\bar{S}_u, \bar{M}_u, u\right)}{\partial x^2}\left(\bar{\sigma}_u^2 - \sigma^2\left(\bar{S}_u, u\right)\right) = \frac{\partial^2 C\left(\bar{S}_u, \bar{M}_u, u\right)}{\partial x^2}\sigma^2\left(\bar{S}_u, u\right)\left(\epsilon \operatorname{sgn}\left(\frac{\partial^2 C\left(\bar{S}_u, \bar{M}_u, u\right)}{\partial x^2}\right)\right), \tag{1.2.6}$$

which is positive for any scenario. While the definition of $\bar{\sigma}$ depends on K and B through

$$\frac{\partial^2 C\left(\bar{S}_t, \bar{M}_t, t\right)}{\partial x^2},$$

the sign of this quantity is mainly driven by the position of the initial spot and running maximum. This toy example then becomes applicable to a large set of different up-and-out call contracts. To understand what happens graphically, let $S_0 = 110$ and consider an up-and-out call option of strike 100, barrier 120 and maturity T. In Figure 1.2.1, we graph the payout function at maturity without the dependence on the running maximum for better intuition. We note that the gamma of the barrier option, i.e. $\Gamma = \frac{\partial^2 C}{\partial x^2}$ is positive around the strike, where we also have $\bar{\sigma} \geq \sigma$, and negative near the barrier, where we also have $\bar{\sigma} \leq \sigma$. Therefore, from (1.2.5) and (1.2.6), the portfolio wealth is negative in both scenarios, which gives the practitioner entering a long position in an up-and-out call, a strong advantage if trading with a counterparty pricing barrier options with a pure local volatility model. Although this toy example does not reflect reality, nor does it propose a way to guarantee arbitrage, it shows that the real market diffusion can give rise to equation (1.2.6). This emphasises the importance of using the appropriate volatility, calibrated to a relevant set of market data including barrier contracts. The discrepancy is even more pronounced if σ is assumed constant and S follows a Black–Scholes model.

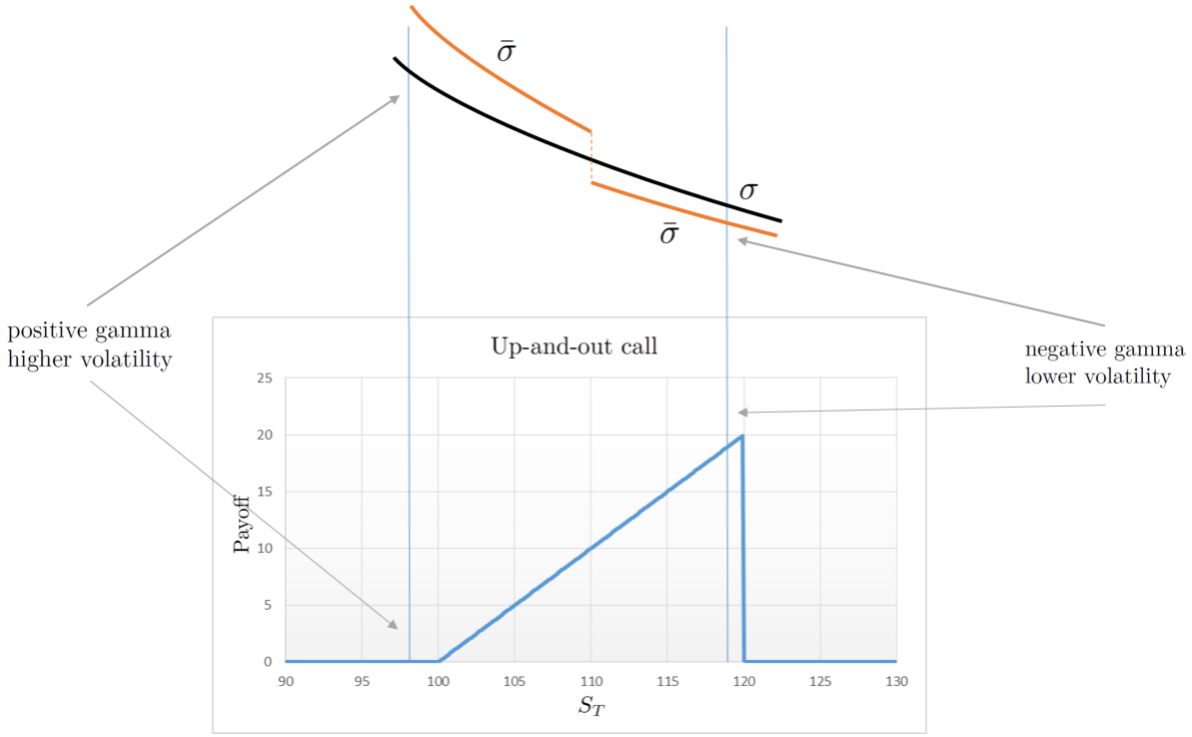


Figure 1.2.1: Volatility impact for up-and-out call options

1.3 Outline and contribution

In Brunick and Shreve [25], the authors extend the mimicking result from [84] (see Section 1.4) to path functionals of the underlying process, of which the running maximum is a special case. In particular, they prove, for any stochastic volatility process, the existence of a mimicking low-dimensional Markovian process with the same joint density for (S, M). The result suggests a natural extension of the local volatility model by considering a spot and running maximum dependent volatility function, a model we will refer to as the Brunick–Shreve model in the remainder of the thesis. By construction, using the Markovian projection of a stochastic volatility model onto S and M, i.e. by defining the squared volatility as the expectation of the variance process conditional on the value of the spot and the running maximum, one obtains the same barrier option prices for a multi-factor model and a mimicking one-dimensional Brunick–Shreve model. This provides us with a first stepping-stone to extend the generic vanilla calibration framework of [54, 129, 82] to barrier options. In a vanilla setting, one can rely on the Dupire formula to encode vanilla options prices, however, no such formula exists for barrier options, a problem that can be summarised with the chart depicted in Figure 1.3.1 for a stochastic volatility process α. Therefore, in Chapter 2, we derive a forward partial-integro differential equation (PIDE) as well as a Dupire-type formula for barrier options in order to complete

the Markovian projection framework for barrier options.

$$
\begin{array}{ccccc}
 & & \text{Gyöngy [84]} & & \text{Dupire [54]} \\
 & & \downarrow & & \downarrow \\
\sigma^2_{|S}(K,T) & = & \mathbb{E}^{\mathbb{Q}}[\alpha_T^2 \mid S_T = K] & = & \frac{\frac{\partial C}{\partial T} + (r-q)K\frac{\partial C}{\partial K} + qC}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}} \\
\sigma^2_{|S,M}(K,B,T) & = & \mathbb{E}^{\mathbb{Q}}[\alpha_T^2 \mid S_T = K\,, M_T = B] & = & ? \\
 & & \uparrow & & \uparrow \\
 & & \text{Brunick \& Shreve [25]} & & \text{Chapter 2}
\end{array}
$$

Figure 1.3.1: Analogy of results for vanilla and barrier options

The PIDE is discretised with a finite-difference scheme where we present validation results. As part of the main result, we also derive the Fokker–Planck equation of the spot and running maximum under a Brunick–Shreve model, which exhibits a peculiar boundary equation along the diagonal $S = M$. These results have the useful property of being forward in time, which allows us to price barrier options for a set of N_K strikes, N_B barrier levels and N_T maturities in one single resolution, a task that would require $N_K \times N_B \times N_T$ PDE resolutions with the backward approach. On the one hand, the speed-up is significant when the calibration relies on the minimisation of the error between model and market prices. On the other hand, the parameters can be assumed flat in time between two maturities, such that a bootstrap-type calibration is possible and efficient.

Alongside exotic contracts with barrier features, industry practitioners are also interested in long-dated contracts such as the Power Reverse Dual Currency Note (PRDC), or long-term AutoCallable options, which are exposed to the stochasticity of the underlying short rates. Assuming constant interest rates is appealing due to its simplicity and does not lead to serious mispricing of options with short maturities. However, empirical results [145] have confirmed that the constant interest rate assumption is inappropriate for long-dated FX products, and the effect of interest rate volatility can be as relevant as that of the FX rate volatility for longer maturities. Extensive research has been carried out in the area of option pricing with stochastic volatility and interest rates in the past few years. Van Haastrecht et al. [145] extended the model of Schöbel and Zhu [134] to currency derivatives by including stochastic interest rates, a model that benefits from analytical tractability even in a full correlation setting due to the processes being Gaussian. On the other hand, Ahlip and Rutkowski [4], Grzelak and Oosterlee [77] and Van Haastrecht and Pelsser [146] examined the Heston–CIR/Vasicek hybrid models and concluded that they give rise to non-affine models even under a partial correlation structure of the driving Brownian motions. In the foreign exchange market, the inclusion of stochastic interest rates results in a 4-factor model associated with a complex calibration routine. A few papers discuss this problem under simpler settings. Deelstra [45] and Clark [35] mainly consider 3-factor hybrid local volatility models and focus on the theoretical rather than the practical aspects of the calibration, whereas Stoep, Grzelak and Oosterlee [143] consider an application to a 2-factor hybrid

local volatility. In [82], Guyon and Labordère discuss an application of Monte Carlo-based calibration methods to a 3-factor LSV equity model with stochastic domestic rate and discrete dividends.

Therefore, in Chapter 3, we consider a 4-factor hybrid LSV model, with a stochastic variance and stochastic domestic and foreign short rate. We extend the calibration condition for the leverage function given in [82] for a 3-factor LV–2Hull–White model and in [45] for a 4-factor Heston-type LSV–2Hull–White model, and give a novel and rigorous derivation. We propose a new calibration approach that builds on the particle method of [82], and combine it with a new and efficient variance reduction technique that takes advantage of PDE calibration to increase its stability and accuracy while keeping the dimensionality under control with Monte Carlo. The proposed variance reduction techniques allow us to evaluate each expectation with a small amount of paths and yield a 125 speed-up factor compared to the plain particle method. Furthermore, numerical experiments suggest that this method almost recovers the calibration speed from the corresponding 2-factor LSV model with deterministic rates.

In Chapter 4, we initially revert to deterministic rates and use the forward PIDE for barriers to calibrate a Brunick–Shreve model by penalised optimisation, and discuss details about the finite-difference scheme. We then extend the Brunick–Shreve model and include a stochastic variance to introduce a Local-Maximum Stochastic Volatility model (LMSV), the natural generalisation of LSV models to fit barrier options. The calibration is performed by extending the particle method to account for a two-dimensional conditional state space (S, M). However, if one prefers to not include the running maximum in the parametrisation, we present a calibration approach for a Heston-type LSV model with a spot dependent vol-of-vol (LSV-LVV). It can easily be generalised to other types of stochastic volatility diffusions. This vol-of-vol local volatility function is also time dependent such that we can perform a best-fit to no-touch options at each quoted maturity while ensuring perfect calibration to vanilla options through the local leverage function. The main drawback of working under a LSV model is the potentially less accurate calibration to no-touch options. However, the calibration results are very promising and the calibrated LSV-LVV model prices no-touch options within the market bid–ask spread for all barrier levels and maturities. On the numerical side, we propose a finite element approach to the solution of the Fokker–Planck PDE for the joint density of (S, M, V). It requires an optimal mesh construction, an efficient handling of the boundary equation when $S = M$ and a difficult point-wise Dirichlet boundary condition at (S_0, S_0, V_0) to incorporate. Validation results are presented for the calibrated models, where we price up-and-out call options by integrating the computed joint density against the path-dependent payoff function.

Finally, in Section 4.9 of Chapter 4, we combine the stochastic rates with path-dependent volatility models to improve the accuracy for long-dated options, while incorporating the touch probability of the market in the model. A novel self-consistency formula is derived for the leverage function and a calibration is possible via our control-variate particle method discussed in Chapter 3. The numerical experiments are left for further research.

1.4 Mimicking an Itô process

The work we present in this thesis is built upon two main concepts called mimicking results. If one considers a spot process which embeds an n-dimensional stochastic variance driven by an Itô process, it is possible to find a one-factor mimicking diffusion with an appropriate local volatility function that will exactly reproduce the marginal density of the spot process. This means that under both the stochastic volatility model and the local volatility model, a call option will have the same price. This finding is proved in the work of [84] and has been re-derived in a financial context, in [53]. First, assume the existence of a filtered probability space $(\chi, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q}^d)$ with a domestic risk-neutral measure $\mathbb{Q}^d$.

Theorem (Gyöngy (1986))**.** *Let S_t be an Itô process such that*

$$S_t = S_0 + \int_0^t \mu_s S_s \, ds + \int_0^t \alpha_s S_s \, dW_s \,,$$

where W is a standard Brownian motion and μ and α are continuous $\mathcal{F}_t$-adapted semi-martingales satisfying, for any $t \geq 0$,

$$\mathbb{E}^{\mathbb{Q}}\left[\int_0^t \left(|\mu_u S_u| + \alpha_u^2 S_u^2\right) \, du\right] < \infty.$$

Then there exist measurable functions defined as

$$\begin{aligned} \hat{\mu}(x,t) &= \mathbb{E}^{\mathbb{Q}}[\mu_t \,|\, S_t = x] \\ \hat{\alpha}(x,t) &= \sqrt{\mathbb{E}^{\mathbb{Q}}[\alpha_t^2 \,|\, S_t = x]}\,, \end{aligned}$$

such that the following SDE

$$\hat{S}_t = \hat{S}_0 + \int_0^t \hat{\mu}\big(\hat{S}_s, s\big)\hat{S}_s \, ds + \int_0^t \hat{\alpha}\big(\hat{S}_s, s\big)\hat{S}_s \, dW_s \,, \hat{S}_0 = S_0 \,,$$

admits a weak solution with the same marginal distribution as S for any time $t \geq 0$.

This result gave rise to the large literature of Markovian projection techniques to calibrate vanilla options [122, 82, 129].

Similarly, Brunick and Shreve [25] extend this approach to path-functionals of which the running maximum is a particular case. If one considers a spot process which embeds an n-dimensional stochastic variance driven by an Itô process, it is possible to find a one-factor two-state mimicking diffusion with an appropriate volatility function that will exactly reproduce the marginal density of the spot and running maximum process. This means that under both the stochastic volatility model and the Brunick–Shreve volatility model, an up-and-out call option will have the same price. More precisely, we have the following.

Theorem (Brunick and Shreve (2013))**.** *Let S_t be an Itô process such that*

$$\begin{cases} S_t &= S_0 + \int_0^t \mu_s S_s \, ds + \int_0^t \alpha_s S_s \, dW_s \,, \\ M_t &= \max\limits_{0 \le u \le t} S_u \,, \end{cases}$$

where W is a standard Brownian motion and μ and α are continuous $\mathcal{F}_t$-adapted semi-martingales satisfying, for any $t \geq 0$,

$$\mathbb{E}^{\mathbb{Q}}\left[\int_0^t \Big(|\mu_u S_u| + \alpha_u^2 S_u^2 \Big) \, du \right] < \infty.$$

Then there exist measurable functions defined as

$$\begin{aligned} \hat{\mu}(x,y,t) &= \mathbb{E}^{\mathbb{Q}}[\mu_t \,|\, S_t = x, \, M_t = y] \\ \hat{\alpha}(x,y,t) &= \sqrt{\mathbb{E}^{\mathbb{Q}}[\alpha_t^2 \,|\, S_t = x \,, M_t = y]} \,, \end{aligned}$$

such that the following SDE

$$\begin{cases} \hat{S}_t &= \hat{S}_0 + \int_0^t \hat{\mu}\Big(\hat{S}_s, \hat{M}_s, s\Big) \hat{S}_s \, ds + \int_0^t \hat{\alpha}\Big(\hat{S}_s, \hat{M}_s, s\Big) \hat{S}_s \, dW_s \,, \hat{S}_0 = S_0 \,, \\ \hat{M}_t &= \max\limits_{0 \le u \le t} \hat{S}_u \,, \end{cases}$$

admits a weak solution such that $\Big(\hat{S}, \hat{M}\Big)$ and $\Big(S, M\Big)$ have the same marginal distribution for any time $t \geq 0$.

The result provides the necessary theoretical framework to extend the work of Dupire [53] to barrier options. In Chapter 4, we prove Theorem 1.4 under a stronger condition and in a financial setting which facilitates the extension to stochastic rates.

2 A Forward Equation for Barrier Options under the Brunick–Shreve Markovian Projection

2.1 Introduction

In our work, we approach the calibration from the Brunick–Shreve mimicking point of view [25] in the general framework of continuous stochastic volatility models. We derive a condition to be satisfied by the expectation of the stochastic variance conditional on the spot and its running maximum, (S_t, M_t), in order to reproduce barrier prices. This conditional expectation is often referred to as a Markovian projection. For simplicity, we focus on up-and-out call options with no rebate and continuous monitoring of the barrier, but the analysis extends to other payoff types. In the case of double barrier options, one would also have to consider the running minimum leading to an additional dimension and boundary term.

The derivation is inspired by the work of Derman & Kani [49] and Dupire [54, 55]. In that context, Gyöngy's mimicking result [84] provides the financial engineer with a recipe to build a low-dimensional Markovian process which reproduces exactly any marginal density of the spot price process for all times t. Brunick and Shreve [25] extend Gyöngy's result to path functionals of the underlying process, of which the running maximum is a special case. In particular, they prove for any stochastic volatility process the existence of a mimicking low-dimensional Markovian process with the same joint density for (S_t, M_t).

In [62], Forde presents a way to retrieve the mimicking coefficient by deriving a forward equation for the characteristic function of (S_t, M_t); computation of the coefficient is possible via an inverse two-dimensional Fourier-Laplace transform. Here, we present an alternative method to retrieve the mimicking coefficient, which lies closer to the well-known Dupire formula and can hence benefit from the earlier work in the field of vanilla calibration.

We derive a partial-integro differential equation (PIDE) in strike, barrier level and maturity for up-and-out calls priced under the Brunick–Shreve model. This forward differential equation has the same useful properties (and drawbacks) as the Dupire forward PDE. As a consequence, it can be used to price a set of up-and-out calls in one single resolution. In this regard, our method shares some similarities with the forward equations derived by Carr and Hirsa [32]. We highlight a few key differences though. First, we consider the class of stochastic volatility models rather than local volatility models with a jump term. In working directly with Brunick and Shreve's Markovian

projection onto (S_t, M_t), we need to consider the running maximum explicitly in our derivation. This complicates the proof and makes the idea used in [32] – employing stopping times – not applicable in our case. For the above type of Markovian projection, one gets an unusual "integro" term in the PIDE involving a second derivative, which requires particular care in the numerical solution. In contrast, the "integro" term in [32] comes from the jump process and is not treated in the same way as ours.

A very recent paper by Guyon [81] explains how path-dependent volatility models, like the Brunick–Shreve one, may be very useful to replicate a market's spot-volatility dynamics, in particular highlighting the running maximum. It is important to note, though, that the diffusion of interest can be a fairly general stochastic process in our framework. More specifically, the variance process does not need to contain the running maximum in its parametrisation, and we do not particularly advocate the dynamic use of such a model here. Indeed, by doing so, one may find oneself with the logical conundrum that the model for the underlying depends on the time of inception of the *option*, i.e., the time the clock starts for the running maximum. The view we take here is that the Brunick–Shreve projection is a "code book" (to borrow a term from [29]) for barrier option prices, to which other models may be calibrated. Additionally, the forward PIDE enables in principle the pricing over a wide set of up-and-out call deals, creating a possible efficient direct solver for the inverse problem, i.e., to retrieve model parameters from any desired model class via the projected volatility $\sigma(S_t, M_t, t)$.

The remainder of this chapter is organised as follows. In Section 2.2, we introduce the modelling setup and hypotheses and derive, as our first main result, a forward equation (in terms of the maturity) for barrier option prices, where the strike and barrier levels are spatial variables. Next, in Section 2.3, we derive a Dupire-type formula for barrier options, leading to a known setup for the reader familiar with vanilla calibration. In order to build the first step of a calibration routine to reprice up-and-out call options, we deduce a forward PIDE in Section 2.4 with better stability properties and develop a numerical solution scheme. Finally, in Section 2.5, we show that the forward PIDE and the "classical" backward pricing PDE agree on the price of barrier options, which validates our approach. Section 2.6 concludes.

2.2 Setup and main result

We consider a market with a risk-free, deterministic and possibly time-dependent short rate, $r(t)$ at time t, and continuously compounded deterministic dividend $q(t)$. Then $D(t) = \exp(-\int_0^t r(u)\,du)$ is the discount factor, and $Q(t) = \exp(\int_0^t q(u)\,du)$ is the dividend capitalisation.

We assume the existence of a filtered probability space $(\chi, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q})$ with a (not necessarily unique) risk-neutral measure $\mathbb{Q}$, under which the price process of a risky asset follows

$$\frac{dS_t}{S_t} = (r(t) - q(t))\,dt + \alpha_t\,dW_t, \tag{2.2.1}$$

where W is a standard Brownian motion and α_t is a continuous and positive $\mathcal{F}_t$-adapted semi-martingale such that

$$\mathbb{E}^{\mathbb{Q}}\left[\int_0^t \alpha_u^2 S_u^2 \, du\right] < \infty. \tag{2.2.2}$$

A practically important example is $\alpha_t = f(S_t, t)\sqrt{V_t}$ with a *local volatility* function f and a CIR process V, which is often referred to as a *local stochastic volatility* (LSV) model.

In this chapter, we consider an up-and-out call option on S with continuously monitored barrier B, strike K, and maturity T. The arbitrage-free price under $\mathbb{Q}$ is

$$C(K, B, T) = \mathbb{E}^{\mathbb{Q}}\Big[D(T)(S_T - K)^+ \mathbf{1}_{M_T < B}\Big],$$

where $\mathbf{1}_A$ is the indicator function of event A and

$$M_t = \max_{0 \le u \le t} S_u$$

the running maximum process of S. Adaptations of our results to other types of barrier options (such as put payoffs, down-and-out barriers etc) are easily possible using a similar derivation to below.

We now derive the first main result which links the barrier option price to the Markovian projection of the stochastic volatility (4.2.1) onto the spot and its running maximum $(S_t, M_t)_{t \ge 0}$, which we define by

$$\sigma^2_{|S,M}(K, B, T) = \mathbb{E}^{\mathbb{Q}}\Big[\alpha_T^2 \mid S_T = K, M_T = B\Big]. \tag{2.2.3}$$

A recent result by Brunick and Shreve [25], which we apply to the process $\log(S_t)$, shows that under the integrability condition (2.2.2) on α, the function $\sigma_{|S,M}$ is measurable, and gives the existence of a one-dimensional "mimicking" process $\widehat{S}$ with running maximum $\widehat{M}$, such that for all t

$$(S_t, M_t) \overset{law}{=} (\widehat{S}_t, \widehat{M}_t), \tag{2.2.4}$$

where $\widehat{S}$ is the weak solution of

$$\frac{d\widehat{S}_t}{\widehat{S}_t} = (r(t) - q(t))\, dt + \sigma\Big(\widehat{S}_t, \widehat{M}_t, t\Big)\, d\widehat{W}_t, \tag{2.2.5}$$

with a standard Brownian motion $\widehat{W}$ on a suitable probability space, and

$$\sigma(K, B, T) = \sigma_{|S,M}(K, B, T).$$

Remark. We can take two views of (2.2.5) from the perspective of derivative pricing. First, we can think of σ in (2.2.5) as a model in its own right, with a local volatility which additionally depends

on the running maximum – see [81] for a discussion of "path-dependent" volatility models. This extra flexibility of matching the path-dependence of the local volatility allows calibration to barrier contracts as well as vanillas, yet keeps the market complete. Second, and more common, a different volatility model may be used, for instance the aforementioned LSV model. In that case, the mimicking Brunick–Shreve model will give the same barrier option prices as the higher-dimensional diffusion as barrier option values are characterised precisely by the joint distribution of spot and running maximum. This property is important for calibration purposes as we will describe later. In either case, σ plays a similar role for barrier options as the Dupire local volatility does for vanillas.

Assumption 1. *We assume that σ in (2.2.5) is bounded and $C^{2,1,0}(\Omega \times \mathbb{R}^+)$. This implies that $(\widehat{S}, \widehat{M})$ is Markovian (see [24]; bounded and locally Lipschitz is sufficient). We also assume that the joint density ϕ under $\mathbb{Q}$ of (S_t, M_t) – or, equivalently, $(\widehat{S}_t, \widehat{M}_t)$ – exist and is twice differentiable, with bounded derivatives.*

Proposition 2.1. *Under Assumption 1, the density ϕ is the classical solution of the following Kolmogorov forward equation in the region $\Omega = \{(x, y) : 0 < x < y, S_0 < y\}$,*

$$\begin{aligned}
\frac{\partial \phi}{\partial t} + (r(t) - q(t))\frac{\partial}{\partial x}(x\phi) - \frac{1}{2}\frac{\partial^2}{\partial x^2}\Big(\sigma^2 x^2 \phi\Big) &= 0, && (x, y) \in \Omega,\ t > 0, && (2.2.6)\\
\frac{\partial}{\partial x}\Big(\sigma^2 x^2 \phi\Big) + \frac{1}{2}\frac{\partial}{\partial y}\Big(\sigma^2 x^2 \phi\Big) &= (r(t) - q(t))\ x\ \phi, && x = y,\ t > 0, && (2.2.7)\\
\phi(S_0, S_0, t) &= 0, && t > 0, && (2.2.8)\\
\phi(x, y, 0) &= \delta(x - S_0)\delta(y - S_0), && (x, y) \in \Omega.
\end{aligned}$$

Proof. We assume that the joint density function ϕ of the process (S_t, M_t) exists and belongs to the space $C^{2,1,1}(\Omega \times \mathbb{R}^+)$. Let $h \in C^{2,1}(\Omega)$ be a *test* function of two (spatial) variables. We also assume that h and its derivatives vanish for $x = 0$. We denote $\mu(t) = r(t) - q(t)$. We apply the Itô-Doeblin lemma and get

$$\begin{aligned}
h(S_t, M_t) &= h(S_0, M_0)\\
&\quad + \int_0^t \left[\mu(u) S_u \frac{\partial h(S_u, M_u)}{\partial x} + \frac{1}{2}\sigma^2(S_u, M_u, u) S_u^2 \frac{\partial^2 h(S_u, M_u)}{\partial x^2}\right] \mathrm{d}u\\
&\quad + \int_0^t \sigma(S_u, M_u, u) S_u \frac{\partial h(S_u, M_u)}{\partial x}\, dW_u + \int_0^t \frac{\partial h(S_u, M_u)}{\partial y}\, dM_u\,.
\end{aligned}$$

By taking expectations we can write

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[h(S_t, M_t)] &= h(S_0, M_0) + \mathbb{E}^{\mathbb{Q}}\left[\int_0^t \mu(t) S_u \frac{\partial h(S_u, M_u)}{\partial x}\, \mathrm{d}u\right]\\
&\quad + \mathbb{E}^{\mathbb{Q}}\left[\int_0^t \frac{1}{2}\sigma^2(S_u, M_u, u) S_u^2 \frac{\partial^2 h(S_u, M_u)}{\partial x^2}\, \mathrm{d}u\right] + \mathbb{E}^{\mathbb{Q}}\left[\int_0^t \frac{\partial h(S_u, M_u)}{\partial y}\, dM_u\right].
\end{aligned}$$

Hence, we can introduce the density function ϕ of (S, M), assumed twice differentiable,

$$\mathbb{E}^{\mathbb{Q}}[h(S_t, M_t)] = \int_{S_0}^{\infty} \int_0^y h(x,y)\phi(x,y,t)\,\mathrm{d}x\,\mathrm{d}y\,,$$

differentiate with respect to t, and write for all $t > 0$

$$\int_{S_0}^{\infty} \int_0^y h\frac{\partial \phi}{\partial t}\,\mathrm{d}x\,\mathrm{d}y = \int_{S_0}^{\infty} I_\mu(y)\,\mathrm{d}y + \frac{1}{2}\int_{S_0}^{\infty} I_\sigma(y)\,\mathrm{d}y + \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}^{\mathbb{Q}}\left[\int_0^t \frac{\partial h(S_u, M_u)}{\partial y}\,dM_u\right], \tag{2.2.9}$$

where, explicitly with all arguments,

$$\begin{aligned} I_\sigma(y) &= \int_0^y \left(\sigma^2(x,y,t)x^2\frac{\partial^2 h(x,y)}{\partial x^2}\right)\phi(x,y,t)\,\mathrm{d}x\,, \\ I_\mu(y) &= \int_0^y \left(\mu(t)x\frac{\partial h(x,y)}{\partial x}\right)\phi(x,y,t)\,\mathrm{d}x\,. \end{aligned}$$

For I_σ, we can perform integration by parts twice to get

$$\begin{aligned} I_\sigma(y) &= \frac{\partial h(x,y)}{\partial x}\Big]_{x=y}\sigma^2(y,y,t)y^2\phi(y,y,t) - h(y,y)\,\frac{\partial\left(\sigma^2(x,y,t)x^2\phi(x,y,t)\right)}{\partial x}\Big]_{x=y} \\ &\quad + \int_0^y h(x,y)\frac{\partial^2\left(\sigma^2(x,y,t)x^2\phi(x,y,t)\right)}{\partial x^2}\,\mathrm{d}x\,. \end{aligned}$$

For the integral of I_σ, since

$$\frac{\partial h(x,y)}{\partial x}\Big]_{x=y} = \frac{\mathrm{d}h(y,y)}{\mathrm{d}y} - \frac{\partial h(x,y)}{\partial y}\Big]_{x=y},$$

we can integrate the first term by parts with respect to y,

$$\begin{aligned} \int_{S_0}^{\infty} I_\sigma(y)\,\mathrm{d}y &= -\int_{S_0}^{\infty} h(y,y)\left[\frac{\partial\left(\sigma^2(x,y,t)x^2\phi(x,y,t)\right)}{\partial x}\Big]_{x=y} + \frac{\mathrm{d}\left(\sigma^2(y,y,t)y^2\phi(y,y,t)\right)}{\mathrm{d}y}\right]\mathrm{d}y \\ &\quad + \int_{S_0}^{\infty}\int_0^y h(x,y)\left[\frac{\partial^2\left(\sigma^2(x,y,t)x^2\phi(x,y,t)\right)}{\partial x^2}\right]\mathrm{d}x\,\mathrm{d}y \\ &\quad - \int_{S_0}^{\infty}\frac{\partial h(x,y)}{\partial y}\Big]_{x=y}\sigma^2(y,y,t)y^2\phi(y,y,t)\,\mathrm{d}y \\ &\quad - h(S_0,S_0)\sigma^2(S_0,S_0,t)S_0^2\phi(S_0,S_0,t). \end{aligned}$$

For I_μ, we integrate by parts once,

$$I_\mu(y) = \mu(t)h(y,y)y\phi(y,y,t) - \int_0^y h(x,y)\left[(r(t)-q(t))\frac{\partial x\phi(x,y,t)}{\partial x}\right]\mathrm{d}x.$$

We insert in (2.2.9),

$$\int_{S_0}^{\infty}\int_0^y h(x,y)\left[\frac{\partial\phi(x,y,t)}{\partial t}+\mu(t)\frac{\partial x\phi(x,y,t)}{\partial x}-\frac{1}{2}\left(\frac{\partial^2\big(\sigma^2(x,y,t)x^2\phi(x,y,t)\big)}{\partial x^2}\right)\right]\mathrm{d}x\,\mathrm{d}y = \tag{2.2.10}$$

$$\int_{S_0}^{\infty} h(y,y)\left[\mu(t)y\phi(y,y,t)-\frac{1}{2}\left(\frac{\partial\big(\sigma^2(x,y,t)x^2\phi(x,y,t)\big)}{\partial y}\bigg\rfloor_{x=y}+2\,\frac{\partial\big(\sigma^2(x,y,t)x^2\phi(x,y,t)\big)}{\partial x}\bigg\rfloor_{x=y}\right)\right]\mathrm{d}y$$

$$-\frac{1}{2}h(S_0,S_0)\sigma^2(S_0,S_0,t)S_0^2\phi(S_0,S_0,t)$$

$$-\int_{S_0}^{\infty}\frac{\partial h(x,y)}{\partial y}\bigg\rfloor_{x=y}\sigma^2(y,y,t)y^2\phi(y,y,t)\,\mathrm{d}y+\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}^{\mathbb{Q}}\left[\int_0^t\frac{\partial h(S_u,M_u)}{\partial y}\,dM_u\right].$$

Let us first consider all functions h of compact support on Ω. Hence h and its derivatives are assumed to additionally vanish for $x=y$. Since M_t only grows when $M_t=S_t$, the term $\int_0^t\frac{\partial h(S_u,M_u)}{\partial y}\,dM_u$ is zero. Then (2.2.10) becomes

$$\int_{S_0}^{\infty}\int_0^y h(x,y)\left[\frac{\partial\phi(x,y,t)}{\partial t}+\mu(t)\frac{\partial x\phi(x,y,t)}{\partial x}-\frac{1}{2}\left(\frac{\partial^2\big(\sigma^2(x,y,t)x^2\phi(x,y,t)\big)}{\partial x^2}\right)\right]\mathrm{d}x\,\mathrm{d}y=0\,,$$

where we conclude that for all (x,y,t) in $\big(\Omega\times\mathbb{R}_*^+\big)$

$$\frac{\partial\phi(x,y,t)}{\partial t}+\mu(t)\frac{\partial x\phi(x,y,t)}{\partial x}-\frac{1}{2}\left(\frac{\partial^2\big(\sigma^2(x,y,t)x^2\phi(x,y,t)\big)}{\partial x^2}\right)=0\,.$$

Let us now consider all functions h which do not vanish at $x=y$ but only depend on the space variable x such that we define $g(x)=h(x,y)$ with g in the space $C^2\big(\mathbb{R}^+\big)$. In that case, the terms $\int_0^t\frac{\partial h}{\partial y}\,dM_u$ and $\int_{S_0}^{\infty}\frac{\partial h(x,y)}{\partial y}\Big\rfloor_{x=y}\sigma^2y^2\phi\big\rfloor_{x=y}\mathrm{d}y$ vanish. From (2.2.10) we can write

$$0=\int_{S_0}^{\infty}g(y)\left[\mu(t)y\phi(y,y,t)-\frac{1}{2}\left(\frac{\partial(\sigma^2(x,y,t)x^2\phi(x,y,t))}{\partial y}\bigg\rfloor_{x=y}+2\,\frac{\partial(\sigma^2(x,y,t)x^2\phi(x,y,t))}{\partial x}\bigg\rfloor_{x=y}\right)\right]\mathrm{d}y-$$

$$\frac{1}{2}g(S_0)\sigma^2(S_0,S_0,t)S_0^2\phi(S_0,S_0,t).$$

If additionally we impose g to vanish at $x=S_0$, then we conclude that for all (x,y,t) in $\big(\Omega\times\mathbb{R}_*^+\big)$

$$\mu(t)y\phi(y,y,t)=\frac{1}{2}\,\frac{\partial(\sigma^2(x,y,t)x^2\phi(x,y,t))}{\partial y}\bigg\rfloor_{x=y}+\frac{\partial(\sigma^2(x,y,t)x^2\phi(x,y,t))}{\partial x}\bigg\rfloor_{x=y}.$$

Finally, for all functions g which do not vanish at $x=S_0$ we are left with

$$g(S_0,t)\sigma^2(S_0,S_0,t)S_0^2\phi(S_0,S_0,t)=0\,,$$

which concludes the proof. □

Remark. It is straightforward to generalise the approach to work directly under (4.2.1), say under a local-stochastic volatility model, to derive forward and backward equations for the joint density of S_t, M_t and the stochastic variance V_t. We find again the dual operator with respect to the L_2 inner product, now with three spatial dimensions.

This initial-boundary value problem is the adjoint to the backward equation satisfied by the option value C as a function of S_t, M_t and t, which has a homogeneous Neumann boundary condition on the diagonal $S = M$.

Remark. The assumption of smoothness of the joint density is non-trivial. For the Black-Scholes model, with constant σ, the joint density function is known explicitly and smooth. In a recent work, [63] show the existence of the joint density – not necessarily differentiable – of an Itô process X and its running minimum $\underline{X}$, where the local volatility is a sufficiently smooth and suitably bounded function of X_t and $\underline{X}_t$. We also note that although the density features in our proof, it does not appear in the final result itself, and we conjecture that the regularity assumption may be weakened. Nonetheless, forward equations involving the transition density are widely used for calibration in practice and we envisage that (2.2.6) can also be a useful building block for calibration in the present framework. Therefore, it seems reasonable to require that the model has enough regularity for the density to satisfy such an equation.

The time zero value of barrier options can then be written as

$$\begin{aligned} C(K,B,T) &= D(T)\int_{S_0}^{\infty}\int_{0}^{\infty}(x-K)^{+}\mathbf{1}_{y<B}\,\phi(x,y,T)\,\mathrm{d}x\,\mathrm{d}y \\ &= D(T)\int_{K\vee S_0}^{B}\int_{K}^{y}(x-K)\,\phi(x,y,T)\,\mathrm{d}x\,\mathrm{d}y. \end{aligned} \tag{2.2.11}$$

We can now extend the argument from Dupire [54] for European calls to derive a forward equation for barriers.

Theorem 2.2. *Under (2.2.2) and Assumption 1, the value $C(K,B,T)$ of an up-and-out barrier call satisfies, for all $0 < K < B$ and $T > 0$,*

$$\begin{aligned} \frac{\partial^2 C(K,B,T)}{\partial B\partial T} + (r(T)-q(T))K\frac{\partial^2 C(K,B,T)}{\partial K\partial B} &= \frac{1}{2}\sigma^2_{|S,M}(K,B,T)K^2\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B} - q(T)\frac{\partial C(K,B,T)}{\partial B} \\ &\quad - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}B}\left((B-K)B^2\sigma^2_{|S,M}(B,B,T)\,\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B}\bigg|_{K=B}\right), \end{aligned} \tag{2.2.12}$$

where $\sigma_{|S,M}$ is given by (2.2.3).

Proof. Because of Brunick and Shreve's mimicking result, (2.2.4) and (2.2.5), and the smoothness assumption, we can work with a density satisfying the forward equation (2.2.6) and (2.2.7). We denote

$\mu(t) = r(t) - q(t)$. Differentiating (2.2.11) with respect to B and T in the first equation, using (2.2.6) in the second, and integrating by parts in the third, we obtain

$$\begin{aligned}
\frac{\partial^2 C(K,B,T)}{\partial T \partial B} &= -r(T)\frac{\partial C(K,B,T)}{\partial B} + D(T)\int_K^B (x-K)\frac{\partial \phi}{\partial t}(x,B,T)\,\mathrm{d}x \\
&= -r(T)\frac{\partial C(K,B,T)}{\partial B} + \\
&\quad D(T)\int_K^B (x-K)\left[-\mu(T)\frac{\partial}{\partial x}(x\phi(x,B,T)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\big(\sigma^2(x,B,T)x^2\phi(x,B,T)\big)\right]\mathrm{d}x \\
&= -r(T)\frac{\partial C(K,B,T)}{\partial B} + \\
&\quad \mu(T)D(T)\left[-B(B-K)\phi(B,B,T) + \int_K^B (x-K)\phi(x,B,T)\,\mathrm{d}x + K\int_K^B \phi(x,B,T)\,\mathrm{d}x\right] + \\
&\quad \frac{1}{2}D(T)\big(\sigma^2(K,B,T)K^2\phi(K,B,T) - \sigma^2(B,B,T)B^2\phi(B,B,T)\big) + \\
&\quad \frac{1}{2}D(T)(B-K)\left[\frac{\partial}{\partial x}\big(\sigma^2(x,B,T)x^2\phi(x,B,T)\big)\right]_{x=B}.
\end{aligned}$$

Using the boundary condition (2.2.7) in the second line of

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}B}\big(\sigma^2(B,B,T)B^2\phi(B,B,T)\big) &= \left[\frac{\partial}{\partial x}\big(\sigma^2(x,B,T)x^2\phi(x,B,T)\big)\right]_{x=B} + \left[\frac{\partial}{\partial B}\big(\sigma^2(x,B,T)x^2\phi(x,B,T)\big)\right]_{x=B} \\
&= -\left[\frac{\partial}{\partial x}\big(\sigma^2(x,B,T)x^2\phi(x,B,T)\big)\right]_{x=B} + 2\mu(T)B\phi(B,B,T),
\end{aligned}$$

we get, by another application of the product rule,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}B}\big((B-K)\sigma^2(B,B,T)B^2\phi(B,B,T)\big) &= -(B-K)\left[\frac{\partial}{\partial x}\big(\sigma^2(x,B,T)x^2\phi(x,B,T)\big)\right]_{x=B} + \\
&\quad 2\mu(T)(B-K)B\phi(B,B,T) + \sigma^2(B,B,T)B^2\phi(B,B,T).
\end{aligned}$$

The result now follows by differentiating (2.2.11) once more to obtain

$$\begin{aligned}
\frac{\partial^2 C(K,B,T)}{\partial K \partial B} &= -D(T)\int_K^B \phi(x,B,T)\,\mathrm{d}x, \\
\frac{\partial^3 C(K,B,T)}{\partial K^2 \partial B} &= D(T)\phi(K,B,T),
\end{aligned} \tag{2.2.13}$$

and substituting everything above. □

Remark. We can think of (2.2.12) in two ways. Given a model, either via σ directly or via a specification that allows computation of the Markovian projection, the forward equation allows computation of barrier option prices across all strikes, maturities and barriers as the solution of a single PDE. Conversely, if barrier call prices are observed on the market, (2.2.12) allows inferences on the Markovian projection (2.2.3) of the volatility. As is the case with the Dupire formula for European calls, a continuum of prices is not available and some sort of interpolation is required and can be

notoriously unstable. We will return to these points in depth in the following sections.

For future reference, we define

$$\widetilde{C}(K,B,T) = \mathbb{E}^{\mathbb{Q}}\Big[D(T)Q(T)(S_T-K)^+\mathbf{1}_{M_T<B}\Big],$$

the capitalisation of the market price with dividends. By doing so, we eliminate the term $q(T)\frac{\partial C(K,B,T)}{\partial B}$ if we replace $C(K,B,T)$ by $\widetilde{C}(K,B,T)$ in (2.2.12). We will work with $\widetilde{C}$ in the remainder of the chapter.

2.3 A Dupire-type formula for barrier options

The program we aim to complete is summarized in Fig. 2.3.1.

		Gyöngy [84]		Dupire [54]
		$\downarrow$		$\downarrow$
$\sigma^2_{\vert S}(K,t)$	$=$	$\mathbb{E}^{\mathbb{Q}}[\alpha_t^2 \mid S_t = K]$	$=$	$\frac{\frac{\partial C}{\partial T}+(r-q)K\frac{\partial C}{\partial K}+qC}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}}$
$\sigma^2_{\vert S,M}(K,B,t)$	$=$	$\mathbb{E}^{\mathbb{Q}}[\alpha_t^2 \mid S_t = K \cap M_t = B]$	$=$	(2.3.4)
		$\uparrow$		
		Brunick & Shreve [25]		

Figure 2.3.1: Analogy of results for European and barrier options.

Equation (2.2.12) cannot be solved for the Markovian projection $\sigma_{|S,M}$ directly, due to the presence of the term $\sigma_{|S,M}(B,B,T)$. We derive a formula for $\sigma_{|S,M}$ as follows.

At $K=0$, (2.2.12) still holds and integration with respect to B gives

$$\frac{\partial\widetilde{C}(0,B,T)}{\partial T} = -\frac{1}{2}\sigma^2_{|S,M}(B,B,T)B^3\,\frac{\partial^3\widetilde{C}(K,B,T)}{\partial K^2\partial B}\bigg|_{K=B} \qquad \forall (B,T)\in\,]S_0,+\infty[\,\times\mathbb{R}_*^+, \tag{2.3.1}$$

noticing that because of (2.2.13) and (2.2.8) no integration constant appears. It links the price of the foreign no-touch option,

$$FNT(B,T) = D(T)\mathbb{E}^{\mathbb{Q}}[S_T\mathbf{1}_{M_T<B}] = \frac{\widetilde{C}(0,B,T)}{Q(T)}\,, \tag{2.3.2}$$

to the market-implied joint density of (S_t,M_t) at $K=B$; see (2.2.13). If we substitute (2.3.1) into the last term of (2.2.12), we get

$$\frac{\partial^2\widetilde{C}(K,B,T)}{\partial B\partial T} + (r(T)-q(T))K\frac{\partial^2\widetilde{C}(K,B,T)}{\partial K\partial B} = \frac{1}{2}\sigma^2_{|S,M}(K,B,T)K^2\frac{\partial^3\widetilde{C}(K,B,T)}{\partial K^2\partial B} + \frac{\partial\Big(\frac{(B-K)^+}{B}\frac{\partial\widetilde{C}(0,B,T)}{\partial T}\Big)}{\partial B}\,.$$

Therefore, as $B \geq K$ in the case of interest,

$$\frac{\partial^2 \widetilde{C}(K,B,T)}{\partial B \partial T} - \frac{K}{B^2}\frac{\partial \widetilde{C}(0,B,T)}{\partial T} - \frac{(B-K)}{B}\frac{\partial^2 \widetilde{C}(0,B,T)}{\partial T \partial B} + (r(T)-q(T))K\frac{\partial^2 \widetilde{C}(K,B,T)}{\partial K \partial B} = \frac{1}{2}\sigma^2_{|S,M}(K,B,T)K^2\frac{\partial^3 \widetilde{C}(K,B,T)}{\partial K^2 \partial B}\,. \tag{2.3.3}$$

By rearrangement of (2.3.3) we get the following Dupire-type formula.

Corollary 2.3. *Under the assumptions of Theorem 2.2, the unique Brunick–Shreve mimicking volatility for up-and-out call options is*

$$\sigma_{|S,M}(K,B,T) = \sqrt{\frac{\frac{\partial^2 \widetilde{C}(K,B,T)}{\partial B \partial T} - \frac{K}{B^2}\frac{\partial \widetilde{C}(0,B,T)}{\partial T} - \frac{(B-K)}{B}\frac{\partial^2 \widetilde{C}(0,B,T)}{\partial T \partial B} + (r(T)-q(T))K\frac{\partial^2 \widetilde{C}(K,B,T)}{\partial K \partial B}}{\frac{1}{2}K^2\frac{\partial^3 \widetilde{C}(K,B,T)}{\partial K^2 \partial B}}}, \quad 0 \leq K \leq B,\, T \geq 0\,. \tag{2.3.4}$$

This formula suffers from similar issues to those one finds with the evaluation of the standard Dupire formula, in that derivatives of the value function over a continuum of strikes and maturities are required. Practical approximations using the sparsely available data are necessarily sensitive to the method of interpolation.

This is exacerbated here as (2.2.12) involves derivatives up to order four. The formula (2.3.4) can be expected to be numerically somewhat more stable since we replaced the higher-order derivative with respect to the strike at barrier level with first and second order derivatives at zero strike. In addition to the numerical improvement, it is also easier in practice to retrieve barrier prices at zero strike with foreign no-touch prices, recalling (2.3.2), for which quotes are often available (e.g., in the FX markets).

2.4 A forward partial-integro differential equation for barrier options

It is well known that ill-posed parameter estimation problems can often be regularised through a penalised optimisation routine, a good example of which is the calibration of local volatility through Tikhonov regularisation as presented, e.g., by Crépey [42], Egger and Engl [56], as well as Achdou and Pironneau [2].

However, in order to achieve this, a suitable forward partial differential equation is required. We now propose a further rearrangement of (2.2.12), which is more suited to the numerical computation of $C(K,B,T)$ taking the Markovian projection $\sigma_{|S,M}$ as input.

2.4.1 Formulation as PIDE

Equation (2.2.12) can also be expressed in PIDE form by integrating with respect to B. We start by integrating the diffusive term of (2.2.12),

$$\begin{aligned}\int_{S_0\vee K}^{B} \tfrac{1}{2}\sigma^2_{|S,M}(K,b,T)K^2\tfrac{\partial^3\widetilde{C}(K,b,T)}{\partial K^2\partial b}\,\mathrm{d}b \;=\;& \tfrac{1}{2}\sigma^2_{|S,M}(K,B,T)K^2\tfrac{\partial^2\widetilde{C}(K,B,T)}{\partial K^2}\\ &-\tfrac{1}{2}\sigma^2_{|S,M}(K,S_0\vee K,T)K^2\tfrac{\partial^2\widetilde{C}(K,S_0\vee K,T)}{\partial K^2}\\ &-\int_{S_0\vee K}^{B}\tfrac{1}{2}K^2\tfrac{\partial^2\widetilde{C}(K,b,T)}{\partial K^2}\tfrac{\partial\sigma^2_{|S,M}(K,b,T)}{\partial b}\,\mathrm{d}b.\end{aligned}$$

The term in the second line vanishes for all K and T. Indeed, when $S_0 \geq K$ this option is already knocked-out. When $S_0 < K$, the term is also zero as explained below in (2.4.5). The other terms can all be directly integrated with respect to B taking into account that, similarly, no integration constant will appear. This allows us to define the following initial boundary value problem.

Corollary 2.4. *Under the assumptions of Theorem 2.2, the up-and-out call price follows the Volterra-type PIDE, expressed as an initial boundary value problem,*

$$\begin{aligned}&\frac{\partial\widetilde{C}(K,B,T)}{\partial T}+(r(T)-q(T))K\frac{\partial\widetilde{C}(K,B,T)}{\partial K}-\frac{1}{2}\sigma^2_{|S,M}(K,B,T)K^2\frac{\partial^2\widetilde{C}(K,B,T)}{\partial K^2}= \qquad (2.4.1)\\ &\quad-\frac{1}{2}\sigma^2_{|S,M}(B,B,T)B^2(B-K)\frac{\partial^3\widetilde{C}(B,B,T)}{\partial K^2\partial B}-\int_{S_0\vee K}^{B}\frac{1}{2}K^2\frac{\partial^2\widetilde{C}(K,b,T)}{\partial K^2}\frac{\partial\sigma^2_{|S,M}(K,b,T)}{\partial b}\,\mathrm{d}b\,,\\ &\qquad\qquad (K,B,T)\in[0,+\infty[\times\,]S_0,+\infty[\times\mathbb{R}^+_*,\end{aligned}$$

$$\widetilde{C}(K,B,0)=(S_0-K)^+\mathbf{1}_{S_0<B},\qquad T=0, \qquad (2.4.2)$$

$$\widetilde{C}(B,B,T)=0,\qquad K=B, \qquad (2.4.3)$$

$$\widetilde{C}(K,S_0,T)=0,\qquad B=S_0. \qquad (2.4.4)$$

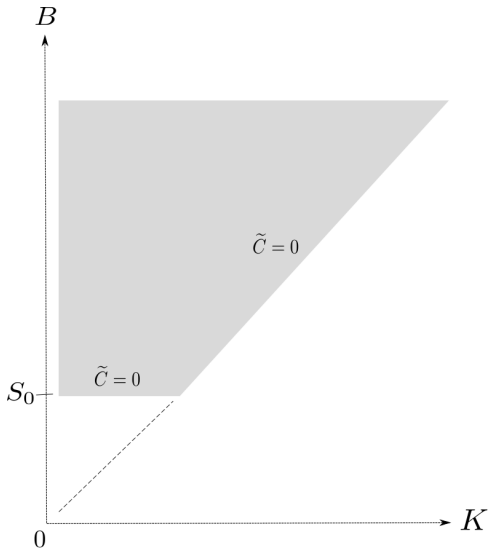


Figure 2.4.1: PIDE Domain and Boundaries

The initial condition (2.4.2) is the payoff obtained at maturity. Condition (2.4.3) expresses the fact that the option gets knocked out before being in-the-money if $K = B$, while (2.4.4) says the option gets knocked out immediately at inception if $S_0 = B$.

Remark. We note that in the models we will consider, no boundary condition needs to be specified at $K = 0$, since the coefficients of the K-derivatives vanish sufficiently fast as $K \to 0$. This is the case, e.g., if $\sigma^2_{|S,M}$ and its B-derivative are bounded when $K \to 0$.

In fact, a boundary condition at $K = 0$ could be derived by differentiating (2.2.11) twice,

$$\frac{\partial^2 \widetilde{C}(0,B,T)}{\partial K^2} = D(T)Q(T)\int_{S_0}^{B} \phi(0,y,T)\,\mathrm{d}y = D(T)Q(T)\mathbb{Q}(S_0 < M_T < B \mid S_T = 0)\psi(0,T),$$

where $\psi(0,T)$ is the density of S_T at 0. The latter is equal to zero since the density of the spot at zero is zero under the log-spot model in (4.2.1). Linearity conditions, also known as "Zero Gamma" for spot PDEs, are commonly used in finance and particularly useful for exotic contracts. They usually lead to stable numerical schemes (see Windcliff [150]).

We also note that

$$\frac{\partial^2 \widetilde{C}(B,B,T)}{\partial K^2} = D(T)Q(T)\mathbb{Q}(S_0 < M_T < B \mid S_T = B)\psi(B,T) = 0, \tag{2.4.5}$$

because if the spot is equal to B at T, then the probability that the running maximum is smaller than B is zero. While we will use the Dirichlet condition on the boundary $K = B$ for computations, (2.4.5) allows us to express the third order cross derivative in the second line of (2.4.1) by a third

order derivative only with respect to K. Indeed,

$$0 = \frac{\mathrm{d}}{\mathrm{d}B}\left(\frac{\partial^2 \widetilde{C}(B,B,T)}{\partial K^2}\right) = \frac{\partial^3 \widetilde{C}(B,B,T)}{\partial K^3} + \frac{\partial^3 \widetilde{C}(B,B,T)}{\partial K^2 \partial B},$$

leading to

$$\frac{\partial^3 \widetilde{C}(B,B,T)}{\partial K^3} = -\frac{\partial^3 \widetilde{C}(B,B,T)}{\partial K^2 \partial B}. \tag{2.4.6}$$

The domain sketched in Fig. 2.4.1 is unbounded for large B. We denote B_{Max} the maximum value of barrier levels we will consider in the numerical solution. Note that because of the structure of (2.4.1), no boundary condition is needed for $B = B_{Max}$.

2.4.2 Finite difference approximation

In this Sub-section we propose a finite difference scheme with uniform spacing for simplicity of implementation. In Chapter 4, we further extend the idea and develop an improved discretisation scheme with non-uniform space grid and backward differentiation formula time stepping. We define a uniform mesh which contains $M+1$ time points, $N+1$ spatial points in the strike and $P+1$ in the barrier variable, leading to the following definition of the step sizes:

$$\begin{array}{rclll} K_i &=& i\Delta_K, & \Delta_K = \frac{B_{Max}-S_0}{N+1}, & i \in [\![0, N]\!], \\ B_j &=& S_0 + j\Delta_B, & \Delta_B = \frac{B_{Max}-S_0}{P+1}, & j \in [\![0, P]\!], \\ T_m &=& m\Delta_T, & \Delta_T = \frac{T_{Max}}{M+1}, & m \in [\![0, M]\!]. \end{array}$$

For simplicity, we impose $\Delta_K = \Delta_B$. This will ensure that for any B_j, the corresponding mesh line will contain at least all (B_u, B_j) for all u smaller than j, which is useful for the following algorithm.

We can identify an interesting property of (2.4.1), especially visible with the substitution (2.4.6), in order to approximate the solution inductively on a discrete lattice. At $B = S_0$, the solution is known to be uniformly zero over the strike and time variable. Assume now an approximate solution is known up to a certain B. Moving from B to $B + \Delta_B$, and approximating the integral by a quadrature rule, gives a PDE in time and strike at level $B + \Delta_B$, which can be solved by finite differences. From now on, we will refer to the PDE at a given barrier level B_j as a "PDE layer". We can then solve the PIDE for each layer at points B_j from S_0 to B_{Max}.

We denote the discrete solution vector in such a layer by

$$\mathbf{u}^m_{\cdot,j} = \left[\widetilde{C}(K_0, B_j, T_m), \ldots, \widetilde{C}(B_j, B_j, T_m)\right]'$$

of size $n_j = \frac{B_j - K_0}{\Delta_K} + 1$, where $'$ denotes the transpose. We also denote by $\mathbf{I}_n$ the identity matrix of

size $n \times n$.

Derivatives

Derivatives are approximated by centered finite differences at each space point except at $K = 0$ and $K = B_j$, where they are computed, respectively, forward and backward. For the time being, the boundary conditions are not taken into account. We assume equally spaced points and define two operators as follows:

$$\begin{aligned} \delta_K \mathbf{u}_{i,j}^m &= \frac{\mathbf{u}_{i+1,j}^m - \mathbf{u}_{i-1,j}^m}{2\Delta_K}, \\ \delta_{KK} \mathbf{u}_{i,j}^m &= \frac{\mathbf{u}_{i+1,j}^m - 2\mathbf{u}_{i,j}^m + \mathbf{u}_{i-1,j}^m}{\Delta_K^2}. \end{aligned} \tag{2.4.7}$$

The usual matrix derivative operator can be defined for both the first and second order derivative:

$$\mathbf{D} = \frac{1}{\Delta_K} \begin{bmatrix} -1 & 1 & 0 & ... & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & ... & 0 & -1 & 1 \end{bmatrix}, \quad \mathbf{D_2} = \frac{1}{\Delta_K^2} \begin{bmatrix} 1 & -2 & 1 & 0 & ... & 0 \\ 1 & -2 & 1 & 0 & (0) & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & (0) & 0 & 1 & -2 & 1 \\ 0 & ... & 0 & 1 & -2 & 1 \end{bmatrix}.$$

We also define the forward time difference operator

$$\delta_T \mathbf{u}_{i,j}^m = \frac{\mathbf{u}_{i,j}^{m+1} - \mathbf{u}_{i,j}^m}{\Delta T}.$$

Integral

The integral term will be computed using the trapezoidal quadrature rule. This yields an order two consistent approximation. The term of interest is

$$F(K_i, B_j, T_m) = \int_{S_0 \vee K}^{B_j} -\frac{1}{2} K_i^2 \frac{\partial^2 \widetilde{C}(K_i, b, T_m)}{\partial K^2} \frac{\partial \sigma_{|S,M}^2(K_i, b, T_m)}{\partial b} \, \mathrm{d}b.$$

Let us assume that we know an approximation to the solution of the PIDE for the discrete set of barriers $(B_n)_{n<j}$. We also know that, when the barrier is at the spot level, the integrand is zero.

Hence,

$$\begin{aligned} F(K_i, B_j, T_m) \;=\; & -\sum_{n=1}^{j-1}\left(\frac{1}{2}K_i^2\frac{\partial^2\widetilde{C}(K_i,B_n,T_m)}{\partial K^2}\frac{\partial\sigma^2_{|S,M}(K_i,B_n,T_m)}{\partial B}\Delta_B\right) \\ & -\frac{1}{4}K_i^2\frac{\partial^2\widetilde{C}(K_i,B_j,T_m)}{\partial K^2}\frac{\partial\sigma^2_{|S,M}(K_i,B_j,T_m)}{\partial B}\Delta_B \\ & +\mathcal{O}\left(\Delta_B^2\right). \end{aligned} \tag{2.4.8}$$

The sum in the first line can be computed for all K_i and T_m in a forward induction over j, as the solution is known for all the barrier levels involved. This sum is then treated as a source function for the PDE layer of level B_j. We define a vector

$$\mathbf{f}^m_{\cdot,j} = \left[\; -\textstyle\sum_{n=1}^{j-1}\left(\frac{1}{2}K_i^2\delta_{KK}\mathbf{u}^m_{i,n}\frac{\partial\sigma^2_{|S,M}(K_i,B_n,T_m)}{\partial B}\Delta_B\right)\;\right]_{i=0,1,\ldots,n_j}.$$

The remaining term in (2.4.8) gives a small correction to the diffusion at B_j and we can incorporate it in the discretisation of the corresponding diffusive term of (2.4.1).

Boundary derivative term

To approximate the "boundary derivative" at (B, B, T) in (2.4.1), we use (2.4.6) and a first order approximation to the third derivative with discretisation matrix written as

$$\mathbf{\Phi} = \frac{1}{\Delta_K^3}\begin{bmatrix} 0 & \ldots & 0 & 0 & -1 & 3 & -3 & 1 \\ \vdots & (0) & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & -1 & 3 & -3 & 1 \end{bmatrix}.$$

As this term is present in the discretised equation for all interior mesh points, this reduces the consistency order in Δ_K of the overall scheme to one. We found higher order finite differences to be unstable.

Remark. A second order accurate stable scheme can be obtained by using (2.2.13) to replace the third order derivative at the boundary by the density, which can be found by solving the Kolmogorov forward equation (2.2.6) numerically.

PIDE in terms of matrix operations

If we take into account the finite difference approximations and quadrature rule for the integral, it is now possible to give a discretised PIDE, for a given triplet $(i,j,m) \in [\![0,N]\!] \times [\![0,M]\!] \times [\![0,P]\!]$ by

$$\begin{aligned}\delta_T \mathbf{u}^m_{i,j} + (r(T_m) - q(T_m))K_i \delta_K \mathbf{u}^m_{i,j} - \frac{1}{2}\left(\sigma^2_{|S,M}(K_i,B_j,T_m) - \frac{1}{2}\frac{\partial \sigma^2_{|S,M}(K_i,B_j,T_m)}{\partial B}\Delta_B\right)K_i^2 \delta_{KK}\mathbf{u}^m_{i,j} \\ + \frac{1}{2}\sigma^2_{|S,M}(B_j,B_j,T_m)B_j^2(B_j-K)^+\delta^-_{KKB}\mathbf{u}^m_{n_j,j} = \\ -\sum_{n=1}^{j-1}\frac{1}{2}K_i^2\delta_{KK}\mathbf{u}^m_{i,j}\frac{\partial \sigma^2_{|S,M}(K_i,B_n,T_m)}{\partial B}\Delta_B,\end{aligned} \tag{2.4.9}$$

and specify the coefficient matrices

$$\begin{aligned}\mathbf{A}^m_{.,j} &= (r(T_m)-q(T_m))\mathrm{diag}\Big(K_0,...,K_{n_j}\Big), \\ \mathbf{B}^m_{.,j} &= -\frac{1}{2}\mathrm{diag}\left(\left(\sigma^2_{|S,M}(K_i,B_j,T_m)K_i^2 - \frac{1}{2}K_i^2\frac{\partial\sigma^2_{|S,M}(K_i,B_j,T_m)}{\partial B}\Delta_B\right)_{i\in[\![0,n_j]\!]}\right), \\ \mathbf{C}^m_{.,j} &= -\frac{1}{2}\mathrm{diag}\Big(\Big(\sigma^2_{|S,M}(B_j,B_j,T_m)B_j^2(B_j-K_i)^+\Big)_{i\in[\![0,n_j]\!]}\Big).\end{aligned}$$

Remark. We can also approximate $\mathbf{B}^m_{.,j}$ further by Taylor expansion,

$$\mathbf{B}^m_{.,j} = -\frac{1}{2}\mathrm{diag}\left(\left(\sigma^2_{|S,M}\left(K_i,B_j-\frac{\Delta_B}{2},T_m\right)K_i^2\right)_{i\in[\![0,n_j]\!]}\right) + \mathcal{O}\Big(\Delta_B^2\Big),$$

which has a negative sign irrespective of the mesh size and does not alter the convergence order.

Under forward Euler time stepping, the complete scheme can be more compactly written as

$$\begin{aligned}\frac{\mathbf{u}^{m+1}_{.,j}-\mathbf{u}^m_{.,j}}{\Delta T} + \mathbf{L}^m_{.,j}\mathbf{u}^m_{.,j} &= \mathbf{f}^m_{.,j}, \\ \mathbf{L}^m_{.,j} &= \mathbf{A}^m_{.,j}\mathbf{D} + \mathbf{B}^m_{.,j}\mathbf{D_2} + \mathbf{C}^m_{.,j}\mathbf{\Phi}.\end{aligned} \tag{2.4.10}$$

Under θ-time-stepping, the scheme becomes

$$\Big(\mathbf{I}_{n_j} + \theta\Delta_T\mathbf{L}^{m+1}_{.,j}\Big)\mathbf{u}^{m+1}_{.,j} = \Big(\mathbf{I}_{n_j} - (1-\theta)\Delta_T\mathbf{L}^m_{.,j}\Big)\mathbf{u}^m_{.,j} + \theta\Delta_T\mathbf{f}^{m+1}_{.,j} + (1-\theta)\Delta_T\mathbf{f}^m_{.,j}.$$

This includes the second-order accurate Crank-Nicolson scheme for $\theta = 0.5$, which is used for our numerical computations.

This section shows that we can adapt the classical tools of finite difference methods in order to solve this PIDE. Each of the layers being a one-dimensional PDE, it is solved by successive roll-forward

performed by a Gaussian elimination. Although the Thomas algorithm cannot be applied directly since $\mathbf{L}^m_{\cdot,j}$ is not tri-diagonal (even after applying the necessary boundary conditions), a sparse Gaussian elimination will still be $\mathcal{O}(n)$ as the dense sub-matrix has a fixed number of columns.

Solution algorithm

We conclude with a summary of a possible algorithm to solve the PIDE numerically:

Algorithm 2.1 PIDE Discretisation

$\mathbf{u}^0_{\cdot,j} = \left((S_0 - K_i)^+ \mathbf{1}_{S_0<B_j}\right)_{i\in[\![0,N]\!]}$
$\mathbf{u}^m_{\cdot,0} = 0$
$\mathbf{f}^m_{\cdot,0} = 0$
for $(\ j = 0\,;\, j \le P\,;\, j++\ )$ **do**
  solve B_j-layer PDE for $\left(\mathbf{u}^m_{\cdot,j}\right)_{m\in[\![0,M]\!]}$:

$$\left(\mathbf{I}_{n_j} + \theta\Delta_T \mathbf{L}^{m+1}_{\cdot,j}\right)\mathbf{u}^{m+1}_{\cdot,j} = \left(\mathbf{I}_{n_j} - (1-\theta)\Delta_T \mathbf{L}^m_{\cdot,j}\right)\mathbf{u}^m_{\cdot,j} + \theta\Delta_T \mathbf{f}^{m+1}_{\cdot,j} + (1-\theta)\Delta_T \mathbf{f}^m_{\cdot,j}$$

  compute $\mathbf{f}^m_{\cdot,j+1}$ from $\mathbf{f}^m_{\cdot,j}$ and $\mathbf{u}^m_{\cdot,j}$ (for $j < P$)
end for

2.5 Numerical results

2.5.1 Pricing under the mimicking Brunick–Shreve model

For validation purposes of the forward equation, we will use a numerical solution of the backward pricing PDE in the Brunick–Shreve model. The backward PDE gives the price over a range of S and M and t, for fixed K, B and T.

Augmented state Feynman-Kac PDE

Recall that the spot diffusion of the underlying under $\mathbb{Q}$ is

$$\frac{dS_t}{S_t} = (r(t) - q(t))\,dt + \sigma(S_t, M_t, t)\,dW_t. \tag{2.5.1}$$

We note that the spot process is not Markovian anymore due to the dependence of the volatility on the running maximum. Hence, we cannot use the standard one dimensional Feynman-Kac PDE with Dirichlet boundary condition, but need to augment the state space even for barrier options (as well as

European options). Assume that strike K, barrier B and maturity T are all fixed. Then, following identical steps to the derivation by Shreve [135] in the Black-Scholes case, we get the following.

Proposition 2.5. *Under a Brunick–Shreve model (2.5.1), the up-and-out call price C defined in the region $\Omega_B = \{(x, y) : 0 < x < y, S_0 < y < B\}$ is the solution to the following initial boundary value problem:*

$$\begin{aligned}
\frac{\partial C}{\partial t} + (r(t) - q(t))x\frac{\partial C}{\partial x} + \frac{1}{2}x^2\sigma^2(x, y, t)\frac{\partial^2 C}{\partial x^2} - r(t)C = 0, &\qquad (x, y) \in \Omega_B,\ 0 < t < T, \quad (2.5.2)\\
C(0, y, t) = 0, &\qquad t \le T,\\
C(x, y, T) = (x - K)^+\mathbf{1}_{y<B}, &\qquad y > S_0,\\
\frac{\partial C(x, y, t)}{\partial y}\Big|_{x=y} = 0, &\qquad y > S_0,\\
C(x, B, t) = 0, &\qquad x > 0,\ t \le T.
\end{aligned}$$

Proof. We first note that the up-and-out call is a traded contract. As a consequence, the discounted price process $D(t)C_t$ is a martingale under the risk-neutral measure $\mathbb{Q}$. Hence

$$D(t)C_t = \mathbb{E}^{\mathbb{Q}}\Big[D(T)(S_T - K)^+\mathbf{1}_{M_T<B} \mid \mathcal{F}_t\Big] \qquad 0 \le t \le T.$$

Furthermore, as (S_t, M_t) is a Markovian vector, there exists a function v such that

$$D(t)v(x, y, t) = \mathbb{E}^{\mathbb{Q}}\Big[D(T)(S_T - K)^+\mathbf{1}_{M_T<B} \mid S_t = x,\ M_t = y\Big] \qquad \forall (x, y) \in \Omega_B,\ 0 \le t \le T.$$

We assume that v is smooth and belongs to the space $C^{2,1,1}(\Omega_B \times \mathbb{R}^+)$. Now write

$$C_t = v(S_t, M_t, t).$$

By the Itô-Doeblin lemma and recalling that the running maximum process M has finite variation, zero quadratic variation and zero cross-variation with S (see [135]),

$$d(D(t)C_t) \;=\; -r(t)D(t)C_t\,dt + D(t)\left[\frac{\partial v}{\partial t}\,dt + \frac{\partial v}{\partial x}\,dS_t + \frac{\partial v}{\partial y}\,dM_t + \frac{1}{2}\frac{\partial^2 v}{\partial x^2}\,d[S]_t\right].$$

The process $D(t)C_t$ is a martingale if and only if

$$\begin{aligned}
\frac{\partial v}{\partial t} - r(t)v + (r(t) - q(t))x\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2x^2\frac{\partial^2 v}{\partial x^2} &= 0, \quad (x, y) \in \Omega_B,\ 0 < t < T, \qquad (2.5.3)\\
\frac{\partial v}{\partial y}\Big|_{x=y} &= 0, \quad y > S_0.
\end{aligned}$$

This gives the desired backward PDE. We note that the boundary condition is equivalent to

$$\frac{\mathrm{d}v(y,y,t)}{\mathrm{d}y} = \frac{\partial v(x,y,t)}{\partial x}\Big]_{x=y}. \tag{2.5.4}$$

□

Finite difference approximation

We briefly explain the numerical solution of (2.5.2).

We first notice that (2.5.2) is a classical Black-Scholes PDE (where the volatility is a function of S, a "parameter" M, and t), and derivatives with respect to y only enter via the Neumann boundary condition (on the diagonal $x = y$). This means that for a fixed y, we have to solve a Black-Scholes type PDE on $[0, y] \times \{y\}$. From now on, the one-dimensional PDE for a given level of y will be referred to as a "PDE layer".

If $y = B$, then $C(x, y, t) = 0$ for all x. We can then use the Neumann boundary condition to bootstrap backwards in y from B to S_0. To illustrate the idea, consider now $y = B - \Delta_y$ for a mesh size Δ_y, then

$$\frac{\partial C(x,y,t)}{\partial y}\Big]_{x=y} = \frac{C(y, y+\Delta_y, t) - C(y,y,t)}{\Delta y} + \mathcal{O}(\Delta_y) = 0 \implies C(y,y,t) = C(y, y+\Delta_y, t) + \mathcal{O}\Big(\Delta_y^2\Big).$$

We can build the next PDE layer at y using an approximate Dirichlet boundary condition at $x = y$, $C(y, y, t) \approx C(y, y + \Delta_y, t)$ (here, 0). Subsequent PDE layers can be constructed similarly by backward reasoning. We will describe a higher-order version below. The solution in each layer will then depend on layers with greater y via this boundary condition. The premium value is retrieved from $C(S_0, S_0, 0)$.

We now describe the construction of meshes of the form $(x_j(y), y)_{j=1,\ldots,N_x(y)}$ for layer y, as well as of the spacing of points in the y-direction. The best accuracy was achieved numerically with a refined running maximum grid close to S_0 and, for each y, a uniformly spaced spot grid (except for nodes close to the diagonal, as detailed below).

Denote by Δ_y a desired target mesh width. Then for the y-mesh we choose $N_y = \left\lceil \frac{B-S_0}{\Delta_y} \right\rceil$ points with an exponential grading as defined in [148], so that $\forall\, i \in [\![0, N_y]\!]$:

$$\begin{aligned} y_i &= (S_0 - \theta) + \theta \, \exp(\lambda z_i), \\ \lambda &= 2, \\ \theta &= \frac{B - S_0}{e^\lambda - 1}, \\ z_i &= \frac{i}{(N_y + 1)}. \end{aligned}$$

The construction of the mesh, illustrated in Figure 2.5.1, ensures that y_i lies on the grid of the PDE layer of levels y_{i+1} and y_{i+2}, and hence no interpolation is needed to retrieve the boundary condition for the next PDE layer as described in the next paragraph.

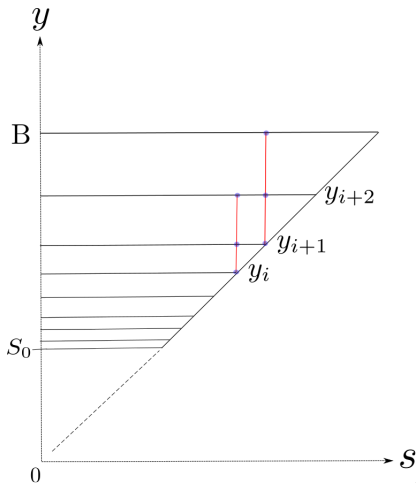


Figure 2.5.1: Mesh construction for backward equation. The spot grid is recomputed for every PDE layer to be $(x_j(y_i), y_i)$, where $x_j = j\Delta_x$, $j < N_x - 2$, $x_{N_x-j} = y_{i-j}$, $j \in \{0, 1, 2\}$, $\Delta_x = y_{i-2}/(N_x - 2)$ and $N_x = \lceil y_i/\Delta_y \rceil$ to keep the spacing approximately constant.

In order to increase accuracy from the numerical boundary condition described at the start of this section, we use second order Taylor expansion in y-direction instead of a simple backward finite difference for all y_i with $i < N_y - 1$. The bootstrap idea stays the same. Indeed, write $c(y) = c(y_i, y, t)$ where y_i is a certain given level of the y-discretisation and $c'(y) = \frac{\partial c(y_i,y,t)}{\partial y}$. We consider a Taylor expansion around $y = y_i$: $c(y) = c(y_i) + \frac{1}{2}(y - y_i)^2 c''(y_i) + o((y - y_i)^2)$ (as $c'(y_i) = 0$ by the boundary condition). Using this at points y_{i+1} and y_{i+2}, we get $c(y_i) = c(y_{i+2}) + \frac{(y_{i+2}-y_i)^2}{(y_{i+2}-y_i)^2-(y_{i+1}-y_i)^2}(c(y_{i+1}) - c(y_{i+2}))$. This is then used as a boundary value at y_i. The first layer at $N_y - 1$ is treated with a first order backward difference as described previously.

2.5.2 Numerical validation

Our numerical validation consists in pricing a set of up-and-out call options for different strikes, maturities and barriers with:

1. the Forward PIDE (2.4.1) and one numerical solution for the whole set of deal parameters;
2. the Backward Feynman-Kac PDE (2.5.2) and as many solutions as triplets of deal parameters.

The goal is to make them match with about one basis point tolerance.

The Brunick–Shreve volatility is generated with an SVI parametrisation in both barrier and strike dimension (details can be found in [68]) defined as

$$\sigma(x,y,t) = \frac{1}{2}\left(\sigma_{SVI}\left(\log\left(\frac{x}{S_0}\right),t+1\right)+\sigma_{SVI}\left(\log\left(\frac{y}{S_0}\right),t+1\right)\right), \tag{2.5.5}$$

$$\sigma_{SVI}(k,t) = \frac{\sqrt{a+b\left(\rho(k-m)+\sqrt{(k-m)^2+\sigma^2}\right)}}{\sqrt{t}}.$$

The Brunick–Shreve volatility surface has a shape as in Figure 2.5.2.

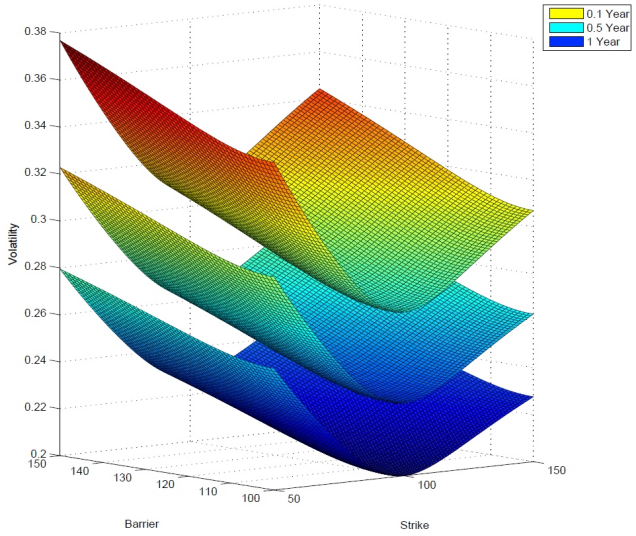


Figure 2.5.2: Assumed Brunick–Shreve volatility surface as per (2.5.5), where $a = 0.04, b = 0.2, \sigma = 0.2, \rho = m = 0$, and $T \in \{0.1, 0.5, 1\}$.

The spot is $S_0 = 100$, the risk-free rate is $r = 0.1$ and the dividend yield is $q = 0.05$.

We use the numerical schemes as described in 2.4.2 and 2.5.1 for the PIDE and PDE solution, respectively, with Crank-Nicolson time stepping. The discretisation parameters are $N = P = N_y = 1000$ space steps (with N_x adjusted as described) and $M = 1000$ time steps.

We compare prices for (K, B, T) covering the set $[0, 120] \times [100, 120] \times \{1\}$ with 120 points in strike and 40 points in barrier levels, see Table 2.5.1. The error is computed as the relative error if the net present value (NPV) is above one and as absolute error otherwise.

Average Difference	Maximum Difference
4.6e-5	3.5e-4

Table 2.5.1: Difference between forward and backward solutions over strikes between 0 and 120, and up-and-out barriers between 100 and 120, all for maturity 1.

We can analyse more precisely the behaviour for a few barrier levels. For example, Figure 2.5.3 shows the difference as a function of strike. The associated values for a barrier fixed at 120 are in Table 2.5.2.

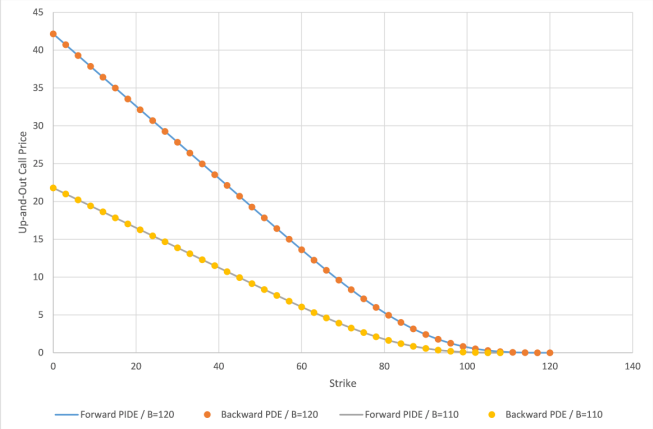


Figure 2.5.3: Forward PIDE vs Backward PDE: up-and-out call option prices with Brunick–Shreve volatility, $S = 100$ and $T = 1$.

Strike	Forward PIDE	Backward PDE	Rel. Diff.
0	42.1486	42.1486	1e-06
9	37.8567	37.8568	1e-06
18	33.5649	33.5650	2e-06
27	29.2731	29.2732	2e-06
36	24.9815	24.9815	3e-06
45	20.6928	20.6929	3e-06
54	16.4263	16.4264	2e-06
63	12.2536	12.2535	5e-06
72	8.3438	8.3436	2e-05
81	4.9680	4.9677	6e-05
90	2.4170	2.4168	9e-05
99	0.8472	0.8472	9e-07
108	0.1546	0.1547	1e-04
117	0.0023	0.0023	8e-05
120	0.0000	0.0000	0

Table 2.5.2: Results and relative errors for different strikes; barrier = 120, forward = 105.13.

The results match with the desired accuracy.

2.6 Conclusions

In this chapter, we provided a new and numerically effective way to work with observed barrier option prices to determine the Brunick–Shreve Markovian projection of a stochastic variance of a semi-martingale S with running maximum M onto (S_t, M_t). The present Dupire-type formula, which provides similar advantages (and disadvantages) as the standard Dupire approach for vanillas, was then re-arranged to obtain a forward PIDE, which is convenient to control the numerical stability of best fit algorithms as in [42], and we present a possible approach in Chapter 4. There is a well-known literature on the calibration of vanilla options through Markovian projection for LSV models (see Guyon and Labordère [82] and Ren, Madan and Qian [129]). We believe that extending these methods

combined with the Forward PIDE we presented in this chapter will lead to novel calibration algorithms for barrier optionsas we discuss in Chapter 4.

3 Calibration of a Four-Factor Hybrid LSV Model with a New Control Variate Particle Method

3.1 Introduction

In order to improve the pricing and hedging of foreign exchange (FX) options, we introduce stochastic domestic and foreign short interest rates into the model. Assuming constant interest rates is appealing due to its simplicity and does not lead to serious mispricing of options with short maturities. However, empirical results [145] have confirmed that the constant interest rate assumption is inappropriate for long-dated FX products, and the effect of interest rate volatility can be as relevant as that of the FX rate volatility for longer maturities. Extensive research has been carried out in the area of option pricing with stochastic volatility and interest rates in the past few years. Van Haastrecht et al. [145] extended the model of Schöbel and Zhu [134] to currency derivatives by including stochastic interest rates, a model that benefits from analytical tractability even in a full correlation setting due to the processes being Gaussian. On the other hand, Ahlip and Rutkowski [4], Grzelak and Oosterlee [77] and Van Haastrecht and Pelsser [146] examined the Heston–CIR/Vasicek hybrid models and concluded that they give rise to non-affine models even under a partial correlation structure of the driving Brownian motions.

In the foreign exchange market, the inclusion of stochastic interest rates results in a 4-factor model associated with a complex calibration routine. A few papers discuss this problem under simpler settings. Deelstra [45] and Clark [35] mainly consider 3-factor hybrid local volatility models and focus on the theoretical rather than the practical aspects of the calibration, whereas Stoep, Grzelak and Oosterlee [143] consider an application to a 2-factor hybrid local volatility. In [82], Guyon and Labordère discuss an application of Monte Carlo-based calibration methods to a 3-factor LSV equity model with stochastic domestic rate and discrete dividends.

The model of Cox et al. [39] is very popular when modeling short rates because the (square-root) CIR process admits a unique strong solution, is mean-reverting and analytically tractable. Cox et al. found the conditional distribution to be non-central chi-squared, on the basis of which Broadie and Kaya [22] proposed an efficient exact simulation scheme for the (square-root) process using acceptance-rejection sampling. As of late, the non-negativity of the CIR process is considered to be less desirable when modeling short rates. On one hand, central banks have significantly reduced the interest rates since the 2008 financial crisis and it is now commonly accepted that interest rates need not be positive. On

the other hand, if interest rates dropped too far below zero, then large amounts of money would be withdrawn from banks and government bonds, putting a severe squeeze on deposits. Hence, we model the domestic and foreign short rates using the shifted CIR (CIR++) process of Brigo and Mercurio [20]. The CIR++ model allows the short rates to become negative and can fit any observed term structure exactly while preserving the analytical tractability of the original model for bonds, caps, swaptions and other basic interest rate products.

According to Ren et al. [129], Tian et al. [140], Van der Stoep et al. [142] and Guyon and Labordère [82], the general LSV model allows for a better calibration to European options and improves the pricing and risk-management performance when compared to pure local volatility or pure stochastic volatility models. We focus on the Heston-type LSV model because the Cox–Ingersoll–Ross (CIR) process for the variance is widely used in the industry due to its desirable properties, such as mean-reversion and non-negativity, and since semi-analytic formulae are available for calls and puts under Heston's model and can help calibrate the parameters easily. The local volatility component allows a perfect calibration to the market prices of vanilla options. At the same time, the stochastic volatility component already provides built-in smiles and skews which give a rough fit, so that a leverage function relatively close to one suffices for a perfect calibration.

In this work, we consider the 4-factor hybrid LSV model defined below in (3.2.1), which is a Heston-type LSV model with two shifted CIR short-rate processes that accounts for negative rates and provides alternative modeling of rates in lieu of the more classical Hull–White process. We extend the calibration condition for the leverage function given in [82] for a 3-factor LV–2Hull–White model and in [45] for a 4-factor Heston-type LSV–2Hull–White model to the 4-factor LSV–2CIR++ model, and give a novel and rigorous derivation.

We propose a new calibration approach that builds on the particle method of [82], and combine it with a novel and efficient variance reduction technique that takes advantage of PDE calibration to increase its stability and accuracy while keeping the dimensionality under control with Monte Carlo.

The proposed variance reduction techniques allow to evaluate each expectation with a small amount of particles/paths (we find that $4000 - 20000$ particles are sufficient in practice). Our numerical experiments suggest that this method almost recovers the calibration speed from the corresponding 2-factor LSV model with deterministic rates defined in (3.2.2).

As a side result, we explain how to effectively deal with the broken Feller condition for the Heston-type LSV Kolmogorov forward *Partial Differential Equation* (PDE) and numerically solve the PDE using the finite element method with a *Backward Differentiation Formula* (BDF) time-stepping scheme and appropriate non-Dirichlet boundary condition. To the best of our knowledge, this represents a new approach that complements the literature on the use of ADI schemes [35, 129, 152] to handle the PDE calibration of an LSV model with deterministic rates. Moreover, a main advantage of the finite element method compared to ADI schemes is that the Dirac delta initial condition can be handled naturally in the weak formulation. This framework yields an accurate calibration of an LSV with

deterministic rates for a broad set of market data. The idea we propose in this chapter is to use the 2-factor LSV model calibrated via the PDE as a "control variate" (generalising this concept from the standard setting to conditional expectations) for a 4-factor LSV model with stochastic rates calibrated via a Monte Carlo particle method as outlined above.

Furthermore, we rely on the forward Dupire PDE to compute the second order derivative of the call price with respect to the strike, which appears in the calibration formula (similar to the Dupire formula), instead of extracting this quantity from the market. By doing so, we prevent instabilities and get a more accurate estimate of the leverage function, which is only possible through a proper calibration of a pure local volatility model that we will also discuss.

Finally, we believe that including stochastic rates is of paramount importance to price some specific exotic derivatives such as the Accumulator, the *Power Reverse Dual-Currency note* (PRDC) or the *AutoCallable Barrier Dual-Currency Note* (ABDC). In particular, in 3.6, we consider the pricing problem for the last contract and show the impact of stochastic rates on the net present value. Both the AutoCallable and the Accumulator can embed a barrier feature that will benefit from a local-stochastic volatility component. Indeed, pure stochastic volatility models tend to under-estimate the knock-out probability while pure local volatility models over-estimate it. Therefore, a combination of both local and stochastic volatility provides the best trade-off for which a mixing factor can eventually be included [35]. Moreover, stochastic rates become necessary for any hybrid product that embeds the rates explicitly. One example is the spread option between the performance of the FX rate and the Libor rate.

The remainder of this chapter is organised as follows. In Section 3.3, we provide a necessary and sufficient condition for a perfect calibration to vanilla quotes and prove this rigorously emphasising the use of local times and possible moment explosions. In Section 3.4, we present our approach to the calibration problem and describe the process step-by-step. First, we calibrate a pure local volatility model (3.2.3) with the forward Dupire PDE. Second, we propose to calibrate a LSV model with deterministic rates using the finite element method. Furthermore, we employ a change-of-variable technique to tackle the broken Feller condition efficiently. In Section 3.5, we combine the calibrated 2-factor LSV model (3.2.2) and variance reduction techniques to calibrate the 4-factor LSV model (3.2.1) with stochastic rates via the particle method. We present numerical results and show that a low number of particles suffice to provide a very good fit to market quotes, which suggests that the method is computationally efficient. Section 3.6 present some pricing results for an ABDC contract and Section 4.11 concludes with a brief discussion.

3.2 Model definition

We consider a domestic and a foreign market with risk-free short rates r^d and r^f. We denote by D^d and D^f the domestic and foreign discount factors associated with their respective money market

accounts. Throughout, T and t are fixed positive times such that $t < T$. We assume the existence of a filtered probability space $(\chi, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q}^d)$ with a domestic risk-neutral measure $\mathbb{Q}^d$, under which the price process of a risky asset follows a system of SDEs

$$\begin{cases} \dfrac{dS_t}{S_t} = \left(r_t^d - r_t^f\right) dt + \alpha(S_t, t)\sqrt{V_t}\, dW_t \\ r_t^d = g_t^d + h^d(t) \\ r_t^f = g_t^f + h^f(t) \\ dg_t^d = \kappa_d\left(\theta_d - g_t^d\right) dt + \xi_d\sqrt{g_t^d}\, dW_t^{g^d} \\ dg_t^f = \left(\kappa_f\left(\theta_f - g_t^f\right) - \rho_{Sf}\xi_f\sqrt{g_t^f}\,\alpha(S_t, t)\sqrt{V_t}\right) dt + \xi_f\sqrt{g_t^f}\, dW_t^{g^f} \\ dV_t = \kappa(\theta - V_t)\, dt + \xi\sqrt{V_t}\, dW_t^V \end{cases} \tag{3.2.1}$$

with the correlation structure

$$d\langle W_t, W_t^V\rangle \;=\; \rho\, dt, \qquad d\langle W_t, W_t^{g^d}\rangle \;=\; \rho_{Sd}\, dt, \qquad d\langle W_t, W_t^{g^f}\rangle \;=\; \rho_{Sf}\, dt,$$

with $\rho, \rho_{Sd}, \rho_{Sf} \in (-1, 1)$, the other correlations being zero, and for given functions $\alpha : \mathbb{R}^+ \times [0, T] \to \mathbb{R}^+$, $h^{d/f} : [0, T] \to \mathbb{R}$, and non-negative numbers $\kappa, \theta, \xi, \kappa_d, \theta_d, \xi_d, \kappa_f, \theta_f, \xi_f$, as well as initial values S_0, g_0^d, g_0^f, V_0.

The domestic and foreign discount factors are then given by

$$D_t^d = \mathrm{e}^{-\int_0^t r_u^d du}, \quad D_t^f = \mathrm{e}^{-\int_0^t r_u^f du},$$

respectively.

We also define two simpler models that we will refer to in the remainder of the chapter. First, we can write the related 2-factor Heston-type LSV model with deterministic rates as

$$\begin{cases} \dfrac{dS_t^{2D}}{S_t^{2D}} = \left(\bar{r}^d(t) - \bar{r}^f(t)\right) dt + \alpha^{2D}\left(S_t^{2D}, t\right)\sqrt{V_t}\, dW_t,\ S_0^{2D} = S_0, \\ dV_t^{2D} = \kappa\left(\theta - V_t^{2D}\right) dt + \xi\sqrt{V_t^{2D}}\, dW_t^V \end{cases} \tag{3.2.2}$$

for a given function $\alpha^{2D} : \mathbb{R}^+ \times [0, T] \to \mathbb{R}^+$, and the pure *Local volatility* (LV) model as

$$\frac{dS_t^{LV}}{S_t^{LV}} = \left(\bar{r}^d(t) - \bar{r}^f(t)\right) dt + \sigma_{LV}\left(S_t^{LV}, t\right) dW_t,\ S_0^{LV} = S_0, \tag{3.2.3}$$

where $\bar{r}^d(t) = -\frac{\partial \ln P^d(0,t)}{\partial t}$ and $\bar{r}^f(t) = -\frac{\partial \ln P^f(0,t)}{\partial t}$, with $P^{d/f}(0, T)$ the market zero coupon bond prices for the domestic and foreign money market accounts, respectively, and a given function $\sigma_{LV} : \mathbb{R}^+ \times [0, T] \to \mathbb{R}^+$. Note that while the volatility is 'local', i.e., a function of spot FX and time,

the short rates are assumed to be a function of time only. We also note for future reference that under the pure LV model (3.2.3) call prices C_{LV} satisfy the forward Dupire PDE

$$\frac{\partial C_{LV}}{\partial T} - \frac{1}{2}\sigma_{LV}(K,T)^2 K^2 \frac{\partial^2 C_{LV}}{\partial K^2} + K\Big(\bar{r}^d(T) - \bar{r}^f(T)\Big)\frac{\partial C_{LV}}{\partial K} + \bar{r}^f(T) C_{LV} = 0\,. \tag{3.2.4}$$

The purpose of this chapter is to calibrate h^d, h^f, κ_d, κ_f, θ_d, θ_f, ξ_d, ξ_f, κ, θ, ξ, ρ and especially α in (3.2.1). We will use calibration of (3.2.2) and (3.2.3) as "stepping stones". In order to illustrate the full calibration process from start to finish and to make the structure clear to the reader, we provide a flowchart of the calibration in Figure 3.2.1.

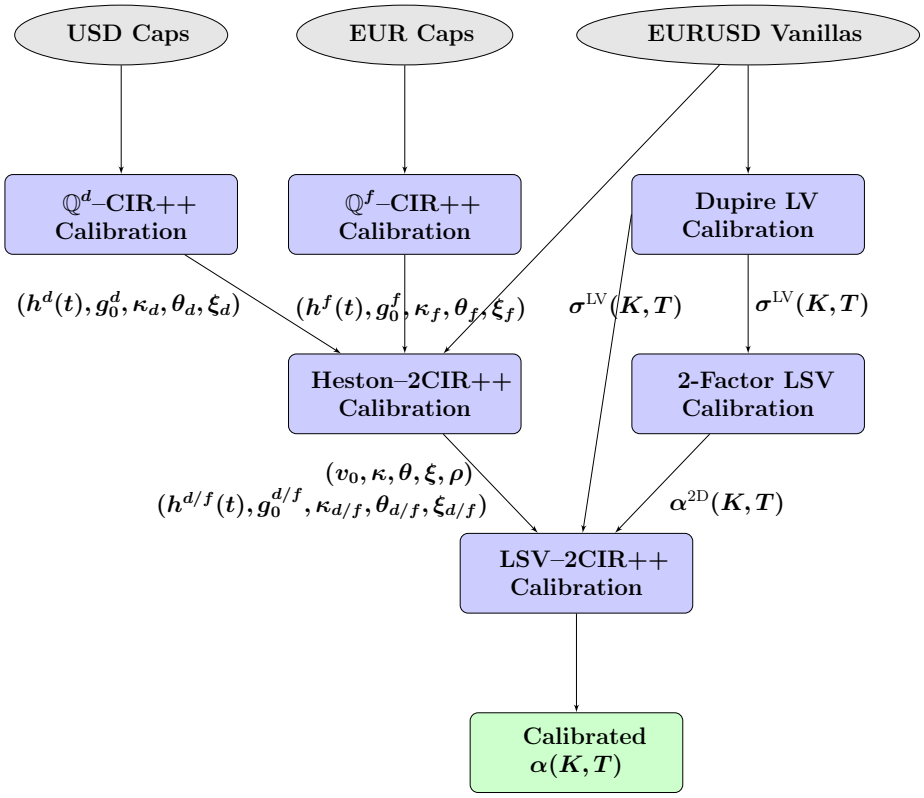


Figure 3.2.1: Full calibration routine flowchart

3.3 A necessary and sufficient condition for exact calibration

In this section, we derive the main formula that links market call prices, via the Dupire local volatility, to call prices under model (3.2.1).

The spot S_T is associated with the currency pair ccy1ccy2 (following the notations in [35]) and denotes the amount of units of ccy2 (domestic currency) needed to buy one unit of ccy1 (foreign currency) at

time T.

Let the discounted call option price under model (3.2.1) for a notional of one unit of *ccy*1 be

$$C(K,T) = \mathbb{E}^{\mathbb{Q}^d}\left[D_T^d (S_T - K)^+\right].$$

The owner of a long call contract, for a notional amount N_{ccy1}, has the right to buy N_{ccy1} in ccy1 in exchange for an amount $N_{ccy2} = KN_{ccy1}$ in ccy2 at expiry. The payout is $(N_{ccy1}S_T - N_{ccy2})^+$ and its discounted price is $N_{ccy1}C(K,T)$. The market standard is to quote call options in ccy2.

We can now state the following result, which we will prove in the remainder of this section.

Theorem 3.1. *Assume that α is uniformly bounded by α_{max}, $h^{d,f}$ are uniformly bounded and that both the marginal density $\phi(T,\cdot)$ of S_T in (3.2.1) and $\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d V_T \,|\, S_T = \cdot\right]$ exist and are continuous. Then, the call price $C(K,T)$ under model (3.2.1) matches the price C^{LV} under the local volatility model (3.2.3) for any strike $K > 0$ and maturity $T < T^*$ if and only if, for all $K, T > 0$,*

$$\alpha^2(K,T) = \frac{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \,|\, S_T = K\right]}{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d V_T \,|\, S_T = K\right]}\left(\sigma_{LV}(K,T)^2 + q(K,T)\right), \tag{3.3.1}$$

where σ_{LV} is a local volatility as in (3.2.3),

$$\begin{aligned} q(K,T) &= \frac{\mathbb{E}^{\mathbb{Q}^d}[Q_T]}{\frac{1}{2}K^2\frac{\partial^2 C_{LV}}{\partial K^2}}, \\ Q_T &= D_T^d\left(r_T^f - \bar{r}^f(T)\right)(S_T - K)^+ - KD_T^d \mathbf{1}_{S_T \geq K}\left[\left(r_T^d - \bar{r}^d(T)\right) - \left(r_T^f - \bar{r}^f(T)\right)\right], \end{aligned} \tag{3.3.2}$$

and, denoting $\varphi = 2 + \sqrt{2}$ and $\zeta = \xi\alpha_{max}$,

$$\begin{cases} T^* = \frac{2}{\sqrt{\varphi^2\zeta^2 - \kappa^2}}\left[\frac{\pi}{2} + \arctan\left(\frac{\kappa}{\sqrt{\varphi^2\zeta^2 - \kappa^2}}\right)\right], & \text{if } \kappa < \varphi\zeta, \\ T^* = \infty, & \text{if } \kappa \geq \varphi\zeta. \end{cases} \tag{3.3.3}$$

Remark. The ratio on the right-hand side of (3.3.1) accounts for the stochastic volatility, such that if there is no stochastic volatility (i.e. $V_T = 1$), we recover the formula in [35]. The term q in (3.3.2) accounts for the stochasticity of rates and, if rates are deterministic, $q = 0$ and we recover the formula derived in [129, 82]

$$\alpha^{2D}(K,T) = \frac{\sigma_{LV}(K,T)}{\sqrt{\mathbb{E}^{\mathbb{Q}^d}\left[V_T^{2D} \,|\, S_T^{2D} = K\right]}}. \tag{3.3.4}$$

At time $T = 0$, since by construction we have $r_0^d - \bar{r}^d(0) = r_0^f - \bar{r}^f(0) = 0$, the leverage function is $\alpha(K,0) = \frac{\sigma_{LV}(K,0)}{\sqrt{v_0}}$.

An alternative form of (3.3.1) can be obtained by subtracting $\alpha(K,T)\frac{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d V_T \,|\, S_T = K\right]}{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \,|\, S_T = K\right]}\times$ (3.2.4) from

$\sigma^2_{LV}(K,T)\times$ (3.3.11) , namely

$$\begin{aligned}\alpha^2(K,T) &= \sigma^2_{LV}(K,T)\frac{\mathbb{E}^{\mathbb{Q}^d}\left[D^d_T \,|\, S_T = K\right]}{\mathbb{E}^{\mathbb{Q}^d}\left[D^d_T V_T \,|\, S_T = K\right]} \\ &\times \left(\frac{\frac{\partial C}{\partial T} + \mathbb{E}^{\mathbb{Q}^d}\left[D^d_T r^f_T (S_T - K)^+\right] - \mathbb{E}^{\mathbb{Q}^d}\left[D^d_T \mathbf{1}_{S_T \geq K} K\left(r^d_T - r^f_T\right)\right]}{\frac{\partial C}{\partial T} + \bar{r}^f(T) C_{LV} + K(\bar{r}^d(T) - \bar{r}^f(T))\frac{\partial C_{LV}}{\partial K}}\right).\end{aligned} \tag{3.3.5}$$

If we divide the denominator by σ^2_{LV} and re-arrange, we recover the formula presented in [45] for an LSV–2Hull–White process, namely

$$\alpha^2(K,T)\frac{\mathbb{E}^{\mathbb{Q}^d}\left[D^d_T V_T \,|\, S_T = K\right]}{\mathbb{E}^{\mathbb{Q}^d}\left[D^d_T \,|\, S_T = K\right]} = \frac{\frac{\partial C}{\partial T} - \mathbb{E}^{\mathbb{Q}^d}\left[D^d_T\left(K r^d_T - r^f_T S_T\right)\mathbf{1}_{S_T > K}\right]}{\frac{1}{2}K^2\frac{\partial^2 C_{LV}}{\partial K^2}}.$$

However, we believe practitioners are more familiar with (3.3.1) as derived in [35] for a 3-factor LV–2Hull–White process, which is why we use the main formulation (3.3.1) in the remainder of the chapter.

Theorem 3.1 generalises the formula presented in [82] to stochastic foreign rates. Here, we consider an LSV–2CIR++ model and derive our result in a rigorous fashion, emphasising the importance of the local time process. Furthermore, in Lemma 3.4 we provide a sufficient condition for the process

$$\int_0^t \mathbf{1}_{S_u \geq K} D^d_u S_u \alpha(S_u, u)\sqrt{V_u}\, dW_u$$

to be a true martingale up to T^*, which is an important step in the proof of Theorem 3.1 that is often omitted in the literature. On the one hand, T^* is a lower bound for the explosion time of the second moment of the discounted spot process $D^d_t S_t$. On the other hand, [8] show that the moment explodes in finite time for the Heston model, a property that is inherited by our Heston-type LSV–2CIR++ model (3.2.1) as well as the Heston-type LSV–2Hull–White in [45]. Therefore, the formula may not hold for some values of the model parameters and for large maturities T. However, in practice, T^* is very large. For instance, following our calibration routine in Section 3.5 we obtain

$$\kappa = 1.4124,\ \xi = 0.2988,\ \alpha_{max} = 1.40,$$

such that

$$T^* = 28.6.$$

An issue may arise for large values of ξ. Indeed, a numerical experiment in [82] raises a question on the existence of a calibrated 2-factor LSV model when $\xi \approx 350\%$ (high values of the vol-of-vol are used to match forward smiles). In this particular case and with the other model parameters kept the same, we find $T^* = 0.20$, which indicates that moment explosions may occur sooner.

As an aside, we note that the continuity of the joint density $\psi(T,\cdot,\cdot,\cdot)$ of $\left(S_T, V_T, D_T^d\right)$ in (3.2.1) is sufficient for $\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d V_T \,|\, S_T = \cdot\right]$ to be continuous in the assumptions of Theorem 3.1.

We first state and prove two auxiliary lemmas necessary for the derivation of Theorem 3.1.

Lemma 3.2. *Given a filtered probability space $(\chi, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q}^d)$, let W be a standard Brownian motion and μ, Y two $\mathcal{F}_t$-adapted processes with Y continuous and of finite second moment and*

$$\int_0^t \left(|\mu_u| + Y_u^2\right) du < \infty\,.$$

Consider a continuous Itô process X given by

$$X_t \;=\; X_0 + \int_0^t \mu_u\, du + \int_0^t Y_u\, dW_u$$

whose marginal density function $\phi(t,\cdot)$ and $\mathbb{E}^{\mathbb{Q}^d}\left[D_t Y_t^2 \,|\, X_t = \cdot\right]$ are assumed to be continuous. Further denote by $(l_t^a)_{t\geq 0}$ the local time of X at level a. Then, for any continuous, integrable and positive $\mathcal{F}_t$-adapted semi-martingale D and any real number a,

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t D_u\, dl_u^a\right] = \int_0^t \left(\mathbb{E}^{\mathbb{Q}^d}\left[D_u Y_u^2 \,|\, X_u = a\right]\phi(a,u)\right) du\,.$$

In order to prove Lemma 3.2, we first state a necessary auxiliary result, which is an adaptation of the Itô-Tanaka formula proof in [132], where the integrand D in the local time integral is 1.

Proposition 3.3. *On a filtered probability space $(\chi, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q}^d)$, let D and X be two $\mathcal{F}_t$-adapted continuous semi-martingales, with D positive and integrable, $(l_t^a)_{t\geq 0}$ the local time of X at level a and, for all $n > 0$,*

$$\delta_n^a(x) = \begin{cases} 0, & |x-a| > \frac{1}{n}, \\ \frac{n}{2}, & |x-a| \leq \frac{1}{n}\,. \end{cases} \tag{3.3.6}$$

Then, for any horizon time $T > 0$ and $0 < t \leq T$, $\int_0^t D_s \delta_n^a(X_s)\, d\langle X\rangle_s$ converges almost surely, and uniformly in time, to $\int_0^t D_s\, d_s l^a$.

Proof. We closely follow Section 45 of [132], but use a more concrete expression of the regularisation function. The local time at level a is defined as a continuous adapted increasing process such that

$$|X_t - a| - |X_0 - a| = \int_0^t sgn(X_s - a)\, dX_s + l_t^a\,, \tag{3.3.7}$$

with $sgn(x) = -1$ for $x \leq 0$ and $sgn(x) = 1$ for $x > 0$. We define a sequence of functions $(f_n)_{n\geq 0}$ for

all $x \in \mathbb{R}$, as in Chapter 4 of [117]

$$f_n(x) = \begin{cases} |x-a|, & |x-a| > \frac{1}{n}, \\ \frac{1}{2}\Big(\frac{1}{n} + n(x-a)^2\Big), & |x-a| \leq \frac{1}{n}\,. \end{cases}$$

Hence, for all $n > 0$, $\frac{1}{2}f_n^{''} = \delta_n^a$ a.e. We recall from the proof of Tanaka's formula (see [132] for more details) that $\int_0^t \delta_n^a(X_s)\, d\langle X\rangle_s$ converges almost surely to l_t^a (uniformly in t). The sequence f_n converges uniformly to $x \to |x-a|$ and $f_n^{'}$ converges point-wise to $sgn(x-a)$ such that by the Itô-Doeblin formula we can write

$$\begin{aligned} D_t f_n(X_t) - D_0 f_n(X_0) &= \int_0^t f_n(X_s)\, dD_s + \int_0^t D_s f_n^{'}(X_s)\, dX_s \\ &+ \frac{1}{2}\int_0^t D_s f_n^{''}(X_s)\, d\langle X\rangle_s + \int_0^t d\langle D, f_n(X)\rangle_s. \end{aligned} \tag{3.3.8}$$

We denote

$$C_t^n = \frac{1}{2}\int_0^t D_s f_n^{''}(X_t)\, d\langle X\rangle_s$$

and, since $f_n^{''}(x) = 0$ for any x such that $|x-a| \geq \frac{1}{n}$, we have

$$\int_0^t \mathbf{1}_{|X_s - a| > \frac{1}{n}}\, dC_s^n = 0\,.$$

Also, from the definition of f_n, for all $x \in \mathbb{R}$,

$$sgn(x-a) - f_n^{'}(x) = \begin{cases} 0, & |x-a| \geq \frac{1}{n}, \\ sgn(x-a) - n(x-a), & |x-a| < \frac{1}{n}, \end{cases}$$

and then for any given x, and $n > 0$, both $D_t\Big|sgn(x-a) - f_n^{'}(x)\Big|$ and $D_t||x-a| - f_n(x)|$ are smaller than D_t which is integrable. Let $X_t = X_0 + M_t + A_t$ be the canonical decomposition of X and $D_t = D_0 + N_t + R_t$ the canonical decomposition of D. Localisation allows us to reduce the problem to the case where M and N are bounded and A and R are of bounded variation. Then,

$$\left\|\int_0^T D_s\Big(sgn(X_s-a) - f_n^{'}(X_s)\Big)\, dM_s\right\|_2^2 = \mathbb{E}\left[\int_0^T \Big(D_s\Big(sgn(X_s-a) - f_n^{'}(X_s)\Big)\Big)^2\, d\langle M\rangle_s\right],$$

for which the right-hand-side goes to zero when n goes to infinity by dominated convergence theorem. By Doob's L^2 martingale inequality, we can write

$$\left\|\sup_{t\in[0,T]}\left|\int_0^t D_s\Big(sgn(X_s-a) - f_n^{'}(X_s)\Big)\, dM_s\right|\right\|_2 \leq 2\left\|\int_0^T D_s\Big(sgn(X_s-a) - f_n^{'}(X_s)\Big)\, dM_s\right\|_2,$$

to conclude that

$$\sup_{t\in[0,T]}\left|\int_0^t D_s\Big(sgn(X_s-a)-f_n^{'}(X_s)\Big)\,dM_s\right|\to 0\,, \tag{3.3.9}$$

in L^2 and in probability. We may then assume that (3.3.9) also holds almost surely, since, we could work with a subsequence on n for which the statement is true instead). Similarly, we have

$$\sup_{t\in[0,T]}\left|\int_0^t (|X_t-a|-f_n(X_s))\,dN_s\right|\to 0\,a.s.$$

Also,

$$\begin{aligned}\left|\int_0^t D_s\Big(sgn(X_s-a)-f_n^{'}(X_s)\Big)\,dA_s\right| &\leq \int_0^t D_s\Big|\Big(sgn(X_s-a)-f_n^{'}(X_s)\Big)\Big|\,|dA_s|\\ &\leq \int_0^t D_s|\,dA_s|\,,\end{aligned}$$

and

$$\sup_{t\in[0,T]}\left|\int_0^t D_s\Big(sgn(X_s-a)-f_n^{'}(X_s)\Big)\,dA_s\right|\leq\int_0^T D_s|\,dA_s|\,.$$

The monotone-convergence theorem allows us to conclude that

$$\sup_{t\in[0,T]}\left|\int_0^t D_s\Big(sgn(X_s-a)-f_n^{'}(X_s)\Big)\,dA_s\right|\to 0\,,$$

in L^1, in probability and almost surely (on passing to a sub-sequence on n if necessary). Hence,

$$\int_0^t D_s f_n^{'}(X_s)\,dA_s\to\int_0^t D_s sgn(X_s)\,dA_s \tag{3.3.10}$$

almost surely and uniformly in time. Similarly,

$$\left|\int_0^t (f_n(X_s)-|X_s-a|)\,dR_s\right| \leq \int_0^t |dR_s|\,,$$

and we get

$$\int_0^t f_n(X_s)\,dR_s \to \int_0^t |X_s-a|\,dR_s$$

almost surely and uniformly in time. Additionally, we can write

$$\int_0^t d\langle D,f_n(X)\rangle_s=\int_0^t f_n^{'}(X_s)\,d\langle N,M\rangle_s\,.$$

From the Kunita-Watanabe inequality,

$$\begin{aligned}\int_0^t \left|\Big(sgn(X_s-a)-f_n^{'}(X_s)\Big)\right| d\langle N,M\rangle_s &\leq \int_0^t \left|\Big(sgn(X_s-a)-f_n^{'}(X_s)\Big)\right| |d\langle N,M\rangle_s| \\ &\leq \sqrt{\int_0^t |(sgn(X_s-a)-f_n^{'}(X_s))|^2\, d\langle M\rangle_s}\sqrt{\int_0^t d\langle N\rangle_s}.\end{aligned}$$

Since $\langle N\rangle_s$ and $\langle M\rangle_s$ are increasing processes of finite variation, we proceed as in (3.3.10) and conclude

$$\int_0^t f_n^{'}(X_s)\, d\langle N,M\rangle_s \to \int_0^t sgn(X_s-a)\, d\langle N,M\rangle_s \, a.s.$$

Hence, from (3.3.8), C^n converges to a limit ζ almost surely (uniformly in time t). Applying integration by parts to the Tanaka formula (3.3.7), we can write

$$\begin{aligned}D_t|X_t-a|-D_0|X_0-a| &= \int_0^t |X_s-a|\, dD_s + \int_0^t D_s sgn(X_s-a)\, dX_s \\ &+ \int_0^t D_s\, d_s l^a + \int_0^t d\langle D,|X-a|\rangle_s\,,\end{aligned}$$

where Tanaka's formula also allows to write

$$\int_0^t d\langle D,|X-a|\rangle_s = \int_0^t sgn(X_s-a)\, d\langle N,M\rangle_s\,,$$

and conclude that

$$\zeta_t = \int_0^t D_s\, d_s l^a\,.$$

□

We can now prove Lemma 3.2.

Proof of Lemma 3.2. From Proposition 3.3, we know that the following holds almost surely:

$$\int_0^t D_u\, dl_u^a = \lim_{n\to\infty}\int_0^t D_u \delta_n^a(X_u) Y_u^2\, du\,.$$

This implies convergence in distribution, so that we can write

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t D_u\, dl_u^a\right] = \lim_{n\to\infty}\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t D_u \delta_n^a(X_u) Y_u^2\, du\right],$$

and, by the stochastic Fubini theorem, we get

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t D_u\, dl_u^a\right] &= \lim_{n\to\infty}\int_0^t \mathbb{E}^{\mathbb{Q}^d}\left[D_u\delta_n^a(X_u)\, Y_u^2\right] du \\ &= \lim_{n\to\infty}\int_0^t \mathbb{E}^{\mathbb{Q}^d}\left[\delta_n^a(X_u)\mathbb{E}^{\mathbb{Q}^d}\left[D_u Y_u^2 \mid X_u\right]\right] du\,.\end{aligned}$$

We denote $\gamma(x,u) = \mathbb{E}^{\mathbb{Q}^d}\left[D_u Y_u^2 \mid X_u = x\right]$, such that

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t D_u\, dl_u^a\right] &= \lim_{n\to\infty}\int_0^t\int_0^\infty \delta_n^a(x)\gamma(x,u)\phi(x,u)\, dxdu \\ &= \lim_{n\to\infty}\int_0^\infty \delta_n^a(x)\left(\int_0^t \gamma(x,u)\phi(x,u)\, du\right) dx,\end{aligned}$$

where we recall that ϕ is the marginal density of X and where we have used Fubini's Theorem in the second line. By the continuity assumptions on ϕ and γ, we deduce that

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t D_u\, dl_u^a\right] = \int_0^t \left(\mathbb{E}^{\mathbb{Q}^d}\left[D_u Y_u^2 \mid X_u = a\right]\phi(a,u)\right) du\,.$$

□

Lemma 3.2 provides a link between the local time of a process and its density function. This plays a key role in a financial setting, where the marginal density of the spot is the second order derivative of the call price with respect to the strike, up to a discount factor.

Lemma 3.4. *Given the set-up of Theorem 3.1,*

$$M_t = \int_0^t \mathbf{1}_{S_u\geq K} D_u^d\alpha(S_u,u)S_u\sqrt{V_u}\, dW_u$$

is a true martingale up to T^* given by (4.3.24).

Proof of Lemma 3.4. Since α and $\mathbf{1}_{S_u\geq K}$ are bounded, the process

$$M_t = \int_0^t \mathbf{1}_{S_u\geq K} D_u^d\alpha(S_u,u)S_u\sqrt{V_u}\, dW_u,$$

is a true martingale if

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t \left(D_u^d S_u\right)^2 V_u\, du\right] < \infty\,.$$

On the one hand, since $t < T^*$, from Proposition 3.13 in [40], we can find $\omega > 2$ such that

$$\sup_{u\in[0,t]} \mathbb{E}^{\mathbb{Q}^d}\Big[\Big(D_u^d S_u\Big)^\omega\Big] < \infty\,.$$

On the other hand, from Theorem 3.1 in [91],

$$\sup_{u\in[0,t]} \mathbb{E}^{\mathbb{Q}^d}\Big[V_u^{\frac{\omega}{\omega-2}}\Big] < \infty\,.$$

Using Hölder's inequality with the pair $\Big(\frac{\omega}{2}, \frac{\omega}{\omega-2}\Big)$,

$$\mathbb{E}^{\mathbb{Q}^d}\Big[\Big(D_u^d S_u\Big)^2 V_u\Big] \le \mathbb{E}^{\mathbb{Q}^d}\Big[\Big(D_u^d S_u\Big)^\omega\Big]^{\frac{2}{\omega}} \mathbb{E}^{\mathbb{Q}^d}\Big[V_u^{\frac{\omega}{\omega-2}}\Big]^{\frac{\omega-2}{\omega}} < \infty.$$

Finally, using the Fubini theorem, $\mathbb{E}^{\mathbb{Q}^d}\Big[\int_0^t \Big(D_u^d S_u\Big)^2 V_u\, du\Big] < \infty$ and hence M is a true martingale of zero expectation. □

Combining Lemmas 3.2 and 3.4, we can derive the following proposition.

Proposition 3.5. *Under the assumptions of Theorem 3.1, and for $T^* > 0$ given by (4.3.24), the call price $C(K,T)$ under model (3.2.1) satisfies the forward equation*

$$\begin{aligned}\frac{\partial C(K,T)}{\partial T} - \frac{1}{2}\alpha(K,T)^2 K^2 \frac{\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d V_T \,|\, S_T = K\Big]}{\mathbb{E}^{\mathbb{Q}^d}\big[D_T^d \,|\, S_T = K\big]} \frac{\partial^2 C(K,T)}{\partial K^2} \\ +\ \mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d r_T^f (S_T - K)^+\Big] - \mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d \mathbf{1}_{S_T\ge K} K\Big(r_T^d - r_T^f\Big)\Big] = 0\end{aligned} \tag{3.3.11}$$

for any strike $K > 0$ and maturity $T < T^*$.

Proof. Let $K \in \mathbb{R}^+$, $0 < t \le T < T^*$ and $H_t = (S_t - K)^+$. The Trotter-Meyer theorem [132] gives

$$(S_t - K)^+ - (S_0 - K)^+ \;=\; \int_0^t \mathbf{1}_{S_u\ge K}\, dS_u + \frac{1}{2} l_t^K\,,$$

which we can write in differential form as

$$dH_t \;=\; \mathbf{1}_{S_t\ge K} S_t\Big(r_t^d - r_t^f\Big)\, dt + \frac{1}{2} dl_t^K + \mathbf{1}_{S_t\ge K}\alpha(S_t,t) S_t \sqrt{V_t}\, dW_t\,.$$

Also,

$$\begin{aligned} d\Big(D_t^d H_t\Big) \;&=\; D_t^d\Big[-r_t^d H_t + \mathbf{1}_{S_t\ge K} S_t\Big(r_t^d - r_t^f\Big)\Big]\, dt + \frac{1}{2} D_t^d\, dl_t^K \\ &+\; \mathbf{1}_{S_t\ge K} D_t^d \alpha(S_t,t) S_t \sqrt{V_t}\, dW_t\,. \end{aligned} \tag{3.3.12}$$

Hence, by applying Lemma 3.2 with $X_t = S_t$, $D_t = D_t^d$ and $Y_t = \alpha(S_t, t)S_t\sqrt{V_t}$, we can write

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t D_u^d \, dl_u^K\right] = \int_0^t \Big(\alpha(K,u)^2 K^2 \mathbb{E}^{\mathbb{Q}^d}\Big[D_u^d V_u \,|\, S_u = K\Big]\phi(K,u)\Big)\, du\,, \tag{3.3.13}$$

where ϕ is the marginal density function of S at time t. Furthermore, one can define $\bar{\phi}_n$ as

$$\bar{\phi}_n(K,u) = \mathbb{E}^{\mathbb{Q}^d}\Big[D_u^d \delta_n^K(S_u)\Big] = \int_0^\infty \delta_n^K(x)\mathbb{E}^{\mathbb{Q}^d}\Big[D_u^d \,|\, S_u = x\Big]\phi(x,u)\, dx\,,$$

with δ_n^K defined as in (4.3.3) by

$$\delta_n^K(x) = \begin{cases} 0, & |x-K| > \frac{1}{n}, \\ \frac{n}{2}, & |x-K| \le \frac{1}{n}\,, \end{cases}$$

and by a similar reasoning to that Lemma 3.2 we get

$$\lim_{n\to\infty} \bar{\phi}_n = \mathbb{E}^{\mathbb{Q}^d}\Big[D_u^d \,|\, S_u = K\Big]\phi(K,u)\,.$$

And since

$$\frac{\partial^2 C(K,u)}{\partial K^2} = \lim_{n\to\infty} \mathbb{E}^{\mathbb{Q}^d}\Big[D_u^d \delta_n^K(S_u)\Big]\,,$$

we write

$$\frac{\partial^2 C(K,u)}{\partial K^2} = \mathbb{E}^{\mathbb{Q}^d}\Big[D_u^d \,|\, S_u = K\Big]\phi(K,u). \tag{3.3.14}$$

Combining (3.3.13) and (3.3.14) allows to write

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t D_u^d \, dl_u^K\right] = \int_0^t \alpha(K,u)^2 K^2 \frac{\mathbb{E}^{\mathbb{Q}^d}\Big[D_u^d V_u \,|\, S_u = K\Big]}{\mathbb{E}^{\mathbb{Q}^d}[D_u^d \,|\, S_u = K]} \frac{\partial^2 C(K,u)}{\partial K^2}\, du\,.$$

Hence, integrating (3.3.12),

$$\begin{aligned} C(K,t) = \mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d H_t\Big] &= \int_0^t \Big(-\mathbb{E}^{\mathbb{Q}^d}\Big[D_u^d r_u^d (S_u - K)^+\Big] + \mathbb{E}^{\mathbb{Q}^d}\Big[D_u^d \mathbf{1}_{S_u \ge K} S_u\Big(r_u^d - r_u^f\Big)\Big]\Big)\, du \\ &+ \frac{1}{2}\int_0^t \alpha(K,u)^2 K^2 \frac{\mathbb{E}^{\mathbb{Q}^d}\Big[D_u^d V_u \,|\, S_u = K\Big]}{\mathbb{E}^{\mathbb{Q}^d}[D_u^d \,|\, S_u = K]} \frac{\partial^2 C(K,u)}{\partial K^2}\, du \\ &+ \mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t \mathbf{1}_{S_u \ge K} D_u^d \alpha(S_u,u) S_u \sqrt{V_u}\, dW_u\right]. \end{aligned} \tag{3.3.15}$$

Furthermore, on a fixed time interval $[0, T^*]$, D_t^d is uniformly bounded by $e^{T^* \max_{u\in[0,T^*]}|h^d(u)|}$. Then, from Lemma 3.4 we know that $\int_0^t \mathbf{1}_{S_u \ge K} D_u^d \alpha(S_u,u) S_u \sqrt{V_u}\, dW_u$ is a true martingale of zero expecta-

tion.

We write (3.3.15) at time T, differentiate with respect to T and, upon noticing that $\mathbf{1}_{S_T \geq K} S_T = (S_T - K)^+ + \mathbf{1}_{S_T \geq K} K$, we get (3.3.11). □

Therefore, we now have all the tools to provide a proof for Theorem 3.1.

Proof of Theorem 3.1. First, we want to ensure that (3.3.1) is a necessary condition for

$$C(K,T) = C_{LV}(K,T)\,. \tag{3.3.16}$$

Hence, by subtracting the Dupire PDE (3.2.4) from (3.3.11), we obtain

$$\begin{aligned}\frac{1}{2}K^2\left(\alpha(K,T)^2\frac{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d V_T \,|\, S_T = K\right]}{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \,|\, S_T = K\right]} - \sigma_{LV}(K,T)^2\right)\frac{\partial^2 C_{LV}}{\partial K^2} &= \mathbb{E}^{\mathbb{Q}^d}\left[D_T^d r_T^f (S_T - K)^+\right] \\ &\quad -\bar{r}^f(T)C_{LV} - \mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \mathbf{1}_{S_T \geq K} K\left(r_T^d - r_T^f\right)\right] - K\left(\bar{r}^d(T) - \bar{r}^f(T)\right)\frac{\partial C_{LV}}{\partial K}\,,\end{aligned}$$

so

$$\alpha^2(K,T) = \frac{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \,|\, S_T = K\right]}{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d V_T \,|\, S_T = K\right]}\Big(\sigma_{LV}(K,T)^2 + \bar{q}(K,T)\Big),$$

where

$$\overline{Q}_T = D_T^d r_T^f (S_T - K)^+ - \bar{r}^f(T)C_{LV} - K\left(D_T^d \mathbf{1}_{S_T \geq K}\left(r_T^d - r_T^f\right) + \left(\bar{r}^d(T) - \bar{r}^f(T)\right)\frac{\partial C_{LV}}{\partial K}\right).$$

It remains to show that $\bar{q} = q$. First, $D_T^d (S_T - K)^+$ is weakly differentiable with respect to K with $\frac{\partial D_T^d (S_T - K)^+}{\partial K} = D_T^d \mathbf{1}_{S_T \geq K}$, which is bounded by the integrable process D_T^d. We can interchange differentiation and expectation to get $\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \mathbf{1}_{S_T \geq K}\right] = -\frac{\partial C}{\partial K}$. Since the models agree, $\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \mathbf{1}_{S_T \geq K}\right] = -\frac{\partial C_{LV}}{\partial K}$ and $\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d (S_T - K)^+\right] = C_{LV}$, and (3.3.1) holds. Moreover, it is easy to see that (3.3.1) is also a sufficient condition for (3.3.16). □

Remark. In this setting, we rely on the fact that a pure LV model, as defined in (3.2.3), can reproduce any arbitrage-free market call prices. In [59], no-arbitrage constraints on call prices are given. Among others, one needs price convexity (to ensure existence of a non-negative risk-neutral density by the Breeden and Litzenberger formula), and calendar spreads to be positive. Namely for a market call price C and all $0 < T_1 < T_2$,

$$\begin{aligned}\frac{\partial^2 C}{\partial K^2} &\geq 0\,, \\ C\left(K e^{\int_{T_1}^{T_2}\left(\bar{r}^d(t) - \bar{r}^f(t)\right)dt}, T_2\right) &\geq e^{-\int_{T_1}^{T_2}\bar{r}^f(t)\,dt} C(K,T_1)\,.\end{aligned} \tag{3.3.17}$$

In [30], it is shown that an equivalent formulation of (3.3.17) is

$$\frac{\partial C(K,T)}{\partial T} + \bar{r}^f(T)C(K,T) + K\Big(\bar{r}^d(T) - \bar{r}^f(T)\Big)\frac{\partial C(K,T)}{\partial K} \geq 0\,.$$

Additionally, the Dupire formula

$$\sigma_{LV}(K,T)^2 = \frac{\frac{\partial C(K,T)}{\partial T} + \bar{r}^f(T)C(K,T) + K\Big(\bar{r}^d(T) - \bar{r}^f(T)\Big)\frac{\partial C(K,T)}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C(K,T)}{\partial K^2}} \tag{3.3.18}$$

provides a necessary and sufficient condition for the pure LV model (3.2.3) to match market call prices. Existence of a local volatility function is then guaranteed by the (static) no-arbitrage conditions that preclude the numerator or denominator to become negative in (3.3.18) (see also [29] for details).

We also note that for a fixed set of stochastic rates parameters, q can become negative and potentially larger than σ_{LV}^2 in (3.3.1). In that case, there exists no leverage function α such that model (3.2.1) fits the market call prices. However, in practice we did not encounter any problem by calibrating the stochastic rates parameters directly to the market. Additionally, since model (3.2.1) is a generalisation of (3.2.3), there always exists a set of stochastic rates parameters such that α exists and satisfies (3.3.1).

3.4 Fast calibration with a new control variate particle method

In this section we describe the full calibration routine step-by-step. First, we give the market data. Then we present the various control variates that we use in order to drastically reduce the computational cost of the particle method for α in the 4-factor model (3.5.1), which build on intermediary calibration steps. Second, we detail each of the these steps, including calibration of the LV, LSV, and interest rate models.

3.4.1 Market data set

We use a set of market data from Bloomberg from the 18/03/2016 for the currency pair EURUSD. The spot value is denominated as the price of one EUR in USD units,

$$\text{Spot} = 1.1271\,(\$\,\text{per}\,€)\,,$$

and the implied volatility is retrieved for the Deltas

$$\text{10D-Put}, \text{25D-Put}, \text{50D}, \text{25D-Call}, \text{10D-Call}$$

and for the maturity pillars

$$\text{3W, 1M, 2M, 3M, 6M, 1Y, 1Y6M, 2Y, 3Y, 5Y.}$$

As an aside, we skip the 7Y and 10Y pillars as the quotes we retrieved were not liquid enough. We plot the market volatility surface on a strike scale in Figure 3.4.1 (from 6 months onward for a better illustration of the shape). The volatility is interpolated linearly in variance on a log-moneyness scale and extrapolated linearly.

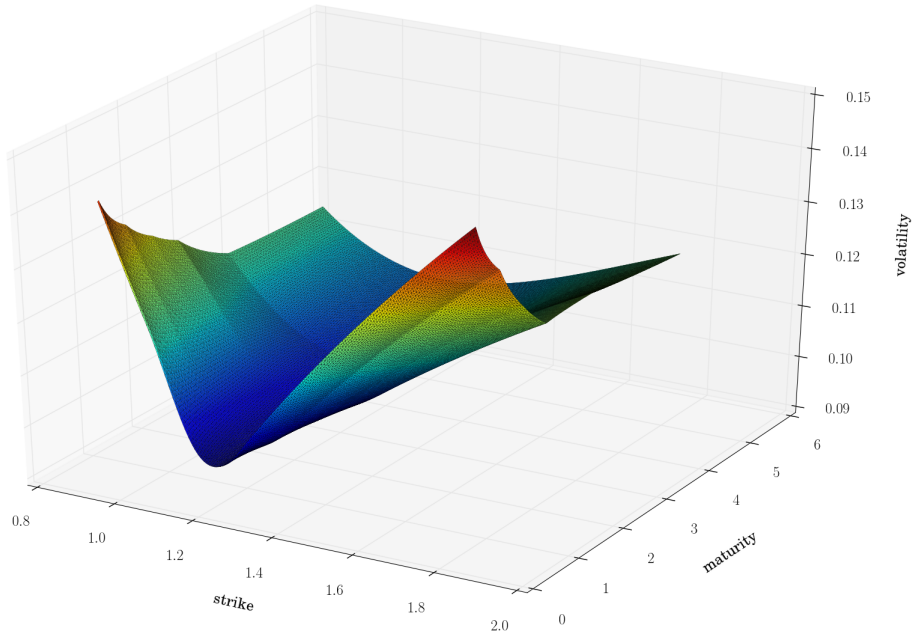


Figure 3.4.1: Market volatility surface for the EUR-USD (18/03/2016)

The domestic risk-neutral measure $\mathbb{Q}^d$ is the one associated with the USD money market account. The foreign risk-neutral measure $\mathbb{Q}^f$ is the one associated to the EUR money market account. The net present value of any financial contract will be expressed in units of domestic numéraire unless stated otherwise.

3.4.2 Variance reduction for standard expectations

If we compare formula (3.3.4) and the formula in Theorem 3.1, we notice that when the rates are stochastic, an additional term appears in the leverage function, namely

$$q(K,T) = \frac{\mathbb{E}^{\mathbb{Q}^d}[Q_T]}{\frac{1}{2}K^2 \frac{\partial^2 C_{LV}}{\partial K^2}},$$

where

$$Q_T = \underbrace{D_T^d\Big(r_T^f - \bar{r}^f(T)\Big)(S_T - K)^+}_{=:X_T} - K\underbrace{D_T^d \mathbf{1}_{S_T \geq K}\Big[\Big(r_T^d - \bar{r}^d(T)\Big) - \Big(r_T^f - \bar{r}^f(T)\Big)\Big]}_{=:X'_T}. \tag{3.4.1}$$

In this section we assume that a pure local volatility model (3.2.3) has been previously calibrated to market call option prices such that for any quoted strike K and maturity T,

$$C_{LV}(K,T) = C_{Market}(K,T)\,.$$

First, we note that the function $\frac{\partial^2 C_{LV}}{\partial K^2}$ can become very small, which can lead to numerical instabilities. To the best of our knowledge, in the literature ([82, 129, 35]) the Dupire formula is typically used to compute this quantity, basically, by differentiating the market prices twice. However, this approach leads to instabilities as a small error in the computation of this second derivative can lead to a big difference in the estimated value of q. Therefore, we propose to calibrate a pure local volatility model with the forward PDE as explained in Subsection 3.4.6 and use $\frac{\partial^2 C_{LV}}{\partial K^2}$ obtained from the PDE solution with a smoothing scheme. Then this quantity will be smooth and stable.

Second, for the random variable Q and for an associated N-sample $(Q_i)_{i\in[\![1,N]\!]}$, we use the sample mean $\frac{1}{N}\sum_{i=1}^N Q_i$ as the Monte Carlo estimator of $\mathbb{E}^{\mathbb{Q}^d}[Q]$ with control variates

$$Y_T = D_T^d(S_T - K)^+, \qquad Z_T = r_T^f - \bar{r}^f(T)\,,$$

for X_T defined in (3.4.1), and

$$Y'_T = D_T^d \mathbf{1}_{S_T \geq K}, \qquad Z'_T = \Big(r_T^d - \bar{r}^d(T)\Big) - \Big(r_T^f - \bar{r}^f(T)\Big)\,,$$

for X'_T. We know that the model (3.2.1) is perfectly calibrated to call option prices if and only if

$$\begin{aligned} \mu_1 \coloneqq \mathbb{E}^{\mathbb{Q}^d}[Y_T] &= C_{LV}(K,T), \\ \mu_2 \coloneqq \mathbb{E}^{\mathbb{Q}^d}[Y'_T] &= -\frac{\partial C_{LV}}{\partial K}(K,T). \end{aligned}$$

Let us also define

$$\gamma_1 := \mathbb{E}^{\mathbb{Q}^d}[X_T], \quad \gamma_2 := \mathbb{E}^{\mathbb{Q}^d}\Big[X'_T\Big]$$

and

$$\begin{aligned} \zeta_1 &\coloneqq \mathbb{E}^{\mathbb{Q}^d}[Z_T] = g_0^f e^{-\kappa_f T} + \theta_f\Big(1 - e^{-\kappa_f T}\Big) + h^f(T) - \bar{r}^f(T), \\ \zeta_2 &\coloneqq \mathbb{E}^{\mathbb{Q}^d}\Big[Z'_T\Big] = g_0^d e^{-\kappa_d T} + \theta_d\Big(1 - e^{-\kappa_d T}\Big) + h^d(T) - \bar{r}^d(T) - \zeta_1. \end{aligned}$$

We denote the corresponding Monte Carlo estimators by $\hat{\mu}_{1,N}$, $\hat{\mu}_{2,N}$, $\hat{\gamma}_{1,N}$, $\hat{\gamma}_{2,N}$, $\hat{\zeta}_{1,N}$ and $\hat{\zeta}_{2,N}$,

respectively, using the same Brownian paths for W, W^V, W^{g^d}, W^{g^f} in all estimators. We can define a new Monte Carlo estimator γ_N^* for $\gamma_1 - K\gamma_2$,

$$\gamma_N^* = \left(\gamma_{1,N}^* - K\gamma_{2,N}^*\right), \tag{3.4.2}$$

with

$$\begin{aligned}
\gamma_{1,N}^* &= \hat{\gamma}_{1,N} + \lambda_1(\hat{\mu}_{1,N} - C_{LV}(K,T)) + \eta_1\left(\hat{\zeta}_{1,N} - \zeta_1\right), \\
\gamma_{2,N}^* &= \hat{\gamma}_{2,N} + \lambda_2\left(\hat{\mu}_{2,N} + \frac{\partial C_{LV}(K,T)}{\partial K}\right) + \eta_2\left(\hat{\zeta}_{2,N} - \zeta_2\right).
\end{aligned}$$

We define a random variable

$$G_T = X_T + \lambda_1(Y_T - C_{LV}(K,T)) + \eta_1(Z_T - \zeta_1)$$

such that $\gamma_{1,N}^*$ is an estimator of $\mathbb{E}[G_T] = \mathbb{E}[X_T]$, and we similarly define a random variable G' involving Y' and Z', with an estimator $\gamma_{2,N}^*$. Each expectation now has two control variates (see Section 4.1 in [73]). The variance of G_T is

$$\begin{aligned}
\mathbb{V}[X_T + \lambda_1(Y_T - C_{LV}(K,T)) + \eta_1(Z_T - \zeta_1)] &= \mathbb{V}[X_T] + \lambda_1^2\mathbb{V}[Y_T] + \eta_1^2\mathbb{V}[Z_T] \\
&\quad + 2\lambda_1\mathrm{Cov}(X_T, Y_T) + 2\eta_1\mathrm{Cov}(X_T, Z_T) + 2\lambda_1\eta_1\mathrm{Cov}(Y_T, Z_T).
\end{aligned}$$

A minimum is obtained (see [73]) when $\lambda_1 = \bar{\lambda}(X_T, Y_T, Z_T)$ and $\eta_1 = \bar{\lambda}(X_T, Z_T, Y_T)$, where

$$\bar{\lambda}(X,Y,Z) = \frac{\mathrm{Cov}(X,Y)\mathrm{Cov}(Y,Z) - \mathrm{Cov}(X,Y)\mathbb{V}[Z]}{\mathbb{V}[Z]\mathbb{V}[Y] - \mathrm{Cov}(Y,Z)^2},$$

which can be estimated from the Monte Carlo samples. Similar expressions are obtained for X', where the corresponding optimisers are $\lambda_2 = \bar{\lambda}(X_T', Y_T', Z_T')$, $\eta_2 = \bar{\lambda}(X_T', Z_T', Y_T')$. Then

$$q(K,T) = \lim_{N\to\infty} \frac{\gamma_N^*}{\frac{1}{2}K^2\frac{\partial^2 C_{LV}}{\partial K^2}}.$$

This approach is particularly useful for out-of-the-money options and digital options as the Monte Carlo estimator will exhibit higher variance in these settings. The method will help reduce the variance while working with a sufficiently low number of particles (i.e., paths).

One should be careful in the eventuality where there is no particle S_T above K. In that case, the

estimated variances $\mathbb{V}\Big[D_T^d \mathbf{1}_{S_T \geq K}\Big]$ and $\mathbb{V}\Big[D_T^d (S_T - K)^+\Big]$ are zero, and we pick

$$\begin{aligned} \lambda_1 &= -\zeta_1, \quad \eta_1 = -C_{LV}(K,T)\,, \\ \lambda_2 &= -\zeta_2, \quad \eta_2 = \frac{\partial C_{LV}(K,T)}{\partial K}\,, \end{aligned}$$

such that both control variates are of the same order of magnitude. Similarly, if all particles S_T are above K, the estimated variance $\mathbb{V}\Big[D_T^d \mathbf{1}_{S_T \geq K}\Big]$ is zero. We then choose

$$\lambda_2 = -\zeta_2, \quad \eta_2 = \frac{\partial C_{LV}(K,T)}{\partial K}\,.$$

3.4.3 Variance reduction for the Markovian projection

Our goal is to reduce the variance of an estimator p_N for

$$p(K,T) = \frac{\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d V_T \,|\, S_T = K\Big]}{\mathbb{E}^{\mathbb{Q}^d}\big[D_T^d \,|\, S_T = K\big]}.$$

Let us assume that we have simulated N Monte Carlo paths for the system $\Big(D_t^d, S_t, V_t\Big)$ up to time $T > 0$. We denote the state of this system at time T for the i-th path by (D_T^i, S_T^i, V_T^i). Following the particle method used for LSV model calibration in [82], an estimator for p is defined as

$$\begin{aligned} \hat{p}_N(K,T) &= \frac{\sum_{i=1}^N D_T^i V_T^i \delta_N(S_T^i - K)}{\sum_{i=1}^N \delta_N(S_T^i - K)} \frac{\sum_{i=1}^N \delta_N(S_T^i - K)}{\sum_{i=1}^N D_T^i \delta_N(S_T^i - K)} \\ &= \frac{\sum_{i=1}^N D_T^i V_T^i \delta_N(S_T^i - K)}{\sum_{i=1}^N D_T^i \delta_N(S_T^i - K)}\,, \end{aligned}$$

where δ_N is a kernel function with a bandwidth that "tightens" with the number of particles. We assume that the 2-factor LSV model (3.2.2) is perfectly calibrated to market call prices. We know from [129, 54, 82] that this is the case if and only if

$$\alpha^{2D}(K,T) = \frac{\sigma_{LV}(K,T)}{\sqrt{\mathbb{E}^{\mathbb{Q}^d}\big[V_T^{2D} \,|\, S_T^{2D} = K\big]}}\,.$$

We can use

$$p^{2D} = \mathbb{E}^{\mathbb{Q}^d}\Big[V_T^{2D} \,|\, S_T^{2D} = K\Big]$$

and its Monte Carlo estimator,

$$\hat{p}_N^{2D}(K,T) = \sum_{i=1}^N \frac{V_T^{2D,i} \delta_N\Big(S_T^{2D,i} - K\Big)}{\sum_{i=1}^N \delta_N\Big(S_T^{2D,i} - K\Big)}\,, \tag{3.4.3}$$

as a control variate for $\hat{p}_N$. We define a new estimator p_N^* such that

$$p_N^* = \hat{p}_N(K,T) + \lambda\Big(\hat{p}_N^{2D}(K,T) - p^{2D}(K,T)\Big). \tag{3.4.4}$$

The latter has an asymptotically diminishing bias if we assume the particle method to converge in distribution (and neglect the time stepping bias). In order to get a good estimate for λ, we can rewrite the above estimator as follows

$$p_N^* = \frac{1}{N}\sum_{i=1}^{N} \frac{D_T^i V_T^i \delta_N(S_T^i - K)}{\frac{1}{N}\sum_{i=1}^{N} D_T^i \delta_N(S_T^i - K)} + \lambda\left(\frac{1}{N}\sum_{i=1}^{N} \frac{V_T^{2D,i}\delta_N\Big(S_T^{2D,i} - K\Big)}{\frac{1}{N}\sum_{i=1}^{N}\delta_N\Big(S_T^{2D,i} - K\Big)} - p^{2D}(K,T)\right),$$

which can be also written as

$$p_N^* = \frac{1}{N}\sum_{i=1}^{N} m_i + \lambda\left(\frac{1}{N}\sum_{i=1}^{N} m_i^{2D} - p^{2D}(K,T)\right)$$

with

$$m_i = \frac{D_T^i V_T^i \delta_N\big(S_T^i - K\big)}{\frac{1}{N}\sum_{i=1}^{N} D_T^i \delta_N\big(S_T^i - K\big)}, \quad m_i^{2D} = \frac{V_T^{2D,i}\delta_N\big(S_T^{2D,i} - K\big)}{\frac{1}{N}\sum_{i=1}^{N}\delta_N\big(S_T^{2D,i} - K\big)}$$

which is in the standard Monte Carlo control variate form and where the convergence of m_i and m_i^{2D} in N is ensured by propagation of chaos (see Section 3.5.2 for more details). We can think of m_i and m_i^{2D} as samples of two random variables m and m^{2D} respectively, and for the best variance reduction (see Section 4.1 in [73]), we take

$$\lambda = -\frac{\mathrm{Cov}\Big(m, m^{2D}\Big)}{\mathrm{Var}(m^{2D})},$$

which we can estimate by

$$\hat{\lambda}_N = -\frac{\sum_{i=1}^{N}(m_i - \overline{m})\Big(m_i^{2D} - \overline{m}^{2D}\Big)}{\sum_{i=1}^{N}\big(m_i^{2D} - \overline{m}^{2D}\big)^2}, \tag{3.4.5}$$

where $\overline{m}$ and $\overline{m}^{2D}$ are the sample means. We recall that the expected variance reduction factor is

$$\frac{1}{1 - \mathrm{Corr}(m, m^{2D})^2}. \tag{3.4.6}$$

Hence, if the stochastic rates are not highly volatile, i.e. if ξ_f and ξ_d are small enough, we expect a very good variance reduction as the correlation between the particles generated by the 4-factor hybrid LSV model (3.2.1) and by the 2-factor LSV model (3.2.2) will be high. Our numerical tests performed on a model calibrated to recent EUR and USD market data exhibit a correlation (measured for random variables m and m^{2D}) between model (3.2.1) and (3.2.2) of 95% up to 1.5 years and 50% around 5 years. Additionally, our stress test in Subsection 3.5.4 suggests that even under high volatility regimes for the rate processes, i.e. when ξ_f and ξ_d are very big, the variance reduction

brought by this control variate is significant.

3.4.4 Shifted CIR model

The domestic and foreign short interest rates are modeled by the shifted CIR (CIR++) process [21]. On the one hand, this model preserves the analytical tractability of the CIR model for bonds, caps and other basic interest rate products. On the other hand, it is flexible enough to fit the initial term structure of interest rates exactly. For $i \in \{d, f\}$, the short rate dynamics under their respective spot measures, i.e., $\mathbb{Q}^d$ – domestic and $\mathbb{Q}^f$ – foreign, are given by

$$\begin{cases} r_t^i = g_t^i + h^i(t), \\ dg_t^i = \kappa_i(\theta_i - g_t^i)dt + \xi_i\sqrt{g_t^i}\, dB_t^{g^i}, \quad g_0^i > 0, \end{cases} \tag{3.4.7}$$

where B^{g^d} and B^{g^f} are Brownian motions under $\mathbb{Q}^d$ and $\mathbb{Q}^f$, respectively. The mean-reversion parameters κ_i, the long-term mean parameters θ_i and the volatility parameters ξ_i are the same as in (3.2.1). The calibration of the short rate model (3.4.7) follows the same approach for both the domestic and the foreign interest rate. For simplicity, we drop the subscripts and superscripts "d" and "f" in the remainder of the subsection and define the vector of parameters $\beta_1 = (g_0, \kappa, \theta, \xi)$. According to Brigo and Mercurio [20], an exact fit to the initial term structure of interest rates is equivalent to $h(t) = \varphi^{\text{CIR}}(t; \beta_1)$ for all $t \in [0, T]$, where

$$\begin{aligned} \varphi^{\text{CIR}}(t;\beta_1) &= \bar{r}(0,t) - \bar{r}^{\text{CIR}}(0,t;\beta_1), \\ \bar{r}^{\text{CIR}}(0,t;\beta_1) &= \frac{2\kappa\theta(\exp\{t\nu\}-1)}{2\nu + (\kappa+\nu)(\exp\{t\nu\}-1)} + g_0 \frac{4\nu^2 \exp\{t\nu\}}{[2\nu + (\kappa+\nu)(\exp\{t\nu\}-1)]^2}, \end{aligned} \tag{3.4.8}$$

$\nu = \sqrt{\kappa^2 + 2\xi^2}$ and $\bar{r}(0,t)$ is the market instantaneous forward rate at time 0 for a maturity t, i.e.,

$$\bar{r}(0,t) = \bar{r}(t) = -\frac{\partial \ln P(0,t)}{\partial t}, \tag{3.4.9}$$

where $P(0,t)$ is the market zero coupon bond price at time 0 for a maturity t. The value of the zero coupon bond is given by

$$P(0,t) = \frac{1}{1 + \Delta(0,t)R(0,t)}, \tag{3.4.10}$$

where $\Delta(0,t)$ is the year fraction from 0 to time t and $R(0,t)$ is the current (simply-compounded) deposit rate with maturity date t which is quoted in the market. As an aside, note that the standard day count convention for USD and EUR is Actual 360.

In order to estimate the zero coupon curve (also known as the term structure of interest rates or the yield curve), we assume that the instantaneous forward rate is piecewise-flat. Consider the time nodes $t_0 = 0, t_1, \ldots, t_n$ and the set of estimated instantaneous forward rates $f_1, f_2, \ldots, f_n$ from which the

curve is constructed, and define

$$\bar{r}(t) = f_i \quad \text{if} \quad t_{i-1} \leq t < t_i, \quad \text{for} \quad i = 1, 2, \ldots, n. \tag{3.4.11}$$

Using (3.4.9) – (3.4.11) and solving the resulting linear system of equations, we get

$$f_i = \frac{1}{\Delta(t_{i-1}, t_i)} \ln\left(\frac{1 + \Delta(0, t_i) R(0, t_i)}{1 + \Delta(0, t_{i-1}) R(0, t_{i-1})}\right) \quad \text{for} \quad i = 1, 2, \ldots, n. \tag{3.4.12}$$

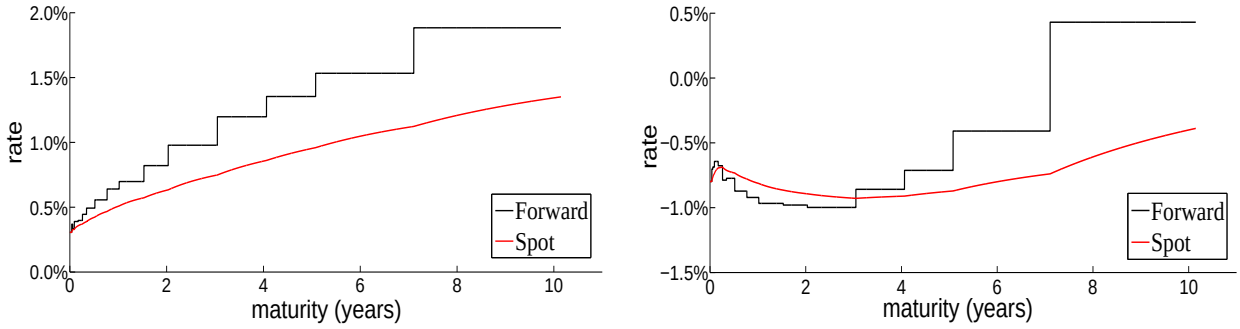
The continuously-compounded spot rate, i.e., the constant rate at which the value of a pure discount bond must grow to yield one unit of currency at maturity, is defined as

$$R_0(0, t) = \frac{1}{\Delta(0, t)} \int_0^t \bar{r}(s)\, ds. \tag{3.4.13}$$

Using (3.4.11) – (3.4.13), we deduce that

$$R_0(0, t) = \frac{\Delta(t, t_i)}{\Delta(0, t)\Delta(t_{i-1}, t_i)} \ln\left(1 + \Delta(0, t_{i-1}) R(0, t_{i-1})\right) + \frac{\Delta(t_{i-1}, t)}{\Delta(0, t)\Delta(t_{i-1}, t_i)} \ln\left(1 + \Delta(0, t_i) R(0, t_i)\right) \tag{3.4.14}$$

whenever $t_{i-1} \leq t < t_i$. In Figure 3.4.2, we plot the USD and EUR zero coupon curves $t \mapsto R_0(0, t)$, $t > 0$, estimated from the quoted deposit rates from March 18, 2016, together with the flat-forward instantaneous forward rates.



(a) USD Currency **(b)** EUR Currency

Figure 3.4.2: The instantaneous forward rates and the continuously-compounded spot rates.

A choice of the shift function h as in (3.4.8) results in an exact fit to the initial term structure of interest rates independent of the value of the parameter vector β_1.

Next, we determine β_1 by calibrating the CIR++ model to the current term structure of volatilities, in particular, by fitting at-the-money (ATM) cap volatilities. We consider caps with integer maturities ranging from 1 to 10 years for both currencies, with an additional 18 month cap for EUR. For USD,

all caps have quarterly frequency, whereas for EUR the 1 year and 18 month caps have quarterly frequency and the 2 to 10 year caps have semi-annual frequency. A cap is a set of spanning caplets with a common strike so the value of the cap is simply the sum of the values of its caplets. It is market standard to price caplets with the Black formula, in which case the fair value of the cap at time 0 with rate (strike) K, reset times $T_a, T_{a+1}, \ldots, T_{b-1}$ and payment times $T_{a+1}, \ldots, T_{b-1}, T_b$ is:

$$\mathrm{Cap}_{\mathrm{Black}}(K, \sigma_{a,b}) = \sum_{i=a+1}^{b} P(0, T_i)\Delta(T_{i-1}, T_i)\mathrm{Black}\big(K, F(0, T_{i-1}, T_i), \sigma_{a,b}\sqrt{T_{i-1}}\big), \tag{3.4.15}$$

where $F(0, T, S)$ is the simply-compounded forward rate at time 0 for the expiry T and maturity S defined as

$$F(0, T, S) = \frac{1}{\Delta(T, S)}\left(\frac{P(0, T)}{P(0, S)} - 1\right) \tag{3.4.16}$$

and the Black volatility $\sigma_{a,b}$ corresponding to a strike K is retrieved from market quotes. Denoting by ϕ_0 and Φ_0 the standard normal probability density function (PDF) and cumulative distribution function (CDF), respectively, Black's formula is:

$$\begin{aligned} \mathrm{Black}(K, F, v) &= F\Phi_0(d_1) - K\Phi_0(d_2), \\ d_{1,2} &= \frac{\ln(F/K) \pm v^2/2}{v}. \end{aligned} \tag{3.4.17}$$

However, Black's formula cannot cope with negative forward rates F or strikes K, in which case we switch to Bachelier's (normal) formula in (3.4.15):

$$\begin{aligned} \mathrm{Normal}(K, F, v) &= (F - K)\Phi_0(d) + v\phi_0(d), \\ d &= \frac{F - K}{v}. \end{aligned} \tag{3.4.18}$$

The data in Figure 3.4.2b suggest that the instantaneous forward rate for EUR takes negative values. Therefore, we use Black cap volatility quotes for USD and Normal cap volatility quotes for EUR. The market prices of at-the-money caps are computed by inserting the forward swap rate

$$S_{a,b} = \frac{P(0, T_a) - P(0, T_b)}{\sum_{i=a+1}^{b} \Delta(T_{i-1}, T_i)P(0, T_i)} \tag{3.4.19}$$

as strike and the quoted cap volatility as $\sigma_{a,b}$ in (3.4.15), using either Black's or Bachelier's formula.

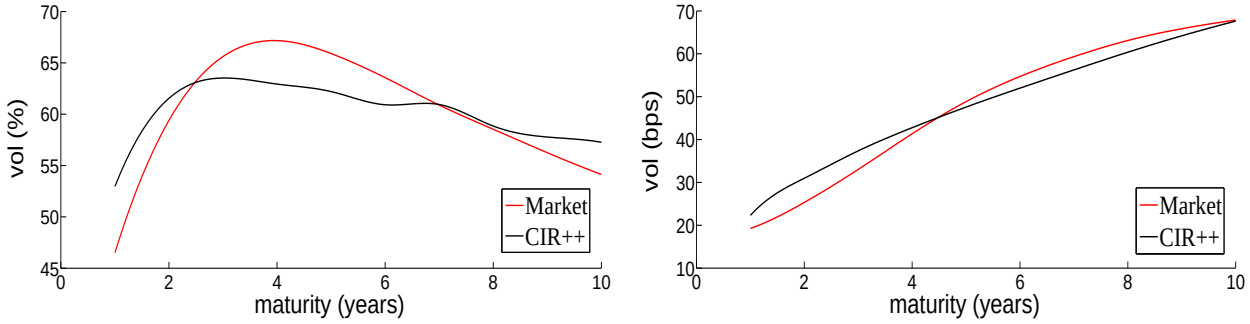
Fitting the CIR++ model to cap volatilities means finding the value of β_1 for which the model cap prices, which are available in closed-form [21], best match the market cap prices. The calibration is performed by minimising the sum of the squared differences between model- and market-implied cap volatilities:

$$\min_{\beta_1 \in \mathbb{R}_+^4} \sum_{1 \le i \le n} \left[\sigma^{\mathrm{CIR}}(T_i; \beta_1) - \sigma^{\mathrm{M}}(T_i)\right]^2, \tag{3.4.20}$$

where σ^{CIR} and σ^{M} stand for the model- and the market-implied cap volatilities, respectively, and $T_1, \ldots, T_n$ are the cap maturities. Model-implied cap volatilities are obtained by pricing market caps with the CIR++ model and then inverting the formula (3.4.15) in order to retrieve the implied volatility associated with each maturity. We choose to calibrate the model to cap volatilities since they are of similar magnitude, unlike cap prices which can differ by a few orders of magnitude. The calibration results are displayed in Table 3.4.1.

On the one hand, (3.4.20) is a highly nonlinear and non-convex optimisation problem, and the objective function may have multiple local minima. On the other hand, global optimisation algorithms require a very high computation time and do not scale well with complexity, as opposed to local optimisation methods. A fast calibration is important in practice since option pricing models may need to be re-calibrated several times within a short time span. Therefore, we used a nonlinear least-squares solver, in particular the trust-region-reflective algorithm [37], for the calibration and a global optimisation method, in particular a genetic algorithm [138], for verification purposes only.

Figure 3.4.3 shows the fitting capability of the CIR++ model, and the implied cap volatility curve is compared to the market curve for each currency. Taking into account that the model has only 4 parameters to fit between 10 and 11 data points, we conclude that the CIR++ model provides a fairly reasonable fit to the term structure of cap volatilities $T_i \mapsto \sigma^{\mathrm{M}}(T_i)$, $1 \leq i \leq n$.



(a) USD Currency **(b)** EUR Currency

Figure 3.4.3: The market- and model-implied term structures of cap volatilities.

The calibration results are displayed in Table 3.4.1.

3.4.5 Four-factor hybrid stochastic volatility model calibration

Consider a "purely stochastic" version of the model (3.2.1) – the Heston-2CIR++ model with leverage function $\equiv 1$ – and additionally suppose that the domestic and the foreign short interest rate dynamics are independent of the dynamics of the spot FX rate. The model is governed by the following system

CCY	g_0	κ	θ	ξ
USD	0.0001	0.0837	0.5469	0.0274
EUR	0.0001	0.0110	1.1656	0.0370

Table 3.4.1: The calibrated CIR parameters

of SDEs under the domestic risk-neutral measure $\mathbb{Q}^d$:

$$\begin{cases} \dfrac{dS_t^{SV}}{S_t^{SV}} = \left(r_t^d - r_t^f\right) dt + \sqrt{V_t}\, dW_t,\ S_0^{SV} = S_0, \\ r_t^d = g_t^d + h^d(t) \\ r_t^f = g_t^f + h^f(t) \\ dg_t^d = \kappa_d\left(\theta_d - g_t^d\right) dt + \xi_d \sqrt{g_t^d}\, dW_t^{g^d} \\ dg_t^f = \kappa_f\left(\theta_f - g_t^f\right) dt + \xi_f \sqrt{g_t^f}\, dW_t^{g^f} \\ dV_t = \kappa(\theta - V_t)\, dt + \xi \sqrt{V_t}\, dW_t^V\,, \end{cases} \tag{3.4.21}$$

where W^s and W^V are correlated Brownian motions with correlation coefficient ρ. Note that the quanto correction term in the drift of the foreign short rate vanishes due to the postulated independence assumption between the spot FX rate and foreign short rate dynamics.

Define the vector of parameters $\beta_2 = (v_0, \kappa, \theta, \xi, \rho)$. The next step in our calibration is to find the values of these 5 model parameters for which European call option prices best match the market call prices retrieved from volatility quotes for different strikes and maturities.

For a EURUSD transaction, the market standard is to choose USD as the domestic currency and EUR as the foreign currency. The forward FX rate for a payment date T is defined as

$$F_T = \frac{P^f(0,T)}{P^d(0,T)} S_0, \tag{3.4.22}$$

where $P^d(0,T)$ and $P^f(0,T)$ are the domestic and foreign discount factors at time 0 for a maturity T, respectively.

Under the postulated simple correlation structure of the Brownian drivers and when the short rates are driven by the CIR process, i.e., when $h^{d,f} = 0$, Ahlip and Rutkowski [4] derive an efficient closed-form formula for the European call option price. Hence, we denote by $C_{\mathrm{A}}(K,T)$ the fair value under the Heston–2CIR model of a European call option with strike K and maturity T computed with the aforementioned formula, and by $C_{\mathrm{H}}(K,T)$ the fair value of the same option but under the

Heston–2CIR++ model. For $i \in \{d, f\}$, we define for brevity

$$H_i = \exp\left\{\int_0^T h^i(t)\,dt\right\},$$

where the shift functions $h^{d,f}$ were calibrated in the previous subsection. Then we can extend the pricing formula of Ahlip and Rutkowski [4] as follows.

$$\begin{aligned} C_{\mathrm{H}}(K,T) &= \mathbb{E}^{\mathbb{Q}^d}\left[\exp\left\{-\int_0^T r_t^d\,dt\right\}(S_T^{SV}-K)^+\right] \\ &= H_f^{-1}\,\mathbb{E}^{\mathbb{Q}^d}\left[\exp\left\{-\int_0^T g_t^d\,dt\right\}(H_f H_d^{-1} S_T^{SV} - H_f H_d^{-1} K)^+\right]. \end{aligned}$$

Therefore, $C_{\mathrm{H}}(K,T) = H_f^{-1} C_{\mathrm{A}}(\tilde{K},T)$, where $\tilde{K} = H_f H_d^{-1} K$. We now calibrate the Heston–2CIR++ model by minimising the sum of the squared differences between model and market call prices:

$$\min_{\beta_2 \in \mathbb{R}_+^4 \times [-1,1]} \sum_{\substack{1\le i\le n \\ 1\le j\le m}} \left[C_{\mathrm{H}}(K_j,T_i;\beta_2) - C_{\mathrm{BS}}(K_j,T_i,\sigma_{i,j})\right]^2, \tag{3.4.23}$$

where $\sigma_{i,j}$ is the quoted volatility corresponding to a strike K_j and a maturity T_i, for $j = 1,\dots,m$ and $i = 1,\dots,n$. There are many ways to choose the objective function (error measure) in (3.4.23). For instance, we may consider either call prices or Black–Scholes implied volatilities and minimise the sum of either absolute or relative (squared) differences between model and market values, using either uniform or non-uniform weights. We choose this particular error measure, which assigns more weight to more expensive options (in-the-money, long-term) and less weight to cheaper options (out-of-the-money, short-term), for two reasons. First, the Heston model, and hence the Heston–2CIR++ model by extension, cannot reproduce the smiles or skews typically observed for short maturities that well and a more careful calibration to these smiles would result in a larger overall model error due to the inherent poor fit of the model to the short-term. Second, market data becomes scarce as the maturity increases, and hence we already assigned more weight to the short- and mid-term sections of the volatility surface; for instance, we have more maturities up to 1 year than between 1 and 5 years. The calibration results are displayed in Table 3.4.2.

v_0	κ	θ	ξ	ρ
0.0094	1.4124	0.0137	0.2988	−0.1194

Table 3.4.2: The calibrated Heston parameters

As before, we employ a nonlinear least-squares solver (the trust-region-reflective algorithm, see [36]) for the calibration and a global optimisation method (a genetic algorithm) for verification purposes.

Due to the non-linearity and non-convexity of the problem, the calibrated model parameters may end up in a local rather than a global minimum of the objective function. Hence, a good initial parameter guess may significantly improve the quality of the calibration. Practitioners usually use variance swap prices to calibrate v_0, κ and θ. In our case, we found the squared ATM 3-week and 5-year volatilities to provide good initial guesses for v_0 and θ, respectively.

3.4.6 One-factor local volatility model calibration

The calibration routine for a pure local volatility model is run with a standard algorithm forward in maturity. We recall that model (3.2.3) is written as

$$\frac{dS_t^{LV}}{S_t^{LV}} = \Big(\bar{r}^d(t) - \bar{r}^f(t)\Big)\, dt + \sigma_{LV}\Big(S_t^{LV}, t\Big)\, dW_t\,,$$

and we want to find the function σ_{LV} for which the call prices under the local volatility model match the quoted market prices exactly. This is crucial as both σ_{LV} and $\frac{\partial^2 C_{LV}}{\partial K^2}$ appear in the leverage function formula (3.3.1). The forward Dupire PDE (3.2.4),

$$\begin{cases} \frac{\partial C_{LV}}{\partial T} + \Big(\bar{r}^d(T) - \bar{r}^f(T)\Big) K \frac{\partial C_{LV}}{\partial K} + \bar{r}^f(T) C_{LV} - \frac{1}{2} K^2 \frac{\partial^2 C_{LV}}{\partial K^2} \sigma_{LV}^2(K,T) = 0\,, \\ C_{LV}(K,0) = (S_0 - K)^+, \quad C_{LV}(0,T) = S_0, \quad C_{LV}(S_{max},T) = 0\,, \end{cases} \tag{3.4.24}$$

provides an efficient way to calibrate and, eventually, regularise the problem. Denote by Φ the map from the local volatility to the model implied volatility Σ_{Model}, i.e.,

$$\Phi\Big(\Big\{\sigma_{LV}(s,t)\,|\,(s,t)\in\mathbb{R}_+^2\Big\}\Big) = \Big\{\Sigma(K,T)\,|\,(K,T)\in\mathbb{R}_+^2\Big\}\,.$$

Furthermore, the PDE solution for a guess σ_S of the local volatility gives call prices for the whole set of strikes and maturities. Inverting the Black formula allows to retrieve the model implied volatilities Σ. Hence, as proposed in [141], we can use the forward Dupire PDE (3.4.24) combined with an efficient implied volatility inverter [97] as the mapping function Φ. A very useful property of this PDE is that it can be solved forward in maturity. Let a set of maturities quoted on the market be $(T_1, ..., T_{N_{Mat}})$ and a set of M_i strikes for a given maturity T_i be $(K_{T_i,1}, ..., K_{T_i,M_i})$. It is possible to solve the PDE on $[0, T_1]$, then on $[T_1, T_2]$ and so forth.

3.4.6.1 Computation of the target volatility surface

For the calibration routine, we will compute the solution of the PDE (3.4.24) by a finite difference method. The spot grid is defined on $[0, S_{max}]$, where $S_{max} = S_0 e^{\frac{6}{2}\sigma_{ATM}\left(\frac{T_{max}}{2}\right)\sqrt{\frac{T_{max}}{2}}}$. In order to speed up the calibration routine, we prefer not to use too many spot steps and time steps (150 steps in space and 20 time steps per year). Hence, the scheme will not have converged to the solution of

the PDE at this point. In order to tackle this problem and still benefit from a good speed-up, we will compute a "target volatility surface": instead of calibrating the market volatility surface, we will calibrate a volatility surface that takes into account the discretisation error of the numerical PDE solution. Industry practitioners like Murex use this approach [111]. The algorithm to build the target surface is explained below.

Algorithm 3.1 Computation of the target volatility surface

```
for ( i = 1 ; i ≤ N_Mat ; i++ ) do
  for ( j = 1 ; j ≤ M_i ; j++ ) do
    define σ_Market = Σ_Market(K_{i,j}, T_i) from the market volatility surface
    solve the PDE (3.4.24) with constant local vol σ_LV = σ_Market
    get C(K_{i,j}, T_i) from the numerical solution
    get Σ_Target(K_{i,j}, T_i) by inverting the price with the Black-Scholes formula
  end for
end for
```

3.4.6.2 Calibration by fixed-point algorithm and forward induction

The calibration can be performed with a regularised minimisation using a penalised function. We refer the reader to [42, 111] for this approach. In this work, we choose a fixed-point algorithm for time efficiency, but want to highlight that a penalised minimisation yields to the same calibrated function σ_{LV}. The local volatility function is defined on a grid of points interpolated with cubic splines in spot and backward flat in time. In the FX case, where there are 5 quoted strikes per maturity (10 maturities), the local volatility is defined on a grid of 50 points. Each one of the points $\sigma_{LV}^{i,j} = \sigma_{LV}(K_{T_i,j}, T_i)$, with $i \in [\![1, 10]\!]$ and $j \in [\![1, 5]\!]$, can be seen as a parameter of the local volatility surface. For a given maturity T_i, the local volatility is defined on the interval $[K_{T_i,1}, K_{T_i,5}]$ and is extrapolated flat outside those bounds.

In order to define a first guess for the calibration routine, we use a smoothed bi-variate cubic spline following the algorithm in [50] to interpolate in strike and maturity the call prices on the market. This allows us to use the Dupire formula to define a first guess for the first maturity $T = T_1$. After the calibration of the first maturity pillar T_1, the first guess for the next pillar is the current maturity local volatility. This approach has shown the best stability and speed in our tests.

As we now have a way to get the model implied volatility from the local volatility (with Φ), one can follow a Picard fixed-point algorithm as proposed in [127, 141] that we describe below. The main idea is to start with a local volatility guess σ_{LV}, solve the Dupire PDE forward in maturity up to the first maturity pillar, and store the Call prices under the local volatility model obtained from the solution. We then retrieve the model implied volatilities Σ_{Model} and update the local volatility guess according to (3.4.25) and iterate this process until convergence.

Algorithm 3.2 Fixed-point forward induction

for $(\ i = 1\,;\, i \leq N_{Mat}\,;\, i++)$ **do**
  while it $<$ maxIter **do**
    solve PDE (3.4.24) on $[T_{i-1}, T_i]$
    compute model implied vol Σ_{Model} for maturity T_i from the computed call prices
    compute $error = \sum_{m=1}^{M_i} (\Sigma_{Model}(K_{T_i,m}, T_i) - \Sigma_{Target}(K_{T_i,m}, T_i))^2$
    if error $<$ tol **then**
      endwhile
    else
      for $(\ j = 1\,;\, j \leq M_i\,;\, j++)$ **do**
        update local volatility guess

$$\sigma_{LV}(K_{T_i,j}, T_i) = \sigma_{LV}(K_{T_i,j}, T_i) \frac{\Sigma_{Target}(K_{T_i,j}, T_i)}{\Sigma_{Model}(K_{T_i,j}, T_i)} \tag{3.4.25}$$

      end for
    end if
    it++
  end while
end for

Remark. As stated in [127], the map Φ is contracting and so is $f(\{\sigma_{LV}\}) \to \{\sigma_{LV}\} * \frac{\{\Sigma_{Target}\}}{\Phi(\{\sigma_{LV}\})}$. Then f admits a unique fixed point that is the limit of the sequence of local volatility guesses $(\sigma_{LV}^n)_{n\in\mathbb{N}}$ defined as $\left\{\sigma_{LV}^{n+1}\right\} = f(\{\sigma_{LV}^n\})$. Convergence is achieved particularly fast (between 10 and 20 iterations in practice).

The calibrated local volatility is shown in Figure 3.4.4, where we plot it on a time scale to T_{max} for a better illustration of its shape.

We push the discretisation settings for the calibration to 800 space steps and 100 time steps per year (we use finite element method with quadratic basis functions to solve the PDE) and price quoted vanilla contracts with the backward Feynman-Kac PDE under the calibrated local volatility model. Then, we get a maximum error in implied volatility smaller than 0.01% (i.e., for a market volatility of 20%, the calibrated volatility could be $20.00 \pm 0.01\%$ in the worst case scenario).

Additionally, we plot the discounted marginal density of the spot extracted from the market. As mentioned before, this quantity is $\frac{\partial^2 C_{LV}}{\partial K^2}$ and can be computed from the PDE solution immediately and accurately. As we will use the density in the calibration formula in Theorem 3.3.1, we want it to be smooth and accurate. Figure 3.4.5 shows that quantity.

3.4.7 Two-factor Heston-type LSV model calibration by PDE

The next step in our algorithm is to calibrate the 2-factor sub-model of (3.2.1) defined in (3.2.2), which we recall below. In this subsection only, we write S,V and α in lieu of S^{2D}, V^{2D} and α^{2D},

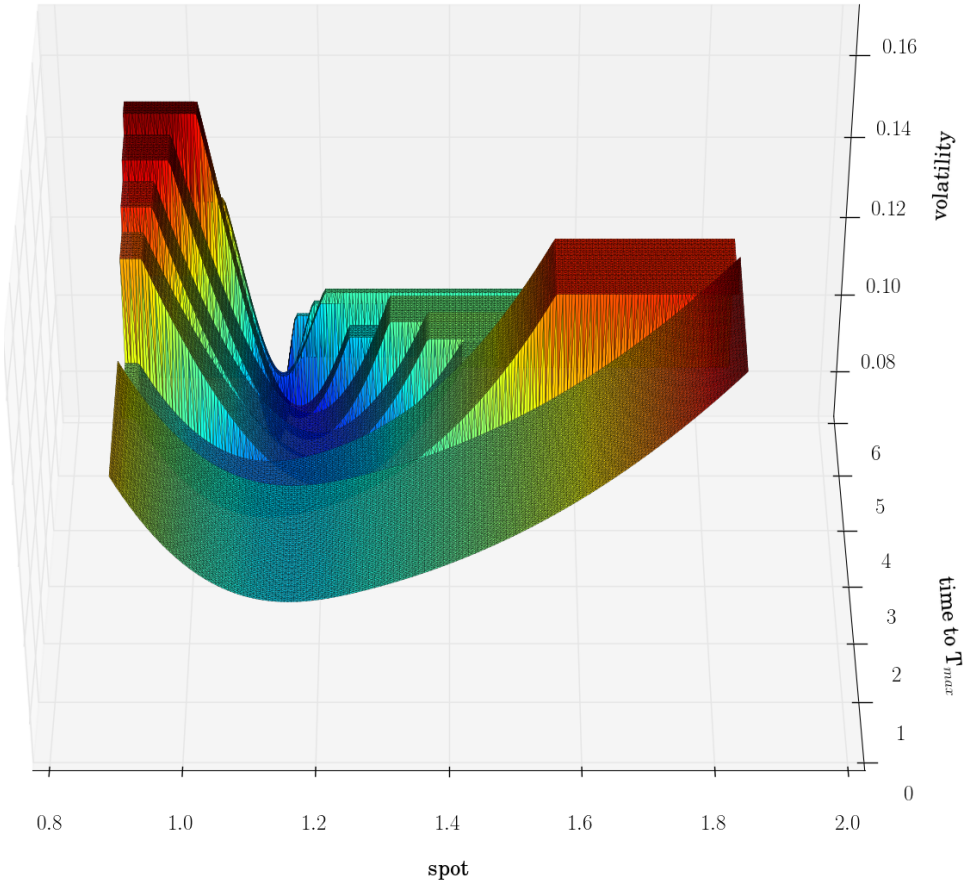


Figure 3.4.4: EURUSD Local volatility function calibrated by forward PDE and fixed-point algorithm

respectively, to ease notation. Hence,

$$\begin{cases} \dfrac{dS_t}{S_t} = \left(\bar{r}^d(t) - \bar{r}^f(t)\right) dt + \alpha(S_t, t)\sqrt{V_t}\, dW_t\,, \\ dV_t = \kappa(\theta - V_t)\, dt + \xi\sqrt{V_t}\, dW_t^V\,. \end{cases}$$

We use the Heston parameters calibrated in Subsection 3.4.5, i.e.,

$$v_0 = 0.0094, \quad \theta = 0.0137, \quad \kappa = 1.4124, \quad \rho = -0.1194, \quad \xi = 0.2988\,,$$

where the Feller ratio is

$$\frac{2\kappa\theta}{\xi^2} \approx 0.4335 < 1\,,$$

which clearly breaks the Feller condition. The calibration of these models by PDEs is extensively discussed in [129, 35, 152] and the approach proves to be very effective when the variance process has almost surely strictly positive paths. However, for Heston type models, the CIR variance process

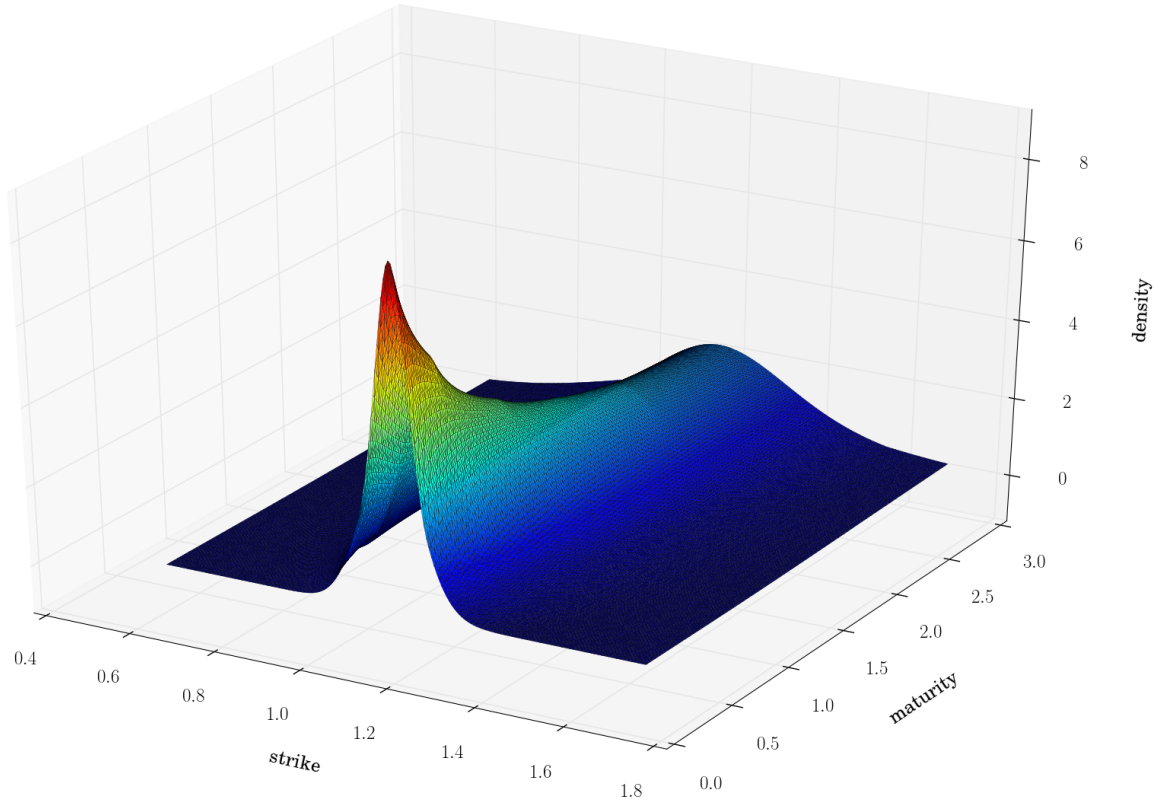


Figure 3.4.5: Market spot marginal density computed from the Dupire forward PDE with calibrated σ_{LV}

can eventually reach zero if the Feller condition is violated. And as a matter of fact, in FX markets, this condition is very often violated. One can refer to Table 6.5 in [35] for examples on a large range of currency pairs and maturities. As a consequence, mass of the variance process will accumulate towards $V = 0$ and the density will explode at this boundary. In [140] the authors propose to reduce the problem by considering $\log(V_t/v_0)$, whereas [35] suggests to refine the mesh near $V = 0$. While these methods can provide acceptable numerical results, they do not entirely solve the problem since $\log(V_t/v_0)$ can reach $-\infty$ with non-zero probability. Hence, in this work, we propose to use a different boundary condition as well as a change of variable to control the solution in the neighbourhood of $V = 0$ such that the density is finite on the boundary.

Transformation and weak formulation

The following is a small variation of the main result in [107], and the proof is therefore omitted. Note that a new non-Dirichlet boundary condition appears at $z = 0$.

Theorem 3.6. *Define the region* $\Omega = \{(x, z) \in \mathbb{R}^2_+\}$ *and assume that the density* ϕ *(under* $\mathbb{Q}^d$*) of*

the Markovian process (S_t, V_t) *started at* (S_0, v_0) *at time 0 exists and is* $C^{1,2,2}(\mathbb{R}^+ \times \Omega)$*. Then* ϕ *is the solution to the Kolmogorov forward equation*

$$\begin{cases} \frac{\partial \phi}{\partial t} + \left(\bar{r}^d(t) - \bar{r}^f(t)\right)\frac{\partial (x\phi)}{\partial x} + \frac{\partial(\kappa(\theta - z)\phi)}{\partial z} \\ \qquad - \frac{1}{2}\left(\frac{\partial^2\left(x^2\alpha^2(x,t)z\phi\right)}{\partial x^2} + \frac{\partial^2\left(\xi^2 z\phi\right)}{\partial z^2} + 2\frac{\partial^2(\rho\xi x\alpha(x,t)z\phi)}{\partial x \partial z}\right) = 0, & (x,z) \in \Omega,\, t > 0\,, \\ \left(\frac{\xi^2}{2}\frac{\partial(z\phi)}{\partial z} - \kappa(\theta - z)\phi + \rho\xi z\frac{\partial(x\alpha(x,t)\phi)}{\partial x}\right)\Big|_{z=0} = 0, & z = 0,\, x \geqslant 0,\, t > 0\,, \\ \lim_{z\to\infty}\phi(x,z,t) = \lim_{x\to\infty}\phi(x,z,t) = \phi(0,z,t) = 0, & (x,z) \in \Omega, t > 0\,, \\ \lim_{t\to 0}\phi(x,y,z,t) = \delta(x - S_0, z - v_0), & (x,z) \in \Omega\,. \end{cases} \tag{3.4.26}$$

Proof. Similar to the proof of Lemma 4.1 and Theorem 4.1 in [107]. □

The density function for a CIR process is written as follows (see [39]),

$$\phi_v(z,t) = ce^{-u-cz}\left(\frac{cz}{u}\right)^{\beta/2} I_\beta(2\sqrt{cuz})\,,$$

with $c = \frac{2\kappa}{(1-e^{-\kappa t})\xi^2}$, $u = cv_0e^{-\kappa t}$ and $\beta = \frac{2\kappa\theta}{\xi^2} - 1$, where I_β is the modified Bessel function of the first kind of order β. We can write the asymptotic expression for small z using the asymptotic formula for the modified Bessel function found in [1]. We get, for z around 0^+,

$$\phi_v(z,t) \quad \sim \quad \frac{c^{(\beta+1)}e^{-u-cz}}{\Gamma(\beta+1)}z^\beta\,, \tag{3.4.27}$$

such that $\phi_v(z,t)$ diverges for $z = 0^+$ when $2\kappa\theta < \xi^2$. This is in line with the well-known stationary distribution density ϕ_v^∞ (see [39]) defined as

$$\phi_v^\infty(z) = \lim_{t\to\infty}\phi_v(z,t) = \frac{\omega^{(\beta+1)}z^\beta e^{-\omega z}}{\Gamma(\beta+1)}\,, \tag{3.4.28}$$

with $\omega = \frac{2\kappa}{\xi^2}$. This justifies a change of variable for the density ϕ of the form $p = \phi z^{-\beta}$, and to solve a new PDE for p that we define hereafter. We can insert $z^\beta p$ and its associated derivatives in (3.4.26). By straightforward calculus we get the following.

Corollary 3.7. *The dampened "density" $p = \phi z^{-\beta}$ satisfies the initial boundary value problem*

$$\begin{cases} \frac{\partial p}{\partial t} + (\bar{r}^d(t) - \bar{r}^f(t))\frac{\partial xp}{\partial x} - \beta\kappa p \\ \qquad + \frac{\partial\kappa(\theta - z)p}{\partial z} - (\beta+1)\Big(\frac{\partial\xi^2 p}{\partial z} + \frac{\partial\rho\xi x\alpha(x,t)p}{\partial x}\Big) \\ \qquad - \frac{1}{2}z\Big[\frac{\partial^2 x^2\alpha^2(x,t)p}{\partial x^2} + \frac{\partial^2\xi^2 p}{\partial z^2} + 2\frac{\partial^2\rho\xi x\alpha(x,t)p}{\partial x\partial z}\Big] = 0\,, & (x,z)\in\Omega,\, t>0\,, \\ \frac{\xi^2 z}{2}\frac{\partial p}{\partial z}\Big|_{z=0} + \kappa z p\big|_{z=0} + \rho\xi z\,\frac{\partial x\alpha(x,t)p}{\partial x}\Big|_{z=0} = 0, & z=0,\, t>0\,, \\ \lim_{z\to\infty} p(x,z,t) = \lim_{x\to\infty} p(x,z,t) = p(0,z,t) = 0, & (x,z)\in\Omega,\, z\neq 0,\, t>0\,, \\ \lim_{t\to 0} p(x,z,t) = z^{-\beta}\delta(x - S_0, z - v_0), & (x,z)\in\Omega\,, \end{cases} \tag{3.4.29}$$

with $\beta = \frac{2\kappa\theta}{\xi^2} - 1$.

While this PDE is easier to handle numerically, one wants to work with the original density function ϕ for most of the applications. There are two main calculations one would like to achieve:

- compute the price of a derivative with final payoff value $f(S_T)$ for a given function f as

$$\mathbb{E}[D_T f(S_T)] = D_T \int_0^\infty f(x) \int_0^\infty \phi(x,z,T)\,dz\,dx\,;$$

- compute the Markovian projection

$$\mathbb{E}[V_T \,|\, S_T = K] = \frac{\int_0^\infty z\phi(K,z,T)\,dz}{\int_0^\infty \phi(K,z,T)\,dz}\,.$$

As ϕ is still intractable for small z and computing $z^\beta p(x,z,t)$ is not numerically feasible, we perform an integration by parts. The trade-off is a loss of accuracy due to the computation of $\frac{\partial p}{\partial z}$.

Let us consider the first case above where we are provided with a payoff function f. Using the change of variable solution p, and integrating by parts, we get

$$\begin{aligned} \int_0^\infty f(x)\int_0^\infty \phi(x,z,T)\,dz\,dx &= \int_0^\infty f(x)\int_0^\infty z^\beta p(x,z,T)\,dz\,dx\,, \\ &= -\int_0^\infty f(x)\int_0^\infty \frac{z^{\beta+1}}{\beta+1}\frac{\partial p(x,z,T)}{\partial z}\,dz\,dx\,, \end{aligned}$$

since $\beta + 1 > 0$ and $\lim_{z\to 0} z^{\beta+1}p = \lim_{z\to 0} z^{\beta+1}p = 0$ due to p being finite. Then we can compute the two quantities above very efficiently using the numerically stable solution p as follows:

$$\mathbb{E}[D_T f(S_T)] = -D_T \int_0^\infty f(x)\int_0^\infty \frac{z^{\beta+1}}{\beta+1}\frac{\partial p(x,z,T)}{\partial z}\,dz\,dx \tag{3.4.30}$$

and

$$\mathbb{E}[V_T \,|\, S_T = K] = -(\beta+1)\frac{\int_0^\infty z^{\beta+1} p(K,z,T)\,dz}{\int_0^\infty z^{\beta+1}\frac{\partial p(K,z,T)}{\partial v}\,dz}\,. \tag{3.4.31}$$

Finite element method with a two-step BDF time scheme

To the best of our knowledge, most of the literature on the subject (see for instance [35, 129, 152]) relies on ADI schemes to solve the Fokker-Planck PDE in Theorem 3.6. In this work, we combine a finite element solution with a *Backward Differentiation Formula* (BDF) scheme in time. Thanks to the weak form of the problem, this will allow us to incorporate the Dirac delta initial condition exactly (without using a limiting density function or other smoothing procedure). Furthermore, since Crank–Nicolson time-stepping or ADI schemes can give rise to instabilities (see [124, 151]), we use a two-step implicit BDF scheme instead to get the same second-order convergence in time while preserving stability (we refer to [61] for some financial applications of BDF schemes).

Equation (3.4.29) can be written as

$$\frac{\partial \phi}{\partial t} - z\nabla\cdot\mathbf{u} + \nabla\cdot(\mathbf{b}_1 - (\beta+1)\mathbf{b}_2) + \mu_t p = 0\,, \tag{3.4.32}$$

with

$$\mathbf{u} = \frac{1}{2}\begin{bmatrix} \frac{\partial\left(x^2\alpha^2(x,t)p\right)}{\partial x} + \frac{\partial(\rho\xi x\alpha(x,t)p)}{\partial z} \\ \frac{\partial\left(\xi^2 p\right)}{\partial z} + \frac{\partial(\rho\xi x\alpha(x,t)p)}{\partial x} \end{bmatrix},\ \mathbf{b}_1 = \begin{bmatrix} (\bar{r}^d(t) - \bar{r}^f(t))x \\ \kappa(\theta - z) \end{bmatrix},\ \mathbf{b}_2 = \begin{bmatrix} \rho\xi x\alpha(x,t)p \\ \xi^2 p \end{bmatrix},\ \mu_t = -\beta\kappa\,.$$

We denote by $\partial\Omega$ the boundary of Ω, by Γ_D the subset with Dirichlet boundary conditions and by Γ_R the subset with Robin boundary condition, so that $\partial\Omega = \Gamma_D \cup \Gamma_R$ where $\Gamma_R = \{(x,z) \in \partial\Omega :\ z = 0\}$. We derive a weak formulation in the usual way (see, e.g., [126]), i.e., we multiply the PDE by a test function $v \in H^1(\Omega)$, integrate over Ω, using the divergence theorem and boundary conditions, to obtain the weak form of (3.4.32),

$$\int_\Omega \frac{\partial p}{\partial t} v\,d\Omega + a(p,v) \quad = \quad 0\,,$$

with the bi-linear form

$$\begin{aligned} a(p,v) \quad = \quad & \int_\Omega \mathbf{u}\cdot\nabla z v\,d\Omega + \int_\Omega \nabla\cdot(\mathbf{b}_1 - (\beta+1)\mathbf{b}_2)v\,d\Omega + \int_\Omega \mu_t p v\,d\Omega \\ & - \int_{\Gamma_R}\left(\kappa z p + \frac{1}{2}z\frac{\partial(\rho\xi x\alpha(x,t)p)}{\partial x}\right)v\,d\Gamma_R\,, \end{aligned}$$

where the last term contains the new boundary condition.

The time scheme will be handled by a two-step BDF scheme. Let us define a uniform time discretisation $m \in [\![0, M]\!]$ with $t_m = m\Delta_t$. We denote $p_m = p(x, z, t_m)$, in which case the BDF scheme can be written as

$$\int_\Omega \left(p_{m+2} - \frac{4}{3} p_{m+1} + \frac{1}{3} p_m \right) v \, d\Omega + \frac{2}{3} \Delta_t a(p_{m+2}, v) = 0 \,,$$

where the first time step is divided into two standard fully implicit time steps (see [70]). At $t = 0$, the Dirac delta initial condition is handled easily via

$$\begin{aligned} \int_\Omega p_0 v \, d\Omega &= \int_\Omega z^{-\beta} \delta(x - S_0, z - v_0) v \, d\Omega \\ &= v_0^{-\beta} v(S_0, v_0) \,. \end{aligned}$$

Then, for the first time step, we need to solve for p_1 such that

$$\int_\Omega p_1 v \, d\Omega + \Delta_t a(p_1, v) = \int_\Omega p_0 v \, d\Omega = v_0^{-\beta} v(S_0, v_0) \,,$$

which is equivalent to solving a linear system where the right-hand side vector is $v_0^{-\beta}$ for the source point node (to ensure that the point (S_0, v_0) is on the mesh) and zero otherwise.

Finite element mesh construction

For an accurate finite element solution, the mesh is of paramount importance. In order to refine the mesh in the most relevant area, we use an exponential mesh on the variance axis and a hyperbolic mesh (see [148]) in the spot direction. This makes the mesh finer around $z = 0$ and $x = S_0$. In order to build our mesh, we first define the grids in spot $(x_i)_{i \in [\![0, N_S]\!]}$ and variance $(z_j)_{j \in [\![0, N_V]\!]}$ separately. Additionally, to solve the PDE numerically, we need to truncate at the boundary and use $\bar{\Omega} = \{(x, z) \in [0, S_{max}] \times [0, V_{max}]\}$ on a time interval $[0, T]$. We choose

$$S_{max} = S_0 e^{5\alpha(S_0, T)\sqrt{(v_0 e^{-\kappa T} + \theta(1 - e^{-\kappa T}))T}}$$

and recall that the stationary distribution of the CIR process is a gamma distribution of density ϕ_v^∞ as defined in (3.4.28). We compute V_{max} with the inverse cumulative density function such that

$$\mathbb{P}(z > V_{max}) = 0.01\% \,.$$

We write

$$\begin{aligned} x_i &= f_h(g_h(\bar{x}_i)) \,, \\ z_j &= f_e(g_e(\bar{z}_j)) \,, \end{aligned}$$

with

$$\begin{aligned} f_h(x) &= S_0 + b\sinh(\nu x + d), \\ b &= \eta(S_{max} - S_{min}), \\ d &= \operatorname{arcsinh}\left(\frac{S_{min} - S_0}{b}\right), \\ \nu &= \operatorname{arcsinh}\left(\frac{S_{max} - S_0}{b}\right) - d, \\ \bar{x}_i &= \frac{i}{(N_S+1)}, \\ \eta &= 0.02\,, \end{aligned}$$

$$\begin{aligned} f_e(z) &= c + c\,\exp(\lambda z), \\ c &= \frac{V_{max}}{e^{\lambda}-1}\,, \\ \lambda &= \max\left(1, 4 - \frac{3\kappa\theta}{\xi^2}\right), \\ \bar{z}_j &= \frac{j}{(N_V+1)}\,, \end{aligned}$$

where η is defined according to our numerical experiments and g_h is the quadratic polynomial that passes through the points $(0,0)$, $(1,1)$,

$$\left(\frac{\left\lfloor f_h^{-1}(S_0)(N_S+1)+0.5\right\rfloor}{N_S+1}, f_h^{-1}(S_0)\right).$$

where λ is defined according to our numerical experiments and g_h is the quadratic polynomial that passes through the points $(0,0)$, $(1,1)$,

$$\left(\frac{\lfloor f_e^{-1}(v_0)(N_V+1)+0.5\rfloor}{N_V+1}, f_e^{-1}(v_0)\right).$$

The latter intermediate step makes sure that both S_0 and v_0 are vertices of their respective grids. The construction of the finite element triangular mesh can be achieved by creating a vertex at each point (x_i, z_j) and defining two triangular cells (upper left and lower right) in each rectangle. The finite element solution is solved with $\mathbb{P}_2$ elements, i.e., a polynomial of order two on a triangle cell. Each triangle is characterised by 6 local degrees of freedom (nodes) as displayed in Figure 3.4.6 (see [116] for details). An example of a generated mesh with 30 spot steps and 30 variance steps is illustrated in Figure 3.4.7.

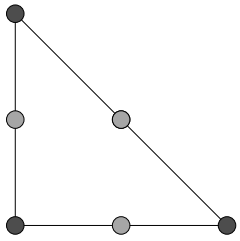


Figure 3.4.6: $\mathbb{P}_2$ element with 6 degrees of freedom

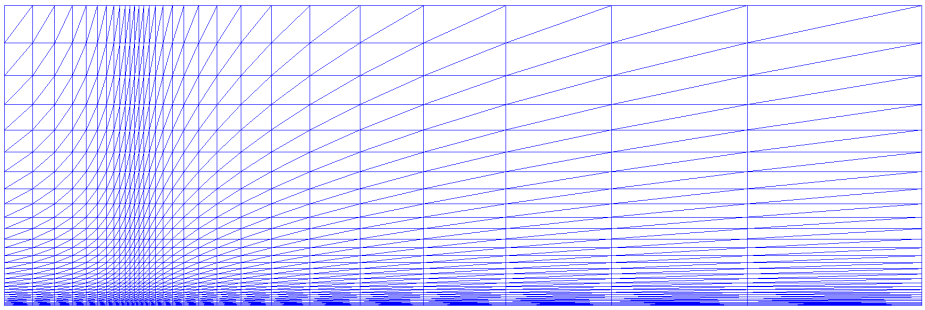


Figure 3.4.7: Finite element triangular mesh refined around $x = S_0$ and $z = 0$

In order to see the improvement due to the changed-of-variable PDE (3.4.29) for p, we plot both

$v^{-\beta}\phi$ and p for a pure Heston model with

$$r = 3\%, \quad q = 1\%, \quad \kappa = 1, \quad \theta = v_0 = 0.04, \quad \rho = -0.3, \quad \xi = 0.5, \quad t = 1\,,$$

which corresponds to a Feller ratio of 0.32. We use 100 time steps as well as 30 spot steps and 30 variance steps. The solution for $p = v^{-\beta}\phi$ computed with no change of variables is presented in Figure 3.4.8 where we notice significant numerical instabilities. The change of variables PDE is solved for the same problem and p is plotted in Figure 3.4.9, where the improvement is clear and the solution is smooth and accurate.

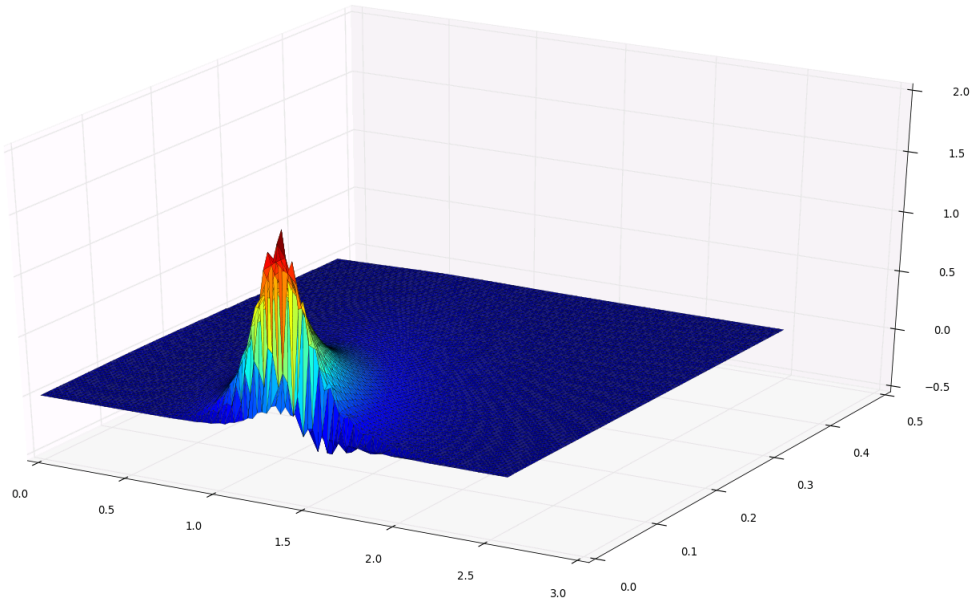


Figure 3.4.8: $p = v^{-\beta}\phi$ using ϕ computed with the standard PDE

Calibration algorithm

The calibration of the 2-factor LSV model (3.2.2) is performed by finding the leverage function α defined in (3.3.1). In the case of deterministic rates, we know from (3.3.4) that

$$\alpha(K,T) = \frac{\sigma_{LV}(K,T)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[V_T \mid S_T = K]}}\,. \tag{3.4.33}$$

Hence, we need to compute $\mathbb{E}^{\mathbb{Q}}[V_T \mid S_T = K]$. The conditional expectation can be written as a function of p, the solution of the dampened PDE (3.4.29) as given in (3.4.31). Both integrals can be computed by double adaptive Clenshaw-Curtis quadrature rules to handle singularities properly when T is small (see [74]).

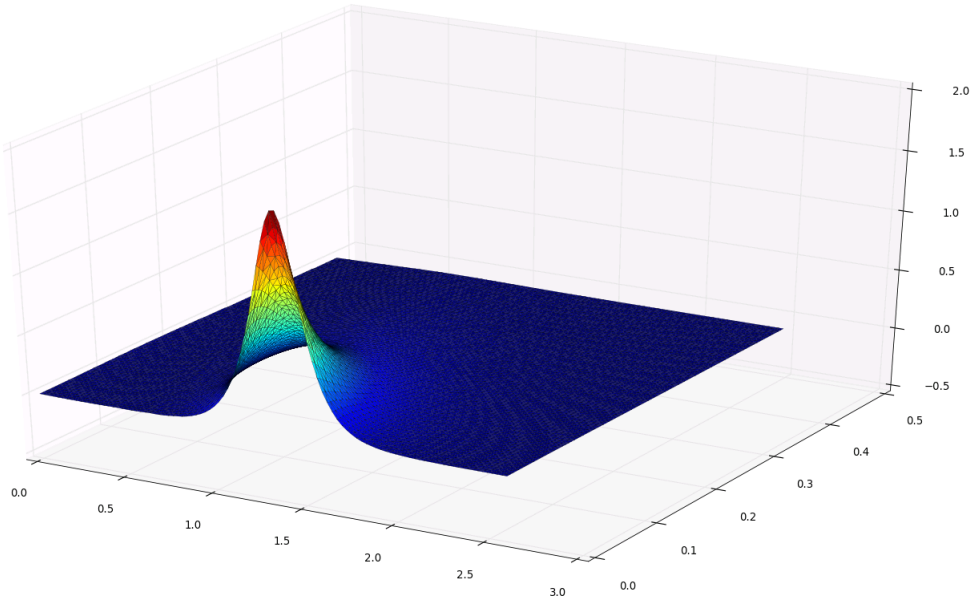


Figure 3.4.9: $p = v^{-\beta}\phi$ computed directly from the change of variables PDE

Furthermore, for very small or very large values of K, both the numerator and denominator will be very small. So we define a smooth extrapolation rule as

$$\mathbb{E}^{\mathbb{Q}}[V_T \,|\, S_T = K] = \frac{(\beta+1)\Big(\int_0^\infty z^{\beta+1} p(K,z,T)\,dz + \epsilon\Big(v_0 e^{-\kappa T} + (\theta - v_0)\Big(1 - e^{-\kappa T}\Big)\Big)\Big)}{\epsilon(\beta+1) - \int_0^\infty z^{\beta+1} \frac{\partial p(K,z,T)}{\partial v}\,dz},$$

where we pick $\epsilon = 10^{-14}$ in our numerical tests.

The calibration will be done forward in time. Denote by $(\Delta T(i))_{i \le N_T}$ the step size and by N_T the number of steps. We use a BDF scheme with constant steps and find that 50 time steps per year and a 80×80 spot-variance finite element mesh give very accurate results.

The leverage function α is defined on a grid of points interpolated with cubic splines in spot and forward-flat in time. There are N_T volatility slices in total such that we denote the i-th slice, i.e., $\alpha(K, T_i)$, by $\alpha_i(K)$. Here, α_i is defined on a grid of N_S points (we find 25 points to provide a sufficiently high numerical accuracy). For a given time slice T_i, the leverage function is defined on the interval $[S^i_{min}, S^i_{max}]$ and is extrapolated flat outside these bounds. This approach allows us to compute $\mathbb{E}^{\mathbb{Q}}[V_T \,|\, S_T = K]$ only on the grid points of α, reducing the computational time considerably. Because we need more grid points around the forward value and fewer around S^i_{min} and S^i_{max}, we will use a hyperbolic grid (with $\eta = 0.05$, see Subsection 3.4.7 for more details) refined around the forward value

$$F_i = S_0 e^{\int_0^{T_i} \left(\bar{r}^d(t) - \bar{r}^f(t)\right) dt}$$

with

$$S^i_{min} = F_i e^{-\frac{8}{2}\sigma_F(T_i)\sqrt{T_i}}, \quad S^i_{max} = F_i e^{\frac{8}{2}\sigma_F(T_i)\sqrt{T_i}},$$

where $\sigma_F(T_i)$ is the at-the-money forward market volatility for maturity T_i (interpolated linearly in variance). Each of the grid values can be seen as a parameter and we denote them by $(\alpha_{i,j})_{i\leq N_T, j\leq N_S}$, with the associated spot grid points $(s_{i,j})_{i\leq N_T, j\leq N_S}$. We can then write the calibration procedure which we present in Algorithm 3.3.

Algorithm 3.3 $\alpha(s,T)$ calibration with the dampened Kolmogorov forward PDE

$\alpha(s,0) = \frac{\sigma_{LV}(s,0)}{\sqrt{v_0}}$
$T = 0$
for $(\ i = 1\,;\, i \leq N_T\,;\, i++)$ **do**
  solve (3.4.29) for p on $[T, T+\Delta T]$ with $\alpha(s,[T,T+\Delta T(i)]) = \alpha(s,T)$
  for $(\ j = 1\,;\, j \leq N_S\,;\, j++)$ **do**

$$E_{V_T} = -(\beta+1)\frac{\int_0^\infty z^{\beta+1} p(s_{i,j}, z, T+\Delta T(i))\, dz}{\int_0^\infty z^{\beta+1} \frac{\partial p(s_{i,j},z,T+\Delta T(i))}{\partial z}\, dz}$$

$$\alpha_{i,j} = \frac{\sigma_{LV}(s_{i,j}, T+\Delta T(i))}{\sqrt{E_{V_T}}}$$

  end for
  $T = T + \Delta T(i)$
end for

In the calibration routine, we use forward-flat interpolation in time for the leverage function to handle the non-linearity of the problem. In between two maturity pillars, where the local volatility is flat, one would expect the exact leverage function to be continuous in time. To account for this, we switch the time interpolation of the leverage function α to linear at the end of the calibration process. The calibrated leverage function α is plotted in Figure 3.4.10.

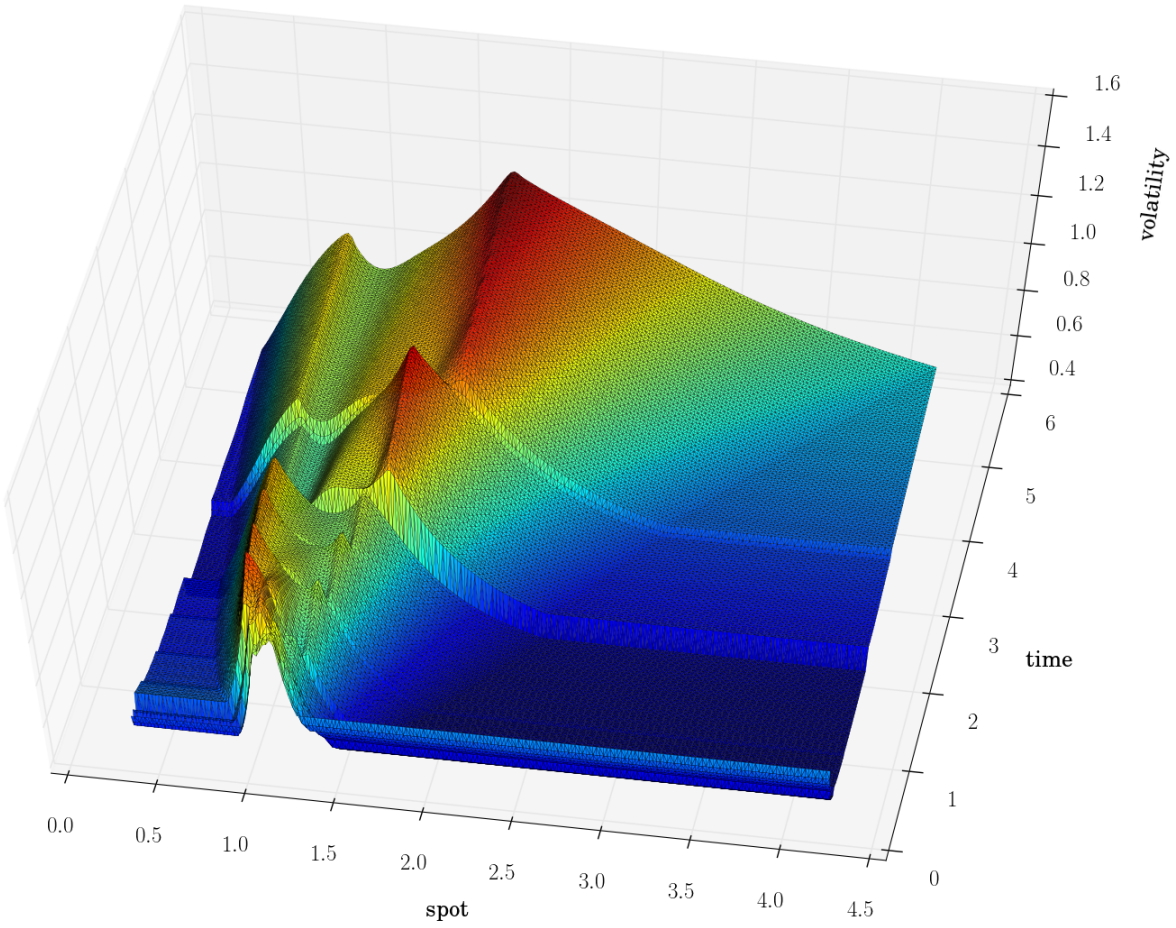


Figure 3.4.10: Calibrated leverage function α for the 2-factor LSV model (3.2.2)

3.5 Four-factor Heston-type LSV 2CIR++ model calibration

In this section, we discuss the calibration algorithm for the main model (3.2.1), which we recall below

$$\begin{cases} \dfrac{dS_t}{S_t} = \left(r_t^d - r_t^f\right) dt + \alpha(S_t, t)\sqrt{V_t}\, dW_t \\ r_t^d = g_t^d + h^d(t) \\ r_t^f = g_t^f + h^f(t) \\ dg_t^d = \kappa_d\left(\theta_d - g_t^d\right) dt + \xi_d\sqrt{g_t^d}\, dW_t^{g^d} \\ dg_t^f = \left(\kappa_f\left(\theta_f - g_t^f\right) - \rho_{Sf}\xi_f\sqrt{g_t^f}\,\alpha(S_t, t)\sqrt{V_t}\right) dt + \xi_f\sqrt{g_t^f}\, dW_t^{g^f} \\ dV_t = \kappa(\theta - V_t)\, dt + \xi\sqrt{V_t}\, dW_t^V\,, \end{cases} \tag{3.5.1}$$

with the correlation structure $d\langle W_t, W_t^V\rangle = \rho\, dt$, $d\langle W_t, W_t^{g^d}\rangle = \rho_{Sd}\, dt$, and $d\langle W_t, W_t^{g^f}\rangle = \rho_{Sf}\, dt$, all other correlations being zero. We will consider the additional model simplification $\rho_{Sd} = \rho_{Sf} = 0$ such that there is no quanto drift adjustment in the foreign short rate dynamics. This allows the use of an extended closed-form formula of [4] presented in Subsection 3.4.5 when no leverage function appears.

Otherwise, one can extend the approximation formulae of [77] from a 3-factor Heston–CIR to a 4-factor Heston–2CIR++ model, and retrieve the historical correlations between the FX spot and the short-term zero-coupon bonds for ρ_{Sd} and ρ_{Sf}. Alternatively, one can assume $\rho_{Sd} = \rho_{Sf} = 0$ for Subsection 3.4.5 and perform an initial guess calibration of the stochastic volatility parameters $b_0 = (v_0, \kappa, \theta, \rho, \xi)$. Now, we compute the exact model price C_{MC} at b_0 with non-zero ρ_{sd} and ρ_{sf} using Monte Carlo with the closed-form call price C_{A} (with $\rho_{Sd} = \rho_{Sf} = 0$) as a control variate, and define the difference between the two prices as $\Delta_0(K,T) = C_{\mathrm{MC}} - C_{\mathrm{A}}$. Assuming the absolute values of the two correlations are small, we calibrate the parameter vector $b = (v_0, \kappa, \theta, \rho, \xi)$ on a small interval around b_0 and approximate the call prices by $C_{\mathrm{A}}(b) + \Delta_0$. The calibration algorithm in this section is generic and can be applied to a non-trivial correlation structure.

We assume that both CIR++ processes are calibrated under their own risk-neutral measure and that the Heston–2CIR++ model is calibrated as in Subsection 3.4.5. The parameters are as follows

$$\begin{cases} v_0 = 0.0094, \quad \theta = 0.0137, \quad \kappa = 1.4124, \quad \rho = -0.1194, \quad \xi = 0.2988\,, \\ g_0^d = 0.0001, \quad \theta_d = 0.5469, \quad \kappa_d = 0.0837, \quad \rho_{Sd} = 0, \quad \xi_d = 0.0274\,, \\ g_0^f = 0.0001, \quad \theta_f = 1.1656, \quad \kappa_f = 0.0110, \quad \rho_{Sf} = 0, \quad \xi_f = 0.0370\,. \end{cases}$$

In order to run the particle method, we use the Monte Carlo QE-scheme time marching for model (3.2.1).

3.5.1 Monte Carlo QE-scheme

The *Quadratic-Exponential (QE)* scheme [7] uses moment-matching techniques and can significantly reduce the Monte Carlo discretisation error. While the full truncation Euler scheme and the QE scheme have been shown to perform well in our tests, we experienced a faster convergence in time for the QE scheme when the Feller condition is broken. Hence, we choose the QE scheme for the variance process and the full truncation Euler for both stochastic rates, as the computational cost will be smaller. We briefly write a generalised QE scheme based on the original scheme from [7] to incorporate a leverage function and stochastic rates in the discretisation.

We follow our time interpolation rule for the calibration of α and interpolate forward-flat in time. We assume for simplicity that each Monte Carlo time step belongs to the α time grid. We can write

$$V_{t+dt} = V_t + \int_t^{t+\Delta t} \kappa(\theta - V_u)\, du + \xi \int_t^{t+\Delta t} \sqrt{V_u}\, dW_u^V\,,$$

and hence

$$\int_t^{t+\Delta t} \sqrt{V_u}\, dW_u^V = \frac{V_{t+dt} - V_t - \int_t^{t+\Delta t} \kappa(\theta - V_u)\, du}{\xi},$$

and

$$d\ln S_t = \left(r_t^d - r_t^f - \frac{1}{2}\alpha^2(S_t,t)V_t\right) dt + \alpha(S_t,t)\rho\sqrt{V_t}\, dW_t^V + \alpha(S_t,t)\sqrt{1-\rho^2}\sqrt{V_t}\, dW_t^S,$$

where W_t^S is a Brownian motion independent of W_t^V. Therefore,

$$\begin{aligned}
\ln S_{t+\Delta t} &= \ln S_t + \int_t^{t+\Delta t} \left(r_u^d - r_u^f\right) du - \frac{1}{2}\alpha^2(S_t,t)\int_t^{t+\Delta t} V_u\, du \\
&+ \frac{\alpha(S_t,t)\rho\Big(V_{t+dt} - V_t - \kappa\theta\Delta t + \kappa\int_t^{t+\Delta t} V_u\, du\Big)}{\xi} \\
&+ \alpha(S_t,t)\sqrt{1-\rho^2}\int_t^{t+\Delta t} \sqrt{V_u}\, dW_u^S.
\end{aligned}$$

We approximate $\int_t^{t+\Delta t} V_u\, du$ by $\left(\frac{V_{t+\Delta t}+V_t}{2}\right)\Delta t$, and note that conditional on V_t and $\int_t^{t+\Delta t} V_u\, du$, since W_u^V and W_u^S are independent, the Itô integral $\int_t^{t+\Delta t} \sqrt{V_u}\, dW_u^S$ is normally distributed with mean zero and variance $\int_t^{t+\Delta t} V_u\, du$. We write the full scheme below

$$\begin{cases}
g^d_{t+\Delta t} & = g^d_t + \kappa_d\Big(\theta_d - \left(g^d_t\right)^+\Big)\Delta t + \xi_d\sqrt{\left(g^d_t\right)^+}\sqrt{\Delta t}\, Z_d \\
g^f_{t+\Delta t} & = g^f_t + \kappa_f\Big(\theta_f - \left(g^f_t\right)^+\Big)\Delta t + \xi_f\sqrt{\left(g^f_t\right)^+}\sqrt{\Delta t}\, Z_f \\
\ln S_{t+\Delta t} & = \ln S_t + \left(\frac{\left(r^d_{t+\Delta t} - r^f_{t+\Delta t}\right)+\left(r^d_t - r^f_t\right)}{2} - \frac{1}{4}\alpha^2(S_t,t)(V_{t+\Delta t}+V_t)\right)\Delta t \\
& \quad + \frac{\alpha(S_t,t)\rho\left(V_{t+\Delta t} - V_t + \kappa\left(\frac{V_{t+\Delta t}+V_t}{2} - \theta\right)\Delta t\right)}{\xi} \\
& \quad + \alpha(S_t,t)\sqrt{1-\rho^2}\sqrt{\frac{V_{t+\Delta t}+V_t}{2}}\sqrt{\Delta t}\, Z \\
V_{t+\Delta t}: & \quad \begin{cases}
\text{if } \psi \le \psi_c & : V_{t+\Delta t} = a(b+Z_v)^2 \\
\text{else} & \begin{cases}
\text{if} \quad U \le p & : V_{t+\Delta t} = 0 \\
\text{else} & : V_{t+\Delta t} = \ln\left(\frac{1-p}{1-U}\right)\frac{m}{1-p},
\end{cases}
\end{cases}
\end{cases} \tag{3.5.2}$$

with

$$\begin{cases}
m & = \theta + (V_t - \theta)e^{-\kappa\Delta t}, \\
\gamma^2 & = \frac{V_t\xi^2 e^{-\kappa\Delta t}}{\kappa}\left(1 - e^{-\kappa\Delta t}\right) + \frac{\theta\xi^2}{2\kappa}\left(1 - e^{-\kappa\Delta t}\right)^2, \\
\psi & = \frac{\gamma^2}{m^2}, \quad p = \frac{\psi-1}{\psi+1}, \quad \beta = \frac{1-p}{m}, \\
b^2 & = \frac{2}{\psi} - 1 + \sqrt{\frac{2}{\psi}}\sqrt{\frac{2}{\psi} - 1}, \quad a = \frac{m}{1+b^2}, \\
\psi_c & = 1.5,
\end{cases}$$

where Z, Z_v, Z_d, Z_f are independent draws from a standard normal distribution and U is a draw from a uniform distribution.

3.5.2 Calibration algorithm with the control variate particle method

We recall from (3.3.1) that the leverage function has to satisfy

$$\alpha^2(K,T) = \frac{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \,|\, S_T = K\right]}{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d V_T \,|\, S_T = K\right]}\left(\sigma_{LV}(K,T)^2 + \frac{\mathbb{E}^{\mathbb{Q}^d}[Q_T]}{\frac{1}{2}K^2\frac{\partial^2 C_{LV}}{\partial K^2}}\right), \tag{3.5.3}$$

where

$$Q_T = D_T^d\Big(r_T^f - \bar{r}^f(T)\Big)(S_T - K)^+ - KD_T^d\mathbf{1}_{S_T\geq K}\Big[\Big(r_T^d - \bar{r}^d(T)\Big) - \Big(r_T^f - \bar{r}^f(T)\Big)\Big],$$

which involves conditional expectations as well as standard expectations. Through these expectations, α depends on the distribution of the joint process $X_t = \left(S_t, V_t, r_t^d, r_t^f\right)$. The process thus falls in the class of McKean-Vlasov processes [108], and the driving SDE is non-linear. Similarly, the joint density of the process X_t is driven by a non-linear 4-dimensional Kolmogorov forward PDE. The particle method was introduced in [108] and is discussed in detail in [139]; it was applied to LSV model calibration in [82, 83] in order to tackle the non-linearity of the problem.

We consider N-sample paths $\left(X_t^i\right)_{i\in[\![1,N]\!]} = \left(S_t^i, V_t^i, r_t^{d,i}, r_t^{f,i}\right)_{i\in[\![1,N]\!]}$ of the process X_t. Let Q_T^i be the i-th sample of Q_T. Then α can be estimated by

$$\hat{\alpha}_N\left(K,T,\left(X_t^i\right)_{i\leq N}\right) = \sqrt{\frac{\sigma_{LV}(K,T)^2}{\hat{p}_N(K,T)} + \frac{\hat{q}_N(K,T)}{\frac{1}{2}\hat{p}_N(K,T)K^2\frac{\partial^2 C_{LV}}{\partial K^2}}},$$

with

$$\hat{q}_N(K,T) = \frac{1}{N}\sum_{i=1}^N Q_T^i$$

and

$$\hat{p}_N(K,T) = \frac{\sum_{i=1}^N D_T^i V_T^i \delta_N(S_T^i - K)}{\sum_{i=1}^N D_T^i \delta_N(S_T^i - K)},$$

which is an estimator for

$$\frac{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d V_T \,|\, S_T = K\right]}{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \,|\, S_T = K\right]},$$

with δ_N a kernel function. The $(4\times N)$-dimensional SDE driving the system $\left(X_t^i\right)_{i\leq N}$ can be written

as

$$\begin{cases} \dfrac{dS_t^i}{S_t^i} = \left(r_t^{d,i} - r_t^{f,i}\right) dt + \hat{\alpha}_N\Big(S_t^i, t, \left(X_t^j\right)_{j\leq N}\Big)\sqrt{V_t^i}\, dW_t^i \\ r_t^{d,i} = g_t^{d,i} + h^d(t) \\ r_t^{f,i} = g_t^{f,i} + h^f(t) \\ dg_t^{d,i} = \kappa_d\Big(\theta_d - g_t^{d,i}\Big)\, dt + \xi_d\sqrt{g_t^{d,i}}\, dW_t^{g^d,i} \\ dg_t^{f,i} = \left(\kappa_f\Big(\theta_f - g_t^{f,i}\Big) - \rho_{Sf}\xi_f\sqrt{g_t^{f,i}}\hat{\alpha}_N\Big(S_t^i, t, \left(X_t^j\right)_{j\leq N}\Big)\sqrt{V_t^i}\right) dt + \xi_f\sqrt{g_t^{f,i}}\, dW_t^{g^f,i} \\ dV_t^i = \kappa(\theta - V_t^i)\, dt + \xi\sqrt{V_t^i}\, dW_t^{V,i}\,, \end{cases}$$

where $\left(W_t^i, W_t^{g^d,i}, W_t^{g^f,i}, W_t^{V,i}\right)_{i\in[\![1,N]\!]}$ are N independent copies of the 4 correlated Brownian motions. One may notice that all the paths of the $4 \times N$ dimensional process $(X_t^i)_{i\leq N}$ are now entangled due to the dependence of $\hat{\alpha}_N$ in $(X_t^j)_{j\leq N}$. The process can be seen as a system of N interacting particles evolving in a 4 dimensional space, where particle i is defined by its position X_t^i. As pointed out in [82], we will use the term "particle" instead of "path" to account for this behaviour. The density function of the process $(X_t^i)_{i\leq N}$ is now driven by a $(4 \times N)$-dimensional linear Kolmogorov forward PDE which approximates asymptotically the non-linear 4-dimensional Kolmogorov forward PDE for X_t. Indeed, by the chaos propagation property (we refer to [139] for details) we get that, for any continuous and bounded function f,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} f\left(X_T^i\right) = \int_{(\mathbb{R}^+)^4} f(x)\phi(x)dx\,,$$

where ϕ is the joint density of (S_t, V_t, r_t^d, r_t^f), the solution to a non-linear 4D Kolmogorov forward PDE. Therefore, $\hat{q}_N$ and $\hat{p}_N$ converge weakly to the true expectations such that $\hat{\alpha}_N(K, T, (X_T^i)_{i\leq N})$ converges in probability to the solution $\alpha(K, T)$ as $N \to \infty$.

Remark. The existence and uniqueness of the solution for the SDE are not established theoretically, to the best of our knowledge. From an empirical perspective, [82, 83] show that for very high values of ξ ($\approx 350\%$, that practitioners may require to generate a particular forward skew), the particle method can only calibrate at-the-money strikes. In our case, the ξ ($\approx 30\%$) is calibrated to the market smiles and we are able to reach a high accuracy without any problems.

We will use the 2-factor LSV model with state vector $\left(S_t^{2D}, V_t^{2D}\right)$ with the associated leverage function α^{2D} calibrated in Subsection 3.4.7 as a conditional control variate for $\hat{p}_N$ defined in Subsection 3.4.3. Additionally, we will use the 1-factor local volatility model with the associated σ_{LV} function calibrated in Subsection 3.4.6 and keep track of the Dupire forward PDE at time T in line with the Monte Carlo time marching. Indeed, $C_{LV}(K, T)$ and its derivatives appear explicitly in α in (4.6.1) as well as in the control variates defined in Subsection 3.4.2.

We denote by N_T the number of time steps and by N the number of particles. The leverage function

α is defined on a grid of points and interpolated with cubic splines in spot and forward flat in time. There are $N_T + 1$ volatility "slices" in total such that we denote the m-th time slice, i.e., $\alpha(\cdot, T_m)$, by α_m, represented numerically as splines on a grid of N_S points. While having N_S too small will lead to accuracy problems, choosing it too large will make the surface rougher due to over-fitting. We find 25-30 points to provide a good trade-off between accuracy and smoothness. For a given time slice T_m, the leverage function is defined on the interval $[S^m_{min}, S^m_{max}]$ and is extrapolated flat outside these bounds. Because we need more grid points around the forward value and less around S^m_{min} and S^m_{max}, we use a hyperbolic grid (with $\eta = 0.05$, see Subsection 3.4.7 for more details) refined around the forward value

$$F_m = S_0 e^{\int_0^{T_m} \left(\bar{r}^d(t) - \bar{r}^f(t)\right) dt}$$

with

$$S^m_{min} = F_m e^{-\frac{6}{2}\sigma_F(T_m)\sqrt{T_m}}, \quad S^m_{max} = F_m e^{\frac{6}{2}\sigma_F(T_m)\sqrt{T_m}}.$$

Each of the grid values can be seen as a parameter and we denote them by $(\alpha_{m,j})_{m \le N_T,\, j \le N_S}$ with the associated spot grid values $(s_{m,j})_{m \le N_T,\, j \le N_S}$.

We can now write the calibration algorithm. We denote the 4-factor particle system at time T for the model (3.2.1) by $\left(S^i_T, V^i_T, r^{d,i}_T, r^{f,i}_T, D^i_T\right)_{i \le N}$, where D_T is an additional state variable for the domestic discount factor. Similarly, we denote the 2-factor particle system at time T for the model (3.2.2) by $\left(S^{2D,i}_T, V^{2D,i}_T\right)_{i \le N}$. We work with an exponential kernel

$$\delta_N(x, T) = \frac{e^{-\frac{1}{2}\left(\frac{x}{h_N(T)}\right)^2}}{h_N(T)\sqrt{2\pi}},$$

with a bandwidth that follows a Silverman-type rule (see [137])

$$h_N(T) = \eta S_0 \sigma_{LV}(S_0, T)\sqrt{\max(T, T_{min})} N^{-\frac{1}{5}},$$

where

$$\eta = 1.5, \quad T = 0.25.$$

The step-by-step calibration is detailed in Algorithm 4.2.

Algorithm 3.4 $\alpha(s,T)$ Calibration with control variate particle method

$\alpha(s,T_1=0)=\frac{\sigma_{LV}(s,0)}{\sqrt{v_0}}$
for $(\ m=1\,;\, m\le N_T\,;\, m++)$ **do**
generate $(Z,Z_v,Z_d,Z_f)_{i\le N}$ and $(U)_{i\le N}$, i.e. $4\times N$ independent draws from $\mathcal{N}(0,1)$ and N draws from $\mathcal{U}([0,1])$, respectively
evolve the 4-factor particle system from T_m to T_{m+1} with $QE-$Scheme (3.5.2) where $\alpha(s,[T_m,T_{m+1}[)=\alpha(s,T_m)$
evolve the 2-factor particle system from T_m to T_{m+1} with $QE-$Scheme (3.5.2) with pre-computed α^{2D} and using $(Z,Z_v,U)_{i\le N}$
solve the Dupire forward PDE (3.4.24) from T_m to T_{m+1} for C_{LV}, $\frac{\partial C_{LV}}{\partial K}$ and $\frac{\partial^2 C_{LV}}{\partial K^2}$
set $T=T_{m+1}$
for $(\ j=1\,;\, j\le N_S\,;\, j++)$ **do**
set $K=s_{m+1,j}$
compute as in (3.4.33)

$$p^{2D}=\mathbb{E}^{\mathbb{Q}^d}\Big[V_T^{2D}\,|\,S_T^{2D}=K\Big]=\left(\frac{\sigma_{LV}(K,T)}{\alpha^{2D}(K,T)}\right)^2$$

compute as in (3.4.3)

$$\hat{p}_N^{2D}(K,T)=\frac{\sum_{i=1}^N V_T^{2D,i}\delta_N\Big(S_T^{2D,i}-K\Big)}{\sum_{i=1}^N \delta_N\Big(S_T^{2D,i}-K\Big)}$$

compute $\hat{\lambda}_N$ as in (3.4.5) and

$$p_N^*(K,T)=\frac{\sum_{i=1}^N D_T^i V_T^i\delta_N(S_T^i-K)}{\sum_{i=1}^N D_T^i\delta_N(S_T^i-K)}+\hat{\lambda}_N\Big(\hat{p}_N^{2D}(K,T)-p^{2D}(K,T)\Big)$$

as in (3.4.4)
compute as in (3.4.2)

$$\gamma_N^* \;=\; \Big(\gamma_{1,N}^*-K\gamma_{2,N}^*\Big)$$

with

$$\gamma_{1,N}^* \;=\; \hat{\gamma}_{1,N}+\lambda_1(\hat{\mu}_{1,N}-C_{LV}(K,T))+\eta_1\Big(\hat{\zeta}_{1,N}-\mathbb{E}^{\mathbb{Q}^d}\Big[\Big(r_T^f-\bar{r}^f(T)\Big)\Big]\Big)$$

$$\gamma_{2,N}^* \;=\; \hat{\gamma}_{2,N}+\lambda_2\left(\hat{\mu}_{2,N}+\frac{\partial C_{LV}(K,T)}{\partial K}\right)+\eta_2\Big(\hat{\zeta}_{2,N}-\mathbb{E}^{\mathbb{Q}^d}\Big[\Big(r_T^d-\bar{r}^d(T)\Big)-\Big(r_T^f-\bar{r}^f(T)\Big)\Big]\Big)$$

compute

$$\alpha_{m+1,j}=\sqrt{\frac{1}{p_N^*(K,T)}\left(\sigma_{LV}(K,T)^2+\frac{\gamma_N^*(K,T)}{\frac{1}{2}K^2\frac{\partial^2 C_{LV}(K,T)}{\partial K^2}}\right)}$$

end for
end for

Our empirical findings suggest that Quasi-Monte Carlo sampling of the random numbers does not provide significant accuracy gains. Also, in order to speed up the computation of the sums involving

kernel functions such as

$$\sum_{i=1}^{N} V_T^{2D,i} \delta_N \left(S_T^{2D,i} - K \right),$$

it is advised (see [82]) to sort the particle state vector by spot value and select only the relevant particles that fall inside an interval $[K - \Delta K,\, K + \Delta K]$, where we choose

$$\Delta K = \sqrt{-2h_N^2(T) \ln\left(\epsilon \sqrt{2\pi} h_N(T)\right)},$$

with $\epsilon = 10^{-5}$.

3.5.3 Calibration results

In this subsection, we discuss the results of our calibration and the relative performance improvement with the use of our control variates for the data described in Subsection 3.4.1. First we compare the benefits of incorporating the control variates (CV) as opposed to the plain particle method (No-CV) as a function of the number of particles. Second, we compare the fit of the 4-factor hybrid LSV model (3.2.1) with the pure 4-factor hybrid SV model with no leverage function. The error measure we use is the absolute error in volatility (in % units). For instance, a maximum error of 0.03% for a 20% market volatility signifies that the calibrated volatility can be $20.00\% \pm 0.03\%$ in the worst case scenario.

We use the calibration routine described in Subsection 3.5.2 with and without control variates for 800, 4 000, 20 000, 100 000, 500 000 and 2 500 000 particles. Additionally, we set the number of time steps to 250 per year (such that, on average, there is one time step per open day). Both the average and maximum error are computed on all quoted deltas and maturities defined in Subsection 3.4.1. The results are presented in Figure 3.5.1.

We infer from the data in Figure 3.5.1 that the use of control variates greatly improves the general calibration routine. On the one hand, convergence[1] in both the maximum error (0.029%) and in the average error (0.011%) is reached for 20 000 particles when using control variates. On the other hand, the plain particle method without the variance reduction can only reach the same accuracy with 2 500 000 particles. In order to emphasise this fact, we calibrate the 4-factor model with 4 000 particles with and without control variates. We then plot, in Figure 3.5.2, the model implied volatility slices 3M, 1Y, 2Y, 5Y.

Hence, from Figure 3.5.1, we infer that the control variate particle method converges approximately 125 times faster than the plain particle method (without control variates). Most practitioners are familiar with the computational cost for the calibration of the 2-factor LSV model (3.2.2). One only needs to add the diffusion of 20 000 particles to get a very accurate calibration to vanilla options for the 4-factor LSV model (3.2.1). This allows us to calibrate a 4-factor model almost as fast as the

[1] i.e., the error from here on is dominated by other sources, such as the time discretisation error

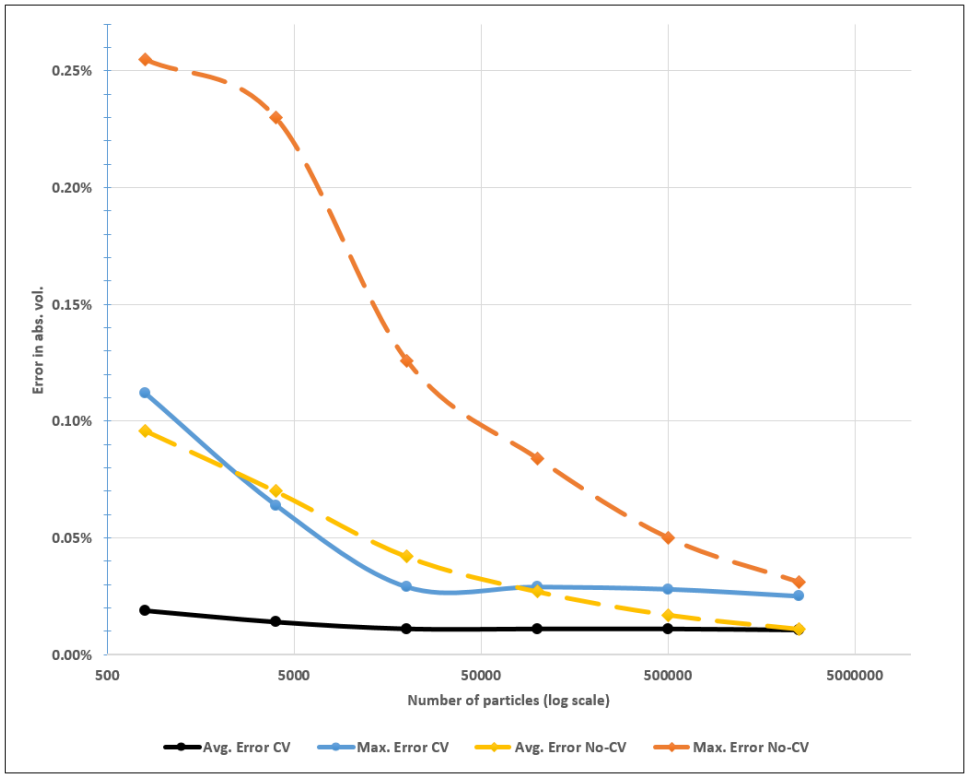


Figure 3.5.1: Error convergence with and without control variates. Error is computed on all quoted deltas and maturities (log-scale)

2-factor version. Moreover, from careful data analysis of Figure 3.5.1, we notice a clear convergence rate of 0.3 in the number of particles for the plain particle method, for both the average and maximum error. The addition of the control variates preserve the convergence rate but reduces the absolute size of the error significantly. We note that the error can be further reduced by increasing the number of time steps per year. However, we decided to work with 250 time steps (per year) as it already yields a very accurate calibration at a reasonable computational cost.

Part of the variance reduction is due to the conditional control variate described in Subsection 3.4.3 which can provide very good results for short-term horizons. In order to understand this, we recall the variance reduction factor in (3.4.6),

$$\frac{1}{1-\mathrm{Corr}(m,m^{2D})^2}.$$

We plot the quantity as a function of time with 500 000 particles in Figure 3.5.3. We notice a trend-line in $1+\frac{C}{T^2}$, for a given constant C, and thus a very good variance reduction for short-term options calibration. For longer term, this still yields a good variance reduction factor of 2.7 for the $3Y$ maturity and 1.44 for the $5Y$ pillar. In addition, the other two control variates, presented in Subsection 3.4.2,

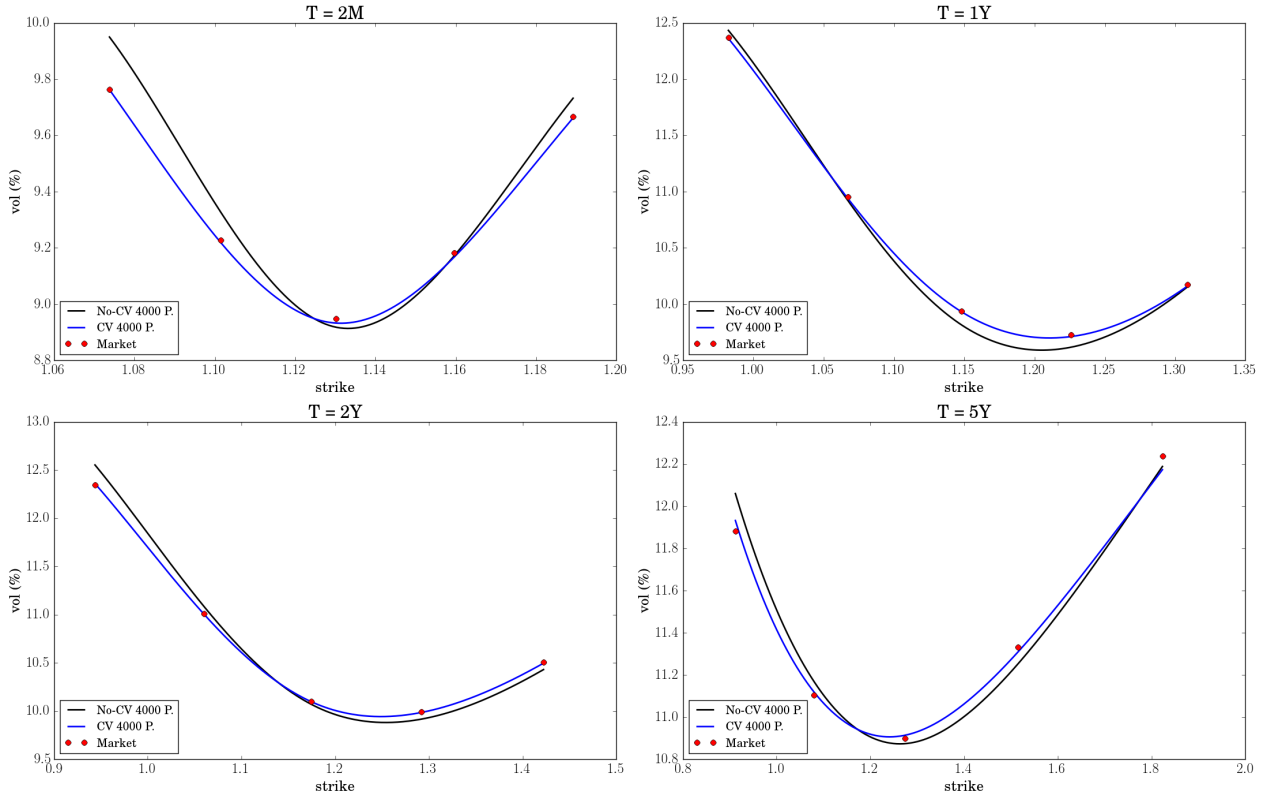


Figure 3.5.2: Calibration fit with 4 000 particles for the 4-factor LSV model with and without control variates

help control the short- to long-term behaviour as well. The variance reduction factor for $\gamma^*_{1,N}$ and $\gamma^*_{2,N}$ seems to reach a steady state in time around 1.6 and 6.5, respectively. We plot both variance reduction factors in Figure 3.5.3.

As a final observation on the calibration fit, we present and compare the accuracy of the 4-factor hybrid LSV model with the 4-factor Hybrid SV (Heston–2CIR++ model without leverage function that was calibrated in Subsection 3.4.5. We use 500 000 particles for that purpose and the results are presented in Figure 3.5.4 as well as in Table 3.5.1. We included a short-term pillar (3M), two mid-term pillars (1Y, 2Y) as well as one long-term pillar (5Y). The re-pricing is performed by Monte Carlo with 500 time steps per year and 5 000 000 paths. The average calibration error is 0.011%, which means that for a market volatility of 20%, the calibrated volatility is $20\% \pm 0.011\%$, on average.

The associated leverage function is plotted in Figure 3.5.5.

We notice that the short- to mid-term leverage function is smooth and well behaved. However, it becomes less stable around $5Y$ for out-of-the-money spot values. This is expected since for short- and mid-term pillars the variance reduction obtained with the 2-factor LSV model is significant. Longer

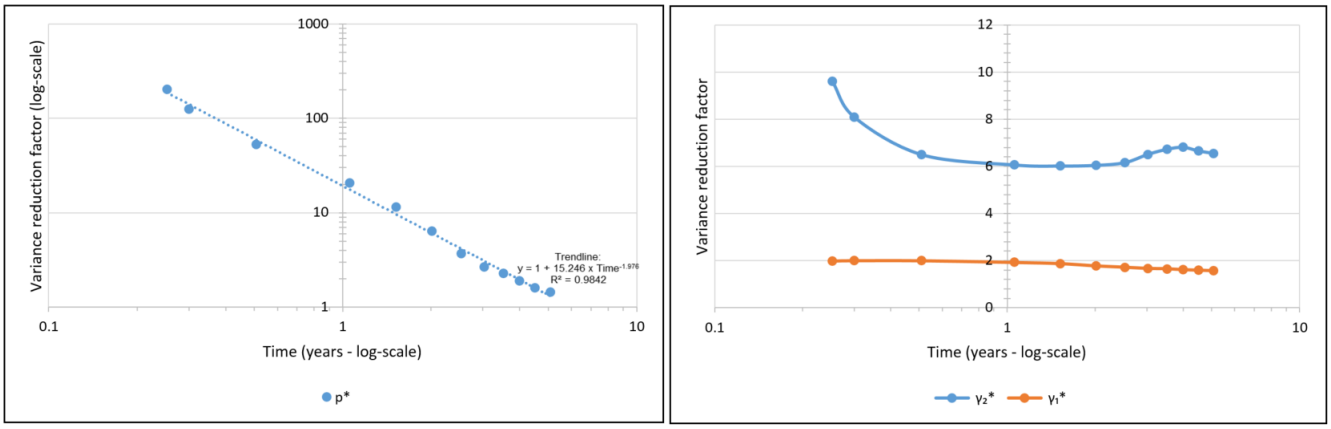


Figure 3.5.3: Variance reduction factor as a function of time for the conditional expectation estimator p_N^* (left, using a log-log scale) and the standard expectations estimators $\gamma_{1,N}^*$ and $\gamma_{2,N}^*$ (right, using a log scale)

term and out-of-the-money spot values are harder to estimate accurately because of the term

$$\frac{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\left(r_T^f - \bar{r}^f(T)\right)(S_T - K)^+\right] - K\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \mathbf{1}_{S_T \geq K}\left[\left(r_T^d - \bar{r}^d(T)\right) - \left(r_T^f - \bar{r}^f(T)\right)\right]\right]}{\frac{1}{2}K^2 \frac{\partial^2 C_{LV}}{\partial K^2}}$$

that appears in $\alpha(K,T)$. Indeed, the computation of the two expectations is less accurate out-of-money (since the relative variances of the two payoffs are bigger) and a small error is emphasised by the very small denominator. Thankfully, the use of control variates from 3.4.2 combined with a good estimation of the denominator (as explained in Subsection 3.4.6) allows to dampen the oscillations. Furthermore, the calibration fit is not very sensitive to far out-of-(or deep in-)the-money values of the leverage function.

3.5.4 Stress scenario

The last calibration result we present is for an extremely stressed set of rates parameters. This will serve as a robustness test of the control-variate particle method for very volatile short rate processes. In order to perform the test, we multiply the calibrated ξ_d and ξ_f by 20 as well as divide θ_d and θ_f by 20. This leads to strongly violated Feller conditions and very high volatilities for the two CIR++ rate processes. The stress scenario parameter values are

$$\begin{cases} v_0 = 0.0094, \quad \theta = 0.0137, \quad \kappa = 1.4124, \quad \rho = -0.1194, \quad \xi = 0.2988\,, \\ g_0^d = 0.0001, \quad \theta_d = \mathbf{0.0273}, \quad \kappa_d = 0.0837, \quad \rho_{Sd} = 0, \quad \xi_d = \mathbf{0.5480}\,, \\ g_0^f = 0.0001, \quad \theta_f = \mathbf{0.0582}, \quad \kappa_f = 0.0110, \quad \rho_{Sf} = 0, \quad \xi_f = \mathbf{0.7400}\,. \end{cases}$$

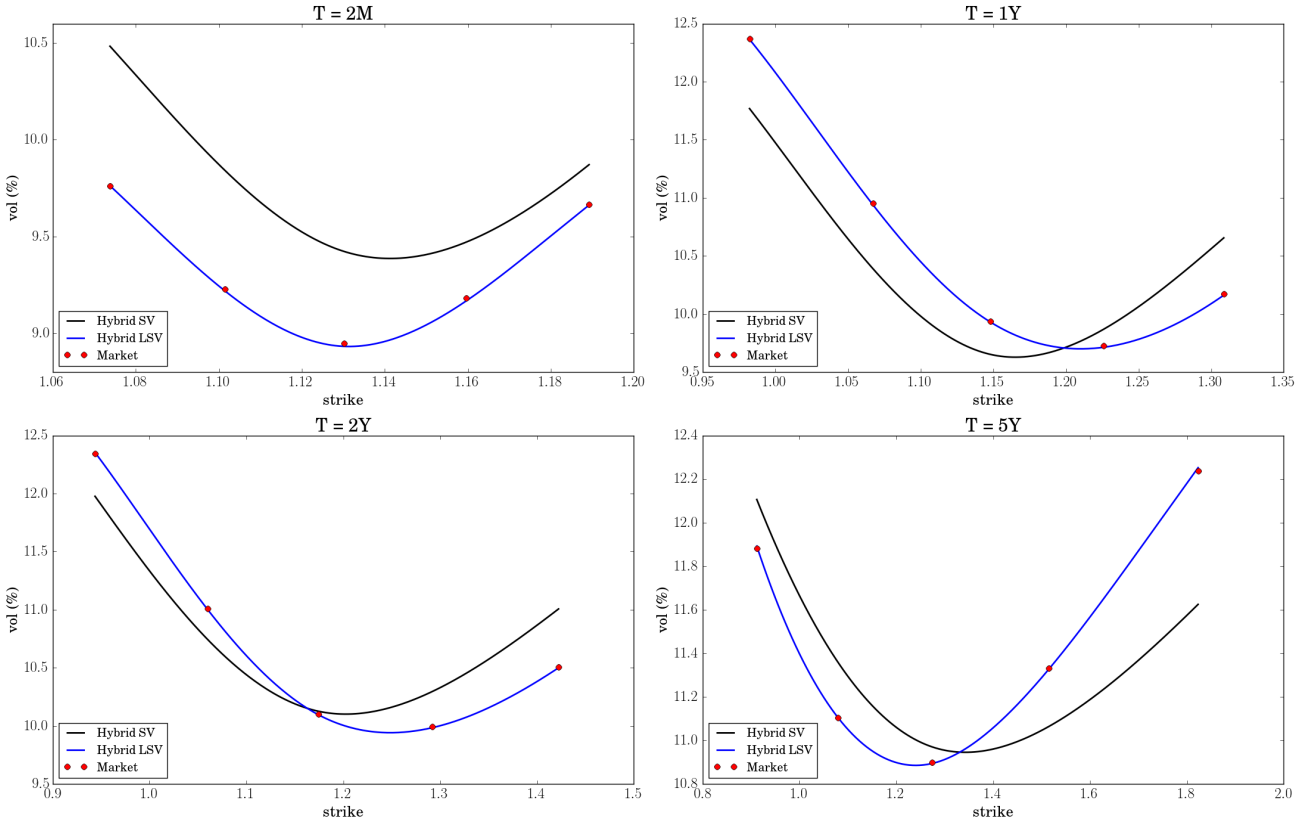


Figure 3.5.4: Calibration fit with 500 000 particles for the 4-factor LSV model vs 4-factor SV model

We emphasise that in practice, the volatility parameters $\xi_{d/f}$ are rarely above 0.06. One can refer to [21] for more details. A calibration summary for the average error in absolute volatility is displayed in Table 3.5.2. On this stress scenario, the calibration via control-variate particle method reaches an error of 0.0165% for 20 000 particles, whereas 2 500 000 particles are required with the plain particle method (without control variates) to reach the same accuracy. This is in line with Figure 3.5.1 for our previous observations with realistic short rate parameters calibrated to the market. Hence, the control-variate particle method shows a consistent improvement over the plain particle method even under stress scenarios. The conditional control variate alone yields a good variance reduction factor of almost 2 for the last maturity (5Y) .

3.6 Pricing of an AutoCallable Barrier Dual-Currency Note

In this section, we will discuss the pricing of more exotic products, more specifically, structured notes embedding barrier features. Some popular FX exotics include the *Power Reverse Dual Currency Note* (PRDC) [35], especially long-term ($\sim$ 30Y), and the Forward Accumulator [153]. Both of these contracts would benefit from the addition of stochastic rates for mid- to long-term expiries. While the

	10D Put		25D Put		50D		25D Call		10D Call		Avg.	
T	**LSV**	**SV**	**LSV**	**SV**	**LSV**	**SV**	**LSV**	**SV**	**LSV**	**SV**	**LSV**	**SV**
3W	0.01	1.14	0.01	1.19	0.02	1.08	0.02	0.69	0.02	0.28	0.02	0.88
1M	0.02	1.33	0.01	1.29	0.01	1.14	0.00	0.78	0.00	0.47	0.01	1.00
2M	0.00	0.72	0.01	0.61	0.01	0.48	0.01	0.29	0.00	0.20	0.01	0.46
3M	0.01	0.12	0.02	0.03	0.02	0.09	0.02	0.08	0.01	0.03	0.02	0.07
6M	0.03	0.68	0.03	0.63	0.02	0.36	0.02	0.05	0.01	0.40	0.02	0.42
1Y	0.01	0.60	0.01	0.56	0.01	0.29	0.01	0.14	0.01	0.48	0.01	0.41
1Y6M	0.01	0.48	0.00	0.40	0.01	0.13	0.01	0.22	0.01	0.49	0.01	0.34
2Y	0.01	0.37	0.00	0.26	0.01	0.03	0.01	0.30	0.01	0.50	0.01	0.29
3Y	0.02	0.20	0.00	0.18	0.00	0.10	0.00	0.01	0.01	0.07	0.01	0.11
5Y	0.01	0.22	0.00	0.25	0.01	0.07	0.00	0.27	0.02	0.61	0.01	0.29

Table 3.5.1: Calibration error in absolute volatility (% unit) for the 4-factor LSV model vs 4-factor Heston–2CIR++ model

N	4 000	20 000	100 000	500 000	2 500 000
Plain particle method	0.0855%	0.0440%	0.0320%	0.0255%	0.0160%
Control variate particle method	0.0312%	0.0165%	0.0133%	0.0133%	0.0133%

Table 3.5.2: Average error in absolute volatility (% unit) for a high-volatility stress scenario on the rate processes. N is the number of particles.

plain PRDC can be seen as a strip of FX vanilla options, its variations such as the Trigger PRDC (with knock-out feature) will be highly impacted by adding a local-stochastic volatility on top of stochastic rates. Here, we consider an *AutoCallable Barrier Dual-Currency Note* (ABDC). While *Barrier Dual-Currency* are usually short term investments, we analyse a mid-term alternative with an embedded AutoCallable Note structure. The contract will be indexed on the EURUSD currency pair, where we use USD as domestic currency and EUR as foreign currency. This product would likely suit an investor who wants to outperform the money market account by accepting a risk of conversion of the nominal to the foreign currency. The structured note is defined on a quarterly period schedule where a coupon is paid. The AutoCall feature reduces the price by early terminating the product if the exchange rate rises above an up barrier. The down-and-in barrier feature activates the conversion at expiry, which gives the owner extra protection against the conversion risk at the expense of increasing the *Net-Present-Value* (NPV). The investor would benefit from a current volatile market if he expects a drop in volatility in the future. Indeed, this would mean a lower NPV due to the increased chance of early redemption and conversion. On a longer time horizon, a very stable market is the best setting for this product; the investor would get all the coupons without converting the nominal at maturity. The contract life cycle, for a nominal N in USD and expiry date T, can be described as follows :

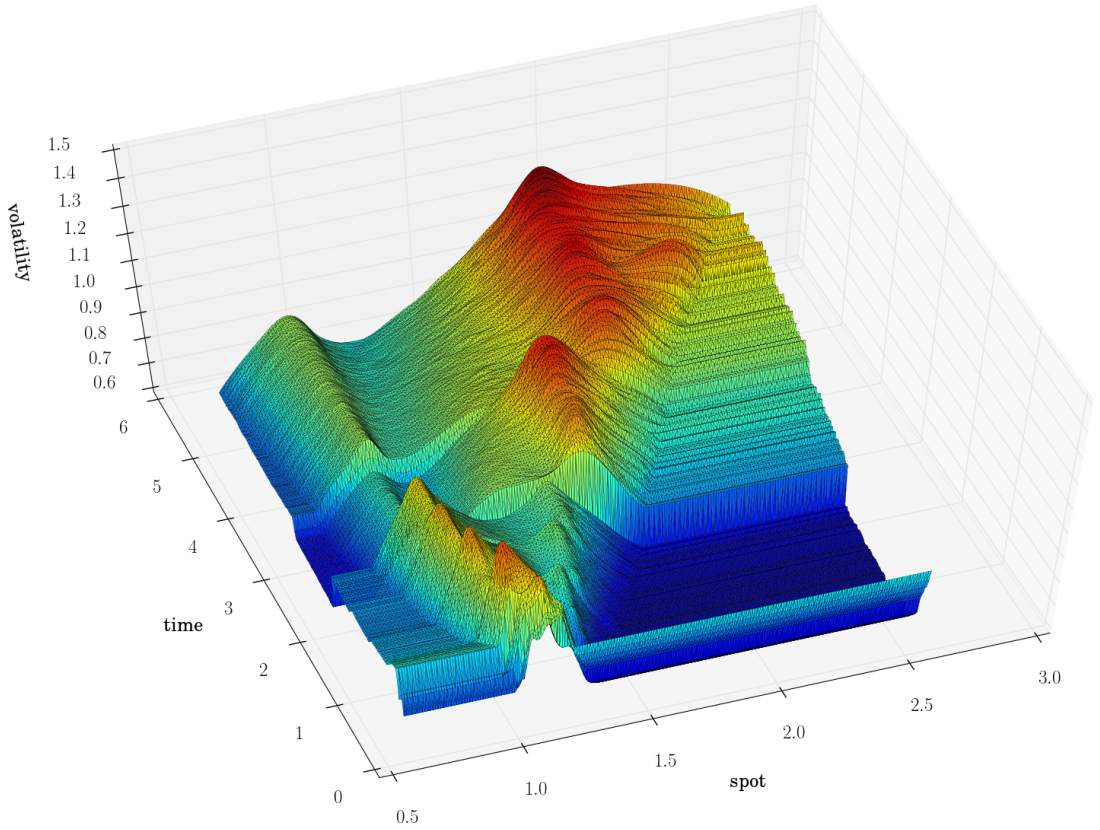


Figure 3.5.5: Calibrated leverage function α for the 4-factor LSV model (3.2.1)

- A monthly fixing schedule is defined for the exchange rate S, such that
  - if S breaches an up-and-out barrier B_{UO}, the contract is early redeemed. The early redemption time is denoted by τ_{ER} and can eventually be infinite in case of no knock-out event. An early redemption coupon cpn_{ER} is paid.
  - if S goes below a down barrier B_{DI}, a short put contract is activated at expiry and we denote the knock-in time by τ_{KI}.
- if S is above the down barrier B_{DI}, a coupon cpn (in % of the nominal) is paid quarterly in USD for a total of M periods, occurring at coupon dates $(t_m)_{1\leq m\leq M}$, such that the total accumulated coupon in USD at expiry T is

$$N\sum_{i=1}^{M}\left(\frac{D^d_{t_i}}{D^d_T}\times cpn\times \mathbf{1}_{t_i\leq\tau_{\mathrm{ER}}}\times \mathbf{1}_{B_{\mathrm{DI}}<S_{t_i}}\right)+N\frac{D^d_{\tau_{\mathrm{ER}}}}{D^d_T}\times cpn_{\mathrm{ER}}\times \mathbf{1}_{\tau_{\mathrm{ER}}\leq T}$$

  (payment occurs at coupon dates).
- At expiry, if no early redemption has occurred, if the down barrier B_{DI} has been activated and

if S is below a pre-defined strike K, then the notional N is converted to EUR at strike price K.

One can see that if the conversion occurs at expiry, the owner receives $\frac{N}{K}$ in EUR in exchange for the nominal N in USD. If the owner converts the amount back to USD, the loss is $N\Big(\frac{S_T}{K}-1\Big)$. Hence, the *profit-and-loss* (PnL) of the product is

$$-\frac{N}{K}(K-S_T)^+\mathbf{1}_{\tau_{\mathrm{KI}}\leq T<\tau_{\mathrm{ER}}}+N\sum_{i=1}^{M}\left(\frac{D^d_{t_i}}{D^d_T}\times cpn\times\mathbf{1}_{t_i\leq\tau_{\mathrm{ER}}}\times\mathbf{1}_{B_{\mathrm{DI}}<S_{t_i}}\right)+N\frac{D^d_{\tau_{\mathrm{ER}}}}{D^d_T}\times cpn_{\mathrm{ER}}\times\mathbf{1}_{\tau_{\mathrm{ER}}\leq T}\,.$$

The accrued coupon is sometimes converted at the exchange rate K (with the notional) at expiry as well. The NPV of the contract is then

$$\begin{aligned}\mathrm{NPV} &= \mathbb{E}^{\mathbb{Q}^d}\left[-\frac{D^d_T}{K}(K-S_T)^+\mathbf{1}_{\tau_{\mathrm{KI}}\leq T<\tau_{\mathrm{ER}}}+\sum_{i=1}^{M}\Big(D^d_{t_i}\times cpn\times\mathbf{1}_{t_i\leq\tau_{\mathrm{ER}}}\times\mathbf{1}_{B_{\mathrm{DI}}<S_{t_i}}\Big)\right]\\ &+ \mathbb{E}^{\mathbb{Q}^d}\Big[D^d_{\tau_{\mathrm{ER}}}\times cpn_{\mathrm{ER}}\times\mathbf{1}_{\tau_{\mathrm{ER}}\leq T}\Big]\,,\end{aligned}$$

in % of the nominal. This product is a yield enhancement contract with no capital guarantee.

Remark. All barriers are monitored discretely in this example. Hence, the Monte Carlo pricing does not require any Brownian bridges techniques as for continuously monitored barriers. Furthermore, the product can be decomposed in linear combinations of first-order exotics for which we proved convergence for the FTE Monte Carlo scheme in [40]. Hence, the convergence of the FTE Monte Carlo scheme follows automatically for this structured product.

The contract parameters are displayed in Table 3.6.1.

N	S_0	T	K	cpn	cpn_{ER}	B_{UO}	B_{DI}
\$ 100 000	1.1271	5Y	105% (of S_0)	2.5%	1.5%	100% (of S_0)	95% (of S_0)

Table 3.6.1: Contract parameters for an AutoCallable Barrier Dual-Currency Note

Barriers are monitored monthly (at fixing dates) and coupons are paid quarterly. The pricing results with 50 000 000 paths and 365 time steps per year, are displayed in Table 3.6.2. We computed the relative difference in premium with respect to the reference model, in this case the LSV–2CIR++ model.

We infer from the data in Table 3.6.2 a 1.13% difference in price between the LSV and LSV–2CIR++ models, which confirms that adding stochastic rates has a significant impact even for a 5Y contract. Both LSV models, with and without stochastic rates, provide the same early redemption and knock-in probabilities. Under the LV model, the product is highly under-priced and early redemption probability is overestimated.

Model	NPV	MC 95% Conf. Int.	ER. Prob.	KI Prob.	Rel. Diff.
LV	1.2079	$(1.2062, 1.2097)$	95.9%	19.5%	−30.05%
LSV	1.7465	$(1.7447, 1.7483)$	95.1%	19.4%	1.13%
LSV–2CIR++	1.7269	$(1.7251, 1.7287)$	95.1%	19.4%	–

Table 3.6.2: Monte Carlo NPVs, in % of N, for the 5Y AutoCallable Barrier Dual-Currency Note

Therefore, on the one hand, we notice that using local-stochastic volatility is of paramount importance when pricing exotic products with barrier features. It is a known fact that pure stochastic volatility model will under-estimate the knock-out probability while local volatility will overestimate it. The "true" price is believed to lie in between pure SV and pure LV models prices, i.e., LSV models (see [35] for details). On the other hand, adding stochastic rates to an LSV model is relevant only for mid- to long-term structured notes where the effect of stochastic rates is accentuated. Furthermore, the impact of stochastic rates is emphasized if one decides to add a correlation between rates and FX spot process. Finally, one could define a more complex model for the rates processes by adding a second factor. This approach is relevant if one needs to price an exotic contract where the rates appear explicitly. An example, is a spread option with a payoff of the form

$$\left(\left(\frac{S_T - S_0}{S_0}\right) - L_T - K\right)^+,$$

with L_T the Libor rate at fixing date T. In that particular case, the calibration algorithm is still applicable at the expense of a higher computational time due to the more complex Monte Carlo scheme.

We conclude this section with an empirical convergence analysis of our Monte Carlo simulation scheme. Since the product can be decomposed into a linear combination of first-order exotics, the theoretical convergence (without a rate) of Monte Carlo estimators follows automatically from the analysis in [40]. The data in Table 3.6.3 suggest a first-order convergence of the time-discretization error. As an aside, note that in the case of continuously monitored barriers, we could use Brownian bridge techniques to recover the first-order convergence.

The empirical findings of this section demonstrate the importance of local-stochastic volatility dynamics as well as stochastic short rate dynamics for the pricing of long-dated exotic FX products. Furthermore, we verified the convergence of our Monte Carlo simulation scheme for one such product.

3.7 Conclusion

In this chapter, we have provided a new and numerically effective method to calibrate a 4-factor LSV model to vanilla options. According to our numerical results and based on real world market data, we

Time steps	NPV	Difference	Rate
12	1.9081	–	–
24	1.8110	0.0971	–
48	1.7623	0.0487	0.996
96	1.7388	0.0234	1.057
192	1.7285	0.0103	1.184

Table 3.6.3: The Monte Carlo estimates of the NPV of the contract specified in Table 3.6.1 for different numbers of time steps (per year) and 5×10^7 sample paths (for a standard deviation of 9.29×10^{-4}), the difference from the previous NPV estimate and the empirical order of convergence.

managed to achieve an approximate 125-fold speed-up for the calibration using control variates, as compared to the plain particle method. We have shown that a high accuracy can be obtained with as few as 20 000 particles (with a maximum error in absolute volatility of 0.03%), and we were able to get a good fit with only 4 000 particles (with a maximum error in absolute volatility of 0.06%). Using the calibrated leverage function from this chapter, we showed that the addition of stochastic rates may have a significant impact on structured products, even more so when barrier features and coupon detachments are combined for longer-dated contracts. Stochastic rates become necessary in the modeling if one wishes to price hybrid products where the rates appear explicitly (for instance, a spread option on the FX performance and the Libor rate). One could use a second factor in the CIR++ processes to improve the fit to caps and use the method presented in this chapter to calibrate the leverage function.

4 Calibration of Local-Stochastic and Path-Dependent Volatility Models to Vanilla and No-Touch Options with Markovian Projections

4.1 Introduction

In addition to call options, practitioners are increasingly interested in including the quotes of touch options in the calibration basket, which will improve the pricing of exotic contracts with barrier features. A few published works already address this question for different model classes. Crosby and Carr [31] consider a particular class of jump models which gives a calibration to both vanillas and barriers, while Pironneau [121] proves that the Dupire equation is still valid for a given barrier level, under the local volatility model.

This chapter addresses the simultaneous calibration to vanilla and barrier (specifically, no-touch) options systematically. The focus is less the calibration of a particular model – although we do calibrate three different models – but on a methodology which allows the efficient calibration of any volatility model. We take the Brunick–Shreve mimicking point of view [25] that the joint distribution of a stock price and its running maximum, or equivalently, barrier prices for all strikes, barrier levels and maturities, can be reproduced by a deterministic volatility function of the spot, the running maximum and time. This is a natural extension of Gyöngy's result [84] that the stock price distribution, or equivalently, call prices for all strikes and maturities, can be reproduced by a deterministic volatility function of the spot and time. In the latter case, this volatility function is the conditional expectation of the variance process given the future spot price, while in path-depedent case the expectation is also conditional on the path-dependent quantity, here the running maximum. These conditional expectations are often referred to as Markovian projections.

In the vanilla case, the calibration is guaranteed if the Markovian projection onto the spot coincides with the local volatility function derived from vanilla quotes by Dupire's formula. Conversely, given the local volatility, model prices can be computed by the forward Dupire PDE, formulated in strike and maturity. The estimation of the local volatility from observed prices is an ill-posed inverse problem, and regularisation approaches have been proposed, e.g., in Crépey [42], Egger and Engl [56] and Murex Analytics [111]. If the underlying model to be calibrated is not itself a local volatility

model, the Dupire PDE can still be used to compute the model prices by utilising the mimicking result, i.e., with the diffusion coefficient in the PDE replaced by the Markovian projection of the squared stochastic volatility onto the spot. The conditional expectation of the stochastic variance under the desired model can be estimated, e.g., by the particle method [83, 82].

The natural extension of the Dupire PDE for calls to a forward equation for barrier option prices, with strike, barrier level and maturity as independent variables, is the forward PIDE derived in Chapter 2. In this chapter, we give a rigorous proof of the PIDE under weaker regularity conditions. Although the approach is designed to handle all types of up-and-out call options, we reduce the calibration scope to vanilla and no-touch options, in line with what the foreign-exchange (FX) market usually quotes. In order to handle down-and-out call/put type contract, one would need to re-derive the forward PIDE in Chapter 2 and apply a similar calibration process as described in this work. The calibration of the local volatility function of spot and running maximum using this forward PIDE is inspired by, and extends, the literature on local volatility model calibration and regularisation cited above.

We also consider an extension of the model in the spirit of LSV models, i.e., a local maximum-dependent volatility function multiplied by a stochastic volatility. These two models fall into the class of path-dependent volatility models and can be very useful to replicate a market's spot-volatility dynamics as explained in [81].

It is important to note, though, that the diffusion of interest in this work can be a general continuous stochastic process. More specifically, the variance process does not need to contain the running maximum in its parametrisation. In the case of models other than the Brunick-Shreve model, we view the calibrated volatility function (of spot, running maximum, and time) as a 'code book' for barrier option prices (a name coined in [29] for local volatility) which implicitly defines the model parameters. We investigate as example a Heston-type LSV model with a spot dependent vol-of-vol (LSV-LVV) that was first introduced in McNamara [109]. The approach can easily be generalised to other types of stochastic volatility diffusions. This vol-of-vol local volatility function is also time dependent such that we can perform a best-fit to no-touch options at each quoted maturity while ensuring perfect calibration to vanilla options through the local leverage function. The main drawback of working under a LSV model is the potentially less accurate calibration to no-touch options. The calibrated LSV-LVV model prices no-touch options well within the market bid–ask spread for all barrier levels and maturities.

The calibration of the LSV-LVV and LMSV models relies on the computation of the Markovian projection of the stochastic variance V_t by either extending the particle method introduced in Guyon and Labordère [83, 82] or by solving the (S_t, M_t, V_t) joint density PDE numerically. In our approach, we rely on a two-dimensional particle method as it offers a straightforward extension to additional stochastic factors.

On the numerical side, we propose a finite element approach to the solution of the Fokker–Planck

PDE for the joint density of (S, M, V). It requires an optimal mesh construction, an efficient handling of the boundary equation when $S = M$ and a difficult point-wise Dirichlet boundary condition at (S_0, S_0, V_0) to incorporate. Validation results are presented for the calibrated models, where we price up-and-out call options by integrating the computed joint density against the path-dependent payoff function.

Finally, we extend the Markovian projection formula to account for stochastic rates, discuss how to adapt the calibration algorithms and suggest that a control-variate particle method, as presented Chapter 3 would be a good candidate to calibrate such a model.

The remainder of this chapter is organised as follows. In Section 4.2, we define the models as well as discussing the market data included in the calibration basket. In Section 4.3, we provide a necessary and sufficient condition for a perfect calibration to up-and-out barrier option quotes and prove this rigorously using local times. Section 4.4 presents the efficient numerical solution of the forward PIDE for barrier options which is central for the algorithms in the remainder of the chapter. In Section 4.5, we present our approach to the calibration problem of the Brunick–Shreve model by forward PIDE and regularised gradient-based optimisation. Then, in Section 4.6, we extend the particle method to a two-dimensional conditional state space and calibrate the LMSV model by Markovian projection. In the same spirit, Section 4.7 makes use of the Markovian projection for the LSV-LVV model and combines it with the forward PIDE for barrier options in order to best-fit no-touch quotes while perfectly calibrating vanilla market prices. The calibration results for all these models are presented and compared in Section 4.8. Section 4.9 provides results for the extension to stochastic rates, Section 4.10 discusses a finite element approach to the solution of the Fokker–Planck PDE for the joint density of (S, M, V) and finally, Section 4.11 concludes with a brief discussion.

4.2 Models and market

4.2.1 Model definition

We consider a domestic and a foreign market with risk-free short rates r^d and r^f, respectively. The spot S_t associated with the currency pair ccy1ccy2 (following the notations in [35]) denotes the amount of units of ccy2 (domestic currency) needed to buy one unit of ccy1 (foreign currency) at time t. We assume the existence of a filtered probability space $(\chi, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q}^d)$ with a domestic risk-neutral measure $\mathbb{Q}^d$, under which the price process S follows the SDE

$$\frac{dS_t}{S_t} = \left(r_t^d - r_t^f\right) dt + \gamma_t \, dW_t \,, \tag{4.2.1}$$

where W is a one-dimensional $\mathcal{F}_t$-adapted Brownian motion and γ_t is a continuous and positive $\mathcal{F}_t$-adapted semi-martingale such that

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t \gamma_u^2 S_u^2 du\right] < \infty\,. \tag{4.2.2}$$

We assume deterministic rates for the time being, such that

$$\begin{aligned} r_t^d &= \bar{r}^d(t)\,, \\ r_t^f &= \bar{r}^f(t)\,, \end{aligned}$$

where $\bar{r}^d(t) = -\frac{\partial \ln P^d(0,t)}{\partial t}$ and $\bar{r}^f(t) = -\frac{\partial \ln P^f(0,t)}{\partial t}$, with $P^{d/f}(0,T)$ being the market zero-coupon bond prices for the domestic and the foreign money market accounts, respectively. We will discuss an extension to stochastic rates in Section 4.9. The domestic and foreign discount factors are defined as

$$D_t^d = \mathrm{e}^{-\int_0^t r_u^d du}, \quad D_t^f = \mathrm{e}^{-\int_0^t r_u^f du}\,. \tag{4.2.3}$$

A model widely used in the industry is the Heston-type LSV model

$$\begin{cases} \dfrac{dS_t}{S_t} = \left(\bar{r}^d(t) - \bar{r}^f(t)\right) dt + \alpha_S(S_t,t)\sqrt{V_t}\, dW_t \\ dV_t = \kappa(\theta - V_t)\, dt + \eta\xi\sqrt{V_t}\, dW_t^V\,, \end{cases} \tag{4.2.4}$$

where W and W^V are standard Brownian motions with constant correlation ρ, moreover $v_0, \kappa, \theta, \xi, \eta$ are scalar parameters, and $\alpha_S : \mathbb{R}^+ \times [0,T] \to \mathbb{R}^+$ the "leverage" function assumed to be bounded and locally Lipschitz in S. Here, the parameter η is redundant (as only the product $\eta\xi$ appears) and can be set to 1 for the time being. It will be used in Section 4.7 to interpolate between pure local volatility ($\eta = 0$, $\kappa = 0$,$v_0 = 1$) and the Heston model ($\eta = 1$, $\alpha_S \equiv 1$).

The generic model definition (4.2.1) gives rise to three further models which will be the focus of this chapter. Let

$$\Omega = \left\{(s,m) \in \mathbb{R}^2 : 0 \le s \vee S_0 \le m\right\}.$$

First, define a Brunick–Shreve (BSV) model with a path-dependent volatility (and no additional stochastic factor) by

$$\begin{cases} \dfrac{dS_t}{S_t} = \left(\bar{r}^d(t) - \bar{r}^f(t)\right) dt + \sigma_{BSV}(S_t, M_t, t)\, dW_t \\ M_t = \max\limits_{0\le u\le t} S_u\,, \end{cases} \tag{4.2.5}$$

where $\sigma_{BSV} : \Omega \times [0,T] \to \mathbb{R}^+$ is assumed to be bounded, locally Lipschitz in S and continuously

differentiable in M. This implies that (S, M) is Markovian (see [24]).

Secondly, we propose the local-maximum-stochastic volatility (LMSV) model defined as

$$\begin{cases} \dfrac{dS_t}{S_t} = \left(\bar{r}^d(t) - \bar{r}^f(t)\right) dt + \alpha(S_t, M_t, t)\sqrt{V_t}\, dW_t \\ dV_t = \kappa(\theta - V_t)\, dt + \xi\sqrt{V_t}\, dW_t^V \\ M_t = \max\limits_{0 \le u \le t} S_u\,, \end{cases} \tag{4.2.6}$$

with $\alpha : \Omega \times [0, T] \to \mathbb{R}^+$ a leverage function that also depends on the running maximum. Finally, we define a Heston-type local-stochastic volatility model with local vol-of-vol (LSV-LVV),

$$\begin{cases} \dfrac{dS_t}{S_t} = \left(\bar{r}^d(t) - \bar{r}^f(t)\right) dt + \alpha_S(S_t, t)\sqrt{V_t}\, dW_t \\ dV_t = \kappa(\theta - V_t)\, dt + \xi(S_t, t)\sqrt{V_t}\, dW_t^V, \end{cases} \tag{4.2.7}$$

with $\alpha_S : \mathbb{R}^+ \times [0, T] \to \mathbb{R}^+$, $\xi : \mathbb{R}^+ \times [0, T] \to \mathbb{R}^+$.

4.2.2 Market data

We use a set of market data from 28/03/2013 for the EURUSD currency pair.[1] The EUR zero-coupon bond prices are plotted in Figure 4.2.2 and the USD zero-coupon bond (ZCB) prices are plotted in Figure 4.2.1.

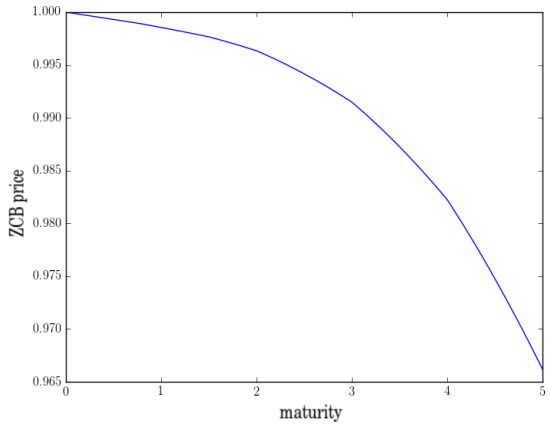


Figure 4.2.1: USD Zero-coupon bond prices.

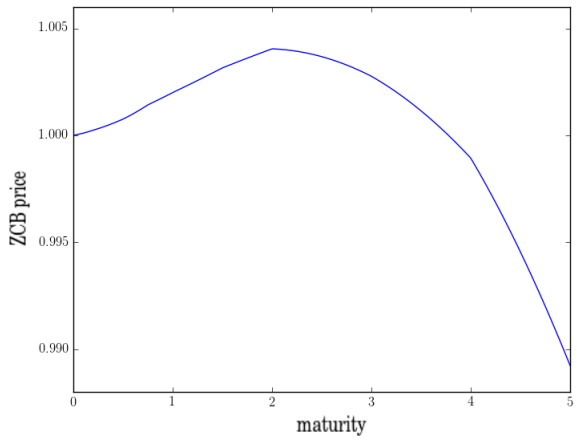


Figure 4.2.2: EUR Zero-coupon bond prices.

[1] The market data for both call option prices and no-touch prices were provided by Markit. The zero-coupon rates for both EUR and USD curves were retrieved from Bloomberg.

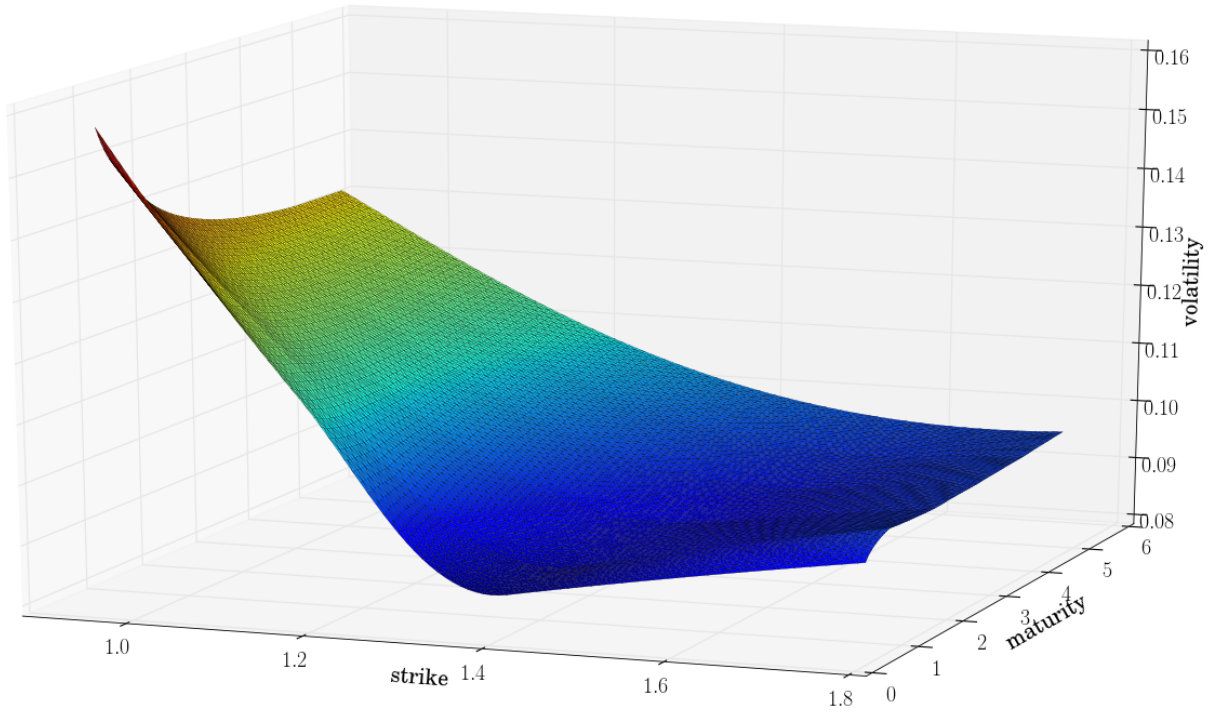


Figure 4.2.3: Market volatility surface for EUR-USD (28/03/2013).

The EURUSD spot value, which is denominated as the price of one EUR in USD units, is

$$S_0 = 1.2837\,(\$\,\text{per}\,€)\,.$$

We use in the calibration at-the-money (spot or forward) volatility, 10 and 25 delta smile-strangles and risk-reversals, i.e. for each maturity the following 5 volatilities on a delta scale,

$$\text{10D-Put}, \text{25D-Put}, \text{50D}, \text{25D-Call}, \text{10D-Call},$$

and the following maturities

$$\text{3M, 6M, 1Y, 2Y, 3Y, 4Y, 5Y.}$$

The volatility surface is plotted in Figure 4.2.3 on a strike scale and is interpolated and extrapolated linearly in variance on a log-moneyness scale.

Additionally, we calibrate to quotes for one-touch options, which pay 1 at maturity exactly if the FX rate breaches the up-barrier B during the product lifespan (continuous monitoring). We convert the market quotes for foreign one-touches in the foreign currency (EUR), $FOT(B,T)$, to foreign no-touch

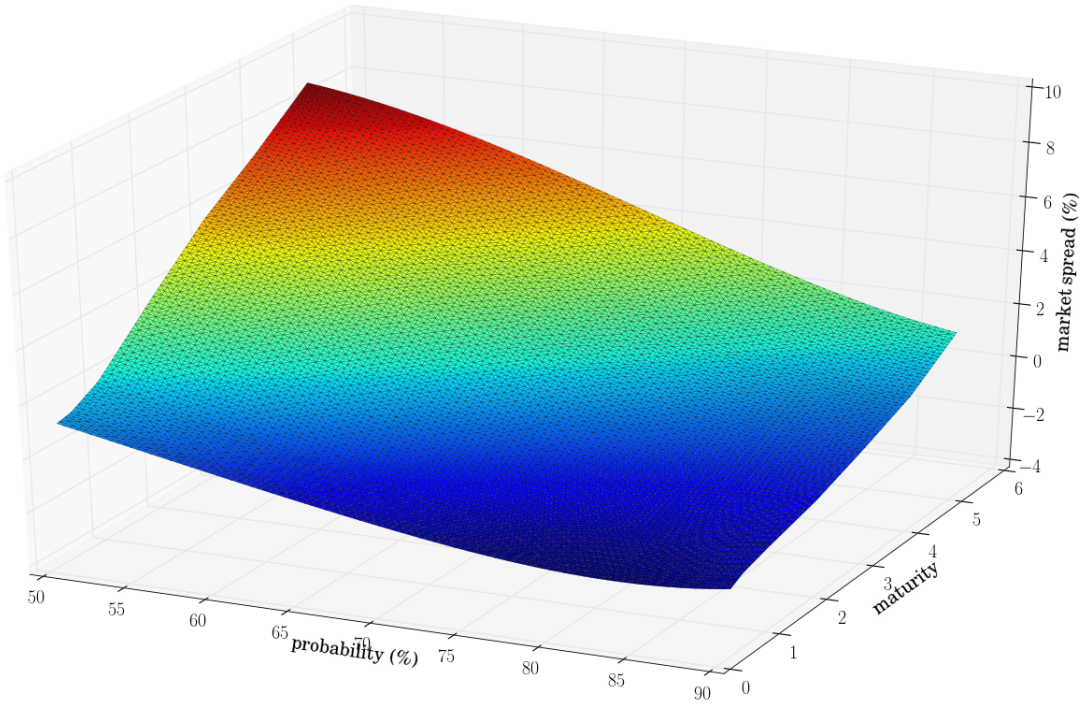


Figure 4.2.4: Foreign no-touch up market prices

options with the following formula (under deterministic rates),

$$FNT(B,T) \quad = \quad D_T^f - FOT(B,T). \tag{4.2.8}$$

By a change of numéraire, the foreign no-touch price under an arbitrage-free model with domestic risk-neutral measure $\mathbb{Q}^d$, and expressed in foreign units (EUR) is

$$FNT(B,T) = \frac{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d S_T \mathbf{1}_{M_T<B}\right]}{S_0}\,.$$

The net present value of any financial contract will be expressed in units of domestic numéraire unless stated otherwise. In the remainder of the chapter we use the term "no-touch" to denote a foreign no-touch contract as defined in (4.2.8). For calibration purposes, we consider the market prices of foreign no-touches for the following maturities

$$\text{3M, 6M, 1Y, 2Y, 3Y, 4Y, 5Y}\,.$$

Each time slice is made of 5 up-and-out barrier levels B, corresponding to no-touch-up probability labels

$$x \approx 50\%,\ 60\%,\ 70\%,\ 80\%,\ 90\%.$$

Those probability labels are an approximate estimation of the no-touch-up probabilities quoted by the market, and do not carry significant meaning (in a similar way, the maturity pillar labelled 1Y does not translate in an exact value for the maturity). For each barrier $B(x)$ corresponding to a reference probability label x, we plot in Figure 4.2.4 the market foreign no-touch-up probabilities as a spread

$$p(x) = x - FNT^{Market}(B(x), T).$$

4.3 Condition for exact calibration to up-and-out call prices

In this section, we derive the main formula that links market up-and-out call prices, via a calibrated Brunick–Shreve volatility under model (4.2.5), to up-and-out call prices (which includes vanilla and no-touch prices) under model (4.2.1). We give an alternative proof to the one proposed in [25] that will ease the generalisation to stochastic rates in Section 4.9.

Let the discounted up-and-out call option price under model (4.2.1) for a notional of one unit of *ccy*1 be

$$C_{\mathrm{UO}}(K, B, T) = \mathbb{E}^{\mathbb{Q}^d}\left[D_T^d (S_T - K)^+ \mathbf{1}_{M_T < B}\right]. \tag{4.3.1}$$

4.3.1 Mimicking result and forward equation

We can now state the following result, which we will prove in the remainder of this section.

Theorem 4.1. *Assume that the marginal density $\phi(T, \cdot, \star)$ of (S_T, M_T) in (4.2.1) exists, is continuous and is strictly positive. Further assume that $\mathbb{E}^{\mathbb{Q}^d}\left[\gamma_T^2 \,|\, S_T = \cdot,\, M_T = \star\right]$ is continuous. Then, the up-and-out call price $C_{UO}(K, B, T)$ under model (4.2.1) satisfies the equation*

$$\begin{aligned}
& C_{UO}(K, B, T) - (S_0 - K)^+ \mathbf{1}_{M_0 < B} \\
= \; & -\int_0^T \bar{r}^f(t) C_{UO}(K, B, t)\, dt - \int_0^T \left(\bar{r}^d(t) - \bar{r}^f(t)\right) K \frac{\partial C_{UO}(K, B, t)}{\partial K} dt \\
+ \; & \frac{1}{2}\int_0^T \int_K^B K^2 \mathbb{E}^{\mathbb{Q}^d}\left[\gamma_t^2 \,|\, S_t = K,\, M_t = b\right] \frac{\partial^3 C_{UO}(K, b, t)}{\partial K^2 \partial b}\, db dt \\
- \; & \frac{1}{2}\int_0^T (B - K)^+ B^2 \mathbb{E}^{\mathbb{Q}^d}\left[\gamma_t^2 \,|\, S_t = B,\, M_t = B\right] \left.\frac{\partial^3 C_{UO}(K, B, t)}{\partial K^2 \partial B}\right|_{K=B} dt
\end{aligned} \tag{4.3.2}$$

for any strike $K > 0$, barrier $B > (K \vee S_0)$ and maturity T.

Equation (4.3.2) is similar to the equation derived in Chapter 2. Here, by using stochastic calculus (specifically, the Trotter-Meyer theorem) rather than PDE arguments, we are able to relax the smoothness condition on the joint density function of (S, M). We first state and prove auxiliary results necessary for the derivation of Theorem 4.1.

Lemma 4.2. *Given a filtered probability space $(\chi, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q}^d)$, let W be a standard Brownian motion and μ, Y two $\mathcal{F}_t$-adapted processes with Y continuous, of finite second moment, and a. s.*

$$\int_0^t \Big(|\mu_u| + Y_u^2\Big)\, du < \infty\,.$$

Consider a strictly positive continuous Itô process X and its running maximum M given by

$$\begin{cases} X_t = & X_0 + \int_0^t \mu_u\, du + \int_0^t Y_u\, dW_u \\ M_t = & \max\limits_{0\leq u\leq t} X_u\,, \end{cases}$$

and let $(\kappa_1, \kappa_2) \in \mathbb{R}_*^+ \times \mathbb{R}^+$ or $(\kappa_1, \kappa_2) \in \mathbb{R}_*^- \times \mathbb{R}_*^+$, where $\mathbb{R}_*^\pm = \mathbb{R}^\pm \backslash \{0\}$, such that

$$Z_t = \kappa_1 X_t + \kappa_2 M_t\,.$$

The joint density function $\phi(\cdot, \star, t)$ of (X_t, M_t), and $\mathbb{E}^{\mathbb{Q}^d}\big[D_t Y_t^2 \,|\, Z_t = \cdot,\, M_t = \star\big]$ are assumed to exist and be continuous. Further denote by $(l_t^a)_{t\geq 0}$ the local time of the process Z at level a. Then, for any continuous, integrable and positive $\mathcal{F}_t$-adapted semi-martingale G and any positive real number a,

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t G_u\, dl_u^a\right] = \frac{1}{|\kappa_1|}\int_0^t \int_{X_0\vee\alpha}^{\beta} \Big(\mathbb{E}^{\mathbb{Q}^d}\Big[G_u Y_u^2 \,|\, Z_u = a,\, M_u = b\Big]\phi\Big(\frac{a-\kappa_2 b}{\kappa_1}, b, u\Big)\Big)\, dbdu\,,$$

such that when $\kappa_1 > 0$,

$$\begin{cases} \alpha = \frac{a}{\kappa_1+\kappa_2}, \beta = \frac{a}{\kappa_2}, & \kappa_2 > 0, a > 0, \\ \alpha = \frac{a}{\kappa_1}, \beta = \infty, & \kappa_2 = 0, \end{cases}$$

and when $\kappa_1 < 0$,

$$\begin{cases} \alpha = \frac{a}{\kappa_2}, \beta = \frac{a}{\kappa_1+\kappa_2}, & \kappa_2 > -\kappa_1, a > 0, \\ \alpha = \frac{a}{\kappa_2}, \beta = \infty, & 0 < \kappa_2 \leq -\kappa_1\,. \end{cases}$$

Proof. For all integer $n > 0$, let

$$\delta_n^a(x) = \begin{cases} 0, & |x-a| > \frac{1}{n}, \\ \frac{n}{2}, & |x-a| \leq \frac{1}{n}\,. \end{cases} \tag{4.3.3}$$

Z is a semi-martingale since M is of finite variation. Then, following Proposition 3.3 in Chapter 3 we can write for any horizon time $T > 0$ and $0 < t \leq T$, that $\int_0^t G_s \delta_n^a(Z_s)\, d\langle Z\rangle_s$ converges almost surely,

and uniformly in time, to $\int_0^t G_s\, d_s l^a$. Therefore, we know that the following holds almost surely:

$$\int_0^t G_u\, dl_u^a = \lim_{n\to\infty} \int_0^t G_u \delta_n^a(Z_u) Y_u^2\, du\,.$$

This implies convergence in distribution, so that we can write

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t G_u\, dl_u^a\right] = \lim_{n\to\infty} \mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t G_u \delta_n^a(Z_u) Y_u^2\, du\right],$$

and, by the stochastic Fubini theorem, we get

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t G_u\, dl_u^a\right] &= \lim_{n\to\infty} \int_0^t \mathbb{E}^{\mathbb{Q}^d}\left[G_u \delta_n^a(Z_u)\, Y_u^2\right] du \\ &= \lim_{n\to\infty} \int_0^t \mathbb{E}^{\mathbb{Q}^d}\left[\delta_n^a(Z_u)\mathbb{E}^{\mathbb{Q}^d}\left[G_u Y_u^2 \mid Z_u,\, M_u\right]\right] du\,.\end{aligned}$$

We denote $\gamma(x,m,u) = \mathbb{E}^{\mathbb{Q}^d}\left[G_u Y_u^2 \mid Z_u = \kappa_1 x + \kappa_2 m,\, M_u = m\right]$, and since $Z_t = \kappa_1 X_t + \kappa_2 M_t$ we can write

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t G_u\, dl_u^a\right] &= \lim_{n\to\infty} \int_0^t \int_{X_0}^\infty \int_0^m \delta_n^a(\kappa_1 x + \kappa_2 m)\gamma(x,m,u)\phi(x,m,u)\, dxdmdu \\ &= \lim_{n\to\infty} \int_{X_0}^\infty \int_0^m \delta_n^{a-\kappa_2 m}(\kappa_1 x)\left(\int_0^t \gamma(x,m,u)\phi(x,m,u)\, du\right) dxdm \\ &= \lim_{n\to\infty} \int_{X_0}^\infty \frac{n}{2} \int_{\frac{a-\kappa_2 m}{\kappa_1} - \frac{1}{|\kappa_1| n}}^{\frac{a-\kappa_2 m}{\kappa_1} + \frac{1}{|\kappa_1| n}} \left(\int_0^t \gamma(x,m,u)\phi(x,m,u)\, du\right) dxdm,\end{aligned}$$

where we have used Fubini's Theorem in the second line. By the continuity assumptions on ϕ and γ, such that for large n, we deduce that

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t G_u\, dl_u^a\right] = \frac{1}{|\kappa_1|}\int_{X_0\vee\alpha}^\beta \int_0^t \left(\mathbb{E}^{\mathbb{Q}^d}\left[G_u Y_u^2 \mid Z_u = a,\, M_u = m\right]\phi\left(\frac{a-\kappa_2 m}{\kappa_1}, m, u\right)\right) dudm\,,$$

where the following condition needs to be satisfied

$$\begin{cases}\kappa_1 x + \kappa_2 m = a \\ 0 \le x \le m\end{cases},$$

which is equivalent to

$$\begin{cases} \frac{a}{\kappa_1+\kappa_2} \leq m \leq \frac{a}{\kappa_2}, & \kappa_1 \geq 0, \kappa_2 > 0, a > 0 \\ \frac{a}{\kappa_1} \leq m, & \kappa_1 > 0, \kappa_2 = 0 \\ \frac{a}{\kappa_2} \leq m \leq \frac{a}{\kappa_1+\kappa_2}, & \kappa_1 < 0, \kappa_2 \geq -\kappa_1, a > 0 \\ \frac{a}{\kappa_2} \leq m, & \kappa_1 < 0, 0 < \kappa_2 < -\kappa_1\,, \end{cases}$$

such that we are able to conclude. □

Lemma 4.2 provides a link between the local time of a linear combination of the spot and its maximum process and their joint density function. This will play a key role later, where we use the fact that the joint density is the third order derivative of the up-and-out call price with respect to the strike twice and the barrier once, up to a discount factor.

Lemma 4.3. *Given a filtered probability space $(\chi, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q}^d)$, let W be a standard $\mathcal{F}_t$-adapted Brownian motion and X_t an Itô process with volatility Y_t and drift μ_t, which are $\mathcal{F}_t$-adapted processes such that, a.s.,*

$$\int_0^t \left(|\mu_u| + Y_u^2\right) du < \infty$$

and Y_t strictly positive almost everywhere. Then, the running maximum process M_t can be written in terms of the local time l_t^0 of the draw-down process $Z_t = M_t - X_t$ at level zero as

$$M_t = M_0 + \frac{1}{2} l_t^0\,. \tag{4.3.4}$$

Proof. Z is positive by construction and we impose $Z_0 = 0$. Consider, for any $\epsilon > 0$, the function $g:\, z \to (\epsilon - z)^+$ defined on $\mathbb{R}^+$. Z is a semi-martingale since M is of finite variation, then by the Trotter-Meyer theorem [132] we can write

$$(\epsilon - Z_t)^+ - \epsilon = \int_0^t -\mathbf{1}_{Z_s \leq \epsilon}\, dM_s + \int_0^t \mathbf{1}_{Z_s \leq \epsilon}\, dX_s + \frac{1}{2} l_t^\epsilon \tag{4.3.5}$$

where $(l_t^a)_{t\geq 0}$ is the local time of Z_t at level a. Since dM_t is strictly positive if and only if Z_t is zero, i.e. the running maximum only grows when it is equal to the spot, we have

$$\int_0^t \mathbf{1}_{Z_s \leq \epsilon}\, dM_s = \int_0^t dM_s = M_t - M_0\,.$$

Also,

$$\int_0^t \mathbf{1}_{Z_s \leq \epsilon}\, dX_s = \int_0^t Y_s \mathbf{1}_{Z_s \leq \epsilon}\, dW_s + \int_0^t \mu_s \mathbf{1}_{Z_s \leq \epsilon}\, ds\,.$$

Since for any $0 \leq s \leq t$, $|\mu_s|$ is integrable by assumption, and

$$|\mu_s \mathbf{1}_{Z_s \leq \epsilon}| \leq |\mu_s|\,,$$

the dominated convergence theorem gives

$$\lim_{\epsilon\to 0}\int_0^t \mu_s \mathbf{1}_{Z_s\leq\epsilon}\,ds = \int_0^t \mu_s \mathbf{1}_{Z_s=0}\,ds\,,$$

where the time set where M grows is of zero Lebesgue measure (see [135]) such that we can write

$$\lim_{\epsilon\to 0}\int_0^t \mu_s \mathbf{1}_{Z_s\leq\epsilon}\,ds = 0\,.$$

Furthermore, the quadratic variation of $\int_0^t \mathbf{1}_{Z_s\leq\epsilon}\,dX_s$ is

$$\int_0^t Y_s^2 \mathbf{1}^2_{Z_s\leq\epsilon}\,ds\,,$$

such that the process $\lim_{\epsilon\to 0}\int_0^t \mathbf{1}_{Z_s\leq\epsilon}\,dX_s$ is a local martingale of quadratic variation

$$\lim_{\epsilon\to 0}\int_0^t Y_s^2 \mathbf{1}^2_{Z_s\leq\epsilon}\,ds = 0\,,$$

by another application of the dominated convergence theorem. Hence, since $\lim_{\epsilon\to 0}\int_0^t \mathbf{1}_{Z_s\leq\epsilon}\,dX_s$ is $\mathcal{F}_t$-adapted, the process is constant and its value is zero. By taking limit on left and right hand side of (4.3.5), we get

$$\lim_{\epsilon\to 0}\Big((\epsilon - Z_t)^+ - \epsilon\Big) = 0 = M_0 - M_t + \frac{1}{2}l_t^0$$

almost surely. □

Remark. We note that if the volatility is not almost surely strictly positive there would exist time intervals where M_t sticks to X_t, leading to $\int_0^t \mathbf{1}_{Z_s\leq\epsilon}\,dS_s$ being non-zero, which would prevent us from concluding. This result can also be found in the path-wise sense in Dupire [55] with the use of functional Itô calculus. Furthermore, in the case of Brownian motion, the result is strongly related to Lévy's Theorem (see [130]), namely, for a Brownian motion B_t , L_t^0 the local time of B_t at level zero and $\overline{B_t}$ its running maximum up to t, the following two-dimensional processes are equal in law:

$$\Big(\overline{B}_t - B_t, \overline{B}_t\Big) \overset{\text{law}}{=} \Big(|B_t|, L_t^0\Big)\,.$$

Finally, by combining Lemma 4.2 and Lemma 4.3 we can give the proof of Theorem 4.1.

Proof of Theorem 4.1. Let $K \in \mathbb{R}^+$, $B > (K \vee S_0)$, $0 < t \leq T$ and define a sequence of almost everywhere C^1 functions which converges point-wise to the step function $\mathbf{1}_{x<B}$, for any real x and a positive integer n

$$I_n^B(x) = \begin{cases} \frac{1+\mathrm{sgn}(B-x)}{2}, & |B-x| \geq \frac{1}{n}, \\ \frac{1+n(B-x)}{2}, & |B-x| < \frac{1}{n}, \end{cases}$$

with $\mathrm{sgn}(x) = -1$ for $x \leq 0$ and $\mathrm{sgn}(x) = 1$ for $x > 0$. The Trotter-Meyer theorem [132] gives

$$(S_t - K)^+ - (S_0 - K)^+ \quad = \quad \int_0^t \mathbf{1}_{S_u \geq K}\, dS_u + \frac{1}{2} L_t^K \,,$$

where L_t^K is the local time at level K of the process S_t. Additionally, since I_n^B has a uniformly bounded derivative $-\delta_n^B$ as defined in (4.3.3) and M is of finite variation, we can write

$$I_n^B(M_T) - I_n^B(M_0) \quad = \quad -\int_0^T \delta_n^B(M_t)\, dM_t \,. \tag{4.3.6}$$

By Itô's lemma,

$$\begin{aligned} D_t^d (S_T - K)^+ I_n^B(M_T) \quad &= \quad (S_0 - K)^+ I_n^B(M_0) - \int_0^T I_n^B(M_t) r^d(t) D_t^d (S_t - K)^+ \, dt \\ &+ \quad \int_0^T I_n^B(M_t)\, D_t^d d(S_t - K)^+ + \int_0^T D_t^d (S_t - K)^+ \, dI_n^B(M_t) \\ &+ \quad \left\langle D_t^d (S_t - K)^+,\, I_n^B(M_T) \right\rangle, \end{aligned}$$

where the cross-variation term is zero (see [135]). We use (4.3.6) and Lemma 4.3, take expectations and write

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}^d}\left[D_t^d (S_T - K)^+ I_n^B(M_T) \right] \quad &= \quad (S_0 - K)^+ I_n^B(M_0) - \mathbb{E}^{\mathbb{Q}^d}\left[\int_0^T I_n^B(M_t) r^d(t) D_t^d (S_t - K)^+ \, dt \right] \\ &+ \quad \mathbb{E}^{\mathbb{Q}^d}\left[\int_0^T D_t^d I_n^B(M_t) \mathbf{1}_{S_t \geq K}\, dS_t \right] + \frac{1}{2} \mathbb{E}^{\mathbb{Q}^d}\left[\int_0^T D_t^d I_n^B(M_t)\, d_t L^K \right] \\ &- \quad \frac{1}{2} \mathbb{E}^{\mathbb{Q}^d}\left[\int_0^T D_t^d (S_t - K)^+ \delta_n^B(M_t)\, d_t l^0 \right], \end{aligned}$$

where $(l_t^0)_{t \geq 0}$ is the local time of $M_t - S_t$ at level 0. As $D_t^d I_n^B(M_t)$ is a semi-martingale since M is a finite variation process, we can use Lemma 4.2 with $G_t = D_t^d I_n^B(M_t)$, $X_t = S_t$, $Y_u = S_u \gamma_u$, $\kappa_1 = 1$

and $\kappa_2 = 0$ such that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^T D_t^d I_n^B(M_t)\, d_t L^K\right] \\
=\ & \int_0^T \int_{S_0\vee K}^{\infty} \Big(\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d I_n^B(b) K^2 \gamma_t^2 \,|\, S_t = K,\, M_t = b\Big]\phi(K,b,t)\Big) dbdt \\
=\ & \int_0^T \int_{S_0\vee K}^{B-\frac{1}{n}} \Big(K^2 \mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d \gamma_t^2 \,|\, S_t = K,\, M_t = b\Big]\phi(K,b,t)\Big) dbdt \\
+\ & \int_0^T \int_{B-\frac{1}{n}}^{B+\frac{1}{n}} \frac{1+n(B-b)}{2} \Big(K^2 \mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d \gamma_t^2 \,|\, S_t = K,\, M_t = b\Big]\phi(K,b,t)\Big) dbdt.
\end{aligned}$$

By taking the limit and using continuity of the integrand,

$$\lim_{n\to\infty} \mathbb{E}^{\mathbb{Q}^d}\left[\int_0^T D_t^d I_n^B(M_t)\, d_t L^K\right] = \int_0^T \int_{S_0\vee K}^{B} \Big(K^2 \mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d \gamma_t^2 \,|\, S_t = K,\, M_t = b\Big]\phi(K,b,t)\Big) dbdt\,.$$

Similarly, as $D_t^d (S_t - K)^+ \delta_n^B(M_t)$ is a semi-martingale as a product of two semi-martingales, we then use Lemma 4.2 with $G_t = D_t^d (S_t - K)^+ \delta_n^B(M_t)$, $X_t = S_t$, $Y_u = S_u \gamma_u$, $\kappa_1 = -1$ and $\kappa_2 = 1$ such that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^T D_t^d (S_t - K)^+ \delta_n^B(M_t)\, dl_t^0\right] \\
=\ & \int_0^t \int_{S_0}^{\infty} \Big(\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d (b-K)^+ \delta_n^B(b) S_t^2 \gamma_t^2 \,|\, S_t = M_t,\, M_t = b\Big]\phi(b,b,t)\Big)\, dbdt \\
=\ & \int_0^t \int_{B-\frac{1}{n}}^{B+\frac{1}{n}} \left(\frac{n}{2}(b-K)^+ b^2 \mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d \gamma_t^2 \,|\, S_t = b,\, M_t = b\Big]\phi(b,b,t)\right) dbdt\,.
\end{aligned}$$

By continuity of the integrand, we also have

$$\lim_{n\to\infty} \mathbb{E}^{\mathbb{Q}^d}\left[\int_0^T D_t^d (S_t - K)^+ \delta_n^B(M_t)\, dl_t^0\right] = \int_0^t (B-K)^+ B^2 \mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d \gamma_t^2 \,|\, S_t = B,\, M_t = B\Big]\phi(B,B,t)\, dt\,.$$

Furthermore,

$$\int_0^T D_t^d I_n^B(M_t) \mathbf{1}_{S_t \geq K} S_t^2 \gamma_t^2\, dW_t$$

is a true martingale since we assumed $\mathbb{E}^{\mathbb{Q}^d}\Big[\int_0^T S_t^2 \gamma_t^2\, dt\Big] < \infty$, such that we can write, using the

dominated convergence theorem,

$$\begin{aligned}
\lim_{n\to+\infty}\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d(S_T-K)^+I_n^B(M_T)\Big] &= \mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d(S_T-K)^+\mathbf{1}_{M_T<B}\Big]\\
&= (S_0-K)^+\mathbf{1}_{M_0<B}-\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^T r^d(t)D_t^d(S_t-K)^+\mathbf{1}_{M_T<B}\,dt\right]\\
&+ \int_0^T\mathbb{E}^{\mathbb{Q}^d}\Big[\mathbf{1}_{M_t<B}\mathbf{1}_{S_t\geq K}D_t^dS_t\Big(r^d(t)-r^f(t)\Big)\Big]\,dt\\
&+ \frac{1}{2}\int_0^T\int_{S_0\vee K}^B\Big(K^2\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d\gamma_t^2\,|\,S_t=K,\,M_t=b\Big]\phi(K,b,t)\Big)\,dbdt\\
&- \frac{1}{2}\int_0^T(B-K)^+B^2\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d\gamma_t^2\,|\,S_t=B,\,M_t=B\Big]\phi(B,B,t)\,dt\,.
\end{aligned}$$

And since

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\Big[\mathbf{1}_{S_t\geq K}\mathbf{1}_{M_t<B}S_t-(S_t-K)^+\mathbf{1}_{M_t<B}\Big] &= \mathbb{E}^{\mathbb{Q}}\Big[\mathbf{1}_{S_t\geq K}\mathbf{1}_{M_t<B}(S_t-K+K)-(S_t-K)^+\mathbf{1}_{M_t<B}\Big]\\
&= K\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{S_t\geq K}\mathbf{1}_{M_t<B}]\,, \qquad (4.3.7)
\end{aligned}$$

we get

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d(S_T-K)^+\mathbf{1}_{M_T<B}\Big] &= (S_0-K)^+\mathbf{1}_{M_0<B}-\int_0^T r^f(t)\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d(S_t-K)^+\mathbf{1}_{M_T<B}\Big]\,dt\\
&+ \int_0^T K\Big(r^d(t)-r^f(t)\Big)\mathbb{E}^{\mathbb{Q}^d}\Big[\mathbf{1}_{M_t<B}\mathbf{1}_{S_t\geq K}D_t^d\Big]\,dt\\
&+ \frac{1}{2}\int_0^T\int_{S_0\vee K}^B\Big(K^2\mathbb{E}^{\mathbb{Q}^d}\Big[\gamma_t^2\,|\,S_t=K,\,M_t=b\Big]D_t^d\phi(K,b,t)\Big)\,dbdt\\
&- \frac{1}{2}\int_0^T(B-K)^+B^2\mathbb{E}^{\mathbb{Q}^d}\Big[\gamma_t^2\,|\,S_t=B,\,M_t=B\Big]D_t^d\phi(B,B,t)\,dt\,.
\end{aligned}$$

Furthermore, writing (4.3.1) as integral of the payoff over the joint density $\phi(\cdot,\star,T)$ of S_T and M_T, by differentiation,

$$\begin{aligned}
\frac{\partial C_{\mathrm{UO}}(K,B,T)}{\partial K} &= -D_T^d\int_K^{+\infty}\int_{S_0\vee K}^B\phi(x,y,T)\,dxdy\\
&= -\mathbb{E}^{\mathbb{Q}}\Big[D_T^d\mathbf{1}_{S_T\geq K}\mathbf{1}_{M_T<B}\Big]\,,\\
\frac{\partial C_{\mathrm{UO}}(K,B,T)}{\partial B} &= D_T^d\int_K^{+\infty}(x-K)\phi(x,B,T)\,dx\,,\\
\frac{\partial^3 C_{\mathrm{UO}}(K,B,T)}{\partial K^2\partial B} &= D_T^d\phi(K,B,T)\,, \qquad (4.3.8)
\end{aligned}$$

which allows us to conclude the proof by substitution. □

Remark. An interesting application of Theorem 4.1 is the Markovian projection of model (4.2.1) onto a Brunick–Shreve model (4.2.5) to mimic up-and-out call prices of model (4.2.5).

Corollary 4.4. *Under the assumptions of Theorem 4.1, the up-and-out call price $C_{UO}(K,B,T)$ under model (4.2.1) matches the price C_{UO}^{BSV} under the Brunick–Shreve volatility model (4.2.5) for any strike $K>0$, barrier $B>(K\vee S_0)$ and maturity T if and only if, for all $K,T>0$ and all $B>(K\vee S_0)$,*

$$\mathbb{E}^{\mathbb{Q}^d}\Big[\gamma_T^2 \,|\, S_T=K,\, M_T=B\Big]=\sigma_{BSV}^2(K,B,T)\,, \tag{4.3.9}$$

where σ_{BSV} is a calibrated Brunick–Shreve volatility as in (4.2.5).

Corollary 4.4 proves in an alternative manner the result derived in [25] for the condition (4.3.9). The present approach lies closer to the one for local volatility in [54], but requires stronger conditions, i.e. existence of a joint density function, than the result presented in [25].

Proof. First, we want to ensure that (4.3.9) is a necessary condition for

$$C_{\mathrm{UO}}(K,B,T)=C_{\mathrm{UO}}^{BSV}(K,B,T)\,. \tag{4.3.10}$$

Let $K=0$ and apply Proposition 4.1 to both model (4.2.1) and model (4.2.5) and subtracting the two equations, noting that both up-and-out call prices and their derivatives match under both models, we have for any $T>0$, $B>S_0$,

$$\int_0^T B^3\Big(\sigma_{BSV}^2(B,B,t)-\mathbb{E}^{\mathbb{Q}^d}\Big[\gamma_t^2 \,|\, S_t=B,\, M_t=B\Big]\Big)\left.\frac{\partial^3 C_{\mathrm{UO}}(K,B,t)}{\partial K^2\partial B}\right|_{K=B} dt \quad=0\,. \tag{4.3.11}$$

Furthermore, since $B>S_0$,

$$\frac{\partial^3 C_{\mathrm{UO}}(K,B,T)}{\partial K^2\partial B}=D_T^d\phi(K,B,T)>0$$

and since (4.3.11) is true for any time T we have

$$\sigma_{BSV}^2(B,B,T)=\mathbb{E}^{\mathbb{Q}^d}\Big[\gamma_T^2 \,|\, S_T=B,\, M_T=B\Big]. \tag{4.3.12}$$

Similarly, when $K>0$, we apply Proposition 4.1 to both model (4.2.1) and model (4.2.5), subtract the two equations, combine it with (4.3.12) and get

$$\int_0^T\int_K^B K^2\Big(\sigma_{BSV}^2(K,b,t)-\mathbb{E}^{\mathbb{Q}^d}\Big[\gamma_t^2 \,|\, S_t=K,\, M_t=b\Big]\Big)\frac{\partial^3 C_{\mathrm{UO}}(K,b,t)}{\partial K^2\partial b}\,dbdt=0$$

which has to be true for any $K > 0$, $T > 0$, and $B > S_0 \vee K$ such that we conclude

$$\sigma^2_{BSV}(K, B, T) \quad = \quad \mathbb{E}^{\mathbb{Q}^d}\left[\gamma^2_T \,|\, S_T = K,\, M_T = B\right],$$

and (4.3.9) holds. The above statements are true since we assumed strictly positive density function, i.e.

$$\frac{\partial^3 C_{\mathrm{UO}}(K, B, T)}{\partial K^2 \partial B} > 0\,.$$

Moreover, it is easy to see that (4.3.9) is also a sufficient condition for (4.3.10). □

4.3.2 Heston-type LSV and LMSV models

The main assumption (4.2.2) made on the generic model (4.2.1) ensures the true martingale property of the local martingale

$$\int_0^t S_u \gamma_u \, dW_u.$$

In the case of the variance process being driven by a CIR diffusion, we get $\gamma_u = \sqrt{V_u}$ as defined for models (4.2.7) and (4.2.6) and, as explained in Chapter 3, the moment bound condition is only true up to a critical time T^*. Indeed, [8] show that the moment explodes in finite time for the Heston model, a property that is inherited by our Heston-type LSV and LMSV model. While T^* is usually large enough for practical applications, it is important to keep in mind that a large vol-of-vol might drastically reduce the explosion time.

4.3.2.1 Second order moment explosion with uncertain volatility and vol-of-vol

We define a Heston-type local-stochastic volatility model with uncertain volatilities. Our goal is to derive a general result about moment explosions for Heston-type LSV models where the leverage function and vol-of-vol are general bounded adapted processes. In particular, this includes models where both the leverage function and the vol-of-vol are bounded functions of the state variables and time as (4.2.6) and (4.2.7).

Let the diffusion of interest be

$$\begin{cases} \dfrac{d\bar{S}_t}{\bar{S}_t} = \left(\bar{r}^d(t) - \bar{r}^f(t)\right) dt + \alpha_t \sqrt{V_t}\, dW_t \\ d\bar{V}_t = \kappa(\theta - V_t)\, dt + \xi_t \sqrt{V_t}\, dW^V_t, \end{cases} \tag{4.3.13}$$

where α is an uncertain volatility process and ξ is an uncertain vol-of-vol process such that

$$(\alpha_t, \xi_t) \in A \equiv [\alpha_{min}, \alpha_{max}] \times [\xi_{min}, \xi_{max}] \tag{4.3.14}$$

and where (W, W^V) is a two-dimensional standard Brownian motion with constant correlation ρ.

The approach here could be extended to uncertain correlation (it is easy to see from the argument below that the worst case is always attained for the largest correlation), but this is not relevant to our application.

Let $\bar{X}_t = \ln \bar{S}_t$ and let the moment generating function of $\bar{X}_T$ be defined as

$$f(u, \tau = T - t, x, v) \equiv \mathbb{E}\Big[e^{u\bar{X}_T} \,|\, \bar{X}_t = x,\, \bar{V}_t = v\Big] = \mathbb{E}\Big[\bar{S}_T^u \,|\, \bar{S}_t = e^x,\, \bar{V}_t = v\Big],$$

where $f(u, 0, x, v) = e^{ux}$, so that $f(u, \ldots)$ describes the u-th moment of S_T for some $u > 1$.

For constant volatility parameters $\bar{\alpha}$ and $\bar{\xi}$, we know by the Feynman–Kac theorem that f follows the Heston PDE.

Recall the moment explosion time for constant parameters $\bar{\alpha}$ and $\bar{\xi}$, i.e. under the classical Heston model. Let

$$\begin{aligned} b(u) &= \rho\bar{\alpha}\bar{\xi}u - \kappa \\ \Delta(u) &= b^2(u) - 2\bar{\xi}^2\Big(\frac{1}{2}\bar{\alpha}^2\Big(u^2 - u\Big)\Big), \end{aligned}$$

and from [66, 104, 8] we know that for a pure Heston model the solution to the Riccati ODE (4.3.19) and consequently the moments of order u of S are finite up to a critical time T^* defined as, for any $u > 1$,

$$T^* = \begin{cases} +\infty & \Delta(u) \geq 0,\, b(u) < 0\,, \\ \Big(\sqrt{\Delta(u)}\Big)^{-1} \ln\Big(\frac{b(u)+\sqrt{\Delta(u)}}{b(u)-\sqrt{\Delta(u)}}\Big) & \Delta(u) \geq 0,\, b(u) > 0\,, \\ \frac{2}{\sqrt{-\Delta(u)}}\Big(\arctan\frac{\sqrt{-\Delta(u)}}{b(u)} + \pi\mathbf{1}_{b(u)<0}\Big) & \Delta(u) < 0\,. \end{cases} \tag{4.3.15}$$

We now return to the general case. Since we are looking for a lower bound for the moment explosion of order u, we are interested in the largest possible value that the moments can attain under the box constraint on the volatility processes,

$$\bar{f}(u, \tau = T - t, x, v) \equiv \sup_{(\alpha,\xi)\in\mathcal{A}} \mathbb{E}\Big[e^{u\bar{X}_T} \,|\, \bar{X}_t = x,\, \bar{V}_t = v\Big],$$

where $\mathcal{A}$ is the set of admissible processes satisfying (4.3.14).

By a standard dynamic programming argument, this leads to the HJB equation

$$\begin{aligned} \frac{\partial \bar{f}}{\partial \tau} &= \kappa(\theta - v)\frac{\partial \bar{f}}{\partial v} + \Big(\bar{r}^d(t) - \bar{r}^{\bar{f}}(t)\Big)\frac{\partial \bar{f}}{\partial x} \\ &\quad + \sup_{(\bar{\alpha},\bar{\xi})\in A}\left\{\frac{1}{2}\bar{\alpha}^2 v\left(\frac{\partial^2 \bar{f}}{\partial x^2} - \frac{\partial \bar{f}}{\partial x}\right) + \rho\bar{\alpha}^2\bar{\xi}^2 v\frac{\partial^2 \bar{f}}{\partial x \partial v} + \frac{1}{2}\bar{\xi}^2 v\frac{\partial^2 \bar{f}}{\partial v^2}\right\}. \end{aligned} \tag{4.3.16}$$

We make an ansatz for the solution in exponential-affine form like in the Heston case [87],

$$\bar{f}(u,\tau,x,v) = e^{q(u,\tau)+v\psi(u,\tau)+ux} \tag{4.3.17}$$

and by inserting in (4.3.16) we get a system of non-linear ODEs which resemble the classical Heston linear ODEs

$$\frac{\partial q(u,\tau)}{\partial \tau} = \kappa\theta\psi(u,\tau) + \Big(\bar{r}^d(T-\tau) - \bar{r}^f(T-\tau)\Big)u \tag{4.3.18}$$

$$\begin{aligned}\frac{\partial \psi(u,\tau)}{\partial \tau} &= -\kappa\psi(u,\tau) \\ &\quad + \sup_{(\bar{\alpha},\bar{\xi})\in A}\left\{\psi^2(u,\tau)\left(\frac{1}{2}\bar{\xi}^2\right) + \psi(u,\tau)\Big(\rho\bar{\alpha}\bar{\xi}u\Big) + \frac{1}{2}\bar{\alpha}^2\Big(u^2-u\Big)\right\},\end{aligned} \tag{4.3.19}$$

with initial conditions $\psi(u,0) = q(u,0) = 0$.

We note that $\bar{\xi}$ and ψ only enter the second part of (4.3.19) via their product, and that the quadratic form in $(\bar{\alpha},\bar{\xi})$ is positive definite if $\rho^2 \leq 1 - u^{-1}$. We are interested in the case $u > 2$ so that this condition is safely satisfied for our parameter regimes and we will assume in the following that this is the case.

Then the supremum in (4.3.19) is always attained for the corners of A.

$\bar{\alpha}$	v_0	θ	κ	ρ	$\bar{\xi}$
$\in [0.735, 1.229]$	0.00827	0.01564	0.7147	-0.4429	$\in [0.003, 0.169]$

Table 4.3.1: Uncertain volatility bounds and Heston parameters.

Figure 4.3.1 shows the results for $u = 2$ and calibrated parameters as given in Table 4.3.1, while for Figure 4.3.2 a range

$$\bar{\xi} \in [0.9, 1.5]$$

was used. The non-linear Riccati ODE (4.3.19) is solved numerically using an explicit Runge–Kutta [26] method combined with a BFGS optimiser with box constraints [156] for generality [2]. The solution ψ is computed accurately and fast with 2000 time steps. Shown are also the optimising parameters $(\bar{\alpha},\bar{\xi})$ as a function of time as well as the solution for (α_{max},ξ_{min}), (α_{min},ξ_{max}) and (α_{max},ξ_{max}).

[2] It would suffice to take the maximum over the four extremal parameter pairs. Indeed, we will see that no more than two regime switches occur, i.e. when ψ reaches certain thresholds, such that the ODE could be solved semi-analytically by a sequence of linear Riccati ODEs.

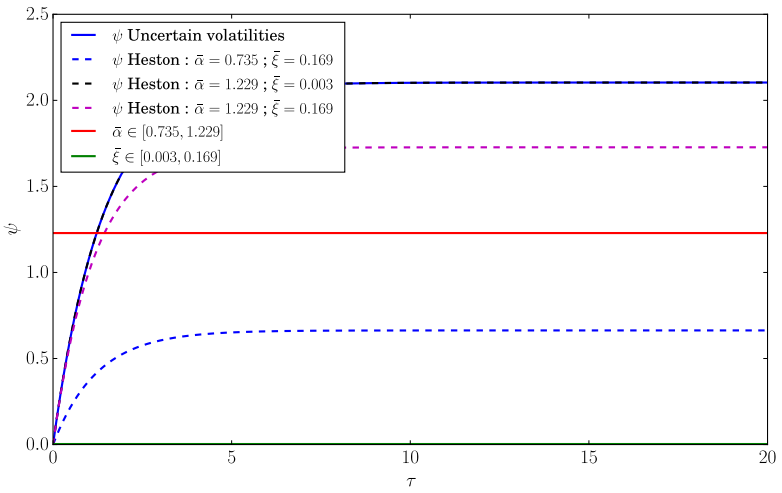


Figure 4.3.1: Steady state of ψ for the uncertain volatility model (4.3.13) in accordance with the LSV-LVV calibrated parameters from Table 4.3.1.

It is seen from (4.3.19), and from Figures 4.3.1 and 4.3.2, that for small Ψ (small τ) the optimal pair of parameters is $(\alpha_{max}, \xi_{min})$.

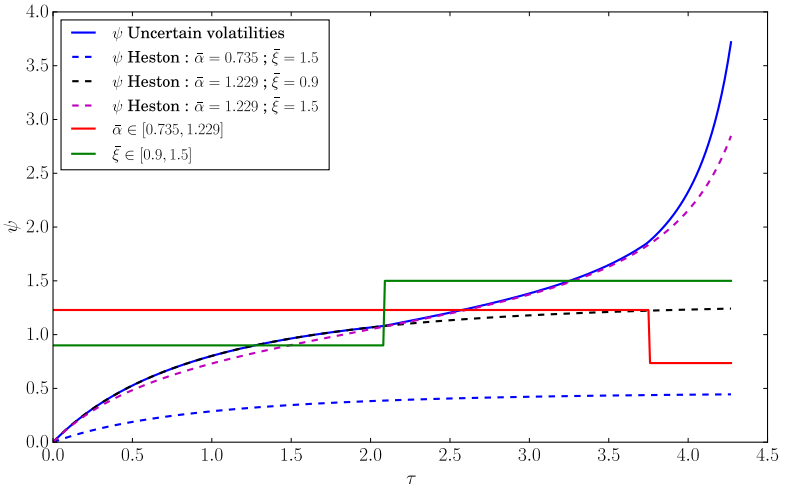


Figure 4.3.2: Explosion of ψ for the uncertain volatility model with $\bar{\xi} \in [0.9, 1.5]$.

If large enough ψ is reached, as in Figure 4.3.2 (this is always the case for τ nearing the explosion time), the optimal pair of parameters is $(\alpha_{min}, \xi_{max})$. This transition goes via $(\alpha_{max}, \xi_{max})$.

Figure 4.3.1 shows that the steady state solution of ψ is reached before $T = 20$ and therefore, the

LSV-LVV model (4.2.7) calibrated on vanilla and no-touch options has finite moment of order two for any positive time. This is in line with the fact that for any pair of parameters $\left(\bar{\alpha}, \bar{\xi}\right)$, we have

$$\begin{aligned} 2\rho\bar{\alpha}\bar{\xi} - \kappa &< 0 \\ \left(2\rho\bar{\alpha}\bar{\xi} - \kappa\right)^2 - 2\bar{\xi}^2\bar{\alpha}^2 &\geq 0\,, \end{aligned}$$

such that according to (4.3.15), $T^* = +\infty$.

For the example in Figure 4.3.2, the moment explosion of order two occurs at $T^* = 4.27$. We also note that none of the solutions with $(\alpha_{max}, \xi_{min})$ or $(\alpha_{min}, \xi_{max})$ shows blow-up, and it occurs only through the regime-switch or the solution with $(\alpha_{max}, \xi_{max})$.

A simpler procedure allows a similar conclusion for the variance process. The u-th moment of V can be written

$$\bar{g}(u, \tau = T - t, v) \equiv \sup_{\xi \in [\xi_{min}, \xi_{max}]} \mathbb{E}\Big[\bar{V}_T^u \,|\, \bar{V}_t = v\Big],$$

with $\bar{g}(u, 0, v) = v^u$. By a standard dynamic programming argument, this leads to the HJB equation

$$\frac{\partial \bar{g}}{\partial \tau} = \kappa(\theta - v)\frac{\partial \bar{g}}{\partial v} + \frac{1}{2}\sup_{\bar{\xi} \in [\xi_{min}, \xi_{max}]} \left\{ \bar{\xi}^2 v \frac{\partial^2 \bar{g}}{\partial v^2} \right\}. \tag{4.3.20}$$

The u-th moment of the classical CIR process with constant volatility is available in closed-form in [**?**], is convex in the initial variance v_0 and is finite for any $T > 0$ as discussed in [91]. Therefore, the convexity property allows to re-write (4.3.20) as

$$\frac{\partial \bar{g}}{\partial \tau} = \kappa(\theta - v)\frac{\partial \bar{g}}{\partial v} + \frac{1}{2}\xi_{max}^2 v \frac{\partial^2 \bar{g}}{\partial v^2}\,, \tag{4.3.21}$$

and since the moment of any order of v with a vol-of-vol ξ_{max} is finite, we can conclude that $\bar{g}$ is finite for any $u > 0$ and $T > 0$.

4.3.2.2 Consistency condition for exact calibration

Under the LSV-LVV model (4.2.7), Corollary 4.4 only applies if condition (4.2.2),

$$\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t S_s^2 V_s ds\right] < \infty\,,$$

is satisfied. In that respect, the LSV-LVV model (4.2.7) can be regarded as an uncertain volatility model (4.3.13), where calibrated box constraints and parameters are displayed in Table 4.3.1, following the procedure presented in Section 4.7. Therefore, from the HJB Riccati ODE (4.3.19), we know that

for any $\epsilon > 0$ and $u = 2 + \epsilon$,

$$\sup_{u\in[0,t]} \mathbb{E}^{\mathbb{Q}^d}[S_s^u]\,,$$

is finite if the solution ψ to the ODE (4.3.19) up to time $t > 0$ does not blow-up. We choose an arbitrarily small $\epsilon > 0$, and combined with parameters from Table 4.3.1, we have shown numerically in Figure 4.3.1 that $\mathbb{E}^{\mathbb{Q}^d}[S_s^u]$ is finite for any $t > 0$. On the other hand, from (4.3.21) and [91], we know that

$$\sup_{s\in[0,t]} \mathbb{E}^{\mathbb{Q}^d}\left[V_s^{\frac{u}{u-2}}\right] < \infty\,.$$

Using Hölder's inequality with the pair $\left(\frac{\omega}{2}, \frac{\omega}{\omega-2}\right)$,

$$\mathbb{E}^{\mathbb{Q}^d}\left[S_s^2 V_s\right] \leq \mathbb{E}^{\mathbb{Q}^d}[S_s^u]^{\frac{2}{u}} \mathbb{E}^{\mathbb{Q}^d}\left[V_s^{\frac{u}{u-2}}\right]^{\frac{u-2}{u}} < \infty. \tag{4.3.22}$$

Finally, using the Fubini theorem, $\mathbb{E}^{\mathbb{Q}^d}\left[\int_0^t S_s^2 V_s\, ds\right] < \infty$ and therefore, condition (4.2.2) is satisfied.

A similar reasoning can be applied to the LMSV model (4.2.6), where an analytical and conservative lower bound can be found as well.

Corollary 4.5. *In model (4.2.6), assume that α is uniformly bounded inside the interval $[\alpha_{min}, \alpha_{max}]$, and that the marginal density $\phi(T, \cdot, \star)$ of (S_T, M_T) and $\mathbb{E}^{\mathbb{Q}^d}[V_T \,|\, S_T = \cdot,\, M_T = \star]$ are continuous. Then, the up-and-out call price $C_{UO}(K, B, T)$ under model (4.2.7) matches the price C_{UO}^{BSV} under the Brunick–Shreve volatility model (4.2.5) for any strike $K > 0$, barrier $B > (K \vee S_0)$ and maturity $T < T^*$ if and only if, for all $K, T > 0$ and all $B > (K \vee S_0)$,*

$$\alpha^2(K, B, T)\mathbb{E}^{\mathbb{Q}^d}[V_T \,|\, S_T = K,\, M_T = B] = \sigma_{BSV}^2(K, B, T)\,, \tag{4.3.23}$$

where σ_{BSV} is a calibrated Brunick–Shreve volatility as in (4.2.5) and, T^* is the largest time such that the solution to,

$$\frac{\partial\psi(\tau)}{\partial\tau} = -\kappa\psi(\tau) + \sup_{\bar\alpha\in[\alpha_{min},\alpha_{max}]}\left\{\frac{1}{2}\bar\xi^2\psi^2(\tau) + 2\rho\bar\alpha\bar\xi\psi(\tau) + \bar\alpha^2\right\},$$

with initial condition $\psi(0) = 0$, is finite. Denoting $\varphi = 2 + \sqrt{2}$ and $\zeta = \xi\alpha_{max}$, a conservative lower bound for T^* is

$$\begin{cases} T_0^* = \frac{2}{\sqrt{\varphi^2\zeta^2 - \kappa^2}}\left[\frac{\pi}{2} + \arctan\left(\frac{\kappa}{\sqrt{\varphi^2\zeta^2-\kappa^2}}\right)\right], & \text{if } \kappa < \varphi\zeta, \\ T_0^* = \infty, & \text{if } \kappa \geq \varphi\zeta. \end{cases} \tag{4.3.24}$$

Proof. Similarly to the LSV-LVV model, we can use (4.3.22) to conclude on the martingale property (4.2.2) from the HJB Riccati ODE (4.3.19). Additionally, for any $t \leq T_0^*$, the true martingale property

of $\int_0^t S_u\sqrt{V_u}\,dW_u$ can be obtained as in Chapter 3, which leads to the conservative explosion time formula (4.3.24). Finally, a direct application of Corollary 4.4 gives directly (4.3.23). □

To ensure that (4.2.2) is valid for the LMSV model (4.2.6), we use the uncertain volatility model (4.3.13) with parameters in Table 4.3.1, a constant vol-of-vol $\xi = 0.0947$ and box constraints on the leverage function

$$\bar{\alpha} \in [0.307, 1.465]\,,$$

obtained from the calibration procedure in Section 4.6. This yields finite moments of order two for any time $t > 0$ by solving the ODE (4.3.19) numerically.

4.3.3 Fokker–Planck PDE for the joint density of (S, M, V)

The calibration by Markovian projection presented in this work can be facilitated by the solution of the Kolmogorov PDE for the joint density of (S, M, V) under a stochastic volatility model. Both the LSV-LVV and the LMSV model can be written in a more generic form as

$$\begin{cases} \dfrac{dS_t}{S_t} = (\bar{r}^d(t) - \bar{r}^f(t))\,dt + \sigma(S_t, M_t, V_t, t)\,dW_t \\ dV_t = f(S_t, V_t, t)\,dt + g(S_t, V_t, t)\,dW_t^V \\ M_t = \max\limits_{0\le u\le t} S_u, \end{cases} \tag{4.3.25}$$

where f and g are assumed to be smooth continuous functions. Let the region

$$\mathcal{H} = \{(x, y, z) : 0 < x \vee S_0 < y,\ z > 0\}\,.$$

We assume that the density ϕ (under $\mathbb{Q}^d$) of the Markovian process (S_t, M_t, V_t) started at (S_0, S_0, v_0) at time 0 exists in the region $\mathcal{H}$. We acknowledge that this assumption is not straightforward to verify and refer to [63] for the existence in the Brunick–Shreve case without stochastic volatility.

Theorem 4.6. *Assume that ϕ belongs to $C^{2,1,2,1}(\mathcal{H} \times \mathbb{R}^+)$ and vanishes for $z \to +\infty$. Additionally, assume that g vanishes for $z = 0$, then ϕ is solution of the following Kolmogorov forward equation*

$$\begin{cases} \frac{\partial\phi}{\partial t} + (\bar{r}^d(t) - \bar{r}^f(t))\frac{\partial(x\phi)}{\partial x} + \frac{\partial(f\phi)}{\partial z} - \frac{1}{2}\left(\frac{\partial^2\left(x^2\sigma^2\phi\right)}{\partial x^2} + \frac{\partial^2\left(g^2\phi\right)}{\partial z^2} + 2\frac{\partial^2(\rho x\sigma g\phi)}{\partial x\partial z}\right) = 0, & (x,y,z) \in \mathcal{H},\, t > 0 \\ \frac{1}{2}\frac{\partial\left(g^2\phi\right)}{\partial z} - f\phi + \rho\frac{\partial(\sigma g x\phi)}{\partial x}\Big|_{z=0} = 0, & t > 0 \\ \bar{r}^d(t) - \bar{r}^f(t))y\phi\big|_{x=y} - \frac{1}{2}\left(\frac{\partial\left(\sigma^2x^2\phi\right)}{\partial y}\Big|_{x=y} + 2\,\frac{\partial\left(\sigma^2x^2\phi\right)}{\partial x}\Big|_{x=y} + 2\,\frac{\partial(\rho x\sigma g\phi)}{\partial z}\big|_{x=y}\right) = 0, & x = y,\, t > 0 \\ \phi(S_0, S_0, z, t) = 0, & z \in [0, +\infty[,\, t > 0 \\ \lim_{z\to\infty}\phi(x,y,z,t) = 0, & (x,y) \in \Omega, t > 0 \\ \lim_{t\to 0}\phi(x,y,z,t) = \delta(x - S_0, y - S_0, z - v_0), & (x,y,z) \in \mathcal{H} \end{cases} \tag{4.3.26}$$

Proof. The derivation can be obtained by extending the proof in Chapter 2 to stochastic volatility and is provided in Sub-section 4.10. □

Remark. The boundary condition can be written as

$$\frac{1}{2}\frac{\partial(\sigma^2x^2\phi)}{\partial y}\Big|_{x=y} = (r(t) - q(t))y\phi\big|_{x=y} - \frac{\partial(x^2\sigma^2\phi)}{\partial x}\Big|_{x=y} - \frac{\partial(\rho x\sigma g\phi)}{\partial z}\Big|_{x=y}.$$

We also note that we have not assumed any Dirichlet boundary condition for $z = 0$ but instead obtained an oblique derivative boundary condition which is in line with the one presented in [107] and Chapter 3.

4.4 Improved numerical solution of the forward PIDE

In this section, we provide an accurate and stable numerical solution to the forward PIDE for barrier options introduced in Chapter 2. The discretisation is an extension of the idea presented in Chapter 2, where more particularly, we construct a tailored non-uniform spatial mesh and a backward differentiation formula (BDF) on a non-uniform time mesh, which performs better than the usual Crank-Nicolson scheme. We will work under model (4.2.5) and denote by $C(K, B, T)$ the up-and-out call price under the model. By differentiating (4.3.2) with respect to T and integrating by parts in the second line, we can write the Volterra-type PIDE, expressed as an initial boundary value problem,

for any $B \geq S_0$, $0 \leq K \leq B$ and $T \geq 0$,

$$\frac{\partial C(K,B,T)}{\partial T} + \bar{r}^f(T)C(K,B,T) = -\big(\bar{r}^d(T) - \bar{r}^f(T)\big)K\frac{\partial C(K,B,T)}{\partial K} + \frac{1}{2}\sigma^2_{BSV}(K,B,T)K^2\frac{\partial^2 C(K,B,T)}{\partial K^2} \tag{4.4.1}$$
$$- \frac{1}{2}\sigma^2_{BSV}(B,B,T)B^2(B-K)\left.\frac{\partial^3 C(K,B,T)}{\partial K^2 \partial B}\right|_{K=B}$$
$$- \int_{S_0 \vee K}^{B} \frac{1}{2}K^2\frac{\partial^2 C(K,b,T)}{\partial K^2}\frac{\partial \sigma^2_{BSV}(K,b,T)}{\partial b}\,\mathrm{d}b\,,$$

where

$$\begin{aligned} C(K,B,0) = (S_0-K)^+\mathbf{1}_{S_0<B}, \qquad & (K,B) \in \Omega, \\ C(B,B,T) = 0, \qquad & B \geq S_0, T > 0 \\ C(K,S_0,T) = 0, \qquad & K \leq S_0, T > 0. \end{aligned}$$

The equation is degenerate at $K = 0$ and no boundary condition is needed. Moreover, the solution $V(\cdot,B,T)$ is fully determined by $(V(\cdot,b,T)_{b<B}$ and $(V(\cdot,\cdot,t)_{t<T}$, hence (4.4.1) is solved up to the largest barrier level B_{max} needed.

4.4.1 Pricing vanilla options

Prices of vanilla contracts can either be obtained by a two-dimensional backward-PDE (the second dimension resulting from the dependence of the volatility on the running maximum; see Chapter 2) or the forward PIDE (4.4.1) by setting the maximum up-and-out barrier B_{max} very high. The latter normally requires a significant amount of PIDE layers for each barrier level up to B_{max}, which increases the computational time drastically. Therefore, we use the assumption that the volatility becomes flat in the running maximum dimension above a given level. Let B_{last} be this level, then subsequent layers of the PIDE (4.4.1) will not contribute to the integral term since

$$\int_{B_{last}}^{\infty} \frac{1}{2}K^2\frac{\partial^2 C(K,b,T)}{\partial K^2}\frac{\partial \sigma^2_{BSV}(K,b,T)}{\partial b}\,db = 0\,. \tag{4.4.2}$$

In other words, no error is made by jumping from layer $B = B_{last}$ to a large $B = B_{max}$. Indeed, it seems reasonable to assume that

$$\lim_{B\to\infty} -\frac{1}{2}\sigma^2_{BSV}(B,B,T)B^2(B-K)\left.\frac{\partial^3 C(K,B,T)}{\partial K^2 \partial B}\right|_{K=B} = 0\,,$$

since by (4.3.8) the term in the limit is proportional to the joint density function of (S_T, M_T) at

(B, B), and we conjecture here that it goes to 0 faster than B^{-3}. The PIDE (4.4.1) then becomes

$$\begin{aligned}\frac{\partial C(K,\infty,T)}{\partial T} + \bar{r}^f(T)C(K,\infty,T) &= -\Big(\bar{r}^d(T) - \bar{r}^f(T)\Big)K\frac{\partial C(K,\infty,T)}{\partial K} \qquad (4.4.3)\\ &\quad + \frac{1}{2}\sigma^2_{BSV}(K, B_{last}, T)K^2\frac{\partial^2 C(K,\infty,T)}{\partial K^2}\\ &\quad - \int_0^{B_{last}} \frac{1}{2}K^2\frac{\partial^2 C(K,b,T)}{\partial K^2}\frac{\partial \sigma^2_{BSV}(K,b,T)}{\partial b}\,db \qquad (4.4.4)\end{aligned}$$

with boundary conditions

$$\begin{aligned} C(K,\infty,0) &= (S_0 - K)^+, \qquad && T = 0,\\ \frac{\partial^2 C(\infty,\infty,T)}{\partial K^2} &= 0, && K = \infty. \end{aligned}$$

One will then compute all the layers up to B_{last}, and compute one additional "vanilla" layer for an arbitrary level B_{max} which is driven by the PDE (4.4.3).

4.4.2 Space discretisation

We generalise the space discretisation from Chapter 2 to non-uniform grids to increase convergence speed. We define a non-uniform mesh which contains $M+1$ time points, $N+1$ spatial points in the strike and $P+1$ in the barrier variable, leading to the following definition of the step sizes:

$$\begin{aligned} K_i &= \textstyle\sum_{i'=0}^{i-1}\Delta_K(i'), && i \in [\![0, N]\!],\\ B_j &= S_0 + \textstyle\sum_{j'=0}^{j-1}\Delta_B(j'), && j \in [\![0, P]\!],\\ T_m &= \textstyle\sum_{m'=0}^{m-1}\Delta_T(m'), && m \in [\![0, M]\!]. \end{aligned}$$

We denote by N_0 the node such that $K_{N_0} = S_0$. For simplicity, we impose $\Delta_K(i) = \Delta_B(i - N_0)$ for any $i \in [\![N_0, N]\!]$. This will ensure that for any B_j, the corresponding mesh line will contain at least all (B_u, B_j) for all u smaller than j, which is useful for the following algorithm. We denote the discrete solution vector in such a layer by

$$\mathbf{u}^m_{\cdot,j} = \Big[C(K_0, B_j, T_m), \ldots, C(B_j, B_j, T_m)\Big]'$$

of size $n_j = N_0 + j + 1$, where $'$ denotes the transpose, and define $\mathbf{I}_n$ the identity matrix of size $n \times n$.

Derivatives are approximated by centered finite differences at each space point except at $K = 0$ and $K = B_j$, where they are computed, respectively, by three-point forward and backward one-sided differences. For the time being, the boundary conditions are not taken into account. We assume non-equally spaced points and rely on the algorithm in [64] to define two derivative operators $\delta_K \mathbf{u}^m_{i,j}$ and $\delta_{KK}\mathbf{u}^m_{i,j}$ as well as the corresponding matrix derivative operators $\mathbf{D}$ and $\mathbf{D_2}$ respectively. We also

define the forward time difference operator

$$\delta_T \mathbf{u}_{i,j}^m = \frac{\mathbf{u}_{i,j}^{m+1} - \mathbf{u}_{i,j}^m}{\Delta_T(m)}.$$

The integral term is computed using the trapezoidal quadrature rule where we detail how to handle the non-uniform grid. The term of interest is

$$F(K_i, B_j, T_m) \quad = \quad \int_{S_0 \vee K}^{B_j} g(K_i, b, T_m)\, \mathrm{d}b$$

with

$$g(K, b, T) = -\frac{1}{2} K^2 \frac{\partial^2 C(K, b, T)}{\partial K^2} \frac{\partial \sigma_{BSV}^2(K, b, T)}{\partial b}.$$

We define

$$\bar{g}(K_i, B_j, T_m) = \begin{cases} -\frac{1}{2} K_i^2 \delta_{KK} \mathbf{u}_{i,j}^m \frac{\partial \sigma_{BSV}^2(K_i, B_j, T_m)}{\partial B}\,, & K_i < B_j \\ 0\,, & K_i = B_j\,, \end{cases}$$

since (see Chapter 2)

$$\frac{\partial^2 C(B, B, T)}{\partial K^2} = 0\,.$$

Let us assume that we know an approximation to the solution of the PIDE for the discrete set of barriers $(B_{j'})_{j'<j}$. Also, when the barrier is at the spot level, the integrand is zero. Hence, we can write for any $j \geq 1$

$$\begin{aligned} F(K_i, B_j, T_m) \quad &\approx \quad \frac{1}{2} \sum_{j'=1}^{j-1} \Delta_B(j') (\bar{g}(K_i, B_{j'}, T_m) + \bar{g}(K_i, B_{j'-1}, T_m)) \\ &+ \quad \frac{1}{2} \Delta_B(j) (\bar{g}(K_i, B_j, T_m) + \bar{g}(K_i, B_{j-1}, T_m)) \\ &=: \quad f(K_i, B_j, T_m) + \frac{1}{2} \Delta_B(j) \bar{g}(K_i, B_j, T_m)\,, \end{aligned} \tag{4.4.5}$$

and f can be updated recursively from one layer j to the next $j+1$ with

$$f(K_i, B_{j+1}, T_m) = f(K_i, B_j, T_m) + \frac{1}{2} \bar{g}(K_i, B_j, T_m)(\Delta_B(j) + \Delta_B(j+1))\,.$$

The sum in (4.4.5) can be computed for all K_i and T_m in a forward induction over j, as the solution is known for all the barrier levels involved. This sum is then treated as a source term for the PDE layer of level B_j. We define a vector

$$\mathbf{f}_{.,j}^m = [f(K_i, B_j, T_m)]_{i=0,1,\ldots,n_j-1}.$$

The remaining term

$$\frac{1}{2}\Delta_B(j)\bar{g}(K_i, B_j, T_m)$$

gives a small correction to the diffusion at B_j and we can incorporate it in the discretisation of the corresponding diffusive term of (4.4.1); see (4.4.6) below.

To approximate the "boundary derivative" at (B, B, T) in (4.4.1), we recall that (see Chapter 2) for any $T > 0$

$$\left.\frac{\partial^3 C(K,B,T)}{\partial K^2 \partial B}\right|_{K=B} = -\left.\frac{\partial^3 C(K,B,T)}{\partial K^3}\right|_{K=B}.$$

We can compute a second order, five point approximation to the third derivative with the algorithm from [64] and denote the (numerically computed) discretisation matrix by $\mathbf{\Phi}$.

As a requirement for stability, we found empirically that the five points used for the approximation need to be on the right-hand side of S_0. In order to ensure this, we start the PIDE layers with $j = 0$, skip the barrier layers $j = (1, 2, 3, 4)$ and carry on with layer $j = 5$. No error is introduced if $\frac{\partial \sigma_{BSV}}{\partial m}(K, m, T)$ is zero for any $m \in [S_0, B_5]$ (which will be satisfied by our construction of σ_{BSV}), similar to the situation in Section 4.4.1 for large B. In fact, this allows to start the first layer at the first barrier $B_{first} = \inf\{B : \frac{\partial \sigma_{BSV}}{\partial m}(\cdot, B, t) \neq 0\}$, since

$$\int_{S_0 \vee K}^{B_{first}} \frac{1}{2} K^2 \frac{\partial^2 C(K,b,T)}{\partial K^2} \frac{\partial \sigma^2_{BSV}(K,b,T)}{\partial b}\, \mathrm{d}b = 0.$$

If we need to compute the price of a barrier option with a barrier $B < B_{first}$, we use a cubic spline interpolation between $B = S_0$ (where $C(K, S_0, T) = 0$) and $B = B_{first}$. We will refer to the skipped layers as the "blank layers" in the remainder of the section. We recall that in Chapter 2 we were only able to use a first order accurate approximation of the boundary derivative due to stability issues. As this term is present in the discretised equation for all interior mesh points, the scheme we proposed in Chapter 2 had a consistency order reduced to one in Δ_K. We find that using a second-order BDF time scheme as described below in Section 4.4.3 instead of Crank-Nicolson allows to use a second order accurate approximation and conserve stability. Overall, we obtain an order two consistent spatial approximation for smooth enough meshes.

Finally, the surface of barrier option prices can be retrieved by cubic spline interpolation in both strikes and barriers with flat extrapolation for large barrier levels. It is a consistent assumption since, for any $B > B_{max}$, the contract should be a vanilla option.

4.4.3 BDF2 scheme with variable step size

The main difficulty in the time discretisation of the forward PIDE (4.4.1) arises from the term

$$\frac{1}{2}\sigma_{BSV}^2(B,B,T)B^2(B-K)\left.\frac{\partial^3 C(K,B,T)}{\partial K^2 \partial B}\right|_{K=B},$$

which, as per (4.3.8), contains the joint density ϕ of the process (S_T, M_T) and becomes a Dirac delta point source at (S_0, S_0) for $T = 0$. This is an issue for all B close to S_0 for short-term options.

In order to tackle the problem, we first subdivide the initial time step and perform 4 backward Euler steps of a quarter step-size, and make use of the "blank layers" as described in Section 4.4.2. For the BDF2 scheme, a single initial backward Euler step would suffice, but for better comparison with the Crank–Nicolson scheme we adopt the Rannacher startup with four steps (see [70]) in both cases.

If we take into account the finite difference approximations and quadrature rule for the integral, it is now possible to give a discretised PIDE, for a given triplet $(i,j,m) \in [\![0,N]\!] \times [\![0,M]\!] \times [\![0,P]\!]$ by

$$\begin{aligned}\delta_T \mathbf{u}_{i,j}^m + \bar{r}^f(T_m)K_i\mathbf{u}_{i,j}^m + \Big(\bar{r}^d(T_m) - \bar{r}^f(T_m)\Big)K_i\delta_K\mathbf{u}_{i,j}^m \\ -\frac{1}{2}\left(\sigma_{BSV}^2(K_i,B_j,T_m) - \frac{1}{2}\frac{\partial\sigma_{BSV}^2}{\partial B}(K_i,B_j,T_m)\Delta_B\right)K_i^2\delta_{KK}\mathbf{u}_{i,j}^m \\ -\frac{1}{2}\sigma_{BSV}^2(B_j,B_j,T_m)B_j^2(B_j-K_i)^+\delta_{KKK}^-\mathbf{u}_{n_j,j}^m = \mathbf{f}_{\cdot,j}^m,\end{aligned} \tag{4.4.6}$$

where only the right-hand side differs from Chapter 2, and specify the coefficient matrices

$$\begin{aligned}\mathbf{A}_{\cdot,j}^m &= (r(T_m)-q(T_m))\mathrm{diag}\Big(K_0,...,K_{n_j}\Big),\\ \mathbf{B}_{\cdot,j}^m &= -\frac{1}{2}\mathrm{diag}\left(\left(\sigma_{BSV}^2(K_i,B_j,T_m)K_i^2 - \frac{1}{2}K_i^2\frac{\partial\sigma_{BSV}^2(K_i,B_j,T_m)}{\partial B}\Delta_B(j)\right)_{i\in[\![0,n_j-1]\!]}\right),\\ \mathbf{C}_{\cdot,j}^m &= -\frac{1}{2}\mathrm{diag}\Big(\Big(\sigma_{BSV}^2(B_j,B_j,T_m)B_j^2(B_j-K_i)^+\Big)_{i\in[\![0,n_j-1]\!]}\Big).\end{aligned}$$

Under backward Euler time stepping, the complete scheme can be more compactly written as

$$\begin{aligned}\frac{\mathbf{u}_{\cdot,j}^m - \mathbf{u}_{\cdot,j}^{m-1}}{\Delta T} + \mathbf{L}_{\cdot,j}^m\mathbf{u}_{\cdot,j}^m &= \mathbf{f}_{\cdot,j}^m,\\ \mathbf{L}_{\cdot,j}^m &= \mathbf{A}_{\cdot,j}^m\mathbf{D} + \mathbf{B}_{\cdot,j}^m\mathbf{D_2} + \mathbf{C}_{\cdot,j}^m\mathbf{\Phi}.\end{aligned} \tag{4.4.7}$$

To define the BDF scheme for variable step size, we denote $C(K,B,T_m)$ by C_m and write Newton's interpolation polynomial in time as

$$C(T) = C_m + [C_m, C_{m-1}](T-T_m) + [C_m, C_{m-1}, C_{m-2}](T-T_m)(T-T_{m-1}),$$

where $[T_1, T_2, ..., T_M]$ is the divided difference operator on the function C for the time points $(T_m)_{m\leq M}$. Taking the derivative and writing the polynomial coefficient in full,

$$\Delta_T(m)q_m = \Delta_T(m)\mathcal{L}_m C_m + C_m - C_{m-1} + \frac{(\Delta_T(m))^2}{\Delta_T(m)+\Delta_T(m-1)}\left[\frac{C_m - C_{m-1}}{\Delta_T(m)} - \frac{C_{m-1}-C_{m-2}}{\Delta_T(m-1)}\right],$$

where $\mathcal{L}_m$ is the linear operator at time T_m corresponding to the matrix $\mathbf{L}^m_{\cdot,j}$, and q_m the right-hand-side function at time T_m corresponding to the vector $\mathbf{f}^m_{\cdot,j}$. More compactly and as a linear system of equations,

$$\left(\mathbf{I}_{n_j} + \frac{\Delta_T(m)}{1+\gamma_m}\mathbf{L}^m_{\cdot,j}\right)\mathbf{u}^m_{\cdot,j} = \left(1 + \frac{\Delta_T(m)}{\Delta_T(m-1)}\frac{\gamma_m}{1+\gamma_m}\right)\mathbf{u}^{m-1}_{\cdot,j} - \left(\frac{\Delta_T(m)}{\Delta_T(m-1)}\frac{\gamma_m}{1+\gamma_m}\right)\mathbf{u}^{m-2}_{\cdot,j} + \frac{\Delta_T(m)}{1+\gamma_m}\mathbf{f}^m_{\cdot,j},$$

with

$$\gamma_m = \frac{\Delta_T(m)}{\Delta_T(m)+\Delta_T(m-1)},$$

which gives an implicit second-order linear multi-step method. Assuming there exists a smooth bijective mapping between the non-uniform and a uniform time mesh, the method is consistent of order 2 and stability is preserved if the step-size ratio is bounded as follows (see [85])

$$0 < \frac{\Delta_T(m)}{\Delta_T(m-1)} < 1+\sqrt{2}\,,$$

which is guaranteed by a smooth change of the step-size.

4.4.4 Non-uniform mesh construction and numerical tests

In order to get the best accuracy, we refine the grid around $(K, B) = (S_0, S_0)$ for two reasons. First, because this will add more barrier levels where $\frac{\partial C}{\partial B}$ is high in order to efficiently capture the change in call price as well as reduce an eventual error generated by the "blank layers". Second, on the strike scale, the inflection point will be around $K = S_0$ and higher accuracy becomes important in that specific zone.

A hyperbolic mesh as in Chapter 3 is chosen with $\eta = 0.05$ chosen according to our numerical experiments, where we require S_0 to be a node. In Figure 4.4.1, we plot the generated grid with $N = 50$ in strikes and barrier levels, including the initial "blank layers".

In order to perform our numerical tests we calibrate a pure local volatility model to the set of vanilla options presented in Section 4.2.2 and obtain a local volatility function σ_{LV}. The calibration routine follows a fixed-point algorithm as described in Chapter 3 based on the work of [141]. From this calibrated local volatility we define a hypothetical Brunick–Shreve volatility of the form

$$\sigma_{BSV}(s,m,t) = \sqrt{\sigma_{LV}(s,t)\sigma_{LV}(m,t)}\,, m \geq s\,,$$

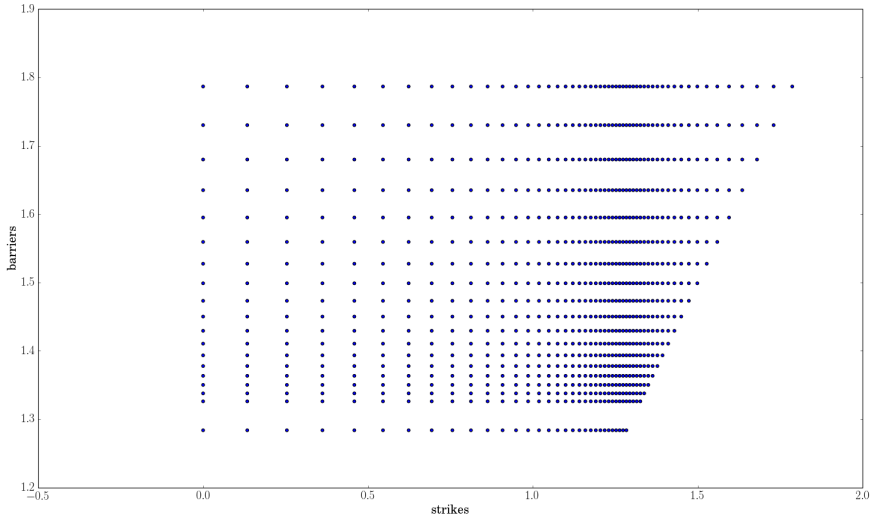


Figure 4.4.1: Finite difference hyperbolic grid for the forward PIDE (4.4.1) initialised with "blank layers".

defined on a grid of strikes, barriers and maturities $\left(K_{T_i,j}, B_{T_{i,},k}, T_i\right)$, with $i \in [\![1, 10]\!]$, $j \in [\![1, 5]\!]$ and $k \in [\![1, 5]\!]$ such that $B_{T_{i,},k} = K_{T_i,k}$ and interpolated with cubic splines in space and backward-flat in time. Since an integration by parts is used to derive the forward PIDE in Chapter 2, we need to ensure continuous differentiability of σ_{BSV} in m. We make σ_{BSV} continuously differentiable in s as well. An issue arises with the flat extrapolation on the boundary where the derivative jumps to zero. In order to tackle this problem, we propose a smooth transformation by a change of coordinates $x \to u_{.}(x)$. Let a rectangle be $[x_{min}, x_{max}] \times [y_{min}, y_{max}]$ outside which we need to extrapolate the function σ_{BSV} of the variable (x, y). This change of variable is based on the fact that a flat extrapolation can be performed by setting

$$\sigma_{BSV}(x, y, t) = \sigma_{BSV}(\bar{u}_x(x), \bar{u}_y(y), t)$$

with

$$\begin{aligned} \bar{u}_x(x) &= x_{max}\mathbf{1}_{x \geq x_{max}} + \mathbf{1}_{x < x_{max}}[x_{min}\mathbf{1}_{x \leq x_{min}} + x\mathbf{1}_{x > x_{min}}] \\ \bar{u}_y(y) &= y_{max}\mathbf{1}_{x \geq y_{max}} + \mathbf{1}_{y < y_{max}}[y_{min}\mathbf{1}_{y \leq y_{min}} + y\mathbf{1}_{y > y_{min}}]. \end{aligned}$$

The first step is to extrapolate the 2-dimensional function linearly in space. We define $\bar{\sigma}_{BSV}$ the

linearly extrapolated volatility as

$$\begin{aligned}\bar{\sigma}_{BSV}(x,y,t) &= \sigma_{BSV}(\bar{u}_x(x),\bar{u}_y(y),t)\\ &+ \mathbf{1}_{x>x_{max}}\frac{\partial\sigma_{BSV}(x_{max},\bar{u}_y(y),t)}{\partial x}(x-x_{max})\\ &+ \mathbf{1}_{x<x_{min}}\frac{\partial\sigma_{BSV}(x_{min},\bar{u}_y(y),t)}{\partial x}(x-x_{min})\\ &+ \mathbf{1}_{y>y_{max}}\frac{\partial\sigma_{BSV}(\bar{u}_x(x),y_{max},t)}{\partial y}(y-y_{max})\\ &+ \mathbf{1}_{y<y_{min}}\frac{\partial\sigma_{BSV}(\bar{u}_x(x),y_{min},t)}{\partial y}(y-y_{min})\,.\end{aligned}$$

We then define a smoothed transition of the coordinate x at both x_{max} and x_{min}, such that u_x is as follows

$$u_x(x) = x_{max}w(x,x_{max},\eta_0) + (1-w(x,x_{max},\eta_0))[x_{min}(1-w(x,x_{min},-\eta_0)) + xw(x,x_{min},-\eta_0)] \tag{4.4.8}$$

with

$$\begin{cases} w(x,x_0,\eta_0) &= \frac{1}{2}(1+\tanh\Big(\frac{2x_0}{\epsilon}\frac{(x-\bar{x}(\eta_0))}{\bar{x}(\eta_0)}\Big)\\ \bar{x}(\eta_0) &= \frac{2x_0^2}{2x_0+\epsilon\,\mathrm{arctanh}\big(\eta_0(1-\frac{\epsilon}{2})\big)}\\ \epsilon &= \frac{S_0}{10}\,.\end{cases}$$

In the same way we can define a smooth transition of the coordinate y at y_{max} and y_{min} that we denote u_y. The idea of the smoothing is to be able to control the impact of the transition on the inside of the domain $[x_{min},x_{max}]$ by means of the parameter η_0. If $\eta_0=1$, most of the transition happens outside of the domain which allows to match the values of the flat extrapolated function inside the domain. In contrast, if $\eta_0=-1$, most of the transition will happen inside the domain. This behaviour can seem attractive at first, however, it will give rise to issues if the function needs to take a specific shape inside the domain, a local volatility for instance. In both cases, we have $u_x(x_{max})\approx x_{max}$ and $u_x(x_{min})\approx x_{min}$. Finally, we reach a trade-off when $\eta=0$. This is the value we pick in our implementation. The new volatility is then

$$\sigma_{BSV}(x,y,t) = \bar{\sigma}_{BSV}(u_x(x),u_y(y),t)\,.$$

We plot in Figure 4.4.6 the assumed Brunick–Shreve volatility with smooth transition on the boundary.

In order to demonstrate the importance of "blank layers", we plot the value of

$$D_T^d\phi(B,B,T) = \left.\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B}\right|_{K=B},$$

with $N=700$ and $M=100$, as a function of the barrier level for $T=1$ with and without "blank

layers" in Figure 4.4.2 and 4.4.3 respectively. Evaluating the term $\phi(B, B, T)$ accurately is necessary since the value of the foreign no touch option is directly linked to it by

$$\frac{\partial C(0,B,T)}{\partial T} + \bar{r}^f(T)C(0,B,T) = -\frac{1}{2}\sigma_{BSV}^2(B,B,T)B^3 \frac{\partial^3 C(K,B,T)}{\partial K^2 \partial B}\bigg|_{K=B} \qquad \forall (B,T) \in]S_0, +\infty[\times\mathbb{R}_*^+.$$

Finally, to highlight the importance of a smoothing time stepping scheme, we plot $\phi(B, B, T)$ with "blank layers" combined with BDF2 in Figure 4.4.4 and Crank–Nicolson in Figure 4.4.5, which shows that Crank–Nicolson can generate instabilities if the number of time steps is too small, even when using Rannacher initialisation.

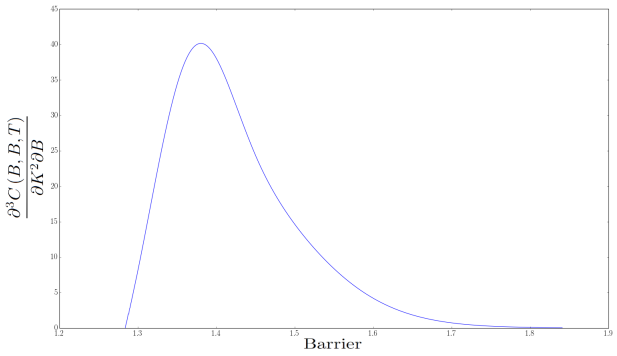


Figure 4.4.2: $\phi(B, B, T)$ as a function of the barrier level for $T = 1$ initialised with "blank layers". 700 strike steps, 100 BDF2 time steps + Rannacher initialisation.

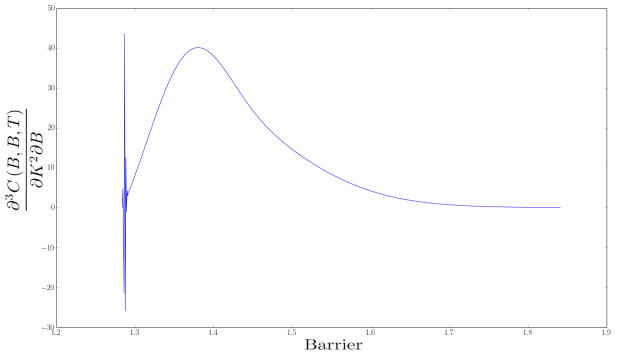


Figure 4.4.3: $\phi(B, B, T)$ as a function of the barrier level for $T = 1$ initialised without "blank layers". 700 strike steps, 100 BDF2 time steps + Rannacher initialisation.

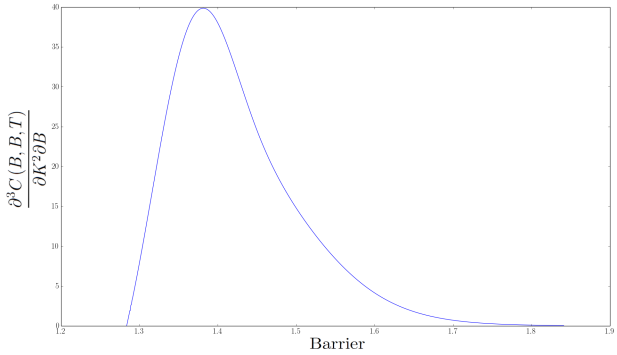


Figure 4.4.4: $\phi(B, B, T)$ as a function of the barrier level for $T = 1$ initialised with "blank layers". 700 strike steps, 10 BDF2 time steps + Rannacher initialisation.

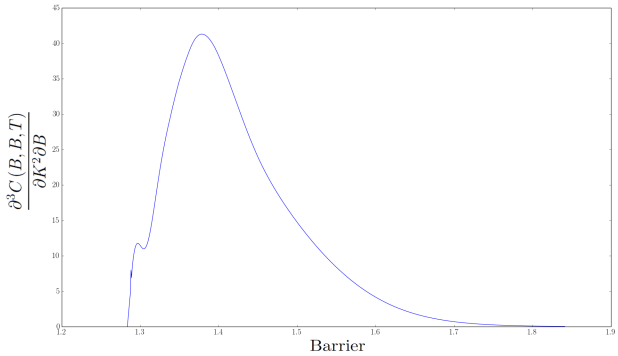


Figure 4.4.5: $\phi(B, B, T)$ as a function of the barrier level for $T = 1$ initialised with "blank layers". 700 strike steps, 10 Crank–Nicolson time steps + Rannacher initialisation.

In order to numerically verify the PIDE price, we compute the price of an up-and-out call option for $K = 80\% \times S_0$, $B = 110\% \times S_0$ and $T = 1$. Prices are computed with both the forward PIDE (700 spot steps and 100 times steps) and crude Monte Carlo combined with the Brownian bridge technique as in Chapter 6 of [73] (with 5×10^7 paths and 500 time steps). Results are presented in Table 4.4.1.

Forward PIDE	Monte Carlo (with 95% conf. int.)
0.15823	0.15825 (0.15821, 0.15828)

Table 4.4.1: Price of an up-and-out call option for $K = 80\% \times S_0$, $B = 110\% \times S_0$ and $T = 1$.

In Table 4.4.2, we plot the convergence order as a function of the number of strike points. More precisely, the error in row n is the absolute difference between the net-present-value with $2 \times N_K$ and N_K strike points. The order displayed in row n is $\ln(e_n/e_{n+1})/\ln 2$, where e_n is the error in row n. We notice that while the convergence order is close to 3 for smaller number of strike points, the asymptotic behaviour is indeed of order 2. A similar approach is used for Table 4.4.3 where the convergence of order 2 is obtained after 240 time steps per year. However, even for smaller number of time steps, the error is small and the price is accurate.

N_K	Error	Order
300	9.43×10^{-6}	2.93
600	1.38×10^{-6}	2.03
1200	3.23×10^{-7}	1.98
2400	6.53×10^{-8}	–

Table 4.4.2: Convergence order in number of strike steps N_K for 60 time steps

N_T	Error	Order
60	7.58×10^{-6}	1.12
120	3.48×10^{-6}	-0.99
240	6.91×10^{-6}	2.07
480	1.64×10^{-6}	2.04
960	3.99×10^{-7}	–

Table 4.4.3: Convergence order in number of time steps N_T for 1200 strike steps

Finally, we show in Figure 4.4.7 up-and-out call prices for $T = 1$ as a function of strike and barrier level.

4.5 Calibration of the Brunick–Shreve model

In this section, we discuss the simultaneous calibration of the Brunick–Shreve model by forward PIDE to vanilla and no-touch options. The model represents the natural extension of the Dupire local volatility model to handle up-and-out barriers.

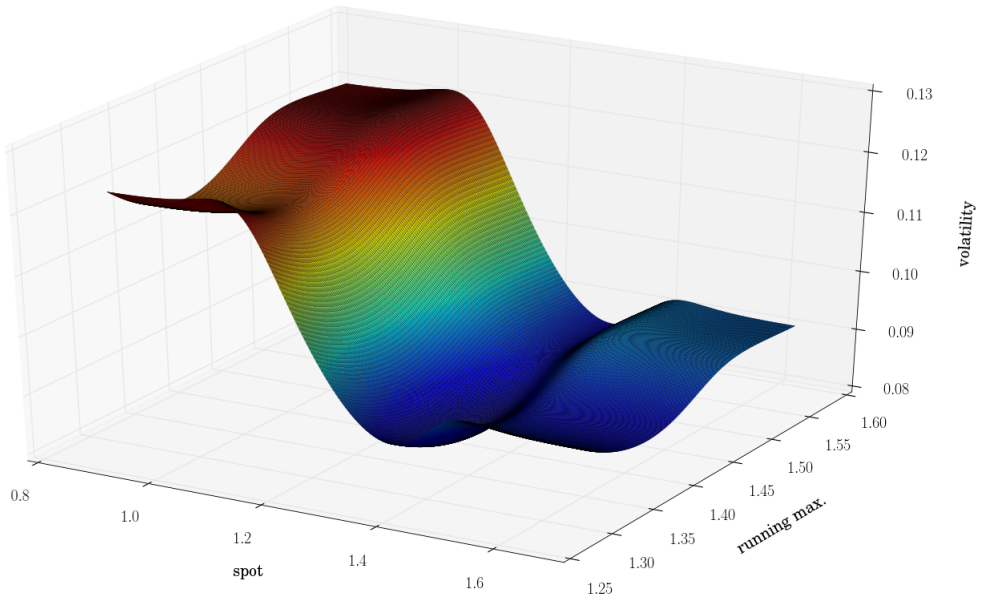


Figure 4.4.6: Brunick–Shreve volatility used for numerical test with smooth transition at the boundary of the domain. The construction is such that $\sigma_{BSV}(s,m,t) = \sqrt{\sigma_{LV}(s,t)\sigma_{LV}(m,t)}$.

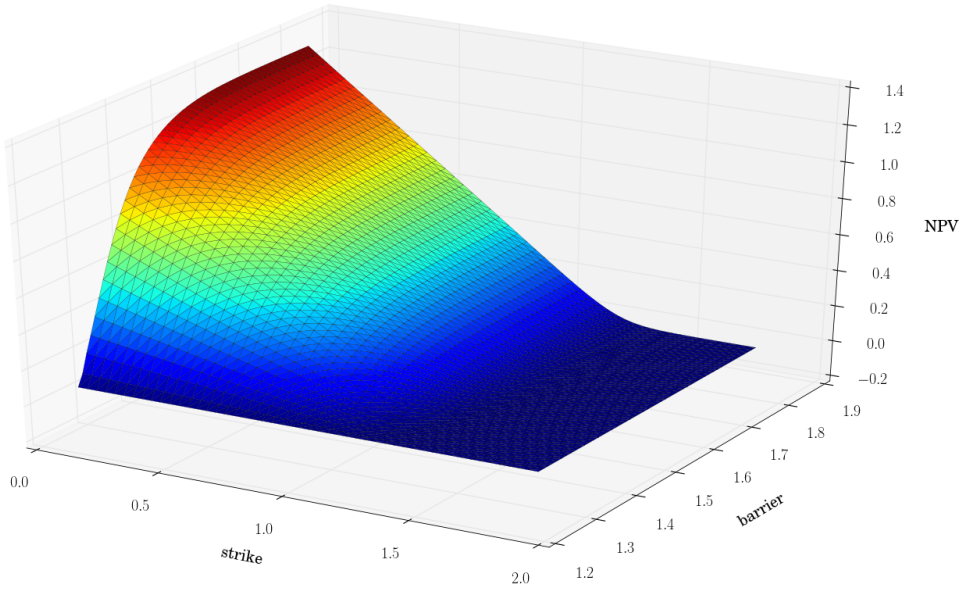


Figure 4.4.7: Up-and-out call prices computed with the forward PIDE for different values of strikes and barrier levels and $T = 1$.

4.5.1 Gradient operator with respect to the volatility parameters

In order to perform a best-fit algorithm, knowing the gradient with respect to the model parameters is of paramount importance for the gradient-based optimisation process. Assume that the Brunick–Shreve volatility is a function of N parameters $\Lambda_i = (\sigma_{i,1}, ..., \sigma_{i,N})$ that depends only on the considered quoted maturities $(T_i)_{i \leq N_{Mat}}$ but is flat in between quoted pillars. So, we need to compute

$$\nabla C(\sigma_{i,1}, ..., \sigma_{i,N})\,.$$

The above equation (4.4.1) can be written as

$$\frac{\partial C(\Lambda_i)}{\partial T} + \bar{r}^f(T)C(\Lambda_i) + \big(\bar{r}^d(T) - \bar{r}^f(T)\big)K\frac{\partial C(\Lambda_i)}{\partial K} - \frac{1}{2}\sigma^2_{BSV}(\Lambda_i)K^2\frac{\partial^2 C(\Lambda_i)}{\partial K^2} = -\frac{1}{2}\sigma^2_{BSV}(\Lambda_i)\big\rfloor_{K=B}B^2(B-K)\,\frac{\partial^3 C(\Lambda_i)}{\partial K^2 \partial B}\bigg\rfloor_{K=B} - \int_{S_0 \vee K}^{B} K^2\frac{\partial^2 C(\Lambda_i)}{\partial K^2}\sigma_{BSV}(\Lambda_i)\frac{\partial \sigma_{BSV}(\Lambda_i)}{\partial b}\,\mathrm{d}b\,.$$

We can differentiate with respect to each of the parameters $(\sigma_i)_{i \in [\![1,N]\!]}$, which gives

$$\begin{aligned}\frac{\partial \nabla C(\Lambda_i)}{\partial T} + \bar{r}^f(T)\nabla C(\Lambda_i) &= -\big(\bar{r}^d(T) - \bar{r}^f(T)\big)K\frac{\partial \nabla C(\Lambda_i)}{\partial K} + \frac{1}{2}\sigma^2_{BSV}(\Lambda_i)K^2\frac{\partial^2 \nabla C(\Lambda_i)}{\partial K^2} \\ &- \frac{1}{2}\sigma^2_{BSV}(\Lambda_i)\big\rfloor_{K=B}B^2(B-K)\,\frac{\partial^3 \nabla C(\Lambda_i)}{\partial K^2 \partial B}\bigg\rfloor_{K=B} \\ &- \int_{S_0 \vee K}^{B} K^2\frac{\partial^2 \nabla C(\Lambda_i)}{\partial K^2}\sigma_{BSV}(\Lambda_i)\frac{\partial \sigma_{BSV}(\Lambda_i)}{\partial b}\,\mathrm{d}b + R(\Lambda_i)\,,\end{aligned}$$

with

$$\begin{aligned}R(\Lambda_i) &= \sigma_{BSV}(\Lambda_i)(\nabla\sigma_{BSV}(\Lambda_i))K^2\frac{\partial^2 C(\Lambda_i)}{\partial K^2} - (\sigma_{BSV}(\Lambda_i)(\nabla\sigma_{BSV}(\Lambda_i)))\big\rfloor_{K=B}B^2(B-K)\,\frac{\partial^3 C(\Lambda_i)}{\partial K^2 \partial B}\bigg\rfloor_{K=B} \\ &- \int_{S_0 \vee K}^{B} K^2\frac{\partial^2 C(\Lambda_i)}{\partial K^2}\left(\nabla(\sigma_{BSV}(\Lambda_i))\frac{\partial \sigma_{BSV}(\Lambda_i)}{\partial b} + \sigma_{BSV}(\Lambda_i)\nabla\left(\frac{\partial \sigma_{BSV}(\Lambda_i)}{\partial b}\right)\right)\mathrm{d}b\,.\end{aligned}$$

Hence, $\nabla C(\Lambda_i)$ follows the same PDE as C, but with an inhomogeneous term which is a function of C and its spatial derivatives, with initial condition

$$\nabla C(\Lambda_i)(K, B, 0) = 0 \qquad 0 \leq K, \quad K \vee S_0 \leq B,$$

and with boundary conditions

$$\begin{aligned}\nabla C(\Lambda_i)(B, B, T) = 0, &\qquad S_0 \leq B, \\ \nabla C(\Lambda_i)(K, S_0, T) = 0, &\qquad K \leq S_0,\end{aligned}$$

which match the Dirichlet boundary conditions for $C(K, B, T)$. This useful property confirms that we can use the same discretised linear operator for both C and $\nabla C(\Lambda_i)$. The additional source term R on the right-hand side is fully known since the solution for C is computed beforehand. Only one

costly LU factorisation is needed to compute both C and $\nabla C(\Lambda_i)$ at each implicit time step. Solving the $N+1$ linear systems will be fast by forward and backward substitution. Additionally, since the volatility is assumed backward-flat in maturity, the set of parameters Λ_{i+1} has no impact on the values of $C(K,B,T)$ for any $T \leq T_i$. Hence we also have

$$\nabla C(\Lambda_i)(K,B,T_{i-1}) = 0,\ i \geq 1\,.$$

Some numerical experiments for a volatility defined on a grid of 5×5 points, i.e. 25 parameters, showed that the additional evaluation of the gradient (with respect to each of the parameters) requires only twice the time needed to solve the PDE. As a comparison, the powerful adjoint algorithmic differentiation (AAD) technique (see [71] for applications in derivative pricing) can achieve the same task for a computational time between three to four times the time needed to solve the original PDE independent of the number of parameters (see Section 4.6 in [75]). Therefore, the chosen approach leads to a competitive computational time in the present setting. Additionally, the gradient components can be computed in parallel which would further reduce the computational cost.

4.5.2 Calibration with regularised best-fit algorithm

The Brunick–Shreve volatility function is defined on a grid of points interpolated with quadratic splines in spot and running maximum and backwards flat in time. In the market setting, and for each quoted maturity, we choose to define the volatility surface on a grid of points formed by the $N_K = 5$ quoted strikes in the spot direction and $N_B = 4$ levels $(M_{T_i,k})_{k\in[\![1,N_B]\!]}$ in the running maximum direction uniformly spaced on the interval

$$\left[S_0 + \frac{(B_{T_i,50\%} - S_0)}{4},\ B_{T_i,90\%}\right],$$

where $B_{T_i,50\%}$ and $B_{T_i,90\%}$ are, respectively, the quoted up-and-out barriers for the corresponding 50% and 90% no-touch probabilities for maturity T_i. Hence, for each maturity slice the volatility surface grid is parametrised by $N_K \times N_B$ points interpolated by quadratic splines. The optimisation will be performed over these $N_K \times N_B$ parameters that we write as a $(N_K \times N_B)$ matrix

$$\Lambda_i = \left[\sigma_{\text{BSV}}^{i,j,k}\right]_{j,k},$$

with $\sigma_{\text{BSV}}^{i,j,k} = \sigma_{\text{BSV}}(K_{T_i,j}, M_{T_i,k}, T_i)$, with $i \in [\![1, N_{Mat}]\!]$, $j \in [\![1, N_K]\!]$, $k \in [\![1, N_B]\!]$. For a given maturity T_i, the volatility is defined on the interval $[K_{T_i,1}, K_{T_i,N_K}] \times [M_{T_i,1}, K_{T_i,N_B}]$ and is extrapolated smooth-flat outside those bounds.

We define the objective function $\bar{e}$ for Q_K quoted strikes, Q_B quoted barrier levels and each maturity T_i as

$$\begin{aligned}
\bar{e}(\Lambda_i) &= e(\Lambda_i)(1+\mathcal{P}(\Lambda_i)) \\
e(\Lambda_i) &= \sum_{l=1}^{Q_B} e_{FNT}(\Lambda_i)^2 + \eta^2 \sum_{l=1}^{Q_K} e_\Sigma(\Lambda_i)^2 \qquad (4.5.1)\\
e_{FNT}(\Lambda_i) &= FNT^{\text{Model}}(B_{T_i,l}, T_i, \Lambda_i) - FNT^{\text{Market}}(B_{T_i,l}, T_i) \qquad (4.5.2)\\
e_\Sigma(\Lambda_i) &= \Sigma^{\text{Model}}(K_{T_i,l}, T_i, \Lambda_i) - \Sigma^{\text{Market}}(K_{T_i,l}, T_i)\,, \qquad (4.5.3)
\end{aligned}$$

with Σ^{Market} the market Black–Scholes implied volatility, Σ^{Model} the model implied volatility, $\eta = 5$ and

$$\mathcal{P}(\Lambda_i) = \frac{1}{N_K N_B} \sum_{l=1}^{N_K} \sum_{m=1}^{N_B} \left(\frac{\partial^2 \sigma_{BSV}(K_{T_i,l}, M_{T_i,m}, T_i)}{\partial K^2} \right)^2 h\left(-\frac{\partial^2 \sigma_{BSV}(K_{T_i,l}, M_{T_i,m}, T_i)}{\partial K^2}, 0.5, 0.5 \right) \qquad (4.5.4)$$

$$h(x, x_0, \epsilon) = \frac{1 + \tanh(2\frac{(x-x_0)}{\epsilon})}{2}\,,$$

where the second derivative is obtained by differentiation of the interpolant. The penalisation function $\mathcal{P}$ is a Tikhonov-type regularisation that reduces the number of local minima and improves the stability of the volatility surface. Tikhonov regularisation has been shown to provide well-posedness, under certain conditions, of the vanilla calibration problem for a local volatility model in [42] and [56]. The parametric form (4.5.4) for the penalisation was chosen empirically. We acknowledge that the penalised error $\bar{e}(\Lambda_i)$ as defined in (4.5.1) does not prevent over-parametrisation if the market data are fit perfectly, i.e. when $e(\Lambda_i) = 0$. However, during the iterative optimisation we found $e(\Lambda_i)$ to be always strictly positive and the penalisation will favour smoother, convex shapes of the volatility in the strike direction; see also Section 4.5.3. Here, h acts as a smoothed step function to ensure differentiability with respect to the Λ_i parameters (for the BFGS routine).

The calibration algorithm uses the bounded L-BFGS routine [156] where the gradient of the objective function $\bar{e}$ needs to be computed. For instance, for parameter $\sigma_{\text{BSV}}^{i,j,k}$ we can write

$$\begin{aligned}
\frac{\partial e(\Lambda_i)}{\partial \sigma_{\text{BSV}}^{i,j,k}} &= 2 \sum_{l=1}^{Q_B} e_{FNT}(\Lambda_i) \frac{\partial FNT^{\text{Model}}(T_i, B_{T_i,l}, \Lambda_i)}{\partial \sigma_{BSV}^{i,j,k}} \\
&+ 2\eta^2 \sum_{l=1}^{Q_K} \frac{e_\Sigma(\Lambda_i)}{\mathcal{V}\left(T_i, K_{T_i,l}, \sigma_{BSV}^{i,j,k}\right)} \frac{\partial Call^{\text{Model}}(T_i, K_{T_i,l}, \Lambda_i)}{\partial \sigma_{BSV}^{i,j,k}},
\end{aligned}$$

where $\mathcal{V}(T, K, \sigma)$ is the standard Black–Scholes vega for maturity T, strike K and volatility σ. The calibration algorithm is described in Algorithm 4.1 for N_{Mat} maturity pillars where we use the calibrated local volatility as a first guess for the first pillar.

Algorithm 4.1 Calibration of the Brunick–Shreve model

$\sigma_{\mathrm{BSV}}(\cdot, \circ, T_1) = \sigma_{\mathrm{LV}}(\cdot, T_1)$
for $(\ i = 1\,;\, i \leq N_{Mat}\,;\, i++)$ **do**
  while $\bar{e} > 10^{-10}$ and $\|\nabla \bar{e}\| > 10^{-8}$ **do**
    solve forward PIDE (4.4.1) on $[T_{i-1}, T_i]$
    compute model implied vol Σ^{Model} for T_i from computed up-and-out call prices $C(K, \infty, T_i)$

    compute model foreign no-touch price FNT^{Model} for maturity T_i from the computed up-and-out call prices $C(0, B, T_i)/S_0$
    compute the objective function

$$\bar{e}(\Lambda_i) = e(\Lambda_i)(1 + \mathcal{P}(\Lambda_i))$$

    as in (4.5.1)
    compute the objective function gradient $\nabla \bar{e}(\Lambda_i)$ using the forward PIDE solution for $\nabla C(\Lambda_i)$

    update volatility surface points Λ_i with the L-BFGS-B algorithm
  end while
  set $\sigma_{\mathrm{BSV}}(\cdot, \circ, T_{i+1}) = \sigma_{BSV}(\cdot, \circ, T_i)$
end for

The model is calibrated using the PIDE (4.4.1) discretised with 1200 strike steps and 40 time steps in between each quoted maturity T_i. The calibrated Brunick–Shreve volatility is plotted in Figures 4.5.1 and 4.5.2.

We will see in Section 4.8, specifically the first column of Table 4.8.2, that the calibration is highly precise, with an absolute error for no-touches of never higher than 0.03%.

4.5.3 The importance of regularisation

In order to emphasise the importance of the penalisation function, we calibrate the Brunick–Shreve model without regularisation, i.e. $\mathcal{P} \equiv 0$, and plot the resulting Brunick–Shreve volatility in Figures 4.5.3 and 4.5.4, with the same axis range as for the regularised solution in Figures 4.5.1 and 4.5.2, which highlights the strong variation especially for longer maturities.

The calibration process shows the existence of a few local minima in the objective function. This is not surprising as the Brunick–Shreve model is, in principle, able to calibrate perfectly a discrete set of up-and-out call options (which includes vanilla options), hence by only providing call and no-touch prices, there is expected to exist more than one solution to the calibration problem. Additionally, we note that even if a set of market up-and-out call options were available, uniqueness of the Brunick–Shreve volatility is still not guaranteed due to the discreteness of the data.

In the present setting, with five vanilla and five no-touch quotes per maturity, and 20 parameters, we find many surfaces which fit the data. However, starting the iterative optimisation procedure from

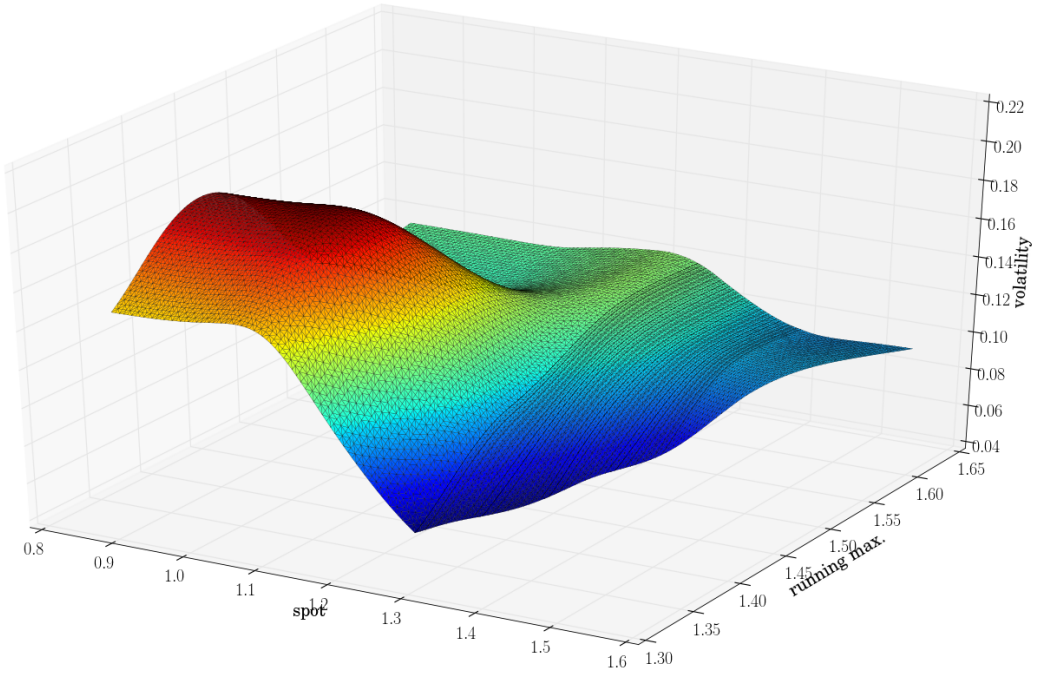


Figure 4.5.1: Brunick–Shreve volatility function $T = 1Y$ with regularisation.

the Dupire volatility (i.e., no dependence on the maximum), the regularisation steers the approximate minimiser towards a calibrated surface with small penalty term.

4.5.4 The underlying local volatility

A sanity check for the calibration, other than the mean square error, is to compute the underlying local volatility implied by the Brunick–Shreve model. Indeed, since the calibration basket was composed of both vanilla and no-touch prices, the Markovian projection of the volatility onto S_t is a calibrated local volatility function. The latter can be compared to the one calibrated on vanilla options only. The Brunick–Shreve implied local volatility is retrieved by Markovian projection as follows

$$\begin{aligned} \sigma_{BSV\,|\,S}(K,T) &= \sqrt{\mathbb{E}^{\mathbb{Q}^d}\left[\sigma^2_{BSV}(K,M_T,T)\,|\,S_T=K\right]} \\ &= \sqrt{\frac{\int_{S_0\vee K}^{\infty}\sigma^2_{BSV}(K,y,T)\phi_{SM}(K,y,T)\,dy}{\int_{S_0\vee K}^{\infty}\phi_{SM}(K,y,T)\,dy}}, \end{aligned}$$

where ϕ_{SM} is the joint density of the spot and running maximum as presented in Corollary 4.6, but here without stochastic volatility. We solve for ϕ_{SM} by a finite element method [105] for the forward Kolmogorov equation (4.3.26), with no stochastic volatility, and compute, at each time point, the projection $\sigma_{BSV\,|\,S}$ for a set of strike points and interpolate in between. The mesh is hand-made and

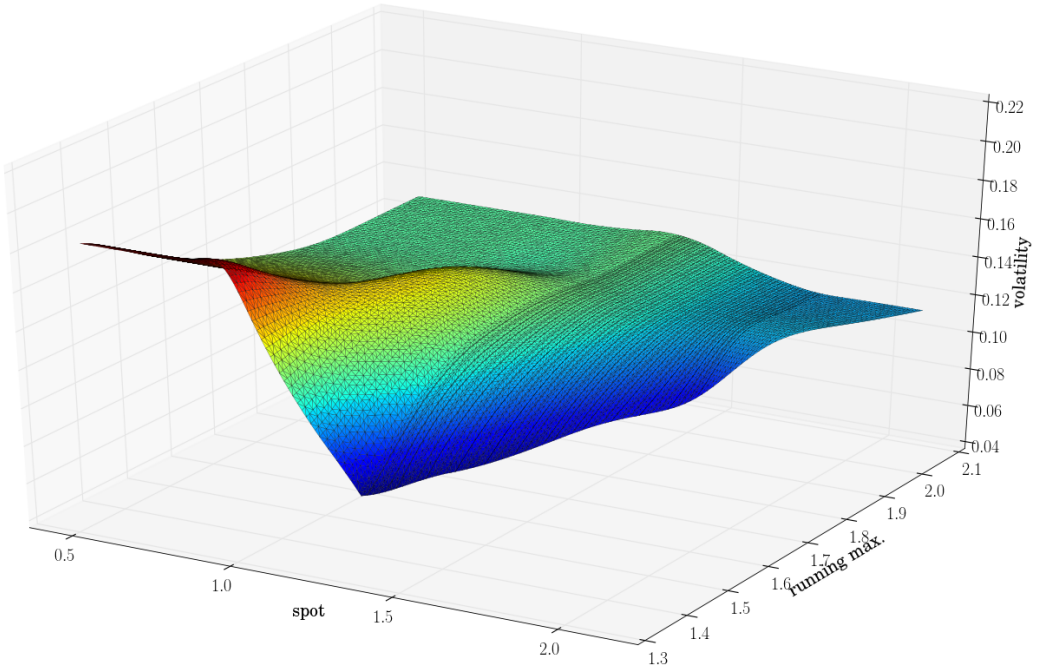


Figure 4.5.2: Brunick–Shreve volatility function $T = 5Y$ with regularisation.

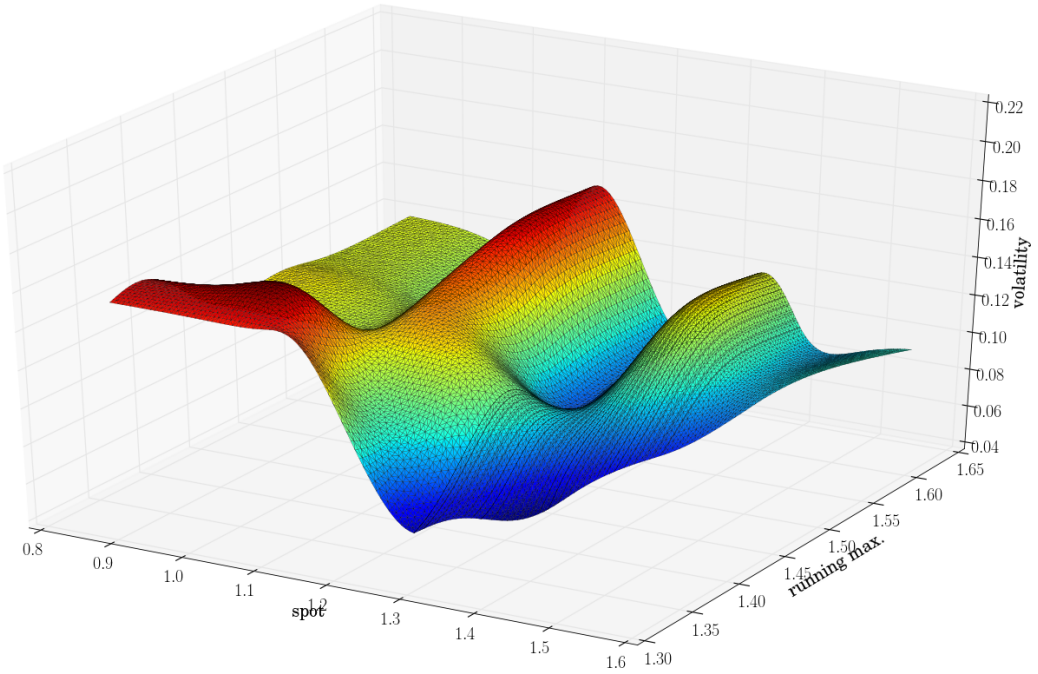


Figure 4.5.3: Brunick–Shreve volatility function $T = 1Y$ with no regularisation.

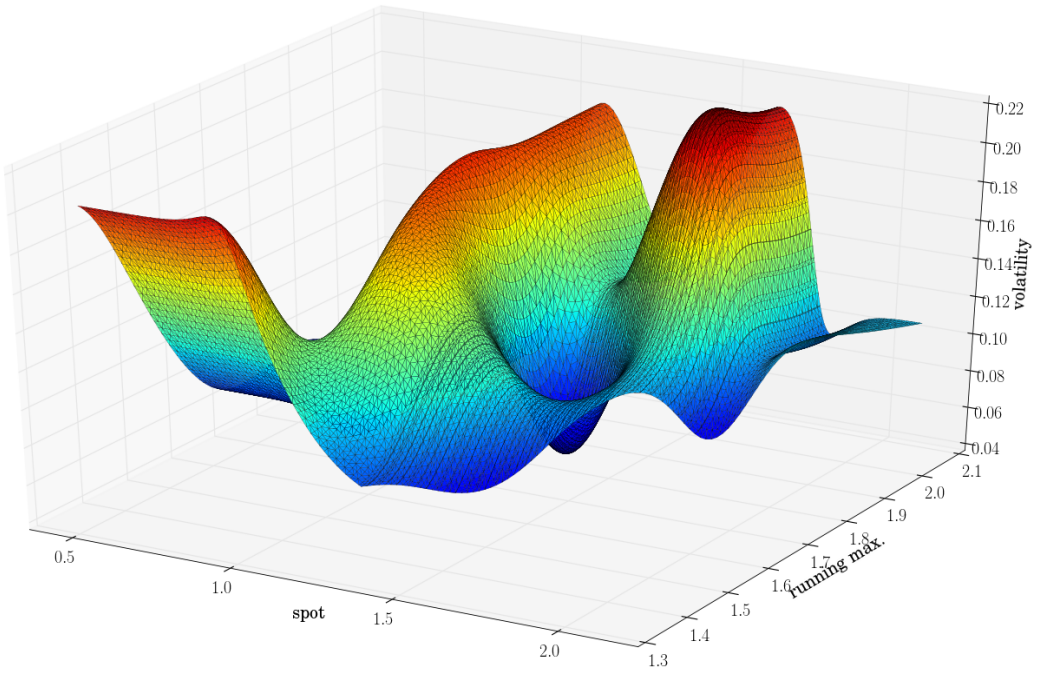


Figure 4.5.4: Brunick–Shreve volatility function $T = 5Y$ with no regularisation.

is refined around the point (S_0, S_0), a region where the solution changes rapidly in space, and coarser for small values of the spot or large values of the running maximum. The resulting projected local volatility is plotted in Figure 4.5.6, next to the local volatility calibrated by the standard Dupire PDE[3] in Figure 4.5.5. The two look very similar as desired.

4.6 Calibration of the LMSV model by 2D particle method

In this section, we discuss the calibration algorithm for the LMSV model (4.2.6). We assume that a calibrated Brunick–Shreve volatility σ_{BSV} is at our disposal.

First we calibrate a pure Heston model to vanilla options by minimising the mean square error for the difference between market and Heston implied volatility. For research purposes only, we use a Basin-Hoppin global optimisation [147], however, a variance swap-based calibration may be preferable in practice (see [110, 79] for details).

We choose a particle method to compute the Markovian projection, since with higher-dimensional models, such as stochastic rates, stochastic vol-of-vol or stochastic correlation, the particle method yields a straightforward extension. However, for the two-factors, three-states model (4.2.6), it is also

[3] using cubic splines with five nodes in strike direction and flat interpolation between maturities, by a fixed-point algorithm [127] without regularisation

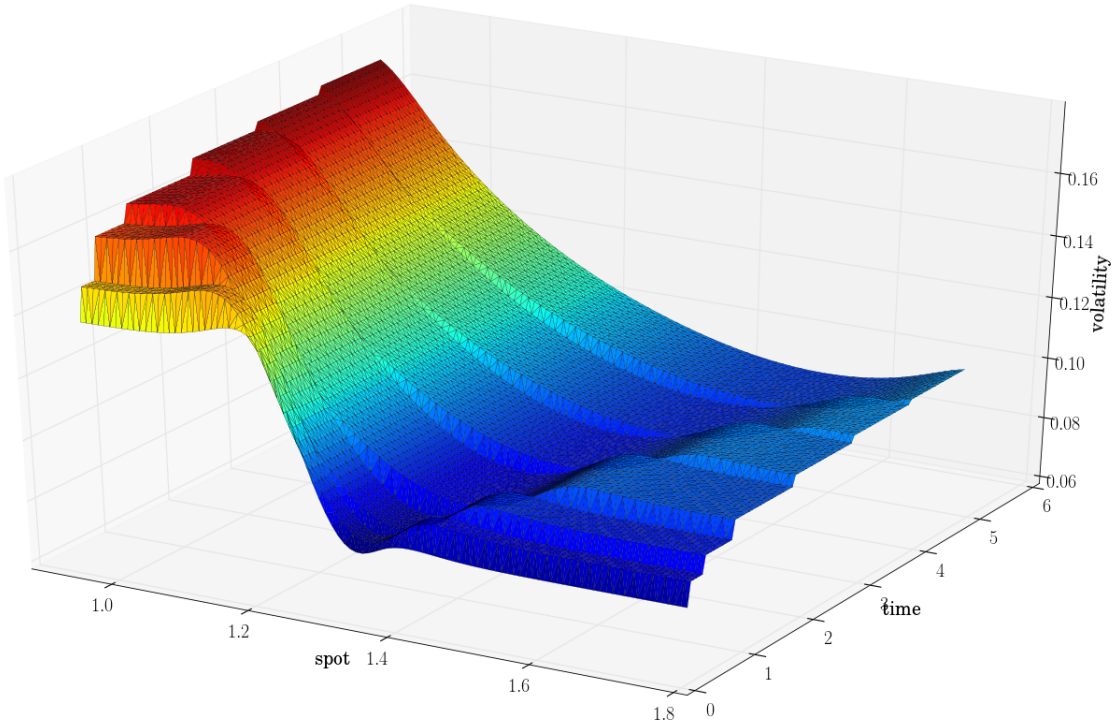


Figure 4.5.5: EURUSD local volatility calibrated by forward Dupire PDE.

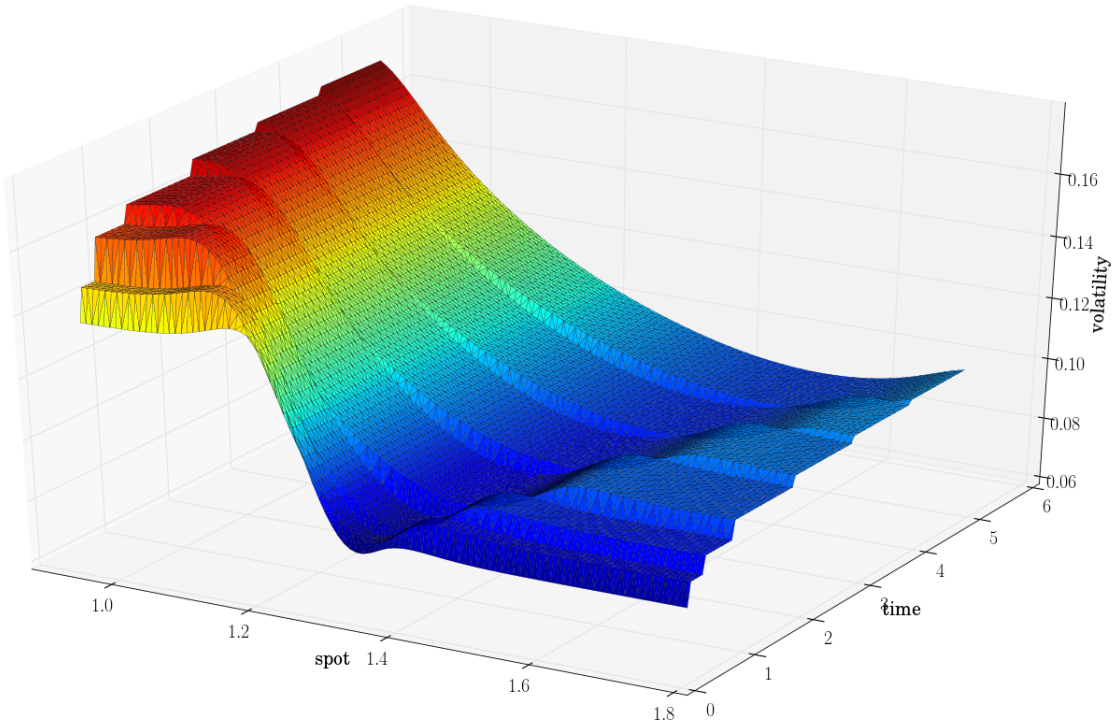


Figure 4.5.6: EURUSD local volatility projected from the calibrated Brunick–Shreve volatility.

possible to numerically solve the Fokker–Plank PDE (4.3.26) for ϕ, and compute the Markovian projection by quadrature as

$$\mathbb{E}^{\mathbb{Q}^d}[V_T \,|\, S_T = K,\, M_T = B] = \frac{\int_0^\infty v\phi(K,B,v,T)\,dv}{\int_0^\infty \phi(K,B,v,T)\,dv}\,.$$

As shown by the cross-model calibration fit in Section 4.8 as well as the results in [35], the Heston-type LSV model with the Heston parameters calibrated on vanilla options overestimates the no-touch probability. A rule-of-thumb [35] suggests that dividing ξ calibrated to vanillas by two, i.e. taking the so-called "mixing factor" in (4.2.4) to be 0.5, provides a good starting point to fitting no-touch options; see also the beginning of Section 4.7. The parameters used are thus as displayed in Table 4.6.1.

v_0	θ	κ	ρ	ξ
0.00827	0.01564	0.7147	-0.4429	0.0947

Table 4.6.1: The calibrated Heston parameters.

In order to run the particle method, we use the Monte Carlo QE-scheme [7] time marching for model (4.2.6). The scheme is adapted from the one described in Chapter 3 where the volatility now depends on the running maximum as well. The running maximum evolution is performed with a Brownian bridge technique as described in Chapter 6 of [73] where we sample the maximum of the Brownian bridge that starts at S_t and ends at $S_{t+\Delta t}$, with α kept constant, and update the maximum process. Namely, the maximum process is approximated as follows

$$\begin{aligned} G_t &= \frac{S_t + S_{t+\Delta t} + \sqrt{(S_t + S_{t+\Delta t})^2 - 2\left(S_t\alpha(S_t, M_t, t)\sqrt{V_t}\right)^2 \Delta t\, \log(U_t)}}{2} \\ M_{t+\Delta t} &= \max(G_t, M_t)\,, \end{aligned}$$

where U_t is an independent draw from the uniform distribution $\mathcal{U}(0,1)$, i.i.d. across t.

We recall from (4.5) that the leverage function has to satisfy

$$\alpha^2(K,B,T)\mathbb{E}^{\mathbb{Q}^d}[V_T \,|\, S_T = K,\, M_T = B] \;=\; \sigma^2_{BSV}(K,B,T)\,, \tag{4.6.1}$$

to ensure perfect calibration to both vanilla and no-touch options for all maturities as long as α is bounded by 2.2, which is the value that gives $T^* = \infty$ in (4.3.24) and is easily verified in our case. Here, σ_{BSV} was found in Section 4.5 and it remains to estimate the conditional expectation.

Through the conditional expectation, α depends on the distribution of the joint process $X_t = (S_t, V_t, M_t)$. The process thus falls in the class of McKean-Vlasov processes [108]. The particle method was introduced in [108] and is discussed in detail in [139]; it was applied to LSV model calibration in

[82, 83] for a projection onto a one-dimensional state S_t. In this work, we extend this approach and consider a projection, i.e. a conditional expectation, onto the two-dimensional state space (S_t, M_t).

We consider N-sample paths $(X_t^i)_{i\in[\![1,N]\!]} = (S_t^i, V_t^i, M_t^i)_{i\in[\![1,N]\!]}$ of X. Then α can be estimated by

$$\hat{\alpha}_N\Big(K, B, T, \big(X_t^j\big)_{j\leq N}\Big) = \sqrt{\frac{\sigma_{BSV}^2(K,B,T)}{\hat{p}_N(K,B,T)}},$$

with

$$\hat{p}_N(K,B,T) = \frac{\frac{1}{N}\sum_{i=1}^N V_T^i \delta_N(S_T^i - K,\, M_T^i - B, T) + 2\theta\xi\epsilon}{\frac{1}{N}\sum_{i=1}^N \delta_N(S_T^i - K,\, M_T^i - B, T) + \xi\epsilon},$$

with δ_N a two-dimensional kernel function and for which the extra terms $2\theta\xi\epsilon$ and $\xi\epsilon$ serves as a smooth extrapolation rule for areas containing a small amount of particles. These extra terms include ξ to remove the smoothing if there is no stochastic volatility. In our test we pick $\epsilon = 10^{-4}$. The $(3\times N)$-dimensional SDE driving the system $(X_t^i)_{i\leq N}$ can be written as

$$\begin{cases} \dfrac{dS_t^i}{S_t^i} = \big(r^{d,i}(t) - r^{f,i}(t)\big)\,dt + \hat{\alpha}_N\Big(S_t^i, M_t^i, t, \big(X_t^j\big)_{j\leq N}\Big)\sqrt{V_t^i}\,dW_t^i \\ dV_t^i = \kappa(\theta - V_t^i)\,dt + \xi\sqrt{V_t^i}\,dW_t^{V,i} \\ M_t^i = \max\limits_{0\leq u\leq t} S_u^i\,, \end{cases}$$

where $(W_t^i, W_t^{V,i})_{i\in[\![1,N]\!]}$ are N independent copies of the 2 correlated driving Brownian motions. We refer to Chapter 3, [82] and [139] for more extensive details about the particle method and convergence results. Similarly, as in the aforementioned works, $\hat{p}_N$ converges weakly to the true expectation such that $\hat{\alpha}_N(K,B,T,(X_T^j)_{j\leq N})$ converges in probability to the solution $\alpha(K,B,T)$ as $N\to\infty$.

We denote by N_T the number of time steps. The leverage function α is defined on a grid of points and interpolated with bi-variate quadratic splines in the spot and running maximum directions and forward flat in time. There are N_T+1 volatility "slices" in total such that we denote the m-th time slice, i.e., $\alpha(\cdot, \blacksquare, T_m)$, by α_m, represented numerically as splines on a grid of $N_S \times N_M$ points. While having N_S and N_M too small will lead to accuracy problems, choosing them too large will make the surface rougher due to over-fitting. We found a grid of 15×10 points to provide a good trade-off between accuracy and smoothness. We do not use any regularisation for α here. For a given time slice T_m, the leverage function is defined on the spatial domain $[S_{min}^m,\, S_{max}^m] \times [S_0,\, S_{max}^m]$ and is extrapolated flat outside these bounds. The surface construction starts by defining a spot grid where we need more grid points around the forward value and less around S_{min}^m and S_{max}^m. We then use a hyperbolic grid (with $\eta = 0.05$, see Chapter 3 for more details) refined around the forward value

$$F_m = S_0 e^{\int_0^{T_m}\left(\bar{r}^d(t) - \bar{r}^f(t)\right)dt}$$

with

$$S^m_{min} = F_m e^{-\frac{6}{2}\sigma_F(T_m)\sqrt{T_m}}, \quad S^m_{max} = F_m e^{\frac{6}{2}\sigma_F(T_m)\sqrt{T_m}}.$$

The spot grid is denoted by $(S_{m,j})_{m\leq N_T,\, j\leq N_S}$. The creation of the running maximum grid is done selecting the nodes of the spot grid above the initial spot S_0 which leads to $N_B(m) \leq N_s$, where N_B is now time dependent (around 9-10 in our case). This particular construction is crucial for accuracy as it ensures that the diagonal where $S = M$ is part of the grid. Each of the grid values can be seen as a parameter and we denote them by $(\alpha_{m,j,k})_{m\leq N_T,\, j\leq N_S, k\leq N_M(m)}$ with the associated spot-maximum grid values $(S_{m,j})_{m\leq N_T,\, j\leq N_S}$ and $(M_{m,k})_{m\leq N_T,\, k\leq N_M(m)}$. We denote the 2-factor, 3-states, particle system at time T for the model (4.2.6) by $(S^i_T, V^i_T, M^i_T)_{i\leq N}$.

We work with an anisotropic bi-variate Gaussian kernel

$$\begin{aligned}
\delta_N(x,y,T) &= \frac{\exp\left(-\frac{1}{2}\frac{\zeta(T)}{1-\rho^2_{xy}(T)}\right)}{\gamma(T)}, \\
\zeta(T) &= \frac{x^2}{h^2_x(T)} + \frac{y^2}{h^2_y(T)} - 2\frac{\rho_{xy}(T)xy}{h_x(T)h_y(T)} \\
\gamma(T) &= (2\pi h_x(T)h_y(T)\sqrt{1-\rho^2_{xy}(T)})
\end{aligned}$$

with a specific bandwidth constructed heuristically as follows,

$$\begin{aligned}
h_x(T) &= h_y(T) = \eta S_0 \sigma_{BSV}(S_0, S_0, T)\sqrt{\max(T, T_{min})}N^{-\frac{1}{6}} \\
\rho_{xy}(T) &= (\rho_{max} - \rho_{sample}(T))\exp\left(-k\frac{N_T}{T_{max}}\right) + \rho_{sample}(T) \\
k &= \frac{2T_{max}\ln(2)}{N_0},
\end{aligned}$$

where

$$\eta = 1.5, \quad T_{min} = 0.25,\ \rho_{max} = 0.98,\ N_0 = 180,$$

T_{max} is the last quoted maturity and $\rho_{sample}(T)$ the correlation computed with the sampled particles at time T. To speed up the computation, it is enough to update $\rho_{sample}(T)$ once every year or half-year. The use of an appropriate bandwidth has a significant impact on the accuracy of the method. This bandwidth is inspired by a Silverman-type rule (see [137], [82]) for the values of h_x and h_y and an experimental definition for the correlation part. A heuristic analysis suggests that if the number of time steps is low, the running maximum with the Brownian bridge technique will tend to be underestimated in the important area around (S_0, S_0), since, as seen in Figure 4.6.2, the leverage function is an increasing function of the running maximum in this region. Indeed, in between two time steps t and $t + \Delta t$, the Brownian bridge technique freezes the value of α for a running maximum M at time t smaller than the value of M along $[t, t + \Delta t]$, and therefore, α is underestimated on average such that $M_{t+\Delta t}$ computed by Brownian bridge is smaller than its exact value. The intuition behind

the specification of the correlation $\rho_{xy}(T)$ is as follows: if the number of time steps is low, the running maximum will tend to be underestimated, hence, we give more importance to the particles where the running maximum is higher (which happens when the spot increases), if the number of time steps is large enough, i.e. more than one per day, the running maximum bias is lower and we rely on the sample correlation. It is important to mention that in practice, $\rho_{Sample}(T)$ is around $80\% \pm 5\%$ and does not change significantly over time. Also, the closer we get to the boundary $S = M$, the higher the correlation, which can easily reach 95%. This is due to the fact that for particles where M is large, it is highly probable that S is large as well, therefore, most of the particles will gather around the diagonal boundary. Conversely, if M is around S_0, most particles are for spots going downwards. This is easily observed from the (S_t, M_t) joint density where the mass aggregates around the area (S_0, S_0), the line $S = M$ and the line $M = S_0$ as displayed in Figure 4.6.1 where we plot the density with the finite element method.

Remark. If we were to estimate the (S, M) joint density, due to the fact that the running maximum is always above S_0 and greater than the spot, there is a boundary bias introduced by using a bi-variate Gaussian kernel that has support over $\mathbb{R}^2$, although in our case, this problem is barely noticeable since we are only interested in the conditional expectation. It would be possible to use a gamma kernel in the running maximum direction as discussed in [19] and reflect the mass of the kernel when $S > M$ inside the domain. However, the approach is costly numerically and was found to lead to only minor improvements.

For a given level of spot and running maximum K and B, we only want to compute the kernel function for particles that will have a significant contribution to the sum, i.e. particles that are close enough with respect to the Euclidean metric. This is true when $\delta_N(x, y, T) \leq \epsilon_0$ or in other words all the points contained in the ellipse $\mathcal{E}$ defined by

$$x^2 + y^2 - 2\rho_{xy}(T)xy - H(T) \quad = \quad 0\,,$$

with

$$H(T) = -2h_x^2(T)\ln(\gamma(T)\epsilon_0)\,.$$

The canonical form of ellipse $\mathcal{E}$ expressed in the coordinate system defined by its principal axes is (see [10])

$$\frac{x^2}{-D/(\lambda_1^2\lambda_2)} + \frac{y^2}{-D/(\lambda_1\lambda_2^2)} = 1\,,$$

with $D = -H(T)\Big(1 - \rho_{xy}^2(T)\Big)$, $\lambda_1 = 1 - \rho_{xy}(T)$ and $\lambda_2 = 1 + \rho_{xy}(T)$, the roots of $\lambda \to \lambda^2 - 2\lambda + \Big(1 - \rho_{xy}^2(T)\Big)$. Since $\rho_{xy}(T) \geq 0$, the semi-major axis length is

$$R(T) = \sqrt{-D/(\lambda_1^2\lambda_2)} = \sqrt{\frac{H(T)}{1 - \rho_{xy}(T)}}.$$

Hence, working with the Euclidean distance, the particles with a significant contribution on the value of $\hat{p}_N(K, B, T)$ are in the ball of center (K, B) and radius $R(T)$. We use $\epsilon_0 = \frac{\xi_\epsilon}{10}$ in our tests. A large value of ϵ_0 leads to a fast computation of $\hat{p}_N$ but also reduces the accuracy of the result. In order to perform a distance query efficiently, we build a k-d tree as described in [23], with a worst case complexity of $\mathcal{O}(2n \log n)$ and in which we can perform a binary search with complexity $\mathcal{O}(\log n)$ on average, to find all particles with distance less than ϵ_0. The use of k-d trees for particle method calibration is a novel approach that can easily be generalised to higher-dimensional state spaces. This provides fast access to the nearest neighbours, and, as a consequence, allows to list the particles contained in the ball of centre (K, B) and radius $R(T)$. The construction is based on finding the median particle position with respect to the x-axis by a quick-select algorithm and putting this particle on the first node. Then, we run the k-d tree algorithm on the list of particles smaller than the node switching to the y-axis and put the result on the left leaf. Similarly, perform the same process for the particles greater than the node and put the result on the right leaf. Repeat the algorithm, alternating between median selection on the x- and y-axis until a single particle is left in the list.

Remark. In order to further improve the performance, we could define a distance function between two particles as

$$d(\mathbf{x}_1, \mathbf{x}_2) = \sqrt{(s_1 - s_2)^2 + (m_1 - m_2)^2 - 2\rho(s_1 - s_2)(m_1 - m_2)}\,,$$

where the coordinates of a given particle i are given by $\mathbf{x}_i = [s_i,\, m_i]^T$, with $i \in [\![1, 2]\!]$ in this example. Then d can be combined with a metric tree algorithm [155] for an optimal nearest neighbours search, by only keeping particles within a distance $H(T)$ to the point (K, B) of interest. We note that d is in fact a metric as it is non-negative, symmetric, and verifies the identify of indiscernibles. Additionally, by the following change of variables which transforms $(\mathbf{x}_1, \mathbf{x}_2)$ to $(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2)$, and corresponds to a rotation of angle $\theta = -\pi/4$ and a homothety,

$$\begin{aligned}(\bar{s}_1 - \bar{s}_2) &= ((s_1 - s_2)\cos\theta - (m_1 - m_2)\sin\theta)\sqrt{1-\rho}\,,\\ (\bar{m}_1 - \bar{m}_2) &= ((s_1 - s_2)\sin\theta + (m_1 - m_2)\cos\theta)\sqrt{1+\rho}\,,\end{aligned}$$

d can be written as the Euclidian metric $\bar{d}$ with

$$\bar{d}(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2) = \sqrt{(\bar{s}_1 - \bar{s}_2)^2 + (\bar{m}_1 - \bar{m}_2)^2} = d(\mathbf{x}_1, \mathbf{x}_2)\,,$$

and therefore, the triangle inequality also holds for d.

When we used the normal Silverman rule bandwidth, i.e. $\rho_{xy}(T) = 0$, the accuracy was found to drop drastically and became unsatisfactory for a standard amount of particles, i.e. the range $100\,000 - 2\,000\,000$.

The step-by-step calibration is detailed in Algorithm 4.2.

If the Heston parameters are degenerate, such that $V_t = v_0$ for all t, then the calibration algorithm recovers, as expected, the calibrated Brunick–Shreve volatility. Similarly, if $\xi = 0$, only one particle is sufficient to calibrate the LMSV model.

Algorithm 4.2 Calibration of α with 2D particle method

$\alpha(S, M, T_1 = 0) = \frac{\sigma_{BSV}(S,M,0)}{\sqrt{v_0}}$
for $(\ m = 1\,;\, m \le N_T\,;\, m++)$ **do**
  generate $(Z, Z_v)_{i\le N}$ and $(U_v, U_{max})_{i\le N}$, i.e. $2 \times N$ independent draws from $\mathcal{N}(0,1)$ and $2 \times N$ draws from $\mathcal{U}([0,1])$, respectively
  evolve the 2-factor 3-state particle system from T_m to T_{m+1} with the $QE-$Scheme (Z, Z_v, U_v) where the running maximum is computed by Brownian bridge (U_{max}) and where $\alpha(S, M, [T_m, T_{m+1}[) = \alpha(S, M, T_m)$
  build the particles k-d tree partitioning on the position values $(S^i\,, M^i)$
  set $T = T_{m+1}$
  for $(\ k = 1\,;\, k \le N_M\,;\, k++)$ **do**
    for $(\ j = 1\,;\, j \le N_S\,;\, j++)$ **do**
      set $K = S_{m+1,j}$; $B = M_{m+1,k}$
      select using the k-d tree, let the set $I(K,B)$ of particles i for which $\sqrt{\left(S_T^i - K\right)^2 + \left(M_T^i - B\right)^2} \le R(T)$
      if $(K \le B)$ **then**
        compute and update

$$\hat{p}_N = \frac{\frac{1}{|I(K,B)|}\sum_{i\in I(K,B)} V_T^i \delta_N\big(S_T^i - K,\, M_T^i - B, T\big) + 2\theta\xi\epsilon}{\frac{1}{|I(K,B)|}\sum_{i\in I(K,B)} \delta_N\big(S_T^i - K,\, M_T^i - B, T\big) + \xi\epsilon}$$

      end if
      compute

$$\alpha_{m+1,j,k} = \frac{\sigma_{BSV}(K,B,T)}{\sqrt{\hat{p}_N}}$$

    end for
  end for
end for

The model is calibrated to vanilla and no-touch quotes as detailed in Section 4.2.2. We plot in Figure 4.6.2 the calibrated leverage function for $T = 1$ with 350 time steps per year and 2 000 000 particles. The Brownian increments are generated with a pseudorandom number generator. The calibration fit is compared to other models in Section 4.8.

4.7 Calibration of the Heston-type LSV-LVV model

In this section, we calibrate the LSV-LVV model (4.2.7) to both vanilla and no-touch options. It generalises the Heston-type SLV model, which is state-of-the-art in the industry. For any set of

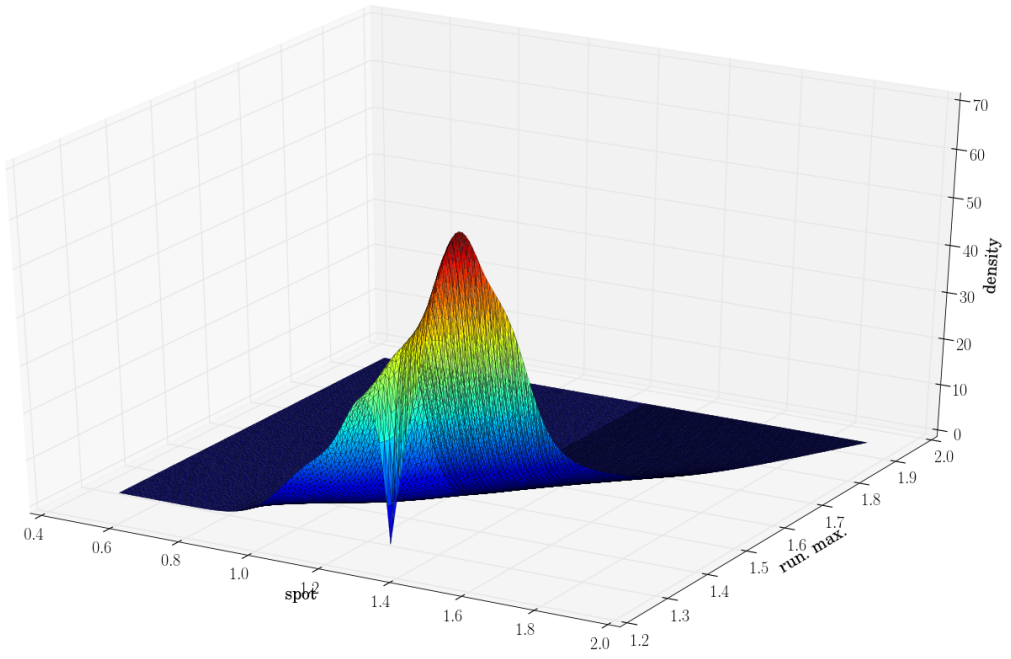


Figure 4.6.1: Joint density of the spot and running maximum after calibration at time $T = 1$.

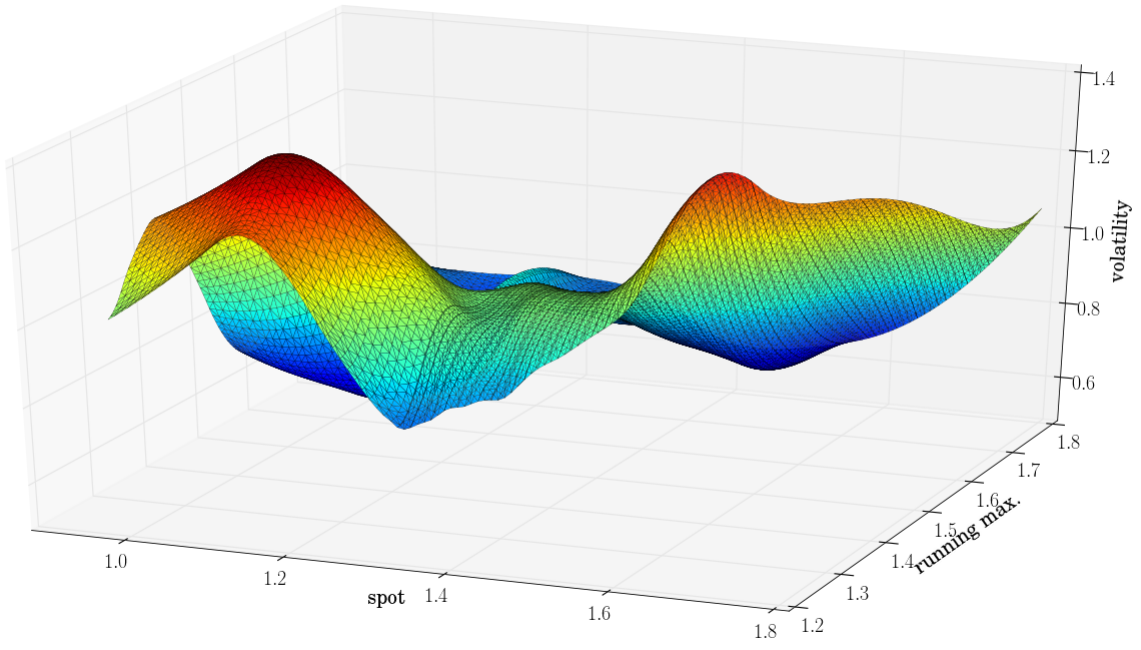


Figure 4.6.2: Calibrated leverage function for the LMSV model at time $T = 1$.

parameters $v_0, \kappa, \theta, \xi, \rho, \eta$ in the Heston-LSV model ($\eta = 1$), a leverage function α_S can be found such that the model gives a perfect fit to arbitrage-free vanilla quotes, by

$$\alpha_S(K,T) \quad = \quad \frac{\sigma_{LV}(K,T)}{\sqrt{\mathbb{E}^{\mathbb{Q}^d}[V_T \,|\, S_T = K]}}\,. \tag{4.7.1}$$

For instance, one can find the Heston parameters that best fit the same vanilla quotes, with the hope that a relatively flat α_S will suffice for a perfect fit.

With $v_0, \kappa, \theta, \xi, \rho, \eta$ fixed, one can then vary η, under the constraint (4.7.1), to find the best fit to no-touch options among this one-parameter family of models perfectly calibrated to vanillas. The parameter η is commonly within the range $[0, 1]$ and called the "mixing factor" (see [35]): the market is believed to stand in between pure local volatility models, i.e. $\eta = 0$, and full Heston LSV models, i.e. $\eta = 1$. We show calibration results in Section 4.8. To add more flexibility, and since no-touch options depend on the running maximum, the vol-of-vol is also made spot- and time-dependent in our model. The idea of defining such a model has been proposed in [109].

4.7.1 Model definition

In this work, we consider a vol-of-vol that is a piecewise linear function of the spot and backward flat in time in between quoted maturity pillars. More precisely, for any time T^{Mat} inside two quoted pillars $[T_n, T_{n+1}]$, B_{max} the biggest quoted no-touch barrier level for maturity the last quoted maturity $T_{N_{Mat}}$, we write

$$\begin{cases} \xi(S,T) = & \bar{\xi}(u_x(S), T) \\ \bar{\xi}(S,T) = & \max((a_{n+1}(S - S_0) + b_{n+1}), 0.01), \end{cases}$$

where we set $T_0 = 0$ and with u_x defined as in (4.4.8) which performs a smooth extrapolation of $\bar{\xi}$ outside the boundary of the domain $[S_0, B_{max}]$. Note that there are (only) two parameters per maturity, compared to one global vol-of-vol parameter for the Heston model.

In our tests, a_n is negative and lies inside $[-1.5, 0]$ and b_n is usually in the interval $[\xi_{heston}/2, 3\,\xi_{heston}/2]$, where ξ_{heston} is the vol-of-vol of a pure Heston model calibrated to vanilla prices; see Figure 4.7.2 for the resulting shape of ξ.

4.7.2 Calibration methodology

We use a particle method [82, 83] for the calibration to vanillas with 'outer' iterations for the best-fit to no-touch quotes. Then the calibration to vanilla options is almost exact thanks to the calibration

of the leverage function α_S with the particle method, while the no-touch prices are best-fitted by minimising the error over the set of parameters (a_n, b_n).

The error function we use in our test, for a given quoted maturity T_n^{Mat}, is the following:

$$\bar{e}(a_n, b_n) = \sum_{l=1}^{Q_B} \left| FNT^{\text{Model}}\left(B_{T_n^{Mat},l}, T_n^{Mat}, a_n, b_n\right) - FNT^{\text{Market}}\left(B_{T_n^{Mat},l}, T_n^{Mat}\right) \right|,$$

where Q_B is the number of quoted barrier levels and $\left(B_{T_n^{Mat},l}\right)_{1\leq l\leq Q_B}$ the vector of quoted barrier levels. If one has access to the bid–ask spread, a more appropriate measure in practice might be

$$\bar{e}(a_n, b_n) = \sum_{l=1}^{Q_B} \left| \frac{FNT^{\text{Model}}\left(B_{T_n^{Mat},l}, T_n^{Mat}, a_n, b_n\right) - FNT^{\text{Market}}_{\text{Mid}}\left(B_{T_n^{Mat},l}, T_n^{Mat}\right)}{FNT^{\text{Market}}_{\text{Bid}}\left(B_{T_n^{Mat},l}, T_n^{Mat}\right) - FNT^{\text{Market}}_{\text{Ask}}\left(B_{T_n^{Mat},l}, T_n^{Mat}\right)} \right|,$$

where in our data, we notice a tighter spread around the 90% no-touch up probability, i.e. the 10% one-touch up probability. We decide to not normalise by the bid–ask spread in our numerical experiments.

At each time step, we compute two Markovian projections by an adaptation of the particle method described in Section 4.6: the classical projection onto the spot for vanilla calibration as in (4.7.1) to define the leverage function, and the Markovian projection of V onto (S, M) to define a projected volatility $\sigma_{|S,M}$ as

$$\sigma_{|S,M}(K, B, T) \quad = \quad \alpha_S(K, T)\sqrt{\mathbb{E}^{\mathbb{Q}^d}[V_T \,|\, S_T = K,\, M_T = B]}\,. \tag{4.7.2}$$

A calibrated local volatility function σ_{LV} is at our disposal and has been calibrated with the procedure presented in Chapter 3. In order to compute the price of no-touch options, we use the forward PIDE for barrier options (4.4.1) combined with the estimated value of $\sigma_{|S,M}$. The Brunick–Shreve model (4.2.5) with this volatility $\sigma_{BSV} = \sigma_{|S,M}$ has the same theoretical joint (S, M) distribution as the LSV-LVV model (4.2.7), and therefore the same no-touch prices. We then minimise $\bar{e}$ under the constraint (4.7.1).

As seen in Figure 4.7.2, both parameters are stable from one maturity to the next and make a good first guess for the next quoted pillar calibration. More precisely, for the shortest quoted maturity, we suggests to start the optimisation with $a_1 = -1$ and $b_1 = \xi_{heston}$ and then use the calibrated (a_n, b_n) of T_n^{Mat} as a first guess for the calibration of no-touch quotes at T_{n+1}^{Mat}. We repeat the process in between each quoted pillar until a local minimum is reached for the calibration error measure to no-touch options. By construction, the vanilla options will be perfectly calibrated, if enough particles are used, since the leverage function α_S is updated at each iteration. For the proof of concept, we use a sequential algorithm coupled with a Nelder–Mead gradient free optimisation (see [113]), for which the convergence is obtained in $10 - 20$ iterations in our tests. There is clearly room for improvement.

We denote by N_{Mat} the number of quoted maturities, with $\left(T_n^{Mat}\right)_{n\leq N_{Mat}}$ the set of quoted pillars, by N_T the number of time steps with $(T_m)_{m\leq N_T}$ a time grid that also contains all the quoted pillars T_n^{Mat} and by N the number of particles.

The step-by-step calibration is detailed in Algorithm 4.3.

Algorithm 4.3 Calibration of (α_S, ξ) with 2D particle method, forward PIDE and 'outer' iterations

$\alpha_S(S,0) = \frac{\sigma_{LV}(S,0)}{\sqrt{v_0}}$
$\sigma_{|S,M}(S,M,0) = \sigma_{LV}(S,0)$
$N_T^{down} = 0$
for $(\ n=1\,;\, n \le N_T^{Mat}\,;\, n++)$ **do**
  find N_T^{up} such that $T_{N_T^{up}} = T_n^{Mat}$
  set the optimisation convergence variable to false: $\mathcal{Z} = 0$
  while $\mathcal{Z} = 0$ **do**
    for $(\ m = N_T^{down}\,;\, m < N_T^{up}\,;\, m++)$ **do**
      generate $(Z, Z_v)_{i\le N}$ and $(U_v, U_{max})_{i\le N}$, i.e. $2\times N$ independent draws from $\mathcal{N}(0,1)$ and $2\times N$ draws from $\mathcal{U}([0,1])$, respectively
      evolve the 2-factor 3-state particle system of model (4.2.7) from T_m to T_{m+1} with the $QE-$Scheme (Z, Z_v, U_v) where the running maximum is computed by Brownian bridge (U_{max}) and where $\alpha_S(S, [T_m, T_{m+1}[) = \alpha_S(S, T_m)$
      build as in Section (4.6), the particles k-d tree partitioning on the position values $(S^i,\, M^i)$, for the evaluation of $I(K,B)$ in (4.7.3)
      set $T = T_{m+1}$
      for each K on a grid, compute as in [82] and Chapter 3 and for a set of selected significant particles $I(K)$

$$\alpha_S(K,T) = \frac{\sigma_{LV}(K,T)\sqrt{\sum_{i\in I(K)} \delta_N^S(S_T^i - K, T)}}{\sqrt{\sum_{i\in I(K)} V_T^i \delta_N^S(S_T^i - K, T)}}\,,$$

      with δ_N^S a one-dimensional kernel function
      for each (K,B) on a grid, compute as in Section (4.6) and for a set of selected significant particles $I(K,B)$

$$\sigma_{|S,M}(K,B,T) = \frac{\alpha(K,T)\sqrt{\frac{1}{|I(K,B)|}\sum_{i\in I(K,B)} V_T^i \delta_N(S_T^i - K,\, M_T^i - B, T) + 2\theta\xi\epsilon}}{\sqrt{\frac{1}{|I(K,B)|}\sum_{i\in I(K,B)} \delta_N(S_T^i - K,\, M_T^i - B, T) + \xi\epsilon}} \quad (4.7.3)$$

      where δ_N is a two-dimensional kernel function
      solve the forward PIDE for barriers (4.4.1) by BDF2 implicit step on $[T_m, T_{m+1}]$ with a volatility of $\sigma_{|S,M}(K,B,T_{m+1})$
    end for
    compute model foreign no-touch price FNT^{Model} for maturity T_n^{Mat} from the computed up-and-out call prices $C\left(0,B,T_n^{Mat}\right)/S_0$
    compute the objective function

$$\bar{e}(a_n,b_n) = \sum_{l=1}^{Q_B}\left|FNT^{\text{Model}}\left(B_{T_n^{Mat},l}, T_n^{Mat}, a_n, b_n\right) - FNT^{\text{Market}}\left(B_{T_n^{Mat},l}, T_n^{Mat}\right)\right|$$

    update (a_n, b_n) guess with the Nelder–Mead algorithm, check convergence based on the difference with the previous guess and update $\mathcal{Z}$
  end while
end for

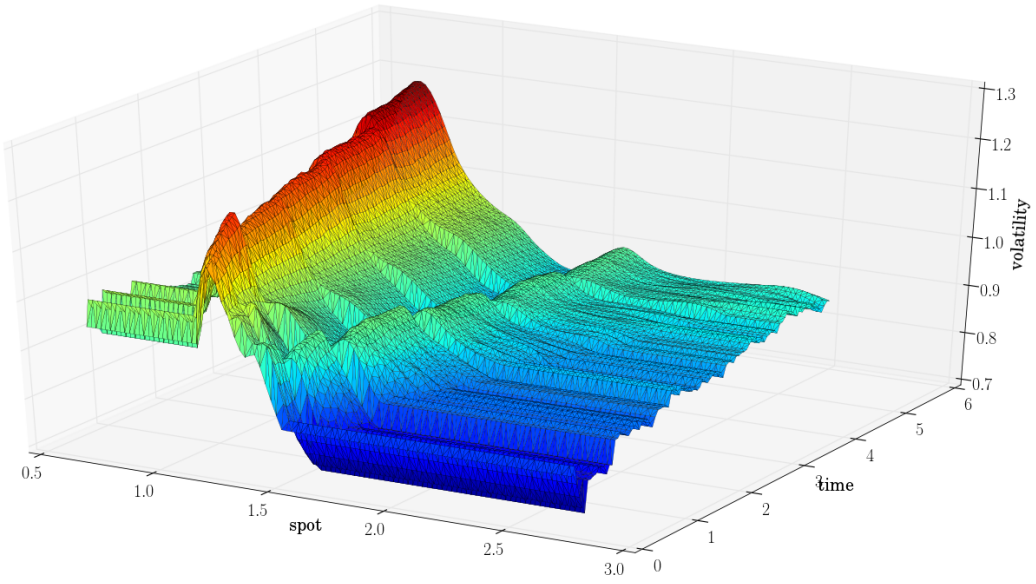


Figure 4.7.1: Calibrated leverage function for the LSV-LVV model.

Finally, we plot in Figure 4.7.1 the calibrated leverage function α_S and in Figure 4.7.2 the calibrated local vol-of-vol ξ from 3 months onward, obtained with 100 time steps per year, 500 000 particles and 900 strike points for the PIDE. For these numerical parameters, the calibration is converged sufficiently. The fit to market data is discussed in Section 4.8.

We can fit the model (4.2.4) similarly to the local vol-of-vol case. We apply Algorithm 4.3 and best-fit the (constant) mixing factor η to no-touch options, with all other parameters best-fitted to vanilla options (see Table 4.6.1, but with $\xi = 0.1894$ twice the value).

The obtained mixing factor is $\eta = 0.5528$ and the associated model will be denominated "LSV mixing=0.55" in the remainder of the chapter. This is in line with [13] where the market realised vol-of-vol is around $50\% - 60\%$ of the Heston calibrated vol-of-vol on vanilla prices. This confirms the market belief [9] that the calibration of touch options, and barriers in general, should provide relevant information about the market smile dynamics.

The calibration results are displayed in Section 4.8 where the model performs very well for longer-term maturities.

Remark (Run times and parallelisation). According to our numerical experiments, the two-dimensional particle method to compute α_S and $\sigma_{|S,M}$ takes approximately two to three times as long as the one-dimensional particle method used for vanilla calibration, i.e. the computation of α_S alone (including the computational time for the particle scheme evolution). Additionally, we also need to solve the forward PIDE at each time step, a task that has a comparable computational time as solving the two-dimensional Heston pricing PDE. Furthermore, the Nelder–Mead optimisation necessitates 'outer'

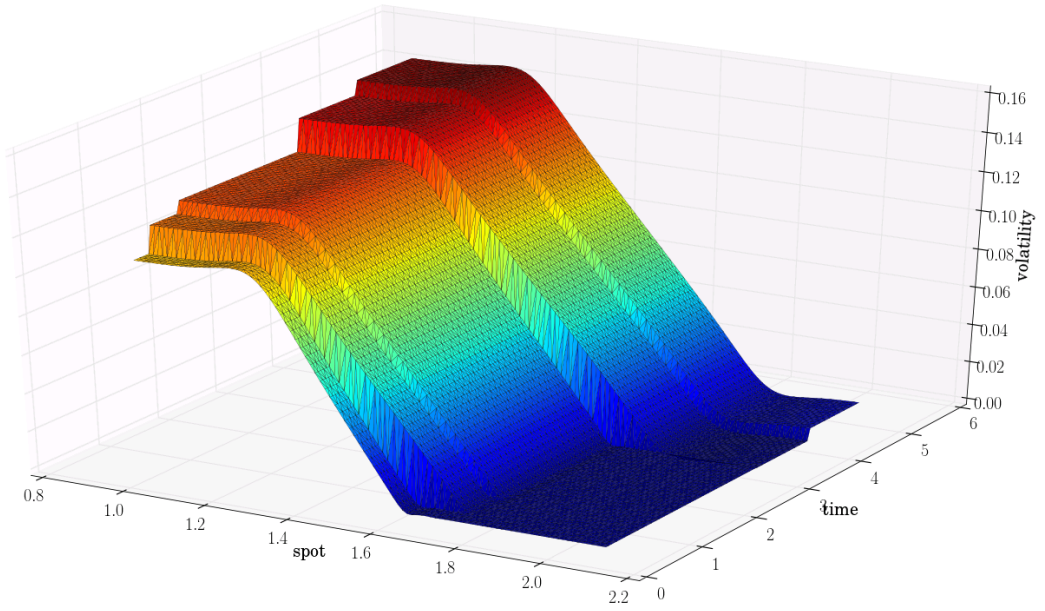


Figure 4.7.2: Calibrated local vol-of-vol $\xi(S,t)$ for the LSV-LVV model.

iterations performed for every guess of (a_n, b_n), where the particle method as well as the forward PIDE have to be solved. If an average of 15 iterations are required for the optimisation algorithm, the computational time for the whole routine would be comparable to 30 particle method runs (for vanilla calibration) plus 15 Heston PDE evaluations. This is unfortunately too slow for practical applications. Hence, we propose to remove the 'outer' iterations bottleneck, i.e. the while loop in Algorithm 4.3, by parallelisation, e.g. by computing the objective function in parallel for different pairs (a, b) and approximating the optimiser by Chebyshev polynomials [67].

4.7.3 Full Monte Carlo approach

In order to compute the price of foreign no-touch options, we rely on the forward PIDE for barrier options combined with the Markovian projection $\sigma_{|S,M}$. However, another possible approach is to compute the no-touch options prices directly with the simulated paths and perform the optimisation process. Indeed, one can remove line 14 and line 15 of Algorithm 4.3 and replace line 17 by "compute model foreign no-touch price FNT^{Model} for maturity T_n^{Mat} from the simulated Monte Carlo particles". We will refer to this approach as "full Monte Carlo" as the forward PIDE is not required anymore. In both cases, the particle method is still needed to compute the leverage function α_S. We performed three Monte Carlo calibrations, one with 100 000 particles, one with 500 000 particles both with 100 time steps per year, and one with 1 000 000 with 300 time steps per year. Additionally, we also calibrated using the forward PIDE with 100 000 particles, 100 time steps per year and 200 strike

	PIDE $N = 100\,000$ $N_K = 200$, $N_T = 100$	MC $N = 100\,000$ $N_T = 100$	MC $N = 500\,000$ $N_T = 100$	MC $N = 1\,000\,000$ $N_T = 300$
Avg. *rel* error	0.474%	1.437%	0.797%	0.621%
Timing factor Θ	1	0.3	1.5	10.6

Table 4.7.1: Full Monte Carlo vs PIDE average one-touch price relative error comparison. "Timing factor" is the time needed for the calibration, relative to the PIDE approach computational time. N_T is the number of time steps per year, N the number of particles and N_K the number of strike points in the PIDE.

points. The error measure we use for this comparison is the relative error for one-touch prices,

$$rel \quad = \quad 100 \times \sum_{l=1}^{Q_B} \left| \frac{FOT^{\text{Model}}\left(B_{T_n^{Mat},l}, T_n^{Mat}\right) - FOT^{\text{Market}}\left(B_{T_n^{Mat},l}, T_n^{Mat}\right)}{FOT^{\text{Market}}\left(B_{T_n^{Mat},l}, T_n^{Mat}\right)} \right|,$$

where the foreign one-touch price FOT can be computed from the foreign no-touch price FNT with (4.2.8). A relative error is better suited for one-touch prices since the market spread gets tighter for higher barrier values, and therefore, the error should be emphasised in that region. Let τ_{PIDE} be the computational time needed to perform the PIDE calibration with 100 000 particles, 200 strike space points and 100 time steps per year.

If we denote by τ_{MC} the timing for a full Monte Carlo calibration with 1 000 000 paths and 300 time steps per year, the "timing factor" Θ is defined as

$$\Theta = \frac{\tau_{MC}}{\tau_{PIDE}}\,.$$

In Table 4.7.1, we display the average *rel* error over all barrier levels and maturities as well as the relative computational times Θ with respect to the PIDE approach. We note that on top of the PIDE numerical resolution, the calibration requires the computation of $\sigma_{|S,M}$. However, in our tests, we find that computing both $\sigma_{|S,M}$ and α_S with 100 000 particles is as fast as computing α_S only with 500 000 particles for the full Monte Carlo approach (excluding the computational time for the particle scheme evolution). Table 4.7.1 allows us to conclude that the full Monte Carlo calibration error becomes comparable to the PIDE calibration error only with more than 1 000 000 particles and 300 time steps per year which is more than ten times slower.

The results plotted in Figure 4.7.3 shows the error *rel* for a short- and a long-term maturity, where we notice that the full Monte Carlo approach, even with 1 000 000 particles and 300 times steps struggles for long-term maturities. Monte Carlo pricing of barrier options is numerically challenging as it translates into integrating a discontinuous payoff function, a problem that becomes more pronounced for higher levels of no-touch barriers (equivalent to lower levels of one-touch barriers) as only a few

particles will breach the barrier. These barrier levels corresponds to the region where the market bid–ask spread is usually tighter and would require more accuracy. Estimating $\sigma_{|S,M}$ and pricing with the forward PIDE does not suffer from this problem as the knock-out feature is treated as a Dirichlet boundary condition on the PDE layer.

Additionally, we plot on the same volatility axis, the calibrated local vol-of-vol functions for both the forward PIDE approach with 100 000 particles in Figure 4.7.4 and the full Monte Carlo approach with 500 000 particles in Figure 4.7.5, both with 100 time steps per year, and notice that the use of the forward PIDE leads to more stable parameters over time. More particularly, we notice that the full Monte Carlo approach leads to a steep calibrated vol-of-vol on the short-term. This translates to the vol-of-vol reaching the lower-bound zero for a large set of spot values; a behavior that is not desirable. Hence, we conclude that combining the forward PIDE with the particle method provides an optimal solution for the calibration problem as opposed to the full Monte Carlo technique (compare to the "fully" converged surface in Figure 4.7.2 with more points and particles). However, the latter provides a good alternative if one is not willing to implement the finite difference discretisation of the forward PIDE.

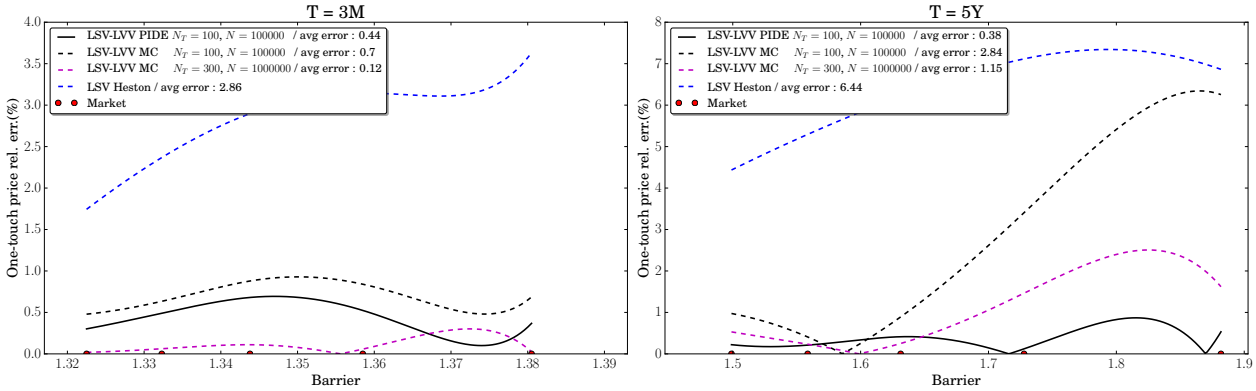


Figure 4.7.3: Full Monte Carlo vs forward PIDE calibration comparison for a short- and long-term maturity pillar, N_T is the number of time steps per year and N the number of particles.

4.8 Calibration results and model comparison

Finally, we present the calibration fit for all models. The calibration fit to vanilla options is given in Table 4.8.1, and the calibration fit to foreign no-touch options is given in Table 4.8.2 and plotted in Figure 4.8.1, for an error defined as

$$e(B,T) = 100 \times \left(FNT^{\text{Model}}(B,T) - FNT^{\text{Market}}(B,T) \right),$$

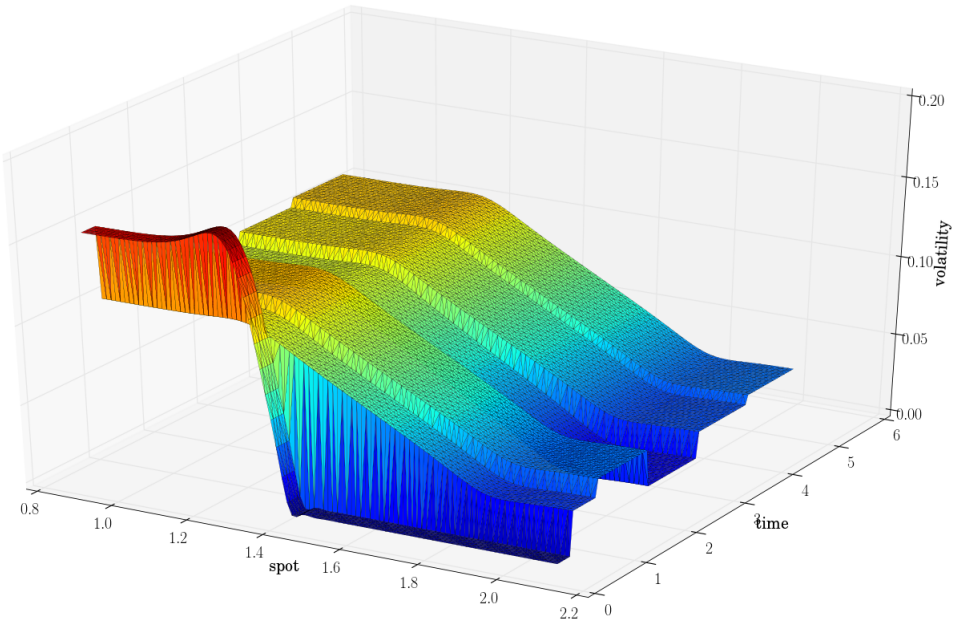


Figure 4.7.4: Local vol-of-vol function calibrated by forward PIDE 200 strike points, 100 000 particles and 100 time steps per year.

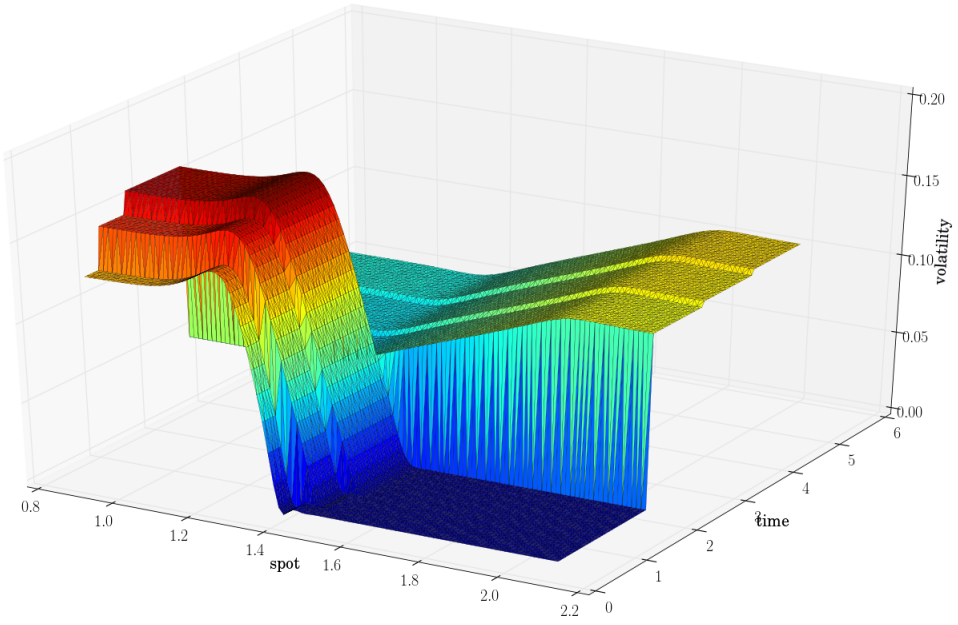


Figure 4.7.5: Local vol-of-vol function calibrated by full Monte Carlo, 500 000 particles and 100 time steps per year.

for both path-dependent models, i.e. the BSV model (4.2.5) and LMSV model (4.2.6), as well as the LSV-LVV model (4.2.7) and the LSV mixing model (4.2.4). As a benchmark, we also include the pure local volatility model and the LSV Heston model, calibrated on vanilla options and where the Heston parameters are also calibrated on call options.

The average error in absolute implied volatility is 0.005% for the BSV model, 0.009% for the LMSV model and 0.014% for the LSV-LVV model and details can be found in Table 4.8.1.

T	**BSV**	**LMSV**	**LSV-LVV**	**LSV** $\eta = 0.55$	**LV**	**LSV Heston**
0.26	0.001	0.008	0.017	0.055	0.001	0.001
0.51	0.003	0.007	0.014	0.040	0.001	0.006
1.01	0.006	0.005	0.016	0.027	0.001	0.005
2.01	0.006	0.005	0.016	0.016	0.000	0.003
3.01	0.005	0.007	0.014	0.011	0.000	0.003
4.01	0.006	0.012	0.013	0.009	0.001	0.005
5	0.006	0.016	0.012	0.005	0.001	0.003

Table 4.8.1: Average absolute error in % for the implied volatility for all models.

For the no-touch prices, we display the average absolute error in % for each maturity as

$$e = 100 \times \frac{1}{Q_B} \sum_{l=1}^{Q_B} \left| FNT^{\text{Model}}\left(B_{T^{Mat},l}, T^{Mat}\right) - FNT^{\text{Market}}\left(B_{T^{Mat},l}, T^{Mat}\right) \right|$$

in Table 4.8.2 where Q_B is the number of quoted barrier levels and $\left(B_{T^{Mat},l}\right)_{1 \le l \le Q_B}$ the set of quoted barrier levels. Hence, for an average absolute error of 0.1% and for a market no-touch probability of 70%, the model will price it at $70\% \pm 0.1\%$ on average. The fit for the two path-dependent models should theoretically be perfect for all vanilla and barrier contracts, and only consists in numerical errors and penalisation, while the LSV-LVV is by construction calibrated to vanilla options but has only two further free parameters per maturity to fit five no-touch prices. All the model prices are computed with 1 000 000 Monte Carlo paths and 365 time steps per year where Brownian increments are generated with Sobol sequences and Brownian bridge construction [28]. The running maximum time scheme is computed with the Brownian bridge technique as in Chapter 6 of [73].

The LSV-LVV, the Brunick–Shreve (BSV) and LMSV models calibrated on both vanilla and foreign no-touch options, perform significantly better for the valuation of no-touch option than the LV or the Heston LSV calibrated on vanilla options only. These results show that the calibration of no-touch options is of paramount importance in order to incorporate the information about the distribution of the running maximum process provided by the market.

As Figure 4.8.1 suggests, if no-touch options are not included in the calibration basket, more classical models like the Heston LSV model can mis-price the no-touch probability by up to 3%. The Brunick–

T	BSV	LMSV	LSV-LVV	LSV $\eta = 0.55$	LV	LSV Heston
0.26	0.012	0.060	0.079	0.834	1.198	0.706
0.51	0.008	0.043	0.140	0.675	1.455	1.454
1.01	0.028	0.047	0.120	0.435	1.605	1.894
2.01	0.029	0.055	0.157	0.123	1.590	2.097
3.01	0.024	0.043	0.148	0.124	1.439	2.186
4.01	0.026	0.055	0.122	0.147	1.420	2.062
5	0.019	0.063	0.166	0.187	1.380	1.960

Table 4.8.2: Average absolute error in % for foreign no-touch quotes for all models; see Fig. 4.8.1 for error plots.

Shreve and LMSV path-dependent models are built to be able to fit no-touch market prices with high precision. The LSV-LVV is designed to not incorporate the running maximum in the parametrisation but to provide the best possible fit to market no-touch prices. We believe the LSV-LVV achieves its main goal by being able to price no-touch option within the market bid–ask spread for every quoted barrier level and maturity.

4.8.1 Smile dynamics

We propose preliminary results to understand how the addition of no-touch options, as a calibration instrument, impacts the underlying smile dynamics. However, our goal is not to provide answers about which model is best suited to reproduce certain market behaviours. A detailed historical analysis, applied to the market of interest, would be required for that matter.

The classical model dynamics were first described in [47] where the sticky-strike, sticky-delta and sticky-implied-tree are compared for different market regimes. In [34], the historical analysis performed on an index suggests that the market behaves with a dynamics in between sticky-strike and sticky-delta. In Chapter 9 of [14], an empirical analysis of equity smiles depict a market dynamics half-way between sticky-implied-tree and sticky-strike. We recall that sticky-strike means that the implied volatility is the same as a function of strike after a movement of the spot, the sticky-delta dynamics suggests that the implied volatility is fixed as a function of the log-moneyness $\ln(K/F)$ and finally, the sticky-implied-tree is the implied volatility generated by a local volatility model, without re-calibration, after a movement of the initial spot. The validity of the local volatility model dynamics, i.e. the sticky-implied-tree, is debated and as discussed in Chapter 2 of [14], practitioners usually prefer to re-calibrate the local volatility for a shifted initial spot using a sticky-strike dynamics, and match the Black–Scholes delta.

In this short analysis, we will shift the spot price upward by 1.26% to $\bar{S}_0 = 1.3$ and recompute the implied volatility surface for all models for a $2Y$ maturity. As explained in [13], for all stochastic

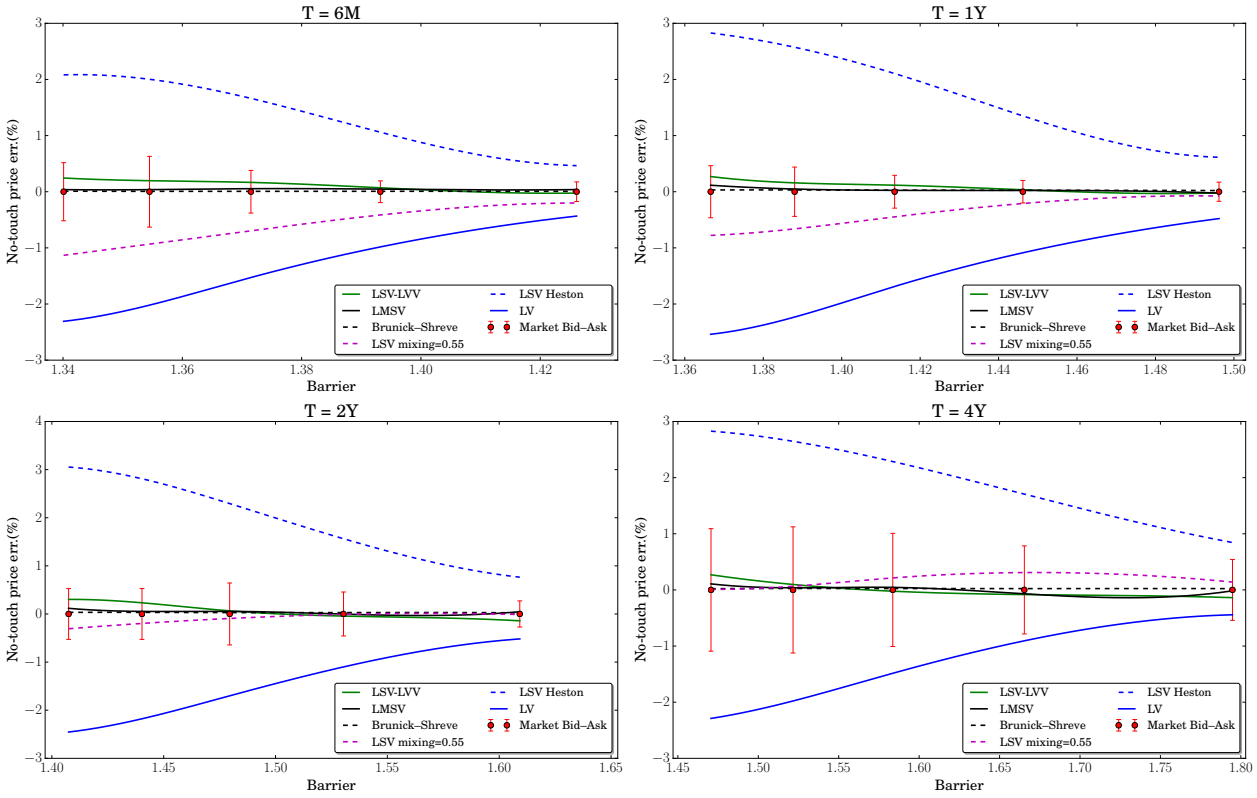


Figure 4.8.1: Calibration fit to foreign no-touch options for all models; see Table 4.8.2 for tabulated errors.

volatility models, the initial variance should be shifted following the instantaneous move in the initial spot, such that the new initial variance $\bar{v}_0$ is

$$\bar{v}_0 = v_0 + \frac{\rho\xi(S_0, 0)}{S\alpha_S(S_0, 0)}\left(\bar{S}_0 - S_0\right), \tag{4.8.1}$$

for the LSV-LVV model (4.2.7). This relates to the fact that the delta of an option, i.e. the amount of shares needed to fully hedge the randomness of the underlying asset, under a stochastic volatility model (4.2.7), carries an additional variance term such that

$$\Delta = \frac{\partial C}{\partial S} + \frac{\rho\xi(S,t)}{S\alpha_S(S,t)}\frac{\partial C}{\partial V}.$$

Since the running maximum does not introduce additional stochastic factor nor does have cross-variation with either S or V, the same hedging arguments are still valid for the LMSV model (4.2.6). We note that if the initial variance is not adjusted, a pure Heston model calibrated on the market volatility surface will have a sticky-delta dynamics as discussed in [47]. However, if the initial variance is adapted with (4.8.1), the Heston model at-the-money volatility will behave as sticky-implied-tree for

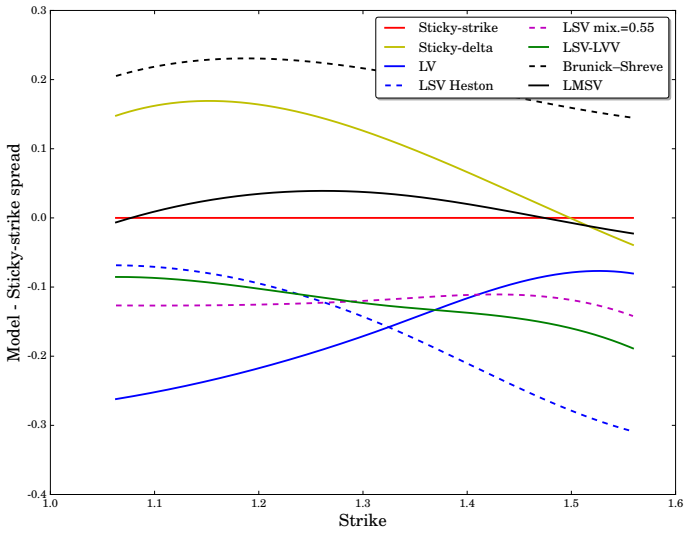


Figure 4.8.2: Implied volatility "model to sticky-strike spread" in % for an upward spot shift of 1.26% for all calibrated models with $T = 2Y$

short maturities and sticky-strike for long-term maturities as depicted in [13]. In the Brunick–Shreve and LMSV path-dependent models, we adapt the starting value of the running-maximum as well. In [9], it is mentioned that the addition of barrier options to the calibration basket should provide information about the smile dynamics. Furthermore, in [81], path-dependent volatility models are shown to be able to generate a very rich diversity of volatility surface dynamics when calibrated to historical data. In Figure 4.8.2, we graph the "model to sticky-strike spread" defined as

$$ms(K,T) = \bar{\sigma}_{model}(K,T) - \bar{\sigma}_{ss}(K,T)\,,$$

where $\bar{\sigma}_{model}(K,T)$ is the Black–Scholes model implied volatility after an upward spot shift of 1.26% and $\bar{\sigma}_{ss}(K,T)$ is the sticky-strike implied volatility (unchanged by the spot shift).

However close to it, the LV smile dynamics graphed in Figure 4.8.2 is not exactly sticky-implied-tree as we did not rely on the Dupire formula for the calibration. Additionally, the Heston LSV model, calibrated on vanilla options only, does not behave as the pure Heston model since the additional leverage function makes the model time inhomogeneous creating a smile dynamics harder to understand theoretically.

As suggested by Figure 4.8.2, the LMSV model (4.2.6) is the only model able to provide a smile dynamics that lies in between sticky-strike and sticky-delta. The LSV-LVV model exhibits an at-the-money volatility dynamics in between sticky-strike and the one generated by the Heston LSV model (calibrated on vanilla options only). Finally, the Brunick–Shreve model, calibrated as per our routine in Section 4.5, does not seem to provide a plausible smile dynamics. Additionally, our numerical tests show that the conclusions are consistent across all maturities.

4.9 Adding stochastic rates

Let us now consider a stochastic volatility model with stochastic rates. We recall that under $\mathbb{Q}^d$, the price process of a risky asset follows a system of SDEs as defined in (4.2.1)

$$\begin{cases} \dfrac{dS_t}{S_t} = \left(r_t^d - r_t^f\right) dt + \gamma_t \, dW_t \\ M_t = \max\limits_{0 \le u \le t} S_u \, . \end{cases} \tag{4.9.1}$$

We now state the extension of Corollary 4.4 to account for both domestic and foreign stochastic rates. We assume that the rates processes are integrable, and perfectly reproduce the zero-coupon bond market prices, under their respective domestic risk-neutral measure.

Theorem 4.7. *Assume that the marginal density $\phi(T, \cdot, \star)$ of (S_T, M_T) in (4.2.1) and*

$$\mathbb{E}^{\mathbb{Q}^d}\left[\gamma_T^2 \,|\, S_T = \cdot,\, M_T = \star\right],$$

are continuous. Then, the up-and-out call price $C_{UO}(K, B, T)$ under model (4.2.1) matches the price C_{UO}^{BSV} under the Brunick–Shreve volatility model (4.2.5) for any strike $K > 0$, barrier $B \ge (K \vee S_0)$ and maturity T if and only if, for all $K, T > 0$ and all $B \ge (K \vee S_0)$,

$$\begin{aligned} \frac{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \gamma_T^2 |\, S_T = K,\, M_T = B\right]}{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d |\, S_T = K,\, M_T = B\right]} &= \sigma_{BSV}^2(K, B, T) \\ &\quad + \frac{y(K, B, T)\phi_M(B, T) - \frac{K}{B^2}\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\left(r_T^f - \bar{r}_T^f\right)\mathbf{1}_{M_T < B} S_T\right]}{\frac{1}{2}K^2 \dfrac{\partial^3 C_{UO}^{BSV}(K, B, T)}{\partial K^2 \partial B}}, \end{aligned} \tag{4.9.2}$$

with

$$\phi_M(B, T) = \int_0^B \frac{\dfrac{\partial^3 C_{UO}^{BSV}(x, B, T)}{\partial K^2 \partial B}}{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d |\, S_T = x,\, M_T = B\right]} \, dx \tag{4.9.3}$$

$$\begin{aligned} y(K, B, T) &= \mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\left(r_T^f - \bar{r}_T^f\right)(S_T - K)^+ \,|\, M_T = B\right] \\ &\quad - K\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d \mathbf{1}_{S_T \ge K}\left(\left(r_T^d - \bar{r}_T^d\right) - \left(r_T^f - \bar{r}_T^f\right)\right) |\, M_T = B\right] \\ &\quad - \frac{(B - K)}{B}\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\left(r_T^f - \bar{r}_T^f\right) S_T \,|\, M_T = B\right], \end{aligned} \tag{4.9.4}$$

and where σ_{BSV}, $\bar{r}_T^d$ and $\bar{r}_T^f$ are deterministic as in (4.2.5), but r_T^d and r_T^f are stochastic with D_T^d defined in (4.2.3).

Proof. First, we want to ensure that (4.9.3) is a necessary condition for

$$C_{\mathrm{UO}}(K,B,T) = C_{\mathrm{UO}}^{BSV}(K,B,T)\,. \tag{4.9.5}$$

Let us denote both prices by the same quantity $C(K,B,T)$ for ease of notation. The proof of Proposition 4.2 holds with stochastic rates until equation (4.3.7) such that we can write

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d(S_T-K)^+\mathbf{1}_{M_T<B}\Big] &= (S_0-K)^+\mathbf{1}_{M_0<B} - \int_0^T \mathbb{E}^{\mathbb{Q}^d}\Big[r_t^f D_t^d(S_t-K)^+\mathbf{1}_{M_T<B}\Big]\,dt \\
&+ \int_0^T K\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d\Big(r_t^d-r_t^f\Big)\mathbf{1}_{M_t<B}\mathbf{1}_{S_t\geq K}\Big]\,dt \\
&+ \frac{1}{2}\int_0^T\int_{S_0\vee K}^B \Big(K^2\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d\gamma_t^2 \,|\, S_t=K,\, M_t=b\Big]\phi(K,b,t)\Big)\,dbdt \\
&- \frac{1}{2}\int_0^T (B-K)^+B^2\mathbb{E}^{\mathbb{Q}^d}\Big[D_t^d\gamma_t^2 \,|\, S_t=B,\, M_t=B\Big]\phi(B,B,t)\,dt\,.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\frac{\partial C(K,B,T)}{\partial B} &= \int_K^{+\infty}\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d|\, S_T=x,\, M_T=B\Big](x-K)\phi(x,B,T)\,dx \\
\frac{\partial^2 C(K,B,T)}{\partial K\partial B} &= -\int_K^{+\infty}\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d|\, S_T=x,\, M_T=B\Big]\phi(x,B,T)\,dx \\
\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B} &= \mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d|\, S_T=K,\, M_T=B\Big]\phi(K,B,T) \\
\int_0^B \frac{\partial^3 C(x,B,T)/\partial K^2\partial B}{\mathbb{E}^{\mathbb{Q}^d}[D_T^d|\, S_T=x,\, M_T=B]}\,dx &= \phi_M(B,T)
\end{aligned}$$

with ϕ_M the marginal-joint density of the running maximum at time T. Hence, by differentiating with respect to T and B,

$$\begin{aligned}
\frac{\partial C(K,B,T)}{\partial T\partial B} &= K\frac{\partial}{\partial B}\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d\mathbf{1}_{M_T\leq B}\mathbf{1}_{S_T\geq K}\Big(r_T^d-r_T^f\Big)\Big] - \frac{\partial}{\partial B}\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d r_T^f\mathbf{1}_{M_T<B}(S_T-K)^+\Big] \\
&+ \frac{1}{2}K^2\frac{\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d\gamma_T^2 \,|\, S_T=K,\, M_T=B\Big]}{\mathbb{E}^{\mathbb{Q}^d}[D_T^d|\, S_T=K,\, M_T=B]}\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B} \\
&- \frac{1}{2}\frac{\partial}{\partial B}\left[(B-K)^+B^2\frac{\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d\gamma_T^2 \,|\, S_T=K,\, M_T=B\Big]}{\mathbb{E}^{\mathbb{Q}^d}[D_T^d|\, S_T=K,\, M_T=B]}\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B}\right|_{K=B}\Bigg]\,.
\end{aligned} \tag{4.9.6}$$

Similarly, for a deterministic rates Brunick–Shreve model (4.2.5) we have

$$\begin{aligned}\frac{\partial C(K,B,T)}{\partial T\partial B} &= -\bar{r}_T^f\frac{\partial C(K,B,T)}{\partial B} - \left(\bar{r}_T^d - \bar{r}_T^f\right)K\frac{\partial C(K,B,T)}{\partial B\partial K} + \frac{1}{2}K^2\sigma_{BSV}^2(K,B,T)\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B}\\ &- \frac{1}{2}\frac{\partial}{\partial B}\left[(B-K)^+B^2\sigma_{BSV}^2(B,B,T)\left.\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B}\right|_{K=B}\right], \qquad (4.9.7)\end{aligned}$$

where C is the up-and out call price under both models, since, as a necessary condition (4.9.5), we want the price and all derivatives to match. By subtracting (4.9.6)–(4.9.7), we get

$$\begin{aligned}0 &= K\frac{\partial}{\partial B}\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\mathbf{1}_{M_T<B}\mathbf{1}_{S_T\geq K}\left(\left(r_T^d-\bar{r}_T^d\right)-\left(r_T^f-\bar{r}_T^f\right)\right)\right]\\ &- \frac{\partial}{\partial B}\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\left(r_T^f-\bar{r}_T^f\right)\mathbf{1}_{M_T<B}(S_T-K)^+\right]\\ &+ \frac{1}{2}K^2\left(p(K,B,T)-\sigma_{BSV}^2(K,B,T)\right)\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B} \qquad (4.9.8)\\ &- \frac{1}{2}\frac{\partial}{\partial B}\left[(B-K)^+B^2\left(p(B,B,T)-\sigma_{BSV}^2(B,B,T)\right)\left.\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B}\right|_{K=B}\right]\end{aligned}$$

with

$$p(K,B,T) = \frac{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\gamma_T^2\,|\,S_T=K,\,M_T=B\right]}{\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d|\,S_T=K,\,M_T=B\right]}.$$

At $K=0$, we get by integrating with respect to B, and multiplying by $B^{-1}(B-K)^+$,

$$\begin{aligned}\frac{(B-K)^+}{B}\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\left(r_T^f-\bar{r}_T^f\right)\mathbf{1}_{M_T<B}S_T\right] &= -\frac{1}{2}\frac{(B-K)^+}{B}B^3 \qquad (4.9.9)\\ &\times\left(p(B,B,T)-\sigma_{BSV}^2(B,B,T)\right)\left.\frac{\partial^3 C(K,B,T)}{\partial K^2\partial B}\right|_{K=B}.\end{aligned}$$

Furthermore,

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\left(r_T^f-\bar{r}_T^f\right)\mathbf{1}_{M_T<B}S_T\right] &= \int_{S_0}^B\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\left(r_T^f-\bar{r}_T^f\right)S_T\,|\,M_T=m\right]\phi_M(m,T)\,dm\\ \mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\mathbf{1}_{M_T<B}\mathbf{1}_{S_T\geq K}\Delta_T^r\right] &= \int_{S_0}^B\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\mathbf{1}_{S_T\geq K}\Delta_T^r\,|\,M_T=m\right]\phi_M(m,T)\,dm\\ \mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\left(r_T^f-\bar{r}_T^f\right)\mathbf{1}_{M_T<B}(S_T-K)^+\right] &= \int_{S_0}^B\mathbb{E}^{\mathbb{Q}^d}\left[D_T^d\left(r_T^f-\bar{r}_T^f\right)(S_T-K)^+\,|\,M_T=m\right]\phi_M(m,T)\,dm,\end{aligned}$$

where

$$\Delta_T^r = \left(r_T^d-\bar{r}_T^d\right)-\left(r_T^f-\bar{r}_T^f\right).$$

Hence, by differentiating (4.9.9) and combining with (4.9.8),

$$
\begin{aligned}
0 &= K\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d \mathbf{1}_{S_T \geq K}\Big(\Big(r_T^d - \bar{r}_T^d\Big) - \Big(r_T^f - \bar{r}_T^f\Big)\Big) \,|\, M_T = B\Big]\phi_M(B,T) \\
&- \Big(\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d\Big(r_T^f - \bar{r}_T^f\Big)(S_T - K)^+ \,|\, M_T = B\Big]\Big)\phi_M(B,T) \\
&+ \frac{(B-K)}{B}\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d\Big(r_T^f - \bar{r}_T^f\Big)S_T \,|\, M_T = B\Big]\phi_M(B,T) \\
&+ \frac{1}{2}K^2\Big(p(K,B,T) - \sigma_{BSV}^2(K,B,T)\Big)\frac{\partial^3 C(K,B,T)}{\partial K^2 \partial B} + \frac{K}{B^2}\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d\Big(r_T^f - \bar{r}_T^f\Big)\mathbf{1}_{M_T<B}S_T\Big].
\end{aligned}
$$

From this we can extract the value of $p(K,B,T)$ and conclude the proof for any $K \leq B$. It is easy to verify that (4.9.3) is also a sufficient condition for (4.9.5). □

Remark. Note that when $K = B$,

$$
\frac{\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d \gamma_T^2 |\, S_T = B,\, M_T = B\Big]}{\mathbb{E}^{\mathbb{Q}^d}\big[D_T^d |\, S_T = B,\, M_T = B\big]} = \frac{\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d\Big(r_T^f - \bar{r}_T^f\Big)\mathbf{1}_{M_T<B}S_T\Big]}{\frac{1}{2}B^3\left.\dfrac{\partial^3 C_B(K,b,T)}{\partial K^2 \partial B}\right|_{K=B,b=B}} + \sigma_{BSV}^2(B,B,T)\,. \tag{4.9.10}
$$

Regarding the use of formula (4.9.2) for $K = 0$, where the numerator vanishes, one could use $K = \epsilon$ for small ϵ and extrapolate flat or linearly to $K = 0$. Additionally, when rates are deterministic, we have $r_T^d = \bar{r}_T^d$ and $r_T^f = \bar{r}_T^f$ such that we recover the deterministic rates formula from Corollary 4.4.

It is then possible to add stochastic rates to an LMSV model and perform the calibration as in Section 4.6 by adjusting the projection formula to account for the rates. This requires an extension of the particle state space to include the stochastic rates as discussed in Chapter 3 for the vanilla case. A control variate particle method can be used to reduce the calibration time by extending Chapter 3. Similarly, one could add stochastic rates to the LSV-LVV model and use the calibration algorithm proposed in Section 4.7, where the leverage function (4.7.1) and projected volatility (4.7.2) would now be

$$
\alpha_S(K,T) = \sqrt{\frac{\mathbb{E}^{\mathbb{Q}^d}\big[D_T^d \,|\, S_T = K\big]}{\mathbb{E}^{\mathbb{Q}^d}\big[D_T^d V_T \,|\, S_T = K\big]}\Big(\sigma_{LV}(K,T)^2 + q(K,T)\Big)},
$$

with q as defined in Chapter 3, and

$$
\begin{aligned}
\sigma_{|S,M}^2(K,B,T) &= \alpha_S^2(K,T)\frac{\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d V_T |\, S_T = K,\, M_T = B\Big]}{\mathbb{E}^{\mathbb{Q}^d}\big[D_T^d |\, S_T = K,\, M_T = B\big]} \\
&- \frac{y(K,B,T)\phi_M(B,T) - \frac{K}{B^2}\mathbb{E}^{\mathbb{Q}^d}\Big[D_T^d\Big(r_T^f - \bar{r}_T^f\Big)\mathbf{1}_{M_T<B}S_T\Big]}{\frac{1}{2}K^2\dfrac{\partial^3 C_{\mathrm{UO}}^{BSV}(K,B,T)}{\partial K^2 \partial B}}.
\end{aligned}
$$

4.10 The Fokker–Planck equation for a two-dimensional diffusion with path-dependence

In this section, we discuss the solution of the Fokker–Planck PDE for the joint density of (S, M, V). First, we prove Theorem 4.6 which emphasises the importance of the boundary $S = M$. Secondly, we propose to solve the PDE with the finite element method and explain its advantages. Finally, we display some validation results and price barrier options by integrating the density function.

4.10.1 Proof of Theorem 4.6

Theorem 4.6 states that if ϕ belongs to $C^{2,1,2,1}(\mathcal{H} \times \mathbb{R}^+)$ and vanishes for $z \to +\infty$ and if we additionally assumes that g vanishes for $z = 0$, then ϕ is solution of the following Kolmogorov forward equation

$$\begin{cases} \frac{\partial \phi}{\partial t} + (\bar{r}^d(t) - \bar{r}^f(t))\frac{\partial (x\phi)}{\partial x} + \frac{\partial (f\phi)}{\partial z} - \frac{1}{2}\left(\frac{\partial^2 \left(x^2\sigma^2\phi\right)}{\partial x^2} + \frac{\partial^2 \left(g^2\phi\right)}{\partial z^2} + 2\frac{\partial^2 (\rho x \sigma g \phi)}{\partial x \partial z} \right) = 0, & (x,y,z) \in \mathcal{H},\, t > 0 \\ \frac{1}{2}\frac{\partial \left(g^2\phi\right)}{\partial z} - f\phi + \rho \frac{\partial (\sigma g x \phi)}{\partial x}\Big|_{z=0} = 0, & t > 0 \\ \bar{r}^d(t) - \bar{r}^f(t))y\phi\Big|_{x=y} - \frac{1}{2}\left(\frac{\partial \left(\sigma^2 x^2 \phi\right)}{\partial y}\Big|_{x=y} + 2\,\frac{\partial \left(\sigma^2 x^2 \phi\right)}{\partial x}\Big|_{x=y} + 2\,\frac{\partial (\rho x \sigma g \phi)}{\partial z}\Big|_{x=y} \right) = 0, & x = y,\, t > 0 \\ \phi(S_0, S_0, z, t) = 0, & z \in [0, +\infty[,\, t > 0 \\ \lim_{z\to\infty} \phi(x,y,z,t) = 0, & (x,y) \in \Omega, t > 0 \\ \lim_{t\to 0} \phi(x,y,z,t) = \delta(x - S_0, y - S_0, z - v_0), & (x,y,z) \in \mathcal{H} \end{cases} \tag{4.10.1}$$

Proof. Let $v \in C^{2,1,2}(\mathcal{H})$, assume that v and its derivatives vanish for $x = 0$ and denote $r(t) - q(t)$ by $\mu(t)$.

We apply the Itô-Doeblin lemma and get

$$\begin{aligned} v(S_T, M_T, V_T) &= v(S_0, M_0, V_0) + \int_0^T \frac{\partial v(S_t, M_t, V_t)}{\partial x}\, dS_t \\ &+ \int_0^T \frac{\partial v(S_t, M_t, V_t)}{\partial y}\, dM_t + \int_0^T \frac{\partial v(S_t, M_t, V_t)}{\partial z}\, dV_t \\ &+ \int_0^T \frac{1}{2}\frac{\partial^2 v(S_t, M_t, V_t)}{\partial x^2}\, dS_t dS_t + \int_0^T \frac{1}{2}\frac{\partial^2 v(S_t, M_t, V_t)}{\partial z^2}\, dV_t dV_t \\ &+ \int_0^T \frac{\partial^2 v(S_t, M_t, V_t)}{\partial x \partial z}\, dS_t dV_t\,, \end{aligned}$$

which gives,

$$\begin{aligned}
v(S_T, M_T, V_T) &= v(S_0, M_0, V_0) + \int_0^T \frac{\partial v(S_t, M_t, V_t)}{\partial x} S_t \sigma(S_t, M_t, V_t, t)\, dW_t \\
&+ \int_0^T \frac{\partial v(S_t, M_t, V_t)}{\partial y}\, dM_t + \int_0^T \frac{\partial v(S_t, M_t, V_t)}{\partial z} g(S_t, V_t, t)\, dW_t^V \\
&+ \int_0^T \left(\frac{\partial v(S_t, M_t, V_t)}{\partial x} \mu(t) S_t + \frac{\partial v(S_t, M_t, V_t)}{\partial z} f(S_t, V_t, t) \right) dt \\
&+ \int_0^T \frac{1}{2} \left(\frac{\partial^2 v(S_t, M_t, V_t)}{\partial x^2} S_t^2 \sigma^2(S_t, M_t, V_t, t) \right) dt \\
&+ \int_0^T \frac{1}{2} \frac{\partial^2 v(S_t, M_t, V_t)}{\partial z^2} g^2(S_t, V_t, t)\, dt \\
&+ \int_0^T \frac{\partial^2 v(S_t, M_t, V_t)}{\partial x \partial z} \rho S_t \sigma(S_t, M_t, V_t, t) g(S_t, V_t, t)\, dt.
\end{aligned}$$

By taking expectations we can write

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[v(S_T, M_T, V_T)] &= v(S_0, M_0, V_0) \\
&+ \mathbb{E}^{\mathbb{Q}}\left[\int_0^T \left(\frac{\partial v(S_t, M_t, V_t)}{\partial x} \mu(t) S_t + \frac{\partial v(S_t, M_t, V_t)}{\partial z} f(S_t, V_t, t) \right) dt \right] \\
&+ \mathbb{E}^{\mathbb{Q}}\left[\int_0^T \left(\frac{1}{2} \frac{\partial^2 v(S_t, M_t, V_t)}{\partial x^2} S_t^2 \sigma^2(S_t, M_t, V_t, t) + \frac{1}{2} \frac{\partial^2 v(S_t, M_t, V_t)}{\partial z^2} g^2(S_t, V_t, t) \right) dt \right] \\
&+ \mathbb{E}^{\mathbb{Q}}\left[\int_0^T \frac{\partial^2 v(S_t, M_t, V_t)}{\partial x \partial z} \rho S_t \sigma(S_t, M_t, V_t, t) g(S_t, V_t, t)\, dt \right] \\
&+ \mathbb{E}^{\mathbb{Q}}\left[\int_0^T \frac{\partial v(S_t, M_t, V_t)}{\partial y}\, dM_t \right],
\end{aligned}$$

since $\mathbb{E}^{\mathbb{Q}}\left[\int_0^T \left(\left(\frac{\partial v}{\partial x} S_t \sigma \right)^2 + \left(\frac{\partial v}{\partial z} g \right)^2 \right) dt \right] < \infty$. Therefore, we can introduce the joint density function ϕ of (S, M, V) at time T, differentiate with respect to T, and write for all $t > 0$

$$\begin{aligned}
\int_{s_0}^{\infty} dy \int_0^y dx \int_0^{\infty} v \frac{\partial \phi}{\partial t}\, dz &= \int_0^{\infty} I_\mu(y)\, dy + \frac{1}{2} \int_0^{\infty} I_\sigma(y)\, dy + \int_0^{\infty} I_\rho(y)\, dy, \qquad (4.10.2) \\
&+ \frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}^{\mathbb{Q}}\left[\int_0^t \frac{\partial v(S_u, M_u, V_u)}{\partial y}\, dM_u \right],
\end{aligned}$$

where explicitly with all arguments,

$$\begin{aligned}
I_\sigma(y) &= \int_0^y dx \int_0^\infty \left(\sigma^2(x,y,z,t) x^2 \frac{\partial^2 v(x,y,z)}{\partial x^2} + g^2(x,z,t) \frac{\partial^2 v(x,y,z)}{\partial z^2} \right) \phi(x,y,z,t)\, dz \\
I_\mu(y) &= \int_0^y dx \int_0^\infty \left(\mu(t) x \frac{\partial v(x,y,z)}{\partial x} + f(x,z,t) \frac{\partial v(x,y,z)}{\partial z} \right) \phi(x,y,z,t)\, dz \\
I_\rho(y) &= \int_0^y dx \int_0^\infty \rho\sigma(x,y,z,t) g(x,z,t) x \frac{\partial^2 v(x,y,z)}{\partial x \partial z} \phi(x,y,z,t)\, dz\,.
\end{aligned}$$

For I_σ, we perform integration by parts twice to get

$$\begin{aligned}
I_\sigma(y) &= \int_0^\infty \left(\frac{\partial v}{\partial x} \sigma^2 y^2 \phi \Big|_{x=y} - \int_0^y \frac{\partial v}{\partial x} \frac{\partial \sigma^2 x^2 \phi}{\partial x}\, dx \right) dz \\
&- \int_0^y \left(\frac{\partial v}{\partial z} g^2 \phi \Big|_{z=0} + \int_0^\infty \frac{\partial v}{\partial z} \frac{\partial g^2 \phi}{\partial z}\, dz \right) dx \\
&= \int_0^\infty \left(\frac{\partial v}{\partial x} \sigma^2 y^2 \phi \Big|_{x=y} - v \frac{\partial \sigma^2 x^2 \phi}{\partial x} \Big|_{x=y} \right) dz \\
&+ \int_0^y \left(- \frac{\partial v}{\partial z} g^2 \phi \Big|_{z=0} + v \frac{\partial g^2 \phi}{\partial z} \Big|_{z=0} \right) dx, \\
&+ \int_0^\infty dz \int_0^y v \left(\frac{\partial^2 \sigma^2 x^2 \phi}{\partial x^2} + \frac{\partial^2 g^2 \phi}{\partial z^2} \right) dx\,.
\end{aligned}$$

Let

$$I_{\sigma,1}(t) = \int_0^\infty dz \int_{S_0}^\infty \left(\frac{\partial v}{\partial x} \sigma^2 y^2 \phi \Big|_{x=y} - v \frac{\partial \sigma^2 x^2 \phi}{\partial x} \Big|_{x=y} \right) dy,$$

and since

$$\frac{\partial v(x,y,z)}{\partial x} \Big|_{x=y} = \frac{\mathrm{d} v(y,y,z)}{\mathrm{d} y} - \frac{\partial v(x,y,z)}{\partial y} \Big|_{x=y},$$

we write

$$\begin{aligned}
I_{\sigma,1}(t) &= \int_0^\infty dz \int_{S_0}^\infty \frac{\mathrm{d} v(y,y,z)}{\mathrm{d} y} \sigma^2 y^2 \phi \Big|_{x=y} dy - \int_0^\infty dz \int_{S_0}^\infty \frac{\partial v}{\partial y} \sigma^2 y^2 \phi \Big|_{x=y} dy \\
&- \int_0^\infty dz \int_{S_0}^\infty v \frac{\partial \sigma^2 x^2 \phi}{\partial x} \Big|_{x=y} dy\,,
\end{aligned}$$

and integrate by part once with respect to y,

$$\begin{aligned} I_{\sigma,1}(t) &= -\int_0^\infty dz \int_{S_0}^\infty v(y,y,z)\left(\frac{\partial \sigma^2 x^2 \phi}{\partial x}\Big\rfloor_{x=y} + \frac{\mathrm{d}\,\sigma^2 y^2 \phi\rfloor_{x=y}}{\mathrm{d}y} \right) dy \\ &- \int_0^\infty dz \int_{S_0}^\infty \frac{\partial v}{\partial y}\sigma^2 y^2 \phi\Big\rfloor_{x=y} dy \\ &- \int_0^\infty v(S_0,S_0,z,t)\sigma^2(S_0,S_0,z,t)S_0^2\phi(S_0,S_0,z,t)\,dz\,. \end{aligned}$$

Moreover,

$$\begin{aligned} \int_{S_0}^\infty I_\sigma(y)\,dy &= I_{\sigma,1}(t) - \int_{S_0}^\infty dy \int_0^y \left(\frac{\partial v}{\partial z} g^2\phi\Big\rfloor_{z=0} - v\frac{\partial g^2\phi}{\partial z}\Big\rfloor_{z=0} \right) dx \\ &+ \int_0^\infty dz \int_{S_0}^\infty dy \int_0^y v\left(\frac{\partial^2 \sigma^2 x^2 \phi}{\partial x^2} + \frac{\partial^2 g^2 \phi}{\partial z^2} \right) dx\,. \end{aligned}$$

For I_ρ, we can perform integration by parts twice,

$$\begin{aligned} I_\rho(y) &= \int_0^y dx \int_0^\infty \rho\sigma g x \frac{\partial^2 v}{\partial x \partial z}\phi\,dz \\ &= -\int_0^\infty dz \int_0^y \frac{\partial \rho\sigma g x\phi}{\partial z}\frac{\partial v}{\partial x}\,dx - \int_0^y \rho\,\sigma g x\phi \frac{\partial v}{\partial x}\Big\rfloor_{z=0} dx \\ &= -\int_0^\infty v\frac{\partial \rho\sigma g x\phi}{\partial z}\Big\rfloor_{x=y} dz + \int_0^y \rho v \frac{\partial \sigma g x\phi}{\partial x}\Big\rfloor_{z=0} dx \\ &- \rho\sigma g y \phi v\rfloor_{x=y,\,z=0} + \int_0^y dx \int_0^\infty v\frac{\partial^2 \rho\sigma g x\phi}{\partial x \partial z}\,dz\,. \end{aligned}$$

For I_μ, we integrate by parts once,

$$\begin{aligned} I_\mu(y) &= \mu\int_0^\infty dz \int_0^y x\frac{\partial v}{\partial x}\phi\,dx + \int_0^y dx \int_0^\infty f\frac{\partial v}{\partial z}\phi\,dz \\ &= \mu\int_0^\infty vy\phi\rfloor_{x=y}\,dz - \int_0^y vf\phi\rfloor_{z=0}\,dx \\ &- \int_0^\infty dz \int_0^y v\Big(\mu\frac{\partial x\phi}{\partial x} + \frac{\partial f\phi}{\partial z}\Big)\,dx\,. \end{aligned}$$

Let us assume that v and its derivatives vanish when $z = 0$, then (4.10.2) is

$$\begin{aligned}
\int_{S_0}^{\infty} dy \int_0^y dx \int_0^{\infty} dz\, v \frac{\partial \phi}{\partial t} &= \int_0^{\infty} dy \int_0^{\infty} \mu v y \phi \big\rfloor_{x=y} dz \\
&- \int_{S_0}^{\infty} dy \int_0^y dx \int_0^{\infty} v \left(\mu \frac{\partial x \phi}{\partial x} + \frac{\partial f \phi}{\partial z} \right) dz \\
&- \frac{1}{2} \int_{S_0}^{\infty} dy \int_0^{\infty} v(y, y, z) \left(\frac{\partial \sigma^2 x^2 \phi}{\partial x} \Big\rfloor_{x=y} + \frac{\mathrm{d}\, \sigma^2 y^2 \phi \big\rfloor_{x=y}}{\mathrm{d}y} \right) dz \\
&- \frac{1}{2} \int_0^{\infty} dz \int_{S_0}^{\infty} \frac{\partial v}{\partial y} \sigma^2 y^2 \phi \Big\rfloor_{x=y} dy \\
&- \frac{1}{2} \int_0^{\infty} v(S_0, S_0, z) \sigma^2(S_0, S_0, z, t) S_0^2 \phi(S_0, S_0, z, t)\, dz \\
&+ \frac{1}{2} \int_{S_0}^{\infty} dy \int_0^y dx \int_0^{\infty} v \left(\frac{\partial^2 \sigma^2 x^2 \phi}{\partial x^2} + \frac{\partial^2 g^2 \phi}{\partial z^2} \right) dz \\
&- \int_0^{\infty} dy \int_0^{\infty} \rho v \frac{\partial \sigma g y \phi}{\partial z} \Big\rfloor_{x=y} dz \\
&+ \int_0^{\infty} dy \int_0^y dx \int_0^{\infty} v \frac{\partial^2 \rho \sigma g x \phi}{\partial x \partial z}\, dz \\
&+ \frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}^{\mathbb{Q}} \left[\int_0^t \frac{\partial v(S_u, M_u, V_u)}{\partial y} dM_u \right].
\end{aligned}$$

By rearranging and noting that

$$\frac{\mathrm{d}\, \sigma^2 y^2 \phi \big\rfloor_{x=y}}{\mathrm{d}y} = \frac{\partial \sigma^2 x^2 \phi}{\partial x} \Big\rfloor_{x=y} + \frac{\partial \sigma^2 x^2 \phi}{\partial y} \Big\rfloor_{x=y},$$

we get

$$\begin{aligned}
\int_{S_0}^{\infty} dy \int_0^y dx \int_0^{\infty} v \left[\frac{\partial \phi}{\partial t} + \mu \frac{\partial x \phi}{\partial x} + \frac{\partial f \phi}{\partial z} - \frac{1}{2} \left(\frac{\partial^2 \sigma^2 x^2 \phi}{\partial x^2} + \frac{\partial^2 g^2 \phi}{\partial z^2} + 2 \frac{\partial^2 \rho \sigma g x \phi}{\partial x \partial z} \right) \right] dz = \qquad (4.10.3) \\
\int_{S_0}^{\infty} dy \int_0^{\infty} dz\, v(y, y, z) \left[\mu y \phi - \frac{1}{2} \left(\frac{\partial \sigma^2 x^2 \phi}{\partial y} + 2 \frac{\partial \sigma^2 x^2 \phi}{\partial x} + 2 \frac{\partial \rho \sigma g x \phi}{\partial z} \right) \right]_{x=y} - \\
\frac{1}{2} \int_0^{\infty} v(S_0, S_0, z) \sigma^2(S_0, S_0, z, t) S_0^2 \phi(S_0, S_0, z, t)\, dz - \\
\int_0^{\infty} dz \int_{S_0}^{\infty} \frac{\partial v}{\partial y} \sigma^2 y^2 \phi \Big\rfloor_{x=y} dy + \\
\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}^{\mathbb{Q}} \left[\int_0^t \frac{\partial v(S_u, M_u, V_u)}{\partial y} dM_u \right].
\end{aligned}$$

Let us additionally assume that v and its derivatives vanish for $x = y$. Since M only grows when

$M = S$, the term

$$\int_0^t \frac{\partial v(S_u, M_u, V_u)}{\partial y} dM_u \,,$$

is zero. Then (4.10.3) becomes

$$\int_{S_0}^{\infty} dy \int_0^y dx \int_0^{\infty} v \left[\frac{\partial \phi}{\partial t} + \mu \frac{\partial x \phi}{\partial x} + \frac{\partial f \phi}{\partial z} - \frac{1}{2} \left(\frac{\partial^2 \sigma^2 x^2 \phi}{\partial x^2} + \frac{\partial^2 g^2 \phi}{\partial z^2} + 2 \frac{\partial^2 \rho \sigma g x \phi}{\partial x \partial z} \right) \right] dz = 0 \,.$$

And we can conclude that for all $(x, y, z, t) \in \mathcal{H} \times \mathbb{R}_*^+$,

$$\frac{\partial \phi}{\partial t} + \mu \frac{\partial x \phi}{\partial x} + \frac{\partial f \phi}{\partial z} - \frac{1}{2} \left(\frac{\partial^2 \sigma^2 x^2 \phi}{\partial x^2} + \frac{\partial^2 g^2 \phi}{\partial z^2} + 2 \frac{\partial^2 \rho \sigma g x \phi}{\partial x \partial z} \right) = 0 \,.$$

Let us now consider all functions v which do not vanish at $x = y$ but only depend on the space variables x and z such that we define $a(x, z) = v(x, y, z)$ with a in $C^{2,2}\Big(\big(\mathbb{R}^+\big)^2\Big)$. In that case, we write

$$\int_0^t \frac{\partial v}{\partial y} dM_u = \frac{\partial v}{\partial y} \sigma^2 y^2 \phi \Big\rfloor_{x=y} = 0 \,.$$

From (4.10.3) and if we also impose a to vanish at $x = S_0$, we conclude

$$\mu y \phi(y, y, z, t) \quad = \quad \left(\frac{1}{2} \frac{\partial \sigma^2 x^2 \phi}{\partial y} + \frac{\partial \sigma^2 x^2 \phi}{\partial x} + \frac{\partial \rho \sigma g x \phi}{\partial z} \right) \Big\rfloor_{x=y} .$$

Furthermore, for all function a which do not vanish at $x = S_0$, we are left, for any $z > 0$, with

$$a(S_0, z) \sigma^2(S_0, S_0, z, t) S_0^2 \phi(S_0, S_0, z, t) = 0 \,,$$

which gives the Dirichlet boundary condition $\phi(S_0, S_0, z, t) = 0$.

Now if v does not vanishes at $z = 0$ but is such that $\frac{\partial v}{\partial z}\Big\rfloor_{z=0} = 0$, and vanishes for $x = y$, we can conclude that

$$\frac{1}{2} \frac{\partial g^2 \phi}{\partial z} \Big\rfloor_{z=0} - f \phi \rfloor_{z=0} + \rho \frac{\partial \sigma g x \phi}{\partial x} \Big\rfloor_{z=0} = 0.$$

If $\frac{\partial v}{\partial z}$ does not vanish but v still vanishes at $x = y$ we get

$$\frac{\partial v}{\partial z} g^2 \phi \Big\rfloor_{z=0} = 0. \tag{4.10.4}$$

And finally, when v does not vanish at $x = y$ we get

$$\rho \sigma(y, y, 0, t) g(y, 0, t) y \phi(y, y, 0, t) = 0 \,,$$

which adds the condition

$$g(\cdot, 0, t) = 0\,.$$

The latter also makes (4.10.4) satisfied if g^2 vanishes fast enough if ϕ diverges at 0. For instance, in the Heston model setting, we can assume, as explained in Chapter 3, that

$$\phi \sim z^{\frac{2\kappa\theta}{\xi^2}-1}\,,$$

and $g = \xi\sqrt{z}$ such that

$$g^2\phi \sim \xi^2 z^{\frac{2\kappa\theta}{\xi^2}}\,,$$

which converges to zero when z goes to zero.

□

4.10.2 Finite element solution

In this section we discuss the solution to the Fokker–Planck PDE presented in Theorem 4.6, using a three-dimensional finite element method. The main advantage of finite elements compared to finite differences is the ability to handle the Robin boundary equation for $x = y$,

$$(r(t) - q(t))y\phi - \frac{1}{2}\left(\frac{\partial x^2\alpha_S^2(t,x)z\phi}{\partial y} + 2\frac{\partial x^2\alpha_S^2(t,x)z\phi}{\partial x} + 2\frac{\partial \rho\xi x\alpha_S(t,x)z\phi}{\partial z}\right) = 0\,.$$

4.10.2.1 Weak formulation

In order to write the variational form of the problem, we use the divergence theorem. Recall that the PDE in the interior of the domain is defined as the classical two-dimensional Fokker–Planck equation

$$\frac{\partial\phi}{\partial t} + (r(t) - q(t))\frac{\partial x\phi}{\partial x} + \frac{\partial f\phi}{\partial z} - \frac{1}{2}\left(\frac{\partial^2 x^2\sigma^2\phi}{\partial x^2} + \frac{\partial^2 g^2\phi}{\partial z^2} + 2\frac{\partial^2 \rho x\sigma g\phi}{\partial x\partial z}\right) = 0\,,$$

which we reformulate as

$$\begin{aligned} 0 = \frac{\partial\phi}{\partial t} - \frac{1}{2}\left(\frac{\partial\left(\frac{\partial x^2\sigma^2\phi}{\partial x} + \frac{\partial\rho x\sigma g\phi}{\partial z}\right)}{\partial x} + \frac{\partial\left(\frac{\partial g^2\phi}{\partial z} + \frac{\partial\rho x\sigma g\phi}{\partial x}\right)}{\partial z}\right) \\ + (r(t) - q(t))x\frac{\partial\phi}{\partial x} + f\frac{\partial\phi}{\partial z} + \phi\left((r(t) - q(t)) + \frac{\partial f}{\partial z}\right). \end{aligned}$$

In vector form we write

$$\frac{\partial\phi}{\partial t} - \nabla\cdot\mathbf{u} + \mathbf{b}\cdot\nabla\phi + \mu_t\phi = 0\,,$$

with

$$\mathbf{u} = \frac{1}{2}\begin{bmatrix} \frac{\partial x^2\sigma^2\phi}{\partial x} + \frac{\partial \rho x\sigma g\phi}{\partial z} \\ 0 \\ \frac{\partial g^2\phi}{\partial z} + \frac{\partial \rho x\sigma g\phi}{\partial x} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} (r(t)-q(t))x \\ 0 \\ f \end{bmatrix}, \quad \mu_t = (r(t)-q(t)) + \frac{\partial f}{\partial z}\,.$$

Denote:

- $\Omega = \{(x,y,z) \in \mathbb{R}^3 : 0 < x < y,\, y > S_0,\, z \geq 0\}$ the support of the solution.
- $\partial\Omega$ the boundary of Ω, where Γ_D is the subset where a Dirichlet boundary condition is given and $\Gamma_{R1} = \{(x,y,z) \in \partial\Omega : x = y\}$ and $\Gamma_{R2} = \{(x,y,z) \in \partial\Omega : z = 0\}$ the subset of Robin boundary conditions, with $\partial\Omega = \Gamma_D \cup \Gamma_{R1} \cup \Gamma_{R2}$.

The divergence theorem gives for any test function v

$$\int_\Omega (\nabla\cdot\mathbf{u})v\, d\Omega + \int_\Omega \mathbf{u}\cdot\nabla v\, d\Omega = \int_{\partial\Omega} (\mathbf{u}\cdot\mathbf{n})v\, d(\partial\Omega)\,,$$

where $\mathbf{n}$ is the unit vector pointing outward of $\partial\Omega$ at each point. On the whole boundary surface,

$$\int_{\partial\Omega} (\mathbf{u}\cdot\mathbf{n})v\, d(\partial\Omega) = \int_{\Gamma_{R1}\cup\Gamma_{R2}} (\mathbf{u}\cdot\mathbf{n})v\, d(\Gamma_{R1}\cup\Gamma_{R2})\,,$$

since we force homogeneous Dirichlet boundary conditions on v over Γ_D. There is a Robin boundary condition over Γ_{R1} and Γ_{R2} such that $\int_{\partial\Omega}(\mathbf{u}\cdot\mathbf{n})v\, d(\partial\Omega)$ reduces to

$$\int_{\Gamma_{R1}} \frac{1}{2\sqrt{2}}\left(\frac{\partial x^2\sigma^2\phi}{\partial x} + \frac{\partial \rho x\sigma g\phi}{\partial z}\right) v d\Gamma_{R1} - \int_{\Gamma_{R2}} \frac{1}{2}\left(\frac{\partial g^2\phi}{\partial z} + \frac{\partial \rho x\sigma g\phi}{\partial x}\right) v d\Gamma_{R2}\,,$$

since the outward vector of the surface $x = y$ is $\mathbf{n}_1 = \frac{1}{\sqrt{2}}(1,-1,0)^T$ and the vector pointing outward of the surface $z = 0$ is $\mathbf{n}_2 = (0,0,-1)^T$. Furthermore, from the boundary conditions,

$$\begin{aligned} (r(t)-q(t))y\phi\big\rfloor_{x=y} - \frac{1}{2}\frac{\partial\sigma^2x^2\phi}{\partial y}\bigg\rfloor_{x=y} &= \left(\frac{\partial x^2\sigma^2\phi}{\partial x} + \frac{\partial\rho x\sigma g\phi}{\partial z}\right)_{(x,y,z)\in\Gamma_{R1}} \\ f\phi\big\rfloor_{z=0} - \frac{1}{2}\frac{\partial\rho\sigma gx\phi}{\partial x}\bigg\rfloor_{z=0} &= \frac{1}{2}\left(\frac{\partial g^2\phi}{\partial z} + \frac{\partial\rho\sigma gx\phi}{\partial x}\right)_{(x,y,z)\in\Gamma_{R2}}. \end{aligned}$$

Hence, we multiply the PDE by a test function v and integrate over Ω. Using the divergence theorem, the equation becomes

$$\begin{aligned}\int_\Omega \frac{\partial \phi}{\partial t} v \, d\Omega + \int_\Omega \mathbf{u}\cdot\nabla v \, d\Omega + \int_\Omega (\mathbf{b}\cdot\nabla\phi) v \, d\Omega + \int_\Omega \mu_t \phi v \, d\Omega \quad &- \\ \int_{\Gamma_{R1}} \left(\frac{(r(t)-q(t))y\phi}{2\sqrt{2}} - \frac{1}{4\sqrt{2}} \frac{\partial \sigma^2 x^2 \phi}{\partial y} \right) v \, d\Gamma_{R1} \quad &+ \\ \int_{\Gamma_{R2}} \left(f\phi\rfloor_{z=0} - \frac{1}{2} \frac{\partial \rho\sigma g x \phi}{\partial x}\Big|_{z=0} \right) v \, d\Gamma_{R2} \quad &= \quad 0\,.\end{aligned}$$

We write

$$\mathbf{u} = \frac{1}{2} \begin{bmatrix} \frac{\partial x^2\sigma^2\phi}{\partial x} + \frac{\partial \rho x \sigma g \phi}{\partial z} \\ 0 \\ \frac{\partial g^2 \phi}{\partial z} + \frac{\partial \rho x \sigma g \phi}{\partial x} \end{bmatrix} = \nabla.(\mathbf{A}\phi)\,, \quad \mathbf{A} = \frac{1}{2} \begin{bmatrix} x^2\sigma^2 & 0 & \rho x \sigma g \\ 0 & 0 & 0 \\ \rho x \sigma g & 0 & g^2 \end{bmatrix},$$

and denote

$$\mathbf{b}^{\Gamma_1} \quad = \quad \begin{bmatrix} 0 \\ \frac{1}{4\sqrt{2}}\sigma^2 y^2 \\ 0 \end{bmatrix}, \quad \mathbf{b}^{\Gamma_2} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2}\rho x \sigma g \end{bmatrix}, \quad \mu_t^{\Gamma_1} = \frac{(r(t)-q(t))y}{2\sqrt{2}}.$$

Finally,

$$\begin{aligned}\int_\Omega \frac{\partial \phi}{\partial t} v \, d\Omega + \int_\Omega (\nabla.(\mathbf{A}\phi))\cdot\nabla v \, d\Omega + \int_\Omega (\mathbf{b}\cdot\nabla\phi) v \, d\Omega + \int_\Omega \mu_t v \phi \, d\Omega \quad &- \\ \int_{\Gamma_{R1}} \Big(\mu_t^{\Gamma_1}\phi - \nabla.\Big(\mathbf{b}^{\Gamma_1}\phi\Big) \Big) v \, d\Gamma_{R1} \quad &+ \\ \int_{\Gamma_{R2}} \Big(f\phi - \nabla.\Big(\mathbf{b}^{\Gamma_2}\phi\Big) \Big) v \, d\Gamma_{R2} \quad &= \quad 0\,.\end{aligned} \tag{4.10.5}$$

Equation (4.10.5) justifies the definition of a Sobolev space $H^1(\Omega)$ of functions u

$$H^1(\Omega) = \Big\{ u \in L^2(\Omega) : (\partial_x u, \partial_y u, \partial_z u) \in \Big(L^2(\Omega)\Big)^3 \Big\},$$

equipped with the natural norm

$$\|u\|_{1,\Omega} = \sqrt{\int_\Omega |u|^2 + \int_\Omega |\nabla u|^2}\,,$$

where partial derivatives are taken in a weak sense. And we additionally define

$$H_0^1(\Omega) = \Big\{ v \in H^1 \,|\, \forall x \in \Gamma_D,\, v(x) = 0 \Big\}.$$

We can reformulate (4.10.5) as

$$\begin{aligned}\frac{d}{dt}\langle\phi, v\rangle + a(\phi, v) &= 0\,,\\ a(u, v) &= \int_\Omega (\nabla.(\mathbf{A}u))\cdot\nabla v\, d\Omega + \int_\Omega (\mathbf{b}\cdot\nabla u)v\, d\Omega + \int_\Omega \mu_t v\phi\, d\Omega\\ &- \int_{\Gamma_{R1}} \Big(\mu_t^{\Gamma_1} u - \nabla.\Big(\mathbf{b}^{\Gamma_1} u\Big)\Big) v\, d\Gamma_{R1}\\ &+ \int_{\Gamma_{R2}} \Big(f u - \nabla.\Big(\mathbf{b}^{\Gamma_2} u\Big)\Big) v\, d\Gamma_{R2}\,,\end{aligned}$$

where $\langle ., .\rangle$ is the standard L^2 inner product on Ω. The additional boundary terms

$$\int_{\Gamma_{R1}} \Big(\mu_t^{\Gamma_1} u - \nabla.\Big(\mathbf{b}^{\Gamma_1} u\Big)\Big) v\, d\Gamma_{R1}\,,$$

preclude from stating if the bilinear form a is coercive. It may still be possible to get error bounds for the finite element solution, although less sharp than usual, but we do not pursue this here.

4.10.2.2 Order 2 BDF scheme with variable step size

The time discretisation will be handled by a Backward Differentiation Formula (BDF) scheme. Let us define a time discretisation $(t_m)_{m\in[\![0,M]\!]}$ with $\Delta_t(m) = t_{m+1} - t_m$. Let

$$\begin{aligned}\left(\int_\Omega \phi_{t_m} v\, d\Omega + \frac{\Delta_t(m)}{1+\gamma_m} a(\phi_{t_m}, v)\right) &= \left(1 + \frac{\Delta_t(m)}{\Delta_t(m-1)}\frac{\gamma_m}{1+\gamma_m}\right)\int_\Omega \phi_{t_{m-1}} v\, d\Omega\\ &- \left(\frac{\Delta_t(m)}{\Delta_t(m-1)}\frac{\gamma_m}{1+\gamma_m}\right)\int_\Omega \phi_{t_{m-2}} v\, d\Omega\,,\end{aligned}$$

with

$$\gamma_m = \frac{\Delta_t(m)}{\Delta_t(m) + \Delta_t(m-1)}\,.$$

The first two time steps are full-implicit. This scheme is of order two in time and exhibits stability advantages compared to Crank–Nicolson. The latter did not show satisfactory results in our tests, even coupled with Rannacher time steps. Furthermore, we would like to ensure that the function found is indeed a density function. Therefore, we want to impose that

1. $\int_\Omega \phi\, d\Omega = 1\,.$
2. $\forall\omega\in\Omega,\ \phi(\omega) \geq 0\,.$

This could be performed by adding a penalisation of the form $\lambda(1 - \int_\Omega \phi\, d\Omega) + \mu\max(-\phi, 0)$ in the initial equation. The term is then multiplied by the test function and integrated leading to an equation

of the form

$$\int_\Omega \frac{\partial \phi}{\partial t} v \, d\Omega + a(\phi, v) + \lambda \int_\Omega \Big(1 - \int_\Omega \phi \, d\Omega\Big) v \, d\Omega + \mu \int_\Omega \max(-\phi, 0) v \, d\Omega = 0.$$

However, the equation is now non-linear and would necessitate a fixed-point routine which would be computationally intensive. Therefore, we overcome this difficulty by setting, at each time step,

$$\begin{aligned} \phi_m(\omega) &\leftarrow max(\phi_m(\omega), 0) \\ \phi_m(\omega) &\leftarrow \frac{\phi_m(\omega)}{\int_\Omega \phi_m \, d\Omega} \, . \end{aligned}$$

The problem to solve can be expressed as finding $u_{m+1} \in H^1$ such that for all $v \in H_0^1$,

$$\begin{cases} \int_\Omega \frac{\partial u_m}{\partial t} v \, d\Omega + a(u_m, v) = 0 \\ u_m = 0, \quad \text{on } \Gamma_D \\ \int_\Omega u_0(\omega) v(\omega) \, d\Omega = \int_\Omega \delta(x - S_0, y - S_0, z - v_0) v \, d\Omega = v(S_0, S_0, v_0) \end{cases}$$

with

$$\begin{aligned} \Gamma_D &= \{(x, y, z) \in \Omega \,|\, x = S_0, \, y = S_0\} \cup \{(x, y, z) \in \Omega \,|\, x = 0\} \\ &\cup \{(x, y, z) \in \Omega \,|\, y = x_{max}\} \cup \{(x, y, z) \in \Omega \,|\, z = \infty\} \, . \end{aligned}$$

4.10.2.3 Galerkin method

Our goal is to discretise the solution space H^1 into H_h^1 such that we can solve a finite-dimensional linear problem. To this end, we will use a truncated support Ω_h that will have the shape of an hexahedron combined with a prism. This geometry will be subdivided into tetrahedra K. The generated triangulation will be denoted by $\mathcal{T}_h$ and an example is depicted Figure 4.10.1.

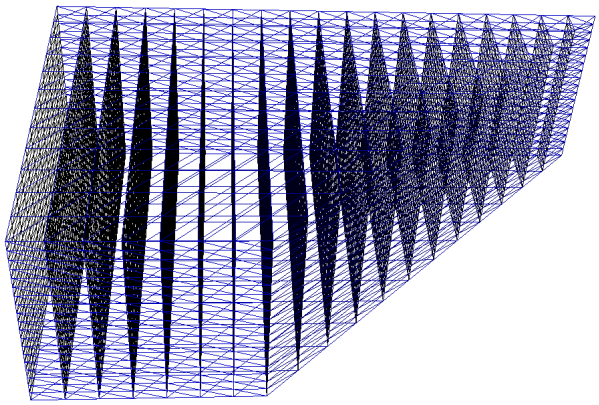


Figure 4.10.1: Possible triangulation of Ω_h.

Let N be the number of nodes, $(\omega_i)_{i\leq N-1}$ the nodes of the mesh and define a set of nodal basis functions $(e_i)_{i\leq N-1}$ associated with the triangulation $\mathcal{T}_h$ defined as

$$\begin{aligned} e_i(\omega_j) &= \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases}, \\ e_i &\in \mathbf{\Pi_n^3}\,, \end{aligned}$$

and where $\mathbf{\Pi_n^3}$ is the space of polynomials of 3 variables and degree n. In H_h^1, the solution is

$$u_{m+1}(x,y,z) = \sum_{i=0}^{N-1} u_{m+1}(x_i,y_i,z_i)e_i(x,y,z)\,.$$

Let $\omega=(x,y,z)\in\Omega_h$ and $u_m^i = u_m(\omega_i)$ for all i in the node numbering $[0,N-1]$. The discrete solution has to satisfy the following equation, where we write a full-implicit scheme for simplicity of notation,

$$\int_\Omega u_{m+1}v\,d\Omega + \Delta_t(m)a(u_{m+1},v_h) = \int_\Omega u_m v_h\,d\Omega\,,$$

or

$$\begin{aligned} \int_\Omega \sum_{i=0}^{N-1} u_{m+1}(\omega_i)e_i(\omega)v_h\,d\Omega &= -\Delta_t(m)a(\sum_{i=0}^{N-1} u_{m+1}(\omega_i)e_i(\omega),v_h) \\ &+ \int_\Omega \sum_{i=0}^{N-1} u_m(\omega_i)e_i(\omega)v_h\,d\Omega\,, \end{aligned}$$

in other words, for any $j<N$,

$$\int_\Omega \sum_{i=0}^{N-1} u_{m+1}^i e_i e_j \, d\Omega + \Delta_t(m) a(\sum_{i=0}^{N-1} u_{m+1}^i e_i, e_j) \quad = \quad \int_\Omega \sum_{i=0}^{N-1} u_m^i e_i e_j \, d\Omega,$$

which gives by linearity of the operator,

$$\begin{aligned} \sum_{i=0}^{N-1} u_{m+1}^i \int_\Omega e_i e_j \, \Omega \quad &= \quad -\Delta_t(m) \sum_{i=0}^{N-1} u_{m+1}^i a(e_i, e_j) \\ &+ \quad \sum_{i=0}^{N-1} u_m^i \int_\Omega e_i e_j \, d\Omega \, . \end{aligned}$$

Furthermore, at $t = 0$ the right-hand side is known by definition of the Dirac delta measure and

$$\sum_{i=0}^{N-1} u_{m+1}^i \int_\Omega e_i e_j \, d\Omega + \Delta_t(m) \sum_{i=0}^{N-1} u_{m+1}^i a(e_i, e_j) \quad = \quad e_j(\omega_{source}) \, ,$$

with $\omega_{source} = (x_0, y_0, 0)$ and $e_{j_{source}}(\omega_{source}) = 1$ where j_{source} is the index number of the source node.

Let us denote by $\mathbf{M}$ the mass matrix and by $\mathbf{K}_m$ the stiffness matrix for time t_m, such that

$$\begin{aligned} (m_{i,j})_{(i,j)\in[\![0,N-1]\!]^2} \quad &= \quad \int_\Omega e_j e_i \, d\Omega \\ (k_{i,j})_{(i,j)\in[\![0,N-1]\!]^2} \quad &= \quad a(e_j, e_i, t_m) \, , \end{aligned}$$

and

$$(\mathbf{u}_m)_{m \le M} = \begin{bmatrix} u_m^{N-1} \\ \vdots \\ u_m^0 \end{bmatrix}.$$

That leads to a linear system of equations,

$$\begin{aligned} \mathbf{M}\mathbf{u}_{m+1} + \Delta_t(m)\mathbf{K}_{m+1}\mathbf{u}_{m+1} \quad &= \quad \mathbf{M}\mathbf{u}_m, \ m > 0 \\ \mathbf{M}\mathbf{u}_1 + \Delta_t(m)\mathbf{K}_1\mathbf{u}_1 \quad &= \quad \mathbf{J}, \ m = 0 \, , \end{aligned} \tag{4.10.6}$$

where $\mathbf{J}$ is zero except for the element number j_{source} which is one. This can be extended to the order 2 BDF time scheme by

$$\mathbf{M}\mathbf{u}_m + \frac{\Delta_t(m)}{1+\gamma_m}\mathbf{K}_{m+1}\mathbf{u}_m = \mathbf{M}\bigg(\bigg(1 + \frac{\Delta_t(m)}{\Delta_t(m-1)}\frac{\gamma_m}{1+\gamma_m}\bigg)\mathbf{u}_{m-1} - \bigg(\frac{\Delta_t(m)}{\Delta_t(m-1)}\frac{\gamma_m}{1+\gamma_m}\bigg)\mathbf{u}_{m-2}\bigg), \ m > 1 \, .$$

Remark. The mass matrix $\mathbf{M}$ can be assembled once, however, $\mathbf{K}_m$ is time-dependent and has to be assembled at each time step.

Finally, the solution is

$$\phi(t_m, x, y, z) = \sum_{i=0}^{N-1} u_m(x_i, y_i, z_i) e_i(x, y, z)\,.$$

4.10.2.4 Initial value

The initial solution at $t = 0$ is the Dirac delta function and usually handled naturally in the weak formulation, as shown in (4.10.6). However, the formulation faces a problem that prevents us from using this property directly. Indeed, the source node (S_0, S_0, v_0) at time $t = 0$ is also a Dirichlet node for any $t > 0$, this means that the Dirichlet boundary condition

$$\phi(S_0, S_0, z, t) = 0\,,$$

is not continuous in time. To the best of our knowledge, it is impossible to write the time-stepping scheme at $t = 0$ and incorporate both the source node and the Dirichlet boundary condition. Explicitly, for a full-implicit time scheme, we define the right-hand side to account for the source node as

$$(\mathbf{M} + \Delta_t(m)\mathbf{K}_1)\mathbf{u}_1 = \mathbf{J},\ m = 0,$$

such that if we want to impose a homogeneous Dirichlet boundary condition on the node (S_0, S_0, v_0), we would need to modify $\mathbf{J}$ and remove the source point.

There are two alternatives to tackle this difficulty. On the one hand, we can use the the joint density of (S, M) under a Black–Scholes model as an approximation for small times, where details of the derivation of (4.10.7) can be found in [135]. Namely, write

$$\phi(x, y, z, \tau) = \phi_{BS}(x, y, \tau)\phi_v(z, \tau)$$

with

$$\phi_{BS}(s, m, t) \;=\; \frac{\sqrt{2}\mathrm{e}^{-\frac{(a-2\,b)^2}{2t} - \frac{a\left(\frac{\sigma^2}{2}+q-r\right)}{\sigma} - \frac{t\left(\frac{\sigma^2}{2}+q-r\right)^2}{2\sigma^2}}(4b-2a)}{2\sqrt{\pi}ms\sigma^2 t^{\frac{3}{2}}}\,, \tag{4.10.7}$$

and

$$a = \frac{\ln(\frac{s}{s_0})}{\sigma}, \quad b = \frac{\ln(\frac{m}{s_0})}{\sigma}\,,$$

where we set

$$\sigma = \sigma(S_0, S_0, v_0)\,,$$

and such that ϕ_v is the density function of the CIR process at time τ starting at v_0. Following [39], a CIR process is distributed as $\frac{Y}{2c}$ where Y is a non-central Chi-square distributed random variable with non-centrality parameter nc and degrees of freedom df, both defined hereafter. Let $\phi_{\mathcal{X}^2}(\cdot, \mathrm{df}, \mathrm{nc})$ be the non-central Chi-square density, then we can write

$$\phi_v(z, \tau) \quad = \quad 2c\phi_{\mathcal{X}^2}(2cz, \mathrm{df}, \mathrm{nc})\,,$$

with

$$c = \frac{2\kappa}{\xi^2(1 - e^{-\kappa\tau})}, \quad \mathrm{df} = \frac{4\kappa\theta}{\xi^2}, \quad \mathrm{nc} = 2cv_0e^{-\kappa\tau}\,.$$

This approach leads to very good results in the two-dimensional case (no stochastic variance) but becomes numerically complex and inexact in the three-dimensional case. Indeed, we see that when τ goes to zero, both c and nc diverge and the density needs to be computed with high floating point accuracy, a procedure that becomes time consuming. It is also possible to approximate $\phi_v(z, \tau)$ with a Gaussian distribution to remove the floating point issue near $\tau = 0$. An additional concern is that the correlation between S_t and V_t is neglected in this approximation, again this problem disappear in two dimensions but leads to non-negligible errors in three dimensions.

Another approach is to approximate the source node to the nearest node in the interior domain and impose the homogeneous Dirichlet boundary condition on the node $\mathbf{p}_0 = (S_0, S_0, v_0)$ as required. This demands to refine the mesh around $\mathbf{p}_0$ such that the source node is as close as the theoretical source point as possible. On the bright side, this refinement is necessary to account for this typical diffusion problem that varies abruptly in the region near $\mathbf{p}_0$, for small times. The construction of such a mesh refinement routine is a two step procedure, given a tetrahedral mesh:

1. Mark each cell containing the point $\mathbf{p}_0$.
2. Refine marked tetrahedra following Plaza and Carey's algorithm [123].
3. Iterate 1 to 7 times (we choose 5 in our tests).

An example of refinement is depicted in Figure 4.10.2.

4.10.3 Mesh construction

4.10.3.1 The two-dimensional case

We first start by proposing a mesh construction and numbering for the two-dimensional case, i.e., the joint density of (S_t, M_t) when S_t is driven by (4.2.5). The mesh is built with a two-step procedure. First of all, we construct a rectangular mesh, as that provides the right structure for the nodes.

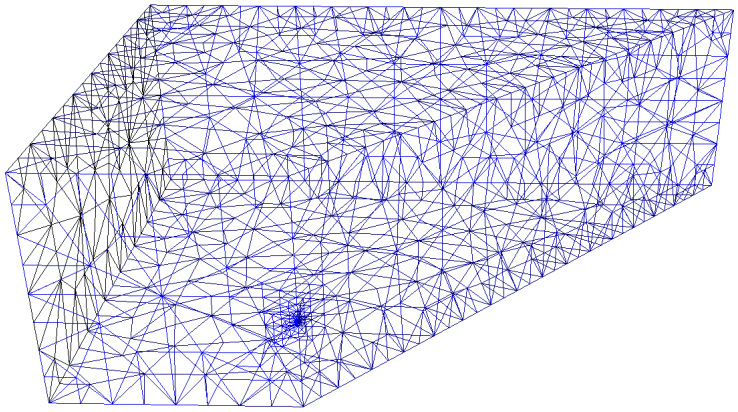


Figure 4.10.2: Mesh refinement near $\mathbf{p}_0 = (S_0, S_0, v_0)$.

Namely, we use a hyperbolic grid in the spot direction which makes the mesh finer around $x = S_0$. More precisely, the grid in spot $(x_i)_{i\in[\![0,N_S]\!]}$ is defined such that

$$x_i \quad = \quad f_h(g_h(\bar{x}_i))$$

with

$$\begin{aligned}
f_h(x) &= S_0 + b\sinh(\nu x + d),\\
b &= \eta(S_{max} - S_{min}),\\
d &= \operatorname{arcsinh}\left(\frac{S_{min} - S_0}{b}\right),\\
\nu &= \operatorname{arcsinh}\left(\frac{S_{max} - S_0}{b}\right) - d,\\
\bar{x}_i &= \frac{i}{(N_S+1)},\\
\eta &= 0.02\,,
\end{aligned}$$

where η is defined according to our numerical experiments and g_h is the quadratic polynomial that passes through the points

$$(0,0), \left(\frac{\left\lfloor f_h^{-1}(S_0)(N_S+1) + 0.5\right\rfloor}{N_S+1}, f_h^{-1}(S_0)\right), (1,1).$$

The latter intermediate step makes sure that S_0 is a vertex of the mesh. The construction of the running maximum grid $(y_j)_{j\in[\![0,N_M^0]\!]}$ is based on the grid in spot as follows

$$\begin{aligned}(y_j)_{j\in[\![0,N_M^0]\!]} = &\quad \{x_i \,|\, i \in N_S \,:\, x_i \geq S_0\}\\ N_M^0 = &\quad \mathrm{card}(\{x_i \,|\, i \in N_S \,:\, x_i \geq S_0\})\,.\end{aligned}$$

The second step of the procedure is to build the mesh from the generated grids $(x_i)_{i\in[\![0,N_S]\!]}$ and $(y_j)_{j\in[\![0,N_M^0]\!]}$. The construction of this finite element triangular mesh can be achieved by creating a vertex at each point (x_i, y_j) and defining two triangular cells (upper left and lower right) in each rectangle, only keeping the upper triangle when on the boundary $x = y$. We explain the exact algorithm in Algorithm 4.4 and show the output mesh in Figure 4.10.3. While in our case it is beneficial to hand-write the mesh, we point out that libraries perform better in most standard cases, generating high quality (unstructured) meshes. One can refer to [69] for extensive details.

Let $(N_M(i))_{i\in[\![0,N_S]\!]}$ be the number of nodes in the spot ladder i. The number of nodes N is the sum over all spot ladders $i \in N_S$. In regard to the number of cells C, if we look at spot ladder i, there are $(N_M(i+1)-1)$ full rectangles (with an upper and lower triangle) since $N_M(i) \geq N_M(i+1)$. If $N_M(i) - N_M(i+1) = 1$, then the last node of ladders i and $i+1$ belong to the boundary $x = y$ and we add the upper boundary triangle to the count. Otherwise, $N_M(i) - N_M(i+1) = 0$ and none of the ladders i, nor $i+1$ hold the last node on the boundary $x = y$. Equivalently, we write

$$N = \sum_{i=0}^{N_S} N_M(i), \quad C = \sum_{i=0}^{N_S-1} 2(N_M(i+1)-1) + (N_M(i) - N_M(i+1))\,.$$

Denote $n(i,j)$ the vertex number for a position (x_i, y_j) such that

$$n(i,j) = \sum_{m=0}^{i-1} N_M(m) - \Big(N_M^0 - N_M(i)\Big) + j\,.$$

Furthermore, to solve the PDE numerically, we need to truncate at the boundary and use $\bar{\Omega} = \{(x,y) \in [0, S_{max}] \times [S_0, S_{max}] \,|\, x \leq y\}$ on a time interval $[0,T]$. We choose

$$\begin{aligned}S_{max} &= S_0 e^{5\overline{\sigma}\sqrt{T}}\,,\\ \overline{\sigma} &= \frac{\int_{S_0}^{2S_0}\int_0^y \sigma\Big(x,y,\frac{T}{2}\Big)\,dxdy}{\frac{3}{2}S_0^2}\,,\end{aligned}$$

where $\overline{\sigma}$ is the average volatility over the domain $\{(x,y) \in [0,2S_0] \times [S_0, 2S_0] \,|\, x \leq y\}$ which has area $\frac{3}{2}S_0^2$.

Algorithm 4.4 Mesh construction for the joint density of (S_t, M_t) under model (4.2.5)

```
for ( i = 0 ; i ≤ N_S ; i++ ) do
  for ( j = 0 ; j ≤ N_M(i) ; j++ ) do
    add Triangle of vertices (n(i,j), n(i,j+1), n(i+1,j+1))
    if y_j ≥ x_{i+1} then
      add Triangle of vertices (n(i,j), n(i+1,j+1), n(i+1,j))
    end if
  end for
end for
```

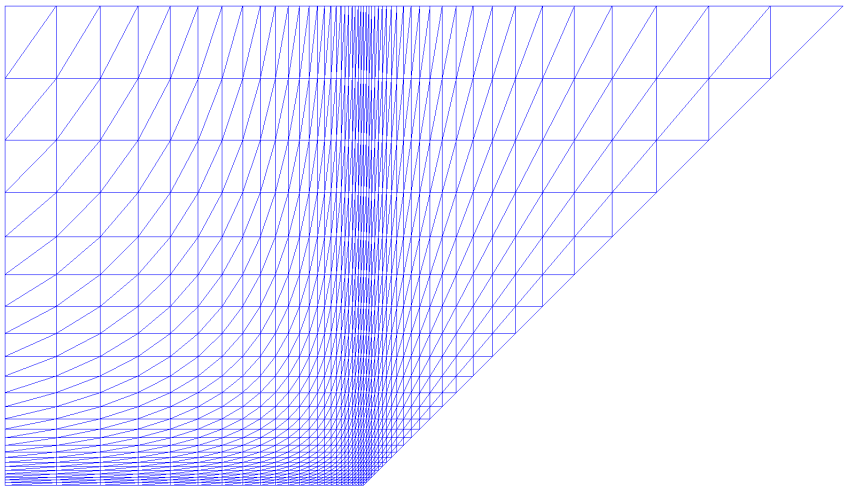


Figure 4.10.3: FEM mesh for the joint density of (S_t, M_t) under model (4.2.5).

4.10.3.2 The three-dimensional case

One possibility to construct a mesh for the joint density function (S_t, M_t, V_t) under model (4.3.25) is to extrude each triangle in Figure 4.10.3 in the variance direction according to a discretisation grid. The result is a 3D mesh made of parallel wedges, i.e. triangular prisms, which we can divide in exactly three tetrahedra, see [51] for details. More precisely, if (a, b, c, d, e, f) are the vertices of the wedge, one possible subdivision are the tetrahedra of vertices

$$\begin{aligned} \text{tet1} &\equiv (a, f, e, d) \\ \text{tet2} &\equiv (a, b, f, e) \\ \text{tet3} &\equiv (a, c, b, f)\,. \end{aligned}$$

An example of mesh generated with this technique is shown in Figure 4.10.1 where we used uniform spacing for visualisation purposes. Unfortunately, this does not provide satisfactory results as the

number of cells increases rapidly and some elements are of poor quality. Therefore, we define the geometry (hexahedron and wedge) and use Constructive Solid Geometry (CSG) alongside TetGen [136] for the meshing. Afterwards, we apply mesh refinement techniques on the boundaries $x = y$ and $y = S_0$. The first one is subject to the degenerated boundary equation when $S = M$, the second one does not hold any boundary condition. We perform the mesh refinement as follows: first, we identify the cells containing a boundary facet that belongs to the boundary of interest, second we refine each of these cells following [123], and finally we iterate the process as many times as desired. In our tests, we iterated 5 times for $x = y$ and 3 times for $y = S_0$. The final mesh is displayed in Figure 4.10.4.

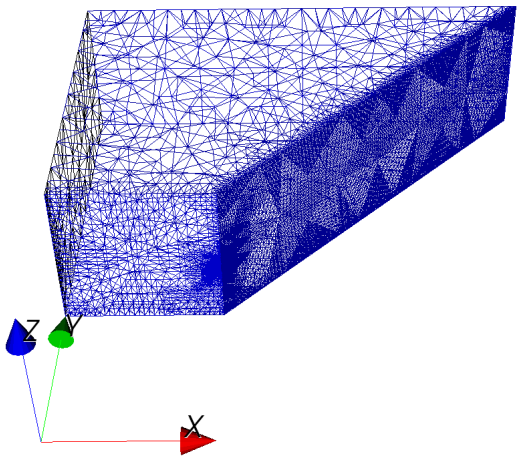


Figure 4.10.4: FEM mesh for the joint density of (S_t, M_t, V_t) under model (4.3.25).

4.10.4 Solving the Fokker–Planck PDE for calibrated models

In this part, we solve numerically both the two- and tree-dimensional PDEs that drives the joint density of (S_t, M_t) and (S_t, M_t, V_t), respectively. In both cases we compute the up-and-out call prices to show that the solutions are sensible and accurate.

4.10.4.1 FEM solution for a Brunick–Shreve model calibrated to both vanilla and foreign no-touch options

We use the calibrated Brunick–Shreve model following our method in Section 4.5. Therefore, if all up-and-out call option quotes were available, the density function, solution to the PDE (4.6), would be solution to the implied density of the market. For validation purposes, we price up-and-out call options with a two-dimensional quadrature, using the density function, and the forward PIDE for barriers (2.4.1), and we choose

$$B = 110\% \times S_0, \quad T = 1,$$

and the results are displayed in Figure 4.10.5. The FEM solution is computed with 50 time steps, Lagrange polynomial of order three for the basis functions and 200 discretisation points in the spot direction. The quadrature is performed on both a rectangle of vertices

$$(0, S_0), (S_0, S_0), (S_0, B), (0, B)\,,$$

and triangle of vertices

$$(0, S_0), (B, B), (S_0, B)\,,$$

where the quadrature rule on the rectangle is a product of Gauss-Legendre line points and the quadrature rule on the triangle is performed according to [154]. Finally, we graph the density function in Figure 4.10.6.

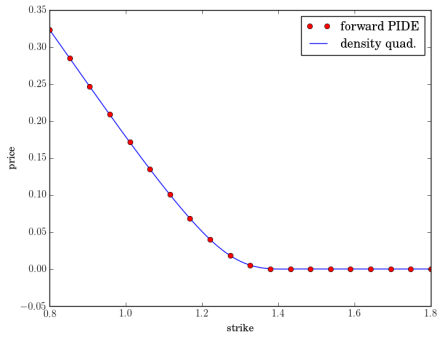


Figure 4.10.5: Forward PIDE vs quadrature price for up-and-out call options with $B = 110\% \times S_0$ on a ladder of strikes. Average relative difference over all plotted points is 0.042%.

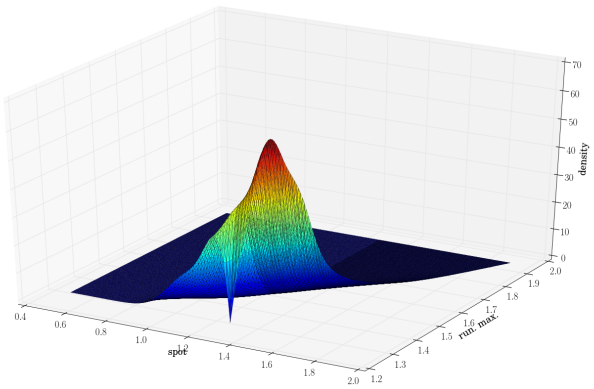


Figure 4.10.6: FEM solution of the PDE (4.3.26) with no variance term for $t = 1$.

4.10.4.2 FEM solution for an LSV model calibrated to vanilla options

For the three-dimensional case, we will work under a Heston-type LSV model calibrated to vanilla options following the procedure we presented in Chapter 3 where we choose a mixing-factor $\eta = 0.5$. In this case, the Feller condition is satisfied and we apply a homogeneous Dirichlet boundary condition $\phi = 0$ for $z = 0$ in (4.3.26). The calibrated Heston parameters are

$$v_0 = 0.00828, \quad \theta = 0.0156, \quad \kappa = 0.7146, \quad \rho = -0.4429, \quad \eta = 0.5, \quad \xi = 0.1893\,.$$

For validation purposes, we price up-and-out call options with both a tree-dimensional quadrature (using the density function) and Monte Carlo. We choose

$$B = 120\% \times S_0, \quad T = 1,$$

and the results are displayed in Figure 4.10.7. The FEM solution is computed with 50 time steps, Lagrange polynomial of order two for the basis functions and 69136 cells. The quadrature is performed on the finite-element space. The Monte Carlo prices are computed with 250 time steps and 500 000 paths using a QE-scheme [7] combined with the Brownian bridge technique (see Chapter 6 of [73]) for the running-maximum estimation. Finally, we plot the solution for a given value in the variance direction $z = v_0$, i.e. the function $\phi(x, y, v_0)$, in Figure 4.10.8.

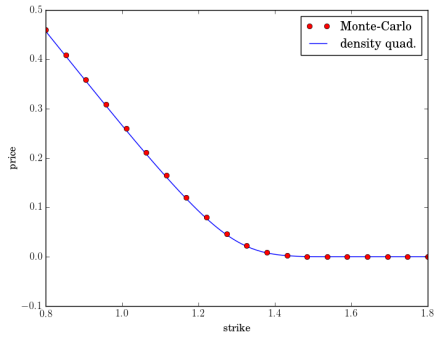


Figure 4.10.7: Monte Carlo vs quadrature price for up-and-out call options with $B = 120\% \times S_0$ on a ladder of strikes. Average relative difference over all plotted points is 0.079%.

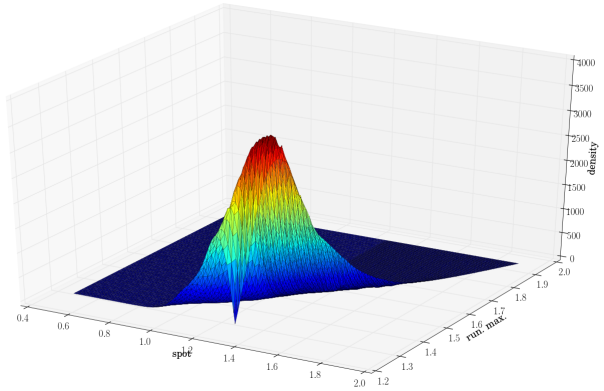


Figure 4.10.8: FEM solution of the PDE (4.3.26) for $t = 1$ and $z = v_0$.

4.11 Conclusion

In this work, we presented three models and their calibration to two traded products, namely, vanilla and no-touch options. First, a Brunick–Shreve model was introduced as the natural extension to the Dupire local volatility framework, and its calibration was performed combining the forward PIDE, a fast and accurate method to compute the high-dimensional gradient as well as an effective regularisation technique. Second, the addition of a Heston-type stochastic volatility on top of the Brunick–Shreve model allowed us to discuss and introduce a new model with a potentially more interesting spot-vol dynamics. The calibration was performed with a new two-dimensional particle method to account for the Markovian projection onto the two-dimensional state space (S, M). Both the Brunick–Shreve and LMSV path-dependent models showed a very accurate calibration to both

vanilla and foreign no-touch options. Finally, we introduced an LSV-LVV model, an extension of the classic Heston-type LSV model with a local vol-of-vol, a model that we believe is of interest to practitioners who do not want to incorporate the running maximum in the parametrisation. Due to the small number of degrees-of-freedom, the chosen LSV-LVV parametrisation cannot match no-touch options prices perfectly, however, the fit is very satisfactory as the model lies well within the market bid–ask spread across all quoted barrier levels and maturities. A finite element solution to the Fokker–Planck PDE for the joint density of (S, M, V) is discussed and provides an accurate and efficient method to compute the peculiar density function.

An interesting addition will be to compare the volatility dynamics implied by the three models through the pricing of forward start options. One would then be able to understand to which extent the calibration of touch options, and barriers in general, can explain the market smile dynamics.

5 Conclusion

In Chapter 2, we provide a forward equation for barrier options, namely, up-and-out call options that allows to retrieve prices at every strikes, barrier levels and maturities in one PDE resolution. A calibration algorithm for the Brunick–Shreve model is proposed in Chapter 4, based on the forward PIDE where the finite-difference scheme is further extended to a non-uniform grid. The forward PIDE can be written in terms of up-and-out call prices and no-touch prices to exhibit a Dupire-type formula for barrier options. However, we do not believe such a formula can be used on market prices directly, due to high-order derivatives and potential instabilities. This chapter creates the foundation for extensions of the calibration literature to vanilla options as presented in Guyon and Labordère [82] and Ren, Madan and Quian [129]. In Chapter 4, we discuss a generic calibration framework and detail the extensions of Markovian projection calibration to barrier options.

In Chapter 3, we used Markovian projection techniques, and more particularly developed variance reduction techniques for the particle method of Guyon and Labordère [82] to calibrate a 4-factor LSV model for vanilla options. This provides an efficient method for calibration of an LSV model with both stochastic domestic and foreign rates, a subject rarely discussed in the literature. A very high accuracy can be obtained with as few as 20 000 particles (with a maximum error in absolute volatility of 0.03%), and we were able to get a good fit with only 4 000 particles (with a maximum error in absolute volatility of 0.06%), leading to a 125-fold speed-up for the calibration using control variates, as compared to the plain particle method. We showed that the addition of stochastic rates may have a significant impact on structured products, even more so when barrier features and coupon detachments are combined for longer-dated contracts. This naturally leads to the question of barrier calibration under a long-term setting, a question that we discuss in Chapter 4, where we present a self-consistency formula for calibration to barrier options under a stochastic rate framework. Extensions to higher dimensional CIR++ processes, where the volatility is made stochastic in order to improve the fit to caps, is possible, and the calibration method presented in this chapter still applies.

In Chapter 4, we built on the forward PIDE of Chapter 2 and discussed three models and their calibration to two traded products, namely, vanilla and no-touch options. First, a Brunick–Shreve model was calibrated combining the forward PIDE, a fast and accurate method to compute the high-dimensional gradient as well as an effective regularisation technique. Second, the addition of a Heston-type stochastic volatility on top of the Brunick–Shreve model allowed us to introduce a stochastic path-dependent volatility model with a potentially more interesting spot-vol dynamics. A

new two-dimensional particle method to account for the Markovian projection onto the two-dimensional state space (S, M) was presented. Both path-dependent volatility models showed a very accurate calibration to both vanilla and foreign no-touch options. Finally, we introduced an extension of the classic Heston-type LSV model including a local vol-of-vol, a model that does not include the running maximum in the parametrisation while providing a better fit to no-touch options than standard LSV models. The fit is very satisfactory as the model lies well within the market bid–ask spread across all quoted barrier levels and maturities. However, the small number of degrees-of-freedom prevents the LSV-LVV model to match no-touch options prices perfectly. A finite element solution to the Fokker–Planck PDE is proposed and provides an efficient method to retrieve the joint density of the path-dependent diffusion (S, M, V).

The computation of accurate, stable and reliable sensitivities would be a natural extension of the calibration framework. The classical approach is to calculate spot sensitivities by finite differences, where the spot value is shifted up and the net-present-value recomputed accordingly. However, a movement of spot implies a movement of market data and particularly the implied volatility surface the model is calibrated onto. On the one hand, a practitioner could believe the model implied smile dynamics is right and compute the Greeks without re-calibration. On the other hand, one could modify the market data according to personal beliefs or statistical analysis, recalibrate the model and compute the sensitivities by finite differences. Additional studies of the models' implied volatility dynamics and its impact on hedging is then important and left for further research.

Bibliography

[1] M. Abramowitz. *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*. Dover Publications, 1974.

[2] Y. Achdou and O. Pironneau. *Computational Methods for Option Pricing*. Frontiers in applied mathematics. Society for Industrial and Applied Mathematics, Philadelphia, 2005.

[3] R. Aebi. Itô formula for non-smooth functions. *RIMS, Kyoto Univ*, 28:595–602, 1992.

[4] R. Ahlip and M. Rutkowski. Pricing of foreign exchange options under the Heston stochastic volatility model and CIR interest rates. *Quantitative Finance*, 13(6):955–966, 2013.

[5] A. Alfonsi. On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods and Applications*, 11(4):355–384, 2005.

[6] A. Alfonsi. Strong order one convergence of a drift implicit Euler scheme: application to the CIR process. *Statistics and Probability Letters*, 83(2):602–607, 2013.

[7] L. Andersen. Simple and efficient simulation of the Heston stochastic volatility model. *The Journal of Computational Finance*, 11(3):1–42, 2008.

[8] L. Andersen and V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11(1):29–50, 2007.

[9] E. Ayache, P. Henrotte, S. Nassar, and X. Wang. Can anyone solve the smile problem. *Wilmott Journal*, January, January 2004.

[10] A. B. Ayoub. The central conic sections revisited. *Mathematics Magazine*, 66(5):322–325, 1993.

[11] M. Barlow and M. Yor. Semi-martingale inequalities via the garsia-rodemich-rumsey lemma, and applications to local times. *Journal of Functional Analysis*, 49(2):198 – 229, 1982.

[12] E. Benhamou, A. Gruz, and A. Rivoira. Stochastic interest rates for local volatility hybrids models. SSRN preprint, SSRN 1107711.

[13] L. Bergomi. Smile dynamics. *Risk Magazine*, Septembre, September 2004.

[14] L. Bergomi. *Stochastic Volatility Modeling*. Chapman and Hall/CRC Financial Mathematics. Chapman and Hall/CRC, 2016.

[15] A. Berkaoui, M. Bossy, and A. Diop. Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM: Probability and Statistics*, 12:1–11, 2008.

[16] T. Björk. The pedestrian's guide to local time. arxiv preprint, arXiv:1512.08912, 2015.

[17] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):pp. 637–654, 1973.

[18] A. Bluemke. *How to Invest in Structured Products: A Guide for Investors and Asset Managers (The Wiley Finance Series)*. The Wiley Finance Series. Wiley, 1 edition, 2009.

[19] T. Bouezmarni and J. V. Rombouts. Nonparametric density estimation for multivariate bounded data. *Journal of Statistical Planning and Inference*, 140(1):139–152, 2010.

[20] D. Brigo and F. Mercurio. A deterministic-shift extension of analytically-tractable and time-homogeneous short-rate models. *Finance and Stochastics*, 5(3):369–387, 2001.

[21] D. Brigo and F. Mercurio. *Interest Rate Models: Theory and Practice*. Springer, 2006.

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

[23] R. A. Brown. Building a balanced k-d tree in $o(kn \log n)$ time. *Journal of Computer Graphics Techniques*, 4(1):50–68, 2015.

[24] G. Brunick. Uniqueness in law for a class of degenerate diffusions with continuous covariance. *Probability Theory and Related Fields*, 155(1):265–302, 2013.

[25] G. Brunick and S. Shreve. Mimicking an Itô process by a solution of a stochastic differential equation. *Ann. Appl. Probab.*, 23(4):1584–1628, 2013.

[26] J. C. Butcher. *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods*. Wiley-Interscience, New York, NY, USA, 1987.

[27] J. Céa. Approximation variationnelle des problèmes aux limites. *Annales de l'institut Fourier*, 14:345–444, 1964.

[28] R. E. Caflisch, W. Morokoff, and A. B. Owen. Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension. *Journal of Computational Finance*, 1:27–46, 1997.

[29] R. Carmona and S. Nadtochiy. Local volatility dynamic models. *Finance and Stochastics*, 13(1):1–48, 2009.

[30] P. Carr. Implied vol constraints. Technical report, Bloomberg Working Paper, 2004. Available at www.javaquant.net/papers/impvolconstrs3.pdf.

[31] P. Carr and J. Crosby. A class of Lévy process models with almost exact calibration to both barrier and vanilla FX options. *Quant. Finance*, 10(10):1115–1136, 2010.

[32] P. Carr and A. Hirsa. Forward evolution equations for knock-out options. In M. C. Fu, R. A. Jarrow, J.-Y. J. Yen, and R. J. Elliott, editors, *Advances in Mathematical Finance*, Applied and Numerical Harmonic Analysis, pages 195–217. Birkhäuser Boston, 2007.

[33] T.-M. Chao and C. Ching-Sung. On some inequalities of local times for Azéma martingales. *Bernoulli*, 6, 2000.

[34] S. Ciliberti, J.-P. Bouchaud, and M. Potters. Smile dynamics: A theory of the implied leverage effect. *Wilmott Journal*, April:87–94, May 2009.

[35] I. J. Clark. *Foreign Exchange Option Pricing: A Practitioner's Guide*. John Wiley & Sons, Chichester, UK, 2010.

[36] T. Coleman and Y. Li. On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds. *Mathematical Programming*, 67(1-3):189–224, 1994.

[37] T. F. Coleman and Y. Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6(2):418–445, 1996.

[38] A. M. Cox and J. Oblój. Robust pricing and hedging of double no-touch options. *Finance and Stochastics*, 15:573–605, Sep 2011.

[39] J. C. Cox, J. Ingersoll, Jonathan E, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, March 1985.

[40] 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. arxiv preprint, arXiv:1501.06084.

[41] 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, arXiv:1701.06001, 2017.

[42] S. Crépey. *Tikhonov regularization*, pages 1807–1812. John Wiley & Sons, Chichester, UK, 2010.

[43] B. Davis. On the barlow-yor inequalities for local time. *Séminaire de probabilités de Strasbourg*, 21:218–220, 1987.

[44] G. Deelstra and F. Delbaen. Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Applied Stochastic Models and Data Analysis*, 14(1):77–84, 1998.

[45] G. Deelstra and G. Rayee. Local volatility pricing models for long-dated FX derivatives. *Applied Mathematical Finance*, 20(4):380–402, 2013.

[46] S. Dereich, A. Neuenkirch, and L. Szpruch. An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. *Proceedings of the Royal Society of London A*, 468(2140):1105–1115, 2012.

[47] E. Derman. Regimes of volatility. *Risk Magazine*, April, April 1999.

[48] E. Derman and I. Kani. The volatility smile and its implied tree. Goldman Sachs Quantitative Strategies Research Notes, 1994.

[49] E. Derman, I. Kani, and M. Kamal. Trading and hedging local volatility. Goldman Sachs Quantitative Strategies Research Notes, 1996.

[50] P. Dierckx. An algorithm for surface-fitting with spline functions. *IMA Journal of Numerical Analysis*, 1(3):267–283, 1981.

[51] J. Dompierre, P. Labbé, M. gabrielle Vallet, and R. Camarero. How to subdivide pyramids, prisms and hexahedra into tetrahedra, 1999.

[52] S. S. Dragomir. *Some Gronwall Type Inequalities and Applications*. Nova Science Publishers, 2003.

[53] B. Dupire. Pricing and hedging with smiles. In *Mathematics of derivative securities (Cambridge, 1995)*, volume 15 of *Publ. Newton Inst.*, pages 103–111. Cambridge Univ. Press, Cambridge, 1997.

[54] B. Dupire. *A Unified Theory of Volatility*, chapter 6. Risk Books, London, 2004.

[55] B. Dupire. Functional Itô calculus. *Bloomberg Education & Quantitative Research and Education Paper Series*, 2009-04-FRONTIERS, 2009.

[56] H. Egger and H. W. Engl. Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates. *Inverse Problems*, 21(3):1027, 2005.

[57] L. C. Evans. *Partial differential equations*. Graduate studies in mathematics, v. 19. American Mathematical Society, 1998.

[58] F. Fang and C. W. Oosterlee. A Fourier-based valuation method for Bermudan and barrier options under Heston's model. *SIAM Journal on Financial Mathematics*, 2:439 – 463, 2011.

[59] M. R. Fengler. Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance*, 9(4):417–428, 2009.

[60] R. Fletcher. A new approach to variable metric algorithms. *The Computer Journal*, 13(3):317–322, Jan. 1970.

[61] F. L. Floch. TR-BDF2 for stable American option pricing. *Journal of Computational Finance*, 17(3), 2014.

[62] M. Forde. On the Markovian projection in the Brunick-Shreve mimicking result. *Statistics & Probability Letters*, 85:98–105, 2013.

[63] M. Forde, A. Pogudin, and H. Zhang. Hitting times, occupation times, tri-variate laws and the forward Kolmogorov equation for one-dimensional diffusion with memory. *Adv. in Appl. Probab.*, 45(3):595–893, 2013.

[64] B. Fornberg. Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of Computation*, 51(184):699–706, 1988.

[65] P. Frey. The finite element approximation.

[66] P. Friz and M. Keller-Ressel. Moment explosions in stochastic volatility models. In *Encyclopedia of Quantitative Finance*, pages 1247–1253. John Wiley, Chichester, UK, 2010.

[67] M. Gaß, K. Glau, M. Mahlstedt, and M. Mair. Chebyshev interpolation for parametric option pricing. preprint, arXiv:1505.04648, 2015.

[68] J. Gatheral. *The Volatility Surface: a Practitioner's Guide*. John Wiley & Sons, Hoboken, N.J, 2006.

[69] C. Geuzaine and J.-F. Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79:1309–1331, 2009.

[70] M. Giles and R. Carter. Convergence analysis of Crank-Nicolson and Rannacher time-marching. *Journal of Computational Finance*, 9(4):89–112, 2006.

[71] M. Giles and P. Glasserman. Smoking adjoints: Fast monte carlo greeks. *Risk*, 19(1):88–92, 2006.

[72] M. B. Giles, D. J. Higham, and X. Mao. Analysing multi-level Monte Carlo for options with non-globally Lipschitz payoff. *Finance and Stochastics*, 13(3):403–413, 2009.

[73] P. Glasserman. *Monte Carlo Methods in Financial Engineering*, volume 53 of *Stochastic Modelling and Applied Probability*. Springer, 2003.

[74] B. Gough. *GNU Scientific Library Reference Manual - Third Edition*. Network Theory Ltd., 3rd edition, 2009.

[75] A. Griewank and A. Walther. *Evaluating derivatives: principles and techniques of algorithmic differentiation*. SIAM, second edition, 2008.

[76] M. Grigoriu. *Stochastic Calculus: Applications in Science and Engineering*. Birkhäuser, New York, third edition, 2013.

[77] L. A. Grzelak and C. W. Oosterlee. On the Heston model with stochastic interest rates. *SIAM Journal on Financial Mathematics*, 2:255–286, 2011.

[78] 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.

[79] F. Guillaume and W. Schoutens. Heston model: The variance swap calibration. *Journal of Optimization Theory and Applications*, 161(1):76–89, 2014.

[80] 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.

[81] J. Guyon. Path-dependent volatility. *Risk Magazine*, September, 2014.

[82] J. Guyon and P. Henry-Labordere. The smile calibration problem solved. SSRN.1885032, 2011.

[83] J. Guyon and P. Henry-Labordere. *Nonlinear Option Pricing*. Chapman and Hall/CRC Financial Mathematics. Chapman and Hall/CRC, Boca Raton, USA, 2013.

[84] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probability Theory and Related Fields*, 71(4):501–516, 1986.

[85] E. Hairer, G. Wanner, and S. P. Nørsett. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer Series in Computational Mathematics 8. Springer-Verlag, Berlin Heidelberg, 2 edition, 1993.

[86] B. Hambly, M. Mariapragassam, and C. Reisinger. A forward equation for barrier options under the Brunick & Shreve Markovian projection. *Quantitative Finance*, 16(6):827–838, 2016.

[87] 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(2):pp. 327–343, 1993.

[88] D. J. Higham and X. Mao. Convergence of Monte Carlo simulations involving the mean-reverting square root process. *The Journal of Computational Finance*, 8(3):35–62, 2005.

[89] D. J. Higham, X. Mao, and A. M. Stuart. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM Journal on Numerical Analysis*, 40(3):1041–1063, 2002.

[90] J. Hull. *Options, Futures, and other Derivatives*. Pearson Prentice Hall, 6. ed., Pearson Internat. ed edition, 2006.

[91] T. R. Hurd and A. Kuznetsov. Explicit formulas for Laplace transforms of stochastic integrals. *Markov Processes and Related Fields*, 14:277–290, 2008.

[92] A. S. Hurn, K. A. Lindsay, and A. J. McClelland. Estimating the parameters of stochastic volatility models using option price data. *Journal of Business and Economic Statistics*, 33(4):579–594, 2014.

[93] M. Hutzenthaler and A. Jentzen. *Numerical Approximations of Stochastic Differential Equations with Non-Globally Lipschitz Continuous Coefficients*, volume 236 of *Memoirs of the American Mathematical Society*. American Mathematical Society, Providence, RI, 2015.

[94] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proceedings of the Royal Society A*, 467:1563–1576, 2011.

[95] M. Hutzenthaler, A. Jentzen, and M. Noll. Strong convergence rates and temporal regularity for Cox–Ingersoll–Ross processes and Bessel processes with accessible boundaries. arxiv preprint, arxiv:1403.6385v1, 2014.

[96] M. Hutzenthaler, A. Jentzen, and X. Wang. Exponential integrability properties of numerical approximation processes for nonlinear stochastic differential equations. arxiv preprint, arxiv:1309.7657, 2014.

[97] P. Jäckel. Let's be rational. *Wilmott*, 2015(75):40–53, 2015.

[98] M. Jex, R. Henderson, and D. Wang. Pricing exotics under the smile. *Risk Magazine*, 11:72–75, 1999.

[99] C. Kahl and P. Jäckel. Fast strong approximation Monte-Carlo schemes for stochastic volatility models. *Quantitative Finance*, 6(6):513–536, 2006.

[100] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*. Graduate texts in mathematics. Springer-Verlag, New York, Berlin, Heidelberg, 1991. Autres tirages corrigés : 1996, 1997, 1999, 2000, 2005.

[101] N. E. Karoui, M. Jeanblanc-Picquè, and S. E. Shreve. Robustness of the black and scholes formula. *Mathematical Finance*, 8(2):93–126, 1998.

[102] P. Kloeden and A. Neuenkirch. Convergence of numerical methods for stochastic differential equations in mathematical finance. In T. Gerstner and P. Kloeden, editors, *Recent Developments in Computational Finance: Foundations, Algorithms and Applications*. World Scientific, 2012.

[103] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, third edition, 1999.

[104] S. R. Liberty and L. Mou. Existence of solutions of a riccati differential system from a general cumulant control problem. *International Journal of Differential Equations*, 2011:1–13, 2011.

[105] A. Logg, K. Mardal, and G. Wells. *Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book*. Lecture Notes in Computational Science and Engineering. Springer Berlin Heidelberg, 2012.

[106] R. Lord, R. Koekkoek, and D. van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2):177–194, 2010.

[107] V. Lucic. Boundary conditions for computing densities in hybrid models via PDE methods. *Stochastics: An International Journal of Probability and Stochastic Processes*, 84(5-6):705–718, 2012.

[108] H. P. McKean. A class of Markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences of the United States of America*, 56:1907–1911, 1966.

[109] S. McNamara. Efficient calibration of FX smile models. *BNP Paribas*, Internal, 2013.

[110] Murex Analytics. Heston model calibration. *Murex*, Internal, 2007.

[111] Murex Analytics. Local volatility model calibration. *Murex*, Internal, 2007.

[112] J. A. Nelder and R. Mead. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, Jan. 1965.

[113] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308, 1965.

[114] A. Neuenkirch and L. Szpruch. First order strong approximations of scalar SDEs defined in a domain. *Numerische Mathematik*, 128(1):103–136, 2014.

[115] S. Ninomiya and N. Victoir. Weak approximation of stochastic differential equations and application to derivative pricing. *Applied Mathematical Finance*, 15(2):107–121, 2008.

[116] R. T. O.C. Zienkiewicz and J. Zhu. *The Finite Element Method Set. Its Basis and Fundamentals.* Butterworth-Heinemann, 6 edition, 2005.

[117] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications (Universitext).* Springer, Jan 2014.

[118] M. Overhaus, A. Ferraris, T. Knudsen, F. Mao, L. Nguyen-Ngoc, and G. Schindlmayr. *Equity Derivatives: Theory and Applications.* Wiley Finance. Wiley, 2003.

[119] E. P. P. Kloeden. *Numerical Solution of Stochastic Diffeential Equations.* Springer, 1992.

[120] O. Pironneau. Calibration of barrier options. *Advances in Numerical mathematics*, 2006.

[121] O. Pironneau. Dupire-like identities for complex options. *Compte rendu de l'académie des sciences I*, 344:127–133, 2007.

[122] V. Piterbarg. Markovian projection method for volatility calibration. *Risk Magazine*, April, 2007.

[123] A. Plaza and G. Carey. Local refinement of simplicial grids based on the skeleton. *Applied Numerical Mathematics*, 32(2):195 – 218, 2000.

[124] D. M. Pooley, K. R. Vetzal, and P. A. Forsyth. Convergence remedies for non-smooth payoffs in option pricing. *Journal of Computational Finance*, 6(4):25–40, 2003.

[125] P. Protter. *Stochastic Integration and Differential Equation.* Springer-Verlag, Berlin, Heidelberg, second edition, 1992.

[126] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*, volume 23. Springer Science & Business Media, 2008.

[127] A. Rehai. The hybrid most likely path. *Risk Magazine*, April, 2006.

[128] C. Reisinger and A. Whitley. The impact of a natural time change on the convergence of the Crank-Nicolson scheme. *IMA Journal of Numerical Analysis*, 34(3):1156, 2014.

[129] Y. Ren, D. Madan, and M. Qian Qian. Calibrating and pricing with embedded local volatility models. *Risk Magazine*, September, 2007.

[130] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[131] C. L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and Martingales. Vol. 1*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Cambridge University Press, Cambridge, second edition, 2000.

[132] C. L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and Martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000.

[133] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill Book Co., New York, third edition, 1976.

[134] R. Schöbel and J. Zhu. Stochastic volatility with an Ornstein–Uhlenbeck process: an extension. *European Finance Review*, 3(1):23–46, 1999.

[135] S. E. Shreve. *Stochastic Calculus for Finance II: Continuous-time Models*. Springer Finance. Springer-Verlag, New York, 2004.

[136] H. Si. TetGen, a Delaunay-based quality tetrahedral mesh generator. *ACM Trans. Math. Softw.*, 41(2):11:1–11:36, Feb. 2015.

[137] B. W. Silverman. Density estimation for statistics and data analysis. *Biometrical Journal*, 30(7):876–877, 1988.

[138] R. Storn and K. Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341–359, 1997.

[139] A.-S. Sznitman. *Ecole d'Eté de Probabilités de Saint-Flour XIX*, chapter Topics in propagation of chaos, pages 165–251. Springer, 1991.

[140] Y. Tian, Z. Zhu, G. Lee, F. Klebaner, and K. Hamza. Calibrating and pricing with a stochastic-local volatility model. *Journal of Derivatives*, 22(3):21–39, 2015.

[141] L. Tur. Local volatility calibration with fixed-point algorithm. *GDF Suez Trading*, Internal, 2014.

[142] 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):1–30, 2014.

[143] A. W. Van der Stoep, L. A. Grzelak, and C. W. Oosterlee. A novel Monte Carlo approach to hybrid local volatility models. *Quantitative Finance, forthcoming*, 2017.

[144] A. Van Haastrecht. *Pricing long-term options with stochastic volatility and stochastic interest rates*. PhD thesis, University of Amsterdam, 2010.

[145] A. Van Haastrecht, R. Lord, A. Pelsser, and D. Schrager. Pricing long-maturity equity and FX derivatives with stochastic interest rates and stochastic volatility. *Insurance: Mathematics and Economics*, 45(3):436–448, 2009.

[146] 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.

[147] D. J. Wales and J. P. K. Doye. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atom. *The Journal of Physical Chemistry A*, 101(28):5111–5116, 1997.

[148] R. White. Numerical solution to PDEs with financial applications. OpenGamma Quantitative Research, https://developers.opengamma.com/quantitative-research/numerical-solutions-to-pdes-with-financial-applications-opengamma.pdf, 2013.

[149] P. Willmott. Where quants go wrong. *Willmott Journal*, 1:1–12, 2009.

[150] H. Windcliff, P. A. Forsyth, and K. R.Vetzal. Analysis of the stability of the linear boundary condition for the Black-Scholes equation. *Journal of Computational Finance*, 8:65–92, 2004.

[151] M. Wyns. Convergence analysis of the Modified Craig-Sneyd scheme for two-dimensional convection-diffusion equations with nonsmooth initial data. arXiv preprint arXiv:1508.04296, 2015.

[152] M. Wyns and J. D. Toit. A finite volume - alternating direction implicit approach for the calibration of stochastic local volatility models. *arXiv preprint arXiv:1611.02961*, 2016.

[153] U. Wystup. *FX Options and Structured Products (The Wiley Finance Series)*. The Wiley Finance Series. Wiley, 2007.

[154] H. Xiao and Z. Gimbutas. A numerical algorithm for the construction of efficient quadrature rules in two and higher dimensions. *Computers and Mathematics with Applications*, 59(2):663 – 676, 2010.

[155] P. N. Yianilos. Data structures and algorithms for nearest neighbor search in general metric spaces. In *Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 1993.

[156] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software*, 23(4):550–560, 1997.