The Heston Stochastic-Local Volatility Model: Efficient Monte Carlo Simulation

Anthonie W. van der Stoep[b,c*]   Lech A. Grzelak[b,c]
Cornelis W. Oosterlee[a,c]

[a] Delft Institute of Applied Mathematics, Delft University of Technology,
Mekelweg 4, 2628 CD, Delft, the Netherlands

[b] Derivatives Research and Validation Group, Rabobank,
Graadt van Roggenweg 400, 3531 AH, Utrecht, the Netherlands

[c] CWI - National Research Institute for Mathematics and Computer Science,
Science Park 123, 1098 XG, Amsterdam, the Netherlands

first version: June 11, 2013
this version: June 11, 2013

Abstract

In this article we propose an efficient Monte Carlo scheme for simulating the stochastic volatility model of Heston [14] enhanced by a non-parametric local volatility component. This hybrid model combines the main advantages of the Heston model and the local volatility model introduced by Dupire [8] and Derman & Kani [7]. In particular, the additional local volatility component acts as a "compensator" that bridges the mismatch between the non-perfectly calibrated Heston model and the market quotes for European-type options. By means of numerical experiments we show that our scheme enables a consistent and fast pricing of products that are sensitive to the forward volatility skew. Detailed error analysis is also provided.

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

1 Introduction

For many years the local volatility (LV) model presented by Dupire [8] and Derman & Kani [7] is considered to be a standard model for pricing and managing risks of structured products. The main advantage of the LV model is its natural modeling of the market implied volatilities for plain vanilla options. As the input for the LV model is an implied Black-Scholes volatility surface, it can be calibrated *exactly* to any given set of arbitrage-free European vanilla option prices. Although well-accepted the LV model has certain limitations, for example, it generates flattening implied forward volatilities [21]. This may lead to a mispricing of financial products like forward-starting or cliquet options. This problem is often addressed by adopting a so-called *sticky-skew technique* which is based on the forward volatilities "as seen today". Alternatives for pricing such products are stochastic volatility (SV) models, like the Heston [14] and Schöbel-Zhu [24] models, where volatility is driven by an additional stochastic differential equation. The SV models are considered to be more accurate choices [9, 11] for pricing forward volatility sensitive derivatives. Additionally, the volatilities in the SV models "move", to a certain extent,

---

*Corresponding author. E-mail address: `A.W.van.der.Stoep@cwi.nl`

independently of a change in spot prices– a property which local volatility models do not have [22]. Although the SV models have desired features for pricing they often cannot be very well calibrated to a given set of arbitrage-free European vanilla option prices. In particular, the accuracy of the Heston model for pricing short-maturity options in the equity market is typically unsatisfactory [9].

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

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

Although the SLV hybrid models are rather new in the financial industry a number of attempts for efficient model evaluation have been made already. Ren *et al.* [22] proposed a stochastic volatility model driven by a lognormal volatility process and developed a tailor-made algorithm for solving the corresponding Kolmogorov forward PDE. An extension of this technique to the Heston SLV was presented in [9] where a finite volume scheme for the model evaluation was used. Although the PDE-discretization techniques are common practice in the financial industry in the context of the local-volatility component, explicit in time discretization methods are typically stable for a very large number of time-grid points requiring significant computational burden. Tian *et al.* engaged a parallel GPU platform to accelerate these computations [27].

The author in [19] moved away from the direct solution of the SLV model and derived via the Markovian projection closed-form approximations to prices of European options on various underlyings. Work on Markovian projections in the context of the SLV models has also been presented in [13], where a so-called "effective local volatility" was derived. The Markovian projections can be widely applied but require a number of conditional expectations to be determined. Very often these expectations are not available analytically and brute-force assumptions need to be imposed so that approximations can be defined [18]. Although mathematically appealing the Markovian projection technique preserves only marginal densities and does not keep marginal distributions of orders higher than one intact. Due to this, prices of securities depending on stock values at multiple times, such as American options and barriers, may significantly differ between the original model and the projected model.

Another attempt for solving the SLV model was presented in [26] where a Levenberg-Marquardt optimization technique for a non-linear Fokker-Planck equation was applied. Another approach for simulation was proposed in [6] by Deelstra and Rayée. By assuming zero correlation between the volatility process and the underlying asset it is possible to efficiently simulate the extended Schöbel-Zhu model.

In this article we solve all necessary model components numerically.

We present a Monte Carlo approach for efficient calibration of the Heston SLV model. In particular, we develop a non-parametric numerical scheme for efficient model evaluation. The scheme is model independent and can be applied to all SLV hybrids, including those based on the SABR model. The technique introduced does not require any advanced methods which makes it intuitive and easy to implement.

The outline of this paper is as follows. In Section 2 we derive the full-scale SLV model and highlight the issues related to efficient model evaluation. Section 3 constitutes the core of this article. We show there how, for a Monte Carlo simulation scheme, *nontrivial* conditional expectations can be evaluated efficiently. We also discuss the simulation of the full-scale model and present how the "unbiased" Monte Carlo scheme for the Heston model [1] can be adopted to the Heston SLV (HSLV) model. In Section 4 some numerical

examples are presented. We particularly concentrate on forward-volatilities implied by the Heston SLV model. Section 5 focuses on the theoretical assessments of the model error and Section 6 concludes.

2 Stochastic-Local Volatility model

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

$$
\begin{aligned}
\mathrm{d}S(t)/S(t) &= r\mathrm{d}t + \sigma(t, S(t))\psi(V(t))\mathrm{d}W_x(t), & (2.1)\\
\mathrm{d}V(t) &= a_v(t, V(t))\mathrm{d}t + b_v(t, V(t))\mathrm{d}W_v(t), & (2.2)\\
\mathrm{d}W_x(t)\mathrm{d}W_v(t) &= \rho_{x,v}\mathrm{d}t, & (2.3)
\end{aligned}
$$

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

Note that the general SLV model described by the system of the SDEs in (2.1) and (2.2) can collapse to either the pure SV model or to the LV model. If we set the local volatility component $\sigma(t, S(t)) = 1$, then the model boils down to a *pure* stochastic volatility model. On the other hand, if the stochastic component of the variance $b_v(t, V(t))$ is equal to 0, the model reduces to a local volatility model. Two popular stochastic volatility models which fit into our framework are the Heston SV model [14] with the variance process driven by the CIR dynamics [4], $\psi(V(t)) = \sqrt{V(t)}$ with $a_v(t, V(t)) = \kappa(\bar{v} - V(t))$ and $b_v(t, V(t)) = \gamma\sqrt{V(t)}$, and the Schöbel-Zhu model [24] with $\psi(V(t)) = V(t)$ and $a_v(t, V(t)) = \kappa(\bar{v} - V(t))$, $b_v(t, V(t)) = \gamma$. Parameter κ controls the speed of mean-reversion, $\bar{v}$ controls a long-term mean and γ determines the volatility of the process $V(t)$.

The SLV model described by Equations (2.1) and (2.2) is yet incomplete as $\sigma(t, S(t))$ is left unspecified. This function can take different forms. It can be, for example, given by the constant elasticity of variance model, i.e. $\sigma(t, S(t)) = \hat{\sigma}S^\beta(t)$, which is a well-known parametric form for describing the volatility movements in terms of the underlying asset $S(t)$. Choosing a parametric form for the local volatility, $\sigma(t, S(t))$, although very flexible and well-accepted, has an undesired feature which is the need for model calibration, i.e. one needs to determine the SV parameters and the LV parameters in the calibration procedure. As the calibration may not always guarantee a sufficiently-well fit to market data we concentrate on non-parametric forms for $\sigma(t, S(t))$ here.

The main concept for deriving a non-parametric LV component $\sigma(t, S(t))$ is as follows: it is well-known that from market data for the European-style options one can determine the market implied density[1], $\hat{f}_S(x)$, of the stock $S(T)$. Furthermore, by deriving the Kolmogorov forward equation for the underlying model we are able to determine the density, $f_S(x)$, of the stock driven by the SDEs (2.1) and (2.2). In a general setting these densities differ and only for a *perfectly* calibrated model they are identical. As in the SLV framework we have one free parameter available, namely $\sigma(t, S(t))$, we may choose the local component so that the densities implied from the market and the model are equal. In the following we derive an expression for the local volatility component $\sigma(t, S(t))$ in the stochastic-local volatility model.

2.1 Specifying $\sigma(t, S(t))$

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

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

---

[1]This result is shown in Lemma 2.1

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

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

$$
\begin{aligned}
\mathrm{d}C(t, K) &= \left(\mathrm{d}\frac{1}{M(t)}\right)\mathbb{E}\left[(S(t) - K)^+\right] + \frac{1}{M(t)}\mathrm{d}\mathbb{E}\left[(S(t) - K)^+\right] \\
&= -\frac{r}{M(t)}\mathbb{E}\left[(S(t) - K)^+\right]\mathrm{d}t + \frac{1}{M(t)}\mathbb{E}\left[\mathrm{d}(S(t) - K)^+\right], \qquad (2.4)
\end{aligned}
$$

where Fubini's theorem justifies the equality $\mathrm{d}\mathbb{E}\left[(S(t) - K)^+\right] = \mathbb{E}\left[\mathrm{d}(S(t) - K)^+\right]$. Regarding the right-hand side in (2.4), we cannot apply Itô's lemma for the evaluation of $\mathrm{d}(S(t) - K)^+$, as the convex function $h(x) = (x - a)^+$ is not differentiable at point $x = a$. Therefore, we will make use of a generalized version of Itô's lemma known as the Tanaka-Meyer formula [15, 20]:

Theorem 2.1 (Tanaka-Meyer formula). *Given a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, $t_0 \leq t < \infty$, let $X(t) = X(t_0) + \widetilde{M}(t) + V(t)$ be a continuous semimartingale, where $\widetilde{M} = \{\widetilde{M}(t), \mathcal{F}(t)\}$ is a continuous local martingale[2], $V = \{V(t), \mathcal{F}(t)\}$ is a càdlàg adapted process[3] of locally bounded variation. Then, for $h(x) = (x - a)^+$ with $a \in \mathbb{R}$:*

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

Proof. A full proof can be found in Tanaka [25]. $\qquad\qquad\square$

Applying the Tanaka-Meyer formula[4], we get

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

which in a differential form is given by:

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

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

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

The dynamics of the call price can be written as:

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

---

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

[3]$V(t)$ is defined on the real numbers (or a subset of them) and is everywhere right-continuous and has left limits everywhere.

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

We simplify this equation by using the equality

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

This gives us the following preliminary result:

Result 2.1. *The dynamics of the European call option price $C(t, K) := C(t_0, t, S(t_0), K)$ with $S(t)$ and $V(t)$ following the dynamics as given in (2.1) and (2.2), respectively, are given by*

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

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

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

Lemma 2.1. *The European call option price $C(t, K) := C(t_0, t, S(t_0), K)$ with $S(t)$ and $V(t)$ following dynamics as given in (2.1) and (2.2), respectively, satisfies*

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

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

Proof. For a proof, see Appendix A. □

We return to the dynamics of the call price given in Result 2.1 where we include the results from Lemma 2.1, i.e.

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

which is equivalent to:

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

We denote by $A(t)$,

$$
\begin{aligned}
A(t) &= \iint_{\mathbb{R}}\delta\left(s - K\right)\sigma^2(t, s)\psi^2(u)s^2 f_{V,S}(u, s)dsdu \\
&= \int_{\mathbb{R}}\psi^2(u)\left(\int_{\mathbb{R}}\delta(s - K)s^2\sigma^2(t, s)f_{V,S}(u, s)ds\right)du. \qquad (2.5)
\end{aligned}
$$

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

$$
\int_{\mathbb{R}}\delta(s - K)s^2\sigma^2(t, s)f_{V,S}(u, s)ds = K^2\sigma^2(t, K)f_{V,S}(u, K). \qquad (2.6)
$$

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

$$
A(t) = K^2\sigma^2(t, K)\int_{\mathbb{R}}\psi^2(u)f_{V,S}(u, K)du, \qquad (2.7)
$$

which is equivalent to:

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

---

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

The dynamics are given by:

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

Using the second equation in Lemma 2.1, we obtain:

$$\mathrm{d}C(t,K) = \left(-rK\frac{\partial C(t,K)}{\partial K} - \frac{1}{2}K^2\sigma^2(t,K)\mathbb{E}\left[\psi^2(V(t))|S(t)=K\right]\frac{\partial^2 C(t,K)}{\partial K^2}\right)\mathrm{d}t, \quad (2.8)$$

which can be expressed as:

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

where $\sigma_{\mathrm{LV}}(t,K)$ denotes Dupire's local volatility [8]. We eventually find the following relation:

$$\sigma^2(t,K) = \frac{\sigma_{\mathrm{LV}}^2(t,K)}{\mathbb{E}\left[\psi^2(V(t))|S(t)=K\right]}. \quad (2.9)$$

The local volatility component $\sigma^2(t,K)$ consists of two ingredients: the deterministic local volatility $\sigma_{\mathrm{LV}}(t,K)$ and the conditional expectation $\mathbb{E}[\psi^2(V(t))|S(t)=K]$. Numerical evaluation of $\sigma_{\mathrm{LV}}(t,K)$ is already well-established in the literature, see for example [2,3,5]. On the other hand, the efficient computation of the conditional expectation in (2.9) is not yet established. The difficulty lies in the fact that the joint distribution of the variance V and the stock S, $f_{V,S}$, is unknown. This is due to the fact that the stock process, $S(t)$, contains a local-volatility component $\sigma_{\mathrm{LV}}(t,S)$ which is also not known analytically. The evaluation of the unknown expectation can be either derived by solving a Kolmogorov forward PDE (e.g. [6,9]) or by applying a Markovian projection approximation [13,18,19]. In this article we concentrate on Monte Carlo evaluation of the stochastic-local volatility model. In the next section we present a numerical method which leads to efficient Monte Carlo model evaluation.

3   Novel technique for $\mathbb{E}\left[\psi^2(V(t))|S(t)=K\right]$

In this section we present a new efficient evaluation of a general stochastic-local volatility model. In particular, by an Euler discretization we simulate the SLV model (2.1)-(2.2) as follows:

$$s_{i+1,j} = s_{i,j} + rs_{i,j}\Delta + \sigma(t_i,s_{i,j})s_{i,j}\psi(v_{i,j})\sqrt{\Delta}Z_x, \quad s_{0,j} = S(t_0), \quad (3.1)$$

$$v_{i+1,j} = v_{i,j} + a_v(t_i,v_{i,j})\Delta + b_v(t_i,v_{i,j})\sqrt{\Delta}Z_v, \quad v_{0,j} = v(t_0), \quad (3.2)$$

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

Using expression (2.9) for $\sigma(t,S)$, System (3.1)-(3.2) becomes:

$$s_{i+1,j} = s_{i,j} + rs_{i,j}\Delta + \sqrt{\frac{\sigma_{\mathrm{LV}}^2(t_i,s_{i,j})}{\mathbb{E}\left[\psi^2(V(t_i))|S(t_i)=s_{i,j}\right]}}s_{i,j}\psi(v_{i,j})\sqrt{\Delta}Z_x, \quad (3.3)$$

$$v_{i+1,j} = v_{i,j} + a_v(t_i,v_{i,j})\Delta + b_v(t_i,v_{i,j})\sqrt{\Delta}Z_v. \quad (3.4)$$

To determine the values of the paths for the *next* time-step, t_{i+1}, one needs to establish two main components, $\sigma_{\mathrm{LV}}^2(t_i,s_{i,j})$ and $\mathbb{E}\left[\psi^2(V(t_i))|S(t_i)=s_{i,j}\right]$. As indicated efficient evaluation of $\sigma_{\mathrm{LV}}^2(t_i,s_{i,j})$ is already well-established in the literature [2,5]. This is not the

case for evaluation of the conditional expectation. The main difficulty in its evaluation is that the conditioning has to be performed on each individual stock realization $s_{i,j}$, i.e. as we simulate a discretized system for (S, V), each realization of $s_{i,j}$ has exactly one corresponding realization of the variance $v_{i,j}$ and this makes the evaluation of the conditional expectation difficult.

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

3.1 Non-parametric method

Suppose that for a given time t_i, $i = 1, \ldots, M$ we have N pairs of Monte Carlo realizations $(s_{i,1}, v_{i,1})$, $(s_{i,2}, v_{i,2})$, $\ldots$, $(s_{i,N}, v_{i,N})$ for which we wish to evaluate the conditional expectation in (3.3). As for each $s_{i,j}$ we have exactly one value $v_{i,j}$ the conditional expectation will always be equal to $\psi^2(v_{i,j})$, which is undesired. Such a problem is a natural consequence of discretization of the continuous system (S, V). Obviously, in order to obtain an accurate estimate we would need to have an infinite set of paths, which is practically unfeasible.

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

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

$$\mathbb{E}\left[\psi^2(V(t_i))|S(t_i) = s_{i,j}\right] \approx \mathbb{E}\left[\psi^2(V(t_i))|S(t_i) \in (b_k, b_{k+1}]\right]. \tag{3.5}$$

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

$$
\begin{aligned}
\mathbb{E}\left[\psi^2(V(t_i))|S(t_i) = s_{i,j}\right] &= \lim_{\varepsilon \to 0^+} \mathbb{E}\left[\psi^2(V(t_i))|S(t_i) \in (s_{i,j} - \varepsilon, s_{i,j} + \varepsilon]\right] \\
&= \lim_{\varepsilon \to 0^+} \frac{\mathbb{E}\left[\psi^2(V(t_i))\mathbb{1}_{S(t_i) \in (s_{i,j}-\varepsilon, s_{i,j}+\varepsilon]}\right]}{\mathbb{Q}\left[S(t_i) \in (s_{i,j} - \varepsilon, s_{i,j} + \varepsilon]\right]}.
\end{aligned} \tag{3.6}
$$

In the limiting case where both boundaries of the bin are equal to $s_{i,j}$ the approximation of the conditional expectation boils down to its exact value. This is an indication for the appropriateness of the approximation in (3.5). The open question that remains is how to choose *proper* bin boundaries b_k for $k = 1, \ldots, \ell + 1$. We consider two following choices in a Monte Carlo simulation framework.

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

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

$$b_{i,k} = \bar{s}_{i,1} + \frac{k-1}{\ell}(\bar{s}_{i,N} - \bar{s}_{i,1}), \quad k = 1 \ldots \ell + 1. \tag{3.7}$$

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

$$b_{i,1} = \bar{s}_{i,1}, \quad b_{i,\ell+1} = \bar{s}_{i,N}, \quad b_{i,k} = \bar{s}_{i,(k-1)N/\ell}, \quad k = 2 \ldots \ell. \tag{3.8}$$

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

$$
\begin{aligned}
\mathbb{E}\left[\psi^2(V(t_i)) \middle| S(t_i) = s_{i,j}\right] &\approx \frac{\mathbb{E}\left[\psi^2(V(t_i)) \mathbb{1}_{S(t_i) \in \left(b_{i,k}, b_{i,k+1}\right]}\right]}{\mathbb{Q}\left[S(t_i) \in (b_{i,k}, b_{i,k+1}]\right]} \\
&\approx \frac{\frac{1}{N} \sum_{j=1}^{N} \psi^2(v_{i,j}) \mathbb{1}_{s_{i,j} \in \left(b_{i,k}, b_{i,k+1}\right]}}{\mathbb{Q}\left[S(t_i) \in (b_{i,k}, b_{i,k+1}]\right]} \\
&= \frac{1}{N\alpha(k)} \sum_{j \in \mathcal{J}_{i,k}} \psi^2(v_{i,j}), \qquad (3.9)
\end{aligned}
$$

where $\alpha(k) := \mathbb{Q}\left[S(t_i) \in (b_{i,k}, b_{i,k+1}]\right]$ represents the probability of the stock being in the kth bin. The second approximation is established by switching between the expectation and the average, which is based on a finite number of $(s_{i,j}, v_{i,j})$-pairs. The value of $\alpha(k)$ depends on the way the bins are chosen:

$$
\alpha(k) = \left\{ \begin{array}{ll} N_k/N, \ k = 1 \ldots \ell & \text{for bins defined as in (3.7),} \\ 1/\ell & \text{for bins defined as in (3.8).} \end{array} \right.
$$

Remark 3.1. *As it will be shown in Section 5.3, the choice of bins affects the convergence of the non-parametric method. If we define the bins according to (3.8), bins close to the mean of the joint density are much smaller than bins in the tails. This is desirable as the region close to the mean contains many more observations, requiring a higher accuracy and thus smaller bin sizes. We will choose bins according to (3.8) in numerical experiments.*

We summarize the non-parametric method in Algorithm 1.

---

for *each time-step $t_i, i = 1 \ldots M$* **do**

1   Generate N pairs of observations $(s_{i,j}, \psi^2(v_{i,j}))$, $j = 1 \ldots N$.
2   Order the elements $\bar{s}_{i,j}$: $\bar{s}_{i,1} \leq \bar{s}_{i,2} \leq \cdots \leq \bar{s}_{i,N}$.
3   Determine the boundaries of ℓ bins $(b_{i,k}, b_{i,k+1}]$, $k = 1 \ldots \ell$ according to either (3.7) or (3.8).
4   For the kth bin approximate the conditional expectation by
     $\mathbb{E}\left[\psi^2(V(t_i)) \middle| S(t_i) \in (b_{i,k}, b_{i,k+1}]\right] \approx \frac{1}{N\alpha(k)} \sum_{j \in \mathcal{J}_{i,k}} \psi^2(v_{i,j})$, where $\mathcal{J}_{i,k}$ is the set of path numbers j for which the observations are in the kth bin at the ith time-step and $\alpha(k)$ represents the probability of the stock being in the kth bin, which is determined by the choice of bins.

end

Algorithm 1: Non-parametric method

We present two illustrative examples where this method for calculating the conditional expectation is applied. Error analysis will be discussed in Section 5.

Experiment 3.1 (Illustrative examples)**.** In order to illustrate how the introduced algorithm works we present two experiments. First, we consider a simple Monte Carlo simulation consisting of 9 paths. In the second experiment we apply the algorithm to calculate the conditional expectation for the Heston model. We start with some assumptions. The initial values for the stock and variance process are $S(t_0) = s_0 = 1$ and $V(t_0) = v_0 = 0.1$, respectively.

In order to obtain path realizations at time t_1 we need to determine the conditional expectation $\mathbb{E}[V(t_0)|S(t_0) = s_0]$, which trivially gives $\mathbb{E}[V(t_0)|S(t_0) = s_0] = v_0 = 0.1$. This holds for all paths in this experiment. To determine the paths at time t_2 we calculate

the expectation $\mathbb{E}[V(t_1)|S(t_1) = s_{1,j}]$, $j = 1 \ldots 9$, as follows. First, we choose the number of bins, $\ell = 3$, and sort the pairs $(s_{1,j}, v_{1,j})$ according to their $s_{1,j}$ values. Then, we assign each pair to a bin – also according to the $s_{1,j}$ values. Finally, we calculate for each bin an approximation of the conditional expectation. The procedure is illustrated in the tables below:

j	t_0	t_1		j	t_1	
	(s_0, v_0)	$\mathbb{E}[V(t_0)]$	$(s_{1,j}, v_{1,j})$		$(s_{1,j}, v_{1,j})$	
1	$(1, 0.1)$	0.1	$(1.9, 0.09)$	9	$(0.4, 0.25)$	
2	$(1, 0.1)$	0.1	$(0.9, 0.15)$	4	$(0.5, 0.20)$	$\Rightarrow \mathbb{E}[V(t_1)|S(t_1) \in (0, 0.9]]$
3	$(1, 0.1)$	0.1	$(1.2, 0.15)$	2	$(0.9, 0.15)$	$\approx \frac{1}{3} \sum_j v_{1,j} = 0.2,$
4	$(1, 0.1)$	0.1	$(0.5, 0.20)$ $\Rightarrow$	6	$(1.1, 0.07)$	
5	$(1, 0.1)$	0.1	$(1.6, 0.06)$	3	$(1.2, 0.15)$	$\Rightarrow \mathbb{E}[V(t_1)|S(t_1) \in (0.9, 1.3]]$
6	$(1, 0.1)$	0.1	$(1.1, 0.07)$	8	$(1.3, 0.08)$	$\approx \frac{1}{3} \sum_j v_{1,j} = 0.1,$
7	$(1, 0.1)$	0.1	$(1.7, 0.05)$	5	$(1.6, 0.06)$	
8	$(1, 0.1)$	0.1	$(1.2, 0.08)$	7	$(1.7, 0.05)$	$\Rightarrow \mathbb{E}[V(t_1)|S(t_1) \in (1.3, 1.9]]$
9	$(1, 0.1)$	0.1	$(0.4, 0.25)$	1	$(1.9, 0.09)$	$\approx \frac{1}{3} \sum_j v_{1,j} = 0.067.$

Let us consider a more practical example. We consider the Heston stochastic volatility model. In this model $\psi(x) = \sqrt{x}$ and the conditional expectation reads $\mathbb{E}[V(t_i)|S(t_i) = s_{i,j}]$. The reason for considering the *pure* Heston model is that we are able to determine the conditional expectation by the 2D-COS method [23]. A discussion about this calculation is provided in Section 5.2.

In Figure 3.1 we compare the results for the conditional expectation from the proposed scheme and the reference obtained by Fourier expansions. Each plot includes a contour plot of the recovered density, the corresponding conditional expectation, and its approximation. In the simulations we considered $1e5$ Monte Carlo paths paths with 2, 5 and 20 bins, respectively. We choose bins that contain equal numbers of realizations. This yields smaller bins close to the mean of the joint density, see also Remark 3.1. The approximation obtained by the algorithm introduced converges to the reference.
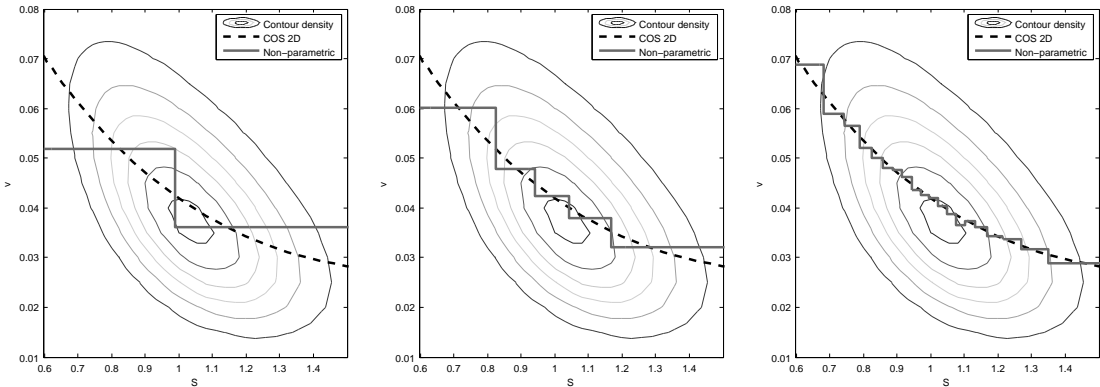

Figure 3.1: The approximation obtained by the non-parametric method converges to the conditional expectation recovered by the COS method as the number of bins increases (2, 5 and 20 respectively).

Remark 3.2. *An alternative method for estimating* $\mathbb{E}\left[\psi^2(V(t_i))|S(t_i) = s_{i,j}\right]$ *is by projecting* $\psi^2(V(t_i))$ *on a set of orthogonal polynomials* $\zeta_k(\cdot)$, $k = 1 \ldots n$: $\psi^2(V(t_i)) = \sum_{k=1}^n \beta_k \zeta_k(S(t_i)) + \epsilon$. *By regressing* $\psi^2(V(t_i))$ *on functions* $\zeta_k(\cdot)$ *the conditional expectation can be calculated, as*

$$\mathbb{E}\left[\psi^2(V(t_i))|S(t_i) = s_{i,j}\right] \approx \mathbb{E}\left[\sum_{k=1}^n \beta_k \zeta_k(S(t_i)) \middle| S(t_i) = s_{i,j}\right] = \sum_{k=1}^n \beta_k \zeta_k(s_{i,j}). \quad (3.10)$$

The above approximation for $\psi^2(V(t_i))$ is based on the assumption that the conditional expectation is an element of the L^2-space of square integrable functions (relative to a particular measure). The conditional expectation can be represented as a linear function of elements of a countable orthonormal basis. Applying the approximation in (3.10) the discrete scheme described by (3.3) and (3.4) becomes:

$$
\begin{aligned}
s_{i+1,j} &= s_{i,j} + r s_{i,j}\Delta + \sqrt{\frac{\sigma_{LV}^2(t_i, s_{i,j})}{\sum_{k=1}^n \beta_k \zeta_k(s_{i,j})}} s_{i,j}\psi(v_{i,j})\sqrt{\Delta}Z_x, \\
v_{i+1,j} &= v_{i,j} + \kappa(\bar{v} - v_{i,j})\Delta + \gamma\sqrt{v_{i,j}}\sqrt{\Delta}Z_v,
\end{aligned} \tag{3.11}
$$

where κ is the speed of mean-reversion, $\bar{v}$ is the long-run variance and γ is the volatility of the variance process ('vol-vol'). Although intuitive and straightforward, the regression-based alternative possesses the drawback that the Feller condition must be satisfied to guarantee a positive conditional expectation for the whole range of arguments. We show a numerical test in Figure 3.2, where we consider a simple quadratic polynomial: $\zeta_1(x) = 1$, $\zeta_2(x) = x$ and $\zeta_3(x) = x^2$.
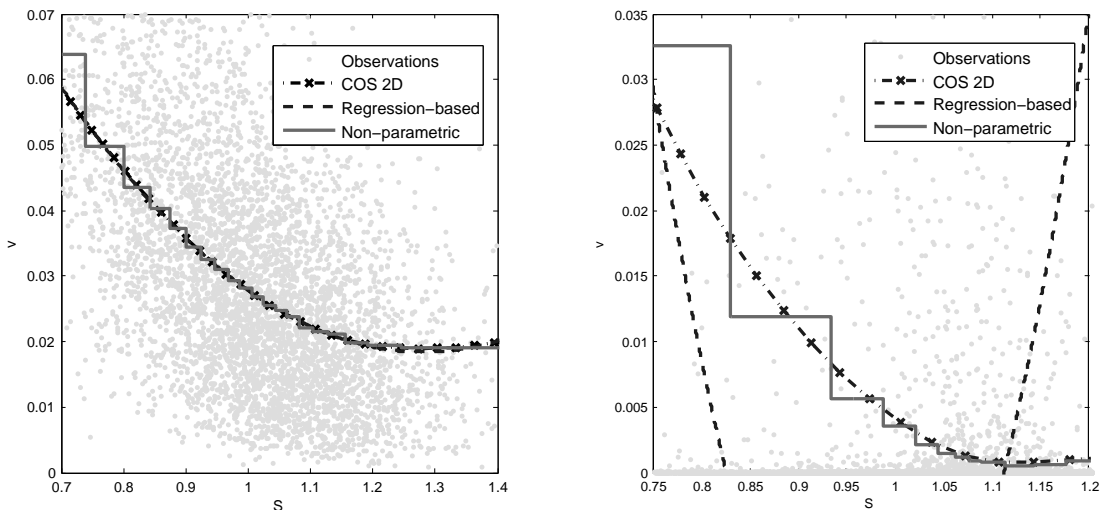

Figure 3.2: The regression-based alternative: Feller satisfied (left) and not satisfied (right). Feller must be satisfied in order to guarantee a non-negative approximation of the conditional expectation. The non-parametric method does not suffer from this restriction, as we see on the right.

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

3.2 Continuous approximation

As the theoretical conditional expectation is continuous, we want its approximation to satisfy this property too. Furthermore, at the right-hand side of Figure 3.2 we observe that at the left boundary of the strike range the fit of the non-parametric approximation to the reference may be improved. In order to obtain a continuous approximation that establishes this, we consider a linearization of the estimated expectation obtained by the non-parametric method. This can be done by connecting the mid-points of the approximations of the non-parametric method, see Figure 3.3.
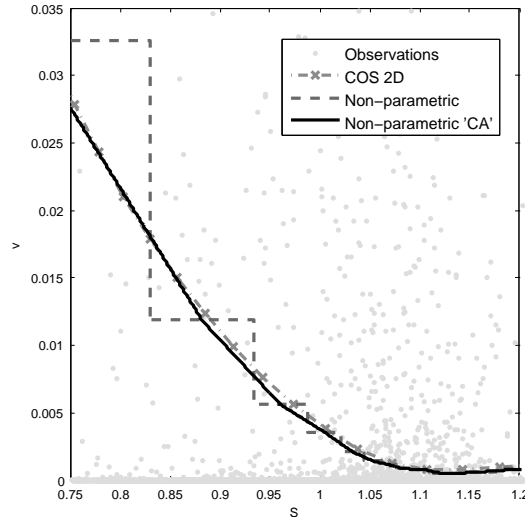
Figure 3.3: The continuous approximation ('CA') gives a better fit to the theoretical conditional expectation, which is recovered by the COS method.

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

3.3 Efficient simulation scheme

The CIR-type process used for the dynamics of the variance in the Heston model does not allow for negative realizations. Unfortunately, when applying the basic Euler discretization scheme the variance process can become negative with non-zero probability. Although several fixes like "absorption at zero" for handling negative variance realizations are known in the literature (see [16] for an overview), these improved methods are typically not free of bias.

In this section we adopt the Quadratic Exponential (QE) scheme introduced by Andersen [1] and apply it for simulating the Heston SLV model. The main difference for Monte Carlo simulation between the pure Heston and the Heston SLV models lies in the fact that the variance of the latter is not only driven by the stochastic volatility, but also by the local volatility component, which is state-dependent. This requires an additional "*freezing approximation*", which is not present in the derivation of the original QE scheme. Numerical experiments show that the additional approximation still yields an accurate simulation scheme.

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

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

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

$$
X(t + \Delta) = X(t) + \int_t^{t+\Delta} \left(r - \frac{1}{2}\widehat{\sigma}^2(s, X(s))V(s)\right)\mathrm{d}s \tag{3.12}
$$

$$
+\rho_{x,v}\int_t^{t+\Delta} \widehat{\sigma}(s, X(s))\sqrt{V(s)}\mathrm{d}\widetilde{W}_v(s) + \sqrt{1 - \rho_{x,v}^2}\int_t^{t+\Delta} \widehat{\sigma}(s, X(s))\sqrt{V(s)}\mathrm{d}\widetilde{W}_x(s).
$$

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

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

11

with

$$c(\Delta) = \frac{\gamma^2}{4\kappa}(1 - \mathrm{e}^{-\kappa\Delta}), \ d = \frac{4\kappa\bar{V}}{\gamma^2}, \ \lambda(t, V(t)) = \frac{4\kappa\mathrm{e}^{-\kappa\Delta}}{\gamma^2(1 - \mathrm{e}^{-\kappa\Delta})}V(t) \qquad (3.13)$$

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

$$\int_t^{t+\Delta} \sqrt{V(s)}\mathrm{d}\widetilde{W}_v(s) = \frac{1}{\gamma}\left(V(t+\Delta) - V(t) - \kappa\bar{V}\Delta + \kappa\int_t^{t+\Delta} V(s)\mathrm{d}s\right). \qquad (3.14)$$

In the last integral in (3.12) the local and stochastic volatilities are coupled. This complicates the simulation as we are not able to directly use the integrated variance from (3.14). As any Monte Carlo simulation involving a local-volatility component requires many time-steps, we perform local-freezing of $\widehat{\sigma}(s, X(s))$ in (3.12), i.e.

$$\int_t^{t+\Delta} \widehat{\sigma}(s, X(s))\sqrt{V(s)}\mathrm{d}\widetilde{W}_v(s) \approx \widehat{\sigma}(t, X(t))\int_t^{t+\Delta} \sqrt{V(s)}\mathrm{d}\widetilde{W}_v(s). \qquad (3.15)$$

Due to the approximation in (3.15) we can use (3.14) in (3.12):

$$
\begin{aligned}
X(t+\Delta) \quad \approx \quad & X(t) + \int_t^{t+\Delta}\left(r - \frac{1}{2}\widehat{\sigma}^2(s, X(s))V(s)\right)\mathrm{d}s \\
& + \frac{\rho_{x,v}\widehat{\sigma}(t, X(t))}{\gamma}\left(V(t+\Delta) - V(t) - \kappa\bar{V}\Delta + \kappa\int_t^{t+\Delta} V(s)\mathrm{d}s\right) \\
& + \sqrt{1 - \rho_{x,v}^2}\int_t^{t+\Delta} \widehat{\sigma}(s, X(s))\sqrt{V(s)}\mathrm{d}\widetilde{W}_x(s).
\end{aligned}
\qquad (3.16)
$$

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

$$
\begin{aligned}
X(t+\Delta) \quad \approx \quad & X(t) + r\Delta - \frac{1}{2}\widehat{\sigma}^2(t, X(t))V(t)\Delta \\
& + \frac{1}{\gamma}\rho_{x,v}\widehat{\sigma}(t, X(t))\left(V(t+\Delta) - V(t) - \kappa\bar{V}\Delta + \kappa V(t)\Delta\right) \\
& + \sqrt{1 - \rho_{x,v}^2}\int_t^{t+\Delta} \widehat{\sigma}(s, X(s))\sqrt{V(s)}\mathrm{d}\widetilde{W}_x(s).
\end{aligned}
$$

Furthermore, by the Itô isometry we have

$$\int_t^{t+\Delta} \widehat{\sigma}(s, X(s))\sqrt{V(s)}\mathrm{d}\widetilde{W}_x(s) \sim \widetilde{Z}_x\sqrt{\int_t^{t+\Delta} \widehat{\sigma}^2(s, X(s))V(s)\mathrm{d}s}, \qquad (3.17)$$

where $\widetilde{Z}_x \sim N(0, 1)$. The integral at the right-side of (3.17) can also be approximated by the Euler discretization, i.e. $\int_t^{t+\Delta} \widehat{\sigma}^2(s, X(s))V(s)\mathrm{d}s \approx \widehat{\sigma}^2(t, X(t))V(t)\Delta$. With this, the discretization scheme becomes

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

with $\rho_1 = (1 - \rho_{x,v}^2)^{1/2}$, $c_1 = \kappa\Delta - 1$, where $c(\Delta)$, d and $\lambda(t, V(t))$ are defined in (3.13) and

$$\widehat{\sigma}^2(t_i, x_{i,j}) \stackrel{\text{def}}{=} \sigma^2(t_i, e^{x_{i,j}}) = \frac{\sigma_{\mathrm{LV}}^2(t_i, s_{i,j})}{\mathbb{E}\left[V(t_i)|S(t_i) = s_{i,j}\right]}. \qquad (3.18)$$

12

In (3.18) we compute Dupire's local volatility,

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

by using the following finite difference approximations:

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

and

$$\frac{\partial^2 C(t_i, s)}{\partial s^2}\bigg|_{s=s_{i,j}} \approx \frac{C(t_i, s_{i,j} + h_2) - 2C(t_i, s_{i,j}) + C(t_i, s_{i,j} - h_2)}{h_2^2}. \tag{3.19}$$

For stability reasons the derivatives are often expressed in terms of implied volatilities [6]. As in practice a continuum of European call prices in time-to-maturity and strike is not available, some interpolation may be required. Detailed discussion on this is provided in the literature (e.g. [2]).

Numerical comparisons between the Euler and the original QE scheme have been provided in the literature [1]. We perform an experiment with our version of the Monte Carlo scheme in the follow-up section, because our scheme is slightly different due to the local volatility component which requires an additional approximation (local freezing of the state-dependent local volatility).

Experiment 3.2 (Efficient simulation scheme). We compare the scheme we propose with the basic Euler discretization scheme. We consider parameter values based on Case III of Andersen [1], i.e. for $T = 5$ we consider the Heston SLV model with $\kappa = 1.05$, $\gamma = 0.95$, $\bar{v} = 0.0855$, $v_0 = 0.0945$, $\rho_{x,v} = -0.315$ and $r = 0$. We perform a Monte Carlo simulation consisting of 20 seeds, $5e4$ paths per seed. The number of bins is set to 20.

For different time-step sizes Δ and strikes K we calculate the absolute error in implied volatilities $|\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{SLV}}|$, where $\bar{\sigma}_{\text{market}}$ and $\bar{\sigma}_{\text{SLV}}$ denote volatilities implied by the market and the HSLV model, respectively.

Results are presented in Table 1. The efficient ('low-bias') simulation scheme outperforms the Euler scheme: it gives a higher accuracy and we observe faster convergence to the reference for a decaying time-step size.

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

Table 1: Average error $|\bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{SLV}}|$ from Monte Carlo simulations of the HSLV model with the Euler and efficient ('low-bias') schemes using 20 random seeds, for multiple time-step sizes Δ and strikes K. Numbers in parentheses are standard deviations over the seeds.

4 Numerical results

Using the simulation scheme introduced in the previous section, we perform Monte Carlo experiments for the pricing of European call and forward-starting options with the Heston SLV model. The results are compared against the Heston and the standard local volatility models.

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

The simulation of the European-style options is performed for maturities (in years) $T = \{0.5, 2, 5, 8, 10\}$, while the pricing of the forward starting options will be done for the following pairs: $\{T_1, T_2\} = \{2, 4\}$ and $\{T_1, T_2\} = \{6, 8\}$. Our Monte Carlo simulation is performed with $5e5$ paths and 100 time-steps per year. The number of bins is set to 20. In Section 5.3 we show that the accuracy of the non-parametric method is already satisfactory for a smaller number of bins.

4.1 European call options

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

In the first experiment we consider the Heston SLV model with the well-calibrated underlying Heston model. For times to maturity 2 and 8 years we display the results in Figure 4.1. As the mismatch between the pure Heston model and the market is small, the contribution of the local-volatility component $\sigma_{LV}(t, S(t))$ is limited in the first test.
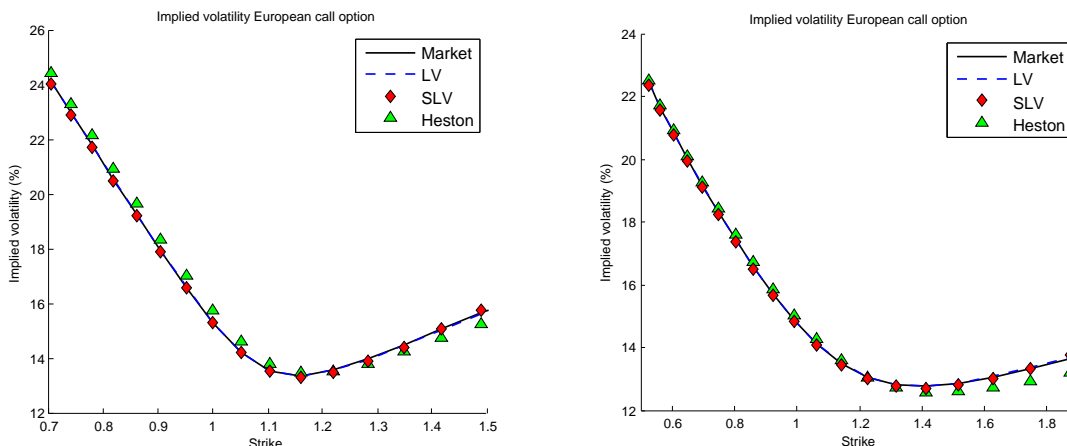

Figure 4.1: Implied volatility European call option. $T = 2$ (left), $T = 8$ (right), well calibrated Heston model.

In Figure 4.2 we display results of the second experiment, in which the Heston model is insufficiently calibrated. We observe that in this case the local volatility term can compensate for the large gap between the market and the Heston model.
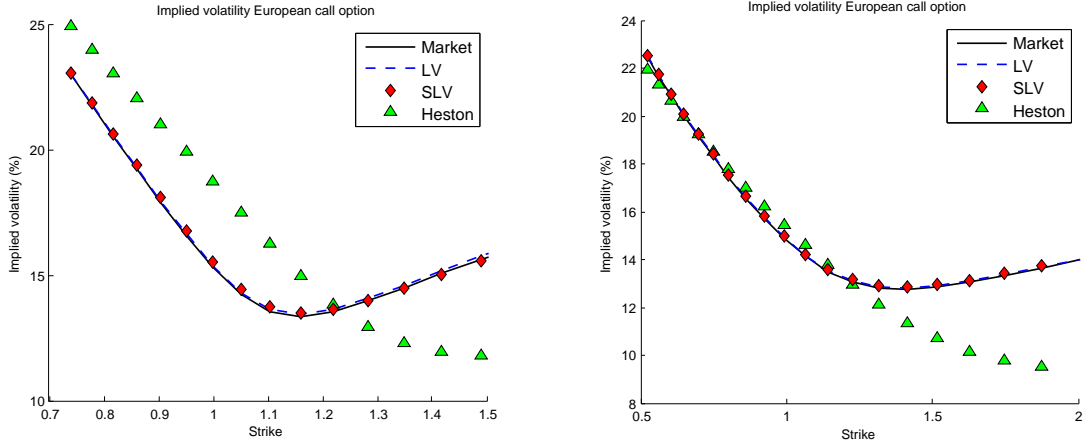
Figure 4.2: Implied volatility European call option. $T = 2$ (left), $T = 8$ (right), insufficiently calibrated Heston model.

For times to maturity 0.5, 2, 5, 8 and 10 years results are given in Table 2 where we display the market implied volatility $\bar{\sigma}_{\text{market}}$ and the error in implied volatilities $\epsilon_{\text{model}} := \bar{\sigma}_{\text{market}} - \bar{\sigma}_{\text{model}}$, where $\bar{\sigma}_{\text{model}}$ denotes the volatility implied by a particular model. We observe a good fit of the LV model as well as the Heston SLV model to the quotes.

4.2 Forward starting options

With the pricing of European call options we see that the local volatility term in the Heston SLV model acts a compensator that bridges the gap between the market and calibrated Heston SV prices – even in the case of an unsatisfactory calibration. In this experiment we price forward starting options that start at time T_1 years and mature at time T_2. As prices of forward starting options are not observable in the market, we discriminate between the LV, SLV and the calibrated Heston models. We first consider the case with $T_1 = 2$, $T_2 = 4$ and a well calibrated Heston model, see the plot on the left side in Figure 4.3. We observe

Figure 4.3: Implied volatility forward starting option. $T_1 = 2$, $T_2 = 4$, well calibrated (left) and insufficiently well calibrated (right) Heston model.

that the volatility implied by the LV model is much flatter than the volatilities implied by the Heston SV and SLV models. Although it has approximately the same value at the lower and upper bounds of the strike range, this does not hold in the ATM region. As the Heston model is almost perfectly calibrated, it is no surprise that the Heston and SLV implied volatilities are almost identical.

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

European call options – error in implied volatilities [%]

Table 2: Errors in implied volatilities for the local volatility (ϵ_{LV}), Heston SLV (ϵ_{SLV}) and the pure Heston (ϵ_{H}) models for a well (Case I) and insufficiently (Case II) calibrated Heston model for multiple times to maturity and strikes.

For the case where the Heston model is insufficiently calibrated, we observe that the SLV model provides a forward smile that is located between the ones implied by the Heston and the LV models. One may consider the results by the SLV model to represent somehow "advanced interpolation" between the Heston and the local volatility models. For $T_1 = 6$ and $T_2 = 8$ similar results are observed, see Figure 4.4.

4.3 Calculation time

Considering the speed of the non-parametric method, we calculate the time it takes to evaluate the conditional expectation for a given number of paths or bins. First, we investigate the relation between calculation time and the number of bins. We fix the number of paths at $1e5$. Table 3 shows that the calculation time behaves linearly in the number of bins. This also holds for the relation between the calculation time and the number of paths (the number of bins is fixed at 20), see Table 4.

In our numerical experiments we choose 20 bins and $1e5$ paths, which implies that the calculation of the conditional expectation by the non-parametric method takes less than $2.5e\text{-}2$ seconds per time step[6]. In Section 5.3 we consider the dependence of the accuracy and the number of bins in more detail for the non-parametric method. A result of this analysis is that the accuracy of the method is quite insensitive to the number of bins.

---

[6]Theoretically, both experiments should result in exactly the same calculation time, i.e., the second time in Table 3 should be equal to the first time in Table 4. This is not the case due to Monte Carlo noise.
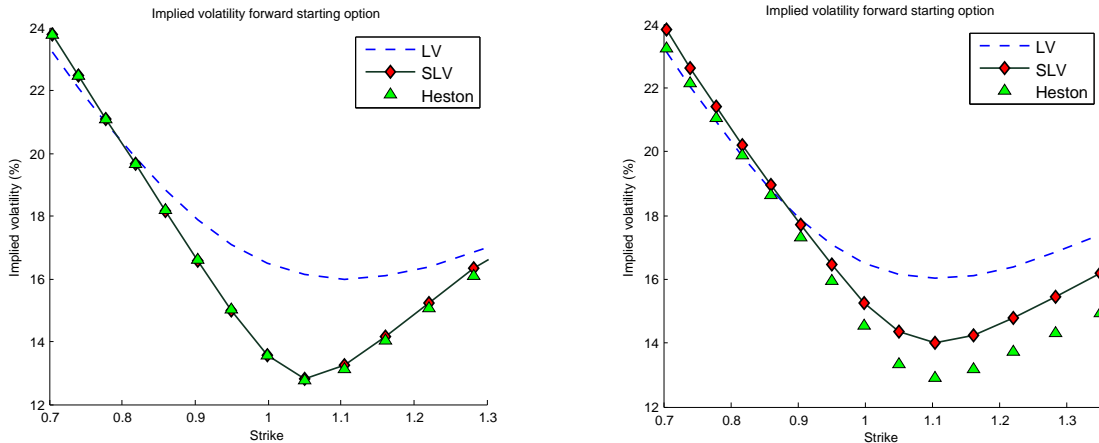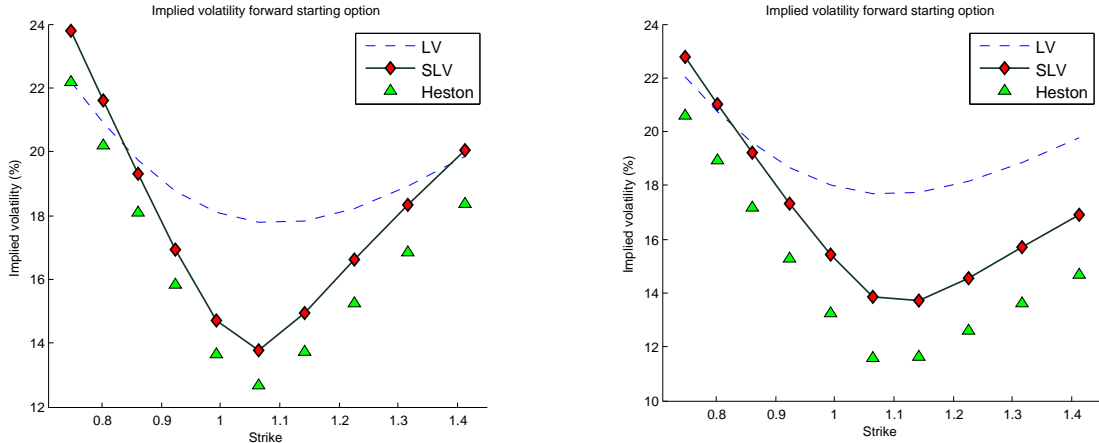
Figure 4.4: Implied volatility forward starting option. $T_1 = 6$, $T_2 = 8$, well calibrated (left) and poorly calibrated (right) Heston model.

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

Table 3: Timing results for different numbers of **bins** (number of paths fixed at $1e5$).

5 Error analysis

In Section 3 we tested the performance of the non-parametric method (and the regression-based alternative) for a pure Heston SV model. We considered this model, as the 2D-COS method [23] provided us with an accurate approximation of the conditional expectation. In particular, we compared an approximation from the non-parametric method to the recovered conditional expectation. However, our main interest lies in the Heston SLV model. In the Monte Carlo pricing of European call options and forward starting options a bias is introduced, which is due to three error sources. The first error originates from approximations in the calculation of Dupire's local volatility term (3.18). In particular, we use finite differences for the three derivatives in (3.19). Further, the discretization of the continuous dynamics to the efficient simulation scheme introduced a discretization error. Last, at each time step we approximate $\mathbb{E}[V(t)|S(t) = s]$ by means of the non-parametric method. These three sources of error generate an error $e = C - \widetilde{C}$, where C is the call price from the "original" Heston SLV model and $\widetilde{C}$ is the price obtained by the discrete Heston SLV model.

Ignoring the bias originating from the finite differences and the discretization (since these errors are well understood), the price mismatch is driven by the difference in conditional expectations $||g - \widehat{g}||$, based on the governing PDEs, as in [12]. Here $g(s) := \mathbb{E}[V(t)|S(t) = s]$ denotes the theoretical conditional expectation (in the Heston SLV model) and $\widehat{g}$ is its piecewise linear continuous approximation we obtain by the non-parametric method. We now provide a pricing error bound that is implied by the mismatch between g and $\widehat{g}$.

5.1 Bound on pricing error

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

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

17

#Paths [1e5]	1	2	3	4	5	6	7	8	9	10
Time [ms]	28.0	51.1	77.9	105.0	131.5	158.9	185.9	214.1	244.3	269.0

Table 4: Timing results for different numbers of **paths** (number of bins fixed at 20).

$$
\begin{aligned}
0 =\ & \frac{\partial C}{\partial t} + rs\frac{\partial C}{\partial s} + \kappa(\bar{V} - V)\frac{\partial C}{\partial V} + \frac{1}{2}Vs^2\frac{\sigma_{\mathrm{LV}}^2(t,s)}{g(s)}\frac{\partial^2 C}{\partial s^2} + \frac{1}{2}\gamma^2 V\frac{\partial^2 C}{\partial V^2} \\
& + \rho_{x,v}\gamma V s\sqrt{\frac{\sigma_{\mathrm{LV}}^2(t,s)}{g(s)}}\frac{\partial^2 C}{\partial s\partial V} - rC,
\end{aligned}
\tag{5.1}
$$

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

Since the expectation in (5.1) is not known analytically we can estimate it by means of Monte Carlo simulation. The resulting, approximating pricing PDE then reads:

$$
\begin{aligned}
0 =\ & \frac{\partial \widetilde{C}}{\partial t} + rs\frac{\partial \widetilde{C}}{\partial s} + \kappa(\bar{V} - V)\frac{\partial \widetilde{C}}{\partial V} + \frac{1}{2}Vs^2\frac{\sigma_{\mathrm{LV}}^2(t,s)}{\widehat{g}(s)}\frac{\partial^2 \widetilde{C}}{\partial s^2} + \frac{1}{2}\gamma^2 V\frac{\partial^2 \widetilde{C}}{\partial V^2} \\
& + \rho_{x,v}\gamma V s\sqrt{\frac{\sigma_{\mathrm{LV}}^2(t,s)}{\widehat{g}(s)}}\frac{\partial^2 \widetilde{C}}{\partial s\partial V} - r\widetilde{C},
\end{aligned}
\tag{5.2}
$$

with the pre-calibrated function $\widehat{g}(s)$ as described in Section 3.1. The PDEs in (5.1) and (5.2) can be written, in shorthand notation, as follows:

$$
\frac{\partial C}{\partial t} + L_1 C = 0, \qquad \frac{\partial \widetilde{C}}{\partial t} + L_2 \widetilde{C} = 0,
\tag{5.3}
$$

with the corresponding operators L_1, as in (5.1) and L_2, as in (5.2). Again, C is the solution from the full-scale HSLV PDE, whereas $\widetilde{C}$ is the solution from the approximating PDE with the estimated function $\widehat{g}(t)$. Both PDEs are accompanied by the same boundary and final conditions. For the error, $e := C - \widetilde{C}$, we find:

$$
\frac{\partial e}{\partial t} + L_1 C - L_2 \widetilde{C} = 0,
\tag{5.4}
$$

which can be re-written as:

$$
\frac{\partial e}{\partial t} + L_1 C - \left(L_1\widetilde{C} + (L_2 - L_1)\widetilde{C}\right) = 0,
\tag{5.5}
$$

and we arrive at the following equation:

$$
\frac{\partial e}{\partial t} + L_1 e = (L_2 - L_1)\widetilde{C},
\tag{5.6}
$$

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

Based on the form in (5.6), multiplying both sides by e, and integration over domain Ω, gives us:

$$
\int_\Omega e\frac{\partial e}{\partial t}\mathrm{d}\Omega + \int_\Omega eL_1 e\mathrm{d}\Omega = \int_\Omega e(L_2 - L_1)\widetilde{C}\mathrm{d}\Omega.
\tag{5.7}
$$

Integration by parts, as follows,

$$
\int_\Omega e\frac{\partial e}{\partial t}\mathrm{d}\Omega = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega e^2\mathrm{d}\Omega = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|e\|_{L^2(\Omega)}^2,
\tag{5.8}
$$

inserted in Equation (5.7), results in:

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|e\|^2_{L^2(\Omega)} = -\int_\Omega eL_1 e\mathrm{d}\Omega + \int_\Omega e(L_2 - L_1)\widetilde{C}\mathrm{d}\Omega. \tag{5.9}$$

Applying classical PDE theory, in particular the Lax-Friedrich inequality and Grönwall's lemma (for more details, see [12, 17]), gives:

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

where α is some positive constant, related to the $V-ellipticity$ of the form $\int_\Omega eL_1 e\mathrm{d}\Omega$ [17].

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

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

Assessing the appropriate norm yields:

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

which can be bounded by:

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

Then, we have:

$$\begin{aligned}
\|e\|_{L^2(\Omega)} &\leq \frac{1}{\alpha}\sup_{s\in(0,t)}\left(\frac{1}{2}s_{\max}^2 V|\sigma_{\mathrm{LV}}^2(t,s)|\left|\frac{\widehat{g}(s) - g(s)}{\widehat{g}(s)g(s)}\right|\|\frac{\partial^2 \widetilde{C}}{\partial s^2}\|_{L^2(\Omega)}\right. \\
&\quad \left. + |\rho_{x,v}|\gamma s_{\max} V|\sigma_{\mathrm{LV}}(t,s)|\left|\frac{\sqrt{\widehat{g}(s)} - \sqrt{g(s)}}{\sqrt{\widehat{g}(s)g(s)}}\right|\|\frac{\partial^2 \widetilde{C}}{\partial s\partial V}\|_{L^2(\Omega)}\right). \tag{5.10}
\end{aligned}$$

This latter inequality bound gives a representation of the parameters and functions that have an impact on the error made when solving for $\widetilde{C}$, as an approximation for C. As both $\frac{\partial^2 \widetilde{C}}{\partial s\partial V}$ and $\frac{\partial^2 \widetilde{C}}{\partial s^2}$ are small for large V-values, the error is governed by the difference between the conditional expectation g from the HSLV model and its approximation $\widehat{g}$, which is obtained from the non-parametric method, as discussed in Section 3.1.

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

$$\lim_{N\to\infty,\,\ell\to\infty}\|g - \widehat{g}\| = 0,$$

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

5.2 Performance

To assess the performance of the non-parametric method, we need to determine a *highly accurate* reference value. As pointed out, it is difficult to find this conditional expectation explicitly. One could of course assume $\rho_{x,v} = 0$, which would immediately imply $\mathbb{E}[V(t)|S(t) = s] = \mathbb{E}[V(t)]$ leading to an analytic solution. On the other hand, with nonzero correlation between the volatility and the underlying asset, we can analyze the performance of approximating the conditional expectation for the case of the *pure* Heston model, i.e. the case where $\sigma_{\mathrm{LV}}(t, s) = 1$.

We make use of the COS method [10, 23]. This introduces a well-understood error between the recovered and the theoretical conditional expectation for the Heston SV model. For a more detailed discussion on this, see [10]. Let $g_{\mathrm{H}}(s)$ be the conditional expectation in the Heston SV model, obtained by the 2D-COS method. We determine an approximation $\widehat{g}_{\mathrm{H}}$ by means of the non-parametric method. To measure the performance of the non-parametric method, we are interested in the mismatch in conditional expectations $||g_{\mathrm{H}} - \widehat{g}_{\mathrm{H}}||$. In the L^2-norm, the mismatch from the Monte Carlo simulation can be written as:

$$\sum_{i=1}^{M} ||g_{\mathrm{H}} - \widehat{g}_{\mathrm{H}}||_{L^2(\Omega)}^2 = \sum_{i=1}^{M} \sum_{k=1}^{\ell} \int_{\mathcal{B}_{i,k}} (g_{\mathrm{H}}(s) - \widehat{g}_{\mathrm{H}}(s))^2 \mathrm{d}s, \qquad (5.11)$$

where Ω is the s domain and $\mathcal{B}_{i,k}$ denotes the kth bin at the ith time-step, $k = 1 \dots \ell$, $i = 1 \dots M$. Note that g_{H} is a smooth function, whereas $\widehat{g}_{\mathrm{H}}$ is piecewise linear.

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

Lemma 5.1. *For an arbitrary bin $\mathcal{B}$ with boundaries $[b_l, b_r]$, the error between g_H and $\widehat{g}_H$ has size*

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

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

Proof. For a proof, see Appendix B. $\qquad\qquad\square$

5.3 Numerical experiment: choice of bins

We now discuss the performance of the non-parametric method with respect to the choice of bins. In particular, we consider the error $|g_{\mathrm{H}}(K) - \widehat{g}_{\mathrm{H}}(K)|$ for $K = 40\%$, $K = 100\%$ and $K = 160\%$. The bins are either chosen with respect to an equidistant grid – see (3.7) – or are equally weighted as in (3.8). Parameter values are $\gamma = 0.2$, $\kappa = 0.2$, $r = 0$, $\rho_{x,v} = -0.6$, $S_0 = 1$, $v_0 = 0.04$, $\bar{v} = 0.04$ and we consider the error at particular time $t = 2$. We choose the number of bins between 1 and 20. Our Monte Carlo simulation is performed with $1e6$ paths. Results are displayed in Figure 5.1.

For deep in-the-money and out-the-money strikes the choice of bins does not affect the performance of the non-parametric method. However, at $K = 100\%$ choosing equally weighted bins yields a faster convergence to the reference $g_{\mathrm{H}}(100)$ than when we choose bins equidistantly. This is due to the natural weighting of the bins defined by (3.8), which provides highest accuracy in the ATM region. Further, for all strikes we note that in order to have a high-quality estimate of the conditional expectation, it is not required to use a large number of bins.
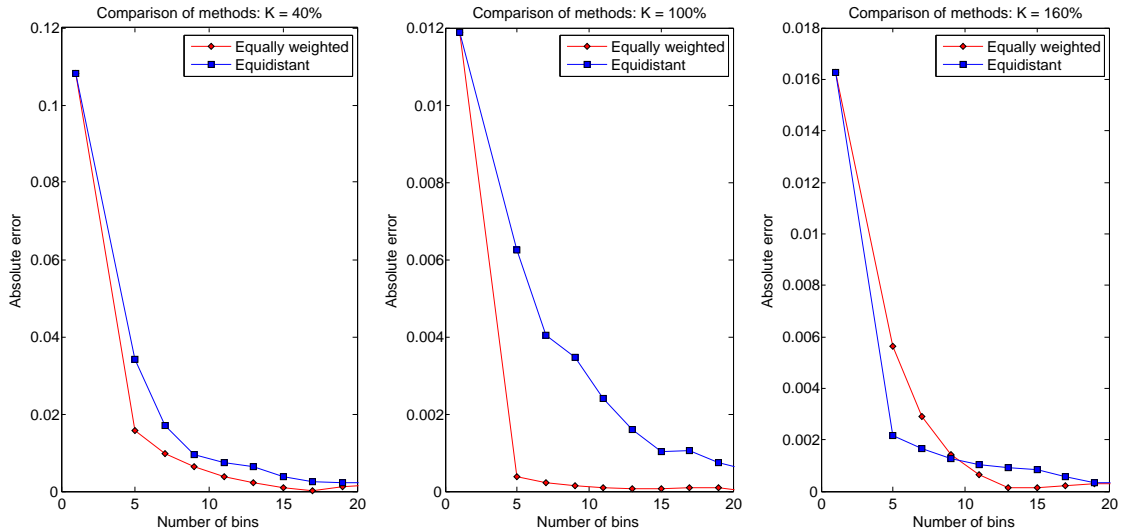
Figure 5.1: Convergence of the non-parametric method for two choices of bins.

In Figure 5.1 we observe highly satisfactory convergence up to 15 bins. With a further increase of the number of bins convergence may stagnate since there is an insufficient number of paths in each bin. In such a case the number of Monte Carlo paths may need to be increased and we should improve the interpolation (continuous approximation, see Section 3.2).

6 Conclusion

In this article we have presented a new Monte Carlo scheme for efficient evaluation of a general Stochastic-Local Volatility model. We have considered the Heston Stochastic-Local Volatility model in numerical experiments. For evaluating this model we have approximated a problematic conditional expectation in a non-parametric way, which is intuitive and easy to implement. This approximation is embedded in a simulation scheme that is strongly based on the QE scheme of Andersen [1] and introduces less bias than more common Euler schemes. By means of numerical experiments and an error analysis we have shown that European-style options can accurately be priced by our method. Furthermore, it enables a consistent and fast pricing of products that are sensitive to the forward volatility smile.

References

[1] L. Andersen. Simple and Efficient Simulation of the Heston Stochastic Volatility Model. *Journal of Computational Finance*, 11(3):1–42, 2008.

[2] J. Andreasen and B. Huge. Volatility Interpolation. *Risk Magazine*, 3:86–89, 2011.

[3] T. Coleman, Y. Li, and A. Verma. Reconstructing the Unknown Local Volatility Function. *Journal of Computational Finance*, 2:77–102, 1998.

[4] J. Cox, J. Ingersoll, and S. Ross. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53:385–407, 1985.

[5] S. De Marco, P. Friz, and S. Gerhold. Rational Shapes of Local Volatility. *Risk Magazine*, 2:82–87, 2013.

[6] G. Deelstra and G. Rayée. Local Volatility Pricing Models for Long-dated FX Derivatives. *Applied Mathematical Finance*, pages 1–23, 2012.

[7] E. Derman and I. Kani. Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility. *International Journal of Theoretical and Applied Finance*, 1(1):61–110, 1998.

[8] B. Dupire. Pricing With a Smile. *Risk Magazine*, 7(1):18–20, 1994.

[9] B. Engelmann, F. Koster, and D. Oeltz. Calibration of the Heston Stochastic Local Volatility Model: A Finite Volume Scheme. *Available at SSRN 1823769*, 2011.

[10] F. Fang and C. W. Oosterlee. A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions. *SIAM Journal on Scientific Computing*, 31(2):826–848, 2008.

[11] J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. John Wiley & Sons, Ltd, 2006.

[12] S. Guo, L. A. Grzelak, and C. W. Oosterlee. Analysis of an Affine Version of the Heston-Hull-White Option Pricing Partial Differential Equation. *Working paper*, February 2013.

[13] P. Henry-Labordère. Calibration of Local Stochastic Volatility Models to Market Smiles: a Monte-Carlo Approach. *Risk Magazine*, pages 112–117, 2009.

[14] S. L. Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6:327–343, 1993.

[15] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1991.

[16] R. Lord, R. Koekkoek, and D. Van Dijk. A Comparison of Biased Simulation Schemes for Stochastic Volatility Models. *Quantitative Finance*, 10:177–194, 2010.

[17] J. Necas. *Direct Methods in the Theory of Elliptic Equations*. Springer, 2012.

[18] V. Piterbarg. Markovian Projection Method for Volatility Calibration. *Available at SSRN 906473*, 2006.

[19] V. Piterbarg and B. Capital. Modern Approaches to Stochastic Volatility Calibration. In *Proceedings of the WBS 3rd Fixed Income Conference, Amsterdam*, volume 41, 2006.

[20] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer, 2005.

[21] R. Rebonato. *Volatility and Correlation in the Pricing of Equity, FX, and Interest-rate Options*. John Wiley & Sons, Ltd, 1999.

[22] Y. Ren, D. Madan, and M. Q. Qian. Calibrating and Pricing with Embedded Local Volatility Models. *Risk Magazine*, 20(9):138–143, 2007.

[23] M. J. Ruijter and C. W. Oosterlee. Two-dimensional Fourier Cosine Series Expansion Method for Pricing Financial Options. *SIAM Journal on Scientific Computing*, 34(5):B642–B671, 2012.

[24] R. Schöbel and J. Zhu. Stochastic Volatility with an Ornstein-Uhlenbeck Process: An Extension. *European Finance Review*, 3:23–46, 1999.

[25] H. Tanaka. Note on Continuous Additive Functionals of the 1-Dimensional Brownian Path. *Z. Wahrscheinlichkeitstheorie*, 1:251–257, 1963.

[26] G. Tataru and T. Fisher. Stochastic Local Volatility. *Bloomberg*, Version 1 (February 5), 2010.

[27] Y. Tian, Z. Zhu, F. Klebaner, and K. Hamza. A Hybrid Stochastic Volatility Model Incorporating Local Volatility. In *Fourth International Conference on Computational and Information Sciences*, pages 333 – 336, 2012.

A   Proof of Lemma 2.1

For notational purposes, we leave out the filtration $\mathcal{F}(t_0)$. We first calculate the derivative of $C(t, K) := C(t_0, t, S(t_0), K)$ with respect to K:

$$\frac{\partial C(t, K)}{\partial K} \quad = \quad \frac{1}{M(t)} \frac{\partial}{\partial K} \mathbb{E}\left[(S(t) - K)^+\right] = \frac{1}{M(t)} \frac{\partial}{\partial K} \int_K^\infty (s - K) f_S(s) \mathrm{d}s, \quad \text{(A.1)}$$

where f_S is the marginal probability density function of the process $S(t)$. Applying Leibniz' integration rule on $(s - K) f_S(s)$ yields:

$$
\begin{aligned}
\frac{\partial}{\partial K} \int_K^\infty (s - K) f_S(s) \mathrm{d}s \quad &= \quad -(K - K) f_S(K) \frac{\partial}{\partial K} K + \int_K^\infty \frac{\partial}{\partial K}(s - K) f_S(s) \mathrm{d}s \\
&= \quad \int_K^\infty \frac{\partial}{\partial K}(s - K) f_S(s) \mathrm{d}s \\
&= \quad \int_\infty^K f_S(s) \mathrm{d}s.
\end{aligned}
$$

Continuing with (A.1) we obtain

$$\frac{\partial C(t, K)}{\partial K} \quad = \quad \frac{1}{M(t)} \frac{\partial}{\partial K} \int_K^\infty (s - K) f_S(s) \mathrm{d}s = \frac{1}{M(t)} \int_\infty^K f_S(s) \mathrm{d}s.$$

Now, for a digital option,

$$\int_\infty^K f_S(s) \mathrm{d}s \quad = \quad -\mathbb{Q}\left[S(t) > K\right] = -\mathbb{E}\left[\mathbb{1}_{S(t) > K}\right]$$

and thus

$$-\frac{\partial C(t, K)}{\partial K} \quad = \quad \frac{1}{M(t)} \mathbb{E}\left[\mathbb{1}_{S(t) > K}\right].$$

So, we have

$$\frac{\partial C(t, K)}{\partial K} \quad = \quad \frac{1}{M(t)} \int_\infty^K f_S(s) \mathrm{d}s.$$

Then, by Leibniz' integration rule, we obtain

$$\frac{\partial^2 C(t, K)}{\partial K^2} = \frac{f_S(K)}{M(t)}.$$

B   Proof of Lemma 5.1

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

$$\widehat{g}_{\mathrm{H}}(s) = \frac{\Delta \widehat{g}_{\mathrm{H}}}{\Delta s} s + \frac{\widehat{g}_{\mathrm{H}}(b_l) b_r - \widehat{g}_{\mathrm{H}}(b_r) b_l}{\Delta s}, \tag{B.1}$$

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

$$g_{\mathrm{H}}(s) = g_{\mathrm{H}}(s_m) + \sum_{n=1}^{\infty} \frac{g_{\mathrm{H}}^{(n)}(s_m)}{n!}(s - s_m)^n. \tag{B.2}$$

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

$$||g_{\mathrm{H}} - \widehat{g}_{\mathrm{H}}||_{L^2(\mathcal{B})}^2 = \int_B (g_{\mathrm{H}}(s) - \widehat{g}_{\mathrm{H}}(s))^2 \mathrm{d}s = \int_{b_l}^{b_r} (g_{\mathrm{H}}(s) - \widehat{g}_{\mathrm{H}}(s))^2 \mathrm{d}s, \tag{B.3}$$

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

$$||g_{\mathrm{H}} - \widehat{g}_{\mathrm{H}}||_{L^2(\mathcal{B})}^2 = \int_{b_l}^{b_r} \left(g_{\mathrm{H}}(s_m) - \widehat{g}_{\mathrm{H}}(s) + \sum_{n=1}^{\infty} \frac{g_{\mathrm{H}}^{(n)}(s_m)}{n!}(s - s_m)^n \right)^2 \mathrm{d}s. \tag{B.4}$$

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

$$g_{\mathrm{H}}(s_m) - \widehat{g}_{\mathrm{H}}(s) = g_{\mathrm{H}}(s_m) - \frac{\Delta\widehat{g}_{\mathrm{H}}}{\Delta s}s - \frac{\widehat{g}_{\mathrm{H}}(b_l)b_r - \widehat{g}_{\mathrm{H}}(b_r)b_l}{\Delta s}. \tag{B.5}$$

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

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

Expanding $g_{\mathrm{H}}(s_m)$ at the boundary points and by plugging this result into (B.5) we find

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

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

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

Now, after some algebraic manipulations we end up with

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

where

$$c_1 := \frac{1}{2}\big((g_{\mathrm{H}}(b_l) - \widehat{g}_{\mathrm{H}}(b_l)) + (g_{\mathrm{H}}(b_r) - \widehat{g}_{\mathrm{H}}(b_r))\big).$$

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

$$||g_{\mathrm{H}} - \widehat{g}_{\mathrm{H}}||_{L^2(\mathcal{B})}^2 = \tag{B.6}$$

$$\int_{b_l}^{b_r} \left(\sum_{n=1}^{6} \frac{g_{\mathrm{H}}^{(n)}(s_m)}{n!}(s - s_m)^n - \sum_{n=2,4,6} \frac{g_{\mathrm{H}}^{(n)}(s_m)}{2^n \cdot n!}(\Delta s)^n - \frac{(s - s_m)\Delta\widehat{g}_{\mathrm{H}}}{\Delta s} + c_1 + \mathcal{O}(\Delta s^8) \right)^2 \mathrm{d}s$$

Evaluating (B.6) yields the result in Lemma 5.1.