Imperial College London

Imperial College London

Department of Mathematics

Simulations of calibrated Local Stochastic Volatility models

Author: Minyuan LI (CID: 02278349)

A thesis submitted for the degree of

MSc in Mathematics and Finance, 2022-2023

Declaration

The work contained in this thesis is my own work unless otherwise stated.

Acknowledgements

First of all, I would like to to express my gratitude to my supervisor Dr. Wolfgang Stockinger, for all the support and guidance throughout the project.

I would also like to thank the QRFX team at JP Morgan & Chase, for helping me build a better understanding of the topic I work in.

Last but not the least, I would like to extend my gratitude towards my family, friends and the Mathematical Finance department at Imperial College of London, for all the encouragement and the support throughout the year.

Abstract

We study local stochastic volatility models (LSVMs), a type of stochastic processes that models market share price dynamics. The volatility term in a LSVM consists of a local volatility component as well as a stochastic volatility process. This makes LSVMs state-of-the-art models in option pricing, since they have a better ability to reproduce the market implied volatility surface than pure local volatility models or stochastic volatility models. In this report, we discuss some theoretical and numerical challenges in the calibrated LSVMs, and focus on the simulation scheme of the calibrated dynamic. Since the calibrated LSVMs can be viewed as a special case of the McKean-Vlasov stochastic differential equations, we introduce the particle method discussed in [1, Chpater 11.6], a numerical scheme that simulates dependent trajectories at each time step. Furthermore, we discuss and compare several existing methods to estimate the volatility term in the calibrated LSVMs by implementing the particle simulations based on these methods. Finally, we include stochastic interest rates in the LSVMs and study a numerical example in the Foreign Exchange market.

Contents

List of Figures

List of Tables

Introduction

In the ground-breaking paper [2], Black and Scholes derive the vanilla option pricing formula assuming that the share price follows the simplified dynamic on a fixed time interval $[0, T]$:

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

with an initial value $S_0 \in \mathbb{R}^+$, a constant risk-free rate $r \in \mathbb{R}$, a volatility component $\sigma \in \mathbb{R}^+$, and $W = (W_t)_{t\in[0,T]}$ is a Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q})$, where $\mathbb{Q}$ is a risk neutral measure.

The simplified assumption on constant volatility, however, fails to capture behaviours observed in the market. When we calibrate the Black-Scholes option pricing formula to the market prices of vanilla options, we obtain a non-constant implied volatility surface with respect to the expiration time and the strike. Furthermore, with a fixed expiration time, the empirical relationship between the implied volatility and the strike resembles an upwards skewed smile, known as the volatility smile. This indicates that deep out-of-the-money options have higher implied volatility values than the at-the-money options with the same option terms.

Extensive research has been done to capture and model the implied volatility smile. A Local volatility model (LVM) assumes that the volatility component can be modelled by a deterministic function of time and share price (assuming zero risk-free rate):

$$dS_t = \sigma(t, S_t)S_t dW_t,$$

where $\sigma : [0, T] \times \mathbb{R}^+ \to: \mathbb{R}^+ \cup \{0\}$ is known as the local volatility function. Several LVMs with parametric forms of σ are considered in practice, including the displaced diffusion model in [3], which is able to include partial volatility smile in the model. A non-parametric form of the local volatility function that allows an exact replication of the market volatility smile was introduced by Dupire in [4]. Such local volatility function can be computed via the Dupire's formula, and the model is able to reproduce market prices of any set of arbitrage-free European vanilla options. However, the downside of such model is that it may introduce some unrealistic behaviours in the share price dynamic, and the smile tends to flatten out for options with longer maturities.

On the other hand, a stochastic volatility model (SVM) introduces volatility smiles by assuming that the volatility component follows some non-negative stochastic process:

$$\begin{aligned} dS_t &= \sqrt{v_t} S_t dW_t, \\ dv_t &= b(t, v_t) + \tilde{\sigma}(t, v_t) dB_t, \end{aligned}$$

for some measurable functions $b : [0, T] \times \mathbb{R}^+ \to \mathbb{R}$ and $\tilde{\sigma} : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+ \cup \{0\}$ which are regular enough so that the volatility process has a unique solution. Some common stochastic volatility models include the Heston model in [5], where the volatility follows a Cox–Ingersoll–Ross (CIR) process

$$\begin{aligned} dS_t &= \sqrt{v_t} S_t dW_t, \\ dv_t &= \mu(\bar{v} - v_t)dt + \xi\sqrt{v_t} dB_t, \end{aligned}$$

with some positive real constant μ, $\bar{v}$, ξ, and $(B_t)_{t\in[0,T]}$ is a Brownian motion (correlated to $(W_t)_{t\in[0,T]}$). Other well-known examples include the Stein-Stein model [6] and the SABR model [7]. Although SVMs are better at capturing volatility smiles with the presence of the stochastic volatility processes [8], it is difficult to achieve a perfect fit to the market given a small finite set of parameters.

A Local Stochastic Volatility Model (LSVM) includes both local volatility and stochastic volatility components when modelling the dynamic of the share price:

$$dS_t = \sigma(t, S_t)\sqrt{v_t}S_t dW_t, \tag{0.0.3a}$$
$$dv_t = b(t, v_t) + \tilde{\sigma}(t, v_t)dB_t, \tag{0.0.3b}$$

where the leverage function $\sigma : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+ \cup \{0\}$ is a non-parametric deterministic function and $v = (v_t)_{t\in[0,T]}$ is a non-negative stochastic process as in SVMs. LSVMs are considered as the state-of-the-art models in option pricing [9]; The stochastic volatility component incorporates a rough volatility smile, whilst the local volatility component allows an exact calibration to any set of arbitrage-free European vanilla options, given that the calibration condition proposed in [10] is satisfied:

$$\sigma^2_{Dup}(t, x) = \sigma^2(t, x)\mathbb{E}^{\mathbb{Q}}[v_t|S_t = x],$$

for $t \in [0, T]$ and $x \in \mathbb{R}^+$, where $\sigma_{Dup} : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+ \cup \{0\}$ is the Dupire local volatility function. LSVMs have become popular modelling tools in financial institutions due to its desirable features in reproducing the market volatility smile, which is crucial for activities such as hedging and exotic option pricing.

We add the condition to (0.0.3) and obtain the calibrated LSVM:

$$dS_t = \sqrt{v_t}\frac{\sigma_{Dup}(t, S_t)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[v_t|S_t]}}S_t dW_t,$$
$$dv_t = b(t, v_t) + \tilde{\sigma}(t, v_t)dB_t.$$

One of the challenges in the calibrated LSVMs is its well-posedness. As pointed out in [11], it is difficult to prove whether there exists a unique solution of the stochastic differential equations in general, and it still remains an open question. Furthermore, it is numerically demanding to simulate the calibrated LSVMs due to the presence of the conditional expectation term $\mathbb{E}^{\mathbb{Q}}[v_t|S_t]$, which cannot be obtained explicitly.

The main focus of this project is to introduce the particle method, a numerical method that can be applied in simulating the calibrated LSVMs, and discuss several methods to approximate the conditional expectation $\mathbb{E}^{\mathbb{Q}}[v_t|S_t]$. The particle method was first introduced in [1, Chpater 11.6, page 280] for the calibrated LSVMs, and it involves simulating N interacting particles and using their empirical distribution to approximate the conditional expectation. However, since at each time step we only observe N particle realisations, the way to approximate the dependency between the share price and the volatility is less straightforward.

Several different methods have been proposed to estimate the conditional expectation. [1] discusses the approximation by the Nadaraya-Watson estimator, where it estimates the joint density of S_t and v_t by some continuous regularising kernel functions. On the other hand, [11] views the conditional expectation as a function $\mathbb{E}^{\mathbb{Q}}[v_t|S_t = \cdot] : \mathbb{R}^+ \cup \{0\} \to \mathbb{R}^+ \cup \{0\}$, and directly approximates it by a weighted sum of exponential kernel functions. Moreover, [12] partitions N realised particles into l bins $B_1, \ldots, B_l$, $l \ll N$, and estimate $\mathbb{E}^{\mathbb{Q}}[v_t|S_t = x]$ by $\mathbb{E}^{\mathbb{Q}}[v_t|S_t \in B_i]$ for $x \in B_i$, $i = 1, \ldots, l$. We will discuss and implement each method in detail, and provide a short comparison among the three methods.

The rest of the report is structured as follows. In Chapter 1, we introduce the LSVM calibration condition in more detail as well as the McKean-Vlasov equation, a tool in the particle simulation. From Chapter 2 to 4, we introduce and discuss the three different simulation methods for the calibrated LSVMs. In Chapter 5, we provide a comparison among the three simulation methods. We conclude the report with Chapter 6, where we discuss the simulation of the calibrated LSVMs with stochastic interest rates in the Foreign Exchange market.

Chapter 1

The LSVM calibration condition and McKean-Vlasov SDEs

Let the stochastic process $S = (S_t)_{t\in[0,T]}$ be the dynamic of a share price defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q})$ for a fixed time $T > 0$. For simplicity, we assume that the market has zero interest rate for the following discussions with the exception in Chapter 6. Recall that in the local stochastic volatility model (LSVM), S evolves with the following dynamic under the risk neutral measure $\mathbb{Q}$:

$$dS_t = \sqrt{v_t}\sigma(t, S_t)S_t dW_t, \tag{1.0.1a}$$

$$dv_t = b(t, v_t)dt + \tilde{\sigma}(t, v_t)dB_t, \tag{1.0.1b}$$

with non-random initial values S_0 and v_0 in $\mathbb{R}^+$, two (correlated) Brownian motions $W = (W_t)_{t\in[0,T]}$ and $B = (B_t)_{t\in[0,T]}$, and some real-valued functions $b : [0, T]\times\mathbb{R}^+ \to \mathbb{R}$, $\tilde{\sigma} : [0, T]\times\mathbb{R}^+ \to \mathbb{R}^+\cup\{0\}$, and $\sigma : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+ \cup \{0\}$. The function σ is known as the leverage function, and it is established in [10] that for an exact calibration to the market, the leverage function needs to satisfy the condition

$$\sigma^2_{Dup}(t, x) = \sigma^2(t, x)\mathbb{E}^{\mathbb{Q}}[v_t|S_t = x],$$

where

$$\sigma^2_{Dup}(t, x) = \frac{\partial_t C(t, x)}{\frac{1}{2}x^2\partial^2_x C(t, x)}, \tag{1.0.2}$$

for all $t \in [0, T]$ and $x \in \mathbb{R}^+$. $C(t, x)$ denotes the market price of a vanilla Call option, with the expiration time t and the strike x. The function $\sigma_{Dup} : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+ \cup \{0\}$ is known as the Dupire local volatility, which can computed using the vanilla option prices in the market with finite difference method.

We write the calibrated LSVM as follows:

$$dS_t = \sqrt{v_t}S_t\frac{\sigma_{Dup}(t, S_t)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[v_t|S_t]}}dW_t, \tag{1.0.3a}$$

$$dv_t = b(t, v_t)dt + \tilde{\sigma}(t, v_t)dB_t. \tag{1.0.3b}$$

We are interested in simulating the above dynamic based on an approach called the particle method, a numerical scheme that involves simulating dependent paths at each time step. The particle method is often used as a numerical approach to approximate the solution to the McKean-Vlasov SDEs governed by for some d-dimensional stochastic process $(X_t)_{t\in[0,T]}$:

$$dX_t = \beta(t, X_t, \mathbb{P}_t)dt + \sigma(t, X_t, \mathbb{P}_t)dW_t, \tag{1.0.4}$$

where $(W_t)_{t\in[0,T]}$ is a n-dimensional Brownian motion, given some initial value $x_0 \in \mathbb{R}^d$, functions $\beta : [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^{d\times n}$, where $\mathcal{P}(\mathbb{R}^d)$ is a set of Borel probability measures in $\mathbb{R}^d$. Note that the drift and the volatility terms are dependent on $(\mathbb{P}_t)_{t\in[0,T]}$, the probability distributions of X_t at some time t. For simplicity, if the drift and the diffusion terms do not explicitly depend on time t, we write $\beta : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^{d\times n}$ instead.

Notice that the calibrated LSVM in (1.0.3) can be seen as a special case of McKean-Vlasov SDEs, with $X_t = (S_t, v_t)$ for $t \in [0, T]$, a 2-dimensional Brownian motion $(W_t)_{t\in[0,T]}$, and

$$\beta(t, (S_t, v_t), \mathbb{P}_t) = [0, b(t, v_t)]^T, \tag{1.0.5a}$$

$$\sigma(t, (S_t, v_t), \mathbb{P}_t) = \begin{bmatrix} \sqrt{v_t} S_t \frac{\sigma_{Dup}(t,S_t)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[v_t|S_t]}} & 0 \\ 0 & \tilde{\sigma}(t, v_t) \end{bmatrix}. \tag{1.0.5b}$$

This observation suggests that the particle method may be considered as a feasible numerical scheme to simulate the calibrated LSVMs. In the following subsection, we first introduce McKean-Vlasov SDEs and the particle method in more detail, and give a numerical example of McKean-Vlasov SDE simulations with the particle method. Finally, we discuss some theoretical properties of the McKean-Vlasov SDEs and the challenges presented in the context of the calibrated LSVMs.

1.1 The particle method and the McKean-Vlasov SDE simulations

McKean-Vlasov SDEs were first introduced in [13]. As shown in (1.0.4), both the drift and the volatility terms have a complex dependency not only on the value of X_t at time $t \in [0, T]$, but also on its distribution. This complex structure creates difficulty in applying classical numerical schemes that simulate each path independently, since the prior knowledge on the distribution of X_t, $t \in (0, T]$ is unknown.

The particle method, on the other hand, offers a feasible way to simulate processes with such dynamics. As opposed to the standard numerical scheme, the main step in the particle method simulates N dependent particles (trajectories) simultaneously, and uses the empirical distribution of the particle realisations at time t as an approximation of the distribution of X_t. More specifically, we approximate $\mathbb{P}_t$ by $\mathbb{P}_t^N$:

$$\mathbb{P}_t^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^{i,N}},$$

where $\delta_x(\cdot)$ is a Dirac function centred at x, and $X_t^{i,N}$ is the value of i^{th} simulated particle at time t. To illustrate the simulation procedure, we present the particle method on McKean-Vlasov SDEs introduced in Chapter 10 of [1] with a numerical example.

1.1.1 Numerical example

Consider the following SDE:

$$dX_t = (X_t + \mathbb{E}X_t)dt + X_t dW_t, X_0 = 1,$$

with a Brownian motion $(W_t)_{t\in[0,T]}$. Consider N particles $\{(X_t^{i,N})_{t\in[0,T]}\}_{i=1}^N$, where each particle follows the dynamic:

$$\begin{aligned} dX_t^{i,N} &= X_t^{i,N} dt + \left(\int x d\mathbb{P}_t^N(x) \right) dt + X_t^{i,N} dW_t^{i,N}, \\ &= \left(X_t^{i,N} + \frac{1}{N} \sum_{j=1}^{N} X_t^{j,N} \right) dt + X_t^{i,N} dW_t^{i,N}, \end{aligned}$$

where $\{(W_t^{i,N})_{t\in[0,T]}\}_{i=1}^N$ are N independent copies of $(W_t)_{t\in[0,T]}$ and $X_0^{i,N} = 1$ for $i = 1, \ldots, N$. Denote $b_N^i(t, \mathbf{X}_t) = X_t^{i,N} + \frac{1}{N} \sum_{j=1}^{N} X_t^{j,N}$, where $\mathbf{X}_t = (X_t^{1,N}, \ldots, X_t^{N,N})$. We simulate the SDE over the time interval $[0, T]$ as follows:

1. Discretise the interval $[0, T]$ into M time steps $\{t_k\}_{k=0}^M$, and set the number of particles N.
2. Initialise $X_0^{i,N} = 1$ for all $i = 1, \ldots, N$.
3. Set $k = 1$ and $b_N^i(t, \mathbf{X}_t) = X_0^{i,N} + \frac{1}{N} \sum_{j=1}^{N} X_0^{j,N}$ for all $t \in [0, t_1]$, $i = 1, \ldots, N$.

4. Update N particles $\{X_t^{i,N}\}_{1=1}^N$ from t_{k-1} to t_k using the Euler-Maruyama scheme:

$$X_{t_k}^{i,N} = X_{t_{k-1}}^{i,N} + b_N^i(t_{k-1}, \mathbf{X}_{t_{k-1}})(t_k - t_{k-1}) + X_{t_{k-1}}^{i,N}\sqrt{t_k - t_{k-1}}Z_k^{i,N},$$

where $Z_k^{i,N}$, $i = 1, \ldots, N$ are i.i.d. standard normal random variables.

5. Compute $b_N^i(t, \mathbf{X}_{t_k}) = X_{t_k}^{i,N} + \frac{1}{N}\sum_{j=1}^N X_{t_k}^{j,N}$ for $i = 1, \ldots, N$. Set $b_N^i(t, \mathbf{X}_t) = b_N^i(t, \mathbf{X}_{t_k})$ for all $t \in [t_k, t_{k+1}]$. Interpolate and extrapolate $b_N^i(t, \mathbf{X}_t)$ for $t \in [t_{k-1}, t_k]$.

6. Increase k by 1. Repeat steps 4, 5 until $k = M$.

In this simplified example, we are able to compute the analytical mean $\mathbb{E}X_t$ by solving an linear ordinary differential equation $d\mathbb{E}X_t = 2\mathbb{E}X_t dt$ with $X_0 = 1$, and obtain $\mathbb{E}X_t = e^{2t}$ for $t \in [0, T]$. Figure 1.1 shows 10 particle trajectories that are randomly selected from $N = 1000$ particles over time interval $[0, 1]$, simulated with $M = 100$ equal-sized time steps. It can be seen that the empirical average trajectory follows the true mean closely.

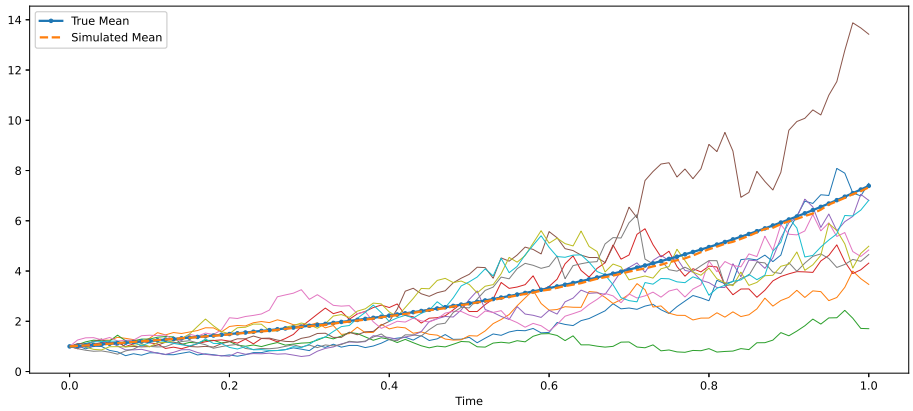

Figure 1.1: 10 randomly selected particle trajectories, the average of simulated trajectories and the trajectory of the true mean.

We see that the particles interact with each other in the simulation steps. Moreover, it is established in [14] that under the conditions that guarantee the well-posedness of the SDEs, as the number of particles N approaches infinity, the particles possess the *chaos propagation property*; that is, at any positive time $t \in (0, T]$, $\{X_t^{i,N}\}_{i=1}^N$ become independent and $\mathbb{P}_t^N$ converges $\mathbb{P}_t$ in distribution, given that each particle is independent at time zero. In the context of the option pricing, the chaos propagation property and together with the Law of Large Number allow us to estimate option prices by taking the average of the (discounted) option payoffs at expiration T:

$$\hat{C}(T, K) = \frac{1}{N}\sum_{i=1}^N (S_T^{i,N} - K)^+ \tag{1.1.2a}$$

$$\hat{P}(T, K) = \frac{1}{N}\sum_{i=1}^N (K - S_T^{i,N})^+. \tag{1.1.2b}$$

In the next section, we discuss some theoretical properties of the McKean-Vlasov SDEs, including the conditions that guarantee the SDEs' well-posedness, and further point out some theoretical challenges in the calibrated LSVMs.

1.2 Challenges in the calibrated LSVMs

Recall the McKean-Vlasov SDE in (1.0.4):

$$dX_t = \beta(t, X_t, \mathbb{P}_t)dt + \sigma(t, X_t, \mathbb{P}_t)dW_t, X_0 = x_0. \tag{1.2.1}$$

Assuming that the drift and the diffusion terms have no explicit dependence on time t, the associated particle system can be written as:

$$dX_t^{i,N} = \beta(X_t^{i,N}, \mathbb{P}_t^N)dt + \sigma(X_t^{i,N}, \mathbb{P}_t^N)dW_t^{i,N}, \tag{1.2.2a}$$

$$\mathbb{P}_t^N = \frac{1}{N}\sum_{i=1}^{N} \delta_{X_t^{i,N}}, \tag{1.2.2b}$$

for $i = 1, \dots, N$, where $\{(W_t^{i,N})_{t\in[0,T]}\}_{i=1}^N$ are independent copies of the Brownian motion $(W_t)_{t\in[0,T]}$. We assume that the initial values of the particles $X_0^{i,N} = x_0$, $i = 1, \dots, N$, for some constant $x_0 \in \mathbb{R}^d$. It is important to understand whether such SDE is well-posed; that is, under which conditions the SDE has a unique solution. We quote the Assumption 3.1 in [15] that guarantees the well-posedness of the above McKean-Vlasov SDE:

Assumption 1.2.1. Assume that the Brownian motions $\{(W_t^{i,N})_{t\in[0,T]}\}_{i=1}^N$ are independent and n-dimensional, and functions $\beta : \mathbb{R}^d \times \mathcal{P}^2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{P}^2(\mathbb{R}^d) \to \mathbb{R}^{d\times n}$ in (1.2.1) and (1.2.2) are Lipschitz, that is, there exists a constant $L \geq 0$ such that

$$|\beta(x, m) - \beta(x', m')| + |\sigma(x, m) - \sigma(x', m')| \leq L(|x - x'| + \mathcal{W}_2(m, m')),$$

for any $x, x' \in \mathbb{R}^d$, $m, m' \in \mathcal{P}^2(\mathbb{R}^d)$. $|\cdot|$ denotes both the Euclidean norm of $\mathbb{R}^d$ and Frobenius norm on $\mathbb{R}^{d\times n}$, depending on the context.

In the assumption, $\mathcal{P}^2(\mathbb{R}^d)$ represents a subset of Borel probability measures $\mathcal{P}$ on $\mathbb{R}^d$ that have finite second moments, and $\mathcal{W}_2$ is defined as the 2-Wasserstein metric, a distance measure between two probability measures in $\mathcal{P}^2(\mathbb{R}^d)$. The technical definitions of $\mathcal{P}^2$ and $\mathcal{W}_2$ are provided in Appendix A.1 and A.2. This assumption guarantees the well-posedness of the McKean-Vlasov SDE in (1.2.1) and its associated particle system in (1.2.2), and we refer readers to Chapter 3 in [15] for more details of the proof.

However, in the case of the calibrated LSVMs with the drift and the diffusion terms defined in (1.0.5), the conditional expectation $\mathbb{E}[v_t|S_t]$ in the denominator of the diffusion term fails to satisfy the Lipschitz condition. This creates difficulty in understanding whether such SDEs are well-posed, and it has been an active research topic. [16] investigates the well-posedness of some two-dimensional McKean-Vlasov SDEs that are similar to (1.0.3):

$$\begin{aligned} dS_t &= b_1(S_t)\frac{h(v_t)}{\mathbb{E}[h(v_t)|S_t]}dt + \sigma_1(S_t)\frac{f(v_t)}{\sqrt{\mathbb{E}[f^2(v_t)|S_t]}}dW_t, \\ dv_t &= b_2(v_t)dt + \sigma_2(v_t)dB_t, \end{aligned}$$

with functions b_1, b_2, σ_1, σ_2, h, f and (correlated) Brownian motions $(W_t)_{t\in[0,T]}$ and $(B_t)_{t\in[0,T]}$. However, the authors only manage to establish the existence and uniqueness of the stationary solution with strong assumptions on b_1, b_2, σ_1, σ_2, h, f. [9] imposes some regularisations on the calibrated LSVMs and proves the well-posedness of the regularised system under some further assumptions, and [17] proves the existence of the calibrated LSVMs with assumptions on the parameters in the stochastic volatility process. Furthermore, [11] discusses the well-posedness of the calibrated LSVMs with the regularised conditional expectation using the Reproducing Kernel Hilbert space, and we will discuss this method in more detail in Chapter 3. However, the well-posedness of the calibrated LSVMs in general remains unclear.

On the other hand, there are further challenges presented from the numerical point of view. Unlike the example (1.1.1) which involves estimating expectations in the drift term, approximating conditional expectations $\mathbb{E}[v_t|S_t]$ in the calibrated LSVMs is less straightforward. It requires the simulation scheme to take into account the dependency structure between v_t and S_t, $t \in [0, T]$ at each time step. In the next three sections, we discuss three different methods to approximate the conditional expectation in the numerical scheme. We then present the simulation results by comparing the implied volatility curve of vanilla options from the simulation of the calibrated LSVMs to the one from the market.

Chapter 2

Simulations of the calibrated LSVM with the regularising kernel method

The simulation method of the calibrated LSVMs presented in this chapter is introduced in Chapter 11.6 of [1]. It approximates the density of the particles (S_t, v_t) for $t \in [0, T]$ by continuous regularising kernel functions. We first introduce the method and the full simulation algorithm, and then present and discuss the simulation results in two different markets. Finally, we give a short discussion on the parameter choices in the regularising kernel functions and its impact on the simulation result.

2.1 The regularising kernel method

Recall the calibrated LSVM in (1.0.3) under the risk neutral measure $\mathbb{Q}$:

$$\begin{aligned} dS_t &= \sqrt{v_t} S_t \frac{\sigma_{Dup}(t, S_t)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[v_t | S_t]}} dW_t, \\ dv_t &= b(t, v_t)dt + \tilde{\sigma}(t, v_t)dB_t, \end{aligned}$$

and its associated particle system:

$$\begin{aligned} dS_t^{i,N} &= \sqrt{v_t^{i,N}} S_t^{i,N} \frac{\sigma_{Dup}(t, S_t^{i,N})}{\sqrt{\mathbb{E}^{\mathbb{Q}_t^N}[v_t | S_t = S_t^{i,N}]}} dW_t^{i,N}, \\ dv_t^{i,N} &= b(t, v_t^{i,N})dt + \tilde{\sigma}(t, v_t^{i,N})dB_t^{i,N}, \end{aligned}$$

with $S_0^{i,N} = S_0$, $v_0^{i,N} = v_0$ for $i = 1, \ldots, N$, and $\{((W_t^{i,N})_{t\in[0,T]}, (B_t^{i,N})_{t\in[0,T]})\}_{i=1}^N$ are independent copies of (W, B), where $W = (W_t)_{t\in[0,T]}$ and $B = (B_t)_{t\in[0,T]}$. To estimate the conditional expectation under the empirical measure $\mathbb{Q}_t^N$ for $t \in (0, T]$, an intuitive approximation could be:

$$\mathbb{E}^{\mathbb{Q}_t}[v_t | S_t = x] \approx \mathbb{E}^{\mathbb{Q}_t^N}[v_t | S_t = x] \text{“=”} \frac{\sum_{i=1}^N v_t^{i,N} \delta(S_t^{i,N} - x)}{\sum_{i=1}^N \delta(S_t^{i,N} - x)},$$

where $\delta(\cdot)$ is a Dirac function centred at zero, and $x \in \mathbb{R}^+$. However, note that the probability measure $\mathbb{Q}_t^N$ is discrete, and the denominator in the above expression has value zero for $x \notin \{S_t^{i,N}\}_{i=1}^N$, therefore it is not well defined.

A solution is to estimate the conditional expectation with the Nadaraya-Watson estimator, that is, we replace the Dirac function $\delta(\cdot)$ by a continuous regularising kernel function $\delta_{t,N}(\cdot)$ with some bandwidth h. A common choice is to use the probability density function of standard normal random variable as the regularising kernel:

$$\delta_{t,N}(x) = \frac{1}{h} K\left(\frac{x}{h}\right) = \frac{1}{\sqrt{2\pi}h} \exp\left(-\frac{x^2}{2h^2}\right).$$

The choice of the bandwidth h typically has an impact on the simulation results. As a starting point, we select $h = S_0 N^{-\frac{1}{5}}$ in reference to [1], and in a later section we provide a more detailed discussion on the impact of h on the simulation results.

For completeness, we briefly explain how to approximate the Dupire function σ_{Dup} from the market vanilla option prices. Recall that as in (1.0.2), the Dupire function is defined as

$$\sigma^2_{Dup}(t,x) = \frac{\partial_t C(t,x)}{\frac{1}{2}x^2\partial_x^2 C(t,x)}.$$

We use a matrix of market vanilla Call option prices with various expiration times and strikes to compute the numerator and denominator through finite difference:

$$\partial_t C(t,x) \approx \frac{C(t+\Delta t,x) - C(t,x)}{\Delta t}, \tag{2.1.3a}$$

$$\partial_x^2 C(t,x) \approx \frac{C(t,x+\Delta x) - 2C(t,x) + C(t,x-\Delta x)}{(\Delta x)^2}. \tag{2.1.3b}$$

For any other strike values and expiration times, we apply cubic interpolation in strike and linear interpolation in expiration time if the strike value and the expiration time falls inside the ranges that are available in the market; otherwise, we apply flat extrapolation.

Based on the above approximation, we derive the following particle system:

$$dS_t^{i,N} = S_t^{i,N}\sigma_{Dup}(t, S_t^{i,N})\sqrt{\frac{\sum_{j=1}^N \delta_{t,N}(S_t^{j,N} - x)}{\sum_{j=1}^N v_t^{j,N}\delta_{t,N}(S_t^{j,N} - x)}}\sqrt{v_t^{i,N}}dW_t^{i,N}, \tag{2.1.4a}$$

$$dv_t^{i,N} = b(t, v_t^{i,N})dt + \tilde{\sigma}(t, v_t^{i,N})dB_t^{i,N}, \tag{2.1.4b}$$

with initial values $S_0^{i,N} = S_0$ and $v_0^{i,N} = v_0$, $i = 1, \dots, N$.

Although we are able to apply the particle method to the above system up to this point, notice that in (2.1.4), we need to recompute the values of every kernel function for each particle at each time step. In other words, it takes $\mathcal{O}(N^2)$ evaluations at each time step to update all N particles, which is untractable for large value N. Therefore, a few considerations are introduced to improve computational efficiency.

First we denote an estimate of the leverage function as:

$$\hat{\sigma}(t,x) := \sigma_{Dup}(t,x)\sqrt{\frac{\sum_{i=1}^N \delta_{t,N}(S_t^{i,N} - x)}{\sum_{i=1}^N v_t^{i,N}\delta_{t,N}(S_t^{i,N} - x)}}. \tag{2.1.5}$$

Notice that at each time step, many particles take values that are close to each other. Therefore, it is reasonable to consider the following: At each time step, instead of computing (2.1.5) at every particle value, we select a grid G_t from $\{S_t^{i,N}\}_{i=1}^N$ with $|G_t| \ll N$, and only compute $\hat{\sigma}(t,x)$ for $x \in G_t$; For particles that are not in G_t, we apply cubic interpolation to obtain $\hat{\sigma}(t,x)$.

The size of G_t is set to be $|G_t| = \max(N_1\sqrt{t}, N_2)$, with $N_1 = 30$, $N_2 = 15$ in [1]. As time increases, the distribution of particles tend to have a larger variation, therefore more grid points are used in the computation. The grid points in G_t can be be chosen as equally-spaced values that covers the range of the set $\{S_t^{i,N}\}_{i=1}^N$.

Furthermore, since the value of $\delta_{i,N}(S_t^{i,N} - x)$ is negligible when the distance between x and $S_t^{i,N}$ is large, in the numerator and the denominator of (2.1.5), we only consider to sum up the terms when the kernel function returns a value greater than a pre-determined small positive threshold η. That is,

$$\hat{\sigma}(t,x) = \sigma_{Dup}(t,x)\sqrt{\frac{\sum_{i\in D_x} \delta_{t,N}(S_t^{i,N} - x)}{\sum_{i\in D_x} v_t^{i,N}\delta_{t,N}(S_t^{i,N} - x)}},$$

where $D_x = \{i \in \{1, \dots, N\} | \delta_{t,N}(S_t^{i,N} - x) > \eta\}$. As a rule of thumb, we set $\eta = 10^{-3}$.

The full simulation algorithm of the calibrated LSVMs using the regularising kernel method is presented in Algorithm 1. The inputs of the algorithms include the number of particles N, the expiry T, the number of time steps M, the Dupire volatility function σ_{Dup}, initial values S_0, v_0,

Algorithm 1: Simulation of the calibrated LSVMs with the regularising kernel method

Data: N, T, M, σ_{Dup}, S_0, v_0, η, ρ, $\delta_{t,N}(\cdot)$
$\Delta t \leftarrow T/M$;
$S_0^{i,N} \leftarrow S_0$ for $i = 1, \dots, N$;
$v_0^{i,N} \leftarrow v_0$ for $i = 1, \dots, N$;
$k \leftarrow 1$;
$\hat{\sigma}(t,x) \leftarrow \sigma_{Dup}(0,x)/\sqrt{v_0}$ for $t \in [0, \Delta t]$;
while $k \leq M$ **do**
  $t_k \leftarrow k\Delta t$;
  for *i in* $1, \dots, N$*:* **do**
    $(Z^i, W^i) \overset{\text{i.i.d.}}{\sim} N\Big([0,0], \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\Big)$;
    $S_{t_k}^{i,N} \leftarrow S_{t_{k-1}}^{i,N} + \sqrt{v_{t_{k-1}}^{i,N}} S_{t_{k-1}}^{i,N} \hat{\sigma}(t_{k-1}, S_{t_{k-1}}^{i,N}) \sqrt{\Delta t} Z^i$;
    $v_{t_k}^{i,N} \leftarrow v_{t_{k-1}}^{i,N} + b(t_{k-1}, v_{t_{k-1}}^{i,N})\Delta t + \tilde{\sigma}(t_{k-1}, v_{t_{k-1}}^{i,N})\sqrt{\Delta t} W^i$;
  $G_{t_k} \leftarrow$ equal-spaced grid covering the set $\{S_{t_k}^{i,N}\}_{i=1}^N$, with $|G_{t_k}| = \max(N_1\sqrt{t_k}, N_2)$, $N_1 = 30$, $N_2 = 15$;
  for *x in* G_{t_k} **do**
    $D_x \leftarrow \{i \in \{1, \dots, N\} | \delta_{t_k,N}(S_{t_k}^{i,N} - x) > \eta\}$;
    $\hat{\sigma}(t_k, x) \leftarrow \sigma_{Dup}(t_k, x) \sqrt{\frac{\sum_{i \in D_x} \delta_{t_k,N}(S_{t_k}^{i,N} - x)}{\sum_{i \in D_x} v_{t_k}^{i,N} \delta_{t_k,N}(S_{t_k}^{i,N} - x)}}$;
  for *i in* $1, \dots, N$*:* **do**
    $\hat{\sigma}(t_k, S_{t_k}^{i,N}) \leftarrow$ cubic interpolation;
    $\hat{\sigma}(t, S_{t_k}^{i,N}) \leftarrow \hat{\sigma}(t_k, S_{t_k}^{i,N})$ for $t \in [t_k, t_{k+1}]$;
  $k \leftarrow k + 1$;

the threshold η, the correlation ρ between Brownian motions W and B, and the kernel function $\delta_{t,N}(\cdot)$ with bandwidth h.

In the next sections, we apply the simulation on the calibrated LSVMs in the Black-Scholes and Heston markets respectively, and assess the quality of simulations by comparing the implied volatility curves from the simulated models to the one in the market.

2.2 Black-Scholes market with the regularising kernel method

In the Black-Scholes market, we assume that the true share price dynamic has a constant volatility:

$$dS_t = \sqrt{v_t} S_t d\tilde{W}_t, \ S_0 = 1, \tag{2.2.1a}$$
$$v_t = 0.1024, \ t \in [0, T], \tag{2.2.1b}$$

where $(\tilde{W}_t)_{t \in [0,T]}$ is a Brownian motion under the risk neutral measure.

Under this setting, the constant volatility implies the flat Dupire function $\sigma_{Dup}(t,x) = \sqrt{v_t} = 0.32$ for $t \in [0,T]$, $x \in \mathbb{R}^+$. We perform simulations on two calibrated LSVMs with different dynamics in the volatility process. The first one has the dynamics

$$dS_t = \sqrt{v_t} S_t \frac{\sigma_{Dup}(t, S_t)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[v_t | S_t]}} dW_t, \ S_0 = 1, \tag{2.2.2a}$$
$$dv_t = 0.1\sqrt{v_t} dB_t, \ v_0 = 0.1024, \tag{2.2.2b}$$

where $(B_t)_{t \in [0,T]}$ and $(W_t)_{t \in [0,T]}$ are independent Brownian motions. We refer this model as the 'simple LSVM', since the dynamic of volatility is similar to the true dynamic in (2.2.1). We define a second calibrated LSVM:

$$dS_t = \sqrt{v_t} S_t \frac{\sigma_{Dup}(t, S_t)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[v_t | S_t]}} dW_t, \ S_0 = 1, \tag{2.2.3a}$$
$$dv_t = \lambda(\bar{v} - v_t)dt + \gamma\sqrt{v_t} dB_t, \ v_0 = 0.1024, \tag{2.2.3b}$$

where $(B_t)_{t\in[0,T]}$ and $(W_t)_{t\in[0,T]}$ are correlated Brownian motions with correlation $\rho = -0.315$. The parameters are assigned with values $\lambda = 1.05$, $\bar{v} = 0.0855$, $\gamma = 0.95$. We refer the second model as the 'complex LSVM', since the volatility follows a rather different dynamic to the one in the market. We are interested to see whether both calibrated LSVMs are able to reproduce the market vanilla option prices through the particle simulations with regularising kernels, regardless of the differences in the underlying models.

Remark 2.2.1. Notice that the volatility in the complex LSVM follows a CIR process with positive mean reversion level, and such CIR process will theoretically stay non-negative for any positive initial values. However, the discretisation in time steps could result a negative volatility with non-zero probability. Therefore, we adopt the full truncation Euler scheme for the volatility process in the numerical implementation proposed in [18]. More specifically, we update particles as follows:

1. Discretise the interval $[0,T]$ into M equal-sized time steps $\{t_k\}_{k=0}^M$, and set $\Delta = \frac{T}{M}$.
2. Initialise $S_0^{i,N} = S_0$, $v^{i,N} = v_0$ and $\hat{\sigma}(0, S_0^{i,N}) = \sigma_{Dup}(0, S_0)/\sqrt{v_0}$ for $i = 1, \dots, N$, and set $k = 1$.
3. Generate N i.i.d. bivariate normal random variables $(W^i, B^i) \sim N\Big([0,0], \left[\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right]\Big)$ for $i = 1, \dots, N$.
4. Update $S_t^{i,N}$ and $v_t^{i,N}$ for $i = 1, \dots, N$ from t_{k-1} to t_k by full truncation Euler scheme:

$$S_{t_k}^{i,N} = S_{t_{k-1}}^{i,N} + \sqrt{(v_{t_{k-1}}^{i,N})^+}\hat{\sigma}(t_{k-1}, S_{t_{k-1}}^{i,N}) S_{t_{k-1}}^{i,N}\sqrt{\Delta}W^{i,N}, \tag{2.2.4a}$$

$$v_{t_k}^{i,N} = v_{t_{k-1}}^{i,N} + \lambda(\bar{v} - (v_{t_{k-1}}^{i,N})^+)\Delta + \gamma\sqrt{(v_{t_{k-1}}^{i,N})^+}\sqrt{\Delta}B^{i,N}, \tag{2.2.4b}$$

   where $(v_{t_{k-1}}^{i,N})^+ = \max(0, v_{t_{k-1}}^{i,N})$.
5. Update the leverage function $\hat{\sigma}$ by the regularising kernel method.
6. Increment k by 1. Repeat steps 3 to 5 until $k = M$.

The convergence rate of the full truncation Euler scheme is further studied in [19], where the authors establish the strong convergence in L^p with order $\frac{1}{2}$ for $2 \le p < \nu - 1$, where $\nu = \frac{2\lambda\bar{v}}{\gamma^2}$ is the Feller ratio and is assumed to be strictly greater than 3. We will use this convention without explicitly stating in all numerical schemes that involve CIR volatility process.

We assess the quality of the simulations by comparing the implied volatility curves of vanilla options obtained from the simulated LSVMs to the market volatility level. The implied volatility is defined as the volatility level in the Black-Scholes vanilla option pricing formula, so that the formula would produce the same price as observed in the market, given the same term conditions.

To compute the implied volatility curves from the simulation of the calibrated LSVMs, we first obtain N share prices $\{S_T^{i,N}\}_{i=1}^N$ at expiry T, and compute the estimates of vanilla option prices with respect to a list of strikes. This is done by taking the average of the (discounted) payoffs at time T as in (1.1.2). Finally, we use out-of-money Call and Put options to compute the implied volatility at a given strike level, where we use root finding algorithms in Python Package.

Figure 2.1 shows the implied volatility curves at expiry $T = 1$ obtained from $N = 10^5$ particles, with $M = 1000$ time steps. In the case of the simple LSVM, the implied volatility curve follows the market curve closely. In the complex LSVM, however, although the implied volatility curve has an overall similar shape, the deviation from the market curve is visually observable at small strike values. Table 2.1 displays the absolute errors in the implied volatility curve in percentage at three strike levels, as well as an overall average absolute error across the entire strike grid. Notice that the error in the complex LSVM is approximately one order larger than the one from the simple LSVM.

Recall that the LSVMs are calibrated to the market through the local volatility component $\sigma : [0,T] \times \mathbb{R}^+ \to \mathbb{R}^+ \cup \{0\}$, known as the leverage function. Figure 2.2 plots the estimated leverage functions at time T, $\hat{\sigma}(T, \cdot) : \mathbb{R}^+ \to \mathbb{R}^+ \cup \{0\}$ in both calibrated LSVMs. We observe two different behaviours: the estimated leverage function of the simple LSVM roughly stays constant at value 1, whilst the one of the complex LSVM has an inverted shape. This difference is mainly

caused by different behaviours in the volatility processes. In the simple LSVM, the volatility process stays roughly constant at the initial values v_0, which is very similar to the one in the underlying market. Therefore, in this case, we expect the estimated leverage function to be close to 1. On the other hand, since the volatility process in the complex LSVM is rather different from the one in the market, we expect this difference is adjusted in the calibration through the leverage function.

One way to improve the simulation quality is by adjusting the bandwidth parameter h in the regularising kernel functions. We will defer the discussion to a later section. At this stage, we can see that the simulation of the calibrated LSVMs with the regularising kernel method is able to reproduce a roughly similar implied volatility curve.

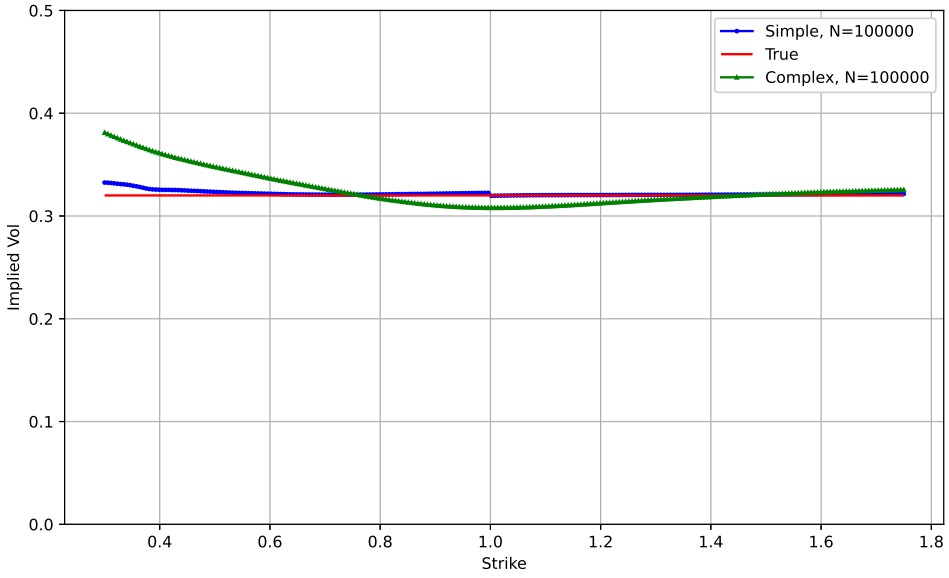

Figure 2.1: Market volatility level and the implied volatility curves from the calibrated LSVMs at $T = 1$ in the Black-Scholes market using the regularising kernel method, with $N = 10^5$, $M = 1$, and kernel bandwidth $h = S_0 N^{-\frac{1}{5}}$.

Absolute error $\lvert\sigma_{Impl}^{True} - \sigma_{Impl}^{LSVM}\rvert$ in %				
Strike	0.7	1	1.3	ave. absolute error
Simple LSVM	0.07717	0.03189	0.05199	0.19035
Complex LSVM	0.6337	1.196842	0.42911	1.22968

Table 2.1: Absolute error in % of implied volatility curves in the Black-Scholes market with the regularising kernel method.

In the Black-Scholes market above, it is known that the market volatility is constant. We now move on to the Heston market, where the market has non-constant volatility.

2.3 Heston market with the regularising kernel method

In the Heston market the volatility follows a CIR process, which introduces more complexity than the previous Black-Scholes market. We assume that the true dynamic of the share price evolves according to the following system of SDEs:

$$dS_t = \sqrt{v_t} S_t d\tilde{W}_t, \ S_0 = 1 \tag{2.3.1a}$$

$$dv_t = \kappa(\theta - v_t)dt + \xi\sqrt{v_t} d\tilde{B}_t, \ v_0 = 0.1024, \tag{2.3.1b}$$

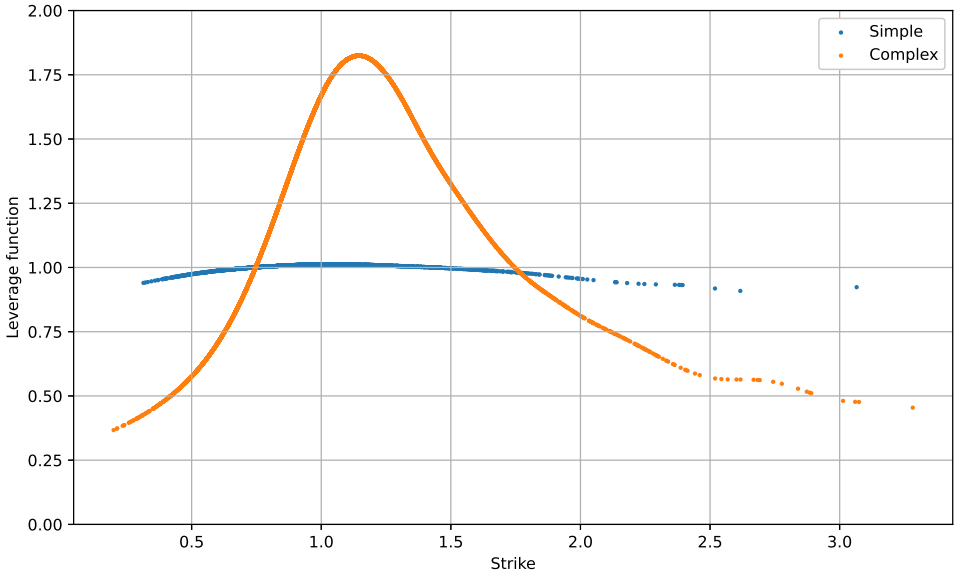

Figure 2.2: Estimated leverage functions in the simple and complex LSVMs at time $T = 1$ over strike K.

with $\kappa = 1.5768$, $\theta = 0.0484$, $\xi = 0.5751$. The Brownian motions $(\tilde{W}_t)_{t\in[0,T]}$ and $(\tilde{B}_t)_{t\in[0,T]}$ are correlated with correlation $\rho = -0.7$.

We first generate market prices of vanilla options with a range of strikes and expiration times using the full truncation Euler scheme. The number of simulations is chosen to be 5×10^5. The synthetic prices are used to approximate σ_{Dup} and the market implied volatility curve at $T = 1$. Note that since we need to approximate second order particle derivatives with respect to strike in σ_{Dup}, the grid of strikes needs to be fine enough for an accurate approximation.

Use the information from the synthetic market, we apply Algorithm 1 to simulate the calibrated LSVMs. As in the Black-Scholes market, we consider two LSVMs: a 'simple LSVM' with similar choices of parameters as in the true market:

$$dS_t = \sqrt{v_t}\frac{\sigma_{Dup}(t, S_t)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[v_t|S_t]}}S_t dW_t,\ S_0 = 1, \tag{2.3.2a}$$

$$dv_t = \lambda(\bar{v} - v_t)dt + \gamma\sqrt{v_t}dB_t,\ v_0 = 0.1024, \tag{2.3.2b}$$

where the parameters $\lambda = \kappa = 1.5768$, $\bar{v} = \theta = 0.0484$ and $\gamma = \xi = 0.5751$. The Brownian motions $(W_t)_{t\in[0,T]}$ and $(B_t)_{t\in[0,T]}$ are correlated with $\rho = -0.7$. The second model 'complex LSVM' has the same form as in (2.3.2), but with a different set of parameter values: $\lambda = 1$, $\bar{v} = 0.0144$, $\gamma = 0.5751$ and $\rho = 0$.

We compare the implied volatility curves from the simulations of the calibrated models to the one from the market at $T = 1$. Figure 2.3 plots the curves, with the number of particles $N = 10^5$ and the number of time steps $M = 1000$.

Both implied volatility curves from the calibrated LSVMs capture the overall shape of the true market curve, although there exists some gaps in between. The absolute errors between the calibrated implied volatility curves and the true one is summarised in Table 2.2. Unlike the case in the Black-Scholes market, the error size is indistinguishable between the simple LSVM and the complex LSVM.

Recall that in the regularising kernel functions, we choose the kernel bandwidth to be $h = S_0 N^{-\frac{1}{5}}$. From the expression (2.1.5), it can be seen that h has an influence on the summations in the conditional expectation, which is a crucial term in the simulation process. We move on to investigate the impact of h on the simulation result in the following subsection.

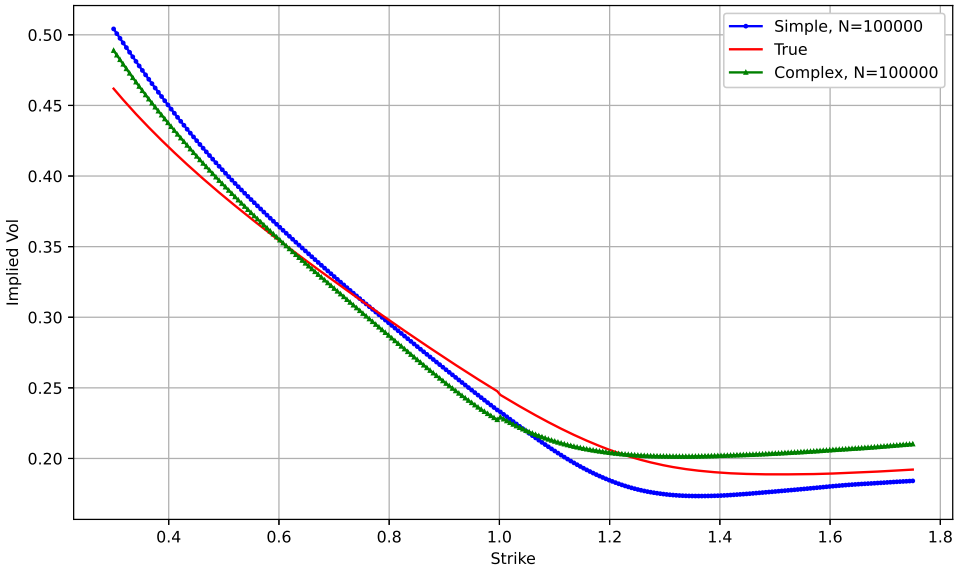

Figure 2.3: Market volatility level and implied volatility curves from the calibrated LSVMs at $T = 1$ in the Heston market using the regularising kernel method, with $N = 10^5$, $M = 1$, and kernel bandwidth $h = S_0 N^{-\frac{1}{5}}$.

Absolute error $\lvert\sigma_{Impl}^{True} - \sigma_{Impl}^{LSVM}\rvert$ in %				
Strike	0.7	1	1.3	Ave. absolute error
Simple LSVM	0.32595	1.20573	2.04098	1.43884
Complex LSVM	0.51804	1.58091	0.63448	1.18096

Table 2.2: Absolute error in % of implied volatility curves in the Heston market with the regularising kernel method.

2.4 Impact of the bandwidth h on the simulation results

From the discussion above, we can see that the particle method with regularising kernels roughly provides a feasible way to simulate the calibrated LSVMs. However, as shown in Figures 2.1 and 2.3, there exists some gap between market implied volatility curves and the ones from the calibrated models. We are interested to see whether it is possible to improve the simulation results by tuning the hyper-parameters in the algorithm. In this subsection, we focus on the choice of bandwidth h in the regularising kernel functions and investigate its impact on the simulation result.

We see that the bandwidth h controls the particle weights in estimating the conditional expectation in (2.1.4). More specifically, when we use regularising kernel method to estimate $\mathbb{E}[v_t|S_t = x]$, with a relatively large h, some weights are distributed to the particles with the components $S_t^{i,N}$ that are further away from x; On the other hand, with a smaller h, only a few particles that have $S_t^{i,N}$ close to x are assigned with majority of the weights. Therefore, it is crucial to see how the simulation results are impacted by the choice of h.

Figure 2.4 graphically shows the quality of the simulations in the Black-Scholes market when we use different values in the bandwidth. We denote $h_0 = S_0 N^{-\frac{1}{5}}$ to be the default bandwidth used in the previous sections. In the left subfigure, the calibrated LSVMs are simulated with $\frac{h_0}{2}$, and the right one is with $\frac{h_0}{5}$. Table 2.3 summarises the absolute errors between true implied volatility curve and the ones from the calibrated models.

From both the figure and the table, we observe that the simple LSVM is less susceptible to the choice of the bandwidth and maintains the relatively low level of average absolute error. However, we see a significant improvement in the complex LSVM when we use a smaller bandwidth.

Similar comparisons are done in the Heston market in Figure 2.5 and Table 2.4. In the left

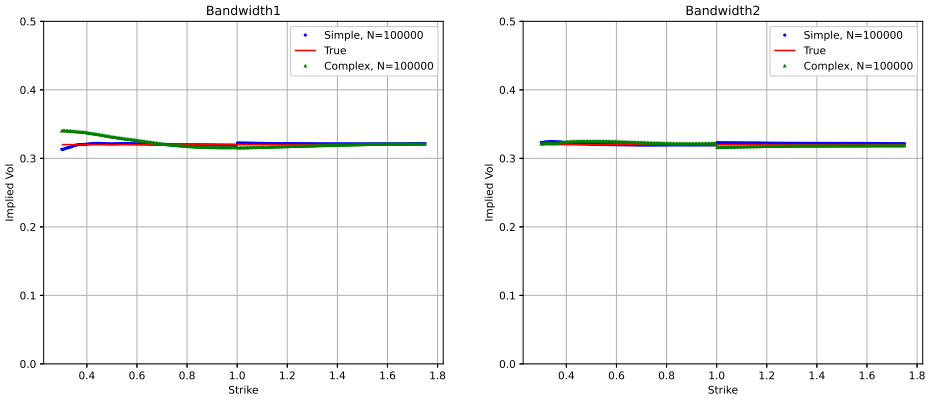

Figure 2.4: Market volatility levels and implied volatility curves from the calibrated LSVMs at $T = 1$ in the Black-Scholes market using the regularising kernel method, with $N = 10^5$, $M = 1$. Left: Bandwidth $h = S_0 N^{-\frac{1}{5}}/2$; Right: Bandwidth $h = S_0 N^{-\frac{1}{5}}/5$.

Absolute error $\lvert\sigma_{Impl}^{True} - \sigma_{Impl}^{LSVM}\rvert$ in %				
Strike	0.7	1	1.3	Ave. absolute error
Simple LSVM, h_0	0.0772	0.0319	0.0520	0.1904
Simple LSVM, $h_0/2$	0.0279	0.2476	0.1272	0.1237
Simple LSVM, $h_0/5$	0.0559	0.2804	0.2001	0.1562
Complex LSVM, h_0	0.6337	1.1968	0.4291	1.2297
Complex LSVM, $h_0/2$	0.0968	0.4667	0.1841	0.4677
Complex LSVM, $h_0/5$	0.2894	0.4035	0.2074	0.2534

Table 2.3: Absolute error in % of implied volatility curves in the Black-Scholes market with various bandwidths in the regularising kernel functions. $N = 10^5$, $M = 1000$.

subplot, the LSVMs are calibrated with $\frac{h_0}{3}$, and the right one with $\frac{h_0}{10}$, and we observe reduction in the average absolute error in both cases. Notice that the gap is still observable between the market implied volatility curve and the simulated ones, and further reducing bandwidth to $\frac{h_0}{10}$ gives a similar average absolute error as in the case of $\frac{h_0}{3}$. However, in the right subplot, we observe that the implied volatility curves from the two calibrated models stays very close to each other.

One possible explanation for the observations above could be the noise in approximating the Dupire function σ_{Dup} from the synthetic market. Recall that the denominator in the Dupire function requires estimations of $\partial_x^2 C(t,x)$, a second order partial derivative of the vanilla Call option prices with respect to the strike x. To obtain an estimate, we used the finite difference method as described in (2.1.3), which may be prone to small noise in the synthetic market prices $C(t,x)$. The two LSVMs may be calibrated to a dynamic that deviates from the true market dynamic due to the noise in the Dupire function estimation. However, since they both use the same estimated Dupire function, we expect to see that the implied volatility curves from the two calibrated models stay close, which is indeed the behaviour in the right subfigure in Figure 2.5.

Absolute error $\lvert\sigma_{Impl}^{True} - \sigma_{Impl}^{LSVM}\rvert$ in %				
Strike	0.7	1	1.3	Ave. absolute error
Simple LSVM, h_0	0.32595	1.20573	2.04098	1.43884
Simple LSVM, $h_0/3$	0.9798	1.3090	0.8382	0.7314
Simple LSVM, $h_0/10$	1.0674	1.2577	0.2237	0.7433
Complex LSVM, h_0	0.51804	1.58091	0.63448	1.18096
Complex LSVM, $h_0/3$	1.2197	1.2075	0.5322	0.7107
Complex LSVM, $h_0/10$	1.2282	1.2152	0.5869	0.7463

Table 2.4: Absolute error in % of implied volatility curves in the Heston market with various bandwidths in the regularising kernel functions.

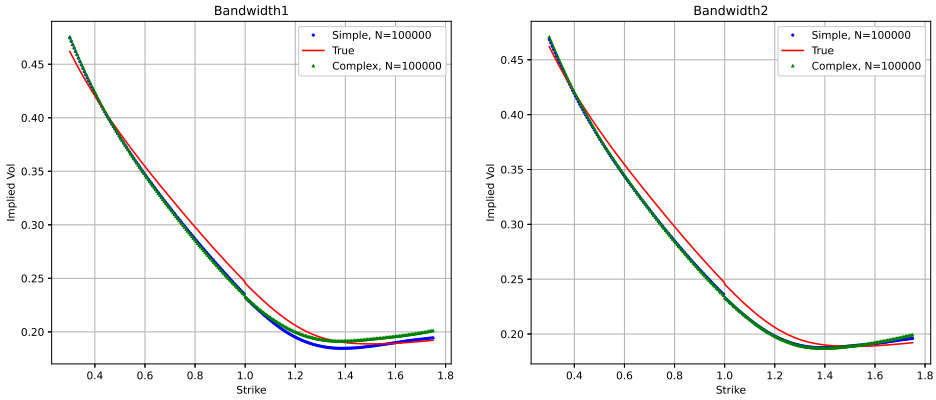

Figure 2.5: Market volatility levels and implied volatility curves from the simple and complex LSVMs at $T = 1$ in the Heston market using regularising kernel method, with $N = 10^5$, $M = 1$. Left: Bandwidth $= S_0 N^{-\frac{1}{5}}/3$; Right: Bandwidth $= S_0 N^{-\frac{1}{5}}/10$.

In summary, from our numerical results, a bandwidth smaller than $S_0 N^{-\frac{1}{5}}$ tends to give a better simulation result. However, if the bandwidth is very close to zero, we effectively turn the kernel functions back into Dirac functions, which should be avoided. Furthermore, a shortcoming of small bandwidths is that it misses the dependence between the volatility and the share prices outside a small region [11]. In other words, when we look at the particles used in the approximation of the conditional expectation $\mathbb{E}^{\mathbb{Q}}[v_t|S_t = x]$, only very few particles with $S_t^{i,N}$ close to the value x are assigned with the most weights, whilst $\delta_{i,N}(S_t^{i,N} - x)$ is close to zero for the others. In the next two sections, we introduce two alternative methods that have been developed for the simulation of the calibrated LSVMs.

Chapter 3

Simulations of the calibrated LSVMs with the RKHS method

Recall that the main difficulty in simulating the calibrated LSVM is to compute a good approximation for the conditional expectation $\mathbb{E}^{\mathbb{Q}}[v_t|S_t = \cdot]$. To address the downside of the regularising kernel method, we present an alternative method that uses Reproducing Kernel Hilbert Spaces (RKHS). This method is proposed and developed in [11], which involves finding a function to directly approximate the conditional expectation.

Without delving into technical details, the RKHS space $\mathcal{H}$ is a Hilbert space of real-valued continuous functions $f : \mathcal{X} \in \mathbb{R}^d \to \mathbb{R}$ for some $d \in \mathbb{N}$. A positive symmetric kernel function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ exists uniquely for each RKHS space, such that for any $x \in \mathcal{X}$, the inner product with the kernel function has the property $\langle f, k_x\rangle_{\mathcal{H}} = f(x)$ for every function $f \in \mathcal{H}$, where $k_x(\cdot) := k(x, \cdot)$.

We are interested in the conditional expectation $\mathbb{E}^{\mathbb{Q}}[v_t|S_t = x]$ for $t \in [0, T]$, and since (S_t, v_t) is a continuous non-degenerating random variable for $t \in (0, T]$, we first define a few quantities using its joint density. Fix $t \in (0, T]$ and denote $q_t(\cdot, \cdot)$ as the joint probability density function of $(S_t, v_t) \in \mathcal{X} \times \mathcal{X}$, where $\mathcal{X} = \mathbb{R}^+$. Furthermore, we assume that (S_t, v_t) has finite second moment for any $t \in (0, T]$ and the function $\mathbb{E}[v_t|S_t = \cdot] \in \mathcal{H}$. Define a function on $\mathcal{X}$:

$$\begin{aligned} c^{q_t}(\cdot) &:= \int_{\mathcal{X}\times\mathcal{X}} k_s(\cdot) v q_t(s, v) dv ds \\ &= \int_{\mathcal{X}\times\mathcal{X}} k_s(\cdot) v q_t(v|s) q_{S,t}(s) dv ds \\ &= \int_{\mathcal{X}} k_s(\cdot) \int_{\mathcal{X}} v q_t(v|s) dv q_{S,t}(s) ds \\ &= \int_{\mathcal{X}} k_s(\cdot) \mathbb{E}^{\mathbb{Q}_t}[v_t|S_t = s] q_{S,t}(s) ds, \end{aligned}$$

where $q_{S,t} : \mathcal{X} \to \mathbb{R}^+ \cup \{0\}$, $q_{S,t}(s) = \int_{\mathcal{X}} q_t(s, v)dv$ and $q_{v,t} : \mathcal{X} \to \mathbb{R}^+ \cup \{0\}$, $q_{v,t}(v) = \int_{\mathcal{X}} q_t(s, v)ds$ are the marginal probability density functions of S_t and v_t respectively, and $q_t(\cdot|s) : \mathcal{X} \to \mathbb{R}^+ \cup \{0\}$, $q_t(v|s) = \frac{q_t(s,v)}{q_{S,t}(s)}$ is the conditional probability density of v_t given $\{S_t = s\}$.

Define an operator $\mathcal{C}^{q_t}$ on functions $f \in \mathcal{H}$:

$$\mathcal{C}^{q_t} f = \int_{\mathcal{X}} k_s(\cdot) f(s) q_{S,t}(s) ds,$$

and we have the following relationship:

$$c^{q_t}(\cdot) = \mathcal{C}^{q_t} \mathbb{E}^{\mathbb{Q}_t}[v_t|S_t = \cdot]. \tag{3.0.2}$$

To simplify the notation, we drop the $\mathbb{Q}_t$ on the expectation and write $\mathbb{E}[v_t|S_t = \cdot]$ instead when there is no ambiguity. Both c^{q_t} and $\mathcal{C}^{q_t}$ are well defined when Assumption K° in [11] is satisfied: the kernel $k(\cdot, \cdot)$ is continuous on $\mathcal{X} \times \mathcal{X}$, and there exists a constant $C > 0$ such that $0 < k(x, x) \leq C(1 + |x|^2)$ for any $x \in \mathcal{X}$. We wish to invert the operator $\mathcal{C}^{q_t}$ to obtain an expression

for $\mathbb{E}[v_t|S_t = \cdot]$, however, this invertability is not guaranteed in general. Notice that the operator $\mathcal{C}^{q_t}$ is symmetric and positive semi-definite, that is, for any functions $f, h \in \mathcal{H}$, we have:

$$\begin{aligned}\langle h, \mathcal{C}^{q_t} f\rangle_{\mathcal{H}} &= \langle h, \int_{\mathcal{X}} k_s(\cdot) f(s) q_{S,t}(s) ds\rangle_{\mathcal{H}} \\ &= \int_{\mathcal{X}} \langle h, k_s(\cdot)\rangle_{\mathcal{H}} f(s) q_{S,t}(s) ds \\ &= \int_{\mathcal{X}} h(s) f(s) q_{S,t}(s) ds,\end{aligned}$$

where we used linearity in the inner product and the properties of kernel functions in the RKHS. Therefore, $\langle h, \mathcal{C}^{q_t} f\rangle_{\mathcal{H}} = \langle f, \mathcal{C}^{q_t} h\rangle_{\mathcal{H}}$, and $\langle f, \mathcal{C}^{q_t} f\rangle_{\mathcal{H}} \geq 0$ for all $f \in \mathcal{H}$. Using the positive semi-definiteness in the operator $\mathcal{C}^{q_t}$, one way to invert the equation (3.0.2) is to apply a similar regularisation as in the ridge regression: we add a regularisation term $\lambda I_{\mathcal{H}}$ with some $\lambda > 0$, where $I_{\mathcal{H}}$ is the identity function in space $\mathcal{H}$. We cite from [11, Section 3] that this adjustment is justified by Hellinger-Teoplitz theorem: for any $\lambda \geq 0$, the spectrum of the operator $\mathcal{C}^{q_t} + \lambda I_{\mathcal{H}}$ is greater that λ. Therefore, for $\lambda > 0$, $(\mathcal{C}^{q_t} + \lambda I_{\mathcal{H}})^{-1}$ exists with norm $\|(\mathcal{C}^{q_t} + \lambda I_{\mathcal{H}})^{-1}\| \leq \lambda^{-1}$.

We would expect that the conditional expectation $\mathbb{E}[v_t|S_t = \cdot]$ is well approximated by $m^{\lambda}(\cdot; q_t) := (\mathcal{C}^{q_t} + \lambda I_{\mathcal{H}})^{-1} c^{q_t}(\cdot)$ when λ stays close to zero.

It turns out that $m^{\lambda}(\cdot; q_t)$ is the optimal solution of the optimisation problem

$$\arg\min_{f \in \mathcal{H}} \left\{ \mathbb{E}(v_t - f(S_t))^2 + \lambda \|f\|^2_{\mathcal{H}} \right\}.$$

Moreover, recall that the conditional expectation $\mathbb{E}[v_t|S_t = \cdot]$ can be interpreted as a function that gives the best estimate of the value v_t given the information of S_t:

$$\mathbb{E}[v_t|S_t = \cdot] = \arg\min_{f \in L^2} \mathbb{E}(v_t - f(S_t))^2.$$

This supports the intuition that $m^{\lambda}(\cdot; q_t)$ would be close to $\mathbb{E}[v_t|S_t = \cdot]$ with small positive λ.

Furthermore, recall in Chapter 1 that one of the challenges in a calibrated LSVM is its well-posedness. A main result achieved in [11, Section 2] is that if we replace the conditional expectation in (1.0.3) with its regularised version

$$\begin{aligned} dS_t &= \sqrt{v_t} S_t \frac{\sigma_{Dup}(t, S_t)}{\sqrt{m^{\lambda}(S_t; q_t)}} dW_t, \\ dv_t &= b(t, v_t) dt + \tilde{\sigma}(t, v_t) dB_t, \end{aligned}$$

for initial values S_0, v_0 and a constant $\lambda > 0$, the Lipschitz continuity can be proved, and well-posedness follows. However, this result is only justified with some positive fixed regularisation parameter λ, and cannot be generalised to the limiting case where $\lambda \to 0$.

Nevertheless, [11, Section 4] proposes that the function $m^{\lambda}(\cdot; q_t)$ can be written as a sum of weighted exponential kernel functions, which can be computed in the simulation. In the next section, we introduce the representation of $m^{\lambda}(\cdot; q_t)$ in the particle system and describe the algorithm used in simulating the calibrated LSVMs with the RKHS method.

3.1 The RKHS method in the calibrated LSVMs

We first write the associated LSVM particle system with the regularised conditional expectation:

$$\begin{aligned} dS_t^{i,N} &= S_t^{i,N} \frac{\sigma_{Dup}(t, S_t^{i,N})}{\sqrt{m^{\lambda}(S_t^{i,N}; \mathbb{Q}_t^N)}} \sqrt{v_t^{i,N}} dW_t^{i,N}, \\ dv_t^{i,N} &= b(t, v_t^{i,N}) dt + \tilde{\sigma}(t, v_t^{i,N}) dB_t^{i,N}, \end{aligned}$$

with $S_0^{i,N} = S_0$ and $v_0^{i,N} = v_0$ for $i = 1, \ldots, N$, where $\mathbb{Q}_t^N$ denotes the empirical distribution of $\{(S_t^{i,N}, v_t^{i,N})\}_{i=1}^N$, and $m^{\lambda}(\cdot; \mathbb{Q}_t^N)$ is the solution of the optimisation problem

$$\arg\min_{f \in \mathcal{H}} \left\{ \frac{1}{N} \sum_{i=1}^{N} (v_t^{i,N} - f(S_t^{i,N}))^2 + \lambda \|f\|^2_{\mathcal{H}} \right\}, \tag{3.1.2}$$

for some $\lambda > 0$. By the RKHS representer theorem in [20], $m^\lambda(\cdot; \mathbb{Q}_t^N)$ can be written as a weighted sum of kernel functions centred at each particle $S_t^{i,N}$:

$$m^\lambda(\cdot; \mathbb{Q}_t^N) = \sum_{i=1}^{N} \alpha_i k_{S_t^{i,N}}(\cdot),$$

where $\alpha_i,\ i = 1, \dots, N$ are the weights, taking values in $\mathbb{R}$.

One issue related with this representation is computational inefficiency. The number of weights α_i to be optimised grows linearly with the number of particles N, and at each time step, it takes $\mathcal{O}(N)$ time complexity to evaluate $k_{S_t^{i,N}}(S_t^{j,N}),\ j = 1, \dots, N$ for each particle. These features make the computation time untractable when we increase the number of particles.

An effective solution to improve computational efficiency is to reduce the number of the kernel functions in the summation. When the number of particles N is large, many realised $\{S_t^{i,N}\}_{i=1}^N$ at each time step are close to each other, and so are their corresponding kernel functions $k_{S_t^{i,N}}(\cdot)$. Therefore, we can choose an positive integer $L \ll N$, select L points $Z_t^j, j = 1, \dots, L$ among $\{S_t^{i,N}\}_{i=1}^N$, and only use the kernel functions centred at these L points to approximate the conditional expectation:

$$\mathbb{E}[v_t | S_t = \cdot] \approx \sum_{j=1}^{L} \beta_j k_{Z_t^j}(\cdot),$$

where $\boldsymbol{\beta} = [\beta_1, \dots, \beta_L]^T$ is a vector of weights, taking values in $\mathbb{R}^L$.

Note that we can further write expressions in vector and matrix forms: Let K be a $N \times L$ matrix, with $K(m,n) = k_{Z_t^n}(S_t^{m,N}) = k(S_t^{m,N}, Z_t^n)$ for $m \in \{1, \dots, N\}$ and $n \in \{1, \dots, L\}$, and we have:

$$K\boldsymbol{\beta} = [\sum_{j=1}^{L} \beta_j k_{Z_t^j}(S_t^{1,N}), \dots, \sum_{j=1}^{L} \beta_j k_{Z_t^j}(S_t^{N,N})]^T.$$

Furthermore, note that

$$\frac{1}{N} \sum_{i=1}^{N} (v_t^{i,N} - f(S_t^{i,N}))^2 = \frac{1}{N} (\mathbf{v}_t^N - K\boldsymbol{\beta})^T (\mathbf{v}_t^N - K\boldsymbol{\beta}),$$

where $\mathbf{v}_t^N = [v_t^{1,N}, \dots, v_t^{1,N}]^T$, and

$$\begin{aligned}
\|\sum_{j=1}^{L} \beta_j k_{Z_t^j}(\cdot)\|_{\mathcal{H}}^2 &= \langle \sum_{j=1}^{L} \beta_j k_{Z_t^j}, \sum_{j=1}^{L} \beta_j k_{Z_t^j} \rangle_{\mathcal{H}} \\
&= \sum_{j=1}^{L} \sum_{l=1}^{L} \beta_j \beta_l \langle k_{Z_t^j}, k_{Z_t^l} \rangle_{\mathcal{H}} && \text{(Linearity of inner product)} \\
&= \sum_{j=1}^{L} \sum_{l=1}^{L} \beta_j \beta_l k_{Z_t^j}(Z_t^l) && \langle f, k_x \rangle = f(x) \\
&= \boldsymbol{\beta}^T R \boldsymbol{\beta},
\end{aligned}$$

where R is a $L \times L$ matrix, $R(m,n) = k(Z_t^m, Z_t^n)$, $m, n \in \{1, \dots, L\}$. The objective function in (3.1.2) can be written in the matrix form

$$\arg\min_{\boldsymbol{\beta} \in \mathbb{R}^L} \left\{ \frac{1}{N} (\mathbf{v}_t^N - K\boldsymbol{\beta})^T (\mathbf{v}_t^N - K\boldsymbol{\beta}) + \boldsymbol{\beta}^T R \boldsymbol{\beta} \right\}.$$

We solve the optimisation problem by the first order differentiation. The optimal $\hat{\boldsymbol{\beta}}$ needs to satisfy the condition:

$$(K^T K + N\lambda R)\hat{\boldsymbol{\beta}} = K^T \mathbf{v}_t^N,$$

and this can be solved using Numpy packages in Python.

To effectively represent the empirical distribution of the simulated particles, we choose the points $Z_1, \dots, Z_L$ to be the equally-spaced quantiles of the set $\{S_{t_k}^{i,N}\}_{i=1}^N$.

Using the expressions and quantities defined above, we summarise the particle simulation for calibrated LSVMs using the RKHS method in Algorithm 2. The inputs of the algorithm are: the number of particle N, the expiry T, the number of time steps M, initial values S_0, v_0, the correlation between share price and volatility process ρ, the Dupire function σ_{Dup}, the regularisation parameter λ, the kernel function $k(\cdot,\cdot)$ and the number of kernel functions L.

Algorithm 2: Simulation of the calibrated LSVMs with the RKHS method

Data: N, T, M, S_0, v_0, σ_{Dup}, λ, k, L, and ρ
$\Delta t \leftarrow T/M$;
$S_0^{i,N} \leftarrow S_0$ for $i = 1, \dots, N$;
$v_0^{i,N} \leftarrow v_0$ for $i = 1, \dots, N$;
$k \leftarrow 1$;
$\hat{\sigma}(t,x) \leftarrow \sigma_{Dup}(0,x)/\sqrt{v_0}$ for $t \in [0, \Delta t]$;
while $k \leq M$ **do**
  $t_k \leftarrow k\Delta t$;
  for *i in* $1, \dots, N$*:* **do**
    $(U^i, W^i) \overset{\text{i.i.d.}}{\sim} N\Big([0,0], \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\Big)$;
    $S_{t_k}^{i,N} \leftarrow S_{t_{k-1}}^{i,N} + \sqrt{v_{t_{k-1}}^{i,N}} S_{t_{k-1}}^{i,N} \hat{\sigma}(t_{k-1}, S_{t_{k-1}}^{i,N}) \sqrt{\Delta t} U^i$;
    $v_{t_k}^{i,N} \leftarrow v_{t_{k-1}}^{i,N} + b(t_{k-1}, v_{t_{k-1}}^{i,N})\Delta t + \tilde{\sigma}(t_{k-1}, v_{t_{k-1}}^{i,N}) \sqrt{\Delta t} W^i$;
  $Z_1, \dots, Z_L \leftarrow L$ points selected among $S_{t_k}^{i,N}, i = 1, \dots, N$ by equal quantiles;
  $R \leftarrow (k(Z_m, Z_n))_{m=1,\dots,L, n=1,\dots,L}$;
  $K \leftarrow (k(S_{t_k}^{i,N}, Z_m))_{i=1,\dots,N, m=1,\dots,L}$;
  $\mathbf{v} \leftarrow [v_{t_k}^{1,N}, \dots, v_{t_k}^{N,N}]^T$;
  $\boldsymbol{\beta} \leftarrow$ Solve $(K^T K + N\lambda R)\boldsymbol{\beta} = K^T \mathbf{v}$;
  for *i in* $1, \dots, N$*:* **do**
    $m^\lambda(S_{t_k}^{i,N}; \mathbb{Q}_t^N) \leftarrow \sum_{j \in L} \beta_j k(Z_j, S_{t_k}^{i,N})$;
    $\hat{\sigma}(t_k, S_{t_k}^{i,N}) \leftarrow \frac{\sigma_{Dup}(t_k, S_{t_k}^{i,N})}{\sqrt{m^\lambda(S_{t_k}^{i,N}; \mathbb{Q}_t^N)}}$;
  $k \leftarrow k + 1$;

In the next two sections, we present the simulation results of the calibrated LSVMs using the RKHS method. As in Chapter 2, we study the quality of the simulations in the Black-Scholes and Heston markets by comparing the implied volatility curves from the simulated models to the one from the market.

3.2 Black-Scholes market with the RKHS method

We use the same Black-Scholes market as in (2.2.1), with constant volatility $v_t = 0.1024$ for $t \in [0, T]$, and consider to simulate two calibrated LSVMs: a simple LSVM

$$dS_t = \sqrt{v_t} S_t \frac{\sigma_{Dup}(t, S_t)}{\sqrt{\mathbb{E}[v_t | S_t]}} dW_t, \ S_0 = 1,$$
$$dv_t = 0.1\sqrt{v_t} dB_t, \ v_0 = 0.1024,$$

with independent $(W_t)_{t \in [0,T]}$ and $(B_t)_{t \in [0,T]}$, and another complex LSVM

$$dS_t = \sqrt{v_t} S_t \frac{\sigma_{Dup}(t, S_t)}{\sqrt{\mathbb{E}[v_t | S_t]}} dW_t, \ S_0 = 1,$$
$$dv_t = \mu(\bar{v} - v_t)dt + \gamma\sqrt{v_t} dB_t, \ v_0 = 0.1024,$$

with $\rho = -0.315$, $\mu = 1.05$, $\bar{v} = 0.0855$ and $\gamma = 0.95$. We wish to see whether the RKHS simulation method is able to cope with different volatility processes and reproduce the implied volatility curves.

We apply the Algorithm 2 with $N = 10^5$ particles, the number of time steps $M = 1000$, $T = 1$, regularisation parameter $\lambda = 10^{-5}$, and we use the default number of kernels $L = 40$ referenced from [11]. The kernel function $k_x(\cdot)$ is selected to be the probability density function of a normal random variable, centred around the mean x with variance 5. We compute the implied volatility curves using out-of-money Call and Put options. Figure 3.1 plots the market implied volatility curve and the ones obtained from the simulated models. We see that the three curves stay very closely to each other. Table 3.1 summarises the absolute errors in the implied volatility curves in percentage. Notice that the average error is comparable with the ones in Table 2.3 using the regularising kernel method with bandwidth $h = h_0/5$.

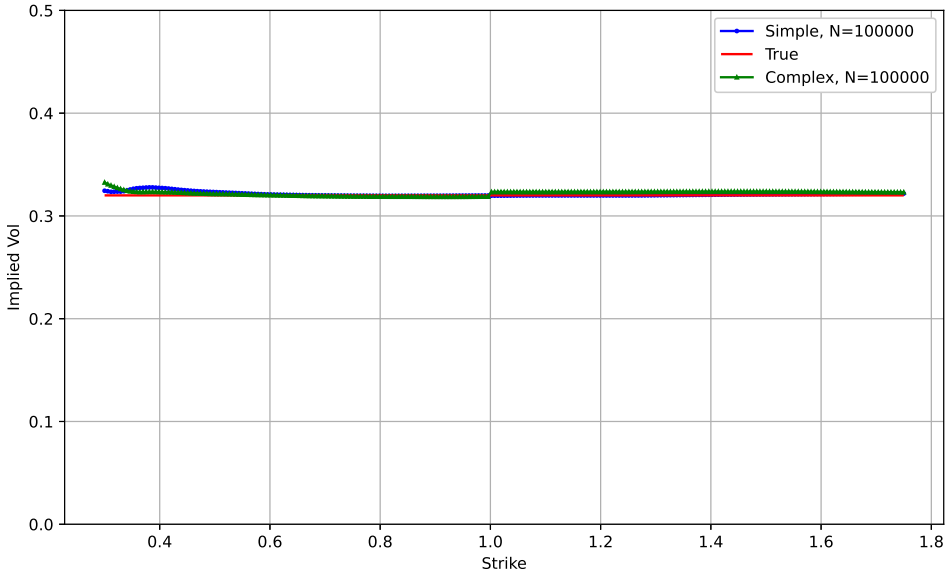

Figure 3.1: Market volatility level and implied volatility curves from the calibrated LSVMs at $T = 1$ in the Black-Scholes market using the RKHS method, with $N = 10^5$, $M = 1$, $L = 40$ and the regularisation parameter $\lambda = 10^{-5}$.

Absolute error $\lvert\sigma_{Impl}^{True} - \sigma_{Impl}^{LSVM}\rvert$ in %				
Strike	0.7	1	1.3	Ave. absolute error
Simple LSVM	0.0077	0.0302	0.0085	0.1318
Complex LSVM	0.0698	0.3369	0.3560	0.2695

Table 3.1: Absolute error in % of implied volatility curves in the Black-Scholes market using the RKHS method.

We see that in the market with flat volatility, the RKHS method manages to simulate the calibrated LSVMs and reproduce the implied volatility curves with decent accuracy, regardless of how the volatility process is different from the market. We now move on to assess the performance in the Heston market with non-constant volatility surface.

3.3 Heston market with the RKHS method

As in Chapter 2, we consider the same Heston market:

$$dS_t = \sqrt{v_t} S_t d\tilde{W}_t,\ S_0 = 1$$
$$dv_t = \kappa(\theta - v_t)dt + \xi\sqrt{v_t} d\tilde{B}_t,\ v_0 = 0.1024,$$

with $\kappa = 1.5768$, $\theta = 0.0484$, $\xi = 0.5751$, $\rho = -0.7$, where ρ represents the correlation between Brownian motions $(\tilde{W}_t)_{\in[0,T]}$ and $(\tilde{B}_t)_{\in[0,T]}$. We simulate two calibrated LSVMs with different

volatility dynamics. Note that the simple LSVM is chosen with a similar but slightly different set of parameter values as the market:

$$dS_t = \sqrt{v_t}\frac{\sigma_{Dup}(t, S_t)}{\sqrt{\mathbb{E}[v_t|S_t]}} S_t dW_t,\ S_0 = 1, \tag{3.3.2a}$$

$$dv_t = \mu(\bar{v} - v_t)dt + \gamma\sqrt{v_t}dB_t, v_0 = 0.1024, \tag{3.3.2b}$$

with $\mu = \kappa = 1.5768$, $\bar{v} = \theta = 0.0484$, $\gamma = \xi = 0.5751$, $v_0 = 0.1024$, and $\rho = -0.5$. The complex LSVM shares the same form as the simple LSVM in (3.3.2), with parameter values: $\mu = 1$, $\bar{v} = 0.0144$, $\gamma = \xi = 0.5751$, $v_0 = 0.0144$ and $\rho = 0$.

We apply Algorithm 2 with the regularisation parameter $\lambda = 10^{-7}$, the number of particles $N = 10^5$, the number of time steps $M = 1000$, $T = 1$, the number of kernels $L = 40$ and the same choice of kernel functions as in the case of the Black-Scholes market. Figure 3.2 plots the implied volatility curves from two simulated models, along with the market implied volatility curve. Table 3.2 summarises the absolute errors in the curves.

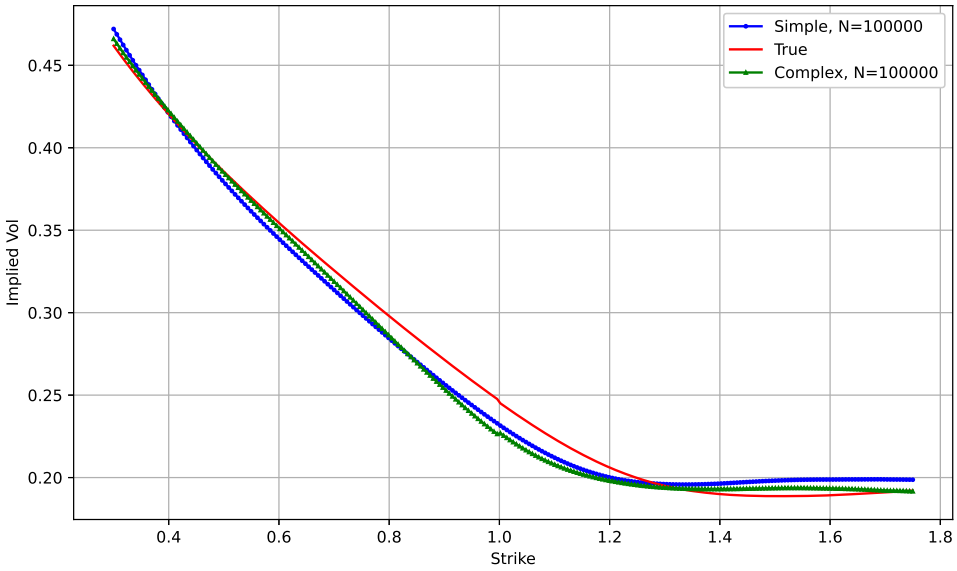

Figure 3.2: Market volatility level and implied volatility curves from the calibrated LSVMs at $T = 1$ in the Heston market using the RKHS method, with $N = 10^5$, $M = 1$, $L = 40$ and regularisation parameter $\lambda = 10^{-7}$.

Absolute error $\lvert\sigma_{Impl}^{True} - \sigma_{Impl}^{LSVM}\rvert$ in %				
Strike	0.7	1	1.3	Ave. absolute error
Simple LSVM	1.1816	1.3517	0.0866	0.8803
Complex LSVM	0.6724	0.8135	0.1226	0.6969

Table 3.2: Absolute error in % of the implied volatility curves in the Heston market using the RKHS method.

From the plot, we see that the implied volatility curves from the calibrated models capture the general shape and trend. Furthermore, it is plausible to see that the average absolute error in the implied volatility curves from the RKHS method has roughly the same size as in the reproducing kernel method with the optimised kernel bandwidth.

Recall that the regularisation parameter λ plays an important role in the RKHS method. We wish to keep it close to zero to obtain a better approximation of the conditional expectation $\mathbb{E}[v_t|S_t]$ for $t \in [0, T]$, whilst the theoretical behaviour of the calibrated LSVMs as $\lambda \to 0^+$ is unclear. In the next section, we investigate the impact of the regularisation term λ on the simulation results.

3.4 Impact of the regularisation parameter on simulation results

We first study how the regularisation parameter λ impacts the average absolute error in implied volatility curves in the Black-Scholes market with the simple LSVM. We select a range of regularisation parameter values $\lambda = 10^{-i}, i = 2, \ldots, 7$, and for each value of λ, we fit the LSVM with number of particles $N = 10^4$ and compute the average absolute error in the implied volatility curve. We then repeat the procedure for 50 times, and compute the mean and standard deviation of the average error for each λ. To reduce the noise in the result, we ensure that for all values of λ in each iteration, the LSVM simulation is run on the same random seed.

Figure 3.3 plots the mean and standard deviation of the average absolute error in percentage of the implied volatility curve over the regularisation parameter λ. We see that the error stays relatively constant when $\lambda \in [10^{-7}, 10^{-5}]$, increases slowly as λ increases from 10^{-5} to 10^{-3}, and eventually takes a jump when λ reaches 10^{-2}.

Furthermore, notice that decreasing λ does not reduce the standard deviations of the error. In the Black-Scholes market, it is reasonable to argue that the error stabilises as λ approaches zero from the right, which coincides with the observations in [11, Section 5].

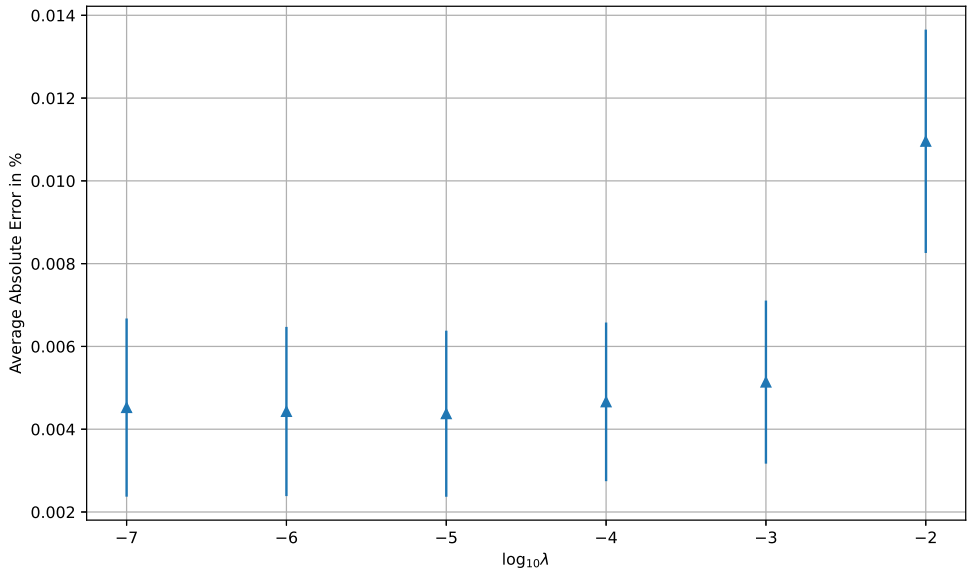

Figure 3.3: Average absolute error vs. the regularisation parameter λ in the Black-Scholes market with the RKHS method.

In the Heston market, where the market implied volatility is no longer flat, we observe some different behaviours. The simulation of the simple LSVM defined in (3.3.2) fails when $\lambda \geq 10^{-6}$, with $N = 10^5$ particles, the number of time steps $M = 1000$ and $T = 1$, whilst the method is able to reproduce the implied volatility curve of the complex LSVM under the same settings with $\lambda \geq 10^{-6}$. By further examining the simulated share prices in the simple LSVM at time $T = 1$, we observe some extreme positive and negative values. Recall that the underlying share price and the volatility processes are correlated with $\rho = -0.5$ in the simple LSVM, whilst there is no correlation in the complex LSVM. This may suggest that the error in estimating the conditional expectation may be magnified when the LSVM has some strong correlation between the share price and the volatility processes, and this could lead to some abnormal behaviours in the share price simulation.

Furthermore, if we strengthen the correlation ρ in the simple LSVM to -0.7 as the one in the underlying market, we fail to find a λ that would allow successful simulations under the same setting ($N = 10^5$, $M = 1000$, $T = 1$). One could try to increase the number of particles and the number of time steps, however this could lead to significant increase in computation time, making the method undesirable comparing to other alternatives.

To summarise, the particle simulation with the RKHS method offers an alternative way to approximate the conditional expectation $\mathbb{E}^{\mathbb{Q}}[v_t|S_t]$ in the calibrated LSVMs. However, as discussed above, the simulation method may encounter some difficulty when the LSVM has strong correlation between the share price dynamic and the volatility process in the Heston market. Furthermore, note that as shown in Algorithm 2, the method involves several matrix multiplications and inversions, which may become computationally expensive with increasing number of particles.

In the next section, we present a third method to approximate $\mathbb{E}^{\mathbb{Q}}[v_t|S_t]$ in the simulation of the calibrated LSVMs. As opposed to the RKHS method, it can be applied with high computational efficiency.

Chapter 4

Simulations of the calibrated LSVMs with the bin Monte Carlo method

Recall that one of the challenges in applying the particle method on the calibrated LSVMs is to estimate the conditional expectation $\mathbb{E}[v_t|S_t]$ appropriately and efficiently. As S_t is a continuous random variable for any $t \in (0, T]$, at any positive time step, the probability of any two particles $(S_t^{i,N}, v_t^{i,N})$ and $(S_t^{j,N}, v_t^{j,N})$ having the event $\{S_t^{i,N} = S_t^{j,N}, i \neq j\}$ is equal to zero. This creates difficulty in using the empirical distribution of the realised particles to estimate the conditional expectation $\mathbb{E}^{\mathbb{Q}}[v_t|S_t = S_t^{i,N}]$, since there will be exactly one value $v_t^{i,N}$ corresponds to the event $\{S_t = S_t^{i,N}\}$.

In the paper [12], the authors propose a simple but effective approach called the bin Monte Carlo method. The method groups the values of S_t into several bins, and computes the conditional expectation of v_t given the bin S_t falls in. More specifically, we partition the range of S_t values into l bins: $(b_1, b_2], (b_2, b_3], \ldots, (b_l, b_{l+1}]$, for some $b_1 \geq 0$ and $b_{l+1} < \infty$. The conditional expectation are approximated as follows:

$$\mathbb{E}^{\mathbb{Q}}[v_t|S_t = x] \approx \mathbb{E}^{\mathbb{Q}}[v_t|S_t \in (b_i, b_{i+1}]],$$

where $x \in (b_i, b_{i+1}]$.

Suppose we set $b_i = x - \epsilon$ and $b_{i+1} = x + \epsilon$ for some small $\epsilon > 0$. As the number of bins grows to infinity, $\epsilon \to 0^+$ and the boundaries of the bin collapse to the value x. This intuitively justifies the approximation in the asymptotic sense. By grouping the particles into bins, it allows us to use the empirical distribution of the realised particles to estimate the conditional expectation, since each bin contains more than one particle:

$$\hat{\mathbb{E}}^{\mathbb{Q}}[v_t|S_t \in (b_i, b_{i+1}]] = \frac{1}{N\hat{\mathbb{Q}}[S_t \in (b_i, b_{i+1}]]} \sum_{i=1}^{N} v_t^{i,N} \mathbb{1}_{S_t^{i,N} \in (b_i, b_{i+1}]}, \tag{4.0.1}$$

where $\hat{\mathbb{Q}}[S_t \in (b_i, b_{i+1}]]$ is an estimate of the probability that the share price falls in the bin $(b_i, b_{i+1}]$ at time t. This probability is determined by the number of bins and the bin type. Two widely used conventions for the bin types are equal bin-width method, where all the bins have the same width, and equal bin-frequency method, where each bin contains approximately the same number of particles. More specifically, at each time t, denote $\bar{S}_t^{1,N}$ and $\bar{S}_t^{N,N}$ as the the smallest and the largest realised particle values for the share price. The equal bin-width method has the following boundary rule and $\hat{\mathbb{Q}}[S_t \in (b_i, b_{i+1}]]$:

$$b_{t,1} = \bar{S}_t^{1,N},\ b_{t,l+1} = \bar{S}_t^{N,N},\ b_{t,i} = b_{t,1} + \frac{i-1}{l}(b_{t,l+1} - b_{t,1}), i = 2, \ldots, l,$$

$$\hat{\mathbb{Q}}[S_t \in (b_i, b_{i+1}]] = \sum_{i=1}^{N} \mathbf{1}_{S_t^{i,N} \in (b_i, b_{i+1}]},$$

whilst the equal bin-frequency method has the boundaries and $\hat{\mathbb{Q}}[S_t \in (b_i, b_{i+1}]]$ as follows:

$$b_{t,i} = \bar{S}_t^{\lfloor \frac{i-1}{l} N \rfloor, N}, \ i = 1, \ldots, l+1,$$
$$\hat{\mathbb{Q}}[S_t \in (b_i, b_{i+1}]] = \frac{1}{l}.$$

Different choice of bins affects the convergence of the simulation result. In the context of the calibrated LSVMs, the distribution of S_t for any $t \in [0, T]$ tends to have higher density around the mid-values, and we wish to obtain a more accurate approximation where the density of S_t is high. Therefore, in the simulation examples, we present the result with equal bin-frequency approach, since it tends to have smaller bin width in the area where S_t is more concentrated.

We present the full algorithm with the bin Monte Carlo method in Algorithm 3. The input of the algorithms are: the number of particles N, the expiry T, the number of time steps M, initial values S_0, v_0, the Dupire function σ_{Dup}, the correlation between share price and volatility processes ρ and the number of bins l.

Algorithm 3: Simulation of the calibrated LSVM with Bin Monte Carlo method

Data: N, T, M, S_0, v_0, σ_{Dup}, l, ρ
$\Delta t \leftarrow T/M$;
$k \leftarrow 1$;
$S_0^{i,N} \leftarrow S_0$ for $i = 1, \ldots, N$;
$v_0^{i,N} \leftarrow v_0$ for $i = 1, \ldots, N$;
$\hat{\sigma}(t, x) \leftarrow \sigma_{Dup}(0, x)/\sqrt{v_0}$ for $t \in [0, \Delta t]$;
while $k \leq M$ **do**
  $t_k \leftarrow k\Delta t$;
  for *i in* $1, \ldots, N$*:* **do**
    $(U^i, W^i) \overset{\text{i.i.d.}}{\sim} N\Big([0,0], \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\Big)$;
    $S_{t_k}^{i,N} \leftarrow S_{t_{k-1}}^{i,N} + \sqrt{v_{t_{k-1}}^{i,N}} S_{t_{k-1}}^{i,N} \hat{\sigma}(t_{k-1}, S_{t_{k-1}}^{i,N}) \sqrt{\Delta t} U^i$;
    $v_{t_k}^{i,N} \leftarrow v_{t_{k-1}}^{i,N} + b(t_{k-1}, v_{t_{k-1}}^{i,N})\Delta t + \tilde{\sigma}(t_{k-1}, v_{t_{k-1}}^{i,N}) \sqrt{\Delta t} W^i$;
  $b_1, \ldots, b_{l+1} \leftarrow \frac{i-1}{l}^{th}$ quantile of $\{S_{t_k}^{i,N}\}_{i=1}^N$;
  Sort particles $\{(S_{t_k}^{i,N}, v_{t_k}^{i,N})\}_{i=1}^N$, and assign each of them to a bin $(b_i, b_{i+1}]$;
  for *each bin* $(b_i, b_{i+1}]$*, i in* $1, \ldots, l$*:* **do**
    $\mathbb{E}^{\mathbb{Q}}[v_t | S_t \in (b_i, b_{i+1}]] = \frac{l}{N} \sum_{i=1}^N v_t^{i,N} \mathbb{1}_{S_{t_k}^{i,N} \in (b_i, b_{i+1}]}$;
  for *i in* $1, \ldots, N$*:* **do**
    $\hat{\sigma}(t_k, S_{t_k}^{i,N}) \leftarrow \frac{\sigma_{Dup}(t_k, S_{t_k}^{i,N})}{\sqrt{\mathbb{E}^{\mathbb{Q}}[v_t | S_t \in (b_j, b_{j+1}]]}}$, $S_{t_k}^{i,N} \in (b_j, b_{j+1}]$;
  $k \leftarrow k + 1$;

In the next two sections, we apply the algorithm in the Black-Scholes and the Heston markets, and assess the simulation performance through the implied volatility curves.

4.1 Black-Scholes Market with the bin Monte Carlo method

We consider the same Black-Scholes market as in Chapter 2, with constant volatility $v_t = 0.1024$ for $t \in [0, T]$, and the same simple and complex LSVMs in (2.2.2) and (2.2.3). We apply Algorithm 3 with input parameters $N = 10^5$, $M = 1000$, $T = 1$ and select the number of bins $l = 20$. Figure 4.1 shows the market implied volatility curve and the ones from the simulations of the calibrated models at time $T = 1$. Table 4.1 summarises the absolute errors in the implied volatility curves. Visually, the simulated implied volatility curves almost perfectly matches the market volatility except some slight deviations at small values of the strike. The result of the average absolute error is remarkably favourable: The error is almost halved in both simple and complex LSVMs compared to case with the RKHS method.

It is also worth noting that the method is very efficient: Unlike the regularising kernel method and the RKHS method, the simple approximation of the conditional expectation in the bin Monte

Carlo method does not involve complex computations, thus make it very fast. We will have a more detailed discussion regarding the computational speed in a later chapter. Despite being fast, the bin Monte Carlo method is able to reproduce the implied volatility curve in the Black-Scholes market with high accuracy. We move on to see its performance when the market has a different volatility dynamic in the Heston market.

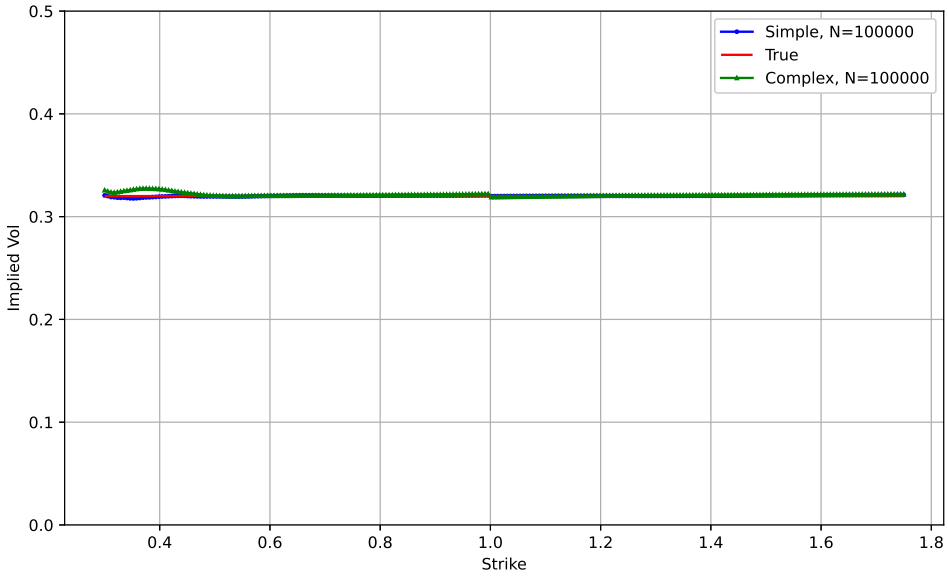

Figure 4.1: Market volatility level and implied volatility curves from the calibrated LSVMs at $T = 1$ in the Black-Scholes market using the bin Monte Carlo method, with $N = 10^5$, $M = 1$ and the number of bins $l = 20$.

Absolute error $\lvert\sigma_{Impl}^{True} - \sigma_{Impl}^{LSVM}\rvert$ in %				
Strike	0.7	1	1.3	Ave. absolute error
Simple LSVM	0.0574	0.0037	0.0147	0.0505
Complex LSVM	0.0509	0.0995	0.0680	0.1428

Table 4.1: Absolute error in % of implied volatility curves in the Black-Scholes market with the bin Monte Carlo method.

4.2 Heston market with the bin Monte Carlo method

We consider the same Heston market, simple and complex LSVMs as defined in Section 2.3, and apply Algorithm 3 with the number of particles $N = 10^5$, $M = 1000$, $T = 1$, and the equal bin-frequency method with the number of bins $l = 20$. Figure 4.2 and Table 4.2 summarise the simulation result. The size of magnitude of the average absolute errors are in the same order as the previous two methods, and the implied volatility curves from the simulated models capture the downwards skewed shape as the one in the market.

Notice that unlike the case with the RKHS method, the bin Monte Carlo method does not experience any difficulty in simulating the calibrated LSVMs with strong correlation between the share price and the volatility processes.

When applying the bin Monte Carlo method, one needs to choose the number of bins. We move on to see whether the simulation results can be improved by a different choice of the number of bins.

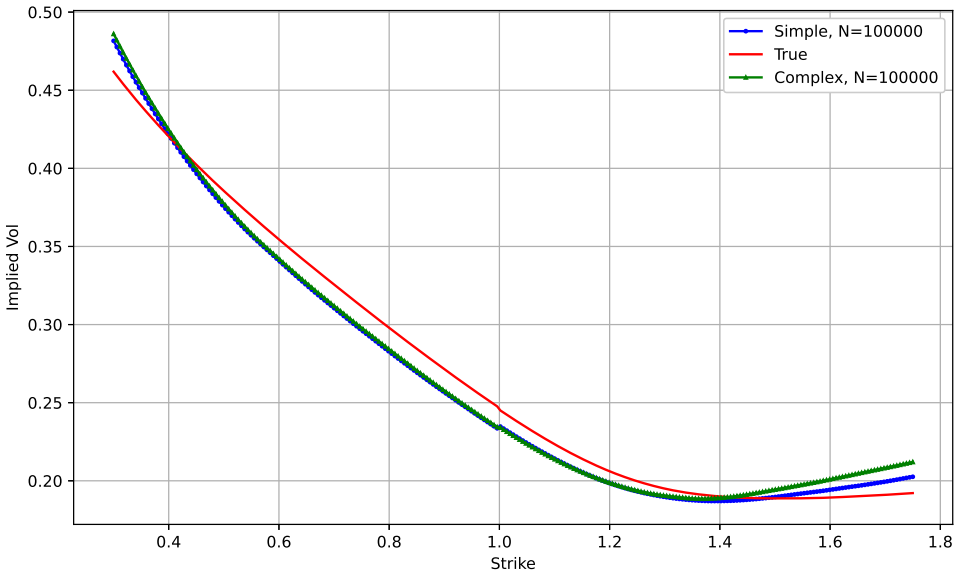

Figure 4.2: Market volatility level and implied volatility curves from the calibrated LSVMs at $T = 1$ in the Heston market using the bin Monte Carlo method, with $N = 10^5$, $M = 1$ and the number of bins $l = 20$.

Absolute error $\lvert\sigma_{Impl}^{True} - \sigma_{Impl}^{LSVM}\rvert$ in %				
Strike	0.7	1	1.3	Ave. absolute error
Simple LSVM	1.5067	1.0294	0.5382	0.9081
Complex LSVM	1.3634	1.0951	0.4564	1.0055

Table 4.2: Absolute error in % of implied volatility curves in the Heston market with the bin Monte Carlo method.

4.3 The impact of the number of bins on simulation results

We focus on the impact of the number of bins on LSVM simulations in the Heston market with equal bin-frequency method. Recall that in the equation (4.0.1), the number of bins l determines the number of estimates we compute on the conditional expectation at each time step. The larger the l, the smaller the bin width, but the fewer particles in each bin. Therefore, it is important find a balance in between: we wish to have a large number of bins so that each bin is smaller in width, but on the other hand, we wish to have a decent amount of particles within each bin for a good estimation. In the extreme case where we select the number of bins to be equal to the size of the particles, each bin will contain exactly one particle. The estimated value of the conditional expectation in each bin is effectively the realised volatility of the single particle, which is very noisy and should be avoided.

Figure 4.3 plots the simulation results with different numbers of bins. The left subplot uses $l = 50$ bins and the right one is with $l = 200$ bins. Table 4.3 summarises the absolute errors in implied volatility curves from the simulations. We see that in both cases the accuracy is improved comparing to the base case $l = 20$. Visually, we observe that the two simulated implied volatility curves from the calibrated LSVMs are very close to each other. Note that similar to Figure 2.4 in the regularising kernel method, there exists a gap with the similar size between calibrated curves and the market curve, which may be caused by the noise in approximating the Dupire function.

It is worth noting that there is a slight difference in the simulation accuracy between the cases with the number of bins $l = 50$ and $l = 200$, and this difference is small. However, the number of bins directly impacts the simulation time. To see this, we record the average simulation time over 7 runs of a calibrated LSVM under the Heston model with $N = 1000$ particles, the number of time steps $M = 1000$ and $T = 1$ over various numbers of bins l. Figure 4.4 shows such relationship. We

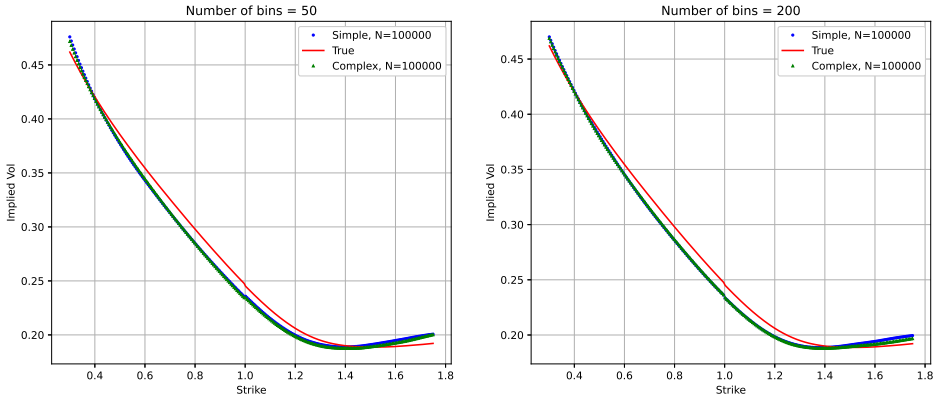

Figure 4.3: Market volatility level and implied volatility curves from the calibrated LSVMs at $T = 1$ in the Heston market using the bin Monte Carlo method, with $N = 10^5$ and $M = 1$. Left: the number of bins $l = 50$; Right: the number of bins $l = 200$.

Absolute error $\lvert\sigma_{Impl}^{True} - \sigma_{Impl}^{LSVM}\rvert$ in %				
Strike	0.7	1	1.3	Ave. absolute error
Simple LSVM, $l = 20$	1.5067	1.0294	0.5382	0.9081
Simple LSVM, $l = 50$	1.2589	0.9241	0.3831	0.7674
Simple LSVM, $l = 200$	1.1528	1.1469	0.4698	0.7133
Complex LSVM, $l = 20$	1.3634	1.0951	0.4564	1.0055
Complex LSVM, $l = 50$	1.2095	1.0622	0.5196	0.7400
Complex LSVM, $l = 200$	1.0657	1.1264	0.5202	0.6601

Table 4.3: Absolute error in % of implied volatility curves in the Heston market with the bin Monte Carlo method.

see that the computation time grows linearly with the number of bins used. Therefore, depending on situations, it may be worth considering to trade a fraction of accuracy for a faster computation.

In the next Chapter, we summarises the three methods we have introduced so far, and provide a short comparisons among them.

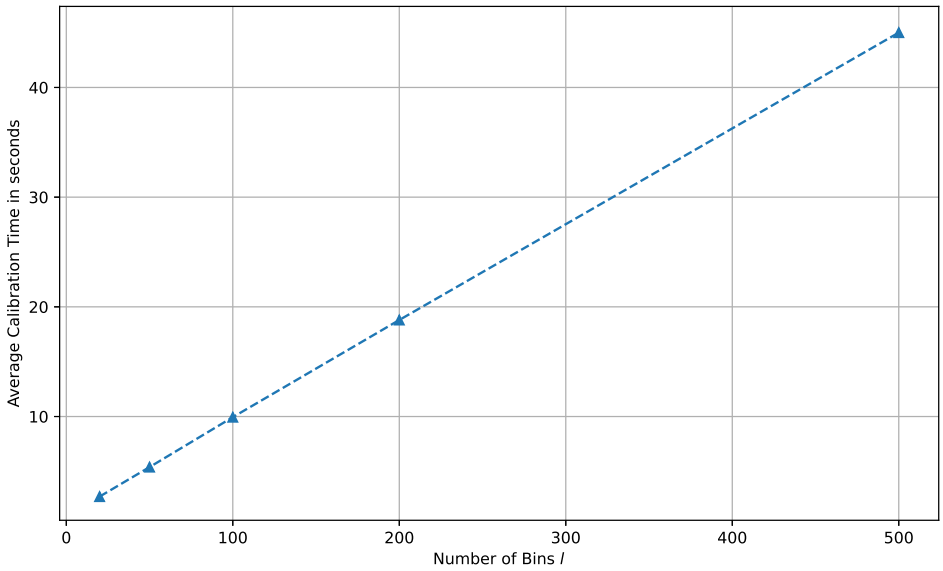

Figure 4.4: Simulation time vs. the number of bins l with the bin Monte Carlo method in the Heston market, with $N = 1000$, $M = 1000$, and $T = 1$.

Chapter 5

Comparisons among the regularising kernel method, the RKHS method and the bin Monte Carlo method

We have introduced three different methods in estimating the conditional expectation term $\mathbb{E}[v_t|S_t]$ in the simulations of the calibrated LSVMs, and in this chapter we give some comparisons among different methods. We focus on computational efficiency, accuracy in the simulation and provide an overall summary of the three methods in the next few sections.

5.1 Computational efficiency vs the number of particles

Computational efficiency is critical when applying an algorithm in practice. In the simulation of the calibrated LSVMs, it is important to understand how the computational power increases with the increasing number of particles N. In the following analysis, we mimic the practical implementations in the real world with the use of efficient tools in Python packages, including vectorisations in Numpy and Pandas packages.

We compare the three methods by varying the number of particles N under the same numerical setting: the same number of time steps $M = 1000$ with expiry $T = 1$. All three methods are tasked to calibrate the simple LSVM in the Black-Scholes market as defined in (2.2.2). Recall that in each method, there exists a hyper-parameter which further determines how to improve the computational efficiency: the size of grid $|G_t|$ where the conditional expectations are estimated in each time step in the regularising kernel method, the number of kernel functions L to approximate the conditional expectation in the RKHS method, and the number of bins l in the bin Monte Carlo method. To control the impact from these parameters, we set $|G_t| = 40$ for all $t \in [0, T]$, $L = l = 40$.

Figure 5.1 plots the simulation time in seconds versus the number of particles N. It can be seen that although all three methods have linear time complexity, the difference among them is significant. It is evident that the RKHS method has a significant larger slope than the others, and this leads to a substantial difference in the efficiency with a large N: If we take 200000 particles, the computation time of the RKHS method is approximately 6 times of the one with the bin Monte Carlo method. Note that the bin Monte Carlo method is roughly twice more efficient than the regularising kernel method with larger values N, making it the most time efficient method among the three.

The high efficiency in the bin Monte Carlo method is mainly attributable to the vectorisation and DataFrame tools in Python packages, which avoid slow Python loops. However, in the RKHS method, the algorithm involves several matrix multiplications and inversions. These operations are computationally heavy and hard to be optimised, which may lead to lower efficiency.

In the next section, we move on to analyse the accuracy in simulations with different number of particles.

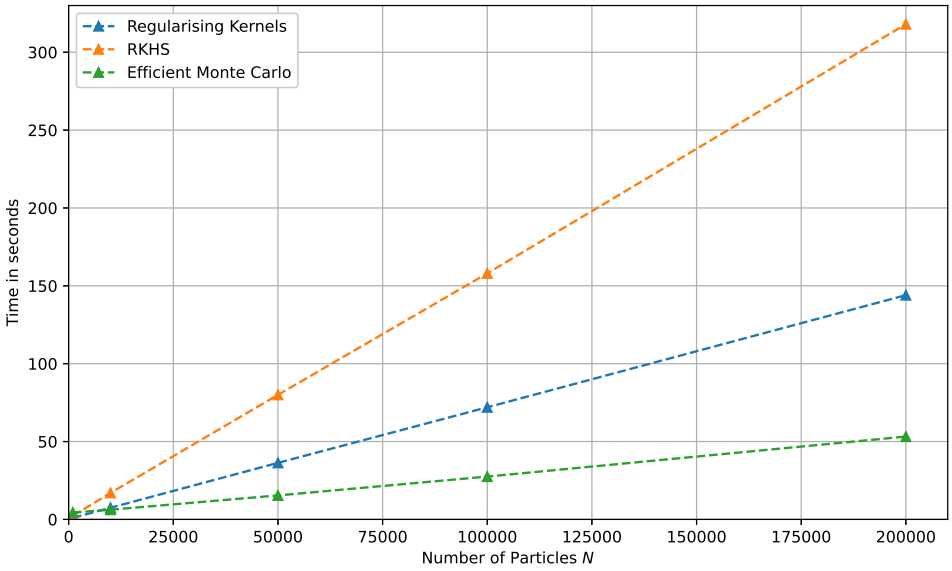

Figure 5.1: Computation time vs. the number of particles N, with $M = 1000$ and $T = 1$.

5.2 Accuracy vs the number of particles

In practice we may consider using a small number of particles N to simulate the LSVMs for a low computation time. In this section, we study the accuracy of the three methods when the number of particles is small.

We select a grid of the number of particles $N \in \{500, 1000, 5000, 10000\}$ and apply the three methods in the Heston market defined in (2.3.1). All three methods are set to simulate the calibrated LSVM

$$dS_t = \sqrt{v_t}\frac{\sigma_{Dup}(t, S_t)}{\sqrt{\mathbb{E}[v_t|S_t]}}S_t dW_t,\ S_0 = 1,$$
$$dv_t = \mu(\bar{\mu} - v_t)dt + \gamma\sqrt{v_t}dB_t,\ v_0 = 0.1024,$$

with independent Brownian motions $(W_t)_{t\in[0,T]}$ and $(B_t)_{t\in[0,T]}$, $\mu = 1$, $\bar{v} = 0.0144$, $\gamma = 0.5751$, $v_0 = 0.0144$. The simulations are carried out with the number of time steps $M = 1000$, the expiry $T = 1$ and various N. The parameters in the three methods are set as follows: the regularising kernel method with the kernel bandwidth $h = S_0 N^{\frac{1}{5}}/10$ and default grid size $|G_t|$ as defined in Chapter 2, the RKHS method with regularisation parameter $\lambda = 10^{-5}$ and default number of kernel functions $L = 40$, and the bin Monte Carlo method with the number of bins l to be 20 for $N \in \{500, 1000\}$, and 40 for $N \in \{5000, 10000\}$. We compare the mean of average absolute errors in the implied volatility curves over 50 simulations for each N.

Figure 5.2 shows the plot of the mean of average absolute errors in the implied volatility curves in percentage versus the number of particles N. As the number of particles increases, the error decreases in all three methods. However, the reduction is more significant in the regularising kernel method and the bin Monte Carlo method than the RKHS method. The accuracy in the RKHS method may be improved by the choosing a smaller regularisation parameters $\lambda < 10^{-5}$, however, in our experiment, the simulation with a smaller λ may fail to reproduce the implied volatility curve in the Heston market given a small number of particles.

In the next section, we summarise the three methods and conclude the chapter.

5.3 Summary

We introduced three different methods to approximate the conditional expectation in the particle simulation scheme of the calibrated LSVMs. The three methods use different intuitions and

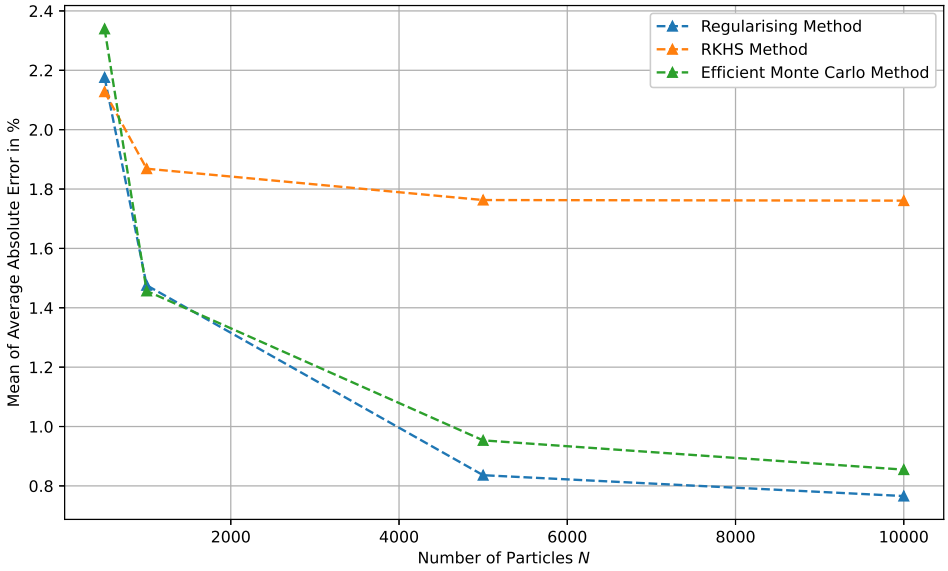

Figure 5.2: The mean of average absolute error in implied volatility curves in % vs. the number of particles N.

techniques, which leads to different implementation complexity and calibration accuracy.

The regularising kernel method approximates the density of (S_t, v_t) for $t \in (0, T]$ based on the empirical distribution of the realised particles. It uses the Nadaraya-Watson estimator for the conditional expectation estimation by replacing Dirac functions with continuous kernel functions $\delta_{t,N}$:

$$\mathbb{E}^{\mathbb{Q}_t}[v_t|S_t = x] \approx \frac{\sum_{i=1}^N v_t^{i,N} \delta_{t,N}(S_t^{i,N} - x)}{\sum_{i=1}^N \delta_{t,N}(S_t^{i,N} - x)}.$$

Furthermore, the implementation of the method is relatively efficient in Python, and the method is able to simulate the calibrated LSVMs and reproduce market implied volatility curves with relatively high accuracy even with a small number of particles. However, it is important to note that the calibration accuracy is sensitive to the choice of the bandwidth in the kernel functions.

On the other hand, instead of approximating any density functions, the RKHS method approaches the calibrated LSVMs by interpreting the conditional expectation $\mathbb{E}^{\mathbb{Q}_t}[v_t|S_t = \cdot]$ as a function and the solution of the optimisation problem

$$\arg\min_{f \in L^2} \mathbb{E}(v_t - f(S_t))^2. \tag{5.3.1}$$

It uses the RKHS space $\mathcal{H}$ to obtain an approximated solution in (5.3.1) by solving an associated regularised problem

$$\arg\min_{f \in \mathcal{H}} \mathbb{E}(v_t - f(S_t))^2 + \lambda \|f\|_{\mathcal{H}}^2.$$

The solution $m^\lambda(\cdot; q_t)$ of the regularised problem can be written as a weighted sum of exponential kernel functions, where q_t is the true bivariate distribution of (S_t, v_t) for $t \in [0, T]$. Moreover, [11] establishes the results that if we replace $\mathbb{E}^{\mathbb{Q}_t}[v_t|S_t = \cdot]$ with $m^\lambda(\cdot; q_t)$ in the calibrated LSVM (1.0.3), the well-posedness of such SDEs can be established with some positive fixed regularisation parameter λ.

However, despite the success in establishing the theoretical results, the RKHS method is more computationally demanding in practice. Furthermore, as discussed in Section 3.4, it could be challenging to apply the RKHS method when the underlying LSVM has a strong correlation between the share price and volatility processes in a market with non-constant volatility surface. However, this difficulty is not encountered by the other two methods.

Finally, the bin Monte Carlo method approximates the conditional expectation by partitioning the particles into l bins:

$$\mathbb{E}^{\mathbb{Q}}[v_t|S_t = x] \approx \mathbb{E}^{\mathbb{Q}}[v_t|S_t \in (b_i, b_{i+1}]],$$

where $x \in (b_i, b_{i+1}]$, for $i = 1, \ldots, l$.

In contrast to the RKHS method, the bin Monte Carlo method is very computationally efficient. Moreover, similar to the regularising kernel method, the bin Monte Carlo method is able to reproduce the implied volatility curve with relatively high accuracy using a small number particles. Together with its high computational efficiency, it makes the bin Monte Carlo method a good choice to be applied in practice.

So far we discussed the simulation of the LSVMs with zero interest rate. In practice, however, interest rates play an important role in option pricing. In the next chapter, we discuss the LSVM simulations with stochastic interest rates and conclude the chapter with a numerical simulation example.

Chapter 6

Simulations of calibrated SIR-LSVMs

In this chapter, we introduce the calibrated LSVMs with stochastic interest rates (SIR-LSVM) in the Foreign Exchange market. We first state the SIR-LSVM calibration condition and briefly explain the particle method for the calibrated dynamic. Finally, we apply the method to simulate the calibrated SIR-LSVM in a synthetic market and compare the simulated implied volatility curves to the one from the market. The model and the method in this chapter is referenced from [21].

6.1 SIR-LSVMs with stochastic rates and the calibration condition

In the Foreign Exchange market (FX), let $(r_t^d)_{t\in[0,T]}$ and $(r_t^f)_{t\in[0,T]}$ be the short rate processes of the domestic currency and the foreign currency between time zero and a fixed time $T > 0$. Define $(S_t)_{t\in[0,T]}$ to be the exchange rate process between the domestic and foreign currencies. By convention, it is interpreted that one unit of the foreign currency is worth S_t units of the domestic currency at time $t \in [0, T]$. We introduce $(D_t^d)_{t\in[0,T]}$ and $(D_t^f)_{t\in[0,T]}$ as discount factors in the domestic and the foreign money markets respectively, where

$$D_t^d = \exp\left(-\int_0^t r_s^d ds\right),$$
$$D_t^f = \exp\left(-\int_0^t r_s^f ds\right).$$

Consider a SIR-LSVM defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q}^d)$, where $\mathbb{Q}^d$ is the risk neutral measure in the domestic money market:

$$dS_t = (r_t^d - r_t^f)S_t dt + \sigma(t, S_t)\sqrt{v_t}S_t dW_t, \tag{6.1.2a}$$
$$r_t^d = g_t^d + h^d(t), \tag{6.1.2b}$$
$$r_t^f = g_t^f + h^f(t), \tag{6.1.2c}$$
$$dg_t^d = \kappa_d(\theta_d - g_t^d)dt + \xi_d\sqrt{g_t^d}dW_t^d, \tag{6.1.2d}$$
$$dg_t^f = \left(\kappa_f(\theta_f - g_t^f) - \rho_{Sf}\xi_f\sqrt{g_t^f}\sigma(t, S_t)\sqrt{v_t}\right)dt + \xi_f\sqrt{g_t^f}dW_t^f, \tag{6.1.2e}$$
$$dv_t = \kappa(\theta - v_t)dt + \xi\sqrt{v_t}dW_t^v, \tag{6.1.2f}$$

with initial values S_0, v_0 take values in $\mathbb{R}^+$, r_0^d, r_0^f in $\mathbb{R}$, non-negative parameters κ_d, θ_d, ξ_d, κ_f, θ_f, ξ_f, κ, θ, ξ and deterministic functions $h^d : [0, T] \to \mathbb{R}$ and $h^f : [0, T] \to \mathbb{R}$. (W, W^v, W^d, W^f) is a four-dimensional Brownian motion, with the correlation structure:

$$d\langle W_t, W_t^v\rangle = \rho dt,\ d\langle W_t, W_t^d\rangle = \rho_{Sd}dt,\ d\langle W_t, W_t^f\rangle = \rho_{Sf}dt,\ d\langle W_t^d, W_t^f\rangle = \rho_{df}dt,$$

where ρ, ρ_{Sd}, ρ_{Sf} and $\rho_{df} \in (-1,1)$. The SIR-LSVM is calibrated to the market when the leverage function $\sigma : [0,T] \times \mathbb{R}^+ \to \mathbb{R}^+$ satisfies the condition:

$$\sigma^2(t,x) = \frac{\mathbb{E}^{\mathbb{Q}^d}[D_t^d | S_t = x]}{\mathbb{E}^{\mathbb{Q}^d}[D_t^d v_t | S_t = x]} \left(\sigma_{Dup}(t,x)^2 + \frac{\mathbb{E}^{\mathbb{Q}^d}[Q_t]}{\frac{1}{2}x^2 \frac{\partial^2 C(t,x)}{\partial x^2}} \right),$$

where

$$Q_t = D_t^d (r_t^f - \bar{r}_t^f)(S_t - x)^+ - x D_t^d \mathbb{1}_{S_t \geq x} [(r_t^d - \bar{r}_t^d) - (r_t^f - \bar{r}_t^f)],$$

and $\bar{r}_t^d = -\frac{\partial}{\partial t} \log P^d(0,t)$ and $\bar{r}_t^f = -\frac{\partial}{\partial t} \log P^f(0,t)$ are the instantaneous forward rates. $P^d(0,t)$ and $P^f(0,t)$ are the zero coupon bond prices in the domestic and foreign money markets. The Dupire local volatility function $\sigma_{Dup} : [0,T] \times \mathbb{R}^+$ with stochastic rates is defined as follows:

$$\sigma_{Dup}(t,x) = \frac{\frac{\partial C(t,x)}{\partial t} + x(\bar{r}_t^d - \bar{r}_t^f)\frac{\partial C(t,x)}{\partial x} + \bar{r}_t^f C(t,x)}{\frac{1}{2}x^2 \frac{\partial^2 C(t,x)}{\partial x^2}},$$

where $C(t,x)$ is the market price of a vanilla Call option on the exchange rate S_t, with maturity t and strike x.

Notice that the domestic and the foreign short rates follow CIR++ dynamics in (6.1.2), and are able to be calibrated to the market and reproduce the zero coupon bond market prices. [21] gives more details in how to calibrate the corresponding parameters in the short rate models, which is beyond the scope of this project. In rest of our discussion, we assume that the parameters $\kappa_d, \theta_d, \xi_d, \kappa_f, \theta_f, \xi_f$ and the functions h^d, h^f are known and have already been calibrated to the market. Given some parameters κ, θ, ξ in the volatility process, we introduce a simulation scheme using the particle method for the calibrated SIR-LSVM.

To achieve this, we consider N particles $\{(S_t^{i,N}, v_t^{i,N}, r_t^{d,i,N}, r_t^{f,i,N}, D_t^{d,i,N})_{t\in[0,T]}\}_{i=1}^N$ with the following dynamics:

$$\begin{aligned}
dS_t^{i,N} &= (r_t^{d,i,N} - r_t^{f,i,N}) S_t^{i,N} dt + \hat{\sigma}(t, S_t^{i,N}) \sqrt{v_t^{i,N}} S_t^{i,N} dW_t^{i,N}, \\
r_t^{d,i,N} &= g_t^{d,i,N} + h^d(t), \\
r_t^{f,i,N} &= g_t^{f,i,N} + h^f(t), \\
dg_t^{d,i,N} &= \kappa_d(\theta_d - g_t^{d,i,N})dt + \xi_d \sqrt{g_t^{d,i,N}} dW_t^{d,i,N}, \\
dg_t^{f,i,N} &= \left(\kappa_f(\theta_f - g_t^{f,i,N}) - \rho_{Sf}\xi_f \sqrt{g_t^{f,i,N}} \hat{\sigma}(t, S_t^{i,N}) \sqrt{v_t^{i,N}} \right) dt + \xi_f \sqrt{g_t^{f,i,N}} dW_t^{f,i,N}, \\
dv_t^{i,N} &= \kappa(\theta - v_t^{i,N})dt + \xi \sqrt{v_t^{i,N}} dW_t^{v,i,N}, \\
dD_t^{d,i,N} &= -r_t^{d,i,N} D_t^{d,i,N} dt,
\end{aligned}$$

where $\{(W^{i,N}, W^{d,i,N}, W^{f,i,N}, W^{v,i,N})\}_{i=1}^N$ are N independent copies of 4-dimensional Brownian motion (W, W^v, W^d, W^f). We use the regularising kernel method in Chapter 2 to approximate the conditional expectation in the calibrated leverage function:

$$\begin{aligned}
\hat{\sigma}(t,x) &= \sqrt{\frac{\mathbb{E}^{\mathbb{Q}^{d,N}}[D_t^{d,N} | S_t = x]}{\mathbb{E}^{\mathbb{Q}^{d,N}}[D_t^d v_t | S_t = x]} \left(\sigma_{Dup}(t,x)^2 + \frac{\mathbb{E}^{\mathbb{Q}^{d,N}}[Q_t]}{\frac{1}{2}x^2 \frac{\partial^2 C}{\partial x^2}} \right)}, \\
&= \sqrt{\frac{1}{\hat{p}(t,x)} \left(\sigma_{Dup}(t,x)^2 + \frac{\frac{1}{N}\sum_{i=1}^N Q_t^{i,N}}{\frac{1}{2}x^2 \frac{\partial^2 C(t,x)}{\partial x^2}} \right)},
\end{aligned}$$

where

$$\begin{aligned}
\hat{p}(t,x) &= \frac{\sum_{i=1}^N D_t^{d,i,N} v_t^{i,N} \delta_{t,N}(S_t^{i,N} - x)}{\sum_{i=1}^N D_t^{d,i,N} \delta_{t,N}(S_t^{i,N} - x)}, \\
Q_t^{i,N} &= D_t^{d,i,N}(r_t^{f,i,N} - \bar{r}_t^{f,i,N})(S_t^{i,N} - x)^+ - x D_t^{d,i,N} \mathbb{1}_{S_t^{i,N} \geq x} [(r_t^{d,i,N} - \bar{r}_t^{d,i,N}) - (r_t^{f,i,N} - \bar{r}_t^{f,i,N})],
\end{aligned}$$

with some kernel functions $\delta_{t,N}(\cdot)$.

Algorithm 4 summarises the full algorithm in simulating the calibrated SIR-LSVMs. The inputs of the algorithm include: the number of particles N, the expiry T, the number of time steps M, the Dupire function σ_{Dup}, initial values S_0, v_0, r_0^d, r_0^f, parameters in the domestic short rate model ξ_d, θ_d, κ_d, $h^d(\cdot)$, parameters in the foreign short rate model ξ_f, θ_f, κ_f, $h^f(\cdot)$, parameters in the volatility process ξ, θ, κ, correlation parameters ρ, ρ_{Sd}, ρ_{Sf}, ρ_{df}, and the parameters in the regularising kernel method $\delta_{t,N}(\cdot)$, h, η.

Given a market with observable vanilla option prices and zero coupon bond prices, we compute the Dupire volatility function using the finite difference method as discussed in Chapter 2, and approximate the instantaneous forward rates similarly:

$$\bar{r}_t^i \approx -\frac{\log P^i(0,t+\Delta t) - \log P^i(0,t)}{\Delta t},$$

for $i \in \{d, f\}$ and some small time interval $\Delta t > 0$.

In the next section, we apply the algorithm in a synthetic Heston market with stochastic interest rates, and compare the implied volatility curves from calibrated SIR-LSVM to the market implied volatility curve.

6.2 Simulation of the calibrated SIR-LSVMs in the Heston market

We first generate a synthetic Heston market with stochastic interest rates:

$$\begin{aligned}
dS_t &= (r_t^d - r_t^f)S_t dt + \sqrt{v_t} S_t dW_t^M, \\
dr_t^d &= \kappa_d^M(\theta_d^M - r_t^d)dt + \xi_d^M \sqrt{r_t^d} dW_t^{d,M}, \\
dr_t^f &= \left(\kappa_f^M(\theta_f^M - r_t^f) - \rho_{Sf}^M \xi_f^M \sqrt{r_t^f}\sqrt{v_t}\right) dt + \xi_f^M \sqrt{r_t^f} dW_t^{f,M}, \\
dv_t &= \kappa^M(\theta^M - v_t)dt + \xi^M \sqrt{v_t} dW_t^{v,M},
\end{aligned}$$

with inital values $S_0 = 1$, $r_0^d = 0.0001$, $r_0^f = 0.0001$ and $v_0 = 0.1024$. $(W^M, W^{d,M}, W^{f,M}, W^{v,M})$ is a four-dimensional Brownian motion, with the correlation structure:

$$d\langle W_t^M, W_t^{v,M}\rangle = \rho^M dt,\ d\langle W_t^M, W_t^{d,M}\rangle = \rho_{Sd}^M dt,\ d\langle W_t^M, W_t^{f,M}\rangle = \rho_{Sf}^M dt,\ d\langle W_t^{d,M}, W_t^{f,M}\rangle = \rho_{df}^M dt,$$

where $\rho^M = -0.7$, $\rho_{Sd}^M = -0.3$, $\rho_{Sf}^M = 0.3$ and $\rho_{df}^M = 0.5$. The parameters with respect to the short rate processes and the volatility process are set to have $\kappa_d^M = 0.0837$, $\theta_d^M = 0.5469$, $\xi_d^M = 0.0274$, $\kappa_f^M = 0.011$, $\theta_f^M = 1.1656$, $\xi_f^M = 0.037$, $\kappa^M = 1.5768$, $\theta^M = 0.0484$, $\xi^M = 0.5751$.

We simulate the calibrated SIR-LSVM as in (6.1.2), assuming that the short rate processes are fully calibrated to the market. That is, we have $\kappa_d = \kappa_d^M$, $\theta_d = \theta_d^M$, $\xi_d = \xi_d^M$, $\kappa_f = \kappa_f^M$, $\theta_f = \theta_f^M$, $\xi_f = \xi_f^M$, and the deterministic functions $h^d(x) = h^f(x) = 0$ for all $x \in \mathbb{R}$.

Similar to the analysis done in the previous chapters, we consider two calibrated SIR-LSVMs. A simple SIR-LSVM has the same correlation structures and the parameter values in the volatility process as with the market, that is, $\rho = \rho^M$, $\rho_{Sd} = \rho_{Sd}^M$, $\rho_{Sf} = \rho_{Sf}^M$, $\rho_{df} = \rho_{df}^M$, and $\kappa = \kappa^M$, $\theta = \theta^M$, $\xi = \xi^M$. In contrast, a complex SIR-LSVM has a different volatility process and correlation structure to the market, where we set $\rho = 0$, $\rho_{Sd} = -0.3$, $\rho_{Sf} = 0.3$, $\rho_{df} = 0.5$, and $\kappa = 1$, $\theta = 0.0144$, $\xi = 0.5751$.

We use the result from Chapter 2 and choose the bandwidth h of the kernel functions to be $h = S_0 N^{-\frac{1}{5}}/10$ to achieve better calibration results. We set the number of particles $N = 10^5$, the number of time steps $M = 1000$ and $T = 1$.

Figure 6.1 plots the implied volatility curves from both calibrated SIR-LSVMs as well as the one from the synthetic market, and Table 6.1 summarises the absolute errors in the implied volatility curves. We see that both curves from the simulated SIR-LSVMs follow each other and the market volatility curve closely. Furthermore, the average absolute error in the implied volatility curves are comparable to the ones in the previous chapters. This suggests that the simulation scheme via the particle method is able to reproduce a decent implied volatility curve from the calibrated SIR-LSVMs, regardless whether the volatility process in the model differs from the market dynamic.

This concludes our discussion on the simulation of the calibrated SIR-LSVMs. The next chapter provides a brief summary and concludes the report.

Algorithm 4: Simulation of the calibrated SIR-LSVM with stochastic interest rates

Data: N, T, M, σ_{Dup}, S_0, v_0, r_0^d, r_0^f, ξ_d, θ_d, κ_d, $h^d(\cdot)$, ξ_f, θ_f, κ_f, $h^f(\cdot)$, ξ, θ, κ, ρ, ρ_{Sd}, ρ_{Sf}, ρ_{df}, $\delta_{t,N}(\cdot)$, h, η

$\Delta t \leftarrow T/M$;

for i *in* $1, \ldots, N$*:* **do**

$S_0^{i,N} \leftarrow S_0$;

$v_0^{i,N} \leftarrow v_0$;

$D_0^{i,N} \leftarrow 1$;

$g_0^{d,i,N} \leftarrow 0$;

$g_0^{f,i,N} \leftarrow 0$;

$k \leftarrow 1$;

$\hat{\sigma}(t,x) \leftarrow \sigma_{Dup}(0,x)/\sqrt{v_0}$ for $t \in [0, \Delta t]$;

while $k \leq M$ **do**

$t_k \leftarrow k\Delta t$;

for i *in* $1, \ldots, N$*:* **do**

$$(W^{i,N}, W^{d,i,N}, W^{f,i,N}, W^{v,i,N}) \overset{\text{i.i.d.}}{\sim} N\left((0,0,0,0), \begin{bmatrix} 1 & \rho_{Sd} & \rho_{Sf} & \rho \\ \rho_{Sd} & 1 & \rho_{df} & 0 \\ \rho_{Sf} & \rho_{df} & 1 & 0 \\ \rho & 0 & 0 & 1 \end{bmatrix}\right);$$

$$S_{t_k}^{i,N} \leftarrow S_{t_{k-1}}^{i,N} + S_{t_{k-1}}^{i,N}(r_{t_{k-1}}^{d,i,N} - r_{t_{k-1}}^{f,i,N})\Delta t + \sqrt{v_{t_{k-1}}^{i,N}} S_{t_{k-1}}^{i,N} \hat{\sigma}(t_{k-1}, S_{t_{k-1}}^{i,N}) \sqrt{\Delta t} W^{i,N};$$

$$D_{t_k}^{i,N} \leftarrow D_{t_{k-1}}^{i,N} - r_{t_{k-1}}^{d,i,N} D_{t_{k-1}}^{i,N} \Delta t;$$

$$g_{t_k}^{d,i,N} \leftarrow g_{t_{k-1}}^{d,i,N} + \kappa_d(\theta_d - g_{t_{k-1}}^{d,i,N})\Delta t + \xi_d \sqrt{g_{t_{k-1}}^{d,i,N}} \sqrt{\Delta t} W^{d,i,N};$$

$$r_{t_k}^{d,i,N} \leftarrow g_{t_k}^{d,i,N} + h^d(t_k);$$

$$g_{t_k}^{f,i,N} \leftarrow g_{t_{k-1}}^{f,i,N} + \left(\kappa^f(\theta_f - g_{t_{k-1}}^{f,i,N}) - \rho_{Sf}\xi_f \sqrt{g_{t_{k-1}}^{f,i,N}} \hat{\sigma}(t_{k-1}, S_{t_{k-1}}^{i,N}) \sqrt{v_{t_{k-1}}^{i,N}}\right)\Delta t + \xi_f \sqrt{g_{t_{k-1}}^{f,i,N}} \sqrt{\Delta t} W^{f,i,N};$$

$$r_{t_k}^{f,i,N} \leftarrow g_{t_k}^{f,i,N} + h^f(t_k);$$

$$v_{t_k}^{i,N} \leftarrow v_{t_{k-1}}^{i,N} + \kappa(\theta - v_{t_{k-1}}^{i,N})\Delta t + \xi \sqrt{v_{t_{k-1}}^{i,N}} \sqrt{\Delta t} W^{v,i,N};$$

$$Q^{i,N}(t_k, x) \leftarrow D_{t_k}^{d,i,N}(r_{t_k}^{f,i,N} - \bar{r}_{t_k}^{f,i,N})(S_{t_k}^{i,N} - x)^+ - x D_{t_k}^{d,i,N} \mathbf{1}_{S_{t_k}^{i,N} \geq x}[(r_{t_k}^{d,i,N} - \bar{r}_{t_k}^{d,i,N}) - (r_{t_k}^{f,i,N} - \bar{r}_{t_k}^{f,i,N})];$$

$G \leftarrow$ equal-spaced grid values covering the set $\{S_{t_k}^{i,N}\}_{i=1}^N$, with $|G| = \max(N_1\sqrt{t_k}, N_2)$, $N_1 = 30$, $N_2 = 15$;

for x *in* G **do**

$$M_x \leftarrow \{i \in \{1, \ldots, N\} | \delta_{t_k,N}(S_{t_k}^{i,N} - x) > \eta\};$$

$$\hat{p}(t_k, x) \leftarrow \frac{\sum_{i \in M_x} D_{t_k}^{i,N} v_{t_k}^{i,N} \delta_{t_k,N}(S_{t_k}^{i,N} - x)}{\sum_{i \in M_x} D_{t_k}^{i,N} \delta_{t_k,N}(S_{t_k}^{i,N} - x)};$$

$$\hat{\sigma}(t_k, x) \leftarrow \sqrt{\frac{\sigma_{Dup}^2(t_k,x)}{\hat{p}(t_k,x)} + \frac{\frac{1}{N}\sum_{i=1}^N Q^{i,N}(t_k,x)}{\frac{1}{2}\hat{p}(t_k,x)x^2 \frac{\partial^2 C(t_k,x)}{\partial x^2}}};$$

for i *in* $1, \ldots, N$*:* **do**

$\hat{\sigma}(t_k, S_{t_k}^{i,N}) \leftarrow$ cubic interpolation;

$\hat{\sigma}(t, S_{t_k}^{i,N}) \leftarrow \hat{\sigma}(t_k, S_{t_k}^{i,N})$ for $t \in [t_k, t_{k+1}]$;

$k \leftarrow k + 1$;

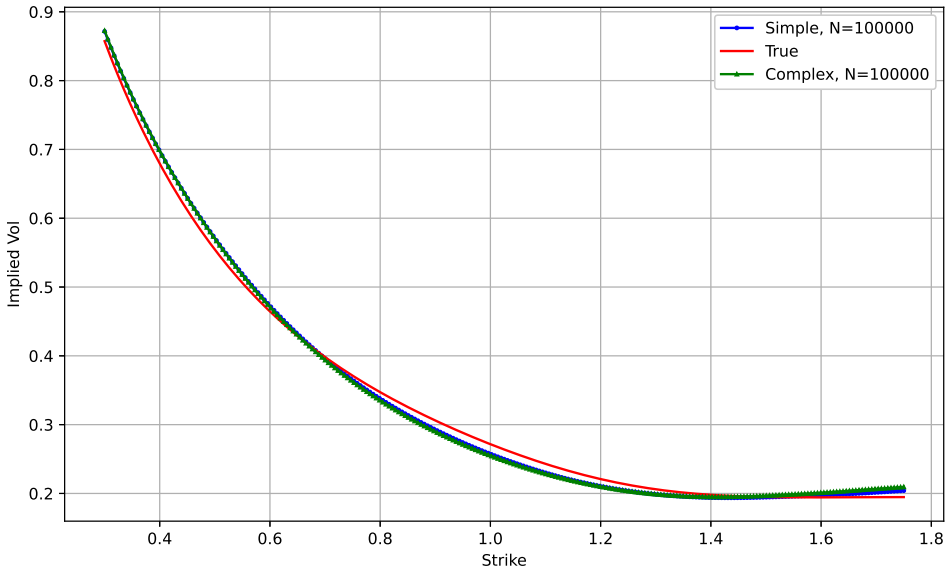

Figure 6.1: Market volatility level and implied volatility curves from the calibrated SIR-LSVMs at $T = 1$ in the Heston market with stochastic interest rates. The simulation is based on the regularising kernel method, with $N = 10^5$ and $M = 1000$.

Absolute error $\lvert\sigma_{Impl}^{True} - \sigma_{Impl}^{LSVM}\rvert$ in %				
Strike	0.7	1	1.3	Ave. absolute error
Simple LSVM	0.2093	1.4227	0.7408	0.9328
Complex LSVM	0.4223	1.6696	0.7256	1.0600

Table 6.1: Absolute error in % of the implied volatility curves in the Heston market with stochastic interest rates.

Conclusion

We introduced and studied the local stochastic volatility model, a state-of-the-art modelling tool in option pricing which allows an exact calibration to the vanilla options in the market. We briefly discussed the challenge in establishing the well-posedness of the calibrated LSVMs, as well as the difficulty in the numerical simulations due to the presence of the conditional expectation $\mathbb{E}[v_t|S_t]$, which is introduced by the calibration condition. Since the calibrated LSVMs can be seen as a special case of the McKean-Vlasov SDEs, we introduced the particle simulation scheme and focused on three different methods to approximate $\mathbb{E}[v_t|S_t]$, which are the regularsing kernel method, the RKHS method and the bin Monte Carlo method.

From our numerical experiments, we see that the calibrated LSVMs are able to reproduce the market volatility curves using the three methods in most scenarios, with some exceptions in the RKHS method. We further provided a discussion on the comparison among the three methods, including the computational efficiency and accuracy.

We conclude the report by discussing the LSVMs with stochastic interest rates (SIR-LSVM) in the FX market, and provided a numerical example to simulate the calibrated SIR-LSVMs in the Heston market with CIR short rates.

In conclusion, despite the challenge on the well-posedness of the calibrated LSVMs, several numerical methods are able to provide feasible ways to simulate the calibrated LSVMs and reproduce the implied volatility curves. This enables the LSVMs to be an effective and practical tool in day-to-day option pricing activities.

Appendix A

Technical Definitions

Both definitions below are referenced from [15, Chapter 2.2] and used in Chapter 1.

A.1 Integrability on the set of Borel probability measures

Definition A.1.1. Let $(\mathcal{X}, d)$ be a metric space, and $\mathcal{P}(\mathcal{X})$ be a set of Borel probability measures on $\mathcal{X}$. For $p \geq 1$, define $\mathcal{P}^p(\mathcal{X})$ to be the set of probability measures $\mu \in \mathcal{P}(\mathcal{X})$ satisfying

$$\int_{\mathcal{X}} d(x, x_0)^p \mu(dx) < \infty,$$

where $x_0 \in \mathcal{X}$ is an arbitrary reference point.

A.2 p-Wasserstein metric

Definition A.2.1. The p-Wasserstein metric on $\mathcal{P}^p(\mathcal{X})$ is defined by

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{X \sim \mu, Y \sim \nu} \mathbb{E}[d(X, Y)^p] \right)^{1/p},$$

where $\mu, \nu \in \mathcal{P}(\mathcal{X})$, and the infimum is over all pairs of $\mathcal{X}$-valued random variables with given marginals μ and ν.

Bibliography

[1] Julien Guyon and Pierre Henry-Labordere. *Nonlinear option pricing*. CRC Press, 2013.

[2] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

[3] Mark Rubinstein. Displaced diffusion option pricing. *The Journal of Finance*, 38(1):213–217, 1983.

[4] Bruno Dupire et al. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[5] Steven 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):327–343, 1993.

[6] Elias M Stein and Jeremy C Stein. Stock price distributions with stochastic volatility: an analytic approach. *The review of financial studies*, 4(4):727–752, 1991.

[7] Patrick S Hagan, Deep Kumar, Andrew S Lesniewski, and Diana E Woodward. Managing smile risk. *The Best of Wilmott*, 1:249–296, 2002.

[8] Eric Renault and Nizar Touzi. Option hedging and implied volatilities in a stochastic volatility model 1. *Mathematical Finance*, 6(3):279–302, 1996.

[9] Christoph Reisinger and Maria Olympia Tsianni. Convergence of the Euler–Maruyama particle scheme for a regularised McKean–Vlasov equation arising from the calibration of local-stochastic volatility models. *arXiv preprint arXiv:2302.00434*, 2023.

[10] Bruno Dupire. A unified theory of volatility. *Derivatives pricing: The classic collection*, pages 185–196, 1996.

[11] Christian Bayer, Denis Belomestny, Oleg Butkovsky, and John Schoenmakers. RKHS regularization of singular local stochastic volatility McKean-Vlasov models. *arXiv preprint arXiv:2203.01160*, 2022.

[12] Anthonie W Van der Stoep, Lech A Grzelak, and Cornelis W Oosterlee. The Heston stochastic-local volatility model: Efficient Monte Carlo simulation. *International Journal of Theoretical and Applied Finance*, 17(07):1450045, 2014.

[13] Henry P McKean Jr. A class of markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences*, 56(6):1907–1911, 1966.

[14] Alain-Sol Sznitman. Topics in propagation of chaos. *Lecture notes in mathematics*, pages 165–251, 1991.

[15] Daniel Lacker. Mean field games and interacting particle systems. *preprint*, 2018.

[16] Daniel Lacker, Mykhaylo Shkolnikov, and Jiacheng Zhang. Inverting the Markovian projection, with an application to local stochastic volatility models. 2020.

[17] Mao Fabrice Djete. Non–regular McKean–Vlasov equations and calibration problem in local stochastic volatility models. *arXiv preprint arXiv:2208.09986*, 2022.

[18] Roger Lord, Remmert Koekkoek, and Dick Van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2):177–194, 2010.

[19] Andrei Cozma and Christoph Reisinger. Strong order 1/2 convergence of full truncation Euler approximations to the Cox–Ingersoll–Ross process. *IMA journal of numerical analysis*, 40(1):358–376, 2020.

[20] Bernhard Schölkopf, Ralf Herbrich, and Alex J Smola. A generalized representer theorem. In *International conference on computational learning theory*, pages 416–426. Springer, 2001.

[21] Andrei Cozma, Matthieu Mariapragassam, and Christoph Reisinger. Calibration of a hybrid local-stochastic volatility stochastic rates model with a control variate particle method. *SIAM Journal on Financial Mathematics*, 10(1):181–213, 2019.