VILNIUS UNIVERSITY

ANTANAS LENKŠAS

WEAK APPROXIMATIONS OF HESTON MODEL BY DISCRETE RANDOM VARIABLES

Doctoral dissertation

Physical sciences, mathematics (01P)

Vilnius, 2016

Doctoral dissertation was written in 2013–2016 at Vilnius University and will be defended externally.

Scientific adviser:

prof. habil. dr. Vigirdas Mackevičius
Vilnius University, Physical sciences, Mathematics – 01P

VILNIAUS UNIVERSITETAS

ANTANAS LENKŠAS

HESTONO MODELIO SILPNOSIOS APROKSIMACIJOS DISKREČIAIS ATSITIKTINIAIS DYDŽIAIS

Daktaro disertacija

Fiziniai mokslai, matematika (01P)

Vilnius, 2016

Disertacija rengta 2013–2016 metais ir ginama eksternu.

Mokslinis vadovas:

prof. habil. dr. Vigirdas Mackevičius

Vilniaus universitetas, fiziniai mokslai, matematika – 01P

Contents

List of Figures

List of Tables

Notation and Abbreviations

Notations	Descriptions
$\mathbb{N}$	the set of positive integers $\{1, 2, \ldots\}$.
$\mathbb{N}_0$	the set of nonnegative integers, $\mathbb{N} \bigcup \{0\}$.
$\mathbb{R}$	the set of real numbers $(-\infty, \infty)$.
$\mathbb{Z}$	the set of integers.
$\mathbb{C}$	the set of complex numbers.
$\mathbb{D}$	the domain of the solution of the two-dimensional SDE. In the context of the Heston model, we take $\mathbb{D} = (0, +\infty) \times [0, +\infty)$, and, in the context of the log-Heston model, we take $\mathbb{D} = \mathbb{R} \times [0, +\infty)$.
$C^{\infty}(\mathbb{D})$	the set of infinitely differentiable functions $f : \mathbb{D} \to \mathbb{R}$.
$C_0^{\infty}(\mathbb{D})$	the set of functions $f : \mathbb{D} \to \mathbb{R}$ of class C^{∞} with compact support.
$C_{pol}^{\infty}(\mathbb{D})$	the set of functions $f : \mathbb{D} \to \mathbb{R}$ of class C^{∞} with all partial derivatives of polynomial growth.
$\mathbb{E}X$	the mean of a random variable X.
$\lvert z \rvert$	$:= \sqrt{z_1^2 + z_2^2 + \cdots + z_n^2}$. The Euclidean norm of $z = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n$.
$N(a, \sigma^2)$	normal distribution with mean a and variance σ^2.
$\mathrm{O}_{p,k}(h^n)$	a function of kth-degree polynomial growth with respect to h^n, i.e., $g(z, h) = \mathrm{O}_{p,k}(h^n)$ if for some $C > 0$ and $h_0 > 0$, $\lvert g(z, h) \rvert \leq C h^n (1 + \lvert z \rvert^k), \quad z \in \mathbb{D},\ h \leq h_0$.

Notations	Descriptions
$\mathrm{O}_p(h^n)$	a function of polynomial growth with respect to h^n, i.e., $g(z,h) = \mathrm{O}_p(h^n)$ if $g(z,h) = \mathrm{O}_{p,k}(h^n)$ for some $k \in \mathbb{N}$.
$\mathcal{O}_p(h^n)$	a function of polynomial growth with respect to h^n when the function g is expressed in terms of another function $f \in C^{\infty}_{pol}(\mathbb{D})$ and the constants C, h_0, and k depend on a good sequence for f only.
i.i.d.	independent identically distributed.
r.v.	random variable.
$[x]$	the integer part of a number x.
x^+	$:= \max(x,0)$ for $x \in \mathbb{R}$ (the positive part of x).
Δ^h	(equidistant) time interval discretization.
∇_k	kth-order term in Taylor's formula.
r_k	kth-order remainder in Taylor's formula.
$B_t, W_t, \widetilde{W}_t$	A standard Brownian motion (Wiener process).
S_t	Asset (interest rate, stock, etc.) price in the Heston model.
X_t	$\log S_t$ asset price in the log-Heston model.
Y_t	Stochastic volatility in the Heston model (CIR equation).
Z_t	$(X_t, Y_t)^T$.
$\hat{X}, \tilde{X}, \overline{X}$	discretizations of X.
$\frac{\partial^{\lvert i \rvert} f(z)}{\partial z^i}$	partial derivatives of two-variable function f. Here $\mathbb{N}_0 = \{0,1,2,\ldots\}$, $i = (i_1, i_2) \in \mathbb{N}_0^2$ are multiindices. and $\lvert i \rvert := i_1 + i_2$, $\partial z^i := \partial z_1^{i_1} \partial z_2^{i_2}$.
$f^{(i)}(z)$	$\frac{\partial^{\lvert i \rvert} f(z)}{\partial z^i}$.
L^k	kth power of the generator of the solution of a stochastic differential equation.
$P_n, Q_n, \hat{P}_n, \hat{Q}_n$	nth-order polynomial.
$C, \hat{C}, \bar{C}, \tilde{C}$	constants, which may vary from line to line.
SDE	Stochastic differential equation.
CIR	Cox–Ingersoll–Ross model.
DVSS	Discrete-variable split-step first-order approximation of the solution of the Heston scheme.

Notations	Descriptions
$DVSS_2$	Discrete-variable split-step second-order approximation of the solution of the Heston scheme.
QE	Andersen's quadratic exponential approximation scheme for the Heston model.
QE-M	Andersen's quadratic exponential approximation scheme with martingale correction.
ALF_2	Second-order approximation scheme for the Heston model, proposed by Alfonsi.
ALF_3	Third-order approximation scheme for the Heston model, proposed by Alfonsi.

Chapter 1

Introduction

In this chapter, we present our research topic, aims and methods used, novelty of main results, list papers, and conferences where our main results were introduced.

1.1 Research topic

While the Black–Scholes (Merton) formula is often considered as one of the most important results in Mathematical Finance, its flaws and biases are also long recognized, and many attempts of generalization and improvement were made in subsequent years. Despite its fame and firm place in textbooks on the subject, nowadays practitioners use it only for simplicity, if use it at all. The assumptions of constant interest rate and constant volatility are the two of the key disadvantages of the Black–Scholes model, especially when it is used over a long-term period. Not surprisingly, much effort was directed to avoid these somewhat unrealistic assumptions. To counteract the first disadvantage, interest rate models where considered in the articles of Vašíček [34] and Cox, Ingersoll, and Ross [10]. The latter resulted in the now famous square-root process (or CIR as it was later named after its authors). Attempts to eliminate the second disadvantage gave birth to the idea of a second source of randomness and led to models allowing stochastic (random) volatility in the papers of Scott [32], Hull and White [16], Stein and Stein [33], and others. A similar approach was taken by Heston [14] in 1993. He offered a model driven by two sources of randomness, where a spot asset price, driven by one source of randomness, was correlated with stochastic volatility, driven by another source of randomness. The model almost instantly became popular, and to this day, more than 20 years after its introduction, remains a number one choice amongst market practitioners and Mathematical Finance theoreticians alike.

In this work, we consider the solution of the stochastic volatility model proposed by Heston [14]:

$$\begin{cases} \mathrm{d}S_t = rS_t\,\mathrm{d}t + \sqrt{Y_t}S_t\,\mathrm{d}\widetilde{W}_t, \quad S_0 = s \geqslant 0, \\ \mathrm{d}Y_t = k(\theta - Y_t)\,\mathrm{d}t + \sigma\sqrt{Y_t}\,\mathrm{d}W_t, \quad Y_0 = y \geqslant 0, \\ \mathrm{d}W_t\,\mathrm{d}\widetilde{W}_t = \rho\,\mathrm{d}t. \end{cases} \tag{1.1.1}$$

Here W and $\widetilde{W}$ are (possibly, dependent) standard Brownian motions, with parameters $\theta, \sigma, k > 0$. Note, that the volatility is controlled by the stochastic square-root process CIR.

Development of stochastic models in Mathematical Finance is inseparable from extensive use of numerical computations as a closed-form solution of the model is usually, except in simple and trivial cases, unknown. Almost the same applies to the Heston model, as the distribution of the Heston process, S in Eq. (1.1.1), is explicitly known only in the form of the characteristic function of $X_t = \log S_t$. A version of such characteristic function was already presented in the original paper by Heston [14] (see formula (2.1.1), where it is used), and various (equivalent) characteristic functions were derived in subsequent studies (e.g. del Baño Rollin et al. [29], Albrecher et al. [1]). For example, del Baño Rollin et al. [29] arrive at the following formula:

$$\begin{aligned} \mathbb{E}(\mathsf{e}^{iuX_t}) = \exp(ixu) & \left(\frac{\mathsf{e}^{kt/2}}{\cosh(dt/2) + \xi \sinh(dt/2)/d} \right)^{2k\theta/\sigma^2} \\ & \times \exp\left(-y \frac{(iu + u^2)\sinh(dt/2)/d}{\cosh(dt/2) + \xi \sinh(dt/2)/d} \right), \end{aligned}$$

where

$$\begin{aligned} x &= \log s, \\ d &= d(u) = \sqrt{(k - \sigma\rho iu)^2 + \sigma^2(iu + u^2)}, \\ \xi &= \xi(u) = k - \sigma\rho ui. \end{aligned}$$

1.2 Aim and problems

The aim of the thesis is to construct "simple" yet "effective" first- and second-order weak approximation schemes for the solution of the Heston model that use, at each step, only generation of one and two discrete random variables, respectively, and to provide rigorous proofs of their accuracy.

As in [22], [21], [20] and other papers (e.g., Andersen [5], Lord et al. [23], and Kloeden and Neuenkirch [18]), to avoid the positivity-preservation problem, we approximate, instead of S_t in Eq. (1.1.1), the logarithm $X_t := \log S_t$ and thus, using Itô's lemma, arrive at the log-Heston model

$$\begin{cases} \mathrm{d}X_t = (r - \frac{1}{2}Y_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}\widetilde{W}_t, \quad X_0 = x := \log(s), \\ \mathrm{d}Y_t = k(\theta - Y_t)\,\mathrm{d}t + \sigma\sqrt{Y_t}\,\mathrm{d}W_t, \quad Y_0 = y \geqslant 0, \\ \mathrm{d}W_t\,\mathrm{d}\widetilde{W}_t = \rho\,\mathrm{d}t. \end{cases} \tag{1.2.1}$$

This also allows us to avoid problems that arise from the fact that moments of the solutions of the Heston model "explode," that is, tend to infinity in finite time (see, e.g., [6]).

1.3 Methods

Methods of calculus, general probability theory, stochastic calculus, statistics, and functional analysis are used in the thesis. Numerical experiments were performed using a number of programming languages, including C programming language and a free software environment for statistical computing and graphics R.

1.4 Actuality and novelty

Today the most popular discretization scheme for the solution of the Heston model is the Andersen's quadratic-exponential algorithm (QE), which, although demonstrates good numerical simulation results in option pricing, has no rigorous proof regarding its accuracy. The same concerns the Alfonsi schemes for the Heston model. In the thesis, we propose schemes that can compete with Andersen's and Alfonsi's and provide a rigorous proof of (strongly potential) weak convergence for them. We construct our schemes using only simple (two- or three-valued) discrete random variables.

The same technique used in this work might be useful constructing discretization schemes for other Mathematical Finance models, for example, CEV-SV or CKLS.

1.5 Main results

We succeeded in constructing simple yet effective first- and second-order weak approximations for the solution of the Heston model that use, at each step, only generation of discrete random variables. They are presented in the following theorems.[1]

Theorem 1.1. Let a one-step approximation $\bar{Z}_h^z = (\bar{X}_h^z, \bar{Y}_h^y)$ of Eq. (1.2.1) be constructed as follows:

(1) Let $\hat{Y}_h^y$ be a random variable taking the values

$$y_{1,2} = y+\sigma^2 h \pm \sqrt{(y+\sigma^2 h)\,\sigma^2 h} \quad \text{with corresponding probabilities } p_{1,2} = \frac{y}{2y_{1,2}}.$$

(2) Let $\tilde{X}_h^z$ be a random variable, independent of $\hat{Y}_h^y$, taking the values

$$x_{1,2} = x \pm \sqrt{yh} \quad \text{with probabilities } \frac{1}{2}.$$

(3) Let a random variable $\hat{X}_h^z$ be defined by

$$\hat{X}_h^z := x + \sqrt{1-\rho^2}(\tilde{X}_h^z - x) + \frac{\rho}{\sigma}(\hat{Y}_h^y - y).$$

(4) Finally, having $\hat{Z}_h^z = (\hat{X}_h^z, \hat{Y}_h^y)$, define the one-step discretization scheme $\bar{Z}_h^z = (\bar{X}_h^z, \bar{Y}_h^y)$ by

$$\bar{Z}_h^z = D\left[\hat{Z}(z,h), h\right], \tag{1.5.1}$$

where $D(z,t) = (D_1^z(t), D_2^y(t))$ is

$$\begin{cases} D_1^z(t) = x + (r - \frac{1}{2}\theta)t + \frac{1}{2k}\left(\mathsf{e}^{-kt} - 1\right)(y-\theta), \\ D_2^y(t) = y\mathsf{e}^{-kt} + \theta(1-\mathsf{e}^{-kt}), \end{cases}$$

[1] For definitions, see Section 3.

i.e., the solution of the so-called deterministic part of Eq. (1.2.1)

$$\begin{cases} \mathrm{d}D_1^z(t) = (r - \frac{1}{2}D_2^y(t))\,dt, & D_1^z(0) = x, \\ \mathrm{d}D_2^y(t) = k(\theta - D_2^y(t))\,\mathrm{d}t, & D_2^y(0) = y. \end{cases}$$

Then the one-step discretization scheme (1.5.1) defines a strongly potential weak first-order approximation of the log-Heston system (1.2.1).

Theorem 1.2. Let a one-step approximation $\bar{Z}_h^z = (\bar{X}_h^z, \bar{Y}_h^y)$ of Eq. (1.2.1) be constructed as follows:

(1) Let $\hat{Y}_h^y$ be a discretization scheme satisfying the following conditions:

(i)

$$\begin{cases} \mathbb{E}(\hat{Y}_h^y - y) = \mathrm{O}_{p,1}(h^3), \\ \mathbb{E}(\hat{Y}_h^y - y)^2 = yh + \mathrm{O}_{p,2}(h^3), \\ \mathbb{E}(\hat{Y}_h^y - y)^3 = \frac{3}{2}yh^2 + \mathrm{O}_{p,3}(h^3), \\ \mathbb{E}(\hat{Y}_h^y - y)^4 = 3y^2h^2 + \mathrm{O}_{p,4}(h^3); \end{cases}$$

(ii) $\mathbb{E}(\hat{Y}_h^y - y)^5 = \mathrm{O}_{p,5}(h^3)$;

(iii) $\mathbb{E}|\hat{Y}_h^y - y|^{2q} = \mathrm{O}_{p,2q}(h^q)$ for all $q \geq 3$.

(2) Let ξ be a discrete random variable, independent of $\hat{Y}_h^y$, with first five moments matching those of a standard random variable. Let, finally, the random variable $\hat{X}_h^z$ be defined by

$$\hat{X}_h^z := x + \xi\sqrt{\frac{1}{2}(y + \hat{Y}_h^y)h}\ .$$

(3) Define the one-step discretization scheme $\tilde{Z}_h^z = (\tilde{X}_h^z, \tilde{Y}_h^y)$ by

$$\tilde{X}_h^z = \sigma\left(\sqrt{1-\rho^2}\,\hat{X}_h^z + \rho\hat{Y}_h^y\right),$$
$$\tilde{Y}_h^y = \sigma^2\tilde{Y}_h^y.$$

(4) Finally, define the one-step discretization scheme $\bar{Z}_h^z = (\bar{X}_h^z, \bar{Y}_h^y)$ by

$$\bar{Z}_h^z = D\big(\tilde{Z}_h^{D(z,h/2)}, h/2\big), \tag{1.5.2}$$

where $D(z,t) = (D_1^z(t), D_2^y(t))$ is the same as in Theorem 1.1 (the solution of the so-called deterministic part of Eq. (1.2.1)).

Then the one-step discretization scheme (1.5.2) defines a strongly potential weak second-order approximation of the log-Heston system (1.2.1).

1.6 Publications

1. A. Lenkšas, V. Mackevičius. *A second-order weak approximation of Heston model by discrete random variables.* Lithuanian Mathematical Journal **55** (2015), 555–572.

2. A. Lenkšas, V. Mackevičius. *Weak approximation of Heston model by discrete random variables.* Mathematics and Computers in Simulation **113** (2015), 1–15.

3. A. Lenkšas, V. Mackevičius. *Option pricing in Heston model by means of weak approximations.* Lietuvos matematikos rinkinys **54** (2013), 27–32.

1.7 Conferences

The results of the thesis were presented in the following conferences:

1. 56th Conference of Lithuanian Mathematical Society, Kaunas, Lithuania, 2015 06 16–17.

2. 11th International Vilnius Conference on Probability Theory and Mathematical Statistics, Vilnius, Lithuania, 2014 06 30–07 04.

3. 55th Conference of Lithuanian Mathematical Society, Vilnius, Lithuania, 2014 06 26–27.

4. 54th Conference of Lithuanian Mathematical Society, Vilnius, Lithuania, 2013 06 19–20.

5. 53rd Conference of Lithuanian Mathematical Society, Klaipėda, Lithuania, 2012 06 11–12.

1.8 Structure of the thesis

The thesis is organized as follows. In Chapter 2, we give an overview of the main results obtained by other authors. Then after some preliminaries and definitions in Chapter 3 Section 3.2, we “split” the approximation problem for the process (X, Y) in Eq. (1.2.1) into the exact solution of the deterministic part and the approximation problem for the stochastic part of the system. In Section 4.2, we construct a potential first-order weak approximation for the stochastic part, and in Section 5.2, we construct a potential second-order weak approximation for the stochastic part. We summarize the constructed algorithms in Sections 4.3 and 5.3, respectively. In Sections 4.4 and 5.4, we illustrate the first-order scheme (DVSS) and the second-order scheme (DVSS_2), respectively, by numerical simulation results, including option pricing and a detailed comparison with the schemes of Andersen [5] and Alfonsi [3, 4].

We finalize the results of the thesis in Chapter 6, and in the Appendix (Chapter 7), we provide additional calculations, which we think would only distract the reader if placed elsewhere in the text.

1.9 Acknowledgments

I would like to take an opportunity to thank here my scientific adviser prof. Vigirdas Mackevičius for his valuable advices and infinite patience. I would also like to thank my colleagues from the Department of Mathematical Analysis for their constant support and encouragement.

Finally, I wish to express my sincere and deep gratitude to my family and friends for inspiring, caring, and love.

Chapter **2**

Historical overview

In this chapter, we discuss the importance of the Heston model and various attempts to construct discretization schemes for the solution of the model.

2.1 Why Heston?

In the Black–Scholes (Merton) model, one of the most famous models in the field of Mathematical Finance, and in many other models as well, the interest rate is assumed to be constant. As this is usually not the case in the real world, soon a new (now classical) approach was taken. In this new approach, an interest rate r_t was modeled by a stochastic process of the form

$$\mathrm{d}r_t = \alpha(t, r_t)\,\mathrm{d}t + \sigma(t, r_t)\,\mathrm{d}B_t,$$

where B_t is a Brownian motion (Wiener process).

One of the first and simplest models of that kind was presented in 1977 by Oldrich Vašíček (see [34]):

$$\mathrm{d}r_t = k(\theta - r_t)\,\mathrm{d}t + \sigma\,\mathrm{d}B_t,$$

where $k > 0$, $\theta > 0$ and $\sigma > 0$.

The greatest advantage of the Vašíček model is that its solution is explicitly known since the model is based on the well-known and relatively simple Ornstein–Uhlenbeck process. Unfortunately, the model can also obtain undesirable negative values (albeit probability of such an event is quite small).

In 1985, John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross presented another, square-root model (now known as CIR), designed to avoid shortcomings of the Vašíček model:

$$\mathrm{d}r_t = (\theta - kr_t)\,\mathrm{d}t + \sigma\sqrt{r_t}\,\mathrm{d}B_t,$$

where, as in the Vašíček model, $k > 0$, $\theta > 0$, and $\sigma > 0$.

In this model, the process never becomes negative if started from a nonnegative value. Besides, it retains the main advantage of the Vašíček model that it is mean-reverting, that is, $\mathbb{E}r_t \to \theta/k$ as $t \to \infty$.

Finally, in 1993, Heston [14] came up with an idea to extend the CIR model to a stock price model by introducing a second source of randomness and assuming that not only underlying asset but its volatility also is driven by the CIR process:

$$\begin{cases} \mathrm{d}S_t = rS_t\,\mathrm{d}t + \sqrt{Y_t}S_t\,\mathrm{d}\widetilde{W}_t, \quad S_0 = s \geqslant 0, \\ \mathrm{d}Y_t = k(\theta - Y_t)\,\mathrm{d}t + \sigma\sqrt{Y_t}\,\mathrm{d}W_t, \quad Y_0 = y \geqslant 0, \\ \mathrm{d}W_t\,\mathrm{d}\widetilde{W}_t = \rho\,\mathrm{d}t. \end{cases}$$

Here W and $\widetilde{W}$ are (possibly, dependent) standard Brownian motions, with parameters $\theta, \sigma, k > 0$. The model was an instant success, mostly because, in this model, European vanilla call (and put) option price is known in the (quasi-)closed form. Denoting the call (put) option price by C (P), i.e., $C = C(s, y, K, T) = \mathrm{e}^{-rT}\,\mathbb{E}(S_T - K)^+$ and $P = P(s, y, K, T) = \mathrm{e}^{-rT}\,\mathbb{E}(K - S_T)^+$, where K is the strike price, from [14] (see also [31]) we know that

$$C(s, y, K, T) = sP_1 - K\mathrm{e}^{-rT}P_2, \tag{2.1.1}$$

where

$$\begin{aligned} P_j(x, y, K, T) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\mathrm{e}^{-\mathrm{i}\phi\log K} f_j(x, y, T, \phi)}{\mathrm{i}\phi}\right)\mathrm{d}\phi, \\ x &= \log s, \\ f_j(x, y, T, \phi) &= \exp\{A_j(T, \phi) + B_j(T, \phi)y + \mathrm{i}\phi x\}, \\ A_j(T, \phi) &= \mathrm{i}r\phi T + \frac{k\theta}{\sigma^2}\left[(b_j - \mathrm{i}\rho\sigma\phi + d)T - 2\log\left(\frac{1 - g_j\mathrm{e}^{d_jT}}{1 - g_j}\right)\right], \\ B_j(T, \phi) &= \frac{b_j - \rho\sigma\phi\mathrm{i} + d_j}{\sigma^2}\left(\frac{1 - \mathrm{e}^{d_jT}}{1 - g_j\mathrm{e}^{d_jT}}\right), \\ g_j &= \frac{b_j - \mathrm{i}\rho\sigma\phi + d_j}{b_j - \mathrm{i}\rho\sigma\phi - d_j}, \end{aligned}$$

$$d_j = \sqrt{(\mathrm{i}\rho\sigma\phi - b_j)^2 - \sigma^2(2\mathrm{i}u_j\phi - \phi^2)},$$

for $j = 1, 2$ and $u_1 = \frac{1}{2}$, $u_2 = -\frac{1}{2}$, $b_1 = k - \rho\sigma$, $b_2 = k$.

The price of the put option can then be calculated using the so-called put–call parity formula

$$P = C - s + K\mathrm{e}^{-rT}. \tag{2.1.2}$$

On the other hand, applications of numerical methods for the Heston model have their own difficulties. Volatility in the Heston model is controlled by the CIR process and in the case of the CIR model, which is based on the square-root diffusion, the main problem is... the square root itself. Any attempt to develop an approximation scheme for the solution of the Heston model runs into the same problem, or, rather, problems as the Heston model has square roots used not only in the CIR process controlling the volatility, but also in the expression of volatility itself. Furthermore, the square roots present not one, but two technical challenges. First, since square-root derivatives near zero are unbounded, any approximation scheme explicitly or implicitly using derivatives of the model coefficients struggles with accuracy near zero. Second is the simple fact that a square root can take only non-negative values. Thus, all classical approximation methods in this case do not provide any useful solutions at all since they either use derivatives (Milstein), or cannot ensure the positivity (Euler), or both.

Maybe that was the main reason why more than 10 years had to pass while papers dealing with discretization of the solution of the Heston model started to appear despite the instant popularity of the Heston model. One of the first worth mentioning was published in 2005 by Higham and Mao [15], who proposed a modified Euler algorithm with a new fix and proved the strong convergence of such modified Euler scheme. To the best of our knowledge, it was the only known example of the rigorous proof of convergence of the Euler scheme in the Heston model setting till 2010 when Lord, Koekkoek, and van Dijk, extending the ideas of Higham and Mao, did the same for other modifications of the Euler scheme for the Heston model (see [23]). In 2006, Broadie and Kaya [7] proposed a so-called exact simulation of the Heston model. Broadie and Kaya devised their algorithm drawing the inspiration from the properties of the variance process in the Heston model. Their result behaves quite well in terms of the mean-square error but is heavy in the sense of time-consuming and can barely be recommended to use in most practical situations. An interesting approach to solve the disadvantage of the exact scheme by focusing on efficient approximation of the scheme was taken by Haastrecht et al. [13] in 2010. Finally, in 2008, Andersen [5] presented an algorithm constructed using a moment-matching technique and exploiting the same property of the variance process as Broadie and Kaya. The scheme constructed by Andersen over time became and up until recently was a de facto standard for numerical simulation of the solution of the Heston model. A promising extension of this scheme was presented by Chan and Joshi [8] in 2013. Finally, Ninomiya and Victoir [28] in 2008 and Alfonsi [3] in 2010 proposed discretization schemes constructed using split-step and moment-matching techniques, which are especially worth mentioning.

2.2 One-step pathwise approximations

We will further look into some of the mentioned schemes in more detail. As is clear from the previous chapter, there are three methods to construct a discretization scheme for the Heston model. The first and simplest method utilizes the Euler discretization scheme. The second method exploits the properties of the variance square-root (CIR) process. Finally, the third one uses a so-called split-step technique. Moment matching is sometimes used in parallel with all three methods.

We will present schemes as one-step weak approximations from time ih to time $(i+1)h$, i.e., from $\overline{Z}_{ih} = (\overline{S}_{ih}, \overline{Y}_{ih}) = (s, y)$ to $\overline{Z}_{(i+1)h} = (\overline{S}_{(i+1)h}, \overline{Y}_{(i+1)h})$. Here $i = \overline{0, [T/h]}$ is the index for the discretization of the fixed time interval $[0, T]$, and h is the discretization step. Note that in most of the cases, in order to preserve the nonnegativity, instead of the asset price S, its logarithm $X = \log(S)$ is discretized.

In most of the schemes (including ours presented further), the Cholesky decomposition is used to decompose correlated Brownian motions $\widetilde{W}_t$ and W_t with correlation coefficient ρ into uncorrelated Brownian motions B_t and W_t:

$$\begin{cases} \mathrm{d}\widetilde{W}_t & = \rho \mathrm{d}W_t + \sqrt{1-\rho^2}\mathrm{d}B_t, \\ \mathrm{d}W_t & = \mathrm{d}W_t. \end{cases} \tag{2.2.1}$$

Also, for simplicity, we will further use the following notations:

$$\Delta\widetilde{W} = \Delta\widetilde{W}_{(i+1)h} = \widetilde{W}_{(i+1)h} - \widetilde{W}_{ih} = \widetilde{W}((i+1)h) - \widetilde{W}(ih),$$
$$\Delta W = \Delta W_{(i+1)h} = W_{(i+1)h} - W_{ih} = W((i+1)h) - W(ih),$$
$$\Delta B = \Delta B_{(i+1)h} = B_{(i+1)h} - B_{ih} = B((i+1)h) - B(ih).$$

In order to simulate ΔW, $\Delta\widetilde{W}$, and ΔB, we can use a standard Gaussian variable, for example, $\Delta W \sim \ N(0,1)\sqrt{h}$.

2.3 Classical schemes and modifications

Euler–Maruyama scheme

The first and simplest idea is borrowing the famous and well-known (first presented in 1768) Euler method from the field of (ordinary) differential equations. Since the idea to reuse the Euler scheme in the field of numerical methods for stochastic calculus was first suggested by Maruyama, in this context the scheme is sometimes called the Euler–Maruyama scheme.

The Euler scheme for (1.2.1) has the following form:

$$\overline{X}_{(i+1)h} = \overline{X}_{ih} + (r - \frac{1}{2}\overline{Y}_{ih})h + \sqrt{\overline{Y}_{ih}}\left(\rho\Delta W + \sqrt{1-\rho^2}\Delta B\right),$$
$$\overline{Y}_{(i+1)h} = \overline{Y}_{ih} + kh(\theta - \overline{Y}_{ih}) + \sigma\sqrt{\overline{Y}_{ih}}\Delta W.$$

Unfortunately, the scheme has two big disadvantages. First, $\overline{Y}_{(i+1)h}$ can take negative values with positive probability - starting with positive $\overline{Y}_{ih} > 0$, we can keep $\overline{Y}_{(i+1)h} > 0$ if and only if

$$\Delta W > \frac{k\Delta t(\overline{Y}_{ih} - \theta) - \overline{Y}_{ih}}{\sigma\sqrt{\overline{Y}_{ih}}}.$$

Therefore, we have a strictly positive probability going negative:

$$P(\overline{Y}_{(i+1)h} < 0) = N\left(\frac{k\Delta t(\overline{Y}_{ih} - \theta) - \overline{Y}_{ih}}{\sigma\sqrt{\overline{Y}_{ih}}}\right),$$

where N is the normal distribution function with zero mean and variance Δt. That immediately disrupts pathwise simulation since in order to calculate the values for the next step, we need to calculate $\sqrt{\overline{Y}_{(i+1)h}}$.

Second, as the square-root is not globally Lipschitz (and not locally Lipschitz near 0), the usual theorems (see [19], [26]) for weak convergence cannot be applied, and, thus, the convergence of the scheme cannot be rigorously guaranteed.

Modifications

Several fixes were proposed to deal with the first problem. In [23], all of the fixes were shown to be particular cases of the following unifying framework:

$$\begin{aligned}
\hat{Y}_{(i+1)h} &= f_1(\overline{Y}_{ih}) + kh(\theta - f_2(\overline{Y}_{ih})) + \sigma\sqrt{f_3(\overline{Y}_{ih})}\Delta W;\\
\overline{Y}_{(i+1)h} &= f_3(\hat{Y}_{(i+1)h}),
\end{aligned}$$

where fixing functions $f_i, i = 1, 2, 3$, satisfy the following conditions:

- $f_i(x) = x$ for $x \geqslant 0$ and $i = 1, 2, 3$;
- $f_3(x) \geqslant 0$ for $x \in \mathbb{R}$.

Note that in Higham and Mao [15], the fixing functions $f_1(x) = f_2(x) = x$ and $f_3(x) = |x|$ were used. However, the best Euler modification for the Heston model in terms of the smallest bias produced is constructed when the fixing functions $f_1(x) = x$ and $f_2(x) = f_3(x) = \max(x, 0)$ are used. This modification was introduced in [23] and named "full truncation" since not only the terms involving Y in the diffusion part, but also terms involving Y in the drift part as well are truncated using the fixing function $\max(x, 0)$.

Nevertheless, only a strong convergence is proved for the "full truncation" scheme, and we still can only rely on good numerical simulation results in examples provided and not on any rigorous proof in case of weak approximations.

Kahl–Jackel scheme

Attempts to apply the Milstein scheme, another classical discretization scheme, encounter the same problems. Same methods as in the case of the Euler approximation to solve these problems can be used, although the Milstein scheme and its modifications are not so popular in the literature for the Heston model as the Euler scheme is. We would like to note an interesting attempt to modify the Milstein scheme (see [17]) by mixing its implicit variation (for variance process) and the Kahl–Jackel discretization scheme (for the asset price):

$$\begin{aligned}
\overline{X}_{(i+1)h} &= \overline{X}_{ih} + rh - \frac{1}{4}\left(\overline{Y}_{(i+1)h} + \overline{Y}_{ih}\right)h + \rho\sqrt{\overline{Y}_{ih}}\Delta W \\
&\quad + \frac{1}{2}\left(\sqrt{\overline{Y}_{(i+1)h}} + \sqrt{\overline{Y}_{(i+1)h}}\right)(\Delta W - \rho\Delta B) + \frac{1}{4}\rho\sigma(\Delta W^2 - h), \\
\overline{Y}_{(i+1)h} &= \frac{\overline{Y}_{ih} + k\theta h + \sigma\sqrt{\overline{Y}_{ih}}\Delta W + \frac{1}{4}\sigma^2(\Delta W^2 - h)}{1 + kh}.
\end{aligned}$$

Unfortunately, the scheme only solves the problem of negative values if $\sigma^2 < 4k\theta$ and is neither accurate nor fast if compared to other schemes in this chapter.

2.4 Broadie–Kaya scheme

Broadie–Kaya [7] composed a bias-free (exact) method to simulate the Heston model. Alas, their method is quite complicated and highly time consuming, which makes it almost impossible to apply in the real-world situations.

Their method exploits the known fact that for the CIR process, $Y((i+1)h)$ conditional upon $Y(ih)$ is a noncentral chi-squared with η degrees of freedom and noncentrality parameter $Y(ih)\lambda(ih,(i+1)h)$, up to a constant scaling factor, that is,

$$Y((i+1)h) \sim \frac{\mathrm{e}^{-kh}}{\lambda(ih,(i+1)h)}\chi^2_\eta\left(Y(ih)\lambda(ih,(i+1)h)\right),$$

with $\eta = \frac{4k\theta}{\sigma^2}$ degrees of freedom. Broadie and Kaya noticed that integrating the CIR equation (equation for variance) leads to the equation

$$Y((i+1)h) = Y(ih) + \int_{ih}^{(i+1)h} k(\theta - Y(u))\,\mathrm{d}u + \sigma\int_{ih}^{(i+1)h}\sqrt{Y(u)}\,\mathrm{d}W(u),$$

and the last integral can be expressed as

$$\int_{ih}^{(i+1)h}\sqrt{Y(u)}\,\mathrm{d}W(u) = \frac{Y((i+1)h) - Y(ih) - k\theta h + k\int_{ih}^{(i+1)h}Y(u)\,\mathrm{d}u}{\sigma},$$

provided that we know $Y((i+1)h)$, $Y(ih)$, and $\int_{ih}^{(i+1)h}Y(u)\,\mathrm{d}u$.

Using the Cholesky decomposition for the asset price process, we get

$$\mathrm{d}X(t) = (r - \frac{1}{2})Y(t)\,\mathrm{d}t + \rho\sqrt{Y(u)}\,\mathrm{d}W(u) + \sqrt{1-\rho^2}\sqrt{Y(u)}\,\mathrm{d}B(u),$$

where B is a Brownian motion independent of W. Finally, knowing $\int_{ih}^{(i+1)h} \sqrt{Y(u)}\mathrm{d}W(u)$ (and other terms), Broadie and Kaya shows that, in the integral form,

$$\begin{aligned} X((i+1)h) &= X(ih) + \frac{\rho}{\sigma}(Y((i+1)h) - Y(ih) - k\theta h) \\ &\quad + \left(r + \frac{k\rho}{\sigma} - \frac{1}{2}\right) \int\limits_{ih}^{(i+1)h} Y(u)\,\mathrm{d}u + \sqrt{1-\rho^2} \int\limits_{ih}^{(i+1)h} \sqrt{Y(u)}\,\mathrm{d}W(u). \end{aligned}$$

Now it is easy to see that, conditionally on $X(ih)$, the distribution of $X((i+1)h)$ is Gaussian:

$$X((i+1)h)\Big|_{X(ih)} \sim N\Big(X(ih) + rh - \frac{1}{2}\int\limits_{ih}^{(i+1)h} Y(u)\,\mathrm{d}u + \rho\int\limits_{ih}^{(i+1)h} \sqrt{Y(u)}\,\mathrm{d}W(u),$$
$$(1-\rho^2)\int\limits_{ih}^{(i+1)h} Y(u)\,\mathrm{d}u\Big).$$

Such a scheme presents two major challenges:

- To sample $Y((i+1)h)$ from $Y(ih)$;
- To generate $\int_{ih}^{(i+1)h} Y(u)\,\mathrm{d}u$. Unfortunately, the distribution of $\int_{ih}^{(i+1)h} Y(u)\,\mathrm{d}u$ conditional on $Y(ih)$ and $Y((i+1)h)$ is not known in a closed form. In [7], a characteristic function is derived, which allows the authors to use Fourier inversion and numerically generate the conditional cumulative distribution function for $\int_{ih}^{(i+1)h} Y(u)\,\mathrm{d}u$, and numerical inversion of this distribution function finally allows to generate a sample of $\int_{ih}^{(i+1)h} Y(u)\,\mathrm{d}u$ itself.

2.5 QE and other Andersen's schemes

The inspiration for the QE (quadratic exponential) and other schemes presented in the article [5] comes from the same source as in the case of Broadie–Kaya scheme, i.e., the fact that the CIR process $Y((i+1)h)$, conditional upon $Y(ih)$, is noncentral chi-squared with $\eta = \frac{4k\theta}{\sigma^2}$ degrees of freedom. Andersen [5] observes that a noncentral chi-square with moderate (or high) noncentrality parameter can be well represented by some power-function applied to a Gaussian variable. Therefore, in [5], it is suggested that

$$\overline{Y}_{(i+1)h} = a(b + N_1)^2,$$

where $N_1 \sim N(0,1)$ is a standard Gaussian random variable, and a and b are some constants determined by moment matching technique. However, this does work only for "sufficiently" large values of $\overline{Y}_{ih}$. For low values of $\overline{Y}_{ih}$, $\overline{Y}_{(i+1)h}$ is drawn from the approximated density

$$\mathbb{P}(\overline{Y}_{(i+1)h} \in [x, x + \mathrm{d}x]) \approx (p\delta(0) + \beta(1-p)\mathsf{e}^{-\beta x})\,\mathrm{d}x,\ \ x \geqslant 0.$$

If

$$\begin{aligned} m &:= \theta + (\overline{Y}_{ih} - \theta)\mathsf{e}^{-kh}, \\ s^2 &:= \frac{\overline{Y}_{ih}\sigma^2\mathsf{e}^{-kh}}{k}(1 - \mathsf{e}^{-kh}) + \frac{\theta\sigma^2}{2k}(1 - \mathsf{e}^{-kh})^2, \end{aligned}$$

and

$$\psi := \frac{s^2}{m^2},$$

then a and b are computed as follows:

$$\begin{aligned} b &= 2\psi^{-1} - 1 + \sqrt{2\psi^{-1}}\sqrt{2\psi^{-1} - 1}\ \geqslant\ 0, \\ a &= \frac{m}{1 + b^2}. \end{aligned}$$

For low values of $\overline{Y}_{ih}$, p and β need to be calculated:

$$\begin{aligned} p &= \frac{\psi - 1}{\psi + 1} \in [0, 1), \\ \beta &= \frac{1 - p}{m} = \frac{2}{m(\psi + 1)} > 0. \end{aligned}$$

If some arbitrary level $\psi_c \in [1, 2]$ is selected (according to [5], the exact choice has a relatively small effect on the quality of the overall simulation scheme; for numerical tests in [5], $\psi_c = 1.5$ is used), then

- if $\psi \leqslant \psi_c$, then

$$\overline{Y}_{(i+1)h} = a(b + N_1)^2;$$

- Otherwise, if $\psi > \psi_c$, then

$$\overline{Y}_{(i+1)h} = \begin{cases} 0, & 0 \leqslant U \leqslant p, \\ \beta^{-1} \log\left(\frac{1-p}{1-U}\right), & p < U \leqslant 1, \end{cases}$$

where U is a uniform random variable from $[0, 1]$.

For the asset price, the following discretization scheme is proposed:

$$\overline{X}_{(i+1)h} = \overline{X}_{ih} + K_0 + K_1\overline{Y}_{ih} + K_2\overline{Y}_{(i+1)h} + \sqrt{K_3\overline{Y}_{ih} + K_4\overline{Y}_{(i+1)h}}N_2,$$

where $N_2 \sim N(0,1)$ is a standard Gaussian random variable independent of $\overline{Y}$, and $K_0, \ldots, K_4$ are given by

$$\begin{aligned} K_0 &= -\frac{k\rho\theta}{\sigma}h; \\ K_1 &= \gamma_1 h(\frac{k\rho}{\sigma} - \frac{1}{2}) - \frac{\rho}{\sigma}; \\ K_2 &= \gamma_2 h(\frac{k\rho}{\sigma} - \frac{1}{2}) + \frac{\rho}{\sigma}; \\ K_3 &= \gamma_1 h(1-\rho^2); \\ K_4 &= \gamma_2 h(1-\rho^2), \end{aligned}$$

for some constants γ_1 and γ_2. The constants are used to handle the time-integral of Y:

$$\int\limits_{ih}^{(i+1)h} Y(u)\,\mathrm{d}u \approx h\left(\gamma_1 Y(ih) + \gamma_2 Y((i+1)h)\right).$$

There are many ways for setting γ_1 and γ_2. In [5], $\gamma_1 = \gamma_2 = 0.5$ is used.

This scheme can be improved by enforcing to comply with martingale condition $\mathbb{E}(\overline{Y}_{(i+1)h}|\overline{Y}_{ih}) = \overline{Y}_{ih}$, but although the resulting scheme QE-M (QE with martingale correction) is slightly more accurate, it appears to be much slower and more complicated.

Today QE is a de facto standard weak approximation scheme for the Heston model, for which test simulations show admiring results in terms of accuracy. Unfortunately, there are no rigorous proof of (weak) convergence of the scheme(s).

2.6 Alfonsi schemes

The first attempt to apply a split-step technique to the Heston model (and other multidimensional SDEs) was presented in 2008 by Ninomiya and Victoir [28].

Unfortunately, their scheme was only defined for $\sigma^2 \leqslant 4k\theta$. Building on the same ideas, in 2010, Alfonsi [3] constructed a second and a third order schemes for the Heston model without any restrictions to the model parameters.

To construct discretization schemes, Alfonsi rewrote the model (1.1.1) using the Cholesky decomposition $\mathrm{d}\widetilde{W}_t = \rho \,\mathrm{d}W_t + \sqrt{1-\rho^2}\,\mathrm{d}B_t$ (here W_t and B_t are independent Brownian motions) and split it into two SDEs

$$\begin{cases} \mathrm{d}S_t &= rS_t\,\mathrm{d}t + \rho\sqrt{Y_t}S_t\,\mathrm{d}W_t, \\ \mathrm{d}Y_t &= k(\theta - Y_t)\,\mathrm{d}t + \sigma\sqrt{Y_t}\,\mathrm{d}W_t, \end{cases}$$

and

$$\begin{cases} \mathrm{d}S_t &= \sqrt{(1-\rho^2)Y_t}S_t\,\mathrm{d}B_t, \\ \mathrm{d}Y_t &= 0. \end{cases}$$

It is easy to integrate the latter SDE exactly, and for the first SDE, the second- and third-order schemes for CIR process described in the same article (see [3]) were used resulting in the second- and third-order discretization schemes for the Heston model.

To compute the discretization for the next step, the following algorithm is proposed:

$$\left(\overline{S}_{(i+1)h}, \overline{Y}_{(i+1)h}\right) = \begin{cases} HZ(HW(\overline{S}_{ih}, \overline{Y}_{ih})) \text{ if } B = 1, \\ HW(HZ(\overline{S}_{ih}, \overline{Y}_{ih})) \text{ otherwise.} \end{cases}$$

Here B is a Bernoulli sample with parameter $1/2$,

$$HZ(s,y) = \left(s\mathrm{e}^{\sqrt{(1-\rho^2)yh}N}, y\right),$$

and

$$HW(s,y) = \left(s\cdot\exp\left[\left(r - \frac{\rho k\theta}{\sigma}\right)h + \frac{\rho\Delta y}{\sigma} + \left(\frac{\rho k}{\sigma} - 0.5\right)(y + 0.5\Delta y)\,h\right], y + \Delta y\right),$$

where $N \sim N(0,1)$ is a standard Gaussian variable, and $\Delta y = CIR_{2,3}(y) - y$. CIR_2 and CIR_3 are second- and third-order discretization schemes for the CIR equation, respectively. Using CIR_2, we construct the second-order Alfonsi scheme for the Heston model (ALF_2), whereas using CIR_3, the third-order one (ALF_3).

Despite the fact that the same framework was proved for the CIR equation, in the case of the Heston model, we have no rigorous proof whatsoever. In fact, the framework was proved to work if SDE has uniformly bounded moments, which is not true in the case of the Heston model. Nevertheless, the numerical simulations of ALF_2 and ALF_3 show very encouraging results.

CHAPTER 3

Preliminaries

In this chapter, we provide all needed definitions and describe a technique used to construct discretization schemes.

3.1 Preliminaries and definitions

Consider the general two-dimensional SDE

$$\mathrm{d}Z_t = b(Z_t,t)\,\mathrm{d}t + \sigma(Z_t,t)\,\mathrm{d}B_t,\ \ t \geq 0,\ Z_0 = z, \tag{3.1.1}$$

for $z \in \mathbb{D} \subset \mathbb{R}^2$ with standard two-dimensional Brownian motion $B_t = (B_t^1, B_t^2)$ and coefficients $b : \mathbb{R}^2 \times I \to \mathbb{R}^2$, and $\sigma = (\sigma_{ij})$, $i,j = 1,2$, where $\sigma_{ij} : \mathbb{R}^2 \times I \to \mathbb{R}^2$ and $I = [0,T] \subset \mathbb{R}$ is a time interval. We assume that the SDE is domain-preserving in the sense that, for every $z \in \mathbb{D}$, the SDE has a unique weak solution Z^z such that $\mathbb{P}\{Z_t^z \in \mathbb{D},\, t \geqslant 0\} = 1$. For example, for the Heston model (1.1.1) one can take $\mathbb{D} = (0,+\infty) \times [0,+\infty)$, and for the log-Heston model (1.2.1) - $\mathbb{D} = \mathbb{R} \times [0,+\infty)$, provided that all the parameters are strictly positive $\theta, \sigma, k > 0$. Note, that further on we will presume the latter in the setting of the log-Heston model and the former in the setting of the Heston model.

On a fixed time interval $[0,T]$, we consider equidistant time discretizations $\Delta^h = \{ih, i = 0,\dots,[T/h]\}$, where $[a]$ denotes the integer part of a number a.

By $C_0^\infty(\mathbb{D})$ we denote the functions on $\mathbb{D}$ of class C^∞ with compact support, and by $C_{pol}^\infty(\mathbb{D})$ the real-valued functions of class C^∞ with all partial derivatives of polynomial growth:

$$C_{pol}^\infty(\mathbb{D}) := \Big\{ f \in C^\infty(\mathbb{D}) : \forall i \in \mathbb{N}_0^2,\ \exists C_i > 0, \exists k_i \geqslant 0, \forall z \in \mathbb{D}, |f^{(i)}(z)| \leqslant C_i(1+|z|^{k_i}) \Big\}.$$

Here $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$, $i = (i_1, i_2) \in \mathbb{N}_0^2$ are multiindices, and

$$f^{(i)}(z) := \frac{\partial^{|i|} f(z)}{\partial z^i}, \quad |i| := i_1 + i_2, \quad \partial z^i := \partial z_1^{i_1} \partial z_2^{i_2}.$$

Following [3], we call the "sequence" $\{(C_i, k_i) : i \in \mathbb{N}_0^2\}$ a good sequence for f.

We shall write $g(z, h) = \mathrm{O}_{p,k}(h^n)$ if for some $C > 0$ and $h_0 > 0$,

$$|g(z, h)| \le Ch^n(1 + |z|^k), \quad z \in \mathbb{D}, \; h \le h_0$$

(the subscript p stands for *polynomial*). We will write $g(z, h) = \mathrm{O}_p(h^n)$ if $g(z, h) = \mathrm{O}_{p,k}(h^n)$ for some $k \in \mathbb{N}$. If, in particular, the function g is expressed in terms of another function $f \in C^\infty_{pol}(\mathbb{D})$ and the constants C, h_0, and k depend on a good sequence for f only, then we will write, instead, $g(z, h) = \mathcal{O}_p(h^n)$. Note the following simple multiplication rules: $\mathrm{O}_p(h^n)\mathrm{O}_p(h^m) = \mathrm{O}_p(h^{n+m})$; $\mathcal{O}_p(h^n)\mathrm{O}_p(h^m) = \mathcal{O}_p(h^{n+m})$; $\mathcal{O}_p(h^n)\mathcal{O}_p(h^m) = \mathcal{O}_p(h^{n+m})$.

Definition 3.1. A discretization scheme $\bar{Z}^h$ is a family of discrete-time $\mathbb{D}$-valued time-homogeneous Markov chains $\bar{Z}^h = \{\hat{Z}^h(z, t),\, z \in \mathbb{D},\, t \in \Delta^h\}$, $h > 0$, with initial values $\hat{Z}^h(z, 0) = z$.

With some abuse of notation, sometimes we omit the superscript h and write $\bar{Z}_t^z$ or $\bar{Z}(z, t)$ instead of $\bar{Z}^h(z, t)$. Note that to define a discretization scheme (in distribution), it suffices to construct one-step "transitions," that is, the (distributions of) random variables $\bar{Z}_h^z = \bar{Z}^h(z, h)$ for all $z \in \mathbb{D}$. Therefore, we shall also call the latter a discretization scheme.

Definition 3.2. A discretization scheme $\bar{Z}^h$ is a weak υth-order approximation for the solution Z^z of Eq. (3.1.1) if, for every $f \in C_0^\infty(\mathbb{D})$, there exists $K > 0$ such that

$$\left|\mathbb{E}f(Z_T^z) - \mathbb{E}f(\bar{Z}_T^z)\right| = \left|\mathbb{E}f(Z_T^z) - \mathbb{E}f\left(\bar{Z}^h(z, T)\right)\right| \leqslant Kh^\upsilon, \;\; h > 0. \qquad (3.1.2)$$

Definition 3.3. Let L be the generator of the solution Z^z of Eq. (3.1.1). The υth-order remainder of a discretization scheme $\bar{Z}^h$ is the operator $R_\upsilon^h : C^\infty_{pol}(\mathbb{D}) \to C(\mathbb{D})$ defined by

$$R_\upsilon^h f(z) := \mathbb{E}f\big(\bar{Z}_h^z\big) - \left[f(z) + \sum_{k=1}^{\upsilon} \frac{L^k f(z)}{k!} h^k\right], \; z \in \mathbb{D}, \; h > 0. \qquad (3.1.3)$$

Definition 3.4. A discretization scheme $\bar{Z}^h$ is a potential υth-order weak approximation of Eq. (3.1.1) if, for every $f \in C^{\infty}_{pol}(\mathbb{D})$,

$$R^h_{\upsilon} f(z) = \mathcal{O}_p(h^{\upsilon+1}). \tag{3.1.4}$$

We will say that $\bar{Z}^h$ is a strongly potential υth-order weak approximation if, in addition, it has uniformly bounded moments of all orders, that is, for all $q \in \mathbb{N}$,

$$\sup_{N\in\mathbb{N}} \sup_{0\le k\le N} \mathbb{E}|\bar{Z}^z_{kh}|^q < +\infty, \quad z \in \mathbb{D}. \tag{3.1.5}$$

Remark 3.1. For the CIR process Y, Alfonsi [2] proved that any strongly potential υth-order weak approximation of the CIR process is, in fact, a υth-order weak approximation. The proof was essentially based on checking that the function $u(t,x) := Ef(Y^x_{T-t})$ is smooth for smooth functions f. We believe that the same applies in our case, but for the moment, we have no rigorous proof of analogous result for the log-Heston model since, in contrast to the CIR process, we do not know a closed-form formula for the density of the log-Heston process (X, Y).[1]

3.2 Split-step technique for the Heston model

In this section we focus on the split-step method for the Heston model. Unfortunately, when trying to apply the method directly to Eq. (1.1.1), we cannot assure the positivity of approximations. Besides, we know (see, for example, [6] and [11]) that moments of the solution of the Heston model explode (i.e., tend to infinity in finite time). Therefore, instead of S_t in Eq. (1.1.1), we approximate $X_t := \log S_t$ and thus consider, instead of the Heston model (1.1.1), the log-Heston model (1.2.1).

Denoting $Z_t = (X_t, Y_t)$ and $z = (x, y)$, we split Eq. (1.2.1) into the "deterministic part"

$$\begin{cases} \mathrm{d}D^z_1(t) = (r - \frac{1}{2}D^y_2(t))\,\mathrm{d}t, & D^z_1(0) = x, \\ \mathrm{d}D^y_2(t) = k(\theta - D^y_2(t))\,\mathrm{d}t, & D^y_2(0) = y, \end{cases} \tag{3.2.1}$$

and the "stochastic part"

$$\mathrm{d}Z_t = \mathrm{d}\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \sqrt{Y_t} & 0 \\ 0 & \sigma\sqrt{Y_t} \end{pmatrix} \begin{pmatrix} \mathrm{d}\widetilde{W}_t \\ \mathrm{d}W_t \end{pmatrix}. \tag{3.2.2}$$

[1] A density formula given by del Baño Rolin et al. [30] in the form of infinite convolution of known densities is rather inconvenient for applications.

Here $Z_0 = z = (x, y) \in \mathbb{D}$, and, as before, $\mathrm{d}\widetilde{W}_t \, \mathrm{d}W_t = \rho \, \mathrm{d}t$.

The solution of the deterministic part, $D(z,t) = (D_1^z(t), D_2^y(t))$, can be easily found:

$$\begin{cases} D_1^z(t) = x + (r - \frac{1}{2}\theta)t + \frac{1}{2k}\left(\mathrm{e}^{-kt} - 1\right)(y - \theta)\,, \\ D_2^y(t) = y\mathrm{e}^{-kt} + \theta(1 - \mathrm{e}^{-kt}). \end{cases} \tag{3.2.3}$$

The solution of the stochastic part is not known in an explicit form. Therefore, for the stochastic part, we need to construct a discretization scheme $\hat{Z}^h = \hat{Z}(z,h) = \left(\hat{X}(z,h), \hat{Y}(y,h)\right)$. Having done this, we can define the first-order split-step approximation for Z_t by the composition

$$\begin{aligned} \bar{Z}^h = \bar{Z}(z,h) = \begin{pmatrix} \bar{X}(z,h) \\ \bar{Y}(y,h) \end{pmatrix} &:= D\left[\hat{Z}(z,h), h\right] \\ &= \begin{pmatrix} \hat{X}(z,h) + (r - \frac{1}{2}\theta)h + \frac{1}{2k}(\mathrm{e}^{-kh} - 1)\big(\hat{Y}(y,h) - \theta\big) \\ \hat{Y}(y,h)\mathrm{e}^{-kh} + \theta(1 - \mathrm{e}^{-kh}) \end{pmatrix} \end{aligned} \tag{3.2.4}$$

and, similarly, the second-order split-step approximation for Z_t by the composition

$$\bar{Z}^h = \bar{Z}(z,h) = \begin{pmatrix} \bar{X}(z,h) \\ \bar{Y}(y,h) \end{pmatrix} := D\left(\tilde{Z}\big(D(z, h/2), h\big), h/2\right). \tag{3.2.5}$$

The following proposition follows from Corollary 1.7 in [3] (see also [24, 25]).

Proposition 3.2. Suppose that $\hat{Z}^h = (\hat{X}^h, \hat{Y}^h)$ is a potential first- or second-order approximation of the stochastic part (3.2.2) of the log-Heston system (1.2.1). Then the compositions (3.2.4) and (3.2.5) define a potential first- and second-order approximations of the log-Heston system (1.2.1) respectively.

Finally, having constructed the split-step approximation (3.2.4) or (3.2.5) for the log-Heston model, we can return to the Heston model (1.1.1) by approximating $S_t = \exp(X_t)$ with the exponent of $\bar{X}$, that is, by

$$\overline{S}^h(s, y, t) = \exp\left\{\bar{X}^h(\log s, y, t)\right\}.$$

Note that the positivity of the approximation $\overline{S}^h$ is trivially assured.

3.3 Moment matching technique for the Heston model

One way to apply the moment matching technique is to follow these steps. In the first step we use Taylor's formula for $\bar{Z}_t^z = (\overline{X}_h^z, \overline{Y}_h^y)$:

$$\begin{aligned}\mathbb{E}f(\bar{Z}_h^z) &= f(z) + \frac{\partial f}{\partial x}(z)\mathbb{E}(\overline{X}_h^z - x) + \frac{\partial f}{\partial y}(z)\mathbb{E}(\overline{Y}_h^y - y) \\ &+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(z)\mathbb{E}(\overline{X}_h^z - x)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(z)\mathbb{E}(\overline{Y}_h^y - y)^2 \\ &+ \frac{\partial^2 f}{\partial x \partial y}(z)\mathbb{E}(\overline{X}_h^z - x)(\overline{Y}_h^y - y) + \cdots .\end{aligned}$$

In the second step, we write Dynkin's formula for $\mathbb{E}f(Z_h)$:

$$\mathbb{E}f(Z_h) = f(z) + Lf(z)h + \frac{h^2}{2!}L^2 f(z) + \frac{h^3}{3!}L^3 f(z) + \cdots ,$$

where

$$\begin{aligned}Lf(z) &= \left(r - \frac{1}{2}y\right)\frac{\partial f}{\partial x}(z) + k\left(\theta - y\right)\frac{\partial f}{\partial y}(z) \\ &+ \frac{y}{2}\left(\frac{\partial^2 f}{\partial x^2}(z) + \sigma^2\frac{\partial^2 f}{\partial y^2}(z)\right) + \rho\sigma y\frac{\partial^2 f}{\partial x \partial y}\ (z)\end{aligned}$$

is the generator of the log-Heston model (1.2.1).

In the next step, we match the coefficients of the $\frac{\partial f}{\partial x}(z)$, $\frac{\partial f}{\partial y}(z)$, $\frac{\partial^2 f}{\partial x^2}(z)$, $\frac{\partial^2 f}{\partial y^2}(z)$ and others in the Taylor's and Dynkin's formulas to get the equations for the moments $\mathbb{E}(\overline{X}_h^z - x)$, $\mathbb{E}(\overline{Y}_h^y - y)$, $\mathbb{E}(\overline{X}_h^z - x)^2$, $\mathbb{E}(\overline{Y}_h^y - y)^2$, and so on. The more equations we write, the more precise, yet more difficult to construct, approximation we aim at.

For example, taking only the coefficients of up to the second-order partial derivatives and allowing some bias ($\mathrm{O}_p(h^2)$) we would arrive at the following system of equations:

$$\begin{cases}\mathbb{E}(\overline{X}_h^z - x) = \left(r - \frac{1}{2}y\right)h + \mathrm{O}_p(h^2), \\ \mathbb{E}(\overline{Y}_h^y - y) = k(\theta - y)h + \mathrm{O}_p(h^2), \\ \mathbb{E}(\overline{X}_h^z - x)^2 = yh + \mathrm{O}_p(h^2), \\ \mathbb{E}(\overline{Y}_h^y - y)^2 = y\sigma^2 h + \mathrm{O}_p(h^2), \\ \mathbb{E}(\overline{X}_h^z - x)(\overline{Y}_h^y - y) = \rho\sigma y h + \mathrm{O}_p(h^2).\end{cases}$$

It is not very difficult to see that the conditions above are satisfied by... the Euler approximation

$$\overline{X}^z_h = \overline{X}(x, y, h, \xi_1, \xi_2) = x + \left(r - \frac{1}{2}y\right) h + \sqrt{y}\left(\rho\xi_1 + \sqrt{1-\rho^2}\xi_2\right),$$

$$\overline{Y}^y_h = \overline{Y}(y, h, \xi_1) = y + k\left(\theta - y\right) h + \sigma\sqrt{y}\xi_1,$$

where independent random variables ξ_1 and ξ_2 are distributed normally with mean 0 and variance h (see Section 2.3).

Chapter **4**

First-order approximation

In this chapter we construct a first-order approximation scheme for the solution of the Heston system (1.1.1). In the first section of the chapter we adapt a split-step technique described earlier to our needs, following section is dedicated to finding an approximation scheme for the (split) stochastic part and finally, in the last section, numerical simulations illustrating our approach are provided.

4.1 A potential first-order approximation

We start to construct our scheme using methods described in Section 3.2. First, in order to assure the positivity and avoid moment explosions we change our model (S_t, Y_t) to the log-Heston model $Z_t = (X_t := \log S_t, Y_t)$ (see (1.2.1)), then split the log-Heston model into the stochastic and deterministic parts. While the solution of the latter $D(z,t) = (D_1^z(t), D_2^y(t))$ is easy to find (see (3.2.3)), we need to construct a discretization scheme for the solution of the stochastic part $\hat{Z}^h = \hat{Z}(z,h) = \Big(\hat{X}(z,h), \hat{Y}(y,h)\Big)$. Having done this, we can define the first-order split-step approximation for Z_t by the composition

$$\begin{aligned}\bar{Z}^h = \bar{Z}(z,h) = \begin{pmatrix}\bar{X}(z,h)\\ \bar{Y}(y,h)\end{pmatrix} &:= D\left[\hat{Z}(z,h), h\right]\\ &= \begin{pmatrix}\hat{X}(z,h) + (r - \frac{1}{2}\theta)h + \frac{1}{2k}(\mathsf{e}^{-kh} - 1)\big(\hat{Y}(y,h) - \theta\big)\\ \hat{Y}(y,h)\mathsf{e}^{-kh} + \theta(1 - \mathsf{e}^{-kh})\end{pmatrix}.\end{aligned} \tag{4.1.1}$$

Proposition 3.2 assures that the composition (4.1.1) defines a potential first-order approximation of the log-Heston system (1.2.1) if we provide a potential first-order approximation scheme of the *stochastic part* of the log-Heston model equation.

4.2 A potential first-order approximation of the stochastic part

Using the Cholesky decomposition, we decompose the correlated Brownian motions $\widetilde{W}_t$ and W_t into uncorrelated Brownian motions B_t and W_t and rewrite the stochastic part of Eq. (1.2.1)

$$\begin{cases} \mathrm{d}X_t = \sqrt{Y_t}\left(\rho\,\mathrm{d}W_t + \sqrt{1-\rho^2}\,\mathrm{d}B_t\right), \\ \mathrm{d}Y_t = \sigma\sqrt{Y_t}\,\mathrm{d}W_t, \end{cases} \tag{4.2.1}$$

where B is a Brownian motion independent of W.

The generator of the stochastic part is

$$Lf(z) = \frac{y}{2}\left(\frac{\partial^2}{\partial x^2} + \sigma^2\frac{\partial^2}{\partial y^2}\right)f(z) + \rho\sigma y\frac{\partial^2 f}{\partial x\partial y}\,(z).$$

Let $\hat{Z}^z_h = (\hat{X}^z_h, \hat{Y}^y_h)$ be any discretization scheme.[1] Using Taylor's formula, we write

$$\begin{aligned} f(\hat{Z}^z_h) = f(z) &+ \frac{\partial f}{\partial x}(z)(\hat{X}^z_h - x) + \frac{\partial f}{\partial y}(z)(\hat{Y}^y_h - y) \\ &+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(z)(\hat{X}^z_h - x)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(z)(\hat{Y}^y_h - y)^2 \\ &+ \frac{\partial^2 f}{\partial x\partial y}(z)(\hat{X}^z_h - x)(\hat{Y}^y_h - y) \\ &+ \nabla_3(\hat{X}^z_h, \hat{Y}^y_h, z) + r_3(\hat{X}^z_h, \hat{Y}^y_h, z), \end{aligned}$$

where

$$\nabla_3(\hat{X}^z_h, \hat{Y}^y_h, z) := \frac{1}{6}\left(\frac{\partial}{\partial x}(\hat{X}^z_h - x) + \frac{\partial}{\partial y}(\hat{Y}^y_h - y)\right)^3 f(x,y) \tag{4.2.2}$$

and

$$r_3(\hat{X}^z_h, \hat{Y}^y_h, z) \tag{4.2.3}$$

are the third-order term and third-order remainder in the formula, respectively. So, the first-order remainder of a discretization scheme $\hat{Z}^z_h$ is

$$R^h_1 f(z) = \mathbb{E}f(\hat{Z}^z_h) - \Big(f(z) + Lf(z)h\Big)$$

[1] Note that we write $\hat{Y}^y_h$ instead of $\hat{Y}^z_h$ since the latter actually depends on y only.

$$
\begin{aligned}
&= \frac{\partial f}{\partial x}(z)\mathbb{E}(\hat{X}_h^z - x) + \frac{\partial f}{\partial y}(z)\mathbb{E}(\hat{Y}_h^y - y) \\
&+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(z)\left(\mathbb{E}(\hat{X}_h^z - x)^2 - yh\right) \\
&+ \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(z)\left(\mathbb{E}(\hat{Y}_h^y - y)^2 - y\sigma^2 h\right) \\
&+ \frac{\partial^2 f}{\partial x \partial y}(z)\left(\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y) - \rho\sigma y h\right) \\
&+ \mathbb{E}\nabla_3(\hat{X}_h^z, \hat{Y}_h^y, z) + \mathbb{E}r_3(\hat{X}_h^z, \hat{Y}_h^y, z).
\end{aligned}
$$

Therefore, a discretization scheme $\hat{Z}_h^z = (\hat{X}_h^z, \hat{Y}_h^y)$ is a potential first-order approximation if the following conditions are satisfied:

$$
\begin{cases}
\mathbb{E}(\hat{X}_h^z - x) = 0, \\
\mathbb{E}(\hat{Y}_h^y - y) = 0, \\
\mathbb{E}(\hat{X}_h^z - x)^2 = yh, \\
\mathbb{E}(\hat{Y}_h^y - y)^2 = y\sigma^2 h, \\
\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y) = \rho\sigma y h,
\end{cases}
\tag{4.2.4}
$$

and for some $h_0 > 0$, $C_3 > 0$, $C_4 > 0$, and $p, q \in \mathbb{N}_0$ depending only on a good sequence for $f \in C_{pol}^\infty(\mathbb{D})$,

$$
|\mathbb{E}\nabla_3(\hat{X}_h^z, \hat{Y}_h^y, z)| \leqslant C_3 h^2(1 + |z|^q), \quad z \in \mathbb{D}, \quad 0 < h < h_0, \tag{4.2.5}
$$

and

$$
E|r_3(\hat{X}_h^z, \hat{Y}_h^y, x, y)| \leqslant C_4 h^2(1 + |z|^p), \quad z \in \mathbb{D}, \quad 0 < h < h_0. \tag{4.2.6}
$$

Proposition 4.1. Let, at each step, $\hat{Y}_h^y$ be a random variable taking the values

$$
y_{1,2} = y + \sigma^2 h \pm \sqrt{(y + \sigma^2 h)\,\sigma^2 h} \quad \text{with corresponding probabilities } p_{1,2} = \frac{y}{2y_{1,2}}, \tag{4.2.7}
$$

and the random variable $\hat{X}_h^z$ be defined by

$$
\hat{X}_h^z := x + \sqrt{1-\rho^2}(\tilde{X}_h^z - x) + \frac{\rho}{\sigma}(\hat{Y}_h^y - y), \tag{4.2.8}
$$

where $\tilde{X}_h^z$ is a random variable, independent of $\hat{Y}_h^y$, taking the values

$$
x_{1,2} = x \pm \sqrt{yh} \quad \text{with probabilities } \frac{1}{2}. \tag{4.2.9}
$$

Then $\hat{Z}_h^z := (\hat{X}_h^z, \hat{Y}_h^y)$ satisfies all the requirements in Eqs. (4.2.4)–(4.2.6).

In the proof, we shall need the following expressions from [27]:

$$\mathbb{E}(\hat{Y}_h^y - y)^3 = 2y(\sigma^2 h)^2 \tag{4.2.10}$$

and

$$\mathbb{E}(\hat{Y}_h^y - y)^{2m} = y(\sigma^2 h)^m \hat{P}_m(y, \sigma^2 h), \tag{4.2.11}$$

where $\hat{P}_m$ and $\hat{Q}_m$ are $(m-1)$th-order homogeneous two-variable polynomials with positive integer coefficients. Also, we shall need the following equation:

$$\mathbb{E}(\tilde{X}_h^z - x)^n = \frac{(\sqrt{yh})^n}{2} + \frac{(-\sqrt{yh})^n}{2} = \begin{cases} (yh)^m, & n = 2m, \\ 0, & n = 2m-1 \end{cases} \tag{4.2.12}$$

and two simple lemmas on the moments of $\hat{X}_h^z$.

Lemma 4.2. For some constant $C > 0$ depending on $n \in \mathbb{N}$, σ, and h_0,

$$\mathbb{E}(\hat{X}_h^z)^n \leqslant C(1 + |z|^n), \quad z \in \mathbb{D}, \quad h \le h_0.$$

Proof. Using Eqs. (4.2.12) and (4.2.11) for even-order moments of $\hat{X}_h^z$ we have

$$\begin{aligned}
\mathbb{E}(\hat{X}_h^z)^{2m} &= \mathbb{E}\left(x + a(\tilde{X}_h^z - x) + b(\hat{Y}_h^y - y)\right)^{2m} \\
&\leqslant C\left(x^{2m} + a^{2m}\mathbb{E}(\tilde{X}_h^z - x)^{2m} + b^{2m}\mathbb{E}(\hat{Y}_h^y - y)^{2m}\right) \\
&= C\left(x^{2m} + a^{2m}(yh)^m + b^{2m}y(\sigma^2 h)^m \hat{P}_m(y, \sigma^2 h)\right) \\
&\leqslant C\left(x^{2m} + a^{2m}y^m h_0^m + b^{2m}y\sigma^{2m}h_0^m \hat{C}(y^{m-1} + (\sigma^2 h_0)^{m-1})\right) \\
&\leqslant C(1 + x^{2m} + y^m) \\
&\leqslant C(1 + |z|^{2m}), \quad z \in \mathbb{D}, \quad h < h_0.
\end{aligned}$$

Note, that constant C above (and below) may vary from line to line. Similarly, for odd-order moments, we have

$$\begin{aligned}
\mathbb{E}(\hat{X}_h^z)^{2m+1} &= \mathbb{E}\left(x + a(\tilde{X}_h^z - x) + b(\hat{Y}_h^y - y)\right)^{2m+1} \\
&\leqslant C\left(x^{2m+1} + a^{2m+1}\mathbb{E}(\tilde{X}_h^z - x)^{2m+1} + b^{2m+1}\mathbb{E}(\hat{Y}_h^y - y)^{2m+1}\right) \\
&= C\left(x^{2m+1} + b^{2m+1}y(\sigma^2 h)^{m+1}\hat{P}_m(y, \sigma^2 h)\right) \\
&\leqslant C\left(x^{2m+1} + b^{2m+1}y\sigma^{2m+2}h_0^{m+1}\hat{C}(y^{m-1} + (\sigma^2 h_0)^{m-1})\right) \\
&\leqslant C(1 + x^{2m+1} + y^m)
\end{aligned}$$

$$\leqslant C(1 + |z|^{2m+1}), \quad z \in \mathbb{D}, \quad h < h_0.$$

□

Lemma 4.3. For some constants $C > 0$ and $\hat{C} > 0$ depending on $m \in \mathbb{N}$, σ, and h_0,

$$\mathbb{E}(\hat{X}_h^z - x)^{2m} \leqslant C h^m (1 + y^m),$$
$$\mathbb{E}(\hat{X}_h^z - x)^{2m+1} \leqslant \tilde{C} h^m (1 + y^m), \quad z = (x, y) \in \mathbb{D}, \quad h < h_0.$$

Proof. Using Eqs. (4.2.12) and (4.2.11) for even-order moments of $\hat{X}_h^z - x$, we have

$$\begin{aligned}
\mathbb{E}(\hat{X}_h^z - x)^{2m} &= \mathbb{E}\left(a(\tilde{X}_h^z - x) + b(\hat{Y}_h^y - y)\right)^{2m} \\
&\leqslant C\left(a^{2m}\mathbb{E}(\tilde{X}_h^z - x)^{2m} + b^{2m}\mathbb{E}(\hat{Y}_h^y - y)^{2m}\right) \\
&= C\left(a^{2m}(yh)^m + b^{2m}y(\sigma^2 h)^m \hat{P}_m(y, \sigma^2 h)\right) \\
&\leqslant C h^m \left(a^{2m}y^m + b^{2m}y\sigma^{2m}\hat{C}(y^{m-1} + (\sigma^2 h_0)^{m-1})\right) \\
&\leqslant C h^m (1 + y^m), \quad z = (x, y) \in \mathbb{D}, \quad h < h_0.
\end{aligned}$$

As in the previous lemma, the constant C varies from line to line. For odd-order moments, we have

$$\begin{aligned}
\mathbb{E}(\hat{X}_h^z - x)^{2m+1} &= \mathbb{E}\left(a(\tilde{X}_h^z - x) + b(\hat{Y}_h^y - y)\right)^{2m+1} \\
&\leqslant C\left(a^{2m+1}\mathbb{E}(\tilde{X}_h^z - x)^{2m+1} + b^{2m+1}\mathbb{E}(\hat{Y}_h^y - y)^{2m+1}\right) \\
&= C\left(b^{2m+1}y(\sigma^2 h)^{m+1}\hat{P}_m(y, \sigma^2 h)\right) \\
&\leqslant C h^m \left(b^{2m+1}y\sigma^{2m+2}h_0\hat{C}(y^{m-1} + (\sigma^2 h_0)^{m-1})\right) \\
&\leqslant C h^m (1 + y^m), \quad z = (x, y) \in \mathbb{D}, \quad h < h_0.
\end{aligned}$$

□

Proof of Proposition 4.1. The first four equalities in (4.2.4) are easily checked directly (cf. [3], Sect. 2.2, [27], Sect. 3). The fifth equality follows from the independence of $\tilde{X}_h^z$ and $\hat{Y}_h^y$ by direct calculation using the second and fourth equalities in (4.2.4):

$$\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y) = \sqrt{1 - \rho^2}\mathbb{E}(\tilde{X}_h^z - x)(\hat{Y}_h^y - y) + \frac{\rho}{\sigma}\mathbb{E}(\hat{Y}_h^y - y)^2 = \rho\sigma y h.$$

Now let us prove estimate (4.2.5). Denote, for short, $a := \sqrt{1-\rho^2}$ and $b := \frac{\rho}{\sigma}$. Using Eqs. (4.2.12) and (4.2.10), the symmetry of $\tilde{X}_h^z - x$, and the independence of $\tilde{X}_h^z$ and $\hat{Y}_h^y$, we have

$$\begin{aligned}
\mathbb{E}(\hat{X}_h^z - x)^3 &= \mathbb{E}\left(a(\tilde{X}_h^z - x) + b(\hat{Y}_h^y - y)\right)^3 \\
&= a^3\mathbb{E}(\tilde{X}_h^z - x)^3 + b^3\mathbb{E}(\hat{Y}_h^y - y)^3 \\
&+ 3a^2b\mathbb{E}(\tilde{X}_h^z - x)^2(\hat{Y}_h^y - y) + 3ab^2\mathbb{E}(\tilde{X}_h^z - x)(\hat{Y}_h^y - y)^2 \\
&= b^3\mathbb{E}(\hat{Y}_h^y - y)^3 = b^3 2y\sigma^4h^2.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y) &= \mathbb{E}\left(a(\tilde{X}_h^z - x) + b(\hat{Y}_h^y - y)\right)^2(\hat{Y}_h^y - y) \\
&= a^2\mathbb{E}(\tilde{X}_h^z - x)^2(\hat{Y}_h^y - y) + 2ab\mathbb{E}(\tilde{X}_h^z - x)(\hat{Y}_h^y - y)^2 + b^2\mathbb{E}(\hat{Y}_h^y - y)^3 \\
&= b^2\mathbb{E}(\hat{Y}_h^y - y)^3 = b^2 2y\sigma^4h^2
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y)^2 &= \mathbb{E}\left(a(\tilde{X}_h^z - x) + b(\hat{Y}_h^y - y)\right)(\hat{Y}_h^y - y)^2 \\
&= a\mathbb{E}(\tilde{X}_h^z - x)(\hat{Y}_h^y - y)^2 + b\mathbb{E}(\hat{Y}_h^y - y)^3 \\
&= b\mathbb{E}(\hat{Y}_h^y - y)^3 = b2y\sigma^4h^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\mathbb{E}\nabla_3(\hat{X}_h^z, \hat{Y}_h^y, z) \\
&\quad = \frac{1}{6}\frac{\partial^3 f}{\partial x^3}(z)\mathbb{E}(\hat{X}_h^z - x)^3 + \frac{1}{2}\frac{\partial^3 f}{\partial x^2\partial y}(z)\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y) \\
&\quad + \frac{1}{2}\frac{\partial^3 f}{\partial x\partial y^2}(z)\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y)^2 + \frac{1}{6}\frac{\partial^3 f}{\partial y^3}(z)\mathbb{E}(\hat{Y}_h^y - y)^3 \\
&\quad = y\sigma^4h^2\left(\frac{b^3}{3}\frac{\partial^3 f}{\partial x^3}(x,y) + b^2\frac{\partial^3 f}{\partial x^2\partial y}(z) + b\frac{\partial^3 f}{\partial x\partial y^2}(z) + \frac{1}{3}\frac{\partial^3 f}{\partial y^3}(z)\right) \\
&\quad \le C_3h^2(1+|z|^q)
\end{aligned}$$

for all $f \in C_{pol}^\infty(\mathbb{D})$, which proves estimate (4.2.5).

Now consider

$$\begin{aligned}
&\mathbb{E}r_3(\hat{X}_h^z, \hat{Y}_h^y, x, y) \\
&\quad = \mathbb{E}\Bigg[\frac{1}{24}\frac{\partial^4 f}{\partial x^4}(z_\theta)(\hat{X}_h^z - x)^4 + \frac{1}{6}\frac{\partial^4 f}{\partial x^3\partial y}(z_\theta)(\hat{X}_h^z - x)^3(\hat{Y}_h^y - y)
\end{aligned}$$

$$+\frac{1}{4}\frac{\partial^4 f}{\partial x^2 \partial y^2}(z_\theta)(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y)^2$$
$$+\frac{1}{6}\frac{\partial^4 f}{\partial x \partial y^3}(z_\theta)(\hat{X}_h^z - x)(\hat{Y}_h^y - y)^3 + \frac{1}{24}\frac{\partial^4 f}{\partial y^4}(z_\theta)(\hat{Y}_h^y - y)^4\Bigg],$$

where $z_\theta = (x_\theta, y_\theta)$ with $x_\theta = x + \theta(\hat{X}_h^z - x)$ and $y_\theta = y + \theta(\hat{Y}_h^y - y)$ for some $\theta \in (0,1)$.

By Lemma 4.2, there are $C_5 > 0$, $h_0 > 0$, and $n \in \mathbb{N}_0$, depending only on a good sequence of f, such that

$$\mathbb{E}(\hat{X}_h^z)^n \leqslant C_5(1 + |z|^n), \quad z \in \mathbb{D}, \quad h < h_0.$$

From [27] we also know that

$$\mathbb{E}(\hat{Y}^h)^n = y\hat{R}_n(y, \sigma^2 h),$$

where $\hat{R}_n$ is an $(n-1)$th-order homogeneous two-variable polynomial with positive integer coefficients.

For any $f \in C_{pol}^\infty(\mathbb{D})$ and multiindex $i \in \mathbb{N}_0^2$, there exist $C > 0$ and $q \in \mathbb{N}_0$ such that

$$\left(f^{(i)}(z)\right)^2 \leqslant C(1 + x^{2q} + y^{2q}).$$

So, we can write the estimate

$$\begin{aligned}
\sqrt{\mathbb{E}\left(f^{(i)}(z_\theta)\right)^2} &\leqslant C\sqrt{1 + \mathbb{E}(\hat{X}_h^z)^{2q} + \mathbb{E}(\hat{Y}^h)^{2q}} \\
&\leqslant C\sqrt{1 + y\hat{R}_{2q}(y, \sigma^2 h) + \hat{C}(1 + |z|^{2q})} \\
&\leqslant C\sqrt{1 + |z|^{2q}} \\
&\leqslant C(1 + |z|^q),
\end{aligned}$$

where (here and below) the constant C varies from line to line. Now, using Lemma 4.3 and the Cauchy–Schwarz inequality, for all $n, m \in \mathbb{N}_0$ such that $n + m = 4$ and any multiindex $i = (n, m) \in \mathbb{N}_0^2$, we have

$$\begin{aligned}
&\mathbb{E}\left(f^{(i)}(z_\theta)(\hat{X}_h^z - x)^m(\hat{Y}_h^y - y)^n\right) \\
&\leqslant \sqrt{\mathbb{E}\left(f^{(i)}(z_\theta)\right)^2}\sqrt{\mathbb{E}(\hat{X}_h^z - x)^{2m}}\sqrt{\mathbb{E}(\hat{Y}_h^y - y)^{2n}} \\
&\leqslant C\sqrt{1 + |z|^{2q}}\sqrt{h^m(1 + |z|^m)}\sqrt{y(\sigma^2 h)^n \hat{P}_n(y, \sigma^2 h)} \\
&\leqslant C\sqrt{h^{n+m}}\sqrt{1 + |z|^{2q}}\sqrt{1 + |z|^m}\sqrt{1 + |z|^n}
\end{aligned}$$

$$\leqslant Ch^2\sqrt{1+|z|^{2q+n+m}} \leqslant Ch^2(1+|z|^{q+2}),$$

which proves Eq. (4.2.6). $\square$

Proposition 4.4. Let $\hat{X}_h^z$ be defined as in Eq. (4.2.8) and $\hat{Y}_h^y$ as in Eq. (4.2.7). Then the approximation $\hat{Z}_t^z := (\hat{X}_t^z, \hat{Y}_t^y)$, $t = kh$, $h = [T/N]$, $N \in \mathbb{N}$, $k = 0, 1, \ldots, N$, defined by the one-step scheme $\hat{Z}_h^z := (\hat{X}_h^z, \hat{Y}_h^y)$ has uniformly bounded moments on the whole time interval $[0, T]$, that is, for all $q \in \mathbb{N}$,

$$\sup_{N\in\mathbb{N}} \sup_{0\le k\le N} \mathbb{E}|\hat{Z}_{kh}^z|^q < +\infty. \tag{4.2.13}$$

Proof. We will prove the equivalent relation

$$\sup_{N\in\mathbb{N}} \sup_{0\le k\le N} \mathbb{E}\Big(|\hat{X}_{kh}^z|^q + \big(\hat{Y}_{kh}^y\big)^q\Big) < +\infty \tag{4.2.14}$$

for all $q \in \mathbb{N}$. Note that it suffices to check the latter for sufficiently large q. By the markovity of the approximation and Proposition 1.5 of [3], in turn, it suffices to prove that, for any $q \in \mathbb{N}$ and any $h_0 > 0$, there exists a constant C_q such that

$$\mathbb{E}\Big(|\hat{X}_h^z|^{2q} + \big(\hat{Y}_h^y\big)^{2q}\Big) \le (|x|^{2q} + y^{2q})(1 + C_q h) + C_q h, \quad 0 < h \le h_0. \tag{4.2.15}$$

From Corollary 2 of [27] we know that there exists a constant C such that, for all $h_0 > 0$ and $q \in \mathbb{N}$,

$$\mathbb{E}\big(\hat{Y}_h^y\big)^{2q} \leqslant y^{2q}(1 + Ch) + Ch, \quad 0 < h \leqslant h_0. \tag{4.2.16}$$

From Lemma 4.3 we know that there exists some constant $\tilde{C}_n$ such that

$$\mathbb{E}(\hat{X}_h^z - x)^n \leqslant \tilde{C}_n h(1 + y^{[\frac{n}{2}]}), \quad 0 < h < h_0.$$

Therefore,

$$\begin{aligned}
\mathbb{E}\big|\hat{X}_h^z\big|^{2q} &= \mathbb{E}\big(\hat{X}_h^z\big)^{2q} \\
&= \mathbb{E}\Big(x + (\hat{X}_h^z - x)\Big)^{2q} \\
&= x^{2q} + \sum_{n=1}^{2q} \tbinom{2q}{n} x^{2q-n} \mathbb{E}\Big(\hat{X}_h^z - x\Big)^n \\
&\le x^{2q} + h\sum_{n=1}^{2q} \tilde{C}_n \tbinom{2q}{n} |x|^{2q-n}(1 + y^{[\frac{n}{2}]}) \\
&= x^{2q} + Chx^{2q} + h\Big(-Cx^{2q} + \sum_{n=0}^{2q-1} \tilde{C}_n \tbinom{2q}{n} |x|^n + \sum_{n=1}^{2q} \tilde{C}_n \tbinom{2q}{n} |x|^{2q-n} y^{[\frac{n}{2}]}\Big)
\end{aligned}$$

$$= x^{2q}(1+Ch) + h\Big(-Cx^{2q} + \sum_{n=0}^{2q-1} a_n|x|^n + \sum_{n=1}^{2q} b_n|x|^{2q-n}y^{[\frac{n}{2}]}\Big), \quad 0 < h \le h_0, \tag{4.2.17}$$

with some constants $a_i > 0$, $i = 0, 1, \ldots, 2q-1$, and $b_i > 0$, $i = 1, \ldots, 2q$.

Joining estimates (4.2.16) and (4.2.17), we have

$$\mathbb{E}\Big(|\hat{X}_h^z|^{2q} + \big(\hat{Y}_h^y\big)^{2q}\Big) \le (|x|^{2q} + y^{2q})(1+2Ch)$$
$$+ h\Big(-C\big(|x|^{2q} + y^{2q}\big) + \sum_{n=0}^{2q-1} a_n|x|^n + \sum_{n=1}^{2q} b_n|x|^{2q-n}y^{[\frac{n}{2}]}\Big), \quad h \le h_0.$$

Since there exists a constant $\overline{C} = \overline{C}(q)$ such that

$$\frac{x^n y^m}{x^{2q} + y^{2q}} \leqslant \frac{(x+y)^{n+m}}{x^{2q} + y^{2q}} \leqslant \overline{C}\frac{(x+y)^{n+m}}{(x+y)^{2q}}, \quad x, y > 0,$$

for any $m, n \in \mathbb{N}_0$ such that $m + n < 2q$, we have that

$$\frac{x^m y^n}{x^{2q} + y^{2q}} \to 0, \quad \text{as } x, y > 0,\ x + y \to \infty. \tag{4.2.18}$$

Therefore,

$$A_{2q} := \sup_{x\in\mathbb{R},\, y>0} \Big(-C\big(|x|^{2q} + y^{2q}\big) + \sum_{n=0}^{2q-1} a_n|x|^n + \sum_{n=1}^{2q} b_n|x|^{2q-n}y^{[\frac{n}{2}]}\Big) < +\infty.$$

Denoting $C_q := \max\{2C, A_{2q}\}$, we finally have

$$\mathbb{E}\Big(|\hat{X}_h^z|^{2q} + \big(\hat{Y}_h^y\big)^{2q}\Big) \le (|x|^{2q} + y^{2q})(1 + C_q h) + C_q h, \quad 0 < h \le h_0,$$

that is, estimate (4.2.15) holds for all $q \in \mathbb{N}$, as required. □

Summarizing Propositions 3.2, 4.1, and 4.4 we get the following theorem.

Theorem 4.5. Let a one-step approximation $\bar{Z}_h^z = (\bar{X}_h^z, \bar{Y}_h^y)$ of Eq. (1.2.1) be constructed as follows:

(1) Let $\hat{Y}_h^y$ be a random variable taking the values

$$y_{1,2} = y + \sigma^2 h \pm \sqrt{(y + \sigma^2 h)\, \sigma^2 h} \quad \text{with corresponding probabilities } p_{1,2} = \frac{y}{2y_{1,2}}.$$

(2) Let $\tilde{X}_h^z$ be a random variable, independent of $\hat{Y}_h^y$, taking the values

$$x_{1,2} = x \pm \sqrt{yh} \quad \text{with probabilities } \frac{1}{2}.$$

(3) Let random variable $\hat{X}_h^z$ be defined by

$$\hat{X}_h^z := x + \sqrt{1-\rho^2}(\tilde{X}_h^z - x) + \frac{\rho}{\sigma}(\hat{Y}_h^y - y).$$

(4) Finally, having $\hat{Z}_h^z = (\hat{X}_h^z, \hat{Y}_h^y)$, define the one-step discretization scheme $\bar{Z}_h^z = (\bar{X}_h^z, \bar{Y}_h^y)$ by Eq. (4.1.1), that is,

$$\begin{aligned}\bar{Z}_h^z &= \bar{Z}(z,h) := D\left[\hat{Z}(z,h), h\right] \\ &= \begin{pmatrix} \hat{X}(z,h) + (r - \frac{1}{2}\theta)h + \frac{1}{2k}(\mathsf{e}^{-kh} - 1)\big(\hat{Y}(y,h) - \theta\big) \\ \hat{Y}(y,h)\mathsf{e}^{-kh} + \theta(1 - \mathsf{e}^{-kh}) \end{pmatrix}. \end{aligned} \tag{4.2.19}$$

Then the one-step discretization scheme (4.2.19) defines a strongly potential weak first-order approximation of the log-Heston system (1.2.1).

Proof. It remains to check that the approximation constructed also has uniformly bounded moments of all orders. Using expressions (3.2.3) for $D(z,t) = (D_1^z(t), D_2^y(t))$ we will show that, for every $q \in \mathbb{N}$, there exists a constant $\tilde{C}_q$, independent of $z \in \mathbb{D}$ and h, such that

$$|D_1^z(h)|^q + \big(D_2^y(h)\big)^q \le (|x|^q + y^q)(1 + \tilde{C}_q h) + \tilde{C}_q h, \quad z \in \mathbb{D}, \quad 0 < h \le h_0.$$

Since there exists a constant C, depending on h_0 and k, such that $1 - \mathsf{e}^{-kh} < Ch$ for $0 < h \le h_0$, we have

$$\begin{aligned}(D_2^y(h))^q &= \big(y\mathsf{e}^{-kh} + \theta(1 - \mathsf{e}^{-kh})\big)^q \\ &\leqslant (y(1 + Ch) + Ch\theta)^q \\ &\leqslant y^q(1 + \tilde{C}h) + \sum_{n=1}^{q} \tbinom{q}{n} y^{q-n}(Ch\theta)^n(1 + Ch)^n \\ &\leqslant y^q(1 + \tilde{C}h) + h\sum_{n=0}^{q-1} a_n y^n, \quad 0 < h \le h_0,\end{aligned}$$

with some $a_i > 0$, $i = 0, \dots, q-1$, and from this we can further estimate:

$$\begin{aligned}(D_2^y(h))^q &\leqslant y^q(1 + \tilde{C}h) + h\sum_{n=0}^{q-1} a_n y^n \\ &= y^q(1 + 2\tilde{C}h) + h\left(-\tilde{C}y^q + \sum_{n=0}^{q-1} a_n y^n\right), \quad 0 < h \le h_0. \end{aligned} \tag{4.2.20}$$

Similarly, for all $h, k > 0$ and some constant $\overline{C}$, depending on r, θ, q, and h_0, we have

$$\begin{aligned}
|D_1^z(h)|^q &= \left| x + (r - \frac{1}{2}\theta)h + \frac{1}{2k}(1 - \mathsf{e}^{-kh})(\theta - y) \right|^q \\
&\leqslant \big(|x| + h(\overline{C} + y) \big)^q \\
&= |x|^q + h \sum_{n=1}^{q} \tbinom{q}{n} |x|^{q-n} h^{n-1} (\overline{C} + y)^n \\
&\leq |x|^q + h \left(\sum_{n=1}^{q} \tbinom{q}{n} |x|^{q-n} h^{n-1} \hat{C} + \sum_{n=1}^{q} \tbinom{q}{n} |x|^{q-n} h^{n-1} \hat{C} y^n \right) \\
&\leq |x|^q + h \left(\sum_{n=0}^{q-1} b_n |x|^n + \sum_{n=1}^{q} b_n |x|^{q-n} y^n \right) \\
&= |x|^q + h \left(b_q y^q + \sum_{n=0}^{q-1} b_n |x|^n + \sum_{n=1}^{q-1} b_n |x|^{q-n} y^n \right) \\
&= |x|^q (1 + \tilde{C}h) + h \left(-\tilde{C}|x|^q + b_q y^q + \sum_{n=0}^{q-1} b_n |x|^n + \sum_{n=1}^{q-1} b_n |x|^{q-n} y^n \right)
\end{aligned} \tag{4.2.21}$$

with some $b_i > 0$, $i = 0, \ldots, q$.

Joining estimates (4.2.20) and (4.2.21) and denoting $a_q = b_q$, for all $0 < h < h_0$, we have

$$\begin{aligned}
|D_1^z(h)|^q + \big(D_2^y(h)\big)^q \leq (|x|^q + y^q)(1 + 2\tilde{C}h) + \\
+ h\Big(- \tilde{C}\big(|x|^q + y^q\big) + \sum_{n=0}^{q} a_n y^n + \sum_{n=0}^{q-1} b_n |x|^n + \sum_{n=1}^{q-1} b_n |x|^{q-n} y^n \Big).
\end{aligned}$$

Using (4.2.18), we notice that

$$A_q := \sup_{x \in \mathbb{R}\ y > 0} \Big(- \tilde{C}\big(|x|^q + y^q\big) + \sum_{n=0}^{q} a_n y^n + \sum_{n=0}^{q-1} b_n |x|^n + \sum_{n=1}^{q-1} b_n |x|^{q-n} y^n \Big) < +\infty.$$

Finally, denoting $C_q := \max\{2\tilde{C}, A_q\}$, we get

$$|D_1^z(h)|^q + \big(D_2^y(h)\big)^q \leq (|x|^q + y^q)(1 + \tilde{C}_q h) + \tilde{C}_q h, \quad 0 < h \leq h_0, \tag{4.2.22}$$

This shows that composition (4.2.19) “preserves” the estimate of type (4.2.15), that is, for every $q \in \mathbb{N}$, there exists a constant $\bar{C}_q$, independent of h, such that

$$\mathbb{E}\big(|\bar{X}_h^z|^q + \big(\bar{Y}_h^y\big)^q\big) \leq (|x|^q + y^q)(1 + \bar{C}_q h) + \bar{C}_q h, \quad 0 < h \leq h_0,$$

which, as before, suffices for the uniform boundedness of all moments of $\bar{Z}_h^z = (\bar{X}_h^z, \bar{Y}_h^y)$. □

We finally have shown that $\hat{Z}^h = (\hat{X}^z_h, \hat{Y}^y_h)$ is a first-order potential approximation of the stochastic part (3.2.2) of the log-Heston model (1.2.1) and, by Propositions 3.2, the composition (4.1.1) defines a potential first-order approximation $\overline{Z}^h$ of the log-Heston model (1.2.1).

4.3 Algorithm

To emphasize the simplicity of the algorithm, we summarize the simulation step from $(\bar{X}_{ih}, \bar{Y}_{ih}) = (x, y)$ to $(\bar{X}_{(i+1)h}, \bar{Y}_{(i+1)h})$:

1. Draw a uniform random number U in the interval $(-1, 1)$, and let $U_1 := |U|$.

2. Given the values x and y, generate a random variable $\tilde{X}$ taking values x_1 and x_2 defined by (4.2.9) with probabilities $1/2$ as follows: if $U < 0$, then set $\tilde{X} := x_1$, otherwise $\tilde{X} := x_2$.

3. Generate a random variable $\hat{Y}$ taking values y_1 and y_2 defined by (4.2.7) with probabilities $p_1 = y/(2y_1)$ and $p_2 = 1 - p_1$, respectively: if $U_1 < p_1$, set $\hat{Y} := y_1$, otherwise $\hat{Y} := y_2$.

4. Calculate (see (4.2.8))
$$\hat{X} := x + \sqrt{1-\rho^2}(\tilde{X} - x) + \frac{\rho}{\sigma}(\hat{Y} - y).$$

5. Calculate (see (4.1.1))
$$\begin{aligned} \bar{X}_{(i+1)h} &= \hat{X} + (r - \tfrac{1}{2}\theta)h + \tfrac{1}{2k}(\mathsf{e}^{-kh} - 1)(\hat{Y} - \theta), \\ \bar{Y}_{(i+1)h} &= \hat{Y}\mathsf{e}^{-kh} + \theta(1 - \mathsf{e}^{-kh}). \end{aligned}$$

Having generated all values $\bar{X}_{ih}$, $i = 0, 1, 2, \ldots$, set $\overline{S}_{ih} := \exp\{\bar{X}_{ih}\}$, $i = 0, 1, 2, \ldots$.

4.4 Simulation examples

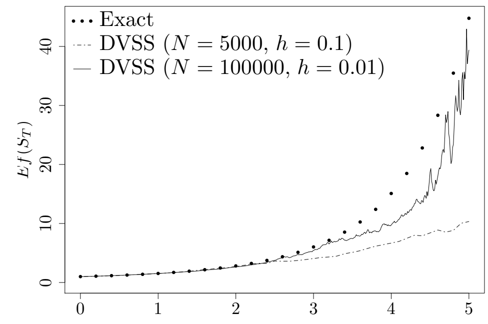
In this section, we give three simulation examples of our approximation scheme, which we call DVSS (discrete-variable split-step) for short.

Third moment

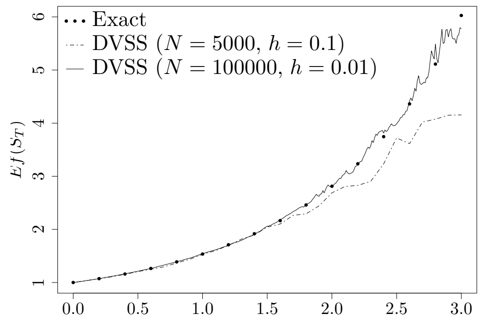
First, we apply the DVSS approximation to the test function $f(x) = x^3$. Although, at first sight, the test function seems to be very simple, but, as is shown in [6] and [11], depending on the set of parameters, the (true) expectations can "explode," providing the possibility to try our method in an "extreme" case, not covered by theoretical results. Unfortunately, neither [6] nor [11] provide explicit formula of $\mathbb{E}S_t^3$; therefore, we were forced to derive the formula ourselves (see Chapter 7). The set of parameters $S_0 = 1$, $Y_0 = 0.09$, $r = 0.02$, $\theta = 0.12$, $\sigma = 0.3$, $k = 1$, $\rho = 0.5$ forces the third moment $\mathbb{E}S_t^3$ to converge to infinity (at about $t = 10$), whereas the set of parameters $S_0 = 1$, $Y_0 = 0.09$, $r = 0.02$, $\theta = 0.12$, $\sigma = 0.2$, $k = 3$, $\rho = -0.5$ provides the third moment without "explosion." We plot approximate expectations by averaging over 5,000 samples with step $h = 0.1$ and by averaging over 100,000 samples with step $h = 0.01$ in one plot, comparing them with true expectations (see, e.g., [29], [6], [11]), marked by bullets "•" at discrete time points with step size 0.2. As can be seen from the plots in Fig. 4.1, in the explosion case, our approximation is rather satisfactory at the beginning but tends to lag behind the true values as time approaches the moment of explosion.

Option pricing

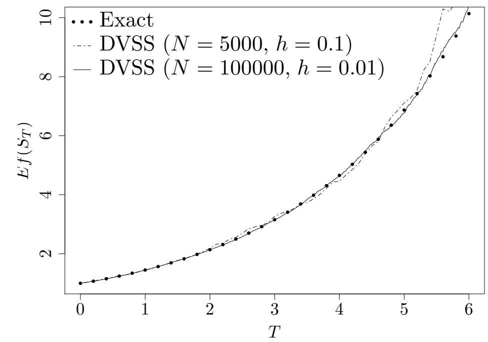
Here we illustrate the DVSS approximation in calculation of European call and put options prices, that is, we apply it to the test functions $f_1(x) = \max\{x - K, 0\}$ and $f_2(x) = \max\{K - x, 0\}$, respectively. For both types of options, we use the same set of fixed parameters taken from [12]: $K = 100$, $r = 0.02$, $Y_0 = 0.09$, $k = 3$, $\theta = 0.12$, $\sigma = 0.2$, and $\rho = -0.5$, together with $S_0 = 80; 100; 120$. To show the dynamics of approximation, we present two approximations of the option price $\mathbb{E}e^{-rT} f_i(S_T)$ in one plot as functions of maturity T (Fig. 4.2). One graph (dashed line) of approximate prices is obtained by averaging over 5,000 samples with step $h = 0.1$, and the other (solid line) by averaging over 100,000 samples with step $h = 0.01$. We compare the obtained values of option prices with the exact option prices calculated using the MATLAB code *HestonFFTVanilla* from [9] (downloaded from www.quantlet.de) and plot the exact values at discrete time points with step size 0.2, denoted by bullets "•".



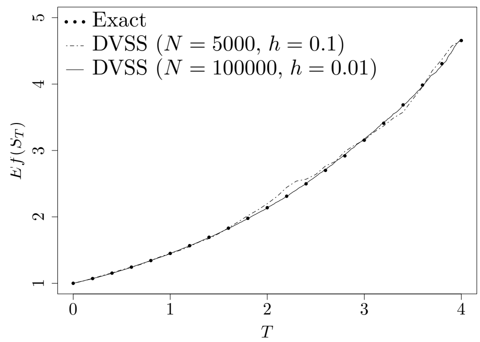
(a) With explosion, $T = 5$



(b) With explosion, $T = 3$



(c) Without explosion, $T = 6$



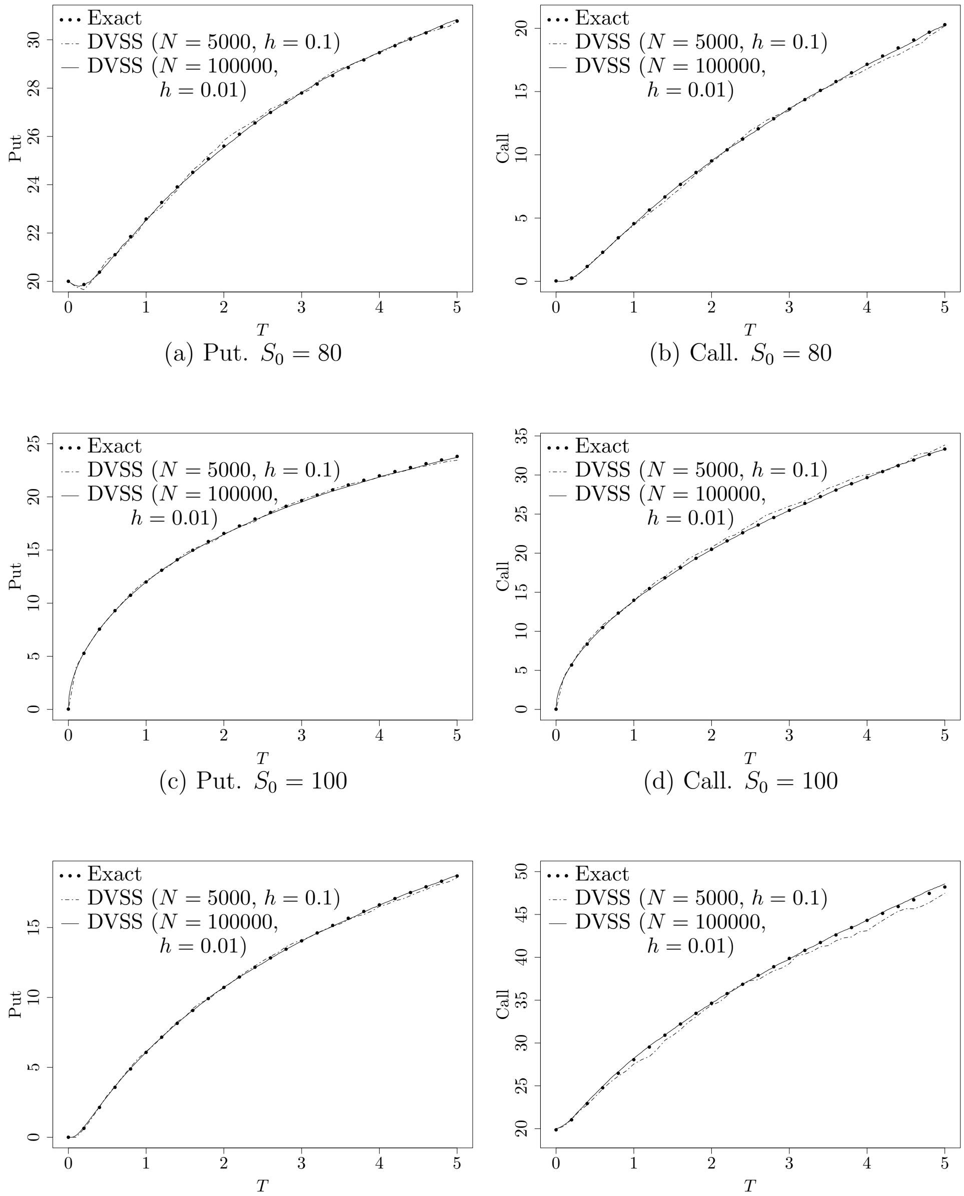
(d) Without explosion, $T = 4$

Figure 4.1: DVSS approximation of $\mathbb{E}S_t^3$ in the Heston model. Fixed parameters: $S_0 = 1$, $Y_0 = 0.09$, $r = 0.02$, $\theta = 0.12$. In the case of $\sigma = 0.3$, $k = 1$, and $\rho = 0.5$, $\mathbb{E}S_t^3$ "explodes" near $T = 10$, whereas $\sigma = 0.2$, $k = 3$, and $\rho = -0.5$ provide $\mathbb{E}S_t^3$ without "explosion".

Comparison with the QE/QE-M (Andersen [5]) and ALF$_2$ (Alfonsi [3]) approximations

Finally, we compare the DVSS scheme with the QE (quadratic-exponential) and the QE-M (QE with martingale correction) schemes of Andersen [5] using the same parameters sets as in [5] and with the second-order scheme proposed by Alfonsi [3], using the same parameters sets as in [3]. To demonstrate better the effect of the step size and number of trajectories, we choose different numbers of trajectories (N) and different steps (h) in each case (see Table 4.1).

We plot the cases from [5] (i.e., cases I, II, and III) in one graph (see Fig. 4.3) setting the time to maturity $T = 10$ and the cases from [3] (i.e., cases IV and V) in another, setting $T = 2$ (see Fig. 4.4). We take $S_0 = 100$ for all the cases, $r = 0$ for cases I, II, and III, and $r = 0.02$ for cases IV and V. As before, in both graphs, the exact option prices are marked by bullets "•" (at discrete time points with step size 1 in cases I, II, and III and with step size 0.2 in cases IV and V). As one



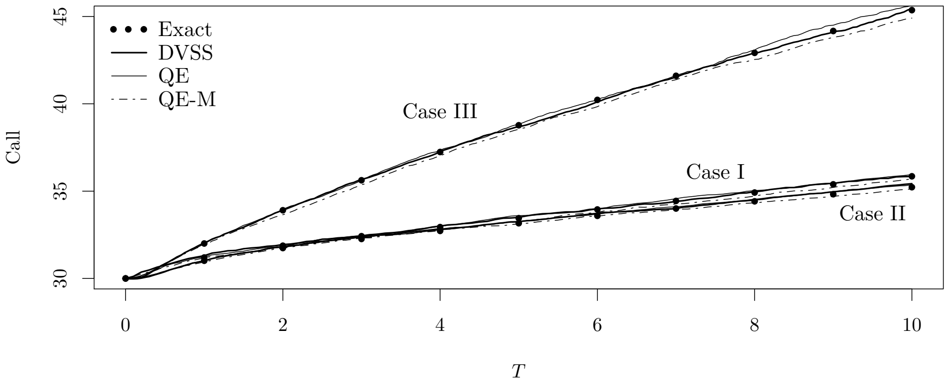
(a) Put. $S_0 = 80$ (b) Call. $S_0 = 80$

(c) Put. $S_0 = 100$ (d) Call. $S_0 = 100$

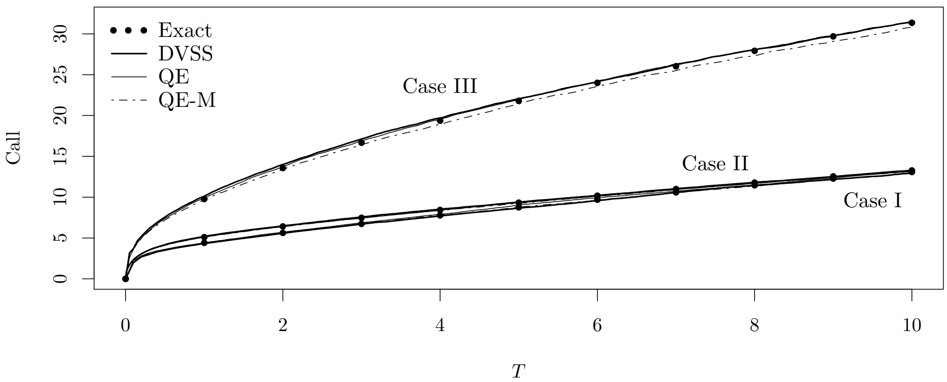
(e) Put. $S_0 = 120$ (f) Call. $S_0 = 120$

Figure 4.2: European option prices in the Heston stochastic volatility model for different asset prices S_0 with fixed parameters $K = 100$, $r = 0.02$, $Y_0 = 0.09$, $k = 3$, $\theta = 0.12$, $\sigma = 0.2$, and $\rho = -0.5$.

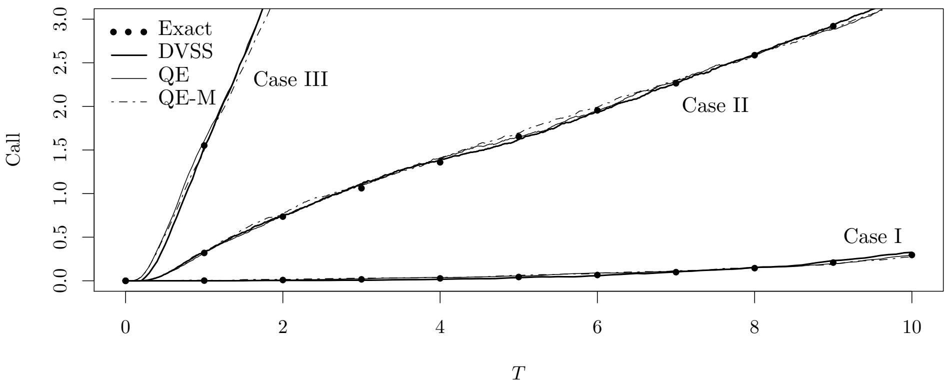
can see from Figs. 4.3 and 4.4, the DVSS approximation behaves rather similarly to the QE(-M) and Alfonsi schemes (and in some cases, even better, for example, in case III with $K = 70$, and case V with $K = 80$). To have a clearer picture, we further provide figures, showing the accuracies of approximations as functions of the step size h. As before, we split the cases into two graphs. Figure 4.5 shows the accuracy of approximations of call price at the time of maturity $T = 10$ in cases I, II, III (with strike prices 70, 100, 140), and Fig. 4.6 shows the put prices at $T = 1$ in cases IV and V (with strike prices 80, 100, 120). In Fig. 4.5, the approximate



(a) Strike price 70

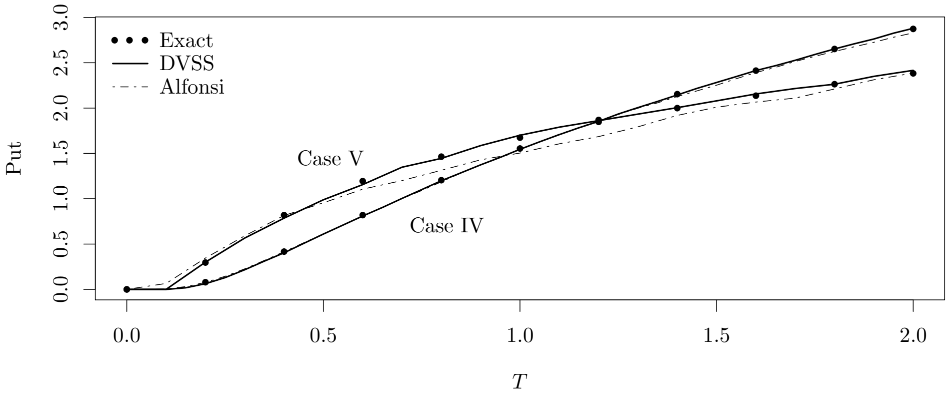


(b) Strike price 100

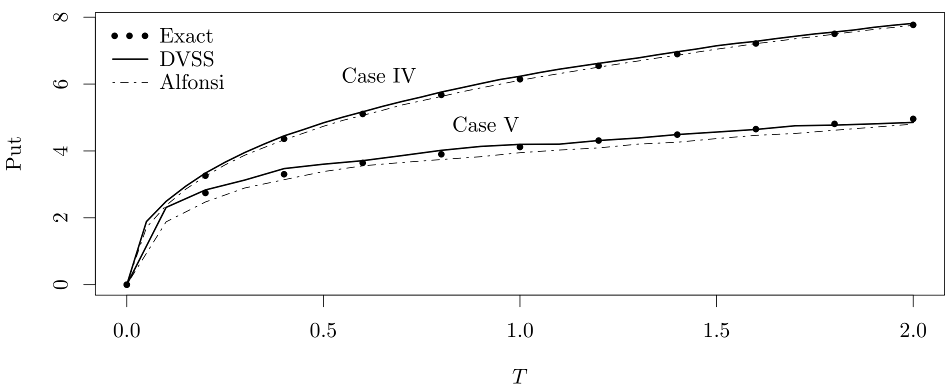


(c) Strike price 140

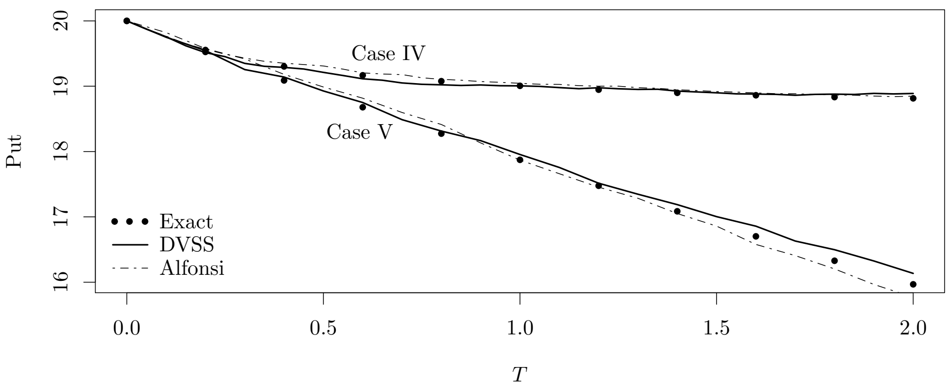
Figure 4.3: Comparison of DVSS and QE(-M) approximations of European option prices in the Heston's stochastic volatility model for different parameters and strike prices (70, 100, 140) with fixed parameters $S_0 = 100$ and $r = 0$.



(a) Strike price 80



(b) Strike price 100



(c) Strike price 120

Figure 4.4: Comparison of DVSS and Alfonsi approximations of European option prices in the Heston's stochastic volatility model for different parameters and strike prices (80, 100, 120) with fixed parameter $S_0 = 100$.

	Case I	Case II (see [5])	Case III	Case IV	Case V (see [3])
σ	1	0.9	1	0.4	1
κ	0.2	0.3	1	0.5	0.5
ρ	-0.9	-0.5	-0.3	-0.5	-0.8
θ	0.04	0.04	0.09	0.04	0.04
r	0	0	0	0.02	0.02
Y_0	0.04	0.04	0.09	0.04	0.04
T	10	10	10	2	2

Table 4.1: Test cases for comparison of DVSS with QE(-M) and Alfonsi schemes. In all cases, $S_0 = 100$.

	Case I	Case II (see [5])	Case III	Case IV	Case V (see [3])
h	0.1	0.01	0.05	0.05	0.1
N	10000	100000	70000	100000	10000

Table 4.2: Step (h) and expiration time (T) parameters used in each test case for comparison of DVSS with QE(-M) and Alfonsi schemes.

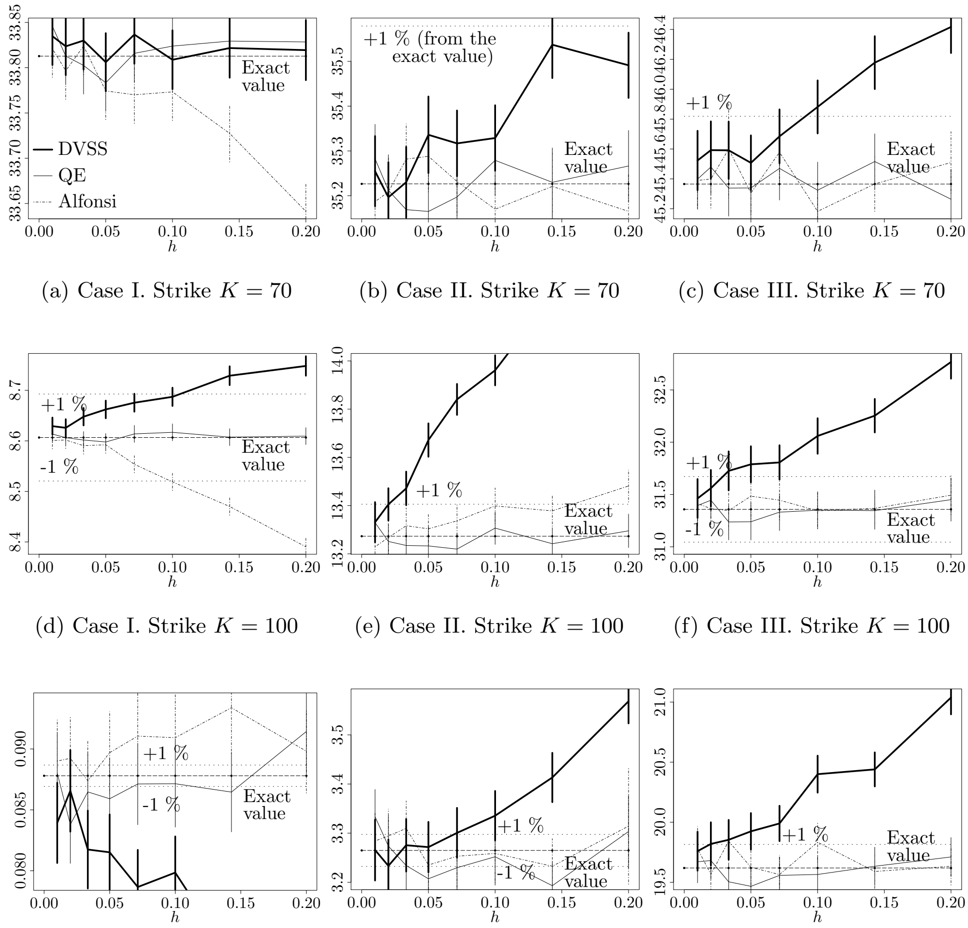
prices are obtained by averaging over 10^6 samples, and in Fig. 4.6 by averaging over 10^8 samples. Dashed lines in both figures represent the exact values of call (Fig. 4.5) or put (Fig. 4.6) prices, whereas dotted lines show the $\pm 1\%$ deviations from the exact value. If one or both dotted lines are missing, this means that all the values in the plot are in the range of -1%, or $+1\%$, or both from the exact value (depending on which lines are missing).

Finally, we take one of the cases from each graph (we choose cases I and V) and examine the errors in the plots more closely. To this end, we calculate the bias

$$\varepsilon_h = \mathbb{E}\mathsf{e}^{-rT}\max(S_T - K, 0) - \mathbb{E}\mathsf{e}^{-rT}\max(\overline{S}_T^h - K, 0),$$

that is, the difference between the exact price of call option and its evaluation using an approximation scheme, and put the results in Tables 4.3 and 4.4. The values in the table are written in bold when the bias is within the window of $\pm 1\%$ from the exact value. We also put the half-length of the 95% confidence interval in parenthesis and mark it with an asterisk if the bias is less than the half-length of the 95% confidence interval, that is, when the error of approximation is not statistically significant.

Examining the figures and Tables 4.3 and 4.4, we can conclude that although, in most cases, the DVSS scheme has a slightly higher bias than the QE and Alfonsi schemes but remains within the window of $\pm 1\%$ (from the exact value), which is a sufficient accuracy for practical applications. Moreover, for small steps (for example, $h = 0.01$), the bias becomes quite similar to that of the QE and Alfonsi schemes.



(a) Case I. Strike $K = 70$ (b) Case II. Strike $K = 70$ (c) Case III. Strike $K = 70$

(d) Case I. Strike $K = 100$ (e) Case II. Strike $K = 100$ (f) Case III. Strike $K = 100$

(g) Case I. Strike $K = 140$ (h) Case II. Strike $K = 140$ (i) Case III. Strike $K = 140$

Figure 4.5: Approximations of call price in function of $h = 1/n$ in cases I, II, and III for different strike prices (70, 100, 140) with fixed parameters $S_0 = 100$, $r = 0$, and $T = 10$. Approximate prices were obtained by averaging over 10^6 samples.

To complete, we compare the computation times of DVSS, Andersen's QE, and Alfonsi schemes. For calculations, we used a laptop PC with Intel(R) CORE(TM) i7-3610QM 2.30 GHz processor and 8 GB RAM and C programming language (the graphs were plotted by using the R programming language with graphics package TikzDevice). The results are listed in Table 4.5. The numbers on the left side of the table show the averages over the runs of 100 000 trajectories with step $h = 0.01$

	DVSS	QE	Alfonsi
h		$K = 70$	
1/5	**-0.00663** (0.03316*)	**-0.01567** (0.03171*)	**0.17145** (0.03176)
1/7	**-0.00892** (0.03275*)	**-0.01661** (0.03173*)	**0.08536** (0.03173)
1/10	**0.00399** (0.03249*)	**-0.01106** (0.03171*)	**0.03948** (0.03173)
1/14	**-0.02387** (0.03230*)	**-0.00311** (0.03168*)	**0.04268** (0.03170)
1/20	**0.00668** (0.03208*)	**0.02940** (0.03170*)	**0.03840** (0.03175)
1/30	**-0.01720** (0.03199*)	**0.01027** (0.03170*)	**-0.01109** (0.03172*)
1/50	**-0.01083** (0.03185*)	**-0.00006** (0.03164*)	**0.01610** (0.03173*)
1/100	**-0.02197** (0.03179*)	**-0.03300** (0.03164)	**-0.00809** (0.03170*)
		$K = 100$	
1/5	-0.14155 (0.01927)	**-0.00273** (0.01680*)	0.21623 (0.01764)
1/7	-0.12227 (0.01860)	**-0.00074** (0.01684*)	0.13652 (0.01732)
1/10	**-0.08051** (0.01804)	**-0.01024** (0.01688*)	0.08783 (0.01707)
1/14	**-0.06883** (0.01771)	**-0.00717** (0.01688*)	**0.05290** (0.01697)
1/20	**-0.05555** (0.01747)	**0.00880** (0.01675*)	**0.01431** (0.01696*)
1/30	**-0.04118** (0.01723)	**0.00496** (0.01678*)	**0.01618** (0.01687*)
1/50	**-0.01918** (0.01702*)	**0.00024** (0.01673*)	**0.00501** (0.01687*)
1/100	**-0.02250** (0.01696)	**-0.00675** (0.01676*)	**0.00588** (0.01680*)
		$K = 140$	
1/5	0.01994 (0.00253)	-0.00360 (0.00346)	-0.00198 (0.00344*)
1/7	0.01503 (0.00274)	0.00138 (0.00327*)	-0.00555 (0.00375)
1/10	0.00798 (0.00297)	**0.00069** (0.00356*)	-0.00309 (0.00356*)
1/14	0.00917 (0.00305)	**0.00071** (0.00334*)	-0.00323 (0.00358*)
1/20	0.00632 (0.00311)	0.00192 (0.00332*)	-0.00186 (0.00336*)
1/30	0.00609 (0.00320)	0.00136 (0.00330*)	**0.00045** (0.00334*)
1/50	0.00127 (0.00335*)	0.00404 (0.00319)	-0.00140 (0.00340*)
1/100	0.00390 (0.00330)	**-0.00019** (0.00337*)	-0.00118 (0.00337*)

Table 4.3: Estimated bias in the test case I for different strike prices ($K = 70$, $K = 100$, and $K = 140$). Values are written in bold when the bias is in the window of $\pm 1\%$ from the exact value. Numbers in parentheses show the half-length of the 95% confidence interval and are marked with an asterisk if the bias is within this interval.

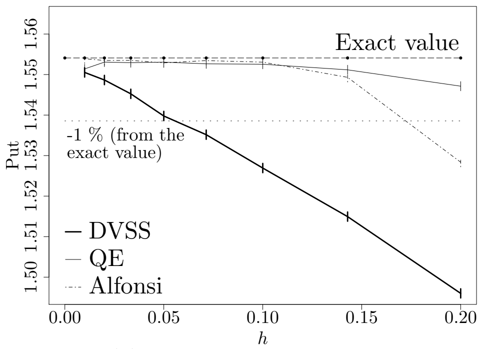
and time to maturity $T = 10$ in all the cases. One can see that the DVSS scheme is much faster in all the cases (see the computational times measured relatively to the DVSS scheme on the right of the table). Note that we tried to optimize the QE and Alfonsi approximation code by caching parts of computations not depending on each time step.

	DVSS	QE	Alfonsi
h		$K = 80$	
1/5	0.06801 (0.00158)	0.01851 (0.00158)	0.08404 (0.00157)
1/7	0.05401 (0.00158)	**0.01264** (0.00158)	0.04318 (0.00158)
1/10	0.03943 (0.00158)	**0.00924** (0.00158)	0.02156 (0.00158)
1/14	0.02929 (0.00158)	**0.00628** (0.00158)	**0.01235** (0.00158)
1/20	0.02016 (0.00158)	**0.00413** (0.00158)	**0.00645** (0.00158)
1/30	**0.01195** (0.00158)	**0.00302** (0.00158)	**0.00262** (0.00158)
1/50	**0.00796** (0.00158)	**0.00386** (0.00158)	**-0.00008** (0.00158*)
1/100	**0.00470** (0.00158)	**0.00034** (0.00158*)	**-0.00195** (0.00159)
		$K = 100$	
1/5	-0,21043 (0,00245)	**0,02198** (0,00249)	0,11304 (0,00244)
1/7	-0,13670 (0,00247)	**0,01264** (0,00250)	0,10136 (0,00247)
1/10	-0,08664 (0,00248)	**0,00994** (0,00250)	0,08052 (0,00249)
1/14	-0,05148 (0,00249)	**0,00797** (0,00250)	0,06121 (0,00250)
1/20	**-0,03059** (0,00249)	**0,00621** (0,00250)	**0,03786** (0,00250)
1/30	**-0,01844** (0,00250)	**0,00560** (0,00250)	**0,02034** (0,00250)
1/50	**-0,00999** (0,00250)	**0,00657** (0,00250)	**0,00475** (0,00250)
1/100	**-0,00333** (0,00250)	**0,00099** (0,00250*)	**0,00030** (0,00251*)
		$K = 120$	
1/5	**0,11991** (0,00316)	**0,00786** (0,00305)	**0,02607** (0,00305)
1/7	**0,10049** (0,00314)	**0,00217** (0,00305*)	**0,03004** (0,00306)
1/10	**0,08482** (0,00312)	**0,00491** (0,00305)	**0,01889** (0,00306)
1/14	**0,06114** (0,00310)	**0,00327** (0,00306)	**0,01053** (0,00306)
1/20	**0,04277** (0,00309)	**0,00297** (0,00306*)	**0,00418** (0,00306)
1/30	**0,02770** (0,00308)	**0,00495** (0,00306)	**0,00137** (0,00306*)
1/50	**0,01679** (0,00307)	**0,00721** (0,00305)	**0,00119** (0,00306*)
1/100	**0,01419** (0,00306)	**0,00338** (0,00306)	**-0,00351** (0,00306)

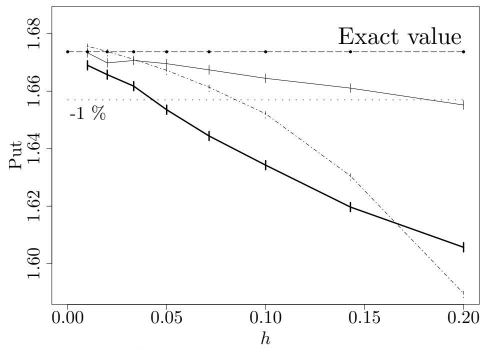
Table 4.4: Estimated bias in the test case V for different strike prices ($K = 80$, $K = 100$, and $K = 120$). Values are written in bold when the bias is in the window of ±1% from the exact value. Numbers in parentheses show the half-length of the 95% confidence interval and are marked with an asterisk if the bias is within this interval.

	100 000 trajectories in sec.			Relative to DVSS	
	DVSS	QE	Alfonsi	QE	Alfonsi
Case I	10.219	39.778	36.219	3.749	3.413
Case II	10.058	27.129	31.203	2.697	3.102
Case III	10.000	25.131	30.663	2.513	3.066
Case IV	10.049	26.546	31.279	2.642	3.113
Case V	10.333	35.690	34.218	3.454	3.311

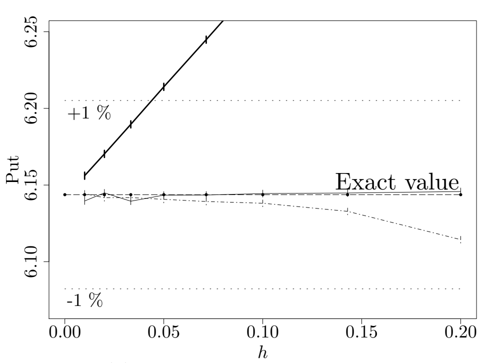
Table 4.5: Average and relative (to DVSS) computation times.



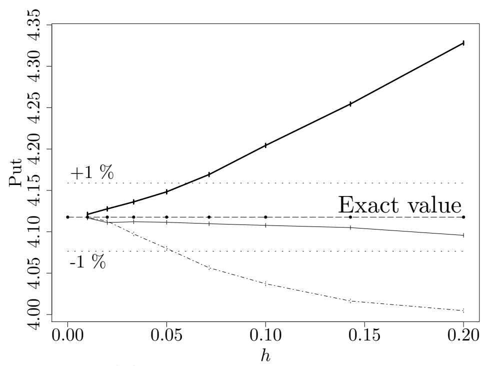
(a) Case IV. Strike $K = 80$



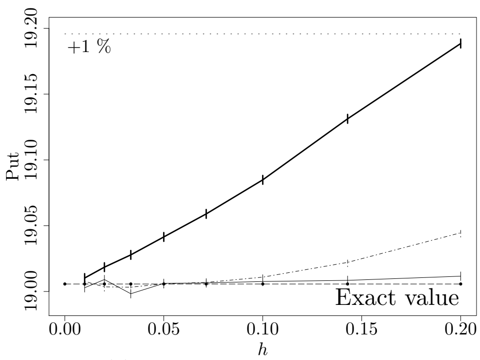
(b) Case V. Strike $K = 80$



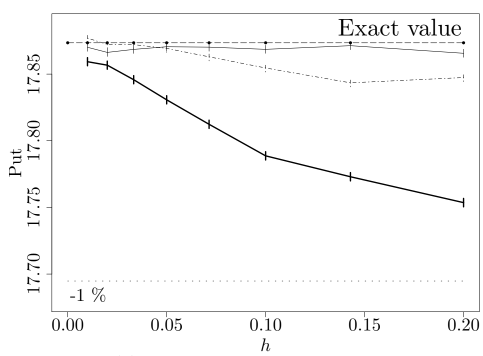
(c) Case IV. Strike $K = 100$



(d) Case V. Strike $K = 100$



(e) Case IV. Strike $K = 120$



(f) Case V. Strike $K = 120$

Figure 4.6: Approximations of put price in function of $h = 1/n$ in cases IV and V for different strike prices (80, 100, 140) with fixed parameters $S_0 = 100$ and $T = 1$. Approximate prices were obtained by averaging over 10^8 samples.

Chapter 5

Second-order approximation

In this chapter we construct a second-order approximation scheme for the solution of the Heston system (1.1.1). Similarly as before, first (in the first section of the chapter) we adapt a split-step technique, then, in the next section find an approximation scheme for the (split) stochastic part and finally, in the last section of the chapter, provide numerical simulations illustrating our scheme and comparing it to other known schemes of the Heston model.

5.1 A potential second-order approximation

As in the case for the first-order approximation, we start to construct our scheme using methods described in Section 3.2. First, in order to assure the positivity and avoid moment explosions we change our model (S_t, Y_t) to the log-Heston model $Z_t = (X_t := \log S_t, Y_t)$ (see (1.2.1)), then split the log-Heston model into the stochastic and deterministic parts. Solution of the latter $D(z,t) = (D_1^z(t), D_2^y(t))$ is known (see (3.2.3)) and we are left to construct a second-order discretization scheme for the solution of the stochastic part $\hat{Z}^h = \hat{Z}(z,h) = \Big(\hat{X}(z,h), \hat{Y}(y,h)\Big)$. Having done this, we can define the second-order split-step approximation for Z_t by the composition

$$\bar{Z}^h = \bar{Z}(z,h) = \begin{pmatrix} \bar{X}(z,h) \\ \bar{Y}(y,h) \end{pmatrix} := D\left(\tilde{Z}\big(D(z,h/2),h\big),h/2\right). \tag{5.1.1}$$

Proposition 3.2 assures that the composition (5.1.1) defines a potential second-order approximation of the log-Heston system (1.2.1) if we provide a potential second-order approximation scheme of the *stochastic part* of the log-Heston model equation.

5.2 A potential second-order approximation of the stochastic part

As before, we rewrite the stochastic part of system (1.2.1) as

$$\begin{cases} \mathrm{d}X_t = \sqrt{Y_t}\left(\rho\,\mathrm{d}W_t + \sqrt{1-\rho^2}\,\mathrm{d}B_t\right), \\ \mathrm{d}Y_t = \sigma\sqrt{Y_t}\,\mathrm{d}W_t, \end{cases} \tag{5.2.1}$$

where B is a Brownian motion independent of W.

Although this equation is simple enough for constructing the first-order (potential) weak approximation (see [22]), as it turns out, it is too difficult to arrive at any simple second-order (potential) weak approximation by discrete random variables.

The generator of such a stochastic part is

$$Lf(z) = \frac{y}{2}\left(\frac{\partial^2}{\partial x^2} + \sigma^2\frac{\partial^2}{\partial y^2}\right)f(z) + \rho\sigma y\frac{\partial^2 f}{\partial x\partial y}\,(z).$$

Therefore,

$$L^2 f(z) = \frac{y}{2}\left(\frac{\partial^2}{\partial x^2} + \sigma^2\frac{\partial^2}{\partial y^2}\right)Lf(z) + \rho\sigma y\frac{\partial^2 Lf}{\partial x\partial y}\,(z).$$

Straightforward and somewhat tedious calculation gives

$$\begin{aligned} L^2 f(z) = {} & \frac{y^2}{4}\frac{\partial^4}{\partial x^4}f(z) + \frac{y^2}{4}\sigma^2\frac{\partial^4}{\partial x^2\partial y^2}f(z) + \rho\sigma\frac{y^2}{2}\frac{\partial^4}{\partial x^3\partial y}f(z) \\ & + \frac{y}{2}\sigma^2\left(\frac{\partial^3}{\partial x^2\partial y} + \sigma^2\frac{\partial^3}{\partial y^3}\right)f(z) + y\sigma^3\rho\frac{\partial^3 f}{\partial x\partial y^2}\,(z) \\ & + \frac{y^2}{2}\sigma^2\left(\frac{1}{2}\left(\frac{\partial^4}{\partial x^2\partial y^2} + \sigma^2\frac{\partial^4}{\partial y^4}\right)f(z) + \rho\sigma\frac{\partial^4}{\partial x\partial y^3}\,(z)\right) \\ & + \frac{\rho\sigma y}{2}\frac{\partial^3}{\partial x^3}f(z) + \frac{\rho\sigma^3 y}{2}\frac{\partial^3}{\partial x\partial y^2}f(z) + \rho^2\sigma^2 y\frac{\partial^3 f(z)}{\partial x^2\partial y} \\ & + \rho\sigma y^2\left(\frac{1}{2}\frac{\partial^4}{\partial x^3\partial y}f(z) + \frac{1}{2}\sigma^2\frac{\partial^4}{\partial x\partial y^3}f(z) + \rho\sigma\frac{\partial^4 f(z)}{\partial x^2\partial y^2}\right). \end{aligned}$$

We easily see that the second-order generator for this equation simply has too many components to possibly arrive at any easy enough discretization scheme.

Therefore, in order to construct an easy-to-implement second-order approximation, we need to simplify the stochastic part. We achieve that by multiplying the first equation by σ, the second by ρ, and then subtracting. This way, we get

$$\mathrm{d}(\sigma X - \rho Y)_t = \sigma\sqrt{1-\rho^2}\sqrt{Y_t}\,\mathrm{d}B_t. \tag{5.2.2}$$

Denoting $\bar{Y}_t = Y_t/\sigma^2$ and $\bar{X}_t := (\sigma X_t - \rho Y_t)/(\sigma^2\sqrt{1-\rho^2})$, we arrive at the system

$$\begin{cases} \mathrm{d}\bar{X}_t = \sqrt{\bar{Y}_t}\,\mathrm{d}B_t, \\ \mathrm{d}\bar{Y}_t = \sqrt{\bar{Y}_t}\,\mathrm{d}W_t. \end{cases} \tag{5.2.3}$$

The solution of system (5.2.1) is then a linear transformation of that system (5.2.3):

$$\begin{aligned} X &= \sigma\left(\sqrt{1-\rho^2}\,\bar{X} + \rho\bar{Y}\right), \\ Y &= \sigma^2\bar{Y}. \end{aligned}$$

The generator of the solution of (5.2.3) is

$$Lf(z) = \frac{y}{2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f(z),$$

and therefore,

$$L^2 f(z) = \frac{y}{2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) Lf(z).$$

Calculating

$$\frac{\partial^2}{\partial x^2} Lf(z) = \frac{y}{2}\left[\frac{\partial^4}{\partial x^4} + \frac{\partial^4}{\partial x^2\partial y^2}\right] f(z),$$

$$\begin{aligned} \frac{\partial}{\partial y} Lf(z) &= \frac{\partial}{\partial y}\left[\frac{y}{2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f(z)\right] \\ &= \frac{1}{2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f(z) + \frac{y}{2}\left(\frac{\partial^3}{\partial x^2\partial y} + \frac{\partial^3}{\partial y^3}\right) f(z), \end{aligned}$$

and

$$\frac{\partial^2}{\partial y^2} Lf(z) = \left(\frac{\partial^3}{\partial x^2\partial y} + \frac{\partial^3}{\partial y^3}\right) f(z) + \frac{y}{2}\left(\frac{\partial^4}{\partial x^2\partial y^2} + \frac{\partial^4}{\partial y^4}\right) f(z),$$

we finally get

$$\begin{aligned} L^2 f(z) &= \frac{y}{2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) Lf(z) \\ &= \frac{y}{2}\frac{\partial^2}{\partial x^2} Lf(z) + \frac{y}{2}\frac{\partial^2}{\partial y^2} Lf(z) \end{aligned}$$

$$= \frac{y^2}{4}\left[\frac{\partial^4}{\partial x^4} + \frac{\partial^4}{\partial x^2 \partial y^2}\right] f(z)$$
$$+ \frac{y}{2}\Bigg[\left(\frac{\partial^3}{\partial x^2\partial y} + \frac{\partial^3}{\partial y^3}\right) + \frac{y}{2}\left(\frac{\partial^4}{\partial x^2\partial y^2} + \frac{\partial^4}{\partial y^4}\right)\Bigg] f(z)$$
$$= \frac{y}{2}\left(\frac{\partial^3}{\partial x^2\partial y} + \frac{\partial^3}{\partial y^3}\right) f(z)$$
$$+ \frac{y^2}{4}\left(\frac{\partial^4}{\partial x^4} + 2\frac{\partial^4}{\partial x^2\partial y^2} + \frac{\partial^4}{\partial y^4}\right) f(z).$$

Let $\hat{Z}_h^z = (\hat{X}_h^z, \hat{Y}_h^y)$ be any discretization scheme.[1] Using Taylor's formula and taking expectations, we write

$$\begin{aligned}
\mathbb{E}f(\hat{Z}_h^z) = f(z) &+ \left(\frac{\partial f}{\partial x}(z)\mathbb{E}(\hat{X}_h^z - x) + \frac{\partial f}{\partial y}(z)\mathbb{E}(\hat{Y}_h^y - y)\right) \\
&+ \Bigg(\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(z)\mathbb{E}(\hat{X}_h^z - x)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(z)\mathbb{E}(\hat{Y}_h^y - y)^2 \\
&\quad + \frac{\partial^2 f}{\partial x \partial y}(z)\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y)\Bigg) \\
&+ \Bigg(\frac{1}{6}\frac{\partial^3 f}{\partial x^3}(z)\mathbb{E}(\hat{X}_h^z - x)^3 + \frac{1}{2}\frac{\partial^3 f}{\partial x^2\partial y}(z)\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y) \\
&\quad + \frac{1}{2}\frac{\partial^3 f}{\partial x\partial y^2}(z)\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y)^2 + \frac{1}{6}\frac{\partial^3 f}{\partial y^3}(z)\mathbb{E}(\hat{Y}_h^y - y)^3\Bigg) \\
&+ \Bigg(\frac{1}{24}\frac{\partial^4 f}{\partial x^4}(z)\mathbb{E}(\hat{X}_h^z - x)^4 + \frac{1}{6}\frac{\partial^4 f}{\partial x^3\partial y}(z)\mathbb{E}(\hat{X}_h^z - x)^3(\hat{Y}_h^y - y) \\
&\quad + \frac{1}{4}\frac{\partial^4 f}{\partial x^2\partial y^2}(z)\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y)^2 \\
&\quad + \frac{1}{6}\frac{\partial^4 f}{\partial x\partial y^3}(z)\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y)^3 + \frac{1}{24}\frac{\partial^4 f}{\partial y^4}(z)\mathbb{E}(\hat{Y}_h^y - y)^4\Bigg) \\
&+ \mathbb{E}\nabla_5(\hat{X}_h^z, \hat{Y}_h^y, z) + \mathbb{E}r_5(\hat{X}_h^z, \hat{Y}_h^y, z),
\end{aligned}$$

where

$$\nabla_5(\hat{X}_h^z, \hat{Y}_h^y, z) := \frac{1}{5!}\left(\frac{\partial}{\partial x}(\hat{X}_h^z - x) + \frac{\partial}{\partial y}(\hat{Y}_h^y - y)\right)^5 f(z) \tag{5.2.4}$$

and

$$r_5(\hat{X}_h^z, \hat{Y}_h^y, z_\theta) \tag{5.2.5}$$

are the fifth-order term and fifth-order remainder in the formula, respectively. So, the second-order remainder of a discretization scheme $\hat{Z}_h^z$ equals

$$R_2^h f(z) = \mathbb{E}f(\hat{Z}_h^z) - \left(f(z) + Lf(z)h + L^2 f(z)\frac{h^2}{2}\right)$$

[1] Note that we again (as in the first-order case) write $\hat{Y}_h^y$ instead of $\hat{Y}_h^z$ since we suppose that the latter depends on y only.

$$
\begin{aligned}
&= \frac{\partial f}{\partial x}(z)\mathbb{E}(\hat{X}_h^z - x) + \frac{\partial f}{\partial y}(z)\mathbb{E}(\hat{Y}_h^y - y) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(z)\left(\mathbb{E}(\hat{X}_h^z - x)^2 - yh\right) \\
&+ \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(z)\left(\mathbb{E}(\hat{Y}_h^y - y)^2 - yh\right) + \frac{\partial^2 f}{\partial x \partial y}(z)\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y) \\
&+ \frac{1}{6}\frac{\partial^3 f}{\partial x^3}(z)\mathbb{E}(\hat{X}_h^z - x)^3 + \frac{1}{6}\frac{\partial^3 f}{\partial y^3}(z)\left(\mathbb{E}(\hat{Y}_h^y - y)^3 - \frac{3}{2}yh^2\right) \\
&+ \frac{1}{2}\frac{\partial^3 f}{\partial x^2 \partial y}(z)\left(\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y) - \frac{1}{2}yh^2\right) + \frac{1}{2}\frac{\partial^3 f}{\partial x \partial y^2}(z)\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y)^2 \\
&+ \frac{1}{24}\frac{\partial^4 f}{\partial x^4}(z)\left(\mathbb{E}(\hat{X}_h^z - x)^4 - 3y^2h^2\right) + \frac{1}{6}\frac{\partial^4 f}{\partial x^3 \partial y}(z)\mathbb{E}(\hat{X}_h^z - x)^3(\hat{Y}_h^y - y) \\
&+ \frac{1}{4}\frac{\partial^4 f}{\partial x^2 \partial y^2}(z)\left(\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y)^2 - y^2h^2\right) + \frac{1}{6}\frac{\partial^4 f}{\partial x \partial y^3}(z)\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y)^3 \\
&+ \frac{1}{24}\frac{\partial^4 f}{\partial y^4}(z)\left(\mathbb{E}(\hat{Y}_h^y - y)^4 - 3y^2h^2\right) \\
&+ \mathbb{E}\nabla_5(\hat{X}_h^z, \hat{Y}_h^y, z) + \mathbb{E}r_5(\hat{X}_h^z, \hat{Y}_h^y, z).
\end{aligned}
$$

Therefore, a discretization scheme $\hat{Z}_h^z = (\hat{X}_h^z, \hat{Y}_h^y)$ is a potential second-order approximation of the solution of system (5.2.3) if the following moment conditions are satisfied:

$$
\begin{cases}
\mathbb{E}(\hat{X}_h^z - x) = 0, & \mathbb{E}(\hat{Y}_h^y - y) = 0, \\
\mathbb{E}(\hat{X}_h^z - x)^2 = yh, & \mathbb{E}(\hat{Y}_h^y - y)^2 = yh, \\
\mathbb{E}(\hat{X}_h^z - x)^3 = 0, & \mathbb{E}(\hat{Y}_h^y - y)^3 = \frac{3}{2}yh^2, \\
\mathbb{E}(\hat{X}_h^z - x)^4 = 3y^2h^2, & \mathbb{E}(\hat{Y}_h^y - y)^4 = 3y^2h^2, \\
\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y) = 0, & \\
\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y) = \frac{1}{2}yh^2, & \\
\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y)^2 = 0, & \\
\mathbb{E}(\hat{X}_h^z - x)^3(\hat{Y}_h^y - y) = 0, & \\
\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y)^2 = y^2h^2, & \\
\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y)^3 = 0, &
\end{cases}
\tag{5.2.6}
$$

$z = (x, y) \in \mathbb{D}$ and for every $f \in C_{pol}^{\infty}(\mathbb{D})$,

$$
\mathbb{E}\nabla_5(\hat{X}_h^z, \hat{Y}_h^y, z) = \mathcal{O}_p(h^3) \tag{5.2.7}
$$

and

$$
\mathbb{E}|r_5(\hat{X}_h^z, \hat{Y}_h^y, z)| = \mathcal{O}_p(h^3). \tag{5.2.8}
$$

Remark 5.1. Important! Note that, in fact, Eqs. (5.2.6) need to be satisfied only up to $\mathrm{O}_p(h^3)$ terms, which means that any expressions of order $\mathrm{O}_p(h^3)$ may be added to the right-hand side of the equations.

Proposition 5.2. Let a one-step approximation $\hat{Z}_h^z = (\hat{X}_h^z, \hat{Y}_h^y)$ be constructed as follows:

(1) Let $\hat{Y}_h^y$ be a discretization scheme satisfying the four conditions on the right in Eqs. (5.2.6) *(*up to $\mathrm{O}_p(h^3)$ terms*)* and, in addition, the following moment conditions:

   (i) $\mathbb{E}(\hat{Y}_h^y - y)^5 = \mathrm{O}_p(h^3)$;

   (ii) $\mathbb{E}|\hat{Y}_h^y - y|^{2q} = \mathrm{O}_p(h^q)$ for all $q \geq 3$.

(2) Let ξ be a discrete random variable, independent from $\hat{Y}_h^y$, with first five moments matching those of a standard normal random variable. Let, finally, the random variable $\hat{X}_h^z$ be defined by

$$\hat{X}_h^z := x + \xi\sqrt{\frac{1}{2}(y + \hat{Y}_h^y)h} \ . \tag{5.2.9}$$

Then the (one-step) approximation scheme $\hat{Z}_h^z := (\hat{X}_h^z, \hat{Y}_h^y)$ satisfies all the requirements in Eqs. (5.2.6)–(5.2.8).

Remark 5.3. An example of approximation scheme $\hat{Y}_h^y$ satisfying conditions (1) of the theorem is a particular case ($\sigma = 1$) of Theorem 15 in [27]. Let the random variables $\hat{Y}_h^y$ take the values y_1, y_2, and y_0 with corresponding probabilities p_1, p_2, and $p_0 = 1 - p_1 - p_2$ defined as follows:

- If $y \geq 2h$, then

$$\begin{aligned} &y_1 = y + (s - \sqrt{\Delta})/2, \quad y_2 = y + (s + \sqrt{\Delta})/2, \quad y_0 = y; \\ &p_1 = \frac{2yh}{\sqrt{\Delta}(\sqrt{\Delta} - s)}, \quad p_2 = \frac{2yh}{\sqrt{\Delta}(\sqrt{\Delta} + s)}, \text{where} \\ &s = \frac{3h}{2}, \quad \Delta = \frac{21}{4}h^2 + 12yh; \end{aligned}$$

- If $0 < y < 2h$, then

$$y_1 = (s - \sqrt{\Delta})/2, \quad y_2 = (s + \sqrt{\Delta})/2, \quad y_0 = 0;$$

$$p_1 = \frac{y(2y+2h-s-\sqrt{\Delta})}{\sqrt{\Delta}(\sqrt{\Delta}-s)}, \quad p_2 = \frac{y(2y+2h-s+\sqrt{\Delta})}{\sqrt{\Delta}(\sqrt{\Delta}+s)}, \quad \text{where}$$
$$s = \frac{4y^2+9yh+3h^2}{2y+h}, \quad \Delta = \frac{h(16y^3+33y^2h+18yh^2+3h^3)}{(2y+h)^2}.$$

Then $\hat{Y}_h^y$ satisfies condition (1) of the theorem.

Remark 5.4. Examples of (distributions of) random variables ξ:

(1) $\mathbb{P}\{\xi = \pm\sqrt{3}\} = \frac{1}{6}$, $\mathbb{P}\{\xi = 0\} = \frac{2}{3}$;

(2) $\mathbb{P}\left\{\xi = \pm\sqrt{1-\sqrt{6}/3}\right\} = \frac{3}{8}$, $\mathbb{P}\left\{\xi = \pm\sqrt{1+\sqrt{6}}\right\} = \frac{1}{8}$.

The first one is simpler, but the second has an advantage: it can be simply generated by three bits of a uniform (pseudo-)random variable in the interval $[0,1]$, leaving the remaining bits for generating $\hat{Y}_h^y$ and, thus, $\hat{X}_h^z$.

Proof of Proposition 5.2. From the condition (i) of the Proposition we have that (up to $\mathrm{O}_p(h^3)$ terms)

$$\begin{aligned}
\mathbb{E}(\hat{Y}_h^y - y) &= 0,\\
\mathbb{E}(\hat{Y}_h^y - y)^2 &= yh,\\
\mathbb{E}(\hat{Y}_h^y - y)^3 &= \frac{3}{2}yh^2,\\
\mathbb{E}(\hat{Y}_h^y - y)^4 &= 3y^2h^2.
\end{aligned}$$

Since $\mathbb{E}\xi^n = 0$, $n = 1,3,5$, $\mathbb{E}\xi^2 = 1$, and $\mathbb{E}\xi^4 = 3$, using the independence of $\hat{Y}_h^y$ and ξ, we easily calculate

$$\mathbb{E}(\hat{X}_h^z - x)^n(\hat{Y}_h^y - y)^k = \left(\sqrt{\frac{h}{2}}\right)^n \mathbb{E}\xi^n \mathbb{E}\left(\sqrt{(y+\hat{Y}_h^y)}\right)^n \mathbb{E}(\hat{Y}_h^y - y)^k = 0$$
$$\text{for } n = 1,3,5 \text{ and all } k \in \mathbb{N}_0,$$

$$\begin{aligned}
\mathbb{E}(\hat{X}_h^z - x)^2 &= \frac{h}{2}\mathbb{E}\left(y+\hat{Y}_h^y\right)\mathbb{E}\xi^2 = hy,\\
\mathbb{E}(\hat{X}_h^z - x)^4 &= \frac{h^2}{4}\mathbb{E}\left(y+\hat{Y}_h^y\right)^2\mathbb{E}\xi^4\\
&= \frac{3h^2}{4}\left((2y)^2 + 4y\mathbb{E}(\hat{Y}_h^y - y) + \mathbb{E}(\hat{Y}_h^y - y)^2\right)\\
&= \frac{3h^2}{4}\left(4y^2 + yh\right) = 3y^2h^2 + \frac{3}{4}yh^3 = 3y^2h^2 + \mathrm{O}_p(h^3),
\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y) &= \frac{h}{2}\mathbb{E}\big[(y + \hat{Y}_h^y)(\hat{Y}_h^y - y)\big]\mathbb{E}\xi^2 \\
&= \frac{h}{2}\big[2y\mathbb{E}(\hat{Y}_h^y - y) + \mathbb{E}(\hat{Y}_h^y - y)^2\big] \\
&= \frac{h}{2}\big[2y \cdot 0 + yh\big] = \frac{yh^2}{2},
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y)^2 &= \frac{h}{2}\mathbb{E}\big[(y + \hat{Y}_h^y)(\hat{Y}_h^y - y)^2\big]\mathbb{E}\xi^2 \\
&= \frac{h}{2}\big[2y\mathbb{E}(\hat{Y}_h^y - y)^2 + \mathbb{E}(\hat{Y}_h^y - y)^3\big] \\
&= \frac{h}{2}\big[2y \cdot yh + \frac{3}{2}yh^2\big] = y^2h^2 + \mathrm{O}_p(h^3).
\end{aligned}
$$

Thus, we have checked all the conditions in Eqs. (5.2.6) up to $\mathrm{O}_p(h^3)$ terms. To prove estimate (5.2.7), we additionally calculate the moments

$$
\begin{aligned}
\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y)^3 &= \frac{h}{2}\mathbb{E}\big[(y + \hat{Y}_h^y)(\hat{Y}_h^y - y)^3\big]\mathbb{E}\xi^2 \\
&= \frac{h}{2}\big[2y\mathbb{E}(\hat{Y}_h^y - y)^3 + \mathbb{E}(\hat{Y}_h^y - y)^4\big] \\
&= \frac{h}{2}\big[2y \cdot \frac{3}{2}yh^2 + 3y^2h^2\big] = 6y^2h^3 = \mathrm{O}_p(h^3)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}(\hat{X}_h^z - x)^4(\hat{Y}_h^y - y) &= \frac{h}{2}\mathbb{E}\big[(y + \hat{Y}_h^y)^2(\hat{Y}_h^y - y)\big]\mathbb{E}\xi^4 \\
&= \frac{3h^2}{4}\big[4y^2\mathbb{E}(\hat{Y}_h^y - y) + 4y\mathbb{E}(\hat{Y}_h^y - y)^2 + \mathbb{E}(\hat{Y}_h^y - y)^3\big] \\
&= \frac{3h^2}{4}\big[4y^2 \cdot 0 + 4y^2h + \frac{3}{2}yh^2\big] \\
&= 3y^2h^3 + \frac{9yh^4}{8} = \mathrm{O}_p(h^3).
\end{aligned}
$$

Therefore, for every function $f \in C_{pol}^\infty(\mathbb{D})$, we have

$$
\begin{aligned}
&\mathbb{E}\nabla_5(\hat{X}_h^z, \hat{Y}_h^y, z) \\
&\quad = \frac{1}{5!}\frac{\partial^5 f}{\partial x^5}(z)\mathbb{E}(\hat{X}_h^z - x)^5 + \frac{1}{4!}\frac{\partial^5 f}{\partial x^4 \partial y}(z)\mathbb{E}(\hat{X}_h^z - x)^4(\hat{Y}_h^y - y) \\
&\qquad + \frac{1}{12}\frac{\partial^5 f}{\partial x^3 \partial y^2}(z)\mathbb{E}(\hat{X}_h^z - x)^3(\hat{Y}_h^y - y)^2 + \frac{1}{12}\frac{\partial^5 f}{\partial x^2 \partial y^3}(z)\mathbb{E}(\hat{X}_h^z - x)^2(\hat{Y}_h^y - y)^3 \\
&\qquad + \frac{1}{4!}\frac{\partial^5 f}{\partial x \partial y^4}(z)\mathbb{E}(\hat{X}_h^z - x)(\hat{Y}_h^y - y)^4 + \frac{1}{5!}\frac{\partial^5 f}{\partial y^5}(z)\mathbb{E}(\hat{Y}_h^y - y)^5 \\
&\quad = \left[\frac{1}{8}\Big(y^2 + \frac{3}{8}yh\Big)\frac{\partial^5 f}{\partial x^4 \partial y}(z) + \frac{1}{4}y^2\frac{\partial^5 f}{\partial x^2 \partial y^3}(z)\right] h^3 + \frac{1}{5!}\frac{\partial^5 f}{\partial y^5}(z)\mathrm{O}_p(h^3) \\
&\quad = \mathcal{O}_p(h^3).
\end{aligned}
$$

Now consider

$$\mathbb{E}r_5(\hat{X}_h^z, \hat{Y}_h^y, x, y) = \frac{1}{6!} \sum_{m+n=6} \tbinom{6}{m} \mathbb{E} \frac{\partial^6 f}{\partial x^m \partial y^n}(z_\theta)(\hat{X}_h^z - x)^m (\hat{Y}_h^y - y)^n, \tag{5.2.10}$$

where $z_\theta = (x_\theta, y_\theta)$ with $x_\theta = x + \theta(\hat{X}_h^z - x)$ and $y_\theta = y + \theta(\hat{Y}_h^y - y)$ for some $\theta \in (0,1)$.

Using condition (i) of the Proposition, the independence of $\hat{Y}_h^y$ and ξ, and the Cauchy–Schwarz inequality, for all $m, n \in \mathbb{N}_0$, we get

$$\begin{aligned}
\mathbb{E}\left|(\hat{X}_h^z - x)^m (\hat{Y}_h^y - y)^n\right| &\leqslant \sqrt{\mathbb{E}(\hat{X}_h^z - x)^{2m}}\sqrt{\mathbb{E}(\hat{Y}_h^y - y)^{2n}} \\
&= \sqrt{\left(\frac{h}{2}\right)^m \mathbb{E}\xi^{2m}\mathbb{E}(y + \hat{Y}_h^y)^m}\sqrt{\mathrm{O}_p(h^n)} \\
&\leqslant \sqrt{Ch^m \mathbb{E}(\hat{Y}_h^y - y + 2y)^m}\sqrt{\mathrm{O}_p(h^n)} \\
&\leqslant \sqrt{Ch^m \left(\mathbb{E}|\hat{Y}_h^y - y|^m + (2y)^m\right)}\sqrt{\mathrm{O}_p(h^n)} \\
&\leqslant \sqrt{Ch^m \left(\sqrt{\mathbb{E}|\hat{Y}_h^y - y|^{2m}} + y^m\right)}\sqrt{\mathrm{O}_p(h^n)} \\
&\leqslant \sqrt{Ch^m \left(\sqrt{\mathrm{O}_p(h^m)} + y^m\right)}\sqrt{\mathrm{O}_p(h^n)} \\
&= \sqrt{\mathrm{O}_p(h^m)}\sqrt{\mathrm{O}_p(h^n)} = \sqrt{\mathrm{O}_p(h^{n+m})} \\
&= \mathrm{O}_p\left(h^{\frac{n+m}{2}}\right),
\end{aligned} \tag{5.2.11}$$

where the constant C depends on m and varies from line to line.

Since for any $f \in C_{pol}^\infty(\mathbb{D})$ and multiindex $i \in \mathbb{N}_0^2$, there exist $C > 0$ and $q \in \mathbb{N}$, depending on a good sequence for f only, such that

$$\left(f^{(i)}(z)\right)^2 \leqslant C(1 + x^{2q} + y^{2q}), \quad z = (x, y) \in \mathbb{D},$$

we can further estimate

$$\begin{aligned}
\sqrt{\mathbb{E}\left(f^{(i)}(z_\theta)\right)^2} &\leqslant C\sqrt{1 + \mathbb{E}(x + \theta(\hat{X}_h^z - x))^{2q} + \mathbb{E}(y + \theta(\hat{Y}_h^y - y))^{2q}} \\
&\leqslant C\sqrt{1 + x^{2q} + y^{2q} + \mathbb{E}(\hat{X}_h^z - x)^{2q} + \mathbb{E}(\hat{Y}_h^y - y)^{2q}} \\
&\leqslant C\sqrt{1 + x^{2q} + y^{2q} + h^q(1 + |z|^{2k})} \\
&\leqslant C(1 + |z|^{q \vee k}), \quad h \leq h_0,
\end{aligned}$$

where the constants $q, k \in \mathbb{N}$, $h_0 > 0$, and $C > 0$ depend on a good sequence for f only (and, as before, C varies from line to line). Thus, shortly,

$$\sqrt{\mathbb{E}\left(f^{(i)}(z_\theta)\right)^2} = \mathcal{O}_p(1).$$

Finally, using the latter and the Cauchy–Schwarz inequality, for any multiindex $i = (n,m) \in \mathbb{N}_0$ such that $n+m=6$, we have

$$\begin{aligned}
&\mathbb{E}\left|f^{(i)}(z_\theta)(\hat{X}_h^z - x)^m(\hat{Y}_h^y - y)^n\right| \\
&\leqslant \sqrt{\mathbb{E}\left(f^{(i)}(z_\theta)\right)^2}\sqrt{\mathbb{E}(\hat{X}_h^z - x)^{2m}(\hat{Y}_h^y - y)^{2n}} \\
&= \mathcal{O}_p(1)\sqrt{\mathrm{O}_p(h^n)\mathrm{O}_p(h^m)} = \mathcal{O}_p(h^{\frac{n+m}{2}}) = \mathcal{O}_p(h^3). \qquad (5.2.12)
\end{aligned}$$

Now inserting estimate (5.2.12) into (5.2.10) proves estimate (5.2.8). □

Proposition 5.5. Let $\hat{Y}_h^y$ be any one-step discretization scheme taking nonnegative values and satisfying the following moment condition:

$$|\mathbb{E}(\hat{Y}_h^y - y)^n| \le c_n(1+y^n)h, \quad y \ge 0,\ h \le h_0,\ n \in \mathbb{N}, \qquad (5.2.13)$$

with some constants c_n, $n \in \mathbb{N}$.
Let $\hat{X}_h^z$ be defined as in Eq. (5.2.9). Then the approximation $\hat{Z}_t^z := (\hat{X}_t^z, \hat{Y}_t^y)$, $t = kh$, $h = [T/N]$, $N \in \mathbb{N}$, $k = 0, 1, \ldots, N$, defined by the one-step scheme $\hat{Z}_h^z := (\hat{X}_h^z, \hat{Y}_h^y)$ has uniformly bounded moments on the whole time interval $[0,T]$, that is, for all $q \in \mathbb{N}$,

$$\sup_{N\in\mathbb{N}} \sup_{0\le k\le N} \mathbb{E}|\hat{Z}_{kh}^z|^q < +\infty. \qquad (5.2.14)$$

Proof. We will prove the equivalent relation

$$\sup_{N\in\mathbb{N}} \sup_{0\le k\le N} \mathbb{E}\Big(|\hat{X}_{kh}^z|^q + \big(\hat{Y}_{kh}^y\big)^q\Big) < +\infty \qquad (5.2.15)$$

for all $q \in \mathbb{N}$. Note that it suffices to check the latter for sufficiently large q. By the markovity of the approximation and Proposition 1.5 of [3], in turn, it suffices to prove that, for any sufficiently large $q \in \mathbb{N}$ and any $h_0 > 0$, there exists a constant C_q such that

$$\mathbb{E}\Big(|\hat{X}_h^z|^q + \big(\hat{Y}_h^y\big)^q\Big) \le (|x|^q + y^q)(1 + C_q h) + C_q h, \quad 0 < h \le h_0. \qquad (5.2.16)$$

We have

$$\begin{aligned}
\mathbb{E}\big(\hat{Y}_h^y\big)^q &= \mathbb{E}\Big(y + (\hat{Y}_h^y - y)\Big)^q \\
&= y^q + \sum_{n=1}^{q} \tbinom{q}{n} y^{q-n}\mathbb{E}\Big(\hat{Y}_h^y - y\Big)^n \\
&\le y^q + \sum_{n=1}^{q} c_n \tbinom{q}{n} y^{q-n}(1+y^n)h
\end{aligned}$$

$$
\begin{aligned}
&= y^q + h\Big(\sum_{n=1}^{q} c_n \tbinom{q}{n} y^q + \sum_{n=1}^{q} c_n \tbinom{q}{n} y^{q-n}\Big) \\
&\le y^q + h\Big(a_q y^q + \sum_{n=0}^{q-1} a_n y^n\Big), \quad 0 < h \le h_0,
\end{aligned}
$$

with some constants $a_i > 0$, $i = 0, 1, \ldots, q$. From this we have

$$
\begin{aligned}
\mathbb{E}\big(\hat{Y}_h^y\big)^q &\le y^q + h\Big(a_q y^q + \sum_{n=1}^{q-1} a_n y^n\Big) \\
&\le y^q(1 + 2a_q h) + h\Big(-a_q y^q + \sum_{n=0}^{q-1} a_n y^n\Big), \quad 0 < h \le h_0. \qquad (5.2.17)
\end{aligned}
$$

For $\mathbb{E}|\hat{X}_h^z|^q$, a rougher estimate suffices:

$$
\begin{aligned}
\mathbb{E}|\hat{X}_h^z|^q &= \mathbb{E}\Big|x + \xi\sqrt{\tfrac{1}{2}(y + \hat{Y}_h^y)h}\,\Big|^q \\
&\le |x|^q + \tfrac{1}{2}\tbinom{q}{2}|x|^{q-2}\mathbb{E}\xi^2 \cdot \mathbb{E}(y + \hat{Y}_h^y)h + \tfrac{1}{4}\tbinom{q}{4}|x|^{q-4}\mathbb{E}\xi^4 \cdot \mathbb{E}(y + \hat{Y}_h^y)^2 h^2 \\
&\quad + \sum_{n=6}^{q} \tfrac{1}{2^{n/2}}\tbinom{q}{n}|x|^{q-n}\mathbb{E}\xi^n \mathbb{E}(y + \hat{Y}_h^y)^{n/2} h^{n/2} \\
&\le |x|^q + b_{q-2}|x|^{q-2}(2y + |\mathbb{E}(\hat{Y}_h^y - y)|)h \\
&\quad + \sum_{n=4}^{q} b_{q-n}|x|^{q-n}((2y)^{n/2} + |\mathbb{E}(\hat{Y}_h^y - y)|^{n/2})h^{n/2} \\
&\le |x|^q + \tilde{b}_{q-2}|x|^{q-2}(y + (1+y)h)h \\
&\quad + \sum_{n=4}^{q} \tilde{b}_{q-n}|x|^{q-n}(y^{n/2} + (1 + y^{n/2})h^{1/2})h^{n/2}, \quad 0 < h \le h_0.
\end{aligned}
$$

The latter estimate can be simplified as follows:

$$
\mathbb{E}|\hat{X}_h^z|^q \le |x|^q + h\Big(\sum_{m+n<q} b_{m,n}|x|^m y^n\Big), \quad 0 < h \le h_0, \qquad (5.2.18)
$$

with some constants $b_{n,m} > 0$ for $m, n \in \mathbb{N}_0$ such that $m + n < q$.

Collecting estimates (5.2.17) and (5.2.18), we have

$$
\begin{aligned}
\mathbb{E}\Big(|\hat{X}_h^z|^q + \big(\hat{Y}_h^y\big)^q\Big) &\le (|x|^q + y^q)(1 + 2a_q h) \\
&\quad + h\Big(-a_q\big(|x|^q + y^q\big) + \sum_{n=0}^{q-1} a_n y^n + \sum_{m+n<q} b_{m,n}|x|^m y^n\Big), \quad h \le h_0.
\end{aligned}
$$

Using (4.2.18), we notice that

$$
B_q := \sup_{x\in\mathbb{R}\, y>0} \Big(-a_q\big(|x|^q + y^q\big) + \sum_{n=1}^{q-1} a_n y^n + \sum_{m+n<q} b_{m,n}|x|^m y^n\Big) < +\infty.
$$

Denoting $C_q := \max\{2a_q, B_q\}$, we finally have

$$\mathbb{E}\Big(|\hat{X}_h^z|^q + \mathbb{E}\big(\hat{Y}_h^y\big)^q\Big) \le (|x|^q + y^q)(1 + C_q h) + C_q h, \quad 0 < h \le h_0,$$

that is, estimate (5.2.16) holds for all $q \in \mathbb{N}$, as required. □

Summarizing Propositions 3.2, 5.2, and 5.5, we get our main result.

Theorem 5.6. Let a one-step approximation $\bar{Z}_h^z = (\bar{X}_h^z, \bar{Y}_h^y)$ of Eq. (1.2.1) be constructed as follows:

(1) Let $\hat{Y}_h^y$ be a discretization scheme satisfying the following conditions:

(i)

$$\begin{cases} \mathbb{E}(\hat{Y}_h^y - y) = \mathrm{O}_{p,1}(h^3), \\ \mathbb{E}(\hat{Y}_h^y - y)^2 = yh + \mathrm{O}_{p,2}(h^3), \\ \mathbb{E}(\hat{Y}_h^y - y)^3 = \frac{3}{2}yh^2 + \mathrm{O}_{p,3}(h^3), \\ \mathbb{E}(\hat{Y}_h^y - y)^4 = 3y^2h^2 + \mathrm{O}_{p,4}(h^3); \end{cases}$$

(ii) $\mathbb{E}(\hat{Y}_h^y - y)^5 = \mathrm{O}_{p,5}(h^3)$;

(iii) $\mathbb{E}|\hat{Y}_h^y - y|^{2q} = \mathrm{O}_{p,2q}(h^q)$ for all $q \ge 3$.

(2) Let ξ be a discrete random variable, independent from $\hat{Y}_h^y$, with first five moments matching those of a standard normal random variable. Let, finally, the random variable $\hat{X}_h^z$ be defined by

$$\hat{X}_h^z := x + \xi\sqrt{\frac{1}{2}(y + \hat{Y}_h^y)h}\ .$$

(3) Define the one-step discretization scheme $\tilde{Z}_h^z = (\tilde{X}_h^z, \tilde{Y}_h^y)$ by

$$\tilde{X}_h^z = \sigma\left(\sqrt{1-\rho^2}\,\hat{X}_h^z + \rho\hat{Y}_h^y\right),$$
$$\tilde{Y}_h^y = \sigma^2\tilde{Y}_h^y.$$

(4) Finally, define the one-step discretization scheme $\bar{Z}_h^z = (\bar{X}_h^z, \bar{Y}_h^y)$ by Eq. (5.1.1), that is,

$$\bar{Z}_h^z = D\big(\tilde{Z}_h^{D(z,h/2)}, h/2\big). \tag{5.2.19}$$

Then the one-step discretization scheme (5.2.19) defines a strongly potential weak second-order approximation of the log-Heston system (1.2.1).

Proof. It remains to check that the approximation constructed also has uniformly bounded moments of all orders. From the previous chapter we know (see 4.2.22) that there exists a constant $\tilde{C}_q$, independent of $z \in \mathbb{D}$ and h, such that

$$|D_1^z(h)|^q + \big(D_2^y(h)\big)^q \le (|x|^q + y^q)(1 + \tilde{C}_q h) + \tilde{C}_q h, \quad 0 < h \le h_0.$$

This shows that composition (5.2.19) "preserves" the estimate of type (5.2.16), that is, for every $q \in \mathbb{N}$, there exists a constant $\bar{C}_q$, independent of h, such that

$$\mathbb{E}\big(|\bar{X}_h^z|^q + \big(\bar{Y}_h^y\big)^q\big) \le (|x|^q + y^q)(1 + \bar{C}_q h) + \bar{C}_q h, \quad 0 < h \le h_0,$$

which suffices for the uniform boundedness of all moments of $\bar{Z}_h^z = (\bar{X}_h^z, \bar{Y}_h^y)$. $\square$

5.3 Algorithm

In this section, we summarize the simulation step from $\bar{Z}_{ih} = (\bar{X}_{ih}, \bar{Y}_{ih}) = (x, y)$ to $\bar{Z}_{(i+1)h} = (\bar{X}_{(i+1)h}, \bar{Y}_{(i+1)h})$:

1. Draw a random number U uniformly distributed in the interval $(0,1)$, and let $U_1 := [8U]$ and $U_2 := \{8U\}$ (the integer and fractional parts of $8U$) (which are independent r.v.s; the former takes the values 0,1,. . . ,7 with probabilities $\frac{1}{8}$, and the latter is uniformly distributed in $(0,1)$).

2. Generate
$$\xi := \begin{cases} \sqrt{1-\sqrt{6}/3} & \text{if } U_1 < 3; \\ -\sqrt{1-\sqrt{6}/3} & \text{if } U_1 > 4; \\ \sqrt{1+\sqrt{6}} & \text{if } U_1 = 3; \\ -\sqrt{1+\sqrt{6}} & \text{if } U_1 = 4. \end{cases}$$

3. Calculate (in view of (3.2.3) with $h/2$ instead of h)
$$\begin{aligned} \bar{x} &= x + (r - \tfrac{1}{2}\theta)\frac{h}{2} + \tfrac{1}{2k}(\mathsf{e}^{-k\frac{h}{2}} - 1)(y - \theta), \\ \bar{y} &= y\mathsf{e}^{-k\frac{h}{2}} + \theta(1 - \mathsf{e}^{-k\frac{h}{2}}). \end{aligned}$$

4. Generate a random variable $\hat{Y}$ taking the values y_1, y_2, and y_0 with probabilities p_1, p_2, and $p_0 = 1 - p_1 - p_2$ as defined in Remark 5.3 (with $\bar{y}$ instead

of y):

$$\hat{Y} := \begin{cases} y_1 & \text{if } U_2 < p_1, \\ y_2 & \text{if } U_2 > 1 - p_2, \\ y_0 & \text{otherwise.} \end{cases}$$

5. Calculate the random variable (see (5.2.9))

$$\hat{X} := \bar{x} + \xi \sqrt{\frac{1}{2}(\bar{y} + \hat{Y})h}\,.$$

6. Calculate

$$\tilde{X} = \sigma\left(\sqrt{1-\rho^2}\hat{X} + \rho\hat{Y}\right),$$
$$\tilde{Y} = \sigma^2\hat{Y}.$$

7. Finally, calculate (using again (3.2.3) with $h/2$ instead of h)

$$\bar{X}_{(i+1)h} = \tilde{X} + (r - \tfrac{1}{2}\theta)\frac{h}{2} + \tfrac{1}{2k}(\mathsf{e}^{-k\frac{h}{2}} - 1)(\tilde{Y} - \theta),$$
$$\bar{Y}_{(i+1)h} = \tilde{Y}\mathsf{e}^{-k\frac{h}{2}} + \theta(1 - \mathsf{e}^{-k\frac{h}{2}}).$$

Having generated all values $\bar{X}_{ih}$, $i = 0, 1, 2, \ldots$, set $\overline{S}_{ih} := \exp\{\bar{X}_{ih}\}$, $i = 0, 1, 2, \ldots$.

5.4 Simulation examples

As in the previous chapter, to illustrate our discretization scheme and to compare it with the quadratic-exponential (QE) scheme proposed by Andersen [5] and the second-order scheme proposed by Alfonsi [3], we use it to calculate European call and put option prices in five cases (see Table 4.1), that is, we apply it to the test functions $f_1(x) = \max\{x - K, 0\}$ and $f_2(x) = \max\{K - x, 0\}$.

We provide two kinds of graphs: first, we plot (put or call) option prices as functions of maturity, and, second, we concentrate ourselves on the accuracies of approximations as functions of the approximation step size h.

We start by illustrating the accuracies of approximations as functions of the step size h. We split the five cases into two graphs. Figure 5.1 shows the accuracy of approximations of put option price at maturity $T = 1$ in the cases from [5], that is, cases I–III (with strike prices 70, 100, 140), and Fig. 5.2 shows the put prices at $T = 1$ in the cases from [3], that is, cases IV and V (with strike prices 80, 100, 120). The approximate prices are obtained by averaging over 10^8 samples. Horizontal dashed lines in both figures show the exact values of put prices, which were calculated using the MATLAB code *HestonFFTVanilla* from [9] (downloaded from www.quantlet.de).

Similarly as in the previous chapter, we take one of the cases from each graph (this time we choose cases I and IV) and examine the errors in the plots more closely. As before, we calculate the bias

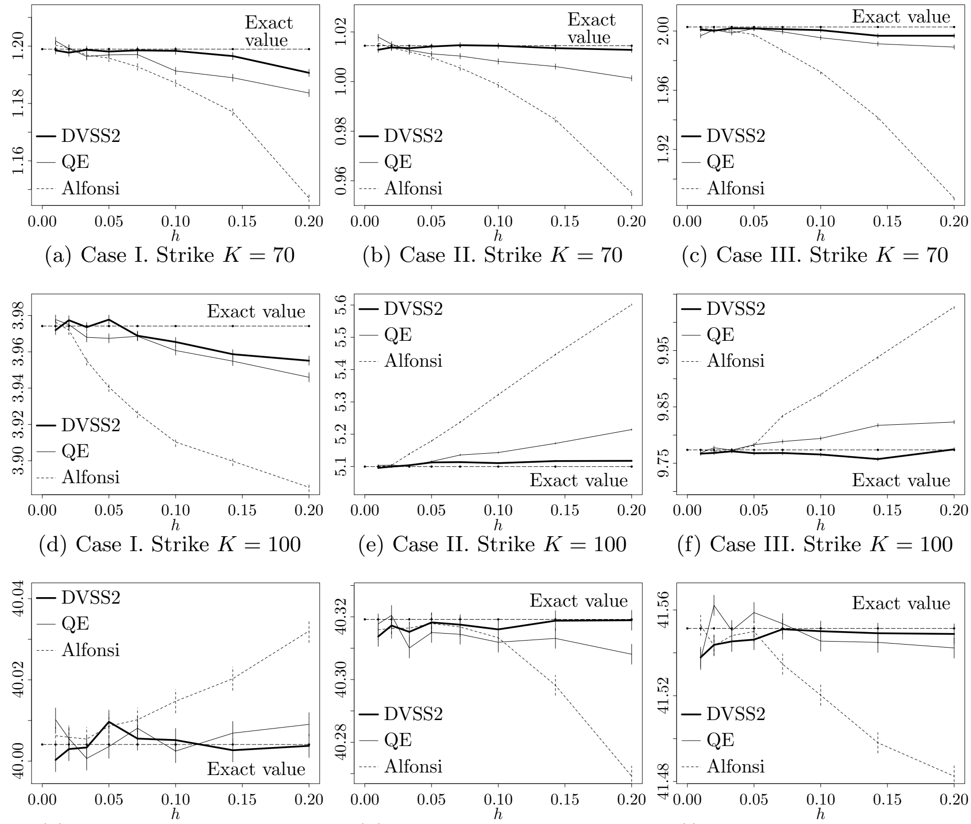
$$\varepsilon_h = \mathbb{E}\mathrm{e}^{-rT}\max(S_T - K, 0) - \mathbb{E}\mathrm{e}^{-rT}\max(\overline{S}_T^h - K, 0),$$

that is, the difference between the exact price of call option and its evaluation using an approximation scheme, and put the results in Tables 5.1 and 5.2. The values in the table are written in bold when the bias is within the window of $\pm 1\%$ from the exact value. We also put the half-length of the 95% confidence interval in parenthesis and mark it with an asterisk if the bias is less than the half-length of the 95% confidence interval, that is, when the error of approximation is not statistically significant.

Next, we plot (call and put) option prices as functions of maturity. As before, we split the cases into two graphs. We plot the cases from [5] (i.e., cases I–III) in Fig. 5.3 with maturity $T = 10$ and the cases from [3] (i.e., cases IV and V) in Fig. 5.4 with maturity $T = 2$. We take $h = 0.2$ and $S_0 = 100$ for all the cases, $r = 0$ for cases I–III, and $r = 0.02$ for cases IV and V. In both graphs, the exact option prices are marked by bullets "•" (at discrete time points with step size 1 in cases I–III and with step size 0.2 in cases IV and V).

From Fig. 5.3 and Fig. 5.4 we see that the DVSS_2 and QE schemes are visually almost indistinguishable, whereas the Alfonsi scheme considerably less accurate. A closer examination of approximations as functions of time step h (see Figs. 5.1 and 5.2 and, also, Table 5.1 and Table 5.2) reveals that, actually, in most of the cases, both QE and Alfonsi schemes have a greater bias for almost all time steps (starting from $h = 0.2$) than DVSS_2.

To complete, we compare the computation times of DVSS_2, Andersen's QE, and Alfonsi schemes. For calculations, we used a laptop PC with Intel(R) CORE(TM)



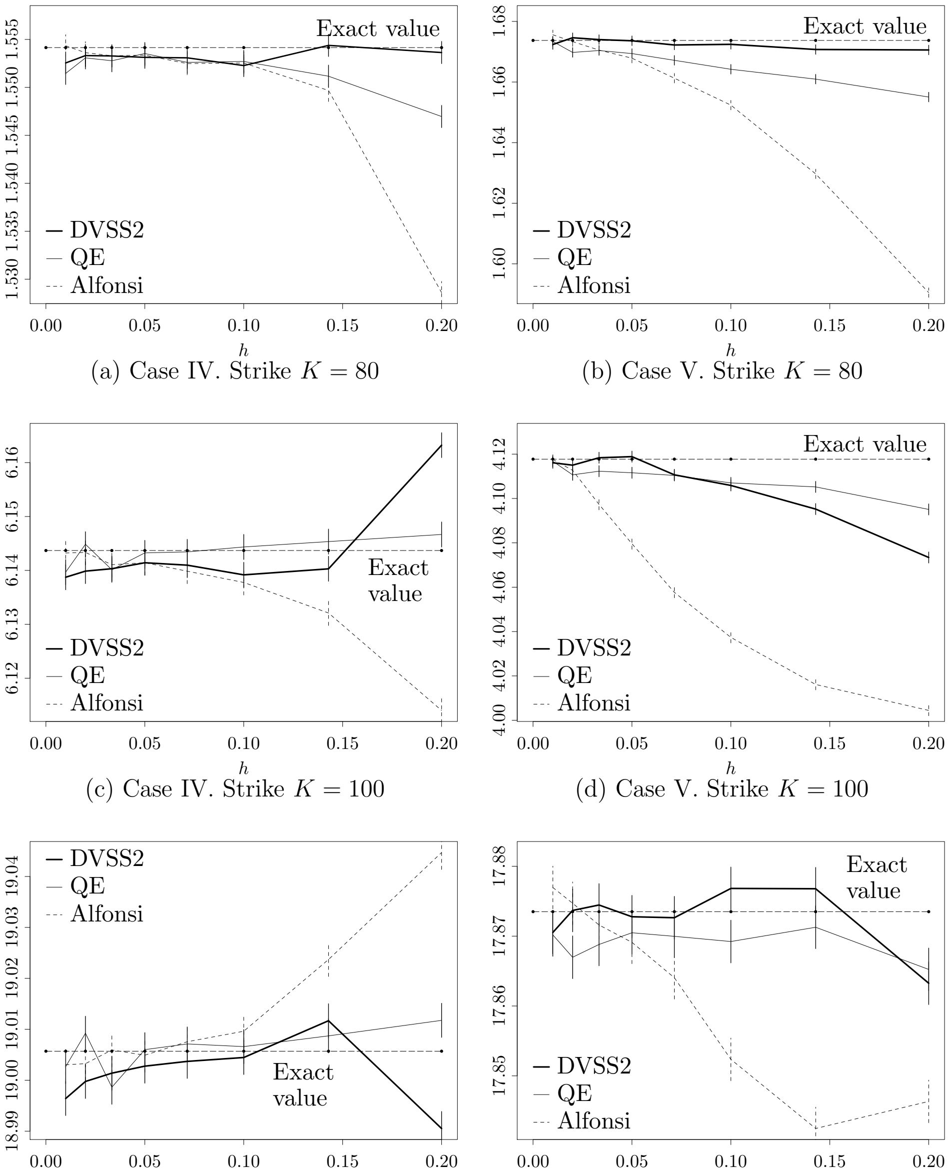
(a) Case I. Strike $K = 70$ (b) Case II. Strike $K = 70$ (c) Case III. Strike $K = 70$

(d) Case I. Strike $K = 100$ (e) Case II. Strike $K = 100$ (f) Case III. Strike $K = 100$

(g) Case I. Strike $K = 140$ (h) Case II. Strike $K = 140$ (i) Case III. Strike $K = 140$

Figure 5.1: Approximations of put option price in function of $h = 1/n$ in cases I, II, and III for different strike prices (70, 100, 140) with fixed parameters $S_0 = 100$, $r = 0$, and $T = 1$.

i7-3610QM 2.30 GHz processor and 8 GB RAM and the C programming language (the graphs were plotted by using the R programming language with graphics package TikzDevice). The results are listed in Table 5.3. The numbers on the left side of the table show the averages over the runs of 100 000 trajectories with step $h = 0.01$ and time to maturity $T = 10$ in all the cases. One can see that the DVSS_2 scheme is faster in all the cases (see the computational times measured relatively to the DVSS_2 scheme on the right of the table). Note that we tried to optimize the QE and Alfonsi approximation code by caching parts of computations not depending on each time step.



(a) Case IV. Strike $K = 80$

(b) Case V. Strike $K = 80$

(c) Case IV. Strike $K = 100$

(d) Case V. Strike $K = 100$

(e) Case IV. Strike $K = 120$

(f) Case V. Strike $K = 120$

Figure 5.2: Approximations of put price in function of $h = 1/n$ in cases IV and V for different strike prices (80, 100, 120) with fixed parameters $S_0 = 100$ and $T = 1$.

	DVSS$_2$	QE	Alfonsi
h		$K = 70$	
1/5	**0,00722** (0,00130)	**0,01547** (0,00130)	**0,05262** (0,00130)
1/7	**0,00283** (0,00130)	**0,00941** (0,00130)	**0,02294** (0,00130)
1/10	**0,00178** (0,00130)	**0,00719** (0,00130)	**0,01221** (0,00130)
1/14	**0,00026** (0,00130*)	**0,00201** (0,00130)	**0,00562** (0,00130)
1/20	**-0,00009** (0,00130*)	**0,00235** (0,00130)	**0,00331** (0,00130)
1/30	**0,00013** (0,00130*)	**0,00248** (0,00130)	**0,00149** (0,00130)
1/50	**-0,00109** (0,00131*)	**-0,00029** (0,00130*)	**-0,00097** (0,00131*)
1/100	**0,00139** (0,00130)	**-0,00286** (0,00131)	**-0,00167** (0,00131)
		$K = 100$	
1/5	**0,01818** (0,00256)	**0,02826** (0,00255)	0,08766 (0,00250)
1/7	**0,01940** (0,00256)	**0,01927** (0,00255)	0,07447 (0,00253)
1/10	**0,01061** (0,00256)	**0,01400** (0,00256)	0,06335 (0,00255)
1/14	**0,00612** (0,00256)	**0,00644** (0,00256)	0,05004 (0,00255)
1/20	**-0,00285** (0,00256)	**0,00656** (0,00256)	**0,03317** (0,00256)
1/30	**-0,00110** (0,00256*)	**0,00648** (0,00256)	**0,02044** (0,00256)
1/50	**-0,00053** (0,00256*)	**-0,00051** (0,00256*)	**0,00215** (0,00256*)
1/100	**0,00208** (0,00256*)	**-0,00355** (0,00256)	**-0,00076** (0,00256*)
		$K = 140$	
1/5	**0,00139** (0,00290*)	**-0,00448** (0,00289)	**-0,02638** (0,00286)
1/7	**-0,00329** (0,00290)	**-0,00271** (0,00290*)	**-0,01744** (0,00288)
1/10	**-0,00142** (0,00290*)	**0,00115** (0,00290*)	**-0,01154** (0,00289)
1/14	**-0,00090** (0,00290*)	**-0,00208** (0,00290*)	**-0,00648** (0,00290)
1/20	**-0,00496** (0,00290)	**0,00098** (0,00290*)	**-0,00451** (0,00290)
1/30	**0,00238** (0,00290*)	**0,00321** (0,00290)	**-0,00075** (0,00290*)
1/50	**-0,00250** (0,00290*)	**-0,00170** (0,00290*)	**-0,00210** (0,00290*)
1/100	**0,00238** (0,00290*)	**-0,00600** (0,00290)	**-0,00349** (0,00290)

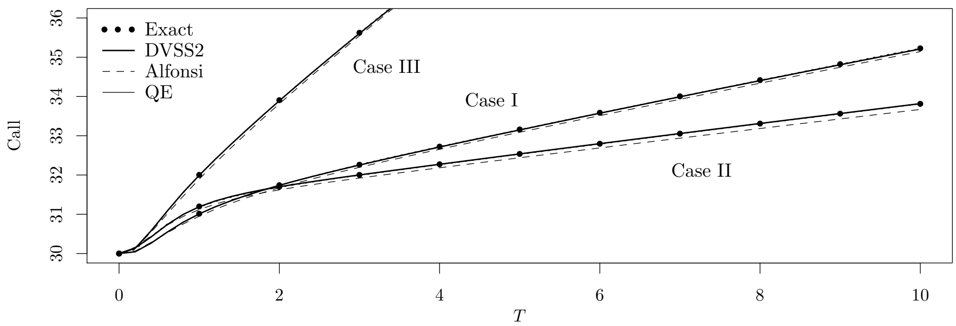
Table 5.1: Estimated bias in the test case I for different strike prices ($K = 70$, $K = 100$, and $K = 140$). Values are written in bold when the bias is in the window of $\pm 1\%$ from the exact value. Numbers in parentheses show the half-length of the 95% confidence interval and are marked with an asterisk if the bias is within this interval.

	$DVSS_2$	QE	Alfonsi
h	$K = 80$		
1/5	**0,00052** (0,00116*)	**0,00719** (0,00116)	0,02554 (0,00116)
1/7	**-0,00023** (0,00116*)	**0,00298** (0,00116)	**0,00447** (0,00116)
1/10	**0,00187** (0,00116)	**0,00145** (0,00116)	**0,00161** (0,00116)
1/14	**0,00107** (0,00116*)	**0,00154** (0,00116)	**0,00167** (0,00116)
1/20	**0,00102** (0,00116*)	**0,00062** (0,00116*)	**0,00079** (0,00116*)
1/30	**0,00087** (0,00116*)	**0,00137** (0,00116)	**0,00085** (0,00116*)
1/50	**0,00083** (0,00116*)	**0,00107** (0,00116*)	**0,00053** (0,00116*)
1/100	**0,00160** (0,00116)	**0,00269** (0,00116)	**-0,00019** (0,00116*)
	$K = 100$		
1/5	**-0,01953** (0,00233)	**-0,00298** (0,00232)	**0,02967** (0,00231)
1/7	**0,00340** (0,00233)	**-0,00165** (0,00233*)	**0,01159** (0,00233)
1/10	**0,00452** (0,00233)	**-0,00066** (0,00233*)	**0,00593** (0,00233)
1/14	**0,00272** (0,00233)	**0,00026** (0,00233*)	**0,00384** (0,00233)
1/20	**0,00227** (0,00233*)	**0,00044** (0,00233*)	**0,00227** (0,00233*)
1/30	**0,00337** (0,00233)	**0,00357** (0,00233)	**0,00263** (0,00233)
1/50	**0,00382** (0,00233)	**-0,00116** (0,00233*)	**0,00035** (0,00233*)
1/100	**0,00496** (0,00233)	**0,00401** (0,00233)	**0,00046** (0,00233*)
	$K = 120$		
1/5	**0,01517** (0,00334)	**-0,00606** (0,00334)	**-0,03900** (0,00333)
1/7	**-0,00598** (0,00334)	**-0,00301** (0,00334*)	**-0,01800** (0,00334)
1/10	**0,00124** (0,00334*)	**-0,00090** (0,00334*)	**-0,00392** (0,00334)
1/14	**0,00200** (0,00334*)	**-0,00143** (0,00334*)	**-0,00186** (0,00334*)
1/20	**0,00293** (0,00334*)	**-0,00031** (0,00334*)	**0,00079** (0,00334*)
1/30	**0,00431** (0,00334)	**0,00707** (0,00334)	**-0,00024** (0,00334*)
1/50	**0,00594** (0,00334)	**-0,00352** (0,00334)	**0,00245** (0,00334*)
1/100	**0,00929** (0,00334)	**0,00303** (0,00334*)	**0,00266** (0,00334*)

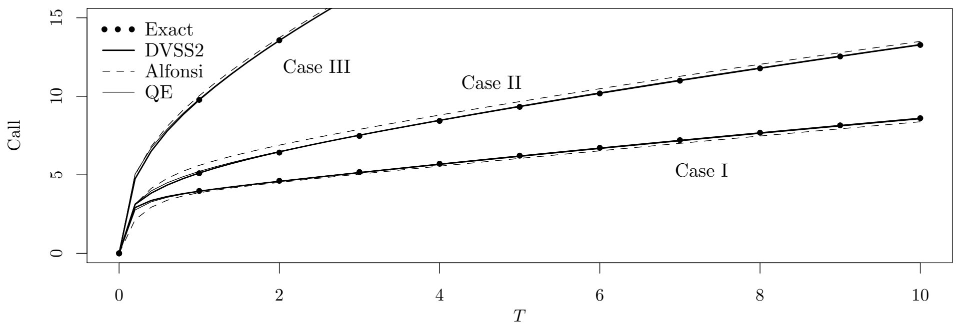
Table 5.2: Estimated bias in the test case IV for different strike prices ($K = 80$, $K = 100$, and $K = 120$). Values are written in bold when the bias is in the window of ±1% from the exact value. Numbers in parentheses show the half-length of the 95% confidence interval and are marked with an asterisk if the bias is within this interval.

	100 000 trajectories in sec.			Relative to $DVSS_2$	
	$DVSS_2$	QE	Alfonsi	QE	Alfonsi
Case I	22.0916	25.0672	30.7702	1.135	1.393
Case II	22.1540	26.4862	31.3246	1.196	1.414
Case III	22.3214	35.6170	33.8316	1.596	1.516
Case IV	22.8598	40.3704	36.2662	1.766	1.586
Case V	22.4612	27.2540	31.6442	1.213	1.409

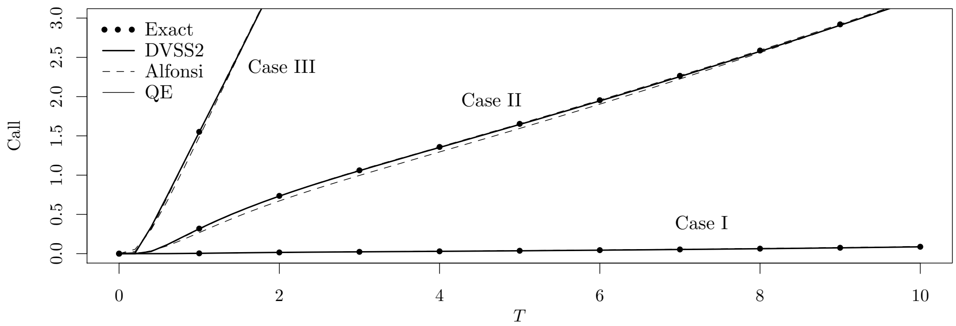
Table 5.3: Average and relative (to $DVSS_2$) computation times



(a) Strike price 70

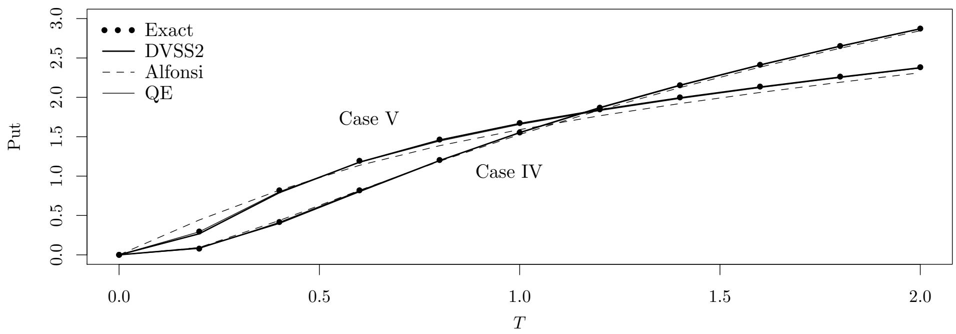


(b) Strike price 100

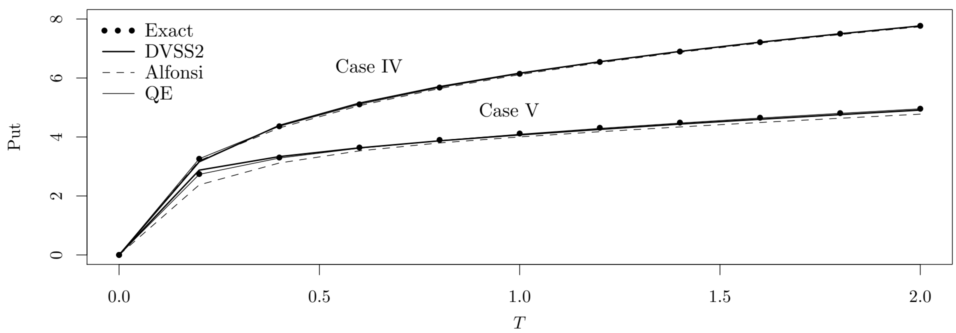


(c) Strike price 140

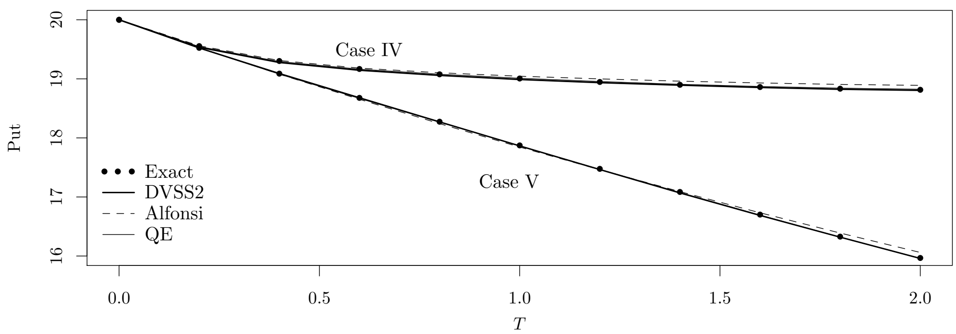
Figure 5.3: Comparison of DVSS_2 with QE and Alfonsi approximations of European option prices in the Heston's stochastic volatility model for different parameters and strike prices (70, 100, 140) with fixed parameters $S_0 = 100$ and $r = 0$. Time step $h = 0.2$.



(a) Strike price 80



(b) Strike price 100



(c) Strike price 120

Figure 5.4: Comparison of DVSS_2 with QE and Alfonsi approximations of European option prices in the Heston's stochastic volatility model for different parameters and strike prices (80, 100, 120) with fixed parameter $S_0 = 100$. Time step $h = 0.2$.

Chapter **6**

Conclusions

In the doctoral dissertation we suggest to construct discretization schemes for the Heston model using “split-step” and moment matching techniques. “Split-step” technique lets to divide the model into deterministic and stochastic parts, so that we need to construct a discretization scheme for the stochastic part only, as the deterministic part is easily solvable. Moment matching helps to construct discrete random variables so that their moments match the moments of weak approximation of the order we aim at.

Although similar ideas can be found in the works of some other authors, for example, Alfonsi [3, 4], to the best of our knowledge, use of only discrete random variables for the construction of the approximation for the Heston model is a completely new idea. Besides, in the works of other authors pursuing similar ideas (see already mentioned [3], or [5]) no rigorous mathematical proof for the convergence is provided.

We have succeeded in constructing simple yet effective first- (DVSS) and second-order (DVSS_2) weak approximations for the solution of the Heston model that use, at each step, only generation of discrete random variables which in many ways outperform other known discretization schemes for the Heston model. To sum up, we compare our approximations to the Andersen QE and Alfonsi ALF_2 and ALF_3, the most popular approximations of the solution of the Heston model to this day, and point out the following advantages:

- *Rigorous proof.* We have rigorously proved that the DVSS approximation scheme is a (strongly potential) first-order and DVSS_2 – a (strongly potential) second-order schemes for the log-Heston model. Although numerical simulations of both QE and Alfonsi schemes show a nice accuracy in option pricing for the Heston model, there are no general theoretical results on the weak convergence of these schemes;

- *Accuracy.* In terms of accuracy, the first-order approximation DVSS in most cases seem to be somewhat lesser when compared to the QE and ALF_2 (or ALF_3) schemes, however, it is performing considerably well as the bias is never outside the window of $\pm 1\%$ from the exact value, which is a sufficient accuracy for most practical applications. Moreover, for small steps (for example, $h = 0.01$), the bias becomes quite similar to that of the QE and Alfonsi schemes. On the other hand, the second-order approximation DVSS_2 clearly outperforms the ALF_2 and ALF_3 and (in many cases) even the QE(-M) approximation schemes;

- *Simplicity.* The approximation schemes (especially the first-order scheme DVSS) presented in this thesis are much simpler to implement than the QE(-M) or Alfonsi schemes, as they use only two- or three-valued discrete random variables (compare Sections 4.3 and 5.3 with [5] or [3]);

- *Lower computational cost.* Numerical examples show that the computation time of the DVSS scheme is from 2.5 to 3.7 times less than that of the QE and Alfonsi schemes (see Table 4.5) and the computation time of the DVSS_2 scheme is from 1.2 to 1.7 times less than that of the QE and Alfonsi schemes (see Table 5.3).

Chapter 7

Appendix

In this chapter, we provide additional calculations which we think would only distract the reader if placed elsewhere in the text.

Moments of Heston model

Some results of this section (but not the formulas of the moments of the Heston model), which we prove for completeness, are known and can be found, for example, in the articles of Andersen and Piterbarg [6] and Fritz and Keller-Ressel [11].

The generator of the log-Heston model (1.2.1) is

$$\begin{aligned} Lu(x,y) &= \frac{y}{2}\left(\frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x}\right)u(x,y) + r\frac{\partial}{\partial x}u(x,y) \\ &+ \frac{\sigma^2 y}{2}\frac{\partial^2}{\partial y^2}u(x,y) + k(\theta - y)\frac{\partial}{\partial y}u(x,y) + \rho\sigma y\frac{\partial^2}{\partial x\partial y}u(x,y). \end{aligned}$$

The function $u(t,x,y) := \mathbb{E}\left(\mathsf{e}^{nX_t}|X_0 = x, Y_0 = y\right)$ satisfies the Kolmogorov backward equation

$$\begin{cases} u'_t = Lu, \\ u(0,x,y) = \mathsf{e}^{nx}. \end{cases}$$

Taking into account that

$$\begin{aligned} u(t,x,y) &= \mathbb{E}\left(\mathsf{e}^{nX_t}|X_0 = x, Y_0 = y\right) \\ &= \mathsf{e}^{nx}\mathbb{E}\left(\mathsf{e}^{nX_t}|X_0 = 0, Y_0 = y\right) = \mathsf{e}^{nx}u(t,0,y) \end{aligned}$$

and denoting

$$f = f_n(t,y) = \mathbb{E}\left(\mathsf{e}^{nX_t}|X_0 = 0, Y_0 = y\right),$$

we notice that

$$\frac{\partial}{\partial x}\left(\mathsf{e}^{nx}f\right) = n\mathsf{e}^{nx}f,$$

$$\left(\frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x}\right)\left(\mathsf{e}^{nx}f\right) = \mathsf{e}^{nx}(n^2-n)f,$$

and

$$\frac{\partial^2 f}{\partial x \partial y}\left(\mathsf{e}^{nx}f\right) = n\mathsf{e}^{nx}\frac{\partial f}{\partial y}.$$

Therefore, we have

$$f'_t = \hat{L}f = \left(\frac{y}{2}\sigma^2\frac{\partial^2}{\partial y^2} + (k(\theta - y) + \rho\sigma n y)\frac{\partial}{\partial y} + \frac{y}{2}(n^2-n) + rn\right)f,$$

$$f_n(0,y) = 1.$$

Presuming that f_n is composed as

$$f_n(t,y) = \mathsf{e}^{\phi(t,n) + y\psi(t,n)},$$

we calculate

$$f'_t = f\phi' + yf\psi',$$

$$\frac{\partial f}{\partial y} = f\psi,$$

and

$$\frac{\partial^2 f}{\partial y^2} = f\psi^2.$$

Comparing both differentiation formulas, we get the equation

$$\phi' + y\psi' = \frac{y}{2}\sigma^2\psi^2 + (k(\theta-y) + \rho\sigma n y)\psi + \frac{y}{2}(n^2-n) + rn.$$

From this we derive the system

$$\begin{cases} \psi' = \frac{\psi^2}{2}\sigma^2 + (\rho\sigma n - k)\psi + \frac{1}{2}(n^2-n), \\ \phi' = k\theta\psi + rn. \end{cases}$$

The first equation is a Riccati differential equation. Denoting $\chi = \rho\sigma n - k$ and $\Delta = \chi^2 - \sigma^2(n^2-n)$, we get:

$$t = \frac{2}{\sigma^2}\int_0^{\psi} \frac{\mathrm{d}\tilde{\psi}}{\left(\tilde{\psi} + \frac{\chi}{\sigma^2}\right)^2 - \frac{\Delta}{\sigma^4}}.$$

Calculating the latter, we get three cases.

- Case $\Delta < 0$. In this case, we have

$$\psi(t) = \frac{\sqrt{-\Delta}}{\sigma^2} \tan\left(\frac{\sqrt{-\Delta}}{2} t + \arctan\left(\frac{\chi}{\sqrt{-\Delta}}\right)\right) - \frac{\chi}{\sigma^2}$$

and

$$\begin{aligned}\phi(t) = &\frac{2k\theta}{\sigma^2} \log\left|\cos\left(\arctan\left(\frac{\chi}{\sqrt{-\Delta}}\right)\right)\right| \\ &- \frac{2k\theta}{\sigma^2} \log\left|\cos\left(\frac{\sqrt{-\Delta}}{2} t + \arctan\left(\frac{\chi}{\sqrt{-\Delta}}\right)\right)\right| \\ &+ \left(rn - \frac{\chi k\theta}{\sigma^2}\right) t.\end{aligned}$$

Note, that in this case, the Heston model moments "explode" at (cf. [6] and [11])

$$t_0 = \frac{2}{\sqrt{-\Delta}} \left(\frac{\pi}{2} - \arctan\left(\frac{\chi}{\sqrt{-\Delta}}\right)\right).$$

- Case $\Delta > 0$. Denoting by ψ_1 and ψ_2 the solutions of the quadratic equation $\frac{\psi^2}{2}\sigma^2 + (\rho\sigma n - k)\psi + \frac{1}{2}(n^2 - n)$, we have

$$\psi(t) = \psi_1 - \frac{\psi_1(\psi_1 - \psi_2)}{\psi_1 - \psi_2 \exp\left\{-\frac{\sigma^2(\psi_1 - \psi_2)}{2} t\right\}}$$

and

$$\begin{aligned}\phi(t) &= k\theta\psi_2 t + rnt - \frac{2k\theta}{\sigma^2} \log\left(\frac{\psi_1}{\psi_2} - \mathrm{e}^{\frac{\sigma^2}{2} t(\psi_2 - \psi_1)}\right)\Bigg|_0^t \\ &= t(k\theta\psi_2 + rn) - \frac{2k\theta}{\sigma^2} \log\left(\frac{\psi_1 - \psi_2 \mathrm{e}^{\frac{\sigma^2}{2} t(\psi_2 - \psi_1)}}{\psi_1 - \psi_2}\right).\end{aligned}$$

Note that, in this case, the moments "explode" when $\chi < 0$ at (cf. [6] and [11])

$$t_0 = \frac{2}{\sigma^2(\psi_1 - \psi_2)} \log\left(\frac{\psi_2}{\psi_1}\right).$$

- Case $\Delta = 0$.

$$\psi(t) = \frac{\chi^2 t}{\sigma^2(2 - \chi t)}$$

and

$$\phi(t) = \frac{2\chi \log 2 - \chi \log (2 - \chi t)^2 - \chi t}{\sigma^2}$$

Note that, in this case, the moments "explode" at (cf. [6] and [11])

$$t_0 = \frac{2}{\chi}.$$

.

Bibliography

[1] H. Albrecher, P. Mayer, W. Schoutens, and J. Tistaert, *The little Heston trap*, Wilmott Magazine (March 2007), 83–92.

[2] A. Alfonsi, *On the discretization schemes for the CIR (and Bessel squared) processes*, Monte Carlo Methods and Applications **11** (2005), no. 4, 355–384.

[3] ———, *High order discretization schemes for the CIR process: Application to affine term structure and Heston models*, Mathematics of Computation **79** (2010), no. 269, 209–237.

[4] ———, *Affine Diffusions and Related Processes: Simulation, Theory and Applications*, 1 ed., Bocconi & Springer Series, vol. 6, Springer, 2015.

[5] L. Andersen, *Simple and efficient simulation of the Heston stochastic volatility model*, Journal of Computational Finance **11** (2008), no. 3, 1–42.

[6] L. B. G. Andersen and V. V. Piterbarg, *Moment Explosions in Stochastic Volatility Models*, Finance and Stochastics **11** (2007), no. 1, 29–50.

[7] M. Broadie and Ö. Kaya, *Exact Simulation of Stochastic Volatility and Other Affine Jump Diffusion Processes*, Operations Research **54** (2006), no. 2, 217–231.

[8] J. H. Chan and M. Joshi, *Fast and accurate long-stepping simulation of the Heston stochastic volatility model*, Journal of Computational Finance **16** (2013), 47–98.

[9] P. Čížek, W. K. Härdle, and R. Weron, *Statistical Tools for Finance and Insurance*, 2 ed., Springer, 2011.

[10] J. C. Cox, J. E. Ingersoll, and S. A. Ross, *A Theory of the Term Structure of Interest Rates*, Econometrica **53** (1985), no. 2, 385–407.

[11] P. Fritz and M. Keller-Ressel, *Moment Explosions in Stochastic Volatility Models*, Encyclopedia of Quantitative Finance (2010), 1247–1253.

[12] R. Frontczak, *Valuing options in Heston's Stochastic Volatility Model: Another Analytical Approach*, Journal of Applied Mathematics **2011** (2011).

[13] A. Van Haastrecht and A. Pelsser, *Efficient, almost exact simulation of the Heston stochastic volatility model*, International Journal of Theoretical and Applied Finance **13** (2010), no. 01, 1–43.

[14] S. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The Review of Financial Studies **6** (1993), no. 2, 327–343.

[15] D. J. Higham and X. Mao, *Convergence of Monte Carlo simulations involving the mean-reverting square root process*, Journal of Computational Finance **8** (2005), no. 3, 35–61.

[16] J. C. Hull and A. White, *The Pricing of Options on Assets with Stochastic Volatilities*, Journal of Finance **42** (1987), no. 2, 281–300.

[17] C. Kahl and P. Jackel, *Fast strong approximation Monte Carlo schemes for Stochastic Volatility Models*, Quantitative Finance **6** (2006), no. 6, 513–536.

[18] P. Kloeden and A. Neuenkirch, *Convergence of numerical methods for stochastic differential equations in Mathematical Finance*, Recent Developments in Computational Finance: Foundations, Algorithms and Applications (T. Gerstner and P. Kloeden, eds.), Interdisciplinary Mathematical Sciences Series, vol. 14, World Scientific, Singapur, 2013, pp. 49–80.

[19] P. Kloeden, E. Platen, and H. Schurz, *Numerical solution of SDE through computer experiments*, 1 ed., Universitext, Springer-Verlag, 1994.

[20] A. Lenkšas and V. Mackevičius, *Option pricing in Heston model by means of weak approximations*, Lietuvos Matematikos Rinkinys **54** (2013), 27–32.

[21] ______, *A second-order weak approximation of Heston model by discrete random variables*, Lithuanian Mathematical Journal **55** (2015), 555–572.

[22] ______, *Weak approximation of Heston model by discrete random variables*, Mathematics and Computers in Simulation **113** (2015), 1–15.

[23] R. Lord, R. Koekkoek, and D. Van Dijk, *A comparison of biased simulation schemes for stochastic volatility models*, Quantit. Finance **10** (2010), no. 2, 177–194.

[24] V. Mackevičius, *On positive approximations of positive diffusions*, Lietuvos Matematikos Rinkinys **47** (2007), no. spec. issue, 58–62.

[25] ______, *On weak approximations of (a, b)-invariant diffusions*, Mathematics and Computers in Simulation **74** (2007), no. 1, 20–28.

[26] ______, *Introduction to Stochastic Analysis: Integrals and Differential Equations*, 1 ed., Wiley, 2011.

[27] ______, *Weak approximation of CIR equation by discrete random variables*, Lith. Math. J. **51** (2011), no. 3, 385–401.

[28] S. Ninomiya and N. Victoir, *Weak approximation of stochastic differential equations and application to derivative pricing*, Applied Mathematical Finance **15** (2008), no. 2, 107–121.

[29] S. Del Baño Rollin, A. Ferreiro-Castilla, and F. Utzet, *A new look at the Heston characteristic function*, arXiv:0902.2154D (2009).

[30] ______, *On the density of log-spot in the Heston volatility model*, Stochastic Processes and their Applications **120** (2010), no. 10, 2037 – 2063.

[31] F. D. Rouah and S. L. Heston, *The Heston Model and its Extensions in Matlab and C#*, Wiley Finance, Wiley, 2013.

[32] L. O. Scott, *Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application*, Journal of Financial and Quantitative Analysis **22** (1987), 419–438.

[33] E. M. Stein and J. C. Stein, *Stock Price Distributions with Stochastic Volatility: An Analytic Approach*, Review of Financial Studies **4** (1991), 727–752.

[34] O. Vasicek, *An equilibrium characterization of the term structure*, Journal of Financial Economics **5** (1977), no. 2, 177–188.