Numerical Methods for Foreign Exchange Option Pricing under Hybrid Stochastic and Local Volatility Models

Andrei Sorin Cozma

Lady Margaret Hall

University of Oxford

Thesis Submitted for the Degree of

Doctor of Philosophy

Hilary 2017

Acknowledgements

First and foremost, I want to thank my supervisor, Prof. Christoph Reisinger, for his excellent guidance, for his patience and encouragement throughout my doctoral studies, and for giving me the confidence to perform research on a challenging topic of interest. His vast knowledge, his passion for research and his genuine care for his students have inspired me greatly and helped me grow as a researcher and as a person, and I consider myself very fortunate to be one of his students.

I would like to thank Prof. Mike Giles, Prof. Ben Hambly, Dr. Gergely Gyurko and Dr. Tigran Nagapetyan for their rigorous feedback and insightful comments from my Transfer and Confirmation of Status.

I would also like to thank my friend, colleague and co-author, Matthieu Mariapragassam, for his enthusiasm, perseverance and extensive knowledge on the valuation of FX options and on model calibration. I am grateful for our interesting and fruitful discussions over coffee and on even more informal occasions, and for the wonderful distractions over the years.

A special thanks goes to my family, who continuously supported and encouraged me during my studies and helped me get through some difficult times.

I also want to thank the Engineering and Physical Sciences Research Council for their financial support, which allowed me to undertake this PhD project and enabled me to attend and present my work at several conferences and workshops.

Last but not least, I want to thank the Mathematical Institute for providing a great working environment.

Statement of Originality

I hereby declare that this thesis contains no material which has been accepted or is currently being submitted for the award of any other degree, diploma, certificate or other qualification in this University or elsewhere and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes two original papers published in peer-reviewed journals [21; 22], which are co-authored with my supervisor, Prof. Christoph Reisinger (in Chapters 2 and 4). Also included in this thesis are two original, unpublished papers [19; 20], which are co-authored with a colleague, Matthieu Mariapragassam, and with my supervisor (in Chapter 3). The ideas, development and writing up of [19] were the principal responsibility of myself, the student, working within the University of Oxford under the supervision of Prof. Christoph Reisinger, and the paper was reproduced in full in this thesis. The ideas, development and writing up of [20] were the principal responsibility of Matthieu Mariapragassam, and a section of the paper was reproduced in part in this thesis. All contributions that are not entirely my own work, namely in Subsection 3.4.3 and Appendix 3.6, are outlined at the beginning of Chapter 3.

Abstract

In this thesis, we study the FX option pricing problem and put forward a 4-factor hybrid stochastic-local volatility model. The model, which describes the dynamics of an exchange rate, its volatility and the domestic and foreign short rates, allows for a perfect calibration to European options and has a good hedging performance. Due to the high-dimensionality of the problem, we propose a Monte Carlo simulation scheme that combines the full truncation Euler scheme for the stochastic volatility component and the stochastic short rates with the log-Euler scheme for the exchange rate. We analyze exponential integrability properties of Euler discretizations for the square-root process driving the stochastic volatility and the short rates, properties which play a key role in establishing the finiteness of moments and the strong convergence of numerical approximations for a large class of stochastic differential equations in finance, including the ones studied in this thesis. Hence, we prove the strong convergence of the exchange rate approximations and the convergence of Monte Carlo estimators for a number of vanilla and exotic options. Then, we calibrate the model to market data and discuss its fitness for pricing FX options. Next, due to the relatively slow convergence of the Monte Carlo method in the number of simulations, we examine a variance reduction technique obtained by mixing Monte Carlo and finite difference methods via conditioning. We consider a purely stochastic version of the model and price vanilla and exotic options by simulating the paths of the volatility and the short rates, and then evaluating the "inner" Black–Scholes-type expectation by means of a partial differential equation. We prove the convergence of numerical approximations and carry out a theoretical variance reduction analysis. Finally, we illustrate the efficiency of the method through a detailed quantitative assessment.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

With the technological advances making the trading of options more accessible to investors, derivative products have become essential tools for corporations and investors alike and the volume of exchange-traded derivatives contracts traded on exchanges worldwide in 2013 reached 21.5 billion, with 12.1 billion futures and 9.3 billion options. Of the latter, 3.78 billion were single stock options and 403 million were foreign exchange (FX) options [88]. In this thesis, we study the pricing problem for FX options and define the underlying spot FX rate as the number of units of domestic currency per unit of foreign currency.

With the Black–Scholes model [10] becoming obsolete as a pricing model after the stock market crash of 1987, the search for a fast and accurate method to price options has been a popular topic in mathematical finance over the past few decades, for practitioners and researchers alike, and the mathematical models employed in derivative pricing have become more and more sophisticated. In particular, the class of stochastic-local volatility (SLV) models, introduced by Jex et al. [60], Lipton [69] and Lipton and McGhee [71], among others, have become very popular in the financial sector in recent years. They contain a stochastic volatility component as well as a local volatility component (called the leverage function) and combine advantages of the two. Derman and Kani [27], Dupire [31] and Rubinstein [81] introduced the local volatility and defined it as the deterministic function of time and spot FX rate that is consistent with observed market prices of European options. While a local volatility model can provide a perfect fit to the implied volatility surface via Dupire's formula, it predicts unrealistic dynamics for the spot FX rate volatility and has a poor hedging performance. According to Ren et al. [80], Tian et al. [89] and van der Stoep et al. [90], the general SLV model allows for a better calibration to European options and improves the pricing and risk-management performance when compared to pure local volatility (LV) or pure stochastic volatility (SV) models, and Lipton et al. [70] recently referred to SLV models as the de facto standard for pricing FX options. We focus on the Heston SLV model because the Cox–Ingersoll–Ross (CIR) process [18] for the variance is widely used in the industry due to its desirable properties, such as mean-reversion and non-negativity, and because semi-analytical formulae are available for calls and puts under the Heston model [46] and can help calibrate the parameters easily. The local volatility component allows for a perfect calibration to the market prices of vanilla options. At the same time, the stochastic volatility component already provides built-in smiles and skews which give a rough calibration, such that a relatively flat leverage function suffices for a perfect calibration.

In order to improve the pricing and hedging of FX options, we introduce stochastic domestic and foreign short interest rates into the model. Assuming constant interest rates is appealing due to its simplicity and does not lead to a serious mispricing of options with short maturities. However, empirical results [93] have confirmed that the constant interest rate assumption is inappropriate for long-dated FX products, and the effect of interest rate volatility can be as relevant as that of the FX rate volatility for longer maturities. There has been a great deal of research carried out in the area of option pricing with stochastic volatility and interest rates in the past couple of years. For instance, Van Haastrecht et al. [93] extended the model of Schöbel and Zhu [82] to FX derivatives by including stochastic interest rates, a model that benefits from analytical tractability even in a full correlation setting due to the processes being Gaussian, whereas Ahlip and Rutkowski [1], Grzelak and Oosterlee [42] and Van Haastrecht and Pelsser [94] examined the Heston–2CIR/Vasicek hybrid models and concluded that a full correlation structure gives rise to a non-affine model even under a partial correlation of the driving Brownian motions. The literature on option pricing with local volatility and stochastic short rates is not as rich. Recently, Deelstra and Rayee [25] and van der Stoep et al. [91] proposed different approaches to evaluate the local volatility function under 1-factor Hull–White short rates [50].

There has also been an increasing interest in the calibration to vanilla options of models with stochastic and local volatility and stochastic interest rate dynamics. For instance, Guyon and Labordére [44] examined Monte Carlo-based calibration methods for a 3-factor SLV equity model with a stochastic domestic short rate and discrete dividends, whereas Deelstra and Rayee [25] examined a 4-factor hybrid SLV model with stochastic domestic and foreign short rates and proposed different approaches to calibrate the leverage function. Hambly et al. [45] took a different path and discussed the calibration of a 2-factor SLV model to barriers.

The model of Cox et al. [18] is popular when modeling short rates because the square-root (CIR) process admits a unique strong solution, is mean-reverting and analytically tractable. Cox et al. found the conditional distribution to be noncentral chi-squared, and Broadie and Kaya [14] proposed an efficient exact simulation scheme for the square-root process based on acceptance-rejection sampling. However, their algorithm presents a number of disadvantages such as complexity and lack of speed, and it is not fit to price strongly path-dependent options that require the value of the FX rate at a large number of time points. Moreover, in the context of a stochastic-local volatility model, the correlations between the underlying processes make it difficult to simulate a noncentral chi-squared increment together with a correlated increment for the FX rate and the short rates, if applicable. As of late, the non-negativity of the CIR process is considered to be less desirable when modeling short rates. On the one hand, the central banks have significantly reduced the interest rates since the 2008 financial crisis and it is now commonly accepted that interest rates need not be positive. On the other hand, if interest rates dropped too far below zero, then large amounts of money would be withdrawn from banks and government bonds, putting a severe squeeze on deposits. Hence, we model the domestic and foreign short rates using the shifted CIR (CIR++) process of Brigo and Mercurio [12]. The CIR++ model allows the short rates to become negative and can fit any observed term structure exactly while preserving the analytical tractability of the original model for bonds, caps, swaptions and other basic interest rate products.

Therefore, in this work, we put forward the Heston–Cox–Ingersoll–Ross++ stochastic-local volatility (Heston–2CIR++ SLV) model to price FX options. Independent of the correlation structure, the model is non-affine and so a closed-form solution to the European option valuation problem is not available. Hence, we see ourselves forced to turn to finite difference or Monte Carlo methods. Finite difference methods are popular in finance and when the evolution of the exchange rate is governed by a complex system of stochastic differential equations (SDEs), it all comes down to solving a higher-dimensional partial differential equation (PDE). This can prove to be difficult due to the curse of dimensionality, because the number of grid points required increases exponentially with the number of dimensions. Monte Carlo algorithms scale linearly with the dimension and can handle path-dependent features easily; also, there are numerous discretization schemes available, like the simple Euler–Maruyama scheme, see, e.g., Glasserman [41]. This makes Monte Carlo methods more attractive for high-dimensional models such as the ones studied here, and the relatively slow convergence $\mathcal{O}\big(M^{-\frac{1}{2}}\big)$ in the number of simulations M raises the question of how to find an efficient variance reduction technique.

The mixed Monte Carlo/partial differential equation solver [72] is a good alternative to the classical Monte Carlo algorithm and, under certain circumstances, can lead to a substantial variance reduction. The idea behind the mixed Monte Carlo/PDE method is to write the option values as nested conditional expectations. Then, the innermost expectation is evaluated analytically or using finite differences, and the outer expectation is evaluated by simulation. The earliest related published work is by Hull and White [49], who considered a European call option under a 2-dimensional stochastic volatility model with uncorrelated asset price and volatility and proved that, conditional on the integral of the variance process, the asset price is lognormally distributed, and hence the option price can be expressed as a Black–Scholes price. Willard [96] extended the analysis of Hull and White to path-independent options under stochastic volatility and instantaneously correlated factors, and used the smoothness of the "conditional price" to calculate price sensitivities (the Greeks). Willard also employed quasi-Monte Carlo (low-discrepancy) methods to further reduce the variance of price estimates, but with no effect on the discretization bias. The mixed Monte Carlo/PDE method develops this conditioning technique, known as conditional Monte Carlo, and allows the combined use of Monte Carlo and finite difference methods for the valuation of path-dependent contracts.

In summary, the focus of this thesis is on FX option pricing and the main topics of interest are: the calibration and the Monte Carlo pricing of FX options under hybrid stochastic and local volatility models; the mixing of Monte Carlo and finite difference methods via conditioning techniques for the valuation of vanilla and exotic options; the convergence and the stability of moments of discretization schemes for high-dimensional systems of SDEs.

In Chapter 2, we analyze exponential integrability properties of the CIR process and its Euler–Maruyama discretizations with various types of truncation and reflection at 0. These properties play a key role in establishing the finiteness of moments and the strong convergence of numerical approximations for a large class of SDEs arising in finance, including the popular Heston stochastic volatility model. We prove that both implicit and explicit Euler–Maruyama discretizations for the CIR process preserve the exponential integrability of the exact solution for a wide range of parameters, and find lower bounds on the explosion time. Then, we look at the Heston model and examine the stability of moments of the process and its approximations numerically.

In Chapter 3, we study the Heston–2CIR++ SLV model in the context of FX markets and propose a Monte Carlo simulation scheme that combines the full truncation Euler scheme for the stochastic volatility component and the stochastic domestic and foreign short interest rates with the log-Euler scheme for the exchange rate. Under a full correlation structure and a realistic set of assumptions on the so-called leverage function, we establish the strong convergence of the exchange rate approximations and deduce the convergence of Monte Carlo estimators for a number of vanilla and path-dependent options. Then, we carry out an efficient model calibration to EURUSD market data and discuss the fitness of the model for pricing FX options.

In Chapter 4, we consider the valuation of European and path-dependent options in FX markets when the currency exchange rate evolves according to the Heston model combined with the CIR dynamics for the stochastic domestic and foreign short interest rates (i.e., the Heston–2CIR model). The mixed Monte Carlo/PDE method requires that we simulate only the paths of the squared volatility and the two interest rates, while an "inner" Black–Scholes-type expectation is evaluated by means of a PDE. This can lead to a substantial variance reduction and complexity improvements under certain circumstances depending on the contract and the model parameters. We establish the uniform boundedness of moments of the exchange rate process and its approximation and prove the strong convergence of the latter. Then, we carry out a variance reduction analysis and find accurate approximations for quantities of interest. All theoretical contributions can be extended to multi-factor short rates with a term structure in a straightforward manner. Finally, we illustrate the efficiency of the method for the 4-factor Heston–2CIR model through a detailed quantitative assessment.

Chapter 2

Exponential Integrability Properties of the Cox–Ingersoll–Ross Process

2.1. Introduction

Let $\big(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0}, \mathbb{P}\big)$ be a filtered probability space satisfying the usual conditions and let $W = \big(W_t\big)_{t\geq 0}$ be a 1-dimensional standard $\mathcal{G}_t$-adapted Brownian motion. The Cox–Ingersoll–Ross process was originally proposed in [18] for short-term interest rate modeling, and is the solution to the following SDE:

$$dy_t = k(\theta - y_t)dt + \xi\sqrt{y_t}\,dW_t, \tag{2.1}$$

where y_0, k, θ and ξ are strictly positive real numbers. According to [62], (2.1) admits a unique strong solution, which is strictly positive when the Feller condition is satisfied, i.e., when $2k\theta \geq \xi^2$. The desirable features of the CIR process under consideration, such as non-negativity and mean-reversion, make it very popular when modeling interest rates or variances, e.g., in Heston's stochastic volatility model [46], which is one of the most popular equity derivatives pricing models among practitioners. The Feller condition is typically satisfied in practice in the former case, but often fails to hold in the latter case.

The conditional distribution of the CIR process is noncentral chi-squared and hence its increments can be simulated exactly. On the other hand, discretization schemes are usually preferred when the entire sample path of the CIR process has to be simulated, or when the process is part of a system of SDEs. For instance, when pricing path-dependent financial derivatives written on an underlying process $S = (S_t)_{t\in[0,T]}$ modeled by a d-dimensional SDE, with CIR dynamics in one or more dimensions, we need to evaluate

$$U = \mathbb{E}\left[f(S)\right], \tag{2.2}$$

where $f : \mathcal{C}([0,T], \mathbb{R}^d) \to \mathbb{R}$ is the discounted payoff. In particular, this class of SDEs contains the popular Heston model and extensions thereof, such as stochastic interest rates [1; 42] or stochastic-local volatility [90]. However, one can rarely find an explicit formula for the quantity in (2.2), in which case we approximate the solution to the SDE via a discretization scheme and employ Monte Carlo simulation methods [41]. Since we cannot use the standard Euler–Maruyama scheme to approximate $(y_t)_{t\in[0,T]}$ defined in (2.1) because of the non-zero probability of the approximation process becoming negative, we set it equal to zero when it turns negative (absorption fix) or reflect it in the origin (reflection fix). An overview of the explicit Euler schemes considered thus far in the literature can be found in [73]. Alternatively, we can use an implicit scheme to discretize the CIR process.

Even though weak convergence is important when estimating expectations of payoffs, strong convergence may be required for complex path-dependent derivatives and plays a crucial role in multilevel Monte Carlo methods [38]. An important step in deriving strong convergence is proving the finiteness of moments of order higher than 1 of the process and its approximation [48]. In addition, for a number of stochastic volatility models, moments of order higher than 1 can explode in finite time [7]. This, however, can cause serious problems in practice when valuing securities whose payoffs have superlinear growth, as is the case with some commonly traded fixed income contracts. For instance, the risk-neutral valuation of CMS swaps and caps or Eurodollar futures contracts involves the evaluation of the second moment [7]. Therefore, moment explosions may lead to infinite prices of derivatives. The same issue can also be observed for Euler approximations of SDEs with superlinearly growing drift or diffusion coefficients [54; 55].

Hence, we need to examine the stability of moments of the actual and the approximated processes, a problem directly related to the exponential integrability of the CIR process and its discretization. The results in the literature on the exponential integrability of numerical schemes for SDEs are few and far in between. Hutzenthaler et al. [57] identified a class of stopped increment-tamed Euler approximations for nonlinear systems of SDEs and proved that they preserve the exponential integrability of the exact solution under some mild assumptions, unlike the explicit, the linear-implicit or some tamed Euler schemes, which rarely do. Here, we consider a different exponential functional of the CIR process. To the best of our knowledge, the work in this chapter is the first to establish that the drift-implicit and a number of explicit Euler discretizations for the CIR process preserve exponential integrability properties (up to a critical time).

This chapter is structured as follows. In Section 2.2, we discuss the discretization schemes. In Section 2.3, we establish the uniform exponential integrability of functionals of the CIR process and its implicit and explicit Euler discretizations. Section 2.4 examines moment stability for a particular model (the Heston model) in more detail. Finally, Section 2.5 summarizes the results and outlines possible future work.

2.2. The discretization schemes

The classical Euler–Maruyama scheme does not preserve the non-negativity of the process and hence it is not well-defined when applied to (2.1) directly because of the square-root diffusion coefficient. A number of corrections have been proposed in the literature, by either setting the process equal to zero when it turns negative, or by reflecting it in the origin. Consider a uniform grid: $\delta t = T/N$, $t_n = n\delta t$ for all $n \in \{0, 1, ..., N\}$. The *partial truncation Euler (PTE) scheme*

$$\tilde{y}_{t_{n+1}} = \tilde{y}_{t_n} + k(\theta - \tilde{y}_{t_n})\delta t + \xi\sqrt{\tilde{y}^+_{t_n}}\,\delta W_{t_n}, \tag{2.3}$$

where $y^+ = \max(0, y)$ and $\delta W_{t_n} = W_{t_{n+1}} - W_{t_n}$, was proposed in [24], whereas the *full truncation Euler (FTE) scheme*

$$\tilde{y}_{t_{n+1}} = \tilde{y}_{t_n} + k(\theta - \tilde{y}^+_{t_n})\delta t + \xi\sqrt{\tilde{y}^+_{t_n}}\,\delta W_{t_n} \tag{2.4}$$

was studied in [73]. The *absorption (ABS) scheme* reads as

$$\tilde{y}_{t_{n+1}} = \tilde{y}^+_{t_n} + k(\theta - \tilde{y}^+_{t_n})\delta t + \xi\sqrt{\tilde{y}^+_{t_n}}\,\delta W_{t_n}. \tag{2.5}$$

For the schemes (2.3) – (2.5), the piecewise constant continuous-time approximation is defined as $\bar{y}_t = \tilde{y}^+_{t_n}$, whenever $t \in [t_n, t_{n+1})$. The *reflection (REF) scheme*

$$\tilde{y}_{t_{n+1}} = \tilde{y}_{t_n} + k(\theta - \tilde{y}_{t_n})\delta t + \xi\sqrt{|\tilde{y}_{t_n}|}\,\delta W_{t_n} \tag{2.6}$$

was introduced in [47], and we define $\bar{y}_t = |\tilde{y}_{t_n}|$, whenever $t \in [t_n, t_{n+1})$. The *symmetrized Euler (SYM) scheme*

$$\tilde{y}_{t_{n+1}} = \left|\tilde{y}_{t_n} + k(\theta - \tilde{y}_{t_n})\delta t + \xi\sqrt{\tilde{y}_{t_n}}\,\delta W_{t_n}\right| \tag{2.7}$$

was studied in [11], and we let $\bar{y}_t = \tilde{y}_{t_n}$, whenever $t \in [t_n, t_{n+1})$. Finally, assuming that the Feller condition holds (i.e., $2k\theta \geq \xi^2$) and applying Itô's formula to $x_t = \sqrt{y_t}$ leads to

$$dx_t = \left(\alpha x_t^{-1} + \beta x_t\right)dt + \gamma\, dW_t, \tag{2.8}$$

where

$$\alpha = \frac{4k\theta - \xi^2}{8}, \ \beta = -\frac{k}{2} \ \text{ and } \ \gamma = \frac{\xi}{2}. \tag{2.9}$$

The *drift-implicit (square-root) Euler scheme*

$$\tilde{x}_{t_{n+1}} = \tilde{x}_{t_n} + \left(\alpha\tilde{x}^{-1}_{t_{n+1}} + \beta\tilde{x}_{t_{n+1}}\right)\delta t + \gamma\delta W_{t_n} \tag{2.10}$$

was proposed in [3] and later studied in [26]. If $4k\theta > \xi^2$, then $\alpha > 0$ and, since $\beta < 0$, (2.10) has the unique positive solution

$$\tilde{x}_{t_{n+1}} = \frac{\tilde{x}_{t_n} + \gamma\delta W_{t_n}}{2(1 - \beta\delta t)} + \sqrt{\frac{(\tilde{x}_{t_n} + \gamma\delta W_{t_n})^2}{4(1 - \beta\delta t)^2} + \frac{\alpha\delta t}{1 - \beta\delta t}}. \tag{2.11}$$

This method is also called the *backward Euler–Maruyama (BEM) scheme* [76]. We let $\bar{y}_t = \tilde{x}^2_{t_n}$, whenever $t \in [t_n, t_{n+1})$.

2.3. Exponential integrability

The goal of this chapter is to establish the exponential integrability of functionals of the CIR process and its approximations. To this end, let $(\bar{y}_t)_{t\in[0,T]}$ be the piecewise constant continuous-time approximation of $(y_t)_{t\in[0,T]}$ corresponding to one of the discretization schemes from (2.3) – (2.7) or (2.11) and, for some $\lambda, \mu \in \mathbb{R}$, define

$$\Theta_t \equiv \exp\left\{\lambda \int_0^t y_u \, du + \mu \int_0^t \sqrt{y_u}\, dW_u\right\}, \quad \forall t \in [0,T], \tag{2.12}$$

$$\bar{\Theta}_t \equiv \exp\left\{\lambda \int_0^t \bar{y}_u \, du + \mu \int_0^t \sqrt{\bar{y}_u}\, dW_u\right\}, \quad \forall t \in [0,T], \tag{2.13}$$

and

$$\varDelta \equiv \lambda + \frac{1}{2}\mu^2. \tag{2.14}$$

Lemma 2.1. *Independent of the discretization scheme $\bar{y}$ employed,*

$$\sup_{t\in[0,T]} \mathbb{E}\left[\bar{\Theta}_t\right] = \begin{cases} 1 & \text{if } \varDelta \le 0, \\ \mathbb{E}\left[\bar{\Theta}_T\right] & \text{if } \varDelta > 0. \end{cases} \tag{2.15}$$

Proof. For convenience of notation, denote $\mathbb{E}_t\big[\,\cdot\,\big] \equiv \mathbb{E}\big[\,\cdot\,|\mathcal{G}_t\big]$. Assuming that $t \in [t_n, t_{n+1}]$ and conditioning on the σ-algebra $\mathcal{G}_{t_n}$, we find that

$$\mathbb{E}_{t_n}\left[\bar{\Theta}_t\right] = \exp\left\{\lambda \int_0^{t_n} \bar{y}_u \, du + \mu \int_0^{t_n} \sqrt{\bar{y}_u}\, dW_u\right\} \exp\left\{(t-t_n)\varDelta \bar{y}_{t_n}\right\}. \tag{2.16}$$

Since $\bar{y}$ is always non-negative, if $\varDelta \le 0$, then $\mathbb{E}_{t_n}\left[\bar{\Theta}_t\right] \le \mathbb{E}_{t_n}\left[\bar{\Theta}_{t_n}\right]$. From the law of iterated expectations,

$$\mathbb{E}\left[\bar{\Theta}_t\right] \le \mathbb{E}\left[\bar{\Theta}_{t_n}\right] \le \mathbb{E}\left[\bar{\Theta}_{t_{n-1}}\right] \le \ldots \le \mathbb{E}\left[\bar{\Theta}_0\right]. \tag{2.17}$$

On the other hand, if $\varDelta > 0$, then $\mathbb{E}_{t_n}\left[\bar{\Theta}_t\right] \le \mathbb{E}_{t_n}\left[\bar{\Theta}_{t_{n+1}}\right]$ and hence

$$\mathbb{E}\left[\bar{\Theta}_t\right] \le \mathbb{E}\left[\bar{\Theta}_{t_{n+1}}\right] \le \mathbb{E}\left[\bar{\Theta}_{t_{n+2}}\right] \le \ldots \le \mathbb{E}\left[\bar{\Theta}_T\right], \tag{2.18}$$

which concludes the proof. ∎

2.3.1. The Cox–Ingersoll–Ross process

The exponential integrability result below is an extension of Proposition 3.1 in [7] to the case $\varDelta \le 0$. When $\varDelta > 0$, we can derive Proposition 2.2 below directly from [7], by writing the moments of the Heston model in terms of $\mathbb{E}\big[\Theta_T\big]$. Our proof takes a fairly different approach and is included for completeness.

Proposition 2.2. *The first moment of the exponential functional of the CIR process defined in* (2.12) *is uniformly bounded up to a critical time* T^**, i.e.,*

$$\sup_{t\in[0,T]} \mathbb{E}\left[\Theta_t\right] < \infty, \quad \forall T < T^*, \tag{2.19}$$

and

$$\mathbb{E}\left[\Theta_T\right] = \infty, \quad \forall T \geq T^*. \tag{2.20}$$

If $\varDelta \leq 0$*, then* $T^* = \infty$*, whereas if* $\varDelta > 0$*, then* T^* *is as given below.*

(1) When $k < \xi(\mu - \sqrt{2\varDelta})$*,*

$$T^* = \frac{1}{\vartheta}\log\left(\frac{\mu\xi - k + \vartheta}{\mu\xi - k - \vartheta}\right), \quad \vartheta = \sqrt{(\mu\xi - k)^2 - 2\xi^2\varDelta}. \tag{2.21}$$

(2) When $k = \xi(\mu - \sqrt{2\varDelta})$*,*

$$T^* = \frac{2}{\mu\xi - k}. \tag{2.22}$$

(3) When $\xi(\mu - \sqrt{2\varDelta}) < k < \xi(\mu + \sqrt{2\varDelta})$*,*

$$T^* = \frac{2}{\hat{\vartheta}}\left[\frac{\pi}{2} - \arctan\left(\frac{\mu\xi - k}{\hat{\vartheta}}\right)\right], \quad \hat{\vartheta} = \sqrt{2\xi^2\varDelta - (\mu\xi - k)^2}. \tag{2.23}$$

(4) When $k \geq \xi(\mu + \sqrt{2\varDelta})$*,*

$$T^* = \infty. \tag{2.24}$$

Proof. Since

$$\int_0^t \sqrt{y_u}\, dW_u = \frac{1}{\xi}y_t - \frac{1}{\xi}(y_0 + k\theta t) + \frac{k}{\xi}\int_0^t y_u\, du \tag{2.25}$$

from (2.1),

$$\Theta_t = \exp\left\{-(y_0 + k\theta t)\hat{\lambda}\right\}\exp\left\{\hat{\lambda}y_t + \hat{\mu}\int_0^t y_u\, du\right\}, \tag{2.26}$$

where

$$\hat{\lambda} = \frac{\mu}{\xi}, \quad \hat{\mu} = \lambda + \frac{\mu k}{\xi}. \tag{2.27}$$

Since the first exponential on the right-hand side of (2.26) is a continuous function of time and since $(y_t)_{t\in[0,T]}$ is a stationary Markov process, it suffices to prove the finiteness of the supremum over τ of

$$\begin{aligned} F(\tau, y) &\equiv \mathbb{E}_t\left[\exp\left\{\hat{\lambda}y_T + \hat{\mu}\int_t^T y_u\, du\right\}\Big|\, y_t = y\right] &(2.28)\\ &= \mathbb{E}\left[\exp\left\{\hat{\lambda}y_\tau + \hat{\mu}\int_0^\tau y_u\, du\right\}\Big|\, y_0 = y\right], &(2.29)\end{aligned}$$

where $\tau = T - t$ and $F(0, y) = \exp\{\hat{\lambda}y\}$. The process $M = (M_t)_{t\in[0,T]}$ defined by

$$M_t \equiv F(\tau, y_t)\exp\left\{\hat{\mu}\int_0^t y_u\, du\right\} \tag{2.30}$$

is a martingale. Applying Itô's formula to the right-hand side and setting the resulting drift term equal to zero, we find a PDE for $F(\tau, y)$:

$$-\partial_\tau F + k(\theta - y)\partial_y F + \frac{1}{2}\xi^2 y \partial_{yy} F + \hat{\mu} y F = 0. \tag{2.31}$$

Suppose that the solution to the PDE is of the form

$$F(\tau, y) = \exp\left\{\hat{\lambda} y + k\theta\hat{\lambda}\tau + G(\tau)y + H(\tau)\right\}, \tag{2.32}$$

with

$$G(0) = 0 \text{ and } H(0) = 0. \tag{2.33}$$

Substituting back into (2.31) with (2.32) results in a system of Riccati ODEs:

$$\partial_\tau G(\tau) = aG(\tau)^2 + bG(\tau) + c, \tag{2.34}$$

$$\partial_\tau H(\tau) = k\theta G(\tau), \tag{2.35}$$

where

$$a = \frac{1}{2}\xi^2, \;\; b = \hat{\lambda}\xi^2 - k = \mu\xi - k \;\text{ and }\; c = \frac{1}{2}\hat{\lambda}^2\xi^2 + \hat{\mu} - \hat{\lambda}k = \varDelta. \tag{2.36}$$

The conditions under which $G(\tau)$ – and so $F(\tau, y)$ – blows up are connected to the position of the roots of the polynomial $f(x) = ax^2 + bx + c$, i.e.,

$$x_{1,2} = -\frac{1}{\xi^2}(\mu\xi - k) \pm \frac{1}{\xi^2}\sqrt{(\mu\xi - k)^2 - 2\xi^2\varDelta}. \tag{2.37}$$

First, if $\varDelta = 0$, then zero is a root of the polynomial. Since $f(x)$ is locally Lipschitz, we conclude that the ODE (2.34) with the initial condition $G(0) = 0$ has a unique solution, namely $G(\tau) = 0$ for all $\tau \geq 0$. From (2.35), we deduce that $H(\tau) = 0$ for all $\tau \geq 0$. Substituting back into (2.32) and making use of (2.29), we deduce that $\mathbb{E}\left[\Theta_t\right] = 1$ for all $t \geq 0$. Alternatively, note that $(\Theta_t)_{t\in[0,T]}$ is a Doléans exponential and a true martingale [15]. Therefore, when $\varDelta = 0$, $\sup_{t\in[0,T]}\mathbb{E}\left[\Theta_t\right] = 1$ for all $T \geq 0$.

Second, if $\varDelta \neq 0$, employing Proposition 3.3 in [67], one can easily show the finiteness of $G(\tau)$, and hence of $H(\tau)$ and $F(\tau, y)$, for all $\tau < T^*$, and the explosion of $G(\tau)$ for all $\tau \geq T^*$. Therefore, $F(\tau, y)$ is continuous and finite on $[0, T]$ for all $T < T^*$ and hence its supremum over the time interval is finite by the boundedness theorem, which concludes the proof. ■

2.3.2. The drift-implicit Euler (BEM) scheme

Suppose that $2k\theta \geq \xi^2$ and let $\bar{y}_t = \tilde{x}^2_{t_n}$ for all $t \in [t_n, t_{n+1})$, with $\tilde{x}$ defined in (2.11). In order to establish the exponential integrability of the BEM scheme, we first prove an auxiliary result.

Lemma 2.3. *Suppose that $T \geq 0$ and $\Delta > 0$. Then*

$$\exists \eta > 1: \quad \forall \omega \in [0,1], \quad 2\eta^2\omega^2\Delta\gamma^2T^2 + 2\eta\omega\gamma\mu T - \eta + 1 < 0, \tag{2.38}$$

if and only if $T < T^*$, where T^* is as given below.

(1) When $\mu < 0$ and $\lambda < \frac{3}{2}\mu^2$,

$$T^* = -\frac{2\mu}{\xi\Delta}. \tag{2.39}$$

(2) When $\mu < 0$ and $\lambda \geq \frac{3}{2}\mu^2$, or when $\mu \geq 0$,

$$T^* = \frac{1}{\xi(\mu + \sqrt{2\Delta})}. \tag{2.40}$$

Proof. Fix any $\eta > 1$ and let $f_\eta : [0,1] \to \mathbb{R}$ be defined by

$$f_\eta(\omega) = 2\omega^2\eta^2\Delta\gamma^2T^2 + 2\omega\eta\gamma\mu T - (\eta - 1), \tag{2.41}$$

with two distinct real roots

$$\omega_{1,2} = \frac{-\mu \pm \sqrt{\mu^2 + 2(\eta-1)\Delta}}{2\eta\Delta\gamma T}. \tag{2.42}$$

Since $\omega_1 > 0 > \omega_2$, we deduce that $f_\eta([0,1]) \subseteq (-\infty, 0)$ if and only if $f_\eta(1) < 0$, i.e.,

$$2\eta^2\Delta\gamma^2T^2 - \eta(1 - 2\gamma\mu T) + 1 < 0. \tag{2.43}$$

However, (2.43) holds for some $\eta > 1$ if and only if the second-order polynomial in η on the left-hand side has two distinct real roots, one of which is greater than one. Substituting back into (2.43) with γ from (2.9), we find the necessary and sufficient conditions:

$$(1 - \mu\xi T)^2 > 2\Delta\xi^2T^2 \;\text{ and }\; 1 - \mu\xi T + \sqrt{(1-\mu\xi T)^2 - 2\Delta\xi^2T^2} > \Delta\xi^2T^2. \tag{2.44}$$

Note that if

$$1 - \mu\xi T > \Delta\xi^2T^2, \tag{2.45}$$

then the conditions specified in (2.44) are equivalent to

$$\xi T(\mu + \sqrt{2\Delta}) < 1, \tag{2.46}$$

whereas if

$$1 - \mu\xi T \leq \Delta\xi^2T^2, \tag{2.47}$$

then the conditions specified in (2.44) are equivalent to

$$\xi T(\mu + \sqrt{2\Delta}) < 1 \;\text{ and }\; 2\Delta\xi T < -4\mu. \tag{2.48}$$

Since

$$1 - \mu\xi T \leq \Delta\xi^2T^2 \;\Leftrightarrow\; \sqrt{\mu^2 + 4\Delta} - \mu \leq 2\Delta\xi T, \tag{2.49}$$

we find an equivalent set of conditions:

$$\xi T(\mu + \sqrt{2\Delta}) < 1, \tag{2.50}$$

and

$$2\Delta\xi T < \sqrt{\mu^2 + 4\Delta} - \mu \quad \text{or} \quad \sqrt{\mu^2 + 4\Delta} - \mu \leq 2\Delta\xi T < -4\mu. \tag{2.51}$$

Next, upon noticing that

$$\mu + \sqrt{2\Delta} \leq 0 \;\;\Leftrightarrow\;\; \mu < 0 \;\;\text{and}\;\; \lambda \leq 0, \tag{2.52}$$

and that

$$\sqrt{\mu^2 + 4\Delta} - \mu < -4\mu \;\;\Leftrightarrow\;\; \mu < 0 \;\;\text{and}\;\; \lambda < \frac{3}{2}\mu^2, \tag{2.53}$$

the conclusion follows relatively easily. ∎

Proposition 2.4. *Let $T \geq 0$. If $\Delta \leq 0$ or otherwise, if $\Delta > 0$ and $T < T^*$, with T^* from* (2.39) – (2.40)*, then there exists $\delta_T > 0$ such that, for all $\delta t \in (0, \delta_T)$, the first moment of the exponential functional from* (2.13) *of the BEM scheme is uniformly bounded, i.e.,*

$$\sup_{\delta t \in (0,\delta_T)} \sup_{t \in [0,T]} \mathbb{E}\left[\bar{\Theta}_t\right] < \infty. \tag{2.54}$$

Proof. If $\Delta \leq 0$, the conclusion follows from Lemma 2.1. If $\Delta > 0$ and $T < T^*$, we know from Lemma 2.3 that $\exists \eta > 1$ independent of δt such that (2.38) holds for all $\omega \in [0, 1]$. Fix any such η. We prove by induction on $0 \leq m \leq N$ that, for sufficiently small values of δt, we have

$$\begin{aligned} \mathbb{E}\left[\bar{\Theta}_T\right] \leq \mathbb{E}\left[\exp\left\{\mu \int_0^{t_{N-m}} \sqrt{\bar{y}_u}\, dW_u + \lambda\delta t \sum_{i=0}^{N-m-1} \bar{y}_{t_i} + \eta m \Delta\delta t\, \bar{y}_{t_{N-m}}\right\}\right] \\ \times \left(1 + 2\eta\Delta\gamma^2 T\delta t\right)^{2m} \exp\left\{\eta\alpha\Delta(\delta t)^2(m-1)m\right\}. \end{aligned} \tag{2.55}$$

Note that when $m = 0$, we have equality. Let us assume that (2.55) holds for $0 \leq m < N$ and prove the inductive step. Conditioning on $\mathcal{G}_{t_{N-m-1}}$, we obtain

$$\begin{aligned} \mathbb{E}\left[\bar{\Theta}_T\right] &\leq \left(1 + 2\eta\Delta\gamma^2 T\delta t\right)^{2m} \exp\left\{\eta\alpha\Delta(\delta t)^2(m-1)m\right\} \\ &\times \mathbb{E}\left[\exp\left\{\mu \int_0^{t_{N-m-1}} \sqrt{\bar{y}_u}\, dW_u + \lambda\delta t \sum_{i=0}^{N-m-1} \bar{y}_{t_i}\right\}\right. \\ &\left.\times\, \mathbb{E}_{t_{N-m-1}}\left[\exp\left\{\eta m\Delta\delta t\, \bar{y}_{t_{N-m}} + \mu\sqrt{\bar{y}_{t_{N-m-1}}}\, \delta W_{t_{N-m-1}}\right\}\right]\right]. \end{aligned} \tag{2.56}$$

For brevity, define $x = \tilde{x}_{t_{N-m-1}}$ and note that if $Z \sim \mathcal{N}(0, 1)$ and independent of $\mathcal{G}_{t_{N-m-1}}$, then $\delta W_{t_{N-m-1}} \overset{\text{law}}{=} \sqrt{\delta t}\, Z$. Let $\mathcal{I}$ be the conditional (inner) expectation in (2.56), then

$$\mathcal{I} = \mathbb{E}_{0,x}\Big[\exp\left\{\eta m\Delta\delta t\, \psi(Z)^2 + \mu x\sqrt{\delta t}\, Z\right\}\Big], \tag{2.57}$$

where

$$\psi(z) = \frac{x+\gamma\sqrt{\delta t}\,z}{2(1-\beta\delta t)} + \sqrt{\frac{(x+\gamma\sqrt{\delta t}\,z)^2}{4(1-\beta\delta t)^2} + \frac{\alpha\delta t}{1-\beta\delta t}}. \tag{2.58}$$

On $\left\{x+\gamma\sqrt{\delta t}Z \le 0\right\}$, since $\beta < 0$, we have that

$$\psi(Z)^2 < \alpha\delta t. \tag{2.59}$$

On $\left\{x+\gamma\sqrt{\delta t}Z > 0\right\}$, one can easily show that

$$\psi(Z)^2 < \left(x+\gamma\sqrt{\delta t}Z\right)^2 + 2\alpha\delta t. \tag{2.60}$$

Hence, if we let $z_0 = -\frac{x}{\gamma\sqrt{\delta t}}$,

$$\begin{aligned}\mathcal{I} \le &\int_{-\infty}^{z_0} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z^2 + \mu x\sqrt{\delta t}\,z + \eta m\alpha\Delta(\delta t)^2\right\} dz \\ &+ \int_{z_0}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z^2 + \mu x\sqrt{\delta t}\,z + \eta m\Delta\delta t\left[\left(x+\gamma\sqrt{\delta t}\,z\right)^2 + 2\alpha\delta t\right]\right\} dz. \end{aligned}\tag{2.61}$$

Suppose that $\delta_T \le \left(2\eta\Delta\gamma^2 T\right)^{-1}$ and define

$$a = \sqrt{1-2\eta m\Delta\gamma^2(\delta t)^2},\; b = \mu x\sqrt{\delta t} + 2\eta m\Delta\gamma x(\delta t)^{\frac{3}{2}},\; c = \sqrt{\eta m\Delta\delta t(x^2+2\alpha\delta t)}. \tag{2.62}$$

Some straightforward calculations lead to the following upper bound:

$$\mathcal{I} \le \exp\left\{\frac{1}{2}\mu^2x^2\delta t + \eta m\alpha\Delta(\delta t)^2\right\}\Phi\left(z_1\right) + \frac{1}{a}\exp\left\{c^2+\frac{b^2}{2a^2}\right\}\Big\{1-\Phi\left(z_2\right)\Big\}, \tag{2.63}$$

where Φ is the standard normal CDF and

$$z_1 = -\frac{x}{\gamma\sqrt{\delta t}} - \mu x\sqrt{\delta t},\;\; z_2 = -\frac{ax}{\gamma\sqrt{\delta t}} - \frac{b}{a}. \tag{2.64}$$

Suppose that $\delta_T \le \left(\gamma\max\left\{0^+, -\mu\right\}\right)^{-1}$. Then $1+\mu\gamma\delta t > 0$ and so $z_1 < 0$. From (2.62),

$$\begin{aligned} a\gamma\sqrt{\delta t}(z_1 - z_2) &= (1-a)x(\mu\gamma\delta t - a) + \gamma\sqrt{\delta t}(b - \mu x\sqrt{\delta t}) \\ &= (1-a)x(1+\mu\gamma\delta t). \end{aligned}\tag{2.65}$$

Therefore, since $x > 0$ and $a \in (0,1]$, $z_2 \le z_1 < 0$. From (2.62), we also see that

$$\begin{aligned} c^2 + \frac{b^2}{2a^2} - \frac{1}{2}\mu^2x^2\delta t - \eta m\alpha\Delta(\delta t)^2 &= \eta m\alpha\Delta(\delta t)^2 + \frac{1}{2a^2}\left(2a^2\eta m\Delta x^2\delta t + b^2 - a^2\mu^2x^2\delta t\right) \\ &= \eta m\alpha\Delta(\delta t)^2 + \frac{1}{a^2}\eta m\Delta x^2\delta t(1+\mu\gamma\delta t)^2, \end{aligned}\tag{2.66}$$

which is non-negative, and hence

$$\mathcal{I} \le \frac{1}{a}\exp\left\{c^2+\frac{b^2}{2a^2}\right\}\Big\{1+\Phi\left(z_1\right)-\Phi\left(z_2\right)\Big\}. \tag{2.67}$$

Applying the mean value (MV) theorem to $\Phi \in C^1$, we can find $z \in [z_2, z_1]$ such that

$$\Phi(z_1) - \Phi(z_2) = (z_1 - z_2)\,\phi(z) \le (z_1 - z_2)\,\phi\left(z_1\right). \tag{2.68}$$

Hence, using (2.65),

$$\Phi(z_1) - \Phi(z_2) \leq \frac{1-a}{a\sqrt{2\pi}} \frac{x(1+\mu\gamma\delta t)}{\gamma\sqrt{\delta t}} \exp\left\{-\frac{x^2(1+\mu\gamma\delta t)^2}{2\gamma^2\delta t}\right\}. \tag{2.69}$$

Consider the function $g : (0, \infty) \to \mathbb{R}$ defined by $g(u) = ue^{-\frac{u^2}{2}}$. Then its global maximum is $e^{-\frac{1}{2}}$, and it is achieved when $u = 1$. We can thus bound the term on the right-hand side of (2.69) from above to get

$$\Phi(z_1) - \Phi(z_2) \leq \frac{1-a}{a\sqrt{2\pi e}} \leq \frac{1}{a} - 1. \tag{2.70}$$

Plugging back into (2.67),

$$\mathcal{I} \leq \frac{1}{a^2} \exp\left\{c^2 + \frac{b^2}{2a^2}\right\}. \tag{2.71}$$

Suppose that $\delta_T \leq \frac{\sqrt{5}-1}{4}\left(\eta\Delta\gamma^2 T\right)^{-1}$, then $a(1+a) > 1$ and hence

$$\frac{1}{a} = 1 + \frac{(1-a)(1+a)}{a(1+a)} \leq 2 - a^2. \tag{2.72}$$

From (2.62), we get

$$\frac{1}{a^2} \leq \left(1 + 2\eta m\Delta\gamma^2(\delta t)^2\right)^2 < \left(1 + 2\eta\Delta\gamma^2 T\delta t\right)^2. \tag{2.73}$$

Combining (2.66), (2.71) and (2.73), we deduce that

$$\mathcal{I} < \left(1 + 2\eta\Delta\gamma^2 T\delta t\right)^2 \exp\left\{2\eta m\alpha\Delta(\delta t)^2 + \frac{1}{2}\mu^2 x^2\delta t + \frac{1}{a^2}\eta m\Delta x^2\delta t(1+\mu\gamma\delta t)^2\right\}. \tag{2.74}$$

Because of our choice of η, the second-order polynomial $f_\eta(\omega)$ defined in (2.41) is negative for all $\omega \in [0, 1]$. In particular, f_η attains its maximum on a closed, bounded interval. Hence, there exists $\omega_0 \in [0, 1]$ independent of δt and m such that $f_\eta(\omega) \leq f_\eta(\omega_0) < 0$ for all $\omega \in [0, 1]$. Suppose that $\delta_T \leq -f_\eta(\omega_0)\left(2\eta(\eta\Delta - \lambda)\gamma^2 T\right)^{-1}$, then

$$\left(1 - a^2\right)\left(\eta - 1 + \frac{\mu^2}{2\Delta}\right) < -f_\eta(\omega_0) \leq -f_\eta(\omega), \quad \forall \omega \in [0, 1]. \tag{2.75}$$

Applying the inequality with $\omega = \frac{m}{N}$ and using (2.41), we get

$$\left(1 - a^2\right)\left(\eta - 1 + \frac{\mu^2}{2\Delta}\right) < \eta - 1 - 2\eta^2 m^2\Delta\gamma^2(\delta t)^2 - 2\eta m\mu\gamma\delta t. \tag{2.76}$$

However, using (2.62), we can rewrite it as

$$\frac{1}{a^2}\eta m(1+\mu\gamma\delta t)^2 < \eta(m+1) - 1, \tag{2.77}$$

so

$$\mathcal{I} < \left(1 + 2\eta\Delta\gamma^2 T\delta t\right)^2 \exp\left\{2\eta m\alpha\Delta(\delta t)^2 - \lambda x^2\delta t + \eta(m+1)\Delta x^2\delta t\right\}. \tag{2.78}$$

Substituting back into (2.56) with (2.78) gives the inductive step. Finally, taking $m = N$ in (2.55) leads to

$$\mathbb{E}\left[\bar{\Theta}_T\right] \leq \left(1 + \frac{2\eta\Delta\gamma^2T^2}{N}\right)^{2N} \exp\left\{\eta\alpha\Delta T^2\left(1 - \frac{1}{N}\right) + \eta\Delta T y_0\right\}$$
$$< \exp\left\{\eta\Delta T^2\left(\alpha + 4\gamma^2\right) + \eta\Delta T y_0\right\}. \tag{2.79}$$

The right-hand side is finite and independent of δt, whence the conclusion. Note that if we follow the same line of proof with $\eta = 1$, we find suboptimal sufficient conditions, whereas with $\eta < 1$, we fail to achieve the inductive step. Also note that δ_T can be computed explicitly from the proof. ■

2.3.3. Explicit Euler schemes with absorption fixes

First, consider the FTE scheme and let $\bar{y}_t = \tilde{y}^+_{t_n}$ for all $t \in [t_n, t_{n+1})$, with $\tilde{y}$ from (2.4).

Lemma 2.5. *Suppose that $T \geq 0$ and $\Delta > 0$. Then*

$$\exists \eta \geq 1: \quad \forall \omega \in [0,1], \quad \eta^2\omega^2\xi^2\Delta T^2 - 2\eta\omega(k - \mu\xi)T - 2\eta + 2 \leq 0, \tag{2.80}$$

if and only if $T \leq T^*$, where T^* is as given below.

(1) When $k \leq \xi(\mu + \frac{1}{2}\sqrt{2\Delta})$,

$$T^* = \frac{1}{\xi(\mu + \sqrt{2\Delta}) - k}. \tag{2.81}$$

(2) When $k > \xi(\mu + \frac{1}{2}\sqrt{2\Delta})$,

$$T^* = \frac{2(k - \mu\xi)}{\xi^2\Delta}. \tag{2.82}$$

Proof. Fix any $\eta \geq 1$ and let $f_\eta : [0,1] \to \mathbb{R}$ be defined by

$$f_\eta(\omega) = \omega^2\eta^2\xi^2\Delta T^2 - 2\omega\eta(k - \mu\xi)T - 2(\eta - 1), \tag{2.83}$$

with real roots

$$\omega_{1,2} = \frac{k - \mu\xi \pm \sqrt{(k - \mu\xi)^2 + 2(\eta - 1)\xi^2\Delta}}{\eta\xi^2\Delta T}. \tag{2.84}$$

Since $\omega_1 \geq 0 \geq \omega_2$, we know that $f_\eta([0,1]) \subseteq (-\infty, 0]$ if and only if $f_\eta(1) \leq 0$, i.e.,

$$\eta^2\xi^2\Delta T^2 - 2\eta\left[1 + (k - \mu\xi)T\right] + 2 \leq 0. \tag{2.85}$$

However, (2.85) holds for some $\eta \geq 1$ if and only if the second-order polynomial in η on the left-hand side has a real root greater or equal to one. Therefore, we find the necessary and sufficient conditions:

$$\left[1 + (k - \mu\xi)T\right]^2 \geq 2\Delta\xi^2T^2 \tag{2.86}$$

and

$$1 + (k - \mu\xi)T + \sqrt{\left[1 + (k - \mu\xi)T\right]^2 - 2\Delta\xi^2T^2} \geq \Delta\xi^2T^2. \tag{2.87}$$

Note that if

$$1 + (k - \mu\xi)T \geq \Delta\xi^2 T^2, \tag{2.88}$$

then the conditions specified in (2.86) – (2.87) are equivalent to

$$\big[\xi(\mu + \sqrt{2\Delta}) - k\big]T \leq 1, \tag{2.89}$$

whereas if

$$1 + (k - \mu\xi)T < \Delta\xi^2 T^2, \tag{2.90}$$

then the conditions specified in (2.86) – (2.87) are equivalent to

$$\big[\xi(\mu + \sqrt{2\Delta}) - k\big]T \leq 1 \quad \text{and} \quad 2\Delta\xi^2 T \leq 4(k - \mu\xi). \tag{2.91}$$

Since

$$1 + (k - \mu\xi)T < \Delta\xi^2 T^2 \quad \Leftrightarrow \quad k - \mu\xi + \sqrt{(k - \mu\xi)^2 + 4\Delta\xi^2} < 2\Delta\xi^2 T, \tag{2.92}$$

we find an equivalent set of conditions:

$$\big[\xi(\mu + \sqrt{2\Delta}) - k\big]T \leq 1, \tag{2.93}$$

and

$$2\Delta\xi^2 T \leq k - \mu\xi + \sqrt{(k - \mu\xi)^2 + 4\Delta\xi^2} \tag{2.94}$$

or

$$k - \mu\xi + \sqrt{(k - \mu\xi)^2 + 4\Delta\xi^2} < 2\Delta\xi^2 T \leq 4(k - \mu\xi). \tag{2.95}$$

Next, upon noticing that

$$k - \mu\xi + \sqrt{(k - \mu\xi)^2 + 4\Delta\xi^2} < 4(k - \mu\xi) \quad \Leftrightarrow \quad k > \xi\Big(\mu + \frac{1}{2}\sqrt{2\Delta}\Big), \tag{2.96}$$

the conclusion follows relatively easily. ■

Proposition 2.6. *Let $T \geq 0$. If $\max\{0, \Delta\}(T^* - T) \geq 0$, with T^* from (2.81) – (2.82), then there exists $\delta_T > 0$ such that, for all $\delta t \in (0, \delta_T)$, the first moment of the exponential functional from (2.13) of the FTE scheme is uniformly bounded, i.e.,*

$$\sup_{\delta t \in (0, \delta_T)} \sup_{t \in [0,T]} \mathbb{E}\left[\bar{\Theta}_t\right] < \infty. \tag{2.97}$$

Proof. If $\Delta \leq 0$, this is a consequence of Lemma 2.1. If $\Delta > 0$ and $T \leq T^*$, we know from Lemma 2.5 that $\exists \eta \geq 1$ independent of δt such that (2.80) holds for all $\omega \in [0, 1]$. Fix any such η. We prove by induction on $0 \leq m \leq N$ that, for sufficiently small values of δt,

$$\begin{aligned} \mathbb{E}\left[\bar{\Theta}_T\right] \leq \mathbb{E}\left[\exp\left\{\mu \int_0^{t_{N-m}} \sqrt{\bar{y}_u}\, dW_u + \lambda\delta t \sum_{i=0}^{N-m-1} \bar{y}_{t_i} + \eta m \Delta \delta t \bar{y}_{t_{N-m}}\right\}\right] \\ \times \exp\left\{\frac{1}{2}\eta\big(k\theta + \vartheta\xi\big)\Delta(\delta t)^2(m-1)m\right\}, \end{aligned} \tag{2.98}$$

where

$$\vartheta = \sqrt{\frac{1}{2\pi}\xi^2 + \frac{1}{2\pi}\sqrt{\xi^4 + 2k^2\theta^2}}. \tag{2.99}$$

Note that when $m = 0$, we have equality. Let us assume that (2.98) holds for $0 \le m < N$ and prove the inductive step. Conditioning on $\mathcal{G}_{t_{N-m-1}}$, we obtain

$$\begin{aligned}\mathbb{E}\left[\bar{\Theta}_T\right] \le \exp&\left\{\frac{1}{2}\eta\big(k\theta + \vartheta\xi\big)\varDelta(\delta t)^2(m-1)m\right\}\\ &\times \mathbb{E}\left[\exp\left\{\mu\int_0^{t_{N-m-1}}\sqrt{\bar{y}_u}\,dW_u + \lambda\delta t\sum_{i=0}^{N-m-1}\bar{y}_{t_i}\right\}\right.\\ &\left.\times\, \mathbb{E}_{t_{N-m-1}}\left[\exp\left\{\eta m\varDelta\delta t\,\bar{y}_{t_{N-m}} + \mu\sqrt{\bar{y}_{t_{N-m-1}}}\,\delta W_{t_{N-m-1}}\right\}\right]\right].\end{aligned} \tag{2.100}$$

Define $\tilde{x} = \tilde{y}_{t_{N-m-1}}$ and $x = \bar{y}_{t_{N-m-1}}$. If $Z \sim \mathcal{N}(0,1)$ and independent of $\mathcal{G}_{t_{N-m-1}}$, then $\delta W_{t_{N-m-1}} \overset{\text{law}}{=} \sqrt{\delta t}\,Z$. Let $\mathcal{I}$ be the conditional expectation in (2.100), then

$$\mathcal{I} \le \mathbb{E}_{0,x}\Big[\exp\Big\{\eta m\varDelta\delta t\max\Big[0,\, x + k(\theta - x)\delta t + \xi\sqrt{x\delta t}\,Z\Big] + \mu\sqrt{x\delta t}\,Z\Big\}\Big]. \tag{2.101}$$

If $x = 0$, then

$$\mathcal{I} \le \exp\Big\{\eta m k\theta\varDelta(\delta t)^2\Big\}. \tag{2.102}$$

If $x > 0$ and

$$z_0 = -\frac{k\theta\delta t + (1 - k\delta t)x}{\xi\sqrt{x\delta t}}, \tag{2.103}$$

then

$$\begin{aligned}\mathcal{I} \le &\int_{z_0}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2 + \mu\sqrt{x\delta t}\,z + \eta m\varDelta\delta t\Big[x + k(\theta - x)\delta t + \xi\sqrt{x\delta t}\,z\Big]\right\}dz\\ &+ \int_{-\infty}^{z_0}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2 + \mu\sqrt{x\delta t}\,z\right\}dz.\end{aligned} \tag{2.104}$$

Suppose that $\delta_T \le \max\left\{0^+, k - \mu\xi\right\}^{-1}$ and define

$$a = \eta m\big[1 - (k - \mu\xi)\delta t\big] + \frac{1}{2}\eta^2 m^2\xi^2\varDelta(\delta t)^2 \ge 0. \tag{2.105}$$

Some straightforward calculations lead to the following upper bound:

$$\mathcal{I} \le \exp\left\{\eta m k\theta\varDelta(\delta t)^2 + \frac{1}{2}\mu^2 x\delta t + a\varDelta x\delta t\right\}\Big\{1 + \Phi(z_1) - \Phi(z_2)\Big\}, \tag{2.106}$$

where

$$z_1 = z_0 - \mu\sqrt{x\delta t} = -\frac{k\theta\delta t + \big[1 - (k - \mu\xi)\delta t\big]x}{\xi\sqrt{x\delta t}} \tag{2.107}$$

and

$$z_2 = z_0 - \mu\sqrt{x\delta t} - \eta m\xi\varDelta(\delta t)^{3/2}\sqrt{x}. \tag{2.108}$$

Clearly $z_2 \le z_1 < 0$, so we can find $z \in [z_2, z_1]$ such that

$$\Phi(z_1) - \Phi(z_2) = (z_1 - z_2)\,\phi(z) \le (z_1 - z_2)\,\phi(z_1). \tag{2.109}$$

Hence, using (2.107) – (2.109),

$$\Phi(z_1) - \Phi(z_2) \leq \frac{1}{\sqrt{2\pi}} \eta m \xi \mathit{\Delta} (\delta t)^{3/2} g(x), \tag{2.110}$$

where $g : (0, \infty) \to \mathbb{R}$ is defined by

$$g(x) = \sqrt{x} \exp \left\{ -\frac{\big[k\theta\delta t + [1 - (k - \mu\xi)\delta t]\, x\big]^2}{2\xi^2 x \delta t} \right\}. \tag{2.111}$$

Suppose that $\delta_T \leq \frac{\sqrt{2}-1}{\sqrt{2}} \max\left\{0^+, k - \mu\xi\right\}^{-1}$. One can easily find the global maximum of the function by computing the first derivative:

$$g'(x) = e^{-z_1^2/2} \left\{ \frac{1}{2\sqrt{x}} + \sqrt{x} \left[\frac{k^2\theta^2\delta t}{2\xi^2 x^2} - \frac{[1 - (k - \mu\xi)\delta t]^2}{2\xi^2 \delta t} \right] \right\}. \tag{2.112}$$

Therefore,

$$g'(x) = 0 \;\; \Leftrightarrow \;\; -[1 - (k - \mu\xi)\delta t]^2\, x^2 + \xi^2 x \delta t + k^2\theta^2\, (\delta t)^2 = 0. \tag{2.113}$$

In order to solve this quadratic, we divide throughout by $(\delta t)^2$ and introduce $\hat{x} = x/\delta t$. Then there exists a unique, positive solution $\hat{x}_0$ for which the derivative is zero, namely

$$\hat{x}_0 = \frac{\xi^2}{2\,[1 - (k - \mu\xi)\delta t]^2} + \sqrt{\frac{\xi^4}{4\,[1 - (k - \mu\xi)\delta t]^4} + \frac{k^2\theta^2}{[1 - (k - \mu\xi)\delta t]^2}}\,. \tag{2.114}$$

From (2.99), we have that

$$\hat{x}_0 < 2\pi\vartheta^2 \;\; \Rightarrow \;\; x_0 < 2\pi\vartheta^2\delta t. \tag{2.115}$$

Since the second root is negative, $g(x)$ must be increasing up to x_0 and decreasing after this point, so the function attains its global maximum at x_0. Hence,

$$g(x) < \vartheta\sqrt{2\pi\delta t}. \tag{2.116}$$

Substituting back into (2.106) with (2.110) and (2.116), we get

$$\mathcal{I} \leq \exp\left\{ \eta m \big(k\theta + \vartheta\xi\big) \mathit{\Delta} (\delta t)^2 + \frac{1}{2}\mu^2 x \delta t + a \mathit{\Delta} x \delta t \right\}. \tag{2.117}$$

Note from (2.102) that this holds when $x = 0$ as well. Applying (2.80) with $\omega = \frac{m}{N}$ leads to

$$\eta^2 m^2 \xi^2 \mathit{\Delta} (\delta t)^2 - 2\eta m (k - \mu\xi)\delta t - 2\eta + 2 \leq 0. \tag{2.118}$$

Hence, from (2.105),

$$a \leq \eta(m + 1) - 1. \tag{2.119}$$

Therefore,

$$\mathcal{I} \leq \exp\left\{ \eta m \big(k\theta + \vartheta\xi\big) \mathit{\Delta} (\delta t)^2 - \lambda x \delta t + \eta(m + 1) \mathit{\Delta} x \delta t \right\}. \tag{2.120}$$

Substituting back into (2.100) with (2.120) gives the inductive step. Finally, taking $m = N$ in (2.98) leads to

$$\mathbb{E}\left[\bar{\Theta}_T\right] < \exp\left\{\frac{1}{2}\eta\Delta T^2\big(k\theta + \vartheta\xi\big) + \eta\Delta T y_0\right\}. \tag{2.121}$$

The right-hand side is finite and independent of δt, whence the conclusion. Note that if we follow the same line of proof with $\eta < 1$, we fail to achieve the inductive step. Also note that δ_T can be computed explicitly from the proof. ■

Second, we consider the partial truncation and the absorption schemes and let $\bar{y}_t = \tilde{y}_{t_n}^+$ for all $t \in [t_n, t_{n+1})$, with $\tilde{y}$ defined in (2.3) and (2.5), respectively.

Proposition 2.7. *Let $T \geq 0$. If $\max\{0, \Delta\}(T^* - T) \geq 0$, with T^* from* (2.81) – (2.82), *then there exists $\delta_T > 0$ such that, for all $\delta t \in (0, \delta_T)$, the first moments of the exponential functionals from* (2.13) *of the partial truncation and absorption schemes are uniformly bounded, i.e.,*

$$\sup_{\delta t \in (0,\delta_T)} \sup_{t \in [0,T]} \mathbb{E}\left[\bar{\Theta}_t\right] < \infty. \tag{2.122}$$

Proof. We follow the argument of Proposition 2.6 closely and note that (2.101) holds for the partial truncation scheme if $\delta_T \leq k^{-1}$, whereas for the absorption scheme we have equality. ■

2.3.4. Explicit Euler schemes with reflection fixes

First, consider the reflection scheme and let $\bar{y}_t = |\tilde{y}_{t_n}|$ for all $t \in [t_n, t_{n+1})$, with $\tilde{y}$ from (2.6).

Proposition 2.8. *Let $T \geq 0$. If $\max\{0, \Delta\}(T^* - T) \geq 0$, then there exists $\delta_T > 0$ such that, for all $\delta t \in (0, \delta_T)$, the first moment of the exponential functional from* (2.13) *of the reflection scheme is uniformly bounded, i.e.,*

$$\sup_{\delta t \in (0,\delta_T)} \sup_{t \in [0,T]} \mathbb{E}\left[\bar{\Theta}_t\right] < \infty, \tag{2.123}$$

where T^* is as given below.

(1) When $k \leq \xi(|\mu| + \frac{1}{2}\sqrt{2\Delta})$,

$$T^* = \frac{1}{\xi(|\mu| + \sqrt{2\Delta}) - k}. \tag{2.124}$$

(2) When $k > \xi(|\mu| + \frac{1}{2}\sqrt{2\Delta})$,

$$T^* = \frac{2(k - |\mu|\xi)}{\xi^2\Delta}. \tag{2.125}$$

Proof. If $\varDelta \leq 0$, this is a consequence of Lemma 2.1. If $\varDelta > 0$ and $T \leq T^*$, we know from Lemma 2.5 that $\exists \eta \geq 1$ independent of δt such that, for all $\omega \in [0,1]$,

$$\eta^2\omega^2\xi^2\varDelta T^2 - 2\eta\omega(k-|\mu|\xi)T - 2\eta + 2 \leq 0. \tag{2.126}$$

Fix any such η. Next, we prove by induction on $0 \leq m \leq N$ that, for sufficiently small values of δt, we have

$$\begin{aligned}\mathbb{E}\left[\bar{\Theta}_T\right] \leq \mathbb{E}&\left[\exp\left\{\mu\int_0^{t_{N-m}}\sqrt{\bar{y}_u}\,dW_u + \lambda\delta t\sum_{i=0}^{N-m-1}\bar{y}_{t_i} + \eta m\varDelta\delta t\,\bar{y}_{t_{N-m}}\right\}\right]\\ &\times\exp\left\{\frac{1}{2}\eta\big(k\theta+2\vartheta\xi\big)\varDelta(\delta t)^2(m-1)m\right\},\end{aligned} \tag{2.127}$$

with ϑ from (2.99). Note that when $m=0$, we have equality. Let us assume that (2.127) holds for $0 \leq m < N$ and prove the inductive step. Conditioning on $\mathcal{G}_{t_{N-m-1}}$, we obtain

$$\begin{aligned}\mathbb{E}\left[\bar{\Theta}_T\right] \leq &\exp\left\{\frac{1}{2}\eta\big(k\theta+2\vartheta\xi\big)\varDelta(\delta t)^2(m-1)m\right\}\\ &\times\mathbb{E}\left[\exp\left\{\mu\int_0^{t_{N-m-1}}\sqrt{\bar{y}_u}\,dW_u + \lambda\delta t\sum_{i=0}^{N-m-1}\bar{y}_{t_i}\right\}\right.\\ &\left.\times\,\mathbb{E}_{t_{N-m-1}}\left[\exp\left\{\eta m\varDelta\delta t\,\bar{y}_{t_{N-m}} + \mu\sqrt{\bar{y}_{t_{N-m-1}}}\,\delta W_{t_{N-m-1}}\right\}\right]\right].\end{aligned} \tag{2.128}$$

Define $\tilde{x} = \tilde{y}_{t_{N-m-1}}$ and $x = \bar{y}_{t_{N-m-1}}$. If $Z \sim \mathcal{N}(0,1)$ and independent of $\mathcal{G}_{t_{N-m-1}}$, then $\delta W_{t_{N-m-1}} \overset{\text{law}}{=} \sqrt{\delta t}\,Z$. Let $\mathcal{I}$ be the conditional expectation in (2.128), then

$$\mathcal{I} = \mathbb{E}_{0,\tilde{x}}\Big[\exp\Big\{\eta m\varDelta\delta t\big|\tilde{x} + k(\theta-\tilde{x})\delta t + \xi\sqrt{x\delta t}\,Z\big| + \mu\sqrt{x\delta t}\,Z\Big\}\Big]. \tag{2.129}$$

If $\tilde{x} = 0$, then

$$\mathcal{I} = \exp\Big\{\eta mk\theta\varDelta(\delta t)^2\Big\}. \tag{2.130}$$

If $\tilde{x} > 0$ and

$$z_0 = -\frac{k\theta\delta t + (1-k\delta t)x}{\xi\sqrt{x\delta t}}, \tag{2.131}$$

since $x = \tilde{x}$, we get

$$\begin{aligned}\mathcal{I} = &\int_{-\infty}^{z_0}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2 + \mu\sqrt{x\delta t}\,z - \eta m\varDelta\delta t\Big[x + k(\theta-x)\delta t + \xi\sqrt{x\delta t}\,z\Big]\right\}dz\\ &+\int_{z_0}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2 + \mu\sqrt{x\delta t}\,z + \eta m\varDelta\delta t\Big[x + k(\theta-x)\delta t + \xi\sqrt{x\delta t}\,z\Big]\right\}dz.\end{aligned} \tag{2.132}$$

Suppose that $\delta_T \leq \max\big\{0^+, k-\mu\xi\big\}^{-1}$. Some straightforward calculations lead to

$$\mathcal{I} \leq \exp\left\{\eta mk\theta\varDelta(\delta t)^2 + \frac{1}{2}\mu^2x\delta t + a\varDelta x\delta t\right\}\Big\{1 + \Phi(z_1) - \Phi(z_2)\Big\}, \tag{2.133}$$

with a defined in (2.105) and

$$z_{1,2} = z_0 - \big(\mu \mp \eta m\xi\varDelta\delta t\big)\sqrt{x\delta t} = -\frac{k\theta\delta t + \big[1 - \big(k - \mu\xi \pm \eta m\xi^2\varDelta\delta t\big)\delta t\big]x}{\xi\sqrt{x\delta t}}. \tag{2.134}$$

Suppose that $\delta_T \leq \max\left\{0^+, k-\mu\xi+\eta\xi^2\mathit{\Delta}T\right\}^{-1}$, then $z_2 \leq z_1 < 0$ and hence we can find $z \in [z_2, z_1]$ such that

$$\Phi(z_1) - \Phi(z_2) = (z_1 - z_2)\,\phi(z) \leq (z_1 - z_2)\,\phi\,(z_1) = \sqrt{\frac{2}{\pi}}\eta m\xi\mathit{\Delta}(\delta t)^{3/2}g(x), \tag{2.135}$$

where $g : (0, \infty) \to \mathbb{R}$ is defined by

$$g(x) = \sqrt{x}\exp\left\{-\frac{\left[k\theta\delta t + \left[1 - \left(k - \mu\xi + \eta\xi^2\mathit{\Delta}T\right)\delta t\right]x\right]^2}{2\xi^2 x\delta t}\right\}. \tag{2.136}$$

Suppose that $\delta_T \leq \frac{\sqrt{2}-1}{\sqrt{2}}\max\left\{0^+, k-\mu\xi+\eta\xi^2\mathit{\Delta}T\right\}^{-1}$. Proceeding as before, we can find an upper bound:

$$g(x) < \vartheta\sqrt{2\pi\delta t}. \tag{2.137}$$

Substituting back into (2.133) with (2.135) and (2.137), we get

$$\mathcal{I} \leq \exp\left\{\eta m\big(k\theta + 2\vartheta\xi\big)\mathit{\Delta}(\delta t)^2 + \frac{1}{2}\mu^2 x\delta t + a\mathit{\Delta}x\delta t\right\}. \tag{2.138}$$

If $\tilde{x} < 0$ and

$$z_0' = -\frac{k\theta\delta t - (1 - k\delta t)x}{\xi\sqrt{x\delta t}}, \tag{2.139}$$

since $x = -\tilde{x}$, we get

$$\begin{aligned}\mathcal{I} = &\int_{-\infty}^{z_0'} \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2 + \mu\sqrt{x\delta t}\,z - \eta m\mathit{\Delta}\delta t\Big[k(\theta + x)\delta t - x + \xi\sqrt{x\delta t}\,z\Big]\right\}dz \\ &+ \int_{z_0'}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2 + \mu\sqrt{x\delta t}\,z + \eta m\mathit{\Delta}\delta t\Big[k(\theta + x)\delta t - x + \xi\sqrt{x\delta t}\,z\Big]\right\}dz.\end{aligned} \tag{2.140}$$

Suppose that $\delta_T \leq \max\left\{0^+, k+\mu\xi\right\}^{-1}$ and define

$$b = \eta m\big[1 - (k + \mu\xi)\delta t\big] + \frac{1}{2}\eta^2 m^2\xi^2\mathit{\Delta}(\delta t)^2 \geq 0. \tag{2.141}$$

Some straightforward calculations lead to the following upper bound:

$$\mathcal{I} \leq \exp\left\{\eta mk\theta\mathit{\Delta}(\delta t)^2 + \frac{1}{2}\mu^2 x\delta t + b\mathit{\Delta}x\delta t\right\}\Big\{1 + \Phi(z_1') - \Phi(z_2')\Big\}, \tag{2.142}$$

where

$$z_{1,2}' = z_0' - \big(\mu \mp \eta m\xi\mathit{\Delta}\delta t\big)\sqrt{x\delta t} = -\frac{k\theta\delta t - \left[1 - \left(k + \mu\xi \mp \eta m\xi^2\mathit{\Delta}\delta t\right)\delta t\right]x}{\xi\sqrt{x\delta t}}. \tag{2.143}$$

Clearly $z_2' \leq z_1'$, and suppose that $\delta_T \leq \frac{\sqrt{2}-1}{\sqrt{2}}\max\left\{0^+, k+\mu\xi+\eta\xi^2\mathit{\Delta}T\right\}^{-1}$. Then we can find $z \in [z_2', z_1']$ such that

$$\Phi(z_1') - \Phi(z_2') = \big(z_1' - z_2'\big)\,\phi(z). \tag{2.144}$$

First, if $x \leq k\theta\delta t\big[1 - \big(k + \mu\xi + \eta\xi^2\mathit{\Delta}T\big)\delta t\big]^{-1}$, we deduce from (2.143) and (2.144) that

$$\Phi(z_1') - \Phi(z_2') \leq 2\eta m\vartheta\xi\mathit{\Delta}(\delta t)^2. \tag{2.145}$$

Second, if $x > k\theta\delta t\big[1 - \big(k + \mu\xi + \eta\xi^2 \mathit{\Delta} T\big)\delta t\big]^{-1}$, then $0 < z_2' \leq z_1'$ and hence

$$\Phi(z_1') - \Phi(z_2') \leq \sqrt{\frac{2}{\pi}}\eta m\xi \mathit{\Delta}(\delta t)^{3/2} h(x), \tag{2.146}$$

where $h : (0, \infty) \to \mathbb{R}$ is defined by

$$h(x) = \sqrt{x}\exp\left\{-\frac{\big[k\theta\delta t - \big[1 - \big(k + \mu\xi + \eta\xi^2 \mathit{\Delta} T\big)\delta t\big]x\big]^2}{2\xi^2 x\delta t}\right\}. \tag{2.147}$$

Proceeding as before, we find

$$h(x) < \vartheta\sqrt{2\pi\delta t}, \tag{2.148}$$

which again leads to the upper bound in (2.145). Substituting back into (2.142), we get

$$\mathcal{I} \leq \exp\left\{\eta m\big(k\theta + 2\vartheta\xi\big)\mathit{\Delta}(\delta t)^2 + \frac{1}{2}\mu^2 x\delta t + b\mathit{\Delta} x\delta t\right\}. \tag{2.149}$$

Combining (2.130), (2.138) and (2.149), we deduce that independent of the sign of $\tilde{x}$,

$$\mathcal{I} \leq \exp\left\{\eta m\big(k\theta + 2\vartheta\xi\big)\mathit{\Delta}(\delta t)^2 + \frac{1}{2}\mu^2 x\delta t + c\mathit{\Delta} x\delta t\right\}, \tag{2.150}$$

where

$$c = \eta m\big[1 - (k - |\mu|\xi)\delta t\big] + \frac{1}{2}\eta^2 m^2 \xi^2 \mathit{\Delta}(\delta t)^2. \tag{2.151}$$

Applying (2.126) with $\omega = \frac{m}{N}$ leads to

$$\eta^2 m^2 \xi^2 \mathit{\Delta}(\delta t)^2 - 2\eta m(k - |\mu|\xi)\delta t - 2\eta + 2 \leq 0, \tag{2.152}$$

and so

$$c \leq \eta(m+1) - 1. \tag{2.153}$$

Hence,

$$\mathcal{I} \leq \exp\left\{\eta m\big(k\theta + 2\vartheta\xi\big)\mathit{\Delta}(\delta t)^2 - \lambda x\delta t + \eta(m+1)\mathit{\Delta} x\delta t\right\}. \tag{2.154}$$

Substituting back into (2.128) with (2.154) gives the inductive step. Taking $m = N$ in (2.127), we deduce that

$$\mathbb{E}\left[\bar{\Theta}_T\right] < \exp\left\{\frac{1}{2}\eta\mathit{\Delta} T^2\big(k\theta + 2\vartheta\xi\big) + \eta\mathit{\Delta} T y_0\right\}. \tag{2.155}$$

The right-hand side is finite and independent of δt, which concludes the proof. Note that δ_T can be computed explicitly from the proof. ∎

Second, consider the symmetrized Euler scheme in (2.7) and let $\bar{y}_t = \tilde{y}_{t_n}$ for all $t \in [t_n, t_{n+1})$.

Proposition 2.9. *Let $T \geq 0$. If $\max\{0, \mathit{\Delta}\}(T^* - T) \geq 0$, with T^* from* (2.81) – (2.82), *then there exists $\delta_T > 0$ such that, for all $\delta t \in (0, \delta_T)$, the first moment of the exponential functional in* (2.13) *of the symmetrized Euler scheme is uniformly bounded, i.e.,*

$$\sup_{\delta t \in (0,\delta_T)} \sup_{t \in [0,T]} \mathbb{E}\left[\bar{\Theta}_t\right] < \infty. \tag{2.156}$$

Proof. We follow the argument of Proposition 2.8 closely and note that $\tilde{x} \geq 0$. ∎

2.4. Moment stability in the Heston model

In this section, we consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q})$ satisfying the usual conditions and suppose that the dynamics of the underlying process are governed by the Heston model under the measure $\mathbb{Q}$:

$$\begin{cases} dS_t = rS_t dt + \sqrt{v_t} S_t \, dW_t^s \\ dv_t = k(\theta - v_t)dt + \xi\sqrt{v_t}\, dW_t^v, \end{cases} \tag{2.157}$$

where W^s and W^v are correlated Brownian motions with constant correlation coefficient $\rho \in [-1, 1]$, r is an arbitrary non-negative number and k, θ, $\xi > 0$ as before. We decompose W^s and write it as a linear combination of independent Brownian motions $\tilde{W}^s$ and W^v. An application of Itô's formula leads to

$$S_T = S_0 \exp\left\{ rT - \frac{1}{2}\int_0^T v_t \, dt + \sqrt{1-\rho^2}\int_0^T \sqrt{v_t}\, d\tilde{W}_t^s + \rho \int_0^T \sqrt{v_t}\, dW_t^v \right\}. \tag{2.158}$$

In particular, we are interested in the evaluation of moments $\mathbb{E}\left[S_T^p\right]$ for $p > 1$. Andersen and Piterbarg [7] give several examples of fixed income securities with superlinear payoffs whose risk-neutral valuation involves the calculation of the second moment. Hence, moment explosions may lead to infinite prices of derivatives. Moreover, establishing the existence of moments of order higher than 1 of the process and its approximation is also important for the convergence analysis.

Conditioning on the σ-algebra $\mathcal{G}_T^v = \sigma(W_t^v;\, 0 \leq t \leq T)$, we find that

$$\mathbb{E}\big[S_T^p\big] = S_0^p\, \mathbb{E}\left[\exp\left\{ prT + \left[\frac{1}{2}p(p-1) - \frac{1}{2}p^2\rho^2\right] \int_0^T v_t \, dt + p\rho \int_0^T \sqrt{v_t}\, dW_t^v \right\}\right]. \tag{2.159}$$

Define the *explosion time* of the moment of order p to be the first time beyond which the moment $\mathbb{E}\left[S_T^p\right]$ will cease to exist, i.e.,

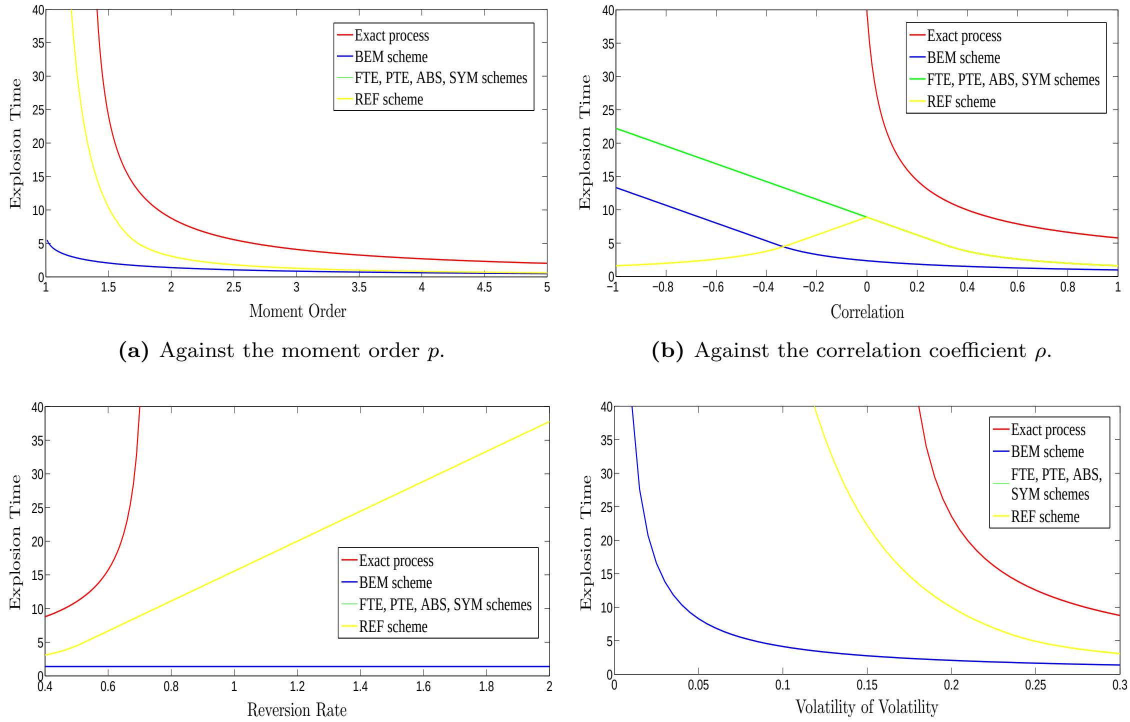
$$T^*(p) \equiv \sup\big\{ t \geq 0 : \, \mathbb{E}\left[S_t^p\right] < \infty \big\}. \tag{2.160}$$

Observe that if the moment does not explode in finite time, then $T^*(p) = \infty$.

For convenience, we fix $S_0 = 1$ and $r = 0$. Proposition 2.2 derives sharp conditions on the finiteness of moments of the process, whereas Propositions 2.4 and 2.6 to 2.9 give lower bounds on the explosion times of moments of the different discretizations in the limit as the time step goes to zero. For illustration, we plot in Figure 2.1 the explosion time and the corresponding lower bounds with different schemes against the model parameters. On the one hand, since

$$\varDelta = \frac{1}{2}p(p-1), \tag{2.161}$$

Propositions 2.2, 2.4 and 2.6 to 2.9 ensure the finiteness of the first moment of the process and its discretizations for all T, i.e., $T^*(1) = \infty$. On the other hand, we infer from the data in Figure 2.1a that both the explosion time for the exact process and the lower bounds with the explicit Euler discretizations approach infinity as p approaches one, i.e., that $\lim_{p\to 1^+} T^*(p) = \infty$. This ensures the uniform boundedness of moments, for p sufficiently close to one, of the explicit schemes even for very long maturities, an important ingredient in proving the strong convergence of the approximation process. Note that the green and the yellow curves in Figure 2.1 overlap when $\rho \geq 0$.



(a) Against the moment order p. **(b)** Against the correlation coefficient ρ.

(c) Against the mean reversion rate k. **(d)** Against the volatility of volatility ξ.

Figure 2.1: The explosion time of moments of the exact process and the lower bounds with the different discretization schemes, plotted against the model parameters when $k = 0.4$, $\xi = 0.3$, $\rho = 0.5$ and $p = 2$, except for the one which is varied.

The data in Figure 2.1b suggest that there exists a critical correlation level ρ^* such that $\mathbb{E}\left[S_T^p\right] < \infty$ for all T, provided $\rho \leq \rho^*$, and $\mathbb{E}\left[S_T^p\right] = \infty$ for some T, provided $\rho > \rho^*$. When $k = 0.4$, $\xi = 0.3$ and $p = 2$, we find $\rho^* = -0.04$. Moreover, we also infer from the data in Figure 2.1b that decreasing the correlation has a damping effect on the second moment of the process, and that for strongly negative correlations between the underlying process and the variance, as is usually the case in equity markets (the so-called *leverage effect*), the lower bounds on the explosion time of the second moment – with all but the reflection scheme – are above the typical maturity range of equity derivatives.

The data in Figures 2.1c and 2.1d indicate that increasing the speed of mean reversion and decreasing the volatility of volatility have a damping effect on the second moment of the process and its explicit Euler discretizations, which is to be expected considering that larger values of k and smaller values of ξ lead to smaller fluctuations in the variance over time. Next, we assign the following values to the underlying model parameters: $k = 0.4$,

$\theta = 0.12$, $\xi = 0.3$, $v_0 = 0.12$ and $\rho = 1$, and note that the Feller condition is satisfied. Henceforth, we estimate the second moment by a standard Monte Carlo estimator.

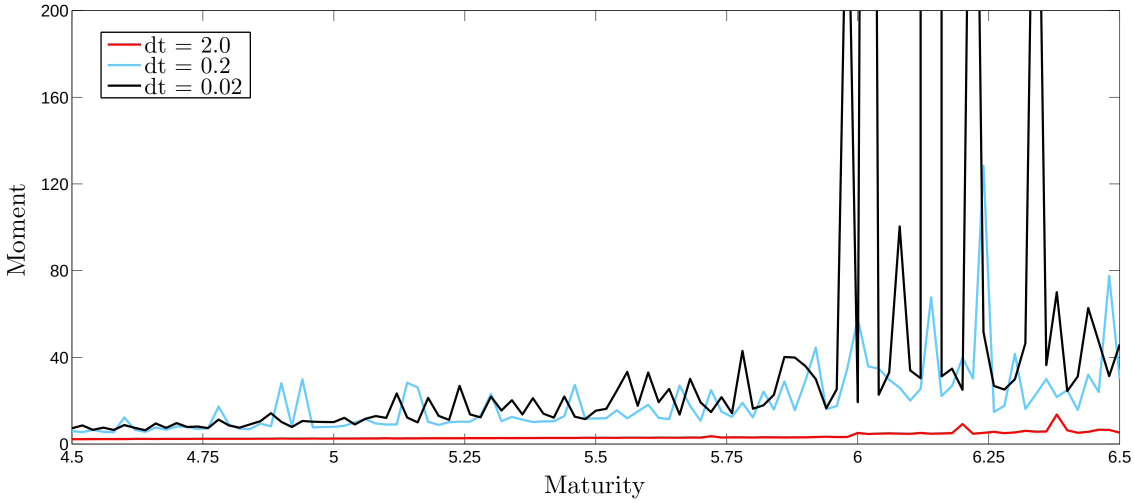


Figure 2.2: The second moment of the Heston model, calculated using the BEM scheme and 3×10^7 simulations, plotted against the maturity T.

From (2.23), the explosion time of the second moment of the process is $T^* = 5.77$. On the other hand, we infer from the data in Figure 2.2 that, for sufficiently small values of the time step – for instance, when $\delta t = 0.02$, the second moment with the BEM scheme will cease to exist after some time T^{BEM} close to T^*. The first sign of moment explosion can be observed when $T^{\text{BEM}} = 5.98$, however this phenomenon is more pronounced when $T = 6.14$, where the moment jumps to 9.7×10^4.

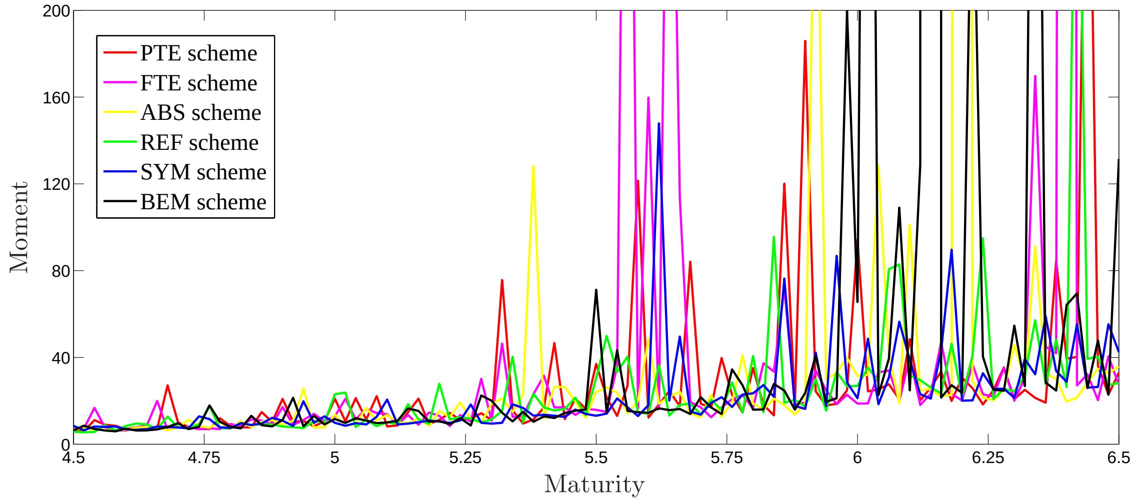


Figure 2.3: The second moment of the Heston model, calculated with different discretization schemes, $\delta t = 0.02$ and 3×10^7 simulations, plotted against the maturity T.

By close inspection of the data in Figure 2.3, we infer that, for sufficiently small values of the time step δt, the approximation to the second moment with either the implicit or one of the explicit schemes considered in this chapter explodes after some critical time close to T^*. Combined with the previous observation, this suggests that the explosion times of the second moments of the process and its discretizations become close as we increase the number of time steps. Therefore, we deduce from the data in Figure 2.1 that the lower bounds for the partial truncation, the full truncation, the absorption and the symmetrized Euler schemes are sharper than the ones for the drift-implicit and the reflection schemes.

2.5. Conclusions

In this chapter, we have established the uniform exponential integrability of functionals of the Cox–Ingersoll–Ross process and a number of its discretization schemes often encountered in the finance literature. One consequence of this result with obvious practical implications is the stability of moments of numerical approximations for a large class of SDEs arising in finance, which in turn is used to prove strong convergence [48]. An open question is whether we can find sharp conditions on the exponential integrability of Euler approximations for the CIR process.

Chapter 3

Monte Carlo Pricing under a Hybrid Stochastic-Local Volatility Model

3.1. Introduction

In this chapter, we study the FX option pricing problem under the Heston–Cox–Ingersoll–Ross++ stochastic-local volatility (Heston–2CIR++ SLV) model, which is a combination of two existing models: the Heston SLV model for the spot FX rate and the CIR++ model for the domestic and foreign short rates. To the best of our knowledge, the work in this chapter is the first to consider Heston SLV dynamics for the spot FX rate together with CIR++ (or even CIR) dynamics for the stochastic short rates.

We examine convergence properties of the Monte Carlo algorithm with the full truncation Euler (FTE) discretization for the squared volatility and the two short interest rates, and the log-Euler discretization for the exchange rate. We prefer the full truncation scheme introduced by Lord et al. [73] because it preserves the positivity of the original process, is easy to implement and is found empirically to produce the smallest bias of all explicit Euler schemes and to outperform the quasi-second order schemes of Kahl and Jäckel [61] and Ninomiya and Victoir [77]. We also choose the FTE scheme over the backward Euler–Maruyama (BEM) scheme since the latter works only in a restricted parameter regime that is often unrealistic, since the BEM scheme is equivalent to the implicit Milstein method employed by Kahl and Jäckel to first order in the time step, and since the quanto correction term in the dynamics of the foreign short rate would lead to technical challenges in the convergence analysis.

The quadratic exponential (QE) scheme of Andersen [6] uses moment-matching techniques to approximate the noncentral chi-squared distributed square-root process and is an efficient alternative to the FTE scheme that typically produces a smaller bias at the expense of a more complex implementation. Empirical results [92] suggest that the QE scheme outperforms the FTE scheme in the case of a severely violated Feller condition, whereas when the Feller condition is satisfied, the two schemes perform comparably well because the bias with the FTE scheme is small even when only a few number of time steps are used per year. Since there is no natural extension to the QE scheme to multi-dimensional Heston-type models that involve several correlated square-root processes such as the ones studied here, we prefer the FTE scheme in the subsequent convergence analysis.

The usual theorems in Kloeden and Platen [64] on the convergence of numerical simulations require the drift and diffusion coefficients to be globally Lipschitz and satisfy a linear growth condition, whereas Higham et al. [48] extended the analysis to a simple Euler scheme for a locally Lipschitz SDE. However, the classical convergence theory does not apply to the CIR process because the square-root diffusion coefficient is not locally Lipschitz around zero. Consequently, alternative approaches have been employed to prove the weak or strong convergence of various discretizations for the square-root process. Alfonsi [3], Deelstra and Delbaen [24], Higham and Mao [47] and Lord et al. [73] examined the strong global approximation and found either a logarithmic convergence rate or none at all. Strong convergence of order $1/2$ of the symmetrized and the BEM schemes was established in Berkaoui et al. [9] and Dereich et al. [26], respectively, albeit in a very restricted parameter regime for the symmetrized scheme. Alfonsi [4] and Neuenkirch and Szpruch [76] recently showed that the BEM scheme for the SDE obtained through a Lamperti transformation is strongly convergent with rate one in the case of an inaccessible boundary point, whereas Hutzenthaler et al. [56] established a positive strong order of convergence in the case of an accessible boundary point.

To the best of our knowledge, the convergence of Monte Carlo algorithms in a stochastic-local volatility context has yet to be established. Higham and Mao [47] considered an Euler simulation of the Heston model with a reflection fix in the diffusion coefficient to avoid negative values. They studied convergence properties of the stopped approximation process and used the boundedness of payoffs to prove weak convergence for a European put and an up-and-out call option. However, Higham and Mao mentioned that the arguments cannot be extended to cope with unbounded payoffs. We work under a different Euler scheme and overcome this problem by proving the uniform boundedness of moments of the true solution and its approximation, and then the strong convergence of the latter. The existence of moment bounds for Euler approximations of the Heston model is important for deriving strong convergence [48] and it has been a long-standing open problem until now. Furthermore, the existence of moment bounds for hybrid Heston-type stochastic-local volatility models plays an important role in the calibration routine [20].

In summary, to the best of our knowledge, the work in this chapter is the first to establish: (1) the boundedness of moments of order higher than 1 of approximation schemes for Heston-type models; (2) the strong convergence of approximation schemes for models with stochastic-local volatility (with a bounded and Lipschitz leverage function) and stochastic CIR++ short rates; (3) the weak convergence for options with unbounded payoffs, in particular, for European, Asian and barrier contracts (up to a critical time).

The remainder of this chapter is structured as follows. In Section 3.2, we introduce the model and discuss the postulated assumptions and an efficient calibration. In Section 3.3, we first define the simulation scheme and discuss the main theorem. Then, we prove the convergence of the exchange rate approximations. We conclude the section by establishing the convergence of Monte Carlo simulations for computing the expected discounted payoffs of European, Asian and barrier options. In Section 3.4, we calibrate the model to market data and include only a brief discussion on the calibration of the leverage function, which is mostly the work of Matthieu Mariapragassam. Section 3.5 summarizes the results and outlines possible future work. Finally, we justify our choice of model and demonstrate convergence in Appendix 3.6. This appendix is joint work with Matthieu Mariapragassam, who carried out the numerical experiments.

3.2. Preliminaries

3.2.1. Model definition

In its most general form, we have in mind a model in an FX market, for the spot FX rate S, the squared volatility of the FX rate v, the domestic short interest rate r^d and the foreign short interest rate r^f. Unless otherwise stated, in this chapter, the subscripts and superscripts "d" and "f" indicate domestic and foreign, respectively. Consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q})$ satisfying the usual conditions and suppose that the dynamics of the underlying processes are governed by the following system of SDEs under the domestic risk-neutral measure $\mathbb{Q}$:

$$\begin{cases} dS_t = \left(r_t^d - r_t^f\right) S_t dt + \sigma(t, S_t)\sqrt{v_t} S_t \, dW_t^s, \quad S_0 > 0, \\ dv_t = k(\theta - v_t)dt + \xi\sqrt{v_t}\, dW_t^v, \quad v_0 > 0, \\ \;\; r_t^d = g_t^d + h_d(t), \\ dg_t^d = k_d(\theta_d - g_t^d)dt + \xi_d\sqrt{g_t^d}\, dW_t^d, \quad g_0^d > 0, \\ \;\; r_t^f = g_t^f + h_f(t), \\ dg_t^f = \left(k_f\theta_f - k_f g_t^f - \rho_{sf}\xi_f\sigma(t, S_t)\sqrt{v_t g_t^f}\right)dt + \xi_f\sqrt{g_t^f}dW_t^f, \quad g_0^f > 0, \end{cases} \tag{3.1}$$

where $\{W^s, W^v, W^d, W^f\}$ are 1-dimensional standard $\mathcal{F}_t$-adapted Brownian motions that are possibly correlated, σ is the leverage function and h_d and h_f are deterministic functions of time. The mean-reversion parameters k, k_d and k_f, the long-term mean parameters θ, θ_d and θ_f, and the volatility parameters ξ, ξ_d and ξ_f are positive real numbers. The quanto correction term in the drift of the foreign short rate in (3.1) comes from changing from the foreign to the domestic risk-neutral measure [16]. As an aside, for Hull–White short rate processes [50], changing from the domestic spot measure to the domestic T-forward measure leads to a dimension reduction of the problem because the diffusion coefficients of the short rates do not depend on the level of the rates [43]. Since this is not the case for CIR short rate processes, we prefer to work under the domestic spot measure. Note that the above system can collapse to the Heston–2CIR++ model if we set $\sigma = 1$, or to a local volatility model with stochastic short rates if we set $k = \xi = 0$. The standard Heston SLV model is the special case $k_d = \xi_d = k_f = \xi_f = h_{d,f} = 0$. We can also think of (3.1) as a model in an equity market with stock price process S, stochastic interest rate r^d and stochastic dividend yield r^f, in which case the quanto correction term vanishes. We consider a full correlation structure between the Brownian drivers $\{W^s, W^v, W^d, W^f\}$, i.e., no assumptions on the constant correlation matrix Σ are made, where

$$\Sigma = \begin{bmatrix} 1 & \rho_{sv} & \rho_{sd} & \rho_{sf} \\ \rho_{sv} & 1 & \rho_{vd} & \rho_{vf} \\ \rho_{sd} & \rho_{vd} & 1 & \rho_{df} \\ \rho_{sf} & \rho_{vf} & \rho_{df} & 1 \end{bmatrix}. \tag{3.2}$$

Fix a time horizon $T > 0$. We work under the following assumptions:

($\mathcal{A}3.1$) There exists a non-negative constant σ_{max} such that, for all $t \in [0, T]$ and $x \in [0, \infty)$, we have

$$0 \leq \sigma(t, x) \leq \sigma_{max}\,. \tag{3.3}$$

($\mathcal{A}3.2$) There exist $N_T \in \mathbb{N}$, a positive real number α and non-negative constants A, B and $(C_j)_{1 \leq j \leq N_T}$ such that, for all $t, u \in [0, T]$ and $x, y \in [0, \infty)$, we have

$$|\sigma(t, x) - \sigma(u, y)| \leq A\,|t - u|^{\alpha} + B\,|x - y| + \sum_{j=1}^{N_T} C_j\, \mathbb{1}_{t \wedge u < \frac{jT}{N_T} \leq t \vee u}. \tag{3.4}$$

($\mathcal{A}3.3$) There exists a non-negative constant h_{max} such that, for all $t \in [0, T]$ and $i \in \{d, f\}$, we have

$$|h_i(t)| \leq h_{max}\,. \tag{3.5}$$

Hence, we assume that σ is bounded, Hölder continuous in time except at a finite number of points and Lipschitz continuous in spot. According to [20], for the leverage function to be consistent with call and put prices, it has to be given by the formula (3.8), which depends on the calibrated Dupire local volatility. In practice, the leverage function is defined on a grid of points (20 points per year in time and 30 points in space usually suffice for an acceptable calibration error [20; 44]), interpolated flat-forward in time and using cubic splines in spot, and extrapolated flat outside an interval. Therefore, there is no loss of generality from a practical point of view in making these assumptions. We also assume that $h_{d,f}$ are bounded. According to [12], for a perfect fit to the initial term structure of interest rates, each of the two shift functions must be given by the difference between the (flat-forward) market instantaneous forward rate and a continuous function of time. Then $h_{d,f}$ are piecewise continuous and the third assumption holds.

3.2.2. Model calibration

The calibration of the 4-factor Heston–2CIR++ SLV model (3.1) is of paramount importance and represents a mandatory step for the efficient pricing of derivative contracts. In Section 3.4, we calibrate the 4-factor SLV model using a new approach [20] that builds on the particle method of [44], combined with a novel and efficient variance reduction technique that takes advantage of PDE calibration to increase its stability and accuracy. The numerical experiments in [20] suggest that this method almost recovers the calibration speed from the associated 2-factor SLV model with deterministic rates. Next, we assume a partial correlation structure where only ρ_{sv}, ρ_{sd} and ρ_{sf} may be non-zero, and denote by D^d and D^f the domestic and foreign discount factors associated with their respective money market accounts, i.e., for any $t \in [0, T]$,

$$D_t^d = e^{-\int_0^t r_u^d du} \;\text{ and }\; D_t^f = e^{-\int_0^t r_u^f du}. \tag{3.6}$$

In addition to (3.1), we consider a pure local volatility (LV) model,

$$dS_t^{\mathrm{LV}} = \big(f_t^d - f_t^f\big) S_t^{\mathrm{LV}} dt + \sigma_{\mathrm{LV}}\big(t, S_t^{\mathrm{LV}}\big) S_t^{\mathrm{LV}} dW_t^s, \quad S_0^{\mathrm{LV}} = S_0, \tag{3.7}$$

where for $i \in \{d, f\}$, $f_t^i = -\frac{\partial}{\partial t} \log P^i(0, t)$ is the market instantaneous forward rate at time 0 for a maturity t and $P^i(0, t)$ is the market zero-coupon bond price at time 0 for a maturity t.

Suppose that the LV model (3.7) has been calibrated and that σ_{LV} has been determined. The main building block of the calibration routine is expressing the leverage function σ in terms of the local volatility function σ_{LV}. We state the necessary and sufficient condition for a perfect calibration to vanilla options in the following theorem, proved in [20].

Theorem 3.1 (Theorem 1 in [20]). *Suppose that both the marginal density of S_T from (3.1) and $\mathbb{E}\big[D_T^d v_T \,|\, S_T = \cdot\big]$ are continuous and that assumptions ($\mathcal{A}3.1$) and ($\mathcal{A}3.3$) hold. Then the call price under model (3.1) matches the call price C_{LV} under model (3.7) for any maturity up to T^* and any strike if and only if, for all $T, K > 0$,*

$$\begin{aligned}\sigma^2(T, K) = \frac{\mathbb{E}\big[D_T^d \,|\, S_T = K\big]}{\mathbb{E}\big[D_T^d v_T \,|\, S_T = K\big]} \Bigg\{ \sigma_{\mathrm{LV}}^2(T, K) + \frac{2}{K^2 \frac{\partial^2 C_{\mathrm{LV}}}{\partial K^2}} \Bigg(\mathbb{E}\Big[D_T^d \big(r_T^f - f_T^f\big)\big(S_T - K\big)^+\Big] \\ - K\, \mathbb{E}\Big[D_T^d \big(r_T^d - f_T^d\big) \mathbb{1}_{S_T \geq K}\Big] + K\, \mathbb{E}\Big[D_T^d \big(r_T^f - f_T^f\big) \mathbb{1}_{S_T \geq K}\Big] \Bigg) \Bigg\},\end{aligned} \tag{3.8}$$

where all expectations are under the domestic risk-neutral measure, σ_{LV} is a local volatility as in (3.7) and, denoting $\varphi = 2 + \sqrt{2}$ and $\zeta = \xi \sigma_{max}$, T^* is as given below.

(1) When $k < \varphi\zeta$,

$$T^* = \frac{2}{\sqrt{\varphi^2 \zeta^2 - k^2}} \left[\frac{\pi}{2} + \arctan\left(\frac{k}{\sqrt{\varphi^2 \zeta^2 - k^2}} \right) \right]. \tag{3.9}$$

(2) When $k \geq \varphi\zeta$,

$$T^* = \infty. \tag{3.10}$$

This result generalizes the formula in [44] to a stochastic foreign short rate. Deelstra and Rayee [25] obtained a similar formula for a 4-factor SLV model with Hull–White short rate processes. An important step in the derivation is that the stochastic integral

$$\int_0^T \mathbb{1}_{S_t \geq K}\, \sigma(t, S_t)\sqrt{v_t} D_t^d S_t \, dW_t^s \tag{3.11}$$

is a true martingale. On the one hand, T^* is a lower bound on the explosion time of the second moment of the discounted spot process $D_t^d S_t$ (see Proposition 3.15). On the other hand, moments of the Heston model can explode in finite time [7], a property that is inherited by the Heston-type model (3.1). Note also Theorem 3.2 and Proposition 3.7 for possible moment explosions of the numerical approximations. Therefore, the formula (3.8) may not hold for some values of the model parameters and for large maturities T. However, in practice, T^* is very large. For instance, following our calibration routine to EURUSD market data from Section 3.4, we find $T^* = 27.9$.

Note that if the volatility in model (3.1) is purely local, i.e., if $v = 1$, we recover the formula in [16], whereas if the two short rates are deterministic, i.e., if $g^d = g^f = 0$, we recover the formula in [44; 80], i.e.,

$$\sigma(T, K) = \frac{\sigma_{\mathrm{LV}}(T, K)}{\sqrt{\mathbb{E}[v_T \mid S_T = K]}}. \tag{3.12}$$

In the computations in Section 3.4, we assume a simple correlation structure where only ρ_{sv} may be non-zero. We use a five-step calibration routine which is explained in detail in [20]. First, we calibrate the LV model (3.7) and find σ_{LV}. We also calibrate the associated 2-factor SLV model with deterministic rates via PDE methods, using formula (3.12) and results from the literature [16; 80]. Independently, we calibrate the two CIR++ processes under their respective markets. Then, we find the Heston parameters by calibrating the associated 4-factor SV model, i.e., the Heston–2CIR++ SV model, using the closed-form solution for the call price in [1], extended to CIR++ processes, with other parameters from the above calibration. Finally, we calibrate σ in the Heston–2CIR++ SLV model (3.1) by working with the particle method of [44] and extending it using the calibrated 2-factor SLV model with deterministic rates as a control variate.

For a partial correlation structure with possibly non-zero ρ_{sd} and ρ_{sf}, we may use the historical values of the correlations between the spot FX rate and the short-term zero-coupon bonds, and extend the approximation formulae of [42] for the call price from a 3-factor Heston–CIR model to a 4-factor Heston–2CIR++ model. Alternatively, we may perform an initial calibration of the stochastic volatility parameters under the assumption that $\rho_{sd} = \rho_{sf} = 0$ to find $\beta_0 = (v_0, k, \theta, \xi, \rho_{sv})$. Then, we compute the exact model price at β_0 with non-zero ρ_{sd} and ρ_{sf}, C_{MC}, using Monte Carlo simulations with the closed-form solution in [1] for the call price when $\rho_{sd} = \rho_{sf} = 0$, C_{A}, as a control variate, and define the difference between the two prices to be $\Delta_0 = C_{\text{MC}} - C_{\text{A}}$. Assuming that the absolute values of the two correlations ρ_{sd} and ρ_{sf} are small, we calibrate the vector of stochastic volatility parameters $\beta = (v_0, k, \theta, \xi, \rho_{sv})$ in a neighborhood of β_0 by approximating the call prices with $C_{\text{A}}(\beta) + \Delta_0$.

3.3. Convergence analysis

3.3.1. The simulation scheme

We employ the full truncation Euler (FTE) scheme from [73] to discretize the variance and the two short rate processes. Consider the CIR process

$$dy_t = k_y(\theta_y - y_t)dt + \xi_y \sqrt{y_t}\, dW_t^y. \tag{3.13}$$

Let $N \in \mathbb{N}$ be a multiple of N_T and create an evenly spaced grid

$$T = N\delta t, \;\; t_n = n\delta t, \quad \forall n \in \{0, 1, ..., N\}. \tag{3.14}$$

First, we introduce the discrete-time auxiliary process

$$\tilde{y}_{t_{n+1}} = \tilde{y}_{t_n} + k_y(\theta_y - \tilde{y}_{t_n}^+)\delta t + \xi_y \sqrt{\tilde{y}_{t_n}^+}\, \delta W_{t_n}^y, \tag{3.15}$$

where $y^+ = \max(0, y)$ and $\delta W_{t_n}^y = W_{t_{n+1}}^y - W_{t_n}^y$, and its continuous-time interpolation

$$\tilde{y}_t = \tilde{y}_{t_n} + k_y(\theta_y - \tilde{y}_{t_n}^+)(t - t_n) + \xi_y \sqrt{\tilde{y}_{t_n}^+}\big(W_t^y - W_{t_n}^y\big), \tag{3.16}$$

for any $t \in [t_n, t_{n+1})$, as suggested in [47]. Then, we define the non-negative processes

$$\begin{cases} \hat{y}_t = \tilde{y}_t^+ & (3.17) \\ \bar{y}_t = \tilde{y}_{t_n}^+ & (3.18) \end{cases}$$

whenever $t \in [t_n, t_{n+1})$. Let $\bar{v}$ and $\bar{g}^d$ be the FTE discretizations of v and g^d, respectively. Taking into account the presence of the quanto correction term in the drift of the foreign short rate, we similarly define

$$\begin{aligned} \tilde{g}_t^f = \tilde{g}_{t_n}^f &+ \Big[k_f\theta_f - k_f\big(\tilde{g}_{t_n}^f\big)^+ - \rho_{sf}\xi_f\sigma\big(t_n, \bar{S}_{t_n}\big)\sqrt{\tilde{v}_{t_n}^+\big(\tilde{g}_{t_n}^f\big)^+}\,\Big](t - t_n) \\ &+ \xi_f\sqrt{\big(\tilde{g}_{t_n}^f\big)^+}\big(W_t^f - W_{t_n}^f\big), \end{aligned} \tag{3.19}$$

where $\bar{S}$ is the continuous-time approximation of S defined below, as well as

$$\begin{cases} \hat{g}_t^f = \big(\tilde{g}_t^f\big)^+ & (3.20) \\ \bar{g}_t^f = \big(\tilde{g}_{t_n}^f\big)^+ & (3.21) \end{cases}$$

whenever $t \in [t_n, t_{n+1})$. For $i \in \{d, f\}$, the domestic and foreign short rate discretizations are

$$\bar{r}^i_t = \bar{g}^i_t + h_i(t). \tag{3.22}$$

Finally, we use an Euler–Maruyama scheme to discretize the log-exchange rate. Let x and $\bar{x}$ be the actual and the approximated log-processes, and let $\bar{S} = e^{\bar{x}}$ be the continuous-time approximation of S. For brevity, define $h = h_d - h_f$. Then the discrete method reads:

$$\bar{x}_{t_{n+1}} = \bar{x}_{t_n} + \int_{t_n}^{t_{n+1}} h(u)du + \Big(\bar{g}^d_{t_n} - \bar{g}^f_{t_n} - \frac{1}{2}\sigma^2\big(t_n, \bar{S}_{t_n}\big)\bar{v}_{t_n}\Big)\delta t + \sigma\big(t_n, \bar{S}_{t_n}\big)\sqrt{\bar{v}_{t_n}}\,\delta W^s_{t_n}. \tag{3.23}$$

However, we find it convenient to work with the continuous-time approximation

$$\bar{x}_t = \bar{x}_{t_n} + \int_{t_n}^{t} h(u)du + \Big(\bar{g}^d_t - \bar{g}^f_t - \frac{1}{2}\bar{\sigma}^2\big(t, \bar{S}_t\big)\bar{v}_t\Big)\big(t - t_n\big) + \bar{\sigma}\big(t, \bar{S}_t\big)\sqrt{\bar{v}_t}\,\Delta W^s_t, \tag{3.24}$$

where $\Delta W^s_t = W^s_t - W^s_{t_n}$ and $\bar{\sigma}\big(t, \bar{S}_t\big) = \sigma\big(t_n, \bar{S}_{t_n}\big)$ whenever $t \in [t_n, t_{n+1})$. Hence,

$$\bar{x}_t = x_0 + \int_0^t \Big(\bar{r}^d_u - \bar{r}^f_u - \frac{1}{2}\bar{\sigma}^2\big(u, \bar{S}_u\big)\bar{v}_u\Big)du + \int_0^t \bar{\sigma}\big(u, \bar{S}_u\big)\sqrt{\bar{v}_u}\,dW^s_u. \tag{3.25}$$

Note that the convergence of the continuous-time approximation ensures that the discrete method approximates the true solution accurately at the gridpoints. Using Itô's formula, we obtain

$$\bar{S}_t = S_0 + \int_0^t \big(\bar{r}^d_u - \bar{r}^f_u\big)\bar{S}_u\,du + \int_0^t \bar{\sigma}\big(u, \bar{S}_u\big)\sqrt{\bar{v}_u}\,\bar{S}_u\,dW^s_u. \tag{3.26}$$

Furthermore, we define $\hat{S}$ to be the piecewise constant continuous-time approximation of S, i.e.,

$$\hat{S}_t = \bar{S}_{t_n} \tag{3.27}$$

whenever $t \in [t_n, t_{n+1})$, approximation which will be of use when pricing path-dependent options. We prefer the log-Euler scheme to the standard Euler scheme to discretize the exchange rate process because the former preserves positivity and produces no discretization bias in the S-direction when σ is constant. Furthermore, if v, g^d and g^f are also constant, then the log-Euler scheme is exact.

3.3.2. The main theorem

Define the arbitrage-free price of an option and its approximation under (3.27),

$$U = \mathbb{E}\Big[e^{-\int_0^T r^d_t dt} f(S)\Big], \tag{3.28}$$

$$\bar{U} = \mathbb{E}\Big[e^{-\int_0^T \bar{r}^d_t dt} f(\hat{S})\Big]. \tag{3.29}$$

The payoff function f may depend on the entire path of the underlying process and the expectation is under the domestic risk-neutral measure. The following theorem combines the statements of Theorems 3.18 – 3.23 and the proof is postponed to Subsection 3.3.5.

Theorem 3.2. *If $2k_f\theta_f > \xi_f^2$ and under assumptions ($\mathcal{A}3.1$) to ($\mathcal{A}3.3$), the following statements hold.*

(1) The approximations to the values of the European put, the up-and-out barrier call and any barrier put option defined in (3.29) *converge as $\delta t \to 0$.*

(2) Denote $\zeta = \xi\sigma_{max}$, with σ_{max} from (3.3)*. If*

$$T < T^* \equiv \frac{4k}{\zeta^2}\,\mathbb{1}_{\zeta<2k} + \frac{1}{\zeta-k}\,\mathbb{1}_{\zeta\geq 2k}, \tag{3.30}$$

then the approximations to the values of the European call, Asian options, the down-and-in/out and the up-and-in barrier call option defined in (3.29) *converge as $\delta t \to 0$.*

Remark 3.3. If assumption ($\mathcal{A}3.1$) holds, then for the purpose of this chapter we choose the smallest upper bound on the leverage function, namely

$$\sigma_{max} = \sup\big\{\sigma(t,x)\,|\,t\in[0,T],\, x\in[0,\infty)\big\}. \tag{3.31}$$

Remark 3.4. If the exchange rate and the foreign short rate dynamics are independent of each other, i.e., if $\rho_{sf} = 0$, then the quanto correction term in the drift of the foreign short rate vanishes and Theorem 3.2 holds independent of the condition $2k_f\theta_f > \xi_f^2$, albeit with a simpler proof. If the domestic and foreign short rates are constant throughout the lifetime of the option and if, moreover, $\sigma(t,x) = 1$ for all $t\in[0,T]$ and $x\in[0,\infty)$, then the system of equations (3.1) collapses to the Heston model and Theorem 3.2 applies with $\zeta = \xi$. This extends the convergence results of Higham and Mao [47] to options with unbounded payoffs.

Remark 3.5. Tian et al. [89] calibrated the term-structure Heston SLV model parameters to the market implied volatility data on EURUSD using 10 term structures and maturities up to 5 years. The data in Table 3.1 suggest that the condition in (3.30) is typically satisfied in practice, both for shorter and longer maturities.

Table 3.1: The calibrated Heston SLV parameters for EURUSD market data from August 23, 2012 [89].

T	k	ξ	σ_{max}	ζ	T^*
1 month	0.885	0.342	1.600	0.547	11.8 years
5 years	0.978	0.499	1.300	0.649	9.3 years

In Section 3.4, the Heston–2CIR++ SLV model is calibrated to the EURUSD market data from March 18, 2016, for maturities up to 5 years. The data in Table 3.2 indicate that the condition in (3.30) is satisfied even for very long maturities. Furthermore, the following parameter values for the foreign short rate process are recovered: $k_f = 0.011$, $\theta_f = 1.166$, $\xi_f = 0.037$, so that the other condition also holds, i.e., $2k_f\theta_f = 0.0257 \gg 0.0014 = \xi_f^2$.

Table 3.2: The calibrated Heston–2CIR++ SLV parameters for EURUSD market data from March 18, 2016.

T	k	ξ	σ_{max}	ζ	T^*
5 years	1.412	0.299	1.399	0.418	32.3 years

In equity markets, the mean-reversion speed is usually several times greater than the volatility of volatility. For instance, Hurn et al. [53] calibrated the Heston model for the S&P 500 index from January 1990 to December 2011 using a combination of two out-of-the-money options, and found that $k = 1.977 \gg 0.456 = \xi$. Furthermore, we typically have $\zeta < 1$, so that the condition in (3.30) holds even for very long maturities.

3.3.3. The square-root process

In order to prove the convergence of the approximation scheme in (3.26), we first need to examine the stability of the moments of order higher than 1 of the actual and the discretized processes. However, this problem is directly related to the exponential integrability of the CIR process and its approximation. Let y be the CIR process in (3.13) and $\bar{y}$ be the piecewise constant FTE interpolant as per (3.18).

Proposition 3.6. *Let $\lambda > 0$ and define the stochastic process*

$$\Theta_t \equiv \exp\left\{\lambda \int_0^t y_u \, du\right\}, \quad \forall t \geq 0. \tag{3.32}$$

If $T < T^*$, then the first moment of Θ_T is bounded, i.e.,

$$\mathbb{E}\left[\Theta_T\right] < \infty, \tag{3.33}$$

where T^* is as given below.

(1) When $k_y^2 < 2\lambda\xi_y^2$,

$$T^* = \frac{2}{\sqrt{2\lambda\xi_y^2 - k_y^2}}\left[\frac{\pi}{2} + \arctan\left(\frac{k_y}{\sqrt{2\lambda\xi_y^2 - k_y^2}}\right)\right]. \tag{3.34}$$

(2) When $k_y^2 \geq 2\lambda\xi_y^2$,

$$T^* = \infty. \tag{3.35}$$

Proof. Follows directly from Proposition 2.2, or from Proposition 3.1 in [7]. ■

Proposition 3.7. *Let $\lambda > 0$ and define the stochastic process*

$$\bar{\Theta}_t \equiv \exp\left\{\lambda \int_0^t \bar{y}_u \, du\right\}, \quad \forall t \geq 0. \tag{3.36}$$

If $T \le T^*$ and $\delta_T < k_y^{-1}$, then the first moment of $\bar{\Theta}_T$ is uniformly bounded, i.e.,

$$\sup_{\delta t \in (0,\delta_T)} \mathbb{E}\left[\bar{\Theta}_T\right] < \infty, \tag{3.37}$$

where T^* is as given below.

(1) When $2k_y^2 \le \lambda\xi_y^2$,

$$T^* = \frac{1}{\sqrt{2\lambda}\xi_y - k_y}. \tag{3.38}$$

(2) When $2k_y^2 > \lambda\xi_y^2$,

$$T^* = \frac{2k_y}{\lambda\xi_y^2}. \tag{3.39}$$

Proof. Following the arguments of Lemma 2.5 and Proposition 2.6 closely, we prove by induction on $0 \le m \le N$ that, for all $\delta t < \delta_T < k_y^{-1}$,

$$\begin{aligned}\mathbb{E}\left[\bar{\Theta}_T\right] \le \mathbb{E}\Bigg[&\exp\left\{\lambda\delta t \sum_{i=0}^{N-m-1} \bar{y}_{t_i} + \eta m\lambda\delta t\bar{y}_{t_{N-m}}\right\}\Bigg] \\ &\times \exp\left\{\frac{1}{2}\eta\big(k_y\theta_y + \vartheta_y\xi_y\big)\lambda(\delta t)^2(m-1)m\right\},\end{aligned} \tag{3.40}$$

where

$$\vartheta_y = \sqrt{\frac{\xi_y^2}{4\pi(1-k_y\delta_T)^2} + \frac{1}{2\pi}\sqrt{\frac{\xi_y^4}{4(1-k_y\delta_T)^4} + \frac{k_y^2\theta_y^2}{(1-k_y\delta_T)^2}}} \tag{3.41}$$

and $\eta \ge 1$ is independent of δt and such that, for all $\omega \in [0,1]$,

$$\eta^2\omega^2\lambda\xi_y^2T^2 - 2\eta\omega k_yT - 2\eta + 2 \le 0. \tag{3.42}$$

Taking $m = N$ in (3.40) leads to the conclusion. ■

The next two results establish uniform moment bounds for the original and the discretized variance and short rate processes. Note that under assumption ($\mathcal{A}3.3$), the moment boundedness of g^i and $\hat{g}^i$, for $i \in \{d,f\}$, extends naturally to $|r^i|$ and $|\bar{r}^i|$.

Proposition 3.8. *The following two statements hold.*

(1) The square-root process in (3.13) *has uniformly bounded moments, i.e.,*

$$\mathbb{E}\left[\sup_{t\in[0,T]} y_t^p\right] < \infty, \quad \forall p \ge 1. \tag{3.43}$$

(2) The FTE scheme in (3.17) *for the square-root process has uniformly bounded moments, i.e.,*

$$\sup_{\delta t\in(0,\delta_T)} \mathbb{E}\left[\sup_{t\in[0,T]} \hat{y}_t^p\right] < \infty, \quad \forall p \ge 1, \ \forall \delta_T > 0. \tag{3.44}$$

Proof. (1) The polynomial moments of the square-root process can be expressed in terms of the confluent hypergeometric function and, according to [26] or to Theorem 3.1 in [52], they are uniformly bounded, i.e.,

$$\sup_{t\in[0,T]} \mathbb{E}\left[y_t^p\right] < \infty, \quad \forall p > -\frac{2k_y\theta_y}{\xi_y^2}. \tag{3.45}$$

On the other hand, using Hölder's inequality, we deduce that

$$\begin{aligned}\sup_{t\in[0,T]} y_t^p \leq 3^{p-1}\big(y_0 + k_y\theta_y T\big)^p + 3^{p-1}k_y^p T^{p-1}\int_0^T y_u^p\,du \\ + 3^{p-1}\xi_y^p \sup_{t\in[0,T]}\left|\int_0^t \sqrt{y_u}\,dW_u^y\right|^p.\end{aligned} \tag{3.46}$$

Using the Burkholder–Davis–Gundy (BDG) inequality (see, e.g., [62]), we can find a constant $C_p > 0$ such that

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \sqrt{y_u}\,dW_u^y\right|^p\right] &\leq C_p\,\mathbb{E}\left[\left(\int_0^T y_u\,du\right)^{p/2}\right] \\ &\leq \frac{1}{2}C_p + \frac{1}{2}C_p T^{p-1}\,\mathbb{E}\left[\int_0^T y_u^p\,du\right].\end{aligned} \tag{3.47}$$

Taking expectations in (3.46) and employing Fubini's theorem, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]} y_t^p\right] \leq 3^{p-1}\big(y_0 + k_y\theta_y T\big)^p + \frac{1}{2}3^{p-1}\xi_y^p C_p \\ + 3^{p-1}\Big(k_y^p + \frac{1}{2}\xi_y^p C_p\Big)T^p \sup_{t\in[0,T]}\mathbb{E}\left[y_t^p\right].\end{aligned} \tag{3.48}$$

The right-hand side is finite by (3.45), whence the conclusion.

(2) Integrating the auxiliary process $\tilde{y}$ defined in (3.16), we deduce that

$$\tilde{y}_t = y_0 + k_y\int_0^t (\theta_y - \bar{y}_u)\,du + \xi_y\int_0^t \sqrt{\bar{y}_u}\,dW_u^y. \tag{3.49}$$

For any $p, \epsilon > 0$, there exists a constant $c(p,\epsilon) > 0$ such that $\max(0,x)^p \leq c(p,\epsilon)e^{\epsilon x}$ for all $x \in \mathbb{R}$. In particular, this implies that $\hat{y}_t^p \leq c(p,\epsilon)e^{\epsilon\tilde{y}_t}$ for all $t \in [0,T]$. Hence,

$$\sup_{t\in[0,T]}\mathbb{E}\left[\hat{y}_t^p\right] \leq c(p,\epsilon)\,e^{\epsilon(y_0+k_y\theta_y T)}\sup_{t\in[0,T]}\mathbb{E}\left[\Gamma_t\right], \tag{3.50}$$

where

$$\Gamma_t \equiv \exp\left\{-\epsilon k_y\int_0^t \bar{y}_u\,du + \epsilon\xi_y\int_0^t \sqrt{\bar{y}_u}\,dW_u^y\right\}. \tag{3.51}$$

Assuming that $t \in [t_n, t_{n+1}]$ and conditioning on the σ-algebra $\mathcal{G}_{t_n}^y$, we get

$$\mathbb{E}_{t_n}^y\left[\Gamma_t\right] = \Gamma_{t_n}\exp\left\{\frac{1}{2}\big(\epsilon^2\xi_y^2 - 2\epsilon k_y\big)\big(t-t_n\big)\bar{y}_{t_n}\right\}. \tag{3.52}$$

However, we can find ϵ sufficiently small such that $2k_y \geq \epsilon\xi_y^2$. Then $\mathbb{E}^y_{t_n}\left[\Gamma_t\right] \leq \mathbb{E}^y_{t_n}\left[\Gamma_{t_n}\right]$ and the law of iterated expectations ensures that

$$\mathbb{E}\left[\Gamma_t\right] \leq \mathbb{E}\left[\Gamma_{t_n}\right] \leq \mathbb{E}\left[\Gamma_{t_{n-1}}\right] \leq \ldots \leq \mathbb{E}\left[\Gamma_0\right] \;\Rightarrow\; \sup_{t\in[0,T]} \mathbb{E}\left[\Gamma_t\right] = 1. \tag{3.53}$$

Substituting back into (3.50), we deduce that

$$\sup_{t\in[0,T]} \mathbb{E}\left[\hat{y}_t^p\right] \leq c(p,\epsilon)\, e^{\epsilon(y_0+k_y\theta_y T)}, \tag{3.54}$$

which is independent of δt. On the other hand, since $\hat{y}_t \leq |\tilde{y}_t|$ and using Hölder's inequality, we get

$$\begin{aligned}\sup_{t\in[0,T]} \hat{y}_t^p \leq 3^{p-1}\big(y_0 + k_y\theta_y T\big)^p + 3^{p-1}k_y^p T^{p-1}\int_0^T \bar{y}_u^p\, du \\ + 3^{p-1}\xi_y^p \sup_{t\in[0,T]}\left|\int_0^t \sqrt{\bar{y}_u}\, dW_u^y\right|^p.\end{aligned} \tag{3.55}$$

Taking expectations on both sides and using the BDG inequality as in (3.47) leads to

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]} \hat{y}_t^p\right] \leq 3^{p-1}\big(y_0 + k_y\theta_y T\big)^p + \frac{1}{2}3^{p-1}\xi_y^p C_p \\ + 3^{p-1}\Big(k_y^p + \frac{1}{2}\xi_y^p C_p\Big)T^p \sup_{t\in[0,T]} \mathbb{E}\left[\bar{y}_t^p\right],\end{aligned} \tag{3.56}$$

for some constant $C_p > 0$. However, $\sup_{t\in[0,T]} \mathbb{E}\left[\bar{y}_t^p\right] \leq \sup_{t\in[0,T]} \mathbb{E}\left[\hat{y}_t^p\right]$, and the conclusion follows from (3.54). ■

Proposition 3.9. *Under assumption ($\mathcal{A}$3.1), the following two statements hold.*

(1) The process g^f has uniformly bounded moments, i.e.,

$$\mathbb{E}\left[\sup_{t\in[0,T]} \big(g_t^f\big)^p\right] < \infty, \quad \forall p \geq 1. \tag{3.57}$$

(2) The process $\hat{g}^f$ from (3.20) *has uniformly bounded moments, i.e.,*

$$\sup_{\delta t\in(0,\delta_T)} \mathbb{E}\left[\sup_{t\in[0,T]} \big(\hat{g}_t^f\big)^p\right] < \infty, \quad \forall p \geq 1,\;\; \forall \delta_T > 0. \tag{3.58}$$

Proof. (1) From (3.1), we know that

$$g_t^f = g_0^f + k_f\theta_f t - k_f\int_0^t g_u^f\, du - \rho_{sf}\xi_f\int_0^t \sigma(u,S_u)\sqrt{v_u g_u^f}\, du + \xi_f\int_0^t \sqrt{g_u^f}\, dW_u^f. \tag{3.59}$$

Since g^f is non-negative and using the fact that $2\sqrt{|ab|} \leq |a| + |b|$, we find an upper bound

$$\begin{aligned}g_t^f \leq g_0^f + k_f\theta_f t + \frac{1}{2}|\rho_{sf}|\xi_f\sigma_{max}\int_0^t v_u\, du + \frac{1}{2}\big(2k_f + |\rho_{sf}|\xi_f\sigma_{max}\big)\int_0^t g_u^f\, du \\ + \xi_f\left|\int_0^t \sqrt{g_u^f}\, dW_u^f\right|.\end{aligned} \tag{3.60}$$

Fix $t \in [0,T]$. Using Hölder's inequality, we get

$$\begin{aligned}\sup_{s\in[0,t]} \big(g_s^f\big)^p &\le 4^{p-1}\big(g_0^f + k_f\theta_f T\big)^p + 2^{p-2}|\rho_{sf}|^p \xi_f^p \sigma_{max}^p T^{p-1} \int_0^T v_u^p\, du \\ &+ 2^{p-2}\big(2k_f + |\rho_{sf}|\xi_f\sigma_{max}\big)^p T^{p-1} \int_0^t \big(g_u^f\big)^p du \\ &+ 4^{p-1}\xi_f^p \sup_{s\in[0,t]} \left| \int_0^s \sqrt{g_u^f}\, dW_u^f \right|^p. \end{aligned} \tag{3.61}$$

Taking expectations on both sides and using the BDG inequality as in (3.47) leads to

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]} \big(g_s^f\big)^p\right] &\le 4^{p-1}\big(g_0^f + k_f\theta_f T\big)^p + 2^{p-2}|\rho_{sf}|^p \xi_f^p \sigma_{max}^p T^p \sup_{u\in[0,T]} \mathbb{E}\left[v_u^p\right] \\ &+ 2^{2p-3}\xi_f^p C_p + \Big(2^{p-2}\big(2k_f + |\rho_{sf}|\xi_f\sigma_{max}\big)^p T^{p-1} \\ &+ 2^{2p-3}\xi_f^p C_p T^{p-1}\Big) \int_0^t \mathbb{E}\left[\sup_{s\in[0,u]} \big(g_s^f\big)^p\right] du, \end{aligned} \tag{3.62}$$

for some constant $C_p > 0$. Applying Gronwall's inequality (see, e.g., [28]), we get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]} \big(g_t^f\big)^p\right] &\le \Big(4^{p-1}\big(g_0^f + k_f\theta_f T\big)^p + 2^{2p-3}\xi_f^p C_p + 2^{p-2}|\rho_{sf}|^p \xi_f^p \sigma_{max}^p T^p \sup_{u\in[0,T]} \mathbb{E}\left[v_u^p\right]\Big) \\ &\times \exp\Big\{2^{p-2}\big(2k_f + |\rho_{sf}|\xi_f\sigma_{max}\big)^p T^p + 2^{2p-3}\xi_f^p C_p T^p\Big\}. \end{aligned} \tag{3.63}$$

The conclusion follows from Proposition 3.8.

(2) Integrating the auxiliary process $\tilde{g}^f$ defined in (3.19), we deduce that

$$\tilde{g}_t^f = g_0^f + k_f\theta_f t - k_f \int_0^t \bar{g}_u^f\, du - \rho_{sf}\xi_f \int_0^t \bar{\sigma}\big(u, \bar{S}_u\big)\sqrt{\bar{v}_u \bar{g}_u^f}\, du + \xi_f \int_0^t \sqrt{\bar{g}_u^f}\, dW_u^f. \tag{3.64}$$

Since $\hat{g}_t^f \le |\tilde{g}_t^f|$, following the previous argument, we deduce that

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]} \big(\hat{g}_s^f\big)^p\right] &\le 4^{p-1}\big(g_0^f + k_f\theta_f T\big)^p + 2^{p-2}|\rho_{sf}|^p \xi_f^p \sigma_{max}^p T^p \sup_{u\in[0,T]} \mathbb{E}\left[\bar{v}_u^p\right] \\ &+ 2^{2p-3}\xi_f^p C_p + \Big(2^{p-2}\big(2k_f + |\rho_{sf}|\xi_f\sigma_{max}\big)^p T^{p-1} \\ &+ 2^{2p-3}\xi_f^p C_p T^{p-1}\Big) \int_0^t \mathbb{E}\left[\big(\bar{g}_u^f\big)^p\right] du. \end{aligned} \tag{3.65}$$

Since $\sup_{u\in[0,T]} \mathbb{E}\left[\bar{v}_u^p\right] \le \sup_{u\in[0,T]} \mathbb{E}\left[\hat{v}_u^p\right]$ and $\bar{g}_u^f \le \sup_{s\in[0,u]} \hat{g}_s^f$, using Gronwall's inequality, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]} \big(\hat{g}_t^f\big)^p\right] &\le \Big(4^{p-1}\big(g_0^f + k_f\theta_f T\big)^p + 2^{2p-3}\xi_f^p C_p + 2^{p-2}|\rho_{sf}|^p \xi_f^p \sigma_{max}^p T^p \sup_{u\in[0,T]} \mathbb{E}\left[\hat{v}_u^p\right]\Big) \\ &\times \exp\Big\{2^{p-2}\big(2k_f + |\rho_{sf}|\xi_f\sigma_{max}\big)^p T^p + 2^{2p-3}\xi_f^p C_p T^p\Big\}. \end{aligned} \tag{3.66}$$

The conclusion follows from Proposition 3.8. ■

Next, we derive the strong mean square convergence of the discretized variance and domestic short rate processes. Lord et al. [73] proved the strong convergence (without a rate) of the continuous-time approximation $\hat{y}$. However, we are rather interested in the behaviour of $\bar{y}$ in the limit as the time step goes to zero. Note that the convergence of $\bar{g}^d$ extends naturally to $\bar{r}^d$.

Proposition 3.10. *The full truncation scheme in* (3.18) *for the square-root process converges strongly in L^2 in the sense that*

$$\lim_{\delta t \to 0} \sup_{t\in[0,T]} \mathbb{E}\left[|y_t - \bar{y}_t|^2\right] = 0. \tag{3.67}$$

Proof. First, fix $t \in [0,T]$. Since $|y_t - \hat{y}_t| \le |y_t - \tilde{y}_t|$, using the Cauchy–Schwarz inequality, we get

$$\sup_{s\in[0,t]} |y_s - \hat{y}_s|^2 \le 2k_y^2 t \int_0^t (y_u - \bar{y}_u)^2 du + 2\xi_y^2 \sup_{s\in[0,t]} \left| \int_0^s \left(\sqrt{y_u} - \sqrt{\bar{y}_u}\right) dW_u^y \right|^2. \tag{3.68}$$

Taking expectations and using Doob's martingale inequality and Fubini's theorem leads to

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]} |y_s - \hat{y}_s|^2\right] &\le 4k_y^2 T \int_0^t \mathbb{E}\left[\sup_{s\in[0,u]} |y_s - \hat{y}_s|^2\right] du + 4k_y^2 T^2 \sup_{u\in[0,T]} \mathbb{E}\left[|\hat{y}_u - \bar{y}_u|^2\right] \\ &\quad + 8\xi_y^2 T \sup_{u\in[0,T]} \mathbb{E}\left[|y_u - \hat{y}_u|\right] + 8\xi_y^2 T \sup_{u\in[0,T]} \mathbb{E}\left[|\hat{y}_u - \bar{y}_u|\right]. \end{aligned} \tag{3.69}$$

Applying Gronwall's inequality, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]} |y_t - \hat{y}_t|^2\right] &\le e^{4k_y^2 T^2} \Bigg(4k_y^2 T^2 \sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|^2\right] + 8\xi_y^2 T \sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|\right] \\ &\quad + 8\xi_y^2 T \sup_{t\in[0,T]} \mathbb{E}\left[|y_t - \hat{y}_t|\right] \Bigg). \end{aligned} \tag{3.70}$$

On the other hand, we know that

$$\sup_{t\in[0,T]} \mathbb{E}\left[|y_t - \bar{y}_t|^2\right] \le 2 \sup_{t\in[0,T]} \mathbb{E}\left[|y_t - \hat{y}_t|^2\right] + 2 \sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|^2\right]. \tag{3.71}$$

Substituting into (3.71) with the upper bound in (3.70), we get

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\left[|y_t - \bar{y}_t|^2\right] &\le 16\xi_y^2 T e^{4k_y^2 T^2} \sup_{t\in[0,T]} \mathbb{E}\left[|y_t - \hat{y}_t|\right] + 16\xi_y^2 T e^{4k_y^2 T^2} \sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|\right] \\ &\quad + 2\Big(1 + 4k_y^2 T^2 e^{4k_y^2 T^2}\Big) \sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|^2\right]. \end{aligned} \tag{3.72}$$

The convergence of the three terms on the right-hand side of (3.72) is a consequence of Lemma A.3 and Theorem 4.2 in [73], which concludes the proof. ■

We now establish a positive strong convergence rate for the FTE scheme in the case of an inaccessible boundary point. This is typically the case in practice when modeling interest rates, but not variances of FX rates (see, e.g., Tables 3.5 and 3.7). Let the Feller ratio be

$$\nu_y = \frac{2k_y\theta_y}{\xi_y^2}. \tag{3.73}$$

First, we prove the following auxiliary result.

Proposition 3.11. *Let $p > 1$, $\kappa > 0$, and define the stopping time*

$$\tau_{p,\kappa} = \inf\big\{t \geq 0 : \gamma_p(t) \geq \kappa\big\}, \tag{3.74}$$

where

$$\gamma_p(t) = \int_0^t \left(\frac{2p}{p-1}k_y^2 T + p^2\xi_y^2 y_u^{-1}\right)du. \tag{3.75}$$

If $2k_y\theta_y > \xi_y^2$, then there exists a positive constant C_p independent of κ and N such that

$$\sup_{t\in[0,T]} \mathbb{E}\left[|y_{t\wedge\tau_{p,\kappa}} - \hat{y}_{t\wedge\tau_{p,\kappa}}|^2\right] \leq C_p e^{\kappa} N^{-1}, \quad \forall N \geq 1. \tag{3.76}$$

Proof. Note that since the Feller condition is satisfied, $(y_t)_{t\geq 0}$ has strictly positive paths and hence $\gamma_p(t)$ is well-defined. Throughout the proof, we will use the fact that, for all $a, b \in \mathbb{R}$ and $q > 1$,

$$(a+b)^2 \leq \frac{q}{q-1}a^2 + qb^2. \tag{3.77}$$

Since $|y_t - \hat{y}_t| \leq |y_t - \tilde{y}_t|$, using (3.77) and the Cauchy–Schwarz inequality, we get

$$\begin{aligned}|y_{t\wedge\tau_{p,\kappa}} - \hat{y}_{t\wedge\tau_{p,\kappa}}|^2 \leq{}& \frac{p}{p-1}k_y^2 T\int_0^{t\wedge\tau_{p,\kappa}} |y_u - \bar{y}_u|^2 du \\ &+ p\xi_y^2\left|\int_0^t \left(\sqrt{y_u} - \sqrt{\bar{y}_u}\right)\mathbb{1}_{u<\tau_{p,\kappa}}\, dW_u^y\right|^2.\end{aligned} \tag{3.78}$$

Taking expectations and using the Itô isometry and the fact that

$$|\sqrt{y_u} - \sqrt{\bar{y}_u}|^2 \leq |y_u - \bar{y}_u|^2 y_u^{-1} \tag{3.79}$$

leads to

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\left[|y_{t\wedge\tau_{p,\kappa}} - \hat{y}_{t\wedge\tau_{p,\kappa}}|^2\right] \leq{}& \frac{p}{p-1}k_y^2 T\,\mathbb{E}\int_0^{T\wedge\tau_{p,\kappa}} |y_u - \bar{y}_u|^2 du \\ &+ p\xi_y^2\,\mathbb{E}\int_0^{T\wedge\tau_{p,\kappa}} |y_u - \bar{y}_u|^2 y_u^{-1} du.\end{aligned} \tag{3.80}$$

Using the triangle inequality and (3.77), we get

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\left[|y_{t\wedge\tau_{p,\kappa}} - \hat{y}_{t\wedge\tau_{p,\kappa}}|^2\right] \leq{}& \frac{p}{p-1}k_y^2 T\,\mathbb{E}\int_0^{T\wedge\tau_{p,\kappa}} \Big(2|y_u - \hat{y}_u|^2 + 2|\hat{y}_u - \bar{y}_u|^2\Big)du \\ &+ p\xi_y^2\,\mathbb{E}\int_0^{T\wedge\tau_{p,\kappa}} \left(p|y_u - \hat{y}_u|^2 + \frac{p}{p-1}|\hat{y}_u - \bar{y}_u|^2\right)y_u^{-1} du.\end{aligned} \tag{3.81}$$

Fix any $p_0 \in (1, \nu_y)$. Using Fubini's theorem and Hölder's inequality with the pair (p_0, q_0) leads to

$$\sup_{t\in[0,T]} \mathbb{E}\left[|y_{t\wedge\tau_{p,\kappa}} - \hat{y}_{t\wedge\tau_{p,\kappa}}|^2\right] \leq \mathbb{E}\int_0^{T\wedge\tau_{p,\kappa}} |y_u - \hat{y}_u|^2 d\gamma_p(u) + \frac{2p}{p-1} k_y^2 T^2 \sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|^2\right]$$
$$+ \frac{p^2}{p-1}\xi_y^2 T \sup_{t\in[0,T]} \mathbb{E}\left[y_t^{-p_0}\right]^{\frac{1}{p_0}} \sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|^{2q_0}\right]^{\frac{1}{q_0}}. \qquad (3.82)$$

Since $\gamma_p(t)$ is strictly increasing, we employ a stochastic time change idea from [9] and let $s = \gamma_p(u)$. Then $u = \tau_{p,s}$ and $\gamma_p(T \wedge \tau_{p,\kappa}) = \gamma_p(T) \wedge \gamma_p(\tau_{p,\kappa}) = \gamma_p(T) \wedge \kappa$, and hence

$$\begin{aligned}
\mathbb{E}\int_0^{T\wedge\tau_{p,\kappa}} |y_u - \hat{y}_u|^2 d\gamma_p(u) &= \mathbb{E}\int_0^{\gamma_p(T)\wedge\kappa} |y_{\tau_{p,s}} - \hat{y}_{\tau_{p,s}}|^2 ds \\
&= \int_0^{\kappa} \mathbb{E}\left[|y_{\tau_{p,s}} - \hat{y}_{\tau_{p,s}}|^2 \mathbb{1}_{\tau_{p,s}<T}\right] ds \\
&\leq \int_0^{\kappa} \sup_{t\in[0,T]} \mathbb{E}\left[|y_{t\wedge\tau_{p,s}} - \hat{y}_{t\wedge\tau_{p,s}}|^2\right] ds. \qquad (3.83)
\end{aligned}$$

Substituting back into (3.82) with the upper bound in (3.83) and then applying Gronwall's inequality, we deduce that

$$\sup_{t\in[0,T]} \mathbb{E}\left[|y_{t\wedge\tau_{p,\kappa}} - \hat{y}_{t\wedge\tau_{p,\kappa}}|^2\right] \leq e^{\kappa}\left\{\frac{p^2}{p-1}\xi_y^2 T \sup_{t\in[0,T]} \mathbb{E}\left[y_t^{-p_0}\right]^{\frac{1}{p_0}} \sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|^{2q_0}\right]^{\frac{1}{q_0}}\right.$$
$$\left. + \frac{2p}{p-1} k_y^2 T^2 \sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|^2\right]\right\}. \qquad (3.84)$$

Let $q \geq 1$ and suppose that $t \in [t_n, t_{n+1})$. Since $|\hat{y}_t - \bar{y}_t| \leq |\tilde{y}_t - \tilde{y}_{t_n}|$, we have

$$\begin{aligned}
|\hat{y}_t - \bar{y}_t|^{2q} &\leq \left(k_y\theta_y\delta t + k_y\delta t\,\hat{y}_{t_n} + \xi_y\sqrt{\hat{y}_{t_n}}\left|W_t^y - W_{t_n}^y\right|\right)^{2q} \\
&\leq 3^{2q-1}k_y^{2q}\theta_y^{2q}(\delta t)^{2q} + 3^{2q-1}k_y^{2q}(\delta t)^{2q}\hat{y}_{t_n}^{2q} + 3^{2q-1}\xi_y^{2q}\hat{y}_{t_n}^{q}\left|W_t^y - W_{t_n}^y\right|^{2q}. \qquad (3.85)
\end{aligned}$$

If $Z \sim \mathcal{N}(0,1)$, then

$$\sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|^{2q}\right] \leq 3^{2q-1}k_y^{2q}\theta_y^{2q}(\delta t)^{2q} + 3^{2q-1}k_y^{2q}(\delta t)^{2q} \sup_{0\leq n\leq N} \mathbb{E}\left[\hat{y}_{t_n}^{2q}\right]$$
$$+ 3^{2q-1}\xi_y^{2q}(\delta t)^{q} \sup_{0\leq n\leq N} \mathbb{E}\left[\hat{y}_{t_n}^{q}\right]\mathbb{E}\left[|Z|^{2q}\right], \qquad (3.86)$$

and hence, from (3.14),

$$\sup_{t\in[0,T]} \mathbb{E}\left[|\hat{y}_t - \bar{y}_t|^{2q}\right]^{\frac{1}{q}} \leq 3^{2-1/q}T\left(k_y^2\theta_y^2 T + k_y^2 T \sup_{\delta t\in(0,T]}\sup_{t\in[0,T]} \mathbb{E}\left[\hat{y}_t^{2q}\right]^{\frac{1}{q}}\right.$$
$$\left. + \xi_y^2\,\mathbb{E}\left[|Z|^{2q}\right]^{\frac{1}{q}} \sup_{\delta t\in(0,T]}\sup_{t\in[0,T]} \mathbb{E}\left[\hat{y}_t^{q}\right]^{\frac{1}{q}}\right)N^{-1}. \qquad (3.87)$$

Substituting back into (3.84) with the upper bound in (3.87) and using Jensen's inequality, we deduce that

$$\sup_{t\in[0,T]} \mathbb{E}\left[|y_{t\wedge\tau_{p,\kappa}} - \hat{y}_{t\wedge\tau_{p,\kappa}}|^2\right] \leq 3^{2-1/q_0}\frac{p}{p-1}T^2\left(2k_y^2 T + p\xi_y^2 \sup_{t\in[0,T]} \mathbb{E}\left[y_t^{-p_0}\right]^{\frac{1}{p_0}}\right)$$

$$\times\left(k_y^2\theta_y^2 T + k_y^2 T \sup_{\delta t\in(0,T]}\sup_{t\in[0,T]}\mathbb{E}\left[\hat{y}_t^{2q_0}\right]^{\frac{1}{q_0}} + \xi_y^2\,\mathbb{E}\left[|Z|^{2q_0}\right]^{\frac{1}{q_0}}\sup_{\delta t\in(0,T]}\sup_{t\in[0,T]}\mathbb{E}\left[\hat{y}_t^{q_0}\right]^{\frac{1}{q_0}}\right)e^{\kappa}N^{-1}. \tag{3.88}$$

The conclusion follows immediately from (3.45) and Proposition 3.8. ■

Proposition 3.12. *If $2k_y\theta_y > \xi_y^2$, then the full truncation scheme in (3.18) for the square-root process converges strongly in L^2 with order $\frac{1}{2}-\frac{4}{8+(\nu_y-1)^2}-$, i.e., for any $\epsilon>0$, there exists a positive constant C independent of N such that*

$$\sup_{t\in[0,T]}\mathbb{E}\left[|y_t-\bar{y}_t|^2\right]^{\frac{1}{2}} \le CN^{-\frac{1}{2}+\frac{4}{8+(\nu_y-1)^2}+\epsilon},\quad \forall N>1. \tag{3.89}$$

Proof. Fix $0<\epsilon<\frac{1}{2}-\frac{4}{8+(\nu_y-1)^2}$ and $N>1$, and let

$$p=\sqrt[3]{1+\frac{\epsilon[8+(\nu_y-1)^2]^2}{4(\nu_y-1)^2-8\epsilon[8+(\nu_y-1)^2]}} \tag{3.90}$$

and

$$\kappa=\frac{8p^3\log N}{8p^3+(\nu_y-1)^2}. \tag{3.91}$$

Then we can easily show that

$$\frac{1}{2}-\frac{4p^3}{8p^3+(\nu_y-1)^2}=\frac{1}{2}-\frac{4}{8+(\nu_y-1)^2}-\epsilon. \tag{3.92}$$

Note that

$$\sup_{t\in[0,T]}\mathbb{E}\left[|y_t-\bar{y}_t|^2\right]\le 2\sup_{t\in[0,T]}\mathbb{E}\left[|y_t-\hat{y}_t|^2\right]+2\sup_{t\in[0,T]}\mathbb{E}\left[|\hat{y}_t-\bar{y}_t|^2\right]. \tag{3.93}$$

On the one hand, we know from (3.87) that

$$\sup_{t\in[0,T]}\mathbb{E}\left[|\hat{y}_t-\bar{y}_t|^2\right]\le 3T\left(k_y^2\theta_y^2T+k_y^2T\sup_{\delta t\in(0,T]}\sup_{t\in[0,T]}\mathbb{E}\left[\hat{y}_t^2\right]+\xi_y^2\sup_{\delta t\in(0,T]}\sup_{t\in[0,T]}\mathbb{E}\left[\hat{y}_t\right]\right)N^{-1}. \tag{3.94}$$

On the other hand, using Hölder's inequality, we find an upper bound for the first expectation on the right-hand side of (3.93):

$$\begin{aligned}\sup_{t\in[0,T]}\mathbb{E}\left[|y_t-\hat{y}_t|^2\right]&\le\sup_{t\in[0,T]}\mathbb{E}\left[|y_t-\hat{y}_t|^2\,\mathbb{1}_{\tau_{p,\kappa}\le T}\right]+\sup_{t\in[0,T]}\mathbb{E}\left[|y_t-\hat{y}_t|^2\,\mathbb{1}_{\tau_{p,\kappa}>T}\right]\\&\le\left\{\sup_{t\in[0,T]}\mathbb{E}\left[y_t^{\frac{2p}{p-1}}\right]^{1-\frac{1}{p}}+\sup_{\delta t\in(0,T]}\sup_{t\in[0,T]}\mathbb{E}\left[\hat{y}_t^{\frac{2p}{p-1}}\right]^{1-\frac{1}{p}}\right\}\mathbb{P}\left(\tau_{p,\kappa}\le T\right)^{\frac{1}{p}}\\&\quad+\sup_{t\in[0,T]}\mathbb{E}\left[|y_{t\wedge\tau_{p,\kappa}}-\hat{y}_{t\wedge\tau_{p,\kappa}}|^2\right].\end{aligned} \tag{3.95}$$

Using Markov's inequality, we deduce that

$$\mathbb{P}\left(\tau_{p,\kappa}\le T\right)=\mathbb{P}\left(\frac{2p}{p-1}k_y^2T^2+p^2\xi_y^2\int_0^T y_t^{-1}dt\ge\kappa\right)$$

$$\leq \mathbb{E}\left[\exp\left\{\frac{1}{4p(p-1)}k_y^2(\nu_y-1)^2T^2+\frac{1}{8}\xi_y^2(\nu_y-1)^2\int_0^T y_t^{-1}dt\right\}\right]$$
$$\times\exp\left\{-\frac{1}{8p^2}(\nu_y-1)^2\kappa\right\}. \tag{3.96}$$

However, we know from Lemma A.2 in [11] or Theorem 3.1 in [52] that

$$\mathbb{E}\left[\exp\left\{\frac{1}{8}\xi_y^2(\nu_y-1)^2\int_0^T y_t^{-1}dt\right\}\right]<\infty. \tag{3.97}$$

Substituting back into (3.93) with the upper bound from Proposition 3.11 and (3.94) – (3.96), and then using Proposition 3.8 and (3.97), we can find positive constants c_1, c_2 and c_3 independent of N such that

$$\sup_{t\in[0,T]}\mathbb{E}\left[|y_t-\bar{y}_t|^2\right]\leq c_1e^{-\frac{1}{8p^3}(\nu_y-1)^2\kappa}+c_2e^{\kappa}N^{-1}+c_3N^{-1}. \tag{3.98}$$

The conclusion follows from (3.91) and (3.92). ■

To the best of our knowledge, Proposition 3.12 is the first result to establish a positive order of strong convergence for the FTE scheme. Next, we perform a numerical analysis of the strong convergence of the scheme and examine the L^1-error

$$e(N)=\mathbb{E}\left[|y_T-\bar{y}_{T,N}|\right]. \tag{3.99}$$

Due to the difficulty in computing this quantity, we estimate as proxy the difference between the values obtained with the FTE scheme for a time step dt ($\bar{y}_{T,2N}$) and the values obtained with the FTE scheme for a time step $2dt$ ($\bar{y}_{T,N}$), and use the fact that, for any $\beta>0$,

$$e(N)=\mathcal{O}\big(N^{-\beta}\big)\;\Leftrightarrow\;\mathbb{E}\left[|\bar{y}_{T,2N}-\bar{y}_{T,N}|\right]=\mathcal{O}\big(N^{-\beta}\big). \tag{3.100}$$

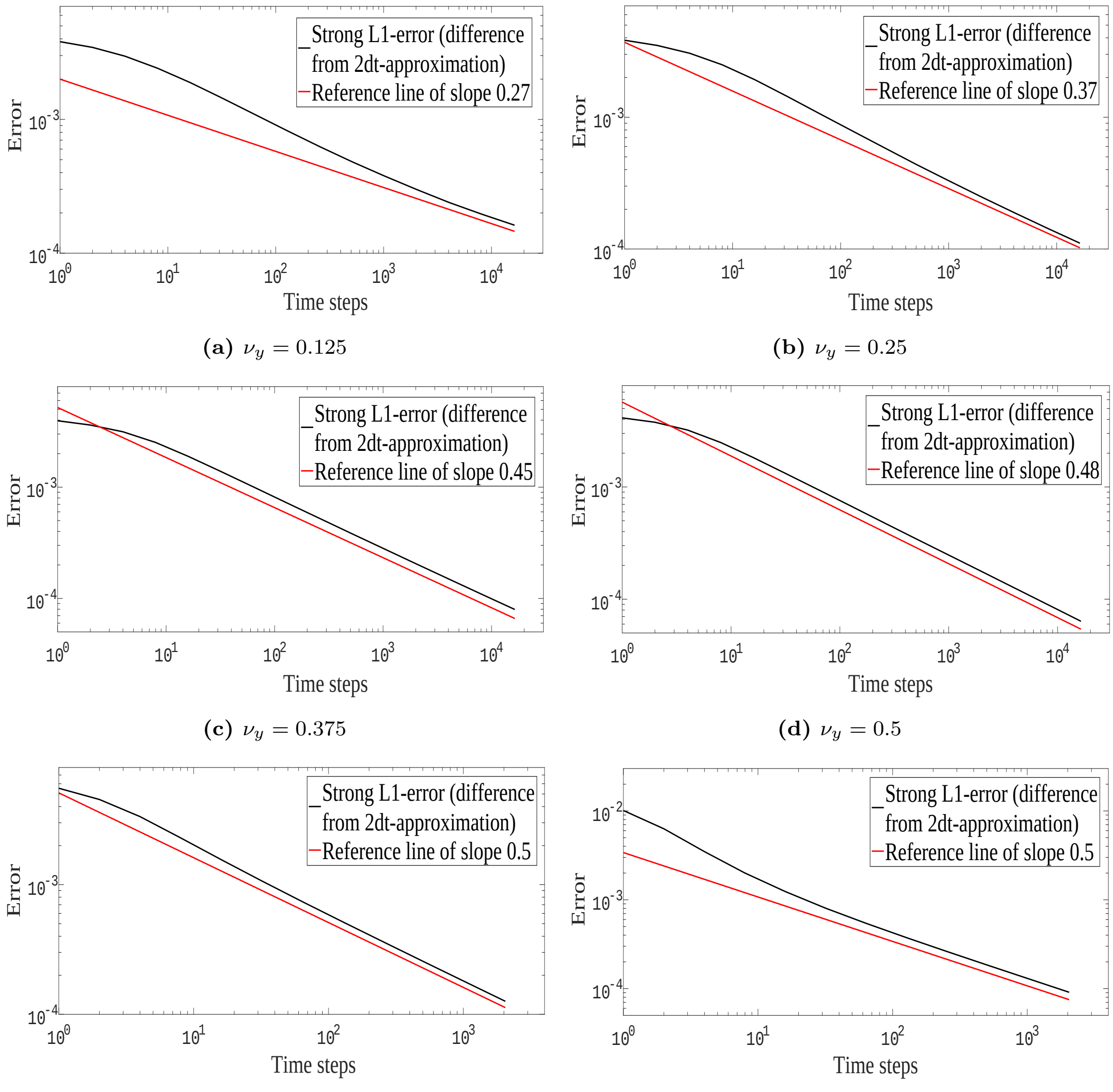
A proof of (3.100) can be found, for instance, in [3]. The data in Figure 3.1 suggest an empirical L^1-order of 1/2 when $\nu_y\geq 1$ and an order between 0 and 1/2 when $\nu_y\in(0,1)$, and hence that the strong convergence rate proved in Proposition 3.12 is not sharp.

3.3.4. The 4-dimensional system

Although weak convergence is important when estimating expectations of payoffs, strong convergence plays a crucial role in multilevel Monte Carlo methods and may be required for complex path-dependent derivatives. First, we prove the uniform mean square convergence of the stopped discretized spot FX rate process in (3.26). Note that the arguments employed in the subsequent convergence analysis do not allow us to establish a positive strong convergence rate for the discretized spot FX rate process.

Proposition 3.13. *Let $L_s>S_0$, $L_v>v_0$, $L_d>g_0^d$, $L_f>g_0^f>\kappa^{-1}$ be constants and define the stopping time*

$$\tau=\inf\big\{t\geq 0:\ S_t\geq L_s\ or\ v_t\geq L_v\ or\ \hat{v}_t\geq L_v\ or\ \hat{g}_t^d\geq L_d\ or\ \hat{g}_t^f\geq L_f\ or\ g_t^f\leq\kappa^{-1}\big\}. \tag{3.101}$$



(a) $\nu_y = 0.125$ **(b)** $\nu_y = 0.25$

(c) $\nu_y = 0.375$ **(d)** $\nu_y = 0.5$

(e) $\nu_y = 1.0$ **(f)** $\nu_y = 2.0$

Figure 3.1: The log-plots of the strong L1-errors against the number of time steps when $y_0 = 0.02$, $k_y \in \{0.125, 0.25, 0.375, 0.5, 1.0, 2.0\}$, $\theta_y = 0.02$, $\xi_y = 0.2$ and $T = 1.0$, computed using 10^7 Monte Carlo simulations.

Under assumptions ($\mathcal{A}3.1$) *to* ($\mathcal{A}3.3$), *the stopped process converges uniformly in* L^2, *i.e.,*

$$\lim_{\delta t \to 0} \mathbb{E}\left[\sup_{t \in [0,T]} \left|S_{t\wedge\tau} - \bar{S}_{t\wedge\tau}\right|^2\right] = 0. \tag{3.102}$$

Proof. The absolute difference between the original and the discretized stopped processes can be bounded from above as follows,

$$\begin{aligned}\left|S_{t\wedge\tau} - \bar{S}_{t\wedge\tau}\right| \leq & \left|\int_0^{t\wedge\tau} h(u)\big(S_u - \bar{S}_u\big)du\right| + \left|\int_0^{t\wedge\tau} \big(g_u^d - \bar{g}_u^d\big)S_u\, du\right| \\ & + \left|\int_0^{t\wedge\tau} \bar{g}_u^d\big(S_u - \bar{S}_u\big)du\right| + \left|\int_0^{t\wedge\tau} \big(g_u^f - \bar{g}_u^f\big)S_u\, du\right| \\ & + \left|\int_0^{t\wedge\tau} \bar{g}_u^f\big(S_u - \bar{S}_u\big)du\right| + \left|\int_0^{t\wedge\tau} \bar{\sigma}\big(u, \bar{S}_u\big)\sqrt{\bar{v}_u}\,\big(S_u - \bar{S}_u\big)dW_u^s\right|\end{aligned}$$

$$+\left|\int_0^{t\wedge\tau}\Big(\sigma\big(u,S_u\big)-\bar{\sigma}\big(u,\bar{S}_u\big)\Big)\sqrt{\bar{v}_u}\,S_u\,dW_u^s\right|$$
$$+\left|\int_0^{t\wedge\tau}\sigma\big(u,S_u\big)\Big(\sqrt{v_u}-\sqrt{\bar{v}_u}\Big)S_u\,dW_u^s\right|. \tag{3.103}$$

Squaring both sides, taking the supremum over all $t\in[0,s]$, where $0\leq s\leq T$, and then employing the Cauchy–Schwarz inequality leads to the upper bound

$$\begin{aligned}\sup_{t\in[0,s]}\big|S_{t\wedge\tau}-\bar{S}_{t\wedge\tau}\big|^2 &\leq 8\big(L_d^2+L_f^2+4h_{max}^2\big)T\int_0^{s\wedge\tau}\big(S_u-\bar{S}_u\big)^2du+8L_s^2T\int_0^T\big(g_u^d-\bar{g}_u^d\big)^2du\\
&+16L_s^2T\int_0^{s\wedge\tau}\big(g_u^f-\hat{g}_u^f\big)^2du+16L_s^2T\int_0^T\big(\hat{g}_u^f-\bar{g}_u^f\big)^2du\\
&+8\sup_{t\in[0,s]}\left|\int_0^{t\wedge\tau}\sigma\big(u,S_u\big)\Big(\sqrt{v_u}-\sqrt{\bar{v}_u}\Big)S_u\,dW_u^s\right|^2\\
&+8\sup_{t\in[0,s]}\left|\int_0^{t\wedge\tau}\bar{\sigma}\big(u,\bar{S}_u\big)\sqrt{\bar{v}_u}\,\big(S_u-\bar{S}_u\big)dW_u^s\right|^2\\
&+8\sup_{t\in[0,s]}\left|\int_0^{t\wedge\tau}\Big(\sigma\big(u,S_u\big)-\bar{\sigma}\big(u,\bar{S}_u\big)\Big)\sqrt{\bar{v}_u}\,S_u\,dW_u^s\right|^2.\end{aligned} \tag{3.104}$$

Note that

$$\int_0^{s\wedge\tau}\big(g_u^f-\hat{g}_u^f\big)^2du\leq\int_0^s\big(g_{u\wedge\tau}^f-\hat{g}_{u\wedge\tau}^f\big)^2du\leq T\sup_{t\in[0,s]}\big|g_{t\wedge\tau}^f-\hat{g}_{t\wedge\tau}^f\big|^2. \tag{3.105}$$

Hence, taking expectations in (3.104), using Fubini's theorem, Doob's martingale inequality and the Itô isometry, and upon noticing that a stopped martingale is also a martingale, we obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,s]}\big|S_{t\wedge\tau}-\bar{S}_{t\wedge\tau}\big|^2\right] &\leq 8L_s^2T^2\sup_{t\in[0,T]}\mathbb{E}\Big[\big|g_t^d-\bar{g}_t^d\big|^2\Big]+32\sigma_{max}^2L_s^2T\sup_{t\in[0,T]}\mathbb{E}\Big[\big|v_t-\bar{v}_t\big|\Big]\\
&+\Big[8\big(L_d^2+L_f^2+4h_{max}^2\big)T+32\sigma_{max}^2L_v\Big]\,\mathbb{E}\left[\int_0^{s\wedge\tau}\big|S_u-\bar{S}_u\big|^2du\right]\\
&+32L_s^2L_v\,\mathbb{E}\left[\int_0^{s\wedge\tau}\big|\sigma\big(u,S_u\big)-\bar{\sigma}\big(u,\bar{S}_u\big)\big|^2du\right]\\
&+16L_s^2T^2\sup_{t\in[0,T]}\mathbb{E}\Big[\big|\hat{g}_t^f-\bar{g}_t^f\big|^2\Big]+16L_s^2T^2\,\mathbb{E}\left[\sup_{t\in[0,s]}\big|g_{t\wedge\tau}^f-\hat{g}_{t\wedge\tau}^f\big|^2\right].\end{aligned} \tag{3.106}$$

First, we bound the last expectation on the right-hand side of (3.106) from above. From (3.59) and (3.64), since $|g_{t\wedge\tau}^f-\hat{g}_{t\wedge\tau}^f|\leq|g_{t\wedge\tau}^f-\tilde{g}_{t\wedge\tau}^f|$, we get

$$\begin{aligned}\big|g_{t\wedge\tau}^f-\hat{g}_{t\wedge\tau}^f\big|\leq\Bigg| &-k_f\int_0^{t\wedge\tau}\big(g_u^f-\hat{g}_u^f\big)du-k_f\int_0^{t\wedge\tau}\big(\hat{g}_u^f-\bar{g}_u^f\big)du\\
&+\xi_f\int_0^{t\wedge\tau}\Big(\sqrt{g_u^f}-\sqrt{\hat{g}_u^f}\Big)dW_u^f+\xi_f\int_0^{t\wedge\tau}\Big(\sqrt{\hat{g}_u^f}-\sqrt{\bar{g}_u^f}\Big)dW_u^f\\
&-\rho_{sf}\xi_f\int_0^{t\wedge\tau}\sigma\big(u,S_u\big)\sqrt{v_u}\Big(\sqrt{g_u^f}-\sqrt{\hat{g}_u^f}\Big)du\end{aligned}$$

$$- \rho_{sf}\xi_f \int_0^{t\wedge\tau} \sigma\big(u, S_u\big)\sqrt{v_u}\Big(\sqrt{\hat{g}_u^f} - \sqrt{\bar{g}_u^f}\Big)du$$
$$- \rho_{sf}\xi_f \int_0^{t\wedge\tau} \sigma\big(u, S_u\big)\sqrt{\bar{g}_u^f}\Big(\sqrt{v_u} - \sqrt{\bar{v}_u}\Big)du$$
$$- \rho_{sf}\xi_f \int_0^{t\wedge\tau} \sqrt{\bar{v}_u \bar{g}_u^f}\Big(\sigma\big(u, S_u\big) - \bar{\sigma}\big(u, \bar{S}_u\big)\Big)du\Bigg|. \tag{3.107}$$

Squaring both sides, taking the supremum over all $t \in [0, s']$, where $0 \leq s' \leq s$, and employing the Cauchy–Schwarz inequality, then taking expectations and using Doob's martingale inequality and Fubini's theorem, we deduce that

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,s']} \big|g_{t\wedge\tau}^f - \hat{g}_{t\wedge\tau}^f\big|^2\right] &\leq 8k_f^2 T \int_0^{s'} \mathbb{E}\left[\big(g_u^f - \hat{g}_u^f\big)^2 \mathbb{1}_{u<\tau}\right]du + 8k_f^2 T \int_0^{T} \mathbb{E}\left[\big(\hat{g}_u^f - \bar{g}_u^f\big)^2\right]du \\
&\quad + 32\xi_f^2 \int_0^T \mathbb{E}\left[\big|\hat{g}_u^f - \bar{g}_u^f\big|\right]du + 8\rho_{sf}^2\xi_f^2\sigma_{max}^2 L_v T \int_0^T \mathbb{E}\left[\big|\hat{g}_u^f - \bar{g}_u^f\big|\right]du \\
&\quad + 32\xi_f^2 \int_0^{s'} \mathbb{E}\left[\Big(\sqrt{g_u^f} - \sqrt{\hat{g}_u^f}\Big)^2 \mathbb{1}_{u<\tau}\right]du \\
&\quad + 8\rho_{sf}^2\xi_f^2\sigma_{max}^2 L_v T \int_0^{s'} \mathbb{E}\left[\Big(\sqrt{g_u^f} - \sqrt{\hat{g}_u^f}\Big)^2 \mathbb{1}_{u<\tau}\right]du \\
&\quad + 8\rho_{sf}^2\xi_f^2\sigma_{max}^2 L_f T \int_0^T \mathbb{E}\left[\big|v_u - \bar{v}_u\big|\right]du \\
&\quad + 8\rho_{sf}^2\xi_f^2 L_v L_f T\, \mathbb{E}\left[\int_0^{s'\wedge\tau} \big|\sigma\big(u, S_u\big) - \bar{\sigma}\big(u, \bar{S}_u\big)\big|^2 du\right].
\end{aligned} \tag{3.108}$$

However,

$$\Big(\sqrt{g_u^f} - \sqrt{\hat{g}_u^f}\Big)^2 \mathbb{1}_{u<\tau} \leq \Big(\sqrt{g_{u\wedge\tau}^f} - \sqrt{\hat{g}_{u\wedge\tau}^f}\Big)^2 \leq \kappa\big(g_{u\wedge\tau}^f - \hat{g}_{u\wedge\tau}^f\big)^2, \tag{3.109}$$

and hence, since $s' \leq s$,

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,s']} \big|g_{t\wedge\tau}^f - \hat{g}_{t\wedge\tau}^f\big|^2\right] &\leq 8\big(k_f^2 T + 4\xi_f^2\kappa + \rho_{sf}^2\xi_f^2\sigma_{max}^2 L_v \kappa T\big) \int_0^{s'} \mathbb{E}\left[\sup_{t\in[0,u]} \big|g_{t\wedge\tau}^f - \hat{g}_{t\wedge\tau}^f\big|^2\right]du \\
&\quad + 8\rho_{sf}^2\xi_f^2 L_v L_f T\, \mathbb{E}\left[\int_0^{s\wedge\tau} \big|\sigma\big(u, S_u\big) - \bar{\sigma}\big(u, \bar{S}_u\big)\big|^2 du\right] \\
&\quad + \big(32\xi_f^2 T + 8\rho_{sf}^2\xi_f^2\sigma_{max}^2 L_v T^2\big) \sup_{t\in[0,T]} \mathbb{E}\left[\big|\hat{g}_t^f - \bar{g}_t^f\big|\right] \\
&\quad + 8\rho_{sf}^2\xi_f^2\sigma_{max}^2 L_f T^2 \sup_{t\in[0,T]} \mathbb{E}\left[\big|v_t - \bar{v}_t\big|\right] \\
&\quad + 8k_f^2 T^2 \sup_{t\in[0,T]} \mathbb{E}\left[\big(\hat{g}_t^f - \bar{g}_t^f\big)^2\right].
\end{aligned} \tag{3.110}$$

Therefore, applying Gronwall's inequality to (3.110), since $s \leq T$, we get

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,s]} \big|g_{t\wedge\tau}^f - \hat{g}_{t\wedge\tau}^f\big|^2\right] &\leq e^{\beta_1 T}\Bigg\{8\rho_{sf}^2\xi_f^2 L_v L_f T\, \mathbb{E}\left[\int_0^{s\wedge\tau} \big|\sigma\big(u, S_u\big) - \bar{\sigma}\big(u, \bar{S}_u\big)\big|^2 du\right] \\
&\quad + 8\rho_{sf}^2\xi_f^2\sigma_{max}^2 L_f T^2 \sup_{t\in[0,T]} \mathbb{E}\Big[\big|v_t - \bar{v}_t\big|\Big] + 8k_f^2 T^2 \sup_{t\in[0,T]} \mathbb{E}\Big[\big(\hat{g}_t^f - \bar{g}_t^f\big)^2\Big]
\end{aligned}$$

$$+\left(32\xi_f^2T+8\rho_{sf}^2\xi_f^2\sigma_{max}^2L_vT^2\right)\sup_{t\in[0,T]}\mathbb{E}\left[\left|\hat{g}_t^f-\bar{g}_t^f\right|\right]\Bigg\},\tag{3.111}$$

where

$$\beta_1=8k_f^2T+32\xi_f^2\kappa+8\rho_{sf}^2\xi_f^2\sigma_{max}^2L_v\kappa T.\tag{3.112}$$

Second, as N is a multiple of N_T, assumption ($\mathcal{A}3.2$) on the leverage function ensures that

$$\mathbb{E}\left[\int_0^{s\wedge\tau}\left|\sigma\big(u,S_u\big)-\bar{\sigma}\big(u,\bar{S}_u\big)\right|^2du\right]\leq 3A^2T(\delta t)^{2\alpha}+3B^2\,\mathbb{E}\left[\int_0^T\left|S_t-S_{\bar{t}}\right|^2\mathbb{1}_{t<\tau}\,dt\right]$$
$$+\,3B^2\,\mathbb{E}\left[\int_0^{s\wedge\tau}\sup_{t\in[0,u]}\left|S_t-\bar{S}_t\right|^2du\right],\tag{3.113}$$

where, for convenience of notation, we defined

$$\bar{t}=\delta t\left\lfloor\frac{t}{\delta t}\right\rfloor.\tag{3.114}$$

Furthermore, we know that

$$\int_0^{s\wedge\tau}\left|S_u-\bar{S}_u\right|^2du\leq\int_0^{s\wedge\tau}\sup_{t\in[0,u]}\left|S_t-\bar{S}_t\right|^2du\leq\int_0^s\sup_{t\in[0,u]}\left|S_{t\wedge\tau}-\bar{S}_{t\wedge\tau}\right|^2du.\tag{3.115}$$

Substituting back into (3.106) with (3.111) – (3.115), we deduce that

$$\mathbb{E}\left[\sup_{t\in[0,s]}\left|S_{t\wedge\tau}-\bar{S}_{t\wedge\tau}\right|^2\right]\leq\beta_2\int_0^s\mathbb{E}\left[\sup_{t\in[0,u]}\left|S_{t\wedge\tau}-\bar{S}_{t\wedge\tau}\right|^2\right]du+\beta_3(\delta t)^{2\alpha}$$
$$+\,\beta_4\sup_{t\in[0,T]}\mathbb{E}\left[\left|v_t-\bar{v}_t\right|\right]+\beta_5\sup_{t\in[0,T]}\mathbb{E}\left[\left|g_t^d-\bar{g}_t^d\right|^2\right]$$
$$+\,\beta_6\sup_{t\in[0,T]}\mathbb{E}\left[\left|\hat{g}_t^f-\bar{g}_t^f\right|\right]+\beta_7\sup_{t\in[0,T]}\mathbb{E}\left[\left|\hat{g}_t^f-\bar{g}_t^f\right|^2\right]$$
$$+\,\beta_8\int_0^T\mathbb{E}\left[\left|S_t-S_{\bar{t}}\right|^2\mathbb{1}_{t<\tau}\right]dt,\tag{3.116}$$

where the constants β_i, for $2\leq i\leq 8$, are defined below.

$$\beta_2=32\sigma_{max}^2L_v+8\big(L_d^2+L_f^2+4h_{max}^2\big)T+96B^2L_s^2L_v\big(1+4\rho_{sf}^2\xi_f^2L_fT^3e^{\beta_1T}\big),$$
$$\beta_3=96A^2L_s^2L_v\big(1+4\rho_{sf}^2\xi_f^2L_fT^3e^{\beta_1T}\big)T,\ \ \beta_4=32\sigma_{max}^2L_s^2T\big(1+4\rho_{sf}^2\xi_f^2L_fT^3e^{\beta_1T}\big),$$
$$\beta_5=8L_s^2T^2,\ \ \beta_6=128\xi_f^2L_s^2\big(4+\rho_{sf}^2\sigma_{max}^2L_vT\big)T^3e^{\beta_1T},\ \ \beta_7=16L_s^2T^2\big(1+8k_f^2T^2e^{\beta_1T}\big),$$
$$\beta_8=96B^2L_s^2L_v\big(1+4\rho_{sf}^2\xi_f^2L_fT^3e^{\beta_1T}\big).\tag{3.117}$$

Therefore, applying Gronwall's inequality to (3.116), we obtain

$$\mathbb{E}\Big[\sup_{t\in[0,T]}\left|S_{t\wedge\tau}-\bar{S}_{t\wedge\tau}\right|^2\Big]\leq e^{\beta_2T}\Bigg\{\beta_3(\delta t)^{2\alpha}+\beta_4\sup_{t\in[0,T]}\mathbb{E}\Big[\left|v_t-\bar{v}_t\right|\Big]+\beta_5\sup_{t\in[0,T]}\mathbb{E}\Big[\left|g_t^d-\bar{g}_t^d\right|^2\Big]$$
$$+\,\beta_6\sup_{t\in[0,T]}\mathbb{E}\Big[\left|\hat{g}_t^f-\bar{g}_t^f\right|\Big]+\beta_7\sup_{t\in[0,T]}\mathbb{E}\Big[\left|\hat{g}_t^f-\bar{g}_t^f\right|^2\Big]$$
$$+\,\beta_8\int_0^T\mathbb{E}\Big[\left|S_t-S_{\bar{t}}\right|^2\mathbb{1}_{t<\tau}\Big]dt\Bigg\}.\tag{3.118}$$

The convergence of the first term on the right-hand side as $\delta t \to 0$ is trivial, whereas the convergence of the next two terms is a consequence of Proposition 3.10. We now show the convergence of the L^2 difference between $\hat{g}^f$ and $\bar{g}^f$.

Suppose that $t \in [t_n, t_{n+1})$. Since $|\hat{g}^f_t - \bar{g}^f_t| \leq |\tilde{g}^f_t - \tilde{g}^f_{t_n}|$ and using that $2\sqrt{|ab|} \leq |a| + |b|$, we can bound the L^2 difference from above as follows:

$$
\begin{aligned}
|\hat{g}^f_t - \bar{g}^f_t|^2 &\leq \frac{1}{4}\Big(2k_f\theta_f\delta t + |\rho_{sf}|\xi_f\sigma_{max}\delta t\,\hat{v}_{t_n} + \big(2k_f + |\rho_{sf}|\xi_f\sigma_{max}\big)\delta t\,\hat{g}^f_{t_n} \\
&\quad + 2\xi_f\sqrt{\hat{g}^f_{t_n}}\,\big|W^f_t - W^f_{t_n}\big|\Big)^2 \\
&\leq 4k_f^2\theta_f^2(\delta t)^2 + \rho_{sf}^2\xi_f^2\sigma_{max}^2(\delta t)^2\hat{v}_{t_n}^2 + \big(2k_f + |\rho_{sf}|\xi_f\sigma_{max}\big)^2(\delta t)^2\big(\hat{g}^f_{t_n}\big)^2 \\
&\quad + 4\xi_f^2\hat{g}^f_{t_n}\big(W^f_t - W^f_{t_n}\big)^2.
\end{aligned}
\tag{3.119}
$$

Hence,

$$
\begin{aligned}
\sup_{t\in[0,T]} \mathbb{E}\Big[\big|\hat{g}^f_t - \bar{g}^f_t\big|^2\Big] &\leq 4k_f^2\theta_f^2(\delta t)^2 + \rho_{sf}^2\xi_f^2\sigma_{max}^2(\delta t)^2 \sup_{0\leq n\leq N} \mathbb{E}\big[\hat{v}_{t_n}^2\big] + 4\xi_f^2\delta t \sup_{0\leq n\leq N} \mathbb{E}\big[\hat{g}^f_{t_n}\big] \\
&\quad + \big(2k_f + |\rho_{sf}|\xi_f\sigma_{max}\big)^2(\delta t)^2 \sup_{0\leq n\leq N} \mathbb{E}\Big[\big(\hat{g}^f_{t_n}\big)^2\Big].
\end{aligned}
\tag{3.120}
$$

The convergence of the term on the left-hand side follows from Propositions 3.8 and 3.9.

Finally, the integrand in the last term in (3.118) can be bounded from above as follows:

$$
\begin{aligned}
\mathbb{E}\Big[\big|S_t - S_{\bar{t}}\big|^2\,\mathbb{1}_{t<\tau}\Big] &= \mathbb{E}\left[\left|\int_{\bar{t}}^t \big[g^d_u - g^f_u + h(u)\big]S_u\,du + \int_{\bar{t}}^t \sigma(u,S_u)\sqrt{v_u}\,S_u\,dW^s_u\right|^2 \mathbb{1}_{t<\tau}\right] \\
&\leq 4\delta t\,\mathbb{E}\left[\int_{\bar{t}}^t \Big[\big(g^d_u\big)^2 + \big(g^f_u\big)^2 + h(u)^2\Big]S_u^2\,\mathbb{1}_{u<\tau}\,du\right] \\
&\quad + 4\,\mathbb{E}\left[\left|\int_{\bar{t}}^t \sigma(u,S_u)\sqrt{v_u}\,S_u\,\mathbb{1}_{u<\tau}\,dW^s_u\right|^2\right] \\
&\leq 4L_s^2(\delta t)^2\left\{\sup_{u\in[0,T]} \mathbb{E}\Big[\big(g^d_u\big)^2\Big] + \sup_{u\in[0,T]} \mathbb{E}\Big[\big(g^f_u\big)^2\Big] + 4h_{max}^2\right\} \\
&\quad + 4\sigma_{max}^2 L_s^2 L_v\delta t.
\end{aligned}
\tag{3.121}
$$

The right-hand side is independent of the time t and tends to zero as $\delta t \to 0$ from Propositions 3.8 and 3.9, which concludes the proof. ∎

Next, we prove the uniform convergence in probability of the discretized spot FX rate process.

Proposition 3.14. *If $2k_f\theta_f > \xi_f^2$ and under assumptions ($\mathcal{A}3.1$) to ($\mathcal{A}3.3$), $\bar{S}$ converges uniformly in probability, i.e.,*

$$
\lim_{\delta t\to 0} \mathbb{P}\left(\sup_{t\in[0,T]} \big|S_t - \bar{S}_t\big| > \epsilon\right) = 0\,, \quad \forall\epsilon > 0.
\tag{3.122}
$$

Proof. Fix $\epsilon > 0$ and note that

$$\left\{ \sup_{t\in[0,T]} \big|S_t - \bar{S}_t\big| > \epsilon \right\} \subseteq \left\{ \sup_{t\in[0,T]} \big|S_{t\wedge\tau} - \bar{S}_{t\wedge\tau}\big| > \epsilon \right\} \cup \big\{\tau < T\big\}, \tag{3.123}$$

with τ from (3.101). In terms of probabilities of events, the previous inclusion implies that

$$\mathbb{P}\left(\sup_{t\in[0,T]} \big|S_t - \bar{S}_t\big| > \epsilon \right) \leq \mathbb{P}\left(\sup_{t\in[0,T]} \big|S_{t\wedge\tau} - \bar{S}_{t\wedge\tau}\big| > \epsilon \right) + \mathbb{P}\left(\tau < T\right). \tag{3.124}$$

The convergence in probability of the stopped process is an immediate consequence of Proposition 3.13 and Markov's inequality. Moreover, from the definition of the stopping time in (3.101), we get

$$\begin{aligned}\big\{\tau < T\big\} \subseteq \left\{ \sup_{t\in[0,T]} v_t \geq L_v \right\} \cup \left\{ \sup_{t\in[0,T]} \hat{v}_t \geq L_v \right\} \cup \left\{ \sup_{t\in[0,T]} \hat{g}^d_t \geq L_d \right\} \cup \left\{ \sup_{t\in[0,T]} \hat{g}^f_t \geq L_f \right\} \\ \cup \left\{ \sup_{t\in[0,T]} S_t \geq L_s \right\} \cup \big\{\tau_\kappa < T\big\},\end{aligned} \tag{3.125}$$

where we defined the stopping time

$$\tau_\kappa = \inf\big\{t \geq 0 : \, g^f_t \leq \kappa^{-1}\big\}. \tag{3.126}$$

However, assuming that $L_s > 1$, we have that

$$\left\{ \sup_{t\in[0,T]} S_t \geq L_s \right\} = \left\{ \sup_{t\in[0,T]} x_t \geq \log L_s \right\} \subseteq \left\{ \sup_{t\in[0,T]} |x_t| \geq \log L_s \right\}. \tag{3.127}$$

Therefore, using (3.125), (3.127) and Markov's inequality, we find an upper bound

$$\begin{aligned}\mathbb{P}\left(\tau < T\right) \leq L_v^{-1}\,\mathbb{E}\left[\sup_{t\in[0,T]} v_t\right] + L_v^{-1}\,\mathbb{E}\left[\sup_{t\in[0,T]} \hat{v}_t\right] + L_d^{-1}\,\mathbb{E}\left[\sup_{t\in[0,T]} \hat{g}^d_t\right] + L_f^{-1}\,\mathbb{E}\left[\sup_{t\in[0,T]} \hat{g}^f_t\right] \\ + \big(\log L_s\big)^{-1}\,\mathbb{E}\left[\sup_{t\in[0,T]} |x_t|\right] + \mathbb{P}\left(\tau_\kappa < T\right).\end{aligned} \tag{3.128}$$

First, recall the formula for the logarithm of the spot FX rate process,

$$x_t = x_0 + \int_0^t \Big(g^d_u - g^f_u + h(u) - \frac{1}{2}\sigma^2\big(u, S_u\big)v_u\Big)du + \int_0^t \sigma\big(u, S_u\big)\sqrt{v_u}\,dW^s_u. \tag{3.129}$$

Taking the supremum over all $t \in [0, T]$ and using the fact that $a < a^2 + 1$, then taking expectations and employing Doob's martingale inequality and the Itô isometry, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]} |x_t|\right] \leq 1 + |x_0| + 2h_{max}T + T\,\mathbb{E}\left[\sup_{t\in[0,T]} g^d_t\right] + T\,\mathbb{E}\left[\sup_{t\in[0,T]} g^f_t\right] \\ + \frac{9}{2}\sigma^2_{max}T\,\mathbb{E}\left[\sup_{t\in[0,T]} v_t\right].\end{aligned} \tag{3.130}$$

Second, we deduce the almost sure positivity of the process g^f when $2k_f\theta_f > \xi^2_f$. To this end, we let $\alpha = (2k_f\theta_f - \xi^2_f)/2\xi^2_f$ and define the function $F : (0,\infty) \to \mathbb{R}$ by $F(x) = x^{-\alpha}$. Using Itô's formula, we have that

$$\mathbb{E}\Big[F\big(g^f_{T\wedge\tau_\kappa}\big)\Big] = F\big(g^f_0\big) - \mathbb{E}\int_0^{T\wedge\tau_\kappa} \alpha\big(g^f_u\big)^{-(1+\alpha)}\big(k_f\theta_f - k_f g^f_u - \rho_{sf}\xi_f\sigma\big(u, S_u\big)\sqrt{v_u g^f_u}\big)du$$

$$+\frac{1}{2}\,\mathbb{E}\int_0^{T\wedge\tau_\kappa}\alpha(1+\alpha)\xi_f^2\big(g_u^f\big)^{-(1+\alpha)}du-\mathbb{E}\int_0^{T\wedge\tau_\kappa}\alpha\xi_f\big(g_u^f\big)^{-(\frac{1}{2}+\alpha)}dW_u^f. \tag{3.131}$$

Note that

$$\alpha k_f\theta_f-\frac{1}{2}\alpha(1+\alpha)\xi_f^2=\frac{1}{4}\alpha\big[4k_f\theta_f-2\xi_f^2-2\alpha\xi_f^2\big]=\frac{1}{2}\alpha^2\xi_f^2, \tag{3.132}$$

and also that

$$\mathbb{E}\int_0^T\alpha^2\xi_f^2\big(g_u^f\big)^{-(1+2\alpha)}\,\mathbb{1}_{u<\tau_\kappa}\,du\le\alpha^2\xi_f^2\kappa^{1+2\alpha}T<\infty, \tag{3.133}$$

so the stochastic integral on the right-hand side of (3.131) is a true martingale. Hence,

$$\begin{aligned}\mathbb{E}\left[F\big(g_{T\wedge\tau_\kappa}^f\big)\right]\le F\big(g_0^f\big)&-\frac{1}{2}\alpha^2\xi_f^2\,\mathbb{E}\int_0^T\big(g_u^f\big)^{-(1+\alpha)}\,\mathbb{1}_{u<\tau_\kappa}\,du+\alpha k_f\,\mathbb{E}\int_0^T\big(g_u^f\big)^{-\alpha}\,\mathbb{1}_{u<\tau_\kappa}\,du\\&+\alpha|\rho_{sf}|\xi_f\sigma_{max}\,\mathbb{E}\int_0^T v_u^{\frac{1}{2}}\big(g_u^f\big)^{-(\frac{1}{2}+\alpha)}\,\mathbb{1}_{u<\tau_\kappa}\,du.\end{aligned} \tag{3.134}$$

Employing Fubini's theorem and Hölder's inequality, we deduce that

$$\begin{aligned}\mathbb{E}\Big[F\big(g_{T\wedge\tau_\kappa}^f\big)\Big]\le&-\frac{1}{2}\alpha^2\xi_f^2\int_0^T\Bigg(\mathbb{E}\Big[\big(g_u^f\big)^{-(1+\alpha)}\mathbb{1}_{u<\tau_\kappa}\Big]-2\alpha^{-1}k_f\xi_f^{-2}\,\mathbb{E}\Big[\big(g_u^f\big)^{-(1+\alpha)}\mathbb{1}_{u<\tau_\kappa}\Big]^{\frac{\alpha}{1+\alpha}}\\&-2\alpha^{-1}|\rho_{sf}|\xi_f^{-1}\sigma_{max}\sup_{t\in[0,T]}\mathbb{E}\Big[v_t^{1+\alpha}\Big]^{\frac{1}{2(1+\alpha)}}\,\mathbb{E}\Big[\big(g_u^f\big)^{-(1+\alpha)}\,\mathbb{1}_{u<\tau_\kappa}\Big]^{\frac{1+2\alpha}{2(1+\alpha)}}\Bigg)du\\&+F\big(g_0^f\big).\end{aligned} \tag{3.135}$$

However, the moments of the square-root process are bounded by Proposition 3.8. Furthermore, if $p,q\in(0,1)$, $c_{p,q}\ge 0$ and $c=c_p+c_q>0$, then for all $x\ge 0$ we have

$$\begin{aligned}x-c_px^p-c_qx^q&=\frac{c_p}{c}\big(x-cx^p\big)+\frac{c_q}{c}\big(x-cx^q\big)\\&=c_pc^{\frac{p}{1-p}}\Big[c^{-\frac{1}{1-p}}x-\Big(c^{-\frac{1}{1-p}}x\Big)^p\Big]+c_qc^{\frac{q}{1-q}}\Big[c^{-\frac{1}{1-q}}x-\Big(c^{-\frac{1}{1-q}}x\Big)^q\Big]\\&\ge-c_pc^{\frac{p}{1-p}}-c_qc^{\frac{q}{1-q}}.\end{aligned} \tag{3.136}$$

Therefore, the integrand in (3.135) is bounded from below by a constant and thus

$$\mathbb{E}\left[F\big(g_{T\wedge\tau_\kappa}^f\big)\right]\le C, \tag{3.137}$$

for some constant C independent of κ. Since g^f has continuous paths, $g_{\tau_\kappa}^f=\kappa^{-1}$ and $F(g_{\tau_\kappa}^f)=\kappa^\alpha$. Hence, from (3.137) and the positivity of F, we deduce that

$$\mathbb{P}\left(\tau_\kappa<T\right)=\kappa^{-\alpha}\,\mathbb{E}\left[F\big(g_{\tau_\kappa}^f\big)\,\mathbb{1}_{\tau_\kappa<T}\right]\le\kappa^{-\alpha}\,\mathbb{E}\left[F\big(g_{T\wedge\tau_\kappa}^f\big)\right]\le C\kappa^{-\alpha}. \tag{3.138}$$

Taking the limit as $\delta t\to 0$ in (3.124), using the upper bounds derived in (3.128), (3.130) and (3.138), and then employing Propositions 3.8 and 3.9, the conclusion follows from the fact that we can choose L_s, L_v, L_d, L_f and κ arbitrarily large. ∎

Many models with stochastic volatility dynamics, including the Heston model, have the undesirable feature of moment instability, i.e., moments of order higher than 1 can explode in finite time [7]. This can cause problems in practice, for instance when computing the arbitrage-free price of an option with a superlinear payoff. Furthermore, establishing the existence of moments of order higher than 1 of the process and its approximation is an important ingredient in the convergence analysis (see [48]). Since the most popular FX contracts grow at most linearly in the FX rate and their risk-neutral valuation involves computing the expected discounted payoff, it is useful to study the finiteness of moments under discounting. Define R to be the discounted exchange rate process,

$$R_t = S_0 \exp\left\{ -\int_0^t r_u^f \, du - \frac{1}{2}\int_0^t \sigma^2(u, S_u) v_u \, du + \int_0^t \sigma(u, S_u)\sqrt{v_u}\, dW_u^s \right\}, \tag{3.139}$$

and let $\bar{R}$ be its continuous-time approximation,

$$\bar{R}_t = S_0 \exp\left\{ -\int_0^t \bar{r}_u^f \, du - \frac{1}{2}\int_0^t \bar{\sigma}^2\big(u, \bar{S}_u\big) \bar{v}_u \, du + \int_0^t \bar{\sigma}\big(u, \bar{S}_u\big)\sqrt{\bar{v}_u}\, dW_u^s \right\}. \tag{3.140}$$

The following result establishes a lower bound on the explosion time of moments of order higher than 1 of the discounted process and is an important step in proving Theorem 3.1 (Theorem 1 in [20]), which, in turn, plays a key role in the calibration of the model.

Proposition 3.15. *Let $\alpha \geq 1$. Under assumptions ($\mathcal{A}3.1$) and ($\mathcal{A}3.3$), if $T < T^*$, then there exists $w_1 > \alpha$ such that, for all $w \in [1, w_1)$, the following holds:*

$$\sup_{t\in[0,T]} \mathbb{E}\left[R_t^w\right] < \infty, \tag{3.141}$$

where T^* is as given below.

(1) When $k < \varphi(\alpha)\zeta$,

$$T^* = \frac{2}{\sqrt{\varphi(\alpha)^2\zeta^2 - k^2}}\left[\frac{\pi}{2} + \arctan\left(\frac{k}{\sqrt{\varphi(\alpha)^2\zeta^2 - k^2}}\right)\right], \tag{3.142}$$

where

$$\varphi(\alpha) = \alpha + \sqrt{(\alpha-1)\alpha} \quad \text{and} \quad \zeta = \xi\sigma_{max}. \tag{3.143}$$

(2) When $k \geq \varphi(\alpha)\zeta$,

$$T^* = \infty. \tag{3.144}$$

Proof. First, we find it convenient to define a new stochastic process L by

$$L_t \equiv S_0 \exp\left\{ h_{max} t - \frac{1}{2}\int_0^t \sigma^2(u, S_u) v_u \, du + \int_0^t \sigma(u, S_u)\sqrt{v_u}\, dW_u^s \right\}. \tag{3.145}$$

As $R_t \leq L_t$ for all $t \in [0, T]$, it suffices to prove the finiteness of the supremum over t of

$$\mathbb{E}\big[L_t^w\big] = S_0^w\, \mathbb{E}\bigg[\exp\left\{ w h_{max} t + w\int_0^t \sigma(u, S_u)\sqrt{v_u}\, dW_u^s - \frac{w}{2}\int_0^t \sigma^2(u, S_u) v_u \, du \right\}\bigg]. \tag{3.146}$$

Second, suppose that $T < T^*$, with T^* defined in (3.142) – (3.144). If $k < \varphi(\alpha)\zeta$, since $\varphi(\cdot)$ is increasing on $[1, \infty)$ and by a continuity argument, we can find $w_1 > \alpha$ such that, for all $w \in (\alpha, w_1)$,

$$k < \varphi(w)\zeta \quad \text{and} \quad T < \frac{2}{\sqrt{\varphi(w)^2\zeta^2 - k^2}}\left[\frac{\pi}{2} + \arctan\left(\frac{k}{\sqrt{\varphi(w)^2\zeta^2 - k^2}}\right)\right]. \tag{3.147}$$

If $k = \varphi(\alpha)\zeta$, since $\varphi(\cdot)$ is strictly increasing on $[1, \infty)$ and

$$\lim_{w \downarrow \alpha^+} \frac{2}{\sqrt{\varphi(w)^2\zeta^2 - k^2}}\left[\frac{\pi}{2} + \arctan\left(\frac{k}{\sqrt{\varphi(w)^2\zeta^2 - k^2}}\right)\right] = \infty, \tag{3.148}$$

we can find $w_1 > \alpha$ such that (3.147) holds for all $w \in (\alpha, w_1)$. Finally, if $k > \varphi(\alpha)\zeta$, by a continuity argument, we can find $w_1 > \alpha$ such that, for all $w \in (\alpha, w_1)$,

$$k > \varphi(w)\zeta. \tag{3.149}$$

Third, fix $w \in (\alpha, w_1)$ and let the functions $f_w, g_w : (1, \infty) \to (1, \infty)$ be defined by

$$f_w(x) = w^2x^2 \tag{3.150}$$

and

$$g_w(x) = \frac{wx}{x-1}(wx - 1) = \left(w + \sqrt{(w-1)w}\right)^2 + \frac{w^2}{x-1}\left(x - 1 - \sqrt{1 - 1/w}\right)^2. \tag{3.151}$$

The first function is strictly increasing, whereas the second function attains a global minimum at $x^* = 1 + \sqrt{1 - 1/w}$. Furthermore, they are equal at x^*, so

$$\inf_{x>1} \max\left\{f_w(x), g_w(x)\right\} = g_w(x^*) = \left(w + \sqrt{(w-1)w}\right)^2. \tag{3.152}$$

Consider the Hölder pair (p, q) satisfying $q = p/(p-1)$ such that

$$p = 1 + \sqrt{\frac{w-1}{w}} \quad \text{and} \quad q = 1 + \sqrt{\frac{w}{w-1}}. \tag{3.153}$$

This particular Hölder pair ensures an optimal lower bound T^* on the explosion time. Next, define the quantity $a = pw^2 - w$ and introduce the stochastic process

$$M_t = pw\int_0^t \sigma(u, S_u)\sqrt{v_u}\, dW_u^s, \tag{3.154}$$

with quadratic variation

$$\langle M\rangle_t = p^2w^2\int_0^t \sigma^2(u, S_u)v_u\, du. \tag{3.155}$$

Then we can rewrite (3.146) as follows:

$$\mathbb{E}\left[L_t^w\right] = S_0^w e^{wh_{max}t}\,\mathbb{E}\left[\exp\left\{\frac{1}{p}\left[M_t - \frac{1}{2}\langle M\rangle_t\right] + \frac{a}{2}\int_0^t \sigma^2(u, S_u)v_u\, du\right\}\right]. \tag{3.156}$$

Applying Hölder's inequality with the pair (p,q) from (3.153) and taking the supremum over $[0,T]$, we get

$$\sup_{t\in[0,T]} \mathbb{E}\left[L_t^w\right] \le S_0^w e^{wh_{max}T} \sup_{t\in[0,T]} \mathbb{E}\left[\exp\left\{M_t - \frac{1}{2}\langle M\rangle_t\right\}\right]^{\frac{1}{p}}$$
$$\times \mathbb{E}\left[\exp\left\{\frac{1}{2} qw(pw-1)\sigma_{max}^2 \int_0^T v_u\, du\right\}\right]^{\frac{1}{q}}. \tag{3.157}$$

The stochastic exponential is a martingale if Novikov's condition (see [62]) is satisfied, i.e.,

$$\mathbb{E}\left[\exp\left\{\frac{1}{2}\langle M\rangle_T\right\}\right] \le \mathbb{E}\left[\exp\left\{\frac{1}{2}p^2w^2\sigma_{max}^2 \int_0^T v_u\, du\right\}\right] < \infty. \tag{3.158}$$

The finiteness of the two expectations in (3.157) follows from (3.147) – (3.153) and Proposition 3.6. The extension to the interval $[1, w_1)$ follows from Jensen's inequality. ■

Proposition 3.16. *Let $\alpha \ge 1$. Under assumptions ($\mathcal{A}3.1$) and ($\mathcal{A}3.3$), if $T < T^*$, then there exists $w_2 > \alpha$ such that, for all $w \in [1, w_2)$ and $\delta_T < k^{-1}$, the following holds:*

$$\sup_{\delta t\in(0,\delta_T)} \sup_{t\in[0,T]} \mathbb{E}\left[(\bar{R}_t)^w\right] < \infty, \tag{3.159}$$

where T^* is as given below.

(1) When $k \le \frac{1}{2}\varphi(\alpha)\zeta$,

$$T^* = \frac{1}{\varphi(\alpha)\zeta - k}, \tag{3.160}$$

where

$$\varphi(\alpha) = \alpha + \sqrt{(\alpha-1)\alpha} \quad \textit{and} \quad \zeta = \xi\sigma_{max}. \tag{3.161}$$

(2) When $k > \frac{1}{2}\varphi(\alpha)\zeta$,

$$T^* = \frac{4k}{\varphi(\alpha)^2\zeta^2}. \tag{3.162}$$

Proof. For convenience, define a new stochastic process $\bar{L}$ by

$$\bar{L}_t \equiv S_0 \exp\left\{h_{max}t - \frac{1}{2}\int_0^t \bar{\sigma}^2\big(u,\bar{S}_u\big)\bar{v}_u\, du + \int_0^t \bar{\sigma}\big(u,\bar{S}_u\big)\sqrt{\bar{v}_u}\, dW_u^s\right\}. \tag{3.163}$$

Since $\bar{R}_t \le \bar{L}_t$ for all $t \in [0,T]$, it suffices to prove the finiteness of the supremum over t and δt of

$$\mathbb{E}\big[\bar{L}_t^w\big] = S_0^w\, \mathbb{E}\left[\exp\left\{wh_{max}t + w\int_0^t \bar{\sigma}\big(u,\bar{S}_u\big)\sqrt{\bar{v}_u}\, dW_u^s - \frac{w}{2}\int_0^t \bar{\sigma}^2\big(u,\bar{S}_u\big)\bar{v}_u\, du\right\}\right]. \tag{3.164}$$

Suppose that $T < T^*$, with T^* defined in (3.160) – (3.162). If $k \le \frac{1}{2}\varphi(\alpha)\zeta$, by a continuity argument, we can find $w_2 > \alpha$ such that, for all $w \in (\alpha, w_2)$,

$$k < \frac{1}{2}\varphi(w)\zeta \quad \text{and} \quad T < \frac{1}{\varphi(w)\zeta - k}. \tag{3.165}$$

On the other hand, if $k > \frac{1}{2}\varphi(\alpha)\zeta$, by a continuity argument, we can find $w_2 > \alpha$ such that, for all $w \in (\alpha, w_2)$,

$$k > \frac{1}{2}\varphi(w)\zeta \quad \text{and} \quad T < \frac{4k}{\varphi(w)^2\zeta^2}. \tag{3.166}$$

Henceforth, we argue as in Proposition 3.15. We fix $w \in (\alpha, w_2)$ and consider the Hölder pair (p, q) from (3.153). Then, we define $a = pw^2 - w$ and introduce the stochastic process

$$\bar{M}_t = pw \int_0^t \bar{\sigma}\big(u, \bar{S}_u\big)\sqrt{\bar{v}_u}\, dW_u^s, \tag{3.167}$$

with quadratic variation

$$\langle \bar{M} \rangle_t = p^2 w^2 \int_0^t \bar{\sigma}^2\big(u, \bar{S}_u\big)\bar{v}_u\, du. \tag{3.168}$$

We can rewrite (3.164) as follows:

$$\mathbb{E}\left[\bar{L}_t^w\right] = S_0^w e^{wh_{max}t}\, \mathbb{E}\left[\exp\left\{\frac{1}{p}\left[\bar{M}_t - \frac{1}{2}\langle \bar{M} \rangle_t\right] + \frac{a}{2}\int_0^t \bar{\sigma}^2\big(u, \bar{S}_u\big)\bar{v}_u\, du\right\}\right]. \tag{3.169}$$

Applying Hölder's inequality and taking the supremum in (3.169), we get

$$\begin{aligned}\sup_{\delta t \in (0, \delta_T)} \sup_{t \in [0,T]} \mathbb{E}\left[\bar{L}_t^w\right] \leq S_0^w e^{wh_{max}T} \sup_{\delta t \in (0, \delta_T)} \sup_{t \in [0,T]} \mathbb{E}\left[\exp\left\{\bar{M}_t - \frac{1}{2}\langle \bar{M} \rangle_t\right\}\right]^{\frac{1}{p}} \\ \times \sup_{\delta t \in (0, \delta_T)} \mathbb{E}\left[\exp\left\{\frac{1}{2} qw\big(pw - 1\big)\sigma_{max}^2 \int_0^T \bar{v}_u\, du\right\}\right]^{\frac{1}{q}}.\end{aligned} \tag{3.170}$$

The stochastic exponential above is a martingale if Novikov's condition is satisfied, i.e.,

$$\mathbb{E}\left[\exp\left\{\frac{1}{2}\langle \bar{M} \rangle_T\right\}\right] \leq \mathbb{E}\left[\exp\left\{\frac{1}{2} p^2 w^2 \sigma_{max}^2 \int_0^T \bar{v}_u\, du\right\}\right] < \infty. \tag{3.171}$$

The conclusion follows from (3.150) – (3.153), (3.165) – (3.166) and Proposition 3.7, and the extension to the interval $[1, w_2)$ follows from Jensen's inequality. ■

Higham et al. [48] proved that for a locally Lipschitz SDE, the boundedness of the pth moments of the exact and numerical solutions, for some $p > 2$, ensures the strong mean square convergence of the Euler–Maruyama method. The existence of moment bounds for explicit Euler approximations, however, remained an open problem. Recently, Hutzenthaler et al. [55] studied SDEs with superlinearly growing coefficients and proved strong and weak divergence in L^p for all $p \geq 1$, and hence that the moment-bound assumption is not satisfied. To the best of our knowledge, the uniform boundedness of moments of order higher than 1 of discretization schemes for the Heston model and extensions thereof has not yet been established – this gap in the literature was also identified in Kloeden and Neuenkirch [63] – and Proposition 3.16 is the first result to address this issue.

Since the typical payoff of a FX contract grows at most linearly in the exchange rate, it suffices to know the strong convergence of the discounted process in L^1 to deduce the convergence of the time-discretization error to zero.

Theorem 3.17. *Under assumptions ($\mathcal{A}$3.1) to ($\mathcal{A}$3.3), if $2k_f\theta_f > \xi_f^2$ and $T < T^*$, where*

$$T^* \equiv \frac{4k}{\zeta^2}\,\mathbb{1}_{\zeta<2k} + \frac{1}{\zeta-k}\,\mathbb{1}_{\zeta\geq 2k} \tag{3.172}$$

and $\zeta = \xi\sigma_{max}$, then the discounted process converges strongly in L^1, i.e.,

$$\lim_{\delta t\to 0}\sup_{t\in[0,T]}\mathbb{E}\left[\left|R_t-\bar{R}_t\right|\right] = 0. \tag{3.173}$$

Proof. Fix $\epsilon > 0$ and define the event $A = \left\{\left|R_t-\bar{R}_t\right| > \epsilon\right\}$, then

$$\sup_{t\in[0,T]}\mathbb{E}\left[\left|R_t-\bar{R}_t\right|\right] \leq \sup_{t\in[0,T]}\mathbb{E}\left[\left|R_t-\bar{R}_t\right|\mathbb{1}_{A^c}\right] + \sup_{t\in[0,T]}\mathbb{E}\left[\left|R_t-\bar{R}_t\right|\mathbb{1}_{A}\right]. \tag{3.174}$$

Hence,

$$\sup_{t\in[0,T]}\mathbb{E}\left[\left|R_t-\bar{R}_t\right|\right] \leq \epsilon + \sup_{t\in[0,T]}\mathbb{E}\left[R_t\,\mathbb{1}_A\right] + \sup_{t\in[0,T]}\mathbb{E}\left[\bar{R}_t\,\mathbb{1}_A\right]. \tag{3.175}$$

Choosing some $1 < w < \min\{w_1, w_2\}$ and applying Hölder's inequality to the two expectations on the right-hand side with the pair $(p,q) = \left(w, \frac{w}{w-1}\right)$ returns the following upper bound:

$$\begin{aligned}\sup_{t\in[0,T]}\mathbb{E}\left[\left|R_t-\bar{R}_t\right|\right] &\leq \epsilon + \left\{\sup_{t\in[0,T]}\mathbb{E}\left[R_t^w\right]^{\frac{1}{w}} + \sup_{t\in[0,T]}\mathbb{E}\left[(\bar{R}_t)^w\right]^{\frac{1}{w}}\right\} \\ &\quad\times \sup_{t\in[0,T]}\mathbb{P}\left(\left|R_t-\bar{R}_t\right| > \epsilon\right)^{1-\frac{1}{w}}.\end{aligned} \tag{3.176}$$

Note that if $\zeta \geq 2k$, then

$$\frac{2}{\sqrt{\zeta^2-k^2}}\left[\frac{\pi}{2}+\arctan\left(\frac{k}{\sqrt{\zeta^2-k^2}}\right)\right] > \frac{\sqrt{3}}{\sqrt{\zeta^2-k^2}} \geq \frac{1}{\sqrt{\zeta^2-k^2}}\sqrt{\frac{\zeta+k}{\zeta-k}} = \frac{1}{\zeta-k}. \tag{3.177}$$

Moreover, if $k < \zeta < 2k$, then

$$\frac{2}{\sqrt{\zeta^2-k^2}}\left[\frac{\pi}{2}+\arctan\left(\frac{k}{\sqrt{\zeta^2-k^2}}\right)\right] > \frac{2}{\sqrt{\zeta^2-k^2}} \geq \frac{4}{\sqrt{\zeta^2-k^2}}\sqrt{\frac{k^2}{\zeta^2}\left(1-\frac{k^2}{\zeta^2}\right)} = \frac{4k}{\zeta^2}. \tag{3.178}$$

Therefore, Propositions 3.15 and 3.16 (with $\alpha = 1$) ensure the boundedness of moments of order w of the discounted process and its approximation. Furthermore, the convergence in probability of the discounted process is a simple consequence of Proposition 3.14, by taking the domestic short rate to be zero. Finally, taking ϵ sufficiently small leads to the conclusion. ■

Note that Theorem 3.17 can be generalized to the L^α case relatively easily for all $\alpha \geq 1$, upon noticing that the critical time T^* from (3.142) – (3.144) is always greater than the one from (3.160) – (3.162), as shown in (3.177) – (3.178) for $\alpha = 1$.

3.3.5. Option valuation

Next, we examine the convergence of Monte Carlo estimators for computing FX option prices when the dynamics of the exchange rate are governed by the Heston–2CIR++ SLV model and assumptions ($\mathcal{A}3.1$) to ($\mathcal{A}3.3$) are satisfied. As an aside, note that we discussed in Section 3.2 how other derivative pricing models, including popular models in equity markets, can be formulated as special cases. First, we consider European options.

Theorem 3.18. *Let $P = \mathbb{E}\left[e^{-\int_0^T r_t^d dt}\big(K - S_T\big)^+\right]$ be the arbitrage-free price of a European put option and $\bar{P} = \mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt}\big(K - \bar{S}_T\big)^+\right]$ its approximation. If $2k_f\theta_f > \xi_f^2$, then*

$$\lim_{\delta t \to 0} \big|P - \bar{P}\big| = 0. \tag{3.179}$$

Proof. A simple string of inequalities gives the following upper bound:

$$\begin{aligned}\big|P - \bar{P}\big| &\leq \mathbb{E}\Big[\big|\big\{e^{-\int_0^T r_t^d dt} - e^{-\int_0^T \bar{r}_t^d dt}\big\}\big(K - S_T\big)^+ + e^{-\int_0^T \bar{r}_t^d dt}\big\{\big(K - S_T\big)^+ - \big(K - \bar{S}_T\big)^+\big\}\big|\Big] \\ &\leq K e^{h_{max}T}\, \mathbb{E}\Big[\big|e^{-\int_0^T g_t^d dt} - e^{-\int_0^T \bar{g}_t^d dt}\big|\Big] + e^{h_{max}T}\, \mathbb{E}\Big[\big|\big(K - S_T\big)^+ - \big(K - \bar{S}_T\big)^+\big|\Big].\end{aligned} \tag{3.180}$$

However, for any non-negative numbers x and y, $|e^{-x} - e^{-y}| \leq |x - y|$, and so we can use Fubini's theorem to obtain an upper bound for the first expectation,

$$\begin{aligned}\sup_{t \in [0,T]} \mathbb{E}\left[\big|e^{-\int_t^T g_u^d du} - e^{-\int_t^T \bar{g}_u^d du}\big|\right] &\leq \sup_{t \in [0,T]} \int_t^T \mathbb{E}\left[|g_u^d - \bar{g}_u^d|\right] du \\ &\leq T \sup_{t \in [0,T]} \mathbb{E}\left[|g_t^d - \bar{g}_t^d|\right].\end{aligned} \tag{3.181}$$

The right-hand side tends to zero by Proposition 3.10. Define the events $A = \big\{S_T < K\big\}$ and $\bar{A} = \big\{\bar{S}_T < K\big\}$, and denote the last expectation in (3.180) by $\mathbb{J}$. Then

$$\mathbb{J} = \mathbb{E}\left[\big|\big(K - S_T\big)^+ - \big(K - \bar{S}_T\big)^+\big|\big(\mathbb{1}_{A\cap\bar{A}} + \mathbb{1}_{A\cap\bar{A}^c} + \mathbb{1}_{A^c\cap\bar{A}} + \mathbb{1}_{A^c\cap\bar{A}^c}\big)\right]. \tag{3.182}$$

Therefore,

$$\begin{aligned}\mathbb{J} &\leq \mathbb{E}\left[\big|S_T - \bar{S}_T\big|\,\mathbb{1}_{A\cap\bar{A}}\right] + \mathbb{E}\left[\big(K - S_T\big)\,\mathbb{1}_{A\cap\bar{A}^c}\right] + \mathbb{E}\left[\big(K - \bar{S}_T\big)\,\mathbb{1}_{A^c\cap\bar{A}}\right] \\ &\leq \mathbb{E}\left[\big|S_T - \bar{S}_T\big|\,\mathbb{1}_{A\cap\bar{A}}\right] + K\,\mathbb{P}\left(A \cap \bar{A}^c\right) + K\,\mathbb{P}\left(A^c \cap \bar{A}\right).\end{aligned} \tag{3.183}$$

Let δ be an arbitrary positive number, then we have the following inclusion of events:

$$\begin{aligned}A \cap \bar{A}^c &= \Big(\big\{S_T \leq K - \delta\big\} \cup \big\{K - \delta < S_T < K\big\}\Big) \cap \big\{\bar{S}_T \geq K\big\} \\ &\subseteq \Big(\big\{S_T \leq K - \delta\big\} \cap \big\{\bar{S}_T \geq K\big\}\Big) \cup \big\{K - \delta < S_T < K\big\} \\ &\subseteq \Big\{\big|S_T - \bar{S}_T\big| \geq \delta\Big\} \cup \big\{K - \delta < S_T < K\big\}.\end{aligned} \tag{3.184}$$

In terms of probabilities of events, we have

$$\mathbb{P}\left(A \cap \bar{A}^c\right) \leq \mathbb{P}\left(\left|S_T - \bar{S}_T\right| \geq \delta\right) + \mathbb{P}\left(K - \delta < S_T < K\right), \quad \forall \delta > 0. \tag{3.185}$$

We can bound the second probability from above in a similar fashion,

$$A^c \cap \bar{A} \subseteq \left\{\left|S_T - \bar{S}_T\right| \geq \delta\right\} \cup \left\{K \leq S_T < K + \delta\right\}, \tag{3.186}$$

so

$$\mathbb{P}\left(A^c \cap \bar{A}\right) \leq \mathbb{P}\left(\left|S_T - \bar{S}_T\right| \geq \delta\right) + \mathbb{P}\left(K \leq S_T < K + \delta\right), \quad \forall \delta > 0. \tag{3.187}$$

For a suitable choice of δ, the last terms on the right-hand side of (3.185) and (3.187) can be made arbitrarily small, whereas the first terms tend to zero by Proposition 3.14. Therefore, the two probabilities in (3.183) converge to zero as $\delta t \to 0$. Finally, fix $\epsilon > 0$ and let $B = \left\{|S_T - \bar{S}_T| > \epsilon\right\}$. We can bound the expectation on the right-hand side of (3.183) as follows:

$$\begin{aligned}\mathbb{E}\left[\left|S_T - \bar{S}_T\right| \mathbb{1}_{A\cap\bar{A}}\right] &= \mathbb{E}\left[\left|S_T - \bar{S}_T\right| \mathbb{1}_{A\cap\bar{A}}\, \mathbb{1}_{B^c}\right] + \mathbb{E}\left[\left|S_T - \bar{S}_T\right| \mathbb{1}_{A\cap\bar{A}}\, \mathbb{1}_{B}\right] \\ &\leq K\, \mathbb{P}\left(\left|S_T - \bar{S}_T\right| > \epsilon\right) + \epsilon. \end{aligned} \tag{3.188}$$

Taking the limit as $\delta t \to 0$, employing Proposition 3.14 and making use of the fact that ϵ can be made arbitrarily small leads to the conclusion. ■

Theorem 3.19. *Let* $C = \mathbb{E}\left[e^{-\int_0^T r_t^d dt}\left(S_T - K\right)^+\right]$ *be the arbitrage-free price of a European call and* $\bar{C} = \mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt}\left(\bar{S}_T - K\right)^+\right]$ *its approximation. If* $2k_f\theta_f > \xi_f^2$ *and* $T < T^*$*, with* T^* *from* (3.172)*, then*

$$\lim_{\delta t \to 0} \left|C - \bar{C}\right| = 0. \tag{3.189}$$

Proof. A simple string of inequalities gives the following upper bound:

$$\begin{aligned}\left|C - \bar{C}\right| &\leq \mathbb{E}\left[\left|\left(R_T - Ke^{-\int_0^T r_t^d dt}\right)^+ - \left(\bar{R}_T - Ke^{-\int_0^T \bar{r}_t^d dt}\right)^+\right|\right] \\ &\leq Ke^{h_{max}T}\, \mathbb{E}\left[\left|e^{-\int_0^T g_t^d dt} - e^{-\int_0^T \bar{g}_t^d dt}\right|\right] + \mathbb{E}\left[\left|R_T - \bar{R}_T\right|\right]. \end{aligned} \tag{3.190}$$

The two expectations on the right-hand side tend to zero as $\delta t \to 0$ from (3.181) and Theorem 3.17, respectively. ■

Asian options depend on the average exchange rate over a predetermined time period. Because the average is less volatile than the underlying rate, Asian options are usually less expensive than their European counterparts and are commonly used in FX and commodity markets, for instance, to reduce the foreign currency exposure of a corporation expecting payments in foreign currency. For any $0 \leq s \leq t \leq T$, define the discount factors:

$$D_{s,t} = e^{-\int_s^t r_u^d du} \quad \text{and} \quad \bar{D}_{s,t} = e^{-\int_s^t \bar{r}_u^d du}. \tag{3.191}$$

Theorem 3.20. *Consider a fixed strike Asian option with arbitrage-free price*

$$U = \mathbb{E}\left[e^{-\int_0^T r_t^d dt}\left[\psi(A(0,T)-K)\right]^+\right], \tag{3.192}$$

$$\bar{U} = \mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt}\left[\psi(\bar{A}(0,T)-K)\right]^+\right]. \tag{3.193}$$

If $2k_f\theta_f > \xi_f^2$ and $T < T^*$, with T^* from (3.172)*, then*

$$\lim_{\delta t\to 0}\left|U-\bar{U}\right| = 0. \tag{3.194}$$

Here, $A(0,T)$ represents the arithmetic average and $\psi = \pm 1$ depending on the payoff (call or put). For continuous monitoring, $A(0,T) = \frac{1}{T}\int_0^T S_t\,dt$ and $\bar{A}(0,T) = \frac{1}{T}\int_0^T \hat{S}_t\,dt = \frac{1}{T}\int_0^T \bar{S}_{\bar{t}}\,dt$.

Proof. The absolute difference can be bounded from above by

$$\left|U-\bar{U}\right| \le \mathbb{E}\left[\left|\left[\psi(D_{0,T}A(0,T)-KD_{0,T})\right]^+ - \left[\psi(\bar{D}_{0,T}\bar{A}(0,T)-K\bar{D}_{0,T})\right]^+\right|\right]. \tag{3.195}$$

Therefore, we end up with the following upper bound:

$$\left|U-\bar{U}\right| \le Ke^{h_{max}T}\,\mathbb{E}\left[\left|e^{-\int_0^T g_t^d dt} - e^{-\int_0^T \bar{g}_t^d dt}\right|\right] + \mathbb{E}\left[\left|D_{0,T}A(0,T) - \bar{D}_{0,T}\bar{A}(0,T)\right|\right]. \tag{3.196}$$

We deduced the convergence of the first expectation in (3.181). Fubini's theorem yields

$$\begin{aligned}\mathbb{E}\left[\left|D_{0,T}A(0,T)-\bar{D}_{0,T}\bar{A}(0,T)\right|\right] &\le \frac{1}{T}\mathbb{E}\left[\int_0^T \left|D_{0,T}S_t - \bar{D}_{0,T}\bar{S}_{\bar{t}}\right| dt\right]\\ &\le \sup_{t\in[0,T]}\mathbb{E}\left[\left|D_{t,T}R_t - \bar{D}_{\bar{t},T}\bar{R}_{\bar{t}}\right|\right].\end{aligned} \tag{3.197}$$

The triangle inequality leads to the following upper bound:

$$\begin{aligned}\sup_{t\in[0,T]}\mathbb{E}\left[\left|D_{t,T}R_t - \bar{D}_{\bar{t},T}\bar{R}_{\bar{t}}\right|\right] &\le \sup_{t\in[0,T]}\mathbb{E}\left[R_t\left|D_{t,T}-D_{\bar{t},T}\right|\right] + \sup_{t\in[0,T]}\mathbb{E}\left[D_{\bar{t},T}\left|R_t - R_{\bar{t}}\right|\right]\\ &+ \sup_{t\in[0,T]}\mathbb{E}\left[R_{\bar{t}}\left|D_{\bar{t},T}-\bar{D}_{\bar{t},T}\right|\right] + \sup_{t\in[0,T]}\mathbb{E}\left[\bar{D}_{\bar{t},T}\left|R_{\bar{t}}-\bar{R}_{\bar{t}}\right|\right],\end{aligned} \tag{3.198}$$

and hence

$$\begin{aligned}e^{-h_{max}T}\sup_{t\in[0,T]}\mathbb{E}\left[\left|D_{t,T}R_t - \bar{D}_{\bar{t},T}\bar{R}_{\bar{t}}\right|\right] &\le \sup_{t\in[0,T]}\mathbb{E}\left[R_t\left|1-D_{\bar{t},t}\right|\right] + \sup_{t\in[0,T]}\mathbb{E}\left[\left|R_t - R_{\bar{t}}\right|\right]\\ &+ \sup_{t\in[0,T]}\mathbb{E}\left[R_t\left|e^{-\int_t^T g_u^d du} - e^{-\int_t^T \bar{g}_u^d du}\right|\right]\\ &+ \sup_{t\in[0,T]}\mathbb{E}\left[\left|R_t - \bar{R}_t\right|\right].\end{aligned} \tag{3.199}$$

First, Hölder's inequality with the pair (w,γ), where $1 < w < w_1$ and $\gamma = w/(w-1)$, yields

$$\sup_{t\in[0,T]}\mathbb{E}\left[R_t\left|1-D_{\bar{t},t}\right|\right] \le \sup_{t\in[0,T]}\mathbb{E}\left[R_t^w\right]^{\frac{1}{w}}\sup_{t\in[0,T]}\mathbb{E}\left[\left|1-D_{\bar{t},t}\right|^\gamma\right]^{\frac{1}{\gamma}}. \tag{3.200}$$

However, using the mean value (MV) and Fubini's theorems and Hölder's inequality, we get

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left|1-D_{\bar{t},t}\right|^{\gamma}\right]^{\frac{1}{\gamma}} \leq e^{h_{max}T} \sup_{t\in[0,T]} \mathbb{E}\left[|r_t^d|^{\gamma}\right]^{\frac{1}{\gamma}}\delta t, \tag{3.201}$$

and hence the convergence of the first term on the right-hand side of (3.199) is a consequence of Propositions 3.8 and 3.15 (with $\alpha = 1$). Second, using the MV theorem and Hölder's inequality, we deduce that

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\left[\left|R_t - R_{\bar{t}}\right|\right] &\leq \sup_{t\in[0,T]} \mathbb{E}\left[(R_t \vee R_{\bar{t}})\left|\log(R_t)-\log(R_{\bar{t}})\right|\right] \\ &\leq 2\sup_{t\in[0,T]} \mathbb{E}\left[R_t^w\right]^{\frac{1}{w}} \sup_{t\in[0,T]} \mathbb{E}\left[\left|\log(R_t)-\log(R_{\bar{t}})\right|^{\gamma}\right]^{\frac{1}{\gamma}}, \end{aligned} \tag{3.202}$$

and hence the convergence of the second term on the right-hand side of (3.199) is a consequence of Propositions 3.8, 3.9 and 3.15 and Theorem 1 in [36]. Third, since both g^d and $\bar{g}^d$ are non-negative processes, for any $\gamma > 1$ we have

$$\left|e^{-\int_t^T g_u^d du} - e^{-\int_t^T \bar{g}_u^d du}\right|^{\gamma} \leq \left|e^{-\int_t^T g_u^d du} - e^{-\int_t^T \bar{g}_u^d du}\right|, \quad \forall t\in[0,T]. \tag{3.203}$$

Applying Hölder's inequality and using (3.203) yields

$$\sup_{t\in[0,T]} \mathbb{E}\left[R_t\left|e^{-\int_t^T g_u^d du} - e^{-\int_t^T \bar{g}_u^d du}\right|\right] \leq \sup_{t\in[0,T]} \mathbb{E}\left[R_t^w\right]^{\frac{1}{w}} \sup_{t\in[0,T]} \mathbb{E}\left[\left|e^{-\int_t^T g_u^d du} - e^{-\int_t^T \bar{g}_u^d du}\right|\right]^{\frac{1}{\gamma}}, \tag{3.204}$$

and hence the convergence of the third term on the right-hand side of (3.199) is a consequence of Proposition 3.15 and (3.181). Finally, the convergence of the fourth term on the right-hand side of (3.199) is due to Theorem 3.17. In case of discrete monitoring or a floating strike, we follow the same argument. ∎

Barrier options continue to gain popularity in many over-the-counter markets, including the FX market. Their popularity can be explained by two key factors. First, barrier options are useful in limiting the risk exposure of an investor in the FX market. Second, they offer additional flexibility and can match an investor's view on the market for a lower price than a vanilla option.

Theorem 3.21. *Consider an up-and-out barrier call with arbitrage-free price*

$$U = \mathbb{E}\left[e^{-\int_0^T r_t^d dt}\left(S_T - K\right)^{+} \mathbb{1}_{\left\{\sup_{t\in[0,T]} S_t \leq B\right\}}\right], \tag{3.205}$$

$$\bar{U} = \mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt}\left(\bar{S}_T - K\right)^{+} \mathbb{1}_{\left\{\sup_{t\in[0,T]} \hat{S}_t \leq B\right\}}\right], \tag{3.206}$$

where K is the strike price and B is the barrier. If $2k_f\theta_f > \xi_f^2$, then

$$\lim_{\delta t\to 0}\left|U-\bar{U}\right| = 0. \tag{3.207}$$

Proof. Define the events $A = \left\{\sup_{t\in[0,T]} S_t \le B\right\}$ and $\bar{A} = \left\{\sup_{t\in[0,T]} \hat{S}_t \le B\right\}$, then

$$\begin{aligned}
\left|U - \bar{U}\right| &\le \mathbb{E}\left[\left|\left(D_{0,T} - \bar{D}_{0,T}\right)\left(S_T - K\right)^+ \mathbb{1}_A + \bar{D}_{0,T}\left\{\left(S_T - K\right)^+ \mathbb{1}_A - \left(\bar{S}_T - K\right)^+ \mathbb{1}_{\bar{A}}\right\}\right|\right] \\
&\le \left(B - K\right)^+ e^{h_{max}T}\, \mathbb{E}\left[\left|e^{-\int_0^T g_t^d dt} - e^{-\int_0^T \bar{g}_t^d dt}\right|\right] \\
&\quad + e^{h_{max}T}\, \mathbb{E}\left[\left|\left(S_T - K\right)^+ \mathbb{1}_A - \left(\bar{S}_T - K\right)^+ \mathbb{1}_{\bar{A}}\right|\right].
\end{aligned} \tag{3.208}$$

The first term tends to zero by (3.181) and we can rewrite the second term as follows:

$$\begin{aligned}
&\mathbb{E}\left[\left|\left(S_T - K\right)^+\left(\mathbb{1}_{A\cap\bar{A}^c} + \mathbb{1}_{A\cap\bar{A}}\right) - \left(\bar{S}_T - K\right)^+\left(\mathbb{1}_{A\cap\bar{A}} + \mathbb{1}_{A^c\cap\bar{A}}\right)\right|\right] \\
&\qquad \le \mathbb{E}\left[\left(S_T - K\right)^+ \mathbb{1}_{A\cap\bar{A}^c}\right] + \mathbb{E}\left[\left(\bar{S}_T - K\right)^+ \mathbb{1}_{A^c\cap\bar{A}}\right] + \mathbb{E}\left[\left|S_T - \bar{S}_T\right| \mathbb{1}_{A\cap\bar{A}}\right] \\
&\qquad \le \left(B - K\right)^+\left\{\mathbb{P}\left(A\cap\bar{A}^c\right) + \mathbb{P}\left(A^c\cap\bar{A}\right)\right\} + \mathbb{E}\left[\left|S_T - \bar{S}_T\right| \mathbb{1}_{A\cap\bar{A}}\right].
\end{aligned} \tag{3.209}$$

We can bound the last expectation from above just as in (3.188) to find

$$\mathbb{E}\left[\left|S_T - \bar{S}_T\right| \mathbb{1}_{A\cap\bar{A}}\right] \le B\,\mathbb{P}\left(\left|S_T - \bar{S}_T\right| > \epsilon\right) + \epsilon, \quad \forall \epsilon > 0. \tag{3.210}$$

Therefore, the expectation converges to zero with the time step by Proposition 3.14. Fix $\delta > 0$. First, note that

$$\begin{aligned}
A\cap\bar{A}^c &\subseteq \left\{\sup_{t\in[0,T]} S_t \le B\right\} \cap \left\{\sup_{t\in[0,T]} \bar{S}_t > B\right\} \\
&\subseteq \left(\left\{\sup_{t\in[0,T]} S_t \le B - \delta\right\} \cap \left\{\sup_{t\in[0,T]} \bar{S}_t > B\right\}\right) \cup \left\{B - \delta < \sup_{t\in[0,T]} S_t \le B\right\} \\
&\subseteq \left\{\sup_{t\in[0,T]} \left|S_t - \bar{S}_t\right| \ge \delta\right\} \cup \left\{B - \delta < \sup_{t\in[0,T]} S_t \le B\right\}.
\end{aligned} \tag{3.211}$$

In terms of probabilities of events, this implies that

$$\mathbb{P}\left(A\cap\bar{A}^c\right) \le \mathbb{P}\left(\sup_{t\in[0,T]} \left|S_t - \bar{S}_t\right| \ge \delta\right) + \mathbb{P}\left(B - \delta < \sup_{t\in[0,T]} S_t \le B\right). \tag{3.212}$$

Second, note that

$$\begin{aligned}
A^c\cap\bar{A} &= \left\{\sup_{t\in[0,T]} S_t > B\right\} \cap \left\{\sup_{t\in[0,T]} \hat{S}_t \le B\right\} \\
&\subseteq \left(\left\{\sup_{t\in[0,T]} S_t > B + 2\delta\right\} \cap \left\{\sup_{t\in[0,T]} \hat{S}_t \le B\right\}\right) \cup \left\{B < \sup_{t\in[0,T]} S_t \le B + 2\delta\right\} \\
&\subseteq \left\{\sup_{t\in[0,T]} \left|S_t - \hat{S}_t\right| > 2\delta\right\} \cup \left\{B < \sup_{t\in[0,T]} S_t \le B + 2\delta\right\} \\
&\subseteq \left\{\sup_{t\in[0,T]} \left|S_t - S_{\bar{t}}\right| > \delta\right\} \cup \left\{\sup_{t\in[0,T]} \left|S_{\bar{t}} - \bar{S}_{\bar{t}}\right| > \delta\right\} \cup \left\{B < \sup_{t\in[0,T]} S_t \le B + 2\delta\right\}.
\end{aligned} \tag{3.213}$$

In terms of probabilities of events, this implies that

$$\mathbb{P}\left(A^c\cap\bar{A}\right) \le \mathbb{P}\left(\sup_{t\in[0,T]} \left|S_t - S_{\bar{t}}\right| > \delta\right) + \mathbb{P}\left(\sup_{t\in[0,T]} \left|S_t - \bar{S}_t\right| \ge \delta\right)$$

$$+\,\mathbb{P}\left(B < \sup_{t\in[0,T]} S_t \le B + 2\delta\right). \tag{3.214}$$

Next, fix $L_s > 1 \vee S_0$ and define the stopping time

$$\tau_s = \inf\big\{t \ge 0 :\ S_t \ge L_s\big\}. \tag{3.215}$$

Using the MV theorem and Markov's inequality, we deduce that

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[0,T]} \big|S_t - S_{\bar{t}}\big| > \delta\right) &\le \mathbb{P}\left(\sup_{t\in[0,T\wedge\tau_s]} \big|S_t - S_{\bar{t}}\big| > \delta\right) + \mathbb{P}\left(\tau_s < T\right)\\ &\le \mathbb{P}\left(\sup_{t\in[0,T]} |x_t - x_{\bar{t}}| > \frac{\delta}{L_s}\right) + \mathbb{P}\left(\sup_{t\in[0,T]} |x_t| \ge \log(L_s)\right)\\ &\le \frac{L_s}{\delta}\,\mathbb{E}\left[\sup_{t\in[0,T]} |x_t - x_{\bar{t}}|\right] + \frac{1}{\log(L_s)}\,\mathbb{E}\left[\sup_{t\in[0,T]} |x_t|\right]. \end{aligned} \tag{3.216}$$

However, the convergence to zero of the first term on the right-hand side of (3.216) is a consequence of Propositions 3.8 and 3.9 and Theorem 1 in [36]. The conclusion follows from Proposition 3.14 and (3.130) since δ can be made arbitrarily small and L_s can be made aribtrarily large independently of one another. ■

Theorem 3.22. *Consider any type of barrier put option with arbitrage-free price*

$$U = \mathbb{E}\left[e^{-\int_0^T r_t^d dt}\big(K - S_T\big)^+\,\mathbb{1}_A\right], \tag{3.217}$$

$$\bar{U} = \mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt}\big(K - \bar{S}_T\big)^+\,\mathbb{1}_{\bar{A}}\right], \tag{3.218}$$

where the events A and $\bar{A}$ depend on the type of barrier:

$$\begin{cases} \textit{up-and-in:} & A = \left\{\sup\limits_{t\in[0,T]} S_t > B\right\},\ \ \bar{A} = \left\{\sup\limits_{t\in[0,T]} \hat{S}_t > B\right\}\\ \textit{up-and-out:} & A = \left\{\sup\limits_{t\in[0,T]} S_t \le B\right\},\ \ \bar{A} = \left\{\sup\limits_{t\in[0,T]} \hat{S}_t \le B\right\}\\ \textit{down-and-in:} & A = \left\{\inf\limits_{t\in[0,T]} S_t < B\right\},\ \ \bar{A} = \left\{\inf\limits_{t\in[0,T]} \hat{S}_t < B\right\}\\ \textit{down-and-out:} & A = \left\{\inf\limits_{t\in[0,T]} S_t \ge B\right\},\ \ \bar{A} = \left\{\inf\limits_{t\in[0,T]} \hat{S}_t \ge B\right\}. \end{cases} \tag{3.219}$$

If $2k_f\theta_f > \xi_f^2$, then

$$\lim_{\delta t\to 0} \big|U - \bar{U}\big| = 0. \tag{3.220}$$

Proof. An upper bound for the absolute difference can be obtained as follows:

$$\begin{aligned}\big|U - \bar{U}\big| &\le \mathbb{E}\left[\big|\big(D_{0,T} - \bar{D}_{0,T}\big)\big(K - S_T\big)^+\,\mathbb{1}_A + \bar{D}_{0,T}\big\{\big(K - S_T\big)^+\,\mathbb{1}_A - \big(K - \bar{S}_T\big)^+\,\mathbb{1}_{\bar{A}}\big\}\big|\right]\\ &\le K e^{h_{max}T}\,\mathbb{E}\left[\left|e^{-\int_0^T g_t^d dt} - e^{-\int_0^T \bar{g}_t^d dt}\right|\right]\\ &+ e^{h_{max}T}\,\mathbb{E}\left[\left|\big(K - S_T\big)^+\,\mathbb{1}_A - \big(K - \bar{S}_T\big)^+\,\mathbb{1}_{\bar{A}}\right|\right]. \end{aligned} \tag{3.221}$$

The first term tends to zero by (3.181) and we can bound the second term as in (3.209):

$$\mathbb{E}\left[\left|\left(K-S_T\right)^+\mathbb{1}_A-\left(K-\bar{S}_T\right)^+\mathbb{1}_{\bar{A}}\right|\right]\leq K\Big\{\mathbb{P}\left(A\cap\bar{A}^c\right)+\mathbb{P}\left(A^c\cap\bar{A}\right)\Big\}$$
$$+\mathbb{E}\left[\left|\left(K-S_T\right)^+-\left(K-\bar{S}_T\right)^+\right|\right]. \qquad (3.222)$$

The events A and $\bar{A}$ differ with the barrier (down-and-in, down-and-out, up-and-in, up-and-out), however one can show in a similar way to (3.212) and (3.214) that

$$\lim_{\delta t\to 0}\mathbb{P}\left(A\cap\bar{A}^c\right)=0 \text{ and } \lim_{\delta t\to 0}\mathbb{P}\left(A^c\cap\bar{A}\right)=0 \qquad (3.223)$$

for any type of barrier. Finally, the convergence of the last term on the right-hand side of (3.222) was derived in Theorem 3.18, which concludes the proof. ■

Theorem 3.23. *Consider a down-and-in/out or an up-and-in barrier call option with arbitrage-free price*

$$U=\mathbb{E}\left[e^{-\int_0^T r_t^d dt}\left(S_T-K\right)^+\mathbb{1}_A\right], \qquad (3.224)$$
$$\bar{U}=\mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt}\left(\bar{S}_T-K\right)^+\mathbb{1}_{\bar{A}}\right], \qquad (3.225)$$

where the events A and $\bar{A}$ depend on the type of barrier as in (3.219). *If $2k_f\theta_f>\xi_f^2$ and $T<T^*$, with T^* from* (3.172), *then*

$$\lim_{\delta t\to 0}\left|U-\bar{U}\right|=0. \qquad (3.226)$$

Proof. An upper bound for the absolute difference can be obtained as follows:

$$\left|U-\bar{U}\right|\leq\mathbb{E}\left[\left|\left(R_T-KD_{0,T}\right)^+-\left(\bar{R}_T-K\bar{D}_{0,T}\right)^+\right|\mathbb{1}_{A\cap\bar{A}}\right]$$
$$+\mathbb{E}\left[\left(R_T-KD_{0,T}\right)^+\mathbb{1}_{A\cap\bar{A}^c}\right]+\mathbb{E}\left[\left(\bar{R}_T-K\bar{D}_{0,T}\right)^+\mathbb{1}_{A^c\cap\bar{A}}\right]. \qquad (3.227)$$

Therefore, we end up with

$$\left|U-\bar{U}\right|\leq\mathbb{E}\left[\left|R_T-\bar{R}_T\right|\right]+Ke^{h_{max}T}\,\mathbb{E}\left[\left|e^{-\int_0^T g_t^d dt}-e^{-\int_0^T \bar{g}_t^d dt}\right|\right]$$
$$+\mathbb{E}\left[R_T\,\mathbb{1}_{A\cap\bar{A}^c}\right]+\mathbb{E}\left[\bar{R}_T\,\mathbb{1}_{A^c\cap\bar{A}}\right]. \qquad (3.228)$$

The convergence of the first two terms on the right-hand side is a consequence of Theorem 3.17 and (3.181), respectively. Applying Hölder's inequality with the pair (w,γ) to the last two terms, where $1<w<\min\{w_1,w_2\}$, we find that:

$$\mathbb{E}\left[R_T\,\mathbb{1}_{A\cap\bar{A}^c}\right]\leq\mathbb{E}\left[\left(R_T\right)^w\right]^{\frac{1}{w}}\mathbb{P}\left(A\cap\bar{A}^c\right)^{\frac{1}{\gamma}} \qquad (3.229)$$

and

$$\mathbb{E}\left[\bar{R}_T\,\mathbb{1}_{A^c\cap\bar{A}}\right]\leq\mathbb{E}\left[\left(\bar{R}_T\right)^w\right]^{\frac{1}{w}}\mathbb{P}\left(A^c\cap\bar{A}\right)^{\frac{1}{\gamma}}. \qquad (3.230)$$

Employing Propositions 3.15 and 3.16 and the limits in (3.223) concludes the proof. ■

Among the most developed exotic derivatives in the FX market are the double barrier options.

Theorem 3.24. *Consider a double knock-out call option with arbitrage-free price*

$$U = \mathbb{E}\left[e^{-\int_0^T r_t^d dt}\left(S_T - K\right)^+ \mathbb{1}_{\left\{\inf_{t\in[0,T]} S_t \geq L,\, \sup_{t\in[0,T]} S_t \leq B\right\}}\right], \tag{3.231}$$

$$\bar{U} = \mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt}\left(\bar{S}_T - K\right)^+ \mathbb{1}_{\left\{\inf_{t\in[0,T]} \hat{S}_t \geq L,\, \sup_{t\in[0,T]} \hat{S}_t \leq B\right\}}\right], \tag{3.232}$$

where K is the strike and L, B are the lower and upper barriers, respectively. If $2k_f\theta_f > \xi_f^2$, then

$$\lim_{\delta t\to 0} \left|U - \bar{U}\right| = 0. \tag{3.233}$$

Proof. Note that we have the following inclusion of events:

$$\begin{aligned}&\left\{\inf_{t\in[0,T]} S_t \geq L,\ \sup_{t\in[0,T]} S_t \leq B\right\} \cap \left\{\inf_{t\in[0,T]} \hat{S}_t \geq L,\ \sup_{t\in[0,T]} \hat{S}_t \leq B\right\}^c \\ &\quad\subseteq \left(\left\{\sup_{t\in[0,T]} S_t \leq B\right\} \cap \left\{\sup_{t\in[0,T]} \hat{S}_t > B\right\}\right) \cup \left(\left\{\inf_{t\in[0,T]} S_t \geq L\right\} \cap \left\{\inf_{t\in[0,T]} \hat{S}_t < L\right\}\right).\end{aligned} \tag{3.234}$$

The rest of the argument follows closely that of Theorem 3.21 and is thus omitted. ■

3.4. Model calibration

In this section, we assume a simple correlation structure and calibrate the Heston–2CIR++ SLV model to EURUSD market data. This is a mandatory step for the efficient pricing of FX options.

3.4.1. The shifted square-root process

The domestic and foreign short rates are modeled by the shifted CIR (CIR++) process [12]. On the one hand, this model preserves the analytical tractability of the square-root (CIR) model for bonds, caps and other basic interest rate products. On the other hand, it is flexible enough to fit the initial term structure of interest rates exactly. For $i \in \{d, f\}$, the short rate dynamics under their respective spot measures, i.e., $\mathbb{Q}^d$ – domestic and $\mathbb{Q}^f$ – foreign, are given by

$$\begin{cases} r_t^i = g_t^i + h_i(t), \\ dg_t^i = k_i(\theta_i - g_t^i)dt + \xi_i\sqrt{g_t^i}\, dB_t^i, \quad g_0^i > 0, \end{cases} \tag{3.235}$$

where B^i is a Brownian motion under $\mathbb{Q}^i$. The mean-reversion parameters k_i, the long-term mean parameters θ_i and the volatility parameters ξ_i are the same as in (3.1). The calibration of the short rate model (3.235) follows the same approach for both the domestic and the foreign interest rate. For simplicity, we drop the subscripts and superscripts "d"

and "f" in the remainder of the subsection and define the parameter vector $\alpha = (g_0, k, \theta, \xi)$. According to Brigo and Mercurio [12], an exact fit to the initial term structure of interest rates is equivalent to $h(t) = \varphi^{\mathrm{CIR}}(t; \alpha)$ for all $t \in [0, T]$, where

$$\varphi^{\mathrm{CIR}}(t;\alpha) = f^{\mathrm{M}}(0,t) - f^{\mathrm{CIR}}(0,t;\alpha), \tag{3.236}$$

$$f^{\mathrm{CIR}}(0,t;\alpha) = \frac{2k\theta(\exp\{t\vartheta\}-1)}{2\vartheta + (k+\vartheta)(\exp\{t\vartheta\}-1)} + g_0 \frac{4\vartheta^2 \exp\{t\vartheta\}}{[2\vartheta + (k+\vartheta)(\exp\{t\vartheta\}-1)]^2}, \tag{3.237}$$

$\vartheta = \sqrt{k^2 + 2\xi^2}$ and $f^{\mathrm{M}}(0,t)$ is the market instantaneous forward rate at time 0 for a maturity t, i.e.,

$$f^{\mathrm{M}}(0,t) = -\frac{\partial}{\partial t} \log P^{\mathrm{M}}(0,t), \tag{3.238}$$

where $P^{\mathrm{M}}(0,t)$ is the market zero-coupon bond price at time 0 for a maturity t. The value of the zero-coupon bond is given by

$$P^{\mathrm{M}}(0,t) = \frac{1}{1 + \Delta(0,t) R^{\mathrm{M}}(0,t)}, \tag{3.239}$$

where $\Delta(0,t)$ is the year fraction from 0 to time t and $R^{\mathrm{M}}(0,t)$ is the current (simply-compounded) deposit rate with maturity date t which is quoted in the market (see Table 3.3). As an aside, note that the standard day count convention for US dollars (USD) and euros (EUR) is Actual 360.

Table 3.3: Deposit rates from the term structure of the EURUSD forward market from March 18, 2016.

Maturity	**USD Deposit Rate**	**EUR Deposit Rate**
2W	0.306%	−0.798%
3W	0.327%	−0.766%
1M	0.328%	−0.737%
2M	0.357%	−0.693%
3M	0.371%	−0.687%
4M	0.389%	−0.711%
6M	0.424%	−0.731%
9M	0.468%	−0.776%
1Y	0.511%	−0.810%
18M	0.575%	−0.859%
2Y	0.638%	−0.885%
3Y	0.757%	−0.915%
4Y	0.876%	−0.894%
5Y	0.983%	−0.852%
7Y	1.169%	−0.720%
10Y	1.448%	−0.381%

In order to estimate the zero-coupon curve (also known as the term structure of interest rates or the yield curve), we assume that the instantaneous forward rate is piecewise-flat.

Consider the time nodes $t_0 = 0, t_1, \ldots, t_n$ and the set of estimated instantaneous forward rates $f_1, f_2, \ldots, f_n$ from which the curve is constructed, and define

$$f^{\mathrm{M}}(0,t) = f_i \quad \text{if} \quad t_{i-1} \leq t < t_i, \quad \text{for} \quad i = 1, 2, \ldots, n. \tag{3.240}$$

Using (3.238) – (3.240) and solving the resulting linear system of equations, we get

$$f_i = \frac{1}{\Delta(t_{i-1}, t_i)} \log\left(\frac{1 + \Delta(0,t_i) R^{\mathrm{M}}(0,t_i)}{1 + \Delta(0,t_{i-1}) R^{\mathrm{M}}(0,t_{i-1})}\right) \quad \text{for} \quad i = 1, 2, \ldots, n. \tag{3.241}$$

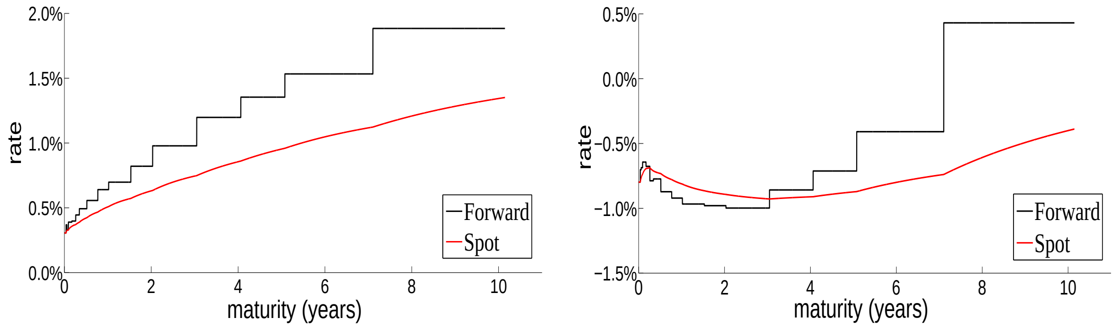
The continuously-compounded spot rate, i.e., the constant rate at which the value of a pure discount bond must grow to yield one unit of currency at maturity, is defined as

$$R_0^{\mathrm{M}}(0,t) = \frac{1}{\Delta(0,t)} \int_0^t f^{\mathrm{M}}(0,s)\, ds. \tag{3.242}$$

Using (3.240) – (3.242), we deduce that

$$\begin{aligned} R_0^{\mathrm{M}}(0,t) = {} & \frac{\Delta(t,t_i)}{\Delta(0,t)\Delta(t_{i-1},t_i)} \log\Big(1 + \Delta(0,t_{i-1}) R^{\mathrm{M}}(0,t_{i-1})\Big) \\ & + \frac{\Delta(t_{i-1},t)}{\Delta(0,t)\Delta(t_{i-1},t_i)} \log\Big(1 + \Delta(0,t_i) R^{\mathrm{M}}(0,t_i)\Big) \end{aligned} \tag{3.243}$$

whenever $t_{i-1} \leq t < t_i$. In Figure 3.2, we plot the USD and EUR zero-coupon curves $t \mapsto R_0^{\mathrm{M}}(0,t)$, $t > 0$, estimated from the quoted deposit rates from March 18, 2016, together with the flat-forward instantaneous forward rates.



(a) USD Currency

(b) EUR Currency

Figure 3.2: The instantaneous forward rates and continuously-compounded spot rates.

A choice of the shift function h as in (3.236) results in an exact fit to the initial term structure of interest rates independent of the value of the parameter vector α.

Next, we determine α by calibrating the CIR++ model to the current term structure of volatilities, in particular, by fitting at-the-money (ATM) cap volatilities. The market quotes used in the calibration are reported in Table 3.4. We consider caps with integer maturities ranging from 1 to 10 years for both currencies, with an additional 18 month cap for EUR. For USD, all caps have quarterly frequency, whereas for EUR the 1 year and 18 month caps have quarterly frequency and the 2 to 10 year caps have semi-annual frequency.

Table 3.4: The ATM USD and EUR cap volatility quotes from March 18, 2016.

Maturity	USD Black Volatility (%)	EUR Normal Volatility (bps)
1Y	46.50	19.24
18M	–	21.95
2Y	59.40	25.33
3Y	65.62	33.03
4Y	67.17	41.35
5Y	65.91	48.85
6Y	63.56	54.72
7Y	60.89	59.34
8Y	58.53	63.12
9Y	56.25	65.86
10Y	54.13	67.90

A cap is a set of spanning caplets with a common strike so the value of the cap is simply the sum of the values of its caplets. It is market standard to price caplets with the Black formula, in which case the fair value of the cap at time 0 with rate (strike) K, reset times $T_a, T_{a+1}, \ldots, T_{b-1}$ and payment times $T_{a+1}, \ldots, T_{b-1}, T_b$ is:

$$\mathrm{Cap}_{\mathrm{Black}}(K, \sigma_{a,b}) = \sum_{i=a+1}^{b} P^{\mathrm{M}}(0, T_i)\Delta(T_{i-1}, T_i)\mathrm{Black}\big(K, F^{\mathrm{M}}(0, T_{i-1}, T_i), \sigma_{a,b}\sqrt{T_{i-1}}\big), \tag{3.244}$$

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

$$F^{\mathrm{M}}(0, T, S) = \frac{1}{\Delta(T, S)}\left(\frac{P^{\mathrm{M}}(0, T)}{P^{\mathrm{M}}(0, S)} - 1\right) \tag{3.245}$$

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

$$\mathrm{Black}(K, F, v) = F\Phi(d_1) - K\Phi(d_2), \tag{3.246}$$

$$d_{1,2} = \frac{\log(F/K) \pm v^2/2}{v}. \tag{3.247}$$

However, Black's formula cannot cope with negative forward rates F or strikes K, in which case we switch to Bachelier's (normal) formula in (3.244):

$$\mathrm{Normal}(K, F, v) = (F - K)\Phi(d) + v\phi(d), \tag{3.248}$$

$$d = \frac{F - K}{v}. \tag{3.249}$$

The data in Figure 3.2b suggest that the instantaneous forward rate for EUR takes negative values. Therefore, we use Black cap volatility quotes for USD and Normal cap volatility quotes for EUR. The market prices of at-the-money caps are computed by inserting the forward swap rate

$$S_{a,b} = \frac{P^{\mathrm{M}}(0,T_a) - P^{\mathrm{M}}(0,T_b)}{\sum_{i=a+1}^{b} \Delta(T_{i-1},T_i) P^{\mathrm{M}}(0,T_i)} \tag{3.250}$$

as strike and the quoted cap volatility as $\sigma_{a,b}$ in (3.244), using either Black's or Bachelier's formula.

Fitting the CIR++ model to cap volatilities means finding the value of α for which the model cap prices, which are available in closed-form [13], best match the market cap prices. The calibration is performed by minimizing the sum of the squared differences between model- and market-implied cap volatilities:

$$\min_{\alpha \in \mathbb{R}_+^4} \sum_{1 \le i \le n} \left[\sigma^{\mathrm{CIR}}(T_i;\alpha) - \sigma^{\mathrm{M}}(T_i)\right]^2, \tag{3.251}$$

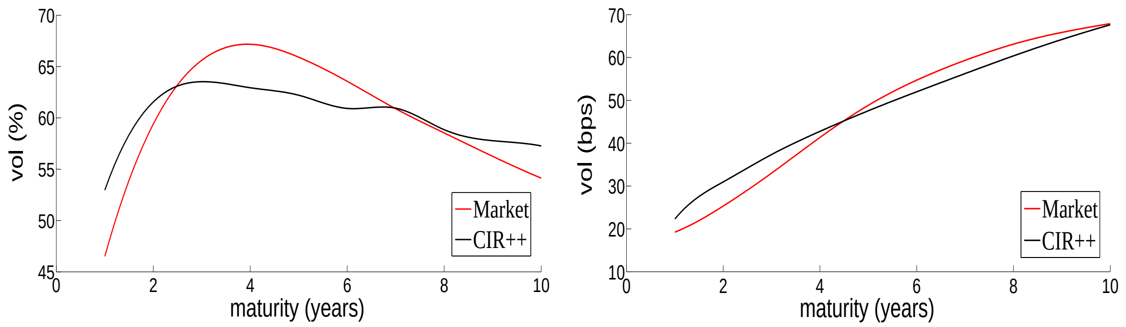
where σ^{CIR} and σ^{M} stand for the model- and the market-implied cap volatilities, respectively, and $T_1, \dots, T_n$ are the cap maturities. Model-implied cap volatilities are obtained by pricing market caps with the CIR++ model and then inverting the formula (3.244) in order to retrieve the implied volatility associated with each maturity. We choose to calibrate the model to cap volatilities since they are of similar magnitude, unlike cap prices which can differ by a few orders of magnitude. The calibration results are displayed in Table 3.5.

Table 3.5: The calibrated Cox–Ingersoll–Ross parameters.

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

On the one hand, (3.251) is a highly nonlinear and non-convex optimization problem, and the objective function may have multiple local minima. On the other hand, global optimization algorithms require a very high computation time and do not scale well with complexity, as opposed to local optimization methods. A fast calibration is important in practice since option pricing models may need to be recalibrated several times within a short time span. Therefore, we used a nonlinear least-squares solver, in particular the trust-region-reflective algorithm [17], for the calibration and a global optimization method, in particular a genetic algorithm [85], for verification purposes only.

Figure 3.3 shows the fitting capability of the CIR++ model, and the implied cap volatility curve is compared to the market curve for each currency. Taking into account that the model has only 4 parameters to fit between 10 and 11 data points, we conclude that the CIR++ model provides a fairly reasonable fit to the term structure of cap volatilities $T_i \mapsto \sigma^{\mathrm{M}}(T_i)$, $1 \le i \le n$.



(a) USD Currency

(b) EUR Currency

Figure 3.3: The market- and model-implied term structures of cap volatilities.

3.4.2. The hybrid stochastic volatility model

Consider a purely stochastic version of the model (3.1), namely the Heston–Cox–Ingersoll–Ross++ (Heston–2CIR++) model, and suppose that the domestic and the foreign short rate dynamics are independent of each other, and also independent of the dynamics of the spot FX rate and its volatility. The model is governed by the following system of SDEs under the domestic risk-neutral measure $\mathbb{Q}^d$:

$$\begin{cases} dS_t^{\mathrm{SV}} = \left(r_t^d - r_t^f\right) S_t^{\mathrm{SV}} dt + \sqrt{v_t} S_t^{\mathrm{SV}} \, dW_t^s, \quad S_0^{\mathrm{SV}} = S_0, \\ dv_t = k(\theta - v_t)dt + \xi \sqrt{v_t} \, dW_t^v, \\ r_t^d = g_t^d + h_d(t), \\ dg_t^d = k_d(\theta_d - g_t^d)dt + \xi_d \sqrt{g_t^d} \, dW_t^d, \\ r_t^f = g_t^f + h_f(t), \\ dg_t^f = k_f(\theta_f - g_t^f)dt + \xi_f \sqrt{g_t^f} \, dW_t^f, \end{cases} \tag{3.252}$$

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

Define the vector of parameters $\beta = (v_0, k, \theta, \xi, \rho)$. The next step in our calibration is to find the values of these 5 model parameters for which European call option prices best match the market call prices retrieved from volatility quotes for different strikes and maturities (see Table 3.6). The spot FX rate is $S_0 = 1.1271$. For a EUR–USD transaction, the market standard is to choose USD as the domestic currency and EUR as the foreign currency. The forward FX rate for a payment date T is defined as

$$F_T = \frac{P_f^{\mathrm{M}}(0,T)}{P_d^{\mathrm{M}}(0,T)} S_0, \tag{3.253}$$

where $P_d^{\mathrm{M}}(0,T)$ and $P_f^{\mathrm{M}}(0,T)$ are the domestic and foreign discount factors at time 0 for a maturity T, respectively. The European call option price (in domestic currency) at time

Table 3.6: The EURUSD volatility quotes (%) from March 18, 2016.

	10D Call		**25D Call**		**50D**		**25D Put**		**10D Put**	
Mat	**Strike**	**Vol**	**Strike**	**Vol**	**Strike**	**Vol**	**Strike**	**Vol**	**Strike**	**Vol**
2W	1.1547	9.405	1.1412	8.876	1.1280	8.497	1.1151	8.569	1.1032	8.880
1M	1.1696	9.224	1.1486	8.711	1.1287	8.375	1.1093	8.494	1.0911	8.831
2M	1.1892	9.667	1.1596	9.182	1.1302	8.947	1.1015	9.227	1.0738	9.762
3M	1.2081	10.026	1.1694	9.575	1.1318	9.465	1.0944	9.950	1.0564	10.689
6M	1.2450	9.916	1.1902	9.550	1.1371	9.760	1.0808	10.745	1.0216	12.034
1Y	1.3087	10.175	1.2259	9.726	1.1482	9.938	1.0672	10.954	0.9823	12.370
18M	1.3679	10.385	1.2602	9.894	1.1610	10.042	1.0615	10.991	0.9591	12.390
2Y	1.4222	10.507	1.2919	9.992	1.1746	10.100	1.0598	11.007	0.9437	12.342
3Y	1.5571	11.370	1.3657	10.600	1.2055	10.365	1.0640	10.790	0.9329	11.755
5Y	1.8230	12.237	1.5149	11.330	1.2739	10.900	1.0791	11.105	0.9119	11.882

0 with strike K and maturity T is computed with the Black–Scholes formula:

$$C_{\mathrm{BS}}(K,T,\sigma) = P_d^{\mathrm{M}}(0,T)\big[F_T\Phi(d_1) - K\Phi(d_2)\big], \tag{3.254}$$

$$d_{1,2} = \frac{\log(F_T/K) \pm \sigma^2 T/2}{\sigma\sqrt{T}}, \tag{3.255}$$

where the volatility σ corresponding to a strike K and a maturity T is retrieved from market quotes.

Under the postulated simple correlation structure of the Brownian drivers and when the short rates are driven by the CIR process, i.e., when $h_{d,f} = 0$, Ahlip and Rutkowski [1] derive an efficient closed-form formula for the European call option price. Hence, we denote by $C_{\mathrm{A}}(K,T)$ the fair value under the Heston–2CIR model of a European call option with strike K and maturity T computed with the aforementioned formula, and by $C_{\mathrm{H}}(K,T)$ the fair value of the same option but under the Heston–2CIR++ model. For $i \in \{d,f\}$, we define for brevity

$$H_i = \exp\left\{\int_0^T h_i(t)\,dt\right\}, \tag{3.256}$$

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

$$\begin{aligned} C_{\mathrm{H}}(K,T) &= \mathbb{E}^{\mathbb{Q}^d}\left[\exp\left\{-\int_0^T r_t^d\,dt\right\}\left(S_T^{\mathrm{SV}} - K\right)^+\right] \\ &= H_f^{-1}\,\mathbb{E}^{\mathbb{Q}^d}\left[\exp\left\{-\int_0^T g_t^d\,dt\right\}\left(H_f H_d^{-1} S_T^{\mathrm{SV}} - H_f H_d^{-1} K\right)^+\right] \\ &= H_f^{-1} C_{\mathrm{A}}(\tilde{K},T), \end{aligned} \tag{3.257}$$

where the shifted strike is

$$\tilde{K} = H_f H_d^{-1} K. \tag{3.258}$$

We now calibrate the Heston–2CIR++ model by minimizing the sum of the squared differences between model and market call prices:

$$\min_{\beta \in \mathbb{R}_+^4 \times [-1,1]} \sum_{\substack{1 \le i \le n \\ 1 \le j \le m}} \left[C_{\mathrm{H}}(K_j, T_i; \beta) - C_{\mathrm{BS}}(K_j, T_i, \sigma_{i,j}) \right]^2, \tag{3.259}$$

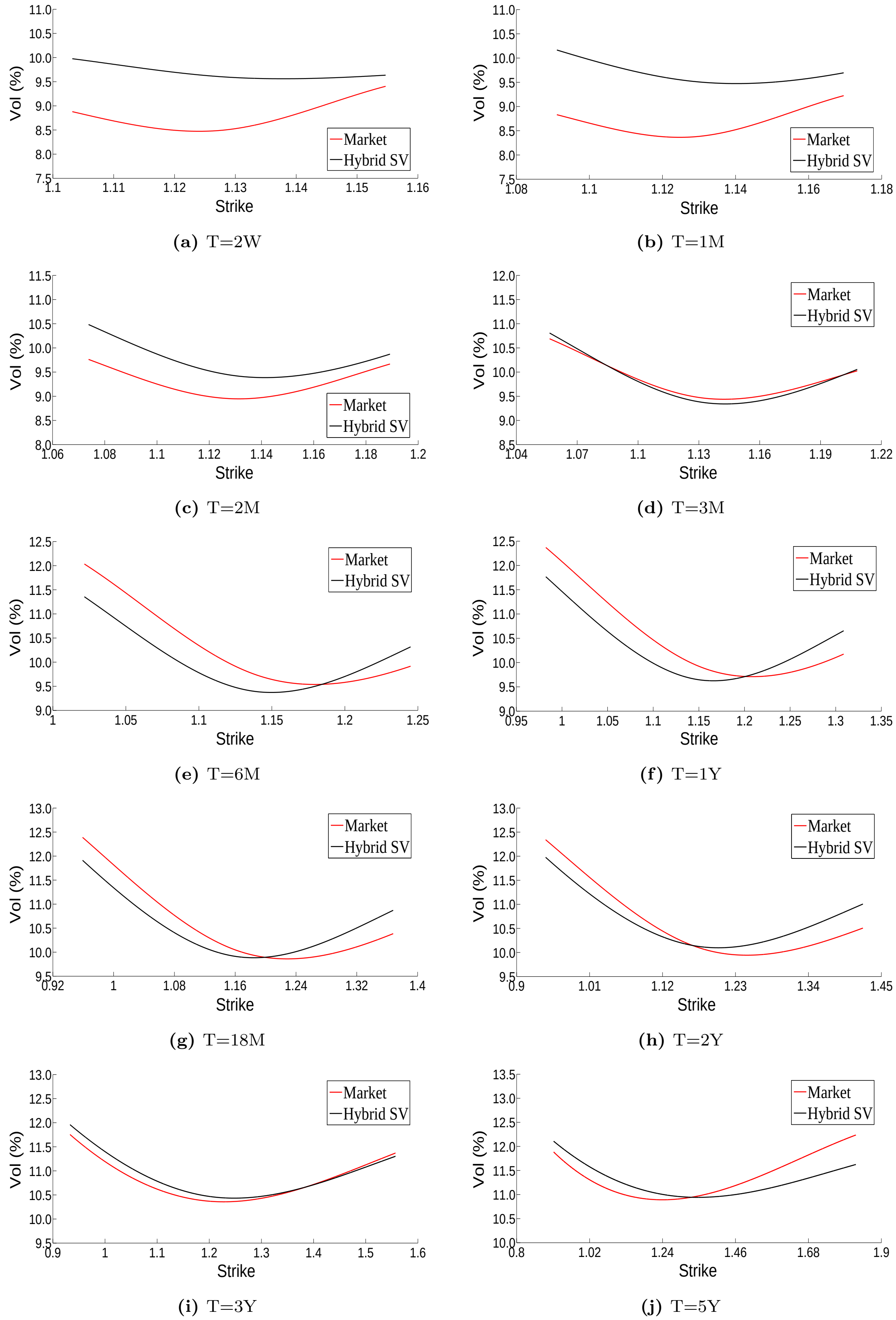
where $\sigma_{i,j}$ is the quoted volatility corresponding to a strike K_j and a maturity T_i, for $j = 1, \ldots, m$ and $i = 1, \ldots, n$. There are many ways to choose the objective function (error measure) in (3.259). For instance, we may consider either call prices or Black–Scholes implied volatilities and minimize the sum of either absolute or relative (squared) differences between model and market values, using either uniform or non-uniform weights. We chose this particular error measure, which assigns more weight to more expensive options (in-the-money, long-term) and less weight to cheaper options (out-of-the-money, short-term), for two reasons. First, the Heston model, and hence the Heston–2CIR++ model by extension, cannot reproduce the smiles or skews typically observed for short maturities that well and a more careful calibration to these smiles would result in a larger overall model error due to the inherent poor fit of the model to the short-term. Second, market data becomes scarce as the maturity increases, and hence we already assigned more weight to the short- and mid-term sections of the volatility surface; for instance, we have more maturities up to 1 year than between 1 and 5 years. The calibration results are displayed in Table 3.7.

Table 3.7: The calibrated Heston parameters.

v_0	k	θ	ξ	ρ
0.0094	1.4124	0.0137	0.2988	−0.1194

As before, we employed a nonlinear least-squares solver (the trust-region-reflective algorithm) for the calibration and a global optimization method (a genetic algorithm) for verification purposes. Due to the nonlinearity and non-convexity of the problem, the calibrated model parameters may end up in a local rather than a global minimum of the objective function. Hence, a good initial parameter guess may significantly improve the quality of the calibration, and we chose the squared ATM 2-week and 5-year volatilities as v_0 and θ, respectively.

The data in Figure 3.4 suggest a poor fit of the Heston–2CIR++ model for short-term options (up to 1 month), which is to be expected considering the well-known difficulty of the Heston model to match the skews observed in the market for short maturities and due to our choice of the objective function which favors long maturities. On the other hand, we infer from Figure 3.4 that for mid- and long-term options, the implied volatility curves display qualitatively correct shapes for all strikes and the Heston–2CIR++ model provides a fairly reasonable fit, especially when taking into account that we only have 5 parameters to fit 50 data points.



(a) T=2W **(b)** T=1M

(c) T=2M **(d)** T=3M

(e) T=6M **(f)** T=1Y

(g) T=18M **(h)** T=2Y

(i) T=3Y **(j)** T=5Y

Figure 3.4: The market- and model-implied volatility curves.

3.4.3. The hybrid stochastic-local volatility model

Next, we briefly discuss the calibration of the leverage function σ in the Heston–2CIR++ SLV model (3.1). We refer to [20] for the full calibration procedure with all the details. First, we calibrate the LV model (3.7) using the Dupire forward PDE and find σ_{LV}. Then, we calibrate the associated 2-factor SLV model with deterministic rates via PDE methods using formula (3.12). Finally, we employ the particle method of [44] and extend it using the calibrated 2-factor SLV model with deterministic rates as a control variate in order to find σ, illustrated in Figure 3.5.

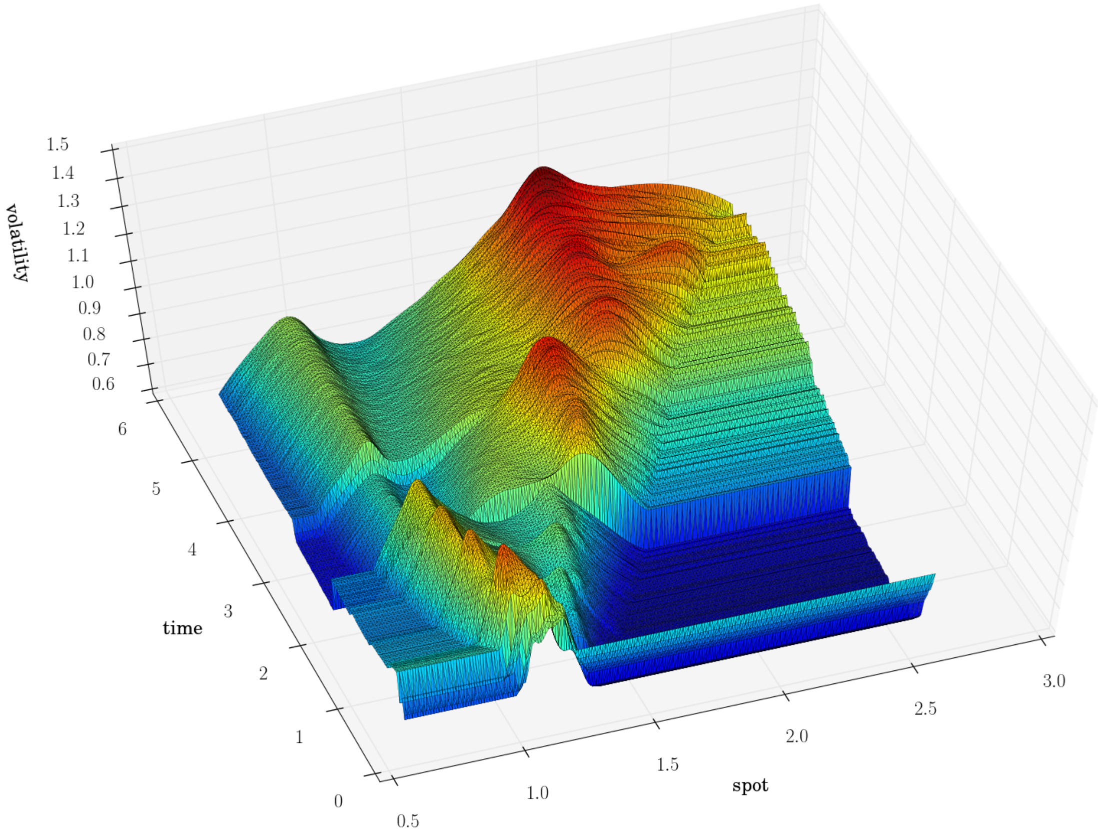


Figure 3.5: The calibrated leverage function for EURUSD market data from March 18, 2016 (taken from [20]).

In [20], the Heston–2CIR++ SLV model (3.1) was calibrated with a very high accuracy with as few as 20 000 particles, for a maximum error in the absolute volatility of only 0.03%. The importance of the local volatility component for a good fit to vanilla options is illustrated in Figure 3.6, where the market-implied volatility curve is plotted together with the model-implied volatility curves corresponding to the 4-factor SLV model (3.1) and the 4-factor SV model (3.252), respectively.

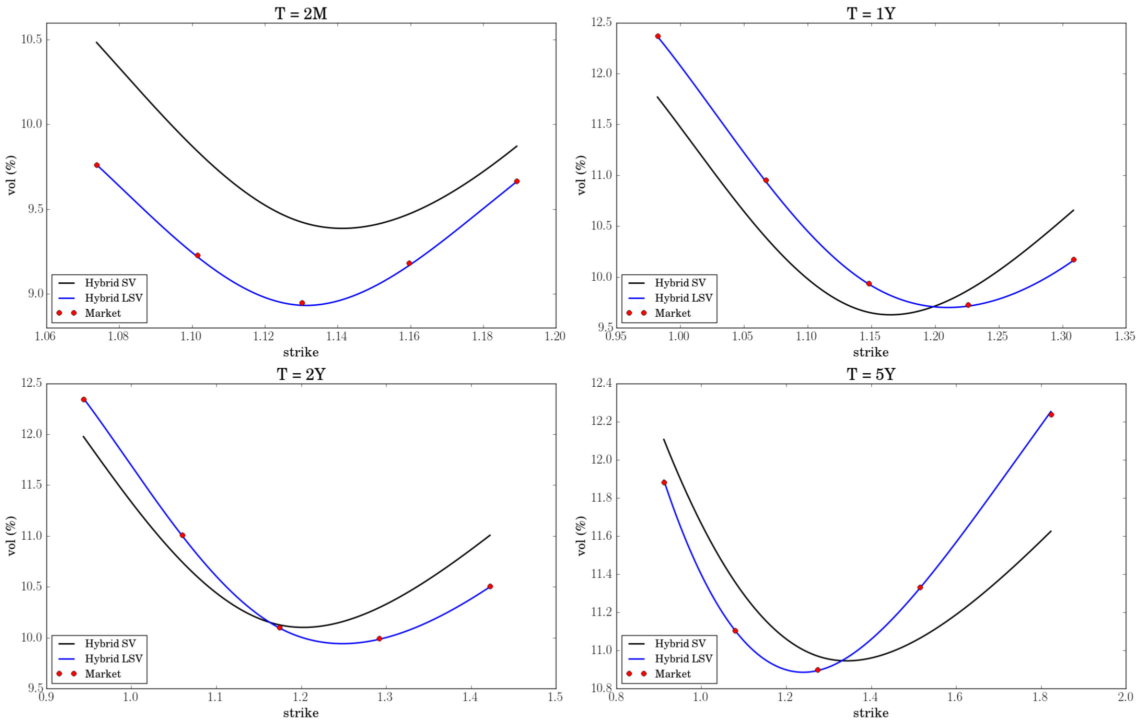


Figure 3.6: Calibration fit with 500 000 particles to EURUSD market data from March 18, 2016 for the 4-factor SLV and SV models. The repricing is performed using Monte Carlo simulations with 500 time steps per year and 5 000 000 paths (taken from [20]).

3.5. Conclusions

In this chapter, we have established the strong L^1-convergence (without a rate) of the discounted exchange rate approximation under a hybrid stochastic-local volatility model with stochastic short rates, a result which can be generalized to the L^p case relatively easily, for all $p \geq 1$, albeit under a stronger condition on the maturity, and which is particularly useful when proving the convergence of Monte Carlo simulations for valuing options with unbounded payoffs. The only previous published work related to this problem that we are aware of is [47], which proves the convergence of an Euler discretization with a reflection fix in the context of the Heston stochastic volatility model and options with bounded payoffs. The analysis carried out in this chapter can be extended to other financial derivatives, including digital options, forward-start options and double-no-touch binary options, to name just a few. Furthermore, we may consider a multi-factor extension of the short rate model, in which case the convergence analysis applies with some slight modifications of the proofs.

However, several unsettled questions remain, like the exact strong convergence order of the full truncation scheme for the CIR process or whether we can find a positive strong order of convergence of schemes for the type of SDEs studied in this chapter. On top of these being interesting and practically relevant questions in their own right, a sufficiently high order enables the use of multi-level simulation, as in [39], with substantial efficiency improvements for the estimation of expected financial payoffs.

3.6. Appendix

In this appendix, we consider the option pricing problem for exotic products, in particular, for structured notes embedding barrier features. Popular FX exotic products include the power reverse dual currency note (PRDC) [16], especially long-term (e.g., 30 years), and also the forward accumulator [97]. Adding stochastic rates improves the pricing of both contracts for mid- to long-term expiries. Moreover, while the plain PRDC can be seen as a strip of vanilla options, adding stochastic-local volatility dynamics improves the pricing of some variations of the contract, like the trigger PRDC (a PRDC with a knock-out feature).

Henceforth, we consider an autocallable barrier dual currency note (ABDC) that is indexed on the EURUSD currency pair, with USD as the domestic currency and EUR as the foreign currency, with quarterly coupon payments. While barrier dual currency notes are typically short-term investments, we examine instead a mid-term alternative with an embedded autocallable note structure. For a nominal N (in USD), a rate (strike) K and an expiry T, the life cycle of the contract is described below.

(1) A monthly fixing schedule is defined for the exchange rate S such that:

- If S crosses the up barrier B_{UO}, the contract is redeemed early and a coupon $\mathcal{C}_{\mathrm{ER}}$ is paid. The early redemption time is denoted by τ_{ER} and can be infinite in case of no knock-out event.
- If S crosses the down barrier B_{DI}, a short put contract is activated at expiry. The knock-in time is denoted by τ_{KI}.

(2) If S is above B_{DI}, a coupon $\mathcal{C}$ (in USD, in % of the nominal) is paid quarterly at coupon dates $\{t_i\,,\, i = 1, \ldots, M\}$, for a total of M periods, such that the accumulated coupon (in USD) at expiry is

$$N \sum_{i=1}^{M} \frac{D^d_{t_i}}{D^d_T} \mathcal{C}\, \mathbb{1}_{t_i < \tau_{\mathrm{ER}}}\, \mathbb{1}_{S_{t_i} > B_{\mathrm{DI}}} + N \frac{D^d_{\tau_{\mathrm{ER}}}}{D^d_T} \mathcal{C}_{\mathrm{ER}}\, \mathbb{1}_{\tau_{\mathrm{ER}} \leq T}\,. \tag{3.260}$$

(3) At expiry, if S is below K, the down barrier has been activated and early redemption has not occurred, then the nominal is converted to EUR at the rate K.

This product would suit an investor who wants to outperform the money market account by taking the risk of conversion of the nominal value to the foreign currency. The autocall feature lowers the price of the product due to early termination in case the exchange rate crosses the up-and-out barrier, whereas the down-and-in barrier feature provides protection against the conversion risk at the expense of increasing the net present value (NPV). Hence, an investor would find this product attractive under the current volatile market conditions if they expected a less volatile market in the future, which would imply a lower NPV due to the increased chance of early redemption and conversion. Over a longer time horizon, a very stable market would be the best outcome for the investor, since they would get all the coupons without converting the nominal at maturity. If conversion occurs at expiry, the investor receives N/K in EUR in exchange for the nominal N in USD. If the investor converts this amount to USD, the loss becomes $N\big(S_T/K - 1\big)$. Hence, the profit-and-loss (PnL) of the product is

$$N \sum_{i=1}^{M} \frac{D^d_{t_i}}{D^d_T} \mathcal{C}\, \mathbb{1}_{t_i<\tau_{\mathrm{ER}}}\, \mathbb{1}_{S_{t_i}>B_{\mathrm{DI}}} + N \frac{D^d_{\tau_{\mathrm{ER}}}}{D^d_T} \mathcal{C}_{\mathrm{ER}}\, \mathbb{1}_{\tau_{\mathrm{ER}}\le T} - \frac{N}{K}(K - S_T)^+\, \mathbb{1}_{\tau_{\mathrm{KI}}\le T<\tau_{\mathrm{ER}}}\,. \tag{3.261}$$

Hence, this product is a yield enhancement contract with no capital guarantee. The accrued coupon is sometimes converted at expiry with the nominal at the rate K. Finally, the NPV of the contract (in % of the nominal) is

$$\begin{aligned}\mathrm{NPV} = \mathbb{E}\Bigg[&\sum_{i=1}^{M} D^d_{t_i} \mathcal{C}\, \mathbb{1}_{t_i<\tau_{\mathrm{ER}}}\, \mathbb{1}_{S_{t_i}>B_{\mathrm{DI}}} + D^d_{\tau_{\mathrm{ER}}} \mathcal{C}_{\mathrm{ER}}\, \mathbb{1}_{\tau_{\mathrm{ER}}\le T} \\ &- \frac{D^d_T}{K} (K - S_T)^+\, \mathbb{1}_{\tau_{\mathrm{KI}}\le T<\tau_{\mathrm{ER}}} \Bigg],\end{aligned} \tag{3.262}$$

where the expectation is taken under the risk-neutral measure. We assume the dynamics of S under this measure as specified in (3.1).

Next, consider the contract parameters from Table 3.8 as well as the calibrated model parameters and leverage function – to the EURUSD market data from March 18, 2016 – from Section 3.4. Suppose that barriers are monitored monthly, at the fixing dates, and coupons are paid quarterly.

Table 3.8: The contract parameters of an autocallable barrier dual currency note, where the nominal N is in USD, the expiry T is in years, the strike K and the barriers B_{UO} and B_{DI} are in % of S_0, and the coupons $\mathcal{C}$ and $\mathcal{C}_{\mathrm{ER}}$ are in USD, in % of N.

N	T	S_0	K	B_{UO}	B_{DI}	$\mathcal{C}$	$\mathcal{C}_{\mathrm{ER}}$
100 000	5Y	1.1271	105%	100%	95%	2.5%	1.5%

We employ the Monte Carlo simulation scheme defined in Section 3.3 with 5×10^7 sample paths and 376 time steps per year to price this contract, and the numerical results are displayed in Table 3.9. Note that we computed the percentage change in premium (NPV) with respect to the reference model, i.e., the Heston–2CIR++ SLV model.

We infer from Table 3.9 that the price of the contract under the 2-factor SLV model is 1.21% higher than under the 4-factor SLV model, whereas the early redemption and the knock-in probabilities are the same under the two models. This suggests that the stochastic rates have a significant impact on the price of the contract even for a 5-year expiry. We

Table 3.9: The NPV of the contract in % of N, the 95% Monte Carlo confidence interval, the early redemption and knock-in probabilities and the percentage change in NPV, for the autocallable barrier dual currency note specified in Table 3.8 under the 4-factor hybrid SLV model (3.1), the 2-factor SLV model with deterministic rates and the LV model (3.7).

Model	NPV	95% CI	ER prob.	KI prob.	Change
HSLV	1.7247	$(1.7228, 1.7265)$	95.1%	19.5%	–
SLV	1.7456	$(1.7438, 1.7474)$	95.1%	19.5%	+1.21%
LV	1.2079	$(1.2062, 1.2097)$	95.9%	19.5%	−29.96%

also infer from Table 3.9 that the contract is highly underpriced and the early redemption probability is overestimated under the LV model (3.7). Furthermore, we calibrated in Section 3.4 the 4-factor SLV and SV models to EURUSD market data from March 18, 2016, and observed an almost perfect fit to vanilla options with the former as opposed to a poor fit with the latter. In conclusion, we notice two things. First, pricing exotic products with barrier features under stochastic-local volatility dynamics is of paramount importance. It is a well-known fact that pure SV models underestimate the knock-out probability whereas pure LV models overestimate it, and the true price of the contract is believed to lie in between pure SV and pure LV model prices. In our case, a small difference in the probability of no knock-out (4.1% versus 4.9%), which is leveraged by the number of coupons detached during the life cycle of the contract, can lead to a big difference in the NPV, and this explains the need to use SLV models to improve the pricing performance. Second, adding stochastic rates to an SLV model is only relevant when pricing mid- to long-term structured notes, where the effect of the stochastic rates is accentuated. The impact of stochastic rates on the contract price is expected to increase in the presence of non-zero correlations between the spot FX rate and the short rates.

We could extend the model (3.1) to multi-factor short rates when pricing exotic products where the rates appear explicitly in the payoff, for example, a spread option with the payoff

$$\left[\left(\frac{S_T - S_0}{S_0}\right) - L_T - K\right]^+, \tag{3.263}$$

where L_T the Libor rate at the fixing date T. In this case, the calibration algorithm described in Section 3.2 can still be applied, with a higher computational cost due to the more complex simulation scheme.

We conclude this appendix with an empirical convergence analysis of our Monte Carlo simulation scheme. Since the product can be decomposed into a linear combination of first-order exotics, the theoretical convergence (without a rate) of Monte Carlo estimators follows automatically from the analysis in Section 3.3. The data in Table 3.10 suggest a first-order convergence of the time-discretization error. As an aside, note that in the case of continuously monitored barriers, we could use Brownian bridge techniques to recover the first-order convergence.

Table 3.10: The Monte Carlo estimates of the NPV of the contract specified in Table 3.8 for different numbers of time steps (per year) and 5×10^7 sample paths (for a standard deviation of 9.29×10^{-4}), the difference from the previous NPV estimate and the empirical convergence order.

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

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

Chapter 4

A Mixed Monte Carlo and Partial Differential Equation Method

4.1. Introduction

In this chapter, we study the FX option pricing problem under the Heston–Cox–Ingersoll–Ross (Heston–2CIR) stochastic volatility model proposed and examined in [1] when the dynamics of the exchange rate and its volatility are correlated, whereas the domestic and foreign short interest rates are pairwise independent and also independent of the exchange rate and the volatility. Under these restrictive assumptions on the independence of the Brownian drivers, Ahlip and Rutkowski [1] argued that the model is affine and derived a semi-analytical formula for the European call option price. The importance of a non-zero correlation between the FX rate and the interest rate(s) was recognized in [51]. However, any other non-zero correlations give rise to a non-affine model, in which case we lose the analytical tractability.

Therefore, we find ourselves forced to turn to numerical algorithms, and the mixed Monte Carlo/partial differential equation solver [72] is a good alternative to the classical Monte Carlo and finite difference (FD) methods. The idea behind the mixed Monte Carlo/PDE method is to write the option values as nested conditional expectations. Then, the innermost expectation is evaluated analytically, in the case of European-style options, or using finite differences, and the outer expectation is evaluated by simulation. The utility of the method can be easily recognized when pricing European-style options under the Heston–2CIR model. Conditional on the entire paths of the squared volatility and the domestic and foreign short rates, the dynamics of the exchange rate are governed by a geometric Brownian motion with time-dependent drift and diffusion coefficients. Combined with the existence of a closed-form solution for the conditional option price, the algorithm reduces the variance in Monte Carlo simulations – by eliminating a source of noise – and the dimension of the problem, from 4 to 3. The mixed Monte Carlo/PDE method has been the subject of several numerical studies in the past few years [8; 23; 68; 74] and various extensions to the original idea have been considered. For instance, Dang et al. [23] used the Hull–White-type dynamics of the interest rates and conditioned on the variance path to find a closed-form solution for the conditional price of a European option. A few of these works examined convergence properties of the algorithm using heuristic arguments, but none of them took into account the error arising from the discretization of the variance and interest rate processes.

To the best of our knowledge, the convergence of the mixed Monte Carlo/PDE method has not yet been established, and even the literature on Monte Carlo methods under stochastic volatility is scarce. Higham and Mao [47] considered an Euler simulation of the Heston model with a reflection fix and proved strong convergence of the stopped approximation process and weak convergence for a European put and an up-and-out call, by using the boundedness of payoffs. In Chapter 3, we extended the results of Higham and Mao to stochastic-local volatility models, stochastic interest rates and derivatives with unbounded and exotic payoffs. The discretization scheme in this chapter coincides with that in Chapter 3 (when the Heston–2CIR++ SLV model collapses to the Heston–2CIR SV model) at the discrete time points, but necessarily differs in the continuous-time interpolation. Some results from Chapter 3 can be utilized, although stronger conditions on the model parameters were needed there and so the results here can also be seen as an improvement for the standard (i.e., non-mixed) scheme. In particular, the new results are always guaranteed for zero or negative correlation in the Heston model. The main conceptual difference, however, is the new continuous-time interpolation using conditional drifts, which is crucial for deriving the conditional PDEs and leads to different technical challenges. Also, for path-dependent options, neither the scheme nor the analysis in Chapter 3 applies to the present work.

The major contributions of this chapter are as follows. First, we derive the uniform boundedness of moments of the 4-dimensional process and its approximation, and prove strong convergence of the discretization scheme in L^p $(p \geq 1)$. We then deduce the convergence of mixed Monte Carlo/PDE estimators for computing option prices and discuss possible extensions to higher-dimensional models. Second, we carry out a thorough theoretical variance reduction analysis of the mixed Monte Carlo/PDE method and employ standard Monte Carlo – with the log-Euler discretization – as the reference method, noting that the analysis applies to general interest rate dynamics. In particular, we investigate how different values of the underlying model parameters affect the variance of mixed estimators. Third, we perform a series of numerical experiments and demonstrate the convergence of the mixed Monte Carlo/PDE method under the 4-factor FX model for two financial derivatives: a European call and an up-and-out put option. In addition, we establish the efficiency of the method by comparison with alternative numerical schemes and examine the sensitivity of the variance reduction factor to changes in the parameters.

The remainder of this chapter is structured as follows. In Section 4.2, we introduce the model, define the mixed simulation scheme and describe the pricing algorithm. In

Section 4.3, we prove strong convergence of the exchange rate approximations and then discuss some extensions. In Section 4.4, we carry out a variance reduction analysis of the mixed Monte Carlo/PDE method for a European option. Various numerical experiments are presented and discussed in Section 4.5. Finally, Section 4.6 summarizes the results and outlines possible future work.

4.2. Preliminaries

4.2.1. Model definition

We have in mind a model in an FX market, for the spot FX rate S, the squared volatility of the FX rate v, the domestic short interest rate r^d and the foreign short interest rate r^f. Unless otherwise stated, the subscripts and superscripts "d" and "f" in this chapter are used to indicate "domestic" and "foreign", respectively. Consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q})$ satisfying the usual conditions and suppose that the dynamics of the underlying processes are governed by the system of SDEs below under the domestic risk-neutral measure $\mathbb{Q}$:

$$\begin{cases} dS_t = \big(r_t^d - r_t^f\big)S_t dt + \sqrt{v_t}S_t dW_t^s, \quad S_0 > 0, \\ dv_t = k\big(\theta - v_t\big)dt + \xi\sqrt{v_t}\, dW_t^v, \quad v_0 > 0, \\ dr_t^d = k_d\big(\theta_d - r_t^d\big)dt + \xi_d\sqrt{r_t^d}\, dW_t^d, \quad r_0^d > 0, \\ dr_t^f = \big(k_f\theta_f - k_f r_t^f - \rho_{sf}\xi_f\sqrt{v_t r_t^f}\big)dt + \xi_f\sqrt{r_t^f}\, dW_t^f, \quad r_0^f > 0, \end{cases} \tag{4.1}$$

where $\{W^s, W^f, W^d, W^v\}$ are 1-dimensional standard $\mathcal{F}_t$-adapted Brownian motions with constant correlation matrix Σ. We consider a full correlation structure between the four Brownian drivers, which reflects movements in the financial markets more accurately and allows for better calibrations. Then, we decouple $\{W^s, W^f, W^d, W^v\}$ and express them as linear combinations of independent Brownian motions $\{W^1, W^2, W^3, W^4\}$. Hence, we define the two vectors $W = [W^s, W^f, W^d, W^v]^{\mathrm{T}}$ and $\tilde{W} = [W^1, W^2, W^3, W^4]^{\mathrm{T}}$. Since Σ is symmetric positive definite, a standard Cholesky factorization gives rise to an upper triangular matrix of coefficients $A = (a_{ij})_{1\leq i,j\leq 4}$ satisfying $\Sigma = AA^{\mathrm{T}}$:

$$\Sigma = \begin{bmatrix} 1 & \rho_{sf} & \rho_{sd} & \rho_{sv} \\ \rho_{sf} & 1 & \rho_{df} & \rho_{vf} \\ \rho_{sd} & \rho_{df} & 1 & \rho_{vd} \\ \rho_{sv} & \rho_{vf} & \rho_{vd} & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.2}$$

This decomposition implies that we can choose $\tilde{W}$ such that $W = A\tilde{W}$. We can determine the matrix of coefficients by solving a system of ten equations. Assuming that $\rho_{vd} \neq \pm 1$, we find

$$
\begin{aligned}
a_{14} &= \rho_{sv}, \quad a_{24} = \rho_{vf}, \\
a_{34} &= \rho_{vd}, \quad a_{33} = \left(1-\rho_{vd}^2\right)^{\frac{1}{2}}, \\
a_{13} &= \left(\rho_{sd}-\rho_{sv}\rho_{vd}\right)\left(1-\rho_{vd}^2\right)^{-\frac{1}{2}}, \\
a_{23} &= \left(\rho_{df}-\rho_{vf}\rho_{vd}\right)\left(1-\rho_{vd}^2\right)^{-\frac{1}{2}}, \\
a_{22} &= \left(1-\rho_{df}^2-\rho_{vf}^2-\rho_{vd}^2+2\rho_{df}\rho_{vf}\rho_{vd}\right)^{\frac{1}{2}}\left(1-\rho_{vd}^2\right)^{-\frac{1}{2}}.
\end{aligned}
\tag{4.3}
$$

Finally, a_{11} and a_{12} can be found in a similar fashion. Note that the quanto correction term in the drift of the foreign interest rate in (4.1) comes from changing from the foreign risk-neutral measure to the domestic one [16]. Alternatively, we can think of (4.1) as a model in an equity market with asset price process S, interest rate r^d and dividend yield r^f, in which case the quanto correction term vanishes.

4.2.2. The mixed simulation scheme

First, we discretize the variance and the two interest rate processes using the FTE scheme of [73]. Consider the square-root process

$$
dy_t = k_y(\theta_y - y_t)dt + \xi_y\sqrt{y_t}\, dW_t^y. \tag{4.4}
$$

Let T be the maturity of the option under consideration and consider a uniform grid

$$
T = N\delta t,\ \ t_n = n\delta t, \quad \forall n \in \{0, 1, ..., N\}. \tag{4.5}
$$

We introduce the discrete-time auxiliary process

$$
\tilde{y}_{t_{n+1}} = \tilde{y}_{t_n} + k_y(\theta_y - \tilde{y}_{t_n}^+)\delta t + \xi_y\sqrt{\tilde{y}_{t_n}^+}\, \delta W_{t_n}^y, \tag{4.6}
$$

where $y^+ = \max(0, y)$ and $\delta W_{t_n}^y = W_{t_{n+1}}^y - W_{t_n}^y$, and the continuous-time interpolation

$$
\tilde{y}_t = \tilde{y}_{t_n} + k_y(\theta_y - \tilde{y}_{t_n}^+)(t-t_n) + \xi_y\sqrt{\tilde{y}_{t_n}^+}\left(W_t^y - W_{t_n}^y\right), \quad \forall t \in [t_n, t_{n+1}), \tag{4.7}
$$

Then, we define the non-negative processes

$$
\begin{cases} \hat{y}_t = \tilde{y}_t^+ & (4.8) \\ \bar{y}_t = \tilde{y}_{t_n}^+ & (4.9) \end{cases}
$$

whenever $t \in [t_n, t_{n+1})$. Let $\bar{v}$ and $\bar{r}^d$ be the FTE discretizations – as defined in (4.9) – of v and r^d, respectively. Taking into account the presence of the quanto correction term in the drift of the foreign interest rate, we similarly define

$$
\begin{aligned}
\tilde{r}_t^f = \tilde{r}_{t_n}^f &+ \left[k_f\theta_f - k_f\left(\tilde{r}_{t_n}^f\right)^+ - \rho_{sf}\xi_f\sqrt{\tilde{v}_{t_n}^+\left(\tilde{r}_{t_n}^f\right)^+}\right](t-t_n) \\
&+ \xi_f\sqrt{\left(\tilde{r}_{t_n}^f\right)^+}\left(W_t^f - W_{t_n}^f\right),
\end{aligned}
\tag{4.10}
$$

as well as

$$
\begin{cases}
\hat{r}^f_t = \left(\tilde{r}^f_t\right)^+ & (4.11)\\
\bar{r}^f_t = \left(\tilde{r}^f_{t_n}\right)^+ & (4.12)
\end{cases}
$$

whenever $t \in [t_n, t_{n+1})$. Next, we define $\bar{S}$, the continuous-time approximation of S, as the solution to the following SDE:

$$
\begin{cases}
d\bar{S}_t = \mu_t \bar{S}_t dt + a_{11}\sqrt{\bar{v}_t}\,\bar{S}_t\, dW^1_t,\\
\mu_t \;= \bar{r}^d_t - \bar{r}^f_t - \dfrac{1}{2}\left(1 - a^2_{11}\right)\bar{v}_t + \displaystyle\sum_{j=2}^{4} a_{1j}\sqrt{\bar{v}_t}\,\frac{\delta W^j_t}{\delta t},
\end{cases}
\tag{4.13}
$$

where $\delta W^j_t = W^j_{t_{n+1}} - W^j_{t_n}$ for all $t \in [t_n, t_{n+1})$, so μ_t is piecewise constant. For convenience, we introduce the actual and the approximated log-processes, $x = \log S$ and $\bar{x} = \log \bar{S}$.

Conditioning on the trajectories of $\left\{W^j, j = 2, 3, 4\right\}$, i.e., on the complete knowledge of the paths of the variance and interest rates, $\bar{S}$ evolves like a geometric Brownian motion with time-dependent drift and diffusion coefficients. It follows from Itô's formula that

$$
\bar{S}_T = S_0 \exp\left\{ \int_0^T \left(\bar{r}^d_u - \bar{r}^f_u - \frac{1}{2}\bar{v}_u\right) du + \sum_{j=2}^{4} a_{1j} \int_0^T \sqrt{\bar{v}_u}\,\frac{\delta W^j_u}{\delta t}\, du + a_{11}\int_0^T \sqrt{\bar{v}_u}\, dW^1_u \right\}.
\tag{4.14}
$$

However, we know the conditional probability law of the stochastic integral on the right-hand side to be that of a normal random variable, namely

$$
\int_0^T \sqrt{\bar{v}_u}\, dW^1_u \;\overset{\text{law}}{=\!=}\; \sqrt{\int_0^T \bar{v}_u\, du}\; Z,
\tag{4.15}
$$

where $Z \sim \mathcal{N}(0, 1)$, and hence we can think of $\bar{S}_T$ as a function of Z. Therefore, we can express it as

$$
\bar{S}_T \overset{\text{law}}{=\!=} S_0 \exp\left\{ \left(r - q - \frac{1}{2}\sigma^2\right) T + \sigma\sqrt{T}\, Z \right\},
\tag{4.16}
$$

where

$$
r \;=\; \frac{1}{T}\int_0^T \bar{r}^d_u\, du \;=\; \frac{1}{N}\sum_{i=0}^{N-1} \bar{r}^d_{t_i},
\tag{4.17}
$$

$$
\sigma^2 \;=\; \frac{a^2_{11}}{T}\int_0^T \bar{v}_u\, du \;=\; \frac{a^2_{11}}{N}\sum_{i=0}^{N-1} \bar{v}_{t_i}
\tag{4.18}
$$

and

$$
\begin{aligned}
q \;&=\; \frac{1}{T}\int_0^T \bar{r}^f_u\, du + \frac{1 - a^2_{11}}{2T}\int_0^T \bar{v}_u\, du - \frac{1}{T}\sum_{j=2}^{4} a_{1j}\int_0^T \sqrt{\bar{v}_u}\,\frac{\delta W^j_u}{\delta t}\, du\\
&=\; \frac{1}{N}\sum_{i=0}^{N-1} \bar{r}^f_{t_i} + \frac{1 - a^2_{11}}{2N}\sum_{i=0}^{N-1} \bar{v}_{t_i} - \frac{1}{T}\sum_{j=2}^{4} a_{1j}\sum_{i=0}^{N-1} \sqrt{\bar{v}_{t_i}}\,\delta W^j_{t_i}.
\end{aligned}
\tag{4.19}
$$

Conditional on the trajectories of $\{W^j, j = 2,3,4\}$, (4.16) has the same law as a terminal asset price that evolves as a geometric Brownian motion with interest rate r, continuous dividend yield q and volatility σ, all constant. Then the arbitrage-free price at time $t = 0$ of a European-style option with payoff $f(S_T)$ is the discounted expectation under the risk-neutral measure, which can be approximated, using the mixed Monte Carlo/PDE method, by

$$\bar{U} = \mathbb{E}\left[e^{-\int_0^T \bar{r}_u^d \, du} f(\bar{S}_T)\right] = \mathbb{E}\left[\mathbb{E}\left[e^{-rT} f(\bar{S}_T) \middle| \mathcal{G}_T^{f,d,v}\right]\right], \tag{4.20}$$

where $\{\mathcal{G}_t^{f,d,v}, 0 \le t \le T\}$ is the natural filtration generated by the independent Brownian motions $\{W^2, W^3, W^4\}$. Note that the processes v, r^d and r^f are adapted to this filtration. The second equality in (4.20) comes from the "tower property" of conditional expectations. Unless otherwise stated, all expectations are under $\mathbb{Q}$. Let the approximate conditional option price be the inner expectation in (4.20), which is analytically tractable for European contracts,

$$\mathbb{E}\left[e^{-rT} f(\bar{S}_T) \middle| \mathcal{G}_T^{f,d,v}\right] = e^{-rT} \int_{-\infty}^{\infty} f\left(\bar{S}_T(z)\right) \phi(z) \, dz, \tag{4.21}$$

where ϕ and Φ are the standard normal probability density function (PDF) and cumulative distribution function (CDF), respectively. The conditional prices of some popular financial instruments (for a strike K) are given below:

$$\begin{cases} \text{European options:} & \psi S_0 e^{-qT} \Phi(\psi d_1) - \psi K e^{-rT} \Phi(\psi d_2) \\ \text{cash-or-nothing options:} & e^{-rT} \Phi(\psi d_2) \\ \text{asset-or-nothing options:} & S_0 e^{-qT} \Phi(\psi d_1), \end{cases} \tag{4.22}$$

where $\psi = 1$ for a call and $\psi = -1$ for a put, and

$$d_{1,2} = \frac{\log(S_0/K) + \left(r - q \pm \sigma^2/2\right) T}{\sigma \sqrt{T}}. \tag{4.23}$$

The approximate option price, i.e., the outer expectation in (4.20), is estimated by a Monte Carlo average over a sufficiently large number of discrete trajectories of $\{W^2, W^3, W^4\}$.

There are many other derivatives that admit a closed-form solution for the conditional price, e.g., the power option, the chooser option or the forward-start option. However, for most path-dependent derivatives we need to use a different approach in order to compute the conditional price, in which case we will rely on finite difference methods (see Section 4.5).

We conclude this section with a discussion on our choice of conditioning. For European option pricing, an analytical formula for the inner expectation is available only when conditioning on all three factors, which results in a reduction of the problem by one dimension. The high-dimensionality of the Heston–2CIR model makes this the natural choice. For path-dependent option pricing, due to the quanto correction term in the drift of the foreign rate, v and r^f are coupled, and hence we cannot condition on r^f alone. Moreover, the variance of Monte Carlo estimators due to the short rates is typically much lower than the variance due to the instantaneous squared volatility. Hence, we could alternatively condition on r^d and solve a 3-dimensional PDE for the inner expectation. On the one hand, we could reach the same level of accuracy with fewer Monte Carlo sample paths than when simulating v and r^f as well. On the other hand, the computational effort grows linearly with the dimension for Monte Carlo methods and exponentially for finite difference methods. Hence, we believe that conditioning on all three factors is more efficient. However, if the exchange rate and the foreign interest rate dynamics are independent, the quanto correction term vanishes. In this case, conditioning on the two short rates would be an interesting alternative.

4.3. Convergence analysis

Even though weak convergence is very important in financial mathematics when estimating expectations of payoffs, strong convergence may be required for complex path-dependent derivatives and plays a key role in multilevel Monte Carlo methods [38]. In this section, we prove the strong convergence of the approximation scheme defined in (4.13). Hence, we examine first the exponential integrability of the square-root process discretization and then the finiteness of moments of order higher than 1 of the exchange rate process and its approximation.

4.3.1. The square-root process

Let y be the square-root process defined in (4.4) and let $\bar{y}$ be the piecewise constant FTE interpolant from (4.9). The exponential integrability of functionals of the two processes was discussed in Propositions 2.2 and 2.6. However, we need to adjust the second result for our approximation scheme (4.13) in order to establish the convergence.

Lemma 4.1. *Let $T \geq 0$ and $\lambda, \mu \in \mathbb{R}$ be given. Denote $\Delta \equiv \lambda + \frac{1}{2}\mu^2$ and define*

$$\bar{\Theta}_t \equiv \exp\left\{\lambda \int_0^t \bar{y}_u \, du + \mu \int_0^t \sqrt{\bar{y}_u} \, \frac{\delta W_u^y}{\delta t} \, du\right\}, \quad \forall t \in [0, T]. \tag{4.24}$$

If $\max\{0, \Delta\}(T^* - T) \geq 0$, then there exists $\delta_T > 0$ such that

$$\sup_{\delta t \in (0, \delta_T)} \sup_{t \in [0,T]} \mathbb{E}\left[\bar{\Theta}_t\right] < \infty, \tag{4.25}$$

where T^* is as given below.

(1) If $k_y \leq \xi_y(\mu + \frac{1}{2}\sqrt{2\Delta})$, then

$$T^* = \frac{1}{\xi_y(\mu + \sqrt{2\Delta}) - k_y}. \tag{4.26}$$

(2) If $k_y > \xi_y(\mu + \frac{1}{2}\sqrt{2\varDelta})$, then

$$T^* = \frac{2(k_y - \mu\xi_y)}{\xi_y^2 \varDelta}. \tag{4.27}$$

Proof. Let $\{\mathcal{G}_t^y, 0 \le t \le T\}$ be the natural filtration generated by W^y and consider the shorthand notation $\mathbb{E}_t^y\left[\,\cdot\,\right] = \mathbb{E}\left[\,\cdot\,|\mathcal{G}_t^y\right]$. If we assume that $t \in [t_n, t_{n+1}]$ and condition on $\mathcal{G}_{t_n}^y$, we get

$$\mathbb{E}_{t_n}^y\left[\bar{\Theta}_t\right] = \exp\left\{\lambda \int_0^{t_n} \bar{y}_u\, du + \mu \int_0^{t_n} \sqrt{\bar{y}_u}\, \frac{\delta W_u^y}{\delta t}\, du + \left[\lambda + \frac{t - t_n}{2\delta t}\mu^2\right](t - t_n)\, \bar{y}_{t_n}\right\}. \tag{4.28}$$

Upon noticing that

$$\sup_{w \in [0,1]} \left(\lambda w + \frac{1}{2}\mu^2 w^2\right) = \varDelta\, \mathbb{1}_{\varDelta > 0}\,, \tag{4.29}$$

we deduce that $\mathbb{E}_{t_n}^y\left[\bar{\Theta}_t\right] \le \mathbb{E}_{t_n}^y\left[\bar{\Theta}_{t_n}\right]$ if $\varDelta \le 0$ and $\mathbb{E}_{t_n}^y\left[\bar{\Theta}_t\right] \le \mathbb{E}_{t_n}^y\left[\bar{\Theta}_{t_{n+1}}\right]$ if $\varDelta > 0$. Moreover, since $\bar{y}$ is piecewise constant,

$$\int_0^{t_n} \sqrt{\bar{y}_u}\, \frac{\delta W_u^y}{\delta t}\, du \;=\; \int_0^{t_n} \sqrt{\bar{y}_u}\, dW_u^y, \quad \forall 0 \le n \le N. \tag{4.30}$$

Henceforth, we follow the argument of Proposition 2.6. ■

4.3.2. Moment bounds

For many stochastic volatility models, moments of order higher than 1 can blow-up in finite time [7]. This can cause significant problems in practice, for instance, when computing the arbitrage-free price of an option whose payoff function has superlinear growth. The same troublesome behaviour can be observed for the Euler–Maruyama approximation of some SDEs with superlinearly growing drift or diffusion coefficients, where moments diverge in finite time [54]. Next, we prove the boundedness of moments of the exchange rate process and its approximation.

At this point, we assume that $\rho_{vd} \neq \pm 1$ and that a_{13} is non-zero, i.e., $\rho_{sd} \neq \rho_{sv}\rho_{vd}$.

Proposition 4.2. *For $\alpha \ge 1$, define the two quantities*

$$\begin{aligned} q_0(\alpha) \equiv \frac{1}{2\alpha^2\xi^2 a_{13}^2} \Bigg\{ &\sqrt{\left[2\alpha\rho_{sv}\xi k + \alpha^2\xi^2(a_{11}^2 + a_{12}^2) - \alpha\xi^2\right]^2 + 4\alpha^2 a_{13}^2 \xi^2 k^2} \\ &- \left[2\alpha\rho_{sv}\xi k + \alpha^2\xi^2(a_{11}^2 + a_{12}^2) - \alpha\xi^2\right] \Bigg\}, \end{aligned} \tag{4.31}$$

$$q_1(\alpha) \equiv q_0(\alpha)\, \mathbb{1}_{\rho_{sv} \le 0} + \min\left\{q_0(\alpha),\, \frac{k}{\alpha\rho_{sv}\xi}\right\} \mathbb{1}_{\rho_{sv} > 0}\,. \tag{4.32}$$

If the conditions

$$k > \alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha - 1)}\,\xi \quad \textit{and} \quad \frac{k_d^2}{2\xi_d^2} > \frac{\alpha\, q_1(\alpha)}{q_1(\alpha) - 1} \tag{4.33}$$

on the model parameters are satisfied, then there exists $\alpha_1 > \alpha$ such that, for all $w \in [1, \alpha_1)$,

$$\sup_{t \in [0,T]} \mathbb{E}\left[S_t^w\right] < \infty. \tag{4.34}$$

Proof. We find it convenient to define a new stochastic process L by

$$L_t \equiv S_t \exp\left\{\int_0^t r_u^f\,du\right\} \tag{4.35}$$

$$= S_0 \exp\left\{\int_0^t \left(r_u^d - \frac{1}{2}v_u\right)du + \int_0^t \sqrt{v_u}\,dW_u^s\right\}. \tag{4.36}$$

Since $S_t \leq L_t$ for all $t \in [0,T]$, it suffices to prove the finiteness of the supremum over t of

$$\mathbb{E}\left[L_t^w\right] = S_0^w\,\mathbb{E}\left[\exp\left\{w\int_0^t r_u^d\,du - \frac{w}{2}\int_0^t v_u\,du + w\sum_{j=1}^4 a_{1j}\int_0^t \sqrt{v_u}\,dW_u^j\right\}\right]. \tag{4.37}$$

Let $\left\{\mathcal{G}_t^{d,v},\, 0 \leq t \leq T\right\}$ be the natural filtration generated by the Brownian drivers W^3 and W^4 – note that the processes r^d and v are adapted to this filtration – and $\left\{\mathcal{G}_t^v,\, 0 \leq t \leq T\right\}$ be the filtration generated by W^4. Conditioning the expectation on the right-hand side of (4.37) on the σ-algebra $\mathcal{G}_t^{d,v}$ and taking into account that W^1 and W^2 are independent, we can compute the inner expectation using moment generating functions (MGFs):

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{wa_{11}\int_0^t \sqrt{v_u}\,dW_u^1 + wa_{12}\int_0^t \sqrt{v_u}\,dW_u^2\right\}\Big|\,\mathcal{G}_t^{d,v}\right] \\
&\quad= \mathbb{E}\left[\exp\left\{wa_{11}\int_0^t \sqrt{v_u}\,dW_u^1\right\}\Big|\,\mathcal{G}_t^{d,v}\right]\mathbb{E}\left[\exp\left\{wa_{12}\int_0^t \sqrt{v_u}\,dW_u^2\right\}\Big|\,\mathcal{G}_t^{d,v}\right] \\
&\quad= \exp\left\{\frac{w^2}{2}\left(a_{11}^2 + a_{12}^2\right)\int_0^t v_u\,du\right\}.
\end{aligned} \tag{4.38}$$

Substituting back into (4.37) with (4.38) leads to

$$\begin{aligned}
\mathbb{E}\left[L_t^w\right] = S_0^w\,\mathbb{E}\Bigg[\exp\Bigg\{&\left[\frac{w^2}{2}\left(a_{11}^2 + a_{12}^2\right) - \frac{w}{2}\right]\int_0^t v_u\,du + w\sum_{j=3}^4 a_{1j}\int_0^t \sqrt{v_u}\,dW_u^j \\
&+ w\int_0^t r_u^d\,du\Bigg\}\Bigg].
\end{aligned} \tag{4.39}$$

Next, we employ Hölder's inequality with the pair (p,q), where $p,q > 1$ and $q = p/(p-1)$, and then condition one of the two resulting expectations on the σ-algebra $\mathcal{G}_t^v$ to arrive at

$$\begin{aligned}
\mathbb{E}\left[L_t^w\right] \leq S_0^w\,&\mathbb{E}\left[\exp\left\{q\left[\frac{w^2}{2}\left(a_{11}^2 + a_{12}^2 + qa_{13}^2\right) - \frac{w}{2}\right]\int_0^t v_u\,du + qwa_{14}\int_0^t \sqrt{v_u}\,dW_u^4\right\}\right]^{\frac{1}{q}} \\
&\times \mathbb{E}\left[\exp\left\{pw\int_0^t r_u^d\,du\right\}\right]^{\frac{1}{p}}.
\end{aligned} \tag{4.40}$$

All that is left to do is to show that the supremum over t of each of the two expectations on the right-hand side of (4.40) is finite. However, Proposition 2.2 provides the following sufficient conditions:

$$k_d \geq \sqrt{2pw}\,\xi_d\,, \tag{4.41}$$

as well as

$$k \geq qw\rho_{sv}\xi \tag{4.42}$$

and

$$w^2\xi^2 a_{13}^2 q^2 + \left[2w\rho_{sv}\xi k + w^2\xi^2\left(a_{11}^2 + a_{12}^2\right) - w\xi^2\right] q - k^2 \leq 0, \tag{4.43}$$

for all maturities $T > 0$. The first assumption in (4.33) ensures that $q_0(\alpha) > 1$, and hence that $q_1(\alpha) > 1$. This implies that $q_1(\alpha)/(q_1(\alpha) - 1) > 1$. Due to the second assumption in (4.33), we can find $p > 1$ such that

$$\frac{k_d^2}{2\xi_d^2} > \alpha p > \frac{\alpha\, q_1(\alpha)}{q_1(\alpha) - 1}, \tag{4.44}$$

and hence

$$k_d > \sqrt{2p\alpha}\,\xi_d. \tag{4.45}$$

The quadratic equation in x below has roots of different signs, with positive root $q_0(\alpha)$.

$$\alpha^2\xi^2 a_{13}^2 x^2 + \left[2\alpha\rho_{sv}\xi k + \alpha^2\xi^2\left(a_{11}^2 + a_{12}^2\right) - \alpha\xi^2\right]x - k^2 = 0. \tag{4.46}$$

However, the Hölder pair satisfies $p = q/(q-1)$, so $q < q_1(\alpha) \leq q_0(\alpha)$. Therefore, q lies in between the two roots of the quadratic, which implies that

$$\alpha^2\xi^2 a_{13}^2 q^2 + \left[2\alpha\rho_{sv}\xi k + \alpha^2\xi^2\left(a_{11}^2 + a_{12}^2\right) - \alpha\xi^2\right]q - k^2 < 0. \tag{4.47}$$

From (4.32), if $\rho_{sv} > 0$,

$$q < q_1(\alpha) \leq \frac{k}{\alpha\rho_{sv}\xi}, \tag{4.48}$$

and hence

$$k > q\alpha\rho_{sv}\xi, \tag{4.49}$$

and this clearly holds when the correlation coefficient is non-positive. Consider the three continuous maps below, which are strictly positive when $w = \alpha$,

$$\begin{cases} w \mapsto k - qw\rho_{sv}\xi; \quad w \mapsto k_d - \sqrt{2pw}\,\xi_d; \\ w \mapsto k^2 - w^2\xi^2 a_{13}^2 q^2 - \left[2w\rho_{sv}\xi k + w^2\xi^2\left(a_{11}^2 + a_{12}^2\right) - w\xi^2\right]q. \end{cases} \tag{4.50}$$

Then we can find $\alpha_1 > \alpha$ such that all three functions are positive on $[\alpha, \alpha_1)$. We have thus proved that conditions (4.41) – (4.43) are satisfied and the conclusion follows. The extension to the interval $[1, \alpha_1)$ follows immediately from Jensen's inequality. ■

In the special case that $a_{13} = 0$, i.e., $\rho_{sd} = \rho_{sv}\rho_{vd}$, the argument is the same and the only difference appears in condition (4.43), which becomes

$$\left[2w\rho_{sv}\xi k + w^2\xi^2\left(1 - \rho_{sv}^2\right) - w\xi^2\right]q - k^2 \leq 0. \tag{4.51}$$

Henceforth, one can easily show that Proposition 4.2 still holds in this case as long as

$$k > \alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi \tag{4.52}$$

and

$$\frac{k_d^2}{2\xi_d^2} > \alpha \max\left\{1,\ \frac{k}{k - \alpha\rho_{sv}\xi},\ \frac{k^2}{(k - \alpha\rho_{sv}\xi)^2 - \alpha(\alpha-1)\xi^2}\right\}. \tag{4.53}$$

Next, assume that a_{13} and a_{14} are not simultaneously zero, i.e., $\rho_{sv}^2 + \rho_{sd}^2 \neq 0$.

Proposition 4.3. *For $\alpha \geq 1$, define*

$$q_2(\alpha) \equiv \left\{ \sqrt{\left[\alpha\rho_{sv}\xi + \tfrac{1}{4}T\alpha^2\xi^2(a_{11}^2 + a_{12}^2) - \tfrac{1}{4}T\alpha\xi^2\right]^2 + T\alpha^2\xi^2(a_{13}^2 + a_{14}^2)k} \right.$$
$$\left. - \left[\alpha\rho_{sv}\xi + \tfrac{1}{4}T\alpha^2\xi^2(a_{11}^2 + a_{12}^2) - \tfrac{1}{4}T\alpha\xi^2\right] \right\} \frac{2}{T\alpha^2\xi^2(a_{13}^2 + a_{14}^2)}. \tag{4.54}$$

If the conditions

$$k > \alpha\rho_{sv}\xi + \frac{1}{4}\alpha(\alpha-1)T\xi^2 \quad \textit{and} \quad \frac{2k_d}{T\xi_d^2} > \frac{\alpha\, q_2(\alpha)}{q_2(\alpha)-1} \tag{4.55}$$

on the model parameters are satisfied, then there exists $\alpha_2 > \alpha$ such that, for all $w \in [1, \alpha_2)$, we can find $\delta_{w,T} > 0$ such that

$$\sup_{\delta t \in (0, \delta_{w,T})} \sup_{t \in [0,T]} \mathbb{E}\left[\bar{S}_t^w\right] < \infty. \tag{4.56}$$

Proof. For convenience, define a new stochastic process $\bar{L}$ by

$$\bar{L}_t \equiv \bar{S}_t \exp\left\{\int_0^t \bar{r}_u^f \, du\right\} \tag{4.57}$$
$$= S_0 \exp\left\{ \int_0^t \left(\bar{r}_u^d - \frac{1}{2}\bar{v}_u\right) du + a_{11}\int_0^t \sqrt{\bar{v}_u}\, dW_u^1 + \sum_{j=2}^4 a_{1j}\int_0^t \sqrt{\bar{v}_u}\, \frac{\delta W_u^j}{\delta t}\, du \right\}. \tag{4.58}$$

Since $\bar{S}_t \leq \bar{L}_t$ for all $t \in [0, T]$, it suffices to prove the finiteness of the supremum over t and δt of

$$\mathbb{E}\left[\bar{L}_t^w\right] = S_0^w\, \mathbb{E}\left[\exp\left\{ -\frac{w}{2}\int_0^t \bar{v}_u\, du + w a_{11}\int_0^t \sqrt{\bar{v}_u}\, dW_u^1 + w\sum_{j=2}^4 a_{1j}\int_0^t \sqrt{\bar{v}_u}\, \frac{\delta W_u^j}{\delta t}\, du \right.\right.$$
$$\left.\left. + w\int_0^t \bar{r}_u^d\, du \right\}\right]. \tag{4.59}$$

Conditioning the expectation on the right-hand side on $\mathcal{G}_T^{d,v}$ and bearing in mind that W^1 and W^2 are independent, we can split the inner expectation into two parts and use MGFs to compute them. First,

$$\mathbb{E}\left[\exp\left\{ w a_{11}\int_0^t \sqrt{\bar{v}_u}\, dW_u^1 \right\} \Big| \, \mathcal{G}_T^{d,v} \right] = \exp\left\{ \frac{w^2}{2} a_{11}^2 \int_0^t \bar{v}_u\, du \right\}. \tag{4.60}$$

Second, let $t \in [t_n, t_{n+1})$. As $\bar{v}$ is piecewise constant and W^2 has independent increments,

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{ w a_{12}\int_0^t \sqrt{\bar{v}_u}\, \frac{\delta W_u^2}{\delta t}\, du \right\} \Big| \, \mathcal{G}_T^{d,v} \right] \\
&= \mathbb{E}\left[\exp\left\{ w a_{12}\int_0^{t_n} \sqrt{\bar{v}_u}\, dW_u^2 + w a_{12}\frac{t-t_n}{\delta t}\int_{t_n}^{t_{n+1}} \sqrt{\bar{v}_u}\, dW_u^2 \right\} \Big| \, \mathcal{G}_T^{d,v} \right] \\
&= \exp\left\{ \frac{w^2}{2} a_{12}^2 \int_0^{t_n} \bar{v}_u\, du \right\} \exp\left\{ \frac{w^2}{2} a_{12}^2 \frac{(t-t_n)^2}{(\delta t)^2} \int_{t_n}^{t_{n+1}} \bar{v}_u\, du \right\} \\
&\leq \exp\left\{ \frac{w^2}{2} a_{12}^2 \int_0^t \bar{v}_u\, du \right\}.
\end{aligned} \tag{4.61}$$

Substituting back into (4.59) with (4.60) and (4.61) leads to the following upper bound:

$$\mathbb{E}\left[\bar{L}_t^w\right] \leq S_0^w\,\mathbb{E}\left[\exp\left\{w\sum_{j=3}^{4}a_{1j}\int_0^t\sqrt{\bar{v}_u}\,\frac{\delta W_u^j}{\delta t}\,du + \left[\frac{w^2}{2}\big(a_{11}^2+a_{12}^2\big)-\frac{w}{2}\right]\int_0^t\bar{v}_u\,du\right.\right.$$
$$\left.\left.+\,w\int_0^t\bar{r}_u^d\,du\right\}\right]. \tag{4.62}$$

Next, we employ Hölder's inequality with the pair (p,q), where $p,q>1$ and $q=p/(p-1)$, and then condition one of the two resulting expectations on the σ-algebra $\mathcal{G}_T^v$ and proceed as in (4.61) to arrive at

$$\mathbb{E}\big[\bar{L}_t^w\big] \leq S_0^w\,\mathbb{E}\Bigg[\exp\Bigg\{qwa_{14}\int_0^t\sqrt{\bar{v}_u}\,\frac{\delta W_u^4}{\delta t}\,du + q\left[\frac{w^2}{2}\big(a_{11}^2+a_{12}^2+qa_{13}^2\big)-\frac{w}{2}\right]\int_0^t\bar{v}_u\,du\Bigg\}\Bigg]^{\frac{1}{q}}$$
$$\times\,\mathbb{E}\left[\exp\left\{pw\int_0^t\bar{r}_u^d\,du\right\}\right]^{\frac{1}{p}}. \tag{4.63}$$

All that we have left to do is to show that the supremum over t and δt of each of the two expectations on the right-hand side of (4.63) is finite. However, one can easily deduce from Lemma 4.1 the following sufficient conditions:

$$k_d \geq \frac{1}{2}pwT\xi_d^2 \tag{4.64}$$

and

$$k \geq qw\rho_{sv}\xi + \frac{1}{2}\mathit{\Delta}T\xi^2, \tag{4.65}$$

where

$$\mathit{\Delta} = q\left[\frac{1}{2}w(w-1)+\frac{1}{2}w^2(q-1)\big(a_{13}^2+a_{14}^2\big)\right]. \tag{4.66}$$

As an aside, note that we used $\sum_{j=1}^4 a_{1j}^2 = 1$ to rewrite $\mathit{\Delta}$ as in (4.66). The first assumption in (4.55) ensures that $q_2(\alpha)>1$. This, in turn, implies that $q_2(\alpha)/(q_2(\alpha)-1)>1$. Due to the second assumption in (4.55), we can find $p>1$ such that

$$\frac{2k_d}{T\xi_d^2} > \alpha p > \frac{\alpha\,q_2(\alpha)}{q_2(\alpha)-1}, \tag{4.67}$$

and hence

$$k_d > \frac{1}{2}p\alpha T\xi_d^2. \tag{4.68}$$

The quadratic equation in x below has roots of different signs, with positive root $q_2(\alpha)$.

$$\frac{1}{4}T\alpha^2\xi^2\big(a_{13}^2+a_{14}^2\big)x^2 + \left[\alpha\rho_{sv}\xi+\frac{1}{4}T\alpha^2\xi^2\big(a_{11}^2+a_{12}^2\big)-\frac{1}{4}T\alpha\xi^2\right]x - k = 0. \tag{4.69}$$

However, the Hölder pair satisfies $p=q/(q-1)$, so $q<q_2(\alpha)$. Hence, q lies in between the two roots of the quadratic, which implies that

$$\frac{1}{4}T\alpha^2\xi^2\big(a_{13}^2+a_{14}^2\big)q^2 + \left[\alpha\rho_{sv}\xi+\frac{1}{4}T\alpha^2\xi^2\big(a_{11}^2+a_{12}^2\big)-\frac{1}{4}T\alpha\xi^2\right]q - k < 0. \tag{4.70}$$

Rearranging terms in the above inequality, we obtain

$$k > q\alpha\rho_{sv}\xi + \frac{1}{2}T\xi^2 q\left[\frac{1}{2}\alpha(\alpha-1) + \frac{1}{2}\alpha^2(q-1)\big(a_{13}^2 + a_{14}^2\big)\right]. \tag{4.71}$$

From (4.68) and (4.71), employing a continuity argument similar to that used in the proof of Proposition 4.2, we deduce that the two conditions in (4.64) – (4.65) hold on an interval $[\alpha, \alpha_2)$, for some $\alpha_2 > \alpha$, which concludes the proof. The extension to the interval $[1, \alpha_2)$ follows from Jensen's inequality, with $\delta_{w,T} = \delta_{\alpha,T}$ for all $w \in [1, \alpha]$. ■

In the event that a_{13} and a_{14} are simultaneously zero, i.e., $\rho_{sv} = \rho_{sd} = 0$, one can easily show that Proposition 4.3 still holds as long as

$$k > \frac{1}{4}\alpha(\alpha-1)T\xi^2 \tag{4.72}$$

and

$$\frac{k_d}{T\xi_d^2} > \frac{2\alpha k}{4k - \alpha(\alpha-1)T\xi^2}. \tag{4.73}$$

Since the most popular FX and equity contracts grow at most linearly in FX and asset prices, and their valuation requires the computation of the expected discounted payoff under the risk-neutral measure, it is useful to study finiteness of moments under discounting. Let R be the discounted exchange rate process,

$$R_t = S_0 \exp\left\{-\int_0^t \left(r_u^f + \frac{1}{2}v_u\right)du + \int_0^t \sqrt{v_u}\, dW_u^s\right\}, \tag{4.74}$$

and let $\bar{R}$ be its continuous-time approximation,

$$\bar{R}_t = S_0 \exp\left\{-\int_0^t \left(\bar{r}_u^f + \frac{1}{2}\bar{v}_u\right)du + a_{11}\int_0^t \sqrt{\bar{v}_u}\, dW_u^1 + \sum_{j=2}^{4} a_{1j}\int_0^t \sqrt{\bar{v}_u}\,\frac{\delta W_u^j}{\delta t}\, du\right\}. \tag{4.75}$$

Proposition 4.4. *Let $\alpha \geq 1$. If $T < T^*$, then there exists $\alpha_1 > \alpha$ such that, for all $w \in [1, \alpha_1)$,*

$$\sup_{t\in[0,T]} \mathbb{E}\left[R_t^w\right] < \infty. \tag{4.76}$$

If $\alpha > 1$ and $T \geq T^*$, then

$$\mathbb{E}\left[R_T^\alpha\right] = \infty. \tag{4.77}$$

If $\alpha = 1$, then $T^* = \infty$, whereas if $\alpha > 1$, then T^* is as given below.

(1) If $k < \alpha\rho_{sv}\xi - \sqrt{\alpha(\alpha-1)}\,\xi$, then

$$T^* = \frac{1}{\vartheta(\alpha)}\log\left(\frac{\alpha\rho_{sv}\xi - k + \vartheta(\alpha)}{\alpha\rho_{sv}\xi - k - \vartheta(\alpha)}\right), \tag{4.78}$$

where

$$\vartheta(\alpha) = \sqrt{(\alpha\rho_{sv}\xi - k)^2 - \alpha(\alpha-1)\xi^2}. \tag{4.79}$$

(2) If $k = \alpha\rho_{sv}\xi - \sqrt{\alpha(\alpha-1)}\,\xi$, then

$$T^* = \frac{2}{\alpha\rho_{sv}\xi - k}. \tag{4.80}$$

(3) If $\alpha\rho_{sv}\xi - \sqrt{\alpha(\alpha-1)}\,\xi < k < \alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi$, then

$$T^* = \frac{2}{\hat{\vartheta}(\alpha)}\left[\frac{\pi}{2} - \arctan\left(\frac{\alpha\rho_{sv}\xi - k}{\hat{\vartheta}(\alpha)}\right)\right], \tag{4.81}$$

where

$$\hat{\vartheta}(\alpha) = \sqrt{\alpha(\alpha-1)\xi^2 - (\alpha\rho_{sv}\xi - k)^2}. \tag{4.82}$$

(4) If $k \geq \alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi$, then

$$T^* = \infty. \tag{4.83}$$

Proof. We follow the argument of Proposition 4.2 closely and condition on the σ-algebra $\mathcal{G}_t^v$ instead to deduce that

$$\mathbb{E}\left[R_t^w\right] \leq S_0^w\,\mathbb{E}\left[\exp\left\{\left[\frac{1}{2}w^2\big(1-\rho_{sv}^2\big) - \frac{1}{2}w\right]\int_0^t v_u\,du + w\rho_{sv}\int_0^t \sqrt{v_u}\,dW_u^4\right\}\right]. \tag{4.84}$$

First, suppose that $\alpha = 1$ and $T \geq 0$. If $k < \rho_{sv}\xi$, then

$$\lim_{w\downarrow 1^+} \frac{1}{\vartheta(w)}\log\left(\frac{w\rho_{sv}\xi - k + \vartheta(w)}{w\rho_{sv}\xi - k - \vartheta(w)}\right) = \infty. \tag{4.85}$$

Hence, by a continuity argument, we can find $\alpha_1 > 1$ such that, for all $w \in (1, \alpha_1)$,

$$k < w\rho_{sv}\xi - \sqrt{w(w-1)}\,\xi \quad \text{and} \quad T < \frac{1}{\vartheta(w)}\log\left(\frac{w\rho_{sv}\xi - k + \vartheta(w)}{w\rho_{sv}\xi - k - \vartheta(w)}\right). \tag{4.86}$$

If $k = \rho_{sv}\xi$, then $\rho_{sv} \in (0,1]$ and

$$\begin{aligned}\lim_{w\downarrow 1^+} \frac{2}{\hat{\vartheta}(w)}\left[\frac{\pi}{2} - \arctan\left(\frac{w\rho_{sv}\xi - k}{\hat{\vartheta}(w)}\right)\right] &= \lim_{w\downarrow 1^+} \frac{2}{\hat{\vartheta}(w)}\arctan\left(\frac{\sqrt{w-(w-1)\rho_{sv}^2}}{\sqrt{w-1}\,\rho_{sv}}\right) \\ &= \infty. \end{aligned} \tag{4.87}$$

Furthermore, note that for all $w > 1$,

$$w\rho_{sv}\xi - \sqrt{w(w-1)}\,\xi \leq wk - \sqrt{w(w-1)}\,k < k. \tag{4.88}$$

Hence, we can find $\alpha_1 > 1$ such that, for all $w \in (1, \alpha_1)$,

$$w\rho_{sv}\xi - \sqrt{w(w-1)}\,\xi < k < w\rho_{sv}\xi + \sqrt{w(w-1)}\,\xi \tag{4.89}$$

and

$$T < \frac{2}{\hat{\vartheta}(w)}\left[\frac{\pi}{2} - \arctan\left(\frac{w\rho_{sv}\xi - k}{\hat{\vartheta}(w)}\right)\right]. \tag{4.90}$$

If $k > \rho_{sv}\xi$, then we can find $\alpha_1 > 1$ such that, for all $w \in (1, \alpha_1)$,

$$k > w\rho_{sv}\xi + \sqrt{w(w-1)}\,\xi. \tag{4.91}$$

The conclusion follows from Proposition 2.2 and (4.86) – (4.91). Next, suppose that $\alpha > 1$ and $T < T^*$, with T^* defined in (4.78) – (4.83).

If $k < \alpha\rho_{sv}\xi - \sqrt{\alpha(\alpha-1)}\,\xi$, by a continuity argument, we can find $\alpha_1 > \alpha$ such that, for all $w \in (\alpha, \alpha_1)$,

$$k < w\rho_{sv}\xi - \sqrt{w(w-1)}\,\xi \quad \text{and} \quad T < \frac{1}{\vartheta(w)} \log\left(\frac{w\rho_{sv}\xi - k + \vartheta(w)}{w\rho_{sv}\xi - k - \vartheta(w)}\right). \tag{4.92}$$

If $k = \alpha\rho_{sv}\xi - \sqrt{\alpha(\alpha-1)}\,\xi$, then $\rho_{sv} \in (0, 1]$ and, for all $w > \alpha$,

$$w - \alpha < \sqrt{w(w-1)} - \sqrt{\alpha(\alpha-1)}, \tag{4.93}$$

so

$$w\rho_{sv}\xi - \sqrt{w(w-1)}\,\xi < \alpha\rho_{sv}\xi - \sqrt{\alpha(\alpha-1)}\,\xi. \tag{4.94}$$

Furthermore, note that

$$\begin{aligned}\lim_{w\downarrow\alpha^+} \frac{2}{\hat{\vartheta}(w)}\left[\frac{\pi}{2} - \arctan\left(\frac{w\rho_{sv}\xi - k}{\hat{\vartheta}(w)}\right)\right] &= \lim_{w\downarrow\alpha^+} \frac{2}{\hat{\vartheta}(w)} \arctan\left(\frac{\hat{\vartheta}(w)}{\alpha\rho_{sv}\xi - k}\right) \\ &= \frac{2}{\alpha\rho_{sv}\xi - k}.\end{aligned} \tag{4.95}$$

Hence, we can find $\alpha_1 > \alpha$ such that, for all $w \in (\alpha, \alpha_1)$,

$$w\rho_{sv}\xi - \sqrt{w(w-1)}\,\xi < k < w\rho_{sv}\xi + \sqrt{w(w-1)}\,\xi \tag{4.96}$$

and

$$T < \frac{2}{\hat{\vartheta}(w)}\left[\frac{\pi}{2} - \arctan\left(\frac{w\rho_{sv}\xi - k}{\hat{\vartheta}(w)}\right)\right]. \tag{4.97}$$

If $\alpha\rho_{sv}\xi - \sqrt{\alpha(\alpha-1)}\,\xi < k < \alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi$, we can clearly find $\alpha_1 > \alpha$ such that both (4.96) and (4.97) hold for all $w \in (\alpha, \alpha_1)$. If $k = \alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi$, then for all $w > \alpha$,

$$(\alpha - w)\rho_{sv} < \sqrt{w(w-1)} - \sqrt{\alpha(\alpha-1)}, \tag{4.98}$$

so

$$\alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi < w\rho_{sv}\xi + \sqrt{w(w-1)}\,\xi. \tag{4.99}$$

Furthermore, note that

$$\begin{aligned}\lim_{w\downarrow\alpha^+} \frac{2}{\hat{\vartheta}(w)}\left[\frac{\pi}{2} - \arctan\left(\frac{w\rho_{sv}\xi - k}{\hat{\vartheta}(w)}\right)\right] &= \lim_{w\downarrow\alpha^+} \frac{2\pi}{\hat{\vartheta}(w)} \\ &= \infty.\end{aligned} \tag{4.100}$$

Therefore, we can find $\alpha_1 > \alpha$ such that both (4.96) and (4.97) hold for all $w \in (\alpha, \alpha_1)$. Finally, if $k > \alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi$, then we can find $\alpha_1 > \alpha$ such that, for all $w \in (\alpha, \alpha_1)$,

$$k > w\rho_{sv}\xi + \sqrt{w(w-1)}\,\xi. \tag{4.101}$$

The conclusion follows from Proposition 2.2 and (4.92) – (4.101). The extension to the interval $[1, \alpha_1)$ follows from Jensen's inequality. ■

When the domestic and foreign interest rates are constant, Proposition 4.4 examines moment boundedness in the Heston model. It is an extension of Proposition 3.1 in [7] from bounds on moments of order $\alpha > 1$ to bounds on all moments of order $w \in [\alpha, \alpha_1)$, for some $\alpha_1 > \alpha \geq 1$. We employ this result to prove the strong convergence of the discretized discounted spot FX rate process.

Proposition 4.5. *Let $\alpha \geq 1$. If $T < T^*$, then there exist $\alpha_2 > \alpha$ and $\delta_{w,T} > 0$ such that, for all $w \in [1, \alpha_2)$,*

$$\sup_{\delta t \in (0, \delta_{w,T})} \sup_{t \in [0,T]} \mathbb{E}\left[\bar{R}_t^w\right] < \infty, \tag{4.102}$$

where T^* is as given below.

(1) If $k < \alpha\rho_{sv}\xi + \frac{1}{2}\sqrt{\alpha(\alpha-1)}\,\xi$, then

$$T^* = \frac{1}{\alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi - k}. \tag{4.103}$$

(2) If $k \geq \alpha\rho_{sv}\xi + \frac{1}{2}\sqrt{\alpha(\alpha-1)}\,\xi$, then

$$T^* = \begin{cases} \infty & , \text{ if } \alpha = 1 \\ \dfrac{4(k - \alpha\rho_{sv}\xi)}{\alpha(\alpha-1)\xi^2} & , \text{ if } \alpha > 1. \end{cases} \tag{4.104}$$

Proof. We follow the argument of Proposition 4.3 closely and condition on the σ-algebra $\mathcal{G}_T^v$ instead to deduce that

$$\mathbb{E}\left[\bar{R}_t^w\right] \leq S_0^w\, \mathbb{E}\left[\exp\left\{\left[\frac{1}{2}w^2(1-\rho_{sv}^2) - \frac{1}{2}w\right]\int_0^t \bar{v}_u\,du + w\rho_{sv}\int_0^t \sqrt{\bar{v}_u}\,\frac{\delta W_u^4}{\delta t}\,du\right\}\right]. \tag{4.105}$$

Suppose that $T < T^*$, with T^* from (4.103) – (4.104). If $k < \alpha\rho_{sv}\xi + \frac{1}{2}\sqrt{\alpha(\alpha-1)}\,\xi$, then by a continuity argument, we can find $\alpha_2 > \alpha$ such that, for all $w \in (\alpha, \alpha_2)$,

$$k < w\rho_{sv}\xi + \frac{1}{2}\sqrt{w(w-1)}\,\xi \quad \text{and} \quad T < \frac{1}{w\rho_{sv}\xi + \sqrt{w(w-1)}\,\xi - k}. \tag{4.106}$$

If $k = \alpha\rho_{sv}\xi + \frac{1}{2}\sqrt{\alpha(\alpha-1)}\,\xi$, since k, $\xi > 0$ and, for all $w > \alpha$, we have

$$\rho_{sv} > -\frac{1}{2}\sqrt{1 - \frac{1}{\alpha}} > -\frac{1}{2}, \tag{4.107}$$

then

$$\alpha\rho_{sv}\xi + \frac{1}{2}\sqrt{\alpha(\alpha-1)}\,\xi < w\rho_{sv}\xi + \frac{1}{2}\sqrt{w(w-1)}\,\xi. \tag{4.108}$$

Furthermore, note that

$$\frac{4(k-\alpha\rho_{sv}\xi)}{\alpha(\alpha-1)\xi^2} = \frac{1}{\alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi - k} \tag{4.109}$$

and

$$\lim_{w\downarrow\alpha^+} \frac{1}{w\rho_{sv}\xi + \sqrt{w(w-1)}\,\xi - k} = T^*, \tag{4.110}$$

with T^* from (4.104). Hence, we can find $\alpha_2 > \alpha$ such that (4.106) holds for all $w \in (\alpha, \alpha_2)$. Finally, if $k > \alpha\rho_{sv}\xi + \frac{1}{2}\sqrt{\alpha(\alpha-1)}\,\xi$, since

$$\lim_{w\downarrow\alpha^+} \frac{4(k-w\rho_{sv}\xi)}{w(w-1)\xi^2} = T^*, \tag{4.111}$$

with T^* from (4.104), we can find $\alpha_2 > \alpha$ such that, for all $w \in (\alpha, \alpha_2)$,

$$k > w\rho_{sv}\xi + \frac{1}{2}\sqrt{w(w-1)}\,\xi \quad \text{and} \quad T < \frac{4(k-w\rho_{sv}\xi)}{w(w-1)\xi^2}. \tag{4.112}$$

The conclusion follows from Lemma 4.1 and (4.106) – (4.112). The extension to the interval $[1,\alpha_2)$ follows from Jensen's inequality, with $\delta_{w,T} = \delta_{\alpha,T}$ for all $w \in [1,\alpha]$. ■

Note that for the Heston model, Proposition 4.5 can be seen as an improvement of Proposition 3.16 due to the sharper conditions on the critical time.

4.3.3. The 4-dimensional system

The strong mean square convergence of the discretized variance and domestic interest rate processes was deduced in Proposition 3.10. Before deriving an equivalent result for the foreign interest rate, we need to prove a couple of auxiliary lemmas. The following auxiliary result proves the almost sure positivity of the foreign interest rate. The condition below ensures that the process r^f does not hit zero and allows us to control the potential growth of the absolute difference between the original and the discretized processes that comes from the sublinear correction term in the drift.

Lemma 4.6. *Let $\kappa > (r_0^f)^{-1}$ and define the stopping time*

$$\tau_\kappa = \inf\left\{t \ge 0 : \, r_t^f \le \kappa^{-1}\right\}. \tag{4.113}$$

If $2k_f\theta_f > \xi_f^2$, then

$$\lim_{\kappa\to\infty} \mathbb{P}\left(\tau_\kappa \le T\right) = 0. \tag{4.114}$$

Proof. The proof follows an argument from Proposition 3.14 closely, but is included for completeness. Define the function $F:(0,\infty)\to\mathbb{R}$ by

$$F(x)=x^{-\alpha},\quad \alpha=\frac{1}{2\xi_f^2}\big(2k_f\theta_f-\xi_f^2\big). \tag{4.115}$$

By Itô's formula, we have

$$\begin{aligned}\mathbb{E}\Big[F\big(r^f_{T\wedge\tau_\kappa}\big)\Big]&=F\big(r^f_0\big)-\mathbb{E}\int_0^{T\wedge\tau_\kappa}\alpha\big(r^f_s\big)^{-(1+\alpha)}\big(k_f\theta_f-k_fr^f_s-\rho_{sf}\xi_f\sqrt{v_sr^f_s}\big)ds\\&+\frac{1}{2}\,\mathbb{E}\int_0^{T\wedge\tau_\kappa}\alpha(1+\alpha)\xi_f^2\big(r^f_s\big)^{-(1+\alpha)}ds-\mathbb{E}\int_0^{T\wedge\tau_\kappa}\alpha\xi_f\big(r^f_s\big)^{-(\frac{1}{2}+\alpha)}dW^f_s.\end{aligned} \tag{4.116}$$

However,

$$\mathbb{E}\int_0^T\alpha^2\xi_f^2\big(r^f_s\big)^{-(1+2\alpha)}\,\mathbb{1}_{s<\tau_\kappa}\,ds\le\alpha^2\xi_f^2\kappa^{1+2\alpha}T<\infty, \tag{4.117}$$

so the stochastic integral on the right-hand side of (4.116) is a true martingale. Hence,

$$\mathbb{E}\Big[F\big(r^f_{T\wedge\tau_\kappa}\big)\Big]\le F\big(r^f_0\big)-\mathbb{E}\int_0^{T\wedge\tau_\kappa}\Big(a\big(r^f_s\big)^{-(1+\alpha)}-b\big(r^f_s\big)^{-\alpha}-cv_s^{\frac{1}{2}}\big(r^f_s\big)^{-(\frac{1}{2}+\alpha)}\Big)ds, \tag{4.118}$$

where

$$a=\frac{1}{8\xi_f^2}\big(2k_f\theta_f-\xi_f^2\big)^2,\quad b=\frac{k_f}{2\xi_f^2}\big(2k_f\theta_f-\xi_f^2\big),\quad c=\frac{|\rho_{sf}|}{2\xi_f}\big(2k_f\theta_f-\xi_f^2\big). \tag{4.119}$$

Employing Fubini's theorem and Hölder's inequality in (4.118), we get

$$\begin{aligned}\mathbb{E}\Big[F\big(r^f_{T\wedge\tau_\kappa}\big)\Big]&\le F\big(r^f_0\big)-\int_0^T\Bigg(a\,\mathbb{E}\Big[\big(r^f_s\big)^{-(1+\alpha)}\,\mathbb{1}_{s<\tau_\kappa}\Big]-b\,\mathbb{E}\Big[\big(r^f_s\big)^{-(1+\alpha)}\,\mathbb{1}_{s<\tau_\kappa}\Big]^{\frac{\alpha}{1+\alpha}}\\&-c\sup_{u\in[0,T]}\mathbb{E}\Big[v_u^{1+\alpha}\Big]^{\frac{1}{2(1+\alpha)}}\,\mathbb{E}\Big[\big(r^f_s\big)^{-(1+\alpha)}\,\mathbb{1}_{s<\tau_\kappa}\Big]^{\frac{1+2\alpha}{2(1+\alpha)}}\Bigg)ds.\end{aligned} \tag{4.120}$$

The moments of the square-root process are uniformly bounded from Proposition 3.8 and the function $f:[0,\infty)\to\mathbb{R}$ defined by

$$f(x)=ax-bx^{\frac{\alpha}{1+\alpha}}-c\sup_{u\in[0,T]}\mathbb{E}\big[v_u^{1+\alpha}\big]^{\frac{1}{2(1+\alpha)}}x^{\frac{1+2\alpha}{2(1+\alpha)}} \tag{4.121}$$

is bounded from below, see, e.g., (3.136). Hence, we can find a constant C independent of κ such that

$$\mathbb{E}\Big[F\big(r^f_{T\wedge\tau_\kappa}\big)\Big]\le C. \tag{4.122}$$

Since r^f has continuous paths, we have that $r^f_{\tau_\kappa}=\kappa^{-1}$ and $F(r^f_{\tau_\kappa})=\kappa^\alpha$. Therefore, using (4.122) and the fact that F is positive, we deduce that

$$\kappa^\alpha\,\mathbb{P}\left(\tau_\kappa\le T\right)=\mathbb{E}\Big[F\big(r^f_{T\wedge\tau_\kappa}\big)\,\mathbb{1}_{\tau_\kappa\le T}\Big]\le\mathbb{E}\Big[F\big(r^f_{T\wedge\tau_\kappa}\big)\Big]\le C. \tag{4.123}$$

Letting $\kappa\to\infty$ in (4.123) yields the conclusion. ∎

The next lemma establishes moment bounds for the original and the discretized foreign interest rate processes.

Lemma 4.7. *The following two statements hold.*

(1) The process r^f has uniformly bounded moments, i.e.,

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left(r_t^f\right)^p\right]<\infty,\quad \forall p\geq 1. \tag{4.124}$$

(2) The process $\hat{r}^f$ from (4.11) *has uniformly bounded moments, i.e.,*

$$\sup_{\delta t\in(0,\delta_T)}\mathbb{E}\left[\sup_{t\in[0,T]}\left(\hat{r}_t^f\right)^p\right]<\infty,\quad \forall p\geq 1,\ \forall\delta_T>0. \tag{4.125}$$

Proof. A similar result was derived in Proposition 3.9 for different dynamics of the foreign interest rate, and the proof is included for completeness.

(1) Fix any $p\geq 1$. From (4.1),

$$r_t^f=r_0^f+k_f\theta_f t-k_f\int_0^t r_u^f du-\rho_{sf}\xi_f\int_0^t\sqrt{v_u r_u^f}\,du+\xi_f\int_0^t\sqrt{r_u^f}\,dW_u^f. \tag{4.126}$$

Using the fact that $2\sqrt{|ab|}\leq|a|+|b|$ and Hölder's inequality, we deduce that

$$\begin{aligned}\left(r_t^f\right)^p&\leq 2^{2(p-1)}\left(r_0^f+k_f\theta_f t\right)^p+2^{p-2}|\rho_{sf}|^p\xi_f^p\left(\int_0^t v_u du\right)^p\\&\quad+2^{p-2}\left(2k_f+|\rho_{sf}|\xi_f\right)^p\left(\int_0^t r_u^f du\right)^p\\&\quad+2^{2(p-1)}\xi_f^p\left|\int_0^t\sqrt{r_u^f}\,dW_u^f\right|^p.\end{aligned} \tag{4.127}$$

Fix $t\in[0,T]$. Using Hölder's inequality, we get

$$\begin{aligned}\sup_{s\in[0,t]}\left(r_s^f\right)^p&\leq 2^{2(p-1)}\left(r_0^f+k_f\theta_f T\right)^p+2^{p-2}|\rho_{sf}|^p\xi_f^pT^{p-1}\int_0^T v_u^p du\\&\quad+2^{p-2}\left(2k_f+|\rho_{sf}|\xi_f\right)^pT^{p-1}\int_0^t\left(r_u^f\right)^p du\\&\quad+2^{2(p-1)}\xi_f^p\sup_{s\in[0,t]}\left|\int_0^s\sqrt{r_u^f}\,dW_u^f\right|^p.\end{aligned} \tag{4.128}$$

From the Burkholder–Davis–Gundy inequality, we know that there exists a constant $C_p>0$ such that

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]}\left|\int_0^s\sqrt{r_u^f}\,dW_u^f\right|^p\right]&\leq C_p\,\mathbb{E}\left[\left(\int_0^t r_u^f du\right)^{p/2}\right]\\&\leq\frac{1}{2}C_p+\frac{1}{2}C_pT^{p-1}\,\mathbb{E}\left[\int_0^t\left(r_u^f\right)^p du\right].\end{aligned} \tag{4.129}$$

Taking expectations and employing Fubini's theorem in (4.128), we get

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]}\left(r_s^f\right)^p\right] &\leq 2^{2(p-1)}\left(r_0^f+k_f\theta_f T\right)^p+2^{p-2}|\rho_{sf}|^p\xi_f^p T^p\sup_{u\in[0,T]}\mathbb{E}\left[v_u^p\right]\\&\quad+2^{2p-3}\xi_f^p C_p+\Big(2^{p-2}\left(2k_f+|\rho_{sf}|\xi_f\right)^p T^{p-1}\\&\quad+2^{2p-3}\xi_f^p C_p T^{p-1}\Big)\int_0^t\mathbb{E}\left[\sup_{s\in[0,u]}\left(r_s^f\right)^p\right]du. \qquad (4.130)\end{aligned}$$

Applying Gronwall's inequality, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}\left(r_t^f\right)^p\right] &\leq \Big(2^{2(p-1)}\left(r_0^f+k_f\theta_f T\right)^p+2^{2p-3}\xi_f^p C_p+2^{p-2}|\rho_{sf}|^p\xi_f^p T^p\sup_{u\in[0,T]}\mathbb{E}\left[v_u^p\right]\Big)\\&\quad\times\exp\left\{2^{p-2}\left(2k_f+|\rho_{sf}|\xi_f\right)^p T^p+2^{2p-3}\xi_f^p C_p T^p\right\}. \qquad (4.131)\end{aligned}$$

The conclusion follows from Proposition 3.8.

(2) Fix any $p\geq 1$ and $\delta_T>0$. From (4.10),

$$\tilde{r}_t^f=r_0^f+k_f\theta_f t-k_f\int_0^t\bar{r}_u^f du-\rho_{sf}\xi_f\int_0^t\sqrt{\bar{v}_u\bar{r}_u^f}\,du+\xi_f\int_0^t\sqrt{\bar{r}_u^f}\,dW_u^f. \qquad (4.132)$$

Since $\hat{r}_t^f\leq|\tilde{r}_t^f|$, following the previous argument, we deduce that

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]}\left(\hat{r}_s^f\right)^p\right] &\leq 2^{2(p-1)}\left(r_0^f+k_f\theta_f T\right)^p+2^{2p-3}\xi_f^p C_p+2^{p-2}|\rho_{sf}|^p\xi_f^p T^p\sup_{u\in[0,T]}\mathbb{E}\left[\bar{v}_u^p\right]\\&\quad+\Big(2^{p-2}\left(2k_f+|\rho_{sf}|\xi_f\right)^p T^{p-1}+2^{2p-3}\xi_f^p C_p T^{p-1}\Big)\int_0^t\mathbb{E}\left[\left(\bar{r}_u^f\right)^p\right]du. \qquad (4.133)\end{aligned}$$

Since $\sup_{u\in[0,T]}\mathbb{E}\left[\bar{v}_u^p\right]\leq\sup_{u\in[0,T]}\mathbb{E}\left[\hat{v}_u^p\right]$, with $\hat{v}$ defined as in (4.8), and $\bar{r}_u^f\leq\sup_{s\in[0,u]}\hat{r}_s^f$, applying Gronwall's inequality, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}\left(\hat{r}_t^f\right)^p\right] &\leq \Big(2^{2(p-1)}\left(r_0^f+k_f\theta_f T\right)^p+2^{2p-3}\xi_f^p C_p+2^{p-2}|\rho_{sf}|^p\xi_f^p T^p\sup_{u\in[0,T]}\mathbb{E}\left[\hat{v}_u^p\right]\Big)\\&\quad\times\exp\left\{2^{p-2}\left(2k_f+|\rho_{sf}|\xi_f\right)^p T^p+2^{2p-3}\xi_f^p C_p T^p\right\}. \qquad (4.134)\end{aligned}$$

The conclusion follows from Proposition 3.8. ■

Next, we use Lemma 4.7 to prove the convergence of the L^2 difference between the two continuous-time discretizations.

Lemma 4.8. *The L^2 difference between $\hat{r}^f$ and $\bar{r}^f$ converges to zero with δt, i.e.,*

$$\lim_{\delta t\to 0}\sup_{t\in[0,T]}\mathbb{E}\left[\left(\hat{r}_t^f-\bar{r}_t^f\right)^2\right]=0. \qquad (4.135)$$

Proof. The proof follows an argument from Proposition 3.13 closely, but is included for completeness. Suppose that $t\in[t_n,t_{n+1})$. Since $|\hat{r}_t^f-\bar{r}_t^f|\leq|\tilde{r}_t^f-\tilde{r}_{t_n}^f|$ and from (4.10), we can bound the squared absolute difference from above as follows:

$$\left(\hat{r}_t^f-\bar{r}_t^f\right)^2\leq\frac{1}{4}\Big(2k_f\theta_f\delta t+|\rho_{sf}|\xi_f\delta t\,\hat{v}_{t_n}+\left(2k_f+|\rho_{sf}|\xi_f\right)\delta t\,\hat{r}_{t_n}^f$$

$$+ 2\xi_f \sqrt{\hat{r}^f_{t_n}} \left|W^f_t - W^f_{t_n}\right|\Big)^2$$

$$\leq 4k_f^2\theta_f^2(\delta t)^2 + |\rho_{sf}|^2\xi_f^2(\delta t)^2\hat{v}^2_{t_n} + \big(2k_f + |\rho_{sf}|\xi_f\big)^2(\delta t)^2\big(\hat{r}^f_{t_n}\big)^2$$

$$+ 4\xi_f^2\hat{r}^f_{t_n}\big(W^f_t - W^f_{t_n}\big)^2. \tag{4.136}$$

Therefore,

$$\sup_{t\in[0,T]} \mathbb{E}\left[\big(\hat{r}^f_t - \bar{r}^f_t\big)^2\right] \leq 4k_f^2\theta_f^2(\delta t)^2 + |\rho_{sf}|^2\xi_f^2(\delta t)^2 \sup_{t\in[0,T]} \mathbb{E}\left[\hat{v}^2_t\right] + 4\xi_f^2\delta t \sup_{0\leq n\leq N} \mathbb{E}\left[\hat{r}^f_{t_n}\right]$$

$$+ \big(2k_f + |\rho_{sf}|\xi_f\big)^2(\delta t)^2 \sup_{0\leq n\leq N} \mathbb{E}\left[\big(\hat{r}^f_{t_n}\big)^2\right]. \tag{4.137}$$

Using Proposition 3.8 and Lemma 4.7 concludes the proof. ■

The following lemma derives the strong mean square convergence of the stopped process.

Lemma 4.9. *Let $l > v_0$ and define the stopping times*

$$\tau_l = \inf\big\{t \geq 0 :\, v_t \geq l\big\} \quad \textit{and} \quad \tau = \tau_\kappa \wedge \tau_l, \tag{4.138}$$

with τ_κ defined in (4.113). *Then the stopped process converges uniformly in L^2, i.e.,*

$$\lim_{\delta t\to 0} \mathbb{E}\left[\sup_{t\in[0,T]} \big(r^f_{t\wedge\tau} - \hat{r}^f_{t\wedge\tau}\big)^2\right] = 0. \tag{4.139}$$

Proof. From (4.126) and (4.132), since $|r^f_t - \hat{r}^f_t| \leq |r^f_t - \tilde{r}^f_t|$, we have

$$\begin{aligned}
\left|r^f_{t\wedge\tau} - \hat{r}^f_{t\wedge\tau}\right| \leq \Bigg| &- k_f \int_0^{t\wedge\tau} \big(r^f_u - \hat{r}^f_u\big)du - k_f \int_0^{t\wedge\tau} \big(\hat{r}^f_u - \bar{r}^f_u\big)du \\
&+ \xi_f \int_0^{t\wedge\tau} \Big(\sqrt{r^f_u} - \sqrt{\hat{r}^f_u}\Big)dW^f_u + \xi_f \int_0^{t\wedge\tau} \Big(\sqrt{\hat{r}^f_u} - \sqrt{\bar{r}^f_u}\Big)dW^f_u \\
&- \rho_{sf}\xi_f \int_0^{t\wedge\tau} \sqrt{v_u}\Big(\sqrt{r^f_u} - \sqrt{\hat{r}^f_u}\Big)du - \rho_{sf}\xi_f \int_0^{t\wedge\tau} \sqrt{v_u}\Big(\sqrt{\hat{r}^f_u} - \sqrt{\bar{r}^f_u}\Big)du \\
&- \rho_{sf}\xi_f \int_0^{t\wedge\tau} \sqrt{\bar{r}^f_u}\Big(\sqrt{v_u} - \sqrt{\bar{v}_u}\Big)du\Bigg|.
\end{aligned} \tag{4.140}$$

Fix $t \in [0, T]$. Squaring both sides and using the Cauchy–Schwarz inequality, then taking expectations and using Doob's martingale inequality and Fubini's theorem, we get

$$\begin{aligned}
\mathbb{E}\Bigg[\sup_{s\in[0,t]} \big(r^f_{s\wedge\tau} - \hat{r}^f_{s\wedge\tau}\big)^2\Bigg] \leq\ & 7k_f^2T \int_0^t \mathbb{E}\left[\big(r^f_u - \hat{r}^f_u\big)^2 \mathbb{1}_{u<\tau}\right]du + 7k_f^2T \int_0^T \mathbb{E}\left[\big(\hat{r}^f_u - \bar{r}^f_u\big)^2\right]du \\
&+ 28\xi_f^2 \int_0^t \mathbb{E}\left[\Big(\sqrt{r^f_u} - \sqrt{\hat{r}^f_u}\Big)^2 \mathbb{1}_{u<\tau}\right]du + 28\xi_f^2 \int_0^T \mathbb{E}\Big[\big|\hat{r}^f_u - \bar{r}^f_u\big|\Big]du \\
&+ 7\rho_{sf}^2\xi_f^2T \int_0^t \mathbb{E}\left[v_u\Big(\sqrt{r^f_u} - \sqrt{\hat{r}^f_u}\Big)^2 \mathbb{1}_{u<\tau}\right]du \\
&+ 7\rho_{sf}^2\xi_f^2T \int_0^T \mathbb{E}\left[\bar{r}^f_u\big|v_u - \bar{v}_u\big|\right]du
\end{aligned}$$

$$+ 7\rho_{sf}^2\xi_f^2 T \int_0^T \mathbb{E}\left[v_u\left|\hat{r}_u^f - \bar{r}_u^f\right| \mathbb{1}_{u<\tau}\right] du. \tag{4.141}$$

However, we know that $|r_u^f - \hat{r}_u^f| \, \mathbb{1}_{u<\tau} \le |r_{u\wedge\tau}^f - \hat{r}_{u\wedge\tau}^f|$ and

$$\left(\sqrt{r_u^f} - \sqrt{\hat{r}_u^f}\right)^2 \mathbb{1}_{u<\tau} \le \left(\sqrt{r_{u\wedge\tau}^f} - \sqrt{\hat{r}_{u\wedge\tau}^f}\right)^2 \le \kappa\left(r_{u\wedge\tau}^f - \hat{r}_{u\wedge\tau}^f\right)^2. \tag{4.142}$$

Substituting back into (4.141) with (4.142), we arrive at the following inequality:

$$\begin{aligned}
\mathbb{E}\left[\sup_{s\in[0,t]} \left(r_{s\wedge\tau}^f - \hat{r}_{s\wedge\tau}^f\right)^2\right] &\le \left(7k_f^2 T + 28\xi_f^2\kappa + 7\rho_{sf}^2\xi_f^2 T l\kappa\right) \int_0^t \mathbb{E}\left[\sup_{s\in[0,u]} \left(r_{s\wedge\tau}^f - \hat{r}_{s\wedge\tau}^f\right)^2\right] du \\
&\quad + \left(28\xi_f^2 T + 7\rho_{sf}^2\xi_f^2 T^2 l\right) \sup_{u\in[0,T]} \mathbb{E}\left[\left|\hat{r}_u^f - \bar{r}_u^f\right|\right] \\
&\quad + 7\rho_{sf}^2\xi_f^2 T^2 \sup_{u\in[0,T]} \mathbb{E}\left[\left(\hat{r}_u^f\right)^2\right]^{\frac{1}{2}} \sup_{u\in[0,T]} \mathbb{E}\left[\left|v_u - \bar{v}_u\right|^2\right]^{\frac{1}{2}} \\
&\quad + 7k_f^2 T^2 \sup_{u\in[0,T]} \mathbb{E}\left[\left(\hat{r}_u^f - \bar{r}_u^f\right)^2\right].
\end{aligned} \tag{4.143}$$

The convergence to zero of the last three terms on the right-hand side of (4.143) follows from Proposition 3.10 and Lemmas 4.7 and 4.8. The conclusion follows from a simple application of Gronwall's inequality. ■

The following lemma derives the strong mean square convergence of $\hat{r}^f$.

Lemma 4.10. *If $2k_f\theta_f > \xi_f^2$, then the process $\hat{r}^f$ converges strongly in L^2, i.e.,*

$$\lim_{\delta t\to 0} \sup_{t\in[0,T]} \mathbb{E}\left[\left|r_t^f - \hat{r}_t^f\right|^2\right] = 0. \tag{4.144}$$

Proof. Fix $\kappa > (r_0^f)^{-1}$, $l > v_0$, and recall the definition of the stopping time τ from (4.138). Since r^f and $\hat{r}^f$ are non-negative,

$$\begin{aligned}
\sup_{t\in[0,T]} \mathbb{E}\left[\left|r_t^f - \hat{r}_t^f\right|^2\right] &\le \sup_{t\in[0,T]} \mathbb{E}\left[\left|r_t^f - \hat{r}_t^f\right|^2 \mathbb{1}_{\tau\le t}\right] + \sup_{t\in[0,T]} \mathbb{E}\left[\left|r_t^f - \hat{r}_t^f\right|^2 \mathbb{1}_{t<\tau}\right] \\
&\le \sup_{t\in[0,T]} \mathbb{E}\left[\left(r_t^f\right)^2 \mathbb{1}_{\tau\le T}\right] + \sup_{t\in[0,T]} \mathbb{E}\left[\left(\hat{r}_t^f\right)^2 \mathbb{1}_{\tau\le T}\right] \\
&\quad + \sup_{t\in[0,T]} \mathbb{E}\left[\left|r_{t\wedge\tau}^f - \hat{r}_{t\wedge\tau}^f\right|^2\right].
\end{aligned} \tag{4.145}$$

Since $\mathbb{1}_{\tau\le T} \le \mathbb{1}_{\tau_\kappa\le T} + \mathbb{1}_{\tau_l\le T}$ and applying the Cauchy–Schwarz inequality, we get

$$\begin{aligned}
\sup_{t\in[0,T]} \mathbb{E}\left[\left|r_t^f - \hat{r}_t^f\right|^2\right] &\le \left\{\sup_{t\in[0,T]} \mathbb{E}\left[\left(r_t^f\right)^4\right]^{\frac{1}{2}} + \sup_{t\in[0,T]} \mathbb{E}\left[\left(\hat{r}_t^f\right)^4\right]^{\frac{1}{2}}\right\} \mathbb{P}\left(\tau_\kappa \le T\right)^{\frac{1}{2}} \\
&\quad + \left\{\sup_{t\in[0,T]} \mathbb{E}\left[\left(r_t^f\right)^4\right]^{\frac{1}{2}} + \sup_{t\in[0,T]} \mathbb{E}\left[\left(\hat{r}_t^f\right)^4\right]^{\frac{1}{2}}\right\} \mathbb{P}\left(\tau_l \le T\right)^{\frac{1}{2}} \\
&\quad + \sup_{t\in[0,T]} \mathbb{E}\left[\left|r_{t\wedge\tau}^f - \hat{r}_{t\wedge\tau}^f\right|^2\right].
\end{aligned} \tag{4.146}$$

Using Markov's inequality, we find an upper bound

$$\mathbb{P}\left(\tau_l \leq T\right) \leq \mathbb{P}\left(\sup_{t\in[0,T]} v_t \geq l\right) \leq \frac{1}{l}\,\mathbb{E}\left[\sup_{t\in[0,T]} v_t\right]. \tag{4.147}$$

However, the expectation on the right-hand side is finite from Proposition 3.8. Taking the limit as $\delta t \to 0$ in (4.146) and employing Lemmas 4.6, 4.7 and 4.9, since κ and l can be made arbitrarily large, leads to the conclusion. ■

Proposition 4.11. *If $2k_f\theta_f > \xi_f^2$, then the process $\bar{r}^f$ converges strongly in L^2, i.e.,*

$$\lim_{\delta t\to 0}\sup_{t\in[0,T]} \mathbb{E}\left[\left|r_t^f - \bar{r}_t^f\right|^2\right] = 0. \tag{4.148}$$

Proof. Upon noticing that

$$\left|r_t^f - \bar{r}_t^f\right|^2 \leq 2\left|r_t^f - \hat{r}_t^f\right|^2 + 2\left|\hat{r}_t^f - \bar{r}_t^f\right|^2, \quad \forall t \in [0,T], \tag{4.149}$$

the conclusion follows immediately from Lemmas 4.8 and 4.10. ■

Next, we consider the logarithm of the process from (4.13) and examine its convergence properties. The formulae for the log-process, x, and its approximation, $\bar{x}$, are

$$x_t = x_0 + \int_0^t \left(r_u^d - r_u^f - \frac{1}{2}v_u\right)du + \int_0^t \sqrt{v_u}\,dW_u^s, \tag{4.150}$$

$$\bar{x}_t = x_0 + \int_0^t \left(\bar{r}_u^d - \bar{r}_u^f - \frac{1}{2}\bar{v}_u\right)du + a_{11}\int_0^t \sqrt{\bar{v}_u}\,dW_u^1 + \sum_{j=2}^4 a_{1j}\int_0^t \sqrt{\bar{v}_u}\,\frac{\delta W_u^j}{\delta t}\,du. \tag{4.151}$$

Proposition 4.12. *If $2k_f\theta_f > \xi_f^2$, then the log-process converges uniformly in L^2, i.e.,*

$$\lim_{\delta t\to 0}\mathbb{E}\left[\sup_{t\in[0,T]}\left|x_t - \bar{x}_t\right|^2\right] = 0. \tag{4.152}$$

Proof. Note that for all $t \in [t_n, t_{n+1})$ and for all $j \in \{2,3,4\}$, since $\bar{v}$ is piecewise constant,

$$\begin{aligned}\int_0^t \sqrt{\bar{v}_u}\,\frac{\delta W_u^j}{\delta t}\,du &= \sum_{i=0}^{n-1}\sqrt{\bar{v}_{t_i}}\,\frac{W_{t_{i+1}}^j - W_{t_i}^j}{\delta t}\,\delta t + \sqrt{\bar{v}_{t_n}}\,\frac{W_{t_{n+1}}^j - W_{t_n}^j}{\delta t}\,(t - t_n)\\ &= \int_0^t \sqrt{\bar{v}_u}\,dW_u^j + \sqrt{\bar{v}_t}\,Z_t^j,\end{aligned} \tag{4.153}$$

where we define

$$Z_t^j = \frac{t - t_n}{\delta t}\left(W_{t_{n+1}}^j - W_t^j\right) - \frac{t_{n+1} - t}{\delta t}\left(W_t^j - W_{t_n}^j\right). \tag{4.154}$$

Substituting back into (4.151) with (4.153), we obtain

$$\bar{x}_t = x_0 + \int_0^t \left(\bar{r}_u^d - \bar{r}_u^f - \frac{1}{2}\bar{v}_u\right)du + \int_0^t \sqrt{\bar{v}_u}\,dW_u^s + \sum_{j=2}^4 a_{1j}\sqrt{\bar{v}_t}\,Z_t^j. \tag{4.155}$$

The absolute difference between the original and the discretized log-processes is thus

$$\begin{aligned}\left|x_t-\bar{x}_t\right| = \bigg| \int_0^t \left(r_u^d-\bar{r}_u^d\right)du - \int_0^t \left(r_u^f-\bar{r}_u^f\right)du - \frac{1}{2}\int_0^t \left(v_u-\bar{v}_u\right)du \\ + \int_0^t \left(\sqrt{v_u}-\sqrt{\bar{v}_u}\right)dW_u^s - \sum_{j=2}^{4} a_{1j}\sqrt{\bar{v}_t}\, Z_t^j \bigg|. \end{aligned} \tag{4.156}$$

Squaring both sides of (4.156), applying the Cauchy–Schwarz inequality, taking the supremum over all $t \in [0,T]$ and using the Cauchy–Schwarz inequality for all Riemann integrals leads to

$$\begin{aligned}\sup_{t\in[0,T]}\left|x_t-\bar{x}_t\right|^2 \le 7T\int_0^T \left(r_u^d-\bar{r}_u^d\right)^2du + 7T\int_0^T \left(r_u^f-\bar{r}_u^f\right)^2du + \frac{7}{4}T\int_0^T \left(v_u-\bar{v}_u\right)^2du \\ + 7\sup_{t\in[0,T]}\left|\int_0^t \left(\sqrt{v_u}-\sqrt{\bar{v}_u}\right)dW_u^s\right|^2 + 7\sum_{j=2}^{4} a_{1j}^2 \sup_{t\in[0,T]}\hat{v}_t \sup_{t\in[0,T]}\left|Z_t^j\right|^2. \end{aligned} \tag{4.157}$$

We used the fact that $\sup_{t\in[0,T]}\bar{v}_t \le \sup_{t\in[0,T]}\hat{v}_t$, with $\hat{v}$ defined as in (4.8). Taking expectations and employing Fubini's theorem, Hölder's inequality, Doob's martingale inequality and the Itô isometry, we derive

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}\left|x_t-\bar{x}_t\right|^2\right] \le 7T^2\sup_{t\in[0,T]}\mathbb{E}\left[\left|r_t^d-\bar{r}_t^d\right|^2\right] + 7T^2\sup_{t\in[0,T]}\mathbb{E}\left[\left|r_t^f-\bar{r}_t^f\right|^2\right] \\ + \frac{7}{4}T^2\sup_{t\in[0,T]}\mathbb{E}\left[\left|v_t-\bar{v}_t\right|^2\right] + 28T\sup_{t\in[0,T]}\mathbb{E}\left[\left|v_t-\bar{v}_t\right|\right] \\ + 7\sum_{j=2}^{4} a_{1j}^2\,\mathbb{E}\left[\sup_{t\in[0,T]}\hat{v}_t^2\right]^{1/2}\mathbb{E}\left[\sup_{t\in[0,T]}\left|Z_t^j\right|^4\right]^{1/2}. \end{aligned} \tag{4.158}$$

The convergence as $\delta t \to 0$ of the first four terms on the right-hand side of (4.158) follows from Propositions 3.10 and 4.11. The uniform boundedness of the second moment of the FTE discretization for the variance follows from Proposition 3.8. Finally, employing the definition in (4.154), we bound the term inside the last expectation in (4.158) from above.

$$\begin{aligned}\sup_{t\in[0,T]}\left|Z_t^j\right|^4 &= \sup_{0\le n<N}\ \sup_{t_n\le t<t_{n+1}}\left|\frac{t-t_n}{\delta t}\left(W_{t_{n+1}}^j-W_{t_n}^j\right)-\left(W_t^j-W_{t_n}^j\right)\right|^4 \\ &\le 8\sup_{0\le n<N}\ \sup_{t_n\le t<t_{n+1}}\left[\left(\frac{t-t_n}{\delta t}\right)^4\left|W_{t_{n+1}}^j-W_{t_n}^j\right|^4+\left|W_t^j-W_{t_n}^j\right|^4\right] \\ &\le 16\sup_{0\le n<N}\ \sup_{t_n\le t<t_{n+1}}\left|W_t^j-W_{t_n}^j\right|^4 \\ &\le 16\sup_{t\in[0,T]}\left|W_t^j-W_{\delta t\lfloor t/\delta t\rfloor}^j\right|^4. \end{aligned} \tag{4.159}$$

However, the moments of the modulus of continuity of a Brownian motion converge to 0 as $\delta t \to 0$ [36], which concludes the proof. ∎

We now prove the convergence of the discretized spot FX rate process.

Proposition 4.13. *If $2k_f\theta_f > \xi_f^2$, then the process $\bar{S}$ converges uniformly in probability, i.e.,*

$$\lim_{\delta t\to 0} \mathbb{P}\left(\sup_{t\in[0,T]} \left|S_t - \bar{S}_t\right| > \epsilon\right) = 0, \quad \forall \epsilon > 0. \tag{4.160}$$

Proof. For fixed, positive numbers ϵ and γ such that $\log(1+\epsilon) > \gamma$, define the set

$$B_{\epsilon,\gamma} = \left\{x \in \mathbb{R} \,\middle|\, \exists y \in \mathbb{R}\colon\ |x-y| < \gamma \ \text{ and } \ |e^x - e^y| \geq \epsilon\right\}. \tag{4.161}$$

Since the exponential function is strictly increasing,

$$\begin{aligned} x \in B_{\epsilon,\gamma} \ &\Leftrightarrow\ \exists y \in (x-\gamma, x+\gamma)\colon\ e^{\min\{x,y\}}\left(e^{|x-y|} - 1\right) \geq \epsilon \\ &\Leftrightarrow\ e^x\left(e^\gamma - 1\right) > \epsilon. \end{aligned} \tag{4.162}$$

Hence,

$$B_{\epsilon,\gamma} = \left(a(\epsilon,\gamma), +\infty\right), \qquad \text{where} \qquad a(\epsilon,\gamma) = \log\left(\frac{\epsilon}{e^\gamma - 1}\right) > 0. \tag{4.163}$$

We have the following string of inclusions of events,

$$\begin{aligned} \left\{\sup_{t\in[0,T]} \left|S_t - \bar{S}_t\right| > \epsilon\right\} &\subseteq \left\{\sup_{t\in[0,T]} \left|x_t - \bar{x}_t\right| \geq \gamma\right\} \cup \left\{\sup_{t\in[0,T]} \left|x_t - \bar{x}_t\right| < \gamma,\ \sup_{t\in[0,T]} \left|e^{x_t} - e^{\bar{x}_t}\right| > \epsilon\right\} \\ &\subseteq \left\{\sup_{t\in[0,T]} \left|x_t - \bar{x}_t\right| \geq \gamma\right\} \cup \left\{\exists t \in [0,T]\colon\ x_t \in B_{\epsilon,\gamma}\right\} \\ &\subseteq \left\{\sup_{t\in[0,T]} \left|x_t - \bar{x}_t\right| \geq \gamma\right\} \cup \left\{\sup_{t\in[0,T]} x_t > a(\epsilon,\gamma)\right\}. \end{aligned} \tag{4.164}$$

In terms of probabilities of events, the previous inclusion becomes:

$$\mathbb{P}\left(\sup_{t\in[0,T]} \left|S_t - \bar{S}_t\right| > \epsilon\right) \leq \mathbb{P}\left(\sup_{t\in[0,T]} \left|x_t - \bar{x}_t\right| \geq \gamma\right) + \mathbb{P}\left(\sup_{t\in[0,T]} x_t > a(\epsilon,\gamma)\right). \tag{4.165}$$

The convergence in probability of the log-process is a consequence of Proposition 4.12 and Markov's inequality. Therefore, all that we have left to prove is that the second probability on the right-hand side of (4.165) can be made arbitrarily small. However, if we fix $\epsilon > 0$ and vary γ, then $\lim_{\gamma\to 0} a(\epsilon,\gamma) = \infty$. A simple application of Markov's inequality leads to

$$\mathbb{P}\left(\sup_{t\in[0,T]} x_t > a(\epsilon,\gamma)\right) \leq \mathbb{P}\left(\sup_{t\in[0,T]} |x_t| > a(\epsilon,\gamma)\right) \leq \frac{1}{a(\epsilon,\gamma)}\, \mathbb{E}\left[\sup_{t\in[0,T]} |x_t|\right]. \tag{4.166}$$

Then, using Jensen's inequality and Doob's martingale inequality, we get

$$\begin{aligned} \mathbb{E}\left[\sup_{t\in[0,T]} |x_t|\right] \leq |x_0| + T \sup_{t\in[0,T]} \mathbb{E}\left[r_t^d\right] + T \sup_{t\in[0,T]} \mathbb{E}\left[r_t^f\right] + \frac{T}{2} \sup_{t\in[0,T]} \mathbb{E}\left[v_t\right] \\ + 2\sqrt{T} \sup_{t\in[0,T]} \mathbb{E}\left[v_t\right]^{\frac{1}{2}}. \end{aligned} \tag{4.167}$$

The right-hand side is finite from Propositions 3.8 and 4.7, which concludes the proof. ∎

Theorem 4.14. *Let $\alpha \geq 1$ and assume that the following conditions are satisfied:*

$$\begin{cases} k > \alpha\rho_{sv}\xi + \max\left\{\sqrt{\alpha(\alpha-1)}\xi,\ \frac{1}{4}\alpha(\alpha-1)T\xi^2\right\}, \\ \dfrac{k_d^2}{2\xi_d^2} > \dfrac{\alpha\, q_1(\alpha)}{q_1(\alpha)-1}, \\ \dfrac{2k_d}{T\xi_d^2} > \dfrac{\alpha\, q_2(\alpha)}{q_2(\alpha)-1}, \\ 2k_f\theta_f > \xi_f^2. \end{cases} \tag{4.168}$$

Then the process $\bar{S}$ converges strongly in L^α, i.e.,

$$\lim_{\delta t\to 0}\sup_{t\in[0,T]} \mathbb{E}\left[\left|S_t - \bar{S}_t\right|^\alpha\right] = 0. \tag{4.169}$$

Proof. Fix $\epsilon > 0$ and define the event $A = \left\{\left|S_t - \bar{S}_t\right| > \epsilon\right\}$. Since S and $\bar{S}$ are non-negative,

$$\begin{aligned} \sup_{t\in[0,T]} \mathbb{E}\left[\left|S_t - \bar{S}_t\right|^\alpha\right] &\leq \sup_{t\in[0,T]} \mathbb{E}\left[\left|S_t - \bar{S}_t\right|^\alpha \mathbb{1}_{A^c}\right] + \sup_{t\in[0,T]} \mathbb{E}\left[\left|S_t - \bar{S}_t\right|^\alpha \mathbb{1}_{A}\right] \\ &\leq \epsilon^\alpha + \sup_{t\in[0,T]} \mathbb{E}\left[S_t^\alpha \mathbb{1}_A\right] + \sup_{t\in[0,T]} \mathbb{E}\left[\bar{S}_t^\alpha \mathbb{1}_A\right]. \end{aligned} \tag{4.170}$$

Let $\alpha < w < \min\{\alpha_1, \alpha_2\}$ and apply Hölder's inequality to the two expectations on the right-hand side of (4.170) with the pair $(p,q) = \left(\frac{w}{\alpha}, \frac{w}{w-\alpha}\right)$. Hence, we obtain

$$\begin{aligned} \sup_{t\in[0,T]} \mathbb{E}\Big[\left|S_t - \bar{S}_t\right|^\alpha\Big] \leq \epsilon^\alpha + &\left\{\sup_{t\in[0,T]} \mathbb{E}\big[S_t^w\big]^{\frac{\alpha}{w}} + \sup_{t\in[0,T]} \mathbb{E}\big[\bar{S}_t^w\big]^{\frac{\alpha}{w}}\right\} \\ &\times \sup_{t\in[0,T]} \mathbb{P}\left(\left|S_t - \bar{S}_t\right| > \epsilon\right)^{1-\frac{\alpha}{w}}. \end{aligned} \tag{4.171}$$

The convergence of $\bar{S}$ in probability is a consequence of Proposition 4.13. Finally, using Propositions 4.2 and 4.3 and then taking ϵ sufficiently small concludes the proof. ∎

For the purpose of proving convergence of options prices, it suffices to take $\alpha = 1$ above. Henceforth, we define for brevity $q_{1,2} = q_{1,2}(1)$.

Corollary 4.15. *Assume that the following conditions are satisfied:*

$$\begin{cases} k > \rho_{sv}\xi, \\ \dfrac{k_d^2}{2\xi_d^2} > \dfrac{q_1}{q_1-1}, \\ \dfrac{2k_d}{T\xi_d^2} > \dfrac{q_2}{q_2-1}, \\ 2k_f\theta_f > \xi_f^2. \end{cases} \tag{4.172}$$

Then the process $\bar{S}$ converges strongly in L^1, i.e.,

$$\lim_{\delta t\to 0}\sup_{t\in[0,T]} \mathbb{E}\left[\left|S_t - \bar{S}_t\right|\right] = 0. \tag{4.173}$$

Remark 4.16. It is difficult to tell how often the conditions in (4.172) are met in practice, which is why we investigate their restrictiveness next. Define the positive real numbers

$$\beta_v = \frac{k}{\xi} \quad \text{and} \quad \beta_d = \frac{k_d}{\xi_d}. \tag{4.174}$$

At this point, we make no assumptions on the parameters and find sufficient conditions that only involve β_v and β_d. The first condition in (4.172) is always given if $\beta_v > 1$, whereas the second one requires that $\beta_d > \sqrt{2}$ and can be rewritten as

$$\frac{\beta_d^2}{2} > \frac{q_1}{q_1 - 1} \quad \Leftrightarrow \quad q_1 > \frac{\beta_d^2}{\beta_d^2 - 2}. \tag{4.175}$$

Substituting with the definition of q_1 in (4.32) and taking into account that the correlation ρ_{sv} may take positive values, for (4.175) to hold it suffices to know that

$$\beta_v > \frac{\beta_d^2}{\beta_d^2 - 2} \quad \text{and} \quad q_0 > \frac{\beta_d^2}{\beta_d^2 - 2}. \tag{4.176}$$

However, we already know from (4.47) that, for all $q > 1$,

$$\begin{aligned} q_0 > q \quad &\Leftrightarrow \quad \xi^2 a_{13}^2 q^2 + \left[2\rho_{sv}\xi k + \xi^2\left(a_{11}^2 + a_{12}^2\right) - \xi^2\right] q - k^2 < 0 \\ &\Leftrightarrow \quad q(q-1)\xi^2\left(a_{13}^2 + a_{14}^2\right) < \left(k - q\xi\rho_{sv}\right)^2. \end{aligned} \tag{4.177}$$

Hence, the second inequality in (4.176) becomes

$$\frac{2\beta_d^2}{(\beta_d^2 - 2)^2}\xi^2\left(a_{13}^2 + a_{14}^2\right) < \left(k - \frac{\beta_d^2}{\beta_d^2 - 2}\xi\rho_{sv}\right)^2. \tag{4.178}$$

Employing the fact that $a_{13}^2 + a_{14}^2 \leq 1$, a consequence of $\Sigma = AA^{\mathrm{T}}$, and the first inequality in (4.176), taking square roots and dividing throughout by ξ leads to a sufficient condition for (4.178), namely

$$\beta_v > \frac{\beta_d^2}{\beta_d^2 - 2} + \frac{\beta_d\sqrt{2}}{\beta_d^2 - 2} = \frac{\beta_d}{\beta_d - \sqrt{2}}. \tag{4.179}$$

Hence, for the first two conditions in (4.172) to hold, it suffices to know that

$$\beta_d > \sqrt{2} \quad \text{and} \quad \beta_v > \frac{\beta_d}{\beta_d - \sqrt{2}}. \tag{4.180}$$

For instance, one can easily check that (4.180) holds if $\min\{\beta_v, \beta_d\} > \sqrt{2} + 1$. Define

$$\eta_v = \frac{2}{T\xi} \quad \text{and} \quad \eta_d = \frac{2}{T\xi_d}. \tag{4.181}$$

The third condition in (4.172) requires that $\beta_d\eta_d > 1$ and can be rewritten as

$$\beta_d\eta_d > \frac{q_2}{q_2 - 1} \quad \Leftrightarrow \quad q_2 > \frac{\beta_d\eta_d}{\beta_d\eta_d - 1}. \tag{4.182}$$

However, we already know from (4.71) that, for all $q > 1$,

$$q_2 > q \quad \Leftrightarrow \quad k - q\xi\rho_{sv} > \frac{1}{4}q(q-1)T\xi^2\left(a_{13}^2 + a_{14}^2\right). \tag{4.183}$$

Hence, (4.182) becomes

$$k - \frac{\beta_d \eta_d}{\beta_d \eta_d - 1} \xi \rho_{sv} > \frac{1}{4} \frac{\beta_d \eta_d}{(\beta_d \eta_d - 1)^2} T \xi^2 \left(a_{13}^2 + a_{14}^2 \right). \tag{4.184}$$

Dividing throughout by ξ, we derive a sufficient condition for (4.184),

$$\beta_v > \frac{\beta_d \eta_d}{\beta_d \eta_d - 1} + \frac{1}{2\eta_v} \frac{\beta_d \eta_d}{(\beta_d \eta_d - 1)^2}. \tag{4.185}$$

Hence, for the third condition in (4.172) to hold, it suffices to know that

$$\beta_d \eta_d > 1 \quad \text{and} \quad \beta_v > \frac{\beta_d \eta_d}{\beta_d \eta_d - 1} + \frac{1}{2\eta_v} \frac{\beta_d \eta_d}{(\beta_d \eta_d - 1)^2}. \tag{4.186}$$

If $\min\{\eta_v, \eta_d\} > 1$, the following are clearly sufficient conditions for (4.186),

$$\beta_d > 1 \quad \text{and} \quad \beta_v > \frac{1}{2} \frac{\beta_d}{\beta_d - 1} + \frac{1}{2} \left(\frac{\beta_d}{\beta_d - 1} \right)^2. \tag{4.187}$$

For convenience, also define

$$\nu_f = \frac{2 k_f \theta_f}{\xi_f^2}. \tag{4.188}$$

Table 4.1: The calibrated Heston parameters.

Parameter	**Jessen and Poulsen**	**Elices and Giménez**	**Aït–Sahalia and Kimmel**	**Hurn et al.**
k	2.2200	1.1000	5.1300	1.9775
θ	0.0120	0.0097	0.0436	0.0376
ξ	0.1830	0.1400	0.5200	0.4568
ρ_{sv}	0.0634	0.1400	−0.7540	−0.7591

(a) Column 2: for USDEUR market data, January 2, 2004 to September 27, 2005 [59]. Column 3: for EURUSD market data of August 22, 2006 [32]. Column 4: for the S&P 500 index between January 2, 1990 and September 30, 2003, using VIX data [2]. Column 5: for the S&P 500 index between January 2, 1990 and December 30, 2011, using two out-of-the-money options written on the index [53].

Table 4.2: The calibrated Cox–Ingersoll–Ross parameters.

Parameter	**Driffill et al.**	**Erismann**	**Brigo and Mercurio**	**Laffers**	**Amin**
$k_{d,f}$	0.0684	0.1104	0.3945	0.2820	0.1990
$\theta_{d,f}$	0.0161	0.0509	0.2713	0.0411	0.0497
$\xi_{d,f}$	0.0177	0.0498	0.0545	0.0058	0.0354

(a) Column 2: to the three-month US Treasury bill yield between January 1964 and December 1998 [29]. Column 3: to the US Treasury bill yield between October 1982 and April 2011 [33]. Column 4: to the euro at-the-money caps volatility curve on January 17, 2000 [13]. Column 5: to the euro overnight index average (Eonia) between January 1, 2008 and October 6, 2008 [66]. Column 6: using historical data for euro between January 1, 2001 and September 1, 2011 [5].

As before, one can easily check that (4.187) holds when $\min\{\beta_v, \beta_d\} > \sqrt{2} + 1$. Therefore, if $\min\{\eta_v, \eta_d\} > 1$, $\min\{\beta_v, \beta_d\} > \sqrt{2} + 1$ and $\nu_f > 1$, then the conditions in (4.172) are satisfied. The data in Tables 4.3 and 4.4 suggest that, in practice, $\min\{\eta_v, \eta_d\} > 1$ for short- and mid-term options (e.g., $T \leq 3$Y). Furthermore, we notice that k and k_d are significantly larger than their counterparts, ξ and ξ_d, such that $\min\{\beta_v, \beta_d\} > \sqrt{2}+1$ in all cases except for the parameter configuration of [33], where $\beta_d = 2.2169 < \sqrt{2}+1$. Finally, note that $\nu_f > 1$ in all cases.

Table 4.3: Sufficient conditions on the Heston parameters when $T = 3$Y.

Quantity	**Jessen and Poulsen**	**Elices and Giménez**	**Aït–Sahalia and Kimmel**	**Hurn et al.**
β_v	12.1311	7.8571	9.8654	4.3290
η_v	3.6430	4.7619	1.2821	1.4594

Table 4.4: Sufficient conditions on the Cox–Ingersoll–Ross parameters when $T = 3$Y.

Quantity	**Driffill et al.**	**Erismann**	**Brigo and Mercurio**	**Lafférs**	**Amin**
β_d	3.8644	2.2169	7.2385	48.6207	5.6215
η_d	37.6648	13.3869	12.2324	114.9425	18.8324
ν_f	7.0302	4.5317	72.0666	689.0725	15.7846

Since the payoff of a typical FX or equity contract grows at most linearly in the exchange rate or the asset price, we only need to know the strong convergence of the discounted process in L^1 in order to deduce the convergence of the time-discretization error to zero.

Theorem 4.17. *If $2k_f\theta_f > \xi_f^2$ and $T < T^*$, then the discounted process converges strongly in L^1, i.e.,*

$$\lim_{\delta t \to 0} \sup_{t \in [0,T]} \mathbb{E}\left[\left|R_t - \bar{R}_t\right|\right] = 0, \tag{4.189}$$

where

$$T^* = \begin{cases} \dfrac{1}{\rho_{sv}\xi - k}\ , & \text{if } k < \rho_{sv}\xi \\ \infty\ , & \text{if } k \geq \rho_{sv}\xi. \end{cases} \tag{4.190}$$

Proof. The convergence in probability of the discounted process is a consequence of Proposition 4.13, by taking the domestic interest rate to be zero. The rest of the proof follows the argument of Theorem 4.14 closely and makes use of Propositions 4.4 and 4.5. ∎

We can extend the convergence analysis from the 4-dimensional Heston–2CIR model (4.1) to multi-factor short rates with CIR dynamics and a term structure, in which case Propositions 4.4, 4.5 and 4.11 – 4.13, and hence Theorem 4.17 as well, still hold, albeit with slightly modified proofs.

The condition $k \geq \rho_{sv}\xi$, also known as the good correlation regime [58], is almost always satisfied in both FX and equity markets. This is because the speed of mean reversion k is usually larger than the volatility of volatility ξ, a fact clearly illustrated in Table 4.1. And even if this were not the case, a negative correlation between the underlying process and the variance, as is typically the case in equity markets (the so-called leverage effect), or a small absolute value of this correlation, as is typically the case in FX markets, would ensure the validity of the condition. Furthermore, the condition $2k_f\theta_f > \xi_f^2$ is generally satisfied in practice, a fact clearly illustrated in Table 4.2.

Remark 4.18. Note that Theorem 4.17 can be generalized to the L^α case for all $\alpha \geq 1$, upon noticing that the critical time T^* from (4.78) – (4.83) is always greater than the one from (4.103) – (4.104), a result which is important for the valuation of securities whose payoffs have superlinear growth under the constant interest rate assumption. Let $\alpha > 1$. First, suppose that $k < \alpha\rho_{sv}\xi - \sqrt{\alpha(\alpha-1)}\,\xi$. Since

$$\log(1+a) > \frac{a}{1+a} \tag{4.191}$$

for all $a > 0$, we deduce that

$$\begin{aligned}\frac{1}{\vartheta(\alpha)}\log\left(\frac{\alpha\rho_{sv}\xi - k + \vartheta(\alpha)}{\alpha\rho_{sv}\xi - k - \vartheta(\alpha)}\right) &> \frac{2}{\alpha\rho_{sv}\xi - k + \vartheta(\alpha)}\\ &> \frac{1}{\alpha\rho_{sv}\xi - k}\\ &> \frac{1}{\alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi - k}.\end{aligned} \tag{4.192}$$

Second, suppose that $k = \alpha\rho_{sv}\xi - \sqrt{\alpha(\alpha-1)}\,\xi$. Then clearly

$$\frac{2}{\alpha\rho_{sv}\xi - k} > \frac{1}{\alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi - k}. \tag{4.193}$$

Third, suppose that $\alpha\rho_{sv}\xi - \sqrt{\alpha(\alpha-1)}\,\xi < k < \alpha\rho_{sv}\xi$. Since

$$\arctan(a) + \arctan(1/a) = \pi/2 \tag{4.194}$$

for all $a > 0$ and

$$\arctan(a) > \frac{3a}{1+2\sqrt{1+a^2}} > \frac{a}{2} \tag{4.195}$$

for all $a \in (0,1)$, where the first inequality for the inverse tangent function is due to Shafer [83], we deduce that

$$\begin{aligned}\frac{2}{\hat{\vartheta}(\alpha)}\left[\frac{\pi}{2} - \arctan\left(\frac{\alpha\rho_{sv}\xi - k}{\hat{\vartheta}(\alpha)}\right)\right] &= \frac{2}{\hat{\vartheta}(\alpha)}\arctan\left(\frac{\hat{\vartheta}(\alpha)}{\alpha\rho_{sv}\xi - k}\right)\\ &> \frac{2}{\hat{\vartheta}(\alpha)}\arctan\left(\frac{\hat{\vartheta}(\alpha)}{\alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi - k}\right)\end{aligned}$$

$$> \frac{1}{\alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi - k}. \tag{4.196}$$

Fourth, suppose that $\alpha\rho_{sv}\xi \leq k < \alpha\rho_{sv}\xi + \frac{1}{2}\sqrt{\alpha(\alpha-1)}\,\xi$. Then

$$\begin{aligned}\frac{2}{\hat{\vartheta}(\alpha)}\left[\frac{\pi}{2} - \arctan\left(\frac{\alpha\rho_{sv}\xi - k}{\hat{\vartheta}(\alpha)}\right)\right] &> \frac{2}{\hat{\vartheta}(\alpha)} \\ &> \frac{2}{\sqrt{\alpha(\alpha-1)}\,\xi} \\ &> \frac{1}{\alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi - k}. \end{aligned} \tag{4.197}$$

Finally, suppose that $\alpha\rho_{sv}\xi + \frac{1}{2}\sqrt{\alpha(\alpha-1)}\,\xi \leq k < \alpha\rho_{sv}\xi + \sqrt{\alpha(\alpha-1)}\,\xi$. Employing the arithmetic-geometric mean inequality, we deduce that

$$\begin{aligned}\frac{2}{\hat{\vartheta}(\alpha)}\left[\frac{\pi}{2} - \arctan\left(\frac{\alpha\rho_{sv}\xi - k}{\hat{\vartheta}(\alpha)}\right)\right] &> \frac{4(k - \alpha\rho_{sv}\xi)}{2(k - \alpha\rho_{sv}\xi)\hat{\vartheta}(\alpha)} \\ &> \frac{4(k - \alpha\rho_{sv}\xi)}{\alpha(\alpha-1)\xi^2}, \end{aligned} \tag{4.198}$$

which concludes our discussion on the extension of Theorem 4.17.

4.3.4. Option valuation

We conclude this section with a brief study on the convergence of mixed Monte Carlo/PDE estimators for computing FX option prices. Define the fair price of an option written on S,

$$U = \mathbb{E}\left[e^{-\int_0^T r_t^d dt} f(S)\right], \tag{4.199}$$

and its approximation under (4.13),

$$\bar{U} = \mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S})\right], \tag{4.200}$$

where the payoff function f may depend on the entire path of the process and all expectations in this section are under the domestic risk-neutral measure $\mathbb{Q}$. The following theorem is concerned with the convergence of the time-discretization error to zero under the 4-factor FX model, and can be extended straightforwardly to multi-factor CIR short rates. The proof employs Propositions 4.4, 4.5 and 4.13 and Theorem 4.17. However, we omit it here because of its similarity to the proof of Theorem 3.2 once the aforementioned results are established.

Theorem 4.19. *Suppose that $2k_f\theta_f > \xi_f^2$. Then the following two statements hold.*

(1) The approximations to the values of the European put, the up-and-out call and any barrier put option defined in (4.200) converge as $\delta t \to 0$.

(2) If $T < T^*$, with T^* from (4.190)*, then the approximations to the values of the European call, Asian options, the down-and-in/out and the up-and-in call option defined in* (4.200) *converge as $\delta t \to 0$.*

For European contracts, we can evaluate the conditional option price, i.e., the innermost expectation in (4.20), analytically. Hence, consider M simulations of the discrete paths of the variance and the interest rates and, for $1 \leq j \leq M$, let ω_j denote the j-th sample. Then

$$\frac{1}{M}\sum_{j=1}^{M}\mathbb{E}\Big[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S}_T) \,\big|\, \mathcal{G}_T^{f,d,v}, \omega = \omega_j\Big] = \frac{1}{M}\sum_{j=1}^{M}\mathbb{E}\Big[e^{-\int_0^T \bar{r}_t^d(j) dt} f(\bar{S}_T(j)) \,\big|\, \mathcal{G}_T^{f,d,v}\Big] \quad (4.201)$$

is the mixed Monte Carlo/PDE estimator of the European option price at time $t = 0$. The global error can be split into two parts:

$$\begin{aligned}
e(M,N) &= \mathbb{E}\left[e^{-\int_0^T r_t^d dt} f(S_T)\right] - \frac{1}{M}\sum_{j=1}^{M}\mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S}_T) \,\big|\, \mathcal{G}_T^{f,d,v}, \omega = \omega_j\right] \\
&= \left(\mathbb{E}\left[e^{-\int_0^T r_t^d dt} f(S_T)\right] - \mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S}_T)\right]\right) \\
&\quad + \left(\mathbb{E}\left[\mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S}_T) \,\big|\, \mathcal{G}_T^{f,d,v}\right]\right] - \frac{1}{M}\sum_{j=1}^{M}\mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S}_T) \,\big|\, \mathcal{G}_T^{f,d,v}, \omega = \omega_j\right]\right).
\end{aligned} \quad (4.202)$$

The first term is the time-discretization error and the second term is the statistical error. The convergence to zero of the former was derived in Theorem 4.19 for European put and call options, a result that can easily be extended to other financial derivatives, such as binary options. The convergence to zero of the latter follows from the Central Limit Theorem [41] upon noticing the following upper bound on the variance:

$$\begin{aligned}
\mathrm{Var}\left(\mathbb{E}\Big[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S}_T) \,\big|\, \mathcal{G}_T^{f,d,v}\Big]\right) &\leq \mathbb{E}\Big[\mathbb{E}\Big[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S}_T) \,\big|\, \mathcal{G}_T^{f,d,v}\Big]^2\Big] \\
&\leq \mathbb{E}\Big[e^{-2\int_0^T \bar{r}_t^d dt} f(\bar{S}_T)^2\Big].
\end{aligned} \quad (4.203)$$

Assuming f has at most polynomial growth, we can employ Proposition 4.3 to deduce the finiteness of the variance under some conditions on the model parameters. In particular, if f is Lipschitz, for $\alpha = 2$, Proposition 4.5 gives sharper sufficient conditions.

For path-dependent contracts, closed-form solutions are rarely available and we instead rely on finite differences to compute the conditional option price. Conditioning on the j-th realization of the variance and interest rates paths, let $u_j^{f,d,v}\big(t, \bar{S}_t\big)$ and $\bar{u}_j^{f,d,v}\big(t, \bar{S}_t; P, L\big)$ be the solutions to the conditional PDE and to the associated finite difference scheme, when a uniform mesh with P time steps and L spatial steps, respectively, is employed. It can be shown that the conditional PDE has a unique solution (see Section 7.1.2 in [34]) which is, in fact, the conditional option price (see Theorem 7.3.1 in [84]). Then

$$\frac{1}{M}\sum_{j=1}^{M}\bar{u}_j^{f,d,v}\big(0,S_0;P,L\big) \tag{4.204}$$

is the mixed Monte Carlo/PDE estimator of the option price at $t=0$. The global error can be split into three parts:

$$\begin{aligned}
e(M,N,P,L) &= \mathbb{E}\left[e^{-\int_0^T r_t^d dt} f(S)\right] - \frac{1}{M}\sum_{j=1}^{M}\bar{u}_j^{f,d,v}\big(0,S_0;P,L\big)\\
&= \left(\mathbb{E}\left[e^{-\int_0^T r_t^d dt} f(S)\right] - \mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S})\right]\right)\\
&+\left(\mathbb{E}\Big[\mathbb{E}\Big[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S})\,\big|\,\mathcal{G}_T^{f,d,v}\Big]\Big] - \frac{1}{M}\sum_{j=1}^{M}\mathbb{E}\Big[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S})\,\big|\,\mathcal{G}_T^{f,d,v},\omega=\omega_j\Big]\right)\\
&+\frac{1}{M}\sum_{j=1}^{M}\Big(u_j^{f,d,v}\big(0,S_0\big) - \bar{u}_j^{f,d,v}\big(0,S_0;P,L\big)\Big).
\end{aligned} \tag{4.205}$$

The first term is the time-discretization error, the second term is the statistical error and the third term is the finite difference (FD) discretization error. The convergence to zero of the first term was derived in Theorem 4.19 for Asian and barrier options, and the convergence to zero of the second term is a consequence of the Central Limit Theorem (see the above discussion). Finally, the convergence of the third term, i.e., of the finite difference scheme, depends on the contract and the particular scheme employed.

4.4. Variance reduction analysis

In this section, we perform a variance reduction analysis for European option valuation with the mixed Monte Carlo/PDE method under the 4-factor FX model. However, the theory extends naturally to general interest rate dynamics. We use standard Monte Carlo with the log-Euler discretization as the reference method and define $\hat{x}$ and $\hat{S}$ to be the continuous-time approximations of x, defined in (4.150), and S, defined in (4.1), respectively. Then

$$\hat{x}_t = \hat{x}_{t_n} + \Big(\bar{r}_{t_n}^d - \bar{r}_{t_n}^f - \frac{1}{2}\bar{v}_{t_n}\Big)\big(t-t_n\big) + \sqrt{\bar{v}_{t_n}}\,\Delta W_t^s, \tag{4.206}$$

where $\Delta W_t^s = W_t^s - W_{t_n}^s$ whenever $t\in[t_n,t_{n+1})$. Integrating (4.206) leads to

$$\hat{S}_t = S_0\exp\left\{\int_0^t\Big(\bar{r}_u^d - \bar{r}_u^f - \frac{1}{2}\bar{v}_u\Big)du + \int_0^t\sqrt{\bar{v}_u}\,dW_u^s\right\}. \tag{4.207}$$

We prefer the log-Euler scheme to the standard Euler scheme because it preserves positivity. Moreover, if the processes v, r^d, r^f are constant, then the first scheme is exact. Recall that

$$\bar{S}_t = S_0\exp\left\{\int_0^t\Big(\bar{r}_u^d - \bar{r}_u^f - \frac{1}{2}\bar{v}_u\Big)du + a_{11}\int_0^t\sqrt{\bar{v}_u}\,dW_u^1 + \sum_{j=2}^{4}a_{1j}\int_0^t\sqrt{\bar{v}_u}\,\frac{\delta W_u^j}{\delta t}\,du\right\}. \tag{4.208}$$

Since $\bar{v}$ is piecewise constant, we deduce that $\hat{S}_{t_n} = \bar{S}_{t_n}$ for all $0 \leq n \leq N$. The quantity that we want to estimate is the fair price of a European option with payoff function f, i.e.,

$$\Theta = \mathbb{E}\left[e^{-\int_0^T r_t^d dt} f(S_T)\right]. \tag{4.209}$$

Then the corresponding standard and mixed Monte Carlo estimators are

$$\Theta_{\text{stdMC}} = \frac{1}{M} \sum_{j=1}^{M} e^{-\int_0^T \bar{r}_t^d(j) dt} f(\hat{S}_T(j)), \tag{4.210}$$

$$\Theta_{\text{mixMC}} = \frac{1}{M} \sum_{j=1}^{M} \mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d(j) dt} f(\bar{S}_T(j)) \,\middle|\, \mathcal{G}_T^{f,d,v}\right]. \tag{4.211}$$

Define

$$\text{Var}_{\text{stdMC}} = \text{Var}\big(\Theta_{\text{stdMC}}\big) = \frac{1}{M} \text{Var}\left(e^{-\int_0^T \bar{r}_t^d dt} f(\hat{S}_T)\right) \tag{4.212}$$

and

$$\text{Var}_{\text{mixMC}} = \text{Var}\big(\Theta_{\text{mixMC}}\big) = \frac{1}{M} \text{Var}\left(\mathbb{E}\left[e^{-\int_0^T \bar{r}_t^d dt} f(\bar{S}_T) \,\middle|\, \mathcal{G}_T^{f,d,v}\right]\right). \tag{4.213}$$

Let the variance reduction factor (as in [23]) and the standard deviation ratio be

$$\Gamma_{\text{var}} = \frac{\text{Var}_{\text{stdMC}}}{\text{Var}_{\text{mixMC}}} \quad \text{and} \quad \Gamma_{\text{dev}} = \sqrt{\Gamma_{\text{var}}}. \tag{4.214}$$

For convenience, we also define the discount factors

$$D = e^{-\int_0^T r_t^d dt} \quad \text{and} \quad \bar{D} = e^{-\int_0^T \bar{r}_t^d dt}. \tag{4.215}$$

Remark 4.20. From the "tower property" of conditional expectations, since $\hat{S}_T = \bar{S}_T$,

$$\begin{aligned}
\mathbb{E}\left[\Theta_{\text{stdMC}}\right] &= \mathbb{E}\left[\bar{D} f(\hat{S}_T)\right] \\
&= \mathbb{E}\left[\bar{D} f(\bar{S}_T)\right] \\
&= \mathbb{E}\left[\, \mathbb{E}\left[\bar{D} f(\bar{S}_T) \,|\, \mathcal{G}_T^{f,d,v}\right]\right] \\
&= \mathbb{E}\left[\Theta_{\text{mixMC}}\right].
\end{aligned} \tag{4.216}$$

Therefore, the standard and the mixed Monte Carlo estimators have the same discretization bias, i.e.,

$$\text{Bias}\big(\Theta_{\text{stdMC}}\big) = \text{Bias}\big(\Theta_{\text{mixMC}}\big). \tag{4.217}$$

Remark 4.21. From the law of total variance we know that

$$\text{Var}\big(\bar{D} f(\bar{S}_T)\big) = \text{Var}\big(\, \mathbb{E}\left[\bar{D} f(\bar{S}_T) \,|\, \mathcal{G}_T^{f,d,v}\right]\big) + \mathbb{E}\left[\text{Var}\big(\bar{D} f(\bar{S}_T) \,|\, \mathcal{G}_T^{f,d,v}\big)\right]. \tag{4.218}$$

However, since $\hat{S}_T = \bar{S}_T$ and the variance is non-negative, we deduce from (4.212) and (4.213) that the variance of the standard Monte Carlo estimator is greater than or equal to the variance of the mixed estimator, i.e.,

$$\text{Var}_{\text{stdMC}} \geq \text{Var}_{\text{mixMC}}. \tag{4.219}$$

Assuming a non-trivial payoff function f, equality occurs in (4.219) if and only if the second expectation on the right-hand side of (4.218) is zero, i.e., if and only if $\bar{S}_T$ is $\mathcal{G}_T^{f,d,v}$-measurable. Since $\Sigma = AA^{\mathrm{T}}$, after some straightforward calculations we find an equivalent condition:

$$\begin{aligned} a_{11} = 0 \quad \Leftrightarrow \quad 1 - \rho_{vd}^2 - \rho_{vf}^2 - \rho_{df}^2 + 2\rho_{vd}\rho_{vf}\rho_{df} \\ &= \rho_{sv}^2\big(1-\rho_{df}^2\big) + \rho_{sd}^2\big(1-\rho_{vf}^2\big) + \rho_{sf}^2\big(1-\rho_{vd}^2\big) \\ &+ 2\rho_{sv}\rho_{sd}\big(\rho_{vf}\rho_{df} - \rho_{vd}\big) + 2\rho_{sv}\rho_{sf}\big(\rho_{vd}\rho_{df} - \rho_{vf}\big) \\ &+ 2\rho_{sd}\rho_{sf}\big(\rho_{vd}\rho_{vf} - \rho_{df}\big). \end{aligned} \tag{4.220}$$

In particular, if the variance and the two interest rates are pairwise independent, then

$$a_{11} = 0 \quad \Leftrightarrow \quad \rho_{sv}^2 + \rho_{sd}^2 + \rho_{sf}^2 = 1. \tag{4.221}$$

Therefore, apart from this case, combining conditioning with Monte Carlo always reduces the variance of estimates. This is, however, to be expected since we eliminate the additional noise that comes from simulating the Brownian motion W^1.

Remark 4.22. Note that for any $2 \leq j \leq 4$, using the Itô isometry, we have

$$\operatorname{Var}\bigg(\int_0^T \sqrt{\bar{v}_t}\, dW_t^j\bigg) = \mathbb{E}\bigg[\int_0^T \bar{v}_t\, dt\bigg]. \tag{4.222}$$

Moreover, using the Cauchy–Schwarz and Hölder inequalities, Fubini's theorem, Propositions 3.8 and 3.10, we can easily prove that

$$\lim_{\delta t \to 0} \frac{\mathbb{E}\Big[\int_0^T \bar{v}_t\, dt\Big]}{\operatorname{Var}\Big(\int_0^T \bar{v}_t\, dt\Big)} = \frac{\mathbb{E}\Big[\int_0^T v_t\, dt\Big]}{\operatorname{Var}\Big(\int_0^T v_t\, dt\Big)}. \tag{4.223}$$

From [30], we can compute the first two moments of the integrated square-root process explicitly. Hence, we find

$$\mathbb{E}\bigg[\int_0^T v_t dt\bigg] = \theta T + \frac{v_0}{k} - \frac{\theta}{k} + e^{-kT}\left(\frac{\theta}{k} - \frac{v_0}{k}\right) \tag{4.224}$$

and

$$\begin{aligned} \operatorname{Var}\bigg(\int_0^T v_t dt\bigg) &= \frac{\xi^2}{k^2}\left[\theta T + \frac{v_0}{k} - \frac{5\theta}{2k} + 2e^{-kT}\left(\frac{\theta}{k} + \theta T - v_0 T\right) + e^{-2kT}\left(\frac{\theta}{2k} - \frac{v_0}{k}\right)\right] \\ &< \frac{\xi^2}{k^2}\big(1+e^{-kT}\big)\,\mathbb{E}\bigg[\int_0^T v_t dt\bigg] + \frac{\xi^2}{k^2}\left[\frac{\theta}{2k}e^{-kT}\left(4 + 2kT - e^{-kT} - 3e^{kT}\right)\right] \\ &< \frac{\xi^2}{k^2}\big(1+e^{-kT}\big)\,\mathbb{E}\bigg[\int_0^T v_t dt\bigg]. \end{aligned} \tag{4.225}$$

Combining (4.222) – (4.225), we deduce that, for sufficiently small values of δt,

$$\frac{\operatorname{Var}\Big(\int_0^T \sqrt{\bar{v}_t}\, dW_t^j\Big)}{\operatorname{Var}\Big(\int_0^T \bar{v}_t\, dt\Big)} > \frac{k^2}{\xi^2\left(1+e^{-kT}\right)} > \frac{k^2}{2\xi^2}. \tag{4.226}$$

The data in Tables 4.1 and 4.2 suggest that the interest rates have little impact on the variance of the mixed Monte Carlo estimator, and also that $k \gg \xi$ in both FX and equity markets. Hence, the stochastic integral on the left-hand side of (4.226) – part of the dividend yield q defined in (4.19) – contributes to the overall variance of the mixed estimator considerably more than the squared volatility σ^2 defined in (4.18). Therefore, we expect the minimum variance of the mixed estimator to be attained when all but the first term in q vanish, i.e., when $a_{11} = 1$ and $a_{12} = a_{13} = a_{14} = 0$.

In practice, the volatility of interest rates in the CIR model is very small, i.e., $\xi_{d,f} \ll 1$, a fact which can be observed in Table 4.2. Moreover, the volatility of volatility in the Heston model – calibrated to FX market data – is also small, i.e., $\xi \ll 1$, and is significantly smaller than the rate of mean reversion, i.e., $\xi \ll k$, a fact clearly illustrated in Table 4.1. Hence, the drift term in the square-root model for the variance dominates the diffusion term. Therefore, in the subsequent analysis we assume "almost deterministic" dynamics of the variance and interest rates. Let $\bar{\gamma}$, $\bar{\gamma}^d$ and $\bar{\gamma}^f$ be the FTE discretizations of v, r^d and r^f corresponding to $\xi = \xi_d = \xi_f = 0$, i.e., when the volatility of volatility and the volatility of interest rates are equal to zero. Then

$$\begin{aligned} \int_0^T \bar{r}_t^d \, dt \approx \int_0^T \bar{\gamma}_t^d \, dt, \quad \int_0^T \bar{r}_t^f \, dt \approx \int_0^T \bar{\gamma}_t^f \, dt, \\ \int_0^T \bar{v}_t \, dt \approx \int_0^T \bar{\gamma}_t \, dt, \quad \int_0^T \sqrt{\bar{v}_t} \, dW_t^s \approx \int_0^T \sqrt{\bar{\gamma}_t} \, dW_t^s. \end{aligned} \tag{4.227}$$

Remark 4.23. Suppose that W^s is independent of the Brownian motions W^v, W^d and W^f, i.e., that $a_{11} = 1$ and $a_{12} = a_{13} = a_{14} = 0$. Under the above assumption on the dynamics of the variance and the domestic and foreign interest rates, we have

$$\hat{S}_T = \bar{S}_T \approx \tilde{S}_T, \tag{4.228}$$

where

$$\tilde{S}_T = S_0 \exp\left\{ \int_0^T \left(\bar{\gamma}_t^d - \bar{\gamma}_t^f - \frac{1}{2}\bar{\gamma}_t \right) dt + \int_0^T \sqrt{\bar{\gamma}_t} \, dW_t^s \right\}. \tag{4.229}$$

On the one hand, we deduce from (4.212) that

$$\begin{aligned} \mathrm{Var}_{\mathrm{stdMC}} &= \frac{1}{M} \mathrm{Var}\left(e^{-\int_0^T \bar{r}_t^d dt} f(\hat{S}_T) \right) \\ &\approx \frac{1}{M} \mathrm{Var}\left(e^{-\int_0^T \bar{\gamma}_t^d dt} f(\tilde{S}_T) \right) \\ &= \frac{1}{M} e^{-2\int_0^T \bar{\gamma}_t^d dt} \, \mathrm{Var}\big(f(\tilde{S}_T)\big). \end{aligned} \tag{4.230}$$

On the other hand, we deduce from (4.213) that

$$\begin{aligned} \mathrm{Var}_{\mathrm{mixMC}} &= \frac{1}{M} \mathrm{Var}\left(e^{-\int_0^T \bar{r}_t^d dt} \, \mathbb{E}\left[f(\bar{S}_T) \,|\, \mathcal{G}_T^{f,d,v} \right] \right) \\ &\approx \frac{1}{M} \mathrm{Var}\left(e^{-\int_0^T \bar{\gamma}_t^d dt} \, \mathbb{E}\left[f(\tilde{S}_T) \,|\, \mathcal{G}_T^{f,d,v} \right] \right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{M} e^{-2\int_0^T \bar{\gamma}_t^d dt} \operatorname{Var}\big(\mathbb{E}\left[f(\tilde{S}_T)\right]\big) \\
&= 0. \qquad (4.231)
\end{aligned}
$$

Assuming a non-trivial payoff f as well as a non-zero squared volatility v, since the variance of the mixed estimator is close to zero, this results in a substantial variance reduction.

Remark 4.24. Let f be a differentiable payoff function and suppose that the dynamics of the variance and interest rates are "almost deterministic" (i.e., their own volatility parameters are small compared to their mean-reversion speed). Then we can approximate the discounted payoff as follows:

$$
\begin{aligned}
H &= e^{-\int_0^T \bar{r}_t^d dt} f\bigg(S_0 \exp\bigg\{\int_0^T \Big(\bar{r}_t^d - \bar{r}_t^f - \frac{1}{2}\bar{v}_t\Big)dt + \int_0^T \sqrt{\bar{v}_t}\, dW_t^s\bigg\}\bigg) \\
&\approx e^{-\int_0^T \bar{\gamma}_t^d dt} f\bigg(S_0 \exp\bigg\{\int_0^T \Big(\bar{\gamma}_t^d - \bar{\gamma}_t^f - \frac{1}{2}\bar{\gamma}_t\Big)dt + \int_0^T \sqrt{\bar{\gamma}_t}\, dW_t^s\bigg\}\bigg). \qquad (4.232)
\end{aligned}
$$

For convenience, we introduce the notations

$$
\tilde{D} = e^{-\int_0^T \bar{\gamma}_t^d dt} \quad \text{and} \quad \tilde{F} = S_0 \exp\bigg\{\int_0^T \Big(\bar{\gamma}_t^d - \bar{\gamma}_t^f - \frac{1}{2}\bar{\gamma}_t\Big)dt\bigg\}. \qquad (4.233)
$$

Therefore,

$$
H \approx \tilde{D} f\bigg(\tilde{F} \exp\bigg\{\int_0^T \sqrt{\bar{\gamma}_t}\, dW_t^s\bigg\}\bigg). \qquad (4.234)
$$

According to the data in Table 4.1, $\theta \ll 1$ and so $\bar{\gamma} \ll 1$ as well. Hence, the absolute value of the stochastic integral in (4.234) is small for low maturities T and we can use a first-order Taylor series approximation to get

$$
H \approx \tilde{D}\bigg[f(\tilde{F}) + f'(\tilde{F})\tilde{F}\bigg(\exp\bigg\{\int_0^T \sqrt{\bar{\gamma}_t}\, dW_t^s\bigg\} - 1\bigg)\bigg]. \qquad (4.235)
$$

First, since $\tilde{D}$ and $\tilde{F}$ are constant, the variance of the standard Monte Carlo estimator is

$$
\begin{aligned}
\operatorname{Var}_{\text{stdMC}} &\approx \frac{1}{M} \operatorname{Var}\bigg(\tilde{D}\bigg[f(\tilde{F}) + f'(\tilde{F})\tilde{F}\bigg(\exp\bigg\{\int_0^T \sqrt{\bar{\gamma}_t}\, dW_t^s\bigg\} - 1\bigg)\bigg]\bigg) \\
&\approx \frac{1}{M} \tilde{D}^2 \tilde{F}^2 \big[f'(\tilde{F})\big]^2 \operatorname{Var}\bigg(\exp\bigg\{\int_0^T \sqrt{\bar{\gamma}_t}\, dW_t^s\bigg\}\bigg) \\
&\approx \frac{1}{M} \tilde{D}^2 \tilde{F}^2 \big[f'(\tilde{F})\big]^2 \int_0^T \bar{\gamma}_t dt. \qquad (4.236)
\end{aligned}
$$

Second, using MGFs, the variance of the mixed Monte Carlo estimator is

$$
\begin{aligned}
\operatorname{Var}_{\text{mixMC}} &\approx \frac{1}{M} \tilde{D}^2 \tilde{F}^2 \big[f'(\tilde{F})\big]^2 \operatorname{Var}\bigg(\mathbb{E}\bigg[\exp\bigg\{\sum_{j=1}^4 a_{1j} \int_0^T \sqrt{\bar{\gamma}_t}\, dW_t^j\bigg\} \,\bigg|\, \mathcal{G}_T^{f,d,v}\bigg]\bigg) \\
&\approx \frac{1}{M} \tilde{D}^2 \tilde{F}^2 \big[f'(\tilde{F})\big]^2 \exp\bigg\{a_{11}^2 \int_0^T \bar{\gamma}_t dt\bigg\} \operatorname{Var}\bigg(\exp\bigg\{\sum_{j=2}^4 a_{1j} \int_0^T \sqrt{\bar{\gamma}_t}\, dW_t^j\bigg\}\bigg)
\end{aligned}
$$

$$\approx \frac{1}{M}\tilde{D}^2\tilde{F}^2\big[f'(\tilde{F})\big]^2\big(1-a_{11}^2\big)\int_0^T \bar{\gamma}_t dt. \tag{4.237}$$

Finally, we conclude from (4.236) and (4.237) that

$$\Gamma_{\text{dev}} \approx \big(1-a_{11}^2\big)^{-\frac{1}{2}}. \tag{4.238}$$

Note that Remark 4.24 is consistent with Remarks 4.21 – 4.23. However, the method employed to approximate the variance reduction factor fails if f is not differentiable, which is the case with most vanilla options traded in the FX market.

Remark 4.25. Let f be the European call option payoff and suppose that the dynamics of the variance and interest rates are "almost deterministic" (i.e., their own volatility parameters are small compared to their mean-reversion speed). Then we can approximate the discounted payoff as follows:

$$\begin{aligned} H &= e^{-\int_0^T \bar{r}_t^d dt}\bigg(S_0\exp\bigg\{\int_0^T\Big(\bar{r}_t^d-\bar{r}_t^f-\frac{1}{2}\bar{v}_t\Big)dt+\int_0^T\sqrt{\bar{v}_t}\,dW_t^s\bigg\}-K\bigg)^+ \\ &\approx e^{-\int_0^T \bar{\gamma}_t^d dt}\bigg(S_0\exp\bigg\{\int_0^T\Big(\bar{\gamma}_t^d-\bar{\gamma}_t^f-\frac{1}{2}\bar{\gamma}_t\Big)dt+\int_0^T\sqrt{\bar{\gamma}_t}\,dW_t^s\bigg\}-K\bigg)^+. \end{aligned} \tag{4.239}$$

For now, assume that $a_{11}>0$. For convenience, we define the quantities

$$\begin{aligned} \varrho &= \sqrt{1-a_{11}^2}, \quad \tilde{\sigma}=\sqrt{\int_0^T\bar{\gamma}_t\,dt}, \\ \tilde{D} &= e^{-\int_0^T\bar{\gamma}_t^d dt}, \quad F=S_0\exp\bigg\{\int_0^T\big(\bar{\gamma}_t^d-\bar{\gamma}_t^f\big)dt\bigg\}, \end{aligned} \tag{4.240}$$

as well as

$$\begin{aligned} a &= \frac{\varrho}{\sqrt{1-\varrho^2}}, \quad b=\frac{\log(F/K)+\big(\frac{1}{2}+\varrho^2\big)\tilde{\sigma}^2}{\sqrt{1-\varrho^2}\,\tilde{\sigma}}, \\ b_1 &= \frac{\log(F/K)+\big(\frac{1}{2}-\varrho^2\big)\tilde{\sigma}^2}{\sqrt{1-\varrho^2}\,\tilde{\sigma}}, \quad b_2=\frac{\log(F/K)-\frac{1}{2}\tilde{\sigma}^2}{\sqrt{1-\varrho^2}\,\tilde{\sigma}}. \end{aligned} \tag{4.241}$$

Using (4.239) and (4.240), we approximate the conditional option price as follows:

$$\begin{aligned} \mathbb{E}\Big[H\,\big|\,\mathcal{G}_T^{f,d,v}\Big] &\approx \mathbb{E}\bigg[\tilde{D}\bigg(F\exp\bigg\{-\frac{1}{2}\tilde{\sigma}^2+\int_0^T\sqrt{\bar{\gamma}_t}\,dW_t^s\bigg\}-K\bigg)^+\,\big|\,\mathcal{G}_T^{f,d,v}\bigg] \\ &\approx \mathbb{E}\bigg[\tilde{D}\bigg(F\exp\bigg\{-\frac{1}{2}\tilde{\sigma}^2+\varrho\tilde{\sigma}Z+\sqrt{1-\varrho^2}\,\tilde{\sigma}Z'\bigg\}-K\bigg)^+\,\big|\,Z\bigg], \end{aligned} \tag{4.242}$$

where $Z, Z' \sim \mathcal{N}(0,1)$ are independent such that Z is $\mathcal{G}_T^{f,d,v}$-measurable and $Z' \perp\!\!\!\perp \mathcal{G}_T^{f,d,v}$. Some straightforward calculations lead to

$$\mathbb{E}\Big[H\,\big|\,\mathcal{G}_T^{f,d,v}\Big] \approx \tilde{D}Fe^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2+\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)-\tilde{D}K\Phi(aZ+b_2), \tag{4.243}$$

where Φ is the standard normal CDF and ϕ is the standard normal PDF. From (4.213), the variance of the mixed Monte Carlo estimator is thus

$$\mathrm{Var}_{\text{mixMC}} \approx \frac{1}{M}\,\mathbb{E}\bigg[\Big(\tilde{D}Fe^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2+\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)-\tilde{D}K\Phi(aZ+b_2)\Big)^2\bigg]$$

$$- \frac{1}{M}\,\mathbb{E}\left[\tilde{D}Fe^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2+\varrho\tilde{\sigma}Z}\Phi(aZ+b_1) - \tilde{D}K\Phi(aZ+b_2)\right]^2. \tag{4.244}$$

Differentiating the right-hand side of (4.244) with respect to the parameter ϱ, changing the order of differentiation and integration, and upon noticing that

$$\frac{\partial}{\partial\varrho} e^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2+\varrho\tilde{\sigma}Z} = \big(\tilde{\sigma}Z - \varrho\tilde{\sigma}^2\big)e^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2+\varrho\tilde{\sigma}Z}, \tag{4.245}$$

$$\frac{\partial}{\partial\varrho}\Phi(aZ+b_1) = \frac{\varrho\log(F/K) - \frac{1}{2}\varrho\tilde{\sigma}^2 - \varrho(1-\varrho^2)\tilde{\sigma}^2 + \tilde{\sigma}Z}{(1-\varrho^2)^{3/2}\tilde{\sigma}}\phi(aZ+b_1), \tag{4.246}$$

$$\frac{\partial}{\partial\varrho}\Phi(aZ+b_2) = \frac{\varrho\log(F/K) - \frac{1}{2}\varrho\tilde{\sigma}^2 + \tilde{\sigma}Z}{(1-\varrho^2)^{3/2}\tilde{\sigma}}\phi(aZ+b_2), \tag{4.247}$$

as well as the fact that

$$Fe^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2+\varrho\tilde{\sigma}Z}\phi(aZ+b_1) = K\phi(aZ+b_2), \tag{4.248}$$

we obtain

$$\begin{aligned}
\frac{\partial}{\partial\varrho}\mathrm{Var}_{\mathrm{mixMC}} \approx{}& \frac{2}{M}\tilde{D}^2F\tilde{\sigma}\,\mathbb{E}\Big[F(Z-\varrho\tilde{\sigma})e^{-\varrho^2\tilde{\sigma}^2+2\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)^2 \\
&- K(Z-\varrho\tilde{\sigma})e^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2+\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)\Phi(aZ+b_2) \\
&- aFe^{-\varrho^2\tilde{\sigma}^2+2\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)\phi(aZ+b_1) \\
&+ aKe^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2+\varrho\tilde{\sigma}Z}\Phi(aZ+b_2)\phi(aZ+b_1)\Big] \\
&- \frac{2}{M}\tilde{D}^2F\tilde{\sigma}e^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2}\,\mathbb{E}\Big[Fe^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2+\varrho\tilde{\sigma}Z}\Phi(aZ+b_1) - K\Phi(aZ+b_2)\Big] \\
&\times \mathbb{E}\Big[(Z-\varrho\tilde{\sigma})e^{\varrho\tilde{\sigma}Z}\Phi(aZ+b_1) - ae^{\varrho\tilde{\sigma}Z}\phi(aZ+b_1)\Big].
\end{aligned} \tag{4.249}$$

On the one hand, using integration by parts, we find that

$$\begin{aligned}
\mathbb{E}\Big[(Z-\varrho\tilde{\sigma})e^{\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)\Big] &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}(z-\varrho\tilde{\sigma})e^{-\frac{1}{2}z^2+\varrho\tilde{\sigma}z}\Phi(az+b_1)\,dz \\
&= \left[\frac{-1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2+\varrho\tilde{\sigma}z}\Phi(az+b_1)\right]_{-\infty}^{\infty} \\
&\quad + \int_{-\infty}^{\infty}\frac{a}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2+\varrho\tilde{\sigma}z}\phi(az+b_1)\,dz \\
&= \mathbb{E}\Big[ae^{\varrho\tilde{\sigma}Z}\phi(aZ+b_1)\Big],
\end{aligned} \tag{4.250}$$

and hence the last expectation on the right-hand side of (4.249) is zero. Alternatively, this follows from the fact that the second expectation on the right-hand side of (4.244) is the (unconditional) option price, and hence its derivative with respect to the parameter ϱ is zero. On the other hand, using integration by parts, we can rewrite

$$\begin{aligned}
\mathbb{E}\Big[(Z-\varrho\tilde{\sigma})e^{2\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)^2\Big] &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}(z-2\varrho\tilde{\sigma}+\varrho\tilde{\sigma})e^{-\frac{1}{2}z^2+2\varrho\tilde{\sigma}z}\Phi(az+b_1)^2\,dz \\
&= \left[\frac{-1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2+2\varrho\tilde{\sigma}z}\Phi(az+b_1)^2\right]_{-\infty}^{\infty}
\end{aligned}$$

$$
\begin{aligned}
&+ \int_{-\infty}^{\infty} \frac{2a}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2+2\varrho\tilde{\sigma}z} \Phi(az+b_1)\phi(az+b_1)\,dz \\
&+ \int_{-\infty}^{\infty} \frac{\varrho\tilde{\sigma}}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2+2\varrho\tilde{\sigma}z} \Phi(az+b_1)^2\,dz \\
&= 2a\,\mathbb{E}\left[e^{2\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)\phi(aZ+b_1)\right] \\
&+ \varrho\tilde{\sigma}\,\mathbb{E}\left[e^{2\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)^2\right] \qquad (4.251)
\end{aligned}
$$

and, using (4.248),

$$
\begin{aligned}
\mathbb{E}\Big[(Z-\varrho\tilde{\sigma})e^{\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)\Phi(aZ+b_2)\Big] &= \int_{-\infty}^{\infty} \frac{z-\varrho\tilde{\sigma}}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2+\varrho\tilde{\sigma}z} \Phi(az+b_1)\Phi(az+b_2)\,dz \\
&= \left[\frac{-1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2+\varrho\tilde{\sigma}z} \Phi(az+b_1)\Phi(az+b_2)\right]_{-\infty}^{\infty} \\
&+ \int_{-\infty}^{\infty} \frac{a}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2+\varrho\tilde{\sigma}z} \Phi(az+b_1)\phi(az+b_2)\,dz \\
&+ \int_{-\infty}^{\infty} \frac{a}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2+\varrho\tilde{\sigma}z} \phi(az+b_1)\Phi(az+b_2)\,dz \\
&= K^{-1}Fae^{-\frac{1}{2}\varrho^2\tilde{\sigma}^2}\,\mathbb{E}\Big[e^{2\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)\phi(aZ+b_1)\Big] \\
&+ a\,\mathbb{E}\left[e^{\varrho\tilde{\sigma}Z}\Phi(aZ+b_2)\phi(aZ+b_1)\right]. \qquad (4.252)
\end{aligned}
$$

Substituting back into (4.249) with (4.250)–(4.252), we deduce that

$$
\frac{\partial}{\partial\varrho}\mathrm{Var}_{\mathrm{mixMC}} \approx \frac{2}{M}\tilde{D}^2F^2\varrho\tilde{\sigma}^2e^{-\varrho^2\tilde{\sigma}^2}\,\mathbb{E}\left[e^{2\varrho\tilde{\sigma}Z}\Phi(aZ+b_1)^2\right]. \qquad (4.253)
$$

Finally, a change-of-variable technique leads to

$$
\frac{\partial}{\partial\varrho}\mathrm{Var}_{\mathrm{mixMC}} \approx \frac{2}{M}\tilde{D}^2F^2\varrho\tilde{\sigma}^2e^{\varrho^2\tilde{\sigma}^2}\,\mathbb{E}\left[\Phi(aZ+b)^2\right]. \qquad (4.254)
$$

Note that if $a_{11} < 1$, i.e., if $\varrho > 0$, the right-hand side of (4.254) is strictly positive. This implies that the variance of the mixed estimator decreases as a_{11} increases, and attains its minimum when $a_{11} = 1$. In fact, we know from Remark 4.23 that

$$
\mathrm{Var}_{\mathrm{mixMC}}(a_{11} = 1) \approx 0. \qquad (4.255)
$$

Moreover,

$$
\mathbb{E}\left[\Phi(aZ+b)^2\right] = \Phi\left(\frac{b}{\sqrt{1+a^2}}\right) - 2\mathcal{T}\left(\frac{b}{\sqrt{1+a^2}}, \frac{1}{\sqrt{1+2a^2}}\right), \qquad (4.256)
$$

where Owen's $\mathcal{T}$ function [78] is

$$
\mathcal{T}(\beta,\vartheta) = \phi(\beta)\int_0^{\vartheta} \frac{\phi(\beta x)}{1+x^2}\,dx, \qquad (4.257)
$$

with

$$
\mathcal{T}(\beta,1) = \frac{1}{2}\Phi(\beta) - \frac{1}{2}\Phi(\beta)^2. \qquad (4.258)
$$

Note that

$$\frac{b}{\sqrt{1+a^2}} = \underbrace{\frac{1}{\tilde{\sigma}}\log\frac{F}{K}}_{\equiv\beta_1} + \underbrace{\left(\frac{1}{2}+\varrho^2\right)\tilde{\sigma}}_{\equiv\beta_2(\varrho)} \tag{4.259}$$

and

$$\frac{1}{\sqrt{1+2a^2}} = \underbrace{\sqrt{\frac{1-\varrho^2}{1+\varrho^2}}}_{\equiv\vartheta(\varrho)}. \tag{4.260}$$

According to the data in Table 4.1, $\theta \ll 1$ and so $\bar{\gamma} \ll 1$. Hence, for low maturities T we have $\tilde{\sigma} \ll 1$. From Remark 4.21, using (4.255), we know that the variance of the standard Monte Carlo estimator is obtained by integrating (4.254) over $[0,1]$. Therefore,

$$\Gamma_{\text{var}} \approx \frac{\int_0^1 \upsilon\big[\Phi\big(\beta_1+\beta_2(\upsilon)\big) - 2\mathcal{T}\big(\beta_1+\beta_2(\upsilon),\vartheta(\upsilon)\big)\big]\,d\upsilon}{\int_0^\varrho \upsilon\big[\Phi\big(\beta_1+\beta_2(\upsilon)\big) - 2\mathcal{T}\big(\beta_1+\beta_2(\upsilon),\vartheta(\upsilon)\big)\big]\,d\upsilon}. \tag{4.261}$$

Let the function $g:(0,1)\to(0,1)$ be defined by

$$g(\upsilon) = \Phi\big(\beta_1+\beta_2(\upsilon)\big) - 2\mathcal{T}\big(\beta_1+\beta_2(\upsilon),\vartheta(\upsilon)\big). \tag{4.262}$$

Using (4.257) and the Leibniz integral rule, we compute its first derivative,

$$\begin{aligned} g'(\upsilon) &= 2\upsilon\tilde{\sigma}\phi\big(\beta_1+\beta_2(\upsilon)\big) + \frac{2\upsilon}{\sqrt{1-\upsilon^4}}\phi\big(\beta_1+\beta_2(\upsilon)\big)\phi\big((\beta_1+\beta_2(\upsilon))\vartheta(\upsilon)\big) \\ &\quad + 4\upsilon\tilde{\sigma}\phi\big(\beta_1+\beta_2(\upsilon)\big)\big(\beta_1+\beta_2(\upsilon)\big)\int_0^{\vartheta(\upsilon)} \frac{\phi\big((\beta_1+\beta_2(\upsilon))x\big)x^2}{1+x^2}\,dx \\ &\quad + 4\upsilon\tilde{\sigma}\phi\big(\beta_1+\beta_2(\upsilon)\big)\big(\beta_1+\beta_2(\upsilon)\big)\int_0^{\vartheta(\upsilon)} \frac{\phi\big((\beta_1+\beta_2(\upsilon))x\big)}{1+x^2}\,dx. \end{aligned} \tag{4.263}$$

Some straightforward calculations lead to

$$\begin{aligned} g'(\upsilon) &= \frac{2\upsilon}{\sqrt{1-\upsilon^4}}\phi\big(\beta_1+\beta_2(\upsilon)\big)\phi\big((\beta_1+\beta_2(\upsilon))\vartheta(\upsilon)\big) \\ &\quad + 4\upsilon\tilde{\sigma}\phi\big(\beta_1+\beta_2(\upsilon)\big)\Phi\big((\beta_1+\beta_2(\upsilon))\vartheta(\upsilon)\big), \end{aligned} \tag{4.264}$$

and hence the first derivative of g is strictly positive. Let the function $G:(0,1)\to(0,1)$ be defined by

$$G(\varrho) = 2\varrho^{-2}\int_0^\varrho \upsilon g(\upsilon)\,d\upsilon. \tag{4.265}$$

Using integration by parts, we compute its first derivative,

$$G'(\varrho) = 2\varrho^{-3}\int_0^\varrho \upsilon^2 g'(\upsilon)\,d\upsilon, \tag{4.266}$$

which is strictly positive, and so the function G is increasing. In particular, from (4.261), this implies that

$$\frac{1}{\varrho^2} \le \Gamma_{\text{var}}. \tag{4.267}$$

Furthermore, we deduce from (4.257)–(4.260) that, for all $\upsilon\in[0,1]$,

$$\Phi\big(\beta_1+\tfrac{1}{2}\tilde{\sigma}\big)^2 \le \Phi\big(\beta_1+\beta_2(\upsilon)\big) - 2\mathcal{T}\big(\beta_1+\beta_2(\upsilon),\vartheta(\upsilon)\big) \le \Phi\big(\beta_1+\tfrac{3}{2}\tilde{\sigma}\big). \tag{4.268}$$

Therefore, we deduce from (4.261), (4.267) and (4.268) that

$$\frac{1}{\varrho^2} \leq \Gamma_{\text{var}} \leq \frac{\Phi\left(\beta_1 + \frac{3}{2}\tilde{\sigma}\right)}{\Phi\left(\beta_1 + \frac{1}{2}\tilde{\sigma}\right)^2}\,\frac{1}{\varrho^2}. \tag{4.269}$$

The inequalities in (4.269) are approximate in the sense that the variance reduction factor is bounded from above and below by approximated quantities. Since we assumed that $\tilde{\sigma} \ll 1$, (4.269) becomes

$$\frac{1}{\varrho^2} \leq \Gamma_{\text{var}} \leq \frac{1}{\Phi\left(\beta_1\right)\varrho^2}. \tag{4.270}$$

If we also assume that the option is deep in-the-money, then β_1 is relatively large (e.g., $\beta_1 \geq 1$) such that $\sqrt{\Phi(\beta_1)} \approx 1$, and hence

$$\Gamma_{\text{dev}} \approx \left(1 - a_{11}^2\right)^{-\frac{1}{2}}, \quad \forall a_{11} > 0. \tag{4.271}$$

However, we know from Remark 4.21 that the standard deviation ratio is 1 when $a_{11} = 0$. Hence, (4.269) – (4.271) hold for all values of a_{11}. Interestingly, the deep in-the-money case is also the one where the payoff can be treated as smooth, and this also explains Figure 4.4. Note that Remark 4.25 is consistent with Remarks 4.21 – 4.24.

4.5. Numerical results

In this section, we test the efficiency of the mixed Monte Carlo/PDE method by comparison with alternative numerical schemes for two derivatives: a European call and an up-and-out put option. For the former, we use the analytical formula (4.22) to compute the conditional option price and benchmark against standard Monte Carlo with a log-Euler discretization and, under a simple correlation structure, against the semi-analytical formula of Ahlip and Rutkowski [1]. For the latter, we use either finite differences or the perturbative formula of [35] for the conditional option price, and Monte Carlo – with or without Brownian bridge – as the reference method.

First, we demonstrate the convergence of the mixed Monte Carlo/PDE method and find empirical convergence rates. Then, we investigate the sensitivity of the variance reduction factor (4.214) to changes in the model parameters and link the numerical results to the analysis in Section 4.4. Our platform for the numerical implementation was MATLAB 2012a. The machine configuration on which all numerical experiments were conducted is as follows: Intel Core i3 CPU, M370, 2.40 GHz, 8.00 GB memory (RAM), 64-bit operating system running Windows 7 Professional.

Throughout this section, we assign the following values to the underlying model parameters as a base case, and vary a selection individually or jointly:

$$v_0 = 0.0275, \qquad r_0^d = 0.0524, \qquad r_0^f = 0.0291,$$
$$k = 1.70, \qquad k_d = 0.20, \qquad k_f = 0.32,$$

$$\begin{aligned}
\theta &= 0.0232, & \theta_d &= 0.0475, & \theta_f &= 0.0248, \\
\xi &= 0.1500, & \xi_d &= 0.0352, & \xi_f &= 0.0317, \\
\rho_{sv} &= -0.10, & \rho_{sd} &= -0.15, & \rho_{sf} &= -0.15, \\
\rho_{vd} &= 0.12, & \rho_{vf} &= 0.05, & \rho_{df} &= 0.25.
\end{aligned} \tag{4.272}$$

These values are consistent with empirical observations in FX markets and are close to the calibrated values in Tables 4.1 and 4.2. Also, the values of the correlations ρ_{sd}, ρ_{sf} and ρ_{df} are borrowed from [79].

4.5.1. European call option

Let the spot, the strike and the maturity be $S_0 = 105$, $K = 100$ and $T = 1.5$. As an aside, note that FX option quotes are in terms of volatilities for a fixed delta and a fixed time to expiry, and not in terms of strikes. However, the strike price corresponding to a quoted volatility can easily be recovered using a conversion formula [1]. Recall that Θ from (4.209) is the true option price, and Θ_{stdMC} from (4.210) and Θ_{mixMC} from (4.211) are the standard and mixed Monte Carlo estimators, respectively. In Table 4.5, we report the common discretization bias, i.e., the time-discretization error, which is the same for both estimators by Remark 4.20, and the two standard errors, i.e., the two standard deviations of the sample means, which we denote by StDev and estimate using 10000 samples. However, according to [1], a closed-form solution for the European option price is not available under a full correlation structure. Hence, we need to find an accurate reference estimate Θ^* in order to evaluate the bias. We employ the mixed algorithm with $M = 2\times10^9$ simulations and $N = 200$ time steps to find $\Theta^* = 12.11968$. Then, the bias for a specific number of time steps is estimated using the reference estimate Θ^*, whose accuracy is discussed below, and a sufficiently large number of simulations, i.e., $M = 2\times10^9$.

The time needed to obtain a call price estimate with 64000 simulations and 8 time steps is 0.25 seconds for the standard Monte Carlo method and 0.22 seconds for the mixed Monte Carlo method. Hence, the computational cost is 12% lower with the latter. This, however, is to be expected since we only simulate three of the four underlying processes.

Table 4.5: Simulation results for a European call option.

Time steps	Discretization bias	Simulations	StDev (stdMC)	StDev (mixMC)
1	0.37231	1000	0.50348	0.09342
2	0.08039	4000	0.24126	0.04194
4	0.01775	16000	0.12042	0.02023
8	0.00444	64000	0.06071	0.00994
16	0.00160	256000	0.02932	0.00487
32	0.00073	1024000	0.01507	0.00262

The data in Table 4.5 indicate that the bias is small even when only a few time steps are used. This phenomenon can be explained by the fact that the Feller condition is satisfied, i.e., $2k\theta \geq \xi^2$, $2k_d\theta_d \geq \xi_d^2$ and $2k_f\theta_f \geq \xi_f^2$, and hence the discretizations of the variance and interest rate processes rarely hit zero. The simulation results in Table 4.5 confirm the square-root convergence of the statistical error and the first-order convergence of the discretization error. Then, using extrapolation, we obtain an approximate root mean square error (RMSE) of the reference estimate:

$$\begin{gathered} \mathrm{Bias}\,(\Theta^*) \approx 1.168 \times 10^{-4}, \quad \mathrm{StDev}\,(\Theta^*) \approx 0.593 \times 10^{-4}, \\ \Rightarrow \quad \mathrm{RMSE}\,(\Theta^*) \approx 1.310 \times 10^{-4}. \end{gathered} \tag{4.273}$$

This is equivalent to an RMSE of about 0.001% of the actual option price, suggesting that the reference estimate $\Theta^* = 12.11968$ is accurate to three decimal places.

Furthermore, the standard deviation is reduced by 83%, i.e., is approximately six times lower with the mixed method. However, the variance reduction is not consistent across the range of possible values of the model parameters, and is most sensitive to changes in the correlations between the exchange rate and the squared volatility or the interest rates, i.e., ρ_{sv}, ρ_{sd} and ρ_{sf}, the speed of mean reversion k and the volatility of volatility ξ.

The values $\rho_{vd} = 0.12$, $\rho_{vf} = 0.05$, $\rho_{df} = 0.25$ and $\rho_{sv}, \rho_{sd}, \rho_{sf} \in [-0.5, 0.5]$ lead to a valid (i.e., positive definite) correlation matrix. The data in Figure 4.1 suggest that the highest variance reduction is achieved when the absolute values of the correlations ρ_{sv}, ρ_{sd} and ρ_{sf} are small or, more precisely, when $\rho_{sv} \approx -0.05$ and $\rho_{sd} \approx \rho_{sf} \approx 0$, in which case the standard deviation is reduced by a factor of 20. Hence, the best performance of the mixed estimator coincides with the independence of W^s from the Brownian motions W^v, W^d and W^f, an observation which is consistent with Remarks 4.22 and 4.23.

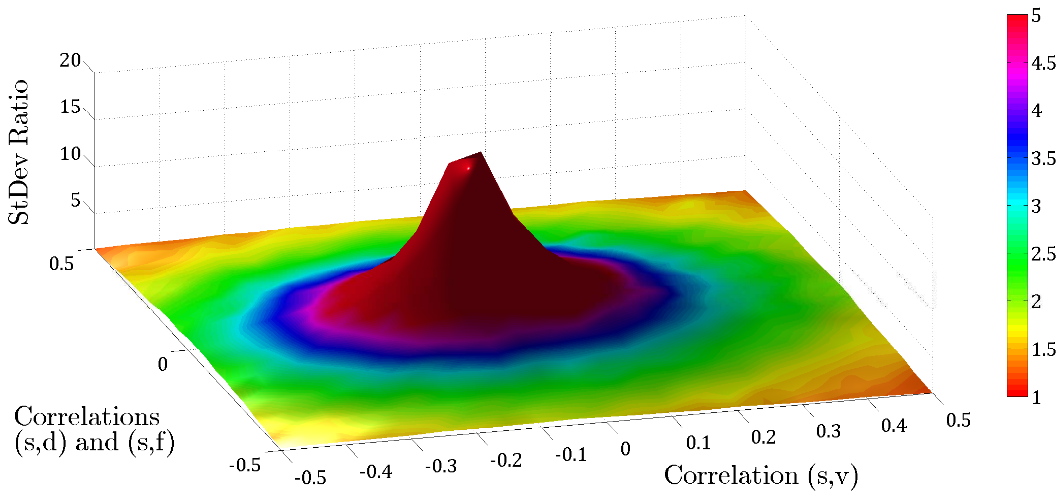


Figure 4.1: The standard deviation ratio Γ_{dev} from (4.214) with 4000 simulations and 10 time steps, plotted against the correlation coefficients ρ_{sv}, ρ_{sd} and ρ_{sf} when the latter two are equal.

Furthermore, we may infer from Figure 4.1 that, in addition to the lower computational cost, the mixed Monte Carlo/PDE method outperforms standard Monte Carlo in terms of accuracy for all $\rho_{sv}, \rho_{sd}, \rho_{sf} \in [-0.5, 0.5]$, and that the variance reduction factor decreases as the absolute values of the correlations between the exchange rate and the squared volatility or the interest rates increase, two facts that were recognized in Remark 4.21.

Assuming that $\rho_{sd} = \rho_{sf}$, since $\rho_{vd} \approx \rho_{vf} \approx 0$, we may approximate

$$a_{11} \approx \sqrt{1 - \rho_{sv}^2 - \frac{2}{1 + \rho_{df}} \rho_{sd}^2} = \sqrt{1 - \rho_{sv}^2 - \frac{8}{5} \rho_{sd}^2}. \tag{4.274}$$

However,

$$\Gamma_{\text{dev}} \approx \left(1 - a_{11}^2\right)^{-\frac{1}{2}} \approx \left(\rho_{sv}^2 + 1.6\rho_{sd}^2\right)^{-\frac{1}{2}} \tag{4.275}$$

from Remark 4.25. Combined with the above observation on the location of the maximum variance reduction factor, we infer that the set of points (ρ_{sv}, ρ_{sd}) corresponding to $\Gamma_{\text{dev}} = \mu$ for some $\mu \geq 1$ approximately takes the form of an ellipse with the equation

$$\mu^2 \left(\rho_{sv} + 0.05\right)^2 + 1.6\mu^2 \rho_{sd}^2 = 1. \tag{4.276}$$

This confirms that the isolines in Figure 4.1 are approximately elliptical, a fact also illustrated in Table 4.6, where the correlation pairs (ρ_{sv}, ρ_{sd}) are chosen so that $\mu = 2.812$. The estimated standard deviation ratio along the contour line is 3.148 ± 0.239, which is close to the theoretical value μ. This attests to the accuracy of our approximations and numerical results.

Table 4.6: The estimated standard deviation ratio for different correlations.

ρ_{sv}	ρ_{sd}	Γ_{dev}
−0.40	−0.05	3.288
−0.40	0.05	3.088
−0.30	−0.20	3.387
−0.30	0.20	2.910
0.20	−0.20	2.951
0.20	0.20	3.127
0.30	−0.05	2.952
0.30	0.05	3.055

Next, we assume a partial correlation structure between the Brownian drivers such that the squared volatility dynamics are independent of the domestic and foreign interest rate dynamics, i.e., $\rho_{vd} = \rho_{vf} = 0$. Using Remark 4.25 and computing a_{11} explicitly from (4.2), we obtain

$$\Gamma_{\text{dev}} \approx \left[\rho_{sv}^2 + \frac{1}{1 - \rho_{df}^2} \left(\rho_{sd}^2 + \rho_{sf}^2 - 2\rho_{sd}\rho_{sf}\rho_{df}\right)\right]^{-\frac{1}{2}}. \tag{4.277}$$

Before, we fixed ρ_{df} and analyzed the variance reduction with respect to ρ_{sv}, ρ_{sd} and ρ_{sf}, when the latter two were equal. Now, we fix ρ_{sv} instead and focus on the effect of varying ρ_{df} on Γ_{dev}. To this end, we can easily show that

$$\frac{\rho_{sd}^2 + \rho_{sf}^2 - 2\rho_{sd}\rho_{sf}\rho_{df}}{1 - \rho_{df}^2} \geq \max\left\{\rho_{sd}^2, \rho_{sf}^2\right\}. \tag{4.278}$$

Assuming that ρ_{sd} and ρ_{sf} are not simultaneously zero, equality in (4.278) holds when

$$\rho_{df} = \rho_{df}^* \equiv \frac{\rho_{sf}}{\rho_{sd}} \mathbb{1}_{|\rho_{sd}| \geq |\rho_{sf}|} + \frac{\rho_{sd}}{\rho_{sf}} \mathbb{1}_{|\rho_{sd}| < |\rho_{sf}|}\,. \tag{4.279}$$

Upon its substitution into (4.277), we find a theoretical standard deviation ratio μ. Table 4.7 shows estimates of Γ_{dev} and compares them with μ when $\rho_{sv} = -0.10$, $\rho_{sf} = -0.20$, $\rho_{sd} \in [-0.25, 0.25]$ and $\rho_{df} \in [-0.85, 0.85]$. Note that these correlation values guarantee a symmetric positive definite correlation matrix. Hence, when ρ_{df}^* lies outside the interval of permitted values, we estimate Γ_{dev} using $\rho_{df} = \pm 0.85$ instead. The data in Table 4.7 suggest that the approximation (4.277) to the standard deviation ratio is fairly accurate, especially for a strongly negative correlation between the two interest rates.

Table 4.7: The empirical and theoretical standard deviation ratios for different correlations.

ρ_{sd}	ρ_{df}^*	Γ_{dev}	μ
-0.25	0.80	4.501	3.714
-0.20	1.00	5.301	4.472
-0.15	0.75	5.305	4.472
-0.10	0.50	5.224	4.472
-0.05	0.25	5.073	4.472
0	0	5.010	4.472
0.05	-0.25	4.850	4.472
0.10	-0.50	4.867	4.472
0.15	-0.75	4.554	4.472
0.20	-1.00	4.467	4.472
0.25	-0.80	3.788	3.714

By close inspection of the data in Figure 4.2, we infer that, for each $\rho_{sd} \in [-0.25, 0.25]$, the highest variance reduction is achieved for ρ_{df}^* as defined in (4.279). Hence, we conclude that μ exhibits the qualitative behaviour of Γ_{dev}, so (4.277) provides a good approximation to the standard deviation ratio. In fact, our observations suggest that μ acts as a lower bound. We can extend these results to a full correlation structure between the Brownian drivers as long as ρ_{vd} and ρ_{vf} are close to zero, as seen before.

Suppose that the exchange rate dynamics are independent of the interest rate dynamics, i.e., that $\rho_{sd} = \rho_{sf} = 0$, and define the optimal correlation ρ_{sv}^* to be the value corresponding to the maximum standard deviation ratio for a given volatility of volatility ξ. Then the data in Figure 4.3 suggest that the highest variance reduction is achieved when the absolute

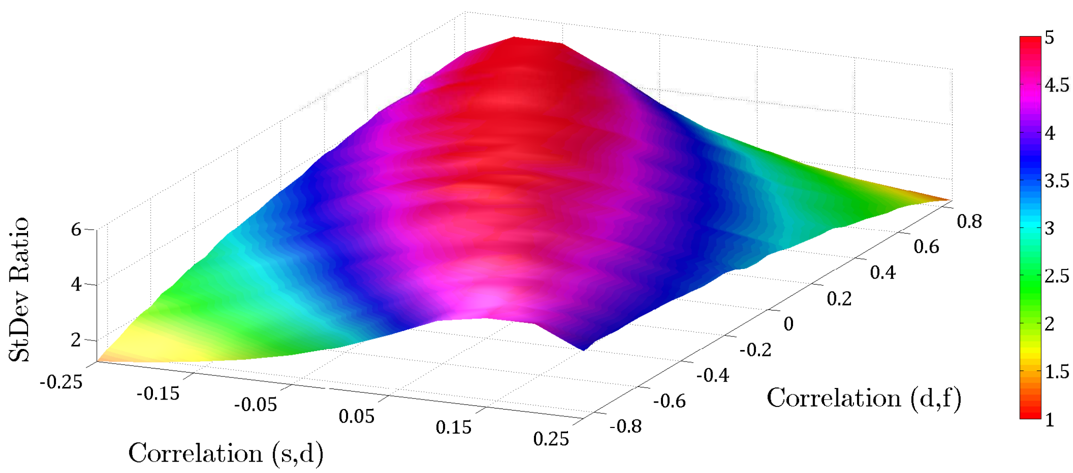


Figure 4.2: The standard deviation ratio with 1000 simulations and 4 time steps, plotted against the correlation coefficients ρ_{sd} and ρ_{df} when $\rho_{sv} = -0.10$, $\rho_{sf} = -0.20$ and $\rho_{vd} = \rho_{vf} = 0$.

value of the correlation is small. In fact, we notice two things. First, the optimal correlation approaches zero as the volatility of volatility decreases, i.e., that $\lim_{\xi \to 0} \rho^*_{sv} = 0$, which is to be expected from Remark 4.22. Second, the standard deviation ratio at ρ^*_{sv} increases as the volatility of volatility decreases. In practice, $\xi_{d,f} \ll 1$ (see Table 4.2), so the two interest rates have little impact on the variance of the mixed Monte Carlo/PDE estimator. Therefore, since ρ^*_{sv} is close to zero, the variance comes mainly from σ defined in (4.16). On the other hand, smaller values of ξ result in a smaller variance of σ, and hence of the mixed estimator as well, which leads to a higher standard deviation ratio.

We observe a similar behaviour when increasing the speed of mean reversion k or the long-run variance θ. Larger values of k result in a smaller variance of σ because of the mean-reverting property of the squared volatility, which ensures that the process returns to the long-run average quickly. On the other hand, larger values of θ produce higher volatilities, which then lead to an increase in the variance of the standard Monte Carlo estimator due to the larger diffusion term in the SDE driving the exchange rate process. We conclude the analysis by noting that values of ρ_{sd} and ρ_{sf} close to zero produce similar results, to some extent. However, when the absolute values of the two correlations are not small, the impact of σ, and hence of k, ξ and θ, on the variance of the mixed estimator is reduced.

The maximum in Figure 4.3 is attained around $\rho_{sv} = 0$ and $\xi = 0.05$, where the standard deviation ratio is $\Gamma_{\text{dev}} = 23$. Therefore, the same level of accuracy with the standard Monte Carlo method requires 529 times more simulations.

Next, we examine the variance reduction for different spots and maturities. The data in Figure 4.4 suggest that, unless the option is far out-of-the-money and the maturity is small, varying S_0 and T has little impact on the standard deviation ratio, which is approx-

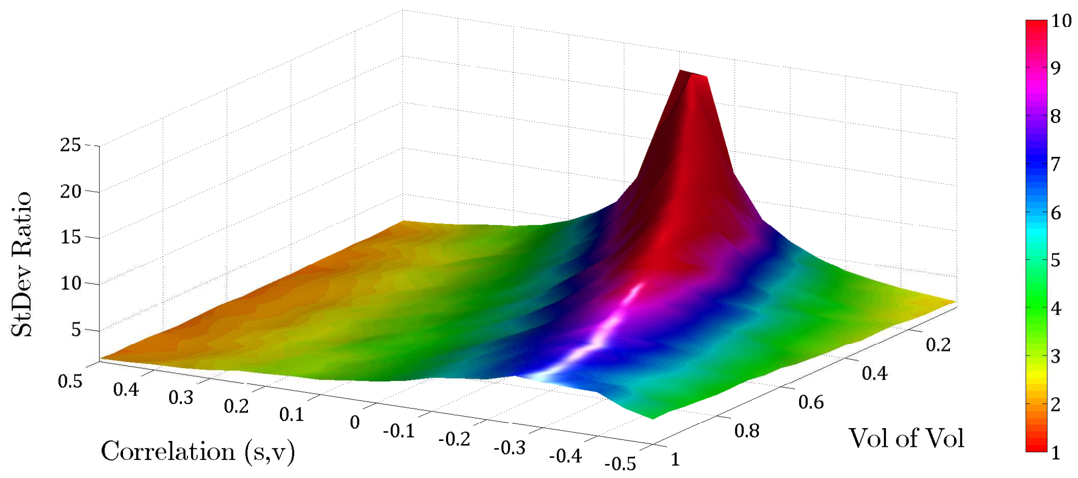


Figure 4.3: The standard deviation ratio with 4000 simulations and 10 time steps, plotted against the correlation ρ_{sv} and the volatility of volatility ξ when $\rho_{sd} = \rho_{sf} = 0$.

imately 2. Computing the option price with standard Monte Carlo means integrating the payoff function, which is not differentiable at the strike, whereas with the mixed Monte Carlo/PDE method we integrate the smooth conditional price. As the probability of a positive payoff decreases with S_0 and T, estimating it accurately with the former requires more simulations. Hence, the relative standard error of the standard Monte Carlo method increases as we go farther out-of-the-money or approach maturity, and the benefit of employing the mixed algorithm becomes clear. For example, for a three-month call with a spot at 75% of the strike, we observe a variance reduction factor of 5275.

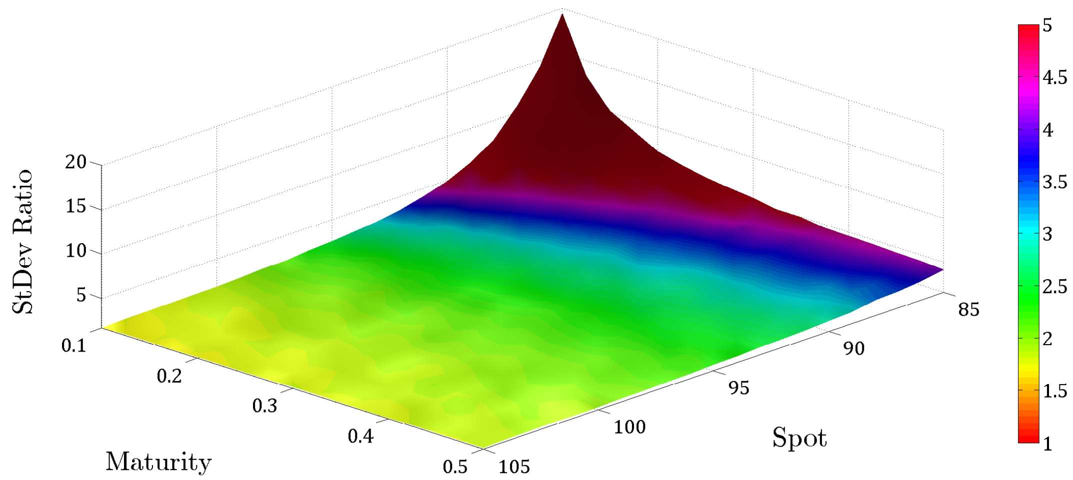


Figure 4.4: The standard deviation ratio with 1000 simulations and 4 time steps, plotted against the spot S_0 and the maturity T when $\rho_{sv} = -0.60$.

Furthermore, Figure 4.4 suggests that the approximation for the standard deviation ratio derived in (4.271) holds when the option is in-the-money, but not when the option is far out-of-the-money. For instance, we compute $a_{11} = 0.7893$, which gives a theoretical standard deviation ratio $\mu = 1.629$. When $S_0 = 105$ and $T = 0.5$, this value is close to the estimated value $\Gamma_{\text{dev}} = 1.832$. However, when $S_0 = 85$ and $T = 0.1$, we observe a much higher standard deviation ratio, namely $\Gamma_{\text{dev}} = 19.863$.

Suppose that the domestic and the foreign short rate dynamics are independent of each other, and also independent of the dynamics of the exchange rate and its volatility, and let $\rho_{sv} = -0.10$, as before. Moreover, suppose that the other model parameters, as well as the spot, the strike and the maturity, take the values listed at the beginning of this section. We use the mixed Monte Carlo/PDE method with $M = 8{\times}10^7$ simulations and $N = 200$ time steps, for an RMSE of about 0.001%, to obtain a call price estimate $\Theta' = 12.13621$, which is about 0.14% higher than the estimate corresponding to a full correlation of the factors, i.e., $\Theta^* = 12.11968$. On the one hand, the postulated independence of the factors is critical from the point of view of analytical tractability [1], but can result in fairly different option prices. On the other hand, a full correlation structure leads to a richer, more realistic model, which is of practical importance for the pricing and hedging of long-term options, and a better fit to the observed market data. We test the accuracy of the mixed method and employ the semi-analytical pricing formula of [1] to find the true option price, $\Theta = 12.13603$, and thus a relative error of Θ' of about 0.0015%, which confirms that mixed Monte Carlo/PDE estimates are correct.

Finally, we briefly discuss how to further reduce the variance of both standard and mixed Monte Carlo estimates using control variates. For the standard Monte Carlo estimator, we may use as a control variate the discounted call payoff under a simple correlation structure (where only ρ_{sv} may be non-zero), since its expectation is known in closed-form [1], which leads to a substantial variance reduction. For the mixed Monte Carlo estimator, we may use as control variates the conditional call price (conditional on the trajectories of the same Brownian motions) under the 4-factor FX model with deterministic interest rates, since its expectation is known in closed-form [46], together with the domestic and foreign discount factors from (3.6), since the corresponding bond prices are known in closed-form [13], which leads to a similar variance reduction. However, the mixed Monte Carlo/PDE method can be applied to a large class of models beyond the one studied in this chapter, and a detailed performance analysis of these control variates is beyond the scope of this thesis.

4.5.2. Up-and-out put option

Let the spot, the strike, the barrier and the maturity be $S_0 = 100$, $K = 105$, $B = 110$ and $T = 0.25$, and consider a continuously monitored up-and-out put option. We will first value the contract using the mixed Monte Carlo/PDE method. Hence, for a specific realization of the variance and interest rates paths, we compute the conditional option price, i.e.,

$$u(t,x) = \mathbb{E}\left[e^{-\int_t^T \bar{r}_u^d\,du}\left(K - \bar{S}_T\right)^+ \mathbb{1}_{\max_{t\le u\le T}\bar{S}_u < B} \,\middle|\, \mathcal{G}_T^{f,d,v},\, \bar{S}_t = x\right]. \tag{4.280}$$

We know that u satisfies the following initial-boundary-value problem:

$$\begin{aligned} \partial_t u + \mu_t x\,\partial_x u + \frac{1}{2}a_{11}^2\bar{v}_t x^2\partial_{xx}u - \bar{r}_t^d u &= 0, && \forall 0 < x < B,\ t < T, \\ u(t,B) &= 0, && \forall t \le T, \\ u(T,x) &= (K-x)^+, && \forall 0 \le x < B, \end{aligned} \tag{4.281}$$

with $\mu = (\mu_t)_{t\ge 0}$ from (4.13). We solve the (random) PDE backward in time from the initial condition, on a rectangular domain with $t \in [0,T]$ and $x \in [0.7S_0, B]$ that is discretized on a uniform grid with $N+1$ temporal nodes and $L+1$ spatial nodes, for a time step $\delta t = T/N$ and a spatial step $\delta x = (B - 0.7S_0)/L$. We selected this particular lower boundary of the spatial computational domain to reduce the number of spatial nodes, while making sure that the truncation error arising from our choice of the domain is negligible. We employ a central difference scheme to approximate the spatial derivatives and a linearity boundary condition [86] stating that $\partial_{xx}u = 0$ at the lower boundary, where the option is deep-in-the-money and the price can be regarded as linear in x. For convenience, we employ the same time grid as used for the discretization of the squared volatility and the interest rates. For a second-order convergence in both space and time, we use the Crank-Nicolson scheme. As an aside, note that an interesting and more stable alternative would be the second-order backward differentiation formula (BDF) scheme [37].

To each pair of indices (i,j), for all $0 \le i \le N$ and $0 \le j \le L$, corresponds a node $(t_i, x_j) = (i\delta t, 0.7S_0 + j\delta x)$ and a value $V_j^i = u(t_i, x_j)$. Hence, the initial and boundary conditions are

$$V_j^N = \left(K - x_j\right)^+,\ \ \forall 0 \le j < L, \tag{4.282}$$

$$V_L^i = 0,\ \ \forall 0 \le i \le N, \tag{4.283}$$

$$V_0^i = 2V_1^i - V_2^i,\ \ \forall 0 \le i < N, \tag{4.284}$$

and the PDE in (4.281) becomes

$$\begin{aligned} \frac{V_j^{i+1} - V_j^i}{\delta t} &+ \frac{1}{2}a_j^i\,\frac{V_{j+1}^i - V_{j-1}^i}{2\delta x} + \frac{1}{2}a_j^{i+1}\,\frac{V_{j+1}^{i+1} - V_{j-1}^{i+1}}{2\delta x} \\ &+ \frac{1}{2}b_j^i\,\frac{V_{j+1}^i - 2V_j^i + V_{j-1}^i}{(\delta x)^2} + \frac{1}{2}b_j^{i+1}\,\frac{V_{j+1}^{i+1} - 2V_j^{i+1} + V_{j-1}^{i+1}}{(\delta x)^2} \\ &+ \frac{1}{2}c^i V_j^i + \frac{1}{2}c^{i+1}V_j^{i+1} = 0, \end{aligned} \tag{4.285}$$

where

$$a_j^i = \mu_{t_i}x_j,\ \ b_j^i = \frac{1}{2}a_{11}^2\bar{v}_{t_i}x_j^2 \quad \text{and} \quad c^i = -\bar{r}_{t_i}^d. \tag{4.286}$$

Rearranging terms in (4.285), we deduce that

$$-A_j^i V_{j-1}^i + \left(1 - B_j^i\right) V_j^i - C_j^i V_{j+1}^i = A_j^{i+1} V_{j-1}^{i+1} + \left(1 + B_j^{i+1}\right) V_j^{i+1} + C_j^{i+1} V_{j+1}^{i+1}, \quad (4.287)$$

for all $0 \leq i \leq N-1$ and $1 \leq j \leq L-1$, where $\nu_1 = \delta t/\delta x$, $\nu_2 = \delta t/(\delta x)^2$, and

$$A_j^i = -\frac{1}{4}\nu_1 a_j^i + \frac{1}{2}\nu_2 b_j^i, \;\; B_j^i = -\nu_2 b_j^i + \frac{1}{2}\delta t\, c^i \;\text{ and }\; C_j^i = \frac{1}{4}\nu_1 a_j^i + \frac{1}{2}\nu_2 b_j^i. \quad (4.288)$$

Using (4.283) and (4.284), and since the latter also holds when $i = N$ for $L \geq 3$, we can rewrite the problem in matrix notation as

$$\begin{pmatrix} 1-2A_1^i-B_1^i & A_1^i - C_1^i & 0 & \cdots\cdots & & 0 \\ -A_2^i & 1-B_2^i & -C_2^i & \cdots\cdots & & 0 \\ 0 & -A_3^i & 1-B_3^i & \cdots\cdots & & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & \cdots\cdots & & -A_{L-2}^i & 1-B_{L-2}^i & -C_{L-2}^i \\ 0 & \cdots\cdots & & 0 & -A_{L-1}^i & 1-B_{L-1}^i \end{pmatrix} \begin{pmatrix} V_1^i \\ V_2^i \\ V_3^i \\ \vdots \\ V_{L-2}^i \\ V_{L-1}^i \end{pmatrix}$$

$$= \begin{pmatrix} 1+2A_1^{i+1}+B_1^{i+1} & C_1^{i+1} - A_1^{i+1} & 0 & \cdots\cdots & & 0 \\ A_2^{i+1} & 1+B_2^{i+1} & C_2^{i+1} & \cdots\cdots & & 0 \\ 0 & A_3^{i+1} & 1+B_3^{i+1} & \cdots\cdots & & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & \cdots\cdots & & A_{L-2}^{i+1} & 1+B_{L-2}^{i+1} & C_{L-2}^{i+1} \\ 0 & \cdots\cdots & & 0 & A_{L-1}^{i+1} & 1+B_{L-1}^{i+1} \end{pmatrix} \begin{pmatrix} V_1^{i+1} \\ V_2^{i+1} \\ V_3^{i+1} \\ \vdots \\ V_{L-2}^{i+1} \\ V_{L-1}^{i+1} \end{pmatrix}. \quad (4.289)$$

The tridiagonal linear system can be solved efficiently with the tridiagonal matrix algorithm (Thomas algorithm), in which case the solution is recovered in $\mathcal{O}(L)$ operations at each timestep. The final estimate of the barrier option price is a Monte Carlo average over a sufficiently large number of discrete trajectories of the Brownian motions W^2, W^3 and W^4.

Alternatively, we can use the perturbative formula of [35] to approximate the conditional option price, and then a simple Monte Carlo average to estimate the outer expectation. Hence, we call this numerical scheme the mixed Monte Carlo/Pert method. Fatone et al. [35] approximated the up-and-out put option price in the Black–Scholes model with time-dependent parameters via a series expansion and provided explicit formulae for the first three terms, which involve some elementary and nonelementary transcendental functions. However, we will focus only on the zeroth-order approximation, because using a first-order correction term results in a hundredfold increase in computation time, and hence in a poor performance of the scheme compared with the mixed Monte Carlo/PDE method. Using $M = 4{\times}10^7$ simulations and $N = 200$ time steps to minimize the sampling and the discretization errors, respectively, we obtain a zeroth-order approximation: $\Theta^0 = 5.7700$.

Since a closed-form solution to the option pricing problem is not available, we need to find an accurate reference estimate Θ^* for the true option price Θ in order to compute the different errors of the numerical methods. Therefore, we use the mixed Monte Carlo/PDE algorithm with the Crank-Nicolson scheme, with $M = 4{\times}10^7$ simulations, $N = 200$ time steps and $L = 20$ spatial steps, to find $\Theta^* = 5.7631$. Hence, the approximation error of the mixed Monte Carlo/Pert method is $\Theta^0 - \Theta^* = 0.0069$, i.e., about 0.1%.

In Figure 4.5, we report the time-discretization errors computed using the reference estimate Θ^* – whose accuracy is discussed below – or Θ^0, and a sufficiently large number of simulations and spatial steps, i.e., $M = 4{\times}10^7$ and $L = 20$. For the standard Monte Carlo and the mixed Monte Carlo/Pert methods, the time-discretization error is defined as the bias, whereas for the mixed Monte Carlo/PDE algorithm, due to our choice of the finite difference grid, it includes the finite difference (FD) time-discretization error. Henceforth, the term "discretization error" stands for the time-discretization error.

As an aside, note that crossings of the barrier are monitored only at discrete times by standard Monte Carlo. This gives rise to a monitoring error, which is included in the discretization error. Moreover, due to the knock-out feature of the option, the true price is smaller than the Monte Carlo estimate, which explains the strong positive bias displayed in Figure 4.5.

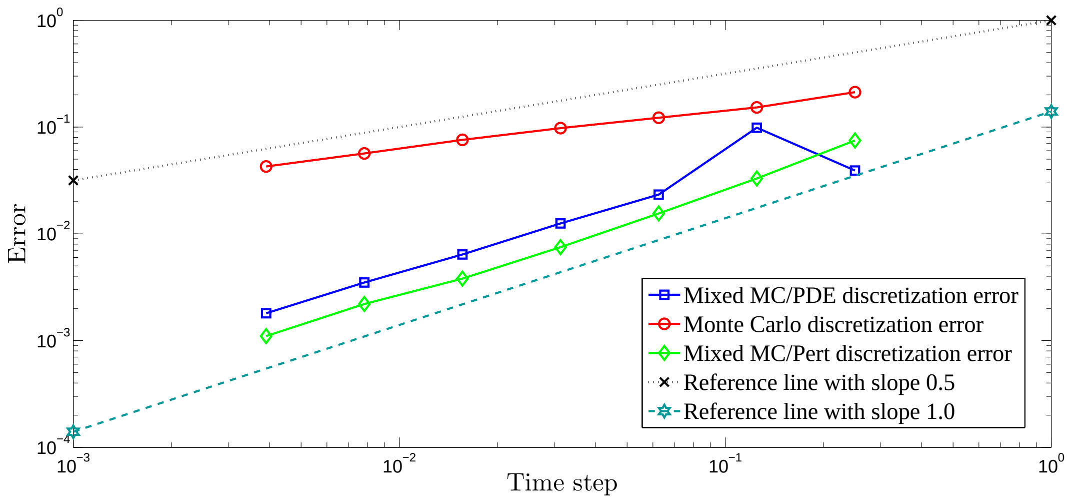


Figure 4.5: The log-plot of time-discretization errors of the mixed Monte Carlo/PDE, the mixed Monte Carlo/Pert and the standard Monte Carlo methods against the time step.

The data in Figure 4.5 suggest a square-root convergence of standard Monte Carlo and a first-order convergence of the mixed algorithms. However, we can use a Brownian bridge technique [41] to improve the first method. For the log-Euler scheme defined in (4.206), the conditional trigger probability is

$$p_n = \begin{cases} \exp\left\{-\dfrac{2(\hat{x}_{t_{n-1}}-b)(\hat{x}_{t_n}-b)}{\bar{v}_{t_{n-1}}\delta t}\right\} & \text{, if } \max\{\hat{x}_{t_{n-1}},\hat{x}_{t_n}\}<b,\ \bar{v}_{t_{n-1}}>0, \\ 0 & \text{, if } \max\{\hat{x}_{t_{n-1}},\hat{x}_{t_n}\}<b,\ \bar{v}_{t_{n-1}}=0, \\ 1 & \text{, if } \max\{\hat{x}_{t_{n-1}},\hat{x}_{t_n}\}\geq b, \end{cases} \tag{4.290}$$

for all $1 \leq n \leq N$, where $b = \log B$, and we approximate the option price by a Monte Carlo average of

$$e^{-\int_0^T \bar{r}_t^d\, dt}\big(K - e^{\hat{x}_T}\big)^+ \prod_{n=1}^{N}\big(1-p_n\big). \tag{4.291}$$

Using the Brownian bridge Monte Carlo method, we recover the first-order convergence. Indeed, the (red) Monte Carlo curve in Figure 4.5 would almost coincide with the (green) mixed Monte Carlo/Pert curve with the Brownian bridge correction. For instance, we calculated the discretization bias with $N = 8$ time steps to be 0.0075 for both Monte Carlo with Brownian bridge and mixed Monte Carlo/Pert.

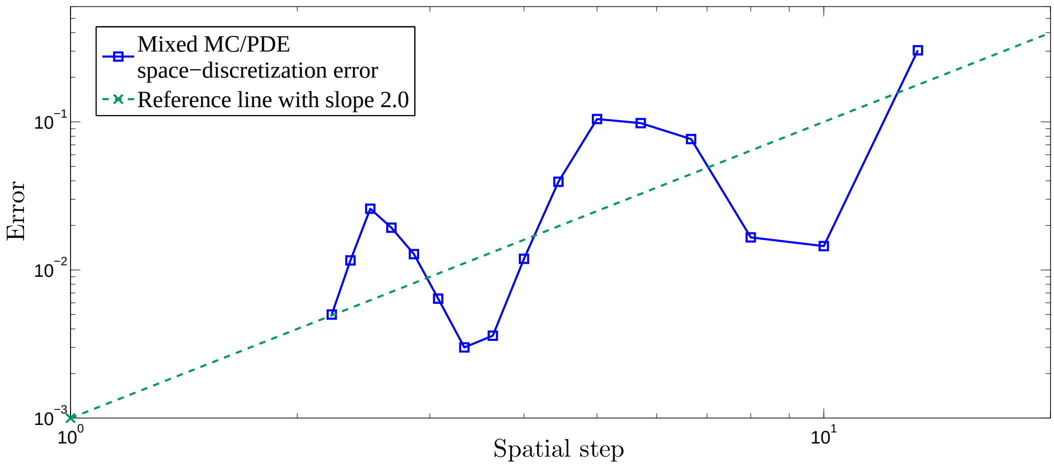


Figure 4.6: The log-plot of the absolute space-discretization error of the mixed Monte Carlo/PDE method against the spatial step.

In Figure 4.6, we report the space-discretization error computed using Θ^* and a sufficiently large number of simulations and time steps, i.e., $M = 4{\times}10^7$ and $N = 200$. The data suggest a second-order convergence as well as a local minimum discretization error when the strike price lies halfway between two adjacent nodes, a technique known as grid-shifting [86]. Hence, considering the first-order convergence of the time-discretization error (T-Err) and the second-order convergence of the finite difference space-discretization error (S-Err) with the mixed Monte Carlo/PDE method, and using extrapolation, we obtain an approximate root mean square error (RMSE) of the reference estimate:

$$\begin{gathered} \text{T-Err} \approx 5.76\times10^{-4},\ \ \text{S-Err} \approx 10.80\times10^{-4},\ \ \text{StDev} \approx 2.09\times10^{-4}, \\ \Rightarrow\ \ \text{RMSE} \approx 1.67\times10^{-3}. \end{gathered} \tag{4.292}$$

This is equivalent to an RMSE of about 0.03% of the actual option price, suggesting that the reference estimate $\Theta^* = 5.7631$ is accurate to two decimal places. Next, using the empirical convergence rates determined above and extrapolation, we compare the three numerical methods in terms of computation time for a given level of accuracy, in particular when the RMSE is at most 0.30% of the option price.

Table 4.8: The number of sample paths, time steps and spatial steps for an RMSE of at most 0.30% of the option price and the CPU time (in seconds) with the standard Monte Carlo, the mixed Monte Carlo/Pert (with a zeroth-order approximation) and the mixed Monte Carlo/PDE methods.

Method	MC paths	Time steps	Spatial steps	CPU time
MC	250000	800	–	61.2
MC/Pert	12000	10	–	3.1
MC/PDE	12000	10	12	2.0

Therefore, when the up-and-out put option price estimates need not be too accurate, e.g., when one decimal place of accuracy is sufficient, the two mixed algorithms are comparable in terms of CPU time and efficiency and are considerably faster than standard Monte Carlo. However, a higher accuracy would require at least a first-order correction term in the mixed Monte Carlo/Pert approximation, making it highly time-consuming. We thus conclude that the mixed Monte Carlo/PDE method is the best of the three schemes.

We mentioned above that the Monte Carlo method with Brownian bridge recovers the observed first-order convergence of the discretization error and the level of the bias from the mixed Monte Carlo/Pert method. The time-discretization error with the mixed Monte Carlo/PDE method is approximately 1.6 times larger, and includes the FD time-discretization error.

For a two-decimal-place accuracy, we fix 100 time steps and 20 spatial steps, such that the space- and time-discretization errors are about 0.02%. The time required to obtain a barrier option price estimate with 40000 simulations is then 26 seconds for the mixed Monte Carlo/PDE method and 2.1 seconds for Monte Carlo with Brownian bridge (1.4 seconds for standard Monte Carlo). Therefore, the computational cost is 92% lower with the latter. In conclusion, due to the square-root convergence of the standard deviation, Γ_{dev} needs to be above 4.5 in order for the mixed Monte Carlo/PDE method to outperform Monte Carlo with Brownian bridge. Just as in the European call option case, the variance reduction is most sensitive to changes in the correlations between the exchange rate and the squared volatility or the interest rates.

Figure 4.7 exhibits similar characteristics to those of Figure 4.1. In particular, the highest variance reduction is achieved when $\rho_{sv} \approx 0.05$ and $\rho_{sd} \approx \rho_{sf} \approx 0$, in which case $\Gamma_{\text{dev}} = 40$. Based on a previous observation, this is equivalent to an 80 times lower computational effort with the mixed algorithm. A careful inspection of the data in Figure

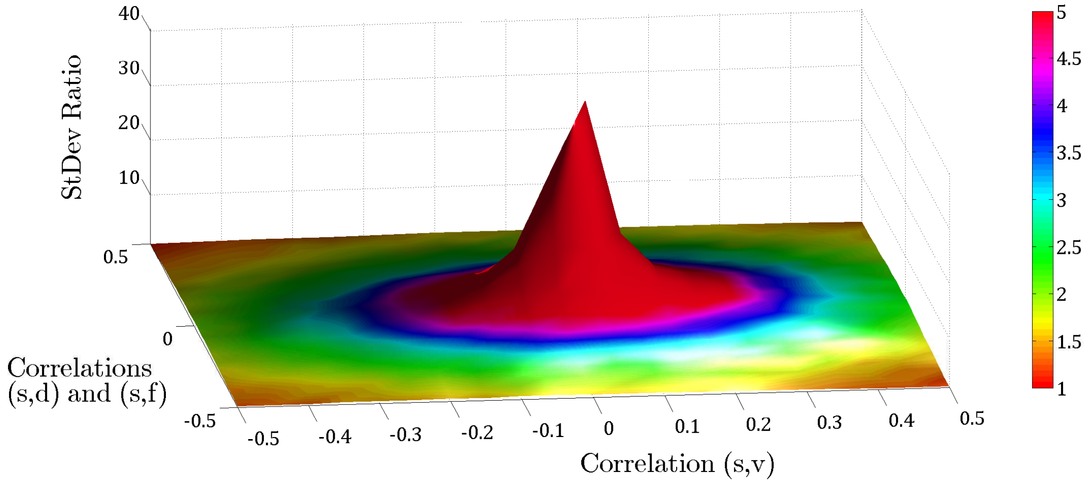


Figure 4.7: The standard deviation ratio with 400 simulations, 10 time steps and 4 spatial steps, plotted against the correlation coefficients ρ_{sv}, ρ_{sd} and ρ_{sf} when the latter two are equal.

4.7 suggests that the set of points (ρ_{sv}, ρ_{sd}) corresponding to $\Gamma_{\text{dev}} > 4.5$ can be described approximately by the following inequality:

$$\left(\rho_{sv} - 0.05\right)^2 + 1.6\rho_{sd}^2 < 4.2^{-2}, \tag{4.293}$$

i.e., the inside of an ellipse, this being the set of parameters where the benefit of variance reduction outweighs the additional complexity of solving the conditional PDE numerically in this instance.

The variance reduction achieved by the mixed method results in computational savings (in the number of samples) by a factor roughly independent of the desired accuracy. Conversely, higher accuracy of the finite difference method can only be achieved by a larger number of mesh points. Hence, it would appear that for high enough accuracy the mixed method can never win over the standard Monte Carlo method. An asymptotic complexity gain of the mixed method for small errors would require that the PDE can be solved with constant effort independent of the desired accuracy (and for this effort to be outweighed by the reduced number of samples required for a given statistical error). Multilevel Monte Carlo methods [40] recently developed for stochastic PDEs are designed precisely to achieve this goal, by concentrating the dominant number of samples on the coarsest meshes while computing corrections on finer meshes with a vanishing number of paths. The application and numerical analysis in the present context is the subject of further research.

4.6. Conclusions

The numerical experiments carried out in Section 4.5 suggest that the mixed method outperforms both standard Monte Carlo and finite difference methods under certain circumstances depending on the contract and the model parameters. When a closed-form solution for the conditional option price is available, we usually see a considerable improvement in both accuracy and computation time. When it is not, the mixed algorithm provides better accuracy at the expense of added computation time, and the set of parameter values where it outperforms the classical schemes is limited.

The analysis carried out in this chapter is not restricted to the 4-factor Heston–2CIR model, but can be extended to higher-dimensional problems. For instance, multi-factor short rates may be considered as in [23], with CIR dynamics and a term structure, in which case the convergence and variance reduction analysis applies with some slight modifications of the proofs.

Stochastic volatility accounts for volatility clustering, dependence in the increments and long-term smiles and skews, but gives rise to unrealistic short-term patterns in the implied volatility. Therefore, in order to improve the behaviour of the implied volatility for short maturities, the Heston–2CIR model could be extended in a number of ways. First, one could consider a stochastic-local volatility model as in (3.1) under a simpler correlation structure (such that the quanto correction term vanishes), which can easily be implemented for barrier option pricing with no extra computational effort. Second, a stochastic correlation (e.g., driven by a modified Jacobi process) between the spot FX rate and its squared volatility could be added [87; 95], in which case the conditional prices of European options (conditional also on the trajectory of the Brownian motion driving the stochastic correlation) are available in closed-form. Third, an independent jump component could be added to the spot FX rate, in which case analytical formulae may be available for the conditional prices of European options. For instance, this is the case when the distribution of the jump size is double-exponential [65] or normal [75].

However, some open problems remain, like the strong convergence order of the discretization scheme. In addition, examining the hedging parameters is also relevant.

Chapter 5

Conclusions

The efficient pricing and hedging of vanilla and exotic FX options requires an adequate model that takes into account both the local and the stochastic features of the volatility dynamics. We have put forward a 4-factor hybrid stochastic and local volatility model that combines state-of-the-art dynamics for the exchange rate with stochastic dynamics for the domestic and foreign short rates. This model allows for: a perfect calibration to the implied volatility surface; a more accurate pricing of exotic derivatives and a better risk-management performance; an exact fit to the initial term structure of interest rates and a more accurate pricing of long-dated interest rate-sensitive derivatives. We have proposed a simple and efficient Monte Carlo simulation scheme for pricing FX options under the 4-factor model.

On the one hand, a number of stochastic volatility models, including the popular Heston model, have the undesirable feature of moment instability, i.e., moments of order higher than 1 can explode in finite time. This can cause problems in practice, for instance when valuing derivatives with superlinear payoffs. On the other hand, proving the existence of moments of order higher than 1 of the process and its approximation is an important ingredient in the calibration routine and in the strong convergence analysis. We have shown, up to a critical time, the exponential integrability of a drift-implicit and several explicit Euler schemes for the square-root process driving the stochastic volatility and the stochastic short rates. This, in turn, plays a critical role in proving the boundedness of moments of numerical approximations for a large class of SDE systems with square-root dynamics in one or more dimensions, like the ones studied here. In particular, we have proved the boundedness of moments of order higher than 1 of approximation schemes for Heston-type models, which has been a long-standing open problem until now.

Weak convergence is important when estimating expectations of payoffs, while strong convergence plays a crucial role in multilevel Monte Carlo methods and may be required for complex path-dependent derivatives. The convergence of Euler schemes in a stochastic-local volatility context has yet to be established, and even the literature on the convergence of Euler schemes under stochastic volatility is scarce. Furthermore, for a large class of SDEs with superlinearly growing coefficients, the explicit Euler method is known to diverge in the strong and in the numerically weak sense in finite time. We have made some realistic assumptions on the leverage function and established, up to a critical time, the strong convergence of numerical approximations under the 4-factor model and the convergence of Monte Carlo estimators for a number of vanilla and exotic products. In practice, option pricing models may need to be recalibrated several times within a short time span. Therefore, a robust and efficient calibration of the 4-factor model is of great importance. We have implemented a fast calibration routine that combines semi-analytical formulae for vanilla option prices under stochastic volatility and short rates with a control variate particle method for the leverage function, and discussed the fitness of the model for pricing long-dated exotic FX products by looking at an autocallable barrier dual currency note.

Due to the slow convergence of the Monte Carlo method in the number of samples, we have analyzed a mixed Monte Carlo/PDE variance reduction method obtained by mixing Monte Carlo and finite difference methods via conditioning. Convergence properties of this algorithm have briefly been examined in the literature using heuristic arguments, without taking into account the error arising from the discretization of the underlying processes. We have proposed a mixed simulation scheme for a purely stochastic version of the model and rigorously proved the strong convergence of the scheme and the convergence of mixed Monte Carlo/PDE estimators for computing option prices, which we have demonstrated numerically for two different contracts. We have performed a variance reduction analysis, using the standard Monte Carlo method as a benchmark, and found accurate approximations for quantities of interest, including the variance reduction factor. Through a detailed numerical analysis carried out for two derivatives, namely a European call option and an up-and-out put option, we have identified under what circumstances the mixed Monte Carlo/PDE method outperforms the classical Monte Carlo and finite difference methods. Furthermore, we have also demonstrated, theoretically and numerically, that the best relative performance of the mixed method is achieved when the exchange rate dynamics are independent of the squared volatility and the domestic and foreign short rate dynamics.

This study has answered a number of questions concerning the rigorous analysis of state-of-the-art numerical methods for pricing FX options, but inevitably also raised some questions, such as: whether we can find sharp conditions on the exponential integrability of Euler approximations for the square-root process; whether we can relax some of the assumptions on the leverage function without losing the convergence properties; or the exact strong convergence order of approximation schemes for the type of SDEs studied in this thesis. Furthermore, it would be interesting and relevant to extend the theoretical and numerical analysis of the mixed Monte Carlo/PDE method to other models of interest, for instance obtained from the original model by adding either a stochastic-local volatility, a jump component to the spot FX rate, or a stochastic correlation between the spot FX rate and its squared volatility.

Bibliography

[1] R. Ahlip and M. Rutkowski. Pricing of foreign exchange options under the Heston stochastic volatility model and CIR interest rates. *Quantitative Finance*, 13(6):955–966, 2013.

[2] Y. Aït-Sahalia and R. Kimmel. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, 83:413–452, 2007.

[3] A. Alfonsi. On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods and Applications*, 11(4):355–384, 2005.

[4] A. Alfonsi. Strong order one convergence of a drift implicit Euler scheme: application to the CIR process. *Statistics and Probability Letters*, 83(2):602–607, 2013.

[5] H. H. N. Amin. Calibration of different interest rate models for a good fit of yield curves. Master's thesis, Delft University of Technology, 2012.

[6] L. Andersen. Simple and efficient simulation of the Heston stochastic volatility model. *The Journal of Computational Finance*, 11(3):1–42, 2008.

[7] L. Andersen and V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11(1):29–50, 2007.

[8] X. X. Ang. A mixed PDE/Monte Carlo approach as an efficient way to price under high-dimensional systems. Master's thesis, University of Oxford, 2013.

[9] A. Berkaoui, M. Bossy, and A. Diop. Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM: Probability and Statistics*, 12:1–11, 2008.

[10] F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.

[11] M. Bossy and A. Diop. An efficient discretization scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^\alpha$, $\alpha \in [1/2, 1)$. INRIA Research Report 5396, 2007.

[12] D. Brigo and F. Mercurio. A deterministic-shift extension of analytically-tractable and time-homogeneous short-rate models. *Finance and Stochastics*, 5(3):369–387, 2001.

[13] D. Brigo and F. Mercurio. *Interest Rate Models: Theory and Practice*. Springer, 2006.

[14] M. Broadie and O. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 54(2):217–231, 2006.

[15] P. Cheridito, D. Filipović, and R. L. Kimmel. Market price of risk specifications for affine models: theory and evidence. *Journal of Financial Economics*, 83(1):123–170, 2007.

[16] I. J. Clark. *Foreign Exchange Option Pricing: A Practitioner's Guide*. Wiley, 2011.

[17] T. F. Coleman and Y. Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6(2):418–445, 1996.

[18] J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[19] A. Cozma, M. Mariapragassam, and C. Reisinger. Convergence of an Euler scheme for a hybrid stochastic-local volatility model with stochastic rates in foreign exchange markets. Preprint, arXiv:1501.06084, 2016.

[20] A. Cozma, M. Mariapragassam, and C. Reisinger. Calibration of a four-factor hybrid local-stochastic volatility model with a new control variate particle method. Preprint, arXiv:1701.06001, 2017.

[21] A. Cozma and C. Reisinger. Exponential integrability properties of Euler discretization schemes for the Cox–Ingersoll–Ross process. *Discrete and Continuous Dynamical Systems – Series B*, 21(10):3359–3377, 2016.

[22] A. Cozma and C. Reisinger. A mixed Monte Carlo and partial differential equation variance reduction method for foreign exchange options under the Heston–Cox–Ingersoll–Ross model. *The Journal of Computational Finance*, 20(3):109–149, 2017. DOI: 10.21314/JCF.2016.318.

[23] D.-M. Dang, K. R. Jackson, and M. Mohammadi. Dimension and variance reduction for Monte Carlo methods for high-dimensional models in finance. *Applied Mathematical Finance*, 22(6):522–552, 2015.

[24] G. Deelstra and F. Delbaen. Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Applied Stochastic Models and Data Analysis*, 14(1):77–84, 1998.

[25] G. Deelstra and G. Rayee. Local volatility pricing models for long-dated FX derivatives. *Applied Mathematical Finance*, 20(4):380–402, 2013.

[26] S. Dereich, A. Neuenkirch, and L. Szpruch. An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. *Proceedings of the Royal Society of London A*, 468(2140):1105–1115, 2012.

[27] E. Derman and I. Kani. Riding on a smile. *Risk*, 7:32–39, 1994.

[28] S. S. Dragomir. *Some Gronwall Type Inequalities and Applications*. Nova Science Publishers, 2003.

[29] J. Driffill, T. Kenc, and M. Sola. An Empirical Examination of Term Structure Models with Regime Shifts. Computing in Economics and Finance 65, Society for Computational Economics, 2003.

[30] D. Dufresne. The integrated square-root process. Research Paper 90, Centre for Actuarial Studies, University of Melbourne, 2001.

[31] B. Dupire. Pricing with a smile. *Risk*, 7:18–20, 1994.

[32] A. Elices and E. Giménez. Applying hedging strategies to estimate model risk and provision calculation. *Quantitative Finance*, 13(7):1015–1028, 2013.

[33] M. Erismann. Analytical propositions to evaluate contingent convertible capital. Master's thesis, University of St. Gallen, 2011.

[34] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[35] L. Fatone, M. C. Recchioni, and F. Zirilli. A perturbative formula to price barrier options with time-dependent parameters in the Black and Scholes world. *The Journal of Risk*, 10(2):131–146, 2008.

[36] M. Fischer and G. Nappo. On the moments of the modulus of continuity of Itô processes. *Stochastic Analysis and Applications*, 28(1):103–122, 2009.

[37] F. L. Floc'h. TR-BDF2 for fast stable American option pricing. *The Journal of Computational Finance*, 17(3):31–56, 2014.

[38] M. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.

[39] M. B. Giles, D. J. Higham, and X. Mao. Analysing multi-level Monte Carlo for options with non-globally Lipschitz payoff. *Finance and Stochastics*, 13(3):403–413, 2009.

[40] M. B. Giles and C. Reisinger. Stochastic finite differences and multilevel Monte Carlo for a class of SPDEs in finance. *SIAM Journal on Financial Mathematics*, 3(1):572–592, 2012.

[41] P. Glasserman. *Monte Carlo Methods in Financial Engineering*, volume 53 of *Stochastic Modelling and Applied Probability*. Springer, 2003.

[42] L. A. Grzelak and C. W. Oosterlee. On the Heston model with stochastic interest rates. *SIAM Journal on Financial Mathematics*, 2:255–286, 2011.

[43] L. A. Grzelak and C. W. Oosterlee. On cross-currency models with stochastic volatility and correlated interest rates. *Applied Mathematical Finance*, 19(1):1–35, 2012.

[44] J. Guyon and P. Henry-Labordére. The smile calibration problem solved. SSRN Working Paper (1885032), 2011.

[45] B. Hambly, M. Mariapragassam, and C. Reisinger. A forward equation for barrier options under the Brunick&Shreve Markovian projection. *Quantitative Finance*, 16(6):827–838, 2016.

[46] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

[47] D. J. Higham and X. Mao. Convergence of Monte Carlo simulations involving the mean-reverting square root process. *The Journal of Computational Finance*, 8(3):35–62, 2005.

[48] D. J. Higham, X. Mao, and A. M. Stuart. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM Journal on Numerical Analysis*, 40(3):1041–1063, 2002.

[49] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2):281–300, 1987.

[50] J. Hull and A. White. One-factor interest-rate models and the valuation of interest-rate derivative securities. *Journal of Financial and Quantitative Analysis*, 28(2):235–254, 1993.

[51] C. Hunter. Hybrid derivatives. In *The Euromoney Derivatives and Risk Management Handbook*. Euromoney Institutional Investor, London, 2005.

[52] T. R. Hurd and A. Kuznetsov. Explicit formulas for Laplace transforms of stochastic integrals. *Markov Processes and Related Fields*, 14:277–290, 2008.

[53] A. S. Hurn, K. A. Lindsay, and A. J. McClelland. Estimating the parameters of stochastic volatility models using option price data. *Journal of Business and Economic Statistics*, 33(4):579–594, 2014.

[54] M. Hutzenthaler and A. Jentzen. *Numerical Approximations of Stochastic Differential Equations with Non-Globally Lipschitz Continuous Coefficients*, volume 236 of *Memoirs of the American Mathematical Society*. American Mathematical Society, Providence, RI, 2015.

[55] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proceedings of the Royal Society A*, 467:1563–1576, 2011.

[56] M. Hutzenthaler, A. Jentzen, and M. Noll. Strong convergence rates and temporal regularity for Cox–Ingersoll–Ross processes and Bessel processes with accessible boundaries. Working paper, arXiv:1403.6385v1, 2014.

[57] M. Hutzenthaler, A. Jentzen, and X. Wang. Exponential integrability properties of numerical approximation processes for nonlinear stochastic differential equations. Working paper, arXiv:1309.7657v3, 2016.

[58] A. Jacquier and C. Martini. Heston 2010. SSRN Working Paper (doi.org/fztnkn), 2011.

[59] C. Jessen and R. Poulsen. Empirical performance of models for barrier option valuation. *Quantitative Finance*, 13(1):1–11, 2013.

[60] M. Jex, R. Henderson, and D. Wang. Pricing exotics under the smile. *Risk*, 12(11):72–75, 1999.

[61] C. Kahl and P. Jäckel. Fast strong approximation Monte-Carlo schemes for stochastic volatility models. *Quantitative Finance*, 6(6):513–536, 2006.

[62] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, second edition, 1991.

[63] P. Kloeden and A. Neuenkirch. Convergence of numerical methods for stochastic differential equations in mathematical finance. In T. Gerstner and P. Kloeden, editors, *Recent Developments in Computational Finance: Foundations, Algorithms and Applications*. World Scientific, 2012.

[64] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, third edition, 1999.

[65] S. G. Kou. A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101, 2002.

[66] L. Lafférs. Empirical likelihood estimation of interest rate diffusion model. Master's thesis, Comenius University, Bratislava, 2009.

[67] S. R. Liberty and L. Mou. Existence of solutions of a Riccati differential system from a general cumulant control problem. *International Journal of Differential Equations*, 2011:13 pages, 2011.

[68] T. Lipp, G. Loeper, and O. Pironneau. Mixing Monte-Carlo and partial differential equations for pricing options. *Chinese Annals of Mathematics B*, 34(2):255–276, 2013.

[69] A. Lipton. The vol smile problem. *Risk*, 15(2):61–65, 2002.

[70] A. Lipton, A. Gal, and A. Lasis. Pricing of vanilla and first-generation exotic options in the local stochastic volatility framework: survey and new results. *Quantitative Finance*, 14(11):1899–1922, 2014.

[71] A. Lipton and W. McGhee. Universal barriers. *Risk*, 15(5):81–85, 2002.

[72] G. Loeper and O. Pironneau. A mixed PDE/Monte-Carlo method for stochastic volatility models. *Comptes Rendus Mathematique*, 347(9):559–563, 2009.

[73] R. Lord, R. Koekkoek, and D. van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2):177–194, 2010.

[74] W. A. McGhee. Pricing path dependent contracts in the presence of stochastic volatility: combining numerical integration, finite difference and conditional Monte Carlo. SSRN Working Paper (doi.org/bhwg), 2014.

[75] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144, 1976.

[76] A. Neuenkirch and L. Szpruch. First order strong approximations of scalar SDEs defined in a domain. *Numerische Mathematik*, 128(1):103–136, 2014.

[77] S. Ninomiya and N. Victoir. Weak approximation of stochastic differential equations and application to derivative pricing. *Applied Mathematical Finance*, 15(2):107–121, 2008.

[78] D. B. Owen. A table of normal integrals. *Communications in Statistics: Simulation and Computation B*, 9(4):389–419, 1980.

[79] V. Piterbarg. Smiling hybrids. *Risk*, 19(5):66–71, 2006.

[80] Y. Ren, D. Madan, and M. Q. Qian. Calibrating and pricing with embedded local volatility models. *Risk*, pages 138–143, September 2007.

[81] M. Rubinstein. Implied binomial trees. *Journal of Finance*, 49(3):771–818, 1994.

[82] R. Schöbel and J. Zhu. Stochastic volatility with an Ornstein–Uhlenbeck process: an extension. *European Finance Review*, 3(1):23–46, 1999.

[83] R. E. Shafer. Elementary problems: E 1867. *The American Mathematical Monthly*, 73(3):309, 1966.

[84] S. E. Shreve. *Stochastic Calculus for Finance 2: Continuous-Time Models*. Springer Finance, 2004.

[85] R. Storn and K. Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341–359, 1997.

[86] D. Tavella and C. Randall. *Pricing Financial Instruments: The Finite Difference Method*. Wiley, 2000.

[87] L. Teng, M. Ehrhardt, and M. Günther. On the Heston model with stochastic correlation. *International Journal of Theoretical and Applied Finance*, 19(6):1–25, 2016.

[88] The World Federation of Exchanges. 2013 WFE annual report. Retrieved from http://www.world-exchanges.org/insight/reports/2013-wfe-annual-report, December 2013.

[89] Y. Tian, Z. Zhu, G. Lee, F. Klebaner, and K. Hamza. Calibrating and pricing with a stochastic-local volatility model. *Journal of Derivatives*, 22(3):21–39, 2015.

[90] A. van der Stoep, L. A. Grzelak, and C. W. Oosterlee. The Heston stochastic-local volatility model: efficient Monte Carlo simulation. *International Journal of Theoretical and Applied Finance*, 17(7):1–30, 2014.

[91] A. van der Stoep, L. A. Grzelak, and C. W. Oosterlee. A novel Monte Carlo approach to hybrid local volatility models. SSRN Working Paper (2766990), 2016.

[92] A. Van Haastrecht. *Pricing long-term options with stochastic volatility and stochastic interest rates*. PhD thesis, University of Amsterdam, 2010.

[93] A. Van Haastrecht, R. Lord, A. Pelsser, and D. Schrager. Pricing long-maturity equity and FX derivatives with stochastic interest rates and stochastic volatility. *Insurance: Mathematics and Economics*, 45(3):436–448, 2009.

[94] A. Van Haastrecht and A. Pelsser. Generic pricing of FX, inflation and stock options under stochastic interest rates and stochastic volatility. *Quantitative Finance*, 11(5):665–691, 2011.

[95] A. Veraart and L. Veraart. Stochastic volatility and stochastic leverage. *Annals of Finance*, 8(2):205–233, 2012.

[96] G. A. Willard. Calculating prices and sensitivities for path-independent derivative securities in multifactor models. *Journal of Derivatives*, 5(1):45–61, 1997.

[97] U. Wystup. *FX Options and Structured Products*. Wiley, 2007.