Efficient Simulation of the Double Heston Model

Article *in* SSRN Electronic Journal · January 2010

DOI: 10.2139/ssrn.1434853

Dylan Possamaï
ETH Zurich

Efficient Simulation of the Double Heston model

Pierre Gauthier [*], Dylan Possamaï [†] [‡]

First version: July 16, 2009
This version: January 28, 2010

Abstract. Stochastic volatility models have replaced Black-Scholes model since they are able to generate a volatility smile. However, standard models fail to capture the smile slope and level movements. The Double-Heston model provides a more flexible approach to model the stochastic variance. In this paper, we focus on numerical implementation of this model. First, following the works of Lord and Kahl, we correct the analytical call option price formula given by Christoffersen et al. Then, we compare numerically the discretization schemes of Andersen, Zhu and Alfonsi to the Euler scheme.

Keywords: double Heston model, Stochastic Volatility, Equity options, Characteristic function, Discretization scheme

JEL Classification: G13
MSC 2000: 60-99, 65C05

[*] pierre.gauthier@pricingpartners.com - Pricing Partners, Paris Cybervillage, 204 Rue de Crimee, 75019 Paris, France - +33155264305

[†] dylan.possamai@pricingpartners.com - Pricing Partners, Paris Cybervillage, 204 Rue de Crimee, 75019 Paris, France

[‡] dylan.possamai@polytechnique.edu - Ecole Polytechnique, Route de Saclay, 91128 Palaiseau, France - +33633873888

1. Introduction

Since Black and Scholes introduced their option valuation model, an extensive literature has been devoted to pointing out its limitations. More precisely, two major features of the equity index options market cannot be captured through Black-Scholes model. First of all, there exists a term structure of implied volatilities. As a matter of fact, the constant volatility parameter is not sufficient to model this behavior. Second, observed market prices for both in-the-money and out-the-money options are higher than Black-Scholes prices. This effect is known as the volatility smile: the volatility depends both on the option expiry and the option strike.

In order to model the smile efficiently, stochastic volatility models are a popular approach. They enable to have distinct processes for the stock return and its variance. Thus, they may generate volatility smiles. Moreover, if the variance process embeds a mean reversion term, these models can capture the term-structure in the variance dynamics. Popular stochastic volatility models include Heston [17], SABR and square-root models to name but a few.

Efficient calibration of stochastic volatility models requires an analytical formula for option prices. For numerous models, including Heston, this is achieved through the Fourier-transform technique described in Carr and Madan [7]. Further refinements have been discussed in Lewis [21] and Lord and Kahl [22].

Numerical simulation may be challenging as convergence speed can be quite low for naive Euler scheme. Therefore, there is a need for more accurate discretization of both stock return and variance processes. Many authors have addressed this issue, including . In [33], van Haastrecht and Pelsser propose an extensive discussion of existing discretization schemes for Heston model.

In order to match precisely the market implied volatility surface, it turns out that Heston model does not have enough parameters. Therefore, we may wish to add degrees of freedom (i.e additional parameters) while retaining analytical tractability. Many academicians and practitioners have tackled this challenge by considering time-dependent extensions of the original Heston model.

Another direction is to model the variance by a variable of higher dimension. Such a path has been followed by Gouriéroux et al [16] and Da Fonseca et al. [13], who replace the Cox-Ingersoll-Ross variance process by a Wishart process. Their model enables to have a tighter control of the covariance dynamics as it is represented by a matrix of size n (2 is enough in practice). In [8], Christoffersen et al. consider a simpler extension, named the Double Heston model, in which they model the variance process by two uncorrelated processes. Thanks to this additional factor, fit quality is much better than with the original Heston model while computing time remains comparable. Empirical studies show that adding a third process does not lead to significant improvements.

In this article, we consider the latter stochastic volatility model. Our objective is two-fold. First, we improve the call option formula given in Christoffersen and we derive a formula for digital options. Second, we study convergence speed of Monte-Carlo simulation using different discretization schemes.

The paper is structured as follows. First, in Section 2, we recall the model dynamics. In

Section 3, we give analytical formulas for vanilla and binary options. Some refinements make these formulas slightly better than those given in the original paper by Christoffersen. Then, in Section 5, we adapt the QE-M scheme introduced by Andersen and the schemes introduced by Alfonsi for the Heston model to the Double Heston model. Section 6 briefly reviews control variate technique. Our numerical results are presented in section 7. Section 8 introduces new research directions and Conclusion concludes.

2. Model presentation

The model studied here is an extension of the Heston Model [17], described notably in [8]. The variance of the risk neutral, ex-dividend price process is determined by two factors

$$\frac{dS_t}{S_t} = rdt + \sqrt{V_t^1}dW_t^1 + \sqrt{V_t^2}dW_t^2 \tag{1}$$

$$dV_t^1 = b_1(\theta_1 - V_t^1)dt + \sigma_1\sqrt{V_t^1}dZ_t^1 \tag{2}$$

$$dV_t^2 = b_2(\theta_2 - V_t^2)dt + \sigma_2\sqrt{V_t^2}dZ_t^2. \tag{3}$$

For the sake of simplicity, we assume the following stochastic structure:

$$\begin{aligned}
\langle dW^1, dZ^1\rangle_t &= \rho_1 dt \\
\langle dW^2, dZ^2\rangle_t &= \rho_2 dt \\
\langle dW^1, dW^2\rangle_t &= 0 \\
\langle dZ^1, dZ^2\rangle_t &= 0.
\end{aligned}$$

In other words, the variance is the sum of two uncorrelated factors that may be individually correlated with stock returns.

3. Option valuation in the Double Heston Model: the characteristic function method

Following previous works by Carr, Madan, Lewis and Lee in particular (see [7], [34], [21] or even [31]), it is now well known that the price of a European option can be easily calculated via Fourier integrals, provided that the characteristic function of the log-spot price has a closed form.

3.1. The Characteristic Function

Just as the standard Heston model, the Double Heston model belongs to the larger class of affine models, for which the computation of the characteristic function is rather straightforward. Let us first note $X_t = \log S_t$ and:

$$\Psi(k,t,x) = \mathbb{E}\left(e^{ikX_t}|X_0 = x, V_0^1 = v_0^1, V_0^2 = v_0^2\right). \tag{4}$$

The dynamic of X_t is given by:

$$dX_t = \left(r - \frac{1}{2}\left(V_t^1 + V_t^2\right)\right)dt + \sqrt{V_t^1}dW_t^1 + \sqrt{V_t^2}dW_t^2. \tag{5}$$

According to the Feynman-Kac theorem, we know that the function Ψ is the solution of the following PDE:

$$\begin{cases} \dfrac{\partial f}{\partial t} & = \mathcal{L}_{X,V^1,V^2} f \\ \\ f(k,0,x) & = e^{ikx}, \end{cases}$$

where $\mathcal{L}_{X,V^1,V^2}$ is the infinitesimal generator of the diffusion of the process (X, V^1, V^2).

Thanks to Itō calculus, it is quite easy to find out that for every $\mathbb{R}^+ \times \mathbb{R}^3$ valued function $f(x, v_1, v_2)$ two times differentiable, we have:

$$\begin{aligned} \mathcal{L}_{X,V^1,V^2} f = & \frac{1}{2}(v_1+v_2)\frac{\partial^2 f}{\partial x^2} + \frac{1}{2}\sigma_1^2 v_1 \frac{\partial^2 f}{\partial v_1^2} + \frac{1}{2}\sigma_2^2 v_2 \frac{\partial^2 f}{\partial v_2^2} + \rho_1\sigma_1 v_1 \frac{\partial^2 f}{\partial x \partial v_1} + \rho_2\sigma_2 v_2 \frac{\partial^2 f}{\partial x \partial v_2} \\ & - \frac{1}{2}(v_1+v_2)\frac{\partial f}{\partial x} + b_1(\theta_1 - v_1)\frac{\partial f}{\partial v_1} + b_2(\theta_2 - v_2)\frac{\partial f}{\partial v_2}. \end{aligned} \tag{6}$$

Given that the model here is affine, we know (see Duffie [10]) that the function Ψ has the following form

$$\Psi(k,t,x) = e^{ikx + A(k,t) + B_1(k,t)v_0^1 + B_2(k,t)v_0^2}, \tag{7}$$

with $A(k,0) = B_1(k,0) = B_2(k,0) = 0$

Then replacing (7) in (6) and identifying the terms containing v_1 and v_2, we obtain the following system of ODEs:

$$\begin{cases} \dfrac{dB_1}{dt} &= \frac{1}{2}\sigma_1^2 B_1^2 - (b_1 - ik\rho_1\sigma_1)B_1 - \frac{k}{2}(k+i) \\ \\ \dfrac{dB_2}{dt} &= \frac{1}{2}\sigma_2^2 B_2^2 - (b_2 - ik\rho_2\sigma_2)B_2 - \frac{k}{2}(k+i) \\ \\ \dfrac{dA}{dt} &= b_1\theta_1 B_1 + b_2\theta_2 B_2. \end{cases}$$

The first two are one-dimensional Ricatti equations, and the third one only requires a straightforward integration. The solutions are given by:

$$\begin{cases} A(k,t) &= \dfrac{b_1\theta_1}{\sigma_1^2}\left((b_1 - \rho_1\sigma_1 ki - d_1)\,t - 2\log\left(\dfrac{1-G_1e^{-d_1t}}{1-G_1}\right)\right) \\ &+ \dfrac{b_2\theta_2}{\sigma_2^2}\left((b_2 - \rho_2\sigma_2 ki - d_2)\,t - 2\log\left(\dfrac{1-G_2e^{-d_2t}}{1-G_2}\right)\right) \\ \\ B_j(k,t) &= \dfrac{b_j - \rho_j\sigma_j ki - d_j}{\sigma_j^2}\dfrac{1-e^{-d_jt}}{1-G_je^{-d_jt}} \text{ for } j = 1,2, \end{cases}$$

where

$$\begin{cases} d_j &= \sqrt{(b_j - \rho_j\sigma_j ki)^2 + \sigma_j^2 k(k+i)} \\ \\ G_j &= \dfrac{b_j - \rho_j\sigma_j ki - d_j}{b_j - \rho_j\sigma_j ki + d_j}. \end{cases}$$

Note that in the definition of $A(k,t)$, some continuity problems could arise, due to the presence of complex logarithms. Nonetheless, in an article by Lord and Kahl [22], it has been proven that with the formulation chosen here (which is different from the original formulation of Heston), the characteristic function remained continuous when one chose the principal branch of the complex logarithm. Therefore, there was no need for us to use the so called "rotation count algorithm" proposed by Kahl and Jäckel in [19].

3.2. The Call Option Price

We know, following Lewis and Lee ([34] and [21]), that the price of a call option of strike K and maturity T is given by:

$$C(K,T) = S_0 - \frac{1}{\pi}\sqrt{S_0K}e^{-\frac{rT}{2}}\int_0^\infty \mathcal{R}e\left[e^{iuk}\Psi\left(u - \frac{i}{2}\right)\right]\frac{du}{u^2+\frac{1}{4}}, \tag{8}$$

where $k = rT - \log K$

Given that we have obtained a closed-form formula for the characteristic function involved in the integration, the only thing remaining in order to numerically evaluate the option price is to compute the integral. Instead of using FFT algorithms proposed by authors like Lewis and Lee, we chose to use the Gauss-Legendre quadrature method developed notably by Abbot in [1], which proves to be faster.
Our formula is different from that proposed by Christoffersen et al. [8]. Readers which are familiar with the characteristic function method are already aware of this recurrent issue. The original call formula requires two numerical integrations while ours, which is exactly that of Lewis [21], involves only one. Thus, we gain not only on computational time but also on numerical precision.

3.3. The Digital Call Option Price

In order to illustrate even more how easily the framework of the Double Heston model allows us to compute efficiently European vanilla options, we derive here the price of a digital Call Option. Given that the formula does not seem to be often quoted in the literature, we will also give a rough sketch of its proof.
We start from the fundamental pricing formula which can be found in [21] or [34], that is to say that the price of a European option with payoff $w(x)$ (as a function of the process of interest, in our case the log-price) and maturity T is:

$$f(S,T) = \frac{e^{-rt}}{2\pi} \int_{i\nu-\infty}^{i\nu+\infty} e^{-izY} \Psi(-z) \widehat{w}(z) dz, \tag{9}$$

where $Y = rT$, $v = Im(z)$, $\widehat{w}$ is the generalized Fourier Transform of the payoff and Ψ the characteristic function of the Lévy process of interest in the model (in our case the log-price), and of course provided that the quantities above all exist.

Then, the generalized Fourier Transform of the payoff of a Digital Call option with strike K is given formally by:

$$\begin{aligned} \widehat{w}(z) &= \int_{-\infty}^{+\infty} e^{izx} \mathbb{1}_{e^x \geq K} dx \\ &= \int_{\log(K)}^{+\infty} e^{izx} dx \\ &= -\frac{K^{iz}}{iz}. \end{aligned}$$

Of course for the formula above to be correct, the imaginary part of the complex z has to remain in the domain $S_w = \{z \in \mathbb{C}; Im(z) > 0\}$.

Then replacing (10) in (9) and setting $\nu = \frac{1}{2}$ (to stay close to the Call formula of the previous section, but any value strictly positive would do) we get:

$$DC(T,K) = -\frac{e^{-rT}}{2\pi} \int_{\frac{i}{2}-\infty}^{\frac{i}{2}+\infty} e^{-izk} \frac{\Psi(-z)}{iz} dz, \tag{10}$$

where $k = rT - \log K$

Given that our integrand has only a simple pole in 0, we can use the residue theorem to change the contour of integration in order to get:

$$DC(T,K) = \frac{e^{-\frac{rT}{2}}}{2\pi} \frac{1}{\sqrt{K}} \int_{-\infty}^{+\infty} e^{-ikx} \Psi\left(-x - \frac{i}{2}\right) \frac{i\left(x - \frac{i}{2}\right)}{x^2 + \frac{1}{4}} dx. \tag{11}$$

We can then finally take the real part of the expression in order to prevent the numerical approximations when computing the integral from rendering the quantity above complex, to finally get:

$$DC(T,K) = \frac{e^{-\frac{rT}{2}}}{2\pi} \frac{1}{\sqrt{K}} \int_{-\infty}^{+\infty} Re\left(e^{-ikx} \Psi\left(-x - \frac{i}{2}\right) \frac{i\left(x - \frac{i}{2}\right)}{x^2 + \frac{1}{4}}\right) dx. \tag{12}$$

4. Improving the closed formula: the control variate method

4.1. The Call Option Case

When we consider options which are deep-in-the-money, the vega is usually very small. Therefore it is necessary to have a very efficient pricing formula, if one wants to be able to calibrate the model to the market. For those extreme strikes, the current pricing formula does not allow to calibrate to the market implied volatility, since it calculates the option price in the money with a too large error. Therefore, we tried to improve it by using a control variate.
The idea here is to use the option price when the two volatility factors are deterministic (i.e $\sigma_1 = \sigma_2 = 0$). In this case we're in a simple Black-Scholes framework, and the price of the option can be computed as:

$$\widetilde{C}(K,T) = S_0 \mathcal{N}(\delta_1) - Ke^{-rT} \mathcal{N}(\delta_0), \tag{13}$$

where

$$
\begin{aligned}
\delta_0 &= \frac{1}{\sigma_T\sqrt{T}}\log\left(\frac{S_0}{Ke^{-rT}}\right) - \frac{1}{2}\sigma_T\sqrt{T}\\
\delta_1 &= \delta_0 + \sigma_T\sqrt{T} \qquad (14)\\
\sigma_T &= \sqrt{\frac{1}{T}\int_0^T \left(V_t^1 + V_t^2\right) dt}\\
&= \sqrt{\theta_1 + \theta_2 + \frac{V_0^1 - \theta_1}{b_1 T}\left(1 - e^{-b_1 T}\right) + \frac{V_0^2 - \theta_2}{b_2 T}\left(1 - e^{-b_2 T}\right)}.
\end{aligned}
$$

Besides, we can apply the characteristic option formula. Putting $\sigma_1 = \sigma_2 = 0$ in (6) we obtain the following PDE for the characteristic function $\widetilde{\Psi}(k,t,x)$ in this case:

$$
\frac{1}{2}(v_1+v_2)\frac{\partial^2 f}{\partial x^2} - \frac{1}{2}(v_1+v_2)\frac{\partial f}{\partial x} + b_1(\theta_1 - v_1)\frac{\partial f}{\partial v_1} + b_2(\theta_2 - v_2)\frac{\partial f}{\partial v_2} - \frac{\partial f}{\partial t} = 0. \qquad (15)
$$

Then we assume again that

$$
\widetilde{\Psi}(k,t,x) = e^{ikx + \widetilde{A}(k,t) + \widetilde{B_1}(k,t)v_0^1 + \widetilde{B_2}(k,t)v_0^2}. \qquad (16)
$$

Then putting (16) in (15) we obtain the three following ODEs:

$$
\begin{cases}
\dfrac{d\widetilde{B}_1}{dt} &= -(b_1 - ik\rho_1\sigma_1)\widetilde{B}_1 - \frac{k}{2}(k+i)\\[2ex]
\dfrac{d\widetilde{B}_2}{dt} &= -(b_2 - ik\rho_2\sigma_2)\widetilde{B}_2 - \frac{k}{2}(k+i)\\[2ex]
\dfrac{d\widetilde{A}}{dt} &= b_1\theta_1\widetilde{B}_1 + b_2\theta_2\widetilde{B}_2.
\end{cases}
$$

Their resolution is trivial and leads to:

$$
\begin{cases}
\widetilde{A}(k,t) &= \dfrac{k(i+k)}{2}\left(\dfrac{\theta_1}{b_1}\left(1 - e^{-b_1 t}\right) + \dfrac{\theta_2}{b_2}\left(1 - e^{-b_2 t}\right) - (\theta_1 + \theta_2)t\right)\\[2ex]
\widetilde{B}_j(k,t) &= -\dfrac{k(i+k)}{2b_j}\left(1 - e^{-b_j t}\right) \text{ for } j = 1, 2.
\end{cases} \qquad (17)
$$

Therefore, the price of the option in this deterministic case can also be given by:

$$\widetilde{C}(K,T) = S_0 - \frac{1}{\pi}\sqrt{S_0 K}e^{-\frac{rT}{2}} \int_0^{\infty} \mathcal{R}e\left[e^{iuk}\widetilde{\Psi}\left(u - \frac{i}{2}\right)\right] \frac{du}{u^2 + \frac{1}{4}}. \tag{18}$$

Thus we have a new formulation for the call option price in the Double Heston Model, using control variate:

$$C(K,T) = S_0\mathcal{N}(\delta_1) - Ke^{-rT}\mathcal{N}(\delta_0) - \frac{1}{\pi}\sqrt{S_0 K}e^{-\frac{rT}{2}} \int_0^{\infty} \mathcal{R}e\left[e^{iuk}\left(\Psi - \widetilde{\Psi}\right)\right] \frac{du}{u + \frac{i}{2}}. \tag{19}$$

The presence of $\widetilde{\Psi}$ in the integral reduces the integrand, preventing the numerical evaluation with the Gauss-Legendre quadrature method to create large errors with extreme strikes.

4.2. The Digital Call Option Case

The ideas here are the same as the previous section, that is to say that we use the option price in the case of deterministic volatilities as a variable control.
In that case the digital Call Option price is also given by a Black-Scholes type formula

$$DC(T,K) = e^{-rT}\mathcal{N}(\delta_0), \tag{20}$$

and therefore we have yet another expression for the price of the digital Call option using control variate:

$$DC(T,K) = e^{-rT}\mathcal{N}(\delta_0) + \frac{e^{-\frac{rT}{2}}}{2\pi}\frac{1}{\sqrt{K}} \int_{-\infty}^{+\infty} \mathcal{R}e\left[e^{-ikx}\left(\Psi - \widetilde{\Psi}\right)\left(-x - \frac{i}{2}\right)\frac{i\left(x - \frac{i}{2}\right)}{x^2 + \frac{1}{4}}\right] dx,$$

where $\widetilde{\Psi}$ is still the characteristic function of the log-price in the case of deterministic variances.

5. Efficient discretization

5.1. The Price Process

5.1.1. The predictor-corrector scheme

Rather than diffusing the logarithm of the spot $\ln S_t$, it may be interesting to consider the logarithm of the discounted spot $\log\left(e^{-rt}S_t\right)$. By applying Itō's formula, we remark

that the rdt terms cancel out. For the discounted asset price, we choose to apply the Predictor-Corrector scheme. This discretization scheme is known to provide better results than the Euler scheme at a low additional effort.
The idea is to start from an exact expression of the logarithm of the discounted price (which can be found in [6] for example). By Itō's formula we have:

$$\begin{aligned} d\left(\log\left(e^{-rt}S_t\right)\right) &= -rdt + \frac{dS_t}{S_t} - \frac{1}{2S_t^2}\left\langle dS, dS\right\rangle_t \\ &= -\frac{V_t^1 + V_t^2}{2}dt + \sqrt{V_t^1}dW_t^1 + \sqrt{V_t^2}dW_t^2. \end{aligned}$$

Then a simple integration on an interval $[t, t+\Delta t]$ leads to:

$$\log\left(e^{-r(t+\Delta t)}S_{t+\Delta t}\right) = \log\left(e^{-rt}S_t\right) - \frac{1}{2}\int_t^{t+\Delta t}\left(V_s^1 + V_s^2\right)ds + \int_t^{t+\Delta t}\sqrt{V_s^1}dW_s^1 + \sqrt{V_s^2}dW_s^2.$$

Then we introduce the correlations between the spot and the variances processes by rewriting the Brownian motions:

$$\begin{cases} W_t^1 = \rho_1 Z_t^1 + \sqrt{1-\rho_1^2}\widetilde{W}_t^1 \\ W_t^2 = \rho_2 Z_t^2 + \sqrt{1-\rho_2^2}\widetilde{W}_t^2, \end{cases} \tag{21}$$

where $\widetilde{W}^1$ and $\widetilde{W}^2$ are Brownian motions independent of W^1 and W^2 respectively.

Therefore we finally have:

$$\begin{aligned} \log\left(e^{-r(t+\Delta t)}S_{t+\Delta t}\right) &= \log\left(e^{-rt}S_t\right) - \frac{1}{2}\int_t^{t+\Delta t} V_s^1 + V_s^2 ds \\ &+ \int_t^{t+\Delta t}\rho_1\sqrt{V_s^1}dZ_s^1 + \int_t^{t+\Delta t}\rho_2\sqrt{V_s^2}dZ_s^2 \\ &+ \int_t^{t+\Delta t}\sqrt{1-\rho_1^2}\sqrt{V_s^1}d\widetilde{W}_s^1 + \int_t^{t+\Delta t}\sqrt{1-\rho_2^2}\sqrt{V_s^2}d\widetilde{W}_s^2. \end{aligned} \tag{22}$$

Then we integrate the SDEs for the variance processes:

$$V_{t+\Delta t}^1 = V_t^1 + \int_t^{t+\Delta t} b_1 \left(\theta_1 - V_s^1\right) ds + \sigma_1 \int_t^{t+\Delta t} \sqrt{V_s^1} dZ_s^1$$

$$V_{t+\Delta t}^2 = V_t^2 + \int_t^{t+\Delta t} b_2 \left(\theta_2 - V_s^2\right) ds + \sigma_2 \int_t^{t+\Delta t} \sqrt{V_s^2} dZ_s^2.$$

Equivalently, and in order to eliminate some of the stochastic integrals in (22), we have:

$$\int_t^{t+\Delta t} \sqrt{V_s^1} dZ_s^1 = \frac{1}{\sigma_1}\left(V_{t+\Delta t}^1 - V_t^1 - b_1\theta_1\Delta t + b_1 \int_t^{t+\Delta t} V_s^1 ds\right)$$

$$\int_t^{t+\Delta t} \sqrt{V_s^2} dZ_s^2 = \frac{1}{\sigma_1}\left(V_{t+\Delta t}^2 - V_t^2 - b_2\theta_2\Delta t + b_2 \int_t^{t+\Delta t} V_s^2 ds\right).$$

In the end we get the following exact representation of the logarithm of the discounted price:

$$\begin{aligned}\log\left(\frac{e^{-r(t+\Delta t)}S_{t+\Delta t}}{e^{-rt}S_t}\right) &= \frac{\rho_1}{\sigma_1}\left(V_{t+\Delta t}^1 - V_t^1 - b_1\theta_1\Delta t\right) + \frac{\rho_2}{\sigma_2}\left(V_{t+\Delta t}^2 - V_t^2 - b_2\theta_2\Delta t\right) \\ &+ \left(\frac{b_1\rho_1}{\sigma_1} - \frac{1}{2}\right)\int_t^{t+\Delta t} V_s^1 ds + \left(\frac{b_2\rho_2}{\sigma_2} - \frac{1}{2}\right)\int_t^{t+\Delta t} V_s^2 ds \\ &+ \sqrt{1-\rho_1^2}\int_t^{t+\Delta t}\sqrt{V_s^1} d\widetilde{W}_s^1 ds + \sqrt{1-\rho_2^2}\int_t^{t+\Delta t}\sqrt{V_s^2} d\widetilde{W}_s^2 ds.\end{aligned}$$

We now have to approximate the integrals in the above formula. The idea of the predictor corrector scheme (see [3]) is first to handle the time integral in a centered manner, that is to say:

$$\int_t^{t+\Delta t} V_s^j ds \approx \frac{\Delta t}{2}\left(V_{t+\Delta t}^j + V_t^j\right) \text{ for } j = 1, 2. \tag{23}$$

Then, given that $\widetilde{W}^1$ is independent of V^1 and $\widetilde{W}^2$ is independent of V^2, the stochastic integrals with respect to $\widetilde{W}^1$ and $\widetilde{W}^2$ in the formula above are (conditionally) Gaussian with zero mean and variance $\int_t^{t+\Delta t} V_s^1 ds$ and $\int_t^{t+\Delta t} V_s^2 ds$, which can be approximated as previously.

Finally, the scheme for the log-price can be written as:

$$\log\left(e^{-r(t+\Delta t)}\widehat{S}_{t+\Delta t}\right) = \log(e^{-rt}\widehat{S}_t) + K_0^1 + K_0^2 + K_1^1\widehat{V}_t^1 + K_1^2\widehat{V}_t^2 + K_2^1\widehat{V}_{t+\Delta t}^1 + K_2^2\widehat{V}_{t+\Delta t}^2$$
$$+\sqrt{K_3^1\left(\widehat{V}_t^1 + \widehat{V}_{t+\Delta t}^1\right)}B_1 + \sqrt{K_3^2\left(\widehat{V}_t^2 + \widehat{V}_{t+\Delta t}^2\right)}B_2, \tag{24}$$

where

$$\begin{cases} K_0^j = -\dfrac{\rho_j b_j \theta_j}{\sigma_j}\Delta t \\ K_1^j = \dfrac{\Delta t}{2}\left(\dfrac{b_j\rho_j}{\sigma_j} - \dfrac{1}{2}\right) - \dfrac{\rho_j}{\sigma_j} \\ K_2^j = \dfrac{\Delta t}{2}\left(\dfrac{b_j\rho_j}{\sigma_j} - \dfrac{1}{2}\right) + \dfrac{\rho_j}{\sigma_j} \\ K_3^j = \dfrac{\Delta t}{2}\left(1-\rho_j^2\right), \end{cases}$$

for $j = 1, 2$ and where B_1 and B_2 are standard Gaussian random variables independent and independent of $\widehat{V}^1$ and $\widehat{V}^2$

5.1.2. Martingale Correction

As noted by Andersen in [3] and Andersen and Piterbarg in [4] the discretized scheme for the discounted price may not always be arbitrage free. That is to say that even though the continuous-time process of discounted price is a martingale, the scheme defined by (24) is not. Even though the practical relevance of this is often minor, it is interesting in our double Heston framework to see how the results of Andersen may be generalized. Therefore, let us examine whether it is possible to modify the scheme (24) to enforce that

$$\mathbb{E}\left[e^{-r\Delta t}\widehat{S}_{t+\Delta t}|\widehat{S}_t\right] = \widehat{S}_t. \tag{25}$$

By iterated conditional expectations and by taking the exponential in (24), we have:

$$
\begin{aligned}
\mathbb{E}\left[e^{-r\Delta t}\widehat{S}_{t+\Delta t}|\widehat{S}_t\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{-r\Delta t}\widehat{S}_{t+\Delta t}|\widehat{S}_t, \widehat{V}^1_{t+\Delta t}, \widehat{V}^2_{t+\Delta t}\right]\right] \\
&= \widehat{S}_t e^{K_0^1+K_0^2+K_1^1\widehat{V}^1_t+K_1^2\widehat{V}^2_t}\mathbb{E}\left[e^{K_2^1\widehat{V}^1_{t+\Delta t}+K_2^2\widehat{V}^2_{t+\Delta t}}\times\right. \\
&\quad \left.\mathbb{E}\left[e^{\sqrt{K_3^1(\widehat{V}^1_t+\widehat{V}^1_{t+\Delta t})}B_1+\sqrt{K_3^2(\widehat{V}^2_t+\widehat{V}^2_{t+\Delta t})}B_2}|\widehat{S}_t, \widehat{V}^1_{t+\Delta t}, \widehat{V}^2_{t+\Delta t}\right]\right] \\
&= \widehat{S}_t e^{K_0^1+K_0^2+(K_1^1+\frac{1}{2}K_3^1)\widehat{V}^1_t+(K_1^2+\frac{1}{2}K_3^2)\widehat{V}^2_t}\times \\
&\quad \mathbb{E}\left[e^{A_1\widehat{V}^1_{t+\Delta t}}|\widehat{V}^1_t\right]\mathbb{E}\left[e^{A_2\widehat{V}^2_{t+\Delta t}}|\widehat{V}^2_t\right],
\end{aligned}
$$

where $A_j = K_2^j + \frac{1}{2}K_3^j$ for $j = 1, 2$.

Let us then note $M_j = \mathbb{E}\left[e^{A_j\widehat{V}^j_{t+\Delta t}}|\widehat{V}^j_t\right]$.

If we were to modify the scheme (24) by replacing K_0^1 and K_0^2 with $\widetilde{K}_0^1$ and $\widetilde{K}_0^2$ so that we have:

$$
e^{\widetilde{K}_0^1+\widetilde{K}_0^2+(K_1^1+\frac{1}{2}K_3^1)\widehat{V}^1_t+(K_1^2+\frac{1}{2}K_3^2)\widehat{V}^2_t}M_1M_2 = 1, \tag{26}
$$

then the martingale condition would be clearly verified.

In order to do so we may for example take:

$$
\widetilde{K}_0^j = -\log M_j - \left(K_1^j + \frac{1}{2}K_3^j\right)\widehat{V_t^j} \text{ for } j = 1, 2. \tag{27}
$$

Of course, for everything we just said to be meaningful, the quantities M_1 and M_2 have to be finite. In most of the variance discretization schemes that we will look at in the following sections, the variance remains positive (only the "full truncation" Euler scheme does not verify this property), it is clear in those cases that if $A_j \leq 0$ then $M_j < +\infty$, since the quantity which expectation is taken is bounded by 0 and 1. Fortunately, it is easy to show that, as a function of ρ_j, A_j is increasing on the interval $[-1, 0]$ and is equal to $-\Delta t/4$ on 0. Therefore, if we assume negative correlations between the stock price and the variance processes (which is almost always verified in practice), the martingale correction can always be done. However, in the case of positive correlation, we will have to further study the discretization scheme chosen for the variance

5.2. The Variance Processes

5.2.1. Full Truncation Euler Scheme

For the variance processes, the first option is to use a slightly modified naive Euler scheme. The most naive discretization scheme here would be to write:

$$\widehat{V}_{t+\Delta t}^{j} = \widehat{V}_t^j + b_j\left(\theta_j - \widehat{V}_t^j\right)\Delta t + \sigma_j\sqrt{\widehat{V}_t^j}G_j\sqrt{\Delta t}, \tag{28}$$

for $j = 1, 2$ and where the G_j are standard Gaussian random variables independent and independent of B_1 and B_2 defined in the previous section.

The fatal flaw of these schemes is that under the condition $2b_j\theta_j < \sigma_j^2$ (which is almost always verified in practical situations), the process V_j can attain the origin with a strictly positive probability (see e.g. [9] and [11]). Therefore the discrete process defined above can become negative with non-zero probability, which would then cause the scheme to fail. In order to solve this problem, several fixes have been proposed in the literature (see [23] for a comprehensive review), and the one which seems to produce the smallest bias is known as the "full truncation" scheme. The idea behind is to allow the process to become negative, but to compensate it by making it deterministic at the same time. It can be written as:

$$\widehat{V}_{t+\Delta t}^{j} = \widehat{V}_t^j + b_j\left(\theta_j - \left(\widehat{V}_t^j\right)^+\right)\Delta t + \sigma_j\sqrt{\left(\widehat{V}_t^j\right)^+}G_j\sqrt{\Delta t}. \tag{29}$$

5.2.2. Zhu scheme

Derivation of the Scheme In [35], Zhu proposes to diffuse the volatility $v_t^j = \sqrt{V_t^j}$ instead of the variance. After application of Itō's lemma (although in this particular case it cannot exactly be applied, we will go back to this issue later), one finds that the volatility processes evolve according to:

$$dv_t^j = \frac{b_j}{2}\left[\left(\theta_j - \frac{\sigma_j^2}{4b_j}\right)\frac{1}{v_t^j} - v_t^j\right]dt + \frac{1}{2}\sigma_j dZ_t^j \tag{30}$$

$$= \kappa_j\left[\lambda_j - v_t^j\right]dt + \epsilon_j dZ_t^j, \tag{31}$$

with evident notations.

This would be a standard Ornstein-Uhlenbeck if the mean-reverting parameters λ_j were constant. Zhu's idea is to approximate the volatility by a true Ornstein-Uhlenbeck process that has the same first two moments. Let us note such a process as X_t^j, which dynamic is given by:

$$dX_t^j = \kappa_j \left[\widehat{\lambda}_j(t) - X_t^j\right] dt + \epsilon_j dZ_t^j$$
$$X_0^j = v_0^j,$$

where $\widehat{\lambda}_j(t)$ is piecewise constant on the discretization grid chosen. That is to say that if we note the discretization grid $\left\{t_k = \frac{kT}{N}; k = 0..N\right\}$, we have:

$$\widehat{\lambda}_j(t) = \sum_{k=0}^{N-1} \widehat{\lambda}_j^k \mathbb{1}_{t_k \le t < t_{k+1}}.$$

It is then very well known that the moments of the process X^j are, for $0 \le k \le N-1$, given by:s X^j are, for $0 \le k \le N-1$, given by:

$$\mathbb{E}\left[X_{t_{k+1}}^j | X_{t_k}\right] = \widehat{\lambda}_j^k + \left(X_{t_k}^j - \widehat{\lambda}_j^k\right) e^{-\kappa_j(t_{k+1}-t_k)} \tag{32}$$

$$\mathbb{V}ar\left[X_{t_{k+1}}^j | X_{t_k}\right] = \frac{\epsilon_j^2}{2\kappa_j}\left(1 - e^{-2\kappa_j(t_{k+1}-t_k)}\right). \tag{33}$$

In order to match those two moments we need to know the mean and the variance of our volatility process. In other words we need to be able to compute:

$$\mathbb{E}\left[v_{t_{k+1}}^j | v_{t_k}\right] = \mathbb{E}\left[\sqrt{V_{t_{k+1}}^j} | V_{t_k}\right]$$
$$\mathbb{V}ar\left[v_{t_{k+1}}^j | v_{t_k}\right] = \mathbb{V}ar\left[\sqrt{V_{t_{k+1}}^j} | v_{t_k}\right].$$

Before going on with the calculations, let us first recall some theoretical results about square-root processes (see [9] and [11] for the proofs), which will prove useful in this section and the following one.

Proposition 1. *Let $F_{\chi'^2}(z; \nu, \lambda)$ be the probability density function for the non-central chi-square distribution with ν degrees of freedom and non-centrality parameter λ*

$$
\begin{aligned}
F_{\chi'^2}(x;\nu,\lambda) &= e^{-\frac{\lambda}{2}} \sum_{j=0}^{+\infty} \frac{\lambda^j}{2^{\nu/2+2j}\Gamma(\nu/2+j)j!} x^{\nu/2+j-1} e^{-x/2} \\
&= \frac{e^{-\frac{\lambda+x}{2}}}{2} \left(\frac{x}{\lambda}\right)^{\frac{\nu}{2}-1} I_{\frac{\nu}{2}-1}\left(\sqrt{\lambda x}\right),
\end{aligned} \tag{34}
$$

where I_δ is the modified Bessel function of the first kind and Γ is the classic Euler Gamma function, respectively defined by:

$$
\begin{aligned}
I_\delta(x) &= \sum_{j=0}^{+\infty} \frac{x^{\delta+2j}}{2^{\delta+2j} j! \Gamma\left(\delta+j+1\right)} \\
\Gamma(x) &= \int_0^{+\infty} e^{-z} z^{x-1} dz.
\end{aligned}
$$

Then if we define for $t < T$ and for $j = 1, 2$

$$
\begin{aligned}
d_j &= \frac{4 b_j \theta_j}{\sigma_j^2} \\
n_j(t) &= \frac{4 b_j e^{-b_j t}}{\sigma_j^2 \left(1 - e^{-b_j t)}\right)},
\end{aligned}
$$

for $j = 1, 2$, conditional on V_t^j, V_T^j is distributed as $e^{-b_j(T-t)}/n_j(T-t)$ times a non-central chi-square distribution with d_j degrees of freedom and non centrality parameter $V_t^j n_j(T-t)$.

From the known properties of this distribution, we can then deduce this proposition, which will be more useful for our purpose

Proposition 2. *For $t \leq T$ and $j = 1, 2$ we have:*

$$
\mathbb{E}\left[V_T^j | V_t^j\right] = \theta_j + \left(V_t^j - \theta_j\right) e^{-b_j(T-t)} \tag{35}
$$

$$
\mathbb{V}ar\left[V_T^j | V_t^j\right] = \frac{\sigma_j^2}{b_j}\left(1 - e^{-b_j(T-t)}\right)\left(V_t^j e^{-b_j(T-t)} + \frac{\theta_j}{2}\left(1 - e^{-b_j(T-t)}\right)\right). \tag{36}
$$

Therefore we have almost obtained one of the two quantities in which we were interested, i.e. the variance of our volatility process. Then, for the sake of simplicity, Zhu in [35] makes an approximation (despite the fact that he claims his scheme to be exact) in order to calculate the mean of the volatility process, and does as if it really were an Ornstein-Uhlenbeck process, thus neglecting the fact that the mean reversion parameter is stochastic. Under these assumptions, he can then write:

$$\mathbb{V}ar\left[v^j_{t_{k+1}}|v_{t_k}\right] \approx \frac{\epsilon_j^2}{2\kappa_j}\left(1 - e^{-2\kappa_j(t_{k+1}-t_k)}\right) \text{ for } j = 1, 2. \tag{37}$$

It then follows that for $j = 1, 2$:

$$\begin{aligned}\left(\mathbb{E}\left[v^j_{t_{k+1}}|v_{t_k}\right]\right)^2 &\approx -\mathbb{V}ar\left[v^j_{t_{k+1}}|v_{t_k}\right] + \mathbb{E}\left[V^j_{t_{k+1}}|V_{t_k}\right] \\ &\approx -\frac{\epsilon_j^2}{2\kappa_j}\left(1 - e^{-2\kappa_j(t_{k+1}-t_k)}\right) + \theta_j + \left(V^j_{t_k} - \theta_j\right)e^{-b_j(t_{k+1}-t_k)}.\end{aligned}$$

Thus, the variance of the processes X^j and v^j are identical (that's precisely why Zhu did not change the parameters κ_j and ϵ_j when he defined the dynamic of X^j), and in order to match the means, remembering that the two processes have to be identical at the instants of discretization (i.e. $\forall k \in [0; N-1]$ $X^j_{t_k} = v^j_{t_k}$), we have to set for $j = 1, 2$:

$$\widehat{\lambda}^j_k = \frac{1}{1 - e^{-\kappa_j(t_{k+1}-t_k)}}\sqrt{-\frac{\epsilon_j^2}{2\kappa_j}\left(1 - e^{-2\kappa_j(t_{k+1}-t_k)}\right) + \theta_j + \left(V^j_{t_k} - \theta_j\right)e^{-b_j(t_{k+1}-t_k)}},$$

where it is understood that the quantity under the square root is set to the zero, were it to become non-positive and where we remind that $\kappa_j = b_j/2$ and $\epsilon_j = \sigma_j/2$.

Finally, the discretization scheme that we just obtained can be written for $j = 1, 2$:

$$\begin{aligned}\widehat{v}^j_{t+\Delta t} &= \widehat{v}^j_t + \frac{b_j}{2}\left(\widehat{\lambda}_j(t) - \widehat{v}^j_t\right) + \frac{\sigma_j}{2}G_j\sqrt{\Delta t} \\ \widehat{V}^j_{t+\Delta t} &= \left(\widehat{v}^j_{t+\Delta t}\right)^2,\end{aligned} \tag{38}$$

where the G_j are standard Gaussian random variables independent and independent of B_1 and B_2 defined in the previous section.

At least theoretically speaking, the scheme proposed by Zhu, does not suffer from the problems related to the computation of square roots of non-positive numbers and he argues that his scheme does not suffer from the "leaking correlation" issue, pointed out by Andersen in [3]. Nonetheless, as we mentioned it at the beginning of this section and as pointed out by Jäckel in [18], the dynamic given by (31) is structurally different from the original square-root process, as soon as the origin is attainable (i.e. as soon as $4b_j\theta_j < \sigma_j^2$). Indeed, the drift term in (31) diverges to $-\infty$, and therefore the transformed equation shows strong absorption into zero near zero, whereas the original variance SDE only exhibits zero as an attainable boundary. This is due to the fact that when the origin is attainable Itō's lemma cannot be applied to the square-root function because it is not twice differentiable at the origin. Despite this rather important technical flaw, the main purpose of (31) is to give us the idea to approximate the volatility process by an Ornstein-Uhlenbeck process moment-matched. Thus, in the end it is as if we were approximating the variance process by the square of an Ornstein-Uhlenbeck. Therefore (31) is never used as such during the simulations, which allows us to avoid the erratic behaviors mentioned by Jäckel. Besides, the results presented by Zhu in [35] and our own results tend to prove that even when the origin is attainable, the scheme does not explode and remains rather accurate.

Martingale Correction We will now try to completely develop the martingale correction technique for this scheme, which was only mentioned by Zhu. As already discussed in the previous section, in order to try to solve the martingale correction problem, we first need to calculate the quantities:

$$M_j = \mathbb{E}\left[e^{A_j V^j_{t+\Delta t}} | V^j_t\right] \text{ for } j = 1, 2.$$

Since we approximated the volatility process by an Ornstein-Uhlenbeck process, conditionally on v^j_t, $v^j_{t+\Delta t}$ follows a Gaussian distribution with mean $\frac{b_j}{2}\left(\widehat{\lambda}_j(t) - v^j_t\right)$ and variance $\frac{\sigma_j^2 \Delta t}{4}$.

Therefore conditionally on V^j_t, $V^j_{t+\Delta t}$ follows a non central χ^2 distribution with one degree of freedom. The calculations are then straightforward and lead to:

$$M_j = \frac{e^{\frac{A_j b_j^2\left(\widehat{\lambda}_j(t) - v^j_t\right)^2}{4\left(1 - A_j \sigma_j^2 \frac{\Delta t}{2}\right)}}}{\sqrt{\left(1 - A_j \sigma_j^2 \frac{\Delta t}{2}\right)}} \text{ for } j = 1, 2, \tag{39}$$

where $A_j = K^j_2 + \frac{1}{2}K^j_3$ and under the condition that $A_j < \frac{2}{\sigma_j^2 \Delta t}$.

That previous condition can, using the definition of A_j, be written:

$$\frac{\Delta t}{2}\left(\frac{b_j\rho_j}{\sigma_j}-\frac{1}{2}\right)+\frac{\rho_j}{\sigma_j}+\frac{\Delta t}{4}\left(1-\rho_j^2\right)<\frac{2}{\sigma_j^2\Delta t}$$

$$\text{i.e. } \frac{\rho_j}{4}\left(\frac{2b_j}{\sigma_j}-\rho_j\right)\Delta t^2+\frac{\rho_j}{\sigma_j}\Delta t-\frac{2}{\sigma_j^2}<0. \tag{40}$$

Given that we have already seen that as long as $\rho_j \leq 0$ the martingale correction always existed, let us suppose $\rho_j > 0$, and let's find the conditions on the time-grid parameter Δt for (40) to be satisfied. We have a second degree polynomial function, which discriminant is given by:

$$\delta_j=\frac{\rho_j}{\sigma_j^2}\left(\frac{4b_j}{\sigma_j}-\rho_j\right). \tag{41}$$

Therefore we have to distinguish between four cases:

1. $0<\rho_j<\frac{2b_j}{\sigma_j}$

   Then both δ_j and the highest degree term in (40) are positive, which leads easily to the condition:

$$\Delta t<\frac{2}{\rho_j}\frac{\sqrt{\rho_j\left(\frac{4b_j}{\sigma_j}-\rho_j\right)}-\rho_j}{2b_j-\rho_j\sigma_j}. \tag{42}$$

2. $\rho_j=\frac{2b_j}{\sigma_j}$

   Then we have just a first order inequation which solution is easily given by:

$$\Delta t<\frac{1}{b_j}. \tag{43}$$

3. $\frac{2b_j}{\sigma_j}<\rho_j<\frac{4b_j}{\sigma_j}$

   Then we have $\delta_j>0$ and the highest degree term in (40) is negative, which leads easily to the condition:

$$\Delta t < \frac{2}{\rho_j} \frac{\sqrt{\rho_j \left(\frac{4b_j}{\sigma_j} - \rho_j\right)} + \rho_j}{\rho_j \sigma_j - 2b_j}. \tag{44}$$

4. $\rho_j \geq \frac{4b_j}{\sigma_j}$

   Then both δ_j and the highest degree term in (40) are negative, and the condition (40) is therefore always verified.

For practical use however, we have roughly $0.1 \leq \sigma_j \leq 1$ and $0.5 \leq b_j \leq 2$ which allows us to think that most of the time we will be in case 1. And for example with $\sigma_j = 0.5$, $b_j = 1$ and $\rho_j = 1$, the condition reads roughly $\Delta t < 10$. Therefore, the restrictions above are very unlikely to prevent the use of the martingale correction when the model is used with common values of the parameters.

That being said, the martingale correction for the Zhu scheme finally consists, for $j = 1, 2$, on replacing the K_0^j in (24) by:

$$\widetilde{K}_0^j = -\frac{A_j b_j^2 \left(\widehat{\lambda}_j(t) - v_t^j\right)^2}{4\left(1 - A\sigma_j^2 \frac{\Delta t}{2}\right)} + \frac{1}{2} \log\left(1 - A_j \sigma_j^2 \frac{\Delta t}{2}\right) - \left(K_1^j + \frac{1}{2} K_3^j\right) \widehat{V_t^j}. \tag{45}$$

5.2.3. Quadratic Exponential scheme

Derivation of the scheme The scheme proposed by Andersen in [3] is based on an approximation of the CIR process of variance by simpler random variables which are easy to simulate. The ideas behind find their origins in observations that a non-central χ^2 distribution with moderate or high non-centrality parameter can be well represented by a power function applied to a Gaussian variable (see [26], [27] and [30]). Even though it seems that a cubic transformation is more likely to provide better results, in our case it will allow for non-positive values of the variances process, which lead Andersen to reject it and instead state that for sufficiently large (a term which will be computed later) values of V_t^j, we can write, that, conditionally on V_t^j:

$$V_{t+\Delta t}^j \approx a_j \left(c_j + Z_j\right)^2 \text{ for } j = 1, 2, \tag{46}$$

where the variables Z_j are independent standard Gaussian random variables, and a_j and c_j constants determined so that the distribution above matches (locally) the first two moments of the corresponding variance process distribution.

Nonetheless, as soon as the non-centrality parameter value, that is to say in our case the value of V_t^j, becomes too low the approximation does not work well. Thanks to (34), it is easy to obtain the following equivalent when λ goes to 0:

$$F_{\chi'^2}(x;\nu,\lambda) \underset{\lambda\to 0^+}{\sim} \frac{e^{-\frac{x}{2}}x^{\nu/2-1}}{2^{\nu/2}\Gamma(\nu/2)}. \tag{47}$$

This is (of course) exactly the probability density function of the corresponding χ^2 distribution in the centered case. Andersen, in order to allow for an easy simulation procedure, chooses to use the following asymptotic approximation for the density of the variance process:

$$\mathbb{P}\left(V_{t+\Delta t}^j \in [x, x+dx]\right) \approx \left(p_j\delta(0) + \beta_j(1-p_j)e^{-\beta_j x}\right) dx \text{ for } j = 1, 2. \tag{48}$$

That is to say a weighted mean between a Dirac mass on 0, which corresponds to the power of x in (47) and an exponential distribution with parameter β_j, which corresponds to the exponential term in (47). The constants p_j and β_j are once more determined by (local) moment-matching.

The choice of those particular approximations is motivated by the fact that it is quite easy to sample from the distributions involved. The Gaussian random variable can be very efficiently simulated with the Ziggurat method due to Marsaglia and Tsang (see [24] for details), whereas sampling according to equation (48) can be done by inversion of the distribution function. Simple and straightforward calculations give us the following inverse distribution function:

$$L_j^{-1}(u) = \mathbb{1}_{p_j < u \leq 1}\beta_j^{-1}\log\left(\frac{1-p_j}{1-u}\right). \tag{49}$$

Using the inverse distribution function sampling methods then amounts to first generate two independent uniform random numbers U_1 and U_2 and then setting for $j = 1, 2$:

$$V_{t+\Delta t}^j \approx L_j^{-1}(U_j). \tag{50}$$

In order to complete the scheme, it remains to precise the meaning of "high and low values" for V_t^j (that is to say finding a switching rule between the two approximations) and the values of the constants defined above. Andersen bases the switching rule on the value of the following quotient which value can be computed thanks to (35) and (36):

$$\Psi_j := \frac{\mathbb{V}ar\left[V_{t+\Delta t}^j | V_t^j\right]}{\left(\mathbb{E}\left[V_{t+\Delta t}^j | V_t^j\right]\right)^2} = \frac{C_1^j V_t^j + C_2^j}{\left(C_3^j V_t^j + C_4^j\right)^2}, \tag{51}$$

where:

$$C_1^j = \frac{\sigma_j^2 e^{-b_j \Delta t}}{b_j}\left(1 - e^{-b_j \Delta t}\right)$$

$$C_2^j = \frac{\theta_j \sigma_j^2}{2 b_j}\left(1 - e^{-b_j \Delta t}\right)^2$$

$$C_3^j = e^{-b_j \Delta t}$$

$$C_4^j = \theta_j \left(1 - e^{-b_j \Delta t}\right).$$

We then have the following propositions (see [3]):

Proposition 3. *For $\Psi_j \leq 2$ the quadratic transformation can be moment-matched with the exact distribution by setting:*

$$c_j^2 = 2\Psi_j^{-1} - 1 + \sqrt{2\Psi_j^{-1}\left(2\Psi_j^{-1} - 1\right)} \tag{52}$$

$$a_j = \frac{\mathbb{E}\left[V_{t+\Delta t}^j | V_t^j\right]}{1 + c_j^2}. \tag{53}$$

Proposition 4. *For $\Psi_j \geq 1$ the distribution of (48) can be moment-matched with the exact distribution by setting:*

$$p_j = \frac{\Psi_j - 1}{\Psi_j + 1} \tag{54}$$

$$\beta_j = \frac{1 - p_j}{\mathbb{E}\left[V_{t+\Delta t}^j | V_t^j\right]}. \tag{55}$$

Fortunately the two domains in the above propositions overlap. Therefore a natural procedure is to introduce a critical level $\Psi_c \in [1, 2]$ and, for $j = 1, 2$, to use (48) if $\Psi_j > \Psi_c$ and (46) if $\Psi_j \leq \Psi_c$. Following Andersen who notes that the exact choice of Ψ_c has a relatively small impact on the quality of the scheme, we set $\Psi_c = 1.5$ in our tests. in the above propositions overlap. Therefore a natural procedure is to introduce a critical level $\Psi_c \in [1,2]$ and, for $j = 1,2$, to use (48) if $\Psi_j > \Psi_c$ and (46) if $\Psi_j \leq \Psi_c$. Following Andersen who notes that the exact choice of Ψ_c has a relatively small impact on the quality of the scheme, we set $\Psi_c = 1.5$ in our tests.

Finally, notice that for the sake of efficiency, it is quite useful to cache the values of the constants $\left\{C_k^j\right\}_{\substack{k=1,..,4\\ j=1,2}}$ which depend only on the time-grid discretization, instead of computing them at each step of the scheme.

Martingale correction As usual, for solving the martingale correction problem, we have to calculate the Laplace transforms of the discretized variance processes. In the case of the QE-scheme, two cases have to distinguished, depending on if we have $\Psi_j \leq \Psi_c$ or not.

1. if $\Psi_j \leq \Psi_c$

   Then the variance has (conditionally) the same distribution as a_j times the square of Gaussian random variable with mean c_j and variance 1. The calculations are therefore quite similar to those of the Zhu scheme and lead to:

$$M_j = \frac{e^{\frac{A_j c_j^2 a_j}{1-2A_j a_j}}}{\sqrt{1-2A_j a_j}} \text{ for } j = 1,2, \tag{56}$$

   under the condition that $A_j < \frac{1}{2a_j}$.

2. if $\Psi_j > \Psi_c$

   Then we have the conditional density of the variance thanks to (48), and easy calculations lead to:

$$M_j = p_j + \frac{\beta_j(1-p_j)}{\beta_j - A_j} \text{ for } j = 1,2, \tag{57}$$

   provided that $A_j < \beta_j$.

As pointed out by Andersen, the first condition is the most restrictive, since we have $\beta_j \approx \frac{1}{\mathbb{E}\left[V_{t+\Delta t}^j | V_t^j\right]}$ where $\mathbb{E}\left[V_{t+\Delta t}^j | V_t^j\right]$ is typically a small number in the case $\Psi_j > \Psi_c$.

Besides, according to the definition of a_j we also know that:

$$a_j \geq \frac{4\sigma_j^2}{b_j}\left(1 - e^{-b_j \Delta t}\right). \tag{58}$$

Thus, the first condition will certainly be satisfied if we have:

$$\frac{\rho_j \Delta t}{4}\left(\frac{2b_j}{\sigma_j} - \rho_j\right) + \frac{\rho_j}{\sigma_j} < \frac{2b_j}{\sigma_j^2\left(1 - e^{-b_j \Delta t}\right)}. \tag{59}$$

And here we can notice that if Δt is sufficiently small (i.e. small enough for us to write $1 - e^{-b_j \Delta t} \approx b_j \Delta t$), we find exactly the same condition as in (40), so much so that the condition is not restrictive here either.

Let us also note that it is not really surprising to have the same conditions for both the Zhu and QE schemes. Indeed, both of them (at least in the case $\Psi_j \leq \Psi_c$ for QE) match the moments of the true variance distribution with the moments of a linear transformation of a non-central χ^2 distribution with one degree of freedom.

Finally, the martingale correction for the QE scheme finally consists, for $j = 1, 2$, on replacing the K_0^j in (24) by:

$$\begin{aligned}\widetilde{K}_0^j = &\left(-\frac{A_j c_j^2 a_j}{1 - 2A_j a_j} + \frac{1}{2}\log\left(1 - 2A_j a_j\right)\right)\mathbb{1}_{\Psi_j \leq \Psi_c} \\ &- \log\left(p_j + \frac{\beta_j(1-p_j)}{\beta_j - A_j}\right)\mathbb{1}_{\Psi_j > \Psi_c} - \left(K_1^j + \frac{1}{2}K_3^j\right)\widehat{V_t^j}.\end{aligned} \tag{60}$$

5.2.4. The Corrected Volatility Transformed Scheme

Derivation of the scheme We have previously seen that in [35], when trying to moment-match an Ornstein-Uhlenbeck process to the variance process, a rather important approximation was used in order not to compute the mean of the volatility. Our idea here is to use the same ideas of this scheme but without this approximation, hoping it improves its accuracy. The first step is to compute the quantity $\mathbb{E}\left[\sqrt{V_{t+\Delta t}^j}|V_t^j\right]$.

A first solution to this problem would be to use the fact pointed out by Gøing-Jaeschke and Yor in [14] that a square-root process is nothing else than a time-changed squared

Bessel process. Thus the square-root of a square-root process would be a time-changed Bessel process, which transition density is known and has a closed form (see [29] for a comprehensive study of Bessel and squared Bessel processes). Nonetheless, for the sake of simplicity and comprehensiveness, we will only use the known form of the density of the square root process given in (34). It then follows:

$$\mathbb{E}\left[\sqrt{V^j_{t+\Delta t}}|V^j_t\right] = \int_0^{+\infty}\sum_{k=0}^{+\infty}\left(n_j(\Delta t)\,e^{b_j\Delta t}\right)^{d_j/2+k}\frac{\left(n_j(\Delta t)\,V^j_t\right)^k y^{(d_j-1)/2+k}}{2^{d_j/2+2k}\Gamma(d_j/2+k)\,k!}\times$$

$$e^{-\frac{n_j(\Delta t)V^j_t}{2}\left(y+V^j_t\right)}dy$$

$$= e^{-\frac{n_j(\Delta t)V^j_t}{2}}\sqrt{\frac{2}{n_j(\Delta t)\,e^{b_j\Delta t}}}\int_0^{+\infty}\sum_{k=0}^{+\infty}\left(\frac{n_j(\Delta t)\,V^j_t}{2}\right)^k\times$$

$$\frac{u^{d_j/2+k-1/2}e^{-u}}{\Gamma(d_j/2+k)\,k!}du$$

$$= e^{-\frac{n_j(\Delta t)V^j_t}{2}}\sqrt{\frac{2}{n_j(\Delta t)\,e^{b_j\Delta t}}}\sum_{k=0}^{+\infty}\left(\frac{n_j(\Delta t)\,V^j_t}{2}\right)^k\frac{\Gamma(d_j/2+k+1/2)}{\Gamma(d_j/2+k)\,k!},$$

where we used the variable change $2u = yn_j(\Delta t)\,V^j_t$ at the second step and Fubini theorem (for positive functions) and the definition of the Euler Gamma function at the third one.

This being done, we now try to find an Ornstein-Uhlenbeck process X^j which has the same mean and variance as the volatility process. We impose that X^j has the same mean reversion parameter $\kappa_j = \frac{b_j}{2}$, and write its dynamic:

$$dX^j_t = \kappa_j\left[\widetilde{\lambda}_j(t) - X^j_t\right]dt + \widetilde{\epsilon}_j(t)dZ^j_t$$

$$X^j_0 = v^j_0,$$

where $\widetilde{\lambda}_j(t)$ and $\widetilde{\epsilon}_j(t)$ are piecewise constant on the discretization grid chosen, like in the Zhu scheme.

With the same notations as in the Zhu scheme section, we have to solve the following system:

$$\begin{cases} \mathbb{E}\left[\sqrt{V^j_{t_{k+1}}}|V^j_{t_k}\right] = \widetilde{\lambda}^k_j\left(1-e^{-\kappa_j(t_{k+1}-t_k)}\right) + v^j_{t_k}e^{-\kappa_j(t_{k+1}-t_k)} \\ \dfrac{1-e^{-b_j(t_{k+1}-t_k)}}{b_j}(\widetilde{\epsilon}^k_j)^2 = \theta_j + (V^j_{t_k}-\theta_j)e^{-b_j(t_{k+1}-t_k)} - \left(\mathbb{E}\left[\sqrt{V^j_{t_{k+1}}}|V^j_{t_k}\right]\right)^2, \end{cases}$$

from which we get:

$$\begin{cases} \widetilde{\lambda}^k_j = \dfrac{\mathbb{E}\left[\sqrt{V^j_{t_{k+1}}}|V^j_{t_k}\right] - v^j_{t_k}e^{-\kappa_j(t_{k+1}-t_k)}}{1-e^{-\kappa_j(t_{k+1}-t_k)}} \\ \widetilde{\epsilon}^k_j = \sqrt{\dfrac{b_j}{1-e^{-b_j(t_{k+1}-t_k)}}\left(\theta_j + (V^j_{t_k}-\theta_j)e^{-b_j(t_{k+1}-t_k)} - \left(\mathbb{E}\left[\sqrt{V^j_{t_{k+1}}}|V^j_{t_k}\right]\right)^2\right)}. \end{cases}$$

Finally, the scheme that we just obtained is given by:

$$\begin{aligned} \widehat{v}^j_{t+\Delta t} &= \widehat{v}^j_t + \frac{b_j}{2}\left(\widetilde{\lambda}_j(t) - \widehat{v}^j_t\right) + \widetilde{\epsilon}_j(t)G_j\sqrt{\Delta t} \qquad (61) \\ \widehat{V}^j_{t+\Delta t} &= \left(\widehat{v}^j_{t+\Delta t}\right)^2, \end{aligned}$$

where the G_j are standard Gaussian random variables independent and independent of B_1 and B_2 defined in the previous section.

Notice what we intentionally kept the mean reversion parameter κ_j constant when we defined the dynamic of X^j so that the system that we have just obtained remained linear and thus easily solvable. Therefore, in order to hope to achieve higher precision in the Zhu Scheme, we only have to use the two parameters defined above instead of those originally proposed. Nonetheless, it is quite clear that the computation of $\mathbb{E}\left[\sqrt{V^j_{t_{k+1}}}|V^j_{t_k}\right]$ can be very time consuming, since it is given by an infinite series expansion. We will now address this issue.

Numerical implementation First of all, it is rather annoying to have to compute the values of both $\Gamma\left(d_j/2+k+1/2\right)$ and $\Gamma\left(d_j/2+k\right)$ when we only need their ratio. Given that there are no simple relations between those two quantities, we chose to use the following Taylor approximation, which is a particular case of a more general formula proposed by Erdelyi and Tricomi in [12]:

$$\frac{\Gamma\left(x+\frac{1}{2}\right)}{\Gamma(x)} \underset{x\to+\infty}{=} \sqrt{x}-\frac{1}{8\sqrt{x}}+O\left(x^{-3/2}\right). \tag{62}$$

In order to understand how good the approximation is, we have represented both functions in the following graph obtained using Mathematica:

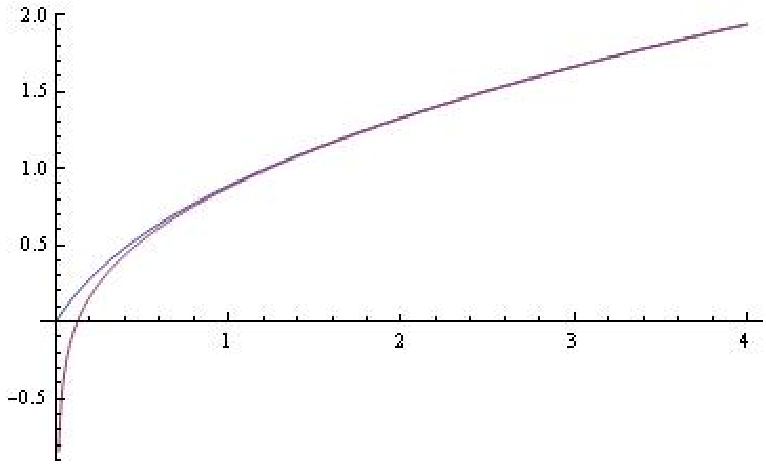


Figure 1: Gamma Ratio Taylor Expansion

It is clear that for the values in which we are interested (i.e. $x>1$) the approximation is really accurate. For example, for $x=2$ the relative error is 0.26%

Let's now have a look at the function which intervenes in the sum defining the mean we are looking at, that is to say:

$$f(k)=e^{-\frac{n_j(\Delta t)V_t^j}{2}}\sqrt{\frac{2}{n_j\left(\Delta t\right)e^{b_j\Delta t}}}\left(\frac{n_j\left(\Delta t\right)V_t^j}{2}\right)^k\frac{\Gamma\left(d_j/2+k+1/2\right)}{\Gamma\left(d_j/2+k\right)k!}. \tag{63}$$

A numerical study of this function showed us that its non-negligible values are extremely concentrated around a certain value of k depending on the values of the model parameters. As an illustration we represent it for $V_t^j=0.5$, $\Delta t=0.01$, $b_j=1$; $\sigma_j=0.5$ and $\theta_j=0.4$:

It is really fortunate because it means that for practical purposes, the sum we have to compute does not have to be infinite. Still, remains the problem of finding the interval

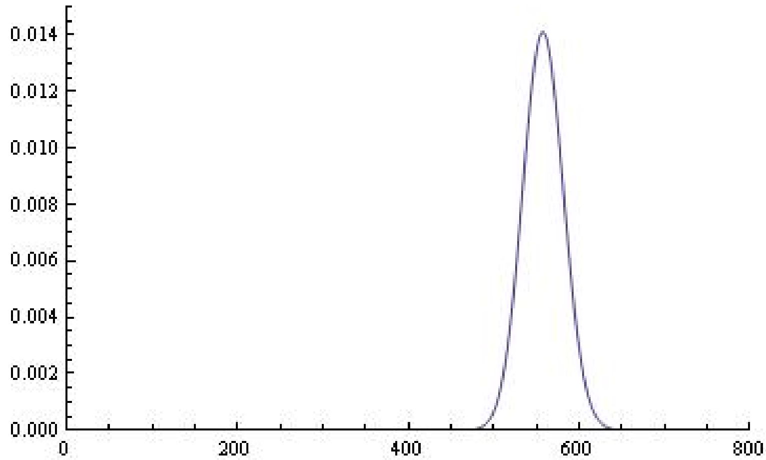


Figure 2: f(k)

on which we have to sum depending on the the model parameters values. In this regard, we tried to numerically find a relation between them but there does not seem to be a simple one which would be valid for all the parameters values. Therefore, thanks to a rather simple algorithm, we decided to cache, for a wide range of usual values of the parameters, the bounds of the interval, so that at each step of the simulation we only have to get the correct values from the cache.

Finally, we have to address the issue of computing the function $f(k)$ itself. Indeed, as we can see it in the graphic above, as we realized it when we made the cache, if implemented naively, the computation of $f(k)$ involves the calculation of enormous quantities such as 800! for example, as well as calculations of very important powers of real numbers. Even though some softwares like Mathematica or Maple are able to compute those quantity accurately, they are far greater than the capacity of usual C++ compilers for instance. Therefore we had to find a way to compute them.

First of all, it is easy to reduce this problem of computation to the first term of the sum (which index is obtained according to the value of the bound given by the cache). Indeed, let us note k_0 the first index of the sum, then the following recurrence relation gives us $f(k)$ for $k \geq k_0$:

$$\begin{cases} f(k_0) &= e^{-\frac{n_j(\Delta t)V_t^j}{2}} \sqrt{\dfrac{2}{n_j\left(\Delta t\right)e^{b_j \Delta t}}} \left(\dfrac{n_j\left(\Delta t\right)V_t^j}{2}\right)^{k_0} \dfrac{\Gamma\left(d_j/2+k_0+1/2\right)}{\Gamma\left(d_j/2+k_0\right)k_0!} \\ f(k+1) &= \dfrac{n_j\left(\Delta t\right)V_t^j}{2(k+1)} \dfrac{\Gamma\left(d_j/2+k+3/2\right)}{\Gamma\left(d_j/2+k+1/2\right)} \dfrac{\Gamma\left(d_j/2+k\right)}{\Gamma\left(d_j/2+k+1\right)} f(k) \\ &= \dfrac{n_j\left(\Delta t\right)V_t^j}{2(k+1)} \dfrac{d_j/2+k+1/2}{d_j/2+k} f(k). \end{cases}$$

Thus the computation of the terms of the sum, provided that $f(k_0)$ has been computed, does not involve any complex calculations and can be handled easily by any compiler, all the more since the value of $n_j\left(\Delta t\right)$ depends only on the model parameters and can therefore be cached before beginning the simulations.

As far as the computation of $f(k_0)$ is concerned, we have already solved the problem posed by the ratio of the Gamma functions. From our point of view the easiest way to deal with the rest of the expression is to compute first its logarithm and write:

$$f(k_0) = e^{-\frac{n_j(\Delta t)V_t^j}{2}+\frac{1}{2}(\log 2-\log(n_j(\Delta t))-b_j\Delta t)+k_0\log\left(\frac{n_j(\Delta t)V_t^j}{2}\right)-\log(k_0!)} \frac{\Gamma\left(d_j/2+k_0+1/2\right)}{\Gamma\left(d_j/2+k_0\right)}.$$

The term in the exponential can be easily computed, the logarithm of the factorial being dealt with using the following Stirling-like approximation due to Ramanujan (far better than the original one, see [28] for a proof):

$$\log(n!) \approx n\log n - n + \frac{1}{6}\log\left[n\left(1+4n\left(1+2n\right)\right)\right] + \frac{\log\pi}{2},$$

whereas the original Stirling approximation reads:

$$\log(n!) \approx n\log n - n + \frac{\log n}{2} + \frac{\log 2\pi}{2}.$$

We represented the two approximations(in red for Stirling and in green for Ramanujan) and the real values (in black). It is clear that the Ramanujan approximation is better.

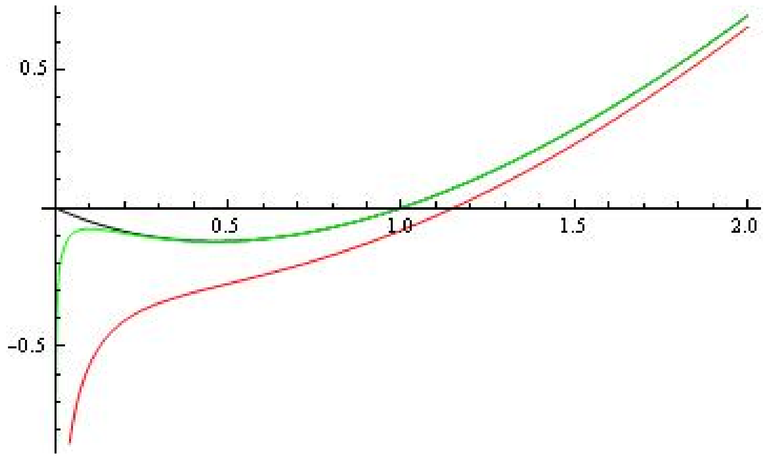


Figure 3: Approximations of $\log(n!)$

Martingale Correction The martingale correction for this scheme is almost identical to that of the Zhu scheme, since in both cases the variances are squares of Ornstein-Uhlenbeck processes. Here we thus have:

$$M_j = \frac{e^{\frac{A_j b_j^2 \left(\widetilde{\lambda}_j(t) - v_t^j\right)^2}{4\left(1-2A_j\left(\widetilde{\epsilon}_j(t)\right)^2 \Delta t\right)}}}{\sqrt{\left(1 - 2A_j\left(\widetilde{\epsilon}_j(t)\right)^2 \Delta t\right)}} \text{ for } j = 1, 2, \tag{64}$$

under the condition $A_j < \frac{1}{2(\widetilde{\epsilon}_j(t))^2 \Delta t}$.

Given that by construction we have $\widetilde{\epsilon}_j \approx \frac{\sigma_j}{2}$, the condition above is not very different from that of the Zhu Scheme, and therefore is not restrictive even with rather high positive values of ρ_j.

That being said, the martingale correction for the corrected transformed-volatility scheme finally consists, for $j = 1, 2$, on replacing the K_0^j in (24) by:

$$\widetilde{K}_0^j = -\frac{A_j b_j^2 \left(\widetilde{\lambda}_j(t) - v_t^j\right)^2}{4\left(1 - 2A_j\left(\widetilde{\epsilon}_j(t)\right)^2 \Delta t\right)} + \frac{1}{2}\log\left(1 - 2A_j\left(\widetilde{\epsilon}_j(t)\right)^2 \Delta t\right) - \left(K_1^j + \frac{1}{2}K_3^j\right)\widehat{V_t^j}. \tag{65}$$

5.2.5. Alfonsi Schemes

In his paper [2], Alfonsi derived second and third order schemes for square-roots processes, improving previous ideas of Ninomiya and Victoir (see [25]). The main idea behind those schemes is quite different from those that we have already studied, and that reason alone justifies that we mention them. Since the original paper is quite technical, we will try here to focus on the main ideas and we will not go very far into the proofs. The interested readers can find all the details in the original papers. The schemes considered here are based on a quite ancient method that dates back to Strang [32] and consists in composing different schemes. Let us remain in a general framework for the time being, and let us introduce the following general $\mathbb{R}$ valued SDE, which is supposed to admit an unique weak solution

$$X_t^x = x + \int_0^t b\left(X_s^x\right) ds + \int_0^t \sigma\left(X_s^x\right) dW_s. \tag{66}$$

The differential operator associated with the SDE is classically given by

$$\forall f \in \mathcal{C}^2(\mathbb{D}, \mathbb{R}),\ Lf(x) = b(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x), \tag{67}$$

where $\mathbb{D}$ is a subset of $\mathbb{R}$.

Suppose now that the operator L can be split into two (or more) operators L_1 and L_2 which can be associated to two other SDEs, and that we have discretization schemes of order ν for L_1 and L_2. Then we know (see [2]) that if L_1 and L_2 commute, then the composed scheme (that is to say first using the scheme for L_1 starting from x with a time-step Δt then using the scheme for L_2 with the same time-step) is going to be (under the correct technical assumptions in which we are not interested here) a discretization scheme of order ν for L. Therefore, the problem of simulation can, thanks to this method, be reduced to several simulation problems, but which can be much simpler, as we will see in the case of the square root process.

Derivation of the Second Order Scheme Turning back to our purpose and only considering for the moment a general CIR process, a "good" way (since it allows exact calculations) to split the corresponding differential operator has been proposed in [25] $\forall f \in \mathcal{C}^2(\mathbb{R}_+, \mathbb{R})$,

$$\begin{aligned} L^{CIR}f(x) &= b(\theta - x)f'(x) + \frac{1}{2}\sigma^2 x f''(x) \\ &= V_0^{CIR}f(x) + \frac{1}{2}\left(V_1^{CIR}\right)^2 f(x), \end{aligned}$$

where

$$\begin{aligned} V_0^{CIR}f(x) &= \left(b(\theta - x) - \frac{\sigma^2}{4}\right) f'(x) \\ V_1^{CIR}f(x) &= \sigma\sqrt{x}f'(x). \end{aligned}$$

Let us then note

$$\Psi_b(t) = \frac{1-e^{-bt}}{b}.$$

We then have the following proposition of [2]

Proposition 5. *Let us note*

$$K_2(\Delta t) = \mathbb{1}_{\sigma^2 > 4b\theta} e^{\frac{b\Delta t}{2}} \left(\left(\frac{\sigma^2}{4} - \theta b \right) \Psi_b \left(\frac{\Delta t}{2} \right) + \left[\sqrt{e^{\frac{b\Delta t}{2}} \left(\frac{\sigma^2}{4} - \theta b \right) \Psi_b \left(\frac{\Delta t}{2} \right)} + \frac{\sigma}{2}\sqrt{3\Delta t} \right]^2 \right),$$

and let us define the discrete random variable Y so that

$$\mathbb{P}\left(Y=0\right) = \frac{2}{3}$$

$$\mathbb{P}\left(Y=\sqrt{3}\right) = \frac{1}{6}$$

$$\mathbb{P}\left(Y=-\sqrt{3}\right) = \frac{1}{6}.$$

Then the discretization scheme

$$\widehat{V}_{t+\Delta t} = e^{-\frac{b\Delta t}{2}} \left(\sqrt{\left(b\theta - \frac{\sigma^2}{4}\right) \Psi_b\left(\frac{\Delta t}{2}\right) + e^{-\frac{b\Delta t}{2}} \widehat{V}_t} + \frac{\sigma}{2}\sqrt{\Delta t} Y \right)^2 + \left(b\theta - \frac{\sigma^2}{4}\right) \Psi_b \left(\frac{\Delta t}{2}\right), \tag{68}$$

is a second order discretization scheme for the CIR process when $\widehat{V}_t \geq K_2(\Delta t)$.

The scheme above is (almost) the one proposed by Ninomiya and Victoir in [25], but as we can see, it is no longer defined when the CIR process becomes too close to zero. The idea of Alfonsi is then to approximate in this case the CIR process by a discrete random variable which takes only two values and which matches the first two moments of the CIR process. It is the same approach as Andersen but with much simpler random variables. The calculations are then straightforward and lead Alfonsi to the following proposition

Proposition 6. *Let $U \sim \mathcal{U}[0,1]$ (independent of Y of the previous proposition) and let us note*

$$u_q(\Delta t, V_t) = \mathbb{E}\left[\, (V_{t+\Delta t})^q \mid V_t \right],$$

the values of u_1 and u_2 being given by proposition 1.

If we note

$$\pi(\Delta t, \widehat{V}_t) = \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{u_1^2(\Delta t, \widehat{V}_t)}{u_2(\Delta t, \widehat{V}_t)}},$$

then the following scheme

$$\widehat{V}_{t+\Delta t} = \mathbb{1}_{U \leq \pi(\Delta t, \widehat{V}_t)} \frac{u_1(\Delta t, \widehat{V}_t)}{2\pi(\Delta t, \widehat{V}_t)} + \mathbb{1}_{U > \pi(\Delta t, \widehat{V}_t)} \frac{u_1(\Delta t, \widehat{V}_t))}{2(1 - \pi(\Delta t, \widehat{V}_t))}, \tag{69}$$

is a second order discretization scheme for the CIR process when $\widehat{V}_t < K_2(\Delta t)$.

Finally, in order to use this scheme, it suffices, at each step of the discretization, to use one of the two previous propositions, depending on the value of $\widehat{V}_t$ with respect to $K_2(\Delta t)$. Its use in the framework of the Double Heston model is then clear, since the two volatility processes are not correlated, and therefore each can be simulated thanks to the above method.

5.2.6. Derivation of the Third Order Scheme

The main ideas in this case are the same as with the second order scheme. However the calculations are more tedious and the scheme is a lot more complicated. Therefore, we will only give the scheme itself without further explanations, and we urge the interested readers to check the details for themselves in [2], which lecture is very rewarding.

Proposition 7. *Let ϵ and ζ be respectively independent uniform random variables on $\{-1, 1\}$ and $\{1, 2, 3\}$, let Z be a discrete random variable so that*

$$\mathbb{P}\left(Z = \sqrt{3+\sqrt{6}}\right) = \mathbb{P}\left(Z = -\sqrt{3+\sqrt{6}}\right) = \frac{\sqrt{6}-2}{4\sqrt{6}}$$

$$\mathbb{P}\left(Z = \sqrt{3-\sqrt{6}}\right) = \mathbb{P}\left(Z = -\sqrt{3-\sqrt{6}}\right) = \frac{1}{2} - \frac{\sqrt{6}-2}{4\sqrt{6}},$$

and define

$$\begin{aligned} K_3(\Delta t) = \Psi_{-b}(\Delta t) \Bigg[& \mathbb{1}_{\frac{4b\theta}{3} < \sigma^2 < 4b\theta} \left(\sqrt{\frac{\sigma^2}{4} - b\theta + \frac{\sigma}{\sqrt{2}}\sqrt{b\theta - \frac{\sigma^2}{4}}} + \frac{\sigma}{2}\sqrt{3+\sqrt{6}} \right)^2 \\ & + \mathbb{1}_{\sigma^2 > 4b\theta} \left(\frac{\sigma^2}{4} - b\theta + \left(\sqrt{\frac{\sigma}{\sqrt{2}}\sqrt{\frac{\sigma^2}{4} - b\theta}} + \frac{\sigma}{2}\sqrt{3+\sqrt{6}} \right)^2 \right) \\ & + \mathbb{1}_{\sigma^2 \leq \frac{4b\theta}{3}} \frac{\sigma}{\sqrt{2}}\sqrt{b\theta - \frac{\sigma^2}{4}} \Bigg]. \end{aligned}$$

Define also the three following functions

$$\varphi_0(x) = x + \left(b\theta - \frac{\sigma^2}{4}\right)\Psi_{-b}(\Delta t)$$

$$\varphi_1(x) = \left[\left(\sqrt{x} + \frac{\sigma}{2}\sqrt{\Psi_{-b}(\Delta t)}Y\right)^+\right]^2$$

$$\varphi_2(x) = x + \frac{\sigma\epsilon}{\sqrt{2}}\sqrt{\left|b\theta - \frac{\sigma^2}{4}\right|\Psi_{-b}(\Delta t)}.$$

Then the following scheme

$$\begin{aligned}\widehat{V}_{t+\Delta t} = e^{-b\Delta t}\Big[&\mathbb{1}_{\zeta=1}\left(\mathbb{1}_{\sigma^2\leq 4b\theta}\varphi_1\circ\varphi_0\circ\varphi_2(\widehat{V}_t) + \mathbb{1}_{\sigma^2>4b\theta}\varphi_0\circ\varphi_1\circ\varphi_2(\widehat{V}_t)\right)\\ &+\mathbb{1}_{\zeta=2}\left(\mathbb{1}_{\sigma^2\leq 4b\theta}\varphi_1\circ\varphi_2\circ\varphi_0(\widehat{V}_t) + \mathbb{1}_{\sigma^2>4b\theta}\varphi_0\circ\varphi_2\circ\varphi_1(\widehat{V}_t)\right)\\ &+\mathbb{1}_{\zeta=3}\left(\mathbb{1}_{\sigma^2\leq 4b\theta}\varphi_2\circ\varphi_1\circ\varphi_0(\widehat{V}_t) + \mathbb{1}_{\sigma^2>4b\theta}\varphi_2\circ\varphi_0\circ\varphi_1(\widehat{V}_t)\right)\Big],\end{aligned}$$

is a third order scheme for the CIR process when $\widehat{V}_t \geq K_3(\Delta t)$.

The idea to find a third order scheme for the CIR process when we are in the neighborhood of zero is the same as with the second order scheme. We approximate the (conditional) CIR process by a discrete random variable which can only take two values and matches the first three moments. The calculations are longer but lead eventually to the following proposition (see [2] for a proof)

Proposition 8. *Let* $U \sim \mathcal{U}[0,1]$ *(independent of* Y, ϵ *and* ζ *of the previous proposition), and define*

$$s(\Delta t, \widehat{V}_t) = \frac{u_3(\Delta t, \widehat{V}_t) - u_1(\Delta t, \widehat{V}_t)u_2(\Delta t, \widehat{V}_t)}{u_2(\Delta t, \widehat{V}_t) - u_1^2(\Delta t, \widehat{V}_t)}$$

$$p(\Delta t, \widehat{V}_t) = \frac{u_1(\Delta t, \widehat{V}_t)u_3(\Delta t, \widehat{V}_t) - u_2^2(\Delta t, \widehat{V}_t)}{u_2(\Delta t, \widehat{V}_t) - u_1^2(\Delta t, \widehat{V}_t)}$$

$$\delta(\Delta t, \widehat{V}_t) = \sqrt{s^2(\Delta t, \widehat{V}_t) - 4p(\Delta t, \widehat{V}_t)}$$

$$\pi(\Delta t, \widehat{V}_t) = \frac{1}{\delta(\Delta t, \widehat{V}_t)}\left(u_1(\Delta t, \widehat{V}_t) - \frac{s(\Delta t, \widehat{V}_t) - \delta(\Delta t, \widehat{V}_t)}{2}\right),$$

where the third conditional moment of the CIR process can be calculated thanks to its known density and is equal to

$$u_3(\Delta t, V_t) = \sigma^2 \Psi_b(\Delta t) \left[2V_t^2 e^{-2b\Delta t} + \Psi_b(\Delta t) \left(b\theta + \frac{\sigma^2}{2} \right) \left(3V_t e^{-b\Delta t} + b\theta \Psi_b(\Delta t) \right) \right]$$

$$+ u_1(\Delta t, V_t) u_2(\Delta t, V_t).$$

Then the scheme

$$\widehat{V}_{t+\Delta t} = \frac{1}{2} \left[\mathbb{1}_{U < \pi(\Delta t, \widehat{V}_t)} \left(s(\Delta t, \widehat{V}_t) + \delta(\Delta t, \widehat{V}_t) \right) + \mathbb{1}_{U \geq \pi(\Delta t, \widehat{V}_t)} \left(s(\Delta t, \widehat{V}_t) - \delta(\Delta t, \widehat{V}_t) \right) \right],$$

is a third order scheme for the CIR process when $\widehat{V}_t < K_3(\Delta t)$.

5.2.7. Martingale Correction

We will only treat the martingale correction for the Alfonsi schemes in the second-order case, the calculations being far too complicated with the third-order scheme for the eventual gain that could come from it. That being said, since only independent discrete random variables intervene in (68) and (69), the conditional Laplace transform is easy to get, and after some calculations we have for $j = 1, 2$

$$M_j = \mathbb{1}_{\widehat{V}_t^j \geq K_2^j(\Delta t)} \frac{2}{3} e^{-A_j \left(\beta_j + e^{-\frac{b_j \Delta t}{2}} \alpha_j \right)} \left[2 + e^{-\frac{3\sigma_j^2}{4} \Delta t A_j e^{-\frac{b_j \Delta t}{2}}} \operatorname{ch} \left(A_j e^{-\frac{b_j \Delta t}{2}} \sigma_j \sqrt{3\alpha_j \Delta t} \right) \right]$$

$$+ \mathbb{1}_{\widehat{V}_t^j < K_2^j(\Delta t)} \left[\pi_j(\Delta t, \widehat{V}_t^j) e^{-A_j \frac{u_1^j(\Delta t, \widehat{V}_t^j)}{2\pi_j(\Delta t, \widehat{V}_t^j)}} + (1 - \pi_j(\Delta t, \widehat{V}_t^j)) e^{-A_j \frac{u_1^j(\Delta t, \widehat{V}_t^j)}{2(1 - \pi_j(\Delta t, \widehat{V}_t^j))}} \right], \quad (70)$$

where

$$\beta_j = \left(b_j \theta_j - \frac{\sigma_j^2}{4} \right) \Psi_{b_j} \left(\frac{\Delta t}{2} \right)$$

$$\alpha_j = \beta_j + e^{-b_j \frac{\Delta t}{2}} \widehat{V}_t^j.$$

The Laplace transform above always exists, and the martingale correction for the second order Alfonsi scheme finally consists, for $j = 1, 2$, on replacing the K_0^j in (24) by:

$$\widetilde{K}_0^j = -\log M_j - \left(K_1^j + \frac{1}{2} K_3^j \right) \widehat{V_t^j}. \quad (71)$$

5.3. Convergence Considerations

A formal analysis of the convergence properties for the schemes proposed in this paper is difficult and complicated by the fact that the X process may not have any high-order moments. As such, the usual examination of (weak) convergence of expectations of polynomials of X is not always meaningful. While we could, in principle, undertake an examination of the convergence of expectations on selected slow-growing payouts (e.g. call options), the technicalities of such an analysis are considerable and we skip it. (See [23] for examples of this type of analysis). Instead, as pointed out by Andersen in [3] we focus on a simpler concept, namely that of weak consistency. As shown in [20] p. 328, there is a strong link between weak consistency and weak convergence. We will not deal here with the schemes of Alfonsi, since they have been proven to converge in [2], and with the full truncation Euler scheme, since it has also been proven to converge (see for example [5]).

Proposition 9. *The QE-M, Zhu and corrected volatility transformed schemes are weakly consistent. That is to say that, conditional on $\widehat{S}_t$, $\widehat{V}_t^1$ and $\widehat{V}_t^2$, we have for all three schemes*

$$\mathbb{E}_t\left[\frac{1}{\Delta t}\log\frac{\widehat{S}_{t+\Delta t}}{\widehat{S}_t}\right]\underset{\Delta t\to 0}{\rightarrow} r\Delta t-\frac{1}{2}\left(\widehat{V}_t^1+\widehat{V}_t^2\right),\ \mathbb{V}ar_t\left[\frac{1}{\sqrt{\Delta t}}\log\frac{\widehat{S}_{t+\Delta t}}{\widehat{S}_t}\right]\underset{\Delta t\to 0}{\rightarrow}\widehat{V}_t^1+\widehat{V}_t^2$$

$$\mathbb{E}_t\left[\frac{\widehat{V}_{t+\Delta t}^j-\widehat{V}_t^j}{\Delta t}\right]\underset{\Delta t\to 0}{\rightarrow} b_j\left(\theta_j-\widehat{V}_t^j\right),\ \mathbb{V}ar_t\left[\frac{\widehat{V}_{t+\Delta t}^j-\widehat{V}_t^j}{\sqrt{\Delta t}}\right]\underset{\Delta t\to 0}{\rightarrow}\sigma_j^2\widehat{V}_t^j$$

$$\mathbb{C}ov_t\left[\frac{\widehat{V}_{t+\Delta t}^j-\widehat{V}_t^j}{\sqrt{\Delta t}},\frac{\log\widehat{S}_{t+\Delta t}-\log\widehat{S}_t}{\sqrt{\Delta t}}\right]\underset{\Delta t\to 0}{\rightarrow}\rho_j\sigma_j\widehat{V}_t^j.$$

Proof. Using (24) and proposition 2 we have, conditional on $\widehat{S}_t$, $\widehat{V}_t^1$ and $\widehat{V}_t^2$

$$\mathbb{E}\left[\frac{\log\widehat{S}_{t+\Delta t}-\log\widehat{S}_t}{\Delta t}\right]=\frac{K_0^1+K_0^2}{\Delta t}+\frac{\widehat{V}_t^1}{\Delta t}\left(K_1^1+K_2^1e^{-b_1\Delta t}\right)+\frac{\widehat{V}_t^2}{\Delta t}\left(K_1^2+K_2^2e^{-b_2\Delta t}\right)$$
$$+\theta_1\frac{1-e^{-b_1\Delta t}}{\Delta t}K_2^1+\theta_2\frac{1-e^{-b_2\Delta t}}{\Delta t}K_2^2+r\Delta t.$$

Then using the values of the constants $\left(K_i^j\right)_{1\leq i,j\leq 2}$

$$
\begin{aligned}
\mathbb{E}\left[\frac{\log \widehat{S}_{t+\Delta t} - \log \widehat{S}_t}{\Delta t}\right] &= \widehat{V}_t^1 \left(\frac{1}{2}\left(\frac{b_1\rho_1}{\sigma_1} - \frac{1}{2}\right)\left(1 + e^{-b_1\Delta t}\right) - \frac{\rho_1}{\sigma_1}\frac{1 - e^{-b_1\Delta t}}{\Delta t}\right) \\
&+ \widehat{V}_t^2 \left(\frac{1}{2}\left(\frac{b_2\rho_2}{\sigma_2} - \frac{1}{2}\right)\left(1 + e^{-b_2\Delta t}\right) - \frac{\rho_2}{\sigma_2}\frac{1 - e^{-b_2\Delta t}}{\Delta t}\right) \\
&- \frac{\rho_1 b_1 \theta_1}{\sigma_1} - \frac{\rho_2 b_2 \theta_2}{\sigma_2} + \theta_1 \frac{1 - e^{-b_1\Delta t}}{\Delta t} K_2^1 + \theta_2 \frac{1 - e^{-b_2\Delta t}}{\Delta t} K_2^2 + r\Delta t \\
&\underset{\Delta t \to 0}{\longrightarrow} r\Delta t - \frac{1}{2}\left(\widehat{V}_t^1 + \widehat{V}_t^2\right).
\end{aligned}
$$

The calculations for $\mathbb{V}ar\left[\frac{\log \widehat{S}_{t+\Delta t} - \log \widehat{S}_t}{\sqrt{\Delta t}}\right]$ are identical.

Then the results for the expectation and the variance of the discretization schemes for $\widehat{V}_t^1$ and $\widehat{V}_t^2$ are clearly satisfied, since by construction all three schemes match exactly the first two moments of the real variance process given in proposition 2.

Finally, for the last equality of the proposition we have, conditional on $\widehat{S}_t$, $\widehat{V}_t^1$ and $\widehat{V}_t^2$ and using (24)

$$
\begin{aligned}
\mathbb{C}ov\left[\frac{\widehat{V}_{t+\Delta t}^j - \widehat{V}_t^j}{\sqrt{\Delta t}}, \frac{\log \widehat{S}_{t+\Delta t} - \log \widehat{S}_t}{\sqrt{\Delta t}}\right] &= \frac{1}{\Delta t}\mathbb{C}ov\left[\widehat{V}_{t+\Delta t}^j, \log \widehat{S}_{t+\Delta t}\right] \\
&= \frac{K_2^j}{\Delta t}\mathbb{V}ar\left[\widehat{V}_{t+\Delta t}^j\right] \\
&\underset{\Delta t \to 0}{\longrightarrow} \rho_j \sigma_j \widehat{V}_t^j.
\end{aligned}
$$

□

6. Control variate

Control variate is one of the most effective tool for improving Monte Carlo convergence. The core idea of control variate is to exploit the error made on estimates of known quantities to correct estimates of unknown quantities. Let $(X_i)_{1 \le i \le n}$ and $(Y_i)_{1 \le i \le n}$ be the simulations of two payoffs X and Y. Furthermore, let assume that $\mathbb{E}[X]$ is known. The standard Monte-Carlo estimators of $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{72}$$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i. \tag{73}$$

Both estimators are unbiased. Therefore, for any fixed β, we can construct a second unbiased estimator for $\mathbb{E}(Y)$:

$$\bar{Y}(\beta) = \bar{Y} + \beta(\mathbb{E}[X] - \bar{X}). \tag{74}$$

This estimator is unbiased because

$$\begin{aligned}
\mathbb{E}[\bar{Y}(\beta)] &= \mathbb{E}[\bar{Y} + \beta(\mathbb{E}[X] - \bar{X})] \\
&= \mathbb{E}[\bar{Y}] + \beta(\mathbb{E}[\mathbb{E}[X]] - \mathbb{E}[\bar{X}])] \\
&= \mathbb{E}[\bar{Y}] + \beta(\mathbb{E}[X] - \mathbb{E}[X])] \\
&= \mathbb{E}[\bar{Y}].
\end{aligned}$$

Now, in order to apply control variate, we still have to choose appropriately the β coefficient. For the moment, our new estimator matches the expectation of Y. A desirable property is to match also the variance of Y.

$$\begin{aligned}
\mathbb{V}ar[\bar{Y}(\beta)] &= \mathbb{V}ar[\bar{Y} + \beta(\mathbb{E}[X] - \bar{X})] \\
&= \sigma_Y^2 - 2\beta\sigma_X\sigma_Y\rho_{XY} + \beta^2\sigma_X^2, \qquad (75)
\end{aligned}$$

where $\sigma_X^2 = \mathbb{V}ar[X]$, $\sigma_Y^2 = \mathbb{V}ar[Y]$ and $\langle X, Y\rangle = \rho_{XY}$.

The control variate estimator has smaller variance than the standard estimator, which corresponds to the case $\beta = 0$ if

$$\beta^2\sigma_X^2 - 2\beta\sigma_X\sigma_Y\rho_{XY} < 0.$$

The optimal coefficient β minimizes the variance (75) and is given by

$$\beta^* = \frac{\sigma_Y}{\sigma_X}\rho_{XY} = \frac{\mathbb{C}ov[X,Y]}{\mathbb{V}ar[X,Y]}. \tag{76}$$

The effectiveness of control variates depends strongly on the choice of the reference payoff. It can indeed be shown that control variate works better when X and Y are strongly correlated, which is quite intuitive. For a rigorous proof and further insights on control variate, Chapter 4 of Glasserman's book [15] is an excellent reference.

In our numerical examples, since we are pricing a call using Monte-Carlo, a good reference is the spot value. The call price and the spot value are indeed strongly correlated.

7. Numerical results

7.1. Test cases

We consider a call option of strike K with maturities 1 year (case I) and 10 years (case II). The other parameters values are given in appendix A. They have been chosen to match roughly market conditions which are not restrictive. In order to evaluate the Call option price using a specific discretization scheme, we use the so-called Monte-Carlo method, with or without the control-variate techniques described above. All the numerical results are given in the appendix.

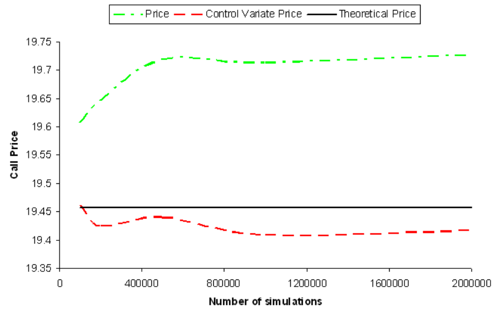
7.2. Conclusions of the Numerical Tests

Numerical results First of all, it is rather evident when considering our results that the Monte-Carlo simulations really benefit from the use of the control variate method. Indeed, with our four schemes, the gain is often quite spectacular, ranging almost always between a factor 2 and a factor 10.

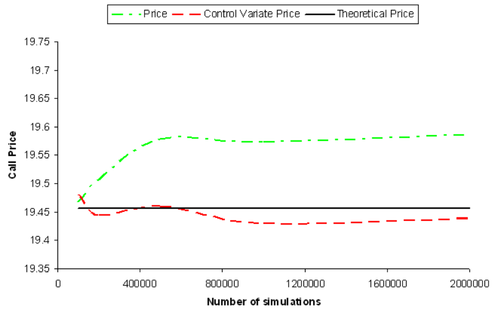
When it comes to the efficiency of the schemes, a quick examination of the results in the appendix shows without any doubt that the QE scheme of Andersen and the two schemes of Alfonsi outperform by far all the other schemes. In both cases I and II, even with a rather small number of simulations (namely 100000) the Monte-Carlo simulation using control variate almost always gives a price with a bias non-significantly different from zero. Then, it seems that the Zhu scheme and the corrected volatility transformed scheme give equivalent results. Therefore, the theoretical correction to the original Zhu scheme does not seem to improve the accuracy of the Monte-Carlo simulations. Besides, even though they are not as accurate as the QE or Alfonsi schemes, they still perform quite well, the relative errors almost always remaining below 1%. Finally, as pointed out in the literature, the full truncation Euler scheme presents a quite important bias which cannot be completely eliminated using control variate.
Now when comparing the Alfonsi schemes and QE-M schemes, with all our simulations (which are not all reported in the present paper) we came to the conclusions that even

though for all three schemes, the biases obtained were not significantly different from zero, the QE-M scheme performed better than the second-order scheme of Alfonsi 75% of the time, whereas the third order Alfonsi scheme performs better than the QE-M scheme 75% of the time, and almost 95% of the time when considering out-of-the-money call options.

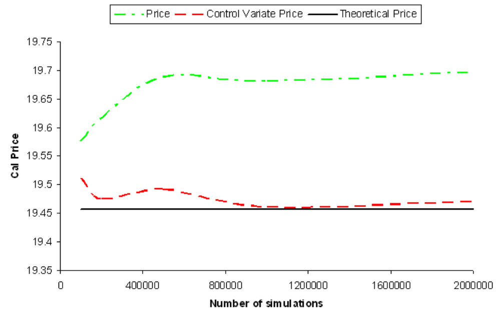
Convergence We represent below the behavior of our schemes in terms of convergence w.t.r. the number of simulations used (figure 4) and the number of steps (figure 5)



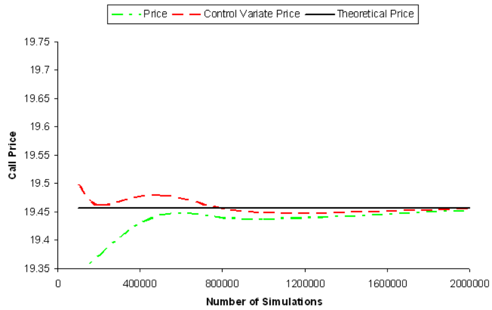
(a) Euler scheme



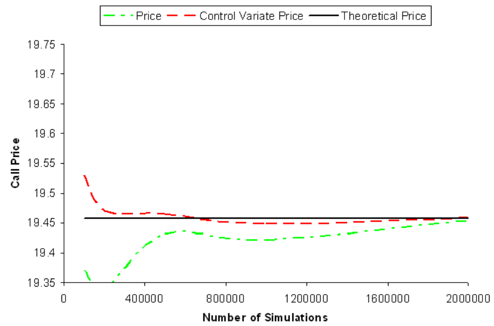
(b) Zhu scheme



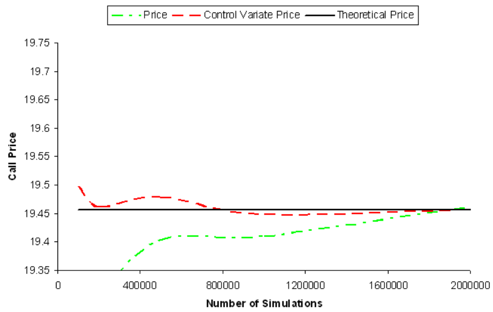
(c) Corrected volatility transformed scheme



(d) QE-M scheme

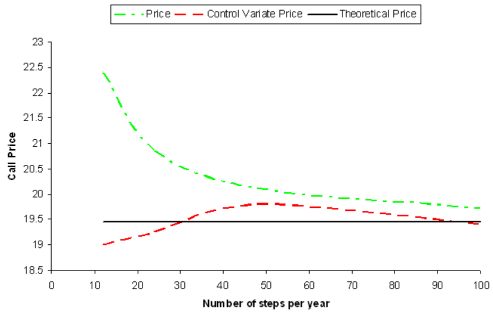


(e) Alfonsi second-order scheme

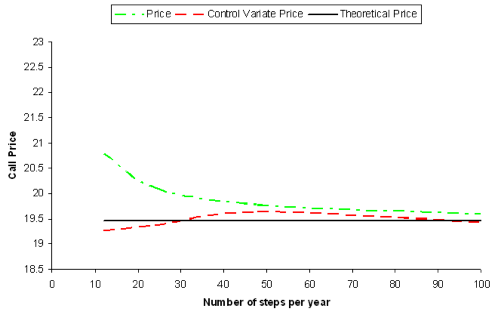


(f) Alfonsi third-order Scheme scheme

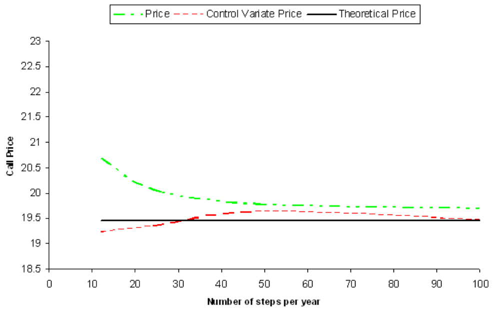
Figure 4: Convergence of the schemes with the number of simulations



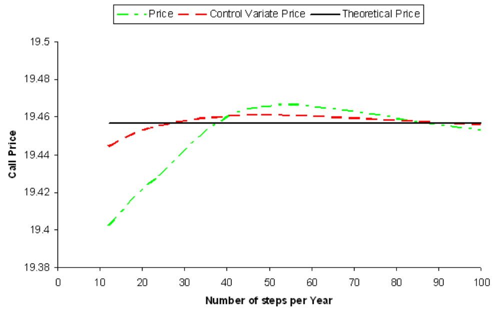
(a) Euler scheme



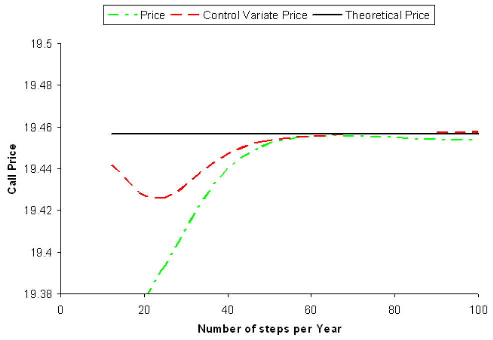
(b) Zhu scheme



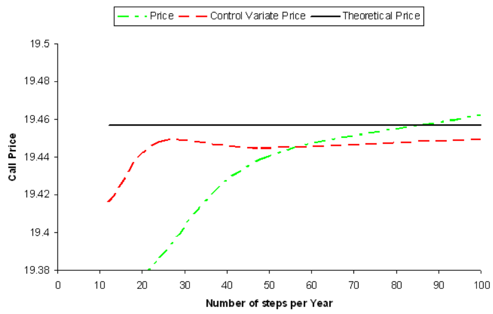
(c) Corrected volatility transformed scheme



(d) QE-M scheme



(e) Alfonsi second-order scheme



(f) Alfonsi third-order Scheme scheme

Figure 5: Convergence of the schemes with the number of steps

It is then quite clear that once more the QE-M and Alfonsi schemes converge much faster than the other schemes, as well with respect to the number of simulations as with respect to the number of steps per year used. The Zhu-scheme and the corrected volatility transformed schemes have the same behavior, even though the convergence with the number of steps seems a little bit faster with the corrected scheme. Finally, the Euler scheme bias remains important even for large numbers of steps and simulations.

Computation time Of course the results above would not be comprehensive without mentioning the computational time of all the schemes. In case I we have the following results for 100000 simulations:

	Euler	**Zhu**	**ExactZhu**	**QE-M**	**Alfonsi 2nd**	**Alfonsi 3rd**
Time	0.797	0.891	14.843	2.547	1.204	2.010

Table 1: Computational Times for the different schemes (s)

The Zhu and Euler Scheme take approximately the same amount of time to compute the call option price, making the Zhu scheme quite efficient in this regard. On the other hand, the corrected Volatility transformed scheme takes a quite large amount of time without any significant improvement in the accuracy of the Monte-Carlo pricing. As expected, the QE-M scheme is rather more time-consuming than the Zhu or the Euler Scheme, but with a result more than 5 times more precise, it remains highly competitive. Finally, the Alfonsi schemes are both faster than the QE-M scheme -roughly 50% faster for the second order scheme and 25% for the third order one- confirming that, from our results, the best scheme to simulate efficiently the Heston and Double Heston models is the third-order scheme of Alfonsi.

8. Further work

After addressing the issue of a double variance process, it may be interesting to consider the discretization of more complex variance processes. For instance, we can consider Wishart variance processes, as described in Da Fonseca et al [13]. Is there any means to adapt the QE-M scheme ? Or the Alfonsi schemes ? How do they compare in terms of convergence speed with the Euler scheme ? And to Wishart specific discretization schemes ?

9. Conclusion

In this paper we have addressed the simulation and discretization issues of the Double Heston model introduced by Christoffersen et al in [8]. They have shown that not only is this model analytically tractable, but the addition of a volatility factor greatly improves the model's flexibility to capture the volatility term structure. The model can then be calibrated more easily and does not suffer from the over-parametrization of a, say, time-dependent Heston model. Our purpose was then to extend the most interesting results concerning the original Heston model to this new two-factor framework. After addressing the issue of European options pricing using the characteristic function method, we have generalized some of the most efficient discretization schemes created for Monte-Carlo pricing with square-root processes, namely the QE-M, Alfonsi and Zhu schemes, and we have explored the eventual benefit of a theoretical correction to the latter. Our conclusions are that in any case the Alfonsi or QE-M schemes should be used. Even though the Zhu scheme performs quite well in comparison to the Euler scheme, especially since they both have approximately identical computational times, the precision for the QE-M or Alfonsi schemes outperforms them by far. Besides, the third-order Alfonsi scheme seems to perform better than the QE-M scheme in most cases, and is also quite faster, making it without a doubt the scheme to use when dealing with Heston or Double Heston models. Unfortunately, our correction to the Zhu scheme does not really improve its convergence properties and is still too time consuming, but it has at least the merit to show that other ways should be explored for improving the original approximation of Zhu. Finally, we are sure that any future improvements over the existing schemes for the Heston Model would be easily adapted to the Double Heston framework.

Acknowledgments

Both authors are very grateful to Aurélien Alfonsi and Abdelkodousse Ahdida for pointing out several very interesting references and for their insightful remarks.

References

[1] Phillip Abbot. Tricks of the trade : Legendre-Gauss quadrature. *Mathematica Journal*, 9:689–691, 2005.

[2] Aurélien Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Mathematics of Computation*, 2008.

[3] Leif Andersen. Efficient simulation of the Heston stochastic volatility model. *Working Paper*, December 2005.

[4] Leif Andersen and Vladimir Piterbarg. Efficient simulation of the Heston stochastic volatility model. *Working Paper*, July 2005.

[5] A. Berkaoui, M. Bossy, and A. Diop. Euler scheme for SDEs with non-lipschitz diffusion coefficient: strong convergence. *RR-5637, INRIA*, July 2005.

[6] Mark Broadie and Ozgur Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 54(2), 2006.

[7] Peter Carr and Dilip B. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2:61:73, 1999.

[8] Peter Christoffersen, Steven Heston, and Kris Jacobs. The shape and term structure of the index option smirks : why multifactor stochastic volatility models work so well. *Working paper,McGill University*, February 2009.

[9] John Cox, Jonathan Ingersoll, and Stephen Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, March 1985.

[10] D.Duffie, D.Filipovic, and W. Schachermayer. Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053, 2003.

[11] Daniel Dufresne. The integrated square-root process. *Working Paper, University of Montréal*, 2001.

[12] A. Erdelyi and F. G. Tricomi. The asymptotic expansion of a ratio of Gamma functions. *Pacific Journal of Mathematics*, 1(1):133–142, 1951.

[13] José Da Fonseca, Martino Grasselli, and Claudio Tebaldi. Option pricing when correlations are stochastic : an analytical framework. *Research report, ESILV*, September 2006.

[14] Anna Göing-Jaeschke and Marc Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[15] Paul Glasserman. *Monte Carlo Methods in Financial Engineering.* Springer-Verlag, 2004.

[16] Christian Gouriéroux. Continuous time Wishart process for stochastic risk. *CREF*, October 2006.

[17] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

[18] Peter Jäckel. Stochastic volatility models : Past, present and future. *The Best of Wilmott 1 : Incorporating the Quantitative Finance Review*, pages 379–390, 2004.

[19] Christian Kahl and Peter Jäckel. Not-so-complex logarithms in the Heston model. *Wilmott magazine*, pages 94–103, September 2005.

[20] P. Kloeden and E. Platen. *Numerical solution of stochastic differential equations.* 3rd ed. Springer Verlag, 1999.

[21] Alan L. Lewis. A simple option formula for general jump-diffusion and other exponential Lévy processes. *Working paper, Envision Financial Systems*, August 2001.

[22] Richard Lord and Christian Kahl. Complex logarithms in Heston-like models. *SSRN Working Paper Series*, February 2008.

[23] Roger Lord, Remmert Koekkoek, and Dick van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Working Paper, Tinbergen Institute*, May 2006.

[24] George Marsaglia and Wai Wan Tsang. The Ziggurat method for generating random variables. *Journal of Statistical Software*, 5(8), October 2000.

[25] S. Ninomiya and N. Victoir. Weak approximation of stochastic differential equations and applications to derivatives pricing. *Applied Mathematical Finance*, 15(2):107–121, 2008.

[26] P. Patnaik. The non-central χ^2 and F distributions and their applications. *Biometrika*, 36:202–232, 1949.

[27] E.S. Pearson. Note on an approximation to the distribution of non-central χ^2. *Biometrika*, 46:364, 1959.

[28] S. Raghavan and S. S. Rangachari. *S. Ramanujan : The Lost Notebook and Other Unpublished Papers*. Springer, p. 339, 1988.

[29] Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion.* Springer-Verlag, third edition, 1999.

[30] J. Roy and Jan Mohamad. An approximation to the non-central χ^2 distribution. *The Indian Journal of Statistics, Series A*, 26(1):81–84, July 1964.

[31] Arthur Sepp. Fourier transform for option pricing under affine jump-diffusions: An overview. *Working paper,Tartu University*, September 2003.

[32] G. Strang. On the construction and comparison of difference schemes. *SIAM Journal on numerical analysis*, 5:506–517, 1968.

[33] Pelsser Van Haastrecht, Lord and Schrager. Pricing long-maturity equity and fx derivatives with stochastic interest rates and stochastic volatility. *SSRN*, 2008.

[34] Roger W.Lee. Option pricing by transform methods : Extensions, unification and error control. *Journal of Computational Finance*, 7(3):51–86, 2004.

[35] Jianwei Zhu. A simple and exact simulation approach to Heston model. *SSRN Working Paper Series*, July 2008.

A. Numerical results

All our numerical results have been performed with the following parameters.

Parameter	Value
Spot	61.90
Short rate	3%
Times Steps per Year	24
Initial Vol 1	60%
Initial Vol 2	70%
Vol of Vol 1	10%
Vol of Vol 2	20%
Mean Reversion 1	90%
Mean Reversion 2	120%
Correlation 1	−50%
Correlation 2	−50%
Long Term Vol 1	10%
Long Term Vol 2	15%

Table 2: Parameters used for Monte-Carlo simulation

For each discretization scheme, we have computed call prices for various Monte Carlo simulations number (ranging from 100000 to 2000000), strikes (70%, 100% and 130%) and maturities (1 year and 10 year). In each table, we provide the price obtained by Monte Carlo simulation, the corresponding standard deviation, the corresponding relative error, the price obtained by Control Variate, the corresponding standard deviation and the corresponding relative error.
The error is computed with respect to the theoretical price given by formula 19. Table 3 lists the numerical prices for our different test cases.

Strikes	Term	Call Price
100%	1	19.4569
100%	10	41.4006
70%	1	27.6092
70%	10	45.2866
130%	1	13.9299
130%	10	38.2780

Table 3: Theoretical call prices

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
100000	100%	1	20.80	0.15	6.90%	19.24	0.04	1.10%
200000	100%	1	20.85	0.10	7.14%	19.28	0.03	0.92%
500000	100%	1	20.80	0.07	6.91%	19.25	0.02	1.07%
1000000	100%	1	20.84	0.05	7.11%	19.25	0.01	1.05%
2000000	100%	1	20.86	0.03	7.20%	19.25	0.01	1.05%
100000	100%	10	42.77	0.66	3.31%	41.01	0.05	0.93%
200000	100%	10	43.59	0.49	5.29%	41.03	0.04	0.89%
500000	100%	10	43.70	0.36	5.54%	40.98	0.02	1.01%
1000000	100%	10	43.69	0.25	5.52%	41.00	0.02	0.97%
2000000	100%	10	43.82	0.17	5.85%	41.01	0.01	0.95%
100000	70%	1	29.25	0.16	5.95%	27.49	0.03	0.42%
200000	70%	1	29.29	0.12	6.08%	27.51	0.02	0.34%
500000	70%	1	29.24	0.07	5.92%	27.49	0.01	0.43%
1000000	70%	1	29.27	0.05	6.03%	27.48	0.01	0.46%
2000000	70%	1	29.30	0.04	6.11%	27.48	0.01	0.45%
100000	70%	10	46.79	0.67	3.32%	45.01	0.04	0.61%
200000	70%	10	47.62	0.50	5.15%	45.03	0.03	0.56%
500000	70%	10	47.72	0.37	5.38%	44.99	0.02	0.66%
1000000	70%	10	47.71	0.25	5.36%	45.00	0.01	0.63%
2000000	70%	10	47.84	0.18	5.65%	45.01	0.01	0.62%
100000	130%	1	15.00	0.13	7.66%	13.65	0.05	2.00%
200000	130%	1	15.05	0.09	8.01%	13.69	0.04	1.73%
500000	130%	1	15.01	0.06	7.72%	13.66	0.02	1.92%
1000000	130%	1	15.05	0.05	8.05%	13.68	0.02	1.81%
2000000	130%	1	15.07	0.03	8.16%	13.68	0.01	1.82%
100000	130%	10	39.53	0.65	3.27%	37.79	0.07	1.27%
200000	130%	10	40.35	0.49	5.41%	37.82	0.05	1.20%
500000	130%	10	40.45	0.36	5.69%	37.76	0.03	1.35%
1000000	130%	10	40.44	0.25	5.66%	37.78	0.02	1.31%
2000000	130%	10	40.58	0.17	6.01%	37.79	0.02	1.28%

Table 4: Call prices for Euler discretization scheme

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
100000	100%	1	20.06	0.14	3.08%	19.35	0.04	0.53%
200000	100%	1	20.10	0.10	3.32%	19.39	0.03	0.36%
500000	100%	1	20.06	0.07	3.08%	19.36	0.02	0.50%
1000000	100%	1	20.10	0.05	3.28%	19.36	0.01	0.48%
2000000	100%	1	20.11	0.03	3.38%	19.36	0.01	0.48%
100000	100%	10	41.59	0.65	0.46%	41.24	0.05	0.39%
200000	100%	10	42.41	0.49	2.44%	41.27	0.04	0.31%
500000	100%	10	42.51	0.35	2.68%	41.22	0.02	0.44%
1000000	100%	10	42.51	0.24	2.67%	41.23	0.02	0.40%
2000000	100%	10	42.64	0.17	3.00%	41.24	0.01	0.38%
100000	70%	1	28.36	0.16	2.72%	27.56	0.03	0.17%
200000	70%	1	28.40	0.11	2.85%	27.58	0.02	0.09%
500000	70%	1	28.35	0.07	2.68%	27.56	0.01	0.18%
1000000	70%	1	28.38	0.05	2.80%	27.55	0.01	0.21%
2000000	70%	1	28.40	0.04	2.88%	27.55	0.01	0.20%
100000	70%	10	45.53	0.65	0.54%	45.18	0.04	0.24%
200000	70%	10	46.37	0.50	2.38%	45.21	0.03	0.16%
500000	70%	10	46.47	0.36	2.60%	45.16	0.02	0.27%
1000000	70%	10	46.46	0.25	2.59%	45.17	0.01	0.25%
2000000	70%	10	46.59	0.17	2.88%	45.18	0.01	0.24%
100000	130%	1	14.39	0.13	3.31%	13.79	0.05	1.04%
200000	130%	1	14.44	0.09	3.66%	13.82	0.04	0.77%
500000	130%	1	14.40	0.06	3.36%	13.80	0.02	0.94%
1000000	130%	1	14.44	0.04	3.70%	13.81	0.02	0.83%
2000000	130%	1	14.46	0.03	3.81%	13.81	0.01	0.83%
100000	130%	10	38.41	0.64	0.34%	38.06	0.07	0.56%
200000	130%	10	39.23	0.49	2.50%	38.11	0.05	0.45%
500000	130%	10	39.33	0.35	2.76%	38.05	0.03	0.59%
1000000	130%	10	39.33	0.24	2.74%	38.07	0.02	0.56%
2000000	130%	10	39.46	0.17	3.09%	38.07	0.02	0.53%

Table 5: Call prices for Zhu discretization scheme

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
100000	100%	1	20.02	0.14	2.88%	19.34	0.04	0.59%
200000	100%	1	20.06	0.10	3.12%	19.37	0.03	0.42%
500000	100%	1	20.02	0.06	2.88%	19.35	0.02	0.56%
1000000	100%	1	20.06	0.05	3.08%	19.35	0.01	0.54%
2000000	100%	1	20.07	0.03	3.18%	19.35	0.01	0.54%
100000	100%	10	41.22	0.64	0.44%	41.23	0.05	0.40%
200000	100%	10	42.03	0.49	1.53%	41.27	0.04	0.33%
500000	100%	10	42.13	0.35	1.77%	41.22	0.02	0.44%
1000000	100%	10	42.13	0.24	1.76%	41.23	0.02	0.40%
2000000	100%	10	42.26	0.17	2.08%	41.24	0.01	0.38%
100000	70%	1	28.32	0.16	2.58%	27.55	0.03	0.20%
200000	70%	1	28.36	0.11	2.70%	27.57	0.02	0.13%
500000	70%	1	28.31	0.07	2.53%	27.55	0.01	0.21%
1000000	70%	1	28.34	0.05	2.65%	27.54	0.01	0.24%
2000000	70%	1	28.36	0.04	2.73%	27.55	0.01	0.23%
100000	70%	10	45.16	0.65	0.29%	45.17	0.04	0.25%
200000	70%	10	45.98	0.50	1.53%	45.20	0.03	0.18%
500000	70%	10	46.08	0.35	1.76%	45.16	0.02	0.28%
1000000	70%	10	46.08	0.24	1.74%	45.17	0.01	0.26%
2000000	70%	10	46.20	0.17	2.02%	45.18	0.01	0.25%
100000	130%	1	14.35	0.13	3.05%	13.77	0.05	1.12%
200000	130%	1	14.40	0.09	3.40%	13.81	0.04	0.86%
500000	130%	1	14.36	0.06	3.11%	13.79	0.02	1.03%
1000000	130%	1	14.41	0.04	3.44%	13.80	0.02	0.91%
2000000	130%	1	14.42	0.03	3.55%	13.80	0.01	0.92%
100000	130%	10	38.05	0.63	0.60%	38.06	0.07	0.56%
200000	130%	10	38.86	0.49	1.53%	38.10	0.05	0.45%
500000	130%	10	38.97	0.35	1.80%	38.06	0.03	0.58%
1000000	130%	10	38.96	0.24	1.78%	38.07	0.02	0.54%
2000000	130%	10	39.09	0.17	2.12%	38.08	0.01	0.52%

Table 6: Call prices for Exact Zhu discretization scheme

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
100000	100%	1	19.37	0.14	0.44%	19.44	0.04	0.08%
200000	100%	1	19.42	0.10	0.21%	19.48	0.03	0.09%
500000	100%	1	19.37	0.06	0.44%	19.45	0.02	0.03%
1000000	100%	1	19.41	0.05	0.24%	19.45	0.01	0.01%
2000000	100%	1	19.43	0.03	0.15%	19.46	0.01	0.01%
100000	100%	10	40.23	0.63	2.83%	41.38	0.05	0.05%
200000	100%	10	41.03	0.48	0.90%	41.40	0.04	0.00%
500000	100%	10	41.13	0.35	0.65%	41.38	0.02	0.06%
1000000	100%	10	41.13	0.24	0.66%	41.39	0.02	0.02%
2000000	100%	10	41.26	0.17	0.35%	41.40	0.01	0.01%
100000	70%	1	27.54	0.16	0.26%	27.62	0.03	0.04%
200000	70%	1	27.57	0.11	0.13%	27.64	0.02	0.11%
500000	70%	1	27.53	0.07	0.30%	27.62	0.01	0.03%
1000000	70%	1	27.56	0.05	0.19%	27.61	0.01	0.00%
2000000	70%	1	27.58	0.04	0.10%	27.61	0.01	0.01%
100000	70%	10	44.11	0.64	2.59%	45.28	0.04	0.02%
200000	70%	10	44.92	0.49	0.80%	45.30	0.03	0.02%
500000	70%	10	45.03	0.35	0.58%	45.27	0.02	0.03%
1000000	70%	10	45.02	0.24	0.59%	45.28	0.01	0.00%
2000000	70%	10	45.15	0.17	0.31%	45.29	0.01	0.01%
100000	130%	1	13.83	0.13	0.69%	13.89	0.05	0.25%
200000	130%	1	13.88	0.09	0.35%	13.93	0.04	0.01%
500000	130%	1	13.84	0.06	0.63%	13.91	0.02	0.14%
1000000	130%	1	13.89	0.04	0.30%	13.93	0.02	0.03%
2000000	130%	1	13.90	0.03	0.19%	13.93	0.01	0.02%
100000	130%	10	37.10	0.62	3.07%	38.24	0.07	0.09%
200000	130%	10	37.91	0.48	0.97%	38.27	0.05	0.01%
500000	130%	10	38.01	0.34	0.70%	38.25	0.03	0.06%
1000000	130%	10	38.00	0.24	0.71%	38.27	0.02	0.03%
2000000	130%	10	38.13	0.17	0.38%	38.27	0.01	0.01%

Table 7: Call prices for QE-M discretization scheme

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
100000	100%	1	19.19	0.14	1.39%	19.46	0.04	0.01%
200000	100%	1	19.20	0.10	1.30%	19.43	0.03	0.12%
500000	100%	1	19.27	0.06	0.95%	19.43	0.02	0.12%
1000000	100%	1	19.32	0.04	0.69%	19.43	0.01	0.11%
2000000	100%	1	19.39	0.04	0.34%	19.43	0.01	0.16%
100000	100%	10	41.01	0.70	0.94%	41.47	0.05	0.17%
200000	100%	10	41.31	0.54	0.22%	41.44	0.04	0.10%
500000	100%	10	41.37	0.35	0.07%	41.41	0.02	0.03%
1000000	100%	10	41.44	0.24	0.09%	41.42	0.02	0.04%
2000000	100%	10	41.35	0.17	0.13%	41.41	0.01	0.03%
100000	70%	1	27.30	0.16	1.11%	27.61	0.03	0.02%
200000	70%	1	27.33	0.11	1.01%	27.59	0.02	0.06%
500000	70%	1	27.41	0.07	0.71%	27.60	0.01	0.05%
1000000	70%	1	27.47	0.05	0.52%	27.59	0.01	0.06%
2000000	70%	1	27.54	0.05	0.25%	27.58	0.01	0.11%
100000	70%	10	44.87	0.70	0.92%	45.33	0.04	0.10%
200000	70%	10	45.19	0.54	0.22%	45.32	0.03	0.07%
500000	70%	10	45.25	0.35	0.07%	45.29	0.02	0.02%
1000000	70%	10	45.32	0.25	0.08%	45.30	0.01	0.03%
2000000	70%	10	45.23	0.17	0.13%	45.30	0.01	0.02%
100000	130%	1	13.70	0.13	1.69%	13.93	0.05	0.01%
200000	130%	1	13.71	0.09	1.59%	13.91	0.04	0.17%
500000	130%	1	13.76	0.06	1.25%	13.90	0.02	0.25%
1000000	130%	1	13.80	0.04	0.90%	13.90	0.02	0.21%
2000000	130%	1	13.87	0.04	0.43%	13.90	0.01	0.21%
100000	130%	10	37.91	0.69	0.96%	38.36	0.07	0.23%
200000	130%	10	38.19	0.53	0.23%	38.32	0.05	0.11%
500000	130%	10	38.25	0.35	0.07%	38.29	0.03	0.03%
1000000	130%	10	38.31	0.24	0.10%	38.29	0.02	0.04%
2000000	130%	10	38.22	0.16	0.14%	38.29	0.01	0.03%

Table 8: Call prices for Second-Order Alfonsi discretization scheme

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
100000	100%	1	19.19	0.14	1.35%	19.40	0.04	0.29%
200000	100%	1	19.31	0.10	0.78%	19.44	0.03	0.08%
500000	100%	1	19.29	0.06	0.88%	19.45	0.02	0.04%
1000000	100%	1	19.36	0.05	0.52%	19.40	0.02	0.28%
2000000	100%	1	19.39	0.03	0.36%	19.45	0.01	0.04%
100000	100%	10	40.61	0.69	1.90%	41.42	0.05	0.05%
200000	100%	10	41.03	0.53	0.89%	41.41	0.04	0.03%
500000	100%	10	41.01	0.32	0.94%	41.37	0.02	0.07%
1000000	100%	10	41.27	0.23	0.32%	41.38	0.02	0.05%
2000000	100%	10	41.39	0.17	0.03%	41.40	0.01	0.01%
100000	70%	1	27.33	0.16	1.03%	27.56	0.03	0.17%
200000	70%	1	27.44	0.11	0.62%	27.59	0.02	0.05%
500000	70%	1	27.42	0.07	0.69%	27.60	0.01	0.02%
1000000	70%	1	27.50	0.06	0.40%	27.55	0.03	0.20%
2000000	70%	1	27.53	0.04	0.27%	27.61	0.01	0.01%
100000	70%	10	44.48	0.69	1.77%	45.30	0.04	0.03%
200000	70%	10	44.91	0.53	0.83%	45.30	0.03	0.03%
500000	70%	10	44.90	0.32	0.84%	45.27	0.02	0.04%
1000000	70%	10	45.16	0.23	0.28%	45.28	0.01	0.02%
2000000	70%	10	45.28	0.17	0.02%	45.28	0.01	0.00%
100000	130%	1	13.70	0.13	1.67%	13.87	0.05	0.40%
200000	130%	1	13.80	0.09	0.96%	13.91	0.04	0.11%
500000	130%	1	13.78	0.06	1.05%	13.92	0.02	0.04%
1000000	130%	1	13.84	0.05	0.62%	13.88	0.02	0.33%
2000000	130%	1	13.86	0.03	0.47%	13.92	0.01	0.09%
100000	130%	10	37.51	0.68	2.01%	38.31	0.07	0.08%
200000	130%	10	37.92	0.52	0.93%	38.30	0.05	0.06%
500000	130%	10	37.89	0.31	1.03%	38.24	0.03	0.09%
1000000	130%	10	38.14	0.22	0.36%	38.25	0.02	0.06%
2000000	130%	10	38.26	0.17	0.04%	38.27	0.01	0.01%

Table 9: Call prices for Third-Order Alfonsi discretization scheme