A comparison of biased simulation schemes for stochastic volatility models

Roger Lord , Remmert Koekkoek & Dick Van Dijk

A comparison of biased simulation schemes for stochastic volatility models

ROGER LORD†, REMMERT KOEKKOEK‡ and DICK VAN DIJK*§

†Financial Engineering, Cardano, St Clements House,
5th Floor, 27–28 Clements Lane, London EC4N 7AE, UK
‡Robeco Alternative Investments, Coolsingel 120, 3011 AG
Rotterdam, The Netherlands
§Erasmus University Rotterdam, Econometric Institute, P.O. Box 1738,
3000 DR Rotterdam, The Netherlands

Using an Euler discretization to simulate a mean-reverting CEV process gives rise to the problem that while the process itself is guaranteed to be nonnegative, the discretization is not. Although an exact and efficient simulation algorithm exists for this process, at present this is not the case for the CEV-SV stochastic volatility model, with the Heston model as a special case, where the variance is modelled as a mean-reverting CEV process. Consequently, when using an Euler discretization, one must carefully think about how to fix negative variances. Our contribution is threefold. Firstly, we unify all Euler fixes into a single general framework. Secondly, we introduce the new full truncation scheme, tailored to minimize the positive bias found when pricing European options. Thirdly and finally, we numerically compare all Euler fixes to recent quasi-second order schemes of Kahl and Jäckel, and Ninomiya and Victoir, as well as to the exact scheme of Broadie and Kaya. The choice of fix is found to be extremely important. The full truncation scheme outperforms all considered biased schemes in terms of bias and root-mean-squared error.

Keywords: Stochastic volatility; Heston; Square root process; CEV process; Euler–Maruyama; Discretization; Strong convergence; Weak convergence; Boundary behaviour

1. Introduction

Within the area of mathematical finance, most models used for the pricing of derivatives start from a set of stochastic differential equations (SDEs) that describe the evolution of certain financial variables, such as the stock price, interest rate or volatility of an asset. For the valuation of exotic derivatives Monte Carlo simulation is often the method of choice, due to its ability to handle both early exercise and path dependent features with relative ease. In such cases it is important to know exactly how to simulate the evolution of the variables of interest. Obviously, if the SDEs can be solved such that the relevant variables can be expressed as a function of a finite set of state variables for which we know the joint distribution, the problem is reduced to sampling from this distribution. This is for example the case with the Black–Scholes model.

Unfortunately not all models allow for such simple representations. For these models the conceptually straightforward Euler–Maruyama (Euler for short) discretization can be used, see e.g. Kloeden and Platen (1999), Jäckel (2002) or Glasserman (2003). The Euler scheme discretizes the time interval of interest, such that the financial variables are simulated on this discrete time grid. Under certain conditions it can be proven that the Euler scheme converges to the true process as the time discretization is made finer and finer. Nevertheless, the disadvantages of such a discretization are clear. Firstly, the magnitude of the bias is unknown for a certain time discretization, so that one will have to rerun the same simulation with a finer discretization to check whether the result is sufficiently accurate. Secondly, the time grid required for a given accuracy may be much finer than is

*Corresponding author. Email: djvandijk@few.eur.nl

strictly necessary for the derivative under consideration—many trades only depend on the realization of the processes at a small number of dates. Clearly, if exact and efficient simulation methods can be devised for a model, they should be preferred.

In this paper we consider simulation schemes based on Euler discretization for the class of models generally referred to as CEV-SV models, see e.g. Andersen and Brotherton-Ratcliffe (2005) and Andersen and Piterbarg (2007). The asset price process (S) and the variance process (V) evolve according to the following SDEs, specified under the risk-neutral probability measure:

$$\begin{aligned}
\mathrm{d}S(t) &= \mu S(t)\mathrm{d}t + \lambda\sqrt{V(t)}S(t)^{\beta}\,\mathrm{d}W_S(t), \\
\mathrm{d}V(t) &= -\kappa(V(t)-\theta)\mathrm{d}t + \omega V(t)^{\alpha}\,\mathrm{d}W_V(t).
\end{aligned} \tag{1}$$

Here μ is the risk neutral drift of the asset price, $\kappa \geq 0$ is the speed of mean-reversion of the variance, $\theta \geq 0$ is the long-term average variance, and $\omega \geq 0$ is the so-called volatility of variance or volatility of volatility. Finally, λ is a scaling constant and W_S and W_V are correlated Brownian motions, with instantaneous correlation coefficient ρ.

To simplify the exposition, we will mainly concentrate on the special case $\alpha = \frac{1}{2}$ and $\beta = 1$, leading to the popular Heston (1993) model. The best performing simulation schemes will however also be tested in a more general example. The Heston model was heavily inspired by the interest rate model of Cox *et al.* (1985), who used the same mean-reverting square root process to model the spot interest rate. It is well known that, given an initial nonnegative value, a square root process cannot become negative, see e.g. Feller (1951), giving the process some intuitive appeal for the modelling of interest rates or variances. The Heston model is often used as an extension of the Black–Scholes model to incorporate stochastic volatility, and is often used for product classes such as equity and foreign exchange, although extensions to an interest rate context also exist, see e.g. Andersen and Andreasen (2002) and Andersen and Brotherton-Ratcliffe (2005).

Although pricing in the Cox–Ingersoll–Ross (CIR) and Heston models is a well-documented topic, most textbooks seem to avoid the issue of how to simulate these models. If we focus purely on the mean-reverting square-root component of (1), there is not a real problem, as Cox *et al.* (1985) found that the conditional distribution of $V(t)$ given $V(s)$ is noncentral chi-squared. Both Glasserman (2003) and Broadie and Kaya (2006) provide a detailed description of how to simulate from such a process. Combining this algorithm with recent advances on the simulation of gamma random variables by Marsaglia and Tsang (2000) will lead to a fast and efficient simulation of the mean-reverting square root process.

Complications arise, however, when we superimpose a correlated asset price, as in (1). As there is no straightforward way to simulate a noncentral chi-squared increment together with a correlated normal increment for the asset price process, the next idea that springs to mind is an Euler discretization. This involves two problems, the first of which is of a practical nature. Despite the domain of the square root process being the nonnegative real line, for any choice of the time grid the probability of the variance becoming negative at the next time step is strictly greater than zero. As we will see, this is much more of an issue in a stochastic volatility context than in the CIR interest rate model, due to the much higher values typically found for the volatility of variance ω. Practitioners have therefore often opted for a quick 'fix' by either setting the process equal to zero whenever it attains a negative value, or by reflecting it in the origin, and continuing from there on. These fixes are often referred to as *absorption* or *reflection*, see e.g. Gatheral (2006). Interestingly this problem also arises in a discrete time setting, a lead we follow up on in the final section.

The second problem is of both a theoretical and practical nature. The usual theorems leading to strong or weak convergence in Kloeden and Platen (1999) require the drift and diffusion coefficients to satisfy a linear growth condition, as well as being globally Lipschitz. Since the square root is not globally Lipschitz, convergence of the Euler scheme is not guaranteed. Although the global Lipschitz condition on the diffusion coefficient can be relaxed to a local one, see Gyöngy (1998), the square root is not locally Lipschitz around zero. For this reason, various alternative methods have been used to prove convergence of particular discretizations for the square root process. We mention Deelstra and Delbaen (1998), Diop (2003), Bossy and Diop (2004), Alfonsi (2005), and Berkaoui *et al.* (2008), who deal with the square root process in isolation.

It is only recently that papers dealing with the simulation of the Heston model in its full glory have started appearing. Andersen and Brotherton-Ratcliffe (2005) were among the first to suggest an approximation scheme for (1) which preserves the positivity of both S and V for general values of α and β. In Broadie and Kaya (2004, 2006) an exact simulation algorithm has been devised for the Heston model. In numerical comparisons of their algorithm to an Euler discretization with the absorption fix, they find that for the pricing of European options in the Heston model and variations thereof, the exact algorithm compares favourably in terms of root-mean-squared (RMS) error. Their algorithm is however highly time-consuming, as we will see, and therefore certainly not recommendable for the pricing of strongly path dependent options that require the value of the asset price on a large number of time instants. Higham and Mao (2005) considered an Euler discretization of the Heston model with a novel fix, for which they prove strong convergence. To the best of our knowledge they are the first to rigorously prove that using an Euler discretization in the Heston model is theoretically correct, by proving that the sample averages of certain options converge to the true values. Unfortunately they do not provide numerical results on the convergence of their fix compared to other Euler fixes. The recent paper of Kahl and Jäckel (2006) considers a number of discretization

methods for a wide range of stochastic volatility models. For the Heston model they find that their IJK-IMM scheme, a quasi-second order scheme tailored specifically toward stochastic volatility models, gives the best results. Their numerical results are however not comparable to those of Broadie and Kaya, as they use a strong convergence measure which cannot directly be related to an RMS error. Finally we should mention the simulation schemes recently constructed by Andersen (2008). As this paper compares to our full truncation scheme and as it postdates an initial version of our paper, we chose not to include these schemes in our comparison. The schemes, specifically tailored for the Heston model, seem to produce a smaller bias than any scheme considered in this paper, at the cost of a more complex implementation.

The contribution of this article is threefold. Firstly, we unify all Euler discretizations corresponding to the different fixes for the problem of negative variance known thus far under a single framework. Secondly, we propose a new fix, called the full truncation scheme. Full truncation is a modification of the Euler scheme of Deelstra and Delbaen (1998), which we will refer to as the partial truncation method. The difference between both methods lies in the treatment of the drift. Whereas partial truncation only truncates terms involving the variance in the diffusion of the variance, full truncation also truncates within the drift. In both schemes however the variance process itself remains negative. Both schemes are extended to (1). Following the train of thought of Higham and Mao, we are able to prove strong convergence for both of these fixes. With this proof in hand the pricing of plain vanilla options and certain exotics via Monte Carlo is justified, as we can then appeal to the results of Higham and Mao. Thirdly and finally, we numerically compare all Euler fixes to the other schemes mentioned above in terms of the size of the bias, as well as RMS error given a certain computational budget.

The article is structured as follows. Section 2 deals with the CEV-SV model and its properties. Section 3 considers simulation schemes for the Heston model. In section 4 we consider Euler schemes for the CEV-SV model and introduce the full truncation scheme, for which we prove strong convergence. Section 5 provides numerical results, whereas section 6 concludes.

2. The CEV-SV model and its properties

For reasons of clarity, we repeat equation (1) here, which specifies the dynamics of the asset price and variance process in the CEV-SV model under the risk neutral probability measure:

$$\begin{aligned}
\mathrm{d}S(t) &= \mu S(t)\mathrm{d}t + \lambda \sqrt{V(t)}S(t)^{\beta}\,\mathrm{d}W_S(t),\\
\mathrm{d}V(t) &= -\kappa(V(t)-\theta)\mathrm{d}t + \omega V(t)^{\alpha}\,\mathrm{d}W_V(t).
\end{aligned} \tag{2}$$

We restrict β to lie in $(0,1]$ and α to be positive. This model is analysed in great detail in Andersen and Piterbarg (2007). Before turning to the issue of the

simulation of (2) in general and the Heston model in particular, we briefly mention some well-known properties of the process $V(t)$ and $S(t)$ that we require in the remainder of this paper. The mean-reverting CEV process $V(t)$ has the following properties:

  (i)   0 is always an attainable boundary for $0 < \alpha < \frac{1}{2}$;
  (ii) 0 is an attainable boundary when $\alpha = \frac{1}{2}$ and $\omega^2 > 2\kappa\theta$. The boundary is strongly reflecting;
  (iii) 0 is unattainable for $\alpha > \frac{1}{2}$;
  (iv) ∞ is an unattainable boundary.

Via the Yamada condition it can be verified that the SDE for $V(t)$ has a unique strong solution when $\alpha \geq \frac{1}{2}$. For $\alpha < \frac{1}{2}$ we impose that the process for $V(t)$ is reflected in the origin. All properties follow from the classical Feller boundary classification criteria (see e.g. Karlin and Taylor 1981). Turning to the condition $\omega^2 > 2\kappa\theta$, we mention that to calibrate the Heston model to the skew observed in equity or FX markets, one often requires large values for the volatility of variance ω, see e.g. the calibration results in Duffie *et al.* (2000) where $\omega \approx 60\%$. In the CIR model ω, then representing the volatility of interest rates, is markedly lower, see e.g. the calibration results in Brigo and Mercurio (2001, p. 115) where this parameter is around 5%. Moreover, the product $\kappa\theta$ is usually of the same magnitude in both models if we use a deterministic shift extension to fit the initial term structure in the CIR model, so that it is safe to say that for typical parameter values the origin will be attainable within the Heston model, whereas in the CIR interest rate model it will not. Concerning (ii) we mention that strongly reflecting here means that the time spent in the origin is zero—$V(t)$ can touch zero, but will leave it immediately. The interested reader is referred to Revuz and Yor (1991) for more details.

Turning to the asset price process in the CEV-SV model, Andersen and Piterbarg (2007) prove that the process S can reach 0 with a positive probability. To ensure that the SDE in (2) has a unique solution, they impose the natural boundary condition that:

  (v) $S(t)$ has an absorbing barrier at 0.

We do the same here, and mention that (v) seems to be consistent with the asymptotic expansion derived for the SABR model in Hagan *et al.* (2002). The SABR model is a special case of an CEV-SV model with $\theta = 0$, $\kappa = -\omega^2/4$ and $\alpha = 1$.

The following section specifically considers the simulation of the Heston model as this model is of great practical importance.

3. Simulation schemes for the Heston model

We now turn to the simulation of (2) when $\alpha = \frac{1}{2}$ and $\beta = 1$, i.e. the Heston model. Obviously there are myriads of schemes one could use to simulate the Heston model. Though we by no means aim to be complete, we outline some schemes here that yield promising results or are frequently cited. We postpone the treatment of Euler

schemes to the next section. Firstly, we demonstrate why in the case of the Heston model it is not wise to change coordinates to the volatility, i.e. the square root of V. Secondly, we briefly discuss the exact simulation method of Broadie and Kaya (2006). Finally, we take a look at alternative discretizations, in particular the quasi-second order schemes of Ninomiya and Victoir (2004) and Kahl and Jäckel (2006).

Apart from the schemes considered in this section, lately a number of papers have appeared in which splitting schemes are considered for mean-reverting CEV processes, see e.g. Moro (2004) and Dornic *et al.* (2005) and Moro and Schurz (2007). The schemes in these papers heavily rely on an exact solution being known for a subsystem of the original SDE. Whilst this is certainly the case for univariate mean-reverting CEV processes, it does not seem likely that such a splitting can be found for the full-blown CEV-SV model. For this reason we do not further consider these schemes here, though the topic does warrant further study.

3.1. Changing coordinates

For reasons of increased speed of convergence it is often preferable to transform an SDE in such a way that it obtains a constant volatility term, see e.g. Jäckel (2002, section 4.2.3). If we do this for the process $V(t)$ in (2) with $\alpha = \frac{1}{2}$, we can achieve this by considering volatility itself:

$$d\sqrt{V(t)} = \left(\frac{\kappa\theta - \frac{1}{2}\omega^2}{2\sqrt{V(t)}} - \frac{1}{2}\kappa\sqrt{V(t)} \right) dt + \frac{1}{2}\omega \, dW_V(t). \quad (3)$$

Although this transformation is seemingly correct, we are only allowed to apply Itō's lemma if the square root is twice differentiable on the domain of $V(t)$. However, since the origin is attainable for $\omega^2 > 2\kappa\theta$, and the square root is not differentiable in zero, the process obtained by incorrectly applying Itō's lemma is structurally different, as is also mentioned in Jäckel (2004). Even when the origin is inaccessible, the numerical behaviour of the transformed equation is rather unstable. Unless $\omega^2 = 2\kappa\theta$, when $V(t)$ is sufficiently small, the drift term in (3) will blow up, temporarily assigning a much too high volatility to the stock price, in turn greatly distorting the sample average of the Monte Carlo simulation. Luckily, anyone trying to implement (3) will pick up this feature rather quickly, as will be illustrated in the numerical results in section 4. We mention that similar issues arise with other coordinate transformations, such as switching to the logarithm of $V(t)$.

3.2. Exact simulation of the Heston model

As mentioned, Broadie and Kaya (2004, 2006) have recently derived a method to simulate without bias from the Heston stochastic volatility model in (2). Although we refer to their papers for the exact details, we outline their algorithm here to motivate why it is highly time-consuming. First of all a large part of their algorithm

relies on the result that for $s \leq t$, $V(t)$ conditional upon $V(s)$ is, up to a constant scaling factor, noncentral chi-squared:

$$V(t) \sim \frac{\omega^2(1 - e^{-\kappa(t-s)})}{4\kappa} \chi_\nu^2 \left(\frac{4\kappa e^{-\kappa(t-s)} V(s)}{\omega^2(1 - e^{-\kappa(t-s)})} \right), \quad (4)$$

where $\chi_\nu^2(\xi)$ is a noncentral chi-squared random variable with ν degrees of freedom and non-centrality parameter ξ. The degrees of freedom are equal to $\nu = 4\kappa\theta\omega^{-2}$. Glasserman (2003) as well as Broadie and Kaya show how to simulate from a noncentral chi-squared distribution. Combining this with recent advances by Marsaglia and Tsang (2000) on the simulation of gamma random variables (the chi-squared distribution is a special case of the gamma distribution), leads to a fast and efficient simulation of $V(t)$ conditional upon $V(s)$.

Secondly, let us define $V(s, t) = \int_s^t V(u) du$ and $V_a(s, t) = \int_s^t \sqrt{V(u)} dW_a(u)$ for $a = S, V$. First of all Broadie and Kaya recognized that integrating the equation for the variance yields:

$$V(t) = V(s) - \kappa V(s, t) + \kappa\theta(t - s) + \omega V_V(s, t), \quad (5)$$

so that we can calculate $V_V(s, t)$ if we know $V(s)$, $V(t)$ and $V(s, t)$. Knowing all these terms, and solving for $\ln S(t)$ conditional upon $\ln S(s)$ yields the final step:

$$\ln S(t) \sim N \left(\ln S(s) + \mu(t - s) - \frac{1}{2}V(s, t) \right.$$
$$\left. + \rho V_V(s, t), (1 - \rho^2)V(s, t) \right), \quad (6)$$

where N indicates the normal distribution. The algorithm can thus be summarized by:

Algorithm 1: Exact simulation of the Heston model by Broadie and Kaya

1. Simulate $V(t)$, conditional upon $V(s)$ from (4)
2. Simulate $V(s, t)$ conditional upon $V(t)$ and $V(s)$
3. Calculate $V_V(s, t)$ from (5)
4. Simulate $S(t)$ given $V(s, t)$, $V_V(s, t)$ and $S(s)$, by means of (6)

The crucial and time-consuming step is the one we skipped over for a reason—step 2. Broadie and Kaya show how to derive the characteristic function of $V(s, t)$ conditional upon $V(t)$ and $V(s)$. This step utilizes the transform method, so that one has to numerically invert the cumulative distribution function, itself found by the numerical Fourier inversion of the characteristic function. Since the characteristic function non-trivially depends on the two realizations $V(s)$ and $V(t)$ via e.g. modified Bessel functions of the first kind, it is not trivial to cache a major part of the calculations. Hence we must repeat this step at each path and date that is relevant for the derivative at hand. It suffices to say that this makes step 2 very time-consuming and unsuitable for highly path-dependent exotics.

3.3. Quasi-second order schemes

In Glasserman (2003, pp. 356–358), a quasi-second order† Taylor scheme is considered. Its convergence is found to be rather erratic, which is one of the reasons why Broadie and Kaya (2006) chose not to compare their exact scheme to second order Taylor schemes. A closer look at Glasserman's scheme shows the probable cause of this erratic convergence—the discretization contains terms which are very similar to the drift term in (3), and can therefore become quite large when $V(t)$ is small. Since then, two papers have applied second order schemes to either the mean-reverting square root process or the Heston model in its full-fledged form, namely Alfonsi (2005) and Kahl and Jäckel (2006). We start with the latter. After comparing a variety of schemes, Kahl and Jäckel conclude that at least for the Heston model applying the implicit Milstein method‡ (IMM) to the variance, combined with their bespoke IJK scheme for the logarithm of the stock price, yields the best results as measured by a strong convergence measure. Their results indicate that their scheme by far outperforms the Euler schemes with the absorption fix. The IMM method discretizes the variance as follows:

$$V(t + \Delta t) = V(t) - \kappa \Delta t \big(V(t + \Delta t) - \bar{V} \big) + \omega \sqrt{V(t)}$$
$$\cdot \Delta W_V(t) + \frac{1}{4} \omega^2 \cdot \big(\Delta W_V(t)^2 - \Delta t \big). \quad (7)$$

The IMM method actually preserves positivity for the mean-reverting square root process, provided $\omega^2 < 4\kappa\theta$, see Kahl (2004). Unfortunately, this condition is not frequently satisfied in an implied calibration of the Heston model. For values outside this range, a fix is again required. The best scheme for the logarithm of the stock price is their IJK scheme:

$$\ln S(t + \Delta t) = \ln S(t) + \mu \Delta t - \frac{1}{4} \Delta t (V(t) + V(t + \Delta t))$$
$$+ \rho \sqrt{V(t)} \cdot \Delta W_V(t) + \frac{1}{2} \Big(\sqrt{V(t)} + \sqrt{V(t + \Delta t)} \Big)$$
$$\cdot (\Delta W_S(t) - \rho \Delta W_V(t)) + \frac{1}{4} \omega \rho \big(\Delta W_V(t)^2 - \Delta t \big) \quad (8)$$

which is specifically tailored to stochastic volatility models, where typically ρ is highly negative. For more details on both discretizations, we refer the interested reader to Kahl (2004) and Kahl and Jäckel (2006). In the remainder we will refer to (7)–(8) as the IJK-IMM scheme.

Alfonsi (2005) deals with the mean-reverting square root process in isolation, and develops an implicit scheme that also preserves positivity by considering the transformed equation (3). The range of parameters for which the scheme works is again $\omega^2 < 4\kappa\theta$. He also considers Taylor expansions of this implicit scheme, the best of which (his $E(0)$ scheme) is equivalent to (7) to first order in Δt. We therefore purely focus on Kahl and Jäckel's scheme in our numerical results. As an interesting side-note, the $E(0)$ scheme coincides exactly with a special case of the variance equation in the Heston and Nandi (2000, Appendix B) model, which they show converges to the mean-reverting square-root process as the time step tends to zero.

Finally, we consider a second order scheme proposed in Ninomiya and Victoir (2004) for SDEs whose drift and diffusion coefficients are smooth functions with bounded derivatives of any order. Though the scheme converges weakly with order 2, it does not seem applicable to the Heston model—the first derivative of the square root function is already not bounded. The example the authors consider however is based in the Heston model, and does, for their choice of parameters, seem to have a second order convergence. Nevertheless, as the technical conditions on the drift and diffusion coefficients are not satisfied, we will refer to the scheme as a quasi-second order scheme.

Let us first describe their scheme for a fully general SDE in Stratonovich form:

$$d\mathbf{Y}(t) = g_0(\mathbf{Y}(t))dt + \sum_{i=1}^{d} \mathbf{g}_i(\mathbf{Y}(t)) \circ dW_i(t), \quad (9)$$

where $\mathbf{Y} \in \mathbb{R}^n$ and $\mathbf{g}_i \colon \mathbb{R}^n \to \mathbb{R}^n$ for $i = 1, \ldots, d$ are smooth functions whose derivatives of any order are bounded. Starting from $\mathbf{y}(t)$, a discretization of $\mathbf{Y}(t)$, the value at the next time step is:

$$\mathbf{y}(t + \Delta t) = \mathbf{y}_{d+1}\Big(\frac{1}{2} \Delta t \Big), \quad (10)$$

which is found by solving the following $d + 2$ ordinary differential equations (ODEs):

$$\frac{d\mathbf{y}_i}{dt} = \begin{cases} \mathbf{g}_i & \text{if } \Lambda(t) = -1, \\ \mathbf{g}_{d+1-i} & \text{if } \Lambda(t) = 1, \end{cases}$$
$$\text{subject to} \quad \mathbf{y}_i(0) = \mathbf{y}_{i-1}\Big(Z_{i-1}(t)\sqrt{\Delta t} \Big) \quad (11)$$

for $i = 0, \ldots, d+1$. With the exception of $Z_0 = \frac{1}{2}\sqrt{\Delta t}$, all $Z_i(t)$'s for $i = 1, \ldots, d$ are i.i.d. standard normal random variables. Further, $\Lambda(t)$ is an independent Bernoulli random variable of parameter $1/2$, and the initial condition of the last ODE is $\mathbf{y}_0(0) = \mathbf{y}(t)$. Finally, $\mathbf{g}_{d+1} = \mathbf{g}_0$. If available, closed-form solutions to the ODE should be preferred, otherwise one can turn to approximations.

Ninomiya and Victoir's example dealt with the Heston model for $\rho = 0$ and considered the system $\mathbf{Y}(t) = (S(t), V(t))^T$. We consider their scheme for $\mathbf{Y}(t) = (X(t), V(t))^T$,

---

†By quasi-second order we mean schemes that do not simulate the double Wiener integral.
‡Though they consider the balanced Milstein method (BMM), for the square root process their control functions (see their figure 6) coincide with the implicit Milstein method. From now on we will therefore refer to their scheme as the IJK-IMM scheme.

where $X(t)$ is $\ln S(t)$, for general values of ρ. The Stratonovich SDE for this system is:

$$dX(t) = (\mu - \tfrac{1}{2}V(t) - \tfrac{1}{4}\omega\rho)dt + \sqrt{V(t)} \circ dW_1(t),$$

$$dV(t) = \left(-\kappa(V(t) - \theta) - \tfrac{1}{4}\omega^2\right)dt + \omega\rho\sqrt{V(t)} \circ dW_1(t)$$

$$+ \omega\sqrt{(1-\rho^2)V(t)} \circ dW_2(t). \quad (12)$$

Before stating the NV scheme, we first need to deal with one problematic ODE.

Lemma 3.1: *The solution to the ODE $v'(t) = \alpha\sqrt{v(t)}$, with $v(0) \geq 0$ a known constant, is:*

$$v(t) = f(t, \alpha, v(0)) = \max\left(\tfrac{1}{2}\alpha t + \sqrt{v(0)}, 0\right)^2 \quad (13)$$

if we make the choice that $v(t)$ immediately leaves the origin when $v(0) = 0$ and α, $t \geq 0$.

Proof: Let us assume that $t \geq 0$ as by symmetry the solution for $t < 0$ is the same as that for $v(-t)$ from the above ODE with $-\alpha$. The general solution is:

$$v(t) = \left(\tfrac{1}{2}\alpha t + \tfrac{1}{2}C\right)^2 \quad (14)$$

with C an arbitrary constant. In order to satisfy the initial condition, C has to equal $\pm 2\sqrt{v(0)}$. It is clear that $v(t)$ must be monotonically decreasing when $\alpha < 0$, and increasing when $\alpha > 0$. As $v'(0) = \tfrac{1}{2}\alpha C$, C must be positive and thus $C = 2\sqrt{v(0)}$. The solution for $\alpha < 0$ needs to be adapted slightly. The time at which v reaches zero follows as the solution to $v(t^*) = 0$ in (14):

$$t^* = -\frac{2\sqrt{v(0)}}{\alpha}. \quad (15)$$

Hereafter, $v(t)$ must be absorbed in zero, as $v(t)$ must remain nonnegative and its derivative cannot be positive. The only problematic case is when $\alpha > 0$ and $v(0) = 0$. As the square root is not Lipschitz in 0, it follows that the solution to the ODE with $v(0) = 0$ is not guaranteed to be unique. Indeed, both $v(t) = 0$ and $v(t) = \tfrac{1}{4}\alpha^2 t^2$ are valid solutions, and can be combined to create an infinite number of solutions. As the origin is strongly reflecting for the square root process, we choose the latter to remain as close to the SDE as possible. This leads to (13). $\square$

We remark that the ODE in Lemma 3.1 is incorrectly solved in Ninomiya and Victoir's paper. We expect this to be less important in their example, as ω is there 10%. With the aid of Lemma 3.1, the solutions to the ODEs in (11) now follow as:

$$x_0(t) = x_0(0) + (\mu - \tfrac{1}{4}\omega\rho)t - \tfrac{1}{2}v_0(0, t),$$

$$v_0(t) = e^{-\kappa t}v_0(0) + (1 - e^{-\kappa t})\left(\theta - \frac{\omega^2}{4\kappa}\right),$$

$$x_1(t) = x_1(0) + \frac{v_1(t) - v_1(0)}{\omega\rho}, \quad v_1(t) = f(t, \omega\rho, v_1(0)),$$

$$x_2(t) = x_2(0), \quad v_2(t) = f\left(t, \omega\sqrt{1-\rho^2}, v_2(0)\right), \quad (16)$$

where f is the solution in (13), and:

$$v_0(0, t) = \int_0^t v_0(u)du = \frac{1}{\kappa}\left(1 - e^{-\kappa t}\right)\left(\frac{\omega^2}{4\kappa} - \theta + v_0(0)\right)$$

$$+ \left(\theta - \frac{\omega^2}{4\kappa}\right)t. \quad (17)$$

We trust the reader can grasp how the scheme works. As in the schemes of Kahl and Jäckel and Alfonsi, the condition $\omega^2 < 4\kappa\theta$ ensures the variance remains positive, as otherwise $v_0(t)$ becomes negative for

$$t > -\frac{1}{\kappa}\ln\frac{4\kappa\theta - \omega^2}{4\kappa\theta - \omega^2 - 4\kappa v} \equiv t^*(v).$$

When $\omega^2 > 4\kappa\theta$ we fix this by using $v_0(\tau)$ instead of $v_0(t)$, and $v_0(0, \tau)$ in $x_0(t)$ instead of $v_0(0, t)$, where $\tau = \min(t^*(v_0(0)), t)$.

As a final remark, it should be clear that not absorbing v in zero is the right choice. If we would absorb, consider the situation where $\omega^2 < 4\kappa\theta$ and $v(0) = 0$. Then $v(t) = 0$, and

$$\lim_{\Delta t \to 0} S(T) = S(0)\exp\left((\mu - \tfrac{1}{4}\omega\rho)T\right), \quad (18)$$

which clearly is undesirable. As we will see the forward asset price is still far from the correct one, even if we impose that $v(t)$ leaves zero immediately. For this reason we omit numerical results for those configurations where $\omega^2 < 4\kappa\theta$ is violated.

4. Euler schemes for the CEV-SV model

Given that the exact simulation method of Broadie and Kaya can be rather time-consuming, as well as the fact that no exact scheme is likely to be devised for the non-affine CEV-SV model, a simple Euler discretization is certainly not without merit. Even if in future a more efficient exact simulation method for the Heston model would be developed, Euler and higher-order discretizations will remain useful for strongly path-dependent options and stochastic volatility extensions of the LIBOR market model, see e.g. Andersen and Andreasen (2002) and Andersen and Brotherton-Ratcliffe (2005), as it is unlikely that the complicated drift terms in such models will allow for exact simulation methods to be devised.

In section 4.1 we firstly unify all presently known Euler discretizations for the CEV-SV model into one framework. Section 4.2 compares all schemes and makes a case for a new scheme—the full truncation scheme. In section 4.3 we prove strong convergence of this scheme. Finally, section 4.4 takes a look at the Euler scheme of Andersen and Brotherton-Ratcliffe (2005), which preserves positivity of the variance process in an alternative way.

Table 1. Overview of Euler schemes known in the literature.

Scheme	Paper	$f_1(x)$	$f_2(x)$	$f_3(x)$
Absorption	Unknown	x^+	x^+	x^+
Reflection	Diop (2003), Bossy and Diop (2004), Berkaoui *et al.* (2008)	$\lvert x \rvert$	$\lvert x \rvert$	$\lvert x \rvert$
Higham and Mao	Higham and Mao (2005)	x	x	$\lvert x \rvert$
Partial truncation	Deelstra and Delbaen (1998)	x	x	x^+
Full truncation	This article	x	x^+	x^+

4.1. Euler discretizations—unification

Turning to Euler discretizations, a naïve Euler discretization for V in (1) would read:

$$V(t + \Delta t) = (1 - \kappa \Delta t)V(t) + \kappa \theta \Delta t + \omega V(t)^\alpha \cdot \Delta W_V(t) \tag{19}$$

with $\Delta W_V(t) = W_V(t + \Delta t) - W_V(t)$. When $V(t) > 0$, the probability of $V(t + \Delta t)$ going negative is

$$\mathbb{P}(V(t + \Delta t) < 0) = N\left(\frac{-(1 - \kappa \Delta t)V(t) - \kappa \theta \Delta t}{\omega V(t)^\alpha \sqrt{\Delta t}}\right), \tag{20}$$

where N is the standard normal cumulative distribution function. Although the probability decays as a function of the time step Δt, it will be strictly positive for any choice hereof. Furthermore, since ω typically is much higher in a stochastic volatility setting than in an interest rate setting, the problem will be much more pronounced for the Heston model. Without care, the scheme for V will not be defined, so we will have to decide what to do in case V turns negative. Practitioners have often opted for a quick 'fix' by either setting the process equal to zero whenever it attains a negative value, or by reflecting it in the origin, and continuing from there on. These fixes are often referred to as *absorption* and *reflection* respectively, see e.g. Gatheral (2006). We note that this terminology is somewhat at odds with the terminology used to classify the boundary behaviour of stochastic processes, see Karlin and Taylor (1981). In that respect the absorption fix is much more similar to reflection in the origin for a continuous stochastic process, whereas absorption as a boundary classification means that the process stays in the absorbed state for the rest of time. Deelstra and Delbaen (1998) and Higham and Mao (2005) have considered other approaches for fixing the variance when it becomes negative. These are discussed below.

All of these Euler schemes can be unified in a single general framework:

$$\tilde{V}(t + \Delta t) = f_1(\tilde{V}(t)) - \kappa \Delta t \cdot (f_2(\tilde{V}(t)) - \bar{V}) + \omega \cdot f_3(\tilde{V}(t))^\alpha \cdot \Delta W_V(t), \tag{21}$$

$$V(t + \Delta t) = f_3(\tilde{V}(t + \Delta t)),$$

where $\tilde{V}(0) = V(0)$ and the functions f_i, $i = 1$ through 3 have to satisfy:

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

The second condition is a strict requirement for any scheme: we have to fix the volatility term when the variance becomes negative. The first condition seems quite a natural thing to ask from a simulation scheme: if the volatility is not negative, the 'fixing' functions f_1 through f_3 should collapse to the identity function in order not to distort the results. In the remainder we use the identity function x, the absolute value function $\lvert x \rvert$ and $x^+ = \max(x, 0)$ as fixing functions. Obviously only the last two are suitable choices for f_3. The schemes considered thus far in the literature, as well as our new scheme that is introduced below, are summarized in table 1.

While the mentioned papers, apart from Higham and Mao, have dealt with the mean-reverting CEV process in isolation, we also have the asset price S to simulate. For the asset price we switch to logarithms, as in Andersen and Brotherton-Ratcliffe (2005). This guarantees non-negativity:

$$\ln S(t + \Delta t) = \ln S(t) + \left(\mu - \frac{1}{2}\lambda^2 S(t)^{2(\beta-1)} V(t)\right)\Delta t + \lambda S(t)^{\beta-1}\sqrt{V(t)} \cdot \Delta W_S(t) \tag{22}$$

and automatically ensures that the first moment of the asset is matched exactly. In an implementation of (22) one would use the Cholesky decomposition to arrive at $\Delta W_S(t) = \rho \Delta W_V(t) + \sqrt{1 - \rho^2}\Delta Z(t)$, with $Z(t)$ independent of $W_V(t)$. Note that special care has to be taken when $S(t)$ drops to zero, due to property (v).

4.2. Euler discretizations—a comparison and a new scheme

One thing to keep in mind when fixing negative variances is the behaviour of the true process. At the beginning of this section we mentioned that the origin is strongly reflecting if it is attainable, in the sense that when the variance touches zero, it leaves again immediately. If we think of both the reflection and the absorption fixes in a discretization context, the absorption fix seems to capture this behaviour as closely as possible. To analyse the behaviour of all fixes, it is worthwhile to consider the case where an Euler discretization causes the variance to go negative, say $\tilde{V}(t) = -\delta < 0$, whereas the true process would stay positive and close to zero, $V(t) = \varepsilon \geq 0$. In table 2 we have depicted the new starting point $f_1(\tilde{V}(t))$, the effective variance† $f_3(\tilde{V}(t))$ and the drift for all fixes as well for the true process.

---

†By effective variance we mean the instantaneous variance of the stock price.

Table 2. Analysis of the dynamics when $V(t) = \varepsilon \geq 0$, but the Euler discretization equals $-\delta < 0$.

Scheme	New starting point	Effective variance	Drift
True process	ε	ε	$\kappa(\theta - \varepsilon)$
Absorption	0	0	$\kappa\theta$
Reflection	δ	δ	$\kappa(\theta - \delta)$
Higham and Mao	$-\delta$	δ	$\kappa(\theta + \delta)$
Partial truncation	$-\delta$	0	$\kappa(\theta + \delta)$
Full truncation	$-\delta$	0	$\kappa\theta$

A priori we expect that the effect of a misspecified effective variance will be the largest, as this directly affects the stock price on which the options we are pricing depend. From table 2 it seems that reflection has the closest resemblance to the true scheme. However, if $\delta > \varepsilon$, which often is the case, it can be expected that the misspecified variance will cause a larger positive bias than absorption. It is worthwhile to note that in the context of the Heston model it has been numerically demonstrated by Broadie and Kaya (2006) that the absorption fix induces a positive bias in the price of a plain vanilla European call. The Higham and Mao fix tries to lower the bias in the reflection scheme by letting the auxiliary process $\tilde{V}(t)$ remain negative. This however has an undesirable side-effect when at the same time reflecting the variance in the origin to obtain the effective volatility. If $\tilde{V}(t)$ drops even further, the effective variance $f_3(\tilde{V}(t))$ will be much too high, in turn causing larger than intended moves in the stock price.

Both the schemes by Deelstra and Delbaen and ourselves can be interpreted as corrections to the absorption scheme. As in the Higham and Mao scheme, both schemes aim to achieve this by allowing the auxiliary process to attain negative values. Contrary to the Higham and Mao scheme, the side-effect of leaving the auxiliary variance negative is not present here, as the effective variance is set equal to zero. We dub the scheme by Deelstra and Delbaen the partial truncation scheme, as only terms involving V in the diffusion of V are truncated at zero. Note that Glasserman (2003, equation (3.66)) also uses this scheme for the CIR process. As will be demonstrated in the numerical results, partial truncation still causes a positive bias. With a view to lowering the bias, we introduce a new Euler scheme, called full truncation, where the drift of V is truncated as well. By doing this the auxiliary process remains negative for longer periods of time, effectively lowering the volatility of the stock, which helps in reducing the bias.

Though this argumentation is heuristic and hard to prove rigorously, the first moment of all 'fixed' Euler schemes matches the pattern we described above.

Lemma 4.1: *When $\Delta t < 1/\kappa$ the first moments of $\tilde{V}(t)$ in the various 'fixed' Euler schemes in table 1 satisfy the following ordering*:

$$\text{Reflection} > \text{Absorption} > \text{Higham-Mao}$$
$$= \text{Partial truncation} > \text{Full truncation}$$

Proof: We consider a finite time horizon $[0, T]$, discretized on a uniform grid $t_n = n\Delta t$, $n = 1, \ldots, T/\Delta t$. Let us denote all discretizations as:

$$\tilde{v}_{n+1} = f_1(\tilde{v}_n) - \kappa\Delta t(f_2(\tilde{v}_n) - \theta) + \omega f_3(\tilde{v}_n)^\alpha \Delta W_{Vn} \quad (23)$$

with $\tilde{v}_n$ indicating the value of the discretization at t_n and $\Delta W_{Vn} = W_V(t_{n+1}) - W_V(t_n)$. Let us define the first moment as $x_n = \mathbb{E}[\tilde{v}_n]$, where the expectation is taken at time 0. The first moment of the Higham–Mao scheme can be shown to satisfy the difference equation $x_{n+1} = (1 - \kappa\Delta t)x_n + \kappa\Delta t\theta$, which by noting that $x_0 = v_0$ can be solved as:

$$x_n = (1 - \kappa\Delta t)^n(v_0 - \theta) + \theta. \quad (24)$$

The result holds regardless of the chosen function f_3, and therefore also holds for the partial truncation scheme. This is an accurate approximation of the first moment of the continuous process $V(t)$, as it is a well-known result that $\mathbb{E}[V(t)] = (1 - e^{-\kappa t})(V(0) - \theta) + \theta$. Since we initially have $x_0 = v_0$ for all schemes, the remaining results can be found by noting that:

$$(1 - \kappa\Delta t) \cdot |\tilde{v}_n| \geq (1 - \kappa\Delta t) \cdot \tilde{v}_n^+ \geq (1 - \kappa\Delta t) \cdot \tilde{v}_n$$
$$\geq v_n - \kappa\Delta t \cdot \tilde{v}_n^+ \quad (25)$$

which are the drift terms of, from left to right, the reflection, absorption, Higham–Mao, partial and full truncation schemes. As x_{n+1} is exactly the expectation of these terms, the statement follows by induction, starting with $n = 0$. In the second step $(n = 1)$ the inequality already becomes strict, as in each of the schemes v_1 can become negative.   □

Certainly the first moment is not all that matters, but the above lemma does demonstrate that both the Higham–Mao and truncation fixes adjust respectively the reflection and absorption fixes such that the first moment is lowered. Both the partial truncation and the Higham–Mao scheme already obtain an accurate approximation of the true first moment. By truncating the drift, full truncation pulls the first moment down even further, with a view to adjust any remaining bias of the partial truncation scheme.

4.3. Strong convergence of the full truncation scheme

As it is our final goal to price derivatives in the Heston model, we have to be absolutely sure that the sample averages of the realized payoffs converge to the option prices as the time step used in the discretization tends to zero. For European options weak convergence is typically enough to prove this result for Euler discretizations, see e.g. Kloeden and Platen (1999), although for more complex path-dependent derivatives strong convergence may be required. As mentioned earlier though, the non-Lipschitzian dynamics of the CEV-SV model preclude us from invoking the usual theorems on weak and strong convergence of Euler discretizations. Focusing on mean-reverting CEV processes, many authors have proven convergence of their particular discretization.

Recently, Diop (2003) and Bossy and Diop (2004) have proven that an Euler discretization with the reflection fix converges weakly for a variety of mean-reverting CEV processes. For the special case of the mean-reverting square root process, weak convergence of order 1 in the time step is proven, provided that $\omega^2 < \frac{1}{2}\kappa\theta$. This certainly ensures that the origin is not attainable. As the proof may carry over to the general case, we mention that the order of convergence derived is min $(\kappa\theta\omega^{-2}, 1)$. Diop proves strong convergence in the L^p ($p \geq 2$) sense of order $\frac{1}{2}$ under a very restrictive condition, which is relaxed somewhat in Berkaoui *et al.* (2008). For $p = 2$ the condition becomes:

$$\kappa\theta \geq \frac{1}{2}\omega^2 + \max\left\{\omega\sqrt{14\kappa},\ 6\sqrt{2}\omega^2\right\} \qquad (26)$$

One can easily check that, unfortunately, this condition is hardly ever satisfied for any practical values of the parameters. Both Higham and Mao, and Deelstra and Delbaen prove strong convergence for their discretization, without any restrictions on the parameters. As for the absorption scheme, to the best of our knowledge there is no paper dealing with the convergence properties of the absorption fix, although its use in practice is widespread, see e.g. Broadie and Kaya (2004, 2006) and Gatheral (2006).

For the mean-reverting CEV process in isolation, following Deelstra and Delbaen, and Higham and Mao, we use Yamada's (1978) method to find the order of strong convergence. In the proof we restrict α to lie in the interval $[\frac{1}{2}, 1]$. This seems to be the case for most practical applications so that the restriction is not that severe. The big picture of our proof is identical to that of Higham and Mao, but the truncated drift complicates the proofs considerably. The full proof is given in the Appendix, here we merely report the main findings.

First let us introduce some notation. The discretization has already been introduced in equation (23) of Lemma 4.1. For the full truncation scheme we have $f_1(x) = x$ and $f_2(x) = f_3(x) = x^+$. To distinguish between the discretization of the variance and the true process, we will denote the discretization with lowercase letters and the true process with uppercase letters. Following Higham and Mao (2005) we also require the continuous-time approximation of (23):

$$\tilde{v}(t) \equiv \tilde{v}_n - \kappa(t - t_n)(\tilde{v}_n^+ - \theta) + \omega\sqrt{\tilde{v}_n^+} \cdot (W_V(t) - W_V(t_n)). \qquad (27)$$

The convergence of the full truncation scheme is proven in the following theorem.

Theorem 4.2 (Strong convergence of $v(t)$ in the L^1 sense)**:** *The full truncation scheme converges strongly in the L^1 sense, i.e. for sufficiently small values of the time step Δt we have*:

$$\lim_{\Delta t \to 0} \sup_{t \in [0,T]} \mathbb{E}\left[\left|V(t) - v(t)\right|\right] = 0. \qquad (28)$$

Proof: See the appendix. $\qquad \square$

Although the above theorem is only proven for the full truncation scheme, it also holds for the partial truncation scheme, albeit with a slightly easier proof. As the proof of strong convergence for the full CEV-SV process and the proof of convergence for plain-vanilla and barrier option prices are quite similar to those provided by Higham and Mao, we omit them here.

4.4. Euler schemes with moment matching

Before comparing all schemes to each other, we finally mention a moment-matching Euler scheme suggested by Andersen and Brotherton-Ratcliffe (2005). In their discretization, the variance V is locally lognormal, where the parameters are determined such that the first two moments of the discretization coincide with the theoretical moments:

$$V(t + \Delta t) = \left(e^{-\kappa\Delta t}V(t) + (1 - e^{-\kappa\Delta t})\theta\right) \cdot e^{-\frac{1}{2}\Gamma(t)^2\Delta t + \Gamma(t)\cdot\Delta W_V(t)},$$

$$\Gamma(t)^2 = \Delta t^{-1} \cdot \ln\left(1 + \frac{\frac{1}{2}\omega^2\kappa^{-1}V(t)^{2\alpha}(1 - e^{-2\kappa\Delta t})}{(e^{-\kappa\Delta t}V(t) + (1 - e^{-\kappa\Delta t})\theta)^2}\right).$$

$$(29)$$

The advantage of this scheme is that no 'fixes' have to be used to prevent the variance from becoming negative. As mentioned earlier, Andersen (2008) constructs more discretizations for the Heston model along the lines of (29), taking the shape of the Heston density function into account. We only compare to (29) and show that it is already much more effective than many of the Euler fixes mentioned in section 4.1.

5. Numerical results

The previous section established the strong convergence of the full truncation scheme. Though it is certainly useful to theoretically establish the convergence of a scheme, at the end of the day we should be interested in what practitioners really care about: the size of the mispricing given a certain computational budget. It is our goal in this section to compare all mentioned schemes to each other. In our comparisons we take into account both the bias and RMS error, as well as the computation time required. To be clear, if α is the true price of a European call, and $\hat{\alpha}$ is its Monte Carlo estimator, the bias of the estimator equals $\mathbb{E}[\hat{\alpha}] - \alpha$, the variance of the estimator is Var$(\hat{\alpha})$, and finally the root-mean-squared error (RMS error or RMSE) is defined as (bias^2+variance)$^{1/2}$. This fills an important gap in the literature as far as the Euler fixes are concerned, as we do not know of a numerical study that compares the various fixes to one another. In the context of the Heston model, Broadie and Kaya only consider the absorption scheme, and estimate its order of weak convergence to be about $\frac{1}{2}$. Alfonsi (2005) compares both reflection and partial truncation to his scheme, but only for the mean-reverting square root process in isolation.

Table 3. Parameter configurations of the examples used.

Example	κ	ω	ρ	θ	$V(0)$	α
SV-I	2	1	−0.3	0.09	0.09	0.5
SV-II	0.5	1	−0.9	0.04	0.04	0.5
SV-III	0.5	1	0	0.04	0.04	0.5
SVJ	3.99	0.27	−0.79	0.014	0.008836	0.5
CEV-SV	1	1.4	0	1	1	0.75

Table 4. Bias when pricing an ATM call in example SV-I Asset price process: $S(0) = 100$, $\mu = r = 0.05$, $\lambda = 1$, $\beta = 1$ Deal specification: European call option, Maturity 5 yrs. True option price: 34.9998.

Steps/yr.	A	R	HM	PT	FT	ABR	IJK-IMM
20	2.114	4.385	2.732	0.424	0.052	**0.004**	−0.223
40	1.602	3.207	1.680	0.197	**0.031**	**−0.001**	**−0.016**
80	1.225	2.388	1.046	0.096	**0.027**	0.015	0.094
160	0.906	1.759	0.615	**0.020**	**−0.008**	**−0.014**	0.098
$O(\Delta t^p)$	0.41	0.44	0.71	1.42	0.82	−0.94	0.10

The parameter configurations we consider for the variance process are given in table 3. We first focus on the Heston (SV) model, and next consider the Bates (SVJ) model. The latter is an extension of the Heston model to include jumps in the asset price. Clearly all results readily carry over to further extensions of the Heston model, such as the models by Duffie *et al.* (2000) and Matytsin (1999), both of which add jumps to the stochastic variance process. The final subsection considers a non-Heston CEV-SV model.

5.1. *Results for the Heston model*

In this subsection we investigate the performance of the various simulation schemes for the Heston model. As Heston (1993) solved the characteristic function of the logarithm of the stock price, European plain vanilla options can be valued efficiently using the Fourier inversion approach of Carr and Madan (1999). For very recent developments with regard to the evaluation of the multi-valued complex logarithm in the Heston model we refer the interested reader to Lord and Kahl (2009). Among other things, this paper proves how to keep the characteristic function in both the Heston model and Broadie and Kaya's exact simulation algorithm continuous for all possible inputs. Finally, for a very efficient Fourier inversion technique which works for virtually all strike prices and maturities we point the reader to Lord and Kahl (2007).

For the Heston model we consider three parameter configurations, which can be found in table 3. In all three examples $\omega^2 \gg 2\kappa\theta$, implying that the origin of the mean-reverting square root process is attainable. An example where the origin is not attainable is deferred to section 5.2. For the quasi-second order scheme of Kahl and Jäckel this means we have to use a fix. We opted for the absorption fix, which they also use in their examples. The probability of a particular discretization yielding a negative value for $V(t)$ is magnified via the large value of ω, cf. equation (20), so that the way in which each discretization treats the boundary condition will be put to the test. The first example stems from Broadie and Kaya (2006), and is the harder of the two examples they consider. Conveniently, using the example of Broadie and Kaya allows us to compare all biased schemes to their exact scheme. The second example stems from Andersen (2008), where it is used to represent the market for long-dated FX options. The lower level of mean-reversion should make the example more challenging than the first. The third example finally is used to price

a double-no-touch option. The correlation of example SV-II is changed to zero here, as this allows us to use reference values from the literature.

As Broadie and Kaya report computation times for both the Euler scheme with absorption and their exact scheme, we scaled our computation times to match their results. Their results were generated on a desktop PC with an AMD Athlon 1.66 GHz processor, 624 Mb RAM, using Microsoft Visual C++ 6.0 in a Windows XP environment. Relative to the Euler schemes from section 4.2, the IJK-IMM scheme, the Andersen and Brotherton-Ratcliffe (ABR) scheme and the Ninomiya and Victoir (NV) scheme take respectively 14%, 16% and 25% longer to value a European option. One final word should be mentioned on the implementation of the biased simulation schemes. Clearly, the efficiency of the simulations could be improved greatly by using the conditional Monte Carlo techniques of Willard (1997). As Broadie and Kaya point out, this only affects the standard error and the computation time, not the size of the bias, which arises mainly due to the integration of the variance process. We therefore chose to keep the implementation as straightforward as possible.

Starting with the first example, table 4 reports the biases of all biased schemes for an at-the-money (ATM) call. To obtain accurate estimates of the bias we used 10 million simulation paths. If a bias is not significantly different from zero at the 95% confidence level, it is marked bold. The first thing to notice is the enormous difference in the magnitude of the bias, demonstrating the need for an appropriate fix. To relate the size of the bias to implied volatilities, we can glance at figure 1. Even with 20 time steps per year the bias of the full truncation scheme is only 7 basis-points (bp) for the ATM call, i.e. the option has an implied volatility of 28.69% instead of 28.62%. This is already accurate enough for practical purposes. In contrast, the bias for the absorption scheme is 3.02%, and 6.28% for the reflection scheme. The ABR scheme seems to yield the best results for the ATM case, though figure 1 demonstrates that considered over all strikes the bias of the full truncation scheme is much lower and more stable.

For the order of weak convergence, it is worthwhile to note that under suitable regularity conditions, see e.g. Theorem 14.5.2. of Kloeden and Platen (1999), the Euler scheme converges weakly with order 1 in the time step. Though the SDE for the mean-reverting square root
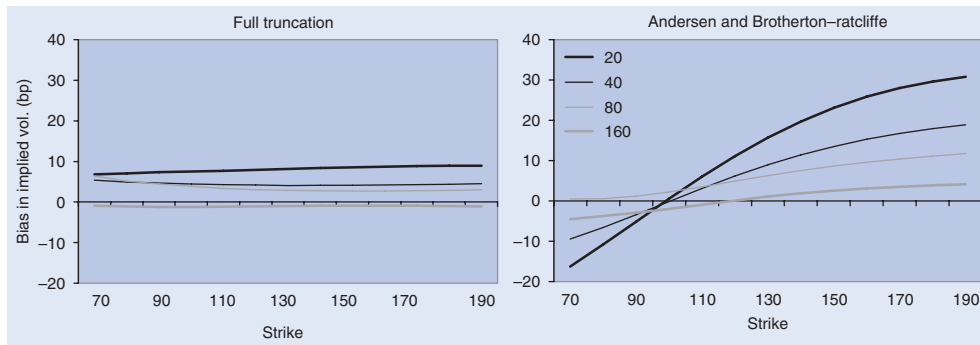
Figure 1. Bias as a function of the strike and the time step in example SV-I.

Table 5. Bias, RMS error and CPU time (in sec.) in the example SV-I for an ATM call.

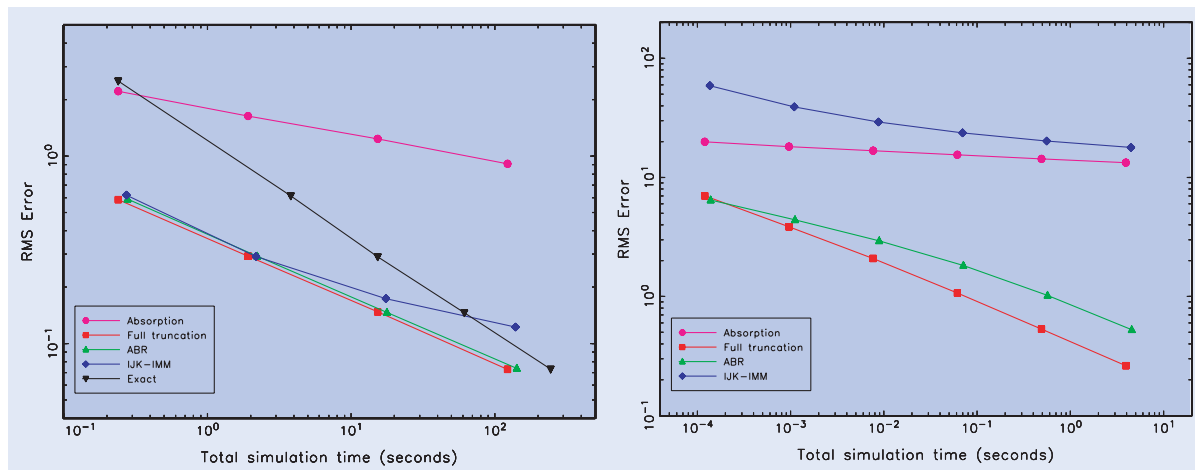| Paths	Steps/yr.	Full truncation			ABR			Exact scheme	
		Bias	RMSE	CPU	Bias	RMSE	CPU	RMSE	CPU
10 000	20	0.052	0.585	0.2	**0.004**	0.590	0.3	0.613	3.8
40 000	40	**0.031**	0.292	1.9	**−0.001**	0.293	2.2	0.290	15.3
160 000	80	**0.027**	0.147	15.4	**0.015**	0.146	17.8	0.146	61.3
640 000	160	**−0.008**	0.073	122.6	**−0.014**	0.074	142.1	0.073	244.5
	$O(\Delta t^p)$	0.95	1.00		−0.94	1.00		1.02	

Figure 2. Convergence of the RMS error in the Heston model for an ATM call Left panel: SV-I example, Right panel: SV-II example.

process does not satisfy these conditions, and it is quite hard to properly estimate the weak order† of convergence with only 10 million paths, both truncation schemes seem to regain this weak order. In contrast, absorption and reflection have a weak order of convergence slightly under $\frac{1}{2}$.

For the quasi-second-order IJK-IMM scheme we note the convergence is somewhat erratic, similar to the aforementioned findings of Glasserman (2003, pp. 356–358). The bias seems to increase when increasing the number of time steps per year from 40 to 80. In contrast,

the absolute value of the bias decreases uniformly for all Euler schemes, neglecting those cases where the bias is statistically indistinguishable from zero.

Finally, let us examine the RMS error and computation time. These are reported in table 5 for full truncation, ABR and the exact scheme. In the left panel of figure 2 the RMSE is plotted as a function of the time step for all schemes. The choice of the number of paths is an important issue here. Duffie and Glynn (1995) have proven that if the weak order of convergence is p, one should increase the number of paths proportional to

---

†The order of weak convergence was estimated here by regressing $\ln(|\text{bias}|)$ on a constant plus $\ln(\Delta t)$.

Table 6. Bias when pricing an ATM call in example SV-II Asset price process: $S(0) = 100$, $\mu = r = 0$, $\lambda = 1$, $\beta = 1$ Deal specification: European call option, Maturity 10 yrs. True option price: 13.0847.

Steps/yr.	A	R	HM	PT	FT	ABR	IJK-IMM
1	18.962	48.472	32.332	12.219	6.371	5.438	57.924
2	17.959	43.321	32.433	8.503	3.710	4.136	38.866
4	16.720	37.842	24.983	5.682	2.041	2.863	29.176
8	15.481	33.161	22.163	3.596	1.055	1.801	23.683
16	14.321	29.200	17.508	2.148	0.525	1.016	20.218
32	13.305	25.987	13.988	1.205	0.259	0.523	17.859
$O(\Delta t^p)$	0.10	0.18	0.25	0.67	0.93	0.68	0.33

Table 7. Bias when pricing a double-no-touch option in example SV-III Asset price process: $S(0) = 100$, $\mu = r = 0$, $\lambda = 1$, $\beta = 1$. Deal specification: 1 yr. double-no-touch option, barriers at 90 and 110. True price: 0.5011.

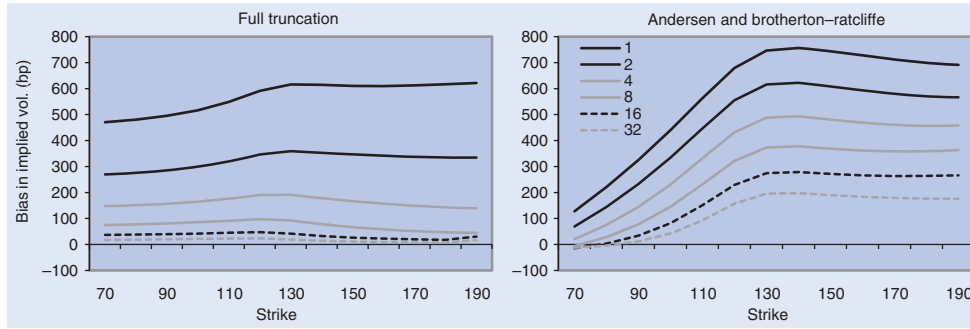| Steps/yr.	A	R	HM	PT	FT	ABR	IJK-IMM
250	−0.190	−0.372	−0.358	0.020	0.022	0.017	−0.235
500	−0.182	−0.346	−0.329	0.016	0.017	0.015	−0.228
1000	−0.174	−0.321	−0.301	0.012	0.013	0.012	−0.218
2000	−0.165	−0.298	−0.275	0.009	0.010	0.009	−0.207

Figure 3. Bias as a function of the strike and the time step in example SV-II.

$(\Delta t)^{-p}$. When $p = 1$, this means that if the time step is halved, we should quadruple the number of paths. Obviously, *a priori* we often do not have an exact value for p, nor do we know the optimal constant of proportionality. We refer the interested reader to the discussion in Broadie and Kaya for the rationale behind the choice of the number of paths in this example. The convergence of the exact scheme is clearly the best. The method produces no bias and hence has $O(N^{-1/2})$ convergence,† N being the number of paths. For a scheme that converges weakly with order p, Duffie and Glynn have proven that for the optimal allocation the RMSE has $O(N^{-p/(2p+1)})$ convergence. Indeed, all biased schemes show a lower rate of convergence than the exact scheme. However, due to the fact that the full truncation scheme already produces virtually no bias with only 20 time steps per year, the RMSEs of both schemes are roughly the same.

For the SV-II example we only report the bias in table 6 as results from the exact scheme are not available to us for this parameter configuration. Again, the truncation

schemes outperform the simple Euler schemes by far. Though the ABR scheme initially has a lower bias, it converges considerably slower than the full truncation scheme. Considered over all strikes in figure 3, the full truncation again generates the least bias, making it the clear winner. Interestingly, the IJK-IMM scheme performs much worse than in the SV-I example—the bias is too large for any practical application. As mentioned in section 3.3 we do not consider the NV scheme for the parameter configurations where $\omega^2 > 2\kappa\theta$, as even the forward is already far from correct. This is particularly evident in this example. If we take e.g. 32 steps per year, the forward price of the asset in the NV scheme equals roughly 179. Considering the fact that the reflection scheme, which at 32 steps per year has the highest bias of the schemes considered, produces a forward price of 101 (the correct answer is 100), it should be clear that the NV scheme is unsuitable when the origin of the square root process is attainable.

So far we have only considered the bias present in European option prices, which reflects the terminal

---

†The discussion here clearly only holds true when using pseudo random numbers, as we do in this paper. In a quasi-Monte Carlo setting the convergence would be $O((\ln N)^2/N)$.

Table 8. Bias when pricing an ATM call in the SVJ example Asset price process: $S(0) = 100$, $\mu = r = 0.0319$, $\lambda = 1$, $\beta = 1$. Jump process: $\xi = 0.11$, $\bar{\mu}_J = -0.12$, $\sigma_J = 0.15$ Deal specification: European call option, Maturity 5 yrs. True option price: 20.1642.

Steps/yr.	A	R	HM	PT	FT	ABR	IJK-IMM	NV	Trans
2	0.836	2.489	5.774	2.790	0.106	−0.146	0.887	2.081	9.043
4	0.400	0.900	0.898	0.399	0.016	−0.096	0.423	0.733	6.844
8	0.179	0.396	0.239	0.083	**−0.013**	−0.070	0.186	0.237	3.725
16	0.083	0.175	0.065	0.019	**−0.005**	−0.037	0.088	0.078	2.518
$O(\Delta t^p)$	1.12	1.27	2.13	2.38	1.36	0.64	1.12	1.58	0.64

distribution of the underlying asset. As a measure of how well these schemes approximate the joint distribution of the asset at various times, we will investigate the bias in double-no-touch prices, which are path-dependent options. A double-no-touch option pays 1 unit of currency if the spot price never hits one of the two barriers. Such options are not uncommon in FX option markets. One reason why we consider them here is that Faulhaber (2002) has shown‡ how to modify Lipton's (2001) eigenfunction expansion approach in order to price double-no-touch options when $\rho = 0$ and the underlying has no drift. This conveniently allows us to generate a reference value with which the simulated values can be compared. Note that both barriers are continuously monitored.

In table 7 the bias of the various schemes is reported. The number of time steps per year coincides with the number of monitoring dates used in the simulation. Though both truncation schemes and the ABR scheme do quite a good job, all other schemes produce a completely wrong price, even for an option with a maturity of 1 year. The need for a scheme which correctly treats the boundary behaviour of the variance process is apparent.

5.2. Results for the Bates model

In the Bates (1996) (SVJ) model, the Heston model is extended with lognormal jumps for the stock price process, where the jumps arrive via a Poisson process:

$$dS(t) = (\mu - \xi\bar{\mu}_J)S(t)dt + \lambda\sqrt{V(t)}S(t)dW_S(t)$$
$$+ J_{N(t)}S(t)dN(t), \tag{30}$$
$$dV(t) = -\kappa(V(t) - \theta)dt + \omega\sqrt{V(t)}\,dW_V(t),$$

where N is a Poisson process with intensity ξ, independent of the Brownian motions. The random variable J_i denotes the ith relative jump size and is lognormally distributed, $\ln J_i \sim N(\mu_J, \sigma_J^2)$. If the ith jump occurs at time t, the stock price right after the jump equals $S(t+) = (1 + J_i)S(t-)$. To ensure no arbitrage, $\bar{\mu}_J$ in (30) has to be the expected relative jump size:

$$1 + \bar{\mu}_J = \mathbb{E}[J_i] = \exp\left(\mu_J + \frac{1}{2}\sigma_J^2\right). \tag{31}$$

The Bates model is often used in an equity or FX context, where the jumps mainly serve to fit the model to the short term skew. Since the jump process is specified independently from the remainder of the model, the same simulation procedure as for the Heston model can be used. If a time step of length T is made until the next relevant date, we draw a random Poisson variable with mean ξT, representing the number of jumps. Subsequently the jump sizes are drawn from the lognormal distribution, and the stock price is adjusted accordingly. In this way the addition of jumps does not add to the discretization error.

The SVJ example stems from Duffie *et al.* (2000), where parameters resulted from a calibration to S&P500 index options. Broadie and Kaya (2006) also use this example, which again allows us to compare the various biased simulation schemes to their exact scheme. We note that the example under consideration satisfies $\omega^2 \ll 2\kappa\theta$, which firstly means that the origin of the square root process is not attainable. Secondly, the low level of ω implies that the probability of any discretization yielding a negative value for V is significantly smaller than in the Heston example. Hence we may expect that the biases are lower than in the previous example.

Thirdly and finally, this combination of parameters is such that the quasi-second order schemes preserve positivity. Contrary to the previous examples this means that the IJK-IMM scheme does not require additional assumptions about the treatment of V at the boundary. Furthermore, the NV scheme should converge.

The bias and RMSE of all schemes, now also including the Euler scheme where we transformed coordinates of the variance as in (3), are reported in table 8 and figure 4 respectively. The number of paths used for the tests in table 8 are 10 000, 40 000, 160 000 and 640 000 respectively. The overall picture is the same as before—the full truncation scheme yields the lowest bias, followed by the ABR scheme and the partial truncation scheme. As the level of bias is so low here, given a fixed computational budget the full truncation scheme by far outperforms the exact scheme. Turning to the transformed scheme, we see its bias is huge compared to the other schemes. Its standard deviation is also much larger, due to the fact that the drift in (3) blows up when V becomes small. Finally, though the quasi-second order schemes automatically preserve positivity for this parameter configuration, they are outperformed in terms of bias

---

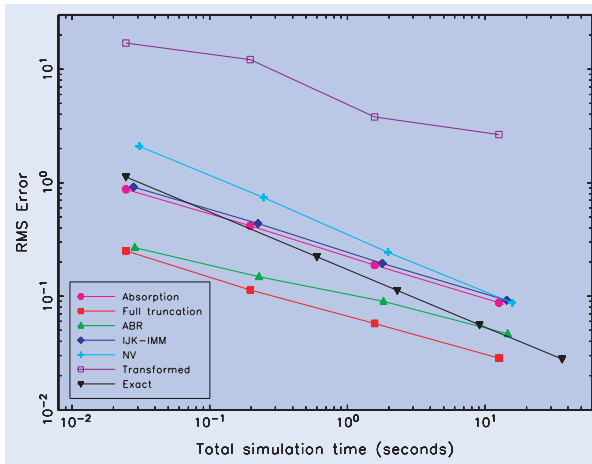‡The author has provided an implementation at http://www.oliverfaulhaber.de.

Figure 4. Convergence of the RMS error in the SVJ example for an ATM call.

Table 9. Bias when pricing an ATM call in the CEV-SV example Asset price process: $S(0) = 100$, $\mu = 0$, $\lambda = 0.04899$, $\beta = 0.5$, discount factor: 2687.74 Deal specification: European call option, Maturity 10 yrs. True option price: 39.22.

Steps/yr.	A	R	HM	PT	FT	ABR
1	5.462	13.007	13.007	5.462	1.278	0.460
2	3.097	6.637	4.887	1.821	0.405	0.273
4	1.381	2.824	1.424	0.513	0.092	0.141
8	0.421	0.844	0.249	0.088	**0.012**	0.073
16	0.062	0.132	**0.010**	−0.002	−0.009	**0.033**
32	**−0.028**	**−0.023**	−0.033	−0.033	−0.033	**−0.011**
$O(\Delta t^p)$	1.62	1.84	2.07	1.94	1.30	1.07

and order of weak convergence by the full truncation scheme.

5.3. *Results for a non-Heston CEV-SV model*

To conclude our extensive numerical analysis, we consider a non-Heston example. The CEV-SV example from table 3 stems from Andersen and Brotherton-Ratcliffe (2005, Appendix A), where their moment-matching Euler scheme is benchmarked to a solution found by solving the corresponding partial differential equation via finite differences. Note that $\alpha = 0.75$, so the origin of the variance process is certainly not attainable.

Table 9 reports the biases of all Euler schemes. Though the schemes in Kahl and Jäckel (2006) and Ninomiya and Victoir (2004) can be used for the more general CEV-SV process, we chose to focus on the Euler schemes as many of them outperformed the quasi-second order schemes in the previous tests. Once again we conclude that all Euler schemes arrive at the correct answer sooner or later, though the truncation and ABR schemes require much less time steps to do so.

6. Conclusions and further research

In this paper we have considered the simulation of the CEV-SV stochastic volatility model and varieties thereof, focusing largely on the Heston model. In the CEV-SV model, the stochastic variance is modelled as a mean-reverting CEV process. When discretizing this process we run into the problem that although the process itself is guaranteed to be nonnegative, any Euler discretization has a nonzero probability of becoming negative in the next time step, regardless of the size of the time step. Hence, we have to 'fix' these negative variances.

Our contribution is threefold. Firstly, we unify all 'fixes' appearing in the literature in a single general framework. Secondly, by analysing the rationale behind the known fixes, we are led to propose a new scheme, the full truncation scheme, designed specifically to minimize

the positive bias one finds when pricing European options using the traditional fixes. Strong convergence is proven for this scheme.

Thirdly and finally, we numerically compare the various Euler schemes to each other, as well as to the quasi-second order schemes by Kahl and Jäckel (2006) and Ninomiya and Victoir (2004), and finally the exact scheme of Broadie and Kaya (2006). All three of these papers compare their schemes to the Euler scheme with an absorption fix and find their scheme to be superior. Our numerical results demonstrate that using the correct fix at the boundary is extremely important, and significantly impacts the magnitude of the bias. In our examples, we find the full truncation scheme produces the smallest bias, closely followed by the moment-matching Euler scheme of Andersen and Brotherton-Ratcliffe (2005) and the partial truncation scheme. The order of weak convergence of the full truncation scheme appears to be close to 1 in the time step, bringing back the order of weak convergence to the theoretical level for an Euler discretization of an SDE with Lipschitzian dynamics. The performance of the quasi-second order schemes is found to be somewhat disappointing. In particular, we demonstrated the NV scheme is only suitable for parameter configurations where $\omega^2 < 2\kappa\theta$, often not the case in practice.

When the volatility of volatility is not too high, the full truncation scheme has relatively small levels of bias and is able to generate a smaller RMS error given a certain computational budget than any other biased or exact scheme considered here. This holds true for both European and path-dependent options. Since an initial version of this paper, Andersen (2008) has specifically designed simulation schemes for the Heston model which mimic its distribution quite closely. These schemes have negligible bias, at the cost of a more complex implementation. On the other hand the full truncation scheme, or indeed that of Andersen and Brotherton-Ratcliffe, is very easy to implement and appears to work fine for a wide variety of processes.

As a final note, we return to the lead mentioned in the introduction, namely that the issues considered here in a continuous time setting can also arise in a discrete time setting. Examples of models where such problems can arise are the model of Heston and Nandi (2000) and the Box–Cox model of Christoffersen and Jacobs (2004). Let us be more specific and look at the first-order version of

the Heston and Nandi model. Here the log-stock price is modelled as:

$$\ln S(t) = \ln S(t - \Delta t) + r + \lambda h(t) + \sqrt{h(t)} z(t),$$

$$h(t + \Delta t) = \tilde{\omega} + \beta h(t) + \alpha \left(z(t) - \gamma \sqrt{h(t)} \right)^2, \quad (32)$$

where $z(t)$ is a standard normal random variable and $h(t)$ is the conditional variance of the log-return between $t - \Delta t$ and t. In this setup $h(t)$ is known at time $t - \Delta t$. Note that all the model parameters will depend on the chosen time step Δt. The process remains stationary with finite first two moments if $\beta + \alpha \gamma^2 < 1$. Without further restrictions on the parameters, $h(t + \Delta t)$ can become negative. In their estimates however ω, β and α are positive and significant at the 95% confidence level, so that there does not seem to be a problem. Turning to their appendix B however, where they prove convergence of (32) to the Heston model with $\rho = -1$ as the time step tends to zero, we see that in their proof they choose† $\tilde{\omega} = (\kappa \theta - \frac{1}{4} \omega^2)(\Delta t)^2$, $\beta = 0$ and $\alpha = \frac{1}{4} \omega^2 (\Delta t)^2$. Positivity of the conditional variance $h(t + \Delta t)$ can thus only be guaranteed provided that $\kappa \theta \geq \frac{1}{4} \omega^2$. This is the same condition under which the schemes of Alfonsi (2005) and Kahl and Jäckel (2006) preserve positivity, and not surprisingly so as we already remarked the equivalence of these three schemes to first order in Δt in section 3.3. Looking in closer detail at their estimation procedure, we see that they only included options with an absolute moneyness less than or equal to 10%, i.e. at or around at-the-money options. In the Heston model $\kappa \theta$ can certainly be smaller than $\frac{1}{4} \omega^2$ when the skew is quite pronounced. This would not be noticed if only options with strikes at or around the at-the-money level would be included in the calibration procedure. Concluding, it may be necessary to introduce restrictions on the parameters in a discrete time setting in order to ensure that the conditional variance process remains positive.

Acknowledgements

References

Alfonsi, A., On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Meth. Appl.*, 2005, **11**, 355–384.

Andersen, L.B.G., Simple and efficient simulation of the Heston stochastic volatility model. *J. Comput. Finan.*, 2008, **11**, 1–42.

Andersen, L.B.G. and Andreasen, J., Volatile volatilities. *Risk*, 2002, **15**, 163–168.

Andersen, L.B.G. and Brotherton-Ratcliffe, R., Extended LIBOR market models with stochastic volatility. *J. Comput. Finan.*, 2005, **9**, 1–40.

Andersen, L.B.G. and Piterbarg, V.V., Moment explosions in stochastic volatility models. *Finan. Stochast.*, 2007, **11**, 29–50.

Bates, D.S., Jumps and stochastic volatility: exchange rate processes implicit in Deutsche Mark options. *Rev. Finan. Stud.*, 1996, **9**, 69–107.

Berkaoui, A., Bossy, M. and Diop, A., Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM Probab. Stat.*, 2008, **12**, 1–11.

Bossy, M. and Diop, A., An efficient discretization scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^\alpha$, $\alpha \in [1/2, 1)$, *INRIA Working Paper No. 5396*, 2004.

Broadie, M. and Kaya, Ö., Exact simulation of option Greeks under stochastic volatility and jump diffusion models, in *Proceedings of the 2004 Winter Simulation Conference*, edited by R.G. Ingalls, M.D. Rossetti, J.S. Smith, and B.A. Peters, pp. 1607–1615, 2004 (The Society for Computer Simulation).

Broadie, M. and Kaya, Ö., Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operat. Res.*, 2006, **54**, 217–231.

Brigo, D. and Mercurio, F., *Interest Rate Models: Theory and Practice*, 2001 (Springer: Berlin).

Carr, P. and Madan, D.B., Option valuation using the fast Fourier transform. *J. Comput. Finan.*, 1999, **2**, 61–73.

Christoffersen, P. and Jacobs, K., Which GARCH model for option valuation? *Manag. Sci.*, 2004, **50**, 1204–1221.

Cox, J.C., Ingersoll, J.E. and Ross, S.A., A theory of the term structure of interest rates. *Econometrica*, 1985, **53**, 385–407.

Deelstra, G. and Delbaen, F., Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Appl. Stochast. Models Data Anal.*, 1998, **14**, 77–84.

Diop, A., Sur la discrétisation et le comportement à petit bruit d'EDS unidimensionelles dont les coefficients sont à derives singulières, PhD thesis, 2003, INRIA.

Dornic, I., Chaté, H. and Muñoz, M.A., Integration of Langevin equations with multiplicative noise and the viability of field theories for absorbing phase transitions. *Phys. Rev. Lett.*, 2005, **94**, 100601-1–100601-4.

Duffie, D. and Glynn, P., Efficient Monte Carlo simulation of security prices. *Ann. Appl. Probab.*, 1995, **5**, 897–905.

Duffie, D., Pan, J. and K. Singleton, K., Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 2000, **68**, 1343–1376.

Faulhaber, O., Analytic methods for pricing double barrier options in the presence of stochastic volatility, MSc thesis, University of Kaiserslautern, 2002. Available online at: http://www.oliverfaulhaber.de/diplomathesis/HestonBarrierAnalytic.pdf

Feller, W., Two singular diffusion problems. *Ann. Math.*, 1951, **54**, 173–182.

Gatheral, J., *The Volatility Surface: A Practitioner's Guide*, 2006 (Wiley: New York).

Glasserman, P., *Monte Carlo Methods in Financial Engineering*, 2003 (Springer Verlag: New York).

Gyöngy, L., A note on Euler approximations. *Potential Anal.*, 1998, **8**, 205–216.

---

†It seems to us that there are different ways to prove this; the conclusion here will however be the same.

Hagan, P.S., Kumar, D., Lesniewski, A.S. and Woodward, D.E., Managing smile risk. *Wilmott Mag.*, 2002, **July**, 84–108.

Heston, S.L., A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Finan. Stud.*, 1993, **6**, 327–343.

Heston, S.L. and Nandi, S., A closed-form GARCH option valuation model. *Rev. Finan. Stud.*, 2000, **13**, 585–625.

Higham, D.J. and Mao, X., Convergence of the Monte Carlo simulations involving the mean-reverting square root process. *J. Comput. Finan.*, 2005, **8**, 35–62.

Jäckel, P., *Monte Carlo Methods in Finance*, 2002 (Wiley: New York).

Jäckel, P., Stochastic volatility models: past, present and future, in *The Best of Wilmott 1: Incorporating the Quantitative Finance Review*, edited by P. Wilmott, pp. 379–390, 2004 (Wiley: New York).

Kahl, C., Positive numerical integration of stochastic differential equations, Diploma thesis, University of Wuppertal and ABN AMRO, 2004.

Kahl, C. and Jäckel, P., Fast strong approximation Monte-Carlo schemes for stochastic volatility models. *Quant. Finan.*, 2006, **6**, 513–536.

Karlin, S. and Taylor, H., *A Second Course in Stochastic Processes*, 1981 (Academic Press: New York).

Kloeden, P.E. and Platen, E., *Numerical Solution of Stochastic Differential Equations*, 3rd ed., 1999 (Springer Verlag: New York).

Lipton, A., *Mathematical Methods for Foreign Exchange— A Financial Engineer's Approach*, 2001 (World Scientific: Singapore).

Lord, R., and Kahl, C., Complex logarithms in Heston-like stochastic volatility models. *Math. Finan.*, 2009, forthcoming.

Lord, R. and Kahl, C., Optimal Fourier inversion in semi-analytical option pricing. *J. Comput. Finan.*, 2007, **10**, 1–30.

Marsaglia, G. and Tsang, W.W., A simple method for generating gamma variables. *ACM Trans. Math. Software*, 2000, **26**, 363–372.

Matytsin, A., Modelling volatility and volatility derivatives. Paper presented at Columbia Practitioners Conference on the Mathematics of Finance, New York, 25 September, 1999. Available online at: http;//www.math.columbia.edu/~smirnov/matytsin.pdf

Moro, E., Numerical schemes for continuum models of reaction-diffusion systems subject to internal noise. *Phys. Rev. E*, 2004, **70**, 045102(R)-1–045102(R)-4.

Moro, E. and Schurz, H., Boundary preserving semi-analytic numerical algorithms for stochastic differential equations. *SIAM J. Sci. Comput.*, 2007, **29**, 1525–1549.

Ninomiya, S., and Victoir, N., Weak approximation of stochastic differential equations and application to derivative pricing Working Paper, Tokyo Institute of Technology and Oxford University, 2004.

Revuz, D. and Yor, M., *Continuous Martingales and Brownian Motion*, 1991 (Springer Verlag: New York).

Willard, G.A., Calculating prices and sensitivities for path-independent derivative securities in multifactor models. *J. Deriv.*, 1997, **5**, 45–61.

Yamada, T., Sur une construction des solutions d'équations différentielles stochastiques dans le cas non-Lipschitzien, in *Séminaire de Probabilité, Lecture Notes in Mathematics 649* Vol. XII, pp. 114–131, 1978 (Springer: New York).

Appendix A: Proof of strong convergence

In this appendix we prove strong convergence of the full truncation scheme applied to the mean-reverting CEV process with $\frac{1}{2} \leq \alpha \leq 1$. We use the same style of proof as Deelstra and Delbaen (1998), and Higham and Mao (2005). As the proof of convergence for the full CEV-SV process follows along the same lines, we only focus on the strong L^1 convergence for the stochastic variance here. Though Lemmas 4.1 and A.2 also hold when $0 < \alpha < \frac{1}{2}$, the proof used for the main theorem no longer seems applicable. Nevertheless, all practical applications seem to use $\alpha \geq \frac{1}{2}$, so that this is no restriction.

For ease of exposure the discretization over a finite time horizon $[0, T]$ is performed on a uniform grid $t_n = n\Delta t$, $n = 1, \ldots, T/\Delta t$. The discretization of the auxiliary process at t_n is given by:

$$\tilde{v}_{n+1} = \tilde{v}_n - \kappa \Delta t (\tilde{v}_n^+ - \theta) + \omega \tilde{v}_n^{+\alpha} \Delta W_{Vn}, \qquad \text{(A1)}$$

where $\Delta W_{Vn} = W_V(t_{n+1}) - W_V(t_n)$. The effective variance is $v_n = \tilde{v}_n^+$. To distinguish between the discretization of the variance and the true process, we will denote the discretization with small letters and the true process with capital letters. Following Higham and Mao (2005) we will consider the continuous-time approximation of (A1):

$$\tilde{v}(t) \equiv \tilde{v}_n - \kappa(t - t_n)(\tilde{v}_n^+ - \theta) + \omega \tilde{v}_n^{+\alpha} \cdot (W_V(t) - W_V(t_n)) \qquad \text{(A2)}$$

or, in integral notation:

$$\tilde{v}(t) = \tilde{v}(0) - \kappa \int_0^t (\tilde{v}_\tau(u)^+ - \theta)\mathrm{d}u + \omega \int_0^t \tilde{v}_\tau(u)^{+\alpha}\, \mathrm{d}W_V(u), \qquad \text{(A3)}$$

where $\tilde{v}(0) = v_0$, $\tilde{v}_\tau(0) = \tilde{v}(\tau(t))$ and $\tau(t)$ equals t_n if $t_n \leq t \leq t_{n+1}$. Obviously $\tilde{v}_\tau(t)$ coincides with $\tilde{v}(t)$ at the grid points of the discretization.

One of the elements required in proving strong convergence of the full truncation scheme, are bounds on the first and second moments of the effective variance v_n. In the remainder we denote the first and second moments by $x_n \equiv \mathbb{E}[\tilde{v}_n]$ and $y_n \equiv \mathbb{E}[\tilde{v}_n^2]$ respectively. In the main text Lemma 4.1 already supplied the following inequality:

$$x_n = \mathbb{E}[\tilde{v}_n] \leq (1 - \kappa \Delta t)^n (v_0 - \theta) + \theta. \qquad \text{(A4)}$$

As we do not require sharp bounds, we will use the following corollary which follows directly.

Corollary A.1: *For $\Delta t < 2/\kappa$ the first moment of $\tilde{v}_n$ in the full truncation scheme is bounded from above by*:

$$x_n \leq |v_0 - \theta| + \theta \equiv U_x. \qquad \text{(A5)}$$

Proof: Follows immediately from Lemma 4.1. □

Secondly, we will find an upper bound on the second moment of $\tilde{v}_n$.

Lemma A.2: (Bounding the second moment of the full truncation scheme): *For any $n = 0, \ldots, N$ where $N\Delta t = T$, and $\Delta t < 2/\kappa$, the second moment of $\tilde{v}_n$ in the full truncation scheme is bounded by*:

$$y_n \leq \gamma^N v_0^2 + \frac{\gamma^N - 1}{\gamma - 1} \cdot \left(2\kappa\theta\Delta t U_x + (\kappa\theta\Delta t)^2 + \omega^2\Delta t\right) \equiv U_y(\Delta t), \qquad \text{(A6)}$$

where $\gamma \equiv \max\{1, (1 - \kappa\Delta t)^2 + 2\alpha\omega^2\Delta t\}$. *Furthermore,* $\lim_{\Delta t \to 0} U_y(\Delta t) < \infty$.

Proof: Clearly, $y_0 = v_0^2$ so that the assertion is true for $n = 0$. Suppose the lemma now holds true for some n. Using (A1) we can then write:

$$y_{n+1} = (\kappa\theta\Delta t)^2 + \mathbb{E}\big[(\tilde{v}_n - \kappa\Delta t\tilde{v}_n^+)^2\big] + 2\kappa\theta\Delta t$$
$$\cdot \mathbb{E}\big[(\tilde{v}_n - \kappa\Delta t\tilde{v}_n^+)\big] + \omega^2\Delta t\mathbb{E}[\tilde{v}_n^{+2\alpha}]. \tag{A7}$$

To bound this expression, we note that, apart from the first constant, the right-hand side can be written as the expectation of the following function:

$$f(\tilde{v}_n) = \begin{cases} \tilde{v}_n^2 + 2\kappa\theta\Delta t\tilde{v}_n, & \tilde{v}_n \leq 0, \\ \tilde{v}_n^2(1 - \kappa\Delta t)^2 + 2\kappa\theta\Delta t(1 - \kappa\Delta t)\tilde{v}_n \\ \quad + \omega^2\Delta t\tilde{v}_n^{+2\alpha}, & \tilde{v}_n \geq 0. \end{cases} \tag{A8}$$

Since $\tilde{v}_n^{+2\alpha} \leq 1 + 2\alpha\tilde{v}_n^2$ as long as $\alpha \leq 1$, (A8) can be bounded from above by:

$$f(\tilde{v}_n) \leq \gamma\tilde{v}_n^2 + 2\kappa\theta\Delta t\tilde{v}_n + \omega^2\Delta t, \tag{A9}$$

where γ is as defined above. Returning to (A7) we then have:

$$y_{n+1} \leq \gamma y_n + 2\kappa\theta\Delta t \cdot x_n + (\kappa\theta\Delta t)^2 + \omega^2\Delta t. \tag{A10}$$

Repeated use of (A10) and our corollary immediately yields (A6). Finally, it follows that:

$$\lim_{\Delta t \to 0} U_y(\Delta t) = \max\left\{1, e^{2(\alpha\omega^2 - \kappa)T}\right\}v_0^2$$
$$+ \frac{e^{2(\alpha\omega^2 - \kappa)T} - 1}{2(\alpha\omega^2 - \kappa)}\big(2\kappa\theta U_x + \omega^2\big) < \infty, \tag{A11}$$

so that the second moment of the discretization does not blow up in finite time. $\qquad\square$

Before addressing the strong L^1 error we need a bound on the L^2 difference between the two continuous-time approximations $v_\tau(t)$ and $v(t)$. The proof entirely depends on Lemmas 4.1 and A.2.

Lemma A.3 (The L^2 difference between $v_\tau(t)$ and $v(t)$)**:** *For $\Delta t < 2/\kappa$ we have:*

$$\sup_{t \in [0,T]} \mathbb{E}\big[(v(t) - v_\tau(t))^2\big] \leq (\kappa\Delta t)^2 \cdot \big(\theta + U_y(\Delta t)\big)$$
$$+ \omega^2\Delta t \cdot U_y(\Delta t)^\alpha \equiv U_{\text{cont}}(\Delta t). \tag{A12}$$

Proof: First of all note that $\mathbb{E}[(v(t) - v_\tau(t))^2] \leq \mathbb{E}[(\tilde{v}(t) - \tilde{v}_\tau(t))^2]$. For $t \in [t_n, t_{n+1})$ we have:

$$\mathbb{E}\big[(\tilde{v}(t) - \tilde{v}_\tau(t))^2\big] = \kappa^2(t - t_n)^2 \cdot \mathbb{E}\big[(\tilde{v}_n^+ - \theta)^2\big]$$
$$+ \omega^2(t - t_n) \cdot \mathbb{E}\big[\tilde{v}_n^{+2\alpha}\big]. \tag{A13}$$

The first term can be bounded from above by:

$$\mathbb{E}[(\tilde{v}_n^+ - \theta)^2] = \theta^2 - 2\theta\mathbb{E}[\tilde{v}_n^+] + \mathbb{E}[\tilde{v}_n^+\tilde{v}_n^+] \leq \theta^2 + y_n, \tag{A14}$$

so that (A14) becomes:

$$\mathbb{E}\big[(\tilde{v}(t) - \tilde{v}_\tau(t))^2\big] \leq \kappa^2(t - t_n)^2 \cdot (\theta + y_n) + \omega^2(t - t_n) \cdot y_n^\alpha$$
$$\leq \kappa^2(t - t_n)^2 \cdot \big(\theta + U_y(\Delta t)\big)$$
$$+ \omega^2(t - t_n) \cdot U_y(\Delta t)^\alpha. \tag{A15}$$

The supremum on $[0, T]$ is then bounded from above by (A12), which completes the proof. $\qquad\square$

Clearly $U_{\text{cont}}(\Delta t)$ is of $O(\Delta t)$, so that the difference between the discrete-time approximation and its continuous extension vanishes when the time step tends to zero. We are now ready to prove strong convergence in the L^1 sense.

Theorem A.4 (Strong convergence of $v(t)$ in the L^1 sense)**:** *The full truncation scheme converges strongly in the L^1 sense:*

$$\lim_{\Delta t \to 0} \mathbb{E}\sup_{t \in [0,T]}\big[|V(t) - v(t)|\big] = 0. \tag{A16}$$

Proof: First note that $\mathbb{E}[|V(t) - v(t)|] \leq \mathbb{E}[|V(t) - \tilde{v}(t)|]$, so that it is sufficient to show (A16) for the latter expression. We will bound it from above in a function of the time step, so that we can prove that this L^1 norm tends to zero as the time step tends to zero. As in Yamada (1978), this is achieved by bounding $\mathbb{E}[\phi_k(V(t) - \tilde{v}(t))]$ for a series of $C^2(\mathbb{R}, \mathbb{R})$ functions ϕ_k which tend to the absolute function. Here we use the same notation as in Higham and Mao (2005). First of all let $a_k = \exp[-k(k+1)/2]$ for $k \geq 0$, so that $\int_{a_k}^{a_{k-1}} u^{-1}\,du = k$. For each integer $k \geq 1$ there exists a continuous function ψ_k with support in (a_{k-1}, a_k) such that $0 \leq \psi_k(u) \leq 2k^{-1}u^{-1}$ and $\int_{a_{k-1}}^{a_k} \psi_k(u)du = 1$. Defining $\phi_k(x) = \int_0^{|x|}\int_0^y \psi_k(u)du\,dy$, it follows that $\phi_k \in C^2(\mathbb{R}, \mathbb{R})$, $\phi_k(0) = 0$, and:

$$|\phi_k'(x)| \leq 1,$$
$$|\phi_k''(x)| = 2k^{-1}|x|^{-1}1_{[a_k < |x| < a_{k-1}]}, \tag{A17}$$
$$|x| - a_{k-1} \leq \phi_k(x) \leq |x|.$$

Consider $\phi_k(V(t) - \tilde{v}(t))$. Using Itō's lemma and taking expectations yields:

$$\mathbb{E}[\phi_k(V(t) - \tilde{v}(t))] = -\kappa M(t) + \frac{1}{2}\omega^2 I(t), \tag{A18}$$

where we defined:

$$M(t) \equiv \mathbb{E}\left[\int_0^t \phi_k'(V(u) - \tilde{v}(u)) \cdot \big(V(u) - \tilde{v}_\tau(u)^+\big)du\right],$$

$$I(t) \equiv \mathbb{E}\left[\int_0^t \phi_k''(V(u) - \tilde{v}(u)) \cdot \big(V(u)^\alpha - \tilde{v}_\tau(u)^{+\alpha}\big)^2 du\right]. \tag{A19}$$

Note that for $\frac{1}{2} \leq \alpha \leq 1$ we can bound:

$$\big(V(u)^\alpha - \tilde{v}_\tau(u)^{+\alpha}\big)^2$$
$$\leq \big|V(u) - \tilde{v}_\tau(u)\big| \cdot \big(1 + (2\alpha - 1) \cdot \big|V(u) - \tilde{v}_\tau(u)\big|\big) \tag{A20}$$

and furthermore we have $|V(u) - \tilde{v}_\tau(u)| \leq |V(u) - \tilde{v}(u)| + |\tilde{v}(u) - \tilde{v}_\tau(u)|$. Using the property of the

second derivative of ϕ in (A17) it follows that, with $\tilde{\alpha} = 2\alpha - 1$:

$$I(t) \leq \frac{2t}{k}\left(1 + 2\tilde{\alpha}\sqrt{U_{\text{cont}}(\Delta t)} + \tilde{\alpha}a_{k-1}\right)$$

$$+ \frac{2t}{ka_k}\left(\sqrt{U_{\text{cont}}(\Delta t)} + \tilde{\alpha}U_{\text{cont}}(\Delta t)\right) \equiv U_{I,k}(t, \Delta t),$$

$$(A21)$$

where we used $\mathbb{E}[|X|] \leq \sqrt{\mathbb{E}[X^2]}$ for any random variable X and Lemma A.3. Turning to $M(t)$, we use the property of the first derivative of ϕ from (A17) to obtain:

$$M(t) \leq \mathbb{E}\left[\int_0^t \left|V(u) - \tilde{v}_\tau(u)^+\right|du\right] \leq \mathbb{E}\left[\int_0^t \left|V(u) - \tilde{v}_\tau(u)\right|du\right]$$

$$\leq \mathbb{E}\left[\int_0^t \left|V(u) - \tilde{v}(u)\right|du\right] + \mathbb{E}\left[\int_0^t \left|\tilde{v}(u) - \tilde{v}_\tau(u)\right|du\right]$$

$$\leq \mathbb{E}\left[\int_0^t \left|V(u) - \tilde{v}(u)\right|du\right] + t\sqrt{U_{\text{cont}}(\Delta t)}. \quad (A22)$$

Combining the bounds on $I(t)$ and $M(t)$ in (A18) with the third property in (A17) yields:

$$\mathbb{E}[\phi_k(V(t) - \tilde{v}(t))] \leq \kappa\mathbb{E}\left[\int_0^T \left|V(u) - \tilde{v}(u)\right|du\right] + T\sqrt{U_{\text{cont}}(\Delta t)}$$

$$+ \frac{1}{2}\omega^2 U_{I,k}(T, \Delta t), \quad (A23)$$

where we also bounded t from above by T. This gives an upper bound of the same form as in Higham and Mao, and allows us to apply Gronwall's inequality:

$$\sup_{t \in [0,T]} \mathbb{E}\left[\left|V(t) - \tilde{v}(t)\right|\right]$$

$$\leq e^{\kappa T}\left[a_{k-1} + T\sqrt{U_{\text{cont}}(\Delta t)} + \frac{1}{2}\omega^2 U_{I,k}(T, \Delta t)\right] \quad (A24)$$

Since (A24) holds for any value of k and $\lim_{\Delta t \to 0} U_{I,k}(t, \Delta t) = 0$ due to (A11) and (A12), it follows that $\lim_{\Delta t \to 0} \mathbb{E}\sup_{t \in [0,T]}[|V(t) - \tilde{v}(t)|] = 0$ as in Corollary 3.1 of Higham and Mao. This immediately implies (A16). The order of convergence unfortunately does not follow from this proof, as $\lim_{k \to \infty} U_{I,k}(t, \Delta t) = \infty$. $\qquad\square$