TI 2006-046/4
Tinbergen Institute Discussion Paper

A Comparison of Biased Simulation Schemes for Stochastic Volatility Models

Roger Lord[1,4]
Remmert Koekkoek[2]
Dick van Dijk[1]

[1] *Econometric Institute, Erasmus University Rotterdam, and Tinbergen Institute;*
[2] *Robeco Alternative Investments, Rotterdam;*
[4] *Rabobank International, Utrecht.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

A comparison of biased simulation schemes for stochastic volatility models

Roger Lord[1]
Remmert Koekkoek[2]
Dick van Dijk[3]

First version: June 23, 2005
This version: May 17, 2006

ABSTRACT

When using an Euler discretisation to simulate a mean-reverting square root process, one runs into the problem that while the process itself is guaranteed to be nonnegative, the discretisation is not. Although an exact and efficient simulation algorithm exists for this process, at present this is not the case for the Heston stochastic volatility model, where the variance is modelled as a square root process. Consequently, when using an Euler discretisation, 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 minimise the upward bias found when pricing European options. Thirdly and finally, we numerically compare all Euler fixes to a recent quasi-second order scheme of Kahl and Jäckel and the exact scheme of Broadie and Kaya. The choice of fix is found to be extremely important. The full truncation scheme by far outperforms all biased schemes in terms of bias, root-mean-squared error, and hence should be the preferred discretisation method for simulation of the Heston model and extensions thereof.

Keywords: Stochastic volatility, Heston, square root process, Euler-Maruyama, discretisation, strong convergence, weak convergence, boundary behaviour.

AMS Classification: 62P05, 65C05, 68U20.
JEL Classification: C63, G13.

Part of this research was carried out while the second author was writing his Master's thesis with the Trading Risk Management department of the ING Group. We thank Christian Kahl for many useful comments and suggestions. We are also grateful to Michel Vellekoop, seminar participants at Rabobank International and the Finance mini-symposium at the 42nd Dutch Mathematical Congress for comments. Any remaining errors are our own.

[1] Erasmus University Rotterdam, Econometric Institute and Tinbergen Institute, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands (e-mail: lord@few.eur.nl, Tel. +31-(0)10-4088935) and Modelling and Research (UC-R-355), Rabobank International, P.O. Box 17100, 3500 HG Utrecht, The Netherlands (e-mail: roger.lord@rabobank.com, Tel. +31-(0)30-2166566).

[2] Robeco Alternative Investments, Coolsingel 120, 3011 AG Rotterdam, The Netherlands (e-mail: r.koekkoek@robeco.nl, Tel. +31-(0)10-2247253).

[3] Erasmus University Rotterdam, Econometric Institute, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands (e-mail: djvandijk@few.eur.nl, Tel. +31-(0)10-4081263).

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. Since Monte Carlo simulation is often the method of choice for the valuation of exotic derivatives, due to its ability to handle both early exercise and path dependent features with relative ease, 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) discretisation can be used, see e.g. Kloeden and Platen [1999], Jäckel [2002] or Glasserman [2003]. The Euler scheme discretises 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 discretisation is made finer and finer. Nevertheless, the disadvantages of such a discretisation are clear. Firstly, the magnitude of the bias is unknown for a certain time discretisation, so that one will have to rerun the same simulation with a finer discretisation to check whether the result is sufficiently accurate. Secondly, the time grid required for a certain accuracy may be much finer than is strictly necessary for the derivative under consideration – many trades only depend on the realisation 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.

The models we consider in this paper are all based on the Heston [1993] stochastic volatility model. In this model, the stock price process (S) and the variance process (V) evolve according to the following SDEs, specified under the risk-neutral probability measure:

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sqrt{V(t)}\,S(t)dW_S(t) \\ dV(t) &= -\lambda\left(V(t) - \overline{V}\right)dt + \eta\sqrt{V(t)}\,dW_V(t) \end{aligned} \tag{1}$$

where μ is the risk neutral drift of the stock price, which may be different from the risk free rate due to dividends, λ is the speed of mean-reversion of the variance, $\overline{V}$ is the long-term average variance, and η is the so-called volatility of variance. Finally, W_S and W_V are correlated Brownian motions, with instantaneous correlation coefficient ρ. The Heston model was heavily inspired by the interest rate model of Cox, Ingersoll and Ross [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 a first extension of the Black-Scholes model that includes stochastic volatility, and is typically 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].

Though pricing in the Cox-Ingersoll-Ross (CIR) and Heston models is a well-documented subject, most textbooks seem to avoid the topic 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 stock 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 stock price process, the next idea that springs to mind is an Euler discretisation. This involves two problems, the first of which is of a practical nature. Despite the domain of the square root process being $[0,\infty)$, 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 an interest rate setting, due to the much higher values one finds for the volatility of variance η. 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 [2005]. 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 recently Albin et al. [2005] relax the global Lipschitz condition somewhat, their results are only applicable to processes on an open interval, whereas the domain of the square root process is $[0,\infty)$, with 0 being an attainable boundary. For this reason, various alternative methods have been used to prove convergence of particular discretisations for the square root process. We mention Deelstra and Delbaen [1998], Diop [2003], Bossy and Diop [2004], Berkaoui, Bossy and Diop [2005] and Alfonsi [2005], 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. 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 discretisation 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. However, their algorithm is 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 stock price at a large number of time instants. It is for this reason that Higham and Mao [2005] considered an Euler discretisation of (1) with a novel fix, for which they prove strong convergence. In addition, and more importantly, they are to the best of our knowledge the first to rigorously prove that using an Euler discretisation in the Heston model for the pricing of plain vanilla and barrier options is theoretically correct, by proving that the sample averages converge to the true values. Unfortunately they do not provide numerical results on the convergence of their fix compared to other Euler fixes. Finally, we mention the recent paper of Kahl and Jäckel [2005a], who compare a number of discretisation methods for a whole host of stochastic volatility models. For the Heston model they find that their IJK-BMM 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.

The contribution of this article is threefold. Firstly, we unify all Euler discretisations 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], to which we from hereon refer 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 the Heston model. We motivate that full truncation can be advantageous if the partial truncation method is found to generate an upward bias on the option price. 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 schemes of Broadie and Kaya and Kahl and Jäckel in terms of the size of the bias, as well an RMS error given a certain computational budget.

The article is structured as follows. Section 2 deals with the Heston model and its simulation, both exact and biased. In section 3 we prove strong convergence for the full truncation scheme, enabling us to invoke the theorems from Higham and Mao that justify the usage of Monte Carlo for the non-Lipschitzian dynamics in the Heston model. Section 4 provides numerical results on the convergence of several Euler schemes. Finally, we conclude in section 5.

2. The Heston model and its simulation

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

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sqrt{V(t)}\,S(t)dW_S(t) \\ dV(t) &= -\lambda\left(V(t) - \overline{V}\right)dt + \eta\sqrt{V(t)}\,dW_V(t) \end{aligned} \tag{2}$$

The stock price S has μ as its risk neutral drift, which may be different from the risk free rate due to dividends. The variance process V is a mean-reverting square root process, where λ is its speed of mean-reversion, $\overline{V}$ is the long-term average variance, and η is the volatility of variance. The correlated Brownian motions W_S and W_V satisfy $dW_S(t) \cdot dW_V(t) = \rho dt$. Via the Yamada condition it can be verified that the SDE in (2) has a unique strong solution.

Before turning to the issue of simulating the dynamics of the Heston model, we briefly specify some well-known properties of the square root process V(t) that we require in the remainder of this paper. These properties are:

i) 0 is an attainable boundary when $\eta^2 > 2\lambda\overline{V}$;
ii) When $\eta^2 > 2\lambda\overline{V}$, the origin is strongly reflecting;
iii) ∞ is an unattainable boundary.

Using the classical Feller boundary classification criteria (see e.g. Karlin and Taylor [1981]) it is easy to establish properties i) and iii). Turning to the condition $\eta^2 > 2\lambda\overline{V}$, 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 η, see e.g. the calibration results in Duffie, Pan and Singleton [2000] where $\eta \approx 60\%$. In the CIR model η, 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 $\lambda\overline{V}$ 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 accessible within the Heston model, whereas in the CIR interest rate model it will be inaccessible. The second property is demonstrated by Revuz and Yor [1991]. Strongly reflecting here means that the time spent in the origin is zero – the variance can touch zero, but will leave it immediately. The interested reader is referred to Andersen and Piterbarg [2004], where the behaviour of the Heston model and related stochastic volatility models is analysed in detail.

We now turn to the simulation of (2). Firstly, we demonstrate why 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]. Thirdly, we unify all Euler discretisations known thus far, and by analysing the boundary behaviour of the square root process, we make a strong case for a new scheme: the full truncation scheme. Finally, we take a brief look at alternative discretisations, in particular the recently proposed scheme by Kahl and Jäckel [2005a].

2.1. 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), we can achieve this by considering volatility itself:

$$d\sqrt{V(t)} = \left(\frac{\lambda\overline{V} - \frac{1}{4}\eta^2}{2\sqrt{V(t)}} - \frac{1}{2}\lambda\sqrt{V(t)} \right) dt + \frac{1}{2}\eta dW_V(t) \tag{3}$$

Although this transformation is seemingly correct, we are only allowed to apply Itô's lemma if the square root is twice differentiable on the domain of V(t). However, since the origin is attainable for $\eta^2 > 2\lambda\overline{V}$, and the square root is not differentiable in zero, the process obtained by incorrectly applying Itô'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 $\lambda\overline{V} = \frac{1}{4}\eta^2$, 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 illustrated in section 2.4 below. We mention that similar issues arise with other coordinate transformations, such as switching to the logarithm of V(t).

2.2. Exact simulation

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 $s \le t$, V(t) conditional upon V(s) is, up to a constant scaling factor, noncentral chi-squared:

$$V(t) \sim \frac{\eta^2(1-e^{-\lambda(t-s)})}{4\lambda}\chi^2_\nu\left(\frac{4\lambda e^{-\lambda(t-s)}V(s)}{\eta^2(1-e^{-\lambda(t-s)})}\right) \tag{4}$$

where $\chi^2_\nu(\xi)$ is a noncentral chi-squared random variable with ν degrees of freedom and non-centrality parameter ξ. The degrees of freedom are equal to $\nu = 4\lambda\overline{V}\eta^{-2}$. 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:

$$V(t)=V(s)-\lambda V(s,t)+\lambda\overline{V}(t-s)+\eta V_V(s,t) \tag{5}$$

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:

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

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

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)

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

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 utilises 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 realisations V(s) and V(t) via e.g. modified Bessel functions of the first kind, we cannot precompute a major part of the calculations, and thus must repeat this step at each path and date that is relevant for the exotic derivative. It suffices to say that this makes step 2 very time-consuming and unsuitable for highly path-dependent exotics.

2.3. Euler discretisations - unification

Given that the exact simulation method of Broadie and Kaya can be rather time-consuming, a simple Euler discretisation may not be without merit. Even if in future a more efficient exact simulation method for the Heston model would be developed, Euler and higher-order discretisations 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. Turning to Euler discretisations, a naïve Euler discretisation for V to get from time t to t+Δt would read:

$$V(t+\Delta t)=\left(1-\lambda\Delta t\right)V(t)+\lambda\overline{V}\Delta t+\eta\sqrt{V(t)}\cdot\Delta W_V(t) \tag{7}$$

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

$$\mathbb{P}\left(V(t+\Delta t)<0\right)=N\left(\frac{-\left(1-\lambda\Delta t\right)V(t)-\lambda\overline{V}\Delta t}{\eta\sqrt{V(t)\Delta t}}\right) \tag{8}$$

where N is the standard normal cumulative distribution function. Although the probability decays as a function of the timestep Δt, it will be strictly positive for any choice hereof. Furthermore, since η 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. If we do not want the volatility to cross over to the imaginary domain, we will have to decide what to do in case the variance 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 [2005]. 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 in more detail below.

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

$$V(t+\Delta t) = f_1(V(t)) - \lambda \Delta t \cdot \left(f_2(V(t)) - \overline{V}\right) + \eta \sqrt{f_3(V(t))} \cdot \Delta W_V(t) \tag{9}$$

where the functions f_i, i = 1, 2, 3 have to satisfy:

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

The second condition is a strict requirement for any scheme: we have to fix the volatility 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_l 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 |x| 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 summarised in Table 1.

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. [2005]	$\|x\|$	$\|x\|$	$\|x\|$
Higham and Mao	Higham and Mao [2005]	x	x	$\|x\|$
Partial truncation	Deelstra and Delbaen [1998]	x	x	x^+
Full truncation	Lord, Koekkoek and Van Dijk [2006]	x	x^+	x^+

Table 1: Overview of Euler schemes known in the literature

While the mentioned papers, apart from Higham and Mao, have dealt with the square root process in isolation, we also have the stock price S to simulate. We can either discretise the stock price process directly, or its logarithm. Though the first is more direct, switching to logarithms can be advantageous. As the solution to the stock price in (2) can be expressed as (cf. (6)):

$$S(t) = S(s) \cdot \exp\left(\mu - \tfrac{1}{2} V(s,t) + V_S(s,t)\right) \tag{10}$$

it should be clear that discretising the logarithm of S will not produce any discretisation error in the stock price direction, whereas the alternative will. For very small timesteps the difference will however be negligible. Discretising the stock price directly yields:

$$S(t+\Delta t) = S(t)\cdot\left(1+\mu\Delta t+\sqrt{f_5(V(t))}\cdot\Delta W_S(t)\right) \tag{11}$$

whereas applying Itô's lemma to ln S(t) and subsequently discretising this equation, yields:

$$\ln S(t+\Delta t) = \ln S(t) + \left(\mu - \tfrac{1}{2} f_4(V(t))\right)\cdot\Delta t + \sqrt{f_5(V(t))}\cdot\Delta W_S(t) \tag{12}$$

In an implementation of (11) or (12) 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 two additional fixing functions have been introduced. Again, $f_5(x)$ should be nonnegative for all possible values of x. Although one is free to explore alternatives, we deemed it sensible to set $f_4 = f_5 = f_3$ in our schemes, thereby treating both the Itô correction term in (12), as well as the volatility of the stock price in (11) and (12) identically to the diffusion of the variance. We note that (11) to first order yields the correct first moment for the stock price, whereas (12) exactly reproduces the first moment when f_4 and f_5 coincide.

2.4. Euler discretisations – a qualitative 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 discretisation 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 discretisation 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 $f_1(\tilde{V}(t))$, effective variance[4] $f_3(\tilde{V}(t))$ and the drift for all fixes as well for the true process.

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

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

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 causes a positive bias in the price of a plain vanilla European call. 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

[4] By effective variance we mean the instantaneous variance of the stock price.

we are pricing depend. It therefore seems evident that reflection will cause a larger positive bias than absorption. The Higham and Mao fix tries to alleviate this problem by letting the process V(t) remain negative (f_1 is here the identity function). This however has an undesirable side-effect, when at the same time reflecting the variance in the origin to obtain the effective volatility. If at some point in time V(t) becomes negative, and the Wiener increment for the next timestep is also negative, V(t) at the next timestep will drop even further. If this happens the effective volatility will be much too high, in turn causing larger than intended moves in the stock price.

The scheme by Deelstra and Delbaen can be interpreted as a first bias correction applied to the absorption scheme. As in the Higham and Mao scheme, it aims to achieve this by leaving V(t) negative. Contrary to the Higham and Mao scheme, the side-effect of leaving the 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, eq. (3.66)] also uses this scheme for the CIR process. As shown below, partial truncation may still be found to cause a positive bias. We therefore propose a new Euler discretisation scheme, called full truncation, where the drift of V is truncated as well. By truncating the drift the process V(t) remains negative for longer periods of time, taking away more volatility from the stock, which helps to reduce the bias. This intuitive analysis is stated more rigorously in the following lemma.

Lemma 1:
When $\Delta t < 1/\lambda$ the first moments of the various "fixed" Euler schemes in Table 1 satisfy:

Reflection > Absorption > Higham-Mao = Partial truncation > Full truncation

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

$$v_{n+1} = f_1(v_n) - \lambda\Delta t\left(f_2(v_n) - \overline{v}\right) + \eta\sqrt{f_3(v_n)}\,\Delta W_{Vn} \tag{13}$$

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

$$x_n = (1-\lambda\Delta t)^n (v_0 - \overline{v}) + \overline{v} \tag{14}$$

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 $\mathbb{E}[V(t)] = (1-e^{-\lambda t})(V(0)-\overline{V}) + \overline{V}$. Since we initially have $x_0 = v_0$ for all schemes, the remaining results can be found by noting that:

$$(1-\lambda\Delta t)\cdot|v_n| \geq (1-\lambda\Delta t)\cdot v_n^+ \geq (1-\lambda\Delta t)\cdot v_n \geq v_n - \lambda\Delta t\cdot v_n^+ \tag{15}$$

which are the drift terms of, from left to right, the reflection, absorption, Higham-Mao and partial truncation and finally the 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.

As the above analysis has been rather qualitative, we will try and visualise some of the effects. In Figure 1 we have depicted a sample path of the effective volatility and the stock price over the period of one day. The "true" sample path was generated with an Euler discretisation of the Heston model, using a timestep of 1E-05 year. Using the same realisations of the Wiener process, we generated sample paths for the effective volatility of the different fixes, i.e. $\sqrt{f_3(V(t))}$, using a timestep which is ten times as large[5]. We also included a sample path of the absolute value of the volatility, obtained from an Euler discretisation of the transformed SDE in (3).

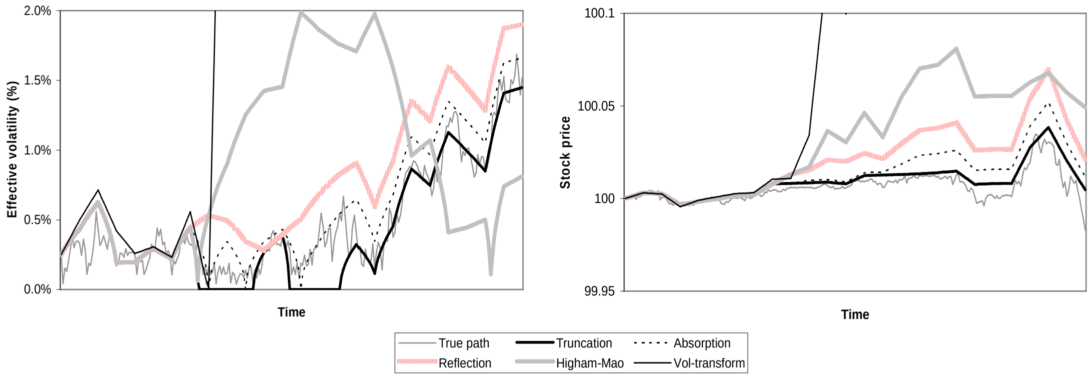

Figure 1: Sample path of the effective volatility, all fixes, as well as the transformed SDE in (3)

The first thing we notice from Figure 1 is that changing coordinates to volatility itself is numerically instable. For small values of V(t) the volatility can explode; since the high effective volatility will persist for a long time, we immediately notice this in the stock price. Secondly, we notice that both the absorption and truncation[6] fixes remain fairly close to the true path, whereas the reflection fix slightly overshoots at the beginning. Finally, the undesirable side-effect of the Higham and Mao fix is evident on this sample path – because V(t) goes negative in the middle of the graph, the effective volatility is too high, causing larger moves in the stock price.

2.5. Alternative discretisations

Obviously there are myriads of schemes other than the Euler scheme one could use for the discretisation of the Heston model. Though we by no means aim to be complete, we briefly consider some schemes here that yield promising results or are frequently cited.

In Glasserman [2003, pp. 356-358], a quasi-second order[7] 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

[5] The graphs for the fixes were created by linearly interpolating between the realisations.
[6] Since the sample paths of partial and full truncation were virtually indistinguishable for this example, we chose to display only one line for both fixes.
[7] By quasi-second order we mean schemes that do not simulate the double Wiener integral.

Glasserman's scheme shows the probable cause of this erratic convergence – the discretisation 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 [2005a]. 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[8] (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 (9) and (12) with the absorption fix. The IMM method discretises the variance as follows:

$$V(t+\Delta t) = V(t) - \lambda \Delta t\left(V(t+\Delta t) - \overline{V}\right) + \eta\sqrt{V(t)} \cdot \Delta W_V(t) + \tfrac{1}{4}\eta^2 \cdot \left(\Delta W_V(t)^2 - \Delta t\right) \quad (16)$$

The IMM method actually preserves positivity for the mean-reverting square root process, provided that $\eta^2 < 4\lambda\overline{V}$, 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:

$$\begin{aligned}\ln S(t+\Delta t) = \ln S(t) + \mu\Delta t - \tfrac{1}{4}\Delta t\left(V(t) + V(t+\Delta t)\right) + \rho\sqrt{V(t)} \cdot \Delta W_V(t) \\ + \tfrac{1}{2}\left(\sqrt{V(t)} + \sqrt{V(t+\Delta t)}\right)\cdot\left(\Delta W_S(t) - \rho\Delta W_V(t)\right) - \tfrac{1}{4}\eta\rho\left(\Delta W_V(t)^2 - \Delta t\right)\end{aligned} \quad (17)$$

which is specifically tailored to stochastic volatility models, where typically ρ is highly negative. For more details on both discretisations, we refer the interested reader to Kahl [2004] and Kahl and Jäckel [2005a]. In the remainder we will refer to (16)-(17) as the IJK-IMM scheme. Note that this scheme too is quasi-second order.

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 $\eta^2 < 4\lambda\overline{V}$. He also considers Taylor expansions of this implicit scheme, the best of which (his E(0) scheme) is equivalent to (17) to first order in Δt. We therefore purely focus on Kahl and Jäckel's scheme in our numerical results. As an interesting sidenote, 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 timestep tends to zero.

Finally, we mention a moment-matching scheme suggested by Andersen and Brotherton-Ratcliffe [2005]. In their discretisation, the variance V is locally lognormal, where the parameters are determined such that the first two moments of the discretisation coincide with the theoretical moments. Kahl and Jäckel have incorporated this scheme into their comparisons, and conclude that for the Heston model it "has practically no convergence advantage over straightforward explicit Euler integration". For this reason we exclude it from our numerical results.

3. Strong convergence of the full truncation scheme

In the previous section we introduced the full truncation scheme and motivated how the truncation of the drift and diffusion of the variance process can help to reduce the bias inherent to

[8] Though they consider the balanced Milstein method (BMM), for the square root process their control functions (see figure 6) coincide with the implicit Milstein method. From now on we will therefore refer to their scheme as the IJK-IMM scheme.

any Euler discretisation of the Heston model. 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 realised payoffs converge to the option prices as the timestep used in the discretisation tends to zero. One may be tempted to think that for sufficiently small values of the timestep the probability of the discretisation in (9) of the square root process yielding negative values is negligible, so that for any choice of fixing functions the fixed Euler discretisation will converge. As plausible as this may seem, it is a fallacy. As an example, consider the case where $f_2(x) = f_3(x) = |x|$ and f_1 is the identity function. A priori nothing seems wrong, but if at any time instant the process becomes negative and such that its absolute value is larger than its mean-reverting value, this scheme will cause the process to drift to even more negative values. Numerical evidence suggests that this discretisation scheme indeed explodes and does not converge. Hence, convergence of any fix must be proven rigorously.

For European options weak convergence is typically enough to prove this result for Euler discretisations, see e.g. Kloeden and Platen [1999], although for more complex path-dependent derivatives strong convergence may be required. As mentioned earlier though, the square root dynamics of the Heston model preclude us from invoking the usual theorems on weak and strong convergence of Euler discretisations. Consequently, various authors have proven convergence of their particular discretisation of the mean reverting square root process in isolation. Recently, Diop [2003] and Bossy and Diop [2004] have proven that an Euler discretisation with the reflection fix converges weakly for a variety of mean-reverting constant elasticity of variance (CEV) processes. For the special case of the mean-reverting square root process, weak convergence of order 1 in the timestep is proven, provided that $\lambda\overline{V} > 2\eta^2$. 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 $\min\left(\lambda\overline{V}\eta^{-2},1\right)$. Diop proves strong convergence in the L^p ($p \geq 2$) sense of order ½ under a very restrictive condition, which is relaxed somewhat in Berkaoui et al. [2005]. For p = 2 the condition becomes:

$$\lambda\overline{V} \geq \tfrac{1}{2}\eta^2 + \max\left\{\eta\sqrt{14\lambda}, 6\sqrt{2}\eta^2\right\} \tag{18}$$

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 of order ½ for their discretisation, 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 [2005].

For the mean-reverting square root process in isolation, following Deelstra and Delbaen and Higham and Mao, we use Yamada's [1978] method to find the order of strong convergence. The big picture of our proofs is virtually identical to that of Higham and Mao, but the truncated drift complicates the proofs considerably. Here we merely report the main findings. First let us introduce some notation. The discretisation has already been introduced in equation (13) of lemma 1. For the full truncation scheme we have $f_1(x) = x$ and $f_2(x) = f_3(x) = x^+$. To distinguish between the discretisation of the variance and the true process, we will denote the discretisation with lowercase letters (v) and the true process with uppercase letters (V). Obviously $\overline{v} = \overline{V}$. Following Higham and Mao [2005] we also require the continuous-time approximation of (13):

$$v(t) \equiv v_n - \lambda(t - t_n)(v_n^+ - \overline{v}) + \eta\sqrt{v_n^+} \cdot \left(W_V(t) - W_V(t_n)\right) \tag{19}$$

The main findings for the mean-reverting square root process now follow from theorems 1 and 2.

Theorems 1 and 2 – Strong convergence of v(t) in the L^1 and L^2 sense
The full truncation scheme converges strongly of order ½ in the L^1 sense and of order ¼ in the L^2 sense, i.e. for sufficiently small values of the timestep Δt and some constants β and γ, we have:

$$\begin{aligned} &\sup_{t\in[0,T]} \mathbb{E}\left[\left|V(t)-v(t)\right|\right] \leq \beta\sqrt{\Delta t} \\ &\mathbb{E}\left[\sup_{t\in[0,T]} \left|V(t)-v(t)\right|^2\right]^{1/2} \leq \gamma(\Delta t)^{1/4} \end{aligned} \tag{20}$$

Proof: See the appendix. □

We note that the true order of L^2 convergence is probably higher, but we were not able to find a sharper estimate. Although the above theorems refer to the full truncation scheme, they also hold for the partial truncation scheme, with an easier proof. To discretise the stock price process, we use the direct discretisation in (11), although the proofs almost certainly carry over to the log-discretisation in (12). Analogous to equations (13) and (19), we now define:

$$\begin{aligned} &s_{n+1} = (1+\mu\Delta t)s_n + s_n\sqrt{v_n^+}\,\Delta W_{Sn} \\ &s(t) \equiv s_n + \mu(t-t_n)s_n + s_n\sqrt{v_n^+}\cdot\left(W_S(t)-W_S(t_N)\right) \end{aligned} \tag{21}$$

In theorem 3, strong L^2 convergence is proven of the stock price. The result is slightly different from the usual proofs of strong convergence, as we consider the stock price stopped by a stopping time. We need this result to prove convergence of plain-vanilla and barrier option prices.

Theorem 3 – Strong L^2 convergence of the stopped stock price
If we define the following stopping time for any i, j > 0:

$$\tau_{ij} \equiv \inf\left\{t \geq 0 : |S(t)| > i \vee |v(t)| > j\right\} \tag{22}$$

and $t_{ij} = \min(t,\tau_{ij})$, then for a sufficiently small timestep and some constant β_{ij}, depending on both i and j, we have:

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|S(t_{ij})-s(t_{ij})\right|^2\right]^{1/2} \leq \beta_{ij}(\Delta t)^{1/4} \tag{23}$$

Proof: See the appendix. □

Theorem 3 was originally proven in Higham and Mao [2005] for their particular discretisation. As before, we only supply the proof for the full truncation scheme, although one could follow the exact same steps to prove convergence for the extension of the partial truncation scheme to the Heston model. To our knowledge, Higham and Mao were the first to consider convergence of Euler discretisations of the Heston model. Moreover, they filled a gap in the literature by showing that the strong convergence in theorems 1-3 is sufficient to guarantee that the sample averages of the payoffs of plain vanilla options and barrier options converge to the true option prices. This is dealt with in the final theorem 4. Its proof is independent of the discretisation, and only depends on theorems 1-3, as well as some auxiliary results.

Theorem 4 – Convergence of plain vanilla options and barriers
Let us define the following payoffs:

$$V(T) = \left(\gamma(S(T) - K)\right)^{+} \cdot 1_{[S(t)\in A, 0\le t\le T]} \qquad V_{\Delta t}(T) = \left(\gamma(s(T) - K)\right)^{+} \cdot 1_{[s(t)\in A, 0\le t\le T]} \tag{24}$$

with $\gamma \in \{-1,1\}$ and A either of the form [0,B] or [B,∞). For any $\varepsilon > 0$ we can find i, j > 0 such that, for sufficiently small timesteps there is a constant β_{ij}, depending on both i and j, such that:

$$\left| \mathbb{E}[V(T)] - \mathbb{E}[V_{\Delta t}(T)] \right| \le \varepsilon + \beta_{ij}\sqrt{\Delta t} \tag{25}$$

Proof: See the appendix. □

4. Numerical results

The previous section established the strong convergence of the full truncation scheme and justified the pricing of plain vanilla and barrier options under this scheme. Obviously it is useful to have a theoretical result on the order of convergence, but several practical problems remain. Firstly, the theoretical results only say something about the limiting behaviour of the discretisation. Secondly, the actual rate of convergence may be faster than the theorem indicates. Thirdly, if we have two schemes that both achieve strong convergence of order ½ in the timestep Δt, anyone would prefer the scheme with the smallest constant in front of $\Delta t^{1/2}$. Though the proofs provide these constants, they are usually much too crude to be of any practical relevance. At the end of the day we should therefore 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 five Euler fixes, as well as the schemes of Broadie and Kaya [2006] and the second order IJK-IMM scheme of Kahl and Jäckel [2005a]. 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 the European call, and $\hat{\alpha}$ is the Monte Carlo estimator, the bias of the estimator equals $\mathbb{E}[\hat{\alpha}] - \alpha$, the variance of the estimator is $\mathrm{Var}(\hat{\alpha})$, and finally the root-mean-squared error (RMS error or RMSE) is defined as $(\text{bias}^2+\text{variance})^{1/2}$. 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 ½. Alfonsi [2005] compares both reflection and partial truncation to his scheme, but only for the mean-reverting square root process.

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, Pan and Singleton [2000] and Matytsin [1999], both of which add jumps to the stochastic variance process. As Broadie and Kaya [2006], we focus purely on European call options, since these can be valued in closed-form in both models. We leave the investigation of the bias in more complex path-dependent options for future research.

4.1. 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

Steps/yr.	Absorption	Reflection	Higham-Mao	Partial truncation	Full truncation
20	2.102	4.360	2.710	0.420	0.050
40	1.595	3.205	1.671	0.188	0.025
80	1.209	2.369	1.021	0.081	0.013
160	0.923	1.771	0.613	0.034	0.007
$O(\Delta t^p)$	0.40	0.43	0.71	1.21	0.95

Table 3: Bias for Euler fixes when pricing an ATM call in the Heston example
For all fixes the log-transformation for the stock price was used
Stochastic variance process: λ = 200%,V(0) = $\overline{V}$ = 9%, η = 100%, ρ = -30%
Stock price process: S = 100, μ = r = 5%
Deal specification: European call option, Maturity 5 yrs. True option price: 34.9998.

Paths	Steps/yr.	Full truncation – log			IJK-IMM-Absorption			Exact scheme	
		Bias	RMSE	CPU	Bias	RMSE	CPU	RMSE	CPU
10,000	20	0.050	0.591	0.2	-0.231	0.615	0.2	0.613	3.8
40,000	40	0.025	0.289	1.9	-0.020	0.286	2.1	0.290	15.3
160,000	80	0.013	0.145	15.4	0.083	0.166	16.7	0.146	61.3
640,000	160	0.007	0.07	122.6	0.113	0.134	131.4	0.073	244.5
	$O(\Delta t^p)$	0.95	1.02		0.10	0.74		1.02	

Table 4: Bias, RMS error and CPU time (in sec.) in the Heston example for an ATM call

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 Kahl and Jäckel [2005b] and Lord and Kahl [2006a]. In the latter paper it is also pointed out how to keep the characteristic function which appears in Broadie and Kaya's exact simulation algorithm continuous. Finally, for a very efficient Fourier inversion technique which works for almost all strike prices and maturities we point the reader to Lord and Kahl [2006b].

The set of parameters and deal specification used here can be found in the caption of Table 3 and stem from the second example of Broadie and Kaya [2006]. Though η is quite high and ρ is somewhat higher than is frequently found in an implied calibration, we feel the example is quite a good test case as $\eta^2 >> 2\lambda\overline{V}$, implying that the origin of the mean-reverting square root process is attainable. Furthermore, the probability of a particular discretisation yielding a negative value for V(t) is magnified via the large value of η, cf. equation (8), so that the way in which each discretisation treats the boundary condition will be put to the test. Conveniently, using the example of Broadie and Kaya allows us to compare all biased schemes to their exact scheme. As they 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 of RAM, using Microsoft Visual C++ 6.0 in a Windows XP environment. 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 mean-reverting square root process. We therefore chose to keep the implementation as straightforward as possible.

In Table 3 we first report the biases of the five Euler fixes for an at-the-money (ATM) call, using the log-transformation for the stock price, i.e. simulating using equations (9) and (12). To obtain accurate estimates of the bias we used 100 million simulation paths. The standard error of each bias is roughly 0.006. Without the log-transformation the biases are uniformly higher, so that it clearly makes sense to use it here. The first thing one notices is the enormous difference in

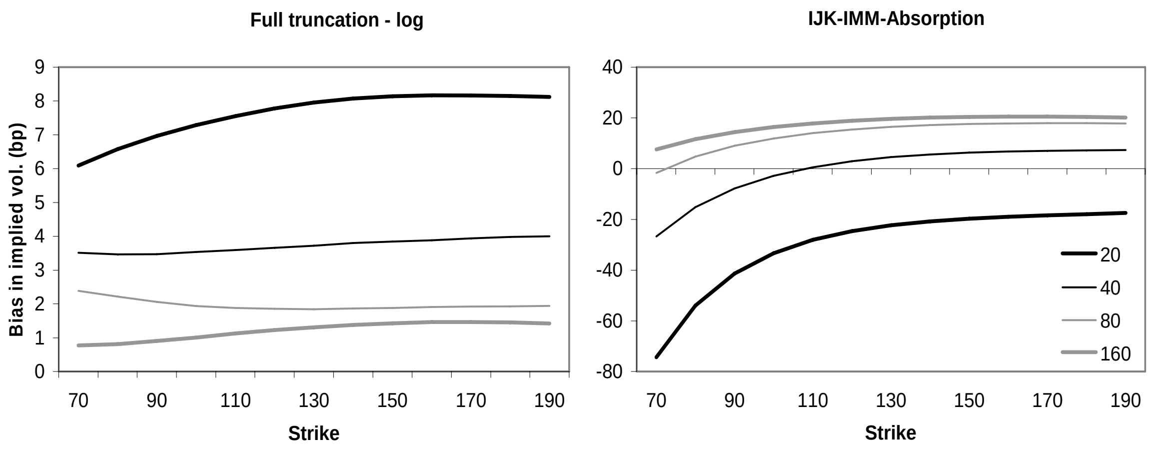

Figure 2: Bias as a function of the strike and the timestep in the Heston example

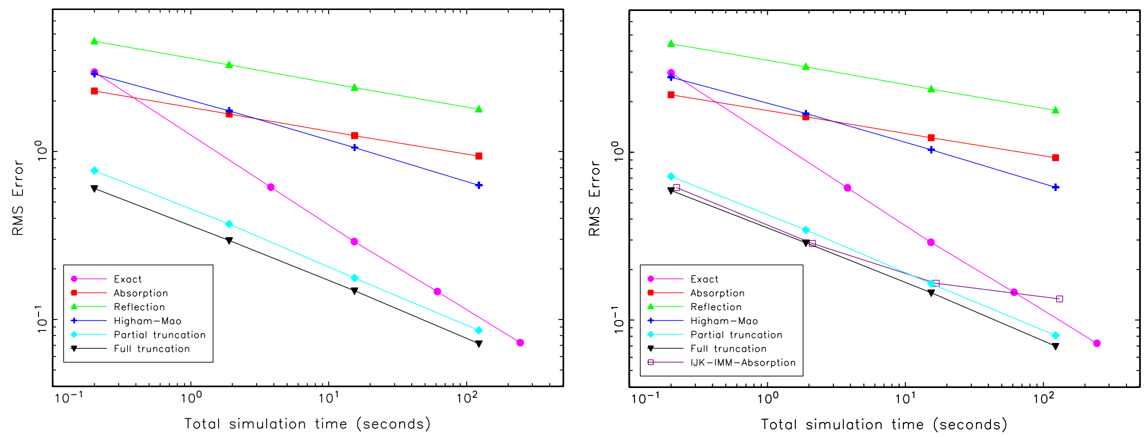

Figure 3: Convergence of the RMS error[9] in the Heston example for an ATM call
Left panel: Direct integration of stock price, Right panel: Log-transformation of stock price

the magnitude of the bias. To relate the size of the bias of the full truncation scheme to implied volatilities, one can glance at Figure 2. Even with twenty timesteps per year the bias is only 7 basispoints (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. It is clear that the combination of truncating the effective volatility and allowing the variance process to remain negative is the most effective in reducing the bias. One possible reason for this is that the boundary behaviour remains as close as possible to the boundary behaviour of the continuous time process. As expected, the truncation of the drift lowers the bias even further.

For the order of weak convergence, it is worthwhile to note that under some regularity conditions, see e.g. Theorem 14.5.2. of Kloeden and Platen [1999], the Euler scheme converges weakly with order 1 in the timestep. Though the SDE for the mean-reverting square root process does not satisfy these conditions, both schemes seem to regain this weak order[10], the partial truncation scheme leading to an order of convergence even slightly above 1. In contrast,

[9] The first point from the left in the exact scheme has been extrapolated.
[10] The order of weak convergence was estimated here by regressing ln(|bias|) on a constant plus ln(Δt).

absorption and reflection have a weak order of convergence slightly under ½. The orders of convergence are roughly the same when considering the first example of Broadie and Kaya.

Turning to the IJK-IMM scheme of Kahl and Jäckel [2005a], the estimated bias of this scheme is reported in Table 4. Note that for the combination of parameters in Table 3 the IMM scheme fails to preserve positivity for the variance process. We chose to implement it in combination with the absorption fix, and hence refer to the full scheme as IJK-IMM-Absorption. Initially the size of the bias compares favourably to the other biased schemes, though full truncation still performs better. The convergence is however rather erratic, similar to the aforementioned findings of Glasserman [2003, pp. 356-358] who considered another second order Taylor scheme for the Heston model. The bias seems to increase when increasing the number of timesteps per year from 40 to 80, and then to 160. This behaviour is visualised in Figure 2 where we plot the bias in terms of implied volatility as a function of the strike and the timestep. Indeed, the sign of the bias switches when the timestep is decreased. In contrast, the absolute value of the bias decreases uniformly for the full truncation scheme. Interestingly, using the full truncation fix in combination with the IMM scheme makes the convergence uniform again, and in addition increases the weak order of convergence to 0.95. These results are not included in Table 4 as the size of the bias was slightly larger than of the IJK-IMM-Absorption scheme.

Finally, let us investigate the RMS error and computation time. These are reported in Table 4 for full truncation, IJK-IMM-Absorption and the exact scheme. In Figure 3 the RMSE is plotted as a function of the timestep for all schemes. The choice of the number of paths is an important issue here. It suffices to say here that Duffie and Glynn [1995] have proven that if the weak order of convergence is p, one should increase the number of paths proportional to $(\Delta t)^{-p}$. When p = 1, this means that if the timestep 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[11], 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 twenty timesteps per year, the RMSEs of both schemes are roughly the same. Given the stark contrast in computation time required for both schemes, combined with the fact that this is an example with an extremely large volatility of variance η, causing large biases, the full truncation scheme should certainly be the discretisation method of choice, even for weakly path-dependent options.

4.2. Results for the Bates model

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

$$\begin{aligned} dS(t) &= (\mu - \xi\overline{\mu}_J)S(t)dt + \sqrt{V(t)}\,S(t)dW_S(t) + J_{N(t)}\,S(t)dN(t) \\ dV(t) &= -\lambda\left(V(t) - \overline{V}\right)dt + \eta\sqrt{V(t)}\,dW_V(t) \end{aligned} \tag{26}$$

[11] 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 $O((\ln N)^2/N)$.

Steps/yr.	Absorption		Partial truncation		Full truncation	
	No log	Log	No log	Log	No log	Log
2	0.588	0.832	2.860	2.780	-0.145	0.103
4	0.281	0.411	0.280	0.410	-0.102	0.028
8	0.134	0.201	0.038	0.104	-0.059	0.008
16	0.055	0.089	-0.009	0.025	-0.033	0.001
$O(\Delta t^p)$	1.13	1.07	2.81	2.24	0.73	2.17

Table 5: Bias for Euler fixes when pricing an ATM call in the Bates example
Stochastic variance process: λ = 399%, V(0) = 0.8836%, $\overline{V}$ = 9%, η = 27%, ρ = -79%
Stock price process: S = 100, μ = r = 3.19%
Jump process: ξ = 11%, $\overline{\mu}_J$ = -12%, σ_J = 15%
Deal specification: European call option, Maturity 5 yrs. True option price: 20.1642.

Paths	Steps/yr.	Full truncation – log			IJK-IMM			Exact scheme	
		Bias	RMSE	CPU	Bias	RMSE	CPU	RMSE	CPU
10,000	2	0.103	0.250	0.1	0.885	0.914	0.1	0.223	0.6
40,000	4	0.028	0.115	0.3	0.432	0.446	0.3	0.112	2.3
160,000	8	0.008	0.057	1.9	0.204	0.211	2.1	0.056	9.1
640,000	16	0.001	0.028	12.6	0.094	0.098	13.7	0.028	36.2
	$O(\Delta t^p)$	2.17	1.05		1.08	1.07		1.00	

Table 6: Bias, RMS error and CPU time (in sec.) in the Bates example for an ATM call

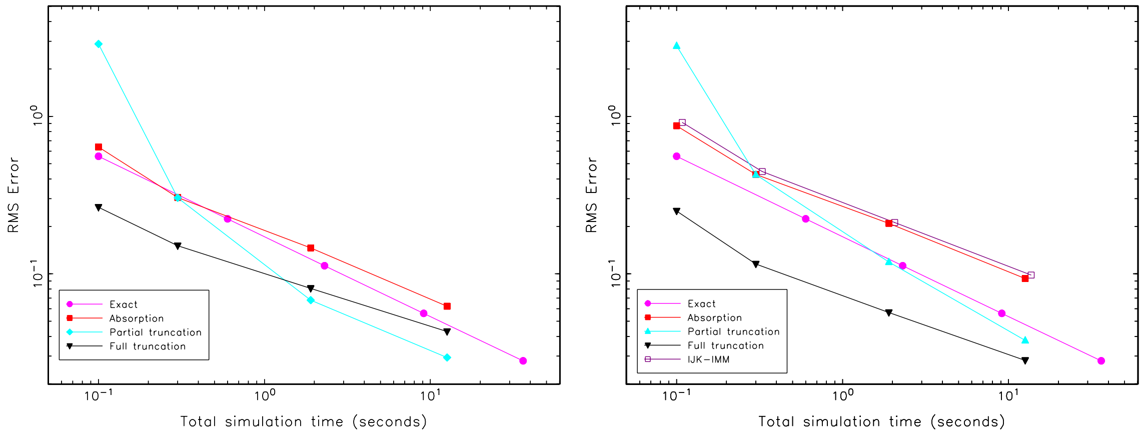

Figure 4: Convergence of the RMS error in the Bates example for an ATM call
Left panel: Direct integration of stock price, Right panel: Log-transformation of stock price

where N is a Poisson process with intensity ξ, independent of the Brownian motions. The random variable J_i denotes the i^{th} relative jump size and is lognormally distributed, $\ln J_i \sim N(\mu_J, \sigma_J^2)$. If the i^{th} jump occurs at time t, the stock price right after the jump equals $S(t+) = J_i S(t-)$. To ensure no arbitrage, $\overline{\mu}_J$ in (26) has to be the expected relative jump size:

$$1+\overline{\mu}_J = \mathbb{E}[J_i] = \exp(\mu_J + \tfrac{1}{2}\sigma_J^2) \tag{27}$$

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 timestep of length T is made till 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 discretisation error.

The parameters and deal specification for our example can be found in the caption of Table 5. This example stems from Duffie, Pan and Singleton [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 $\eta^2 << 2\lambda\overline{V}$, which firstly means that the origin of the square root process is not attainable. Secondly, the low level of η implies that the probability of any discretisation yielding a negative value is significantly smaller than in the Heston example. Hence we may expect that the sizes of the bias are lower than in the previous example. Thirdly and finally, this combination of parameters is such that the IMM scheme preserves positivity. The IJK-IMM scheme, contrary to the previous example, does not require an additional assumption about the treatment of the stochastic variance at the boundary.

Starting with the bias of an ATM call in Table 5, we see an interesting pattern when the logarithmic transformation is not used for the stock price. As expected, the full truncation scheme yields a lower bias than the partial truncation scheme. The bias is however decreased so far that it turns negative. When using the logarithmic transformation, the bias of the full truncation scheme remains positive. How can this be explained? On the one hand, note that not using a logarithmic transformation induces a negative bias in the stock price when the expected rate of return μ is positive. For example, if no jumps were included in this example, we would have had $\mathbb{E}[S(1) \mid S(0)] = 100 \cdot 1.0319 = 103.19$ if (11) were used. In contrast, the same expectation equals $100 \cdot \exp(0.0319) = 103.24$ when (12) is used. Since jumps are used, and the expected relative jump size is negative, the drift rate in (26) is slightly higher. This causes an even larger negative bias in the stock price. On the other hand, fixing negative variances using traditional fixes such as absorption or reflection induces a positive bias in the option price, as was evident in the previous subsection. Due to the combination of parameters of this example, the probability of the Euler discretisations yielding a negative value is significantly reduced. As the full truncation scheme is tailored to minimise an upward bias, the combination of both downward biases drives the bias into the negative domain. By using the logarithmic coordinates in (12) this effect is corrected.

Having explained this behaviour, we can again conclude that the full truncation scheme by far outperforms the other Euler fixes in terms of bias, certainly when applied in logarithmic coordinates. Results for the other fixes, apart from absorption, are not reported here as the previous subsection demonstrated quite convincingly that the truncation schemes by far outperform the other Euler fixes. Absorption was included since it is the most widely used fix. The order of weak convergence is higher than in the Heston example, most likely due to the lesser importance of the fixes here. The relatively high bias found when the partial truncation scheme is used with only two timesteps, also evident in Figure 4 is probably due to the fact that Δt is barely smaller than $2/\lambda$. As can be seen from (14), this is a strict requirement for the first moment of the partial truncation scheme to converge. Still, the order of convergence is, as before, the highest among all biased schemes. The size of the bias is nevertheless higher for the chosen timesteps.

As mentioned, for this example the IJK-IMM scheme automatically preserves positivity. However, it is outperformed both in terms of bias and order of weak convergence by the full truncation scheme. Interestingly, Figure 4 indicates that the level of bias of the IJK-IMM scheme is here roughly comparable to that of the absorption scheme. Comparing to the exact scheme, we see that like before, the RMS errors of both the full truncation scheme and the exact scheme are roughly equivalent. The cause of this is the almost negligible level of bias found by the full truncation scheme, with only 2 steps per year the bias is 14 bp in terms of implied volatility. Clearly, since the Euler full truncation scheme is the computationally least intensive scheme, and produces the smallest bias among all biased schemes, it should be the preferred discretisation.

5. Conclusions and further research

In this paper we have considered the simulation of the Heston stochastic volatility model and varieties thereof. In this model, the stochastic variance is modelled as a mean-reverting square root process. When discretising this process one immediately runs into the problem that although the process itself is guaranteed to be nonnegative, any Euler discretisation has a nonzero probability of becoming negative in the next timestep, regardless of the size of the timestep. Hence, one has 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 up to propose a new scheme, the full truncation scheme, designed specifically to minimise the positive bias one finds when pricing European options using the traditional fixes. Strong convergence of order ½ in the timestep is proven for this scheme. Combined with the recent results of Higham and Mao [2005] this justifies the pricing of European and barrier options.

Thirdly and finally, we numerically compare the various Euler schemes to each other, as well as to the recently developed quasi-second order IJK-IMM scheme by Kahl and Jäckel [2005a] and the exact scheme of Broadie and Kaya [2006]. Both 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, when looking at the error made when pricing a European call, we find the full truncation scheme produces the smallest bias by far, even when compared to the IJK-IMM scheme. The order of weak convergence seems to be around 1 in the timestep. The full truncation fix therefore seems to bring back the order of weak convergence to the theoretical level for an Euler discretisation of an SDE with Lipschitzian dynamics. We mention that the partial truncation scheme seems to have an even higher order of convergence. Nonetheless, the size of the bias is much higher than that of full truncation for practical values of the timestep. Interestingly, using our full truncation fix within the IJK-IMM scheme also seems to greatly improve its order of weak convergence.

Given the almost negligible levels of bias we find with the full truncation scheme, the scheme is able to generate a much smaller RMS error given a certain computational budget than any other biased or exact scheme considered here. As such it should be the preferred discretisation method for the simulation of the Heston stochastic volatility model and varieties hereof, even when valuing European or weakly path-dependent options. In future work we aim to investigate the effectiveness of the full truncation scheme when applied to exotic path-dependent options in more general stochastic volatility models, and perhaps design even more effective schemes. Clearly, although our focus has here been on the Heston model, the schemes presented here can readily be adapted to any stochastic process with reflecting boundaries, such as the class of affine jump-diffusion models and CEV-type 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:

$$\begin{aligned} \ln S(t) &= \ln S(t-\Delta t) + r + \lambda h(t) + \sqrt{h(t)}z(t) \\ h(t+\Delta t) &= \omega + \beta h(t) + \alpha\left(z(t) - \gamma\sqrt{h(t)}\right)^2 \end{aligned} \tag{27}$$

where z(t) is a standard normal random variable and h(t) is the conditional variance of the log-return between t-Δt and t. In this setup h(t) is known at time t-Δt. Note that all the model parameters will depend on the chosen timestep 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 (27) to the Heston model with $\rho = -1$ as the timestep tends to zero, we see that in their proof they choose[12] $\omega = (\lambda\overline{V} - \frac{1}{4}\eta^2)(\Delta t)^2$, $\beta = 0$ and $\alpha = \frac{1}{4}\eta^2(\Delta t)^2$. Positivity of the conditional variance $h(t+\Delta t)$ can thus only be guaranteed provided that $\lambda\overline{V} \geq \frac{1}{4}\eta^2$. This is the same condition under which the schemes of Alfonsi [2005] and Kahl and Jäckel [2005b] preserve positivity, and not surprisingly so as we already remarked the equivalence of these three schemes to first order in Δt in section 2.5. Looking in closer detail at their estimation procedure, we see that they only included options with an absolute moneyness less than or equal to ten percent, i.e. at or around at-the-money options. In the Heston model $\lambda\overline{V}$ can certainly be smaller than $\frac{1}{4}\eta^2$, when the skew is quite pronounced. This would not be noticed however if only options at or around the at-the-money level would be included in the calibration procedure. Concluding, the "fixes" considered in the continuous time setting in this paper certainly also have an application in a discrete time setting.

Bibliography

ALBIN, J.M.P., ASTRUP JENSEN, B., MUSZTA, A. AND M. RICHTER (2005). "A note on the Lipschitz condition for Euler approximation of stochastic differential equations", working paper, Chalmers University of Technology and Copenhagen Business School.

ALFONSI, A. (2005). "On the discretization schemes for the CIR (and Bessel squared) processes", CERMICS report 2005-279.

ANDERSEN, L.B.G. AND J. ANDREASEN (2002). "Volatile volatilities", *Risk*, vol. 15, no. 12, December 2002, pp. 163-168.

ANDERSEN, L.B.G. AND R. BROTHERTON-RATCLIFFE (2005). "Extended LIBOR market models with stochastic volatility", *Journal of Computational Finance*, vol. 9, no. 1, pp. 1-40.

ANDERSEN, L.B.G. AND V.V. PITERBARG (2004). "Moment explosions in stochastic volatility models", working paper, Bank of America.

BATES, D.S. (1996). "Jumps and stochastic volatility: exchange rate processes implicit in Deutsche Mark options", *Review of Financial Studies*, vol.9, no.1., pp. 69-107.

BERKAOUI, A., M. BOSSY AND A. DIOP (2005). "Euler schemes for SDEs with non-Lipschitz diffusion coefficient: strong convergence", INRIA working paper no. 5637.

BOSSY, M. AND A. DIOP (2004). "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.

BROADIE, M. AND Ö. KAYA (2004). "Exact simulation of option greeks under stochastic volatility and jump diffusion models", in R.G. Ingalls, M.D. Rossetti, J.S. Smith and B.A. Peters (eds.), *Proceedings of the 2004 Winter Simulation Conference*.

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

BROADIE, M. AND Ö. KAYA (2006). "Exact simulation of stochastic volatility and other affine jump diffusion processes", *Operations Research*, vol. 54, no. 2.

BRIGO, D. AND F. MERCURIO (2001). *Interest Rate Models: Theory and Practice*, Springer.

CARR, P. AND D.B. MADAN (1999). "Option valuation using the Fast Fourier Transform", *Journal of Computational Finance*, vol. 2, no. 4, pp. 61-73.

CHRISTOFFERSEN, P. AND K. JACOBS (2004). "Which GARCH model for option valuation?", *Management Science*, vol. 50, no. 9, pp. 1204-1221.

COX, J.C., INGERSOLL, J.E. AND S.A. ROSS (1985). "A theory of the term structure of interest rates", *Econometrica*, vol. 53, no. 2, pp. 385-407.

DEELSTRA, G. AND F. DELBAEN (1998). "Convergence of discretized stochastic (interest rate) processes with stochastic drift term", *Applied Stochastic Models and Data Analysis*, vol. 14, no. 1, pp. 77-84.

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

DUFFIE, D. AND P. GLYNN (1995). "Efficient Monte Carlo simulation of security prices", *Annals of Applied Probability*, vol. 5, no. 4, pp. 897-905.

DUFFIE, D., PAN, J. AND K. SINGLETON (2000). "Transform analysis and asset pricing for affine jump-diffusions", *Econometrica*, vol. 68, pp. 1343-1376.

FELLER, W. (1951). "Two singular diffusion problems", *Annals of Mathematics*, vol. 54, pp. 173-182.

GATHERAL, J. (2005). Case studies in financial modelling course notes, Courant Institute of Mathematical Sciences, Fall Term 2005.

GLASSERMAN, P. (2003). *Monte Carlo methods in financial engineering*, Springer Verlag, New York.

HESTON, S.L. (1993). "A closed-form solution for options with stochastic volatility with applications to bond and currency options", *Review of Financial Studies*, vol. 6, no. 2, pp. 327-343.

HESTON, S.L. AND S. NANDI (2000). "A closed-form GARCH option valuation model", *Review of Financial Studies*, vol. 13, no. 3, pp. 585-625.

HIGHAM, D.J. AND X. MAO (2005). "Convergence of the Monte Carlo simulations involving the mean-reverting square root process", *Journal of Computational Finance*, vol. 8, no. 3, pp. 35-62.

JÄCKEL, P. (2002). *Monte Carlo methods in finance*, John Wiley and Sons.

JÄCKEL, P. (2004). "Stochastic volatility models: past, present and future", pp. 379-390 in P. Wilmott (ed). *The Best of Wilmott 1: Incorporating the Quantitative Finance Review*, P. Wilmott (ed.), John Wiley and Sons, pp. 379-390.

KAHL, C. AND P. JÄCKEL (2005A). "Fast strong approximation Monte-Carlo schemes for stochastic volatility models", working paper, ABN·AMRO and University of Wuppertal.

KAHL, C. AND P. JÄCKEL (2005B). “Not-so-complex logarithms in the Heston model”, *Wilmott Magazine*, September 2005.

KAHL, C. (2004). “Positive numerical integration of stochastic differential equations”, Diploma thesis, University of Wuppertal and ABN·AMRO.

KARLIN, S. AND H. TAYLOR (1981). *A second course in stochastic processes*, Academic Press.

KLOEDEN, P.E. AND E. PLATEN (1999). *Numerical solution of stochastic differential equations*, 3rd edition, Springer Verlag, New York.

LORD, R. AND C. KAHL (2006A). “Why the rotation count algorithm works”, working paper, Erasmus University Rotterdam, Rabobank International, University of Wuppertal and ABN·AMRO.

LORD, R. AND C. KAHL (2006B). “Optimal Fourier inversion”, working paper, Erasmus University Rotterdam, Rabobank International, University of Wuppertal and ABN·AMRO.

KLOEDEN, P.E. AND E. PLATEN (1999). *Numerical solution of stochastic differential equations*, 3rd edition, Springer Verlag, New York.

MARSAGLIA, G. AND W.W. TSANG (2000). “A simple method for generating gamma variables”, *ACM Transactions on Mathematical Software*, vol. 26, no. 3, pp. 363-372.

MATYTSIN, A. (1999). “Modelling volatility and volatility derivatives”, Columbia Practitioners Conference on the Mathematics of Finance.

REVUZ, D. AND M. YOR (1991). *Continuous Martingales and Brownian Motion*, Springer Verlag, New York.

WILLARD, G.A. (1997). “Calculating prices and sensitivities for path-independent derivative securities in multifactor models”, *Journal of Derivatives*, vol. 5, no. 1, pp. 45-61.

YAMADA, T. (1978). “Sur une construction des solutions d’équations différentielles stochastiques dans le cas non-Lipschitzien”, in *Séminaire de Probabilité*, vol. XII, LNM 649, pp. 114-131, Springer, Berlin.

Appendix – Proof of strong convergence

In this appendix we prove strong convergence of the full truncation scheme applied to the Heston model. First of all we consider the mean-reverting square root process in isolation. For ease of exposure the discretisation over a finite time horizon [0,T] is performed on a uniform grid $t_n = n\Delta t$, n = 1, …, $T/\Delta t$. The full truncation scheme then reads:

$$v_{n+1} = v_n - \lambda\Delta t(v_n^+ - \overline{v}) + \eta\sqrt{v_n^+}\,\Delta W_{Vn} \tag{A.1}$$

where v_i indicates the value of the discretisation at t_i and $\Delta W_{Vn} = W_V(t_{n+1}) - W_V(t_n)$. To distinguish between the discretisation of the variance and the true process, we will denote the discretisation with small letters (v) and the true process with capital letters (V). Obviously $\overline{v} = \overline{V}$. Following Higham and Mao [2005] we will consider the continuous-time approximation of (A.1):

$$v(t) \equiv v_n - \lambda(t - t_n)(v_n^+ - \overline{v}) + \eta\sqrt{v_n^+}\cdot\big(W_V(t) - W_V(t_n)\big) \tag{A.2}$$

or, in integral notation:

$$v(t) = v(0) - \lambda\int_0^t (v_\tau(u)^+ - \overline{v})\,du + \eta\int_0^t \sqrt{v_\tau(u)^+}\,dW_V(u) \tag{A.3}$$

where $v(0) = v_0$, $v_\tau(t) = v(\tau(t))$ and $\tau(t)$ equals t_n if $t_n \le t \le t_{n+1}$. Obviously $v_\tau(t)$ coincides with v(t) at the gridpoints of the discretisation.

Bounds on the first and second moment of the stochastic variance

One of the elements required in proving strong convergence of the full truncation scheme, are bounds on the first and second moments of v_n in (A.1). In the remainder we denote the first and second moments by $x_n \equiv \mathbb{E}[v_n]$ and $y_n \equiv \mathbb{E}[v_n^2]$ respectively. In the main text lemma 1 already supplied the following inequality:

$$x_n = \mathbb{E}[v_n] \le (1 - \lambda\Delta t)^n (v_0 - \overline{v}) + \overline{v} \tag{A.4}$$

As we do not require very sharp bounds, we will often use the following corollary.

Corollary 1:
For $\Delta t < 2/\lambda$ the first moment of v_n in the full truncation scheme is bounded from above by:

$$x_n \le |\, v_0 - \overline{v}\,| + \overline{v} \tag{A.5}$$

Proof:
Trivial from (A.4). □

Secondly, we will find an upper bound on the second moment of v_n.

Lemma 2 – Bounding the second moment of the full truncation scheme
For any n = 0, …, N where $N\Delta t = T$, and $\Delta t < 2/\lambda$, the second moment of v_n in the full truncation scheme is bounded by:

$$y_n \le v_0^2 + T \cdot \left\{ 2\lambda\bar{v}U_x + (\lambda\bar{v}\Delta t)^2 + \frac{\Delta t(\eta^2 - 2\lambda^2\bar{v}\Delta t)}{1-(1-\lambda\Delta t)^2} \right\} \equiv U_y(\Delta t) \tag{A.6}$$

Proof:
Clearly, $y_0 = v_0^2$ so that the assertion is then true. Suppose the lemma now holds true for some n. Using (A.1) we can then write:

$$\begin{aligned} y_{n+1} &= (\lambda\bar{v}\Delta t)^2 + \mathbb{E}\left[(v_n - \lambda\Delta t v_n^+)^2\right] \\ &\quad + 2\lambda\bar{v}\Delta t \cdot \mathbb{E}\left[(v_n - \lambda\Delta t v_n^+)\right] + \eta^2\Delta t\,\mathbb{E}[v_n^+] \end{aligned} \tag{A.7}$$

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:

$$f(v_n) = \begin{cases} v_n^2(1-\lambda\Delta t)^2 + \left(2\lambda\bar{v}\Delta t(1-\lambda\Delta t) + \eta^2\Delta t\right)v_n & v_n \ge 0 \\ v_n^2 + 2\lambda\bar{v}\Delta t\, v_n & v_n \le 0 \end{cases} \tag{A.8}$$

This means that f(x) is a piecewise parabola:

$$f(x) = \begin{cases} ax^2 + bx & x \le 0 \\ cx^2 + dx & x \ge 0 \end{cases} \tag{A.9}$$

We will bound f(x) from above by a quadratic function with constant coefficients:

$$f(x) \le \alpha x^2 + \beta x + \gamma \tag{A.10}$$

so that if both random variables are ordered, their expectation will also satisfy this ordering. First of all note that since $\lambda < 2/\Delta t$, we have that $c < a = 1$, so that we can safely set $\alpha = a = 1$. Setting β equal to b means that we must choose $\gamma \ge 0$. For $x \ge 0$ (A.10) now implies that:

$$(c-1)x^2 + (d-b)x - \gamma \le 0 \tag{A.11}$$

We will ensure that this holds true by choosing γ such that the maximum of the parabola in (A.11) with respect to x on $[0,\infty)$ equals zero. The maximum occurs in:

$$x^* = \frac{d-b}{2(1-c)} \tag{A.12}$$

When $d < b$ this is smaller than zero, so then we only need to ensure that the condition in (A.11) holds true for $x = 0$. This can be achieved by choosing $\gamma \ge 0$. When $d \ge b$ it can be shown that:

$$\gamma \geq \frac{(d-b)^2}{4(1-c)} \tag{A.13}$$

ensures that the condition in (A.11) always holds true. Returning to (A.7) we then have:

$$y_{n+1} \leq y_n + 2\lambda\bar{v}\Delta t \cdot x_n + (\lambda\bar{v}\Delta t)^2 + \frac{\Delta t(\eta^2 - 2\lambda^2\bar{v}\Delta t)}{1-(1-\lambda\Delta t)^2} \tag{A.14}$$

Repeated use of (A.14) and corollary 1 immediately yields (A.6), if we in addition also bound $n\Delta t$ from above by T. □

It is important to note that $\lim_{\Delta t \to 0} U_y(\Delta t) = v_0^2 + TU_x < \infty$, so that the second moment of the discretisation does not blow up in finite time. We proceed by studying the strong L^1 error, followed by the strong L^2 error.

The strong L^1 error for the mean-reverting square root process

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 1 and 2.

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

$$\sup_{t\in[0,T]} \mathbb{E}\left[(v(t)-v_\tau(t))^2\right] \leq (\lambda\Delta t)^2 \cdot \left(\bar{v} + U_y(\Delta t)\right) + \eta^2 \Delta t \cdot \sqrt{U_y(\Delta t)} \equiv U_{cont}(\Delta t) \tag{A.15}$$

Proof:
For $t \in [t_n, t_{n+1})$ we have:

$$\mathbb{E}\left[(v(t)-v_\tau(t))^2\right] = \lambda^2 (t-t_n)^2 \cdot \mathbb{E}[(v_n^+ - \bar{v})^2] + \eta^2 (t-t_n) \cdot \mathbb{E}[v_n^+] \tag{A.16}$$

The first term can be bounded from above by:

$$\mathbb{E}[(v_n^+ - \bar{v})^2] = \bar{v}^2 - 2\bar{v}\mathbb{E}[v_n^+] + \mathbb{E}[v_n^+ v_n^+] \leq \bar{v}^2 + y_n \tag{A.17}$$

so that (A.16) becomes:

$$\begin{aligned} \mathbb{E}\left[(v(t)-v_\tau(t))^2\right] &\leq \lambda^2 (t-t_n)^2 \cdot (\bar{v} + y_n) + \eta^2 (t-t_n) \cdot \sqrt{y_n} \\ &\leq \lambda^2 (t-t_n)^2 \cdot \left(\bar{v} + U_y(\Delta t)\right) + \eta^2 (t-t_n) \cdot \sqrt{U_y(\Delta t)} \end{aligned} \tag{A.18}$$

Finally, the supremum on [0,T] is then bounded from above by:

$$\sup_{t\in[0,T]} \mathbb{E}\left[(v(t)-v_\tau(t))^2\right] \leq (\lambda\Delta t)^2 \cdot \left(\bar{v} + U_y(\Delta t)\right) + \eta^2 \Delta t \cdot \sqrt{U_y(\Delta t)} \tag{A.19}$$

which completes the proof. □

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

Theorem 1 – Strong convergence of v(t) in the L^1 sense
The full truncation scheme converges strongly of order ½ in the L^1 sense, i.e. for sufficiently small values of the timestep Δt and for some constant β, we have:

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left|V(t)-v(t)\right|\right] \le \beta\sqrt{\Delta t} \tag{A.20}$$

Proof:
In this proof we bound $\mathbb{E}\left[\left|V(t)-v(t)\right|\right]$ from above in a function of the timestep, so that we can prove that this L^1 norm tends to zero as the timestep tends to zero. As in Yamada [1978], this is achieved by bounding $\mathbb{E}\left[\phi_k\left(V(t)-v(t)\right)\right]$ 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 $a_0 = 1$ and define $a_k = e^{-k(k+1)/2}$ for $k \ge 1$, so that $\int_{a_k}^{a_{k-1}} u^{-1}du = k$. For each integer $k \ge 1$ there exists a continuous function ψ_k with support in (a_{k-1}, a_k) such that

$$0 \le \psi_k(u) \le 2k^{-1}u^{-1} \tag{A.21}$$

and $\int_{a_{k-1}}^{a_k} \psi_k(u)du = k$. Defining $\phi_k(x) = \int_0^{|x|}\int_0^y \psi_k(u)\,du\,dy$, then $\phi_k \in C^2(\mathbb{R},\mathbb{R})$, $\phi_k(0) = 0$, and:

$$\begin{aligned} &\left|\phi_k'(x)\right| \le 1 \\ &\left|\phi_k''(x)\right| \begin{cases} \le 2k^{-1}\,|x|^{-1} & a_k < |x| < a_{k-1} \\ = 0 & \text{otherwise} \end{cases} \end{aligned} \tag{A.22}$$

A final property we will use is:

$$|x| - a_{k-1} \le \phi_k(x) \le |x| \tag{A.23}$$

Consider $\phi_k\left(V(t)-v(t)\right)$. Using Itô's lemma and taking expectations yields:

$$\mathbb{E}\left[\phi_k\left(V(t)-v(t)\right)\right] = -\lambda M(t) + \tfrac{1}{2}\eta^2 I(t) \tag{A.24}$$

where we defined:

$$\begin{aligned} M(t) &\equiv \mathbb{E}\left[\int_0^t \phi_k'\left(V(u)-v(u)\right)\cdot\left(V(u)-v_\tau(u)^+\right)du\right] \\ I(t) &\equiv \mathbb{E}\left[\int_0^t \phi_k''\left(V(u)-v(u)\right)\cdot\left(\sqrt{V(u)}-\sqrt{v_\tau(u)^+}\right)^2 du\right] \end{aligned} \tag{A.25}$$

One can show that $\left(\sqrt{x}-\sqrt{y^+}\right)^2 \le |x-y|$. Since $V(u) \ge 0$ a.s., we have:

$$I(t) \le \mathbb{E}\left[\int_0^t \left|\phi_k''\left(V(u)-v(u)\right)\right| \cdot \left| V(u)-v_\tau(u) \right| du\right] \tag{A.26}$$

Secondly, note that $\left| V(u)-v_\tau(u) \right| \le \left| V(u)-v(u) \right| + \left| v(u)-v_\tau(u) \right|$. Thirdly, we use the property of the second derivative of ϕ in (A.22). It follows that:

$$\begin{aligned} I(t) &\le \mathbb{E}\left[\int_0^t \tfrac{2}{k} 1_{[a_k < |V(u)-v(u)| < a_{k-1}]}\, du + \int_0^t \tfrac{2}{ka_k} \left| v(u)-v_\tau(u) \right| du\right] \\ &\le \tfrac{2t}{k} + \tfrac{2t}{ka_k}\sqrt{U_{cont}(\Delta t)} \end{aligned} \tag{A.27}$$

where we used $\mathbb{E}[|X|] \le \sqrt{\mathbb{E}[X^2]}$ for any random variable X and lemma 3. Turning to M(t), we use the property of the first derivative of ϕ from (A.22) and obtain:

$$\begin{aligned} M(t) &\le \mathbb{E}\left[\int_0^t \left| V(u)-v_\tau(u)^+ \right| du\right] \le \mathbb{E}\left[\int_0^t \left| V(u)-v_\tau(u) \right| du\right] \\ &\le \mathbb{E}\left[\int_0^t \left| V(u)-v(u) \right| du\right] + \mathbb{E}\left[\int_0^t \left| v(u)-v_\tau(u) \right| du\right] \end{aligned} \tag{A.28}$$

M(t) can thus be bounded from above by:

$$M(t) \le \mathbb{E}\left[\int_0^t \left| V(u)-v(u) \right| du\right] + t\sqrt{U_{cont}(\Delta t)} \tag{A.29}$$

Combining the bounds on I(t) and M(t) from (A.26) and (A.29) in (A.24) yields:

$$\mathbb{E}\left[\phi_k\left(V(t)-v(t)\right)\right] \le \lambda \mathbb{E}\left[\int_0^t \left| V(u)-v(u) \right| du\right] + t\left(\lambda + \frac{\eta^2}{ka_k}\right)\sqrt{U_{cont}(\Delta t)} + \frac{\eta^2 t}{k} \tag{A.30}$$

From property (A.23) it follows that $\mathbb{E}\left[\phi_k\left(V(t)-v(t)\right)\right] \ge \mathbb{E}\left[\left| V(t)-v(t) \right|\right] - a_{k-1}$, hence:

$$\mathbb{E}\left[\left| V(t)-v(t) \right|\right] \le a_{k-1} + \frac{\eta^2 T}{k} + T\left(\lambda + \frac{\eta^2}{ka_k}\right)\sqrt{U_{cont}(\Delta t)} + \lambda \mathbb{E}\left[\int_0^t \left| V(u)-v(u) \right| du\right] \tag{A.31}$$

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 to find:

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left| V(t)-v(t) \right|\right] \le e^{\lambda T}\left[a_{k-1} + \frac{\eta^2 T}{k} + T\left(\lambda + \frac{\eta^2}{ka_k}\right)\sqrt{U_{cont}(\Delta t)}\right] \tag{A.32}$$

Since (A.32) holds for any value of k, it is easy to show that $\lim_{\Delta t \to 0} \sup_{t\in[0,T]} \mathbb{E}\left[\left| V(t)-v(t) \right|\right] = 0$ as in corollary 3.1 of Higham and Mao. Furthermore, the order of convergence is determined by the order of $U_{cont}(\Delta t)$, which we know to be $O(\Delta t)$ from lemma 3. As we here take its square root, the order of strong convergence is ½. □

The strong L^2 error for the mean-reverting square root process

The following theorem derives a bound on the error using an L^2 measure, and with the supremum inside the expectation operator. The proof uses the L^1 error from theorem 1.

Theorem 2 – Strong convergence of v(t) in the L^2 sense

For sufficiently small values of the timestep Δt and for some constant β, we have:

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|V(t)-v(t)\right|^2\right]^{1/2} \le \beta(\Delta t)^{1/4} \tag{A.33}$$

Proof:

Subtracting v(t) from V(t) yields the following equation:

$$V(t)-v(t) = -\lambda\int_0^t\left(V(u)-v_\tau(u)^+\right)du + \eta\int_0^t\left(\sqrt{V(u)}-\sqrt{v_\tau(u)^+}\right)dW_V(u) \tag{A.34}$$

Squaring both sides allows us to deduce the following inequality:

$$\begin{aligned}\left(V(t)-v(t)\right)^2 &\le 2\lambda^2\left(\int_0^t\left(V(u)-v_\tau(u)\right)du\right)^2 + 2\eta^2\left(\int_0^t\left(\sqrt{V(u)}-\sqrt{v_\tau(u)^+}\right)dW_V(u)\right)^2 \\ &\le 2\lambda^2 t\int_0^t\left(V(u)-v_\tau(u)\right)^2du + 2\eta^2\left(\int_0^t\left(\sqrt{V(u)}-\sqrt{v_\tau(u)^+}\right)dW_V(u)\right)^2\end{aligned} \tag{A.35}$$

where we used $(x-y)^2 \le 2x^2 + 2y^2$ as well as $|x-y^+| \le |x-y|$ for $x \ge 0$ in the first step, and the Cauchy-Schwartz inequality in the second step. For any $s \le T$ we then have:

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,s]}\left(V(t)-v(t)\right)^2\right] &\le 2\lambda^2 T\,\mathbb{E}\int_0^s\left(V(u)-v_\tau(u)\right)^2du \\ &\quad + 8\eta^2\,\mathbb{E}\int_0^s\left(\sqrt{V(u)}-\sqrt{v_\tau(u)^+}\right)^2du \\ &\le 2\lambda^2 T\,\mathbb{E}\int_0^s\left(V(u)-v_\tau(u)\right)^2du + 8\eta^2\,\mathbb{E}\int_0^s\left|V(u)-v_\tau(u)\right|du \\ &\le 2\lambda^2 T\,\mathbb{E}\int_0^s\left(V(u)-v(u)\right)^2du + 2\lambda^2 T\,\mathbb{E}\int_0^s\left(v(u)-v_\tau(u)\right)^2du \\ &\quad + 8\eta^2 T\left(\sqrt{U_{cont}(\Delta t)} + \sup_{t\in[0,T]}\mathbb{E}\left[\left|V(t)-v(t)\right|\right]\right) \\ &\le 2\lambda^2 T\int_0^s\mathbb{E}\left[\sup_{t\in[0,u]}\left(V(t)-v(t)\right)^2\right]du + 2\lambda^2T^2U_{cont}(\Delta t) \\ &\quad + 8\eta^2 T\left(\sqrt{U_{cont}(\Delta t)} + \sup_{t\in[0,T]}\mathbb{E}\left[\left|V(t)-v(t)\right|\right]\right)\end{aligned} \tag{A.36}$$

where we used Doob's martingale inequality in the first step, $\left(\sqrt{x}-\sqrt{y^+}\right)^2 \le |x-y|$ in the second, and elements from the proof of theorem 1 in the remaining steps. Applying Gronwall's inequality yet again leads to the final bound:

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left(V(t)-v(t)\right)^2\right]\le 2\lambda^2 T\exp(2\lambda^2T^2)\cdot \left(TU_{cont}(\Delta t)+4\sqrt{U_{cont}(\Delta t)}+\frac{4\eta^2}{\lambda^2}\sup_{t\in[0,T]}\mathbb{E}\left[\left|V(t)-v(t)\right|\right]\right) \tag{A.37}$$

Since the L^1 error tends to zero when the timestep tends to zero, it is also clear that the L^2 error in (A.37) does. The order of convergence is again $O(\Delta t^{1/2})$, which leads to $O(\Delta t^{1/4})$ in (A.33). □

Convergence of plain vanilla and barrier prices

Before we can establish convergence of the sample averages of the payoffs of European calls, puts and barriers, we need to bound the expectation of the supremum of the absolute value of v(t), and prove strong convergence in the L^2 sense for the stock price, stopped by a stopping time.

Lemma 4 – Bounding the expectation of the supremum of |v(t)|

For any $c \ge 0$ and some constant C_{BDG} we have:

$$\mathbb{E}\left[\sup_{t\in[0,T]}|v(t)|\right]\le\left(v(0)+\lambda\bar{v}T+\tfrac{1}{4c}\eta C_{BDG}\right)\exp\left((\lambda+c\eta C_{BDG})T\right) \tag{A.38}$$

Proof:

First of all we note that for any $c > 0$ we can bound the square root function as $\sqrt{x}\le\frac{1}{4c}+cx$. With this useful inequality in hand, we can use the definition of v(t) in (A.3) and the Burkholder-Davis-Gundy inequality to arrive at:

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,s]}|v(t)|\right]&\le v(0)+\lambda\bar{v}T+\lambda\int_0^s\mathbb{E}\left[|v_\tau(u)|\right]du+\eta C_{BDG}\,\mathbb{E}\left[\left(\int_0^s v_\tau(u)^+du\right)^{1/2}\right]\\&\le v(0)+\lambda\bar{v}T+\lambda\int_0^s\mathbb{E}\left[|v_\tau(u)|\right]du+\eta C_{BDG}\,\mathbb{E}\left[\left(\int_0^s|v_\tau(u)|\,du\right)^{1/2}\right]\\&\le v(0)+\lambda\bar{v}T+\tfrac{1}{4c}\eta C_{BDG}+(\lambda+c\eta C_{BDG})\int_0^s\mathbb{E}\left[|v_\tau(u)|\right]du\\&\le v(0)+\lambda\bar{v}T+\tfrac{1}{4c}\eta C_{BDG}+(\lambda+c\eta C_{BDG})\int_0^s\mathbb{E}\left[\sup_{t\in[0,u]}|v(t)|\right]du\end{aligned} \tag{A.39}$$

where C_{BDG} is a universal constant from the Burkholder-Davis-Gundy inequality. Again, Gronwall's inequality yields the desired result. □

We will now introduce the full truncation discretisation for the stock price. Here we only consider a direct discretisation of the stock price, i.e. equation (11), although the proofs could be carried out using a log-discretisation as in (12). The analogue of (A.1) and (A.3) for the stock price are:

$$\begin{aligned}s_{n+1}&=(1+\mu\Delta t)s_n+s_n\sqrt{v_n^+}\,\Delta W_{Sn}\\ s(t)&=s(0)+\mu\int_0^t s_\tau(u)\,du+\int_0^t\sqrt{v_\tau(u)^+}\,s_\tau(u)\,dW_S(u)\end{aligned} \tag{A.40}$$

We are now ready to prove the strong L^2 convergence of the stopped stock price.

Theorem 3 – Strong L^2 convergence of the stopped stock price
If we define the following stopping time for any i, j > 0:

$$\tau_{ij} \equiv \inf\left\{ t \geq 0 : |S(t)| > i \vee |v(t)| > j \right\} \tag{A.41}$$

and $t_{ij} = \min(t,\tau_{ij})$, then for a sufficiently small timestep and some constant β_{ij}, depending on both i and j, we have:

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left| S(t_{ij}) - s(t_{ij}) \right|^2 \right]^{1/2} \leq \beta_{ij} (\Delta t)^{1/4} \tag{A.42}$$

Proof:
Considering the difference of $S(t_{ij})$ and $s(t_{ij})$, we find:

$$\begin{aligned} S(t_{ij}) - s(t_{ij}) = \mu \int_0^{t_{ij}} \left(S(u) - s_\tau(u) \right) du + \int_0^{t_{ij}} S(u) \left(\sqrt{V(u)^+} - \sqrt{v_\tau(u)^+} \right) dW_S(u) \\ + \int_0^{t_{ij}} \sqrt{v_\tau(u)^+} \left(S(u) - s_\tau(u) \right) dW_S(u) \end{aligned} \tag{A.43}$$

Squaring the left-hand side, noting that $(x+y+z)^2 \leq 3x^2 + 3y^2 + 3z^2$ and applying Hölder's inequality for integrals as well as Doob's martingale inequality, yields, for $s \leq T$:

$$\begin{aligned} \mathbb{E} \sup_{t\in[0,s]} \left(S(t_{ij}) - s(t_{ij}) \right)^2 \leq 3\mu^2 T \cdot \mathbb{E} \int_0^{s_{ij}} \left(S(u) - s_\tau(u) \right)^2 du \\ + 12 \mathbb{E} \int_0^{s_{ij}} S(u)^2 \left(\sqrt{V(u)^+} - \sqrt{v_\tau(u)^+} \right)^2 du \\ + 12 \mathbb{E} \int_0^{s_{ij}} v_\tau(u)^+ \left(S(u) - s_\tau(u) \right)^2 du \end{aligned} \tag{A.44}$$

As before we can apply $\left(\sqrt{x} - \sqrt{y^+} \right)^2 \leq |x - y|$ in the second term, and use $y^+ \leq |y|$ in the third term. The stopping time used implies that for all t, $|X(t_{ij})| \leq i$ and $|v(t_{ij})| \leq j$. Combining, we find:

$$\begin{aligned} \mathbb{E} \sup_{t\in[0,s]} \left(S(t_{ij}) - s(t_{ij}) \right)^2 \leq (3\mu^2 T + 12j) \cdot \mathbb{E} \int_0^{s_{ij}} \left(S(u) - s_\tau(u) \right)^2 du \\ + 12 i^2 \cdot \mathbb{E} \int_0^{s_{ij}} \left| V(u) - v_\tau(u) \right| du \end{aligned} \tag{A.45}$$

The remainder of the proof is identical to that of Higham and Mao, and leads to:

$$\mathbb{E}\left[\sup_{t\in[0,s]} \left(S(t_{ij}) - s(t_{ij}) \right)^2 \right] \leq e^{2CT} \left(4CTi^2 \Delta t \left(\mu^2 \Delta t + j \right) + 12 i^2 T \sup_{t\in[0,T]} \mathbb{E}\left[\left| V(t) - v(t) \right| \right] \right) \tag{A.46}$$

with $C = 3\mu^2 T + 12j$. This clearly tends to zero as the timestep does, and the speed with which it approaches zero is of $O(\Delta t^{1/2})$, as follows from theorem 1. □

The final theorem now follows.

Theorem 4 – Convergence of plain vanilla options and barriers
Let us define the following payoffs:

$$V(T) = \left(\gamma(S(T)-K)\right)^+ \cdot 1_{[S(t)\in A, 0\le t\le T]} \quad V_{\Delta t}(T) = \left(\gamma(s(T)-K)\right)^+ \cdot 1_{[s(t)\in A, 0\le t\le T]} \tag{A.47}$$

with $\gamma \in \{-1,1\}$ and A either of the form [0,B] or [B,∞). For any $\varepsilon > 0$ we can find i, j > 0 such that, for sufficiently small timesteps there is a constant β_{ij}, depending on both i and j, such that:

$$\left| \mathbb{E}[V(T)] - \mathbb{E}[V_{\Delta t}(T)] \right| \le \varepsilon + \beta_{ij}\sqrt{\Delta t} \tag{A.48}$$

Proof:
The proof for both the plain vanilla put and the up-and-out barrier call Higham and Mao consider only depends on lemmas 6.1, 6.2 and 6.3. Only lemmas 6.2 and 6.3 depend on the particular discretisation one chooses, and we have proven their equivalents (lemma 4 and theorem 3 respectively) for the full truncation scheme. The proof for the plain vanilla call follows from put-call parity, as Higham and Mao remark. For the barrier case, one can check that the proof for the up-and-out call does not depend on where the initial stock price is located, so that the proof for a down-and-out call follows immediately. We can use the in-out barrier parity to prove convergence for "in" barriers, and round up the proof by noting that for puts the proof carries over directly. The order of convergence follows by inspecting Higham and Mao's proof. □