Convergence of Monte Carlo simulations involving the mean-reverting square root process

Desmond J. Higham

Department of Mathematics, University of Strathclyde, Glasgow G1 1XH, UK

Xuerong Mao

Department of Statistics and Modelling Science, University of Strathclyde, Glasgow G1 1XH, UK

The mean-reverting square root process is a stochastic differential equation (SDE) that has found considerable use as a model for volatility, interest rate, and other financial quantities. The equation has no general, explicit solution, although its transition density can be characterized. For valuing path-dependent options under this model, it is typically quicker and simpler to simulate the SDE directly than to compute with the exact transition density. Because the diffusion coefficient does not satisfy a global Lipschitz condition, there is currently a lack of theory to justify such simulations. We begin by showing that a natural Euler–Maruyama discretization provides qualitatively correct approximations to the first and second moments. We then derive explicitly computable bounds on the strong (pathwise) error over finite time intervals. These bounds imply strong convergence in the limit of the timestep tending to zero. The strong convergence result can be used to justify the method within Monte Carlo simulations that compute the expected payoff of financial products. We spell this out for a bond with interest rate given by the mean-reverting square root process, and for an up-and-out barrier option with asset price governed by the mean-reverting square root process. We also prove convergence for European and up-and-out barrier options under Heston's stochastic volatility model – here the mean-reverting square root process feeds into the asset price dynamics as the squared volatility.

1 The mean-reverting square root process

We consider the mean-reverting square root process in the form of an Itô stochastic differential equation (SDE)

$$dS(t) = \lambda(\mu - S(t))dt + \sigma\sqrt{S(t)}dW(t) \tag{1}$$

Here λ, μ and σ are positive constants and $W(t)$ is a scalar Brownian motion. We assume that the initial condition $S(0)$ is independent of the Brownian motion and has bounded second moment. We also assume that $S(0) \geq 0$ with probability 1. It

is known that a unique strong solution exists for (1), and that non-negativity of the initial data is preserved:

$$S(t) \geq 0 \text{ for all } t \geq 0, \text{ with probability 1} \qquad (2)$$

Further, the solution may attain the value zero only if $\sigma^2 > 2\lambda\mu$. See, for example, Mao (1997, Section 9.2) or Kwok (1998, Section 7.1.5) for more details.

The SDE (1) is widely used in mathematical finance as an alternative to geometric Brownian motion. Most notably, it was proposed by Cox, Ingersoll and Ross (1985) as an interest rate model and it forms the stochastic volatility component of Heston's asset price model (Heston, 1993). The idea of taking the diffusion coefficient to be proportional to the square root of the solution can be traced as least as far back as Cox and Ross (1976).

There are numerous examples in the literature where authors discretize the mean-reverting square root process, typically with an Euler-type scheme. In the finance context, there are two main motivations for such simulations:

❑ using a Monte Carlo approach to compute the expected value of a function of $S(t)$, for example to value a bond or to find the expected payoff of an option (Andersen *et al*, 2002; Andersen and Lund, 1997; Bacinello, 2000; Barone-Adesi and Sorwar, 2002; Broadie and Kaya, 2003; Gkamas, 2001; Zhang and Shu, 2003); and
❑ generating time series in order to test parameter estimation algorithms (Duan, 2003; Fischer *et al*, 2003; Fischer and Roehrl, 2003).

We also note that Clewlow and Strickland (1998), Cyganowski *et al* (2002) and Glasserman (2004) include examples of numerical simulations on SDEs of the form (1) and Misawa (2001) derives a method that applies to a subclass of (1).

The motivation for our work is that the SDE (1) is non-linear and non-Lipschitzian. Thus we cannot appeal to standard convergence theory for numerical simulations, as typified by Theorem 9.6.2 of Kloeden and Platen (1999), to deduce that the numerically computed paths are accurate for small stepsizes. Nor can we appeal to linear stability analysis, such as that in Higham (2000) and Saito and Mitsui (1996), to obtain qualitative information about the behavior of numerical methods over long time intervals. In this work, we address both issues, giving a customized analysis of the simplest and most widely used numerical method. Our results are positive, and they thus justify the type of numerical simulations that are done by researchers and practitioners, as cited above.

A numerical method applied to (1) may break down due to negative values being supplied to the square root function. A natural fix, which we adopt in this work, is to replace the SDE (1) by the equivalent, but computationally safer, problem

$$dS(t) = \lambda(\mu - S(t))dt + \sigma\sqrt{|S(t)|}\, dW(t) \qquad (3)$$

Given a stepsize $\Delta t > 0$, the Euler–Maruyama (EM) method applied to (3) sets $s_0 = S(0)$ and computes approximations $s_n \approx S(t_n)$, where $t_n = n\Delta t$, according to

$$s_{n+1} = s_n(1 - \lambda \Delta t) + \lambda \Delta t \mu + \sigma \sqrt{|s_n|} \Delta W_n \tag{4}$$

where $\Delta W_n = W(t_{n+1}) - W(t_n)$

The next section looks at qualitative behavior. We analyze the ability of the method to reproduce the mean-reverting property of the SDE and to pick up the correct second moment limit. Section 3 deals with the strong error of the method over finite time intervals. We show in sections 4 and 5 how the results can be used to deduce convergence of Monte Carlo simulations for computing the expected payoff from a bond and an up-and-out call option, respectively. These give clear examples of cases where strong (as opposed to weak) convergence is required from a discretization method. In Section 6 we consider coupled equations where the mean-reverting square root process models a stochastic volatility term, and we show that expected payoffs may also be computed accurately.

Although there appears to be no explicit solution for the SDE (1), its transition density can be characterized. Given $S(u)$, for $t > u$ the distribution of $S(t)$ is, within a scaling, non-central chi-square; see, for example, Glasserman (2004, Section 3.4). This gives an alternative means to simulate the process. We study the direct discretization method for a number of reasons:

❏ It is widely used in practice.
❏ In general it is computationally faster than simulating the transition density in cases where the path must be sampled at finely spaced points in order to approximate a path-dependent payoff. Direct discretization requires a single normal sample per step. The alternative of using the exact transition density with the method in Glasserman (2004, Figure 3.5), for example, involves sampling a chi-square random variable. The number of degrees of freedom of the chi-square random variable, and hence the computational expense, depends upon $4\mu\lambda/\sigma^2$ and so is strongly problem-dependent. This argument carries further weight in the case where the mean-reverting square root process is used within Heston's model. Here, Broadie and Kaya (2003) have shown how to simulate from the exact asset price distribution, but for typical parameter values (see their Table 1) an Euler step is many times faster than an exact sample. Hence, while exact simulation will be superior for European-style options and for options whose payoff depends on the asset price at only a widely-spaced, discrete set of times, for fully path-dependent options, where both approaches require the time horizon to be broken into a large number of subintervals, the Euler version remains attractive.
❏ It forms an interesting test case for proving convergence results where there is no global Lipschitz condition for the diffusion term (as mentioned, for example, in Broadie and Kaya (2003) and Glasserman (2004, page 357).

Further, looking to future work, direct discretization adapts easily to changes in the model, such as time-dependency of λ, μ and σ, where the analytical transition density is not available.

We also note that our convergence results in Section 4, Section 5 and (for

independent $W_1(t)$ and $W_2(t)$) Section 6 apply automatically to the case where the transition density for the mean-reverting square root process is simulated exactly, rather than via discretization; here the discretization errors referred to in Corollaries 3.1 and 3.2 become zero for all Δt.

To conclude this section, we mention that although this work is presented from a numerical simulation viewpoint, it could also be regarded as a contribution to the literature on diffusion limits of discrete models. Nelson (1990) showed that a range of Arch models converge in distribution to SDEs, and the particular case of a mean-reverting square root process has been identified in Ishida and Engle (2002) as the limit for a CEVGarch(1,1) process. Our work deals with strong convergence, in L^1 and L^2 senses, of the discrete process (4) to the SDE (3) as $\Delta t \to 0$.

2 First and second moment stability

We begin this section by stating how the first and second moments of the SDE behave.

THEOREM 2.1 *For SDE (3),* $\mathbb{E}[S(t) - \mu] = e^{-\lambda t}(\mathbb{E}(S_0) - \mu$, *so that*

$$\lim_{t \to \infty} \mathbb{E}[S(t)] = \mu \tag{5}$$

and

$$\mathbb{E}\left[S(t)^2 - \mu^2 - \frac{\sigma^2 \mu}{2\lambda}\right] = \left(2\mu + \frac{\sigma^2}{\lambda}\right)\left(\mathbb{E}(S_0) - \mu\right)e^{-\lambda t}$$
$$+ \left(\mathbb{E}(S_0^2) + \left(\mu + \frac{\sigma^2}{2\lambda}\right)(\mu - 2\mathbb{E}(S_0))\right)e^{-2\lambda t}$$

so that

$$\lim_{t \to \infty} \mathbb{E}[S(t)^2] = \mu^2 + \frac{\sigma^2 \mu}{2\lambda} \tag{6}$$

PROOF The first moment result follows immediately from taking expectations in (1). The second moment result can be obtained by applying the Itô formula to $S(t)^2$ and taking expectations, using the result for $\mathbb{E}(S(t))$. □

The properties (5) and (6) may be used to estimate the type of stepsize needed to obtain qualitatively correct solutions.

THEOREM 2.2 *For the method (4) we have*

$$\mathbb{E}(s_n)(1 - \lambda \Delta t)^n (\mathbb{E}(s_0) - \mu) + \mu$$

and hence

$$\begin{aligned}
&\text{for } \Delta t < 2/\lambda, && \mathbb{E}(s_n) \to \mu \text{ as } n \to \infty \\
&\text{for } \Delta t = 2/\lambda, && \mathbb{E}(s_n) = (-1)^n \mathbb{E}(s_0) + \left((-1)^{n+1} + 1\right)\mu, \\
&\text{for } \Delta t > 2/\lambda, && \left|\mathbb{E}(s_n)\right| \to \infty \text{ as } n \to \infty
\end{aligned}$$

PROOF The proof follows trivially after taking expected values in (4). ☐

Theorem 2.2 shows that in the $t_n \to \infty$ limit, we recover the correct mean if and only if the stepsize satisfies the constraint $\Delta t < 2/\lambda$. This constraint corresponds precisely to the linear stability constraint for deterministic problems (Hairer and Wanner, 1996). We now consider the second moment behavior in this stable regime.

THEOREM 2.3 *For the method (4) with $\Delta t < 2/\lambda$ we have*

$$\liminf_{n\to\infty} \mathbb{E}\left(s_n^2\right) \geq \frac{\mu^2 + \frac{\sigma^2\mu}{2\lambda} - \frac{\lambda\Delta t\mu^2}{2}}{1 - \frac{\lambda\Delta t}{2}} =: L(\lambda,\mu,\sigma;\Delta t)$$

where this limiting lower bound thus satisfies

$$L(\lambda,\mu\,\sigma;\Delta t) = \mu^2 + \frac{\sigma^2\mu}{2\lambda} + O(\Delta t), \text{ as } \Delta t \to 0$$

Further, given any $\alpha > 0$, if $(1 - \lambda\Delta t)^2 + \Delta t\sigma^2\alpha/2 < 1$, then

$$\limsup_{n\to\infty} \mathbb{E}\left(s_n^2\right) \leq \frac{\mu^2 + \frac{\sigma^2\mu}{4\lambda\alpha} - \frac{\lambda\Delta t\mu^2}{2}}{1 - \frac{\lambda\Delta t}{2} - \frac{\sigma^2\alpha}{4\lambda}} =: U(\lambda,\mu,\sigma;\Delta t)$$

where this limiting upper bound thus satisfies

$$U(\lambda,\mu,\sigma;\Delta t) = \frac{\mu^2 + \frac{\sigma^2\mu}{4\lambda\alpha}}{1 - \frac{\sigma^2\alpha}{4\lambda}} + O(\Delta t), \text{ as } \Delta t \to 0$$

PROOF Our proof makes use of the readily established fact that a real-valued sequence of the form

$$y_{k+1} = ay_k + b + cr^k \tag{7}$$

where a, b, c, r are constants with $|a| < 1$ and $|r| < 1$, satisfies

$$\lim_{n\to\infty} y_n = \frac{b}{1-a} \tag{8}$$

Squaring and taking expected values in (4) gives

$$\mathbb{E}\left(s_{n+1}^2\right) = (1 - \lambda\Delta t)^2 \mathbb{E}\left(s_n^2\right) + 2\lambda\Delta t\mu(1 - \lambda\Delta t)\mathbb{E}\left(s_n\right)$$
$$+ \lambda^2\Delta t^2\mu^2 + \sigma^2\Delta t\,\mathbb{E}\left(|s_n|\right) \tag{9}$$

Now we replace $\mathbb{E}(|s_n|)$ in (9) by $\mathbb{E}(s_n)$ to obtain the sequence $\{z_n\}$ with $z_0 = \mathbb{E}(s_0^2)$ and

$$z_{n+1} = (1 - \lambda \Delta t)^2 z_n + 2\lambda \Delta t \mu (1 - \lambda \Delta t) \mathbb{E}(s_n)$$
$$+ \lambda^2 \Delta t^2 \mu^2 + \sigma^2 \Delta t \, \mathbb{E}(s_n)$$

for which $\mathbb{E}(s_n^2) \geq z_n$ for all n. Inserting the expression for $\mathbb{E}(s_n)$ from Theorem 2.2 we obtain a sequence of the form (7) with

$$a = (1 - \lambda \Delta t)^2, \tag{10}$$
$$r = 1 - \lambda \Delta t, \tag{11}$$
$$b = 2\lambda \Delta t \mu^2 (1 - \lambda \Delta t) + \lambda^2 \Delta t^2 \mu^2 + \sigma^2 \Delta t \mu \tag{12}$$

Using the expression (8) for the limit then gives the lim inf bound for $\mathbb{E}(s_n^2)$.

For an upper bound, we note that for any $\alpha > 0$

$$\Delta t \, \mathbb{E}(|s_n|) \leq \Delta t \sqrt{\mathbb{E}(s_n^2)} = 2\sqrt{\frac{\Delta t}{2\alpha}} \sqrt{\frac{\Delta t \alpha \mathbb{E}(s_n^2)}{2}} \leq \frac{\Delta t}{2\alpha} + \frac{\Delta t \alpha \mathbb{E}(s_n^2)}{2}$$

Hence, from (9), letting $\hat{z}_0 = E(s_0^2)$ and

$$\hat{z}_{n+1} =$$

$$(1 - \lambda \Delta t)^2 \hat{z}_n + 2\lambda \Delta t \mu (1 - \lambda \Delta t) \mathbb{E}(s_n) + \lambda^2 \Delta t^2 \mu^2 + \sigma^2 \left(\frac{\Delta t}{2\alpha} + \frac{\Delta t \alpha \hat{z}_n}{2} \right)$$

we have $\mathbb{E}(s_n^2) \leq \hat{z}_n$ for all n. Inserting the expression for $\mathbb{E}(s_n)$ from Theorem 2.2 we obtain a sequence of the form (7) with

$$a = (1 - \lambda \Delta t)^2 + \frac{\sigma^2 \alpha \Delta t}{2}, \tag{13}$$
$$r = 1 - \lambda \Delta t, \tag{14}$$
$$b = 2\lambda \Delta t \mu^2 (1 - \lambda \Delta t) + \lambda^2 \Delta t^2 \mu^2 + \Delta t \frac{\sigma^2}{2\alpha} \tag{15}$$

Using the expression (8) for the limit then gives the lim sup bound for $\mathbb{E}(s_n^2)$. $\square$

Theorem 2.3 shows that in the stable regime $\Delta t < 2/\lambda$ the method has a bounded second moment. The lower bound $L(\lambda, \mu, \sigma; \Delta t)$ is close to the exact limiting second moment in Theorem 2.1 for small Δt. The bounds are tight when the volatility parameter σ is not too large. For example, if $\lambda > \sigma^2/(8\mu)$ then we may take $\alpha = 1/(2\mu)$ for all small Δt to get

$$U(\lambda, \mu, \sigma; \Delta t) = \frac{\mu^2 + \dfrac{\sigma^2 \mu}{2\lambda \alpha}}{1 - \dfrac{\sigma^2}{8\lambda \mu}} + O(\Delta t), \quad \text{as} \quad \Delta t \to 0$$

3 Strong convergence

This section deals with the regime where the integration interval, $[0, T]$, is fixed. We consider the error in the numerical solution, measured in strong L^1 and L^2 senses.

In our convergence analysis we find it convenient to work with the continuous-time approximation $s(t)$ defined by

$$s(t) := s_n + (t - t_n)\lambda(\mu - s_n) + \sigma\sqrt{|s_n|}\big(W(t) - W(t_n)\big),$$
$$\text{for } t \in [t_n, t_{n+1}) \tag{16}$$

A more useful characterization of $s(t)$ for the purpose of analysis is

$$s(t) := s_0 + \int_0^t \lambda(\mu - \bar{s}(r))\mathrm{d}r + \sigma\int_0^t \sqrt{|\bar{s}(r)|}\,\mathrm{d}W(r) \tag{17}$$

where the "step function" $\bar{s}(t)$ is defined by

$$\bar{s}(t) := s_n, \text{ for } t \in [t_n, t_{n+1}) \tag{18}$$

Note that $s(t)$ and $\bar{s}(t)$ coincide with the discrete solution at the gridpoints; $\bar{s}(t_n) = s(t_n) = s_n$. In order to obtain positive results about the ability of the discrete method (4) to approximate the true solution at the discrete points $\{t_n\}$, we will prove positive results about the ability of $s(t)$ to approximate $S(t)$. A similar approach was taken in Higham *et al* (2002) for a different class of SDEs.

Our first step is to bound the second moment of the discrete numerical solution over finite time.

LEMMA 3.1 *For any* $\Delta t < 1/2\lambda,$

$$\mathbb{E}\big(s_k^2\big) \le \mathbb{E}\big(s_0^2\big) + 3\mu^2 + \frac{\sigma^4}{2\lambda^2} + 2\lambda\mu T\big|\mathbb{E}(s_0 - \mu)\big| =: C\big(\lambda, \mu, \sigma; \mathbb{E}(s_0), \mathbb{E}(s_0^2)\big),$$
$$0 \le k\Delta t \le T$$

PROOF Following the proof of Theorem 2.3 we have $\mathbb{E}(s_k^2) \le \hat{z}_k$, where $\hat{z}_k$ satisfies a recurrence of the form (7) with (13), (14), (15) and

$$c = 2\lambda\Delta t\mu(1 - \lambda\Delta t)\big(\mathbb{E}(s_0) - \mu\big)$$

Choosing $\alpha = 2\lambda/\sigma^2$, we have $a = 1 - \lambda\Delta t + \lambda^2\Delta t^2$ and hence $3/4 < a < 1$. Solving the recurrence, we find that

$$\hat{z}_n \le \hat{z}_0 + \frac{b}{1-a} + |c|\big(a^{n-1} + ra^{n-2} + r^2 a^{n-3} + \cdots + r^{n-1}\big)$$
$$\le \hat{z}_0 + \frac{b}{1-a} + |c|\,n \tag{19}$$

Now

$$\frac{b}{1-a} = \frac{2\lambda\Delta t\mu^2(1-\lambda\Delta t) + \frac{\sigma^4\Delta t}{4\lambda} + \lambda^2\Delta t^2\mu^2}{\lambda\Delta t(1-\lambda\Delta t)}$$

$$= 2\mu^2 + \frac{\sigma^4}{(1-\lambda\Delta t)4\lambda^2} + \frac{\lambda\Delta t\mu^2}{1-\lambda\Delta t}$$

$$\leq 3\mu^2 + \frac{\sigma^4}{2\lambda^2} \tag{20}$$

Also,

$$\frac{|c|}{\Delta t}T = 2\lambda\mu T(1-\lambda\Delta t)\big|\mathbb{E}(s_0) - \mu\big| \leq 2\lambda\mu T\big|\mathbb{E}(s_0) - \mu\big| \tag{21}$$

The result follows by inserting (20) and (21) into (19). □

Next, we derive a bound for the L^2 difference between the two continuous-time approximations.

LEMMA 3.2 *For $\Delta t < 1/2\lambda$,*

$$\sup_{0\leq t\leq T} \mathbb{E}\left((s(t) - \overline{s}(t))^2\right)$$

$$\leq \Delta t\left[\frac{1}{2}\lambda\left(\mu^2 + 2\mu\left(\big|\mathbb{E}(s_0)\big| + 2\mu\right)\right) + C\left(\lambda, \mu, \sigma; \mathbb{E}(s_0), \mathbb{E}(s_0^2)\right)\right.$$

$$\left. + \sigma^2\sqrt{C\left(\lambda, \mu, \sigma; \mathbb{E}(s_0), \mathbb{E}(s_0^2)\right)}\right]$$

$$=: \Delta t\, D\left(\lambda, \mu, \sigma; \mathbb{E}(s_0), \mathbb{E}(s_0^2)\right)$$

where C is defined in Lemma 3.1.

PROOF Suppose $t \in [k\Delta t, (k+1)\Delta t)$. Then

$$\mathbb{E}\left((s(t) - \overline{s}(t))^2\right) = \mathbb{E}\left[\left((t - t_k)\lambda(\mu - s_k) + \sigma\sqrt{|(s_k)|}(w(t) - w(t_k))\right)^2\right]$$

$$= (t - t_k)^2\lambda^2\mathbb{E}\left((\mu - s_k)^2\right) + \sigma^2\mathbb{E}\left(|(s_k)|\right)(t - t_k)$$

$$\leq \Delta t^2\lambda^2\mathbb{E}\left((\mu - s_k)^2\right) + \sigma^2\Delta t\sqrt{\mathbb{E}(s_k^2)}$$

$$\leq \Delta t\left[\frac{1}{2}\lambda\left(\mu^2 + 2\mu\big|\mathbb{E}(s_k)\big| + \mathbb{E}(s_k^2)\right) + \sigma^2\sqrt{\mathbb{E}(s_k^2)}\right]$$

The result follows using Lemma 3.1 and the bound $\big|\mathbb{E}(s_k)\big| \leq \big|\mathbb{E}(s_0)\big| + 2\mu$ that follows from Theorem 2.2. □

Our first main result is an explicit bound for the strong L^1 error.

THEOREM 3.1 *For $\Delta t < 1/2\lambda$ and any integer $k \geq 1$ we have*

$$\sup_{0 \leq t \leq T} \mathbb{E}\left|S(t) - s(t)\right| \leq e^{\lambda T}\left[e^{-k(k-1)/2} + \frac{\sigma^2 T}{k} + \left(\frac{\sigma^2 T}{k e^{-k(k+1)/2}} + \lambda T\right)\sqrt{D\Delta t}\right]$$

where $D = D(\lambda, \mu, \sigma; \mathbb{E}(s_0), \mathbb{E}(s_0^2))$ is defined in Lemma 3.2.

PROOF The first step is to construct a sequence of C^2 smooth functions, $\psi_k(x)$, with uniformly bounded first derivative, that approximate $|x|$. As k increases, the approximation quality improves, at the expense of a larger second derivative bound.

Let $a_0 = 1$ and $a_k = e^{-k(k+1)/2}$ for $k \geq 1$, so that

$$\int_{a_k}^{a_{k-1}} \frac{du}{u} = k$$

For each $k \geq 1$, there clearly exists a continuous function $\psi_k(u)$ with support in (a_k, a_{k-1}) such that

$$0 \leq \psi_k(u) \leq \frac{2}{ku} \quad \text{for } a_k < u < a_{k-1}$$

and

$$\int_{a_k}^{a_{k-1}} \psi_k(u)du = 1$$

Define

$$\phi_k(x) = \int_0^{|x|} dy \int_0^y \psi_k(u)du$$

Then

$$\phi_k \in C^2(\mathbb{R}, \mathbb{R}), \quad \phi_k(0) = 0,$$
$$\left|\phi_k'(x)\right| \leq 1, \quad \text{for all } x \in \mathbb{R}, \tag{22}$$

$$\left|\phi_k''(x)\right| \begin{cases} \leq \frac{2}{k|x|}, & \text{for } a_k < |x| < a_{k-1} \\ = 0, & \text{otherwise} \end{cases} \tag{23}$$

and, moreover,

$$|x| - a_{k-1} \leq \phi_k(x) \leq |x|, \quad \text{for all } x \in \mathbb{R}, \tag{24}$$

Now, note that

$$S(t) - s(t) = -\lambda \int_0^t \left(S(t) - \bar{s}(t)\right)dr + \sigma \int_0^t \left(\sqrt{|S(r)|} - \sqrt{|\bar{s}(r)|}\right)dW(r)$$

Applying the Itô formula gives

$$\mathbb{E}\phi_k(S(t)-s(t)) = -\lambda\mathbb{E}\int_0^t \phi_k'\big(S(r)-s(r)\big)\big(S(r)-\bar{s}(r)\big)\mathrm{d}r$$

$$+\frac{1}{2}\sigma^2\mathbb{E}\int_0^t \phi_k''\big(S(r)-s(r)\big)\Big(\sqrt{|S(r)|}-\sqrt{|\bar{s}(r)|}\Big)^2\,\mathrm{d}r$$

$$\leq \lambda\int_0^t \mathbb{E}\big|S(r)-\bar{s}(r)\big|\mathrm{d}r + \frac{1}{2}\sigma^2 I(t)$$

where, using (22), (23) and Lemma 3.2,

$$I(t) = \mathbb{E}\int_0^t \phi_k''\big(S(r)-s(r)\big)\Big(\sqrt{|S(r)|}-\sqrt{|\bar{s}(r)|}\Big)^2\,\mathrm{d}r$$

$$\leq \mathbb{E}\int_0^t \phi_k''\big(S(r)-s(r)\big)\big|S(r)-\bar{s}(r)\big|\mathrm{d}r$$

$$\leq \mathbb{E}\int_0^t \phi_k''\big(S(r)-s(r)\big)\big|S(r)-s(r)\big|\mathrm{d}r$$

$$+\mathbb{E}\int_0^t \phi_k''\big(S(r)-s(r)\big)\big|s(r)-\bar{s}(r)\big|\mathrm{d}r$$

$$\leq \mathbb{E}\int_0^t \frac{2}{k}\mathbf{1}_{\{a_k<|S(r)-s(r)|<a_{k-1}\}}\,\mathrm{d}r + \mathbb{E}\int_0^t \frac{2}{ka_k}\big|s(r)-\bar{s}(r)\big|\mathrm{d}r$$

$$\leq \frac{2T}{k} + \int_0^t \frac{2}{ka_k}\,\mathbb{E}\big|s(r)-\bar{s}(r)\big|\mathrm{d}r$$

$$\leq \frac{2T}{k} + \frac{2T}{ka_k}\sqrt{D\Delta t}$$

Using Lemma 3.2 again, we obtain

$$\mathbb{E}\phi_k\big(S(t)-s(t)\big) \leq \lambda\int_0^t \mathbb{E}\big|S(r)-s(r)\big|\mathrm{d}r + \lambda\int_0^t \mathbb{E}\big|s(r)-\bar{s}(r)\big|\mathrm{d}r$$

$$+\frac{\sigma^2 T}{k} + \frac{\sigma^2 T}{ka_k}\sqrt{D\Delta t}$$

$$\leq \lambda\int_0^t \mathbb{E}\big|S(r)-s(r)\big|\mathrm{d}r + \frac{\sigma^2 T}{k} + \left(\frac{\sigma^2 T}{ka_k}+\lambda T\right)\sqrt{D\Delta t}$$

But, from (24),

$$\mathbb{E}\phi_k\big(S(t)-s(t)\big) \geq \mathbb{E}\big|S(t)-s(t)\big| - a_{k-1}$$

So

$$\mathbb{E}\big|S(t)-s(t)\big| \leq a_{k-1} + \frac{\sigma^2 T}{k}$$

$$+\left(\frac{\sigma^2 T}{ka_k}+\lambda T\right)\sqrt{D\Delta t} + \lambda\int_0^t \mathbb{E}\big|S(r)-s(r)\big|\mathrm{d}r, \quad 0\leq t\leq T$$

Applying the Gronwall inequality (see, for example, Mao, 1997, Chapter 1) gives the required bound. □

Theorem 3.1 provides an explicit, computable bound for the L^1 error. To illustrate this, Table 1 shows the bound in the case $\lambda = 0.1$, $\mu = 1$, $\sigma = 0.1$, $\mathbb{E}(s_0) = 1$ and $\mathbb{E}(s_0^2) = 1$ for a range of T and Δt values. In each case, we optimized over k. (Typically, $k = 3$ gives the minimum.) We note that sampling errors from a Monte Carlo simulation usually restrict accuracy to one or two digits, irrespective of the timestepping accuracy. So, for these parameters, our bound proves that the numerical method offers sufficient accuracy for a practical choice of Δt.

TABLE 1 Error bound for $\sup_{0 \leq t \leq T} \mathbb{E}|S(t) - s(t)|$ from Theorem 3.1 for $\lambda = 0.1$, $\mu = 1$, $\sigma = 0.1$, $\mathbb{E}(s_0) = 1$, $\mathbb{E}(s_0^2) = 1$ and various T and Δt (optimized over k).

	$T = 1/12$	$T = 3/12$	$T = 6/12$	$T = 9/12$	$T = 12/12$
$\Delta t = 10^{-1}$	0.08	0.14	0.24	0.34	0.44
$\Delta t = 10^{-2}$	0.06	0.08	0.11	0.14	0.18
$\Delta t = 10^{-3}$	0.05	0.06	0.07	0.08	0.10
$\Delta t = 10^{-4}$	0.04	0.05	0.06	0.07	0.07

Corollary 3.1 below shows that in addition to providing an explicit bound, Theorem 3.1 implies strong L^1 convergence as $\Delta t \to 0$.

COROLLARY 3.1

$$\lim_{\Delta t \to 0} \sup_{0 \leq t \leq T} \mathbb{E}\left|S(t) - s(t)\right| = 0$$

PROOF Given any $\varepsilon > 0$, we may choose $k \geq 1$ such that

$$e^{\lambda T}\left[a_{k-1} + \frac{\sigma^2 T}{k}\right] < \frac{1}{2}\varepsilon$$

and then choose $\Delta t > 0$ such that

$$e^{\lambda T}\left[\frac{\sigma^2 T}{k\,a_k} + \lambda T\right]\sqrt{D\,\Delta t} < \frac{1}{2}\varepsilon$$

From Theorem 3.1, this ensures that $\sup_{0 \leq t \leq T} \mathbb{E}|S(t) - s(t)| < \varepsilon$, as required. □

Next, we derive a bound for a stronger form of the error. This version uses an L^2 measure and places the supremum over time inside the expectation operation. The resulting bound, in Theorem 3.2 below, involves the L^1 error which is explicitly bounded in Theorem 3.1 and hence is also computable. Corollary 3.2 shows that convergence as $\Delta t \to 0$ also follows.

THEOREM 3.2 *For $\Delta t < 1/2\lambda$ we have*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left(S(t)-s(t)\right)^2\right]$$

$$\leq e^{2\lambda^2 T^2}\left(2\lambda^2 T^2 D\Delta t+8\sigma^2 T\sqrt{D\Delta t}+8\sigma^2 T\sup_{0\leq t\leq T}\mathbb{E}\left|S(t)-s(t)\right|\right)$$

where $D=D(\lambda,\mu,\sigma;\mathbb{E}(s_0),\mathbb{E}(s_0^2))$ is defined in Lemma 3.2.

PROOF For any $0\leq t\leq T$, using the Cauchy–Schwarz inequality we have

$$\left(S(t)-s(t)\right)^2\leq 2\lambda^2 T\int_0^t\left(S(r)-\overline{s}(r)\right)^2 dr$$

$$+2\sigma^2\left(\int_0^t\left(\sqrt{|S(r)|}-\sqrt{|\overline{s}(r)|}\right)dW(r)\right)^2$$

Applying the Doob martingale inequality (see, for example, Mao, 1997, Chapter 1), we find that for any $t_1\in[0,T]$

$$\mathbb{E}\left[\sup_{0\leq t\leq t_1}\left(S(t)-s(t)\right)^2\right]$$

$$\leq 2\lambda^2 T^2\mathbb{E}\int_0^{t_1}\left(S(r)-\overline{s}(r)\right)^2 dr+8\sigma^2\mathbb{E}\int_0^{t_1}\left(\sqrt{|S(r)|}-\sqrt{|\overline{s}(r)|}\right)^2 dr$$

$$\leq 2\lambda^2 T\int_0^{t_1}\left(\mathbb{E}\left(S(r)-s(r)\right)^2+\mathbb{E}\left(s(r)-\overline{s}(r)\right)^2\right)dr$$

$$+8\sigma^2\mathbb{E}\int_0^{t_1}\left|S(r)-\overline{s}(r)\right|dr$$

$$\leq 2\lambda^2 T\int_0^{t_1}\mathbb{E}\left(S(r)-s(r)\right)^2 dr+2\lambda^2 T^2 D\Delta t$$

$$+8\sigma^2\int_0^T\mathbb{E}\left|S(r)-s(r)\right|dr+8\sigma^2\int_0^T\mathbb{E}\left|s(r)-\overline{s}(r)\right|dr$$

$$\leq 2\lambda^2 T\int_0^{t_1}\mathbb{E}\left[\sup_{0\leq t\leq s}\left(S(t)-s(t)\right)^2\right]ds+2\lambda^2 T^2 D\Delta t$$

$$+8\sigma^2 T\sqrt{D\Delta t}+8\sigma^2 T\sup_{0\leq t\leq T}\mathbb{E}\left|S(t)-s(t)\right|$$

An application of the Gronwall inequality completes the proof.

COROLLARY 3.2

$$\lim_{\Delta t\to 0}\mathbb{E}\left[\sup_{0\leq t\leq T}\left(S(t)-s(t)\right)^2\right]=0$$

PROOF The proof follows immediately from Corollary 3.1 and Theorem 3.2.   □

4 A bond

In the case where $S(t)$ in (1) models short-term interest rate dynamics, it is pertinent to consider the expected payoff

$$\text{Bond} := \mathbb{E} \exp\left(-\int_0^T S(t)\mathrm{d}t\right) \tag{25}$$

from a bond. A natural approximation based on (4) is

$$\overline{\text{Bond}}_{\Delta t} := \mathbb{E} \exp\left(-\Delta t \sum_{n=0}^{N-1} |s_n|\right) \tag{26}$$

where $N\Delta t = T$; see, for example, Glasserman (2004, Section 6.2.3). It is convenient to rewrite this as

$$\overline{\text{Bond}}_{\Delta t} = \mathbb{E} \exp\left(-\int_0^T |\bar{s}(t)|\mathrm{d}t\right)$$

using the step function $\bar{s}(t)$ in (18). The following result shows that the strong convergence of the SDE approximation confers convergence in this scenario.

THEOREM 4.1 *In the notation above,*

$$\lim_{\Delta t \to 0} \left| \text{Bond} - \overline{\text{Bond}}_{\Delta t} \right| = 0$$

PROOF Using $e^{-|x|} - e^{-|y|} \le |x-y|$ and the non-negativity of $S(t)$, we have

$$\left|\text{Bond} - \overline{\text{Bond}}_{\Delta t}\right| = \mathbb{E}\left[\exp\left(-\int_0^T S(t)\mathrm{d}t\right) - \exp\left(-\int_0^T |\bar{s}(t)|\mathrm{d}t\right)\right]$$

$$\le \mathbb{E}\left|\int_0^T S(t)\mathrm{d}t - \int_0^T \bar{s}(t)\mathrm{d}t\right|$$

$$\le \mathbb{E}\int_0^T \left| S(t) - |\bar{s}(t)| \right|\mathrm{d}t$$

$$\le \mathbb{E}\int_0^T \left| S(t) - \bar{s}(t) \right|\mathrm{d}t$$

$$\le T \sup_{[0,T]} \mathbb{E}\left| S(t) - \bar{s}(t) \right|$$

But Lemma 3.2 and Corollary 3.2 imply that $\sup_{[0,T]} \mathbb{E}|S(t) - \bar{s}(t)| \to 0$, completing the proof.   □

5 A path-dependent option

We now consider the case where the mean-reverting square root process (1) models a financial quantity, such as the short-term interest rate, on which an option has been written (see, for example, Hull, 2000, Section 21.5). In this case the expected payoff from the option is of relevance.

We consider an up-and-out call option, which, at expiry time T, pays the European value if $S(t)$ never exceeded the fixed barrier, B, and pays zero otherwise. We suppose that the expected payoff is computed from a Monte Carlo simulation based on the method (4). Here, using the discrete numerical solution to approximate the true path gives rise to two distinct sources of error:

❑ a discretization error due to the fact that the path is not followed exactly – the numerical solution may cross the barrier at time t_n when the true solution stays below, or vice versa; and

❑ a discretization error due to the fact that the path is only approximated at discrete time points – for example, the true path may cross the barrier and then return within the interval (t_n, t_{n+1}).

The following theorem uses the strong convergence property to show that the expected payoff from the numerical method converges to the correct expected payoff as $\Delta t \to 0$. Note that using the step function $\bar{s}(t)$ in (28) is equivalent to using the discrete solution.

THEOREM 5.1 *For the mean-reverting square root process (1) and numerical method (18), define*

$$V := \mathbb{E}\left[(S(T) - E)^+ \mathbf{1}_{\{0 \le S(t) \le B,\ 0 \le t \le T\}} \right] \tag{27}$$

$$\bar{V}_{\Delta t} := \mathbb{E}\left[(\bar{s}(T) - E)^+ \mathbf{1}_{\{0 \le \bar{s}(t) \le B,\ 0 \le t \le T\}} \right] \tag{28}$$

where the exercise price, E, and barrier, B, are constant. Then

$$\lim_{\Delta t \to 0} \left| V - \bar{V}_{\Delta t} \right| = 0$$

PROOF Let $A := \{0 \le S(t) \le B,\ 0 \le t \le T\}$ and $\bar{A}_{\Delta t} := \{0 \le \bar{s}(t) \le B,\ 0 \le t \le T\}$. Making use of the inequality

$$\left| (S(T) - E)^+ - (\bar{s}(T) - E)^+ \right| \le \left| S(T) - \bar{s}(t) \right|$$

we have

$$\left| V - \bar{V}_{\Delta t} \right| \le \mathbb{E} \left| (S(T) - E)^+ \mathbf{1}_A - (\bar{s}(T) - E)^+ \mathbf{1}_{\bar{A}_{\Delta t}} \right|$$

$$\le \mathbb{E} \left(\left| (S(T) - E)^+ - (\bar{s}(T) - E)^+ \right| \mathbf{1}_{A \cap \bar{A}_{\Delta t}} \right)$$

$$+ \mathbb{E} \left((S(T) - E)^+ \mathbf{1}_{A \cap \bar{A}_{\Delta t}^c} \right) + \mathbb{E} \left((\bar{s}(T) - E)^+ \mathbf{1}_{A^c \cap \bar{A}_{\Delta t}} \right)$$

$$\leq \mathbb{E}\left(\left|S(T)-\bar{s}(T)\right|\mathbf{1}_{A\cap\bar{A}_{\Delta t}}\right)+(B-E)\mathbb{P}\left(A\cap\bar{A}_{\Delta t}^{c}\right)$$
$$+(B-E)\mathbb{P}\left(A^{c}\cap\bar{A}_{\Delta t}\right)$$

Now, from Corollary 3.1, we have $\lim_{\Delta t\to 0}\mathbb{E}(\left|S(T)-\bar{s}(T)\right|)=0$. Hence, our proof is complete if we can show that

$$\lim_{\Delta t\to 0}\mathbb{P}\left(A\cap\bar{A}_{\Delta t}^{c}\right)=0 \tag{29}$$

and

$$\lim_{\Delta t\to 0}\mathbb{P}\left(A^{c}\cap\bar{A}_{\Delta t}\right)=0 \tag{30}$$

For any sufficiently small δ, we have

$$A=\left\{\sup_{0\leq t\leq T}S(t)\leq B\right\}$$
$$=\left\{\sup_{0\leq t\leq T}S(t)\leq B-\delta\right\}\cup\left\{B-\delta<\sup_{0\leq t\leq T}S(t)\leq B\right\}$$
$$\subseteq\left\{\sup_{0\leq k\Delta t\leq T}S(k\Delta t)\leq B-\delta\right\}\cup\left\{B-\delta<\sup_{0\leq t\leq T}S(t)\leq B\right\}$$
$$=:\bar{A}_{1}\cup\bar{A}_{2}$$

Hence,

$$A\cap\bar{A}_{\Delta t}^{c}\subseteq\left(\bar{A}_{1}\cap\bar{A}_{\Delta t}^{c}\right)\cup\left(\bar{A}_{2}\cap\bar{A}_{\Delta t}^{c}\right)$$
$$\subseteq\left\{\sup_{0\leq k\Delta t\leq T}\left|S(k\Delta t)-\bar{s}(k\Delta t)\right|\geq\delta\right\}\cup\bar{A}_{2}$$

So,

$$\mathbb{P}\left(A\cap\bar{A}_{\Delta t}^{c}\right)\leq\mathbb{P}\left(\sup_{0\leq k\Delta t\leq T}\left|S(k\Delta t)-\bar{s}(k\Delta t)\right|\geq\delta\right)+\mathbb{P}\left(\bar{A}_{2}\right)$$
$$\leq\frac{1}{\delta^{2}}\mathbb{E}\left(\sup_{0\leq k\Delta t\leq T}\left(S(k\Delta t)-\bar{s}(k\Delta t)\right)^{2}\right)+\mathbb{P}\left(\bar{A}_{2}\right)$$

Now, for any $\varepsilon>0$, we may choose δ so small that

$$\mathbb{P}\left(\bar{A}_{2}\right)<\frac{1}{2}\varepsilon$$

and then choose Δt so small that

$$\leq\frac{1}{\delta^{2}}\mathbb{E}\left(\sup_{0\leq k\Delta t\leq T}\left(S(k\Delta t)-\bar{s}(k\Delta t)\right)^{2}\right)<\frac{1}{2}\varepsilon$$

whence

$$\mathbb{P}\left(A\cap\bar{A}_{\Delta t}^{c}\right)<\varepsilon$$

This confirms (29).

Now, for any $\delta > 0$, we write

$$A^c = \left\{ \sup_{0 \le t \le T} S(t) > B \right\}$$

$$= \left\{ \sup_{0 \le t \le T} S(t) > B + \delta \right\} \cup \left\{ \sup_{0 \le t \le T} S(t) \le B + \delta \right\}$$

$$=: \bar{A}_3 \cup \bar{A}_4$$

So

$$\mathbb{P}\left(A^c \cap \bar{A}_{\Delta t}\right) \le \mathbb{P}\left(\bar{A}_3 \cap \bar{A}_{\Delta t}\right) + \mathbb{P}\left(\bar{A}_4 \cap \bar{A}_{\Delta t}\right)$$

$$\le \mathbb{P}\left(\sup_{0 \le t \le T} \left| S(t) - \bar{s}(t) \right| > \delta \right) + \mathbb{P}\left(\bar{A}_4\right) \qquad (31)$$

Define

$$s^*(t) = \sum_{k=0}^{\infty} S(k\,\Delta t) \mathbf{1}_{[k\Delta t, (k+1)\Delta t)}(t), \quad 0 \le t \le T$$

and note that

$$\left\{ \sup_{0 \le t \le T} \left| S(t) - \bar{s}(t) \right| > \delta \right\}$$

$$\subseteq \left\{ \sup_{0 \le t \le T} \left| S(t) - s^*(t) \right| > \tfrac{1}{2}\delta \right\} \cup \left\{ \sup_{0 \le t \le T} \left| s^*(t) - \bar{s}(t) \right| > \tfrac{1}{2}\delta \right\}$$

$$= \left\{ \sup_{0 \le k\Delta t \le T} \sup_{k\Delta t \le t \le (k+1)\Delta t} \left| S(t) - S(k\,\Delta t) \right| > \tfrac{1}{2}\delta \right\}$$

$$+ \left\{ \sup_{0 \le k\Delta t \le T} \left| S(k\,\Delta t) - s(k\,\Delta t) \right| > \tfrac{1}{2}\delta \right\}$$

Thus

$$\mathbb{P}\left\{ \sup_{0 \le t \le T} \left| S(t) - \bar{s}(t) \right| > \delta \right\}$$

$$\le \mathbb{P}\left(\sup_{0 \le k\Delta t \le T} \sup_{k\Delta t \le t \le (k+1)\Delta t} \left| S(t) - S(k\,\Delta t) \right| > \tfrac{1}{2}\delta \right)$$

$$+ \frac{4}{\delta^2} \mathbb{E}\left(\sup_{0 \le k\Delta t \le T} \left(S(k\,\Delta t) - s(k\,\Delta t) \right)^2 \right) \qquad (32)$$

Because $S(t)$ is a continuous process in $t \in [0, T]$, almost every sample path of $S(\cdot)$ is uniformly continuous on $[0, T]$. This immediately implies

$$\lim_{\Delta t \to 0} \mathbb{P}\left(\sup_{0 \le k\Delta t \le T} \sup_{k\Delta t \le t \le (k+1)\Delta t} \left| S(t) - S(k\,\Delta t) \right| > \tfrac{1}{2}\delta \right) = 0$$

We also know from Corollary 3.2 that

$$\lim_{\Delta t \to 0} \mathbb{E}\left[\sup_{0 \leq k \Delta t \leq T} \left(S(k \Delta t) - s(k \Delta t) \right)^2 \right] = 0$$

Hence, from (32), for any $\delta > 0$,

$$\lim_{\Delta t \to 0} \mathbb{P}\left(\sup_{0 \leq t \leq T} \left| S(t) - \bar{s}(t) \right| > \delta \right) = 0$$

Now, recalling the definition of $\bar{A}_4$, we see that for $\varepsilon > 0$ we can find a $\delta > 0$ sufficiently small for $\mathbb{P}(\bar{A}_4) < \frac{1}{2}\varepsilon$ and then choose Δt sufficiently small for $\mathbb{P}(\sup_{0 \leq t \leq T} |S(t) - \bar{s}(t)| > \delta) < \frac{1}{2}\varepsilon$. Substituting this into (31) yields $\mathbb{P}(A^c \cup \bar{A}_{\Delta t}) < \varepsilon$, for sufficiently small Δt, confirming (30).

6 Options under stochastic volatility

In this section we consider the case where the mean-reverting square root process is used to represent volatility. We study the Heston (1993) model

$$dX(t) = \lambda_1 \left(\mu_1 - X(t) \right) dt + \sigma_1 X(t) \sqrt{\left| V(t) \right|} \, dW_1(t) \tag{33}$$

$$dV(t) = \lambda_2 \left(\mu_2 - V(t) \right) dt + \sigma_2 \sqrt{\left| V(t) \right|} \, dW_2(t) \tag{34}$$

Here, $V(t)$ is the squared volatility that feeds into the asset price $X(t)$. The Brownian motions $W_1(t)$ and $W_2(t)$ may be correlated. Because of the application to asset pricing, we make the assumption that $X(0)$ and $V(0)$ are constant and positive. We remark that the results in Theorems 6.1 and 6.2 may be derived in a similar manner for the "additive noise" alternative

$$dX(t) = \lambda_1 \left(\mu_1 - X(t) \right) dt + \sigma_1 \sqrt{\left| V(t) \right|} \, dW_1(t) \tag{35}$$

$$dV(t) = \lambda_2 \left(\mu_2 - V(t) \right) dt + \sigma_2 \sqrt{\left| V(t) \right|} \, dW_2(t) \tag{36}$$

which has been proposed for multi-factor interest rates (Kwok, 1998, Section 7.1.7).

We begin with a lemma showing that the positivity in the initial data is preserved for (33)–(34).

LEMMA 6.1

$$\lim_{k \to \infty} \mathbb{P}\left(k^{-1} \leq X(t) \leq k \ \text{ for all } \ 0 \leq t \leq T \right) = 1 \tag{37}$$

PROOF For sufficiently large k, define the stopping time

$$\rho_k = \inf\left\{ t \geq 0 : X(t) > k \ \text{ or } \ X(t) < \frac{1}{k} \right\}$$

Also define

$$U(x) = x - 1 - \log x, \quad \text{for } x > 0$$

By the Itô formula, we have

$$\mathbb{E}U\big(X(T \wedge \rho_k)\big) = U(X(0)) + \mathbb{E} \int_0^{T \wedge \rho_k} \left[\left[1 - \frac{1}{X(t)} \right] \lambda_1 \left(\mu_1 - X(t) \right) + \frac{\sigma_1^2}{2} |V(t)| \right] dt$$

$$\leq U(X(0)) + \lambda_1 (\mu_1 + 1)T + \frac{\sigma_1^2}{2} \int_0^T \mathbb{E}|V(t)| dt$$

But it is easy to show that

$$\mathbb{E}|V(t)| = \mathbb{E}V(t) = e^{-\lambda_2 t} V(0) + \mu_2 (1 - e^{-\lambda_2 t}) \leq V(0) + \mu_2$$

Thus

$$\mathbb{E}U\big(X(T \wedge \rho_k)\big) \leq U(X(0)) + \lambda_1 (\mu_1 + 1)T + \frac{\sigma_1^2}{2}\big(V(0) + \mu_2\big)T \qquad (38)$$

Now we note that for any $\omega\{\rho_k \leq T\}$, either $X(\rho_k(\omega), \omega) = k$ or $X(\rho_k(\omega), \omega) = 1/k$. Hence,

$$U\big(X(\rho_k(\omega), \omega)\big) \geq (k - 1 - \log k) \wedge (k^{-1} - 1 + \log k)$$

and so, dropping the ω for brevity,

$$\mathbb{E}U\big(X(T \wedge \rho_k)\big) \geq \mathbb{E}\left[\mathbf{1}_{\{\rho_k \leq T\}} U(X(\rho_k)) \right]$$

$$\geq \mathbb{E}\left[\mathbf{1}_{\{\rho_k \leq T\}} \left[(k - 1 - \log k) \wedge (k^{-1} - 1 + \log k) \right] \right]$$

Thus, using (38),

$$\mathbb{P}\big(\rho_k \leq T\big)\left[(k - 1 - \log k) \wedge (k^{-1} - 1 + \log k)\right] \leq U(X(0)) + \lambda_1 (\mu_1 + 1)T$$

$$+ \frac{\sigma_1^2}{2}\big(V(0) + \mu_2\big)T$$

Letting $k \to \infty$ yields

$$\lim_{k \to \infty} \mathbb{P}\big(\rho_k \leq T\big) = 0$$

and the assertion (37) follows.   □

As a by-product, this lemma confirms that $X(t) > 0$ a.s. for all $t \leq 0$. Using the Euler–Maruyama method (4) for the volatility equation (34) gives

$$v_{n+1} = v_n \big(1 - \lambda_2 \Delta t\big) + \lambda_2 \Delta t \mu_2 + \sigma_2 \sqrt{|v_n|} \Delta W_{2,n} \qquad (39)$$

where $\Delta W_{2,n} = W_2(t_{n+1}) - W_2(t_n)$, and using this approximation in the asset equation (33) gives

$$x_{n+1} = x_n \left(1 - \lambda_1 \Delta t\right) + \lambda_1 \Delta t \mu_1 + \sigma_1 x_n \sqrt{\left|v_n\right|} \Delta W_{1,n} \tag{40}$$

where $\Delta W_{1,n} = W_1(t_{n+1}) - W_1(t_n)$. We also use the corresponding step functions

$$\bar{x}(t) := x_n \quad \text{and} \quad \bar{v}(t) := v_n, \quad \text{for } t \in \left[t_n, t_{n+1}\right) \tag{41}$$

to define the continuous-time approximations

$$x(t) := x_0 + \int_0^t \lambda_1 \left(\mu_1 - \bar{x}(r)\right) dr + \sigma_1 \int_0^t \bar{x}(r) \sqrt{\left|\bar{v}(r)\right|} \, dW_1(r) \tag{42}$$

$$v(t) := v_0 + \int_0^t \lambda_2 \left(\mu_2 - \bar{v}(r)\right) dr + \sigma_2 \int_0^t \sqrt{\left|\bar{v}(r)\right|} \, dW_2(r) \tag{43}$$

We now prove two lemmas before giving our convergence results for option valuation.

LEMMA 6.2

$$\mathbb{E}\left(\sup_{0 \le t \le T} \left|v(t)\right|\right) \le \left(v(0) + \lambda_2 \mu_2 T + \frac{9}{2}\sigma_2^2\right) e^{-\left(\lambda_2 + \frac{1}{2}\right)T} \tag{44}$$

PROOF Note that for $0 \le t_1 \le T$,

$$v(t_1) = v(0) + \int_0^{t_1} \lambda_2 \left(\mu_2 - \bar{v}(r)\right) dr + \sigma_2 \int_0^{t_1} \sqrt{\left|\bar{v}(r)\right|} \, dW_2(r)$$

By the Burkholder–Davis–Gundy inequality (Prato and Zabczyk, 1992, Theorem 3.14), we derive that for $0 \le t \le T$,

$$\mathbb{E}\left(\sup_{0 \le t_1 \le t} \left|v(t_1)\right|\right)$$

$$\le v(0) + \lambda_2 \mu_2 T + \lambda_2 \int_0^t \mathbb{E}\left|\bar{v}(r)\right| dr + 3\sigma_2 \mathbb{E}\left[\left(\int_0^t \left|\bar{v}(r)\right| dr\right)^{\frac{1}{2}}\right]$$

$$\le v(0) + \lambda_2 \mu_2 T + \frac{9}{2}\sigma_2^2 + \left(\lambda_2 + \frac{1}{2}\right)\int_0^t \mathbb{E}\left(\sup_{0 \le t_1 \le r} \left|v(t_1)\right|\right) dr$$

An application of the Gronwall inequality yields the assertion (44).   □

LEMMA 6.3 *For any given pair of positive numbers i and j, define the stopping time*

$$\tau_{ij} = \inf\left\{t \ge 0 : \left|X(t)\right| > i \text{ or } \left|v(t)\right| > j\right\} \tag{45}$$

Then

$$\lim_{\Delta t \to 0} \mathbb{E}\left(\sup_{0 \le t \le T} \left(X(t \wedge \tau_{ij}) - x(t \wedge \tau_{ij}) \right)^2 \right) = 0 \tag{46}$$

PROOF For the sake of simplicity, we write $\tau_{ij} = \tau$. For any $0 \le t_1 \le T$, we have

$$X(t_1 \wedge \tau) - x(t_1 \wedge \tau)$$

$$= -\lambda_1 \int_0^{t_1 \wedge \tau} \left(X(r) - \bar{x}(r) \right) dr$$

$$+ \sigma_1 \int_0^{t_1 \wedge \tau} \left(X(r)\sqrt{|V(r)|} - \bar{x}(r)\sqrt{|\bar{v}(r)|} \right) dW_1(r)$$

$$= -\lambda_1 \int_0^{t_1 \wedge \tau} \left(X(r) - \bar{x}(r) \right) dr$$

$$+ \sigma_1 \int_0^{t_1 \wedge \tau} X(r)\left(\sqrt{|V(r)|} - \sqrt{|\bar{v}(r)|} \right) dW_1(r)$$

$$+ \sigma_1 \int_0^{t_1 \wedge \tau} \sqrt{|\bar{v}(r)|}\left(X(r) - \bar{x}(r) \right) dW_1(r)$$

By the Hölder inequality and the Doob martingale inequality, we then derive that for any $0 \le t \le T$,

$$\mathbb{E}\left(\sup_{0 \le t_1 \le T} \left(X(t_1 \wedge \tau) - x(t_1 \wedge \tau) \right)^2 \right)$$

$$\le 3\lambda_1^2 t \mathbb{E} \int_0^{t \wedge \tau} \left(X(r) - \bar{x}(r) \right)^2 dr$$

$$+ 12\sigma_1^2 \mathbb{E} \int_0^{t \wedge \tau} \left(X(r) \right)^2 \left(\sqrt{|V(r)|} - \sqrt{|\bar{v}(r)|} \right)^2 dr$$

$$+ 12\sigma_1^2 \mathbb{E} \int_0^{t \wedge \tau} |\bar{v}(r)|\left(X(r) - \bar{x}(r) \right)^2 dr$$

$$\le \left(3\lambda_1^2 T + 12\sigma_1^2 j \right) \mathbb{E} \int_0^{t \wedge \tau} \left(X(r) - \bar{x}(r) \right)^2 dr$$

$$+ 12\sigma_1^2 i^2 \mathbb{E} \int_0^{t \wedge \tau} |V(r) - \bar{v}(r)| dr$$

$$\le C_1 \mathbb{E} \int_0^{t \wedge \tau} \left(X(r) - \bar{x}(r) \right)^2 dr + C_2(\Delta t) \tag{47}$$

where $C_1 = 3\lambda_1^2 T + 12\sigma_1^2 j$ and

$$C_2(\Delta t) = 12\sigma_1^2 i^2 T \left(\sup_{0 \le t \le T} \mathbb{E}|V(r) - \bar{v}(r)| \right)$$

which tends to zero as $\Delta t \to 0$ by Corollary 3.1. Note that

$$\mathbb{E}\int_0^{t\wedge\tau}\big(X(r)-\overline{x}(r)\big)^2\,\mathrm{d}r$$

$$\leq 2\mathbb{E}\int_0^{t\wedge\tau}\big(X(r)-x(r)\big)^2\,\mathrm{d}r+2\mathbb{E}\int_0^{t\wedge\tau}\big(x(r)-\overline{x}(r)\big)^2\,\mathrm{d}r \qquad (48)$$

But

$$\mathbb{E}\int_0^{t\wedge\tau}\big(X(r)-\overline{x}(r)\big)^2\,\mathrm{d}r\leq\mathbb{E}\int_0^t\big(X(r\wedge\tau)-\overline{x}(r\wedge\tau)\big)^2\,\mathrm{d}r$$

$$\leq\int_0^t\mathbb{E}\bigg(\sup_{0\leq t_1\leq r}\big(X(t_1\wedge\tau)-\overline{x}(t_1\wedge\tau)\big)^2\bigg)\mathrm{d}r \qquad (49)$$

Moreover, for $r\in[0,t\wedge\tau]$, let $[r/\Delta t]$ be the integer part of $r/\Delta t$. By definition,

$$x(r)-\overline{x}(r)=\lambda_1(\mu_1-x_{[r/\Delta t]})\big(r-[r/\Delta t]\big)$$

$$+\sigma_1\sqrt{|v_{[r/\Delta t]}|}\,x_{[r/\Delta t]}\big(W_1(r)-W_1\big([r/\Delta t]\Delta t\big)\big)$$

which yields

$$(x(r)-\overline{x}(r))^2\leq 2\lambda_1^2(\mu_1+i)^2\,2\sigma_1^2\,ji^2\big(W_1(r)-W_1\big([r/\Delta t]\Delta t\big)\big)^2$$

Thus

$$\mathbb{E}\int_0^{t\wedge\tau}\big(x(r)-\overline{x}(r)\big)^2\,\mathrm{d}r$$

$$\leq\mathbb{E}\int_0^{t\wedge\tau}\bigg[2\lambda_1^2(\mu_1+i)^2\Delta t^2+2\sigma_1^2\,ji^2\big(W_1(r)-W_1\big([r/\Delta t]\Delta t\big)\big)^2\bigg]\mathrm{d}r$$

$$\leq\int_0^t\bigg[2\lambda_1^2(\mu_1+i)^2\Delta t^2+2\sigma_1^2\,ji^2\mathbb{E}\big(W_1(r)-W_1\big([r/\Delta t]\Delta t\big)\big)^2\bigg]\mathrm{d}r$$

$$\leq T\bigg[2\lambda_1^2(\mu_1+i)^2\,\Delta t^2+2\sigma_1^2\,ji^2\Delta t\bigg]:=C_3(\Delta t) \qquad (50)$$

Substituting (50) and (49) into (48) and then inserting the resulting inequality into (47) we obtain

$$\mathbb{E}\bigg(\sup_{0\leq t_1\leq t}\big(X(t_1\wedge\tau)-x(t_1\wedge\tau)\big)^2\bigg)$$

$$\leq 2C_1\int_0^t\mathbb{E}\bigg(\sup_{0\leq t_1\leq r}\big(X(t_1\wedge\tau)-\overline{x}(t_1\wedge\tau)\big)^2\bigg)\mathrm{d}r+2C_1C_3(\Delta t)+C_2(\Delta t) \qquad (51)$$

The Gronwall inequality gives

$$\mathbb{E}\bigg(\sup_{0\leq t_1\leq T}\big(X(t_1\wedge\tau)-x(t_1\wedge\tau)\big)^2\bigg)\leq\big[2C_1C_3(\Delta t)+C_2(\Delta t)\big]\mathrm{e}^{2C_1T}$$

The required assertion (46) follows by letting $\Delta t\to 0$. $\qquad\square$

Let us now recall the payoff for the European put option:

$$\Lambda=\mathbb{E}[(E-X(T))^+]$$

where $E > 0$ is the exercise price. Accordingly, the payoff based on the numerical method is

$$\bar{\Lambda}_{\Delta t} = \mathbb{E}\left[\left(E - \bar{x}(T)\right)^+\right]$$

The theorem below shows that the numerical approximation is convergent. We remark that the proof makes use of the fact that the payoff is bounded above as a function of $S(T)$, and it is not clear to us how the proof could be extended to cope with unbounded payoff functions. However, if the appropriately discounted expectations are regarded as option values, then the corresponding call option, whose payoff is unbounded, could be valued from the put via put–call parity (see, for example, Hull, 2000, Section 7.4). We also remark that (Glasserman, 2004, Example 6.2.2) performs a closely related Monte Carlo simulation on the same model (33)–(34). Our result contributes toward filling the gap in the literature that Glasserman identified: namely, that the textbook convergence theory does not apply because "the square-root functions in the model dynamics and the kink in the call option payoff violate the smoothness conditions". A similar observation is made in Broadie and Kaya (2003).

THEOREM 6.1 *In the notation above,*

$$\lim_{\Delta t \to 0}\left|\Lambda - \bar{\Lambda}_{\Delta t}\right| = 0 \tag{52}$$

Proof For $i, j > 0$, set

$$A_i = \left\{X(t) \le i,\ 0 \le t \le T\right\} \quad \text{and} \quad B_j = \left\{\left|v(t)\right| \le j,\ 0 \le t \le T\right\}$$

Given any $\varepsilon > 0$, by Lemmas 6.1 and 6.2 we can find i, j sufficiently large for

$$P\left(A_i^c \cup B_j^c\right) \le \frac{\varepsilon}{4E}$$

Compute

$$\begin{aligned}
\left|\Lambda - \bar{\Lambda}_{\Delta t}\right| &\le \mathbb{E}\left|\left(E - X(T)\right)^+ - \left(E - \bar{x}(T)\right)^+\right| \\
&= \mathbb{E}\left(\left|\left(E - X(T)\right)^+ - \left(E - \bar{x}(T)\right)^+\right|\mathbf{1}_{A_i \cap B_j}\right) \\
&\quad + \mathbb{E}\left(\left|\left(E - X(T)\right)^+ - \left(E - \bar{x}(T)\right)^+\right|\mathbf{1}_{A_i^c \cup B_j^c}\right) \\
&\le \mathbb{E}\left(\left|X(T) - \bar{x}(T)\right|\mathbf{1}_{A_i \cap B_j}\right) + 2E\mathbb{P}\left(A_i^c \cup B_j^c\right) \\
&\le \mathbb{E}\left(\left|X(T) - \bar{x}(T)\right|\mathbf{1}_{(\tau_{ij} > T)}\right) + \frac{\varepsilon}{2} \\
&\le \left|X(T \wedge \tau_{ij}) - x(T \wedge \tau_{ij})\right| + \mathbb{E}\left(\left|X(T) - \bar{x}(T)\right|\mathbf{1}_{(\tau_{ij} > T)}\right) \tag{53}
\end{aligned}$$

where τ_{ij} is the stopping time defined in Lemma 6.3. But

$$
\mathbb{E}\left(\left|x(T)-\bar{x}(T)\right|\mathbf{1}_{(\tau_{ij}>T)}\right)
$$
$$
\leq \mathbb{E}\left(\left|\lambda_1(\mu_1+i)\left(T-[T/\Delta t]\right)+\sigma_1\sqrt{j}\,i\left(W_1(T)-W_1\left([T/\Delta t]\Delta t\right)\right)\right|\mathbf{1}_{(\tau_{ij}>T)}\right)
$$
$$
\leq \lambda_1(\mu_1+i)\Delta t+\sigma_1\sqrt{j}\,i\sqrt{\Delta t}
$$

This, together with Lemma 6.3, shows that there is a $\Delta t^* > 0$ such that for all $\Delta t < \Delta t^*$,

$$
\mathbb{E}\left|X(T\wedge\tau_{ij})-x(T\wedge\tau_{ij})\right|+\mathbb{E}\left(\left|x(T)-\bar{x}(T)\right|\mathbf{1}_{(\tau_{ij}>T)}\right)<\frac{\varepsilon}{2}
$$

In view of (53), this completes the proof. $\square$

As we remarked in Section 1, Broadie and Kaya (2003) have shown how to simulate exactly in the Heston model, and their computations demonstrate that this approach is more effcient than Euler discretization for valuing a European put. Generally, for payoffs depending only on $X(T)$, where a single exact step from $t=0$ to $t=T$ is competing against a large number of Euler steps from t_n to t_{n+1}, the exact version will be superior. However, for a fully path-dependent option the situation reverses, because both approaches require the solution to be evolved over small subintervals $[t_n, t_{n+1}]$ and Table 1 of Broadie and Kaya (2003) indicates that the cost of an Euler step is significantly less than the cost of evolving the exact solution. Our final result uses Theorem 6.1 to show that the Euler approach is convergent in such a case: namely for an up-and-out barrier option. Note that using the step function $\bar{x}(t)$ in (55) is equivalent to using the discrete solution.

THEOREM 6.2 *For the Heston model (33)–(34) and the numerical method (41), define*

$$
U := \mathbb{E}\left[\left(X(T)-E\right)^+\mathbf{1}_{\{0\leq X(t)\leq B,\ 0\leq t\leq T\}}\right] \tag{54}
$$
$$
\bar{U}_{\Delta t} := \mathbb{E}\left[\left(\bar{x}(T)-E\right)^+\mathbf{1}_{\{0\leq \bar{x}(t)\leq B,\ 0\leq t\leq T\}}\right] \tag{55}
$$

where the exercise price, E, and barrier, B, are constant. Then

$$
\lim_{\Delta t\to 0}\left|U-\bar{U}_{\Delta t}\right|=0
$$

PROOF Let $A := \{0\leq X(t)\leq B,\ 0\leq t\leq T\}$ and $\bar{A}_{\Delta t} := \{0\leq \bar{x}(t)\leq B,\ 0\leq t\leq T\}$. In the same way as in the proof of Theorem 5.1, we can show that

$$
\left|U-\bar{U}_{\Delta t}\right|\leq \mathbb{E}\left(\left|X(T)-\bar{x}(T)\right|\mathbf{1}_{A\cap\bar{A}_{\Delta t}}\right)+(B-E)\mathbb{P}\left(A\cap\bar{A}_{\Delta t}^{c}\right)
$$
$$
+(B-E)\mathbb{P}\left(A^{c}\cap\bar{A}_{\Delta t}\right) \tag{56}
$$

Now, for $i,j > B$, set

$$A_i := \left\{ X(t) \leq i,\, 0 \leq t \leq T \right\} \text{ and } B_j := \left\{ |v(t)| \leq j,\, 0 \leq t \leq T \right\}$$

Let $\varepsilon > 0$ be arbitrary. By Lemmas 6.1 and 6.2, we can find a pair of i and j sufficiently large for

$$\mathbb{P}\left(A_i^c \cup B_j^c \right) \leq \frac{\varepsilon}{2(1 \vee B)} \tag{57}$$

Compute

$$
\begin{aligned}
\mathbb{E}\left(\left| X(T) - \bar{x}(T) \right| \mathbf{1}_{A \cap \bar{A}_{\Delta t}} \right) &= \mathbb{E}\left(\left| X(T) - \bar{x}(T) \right| \mathbf{1}_{A \cap \bar{A}_{\Delta t} \cap A_i \cap B_j} \right) \\
&\quad + \mathbb{E}\left(\left| X(T) - \bar{x}(T) \right| \mathbf{1}_{A \cap \bar{A}_{\Delta t} \cap (A_i^c \cup B_j^c)} \right) \\
&\leq \mathbb{E}\left(\left| X(T) - \bar{x}(T) \right| \mathbf{1}_{A_i \cap B_j} \right) + B\,\mathbb{P}\left(A_i^c \cup B_j^c \right) \\
&\leq \mathbb{E}\left(\left| X(T) - \bar{x}(T) \right| \mathbf{1}_{\{\tau_{ij} > T\}} \right) + \frac{\varepsilon}{2}
\end{aligned}
$$

where the stopping time τ_{ij} is defined in (45). Recalling the proof of Theorem 6.1, we observe that there is a $\Delta t^* = \Delta t^*(\varepsilon) > 0$ such that for any $\Delta t < \Delta t^*$,

$$\mathbb{E}\left(\left| X(T) - \bar{x}(T) \right| \mathbf{1}_{\{\tau_{ij} > T\}} \right) < \frac{\varepsilon}{2}$$

whence

$$\mathbb{E}\left(\left| X(T) - \bar{x}(T) \right| \mathbf{1}_{A \cap \bar{A}_{\Delta t}} \right) < \varepsilon$$

In other words, we have shown that

$$\lim_{\Delta t \to 0} \mathbb{E}\left(\left| X(T) - \bar{x}(T) \right| \mathbf{1}_{A \cap \bar{A}_{\Delta t}} \right) = 0 \tag{58}$$

Next we will show that $\mathbb{P}(A \cap \bar{A}_{\Delta t}^c) \to 0$ as $\Delta t \to 0$. Using (57), we note that

$$\mathbb{P}\left(A \cap \bar{A}_{\Delta t}^c \right) \leq \mathbb{P}\left(A \cap \bar{A}_{\Delta t}^c \cap A_i \cap B_j \right) + \frac{\varepsilon}{2} \tag{59}$$

On the other hand, for any sufficiently small δ, we have

$$
\begin{aligned}
A &= \left\{ \sup_{0 \leq t \leq T} X(t) \leq B \right\} \\
&= \left\{ \sup_{0 \leq t \leq T} X(t) \leq B - \delta \right\} \cup \left\{ B - \delta < \sup_{0 \leq t \leq T} X(t) \leq B \right\} \\
&\subseteq \left\{ \sup_{0 \leq k\Delta t \leq T} X(k\Delta t) \leq B - \delta \right\} \cup \left\{ B - \delta < \sup_{0 \leq t \leq T} X(t) \leq B \right\} \\
&=: \bar{A}_1 \cup \bar{A}_2
\end{aligned}
$$

Hence,

$$A \cap \bar{A}_{\Delta t}^{\mathrm{c}} \cap A_i \cap B_j \subseteq \left[\bar{A}_1 \cap \bar{A}_{\Delta t}^{\mathrm{c}} \cap A_i \cap B_j \right] \cup \bar{A}_2$$

$$\subseteq \left[\left\{ \sup_{0 \le k\Delta t \le T} \left| X(k\Delta t) - \bar{x}(k\Delta t) \right| \ge \delta \right\} \cap \left\{ \tau_{ij} > T \right\} \right] \cup \bar{A}_2$$

$$\subseteq \left\{ \sup_{0 \le t \le T} \left| X(t \wedge \tau_{ij}) - \bar{x}(t \wedge \tau_{ij}) \right| \ge \delta \right\} \cup \bar{A}_2$$

So,

$$\mathbb{P}\left(A \cap \bar{A}_{\Delta t}^{\mathrm{c}} \cap A_i \cap B_j \right) \le \mathbb{P}\left(\sup_{0 \le t \le T} \left| X(t \wedge \tau_{ij}) - \bar{x}(t \wedge \tau_{ij}) \right| \ge \delta \right) + \mathbb{P}\left(\bar{A}_2 \right)$$

$$\le \frac{1}{\delta^2} \mathbb{E}\left(\sup_{0 \le t \le T} \left| X(t \wedge \tau_{ij}) - \bar{x}(t \wedge \tau_{ij}) \right|^2 \right) + \mathbb{P}\left(\bar{A}_2 \right)$$

Now, for any $\varepsilon > 0$, we may choose δ so small that

$$\mathbb{P}\left(\bar{A}_2 \right) < \frac{\varepsilon}{4}$$

and then, by Lemma 6.3, choose Δt so small that

$$\frac{1}{\delta^2} \mathbb{E}\left(\sup_{0 \le t \le T} \left| X(t \wedge \tau_{ij}) - \bar{x}(t \wedge \tau_{ij}) \right|^2 \right) < \frac{\varepsilon}{4}$$

whence

$$\mathbb{P}\left(A \cap \bar{A}_{\Delta t}^{\mathrm{c}} \cap A_i \cap B_j \right) < \frac{\varepsilon}{2}$$

Substituting this into (59) we obtain

$$\mathbb{P}\left(A \cap \bar{A}_{\Delta t}^{\mathrm{c}} \right) < \varepsilon$$

for sufficiently small Δt. This shows that

$$\lim_{\Delta t \to 0} \mathbb{P}\left(A \cap \bar{A}_{\Delta t}^{\mathrm{c}} \right) = 0 \tag{60}$$

An argument similar to that used in the proof of Theorem 5.1 shows that

$$\lim_{\Delta t \to 0} \mathbb{P}\left(A^{\mathrm{c}} \cap \bar{A}_{\Delta t} \right) = 0 \tag{61}$$

Substituting (58), (60) and (61) into (56) gives the required result.   □

REFERENCES

Andersen, T. G., Benzoni, L., and Lund, J. (2002). An empirical investigation of continuous-time equity return models. *Journal of Finance* **57**, 1239–84.

Andersen, T. G., and Lund, J. (1997). Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics* **77**, 343–77.

Bacinello, A. R. (2000). Valuation of contingent-claims characterising particular pension schemes. *Insurance: Mathematics and Economics* **27**, 177–88.

Barone-Adesi, G., and Sorwar, G. (2002). Interest rate barrier options. In *Computational Methods in Decision-Making, Economics and Finance: Applied Optimization* **74**, E. J. Kontoghiorghes, B. Rustem, and S. Siokos, eds, Kluwer.

Broadie, M., and Kaya, O. (2003), Exact simulation of stochastic volatility and other affine jump diffusion processes. Working Paper.

Clewlow, L., and Strickland, C. (1998). *Implementing Derivative Models*, Wiley, Chichester, UK.

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

Cox, J. C., and Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics* **3**, 145–66.

Cyganowski, S., Kloeden, P. E., and Ombach, J. (2002). *From Elementary Probability to Stochastic Differential Equations*. Springer.

Da Prato, G., and Zabczyk, J. (1992). *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, Cambridge.

Duan, J.-C. (2003). A specification test for time series models by a normality transformation. Working Paper.

Fischer, T., May, A., and Walther, B. (2003). Simulation of the yield curve: Checking a Cox–Ingersoll–Ross model. Working Paper, TU Darmstadt.

Fischer, T., and Roehrl, A. (2003). Risk and performance optimization for portfolios of bonds and stocks. In *Proceedings of the International AFIR Colloquium*, Maastricht.

Gkamas, D. (2001). Stochastic volatility and option pricing. *Quantitative Finance* **1**, 292–7.

Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*. Springer, Berlin.

Hairer, E., and Wanner, G. (1996). *Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems*. Springer Verlag, Berlin, Second edn.

Heston, S. I. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* **6**, 327–43.

Higham, D. J. (2000). Mean-square and asymptotic stability of the stochastic theta method. *SIAM Journal of Numerical Analysis* **38**, 753–69.

Higham, D. J., Mao, X., and Stuart, A. M. (2002). Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM Journal of Numerical Analysis* **40**, 1041–63.

Hull, J. C. (2000). *Options, Futures, and Other Derivatives*. Prentice Hall, New Jersey, Fourth edn.

Ishida, I., and Engle, R. F. (2002). Modeling variance of variance: The square-root, the affine, and the CEV GARCH models. Working paper.

Kloeden, P. E., and Platen, E. (1999). *Numerical Solution of Stochastic Differential Equations*, Third printing. Springer, Berlin,.

Kwok, Y. K. (1998). *Mathematical Models of Financial Derivatives*. Springer, Berlin.

Mao, X. (1997). *Stochastic Differential Equations and Applications*. Horwood.

Misawa, T. (2001). A Lie algebraic approach to numerical integration of stochastic differential equations. *SIAM Journal on Scientific Computing* **23**, 866–90.

Nelson, D. B. (1990). ARCH models as diffusion approximations. *Journal of Econometrics* **45**, 7–38.

Saito, Y., and Mitsui, T. (1996). Stability analysis of numerical schemes for stochastic differential equations. *SIAM Journal of Numerical Analysis* **33**, 2254–67.

Zhang, J. E., and Shu, J. (2003). Pricing S&P 500 index options with Heston's model. In *Proceedings of 2003 IEEE International Conference on Computational Intelligence for Financial Engineering*, pp 85–92.