Adv. Appl. Prob. **47,** 1132–1156 (2015)
Printed in Northern Ireland

UNIFORM APPROXIMATION OF THE COX–INGERSOLL–ROSS PROCESS

GRIGORI N. MILSTEIN,* *Ural Federal University*
JOHN SCHOENMAKERS,** *Weierstrass Institute for Applied Analysis and Stochastics*

Abstract

The Doss–Sussmann (DS) approach is used for uniform simulation of the Cox–Ingersoll–Ross (CIR) process. The DS formalism allows us to express trajectories of the CIR process through solutions of some ordinary differential equation (ODE) depending on realizations of a Wiener process involved. By simulating the first-passage times of the increments of the Wiener process to the boundary of an interval and solving the ODE, we uniformly approximate the trajectories of the CIR process. In this respect special attention is payed to simulation of trajectories near 0. From a conceptual point of view the proposed method gives a better quality of approximation (from a pathwise point of view) than standard, even exact, simulation of the stochastic differential equation at some deterministic time grid.

Keywords: Cox–Ingersoll–Ross process; Doss–Sussmann formalism; Bessel function; confluent hypergeometric equation

2010 Mathematics Subject Classification: Primary 65C30
Secondary 60H35

1. Introduction

The Cox–Ingersoll–Ross (CIR) process $V(t)$ is determined by the following stochastic differential equation (SDE):

$$\mathrm{d}V(t) = k(\lambda - V(t))\mathrm{d}t + \sigma\sqrt{V(t)}\mathrm{d}w, \qquad V(t_0) = V_0, \tag{1}$$

where k, λ, and σ are positive constants, and w is a scalar Brownian motion. Due to [6] this process has become very popular in financial mathematical applications. The CIR process is used in particular as volatility process in the Heston model [12]. It is known (see [15] and [16]) that for $V_0 > 0$ there exists a unique strong solution $V_{t_0,V_0}(t)$ of (1) for all $t \geq t_0 \geq 0$. The CIR process $V(t) = V_{t_0}, V_0(t)$ is positive in the $2k\lambda \geq \sigma^2$ case and nonnegative in the $2k\lambda < \sigma^2$ case. Moreover, in the latter case the origin is a reflecting boundary.

As a matter of fact, (1) does not satisfy the global Lipschitz assumption. The difficulties arising in a simulation method for (1) are connected with this fact and with the natural requirement of preserving nonnegative approximations. Numerous approximation methods for the CIR processes have been proposed. For an extensive list of articles on this subject, see [3] and [7]. In addition, see also [1], [2], [13], and [14], where a number of discretization schemes for the CIR process can be found. Furthermore, we note that in [19] a weakly convergent fully

Received 12 December 2013; revision received 20 July 2014.
* Postal address: Ural Federal University, Lenin Str. 51, 620083 Ekaterinburg, Russia.
** Postal address: Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, 10117 Berlin, Germany. Email address: schoenma@wias-berlin.de

implicit method is implemented for the Heston model. Exact simulation of (1) is considered in [5] and [9] (see also [3]).

In this paper we consider the uniform pathwise approximation of $V(t)$ on an interval $[t_0, t_0 + T]$ using the Doss–Sussmann (DS) transformation ([8], [22], [21]) which allows for expressing any trajectory of $V(t)$ by the solution of some ordinary differential equation (ODE) that depends on the realization of $w(t)$. The approximation $\overline{V}(t)$ will be uniform in the sense that the pathwise error will be uniformly bounded, i.e.

$$\sup_{t_0 \le t \le t_0+T} |\overline{V}(t) - V(t)| \le r \quad \text{almost surely,} \tag{2}$$

where $r > 0$ is fixed in advance. In fact, by simulating the first-passage times of the increments of the Wiener process to the boundary of an interval and solving this ODE, we construct an approximate generic trajectory of $V(t)$. This kind of simulation is simpler than the one proposed in [5] and, moreover, has the advantage of being uniform in nature. Let us consider the simulation of a standard Brownian motion W on a fixed time grid

$$t_0, t_i, \dots, t_n = T.$$

Although W may be exactly simulated at the grid points, the usual piecewise linear interpolation

$$\overline{W}(t) = \frac{t_{i+1} - t}{t_{i+1} - t_i} W(t_i) + \frac{t - t_i}{t_{i+1} - t_i} W(t_{i+1})$$

is not uniform in the sense of (2). Put differently, for any (large) positive number A, there is always a positive probability (though possibly small) that

$$\sup_{t_0 \le t \le t_0+T} |\overline{W}(t) - W(t)| > A.$$

Therefore, for path-dependent applications, for instance, such a standard, even exact, simulation method may be not desirable and a uniform method preserving (2) may be preferred. Apart from applications, however, uniform simulation of trajectories of an SDE in the sense of (2) may be considered as an interesting mathematical problem in its own right.

We note that the original DS results rely on a global Lipschitz assumption that is not fulfilled for (1). Therefore, we have introduced the DS formalism that yields a corresponding ODE whose solutions are defined on random time intervals. If V gets close to 0, however, the ODE becomes intractable for numerical integration and so, for the parts of a trajectory $V(t)$ that are close to 0, we are forced to use some other (non-DS) approach. For such parts, we here propose a different uniform simulation method. Another restriction is connected with the condition

$$\alpha := \frac{4k\lambda - \sigma^2}{8} > 0. \tag{3}$$

We underline that the $\alpha > 0$ case is more general than the $2k\lambda \ge \sigma^2$ case that ensures positivity of $V(t)$ and stress that in the literature many convergence proofs for numerical integration schemes for (1) are based on the assumption that $2k\lambda \ge \sigma^2$. However, for example, in [1] and [13] convergence without this assumption is obtained with a strong error inverse proportional to the logarithm of the number of time steps (loosely speaking). We expect that the results here obtained for $\alpha > 0$ can be extended to the case where $\alpha \le 0$, however in a highly nontrivial way. This is considered in a subsequent work; see [17].

The paper is organized as follows. The next two sections are devoted to DS formalism in connection with (1) and to some auxiliary propositions. In Section 4 we deal with the one-step approximation and in Section 5 with the convergence of the proposed method. A first simulation algorithm is described in Section 5. Section 6 is dedicated to the uniform construction of trajectories close to 0, resulting in a main simulation algorithm and a corresponding convergence theorem. In Section 7 we present a numerical experiment and discuss some beneficial issues of the main algorithm in certain applications. The more technical parts are deferred to the appendices.

2. The DS transformation

Due to the DS approach [8], [15], [21], [22], the solution of (1) may be expressed in the form

$$V(t) = F(X(t), w(t)), \tag{4}$$

where $F = F(x, y)$ is some deterministic function and $X(t)$ is the solution of some ODE depending on the part $w(s)$, $0 \le s \le t$, of the realization $w(\cdot)$ of the Wiener process $w(t)$.

Let us recall the DS formalism according to [21, Section V.28]. In [21], the authors' considered the Stratonovich SDE:

$$\mathrm{d}V(t) = b(V)\,\mathrm{d}t + \gamma(V) \circ \mathrm{d}w(t). \tag{5}$$

The function $F = F(x, y)$ is found from

$$\frac{\partial F}{\partial y} = \gamma(F), \qquad F(x, 0) = x, \tag{6}$$

and $X(t)$ is found from the ODE:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{1}{\partial F/\partial x(X(t), w(t))} b(F(X(t), w(t))), \qquad X(0) = V(0). \tag{7}$$

It turns out that application of the DS formalism after the Lamperti transformation $U(t) = \sqrt{V(t)}$ (see [7]) leads to simpler equations. The Lamperti transformation yields the following SDE with additive noise:

$$\mathrm{d}U = \left(\frac{\alpha}{U} - \frac{k}{2}U\right)\mathrm{d}t + \frac{\sigma}{2}\,\mathrm{d}w, \qquad U(0) = \sqrt{V(0)} > 0, \tag{8}$$

where α is given in (3). Let us seek the solution of (8) in the form

$$U(t) = G(Y(t), w(t)) \tag{9}$$

in accordance with (4)–(7). Because the Itô and Stratonovich forms of (8) coincide, we have

$$b(U) = \frac{\alpha}{U} - \frac{k}{2}U, \qquad \gamma(U) = \frac{\sigma}{2}.$$

The function $G = G(y, z)$ is found from

$$\frac{\partial G}{\partial z} = \frac{\sigma}{2}, \qquad G(y, 0) = y,$$

i.e.

$$G(y, z) = y + \frac{\sigma}{2} z, \tag{10}$$

and $Y(t)$ is found from the ODE,

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \frac{\alpha}{Y + (\sigma/2)w(t)} - \frac{k}{2}\bigg(Y + \frac{\sigma}{2} w(t)\bigg), \qquad Y(0) = U(0) = \sqrt{V(0)} > 0. \tag{11}$$

From (9), (10), and the solution of (11), we formally obtain the solution $U(t)$ of (8):

$$U(t) = Y(t) + \frac{\sigma}{2} w(t). \tag{12}$$

Hence,

$$V(t) = U^2(t) = \bigg(Y(t) + \frac{\sigma}{2} w(t)\bigg)^2. \tag{13}$$

Since the DS results rely on a global Lipschitz assumption that is not fulfilled for (1), solution (13) has to be considered only formally. In this section we therefore provide a direct proof of the following more precise result.

Proposition 1. *Let* $Y(0) = U(0) = \sqrt{V(0)} > 0$. *Let* τ *be the following stopping time:*

$$\tau := \inf\{t \colon V(t) = 0\}.$$

Then (11) has a unique solution $Y(t)$ *on the interval* $[0, \tau)$, *the solution* $U(t)$ *of (8) is expressed by (12) on this interval, and* $V(t)$ *is expressed by (13).*

Proof. Let $(w(t),\ V(t))$ be the solution of the SDE system

$$\mathrm{d}w = \mathrm{d}w(t), \qquad \mathrm{d}V = k(\lambda - V)\,\mathrm{d}t + \sigma\sqrt{V(t)}\,\mathrm{d}w(t),$$

which satisfies the initial conditions $w(0) = 0$, $V(0) > 0$. Then $U(t) = \sqrt{V(t)} > 0$ is a solution of (8) on the interval $[0, \tau)$. Consider the function $Y(t) = U(t) - (\sigma/2)w(t),\ 0 \le t < \tau$. Clearly, $Y(t) + (\sigma/2)w(t) > 0$ on $[0, \tau)$. Due to Itô's formula, we obtain

$$\mathrm{d}Y(t) = \mathrm{d}U(t) - \frac{\sigma}{2}\,\mathrm{d}w(t) = \frac{\alpha\,\mathrm{d}t}{Y + (\sigma/2)w(t)} - \frac{k}{2}\bigg(Y + \frac{\sigma}{2} w(t)\bigg)\,\mathrm{d}t,$$

i.e. the function $U(t) - (\sigma/2)w(t)$ is a solution of (11). The uniqueness of $Y(t)$ follows from the uniqueness of $V(t)$.

So far we started at the moment $t = 0$. It is useful to consider the DS transformation with an arbitrary initial time $t_0 > 0$ (which may even be a stopping time, for example, $0 \le t_0 < \tau$). In this case, we obtain instead of (11) for

$$Y = Y(t; t_0) = U(t) - \frac{\sigma}{2}(w(t) - w(t_0)) = \sqrt{V(t)} - \frac{\sigma}{2}(w(t) - w(t_0)), \qquad t_0 \le t < t_0 + \tau,$$

the equation

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \frac{\alpha}{Y + (\sigma/2)(w(t) - w(t_0))} - \frac{k}{2}\bigg(Y + \frac{\sigma}{2}(w(t) - w(t_0))\bigg), \tag{14}$$

$$Y(t_0; t_0) = \sqrt{V(t_0)}, \quad t_0 \le t < t_0 + \tau,$$

with α given by (3). Clearly,

$$V(t) = \bigg(Y(t; t_0) + \frac{\sigma}{2}(w(t) - w(t_0))\bigg)^2, \qquad t_0 \le t < t_0 + \tau. \tag{15}$$

3. Auxiliary propositions

Let us consider in view of (14) solutions of the ODEs

$$\frac{\mathrm{d}y^0}{\mathrm{d}t} = \frac{\alpha}{y^0} - \frac{k}{2}y^0, \qquad y^0(t_0) = y_0 > 0, \quad t \ge t_0 \ge 0, \tag{16}$$

which are given by

$$y^0(t) = y^0_{t_0,y_0}(t) = \left[y_0^2 \mathrm{e}^{-k(t-t_0)} + \frac{2\alpha}{k}(1 - \mathrm{e}^{-k(t-t_0)}) \right]^{1/2}, \qquad t \ge t_0. \tag{17}$$

In the $\alpha > 0$ case, i.e. is by (3) $4k\lambda > \sigma^2$, we have the following: if $y_0 > \sqrt{2\alpha/k}$ then $y^0_{t_0,y_0}(t) \downarrow \sqrt{2\alpha/k}$ as $t \to \infty$ and if $0 < y_0 < \sqrt{2\alpha/k}$ then $y^0_{t_0,y_0}(t) \uparrow \sqrt{2\alpha/k}$ as $t \to \infty$. Furthermore, $y^0(t) = \sqrt{2\alpha/k}$ is a solution of (16).

Our next goal is to obtain estimates for solutions of

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\alpha}{y + (\sigma/2)\varphi(t)} - \frac{k}{2}\left(y + \frac{\sigma}{2}\varphi(t) \right), \qquad y(t_0) = y_0, \quad t_0 \le t \le t_0 + \theta, \tag{18}$$

(cf. (14)) for a given continuous function $\varphi(t)$.

Lemma 1. *Let $\alpha \ge 0$. Let $y^i(t)$, $i = 1, 2$, be two solutions of (18) such that $y^i(t) + (\sigma/2)\varphi(t) > 0$ on $[t_0, t_0 + \theta]$ for some θ with $0 \le \theta \le T$. Then*

$$|y^2(t) - y^1(t)| \le |y^2(t_0) - y^1(t_0)|, \qquad t_0 \le t \le t_0 + \theta. \tag{19}$$

Proof. We have

$$\mathrm{d}(y^2(t) - y^1(t))^2 = 2(y^2(t) - y^1(t))\left(\frac{\alpha}{y^2(t) + (\sigma/2)\varphi(t)} - \frac{k}{2}\left(y^2(t) + \frac{\sigma}{2}\varphi(t) \right) - \frac{\alpha}{y^1(t) + (\sigma/2)\varphi(t)} + \frac{k}{2}\left(y^1(t) + \frac{\sigma}{2}\varphi(t) \right) \right) \mathrm{d}t. \tag{20}$$

From here

$$\begin{aligned} (y^2(t) - y^1(t))^2 &= (y^2(t_0) - y^1(t_0))^2 \\ &\quad + 2\int_{t_0}^{t} \left[-\alpha \frac{(y^2(s) - y^1(s))^2}{(y^1(s) + (\sigma/2)\varphi(s))(y^2(s) + (\sigma/2)\varphi(s))} - \frac{k}{2}(y^2(s) - y^1(s))^2 \right] \mathrm{d}s \\ &\le (y^2(t_0) - y^1(t_0))^2, \end{aligned}$$

whence (19) follows.

Proposition 2. *For any $\alpha > 0$ it holds that*

$$\left| \sqrt{V_{t_0,V_0^2}(t)} - \sqrt{V_{t_0,V_0^1}(t)} \right| \le \left| \sqrt{V_0^2} - \sqrt{V_0^1} \right|, \qquad t_0 \le t < \infty. \tag{21}$$

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

Proof. In the $2k\lambda \geq \sigma^2$ case, we have

$$\sqrt{V_{t_0,V_0^i}(t)} = Y^i(t;t_0) + \frac{\sigma}{2}(w(t) - w(t_0)) > 0,$$
$$Y^i(t_0;t_0) = \sqrt{V_0^i}, \qquad i = 1,2,\ t_0 \leq t < \infty,$$

where the $Y^i(t;t_0)$ satisfy (14). So, by Lemma 1 with $\varphi(t) = w(t) - w(t_0)$,

$$|Y^2(t;t_0) - Y^1(t;t_0)| \leq \left| \sqrt{V_0^2} - \sqrt{V_0^1} \right|, \qquad t_0 \leq t < \infty,$$

and (21) follows since $Y^2(t;t_0) - Y^1(t;t_0) = \sqrt{V_{t_0,V_0^2}(t)} - \sqrt{V_{t_0,V_0^1}(t)}$. The general case $\alpha > 0$ is proved in Appendix A.

Now consider (18) for a continuous function φ satisfying

$$|\varphi(t)| \leq r, \qquad t_0 \leq t \leq t_0 + \theta \leq t_0 + T, \tag{22}$$

for some $r > 0$ and $0 \leq \theta \leq T$. Along with (16) and (18) with (22), we further consider the equations

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\alpha}{y + (\sigma/2)r} - \frac{k}{2}\left(y + \frac{\sigma}{2}r\right), \qquad y(t_0) = y_0, \tag{23}$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\alpha}{y - (\sigma/2)r} - \frac{k}{2}\left(y - \frac{\sigma}{2}r\right), \qquad y(t_0) = y_0. \tag{24}$$

Let us assume that $y_0 \geq \sigma r > 0$, and consider an $\eta > 0$, to be specified below, that satisfies

$$y_0 \geq \eta \geq \sigma r > 0. \tag{25}$$

The solutions of (16) and (18) with (22), (23), and (24) are denoted by $y^0(t),\ y(t),\ y^-(t)$, and $y^+(t)$, respectively, where $y^0(t)$ is given by (17). By using (17) we derive straightforwardly that

$$y^-(t) = \left[\left(y_0 + \frac{\sigma}{2}r\right)^2 \mathrm{e}^{-k(t-t_0)} + \frac{2\alpha}{k}(1 - \mathrm{e}^{-k(t-t_0)})\right]^{1/2} - \frac{\sigma}{2}r, \qquad t_0 \leq t \leq t_0 + \theta,$$
$$y^+(t) = \left[\left(y_0 - \frac{\sigma}{2}r\right)^2 \mathrm{e}^{-k(t-t_0)} + \frac{2\alpha}{k}(1 - \mathrm{e}^{-k(t-t_0)})\right]^{1/2} + \frac{\sigma}{2}r, \qquad t_0 \leq t \leq t_0 + \theta.$$

Note that $y^-(t) + \sigma r/2 > 0$ and $y^+(t) > \sigma r/2$, $t_0 \leq t \leq t_0 + \theta$. Due to the comparison theorem for ODEs (see, for example, [11, Chapter 3]), the inequality

$$\frac{\alpha}{y + (\sigma/2)r} - \frac{k}{2}\left(y + \frac{\sigma}{2}r\right) \leq \frac{\alpha}{y + (\sigma/2)\varphi(t)} - \frac{k}{2}\left(y + \frac{\sigma}{2}\varphi(t)\right) \leq \frac{\alpha}{y - (\sigma/2)r} - \frac{k}{2}\left(y - \frac{\sigma}{2}r\right),$$

which is fulfilled in view of (22) for $y > \sigma r/2$, then implies that

$$y^-(t) \leq y(t) \leq y^+(t), \qquad t_0 \leq t \leq t_0 + \theta.$$

The same inequality holds for $y(t)$ replaced by $y^0(t)$. Thus, we obtain

$$|y(t) - y^0(t)| \leq y^+(t) - y^-(t), \qquad t_0 \leq t \leq t_0 + \theta. \tag{26}$$

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

Proposition 3. *Let $\alpha > 0$, the inequalities (22) and (25) be fulfilled for a fixed $\eta > 0$, and let $\theta \le T$. We then have*

$$|y(t) - y^0(t)| \le Cr(t - t_0) \le Cr\theta, \qquad t_0 \le t \le t_0 + \theta \text{ with } C = \frac{\sigma k}{2} + \frac{4\alpha\sigma}{3\eta^2}\mathrm{e}^{(k/2)T}. \quad (27)$$

In particular, C is independent of t_0, y_0, and r (provided (25) holds).

Proof. We estimate the difference $y^+(t) - y^-(t)$. It holds that

$$y^+(t) = z^-(t) + \frac{\sigma}{2}r, \qquad y^-(t) = z^+(t) - \frac{\sigma}{2}r,$$
$$y^+(t) - y^-(t) = \sigma r - (z^+(t) - z^-(t)), \quad (28)$$

where

$$z^\pm(t) = \left[\left(y_0 \pm \frac{\sigma}{2}r\right)^2 \mathrm{e}^{-k(t-t_0)} + \frac{2\alpha}{k}(1 - \mathrm{e}^{-k(t-t_0)})\right]^{1/2}.$$

Furthermore,

$$z^+(t) - z^-(t) = \frac{(z^+(t))^2 - (z^-(t))^2}{z^+(t) + z^-(t)} = \frac{2y_0\sigma r\mathrm{e}^{-k(t-t_0)}}{z^+(t) + z^-(t)}. \quad (29)$$

Using the inequality $(a^2 + b)^{1/2} \le a + b/2a$ for any $a > 0$ and $b \ge 0$, we obtain

$$z^+(t) \le \left(y_0 + \frac{\sigma}{2}r\right)\mathrm{e}^{-(k/2)(t-t_0)} + \frac{\alpha}{k}\left(\frac{(1 - \mathrm{e}^{-k(t-t_0)})}{(y_0 + (\sigma/2)r)\mathrm{e}^{-(k/2)(t-t_0)}}\right),$$
$$z^-(t) \le \left(y_0 - \frac{\sigma}{2}r\right)\mathrm{e}^{-(k/2)(t-t_0)} + \frac{\alpha}{k}\left(\frac{(1 - \mathrm{e}^{-k(t-t_0)})}{(y_0 - (\sigma/2)r)\mathrm{e}^{-(k/2)(t-t_0)}}\right),$$

whence

$$z^+(t) + z^-(t) \le 2y_0\mathrm{e}^{-(k/2)(t-t_0)} + \frac{\alpha}{k}\left(\frac{(1 - \mathrm{e}^{-k(t-t_0)})}{\mathrm{e}^{-(k/2)(t-t_0)}}\right)\left(\frac{2y_0}{(y_0^2 - (\sigma^2/4)r^2)}\right).$$

Therefore,

$$\frac{1}{z^+(t) + z^-(t)} \ge \frac{1}{2y_0\mathrm{e}^{-(k/2)(t-t_0)}}\left(1 - \frac{\alpha}{k(y_0^2 - (\sigma^2/4)r^2)}(\mathrm{e}^{k(t-t_0)} - 1)\right).$$

From (29), we have

$$z^+(t) - z^-(t) \ge \sigma r\mathrm{e}^{-(k/2)(t-t_0)}\left(1 - \frac{\alpha}{k(y_0^2 - (\sigma^2/4)r^2)}(\mathrm{e}^{k(t-t_0)} - 1)\right)$$

and so due to (28), we obtain

$$0 \le y^+(t) - y^-(t) \le \sigma r(1 - \mathrm{e}^{-(k/2)(t-t_0)}) + \frac{\alpha\sigma r}{k(y_0^2 - (\sigma^2/4)r^2)}(\mathrm{e}^{(k/2)(t-t_0)} - \mathrm{e}^{-(k/2)(t-t_0)}).$$

Since $1 - \mathrm{e}^{-q\vartheta} \le q\vartheta$ for any $q \ge 0$, $\vartheta \ge 0$, and $y_0^2 - (\sigma^2/4)r^2 \ge \frac{3}{4}\eta^2$ due to (25), we obtain

$$0 \le y^+(t) - y^-(t) \le \frac{\sigma r k}{2}(t - t_0) + \frac{4\alpha\sigma r}{3k\eta^2}\mathrm{e}^{(k/2)(t-t_0)}k(t - t_0).$$

From this and (26), (27) follows with $C = (\sigma k/2) + (4\alpha\sigma/3\eta^2)\mathrm{e}^{(k/2)T}$.

Corollary 1. *Under the assumptions of Proposition 3, by taking $\eta = y_0$, we obtain*

$$|y(t) - y^0(t)| \le \left(\frac{\sigma k}{2} + \frac{4\alpha\sigma}{3y_0^2}\mathrm{e}^{(k/2)T}\right)r\theta = \left(D_1 + \frac{D_2}{y_0^2}\right)r\theta, \qquad t_0 \le t \le t_0 + \theta,$$

where $D_1 := \sigma k/2$ and $D_2 := 4\alpha\sigma\mathrm{e}^{(k/2)T}/3$ depend only on the parameters of the CIR process under consideration and the time horizon T.

4. One-step approximation

Let us suppose that for $t_m,\ t_0 \le t_m < t_0 + T,\ V(t_m)$ is known exactly. In fact, t_m may be considered as a realization of a certain stopping time. Consider $Y = Y(t; t_m)$ on some interval $[t_m, t_m + \theta_m]$ with $y_m := Y(t_m; t_m) = \sqrt{V(t_m)}$, given by the ODE (cf. (14)),

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \frac{\alpha}{Y + (\sigma/2)(w(t) - w(t_m))} - \frac{k}{2}\left(Y + \frac{\sigma}{2}(w(t) - w(t_m))\right), \tag{30}$$
$$Y(t_m; t_m) = \sqrt{V(t_m)}, \qquad t_m \le t \le t_m + \theta_m.$$

Assume that

$$y_m = \sqrt{V(t_m)} \ge \sigma r.$$

Due to (15), the solution $V(t)$ of (1) on $[t_m, t_m + \theta_m]$ is obtained via

$$\sqrt{V(t)} = Y(t; t_m) + \frac{\sigma}{2}(w(t) - w(t_m)), \qquad t_m \le t \le t_m + \theta_m.$$

Though (30) is (just) an ODE, it is not easy to solve it numerically in a straightforward way because of the nonsmoothness of $w(t)$. Here we are going to construct an approximation $y^m(t)$ of $Y(t; t_m)$ via Proposition 3. To this end we simulate the point $(t_m + \theta_m, w(t_m + \theta_m) - w(t_m))$ by simulating θ_m as being the first-passage (stopping) time of the Wiener process $w(t) - w(t_m)$, $t \ge t_m$, to the boundary of the interval $[-r, r]$. So, $|w(t) - w(t_m)| \le r$ for $t_m \le t \le t_m + \theta_m$ and, moreover, the random variable $w(t_m + \theta_m) - w(t_m)$, which equals either $-r$ or $+r$ with probability $\frac{1}{2}$, is independent of the stopping time θ_m. A method for simulating the stopping time θ_m is given in Appendix B. Proposition 3 and Corollary 1 then yield

$$|Y(t; t_m) - y^m(t)| \le \left(D_1 + \frac{D_2}{y_m^2}\right)r(t_{m+1} - t_m), \qquad t_m \le t \le t_{m+1}, \tag{31}$$

with $t_{m+1} := \min(t_m + \theta_m, t_0 + T)$, where $y^m(t)$ is the solution of the problem

$$\frac{\mathrm{d}y^m}{\mathrm{d}t} = \frac{\alpha}{y^m} - \frac{k}{2}y^m, \qquad y^m(t_m) = Y(t_m; t_m) = \sqrt{V(t_m)}$$

that is given by (17) with $(t_m, y_m) = (t_m, \sqrt{V(t_m)})$. So, we have

$$\sqrt{V(t)} = Y(t; t_m) + \frac{\sigma}{2}(w(t) - w(t_m)) = y^m(t) + \frac{\sigma}{2}(w(t) - w(t_m)) + \rho^m(t),$$

where, due to (31),

$$|\rho^m(t)| \le \left(D_1 + \frac{D_2}{y_m^2}\right)r(t_{m+1} - t_m), \qquad t_m \le t \le t_{m+1}.$$

We next introduce the one-step approximation $\sqrt{\overline{V}(t)}$ of $\sqrt{V(t)}$ on $[t_m, t_{m+1}]$ by

$$\sqrt{\overline{V}(t)} := y^m(t) + \frac{\sigma}{2}(w(t) - w(t_m)), \qquad t_m \leq t \leq t_{m+1}. \tag{32}$$

Since $|w(t_{m+1}) - w(t_m)| = r$ if $t_{m+1} = t_m + \theta_m < t_0 + T$, and $|w(t_{m+1}) - w(t_m)| \leq r$ if $t_{m+1} = t_0 + T$, the one-step approximation (32) for $t = t_{m+1}$ is given by

$$\begin{aligned}\sqrt{\overline{V}(t_{m+1})} &:= y^m(t_{m+1}) + \frac{\sigma}{2}(w(t_{m+1}) - w(t_m)) \\ &= y^m(t_{m+1}) + \begin{cases} \dfrac{\sigma}{2} r\xi_m & \text{with } \mathbb{P}(\xi_m = \pm 1) = \frac{1}{2}, \\ & \text{if } t_{m+1} = t_m + \theta_m < t_0 + T, \\ \dfrac{\sigma}{2}\zeta_m & \text{if } t_{m+1} = t_0 + T, \end{cases}\end{aligned} \tag{33}$$

with $\zeta_m = w(t_0 + T) - w(t_m)$ being drawn from the distribution of

$$W_{t_0+T-t_m} \text{ conditional on } \max_{0 \leq s \leq t_0+T-t_m} |W_s| \leq r, \tag{34}$$

where W is an independent standard Brownian motion. For details see Appendix B. So, we have the following theorem.

Theorem 1. *For the one-step approximation $\overline{V}(t_{m+1})$ due to the exact starting value $\overline{V}(t_m) = V(t_m) = y_m^2$, we have the one-step error*

$$\left| \sqrt{V(t_{m+1})} - \sqrt{\overline{V}(t_{m+1})} \right| \leq \left(D_1 + \frac{D_2}{V(t_m)} \right) r(t_{m+1} - t_m). \tag{35}$$

5. The first convergence theorem

In this section we develop a scheme that generates approximations $\sqrt{\overline{V}(t_0)} = \sqrt{V(t_0)}$, $\sqrt{\overline{V}(t_1)}, \dots, \sqrt{\overline{V}(t_{n+1})}$, where $n = 0, 1, 2, \dots$, and $t_1, \dots, t_{n+1}$ are realizations of a sequence of stopping times, and show that the global error in approximation $\sqrt{\overline{V}(t_{n+1})}$ is in fact an aggregated sum of local errors, i.e.

$$r \sum_{m=0}^{n} \left(D_1 + \frac{D_2}{\overline{V}(t_m)} \right)(t_{m+1} - t_m) \leq rT\left(D_1 + \frac{D_2}{\eta_n^2} \right),$$

with $y_m = \sqrt{\overline{V}(t_m)}$, provided that $y_m \geq \sigma r$ for $m = 0, \dots, n$, and so $\eta_n := \min_{0 \leq m \leq n} y_m \geq \sigma r$.

Let us now describe an algorithm for the solution of (1) on the interval $[t_0, t_0 + T]$ in the $\alpha \geq 0$ case. Suppose we are given $V(t_0)$ and r such that

$$\sqrt{V(t_0)} \geq \sigma r.$$

For the initial step we use the one-step approximation according to the previous section and thus obtain (see (33) and (35))

$$\sqrt{\overline{V}(t_1)} = y^0(t_1) + \frac{\sigma}{2}(w(t_1) - w(t_0)), \qquad \sqrt{V(t_1)} = \sqrt{\overline{V}(t_1)} + \rho^0(t_1),$$

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

where

$$|\rho^0(t_1)| \le \left(D_1 + \frac{D_2}{V(t_0)}\right) r(t_1 - t_0) =: C_0 r(t_1 - t_0).$$

Suppose that

$$\sqrt{\overline{V}(t_1)} \ge \sigma r.$$

We then go to the next step and consider the expression

$$\sqrt{V(t)} = Y(t; t_1) + \frac{\sigma}{2}(w(t) - w(t_1)), \tag{36}$$

where $Y(t; t_1)$ is the solution of the problem (see (30))

$$\begin{gathered} \frac{\mathrm{d}Y}{\mathrm{d}t} = \frac{\alpha}{Y + (\sigma/2)(w(t) - w(t_1))} - \frac{k}{2}\left(Y + \frac{\sigma}{2}(w(t) - w(t_1))\right), \\ Y(t_1; t_1) = \sqrt{V(t_1)}, \qquad t_1 \le t \le t_1 + \theta_1. \end{gathered} \tag{37}$$

Now, in contrast to the initial step, the value $\sqrt{V(t_1)}$ is unknown and we are forced to use $\sqrt{\overline{V}(t_1)}$ instead. Therefore, we introduce $\overline{Y}(t; t_1)$ as the solution of (37) with initial value $\overline{Y}(t_1; t_1) = \sqrt{\overline{V}(t_1)}$. From the previous step, we have $|Y(t_1; t_1) - \overline{Y}(t_1; t_1)| = |\sqrt{V(t_1)} - \sqrt{\overline{V}(t_1)}| = |\rho^0(t_1)| \le C_0 r(t_1 - t_0)$. Hence, due to Lemma 1,

$$|Y(t; t_1) - \overline{Y}(t; t_1)| \le \rho^0(t_1) \le C_0 r(t_1 - t_0), \qquad t_1 \le t \le t_1 + \theta_1. \tag{38}$$

Let θ_1 be the first-passage time of the Wiener process $w(t_1 + \cdot) - w(t_1)$ to the boundary of the interval $[-r, r]$. If $t_1 + \theta_1 < t_0 + T$ then set $t_2 := t_1 + \theta_1$, else set $t_2 := t_0 + T$. In order to approximate $\overline{Y}(t; t_1)$ for $t_1 \le t \le t_2$ let us consider along with (37) the equation

$$\frac{\mathrm{d}y^1}{\mathrm{d}t} = \frac{\alpha}{y^1} - \frac{k}{2}y^1, \qquad y^1(t_1) = \overline{Y}(t_1; t_1) = \sqrt{\overline{V}(t_1)}.$$

Due to Proposition 3 and Corollary 1 it holds that

$$|\overline{Y}(t; t_1) - y^1(t)| \le \left(D_1 + \frac{D_2}{\overline{V}(t_1)}\right) r(t_2 - t_1) =: C_1 r(t_2 - t_1), \qquad t_1 \le t \le t_2, \tag{39}$$

and so by (38), we have

$$|Y(t; t_1) - y^1(t)| \le r(C_0(t_1 - t_0) + C_1(t_2 - t_1)), \qquad t_1 \le t \le t_2. \tag{40}$$

We also have (see (36))

$$\sqrt{V(t)} = Y(t; t_1) + \frac{\sigma}{2}(w(t) - w(t_1)) = y^1(t) + \frac{\sigma}{2}(w(t) - w(t_1)) + R^1(t), \tag{41}$$

where

$$|R^1(t)| \le r(C_0(t_1 - t_0) + C_1(t_2 - t_1)), \qquad t_1 \le t \le t_2. \tag{42}$$

We so define the approximation

$$\sqrt{\overline{V}(t)} := y^1(t) + \frac{\sigma}{2}(w(t) - w(t_1)) \tag{43}$$

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

that satisfies

$$\sqrt{V(t)} = \sqrt{\overline{V}(t)} + R^1(t), \quad t_1 \le t \le t_2, \tag{44}$$

and then set

$$\begin{aligned} \sqrt{\overline{V}(t_2)} &= y^1(t_2) + \frac{\sigma}{2}(w(t_2) - w(t_1)) \\ &= y^1(t_2) + \begin{cases} \frac{\sigma}{2} r\xi_1 & \text{with } \mathbb{P}(\xi_1 = \pm 1) = \frac{1}{2}, \text{ if } t_2 = t_1 + \theta_1 < t_0 + T, \\ \frac{\sigma}{2}\zeta_1 & \text{if } t_2 = t_0 + T, \end{cases} \end{aligned} \tag{45}$$

(cf. (33) and (34)). We thus end up with a next approximation $\sqrt{\overline{V}(t_2)}$ such that

$$\left| \sqrt{V(t_2)} - \sqrt{\overline{V}(t_2)} \right| = |R^1(t_2)| \le r(C_0(t_1 - t_0) + C_1(t_2 - t_1)). \tag{46}$$

From the above description it is obvious how to proceed analogously given a generic approximation sequence of approximations $\sqrt{\overline{V}(t_m)}$, $m = 0, 1, 2, \dots, n$, with $\overline{V}(t_0) = V(t_0)$, that satisfies by assumption

$$\sqrt{\overline{V}(t_m)} \ge \sigma r \quad \text{for } m = 0, \dots, n, \tag{47}$$

$$\begin{aligned} \left| \sqrt{V(t_n)} - \sqrt{\overline{V}(t_n)} \right| &\le r \sum_{m=0}^{n-1} \left(D_1 + \frac{D_2}{\overline{V}(t_m)} \right)(t_{m+1} - t_m) \\ &=: r \sum_{m=0}^{n-1} C_m(t_{m+1} - t_m). \end{aligned} \tag{48}$$

Indeed, consider the expression

$$\sqrt{V(t)} = Y(t; t_n) + \frac{\sigma}{2}(w(t) - w(t_n)),$$

where $Y(t; t_n)$ is the solution of the problem

$$\begin{aligned} \frac{\mathrm{d}Y}{\mathrm{d}t} &= \frac{\alpha}{Y + (\sigma/2)(w(t) - w(t_n))} - \frac{k}{2}\left(Y + \frac{\sigma}{2}(w(t) - w(t_n)) \right), \\ Y(t_n; t_n) &= \sqrt{V(t_n)}, \qquad t_n \le t \le t_n + \theta_n, \end{aligned} \tag{49}$$

for a $\theta_n > 0$ to be determined. Since $\sqrt{V(t_n)}$ is unknown we consider $\overline{Y}(t; t_n)$ as the solution of (49) with initial value $\overline{Y}(t_n; t_n) = \sqrt{\overline{V}(t_n)}$. Due to (48) and Lemma 1 again, we have

$$|Y(t; t_n) - \overline{Y}(t; t_n)| \le r \sum_{m=0}^{n-1} C_m(t_{m+1} - t_m), \qquad t_n \le t \le t_n + \theta_n.$$

In order to approximate $\overline{Y}(t; t_n)$ for $t_n \le t \le t_n + \theta_n$, we consider

$$\frac{\mathrm{d}y^n}{\mathrm{d}t} = \frac{\alpha}{y^n} - \frac{k}{2} y^n, \qquad y^n(t_n) = \overline{Y}(t_n; t_n) = \sqrt{\overline{V}(t_n)}. \tag{50}$$

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

By repeating the procedure (39)–(46), we arrive at

$$\sqrt{\overline{V}(t)} := y^n(t) + \frac{\sigma}{2}(w(t) - w(t_n)), \qquad t_n \le t \le t_{n+1},$$

satisfying

$$\left| \sqrt{V(t)} - \sqrt{\overline{V}(t)} \right| = |R^n(t)| \le r \sum_{m=0}^{n} \left(D_1 + \frac{D_2}{\overline{V}(t_m)} \right)(t_{m+1} - t_m), \qquad t_n \le t \le t_{n+1}, \tag{51}$$

with $R^n(t) := Y(t; t_n) - y^n(t)$, $t_n \le t \le t_{n+1}$, and, in particular,

$$\left| \sqrt{V(t_{n+1})} - \sqrt{\overline{V}(t_{n+1})} \right| \le r \sum_{m=0}^{n} \left(D_1 + \frac{D_2}{\overline{V}(t_m)} \right)(t_{m+1} - t_m). \tag{52}$$

Proposition 4. *Let the initial value $\sqrt{V(t_0)}$ be known with accuracy ε, i.e. the known $\sqrt{\overline{V}(t_0)}$ is such that*

$$\left| \sqrt{V(t_0)} - \sqrt{\overline{V}(t_0)} \right| \le \varepsilon, \tag{53}$$

and let $\sqrt{\overline{V}(t_m)} \ge \eta \ge \sigma r,\ m = 0, 1, \dots, n$. Then

$$\left| \sqrt{V(t_{n+1})} - \sqrt{\overline{V}(t_{n+1})} \right| \le \varepsilon + r \sum_{m=0}^{n} \left(D_1 + \frac{D_2}{\eta^2} \right)(t_{m+1} - t_m), \tag{54}$$

where $V(t_{n+1}) = V_{t_0, V(t_0)}(t_{n+1})$.

Proof. Inequality (54) follows from (52) with $V(t_{n+1}) = V_{t_0, \bar{V}(t_0)}(t_{n+1})$, (53), and (see Proposition 2)

$$\left| \sqrt{V_{t_0, V(t_0)}(t_{n+1})} - \sqrt{V_{t_0, \bar{V}(t_0)}(t_{n+1})} \right| \le \left| \sqrt{V(t_0)} - \sqrt{\overline{V}(t_0)} \right|.$$

Remark 1. In principle it is possible to use the distribution function $\mathcal{Q}$ (see (73)) for constructing $\sqrt{\overline{V}(t)}$ for $t_n < t < t_{n+1}$. However, we would rather consider for $t_n \le t \le t_{n+1}$ the approximation

$$\sqrt{\widetilde{V}(t)} := y^n(t) + \frac{\sigma}{2}\widetilde{w}_n(t), \qquad t_n \le t \le t_{n+1},$$

where

(a) for $t_{n+1} < t_0 + T$, $\widetilde{w}$ is an arbitrary continuous function satisfying

$$\widetilde{w}(t_n) = 0, \qquad \widetilde{w}(t_{n+1}) = w(t_{n+1}) - w(t_n) = r\xi_n, \qquad \max_{t_n \le t \le t_{n+1}} |\widetilde{w}_n(t)| \le r;$$

(b) for $t_{n+1} = t_0 + T$, one may take $\widetilde{w}(t) \equiv 0$. As a result we have, similar to (74), an insignificant increase of the error

$$\left| \sqrt{V(t)} - \sqrt{\widetilde{V}(t)} \right| \le r \sum_{m=0}^{n} \left(D_1 + \frac{D_2}{\overline{V}(t_m)} \right)(t_{m+1} - t_m) + \sigma r, \qquad t_n < t < t_{n+1}.$$

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

Let us consolidate the above procedure in a concise way.

Algorithm 1. (*The first simulation algorithm.*)

- Initialize $n := 0$; $t_n := t_0$; $\sqrt{\overline{V}(t_n)} = \sqrt{V(t_0)}$; $\Delta := \sigma r$;

(*) While $\sqrt{\overline{V}(t_n)} \geq \Delta$ and $t_n < t_0 + T$ do

  - simulate an independent random variable ξ_n with $\mathbb{P}(\xi_n = \pm 1) = \frac{1}{2}$, and θ_n as described in Appendix B. If $t_n + \theta_n < t_0 + T$, set $t_{n+1} = t_n + \theta_n$, else set $t_{n+1} = t_0 + T$;
  - solve (50) on the interval $[t_n, t_{n+1}]$ with solution y^n and set

$$\sqrt{\overline{V}(t_{n+1})} = y^n(t_{n+1}) + \frac{\sigma}{2}\begin{cases} r\xi_n & \text{if } t_{n+1} < t_0 + T, \\ 0 & \text{if } t_{n+1} = t_0 + T; \end{cases}$$

  - $t_n^{\text{new}} := t_{n+1}$; $\sqrt{\overline{V}(t_n^{\text{new}})} := \sqrt{\overline{V}(t_{n+1})}$; $n^{\text{new}} := n + 1$.

So, under assumption (47) we obtain estimate (51) (possibly enlarged with a term σr). The next theorem shows that if a trajectory of $V(t)$ under consideration is positive on $[t_0, t_0 + T]$ then the algorithm is convergent on this trajectory. We recall that in the $2k\lambda \geq \sigma^2$ case almost all trajectories are positive; hence, in this case the proposed method is almost surely convergent.

Theorem 2. *Let $4k\lambda \geq \sigma^2$ (i.e. $\alpha \geq 0$). Then for any positive trajectory $V(t) > 0$ on $[t_0, t_0+T]$ the proposed method is convergent on this trajectory. In particular, there exist $\eta > 0$ depending on the trajectory $V(\cdot)$ only, and $r_0 > 0$ depending on η such that*

$$\sqrt{\overline{V}(t_m)} \geq \eta \geq r\sigma, \qquad m = 0, 1, 2, \ldots, \textit{ for any } r < r_0.$$

So, in particular (47) is fulfilled for all $m = 0, 1, \ldots,$ and the estimate (51) implies that, for any $r < r_0$,

$$\left| \sqrt{V(t)} - \sqrt{\overline{V}(t)} \right| \leq r\left(D_1 + \frac{D_2}{\eta^2}\right)T, \qquad t_0 \leq t \leq t_0 + T,$$

and (see Remark 1)

$$\left| \sqrt{V(t)} - \sqrt{\widetilde{V}(t)} \right| \leq r\left(D_1 + \frac{D_2}{\eta^2}\right)T + \sigma r, \qquad t_0 \leq t \leq t_0 + T.$$

Proof. Let us define $\eta := \frac{1}{2}\min_{t_0 \leq t \leq t_0+T} \sqrt{V(t)}$ and

$$r_0 := \min\left(\frac{\eta}{\sigma}, \frac{\eta}{(D_1 + D_2/\eta^2)T}\right), \tag{55}$$

and let $r < r_0$. We then claim that for all m,

$$\sqrt{\overline{V}(t_m)} \geq \eta \geq r\sigma. \tag{56}$$

For $m = 0$, we trivially have

$$\sqrt{\overline{V}(t_0)} = \sqrt{V(t_0)} \geq 2\eta \geq \eta \geq r_0\sigma \geq r\sigma.$$

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

Now suppose by induction that $\sqrt{\overline{V}(t_j)} \geq \eta$ for $j = 0, \ldots, m$. Then due to (52), we have

$$\left| \sqrt{V(t_{m+1})} - \sqrt{\overline{V}(t_{m+1})} \right| \leq r\left(D_1 + \frac{D_2}{\eta^2}\right)T \leq r_0\left(D_1 + \frac{D_2}{\eta^2}\right)T \leq \eta$$

because of (55). Thus, since $\sqrt{V(t_{m+1})} \geq 2\eta$, it follows that $\sqrt{\overline{V}(t_{m+1})} \geq \eta \geq r\sigma$. This proves (56) and the convergence for $r \downarrow 0$.

Remark 2. In the case where $4k\lambda \geq \sigma^2 > 2k\lambda$ trajectories will reach 0 with positive probability, i.e. convergence on such trajectories is not guaranteed by Theorem 2. So it is important to develop some method for continuing the simulations in cases of very small $\overline{V}(t_m)$. One can propose different procedures, for instance, one can proceed with standard SDE approximation methods relying on some known scheme suitable for small V (see, for example, [3]). However, the uniformity of the simulation would be destroyed in this way. We therefore propose in the next section a uniform simulation method that may be started in a value $\overline{V}(t_m) \geq 0$ close to 0.

6. Simulation of trajectories close to 0 and the main algorithm

The simulation algorithm in Section 5 has a drawback. Even in the $2k\lambda \geq \sigma^2$ case, where all the trajectories $V(t)$ are positive, we cannot ensure that after a choice of r the requirement $\sqrt{\overline{V}(t_m)} \geq \sigma r$ will be fulfilled for all m. Of course, in principle it is possible to decrease r when the trajectory approaches 0 (i.e. using an adaptive algorithm with variable expected time step). However, such an algorithm can be very expensive on the parts of the trajectories that get close to 0. This is because the smaller r, the smaller the expected passage time $\mathbb{E}(\theta_n) = r^2$ (see Remark 3) and so such parts may require a large number of steps. We therefore propose an alternative procedure if we enter at some random step $\mathfrak{m}$ a band $(0, \Delta)$ of width $\Delta > 0$ (to be specified later), i.e.

$$\sqrt{\overline{V}(t_k)} \geq \Delta, \quad k = 0, 1, \ldots, \mathfrak{m} - 1, \qquad \sqrt{\overline{V}(t_{\mathfrak{m}})} < \Delta. \tag{57}$$

Starting from $(t_{\mathfrak{m}}, \overline{V}(t_{\mathfrak{m}}))$ we now make a time step $\vartheta = \vartheta_{\overline{V}(t_{\mathfrak{m}})}$ (that may be comparatively large), such that

$$\sqrt{V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t)} < 2\Delta, \quad t_{\mathfrak{m}} \leq t < t_{\mathfrak{m}} + \vartheta, \qquad \sqrt{V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t_{\mathfrak{m}} + \vartheta)} = 2\Delta. \tag{58}$$

That is, $t_{\mathfrak{m}} + \vartheta_{\overline{V}(t_{\mathfrak{m}})}$ is the first-passage time of the trajectory $\sqrt{V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t)}$ to the upper bound of the band $(0, 2\Delta)$. One may think of Δ being large enough compared to r, but at the same time small enough in order to reach a certain accuracy. For example, $\Delta = Ar^a$, $0 < a < \frac{1}{2}$, where A is a positive constant. Although we do not know the trajectory $V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t)$ on the interval $t_{\mathfrak{m}} < t < t_{\mathfrak{m}} + \vartheta$, we do know that it satisfies (58), and we know its values $\sqrt{\overline{V}(t_{\mathfrak{m}})}$ and $\sqrt{V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t_{\mathfrak{m}} + \vartheta)} = 2\Delta$ at the ends of the interval. So, we take, for example, a straight line $L(t)$ that connects the points $(t_{\mathfrak{m}}, \sqrt{\overline{V}(t_{\mathfrak{m}})})$ and $(t_{\mathfrak{m}} + \vartheta, 2\Delta)$, i.e.

$$L(t) := \frac{t - t_{\mathfrak{m}}}{\vartheta} 2\Delta + \frac{t_{\mathfrak{m}} + \vartheta - t}{\vartheta} \sqrt{\overline{V}(t_{\mathfrak{m}})}, \qquad t_{\mathfrak{m}} \leq t \leq t_{\mathfrak{m}} + \vartheta, \tag{59}$$

as an approximation to the unknown $\sqrt{V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t)}$ on the interval $(t_{\mathfrak{m}}, t_{\mathfrak{m}} + \vartheta)$, and set $\sqrt{\overline{V}(t)} := L(t)$, $t_{\mathfrak{m}} \leq t \leq t_{\mathfrak{m}} + \vartheta$. For this approximation, we have the error estimate

$$\left| \sqrt{\overline{V}(t)} - \sqrt{V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t)} \right| \leq 2\Delta \quad \text{for } t_m < t < t_m + \vartheta, \tag{60}$$

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

while at times t_{m} and $t_{\mathrm{m}} + \vartheta$ this error is 0 due to

$$\sqrt{\overline{V}(t_{\mathrm{m}} + \vartheta)} = \sqrt{V_{t_{\mathrm{m}},\overline{V}(t_{\mathrm{m}})}(t_{\mathrm{m}} + \vartheta)} = 2\Delta. \tag{61}$$

By Proposition 2 (we assume that $\alpha > 0$) on the other hand, we have

$$\left| \sqrt{V_{t_{\mathrm{m}},\overline{V}(t_{\mathrm{m}})}(t)} - \sqrt{V_{t_{\mathrm{m}},V(t_{\mathrm{m}})}(t)} \right| \le \left| \sqrt{\overline{V}(t_{\mathrm{m}})} - \sqrt{V(t_{\mathrm{m}})} \right|, \qquad t_{\mathrm{m}} \le t \le t_{\mathrm{m}} + \vartheta. \tag{62}$$

Combining (60) and (62) yields

$$\left| \sqrt{\overline{V}(t)} - \sqrt{V_{t_{\mathrm{m}},V(t_{\mathrm{m}})}(t)} \right| \le 2\Delta + \left| \sqrt{\overline{V}(t_{\mathrm{m}})} - \sqrt{V(t_{\mathrm{m}})} \right|, \qquad t_{\mathrm{m}} < t < t_{\mathrm{m}} + \vartheta,$$

while at time $t_{\mathrm{m}} + \vartheta$ by (61) and (62), we have

$$\left| \sqrt{\overline{V}(t_{\mathrm{m}} + \vartheta)} - \sqrt{V_{t_{\mathrm{m}},V(t_{\mathrm{m}})}(t_{\mathrm{m}} + \vartheta)} \right| \le \left| \sqrt{\overline{V}(t_{\mathrm{m}})} - \sqrt{V(t_{\mathrm{m}})} \right|.$$

In other words, the error of $\sqrt{\overline{V}}$ at the time $t_{\mathrm{m}} + \vartheta$ of passing the band is not larger than the error at t_{m} when $\sqrt{\overline{V}}$ entered the band. That is, the error does not accumulate when $\sqrt{\overline{V}}$ passes through the band $(0, 2\Delta)$. This property is a key feature in our construction.

6.1. The main simulation algorithm and the main convergence theorem

Algorithm 2. The arguments above result in the following (pseudo) algorithm.

Let r and Δ be numbers such that $\Delta \ge \sigma r$.

- Initialize $n := 0$; $t_n := t_0$; $\sqrt{\overline{V}(t_n)} = \sqrt{V(t_0)}$; choose $\Delta > \sigma r$ properly (see below).

(∗∗) Run the first simulation algorithm of Section 5 from (∗);

- set $\mathrm{m} := n^{\mathrm{new}}$; $t_{\mathrm{m}} = t_n^{\mathrm{new}}$;
- if $t_{\mathrm{m}} = t_0 + T$ then finish the simulation;
- if $t_{\mathrm{m}} < t_0 + T$ simulate $\vartheta_{\overline{V}(t_{\mathrm{m}})}$ according to Appendix C;
- if $t_{\mathrm{m}} + \vartheta_{\overline{V}(t_{\mathrm{m}})} \ge t_0 + T$ set $\sqrt{\overline{V}(t)} = \sqrt{\overline{V}(t_{\mathrm{m}})}$ on $[t_{\mathrm{m}}, t_0 + T]$ and finish;
- if $t_{\mathrm{m}} + \vartheta_{\overline{V}(t_{\mathrm{m}})} < t_0 + T$ set $\sqrt{\overline{V}(t)} = L(t)$ (see (59)) on $[t_{\mathrm{m}}, t_{\mathrm{m}} + \vartheta_{\overline{V}(t_{\mathrm{m}})}]$; set $t_n^{\mathrm{new}} := t_{\mathrm{m}} + \vartheta_{\overline{V}(t_{\mathrm{m}})}$; $\sqrt{\overline{V}(t_n^{\mathrm{new}})} = 2\Delta$; Go to (∗∗);

Let us now consider the convergence properties of the main algorithm. Suppose that for a generic point t_n at (∗), we have $|\sqrt{\overline{V}(t_n)} - \sqrt{V(t_n)}| \le \varepsilon_n$ (obviously we may take $\varepsilon_0 = 0$). Then the aggregated error of $\sqrt{\overline{V}(t_{\mathrm{m}})}$ is estimated by (see Proposition 4)

$$\left| \overline{V}(t_{\mathrm{m}}) - \sqrt{V(t_{\mathrm{m}})} \right| \le \varepsilon_n + r \sum_{k=n}^{\mathrm{m}-1} \left(D_1 + \frac{D_2}{\overline{V}(t_k)} \right) (t_{k+1} - t_k) =: \varepsilon_n^{\mathrm{new}}.$$

Assuming that $t_{\mathrm{m}} + \vartheta_{\overline{V}(t_{\mathrm{m}})} < t_0 + T$ (the other case is similar), the error of $\sqrt{\overline{V}(t)}$ on $(t_{\mathrm{m}}, t_{\mathrm{m}} + \vartheta_{\overline{V}(t_{\mathrm{m}})})$, before executing (∗) the next time, is thus estimated by

$$2\Delta + \left| \sqrt{\overline{V}(t_{\mathrm{m}})} - \sqrt{V(t_{\mathrm{m}})} \right| \le 2\Delta + \varepsilon_n^{\mathrm{new}},$$

while the error at t_n^{new} is estimated by $\varepsilon_n^{\mathrm{new}}$. The following theorem is now obvious from the above constructions.

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

Theorem 3. *Let $\alpha > 0$. The above algorithm constructs $\sqrt{\overline{V}(t)}$ on $[t_0, t_0+T]$. It is completed in a finite number of steps with probability 1. The error on $[t_0, t_0+T]$ is estimated by*

$$\left| \sqrt{\overline{V}(t)} - \sqrt{V(t)} \right| \leq 2\Delta + r\left(D_1 + \frac{D_2}{\Delta^2}\right)T. \tag{63}$$

Moreover, the error in $[t_0, t_0+T] \setminus \cup_{t_\mathfrak{m}} (t_\mathfrak{m}, t_\mathfrak{m} + \vartheta_{\overline{V}(t_\mathfrak{m})})$ is estimated by

$$\left| \sqrt{\overline{V}(t)} - \sqrt{V(t)} \right| \leq r\left(D_1 + \frac{D_2}{\Delta^2}\right)T.$$

By the (in a sense) optimal choice $\Delta = Ar^{1/3}$, error (63) is of $O(r^{1/3})$ and the algorithm converges for $r \downarrow 0$.

7. Numerical implementation and some applications

In this section we discuss the numerical implementation of the main algorithm and its merits in some possible applications (for example, in finance where $\sqrt{V}$ may be interpreted as the volatility of a Heston asset price model). However, we underline that an in-depth numerical treatment is beyond the scope of this paper. Let us assume that we need to evaluate the expectation functional

$$\mathbb{E}(f(V_{t_0,V_0}(t)\colon t_0 \leq t \leq t_0 + T)), \tag{64}$$

where f is a function that depends on the whole trajectory of V_{t_0,V_0}.

Now, for instance, suppose that f in (64) does not depend on the parts of the trajectory that are below a certain level l, $l > 0$. A simple example is

$$f(V_{t_0,V_0}(t)\colon t_0 \leq t \leq t_0 + T) = \max_{t_0 \leq t \leq t_0+T} V_{t_0,V_0}(t)$$

with $l = V_0$. We may then choose $\Delta = \sqrt{V_0}/2$ in Algorithm 2; thus, yielding a uniform convergence rate $O(r)$ in any case of $\alpha > 0$ (cf. Theorem 3) for those parts of the trajectories where the function f is sensitive to. Put differently, the particular (uniform) accuracy of the parts of V_{t_0,V_0} below V_0 is irrelevant for the functional f.

Another (financial) example is a call option with strike K on realized volatility upon a certain level l, $l > 0$,

$$C_{t_0} := \mathbb{E}\left(\int_{t_0}^{t_0+T} V_{t_0,V_0}(s)\, \mathbf{1}_{\{V \geq l\}}\, \mathrm{d}s - K\right)^+,$$

where $\mathbf{1}$ is the indicator function (for simplicity the interest rate is assumed to be 0). Note that in a Heston model the integrated volatility process $\int_{t_0}^t V(s)\,\mathrm{d}s$ being the quadratic variation of the log-asset price process is indeed observable, and so is V (at least in principle). For this example we may fix $\Delta = \sqrt{l}/2$ in Algorithm 2 and then a similar remark as in the previous example regarding accuracy applies.

In the general case, for example in the case of the general f above, it is advantageous to choose A in Algorithm 2 according to Theorem 3 for a given choice of r in an optimal way. That is, with $\Delta = Ar^{1/3}$ we have to minimize the global error

$$2Ar^{1/3} + r\left(D_1 + \frac{D_2}{A^2 r^{2/3}}\right)T = \left(2A + \frac{D_2 T}{A^2}\right)r^{1/3} + rD_1T.$$

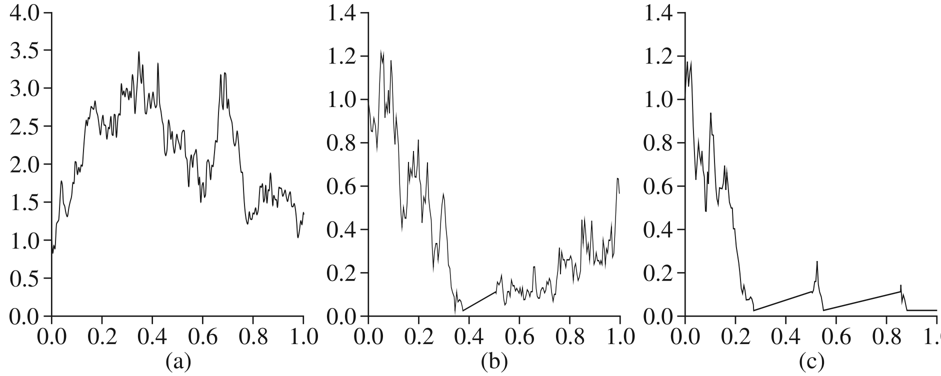

FIGURE 1: Three sample trajectories (a), (b), and (c) that enter the band $(0, \Delta^2) = (0, 0.0283)$ and continue linearly until the level $4\Delta^2 = 0.11318$, zero, one, and three times, respectively.

Thus, $A = (D_2 T)^{1/3} = (4\alpha\sigma \mathrm{e}^{(k/2)T}/3)^{1/3}$ (see Corollary 1) is a suitable choice since $r \ll r^{1/3}$ when r is small.

Some illustrative examples. We have implemented Algorithm 2 for the following CIR parameters:

$$k = \lambda = T = 1, \qquad t_0 = 0, \qquad \sigma = \sqrt{3};$$

hence, $2k\lambda < \sigma^2 < 4k\lambda$, and $\alpha = \frac{1}{8}$. In the algorithm we choose $r = 0.01$ and $\Delta = 0.168\,21$ determined in the above way. In Figure 1 we illustrate some typical trajectories of V. Figure 1(a) depicts a trajectory of V that does not enter the band $(0, \Delta^2) = (0, 0.028\,3)$; hence, it follows Algorithm 1 until T. In Figure 1(b) the trajectory enters the band once, continues linearly until the level $4\Delta^2 = 0.113\,18$, and then follows Algorithm 1 until T. In Figure 1(c), the trajectory enters the band three times.

Appendix A. Addendum to the proof of Proposition 2

It is known that for $\delta > 1$, the Bessel process BES^δ is the unique process that satisfies the integral representation

$$Z(t) = Z(0) + \frac{\delta - 1}{2} \int_0^t \frac{1}{Z(s)}\,\mathrm{d}s + W(t), \qquad 0 \le t < \infty, \tag{65}$$

where W is a standard Brownian motion, $Z(0) > 0$, $Z(t) \ge 0$ almost surely, and that, in particular, $\mathbb{E}(\int_0^t (1/Z(s))\,\mathrm{d}s) < \infty$; see [10, Appendix A1] and [20, Exercise 1.26, Chapter XI]. (For $\delta \le 1$ the representation of BES^δ is less simple and involves the concept of local time.) From this fact we will show that for $\alpha > 0$ the solution of (8) may be represented as

$$U(t) = U(t_0) + \int_{t_0}^t \left(\frac{\alpha}{U(s)} - \frac{k}{2}U(s)\right)\mathrm{d}s + \frac{\sigma}{2}(w(t) - w(t_0)), \qquad U(t_0) > 0,\ t_0 \le t < \infty. \tag{66}$$

Let us consider

$$U(t) = \mathrm{e}^{-k(t-t_0)/2} Z\left(\frac{\sigma^2}{4k}(\mathrm{e}^{k(t-t_0)} - 1)\right), \qquad t_0 \le t < \infty, \tag{67}$$

where Z is the solution of (65) with $\delta = 4k\lambda/\sigma^2 > 1$, and $Z(0) = U(t_0) > 0$ (cf. [10] and the references therein). Note that the function $h\colon t \to \sigma^2(\mathrm{e}^{k(t-t_0)} - 1)/(4k)$ satisfies $h(t_0) = 0$, and that it is smooth and strictly increasing since $k > 0$. Let us further introduce

$$\widetilde{W}(t) - \widetilde{W}(t_0) := \frac{2}{\sigma}\mathrm{e}^{-k(t-t_0)/2}W(h(t)) + \frac{k}{\sigma}\int_{t_0}^{t} \mathrm{e}^{-k(s-t_0)/2}W(h(s))\,\mathrm{d}s. \tag{68}$$

Obviously, $\widetilde{W}(t)$, $t \geq t_0$, is a zero mean Gaussian process and moreover by a straightforward computation it can be shown that $\mathbb{E}(\widetilde{W}(t) - \widetilde{W}(t_0))^2 = t - t_0$ for all $t \geq t_0$. Indeed, by straightforward algebra and using the definition of h, we obtain

$$\begin{aligned}
\mathbb{E}(\widetilde{W}(t) - \widetilde{W}(t_0))^2 &= \frac{4}{\sigma^2}\mathrm{e}^{-k(t-t_0)}h(t) + \frac{4k}{\sigma^2}\mathrm{e}^{-k(t-t_0)/2}\int_{t_0}^{t} \mathrm{e}^{-k(s-t_0)/2}\min(h(t), h(s))\,\mathrm{d}s \\
&\quad + \frac{k^2}{\sigma^2}\int_{t_0}^{t}\int_{t_0}^{t} \mathrm{e}^{-k(s-t_0)/2}\mathrm{e}^{-k(\widetilde{s}-t_0)/2}\min(h(s), h(\widetilde{s}))\,\mathrm{d}s\,\mathrm{d}\widetilde{s} \\
&= \frac{1}{k}(1 - \mathrm{e}^{-k(t-t_0)}) + \frac{4k}{\sigma^2}\mathrm{e}^{-k(t-t_0)/2}\int_{t_0}^{t} \mathrm{e}^{-k(s-t_0)/2}h(s)\,\mathrm{d}s \\
&\quad + \frac{2k^2}{\sigma^2}\int_{t_0}^{t} \mathrm{e}^{-k(\widetilde{s}-t_0)/2}d\widetilde{s}\int_{t_0}^{\widetilde{s}} \mathrm{e}^{-k(s-t_0)/2}h(s)\,\mathrm{d}s \\
&= t - t_0.
\end{aligned}$$

That is, $\widetilde{W}$ is a Brownian motion adapted to its own filtration. Then using the fact that

$$\frac{\delta - 1}{2}\int_0^{h(s)} \frac{1}{Z(u)}\,\mathrm{d}u = \frac{\delta - 1}{2}\int_{t_0}^{s} \frac{1}{Z(h(r))}h'(r)\,\mathrm{d}r = \int_{t_0}^{s} \frac{\alpha}{U(r)}\mathrm{e}^{k(r-t_0)/2}\,\mathrm{d}r,$$

from (65) and (67), we obtain

$$U(s) = \mathrm{e}^{-k(s-t_0)/2}W(h(s)) + \mathrm{e}^{-k(s-t_0)/2}U(t_0) + \mathrm{e}^{-k(s-t_0)/2}\int_{t_0}^{s} \frac{\alpha}{U(r)}\mathrm{e}^{k(r-t_0)/2}\,\mathrm{d}r. \tag{69}$$

Thus, by (68) and (69), it holds that

$$\begin{aligned}
U(t) &= \mathrm{e}^{-k(t-t_0)/2}U(t_0) + \mathrm{e}^{-k(t-t_0)/2}\int_{t_0}^{t} \frac{\alpha}{U(r)}\mathrm{e}^{k(r-t_0)/2}\,\mathrm{d}r \\
&\quad + \frac{\sigma}{2}(\widetilde{W}(t) - \widetilde{W}(t_0)) - \frac{k}{2}\int_{t_0}^{t} \mathrm{e}^{-k(s-t_0)/2}W(h(s))\,\mathrm{d}s \\
&= U(t_0) - \frac{k}{2}\int_{t_0}^{t} U(s)\,\mathrm{d}s + \frac{\sigma}{2}(\widetilde{W}(t) - \widetilde{W}(t_0)) \\
&\quad + \mathrm{e}^{-k(t-t_0)/2}\int_{t_0}^{t} \frac{\alpha}{U(r)}\mathrm{e}^{k(r-t_0)/2}\,\mathrm{d}r + \frac{k}{2}\int_{t_0}^{t} \mathrm{e}^{-k(s-t_0)/2}\,\mathrm{d}s\int_{t_0}^{s} \frac{\alpha}{U(r)}\mathrm{e}^{k(r-t_0)/2}\,\mathrm{d}r.
\end{aligned} \tag{70}$$

In particular, the Lebesgue integral $\int_{t_0}^{t}(\alpha/U(r))\mathrm{e}^{k(r-t_0)/2}\,\mathrm{d}r$ is almost surely an absolutely continuous function in t on $[t_0, t_0 + T]$. Hence, it is everywhere differentiable except for a set of Lebesgue measure 0, and its derivative is equal to $\alpha e^{k(t-t_0)/2}/U(t)$. From this, it follows that the sum of the two last terms in (70) can be expressed as $\int_{t_0}^{t}(\alpha/U(r))\,dr$. So, we arrive at

$$U(t) = U(t_0) + \int_{t_0}^{t}\left(\frac{\alpha}{U(r)} - \frac{k}{2}U(r)\right)dr + \frac{\sigma}{2}\left(\widetilde{W}(t) - \widetilde{W}(t_0)\right). \tag{71}$$

From (71), the representation (66) follows for $w = \widetilde{W}$. In particular, the pair $(U, \widetilde{W})$ in (71) may be considered as a strong SDE solution on the probability space that $\widetilde{W}$ is living on. We now argue that such a solution is unique. If there would exist two different strong solutions $(U_1, \widetilde{W})$ and $(U_2, \widetilde{W})$ with coinciding initial values $U_1(t_0) = U_2(t_0) =: U(t_0) > 0$, then the reverse procedure

$$Z_i(s) := e^{k(h^{-1}(s)-t_0)/2}U_i(h^{-1}(s)), \qquad 0 \le s < \infty,\ i = 1, 2, \tag{72}$$

would similarly yield two strong solutions (Z_1, W) and (Z_2, W) of (65) with $Z_1(0) = Z_2(0) = U(t_0)$ with respect to some (though from different $\widetilde{W}$) Brownian motion W. However, by the uniqueness of the solution to (65), it will follow that $Z_1 = Z_2$, and then from (72) that $U_1 = U_2$.

Finally, with $Y(t) = U(t) - (\sigma/2)(w(t) - w(t_0))$, it holds that

$$Y(t) = Y(t_0) + \int_{t_0}^{t}\left(\frac{\alpha}{Y(s) + (\sigma/2)(w(s) - w(t_0))} - \frac{k}{2}\left(Y(s) + \frac{\sigma}{2}(w(s) - w(t_0))\right)\right)ds$$

for $Y(0) = U(0) > 0$, $0 \le t < \infty$, and that, in particular, Y is an absolutely continuous function. From this, it follows that (20) holds for $t_0 \le t \le t_0 + T$ when $\alpha > 0$ and $\varphi(t) = w(t) - w(t_0)$ is a Brownian trajectory, and then (19) in Lemma 1 goes through for $\theta = T$.

Appendix B. Simulation of θ_m and ζ_m

In simulating θ_m we utilize the distribution function

$$\mathcal{P}(t) := \mathbb{P}(\tau < t),$$

where τ is the first-passage time of the Wiener process $W(t)$ to the boundary of the interval $[-1, 1]$. A very accurate approximation $\tilde{\mathcal{P}}(t)$ of $\mathcal{P}(t)$ is as follows:

$$\mathcal{P}(t) \simeq \tilde{\mathcal{P}}(t) = \int_0^t \tilde{\mathcal{P}}'(s)\,ds$$

with

$$\tilde{\mathcal{P}}'(t) = \begin{cases} \dfrac{2}{\sqrt{2\pi t^3}}(e^{-1/2t} - 3e^{-9/2t} + 5e^{-25/2t}), & 0 < t \le 2/\pi, \\ \dfrac{\pi}{2}(e^{-\pi^2 t/8} - 3e^{-9\pi^2 t/8} + 5e^{-25\pi^2 t/8}), & t > 2/\pi, \end{cases}$$

and it holds that

$$\sup_{t\ge 0}|\tilde{\mathcal{P}}'(t) - \mathcal{P}'(t)| \le 2.13 \times 10^{-16}, \qquad \sup_{t\ge 0}|\tilde{\mathcal{P}}(t) - \mathcal{P}(t)| \le 7.04 \times 10^{-18}$$

(for details, see [18, Chapter 5, Section 3 and Appendix A3]). Now simulate a random variable U uniformly distributed on $[0, 1]$. Then compute $\tau = \mathcal{P}^{-1}(U)$, which is distributed according to $\mathscr{P}$. That is, we have to solve $\tilde{P}(\tau) = U$, for instance, by Newton's method or any other efficient solving routine. Next set $\theta_m = r^2\tau_m$.

For simulating ζ_m in (33), we observe that (34) is equivalent with

$$rW_{r^{-2}(t_0+T-t_m)} \text{ conditional on } \max_{0\le u\le r^{-2}(t_0+T-t_m)} |W_u| \le 1.$$

We next sample ϑ from the distribution function $\mathcal{Q}(x; r^{-2}(t_0+T-t_m))$, where $\mathcal{Q}(x;t)$ is the known conditional distribution function (see [18, Chapter 5, Section 3])

$$\mathcal{Q}(x;t) := \mathbb{P}(W(t) < x \mid \max_{0\le s\le t}|W(s)| < 1), \qquad -1 \le x \le 1, \tag{73}$$

and set $\zeta_m = r\vartheta$. The simulation of the last step looks rather complicated and may be computationally expensive. However, it is possible to take, for $w(t_0+T) - w(t_\nu)$, simply any value between $-r$ and r, for example 0. This may enlarge the one-step error on the last step but does not influence the convergence order of the elaborated method. Indeed, if we set $w(t_0+T) - w(t_\nu)$ to be 0, for instance on the last step, we obtain $\sqrt{\overline{V}(t_0+T)} = y^\nu(t_0+T)$ instead of (33), and

$$|\sqrt{V(t_0+T)} - \sqrt{\overline{V}(t_0+T)}| \le r\sum_{m=0}^{\nu-1}\bigg(D_1 + \frac{D_2}{\overline{V}(t_m)}\bigg)(t_{m+1}-t_m) + \sigma r. \tag{74}$$

Remark 3. We have in any step $\mathbb{E}(\theta_n) = r^2$, the random number of steps before reaching t_0+T, say ν, is finite with probability 1, and $\mathbb{E}(\nu) = O(1/r^2)$. For details see [18, Chapter 5]. In a heuristic sense this means that if we have convergence of order $O(r)$, we obtain accuracy $O(\sqrt{h})$ for an (expected) number of steps $O(1/h)$ similar to the standard Euler scheme.

Appendix C. Simulation of ϑ_x

In order to implement the simulation method for trajectories near 0 we have to find the distribution function of $\vartheta_x = \vartheta_{x,l}$, where $\vartheta_{x,l}$ is the first-passage time of the trajectory $X_{0,x}(s)$ to the level l. For this it is more convenient to change notation and to write (1) in the form

$$\mathrm{d}X(s) = k(\lambda - X(s))\,\mathrm{d}s + \sigma\sqrt{X}\,\mathrm{d}w(s), \qquad X(0) = x, \tag{75}$$

where without loss of generality, we take the initial time to be $s = 0$. The function

$$u(t,x) := \mathbb{P}(\vartheta_{x,l} < t)$$

is the solution of the first boundary value problem of parabolic type ([18, Chapter 5, Section 3])

$$\frac{\partial u}{\partial t} = \frac{1}{2}\sigma^2 x\frac{\partial^2 u}{\partial x^2} + k(\lambda - x)\frac{\partial u}{\partial x}, \qquad t > 0,\ 0 < x < l, \tag{76}$$

with initial data

$$u(0,x) = 0, \tag{77}$$

and boundary conditions

$$u(t,0) \text{ is bounded}, \qquad u(t,l) = 1. \tag{78}$$

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

To obtain homogeneous boundary conditions, we introduce $v = u - 1$. The function v then satisfies

$$\frac{\partial v}{\partial t} = \frac{1}{2}\sigma^2 x \frac{\partial^2 v}{\partial x^2} + k(\lambda - x)\frac{\partial v}{\partial x}, \qquad t > 0,\ 0 < x < l, \tag{79}$$

$$v(0, x) = -1, \qquad v(t, 0) \text{ is bounded}, \qquad v(t, l) = 0. \tag{80}$$

The problem in (79) and (80) can be solved by the method of separation of variables. In this way the Sturm–Liouville problem for the confluent hypergeometric equation (the Kummer equation) arises. This problem is rather complicated however. Below we are going to solve an easier problem as a good approximation to (79) and (80). Along with (75), let us consider the equations

$$\mathrm{d}X^+(s) = k\lambda\,\mathrm{d}s + \sigma\sqrt{X^+}\,\mathrm{d}w(s), \qquad X^+(0) = x,$$

$$\mathrm{d}X^-(s) = k(\lambda - l)\,\mathrm{d}s + \sigma\sqrt{X^-}\,\mathrm{d}w(s), \qquad X^-(0) = x,$$

with $0 \le l < \lambda$. It is not difficult to prove the following inequalities:

$$X^-(s) \le X(s) \le X^+(s). \tag{81}$$

According to (81), we consider three boundary value problems. First, (76)–(78) and then similar ones for the equations

$$\frac{\partial u^+}{\partial t} = \frac{1}{2}\sigma^2 x \frac{\partial^2 u^+}{\partial x^2} + k\lambda\frac{\partial u^+}{\partial x}, \qquad t > 0,\ 0 < x < l,$$

$$\frac{\partial u^-}{\partial t} = \frac{1}{2}\sigma^2 x \frac{\partial^2 u^-}{\partial x^2} + k(\lambda - l)\frac{\partial u^-}{\partial x}, \qquad t > 0,\ 0 < x < l. \tag{82}$$

From (81), it follows that

$$u^-(t, x) \le u(t, x) \le u^+(t, x);$$

hence,

$$v^-(t, x) \le v(t, x) \le v^+(t, x),$$

where $v^- = u^- - 1,\ v^+ = u^+ - 1$.

As the band $0 < x < l = A^2 r^{2a}$, for a certain $a > 0$, is narrow due to small enough r, the difference $v^+ - v^-$ will be small and so we can consider the following problem:

$$\frac{\partial v^+}{\partial t} = \frac{1}{2}\sigma^2 x \frac{\partial^2 v^+}{\partial x^2} + k\lambda\frac{\partial v^+}{\partial x}, \qquad t > 0,\ 0 < x < l, \tag{83}$$

$$v^+(0, x) = -1;\ v^+(t, 0) \text{ is bounded},\ v^+(t, l) = 0 \tag{84}$$

as good approximations of (79) and (80). Henceforth, we write $v := v^+$. By separation of variables we obtain as elementary independent solutions to (83), $\mathcal{T}(t)\mathcal{X}(x)$, where

$$\mathcal{T}'(t) + \mu\mathcal{T}(t) = 0, \quad \text{i.e. } \mathcal{T}(t) = \mathcal{T}_0 e^{-\mu t},\ \mu > 0, \tag{85}$$

$$\tfrac{1}{2}\sigma^2 x\mathcal{X}'' + k\lambda\mathcal{X}' + \mu\mathcal{X} = 0, \qquad \mathcal{X}(0+) \text{ is bounded},\ \mathcal{X}(l) = 0. \tag{86}$$

It can be verified straightforwardly that the solution of (86) can be obtained in terms of Bessel functions of the first kind (see, for example, [4]),

$$\mathcal{X}(x) = \mathcal{X}_\gamma^{\pm}(x) := x^\gamma J_{\pm 2\gamma}(\sigma^{-1}\sqrt{8\mu x}) = x^\gamma O(x^{\pm\gamma}) \qquad \text{if } x \downarrow 0$$

with

$$\gamma := \frac{1}{2} - \frac{k\lambda}{\sigma^2}. \tag{87}$$

Since $\mathcal{X}(x)$ has to be bounded for $x \downarrow 0$, we may take (regardless of the sign of γ)

$$\mathcal{X}(x) = \mathcal{X}_\gamma^-(x) =: \mathcal{X}_\gamma(x) = x^\gamma J_{-2\gamma}(\sigma^{-1}\sqrt{8\mu x}). \tag{88}$$

In our setting, we have $\alpha > 0$, i.e. $\gamma < \frac{1}{4}$.

The following derivation of a Fourier–Bessel series for v is standard but included for convenience of the reader. Denote the positive zeros of J_ν by $\pi_{\nu,m}$, for example,

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sin x, \qquad \pi_{1/2,m} = m\pi, \quad m = 1, 2, \dots. \tag{89}$$

Then the (homogeneous) boundary condition $\mathcal{X}_\gamma(l) = 0$ yields

$$\sigma^{-1}\sqrt{8\mu l} = \pi_{-2\gamma,m}, \quad \text{i.e. } \mu_m := \frac{\sigma^2 \pi^2_{-2\gamma,m}}{8l} \tag{90}$$

and we have

$$\mathcal{X}_{\gamma,m}(x) := x^\gamma J_{-2\gamma}(\sigma^{-1}\sqrt{8\mu_m x}) = x^\gamma J_{-2\gamma}\bigg(\pi_{-2\gamma,m}\sqrt{\frac{x}{l}}\bigg).$$

By the well-known orthogonality relation

$$\int_0^1 z J_{-2\gamma}(\pi_{-2\gamma,k} z) J_{-2\gamma}(\pi_{-2\gamma,k'} z)\,\mathrm{d}z = \frac{\delta_{k,k'}}{2} J^2_{-2\gamma+1}(\pi_{-2\gamma,k}),$$

we obtain, by setting $z = \sqrt{x/l}$,

$$\int_0^l J_{-2\gamma}\bigg(\pi_{-2\gamma,m}\sqrt{\frac{x}{l}}\bigg) J_{-2\gamma}\bigg(\pi_{-2\gamma,m'}\sqrt{\frac{x}{l}}\bigg)\,\mathrm{d}x = l\delta_{m,m'} J^2_{-2\gamma+1}(\pi_{-2\gamma,m});$$

hence,

$$\int_0^l \mathcal{X}_{\gamma,m}(x)\mathcal{X}_{\gamma,m'}(x) x^{-2\gamma}\,\mathrm{d}x = l\delta_{m,m'} J^2_{-2\gamma+1}(\pi_{-2\gamma,m}).$$

Now set

$$v(t,x) = \sum_{m=1}^\infty \beta_m \mathrm{e}^{-\mu_m t}\mathcal{X}_{\gamma,m}(x), \qquad 0 \le x \le l. \tag{91}$$

For $t = 0$ due to the initial condition $v(0,x) = -1$, we have

$$-1 = \sum_{m=1}^\infty \beta_m \mathcal{X}_{\gamma,m}(x).$$

So for any $p = 1, 2, \dots,$

$$-\int_0^l \mathcal{X}_{\gamma,p}(x) x^{-2\gamma}\,\mathrm{d}x = \beta_p l J^2_{-2\gamma+1}(\pi_{-2\gamma_\kappa,p}),$$

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

that is,

$$\beta_p = -\frac{\int_0^l \mathcal{X}_{\gamma,p}(x)x^{-2\gamma}\,\mathrm{d}x}{l J^2_{-2\gamma+1}(\pi_{-2\gamma,p})}. \tag{92}$$

Furthermore, it holds that

$$\begin{aligned}\int_0^l \mathcal{X}_{\gamma,p}(x)x^{-2\gamma}\,\mathrm{d}x &= \int_0^l x^{-\gamma} J_{-2\gamma}\biggl(\pi_{-2\gamma,p}\sqrt{\frac{x}{l}}\biggr)\,\mathrm{d}x \\ &= 2l^{-\gamma+1}\int_0^1 z^{-2\gamma+1} J_{-2\gamma}(\pi_{-2\gamma,p}z)\,\mathrm{d}z \\ &= 2l^{-\gamma+1}\frac{J_{-2\gamma+1}(\pi_{-2\gamma,p})}{\pi_{-2\gamma,p}}\end{aligned}$$

by well-known identities for Bessel functions (see, for example, [4]), and (92), thus, can be written as

$$\beta_p = -\frac{2}{l^{\gamma}\pi_{-2\gamma,p}J_{-2\gamma+1}(\pi_{-2\gamma,p})}, \qquad p = 1, 2, \ldots. \tag{93}$$

So, from $v = u - 1$, (85), (88), (90), (91), and (93), we finally obtain

$$u(t,x) = 1 - 2x^{\gamma}l^{-\gamma}\sum_{m=1}^{\infty}\frac{J_{-2\gamma}(\pi_{-2\gamma,m}\sqrt{x/l})}{\pi_{-2\gamma,m}J_{-2\gamma+1}(\pi_{-2\gamma,m})}\exp\biggl[-\frac{\sigma^2\pi^2_{-2\gamma,m}}{8l}t\biggr], \qquad 0 \le x \le l. \tag{94}$$

It should be noted that, in fact, u from (94) differs from u satisfying (76)–(78). However, this should not lead to any confusion.

Example 1. For $\gamma = -\frac{1}{4}$ we obtain from (94) by (89) straightforwardly,

$$u(t,x) = 1 + \frac{2}{\pi}\sqrt{\frac{l}{x}}\sum_{m=1}^{\infty}\frac{(-1)^m}{m}\sin\biggl(\pi m\sqrt{\frac{x}{l}}\biggr)\exp\biggl[-\frac{\sigma^2\pi^2 m^2}{8l}t\biggr].$$

In order to solve (82) we set $\lambda^- := \lambda - l$ and then apply the Fourier–Bessel series (94) with γ replaced by

$$\gamma^- := \frac{1}{2} - \frac{k\lambda^-}{\sigma^2} = \gamma + \frac{kl}{\sigma^2}. \tag{95}$$

Example 2. We now consider some numerical examples concerning $u^+ = u$ in (94) and u^- given by (94) due to (95). Note that in (94) the function u depends only on σ, l, and γ. That is, u depends on σ, l, and the product $k\lambda$. Let us consider a CIR process with $\sigma = 1$, $\lambda = 1$, $k = 0.75$, and let us take $l = 0.1$. We then compare u^+, which is given by (94) for $\gamma = -0.25$ due to (87) (see Example 1), with u^- given by (94) for $\gamma^- = -0.175$ due to (95). The results are depicted in Figure 2. The sums corresponding to (94) are computed with five terms (more terms did not yield any improvement).

Normalization of $u(t,x)$. For practical applications it is useful to normalize (94) in the following way. Let us treat γ as an essential but fixed parameter and introduce as new parameters

$$\frac{x}{l} = \widetilde{x}, \quad 0 < \widetilde{x} \le 1, \qquad \frac{\sigma^2 t}{8l} = \widetilde{t}, \quad \widetilde{t} \ge 0,$$

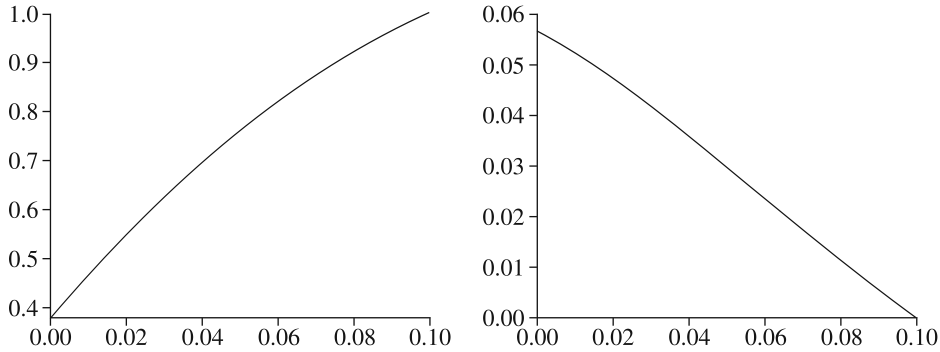

FIGURE 2: The plot of $u^+(0.1, x)$ *(left)*, and the plot of $u^+(0.1, x) - u^-(0.1, x)$ for $0 \le x \le 0.1$ *(right)*.

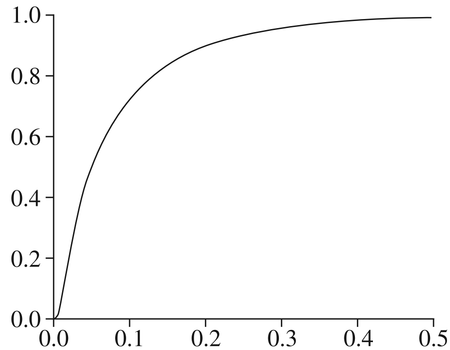

FIGURE 3: Normalized distribution function $\widetilde{u}(\widetilde{t}, 0.5)$ for $\gamma = -\frac{1}{4}$.

and consider the function

$$\widetilde{u}(\widetilde{t}, \widetilde{x}) := 1 - 2\widetilde{x}^{\gamma} \sum_{m=1}^{\infty} \frac{J_{-2\gamma}(\pi_{-2\gamma,m}\sqrt{\widetilde{x}})}{\pi_{-2\gamma,m} J_{-2\gamma+1}(\pi_{-2\gamma,m})} \exp[-\pi^2_{-2\gamma,m}\widetilde{t}], \qquad 0 < \widetilde{x} \le 1,\ \widetilde{t} \ge 0,$$

that is connected to (94) via

$$\widetilde{u}(\widetilde{t}, \widetilde{x}) = \widetilde{u}\left(\frac{\sigma^2 t}{8l}, \frac{x}{l}\right) = u\left(\frac{8l\widetilde{t}}{\sigma^2}, l\widetilde{x}\right).$$

For the simulation of ϑ_x we need to solve the equation

$$u(\vartheta_x, x) = U, \text{ where } U \sim \text{uniform}[0, 1].$$

For this we set $\widetilde{x} = x/l$ and solve the normalized equation $\widetilde{u}(\widetilde{\vartheta}_{\widetilde{x}}, \widetilde{x}) = U$, and then take

$$\vartheta_x = \frac{8l}{\sigma^2}\widetilde{\vartheta}_{\widetilde{x}}.$$

Note that

$$\mathbb{P}(\vartheta_x < t) = \mathbb{P}\left(\widetilde{\vartheta}_{\widetilde{x}} < \frac{\sigma^2 t}{8l}\right) = \widetilde{u}\left(\frac{\sigma^2 t}{8l}, \frac{x}{l}\right).$$

In Figure 3 we have plotted the normalized function $\widetilde{u}(\widetilde{t}, \widetilde{x})$ for $\gamma = -\frac{1}{4}$ and $\widetilde{x} = \frac{1}{2}$.

https://doi.org/10.1239/aap/1449859803 Published online by Cambridge University Press

Acknowledgement

This research was partially supported through EC-Math-MATHEON project SE7.

References

[1] ALFONSI, A. (2005). On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.* **11,** 355–384.

[2] ALFONSI, A. (2010). High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comput.* **79,** 209–237.

[3] ANDERSEN, L. (2008). Simple and efficient simulation of the Heston stochastic volatility model. *J. Comput. Finance* **11,** 1–42.

[4] BATEMAN, H. AND ERDÉLYI, A. (1953). *Higher Transcendental Functions*. McGraw-Hill, New York.

[5] BROADIE, M. AND KAYA, Ö. (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operat. Res.* **54,** 217–231.

[6] COX, J. C., INGERSOLL, J. E., JR. AND ROSS, S. A. (1985). A theory of the term structure of interest rates. *Econometrica* **53,** 385–407.

[7] DEREICH, S., NEUENKIRCH, A. AND SZPRUCH, L. (2012). An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. *Proc. R. Soc. London A* **468,** 1105–1115.

[8] DOSS, H. (1977). Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. H. Poincaré Sect. B (N.S.)* **13,** 99–125.

[9] GLASSERMAN, P. (2004). *Monte Carlo Methods in Financial Engineering*. Springer, New York.

[10] GÖING-JAESCHKE, A. AND YOR, M. (2003). A survey on some generalizations of Bessel processes. *Bernoulli* **9,** 313–349.

[11] HARTMAN, P. (1964). *Ordinary Differential Equations*. John Wiley, New York.

[12] HESTON, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Studies* **6,** 327–343.

[13] HIGHAM, D. J. AND MAO, X. (2005). Convergence of Monte Carlo simulations involving the mean-reverting square root process. *J. Comp. Finance* **8,** 35–61.

[14] HIGHAM, D. J., MAO, X. AND STUART, A. M. (2002). Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.* **40,** 1041–1063.

[15] IKEDA, N. AND WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam.

[16] KARATZAS, I. AND SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd edn. Springer, New York.

[17] MILSTEIN, G. N. AND SCHOENMAKERS, J. G. M. (2015). Uniform approximation of the CIR process via exact simulation at random times. Preprint. *WIAS preprint no. 2113,* ISSN 2198-9855.

[18] MILSTEIN, G. N. AND TRETYAKOV, M. V. (2004). *Stochastic Numerics for Mathematical Physics*. Springer, Berlin.

[19] MILSTEIN, G. N. AND TRETYAKOV, M. V. (2005). Numerical analysis of Monte Carlo evaluation of Greeks by finite differences. *J. Comp. Finance* **8,** 1–34.

[20] REVUZ, D. AND YOR, M. (1991). *Continuous Martingales and Brownian Motion.* Springer, Berlin.

[21] ROGERS, L. C. G. AND WILLIAMS, D. (1987). *Diffusions, Markov Processes, and Martingales,* Vol. 2, *Itô Calculus*. John Wiley, New York.

[22] SUSSMANN, H. J. (1978). On the gap between deterministic and stochastic ordinary differential equations. *Ann. Prob.* **6,** 19–41.