Uniform approximation of the Cox-Ingersoll-Ross process

Grigori N. Milstein* John Schoenmakers†

July 20, 2014

Abstract

The Doss-Sussmann (DS) approach is used for uniform simulation of the Cox-Ingersoll-Ross (CIR) process. The DS formalism allows 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 zero. From a conceptual point of view the proposed method gives a better quality of approximation (from a path-wise point of view) than standard, or even exact simulation of the SDE at some deterministic time grid.

AMS 2010 subject classification. Primary 65C30; secondary 60H35.
Keywords. Cox-Ingersoll-Ross process, Doss-Sussmann formalism, Bessel functions, confluent hypergeometric equation.

1 Introduction

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

$$dV(t) = k(\lambda - V(t))dt + \sigma\sqrt{V}dw(t),\ V(t_0) = V_0, \tag{1}$$

where $k,\ \lambda,\ \sigma$ 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 [14]. It is known ([15], [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 case $2k\lambda \geq \sigma^2$ and nonnegative in the case $2k\lambda < \sigma^2$. Moreover, in the last 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. A lot of approximation

*Ural Federal University, Lenin Str. 51, 620083 Ekaterinburg, Russia; email: Grigori.Milstein@usu.ru
†Weierstrass-Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, 10117 Berlin, Germany; email: schoenma@wias-berlin.de

methods for the CIR processes are proposed. For an extensive list of articles on this subject we refer to [3] and [7]. Besides [3] and [7] we also refer to [1, 2, 12, 13], where a number of discretization schemes for the CIR process can be found. Further we note that in [18] a weakly convergent fully implicit method is implemented for the Heston model. Exact simulation of (1) is considered in [5, 9] (see [3] as well).

In this paper, we consider uniform pathwise approximation of $V(t)$ on an interval $[t_0, t_0+T]$ using the Doss-Sussmann transformation ([8], [21], [20]) which allows for expressing any trajectory of $V(t)$ by the solution of some ordinary differential equation that depends on the realization of $w(t)$. The approximation $\overline{V}(t)$ will be uniform in the sense that the path-wise error will be uniformly bounded, i.e.

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

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 approximately construct a generic trajectory of $V(t)$. Such kind of simulation is more simple than the one proposed in [5] and moreover has the advantage of uniform nature. Let us consider the simulation of a standard Brownian motion W on a fixed time grid

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

Although W may be even 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

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

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). We therefore have introduced the DS formalism that yields a corresponding ODE which solutions are defined on random time intervals. If V gets close to zero however, the ODE becomes intractable for numerical integration and so, for the parts of a trajectory $V(t)$, that are close to zero, 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 $\alpha := \left(4k\lambda - \sigma^2\right)/8 > 0$. We underline that the case $\alpha > 0$ is more general than the case $2k\lambda \ge \sigma^2$ 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 $2k\lambda \ge \sigma^2$. However, for example, in [1] and [12] 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 $\alpha \leq 0$, however in a highly nontrivial way. Therefore, the case $\alpha \leq 0$ will be considered in a subsequent work.

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 5.1. Section 6 is dedicated to the uniform construction of trajectories close to zero 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 Appendix.

2 The Doss-Sussmann transformation

2.1 Due to the Doss-Sussmann approach ([8], [15], [20], [21]), the solution of (1) may be expressed in the form

$$V(t) = F(X(t), w(t)), \tag{3}$$

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

Let us recall the Doss-Sussmann formalism according to [20], V.28. In [20] one considers the Stratonovich SDE

$$dV(t) = b(V)dt + \gamma(V) \circ dw(t). \tag{4}$$

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

$$\frac{\partial F}{\partial y} = \gamma(F),\ \ F(x, 0) = x, \tag{5}$$

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

$$\frac{dX}{dt} = \frac{1}{\partial F/\partial x(X(t), w(t))} b(F(X(t), w(t)),\ \ X(0) = V(0). \tag{6}$$

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

$$dU = (\frac{\alpha}{U} - \frac{k}{2}U)dt + \frac{\sigma}{2}dw,\ \ U(0) = \sqrt{V(0)} > 0, \quad \text{where} \tag{7}$$

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

Let us seek the solution of (7) in the form

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

in accordance with (3)-(6). Because the Ito and Stratonovich forms of equation (7) coincide, we have

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

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

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

i.e.,

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

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

$$\frac{dY}{dt} = \frac{\alpha}{Y + \frac{\sigma}{2} w(t)} - \frac{k}{2}(Y + \frac{\sigma}{2} w(t)), \;\; Y(0) = U(0) = \sqrt{V(0)} > 0. \tag{11}$$

From (9), (10), and solution of (11), we formally obtain the solution $U(t)$ of (7):

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

Hence

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

2.2 Since the Doss-Sussmann 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 give 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:*

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

Then equation (11) has a unique solution $Y(t)$ on the interval $[0, \tau)$, the solution $U(t)$ of (7) is expressed by formula (12) on this interval, and $V(t)$ is expressed by (13).

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

$$dw = dw(t), \;\; dV = k(\lambda - V)dt + \sigma \sqrt{V(t)} dw\,(t)\,,$$

which satisfies the initial conditions $w(0) = 0, \;\; V(0) > 0$. Then $U(t) = \sqrt{V(t)} > 0$ is a solution of (7) on the interval $[0, \tau)$. Consider the function $Y(t) = U(t) - \frac{\sigma}{2} w(t), \;\; 0 \le t < \tau$. Clearly, $Y(t) + \frac{\sigma}{2} w(t) > 0$ on $[0, \tau)$. Due to Itô's formula, we get

$$dY(t) = dU(t) - \frac{\sigma}{2} dw(t) = \frac{\alpha dt}{Y + \frac{\sigma}{2} w(t)} - \frac{k}{2}(Y + \frac{\sigma}{2} w(t))dt,$$

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

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

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

the equation

$$\frac{dY}{dt} = \frac{\alpha}{Y + \frac{\sigma}{2}(w(t) - w(t_0))} - \frac{k}{2}(Y + \frac{\sigma}{2}(w(t) - w(t_0))), \tag{14}$$
$$Y(t_0; t_0) = \sqrt{V(t_0)},\ t_0 \le t < t_0 + \tau,$$

with α given by (8). Clearly,

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

3 Auxiliary propositions

3.1 Let us consider in view of (14) solutions of the ordinary differential equations

$$\frac{dy^0}{dt} = \frac{\alpha}{y^0} - \frac{k}{2}y^0,\ y^0(t_0) = y_0 > 0,\ t \ge t_0 \ge 0, \tag{16}$$

which are given by

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

In the case $\alpha > 0$, i.e., $4k\lambda > \sigma^2$, we have: 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 $t \to \infty$. Further $y^0(t) = \sqrt{2\alpha/k}$ is a solution of (16).

In the case $\alpha = 0$, the solution $y^0_{t_0,y_0}(t) \downarrow 0$ if $t \to \infty$, for any $y_0 > 0$. We note that the case $\alpha \ge 0$ is more general than the case $2k\lambda \ge \sigma^2$ (we recall that in the latter case $V(t) > 0$, $t \ge t_0$).

In the case $\alpha < 0$, the solution $y^0_{t_0,y_0}(t)$ is convexly decreasing for not too large y_0. It attains zero at the moment $\bar{t}$ given by

$$\bar{t} = t_0 + \frac{1}{k}\ln\frac{y_0^2 - 2\alpha/k}{-2\alpha/k} \tag{18}$$

and $y^{0\prime}_{t_0,y_0}(\bar{t}) = -\infty$.

In what follows we deal with the case

$$\alpha = \frac{4k\lambda - \sigma^2}{8} \ge 0. \tag{19}$$

3.2. Our next goal is to obtain estimates for solutions of the equation

$$\frac{dy}{dt} = \frac{\alpha}{y + \frac{\sigma}{2}\varphi(t)} - \frac{k}{2}(y + \frac{\sigma}{2}\varphi(t)),\ y(t_0) = y_0,\ t_0 \le t \le t_0 + \theta, \tag{20}$$

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

Lemma 2 *Let $\alpha \geq 0$. Let $y^i(t),\ i = 1, 2$, be two solutions of (20) such that $y^i(t) + \frac{\sigma}{2}\varphi(t) > 0$ on $[t_0, t_0 + \theta]$, for some θ with $0 \leq \theta \leq T$. Then*

$$\left|y^2(t) - y^1(t)\right| \leq \left|y^2(t_0) - y^1(t_0)\right|,\ t_0 \leq t \leq t_0 + \theta. \tag{21}$$

Proof. We have

$$d(y^2(t) - y^1(t))^2 = 2(y^2(t) - y^1(t)) \tag{22}$$
$$\times \left(\frac{\alpha}{y^2(t) + \frac{\sigma}{2}\varphi(t)} - \frac{k}{2}(y^2(t) + \frac{\sigma}{2}\varphi(t)) - \frac{\alpha}{y^1(t) + \frac{\sigma}{2}\varphi(t)} + \frac{k}{2}(y^1(t) + \frac{\sigma}{2}\varphi(t)) \right) dt.$$

From here

$$(y^2(t) - y^1(t))^2 = (y^2(t_0) - y^1(t_0))^2$$
$$+2\int_{t_0}^{t} [-\alpha \frac{(y^2(s) - y^1(s))^2}{(y^1(s) + \frac{\sigma}{2}\varphi(s))(y^2(s) + \frac{\sigma}{2}\varphi(s))} - \frac{k}{2}(y^2(s) - y^1(s))^2] ds$$
$$\leq (y^2(t_0) - y^1(t_0))^2,$$

whence (21) follows. ∎

Proposition 3 *For any $\alpha > 0$ it holds*

$$\left|\sqrt{V_{t_0,V_0^2}(t)} - \sqrt{V_{t_0,V_0^1}(t)}\right| \leq \left|\sqrt{V_0^2} - \sqrt{V_0^1}\right|,\ t_0 \leq t < \infty. \tag{23}$$

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

$$\sqrt{V_{t_0,V_0^i}(t)} = Y^i(t; t_0) + \frac{\sigma}{2}(w(t) - w(t_0)) > 0, \tag{24}$$
$$Y^i(t_0; t_0) = \sqrt{V_0^i},\ i = 1, 2,\ t_0 \leq t < \infty,$$

where the $Y^i(t; t_0)$ satisfy (14). So, by Lemma 2 with $\varphi(t) = w(t) - w(t_0)$,

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

and (23) 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. ∎

3.3 Now consider (20) for a continuous function φ satisfying

$$|\varphi(t)| \leq r,\ t_0 \leq t \leq t_0 + \theta \leq t_0 + T, \tag{25}$$

for some $r > 0$ and $0 \leq \theta \leq T$. Along with (16), (20) with (25), we further consider the equations

$$\frac{dy}{dt} = \frac{\alpha}{y + \frac{\sigma}{2}r} - \frac{k}{2}(y + \frac{\sigma}{2}r),\ y(t_0) = y_0, \tag{26}$$
$$\frac{dy}{dt} = \frac{\alpha}{y - \frac{\sigma}{2}r} - \frac{k}{2}(y - \frac{\sigma}{2}r),\ y(t_0) = y_0. \tag{27}$$

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

$$y_0 \geq \eta \geq \sigma r > 0. \tag{28}$$

The solutions of (16), (20) with (25), (26), and (27) 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

$$y^-(t) = [(y_0 + \frac{\sigma}{2}r)^2 e^{-k(t-t_0)} + \frac{2\alpha}{k}(1 - e^{-k(t-t_0)})]^{1/2} - \frac{\sigma}{2}r,\ t_0 \leq t \leq t_0 + \theta, \tag{29}$$

$$y^+(t) = [(y_0 - \frac{\sigma}{2}r)^2 e^{-k(t-t_0)} + \frac{2\alpha}{k}(1 - e^{-k(t-t_0)})]^{1/2} + \frac{\sigma}{2}r,\ t_0 \leq t \leq t_0 + \theta. \tag{30}$$

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, e.g., [11], Ch. 3), the inequality

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

which is fulfilled in view of (25) for $y > \sigma r/2$, then implies that

$$y^-(t) \leq y(t) \leq y^+(t),\ t_0 \leq t \leq t_0 + \theta. \tag{31}$$

The same inequality holds for $y(t)$ replaced by $y^0(t)$. We thus get

$$\left|y(t) - y^0(t)\right| \leq y^+(t) - y^-(t),\ t_0 \leq t \leq t_0 + \theta. \tag{32}$$

Proposition 4 *Let* $\alpha = \dfrac{4k\lambda - \sigma^2}{8} \geq 0$, *the inequalities (25) and (28) be fulfilled for a fixed* $\eta > 0$, *and let* $\theta \leq T$. *We then have*

$$\left|y(t) - y^0(t)\right| \leq Cr(t - t_0) \leq Cr\theta,\ t_0 \leq t \leq t_0 + \theta, \quad \textit{with} \tag{33}$$
$$C = \frac{\sigma k}{2} + \frac{4\alpha\sigma}{3\eta^2} e^{\frac{k}{2}T}.$$

In particular, C *is independent of* $t_0,\ y_0$, *and* r *(provided (28) holds).*

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

$$y^+(t) = z^-(t) + \frac{\sigma}{2}r,\ y^-(t) = z^+(t) - \frac{\sigma}{2}r,$$
$$y^+(t) - y^-(t) = \sigma r - (z^+(t) - z^-(t)), \tag{34}$$

where

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

Further,

$$z^+(t) - z^-(t) = \frac{(z^+(t))^2 - (z^-(t))^2}{z^+(t) + z^-(t)} = \frac{2y_0\sigma r e^{-k(t-t_0)}}{z^+(t) + z^-(t)}. \tag{35}$$

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

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

whence

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

Therefore

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

From (35) we have that

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

and so due to (34) we get

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

Since $1-e^{-q\vartheta} \le q\vartheta$ for any $q \ge 0,\ \vartheta \ge 0$, and $y_0^2 - \frac{\sigma^2}{4}r^2 \ge \frac{3}{4}\eta^2$ due to (28), we obtain

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

From this and (32), (33) follows with $C = \frac{\sigma k}{2} + \frac{4\alpha\sigma}{3\eta^2}e^{\frac{k}{2}T}$. ■

Corollary 5 *Under the assumptions of Proposition 4, we get by taking* $\eta = y_0$,

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

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

4 One-step approximation

Let us suppose that for $t_m,\ t_0 \leq t_m < t_0 + T$, $V(t_m)$ 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)),

$$\frac{dY}{dt} = \frac{\alpha}{Y + \frac{\sigma}{2}(w(t) - w(t_m))} - \frac{k}{2}(Y + \frac{\sigma}{2}(w(t) - w(t_m))), \tag{36}$$
$$Y(t_m; t_m) = \sqrt{V(t_m)},\ t_m \leq t \leq t_m + \theta_m.$$

Assume that

$$y_m = \sqrt{V(t_m)} \geq \sigma r. \tag{37}$$

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

$$\sqrt{V(t)} = Y(t; t_m) + \frac{\sigma}{2}(w(t) - w(t_m)),\ t_m \leq t \leq t_m + \theta_m. \tag{38}$$

Though equation (36) is (just) an ODE, it is not easy to solve it numerically in a straightforward way because of the non-smoothness of $w(t)$. We are here going to construct an approximation $y^m(t)$ of $Y(t; t_m)$ via Proposition 4. 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 $w(t) - w(t_m)$, $t \geq t_m$, to the boundary of the interval $[-r, r]$. So, $|w(t) - w(t_m)| \leq r$ for $t_m \leq t \leq t_m + \theta_m$ and, moreover, the random variable $w(t_m + \theta_m) - w(t_m)$, which equals either $-r$ or $+r$ with probability 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 4 and Corollary 5 then yield,

$$|Y(t; t_m) - y^m(t)| \leq \left(D_1 + \frac{D_2}{y_m^2}\right) r\left(t_{m+1} - t_m\right), \quad t_m \leq t \leq t_{m+1} \quad \text{with} \tag{39}$$
$$t_{m+1} := \min(t_m + \theta_m, t_0 + T),$$

where $y^m(t)$ is the solution of the problem

$$\frac{dy^m}{dt} = \frac{\alpha}{y^m} - \frac{k}{2}y^m,\ y^m(t_m) = Y(t_m; t_m) = \sqrt{V(t_m)}$$

that is given by (17) with $(t_m, y_m) = (t_m, \sqrt{V(t_m)})$. We so 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 (39),

$$|\rho^m(t)| \leq \left(D_1 + \frac{D_2}{y_m^2}\right) r\left(t_{m+1} - t_m\right), \quad t_m \leq t \leq t_{m+1}. \tag{40}$$

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

$$\sqrt{\overline{V}(t)} := y^m(t) + \frac{\sigma}{2}(w(t) - w(t_m)), \quad t_m \leq t \leq t_{m+1}. \tag{41}$$

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 (41) for $t = t_{m+1}$ is given by

$$\sqrt{\overline{V}(t_{m+1})} := y^m(t_{m+1}) + \frac{\sigma}{2}(w(t_{m+1}) - w(t_m)) = \tag{42}$$

$$y^m(t_{m+1}) + \frac{\sigma}{2} \cdot \begin{cases} r\xi_m & \text{with } P(\xi_m = \pm 1) = 1/2, \text{ if } t_{m+1} = t_m + \theta_m < t_0 + T, \\ \zeta_m & \text{if } t_{m+1} = t_0 + T, \end{cases}$$

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

$$W_{t_0+T-t_m} \quad \text{conditional on} \quad \max_{0 \leq s \leq t_0+T-t_m} |W_s| \leq r, \tag{43}$$

where W is an independent standard Brownian motion. For details see Appendix B. We so have the following theorem.

Theorem 6 *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*

$$\left| \sqrt{V(t_{m+1})} - \sqrt{\overline{V}(t_{m+1})} \right| \leq \left(D_1 + \frac{D_2}{V(t_m)} \right) r\,(t_{m+1} - t_m)\,. \tag{44}$$

5 The first convergence theorem

In this section we develop a scheme that generates approximations $\sqrt{\overline{V}(t_0)} = \sqrt{V(t_0)}$, $\sqrt{\overline{V}(t_1)}$, ... , $\sqrt{\overline{V}(t_{n+1})}$, where $n = 0, 1, 2, ...$, and $t_1, ..., t_{n+1}$ 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 $m = 0, ..., n$, 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 case $\alpha \geq 0$. 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 (42) and (44))

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

where

$$\left|\rho^0(t_1)\right| \leq \left(D_1 + \frac{D_2}{V(t_0)} \right) r\,(t_1 - t_0) =: C_0 r (t_1 - t_0). \tag{45}$$

Suppose that

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

We then go to the next step and consider the expression

$$\sqrt{V(t)} = Y(t;t_1) + \frac{\sigma}{2}(w(t) - w(t_1)), \tag{46}$$

where $Y(t;t_1)$ is the solution of the problem (see (36))

$$\begin{aligned} \frac{dY}{dt} &= \frac{\alpha}{Y + \frac{\sigma}{2}(w(t) - w(t_1))} - \frac{k}{2}(Y + \frac{\sigma}{2}(w(t) - w(t_1))), \\ Y(t_1;t_1) &= \sqrt{V(t_1)},\ t_1 \leq t \leq t_1 + \theta_1. \end{aligned} \tag{47}$$

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 the equation (47) with initial value $\overline{Y}(t_1;t_1) = \sqrt{\overline{V}(t_1)}$. From the previous step we have that $\left|Y(t_1;t_1) - \overline{Y}(t_1;t_1)\right| = \left|\sqrt{V(t_1)} - \sqrt{\overline{V}(t_1)}\right| = |\rho^0(t_1)| \leq C_0 r(t_1 - t_0)$. Hence, due to Lemma 2,

$$\left|Y(t;t_1) - \overline{Y}(t;t_1)\right| \leq \rho^0(t_1) \leq C_0 r(t_1 - t_0),\ t_1 \leq t \leq t_1 + \theta_1. \tag{48}$$

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 $t_1 \leq t \leq t_2$ let us consider along with equation (47) the equation

$$\frac{dy^1}{dt} = \frac{\alpha}{y^1} - \frac{k}{2}y^1,\ y^1(t_1) = \overline{Y}(t_1;t_1) = \sqrt{\overline{V}(t_1)}.$$

Due to Proposition 4 and Corollary 5 it holds that

$$\left|\overline{Y}(t;t_1) - y^1(t)\right| \leq \left(D_1 + \frac{D_2}{\overline{V}(t_1)}\right) r\,(t_2 - t_1) =: C_1 r\,(t_2 - t_1)\,,\quad t_1 \leq t \leq t_2, \tag{49}$$

and so by (48) we have

$$\left|Y(t;t_1) - y^1(t)\right| \leq r(C_0(t_1 - t_0) + C_1\,(t_2 - t_1)),\quad t_1 \leq t \leq t_2. \tag{50}$$

We also have (see (46))

$$\sqrt{V(t)} = Y(t;t_1) + \frac{\sigma}{2}(w(t) - w(t_1)) = y^1(t) + \frac{\sigma}{2}(w(t) - w(t_1)) + R^1(t), \tag{51}$$

where

$$\left|R^1(t)\right| \leq r(C_0(t_1 - t_0) + C_1\,(t_2 - t_1)),\ t_1 \leq t \leq t_2. \tag{52}$$

We so define the approximation

$$\sqrt{\overline{V}(t)} := y^1(t) + \frac{\sigma}{2}(w(t) - w(t_1)),\quad \text{that satisfies} \tag{53}$$

$$\sqrt{V(t)} = \sqrt{\overline{V}(t)} + R^1(t)\ ,\quad t_1 \leq t \leq t_2, \tag{54}$$

and then set

$$\sqrt{\overline{V}(t_2)} = y^1(t_2) + \frac{\sigma}{2}(w(t_2) - w(t_1)) = \tag{55}$$
$$y^1(t_2) + \frac{\sigma}{2} \cdot \begin{cases} r\xi_1 & \text{with } P(\xi_1 = \pm 1) = 1/2, \text{ if } t_2 = t_1 + \theta_1 < t_0 + T, \\ \zeta_1 & \text{if } t_2 = t_0 + T, \end{cases},$$

cf. (42) and (43). We thus end up with a next approximation $\sqrt{\overline{V}(t_2)}$ such that

$$\left| \sqrt{V(t_2)} - \sqrt{\overline{V}(t_2)} \right| = \left| R^1(t_2) \right| \le r(C_0(t_1 - t_0) + C_1 (t_2 - t_1)). \tag{56}$$

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

$$\sqrt{\overline{V}(t_m)} \ge \sigma r, \quad \text{for } m = 0, ..., n, \quad \text{and} \tag{57}$$

$$\begin{aligned} \left| \sqrt{V(t_n)} - \sqrt{\overline{V}(t_n)} \right| &\le r \sum_{m=0}^{n-1} \left(D_1 + \frac{D_2}{\overline{V}(t_m)} \right) (t_{m+1} - t_m) \qquad (58) \\ &=: r \sum_{m=0}^{n-1} C_m \left(t_{m+1} - t_m \right). \end{aligned}$$

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

$$\begin{aligned} \frac{dY}{dt} &= \frac{\alpha}{Y + \frac{\sigma}{2}(w(t) - w(t_n))} - \frac{k}{2}(Y + \frac{\sigma}{2}(w(t) - w(t_n))), \qquad (59) \\ Y(t_n; t_n) &= \sqrt{V(t_n)}, \ t_n \le t \le t_n + \theta_n, \end{aligned}$$

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 the equation (59) with initial value $\overline{Y}(t_n; t_n) = \sqrt{\overline{V}(t_n)}$. Due to (58) and Lemma 2 again, we have

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

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

$$\frac{dy^n}{dt} = \frac{\alpha}{y^n} - \frac{k}{2} y^n, \ y^n(t_n) = \overline{Y}(t_n; t_n) = \sqrt{\overline{V}(t_n)}. \tag{60}$$

By repeating the procedure (49)-(56) we arrive at

$$\sqrt{\overline{V}(t)} := y^n(t) + \frac{\sigma}{2}(w(t) - w(t_n)), \ t_n \le t \le t_{n+1}, \tag{61}$$

satisfying

$$\left|\sqrt{V(t)}-\sqrt{\overline{V}(t)}\right| = |R^n(t)| \le r\sum_{m=0}^{n}\left(D_1+\frac{D_2}{\overline{V}(t_m)}\right)(t_{m+1}-t_m)\,, \quad t_n\le t\le t_{n+1}, \tag{62}$$

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

$$\left|\sqrt{V(t_{n+1})}-\sqrt{\overline{V}(t_{n+1})}\right| \le r\sum_{m=0}^{n}\left(D_1+\frac{D_2}{\overline{V}(t_m)}\right)(t_{m+1}-t_m)\,. \tag{63}$$

Proposition 7 *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*

$$\left|\sqrt{V(t_0)}-\sqrt{\overline{V}(t_0)}\right| \le \varepsilon, \tag{64}$$

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

$$\left|\sqrt{V(t_{n+1})}-\sqrt{\overline{V}(t_{n+1})}\right| \le \varepsilon + r\sum_{m=0}^{n}\left(D_1+\frac{D_2}{\eta^2}\right)(t_{m+1}-t_m)\,, \tag{65}$$

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

Proof. Inequality (65) follows from (63) with $V(t_{n+1})=V_{t_0,\bar{V}(t_0)}(t_{n+1})$, (64), and (see Proposition 3)

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

■

Remark 8 *In principle it is possible to use the distribution function* $\mathcal{Q}$ *(see (86)) for constructing* $\sqrt{\overline{V}(t)}$ *for* $t_n<t<t_{n+1}$. *However, we rather consider for* $t_n\le t\le t_{n+1}$ *the approximation*

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

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

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

and (b) for $t_{n+1}=t_0+T$, *one may take* $\widetilde{w}(t)\equiv 0$. *As a result we get similar to (87) an insignificant increase of the error,*

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

Let us consolidate the above procedure in a concise way.

5.1 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* $P(\xi_n = \pm 1) = 1/2$, 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 equation (60) on the interval $[t_n, t_{n+1}]$ with solution y^n and set

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

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

So, under the assumption (57) we obtain the estimate (62) (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 case $2k\lambda \geq \sigma^2$ almost all trajectories are positive, hence in this case the proposed method is almost surely convergent.

Theorem 9 *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*

$$\sqrt{\overline{V}(t_m)} \geq \eta \geq r\sigma, \quad m = 0, 1, 2, ...,$$

for any $r < r_0$. So in particular (57) is fulfilled for all $m = 0, 1, ...,$ and the estimate (62) implies that for any $r < r_0$,

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

and (see Remark 8)

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

Proof. Let us define

$$\eta := \frac{1}{2} \min_{t_0 \leq t \leq t_0 + T} \sqrt{V(t)} \quad \text{and}$$
$$r_0 := \min \left(\frac{\eta}{\sigma}, \frac{\eta}{\left(D_1 + \frac{D_2}{\eta^2} \right) T} \right), \tag{66}$$

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

$$\sqrt{\overline{V}(t_m)} \geq \eta \geq r\sigma. \tag{67}$$

For $m = 0$ we trivially have

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

Now suppose by induction that $\sqrt{\overline{V}(t_j)} \geq \eta$ for $j = 0, ..., m$. Then due to (63) 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 (66). 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 (67) and the convergence for $r \downarrow 0$. ■

Remark 10 . *In the case where $4k\lambda \geq \sigma^2 > 2k\lambda$ trajectories will reach zero with positive probability, that is convergence on such trajectories is not guaranteed by Theorem 9. 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 (e.g. see [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 zero.*

6 Simulation of trajectories close to zero and the main algorithm

The simulation algorithm in Section 5.1 has a drawback. Even in the case $2k\lambda \geq \sigma^2$, 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 zero (i.e., using an adaptive algorithm with variable expected time step). However, such an algorithm can be very expensive on parts of trajectories that get close to zero. This is because the smaller r, the smaller the expected passage time $E\theta_n = r^2$ (see Remark 12) and so such parts may require a huge amount 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.

$$\sqrt{\overline{V}(t_k)} \geq \Delta,\ k = 0, 1, ..., \mathfrak{m} - 1,\ \text{and}\ \sqrt{\overline{V}(t_\mathfrak{m})} < \Delta. \tag{68}$$

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

$$\sqrt{V_{t_\mathfrak{m},\overline{V}(t_\mathfrak{m})}(t)} < 2\Delta,\ t_\mathfrak{m} \leq t < t_\mathfrak{m} + \vartheta,\ \sqrt{V_{t_\mathfrak{m},\overline{V}(t_\mathfrak{m})}(t_\mathfrak{m} + \vartheta)} = 2\Delta. \tag{69}$$

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, $\Delta = Ar^a$, $0 < a < 1/2$, 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 inequality (69), 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. We so take, for example, a straight line $L(t)$ that connects the points $(t_{\mathfrak{m}}, \sqrt{\overline{V}(t_{\mathfrak{m}})})$ and $(t_{\mathfrak{m}} + \vartheta, \sqrt{V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t_{\mathfrak{m}} + \vartheta)})$ 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 so have the error estimate

$$\left| \sqrt{\overline{V}(t)} - \sqrt{V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t)} \right| \leq 2\Delta \quad \text{for} \quad t_m < t < t_m + \vartheta, \tag{70}$$

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

$$\sqrt{\overline{V}(t_{\mathfrak{m}} + \vartheta)} = \sqrt{V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t_{\mathfrak{m}} + \vartheta)} = 2\Delta. \tag{71}$$

By Proposition 3 (we assume that $\alpha > 0$) we have on the other hand,

$$\left| \sqrt{V_{t_{\mathfrak{m}},\overline{V}(t_{\mathfrak{m}})}(t)} - \sqrt{V_{t_{\mathfrak{m}},V(t_{\mathfrak{m}})}(t)} \right| \leq \left| \sqrt{\overline{V}(t_{\mathfrak{m}})} - \sqrt{V(t_{\mathfrak{m}})} \right|, \quad t_{\mathfrak{m}} \leq t \leq t_{\mathfrak{m}} + \vartheta. \tag{72}$$

Combining inequalities (70) and (72) then yields,

$$\left| \sqrt{\overline{V}(t)} - \sqrt{V_{t_{\mathfrak{m}},V(t_{\mathfrak{m}})}(t)} \right| \leq 2\Delta + \left| \sqrt{\overline{V}(t_{\mathfrak{m}})} - \sqrt{V(t_{\mathfrak{m}})} \right|, \quad t_{\mathfrak{m}} < t < t_{\mathfrak{m}} + \vartheta, \tag{73}$$

while at time $t_{\mathfrak{m}} + \vartheta$ we have by (71) and (72) that

$$\left| \sqrt{\overline{V}(t_{\mathfrak{m}} + \vartheta)} - \sqrt{V_{t_{\mathfrak{m}},V(t_{\mathfrak{m}})}(t_{\mathfrak{m}} + \vartheta)} \right| \leq \left| \sqrt{\overline{V}(t_{\mathfrak{m}})} - \sqrt{V(t_{\mathfrak{m}})} \right|. \tag{74}$$

In other words, the error of $\sqrt{\overline{V}}$ at the time $t_{\mathfrak{m}} + \vartheta$ of passing the band is not larger than the error at $t_{\mathfrak{m}}$ 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

The arguments above result in the following (pseudo) algorithm. Let r and Δ be numbers such that $\Delta \geq \sigma r$.

- 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 5.1 from $(*)$;

- Set $\mathfrak{m} := n^{\text{new}}$; $t_{\mathfrak{m}} = t_n^{\text{new}}$;
- If $t_{\mathfrak{m}} = t_0 + T$ then finish the simulation;
- If $t_{\mathfrak{m}} < t_0 + T$ simulate $\vartheta_{\overline{V}(t_{\mathfrak{m}})}$ according to Appendix C;
- If $t_{\mathfrak{m}} + \vartheta_{\overline{V}(t_{\mathfrak{m}})} \geq t_0 + T$ set $\sqrt{\overline{V}(t)} = \sqrt{\overline{V}(t_{\mathfrak{m}})}$ on $[t_{\mathfrak{m}}, t_0 + T]$ and finish;
- If $t_{\mathfrak{m}} + \vartheta_{\overline{V}(t_{\mathfrak{m}})} < t_0 + T$ set $\sqrt{\overline{V}(t)} = L(t)$ on $[t_{\mathfrak{m}}, t_{\mathfrak{m}} + \vartheta_{\overline{V}(t_{\mathfrak{m}})}]$; set $t_n^{\text{new}} := t_{\mathfrak{m}} + \vartheta_{\overline{V}(t_{\mathfrak{m}})}$; $\sqrt{\overline{V}(t_n^{\text{new}})} = 2\Delta$; Go to $(**)$;

Let us now consider the convergence properties of the main algorithm. Suppose that for a generic point t_n at $(*)$ we have that $\left|\sqrt{\overline{V}(t_n)} - \sqrt{V(t_n)}\right| \leq \varepsilon_n$ (obviously we may take $\varepsilon_0 = 0$). Then the aggregated error of $\sqrt{\overline{V}(t_{\mathfrak{m}})}$ is estimated by (see Proposition 7)

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

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

$$2\Delta + \left|\sqrt{\overline{V}(t_{\mathfrak{m}})} - \sqrt{V(t_{\mathfrak{m}})}\right| \leq 2\Delta + \varepsilon_n^{\text{new}}, \tag{75}$$

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

Theorem 11 . *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 one. The error on $[t_0, t_0 + T]$ is estimated by*

$$\left|\sqrt{\overline{V}(t)} - \sqrt{V(t)}\right| \leq 2\Delta + r \left(D_1 + \frac{D_2}{\Delta^2} \right) T. \tag{76}$$

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 $\Delta = Ar^{1/3}$, the error (76) 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 (e.g. 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 the present article. Let us assume that we need to evaluate the expectation functional

$$E\,f(V_{t_0,V_0}(t) : t_0 \le t \le t_0 + T), \tag{77}$$

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

Now, for instance, suppose that f in (77) does not depend on the parts of the trajectory that are below a certain level l, $l > 0$. A simple example is

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

with $l = V_0$. We may then choose $\Delta = \sqrt{V_0}/2$ in the main algorithm, thus yielding a uniform convergence rate $O(r)$ in any case of $\alpha > 0$ (cf. Theorem 11) 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, $l > 0$,

$$C_{t_0} := E\left(\int_{t_0}^{t_0+T} V_{t_0,V_0}(s)1_{\{V \ge l\}}ds - K\right)^+$$

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

In the general case, e.g. in the case of general f above, it is advantageous to choose A in the main algorithm according to Theorem 11 for a given choice of r in an optimal way. That is, with $\Delta = Ar^{1/3}$ we have to minimize the global error

$$2Ar^{1/3} + r\left(D_1 + \frac{D_2}{A^2 r^{2/3}}\right)T = \left(2A + \frac{D_2 T}{A^2}\right)r^{1/3} + rD_1 T.$$

Thus, $A = (D_2 T)^{1/3} = \left(4\alpha\sigma e^{\frac{k}{2}T}/3\right)^{1/3}$ (see Corollary 5) is a suitable choice since $r \ll r^{1/3}$ when r is small.

Some illustrative examples

We have implemented the main algorithm for the following CIR parameters,

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

hence $2k\lambda < \sigma^2 < 4k\lambda$, and $\alpha = 1/8$. In the algorithm we choose $r = 0.01$ and $\Delta = 0.16821$ determined in the above way. In Figure 1 we illustrate some typical trajectories of V. The first picture depicts a trajectory of V that does not enter the band $(0, \Delta^2) = (0, 0.0283)$, hence it follows Algorithm 5.1 until T. In the second one the trajectory enters the band once, continues linearly until the level $4\Delta^2 = 0.11318$, and then follows Algorithm 5.1 until T. In the third picture the trajectory enters the band three times.

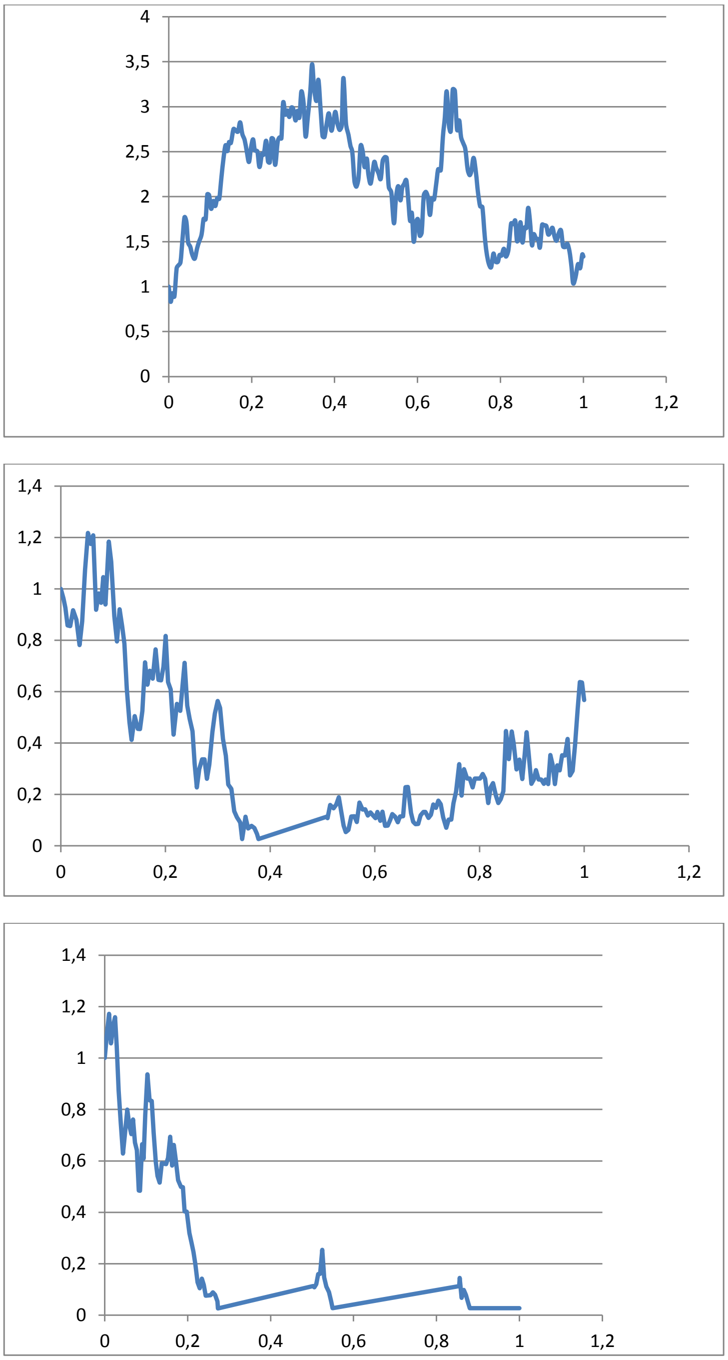

Figure 1: Three sample trajectories that enter the band $(0, \Delta^2) = (0, 0.0283)$ and continue linearly until the level $4\Delta^2 = 0.11318$, 0, 1, and 3 times, respectively.

Appendix

A Addendum to the proof of Proposition 3

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

$$Z(t) = Z(0) + \frac{\delta - 1}{2} \int_0^t \frac{1}{Z(s)} ds + W(t), \quad 0 \leq t < \infty, \tag{78}$$

where W is standard Brownian motion, $Z(0) > 0$, $Z(t) \geq 0$ a.s., and that in particular $E \int_0^t \frac{1}{Z(s)} ds < \infty$, see Appendix A1 in [10], and Ch. XI, Exercise 1.26 in [19]. (For $\delta \leq 1$ the representation of BES$^\delta$ is less simple and involves the concept of local time.) From this fact we will show that for $\alpha > 0$ the solution of (7) may be represented as

$$U(t) = U(t_0) + \int_{t_0}^t (\frac{\alpha}{U(s)} - \frac{k}{2} U(s)) ds + \frac{\sigma}{2} \left(w(t) - w(t_0) \right), \ U(t_0) > 0, \quad t_0 \leq t < \infty. \tag{79}$$

Let us consider

$$U(t) = e^{-k(t-t_0)/2} Z \left(\frac{\sigma^2}{4k} (e^{k(t-t_0)} - 1) \right), \quad t_0 \leq t < \infty, \tag{80}$$

where Z is the solution of (78) with $\delta = 4k\lambda/\sigma^2 > 1$, and $Z(0) = U(t_0) > 0$ (cf. [10] and the references therein). Note that the function $h : t \to \sigma^2(e^{k(t-t_0)} - 1)/(4k)$ satisfies $h(t_0) = 0$, it is smooth and strictly increasing since $k > 0$. Let us further introduce

$$\widetilde{W}(t) - \widetilde{W}(t_0) := \frac{2}{\sigma} e^{-k(t-t_0)/2} W(h(t)) + \frac{k}{\sigma} \int_{t_0}^t e^{-k(s-t_0)/2} W(h(s)) ds. \tag{81}$$

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 $E \left(\widetilde{W}(t) - \widetilde{W}(t_0) \right)^2 = t - t_0$, for all $t \geq t_0$. Indeed, by some algebra and using the definition of h we obtain

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

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

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

we get from (78) and (80),

$$U(s) = e^{-k(s-t_0)/2} W(h(s)) + e^{-k(s-t_0)/2} U(t_0) + e^{-k(s-t_0)/2} \int_{t_0}^{s} \frac{\alpha}{U(r)} e^{k(r-t_0)/2} dr. \tag{82}$$

It thus holds by (81) and (82) that

$$\begin{aligned} U(t) &= e^{-k(t-t_0)/2} U(t_0) + e^{-k(t-t_0)/2} \int_{t_0}^{t} \frac{\alpha}{U(r)} e^{k(r-t_0)/2} dr \\ &\quad + \frac{\sigma}{2} \left(\widetilde{W}(t) - \widetilde{W}(t_0) \right) - \frac{k}{2} \int_{t_0}^{t} e^{-k(s-t_0)/2} W(h(s)) ds \\ &= U(t_0) - \frac{k}{2} \int_{t_0}^{t} U(s) ds + \frac{\sigma}{2} \left(\widetilde{W}(t) - \widetilde{W}(t_0) \right) \\ &\quad + e^{-k(t-t_0)/2} \int_{t_0}^{t} \frac{\alpha}{U(r)} e^{k(r-t_0)/2} dr + \frac{k}{2} \int_{t_0}^{t} e^{-k(s-t_0)/2} ds \int_{t_0}^{s} \frac{\alpha}{U(r)} e^{k(r-t_0)/2} dr. \end{aligned} \tag{83}$$

In particular, the Lebesgue integral $\int_{t_0}^{t} \frac{\alpha}{U(r)} e^{k(r-t_0)/2} dr$ 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 zero, 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 (83) is equal to $\int_{t_0}^{t} \frac{\alpha}{U(r)} dr$. We so arrive at

$$U(t) = U(t_0) + \int_{t_0}^{t} (\frac{\alpha}{U(r)} - \frac{k}{2} U(r)) dr + \frac{\sigma}{2} \left(\widetilde{W}(t) - \widetilde{W}(t_0) \right). \tag{84}$$

From (84) the representation (79) follows for $w = \widetilde{W}$. In particular, the pair $(U, \widetilde{W})$ in (84) may be considered as a strong SDE solution on the probability space where $\widetilde{W}$ 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

$$Z_i(s) := e^{k\left(h^{-1}(s) - t_0\right)/2} U_i(h^{-1}(s)), \quad 0 \leq s < \infty, \quad i = 1, 2, \tag{85}$$

would similarly yield two strong solutions (Z_1, W) and (Z_2, W) of (78) with $Z_1(0) = Z_2(0) = U(t_0)$ with respect to some (though from $\widetilde{W}$ different) Brownian motion W. But, by uniqueness of the solution to (78) it will follow that $Z_1 = Z_2$, and then from (85) that $U_1 = U_2$.

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

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

for $Y(0) = U(0) > 0$, $0 \leq t < \infty$, and that in particular Y is an absolutely continuous function. From this it follows that (22) holds for $t_0 \leq t \leq t_0 + T$ when $\alpha > 0$ and $\varphi(t) = w(t) - w(t_0)$ is a Brownian trajectory, and then inequality (21) in Lemma 2 goes through for $\theta = T$.

B Simulation of θ_m and ζ_m

For simulating θ_m we utilize the distribution function

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

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 the following one:

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

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

and it holds

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

(see for details [17], Ch. 5, Sect. 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 $\mathcal{P}$. That is, we have to solve the equation $\tilde{\mathcal{P}}(\tau) = U$, 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 (42) we observe that (43) is equivalent with

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

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 [17], Ch. 5, Sect. 3)

$$\mathcal{Q}(x;t) := P(W(t) < x \mid \max_{0\le s\le t}|W(s)| < 1), \quad -1 \le x \le 1, \tag{86}$$

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, e.g. zero. 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 zero, for instance, on the last step, we get $\sqrt{\overline{V}(t_0+T)} = y^\nu(t_0+T)$ instead of (42), and

$$\left|\sqrt{V(t_0+T)} - \sqrt{\overline{V}(t_0+T)}\right| \le r\sum_{m=0}^{\nu-1}\left(D_1 + \frac{D_2}{\overline{V}(t_m)}\right)(t_{m+1}-t_m) + \sigma r, \tag{87}$$

Remark 12 *We have in any step $E\theta_n = r^2$, the random number of steps before reaching t_0+T, say ν, is finite with probability one, and $E\nu = O(1/r^2)$. For details see [17], Ch. 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.*

C Simulation of ϑ_x

In order to carry out the simulation method for trajectories near zero 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

$$dX(s) = k(\lambda - X(s))ds + \sigma\sqrt{X}dw(s),\ X(0) = x, \tag{88}$$

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

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

is the solution of the first boundary value problem of parabolic type ([17], Ch. 5, Sect. 3)

$$\frac{\partial u}{\partial t} = \frac{1}{2}\sigma^2 x \frac{\partial^2 u}{\partial x^2} + k(\lambda - x)\frac{\partial u}{\partial x},\ t > 0,\ 0 < x < l, \tag{89}$$

with initial data

$$u(0, x) = 0, \tag{90}$$

and boundary conditions

$$u(t, 0) \text{ is bounded, } u(t, l) = 1. \tag{91}$$

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

$$\frac{\partial v}{\partial t} = \frac{1}{2}\sigma^2 x \frac{\partial^2 v}{\partial x^2} + k(\lambda - x)\frac{\partial v}{\partial x},\ t > 0,\ 0 < x < l, \tag{92}$$

$$v(0, x) = -1;\ v(t, 0) \text{ is bounded, } v(t, l) = 0. \tag{93}$$

The problem (92)-(93) 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 (92)-(93). Along with (88), let us consider the equations

$$dX^+(s) = k\lambda ds + \sigma\sqrt{X^+}dw(s),\ X^+(0) = x, \tag{94}$$

$$dX^-(s) = k(\lambda - l)ds + \sigma\sqrt{X^-}dw(s),\ X^-(0) = x, \tag{95}$$

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

$$X^-(s) \le X(s) \le X^+(s). \tag{96}$$

According to (96), we consider three boundary value problems: first (89)-(91) and next similar ones for the equations

$$\begin{aligned} \frac{\partial u^+}{\partial t} &= \frac{1}{2}\sigma^2 x \frac{\partial^2 u^+}{\partial x^2} + k\lambda\frac{\partial u^+}{\partial x},\ t > 0,\ 0 < x < l, \\ \frac{\partial u^-}{\partial t} &= \frac{1}{2}\sigma^2 x \frac{\partial^2 u^-}{\partial x^2} + k(\lambda - l)\frac{\partial u^-}{\partial x},\ t > 0,\ 0 < x < l. \end{aligned} \tag{97}$$

From (96) 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

$$\frac{\partial v^+}{\partial t} = \frac{1}{2}\sigma^2 x \frac{\partial^2 v^+}{\partial x^2} + k\lambda \frac{\partial v^+}{\partial x},\ t > 0,\ 0 < x < l, \tag{98}$$

$$v^+(0,x) = -1;\ v^+(t,0) \text{ is bounded, } v^+(t,l) = 0, \tag{99}$$

as a good approximation of (92)-(93). Henceforth we write $v := v^+$. By separation of variables we get as elementary independent solutions to (98), $\mathcal{T}(t)\mathcal{X}(x)$, where

$$\mathcal{T}'(t) + \mu\mathcal{T}(t) = 0, \quad \text{i.e.} \quad \mathcal{T}(t) = \mathcal{T}_0 e^{-\mu t}, \quad \mu > 0, \quad \text{and} \tag{100}$$

$$\frac{1}{2}\sigma^2 x\mathcal{X}'' + k\lambda\mathcal{X}' + \mu\mathcal{X} = 0, \quad \mathcal{X}(0+) \text{ is bounded, } \mathcal{X}(l) = 0. \tag{101}$$

It can be verified straightforwardly that the solution of (101) can be obtained in terms of Bessel functions of the first kind (e.g. see [4]),

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

with

$$\gamma := \frac{1}{2} - \frac{k\lambda}{\sigma^2}. \tag{102}$$

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

$$\mathcal{X}(x) = \mathcal{X}_\gamma^-(x) =: \mathcal{X}_\gamma(x) = x^\gamma J_{-2\gamma}\left(\sigma^{-1}\sqrt{8\mu x}\right). \tag{103}$$

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

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,

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sin x \quad \text{and} \quad \pi_{1/2,m} = m\pi, \quad m = 1, 2, ... \tag{104}$$

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

$$\sigma^{-1}\sqrt{8\mu l} = \pi_{-2\gamma,m}, \quad \text{i.e.,} \quad \mu_m := \frac{\sigma^2 \pi^2_{-2\gamma,m}}{8l} \tag{105}$$

and we have

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

By the well-known orthogonality relation

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

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

$$\int_0^l J_{-2\gamma}(\pi_{-2\gamma,m}\sqrt{\frac{x}{l}})J_{-2\gamma}(\pi_{-2\gamma,m'}\sqrt{\frac{x}{l}})dx = l\delta_{m,m'}J^2_{-2\gamma+1}(\pi_{-2\gamma,m}), \quad \text{hence}$$

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

Now set

$$v(t,x) = \sum_{m=1}^{\infty} \beta_m e^{-\mu_m t}\mathcal{X}_{\gamma,m}(x), \qquad 0 \le x \le l. \tag{106}$$

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

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

So for any $p = 1, 2, ...,$

$$-\int_0^l \mathcal{X}_{\gamma,p}(x)x^{-2\gamma}dx = \beta_p l J^2_{-2\gamma+1}(\pi_{-2\gamma_\kappa,p}), \quad \text{i.e.}$$

$$\beta_p = -\frac{\int_0^l \mathcal{X}_{\gamma,p}(x)x^{-2\gamma}dx}{lJ^2_{-2\gamma+1}(\pi_{-2\gamma,p})}. \tag{107}$$

Further it holds that

$$\begin{aligned}\int_0^l \mathcal{X}_{\gamma,p}(x)x^{-2\gamma}dx &= \int_0^l x^{-\gamma}J_{-2\gamma}\left(\pi_{-2\gamma,p}\sqrt{\frac{x}{l}}\right)dx \\ &= 2l^{-\gamma+1}\int_0^1 z^{-2\gamma+1}J_{-2\gamma}\left(\pi_{-2\gamma,p}z\right)dz \\ &= 2l^{-\gamma+1}\frac{J_{-2\gamma+1}\left(\pi_{-2\gamma,p}\right)}{\pi_{-2\gamma,p}}\end{aligned}$$

by well-known identities for Bessel functions (e.g. see [4]), and (107) thus becomes

$$\beta_p = -\frac{2}{l^\gamma \pi_{-2\gamma,p}J_{-2\gamma+1}(\pi_{-2\gamma,p})}, \qquad p = 1, 2, ... \tag{108}$$

So, from $v = u - 1$, (100) (103), (105), (108), and (106) we finally obtain

$$u(t,x) = 1 - 2x^\gamma l^{-\gamma}\sum_{m=1}^{\infty}\frac{J_{-2\gamma}\left(\pi_{-2\gamma,m}\sqrt{\frac{x}{l}}\right)}{\pi_{-2\gamma,m}J_{-2\gamma+1}(\pi_{-2\gamma,m})}\exp\left[-\frac{\sigma^2\pi^2_{-2\gamma,m}}{8l}t\right], \quad 0 \le x \le l. \tag{109}$$

It should be noted that, in fact, u from (109) differs from u satisfying (89)-(91). However, this shouldn't lead to any confusion.

Example 13 *For $\gamma = -1/4$ we get from (109) by (104) straightforwardly,*

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

For solving (97) we set $\lambda^- := \lambda - l$, and then apply the Fourier-Bessel series (109) with γ replaced by

$$\gamma^- := \frac{1}{2} - \frac{k\lambda^-}{\sigma^2} = \gamma + \frac{kl}{\sigma^2}. \tag{110}$$

Example 14 *We now consider some numerical examples concerning $u^+ = u$ in (109) and u^- given by (109) due to (110). Note that actually in (109) the function u only depends 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 (109) for $\gamma = -0.25$ due to (102) (see Example 13), with u^- given by (109) for $\gamma^- = -0.175$ due to (110). The results are depicted in Figure 2. The sums corresponding to (109) are computed with five terms (more terms did not give any improvement).*

Normalization of $u(t, x)$

For practical applications it is useful to normalize (109) in the following way. Let us treat γ as essential but fixed parameter, introduce as new parameters

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

and consider the function

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

that is connected to (109) via

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

For simulation of ϑ_x we need to solve the equation

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

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

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

We have plotted in Figure 3 the normalized function $\widetilde{u}(\widetilde{t}, \widetilde{x})$ for $\gamma = -1/4$ and $\widetilde{x} = 0.5$.

References

[1] A. Alfonsi (2005). On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.,* v. 11, no. 4, 355-384.

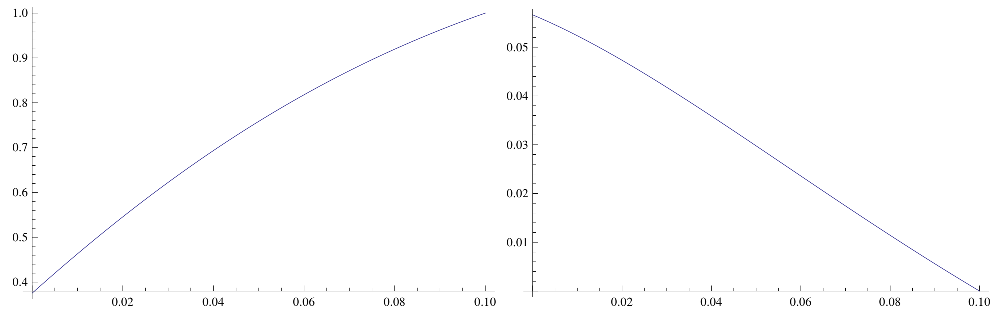

Figure 2: Left panel $u^{+}(0.1, x)$, right panel $u^{+}(0.1, x) - u^{-}(0.1, x)$, for $0 \leq x \leq 0.1$

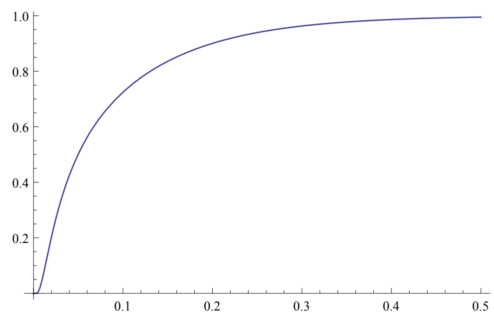

Figure 3: Normalized distribution function $\widetilde{u}(\widetilde{t}, 0.5)$ for $\gamma = -1/4$

[2] A. Alfonsi (2010). High order discretization schemes for the CIR process: Application to affine term structure and Heston models. *Math. Comput.*, v. 79 (269), 209-237.

[3] L. Andersen (2008). Simple and efficient simulation of the Heston stochastic volatility model. *J. of Compute Fin.*, v. 11, 1-42.

[4] H. Bateman, A. Erdélyi (1953). *Higher Transcendental Functions.* MC Graw-Hill Book Company.

[5] M. Broadie, Ö. Kaya (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Oper. Res.*, v. 54, 217-231.

[6] J. Cox, J. Ingersoll, S.A. Ross (1985). A theory of the term structure of interest rates. *Econometrica*, v. 53, no. 2, 385-407.

[7] S. Dereich, A. Neuenkirch, L. Szpruch (2012). An Euler-type method for the strong approximation of the Cox-Ingersoll-Ross process. Proc. R. Soc. A 468, no. 2140, 1105–1115.

[8] H. Doss (1977). Liens entre équations différentielles stochastiques et ordinaries. *Ann. Inst. H. Poincaré Sect B (N.S.),* v. 13, no. 2, 99-125.

[9] P. Glasserman (2003). *Monte Carlo Methods in Financial Engineering.* Springer.

[10] A. Göing-Jaeschke, M. Yor (2003). A survey on some generalizations of Bessel processes. *Bernoulli* v. 9, no. 2, 313-349.

[11] P. Hartman (1964). *Ordinary Differential Equations.* John Willey & Sons.

[12] D.J. Higham, X. Mao (2005). Convergence of Monte Carlo simulations involving the mean-reverting square root process. *J. Comp. Fin.*, v. 8, no. 3, 35-61.

[13] D.J. Higham, X. Mao, A.M. Stuart (2002). Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. anal.*, v. 40, no. 3, 1041-1063.

[14] S.L. Heston (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, v. 6, no. 2, 327-343.

[15] N. Ikeda, S. Watanabe (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland/Kodansha.

[16] I. Karatzas, S.E. Shreve (1991). *Brownian Motion and Stochastic Calculus.* Springer.

[17] G.N. Milstein, M.V. Tretyakov (2004). *Stochastic Numerics for Mathematical Physics.* Springer.

[18] G.N. Milstein, M.V. Tretyakov (2005). Numerical analysis of Monte Carlo evaluation of Greeks by finite differences. *J. Comp. Fin.,* v. 8, no. 3, 1-33.

[19] D. Revuz, M. Yor (1991). *Continuous Maringales and Brownian Motion. Springer.*

[20] L.C.G. Rogers, D. Williams (1987). *Diffusions, Markov Processes, and Martingales, v. 2 : Ito Calculus.* John Willey & Sons.

[21] H.J. Sussmann (1978). On the gap between deterministic and stochastic ordinary differential equations. *Annals of probability*, v. 6, 19-41.