Cox–Ingersoll–Ross and squared Bessel processes driven by Brownian motion: interaction, estimation, applications

Yuliya Mishura

Taras Shevchenko National University of Kyiv

Stochastic models in biomathematics and applications
January 20th - 21th, 2026, University of Salerno, Fisciano (SA), Italy
21 January 2026

Outline

Overview of the results

The talk is based on our common paper
Yuliya Mishura, Kostiantyn Ralchenko and Svitlana Kushnirenko "Driven by Brownian motion Cox–Ingersoll–Ross and squared Bessel processes: Interaction and phase transition". Physics of Fluids, Vol.37, Iss.1 pp. 1 - 15, - 2025
We consider two closely related stochastic processes, namely, Cox–Ingersoll–Ross and Bessel process, both of them being strictly positive solutions of the respective stochastic differential equations. Strictly positive values make them convenient to model real processes in physics, biology, economics. In finances they are used to forecast interest rates and in bond pricing models, see e.g. [Brigo and Mercurio(2001), Di Francesco and Kamm(2022), Maghsoodi(1996), Orlando et al.(2019)]. Similar models are used to simulate changes in the membrane voltage of a neuron [Kelly and Lord(2023)]. In our research we combine the methods of stochastic analysis and methods based on the explicit formulas for probability distributions of CIR and Bessel processes.

In a certain sense, the squared Bessel process can be considered as the result of a phase transition in the Cox–Ingersoll–Ross process. We underline their common and distinct properties. More precisely, we begin by presenting several results that provide upper and lower bounds for the time-asymptotic growth rates of both processes. These bounds exhibit notable similarities between the two models.

Next, we explore the approximation of CIR and squared Bessel processes by a sequence of CIR processes. We prove the convergence of this sequence in integral norms, assuming that the corresponding coefficients converge. Additionally, we establish upper bounds on the rate of convergence. It turns out that the CIR and squared Bessel processes are closely related, as the squared Bessel process can be represented as the limit of a sequence of CIR processes. However, as anticipated, the upper bounds involve coefficients that depend on the length of the time interval and tend to infinity as the interval length increases. In this sense, the processes diverge, or, in other words, they move apart. Nevertheless, the coefficients can be sufficiently close such that, over slowly increasing time intervals, the processes remain comparable.

We then apply this approximation to the problem of parameter estimation for the squared Bessel process using the maximum likelihood method. To establish the strong consistency of the constructed drift parameter estimator, we approximate the squared Bessel process by a sequence of CIR processes, for which the necessary convergence can be derived from their ergodic properties. Furthermore, we show how to estimate the diffusion coefficient of the process based on the realized quadratic variations. Finally, we investigate both processes using the concept of stochastic instability. From this perspective, the properties of the squared Bessel and CIR processes are fundamentally different. We demonstrate that the squared Bessel process exhibits stochastic instability, whereas the CIR process is ergodic and, in this sense, stochastically stable. In addition, we consider an alternative sequence of approximations for the squared Bessel process and show that these approximating processes are also stochastically unstable. Moreover, we prove that, when appropriately normalized, this sequence converges to the (non-squared) Bessel process.

Preliminaries

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and $W = \{W_t, t \geq 0\}$ be a Wiener process on it. We study two stochastic differential equations, namely

$$X_t = x_0 + \int_0^t (a - bX_s)\, ds + \sigma \int_0^t \sqrt{X_s}\, dW_s, \tag{1}$$

and

$$Y_t = y_0 + at + \sigma \int_0^t \sqrt{Y_s}\, dW_s, \tag{2}$$

where $x_0 > 0$, $y_0 > 0$, $a > 0$, $b \geq 0$, and $\sigma > 0$.
It is well known that both equations (1) and (2) admit unique non-negative strong solutions, $X = \{X_t, t \geq 0\}$ and $Y = \{Y_t, t \geq 0\}$, respectively.
The process $X = \{X_t, t \geq 0\}$ was introduced in [Cox et al.(1985)] for the purpose of interest rates modeling. It is commonly referred to as the Cox–Ingersoll–Ross (CIR) process. The process $Y = \{Y_t, t \geq 0\}$ is the squared Bessel process, see, e.g., [Göing-Jaeschke and Yor(2003)] or [Revuz and Yor(1999), Chapter XI] for details.

It follows from the comparison theorem [Karatzas and Shreve(1991), Proposition 2.18, p. 293] that if $x_0 \leq y_0$, then

$$\mathbf{P}\left(X_t \leq Y_t \text{ for all } t \geq 0\right) = 1.$$

In what follows we additionally assume that $2a \geq \sigma^2$. In this case, the trajectories of both processes X and $Z = \sqrt{X}$ with probability 1 remain strictly positive, while in the case $0 < a < \sigma^2/2$, they almost surely hit zero, where the state 0 is instantaneously reflecting (see, e.g., classical paper [Göing-Jaeschke and Yor(2003)] and the more recent ones [Mishura et al.(2024), Mishura and Yurchenko-Tytarenko(2023)] for more details). For the sake of technical simplicity, we assume that

$$2a > \sigma^2.$$

The fraction $d = \frac{4a}{\sigma^2}$ is called the number of degrees of freedom (dimension). Let $Z(t) = \sqrt{X(t)}$.

Case I: $a > \frac{\sigma^2}{4}$

Let $\frac{\sigma^2}{4} < a < \frac{\sigma^2}{2}$, i.e., $1 < d < 2$. This case turns out to be quite similar to $a \geq \frac{\sigma^2}{2}$. Namely, the following theorem holds.

Theorem 1

Let $a > \frac{\sigma^2}{4}$. Then, for any $t \geq 0$,

$$\int_0^t \frac{1}{Z(s)} ds < \infty \quad a.s.$$

and Z a.s. satisfies the SDE of the form

$$Z(t) = \sqrt{x_0} + \frac{1}{2}\left(a - \frac{\sigma^2}{4}\right)\int_0^t \frac{1}{Z(s)} ds - \frac{b}{2}\int_0^t Z(s) ds + \frac{\sigma}{2} W(t). \quad (3)$$

Remark 2

Equation (3) is simply a result of application of Itô formula to $Z(t) = \sqrt{X(t)}$. However, for $a \leq \frac{\sigma^2}{4}$ we should be more careful and apply Itô formula in such a way:

$$\sqrt{X(t)+\varepsilon} = \sqrt{x_0+\varepsilon} + \frac{1}{2}\int_0^t \left(\frac{a}{\sqrt{X(s)+\varepsilon}} - \frac{\sigma^2}{4}\frac{X(s)}{(X(s)+\varepsilon)^{\frac{3}{2}}} \right) ds$$

$$- \frac{1}{2}\int_0^t \frac{bX(s)}{\sqrt{X(s)+\varepsilon}} ds + \frac{\sigma}{2}\int_0^t \frac{\sqrt{X(s)}}{\sqrt{X(s)+\varepsilon}} dW(s).$$

Remark 3

It is possible to prove that for $a > \frac{\sigma^2}{4}$ the process $Z = \sqrt{X}$ is the unique *non-negative* strong solution to the SDE (3). However, if $\frac{\sigma^2}{4} < a < \frac{\sigma^2}{2}$, (3) has other strong solutions; moreover, the uniqueness in law does not hold for (3).

Case II: $a = \frac{\sigma^2}{4}$

The case $a = \frac{\sigma^2}{4}$ turns out to be different from the one described above as the integral $\int_0^t \frac{1}{Z(s)} ds$ is infinite after zero hitting. However, the limit L defined as

$$L(t) := \lim_{n\to\infty} \frac{1}{2} \int_0^t \left(\frac{a}{\sqrt{X(s)+\varepsilon_n}} - \frac{\sigma^2}{4} \frac{X(s)}{(X(s)+\varepsilon_n)^{\frac{3}{2}}} \right) ds$$

has a simple interpretation in terms of Skorokhod reflections as summarized in the following theorem.

Theorem 4

Let $a = \frac{\sigma^2}{4}$ *and denote* $\tau := \inf\{s \geq 0 \mid X(s) = 0\}$*.*

1) *For all* $\gamma > 0$,

$$\int_0^{\tau+\gamma} \frac{1}{Z(s)} ds = \infty \quad a.s.$$

2) *The processes* $Z := \sqrt{X}$ *and* L *defined above is the (unique) solution to Skorokhod problem*

$$Z(t) = \sqrt{x_0} - \frac{b}{2} \int_0^t Z(s) ds + \frac{\sigma}{2} W(t) + L(t), \tag{4}$$

with L *being the corresponding Skorokhod reflection function, i.e. a continuous non-decreasing process starting at 0 with points of growth occurring only at zeros of* Z *and such that* $Z(t) \geq 0$*. In other words,* Z *is a reflected Ornstein-Uhlenbeck process.*

Case II: $a = \frac{\sigma^2}{4}$

The connection between Cox-Ingersoll-Ross and reflected Ornstein-Uhlenbeck processes described on the previous slide can be exploited to obtain an alternative representation of the Skorokhod reflection function. Namely, for an arbitrary sequence $\varepsilon_n \downarrow 0$, denote

$$X_{\varepsilon_n}(t) = x_0 + \int_0^t \left(\frac{\sigma^2}{4} + \varepsilon_n - bX_{\varepsilon_n}(s) \right) ds + \sigma \int_0^t \sqrt{X_{\varepsilon_n}(s)} dW(s).$$

Theorem 5

Let $Z_0(t) = \lim X_{\varepsilon_n}(t)$. *Then*

$$Z_0(t) = \sqrt{x_0} - \frac{b}{2}\int_0^t Z_0(s)ds + \frac{\sigma}{2}W(t) + L_0(t)$$

is a reflected Ornstein-Uhlenbeck process with Skorokhod reflection function L_0*, where, with probability 1,*

$$\sup_{t\in[0,T]}\left|L_0(t) - \frac{1}{2}\int_0^t \frac{\varepsilon_n}{\sqrt{X_{\varepsilon_n}(s)}}ds\right| \to 0, \quad n\to\infty.$$

Case III: $0 < a < \frac{\sigma^2}{4}$

The case $0 < a < \frac{\sigma^2}{4}$ is arguably the most challenging one:

- $\int_0^t \frac{1}{Z(s)} ds$ is infinite after the first moment of zero hitting;
- Z is not a semimartingale;
- L has infinite variation;
- L is decreasing on every interval (τ_1, τ_2) such that $X(t) > 0$ for all $t \in (\tau_1, \tau_2)$;
- in the same time, L is not strictly decreasing on the entire $[0, T]$: otherwise, $Z \leq U$, where U is the standard Ornstein-Uhlenbeck process defined by

$$U(t) = Z(0) - \frac{b}{2} \int_0^t U(s) ds + \frac{\sigma}{2} W(t).$$

However, it is not possible since Y cannot take negative values.

Case III: $0 < a < \frac{\sigma^2}{4}$

In this case the limit

$$L(t) := \lim_{n\to\infty} \frac{1}{2} \int_0^t \left(\frac{a}{\sqrt{X(s)+\varepsilon_n}} - \frac{\sigma^2}{4} \frac{X(s)}{(X(s)+\varepsilon_n)^{\frac{3}{2}}} \right) ds$$

exists but is defined via several time and space transformations of the Wiener process and equals the difference of its local times at two different points.

Distributional and path-wise properties of CIR process

It is well known [Cox et al.(1985)] that X_t follows a non-central chi-squared distribution with the following probability density function:

$$p_t(x) = \frac{1}{c(t)} \left(\frac{x}{x_0 e^{-bt}} \right)^{\nu/2} \exp\left\{ -\frac{x + x_0 e^{-bt}}{c(t)} \right\} I_\nu \left(\frac{2e^{-bt/2}\sqrt{x x_0}}{c(t)} \right) \mathbb{1}_{x>0},$$

where

$$c(t) = \frac{\sigma^2}{2b}\left(1 - e^{-bt}\right), \quad \nu = \frac{2a}{\sigma^2} - 1,$$

and I_ν is the modified Bessel function of the first kind. For $\nu > -1$ and $x \in \mathbb{R}$ this function is defined by the following power series [Oldham et al.(2009), Formula 50:6:1]:

$$I_\nu(x) = \sum_{j=0}^{\infty} \frac{(x/2)^{2j+\nu}}{j!\Gamma(j+1+\nu)},$$

where Γ stands for the Gamma function.

Obviously, $I_\nu(0) = 0$ for all $\nu > 0$, more precisely I_ν has the following behavior as $x \to 0$:

$$I_\nu(x) \sim \frac{(x/2)^\nu}{\Gamma(\nu+1)}.$$

Using this relation, one can show that, as $t \to \infty$,

$$p_t(x) \to \frac{(2b/\sigma^2)^{2a/\sigma^2}}{\Gamma(2a/\sigma^2)} x^{2a/\sigma^2 - 1} e^{-2bx/\sigma^2} \mathbb{1}_{x>0} =: p_\infty(x) \qquad (5)$$

Note that the limiting distribution is a Gamma distribution.
Moreover, the CIR process X is ergodic [Cox et al.(1985)] (see also [Alfonsi(2015), Section 1.2] and [Dehtiar et al.(2022)]). Ergodicity implies that for any function $f \in L^1(\mathbb{R}, p_\infty(x)dx)$, the time average $\frac{1}{T}\int_0^T f(X_t)dt$ converges a.s. to the space average $\int_\mathbb{R} f(x)p_\infty(x)dx$, as $T \to \infty$. In particular, for $a > \frac{\sigma^2}{2}$,

$$\frac{1}{T}\int_0^T \frac{dt}{X_t} \to \int_\mathbb{R} \frac{p_\infty(x)}{x}\,dx = \frac{b}{a - \sigma^2/2}, \quad \text{a. s., when } T \to \infty. \qquad (6)$$

The first two moments of X_t are equal to

$$\mathbf{E}X_t = x_0 e^{-bt} + \frac{a}{b}\left(1 - e^{-bt}\right), \tag{7}$$

and

$$\mathbf{E}X_t^2 = \frac{x_0(\sigma^2 + 2a)}{b}\left(e^{-bt} - e^{-2bt}\right) + \frac{a(\sigma^2 + 2a)}{2b^2}\left(1 - e^{-bt}\right)^2 + x_0^2 e^{-2bt}. \tag{8}$$

The formula for higher moments of the CIR processes is presented in [Okhrin et al.(2023), Proposition 1]. In particular,

$$\mathbf{E}X_t^3 = x_0^3 e^{-3bt} + \left(1 + \frac{3\sigma^2}{2a} + \frac{\sigma^4}{2a^2}\right) \times \left(\frac{a^3}{b^3}\left(1 - e^{-bt}\right)^3 + \frac{3x_0 a^2}{b^2}\left(e^{-bt} - 2e^{-2bt} + e^{-3bt}\right)\right) + \frac{3x_0^2 a}{b}\left(1 + \frac{\sigma^2}{a}\right)\left(e^{-2bt} - e^{-3bt}\right).$$

Distributional and path-wise properties of squared Bessel process

The probability density function of the squared Bessel process Y_t is given by

$$g_t(x) = \frac{1}{2}\left(\frac{x}{y_0}\right)^{\nu/2} \exp\left\{-\frac{2(x+y_0)}{\sigma^2 t}\right\} I_\nu\left(\frac{4\sqrt{xy_0}}{\sigma^2 t}\right) \mathbb{1}_{x>0},$$

where, as before, $\nu = \frac{2a}{\sigma^2} - 1$ (see, e.g., [Revuz and Yor(1999), Chapter XI, Corollary (1.4)]).
Unlike the CIR process X, the squared Bessel process Y is non-ergodic.
For a detailed discussion on the properties of squared Bessel processes, we refer to [Revuz and Yor(1999), Chapter XI]. A comparison of the properties of both ergodic and non-ergodic processes, X and Y, can be found in [Ben Alaya and Kebaier(2013)].
Since $I_\nu(0) = 0$, we see that

$$g_t(x) \to 0, \quad t \to \infty. \tag{9}$$

Therefore, for the squared Bessel process, the limiting distribution does not exist.

The first and second moments of Y_t are equal to:

$$\mathbf{E} Y_t = y_0 + at, \quad \mathbf{E} Y_t^2 = y_0^2 + \left(\frac{\sigma^2}{2} + a\right)\left(2y_0 t + at^2\right). \qquad (10)$$

Both equalities can be derived directly from the equation (2), taking into account that the stochastic integral $\int_0^t \sqrt{Y_s}\, dW_s$ is a square-integrable martingale with zero mean whose second moment equals $\int_0^t \mathbf{E} Y_s\, ds$.

Remark 6

We see from (7) and (8) that the first two moments of the CIR process exist for all t. Moreover, they are totally bounded. Indeed, as it was established in [Ben Alaya and Kebaier(2013), Proposition 3], $\sup_{t\geq 0} \mathbf{E} X_t^p < \infty$ for any $p > -2a/\sigma^2$. In contrast, the first and second moments of the squared Bessel process exhibit linear and quadratic growth with respect to t, respectively.

The general formula for the moments of the Bessel process has the following form: for any $p \geq -\frac{2a}{\sigma^2}$

$$\mathbf{E}\, Y_t^p = \left(\frac{\sigma^2 t}{2}\right)^p \frac{\Gamma\left(\frac{2a}{\sigma^2}+p\right)}{\Gamma\left(\frac{2a}{\sigma^2}\right)} \exp\left\{-\frac{2y_0}{\sigma^2 t}\right\} {}_1F_1\left(\frac{2a}{\sigma^2}+p, \frac{2a}{\sigma^2}, \frac{2y_0}{\sigma^2 t}\right),$$

see the proof of Proposition 3 in [Ben Alaya and Kebaier(2013)]. Here ${}_1F_1$ is the confluent hypergeometric function. We can derive for $p = 3$

$$\begin{aligned}\mathbf{E}\, Y_t^3 &= \left(\frac{a\sigma^4}{2} + \frac{3a^2\sigma^2}{2} + a^3\right) t^3 \\ &\quad + 3\left(\frac{y_0\sigma^4}{2} + \frac{3ay_0\sigma^2}{2} + a^2 y_0\right) t^2 + 3y_0^2\left(\sigma^2 + a\right) t + y_0^3. \qquad (11)\end{aligned}$$

Time-asymptotic growth rate for CIR and quadratic Bessel processes

Now we establish several results that provide a growth rate for the solution to equations (1) and (2), as the function of time and coefficients. As usual, time is included in the constants, since the time interval is fixed in many situations, but for us it is the asymptotic behaviour of functionals of solutions that is most important. We demonstrate what growth rates can be obtained by different methods, and compare the results. The first result follows from the Grönwall inequality therefore gives an exponential growth rate. This growth rate is determined by coefficients a, x_0 and σ and is valid both for solution to (1) or (2).

Proposition 7

Let $Z = \{Z_t, t \geq 0\}$ be a unique solution to (1) or (2) (i.e., $Z = X$ or $Z = Y$). Then, for all $t \geq 0$,

$$\mathbf{E}\left(\sup_{s \leq t} Z_s\right)^2 \leq 2\left((x_0 + at)^2 + 2\sigma^2 t\right) e^{4\sigma^2 t}. \tag{12}$$

Remark 8

Similarly to the above statement, one can establish that for the solution $X = \{X_t, t \geq 0\}$ of the equation (1), the following inequality holds:

$$\mathbf{E} \sup_{s \leq t} \left(X_s + b \int_0^s X_u \, du\right)^2 \leq 2\left((x_0 + at)^2 + 2\sigma^2 t\right) e^{4\sigma^2 t}.$$

The disadvantage of these upper bounds is that being exponential, they grow quickly in time t, but their advantage is that they do not depend on coefficient b.

Now our goal is to obtain the upper bounds for $\mathbf{E}\sup_{s\le t}\left(X_s + b\int_0^s X_u\,du\right)^2$ that will grow not so quickly. We will not apply the Grönwall inequality that always gives exponential growth, but apply directly the values of the moments of CIR-process. Oppositely to previous bounds, the next ones will depend on b and the next goal will be to analyze this dependence when $b \downarrow 0$.

Proposition 9

Let $X = \{X_t, t \ge 0\}$ be a solution to (1). Then, for all $t \ge 0$,

$$\mathbf{E}\sup_{s\le t}\left(X_s + b\int_0^s X_u\,du\right)^2 \le 2(x_0 + at)^2 + \frac{8\sigma^2}{b}\left(x_0 - \frac{a}{b}\right)\left(1 - e^{-bt}\right) + \frac{8\sigma^2 a}{b}t. \quad (13)$$

Remark 10

It follows from (8) that

$$\mathbf{E}\sup_{s\le t}\left(X_s + b\int_0^s X_u\,du\right)^2 \ge \mathbf{E}X_t^2 + b^2\mathbf{E}\left(\int_0^t X_u\,du\right)^2$$
$$\ge \mathbf{E}X_t^2 + b^2\left(\int_0^t \mathbf{E}X_u\,du\right)^2$$
$$= \frac{x_0(\sigma^2+2a)}{b}\left(e^{-bt}-e^{-2bt}\right) + \frac{a(\sigma^2+2a)}{2b^2}\left(1-e^{-bt}\right)^2$$
$$+ x_0^2 e^{-2bt} + \left(\left(x_0-\frac{a}{b}\right)\left(1-e^{-bt}\right)+at\right)^2$$

Comparing this lower bound with the upper bound in (13), we observe that both the upper and lower bounds for $\mathbf{E}\sup_{s\le t}(X_s + b\int_0^s X_u, du)^2$ exhibit a quadratic rate of growth, as $t\to\infty$. The upper bound follows the asymptotic behavior $2a^2t^2 + 4a(x_0+\frac{2\sigma^2}{b})t + O(1)$, while the lower bound behaves as $a^2t^2 + 2a(x_0-\frac{a}{b})t + O(1)$.

Remark 11

(i) Arguing as in the proof of Proposition 9 and using (10) instead of (7), we get the bound

$$\mathbf{E}\sup_{s\le t} Y_s^2 \le 2(y_0+at)^2 + 4\sigma^2\left(2y_0t+at^2\right). \tag{14}$$

Note that the right-hand side of (14) is a limit, as $b\downarrow 0$, of the right-hand side of (13). This can be easily seen from the relation $1-e^{-bt} = bt-\frac{b^2t^2}{2}+o(b^3)$, $b\downarrow 0$. However, if to fix b and consider the asymptotic behaviour, as $t\to\infty$, of the right-hand sides of (13) and (14), we see that the main part of the right-hand side of (13) equals $2a^2t^2$ while the main part of the right-hand side of (14) equals $(2a^2+4a\sigma^2)t^2$. It means (a bit unexpectedly) that the difference between asymptotic behaviour of $\mathbf{E}\sup_{s\le t} Y_s^2$ and $\mathbf{E}\sup_{s\le t} X_s^2$ is very significantly determined by the diffusion coefficient σ as well as (more expected) of the drift coefficient a. Of course, this difference in some latent way depends on b because the value $\frac{8\sigma^2}{b}\left(x_0-\frac{a}{b}\right)\left(1-e^{-bt}\right)$ in the right-hand side of (13) is bounded in t for any $b>0$, however, it is growing as t^2 if we come to the limit, as $b\to 0$.

(ii) The second formula in (10) implies the following lower bound:

$$\mathbf{E}\sup_{s\le t} Y_s^2 \ge \mathbf{E}Y_t^2 = (y_0+at)^2 + \frac{\sigma^2}{2}\left(2y_0t+at^2\right). \tag{15}$$

It can be observed that, compared to the upper bound in (14), this lower bound contains the same terms, but with smaller coefficients. Therefore, we conclude that $\mathbf{E}\sup_{s\le t} Y_s^2$ grows quadratically as a function of t.

(iii) Let us compare two upper bounds for the squared Bessel process, specifically (12) and (14). On the one hand, the bound given by (14) is clearly more advantageous for large values of t. On the other hand, when t is near zero, the bound (12) may provide greater accuracy. More precisely, the bound (12) is superior to (14) if and only if $e^{4\sigma^2 t} \le 2x_0 + at$. If $x_0 > 1/2$, this condition is satisfied for sufficiently small values of t.

Distance between CIR and squared Bessel processes in integral norms

In this section, we establish the rate of convergence of a sequence of CIR processes to a limiting CIR process in two integral norms, $L_1([0, T], \mathbf{P})$ and $L_2([0, T], \mathbf{P})$, over any fixed interval $[0, T]$, under the assumption that the corresponding coefficients converge. Additionally, we analyze the rate of convergence of the CIR processes to a squared Bessel process in both norms, demonstrating a close relationship between these two classes of processes in this context.

However, as expected, the upper bounds of the distance between them contain coefficients depending on the length of the interval and tending to ∞, as the length tends to ∞. In this sense, the processes disperse, or, in other words, move away. Despite this fact, the coefficients can be so close that, under slow growth of time interval, the processes can be still close.

So, consider the following sequence of stochastic differential equations:

$$X_n(t) = x_0 + \int_0^t (a_n - b_n X_n(s))\, ds + \sigma_n \int_0^t \sqrt{X_n(s)}\, dW_s, \tag{16}$$

$n \geq 0$, where $x_0 > 0$, $a_n > 0$, $b_n \geq 0$, and $\sigma_n > 0$ for all $n \geq 0$.
Assume that

$$a_n \to a_0, \quad b_n \to b_0, \quad \sigma_n \to \sigma_0, \quad \text{as } n \to \infty.$$

Note that the equations in (16) satisfy conditions (Y1$_n$)–(Y4$_n$) from Section 4 of [Mishura et al.(2009)], and we have that for any $T > 0$

$$\sup_{t\in[0,T]} \mathbf{E}\,|X_n(t) - X_0(t)| \to 0, \quad \text{as } n \to \infty.$$

In the following theorem, we establish an upper bound for rate of convergence.

Theorem 12

(i) Let $b_0 > 0$*. Then for any* $T > 0$*, the following upper bound holds:*

$$\sup_{t\in[0,T]} \mathbf{E}\,|X_n(t) - X_0(t)| \\ \leq e^{b_n T}\Big(|a_n - a_0|\,T + |b_n - b_0|\,A_0^2(T) + |\sigma_n - \sigma_0|\,A_0(T)\Big), \quad (17)$$

where

$$A_0^2(T) := \frac{1}{b_0}\left(x_0 - \frac{a_0}{b_0}\right)\left(1 - e^{-b_0 T}\right) + \frac{a_0}{b_0}T. \quad (18)$$

(ii) Let $b_0 = 0$*. Then for any* $T > 0$*, the following upper bound holds:*

$$\sup_{t\in[0,T]} \mathbf{E}\,|X_n(t) - X_0(t)| \\ \leq e^{b_n T}\Big(|a_n - a_0|\,T + b_n B_0^2(T) + |\sigma_n - \sigma_0|\,B_0(T)\Big), \quad (19)$$

where $B_0^2(T) = x_0 T + \frac{a_0}{2}\,T^2$.

Remark 13

The function $A_0^2(T)$, defined by (18), is positive, since $A_0^2(T) = \int_0^T \mathbf{E} X_0(s)\,ds$. Moreover, it exhibits linear growth and satisfies the following bounds:

$$\min\left\{x_0, \frac{a_0}{b_0}\right\} T \le A_0^2(T) \le \max\left\{x_0, \frac{a_0}{b_0}\right\} T.$$

To verify these inequalities, we consider two cases.
Case 1: $x_0 - \frac{a_0}{b_0} \ge 0$. Using the inequality $0 \le 1 - e^{-b_0 T} \le b_0 T$, we obtain that

$$\frac{a_0}{b_0} T \le A_0^2(T) \le \frac{1}{b_0}\left(x_0 - \frac{a_0}{b_0}\right) b_0 T + \frac{a_0}{b_0} T = x_0 T. \tag{20}$$

Case 2: $x_0 - \frac{a_0}{b_0} < 0$. In this case, both inequalities in (20) are reversed.

Remark 14

Note that, as $T \to \infty$,

$$A_0^2(T) \sim \frac{a_0}{b_0} T, \quad B_0^2(T) \sim \frac{a_0}{2} T^2.$$

Hence, as $T \to \infty$, the right-hand sides of (17) and (19) are asymptotically equivalent to

$$\frac{1}{b_0}\left(b_0 \left|a_n - a_0\right| + a_0 \left|b_n - b_0\right|\right) T e^{b_n T} \quad \text{and} \quad \frac{1}{2} a_0 b_n T^2 e^{b_n T}$$

respectively.

If, in the above theorem, we take $a_n \equiv a$, $\sigma_n \equiv \sigma$, $b_0 = 0$, the resulting bound simplifies significantly. Specifically, we obtain the following result concerning the approximation of the squared Bessel process by a sequence of CIR processes.

Corollary 15

Let Y be a solution to (2). Consider a sequence of the stochastic differential equations:

$$Y_n(t) = y_0 + \int_0^t (a - b_n Y_n(s))\, ds + \sigma \int_0^t \sqrt{Y_n(s)}\, dW_s, \quad n \geq 1,$$

where $b_n \downarrow 0$, $n \to \infty$. Then, for any $T > 0$, the following bound holds:

$$\sup_{t\in[0,T]} \mathbf{E}\, |Y_n(t) - Y(t)| \leq e^{b_n T} b_n T \left(x_0 + \tfrac{1}{2} a T \right).$$

Corollary 16

Let $b_n \downarrow 0$, $n \to \infty$, and let the sequence $\{T_n, n \geq 1\}$ satisfy

$$T_n \to \infty, \quad e^{b_n T_n} b_n T_n^2 \to 0, \quad \textit{as } n \to \infty. \tag{21}$$

Then

$$\sup_{t \in [0, T_n]} \mathbf{E} \left| Y_n(t) - Y(t) \right| \to 0, \quad n \to \infty.$$

Remark 17

For example, the condition (21) is fulfilled for $b_n = 1/n$, $T_n = \log(\log n)$, $n \geq 2$.

Theorem 18

(*i*) *Let $b_n > 0$ for all $n \geq 0$. Then for any $T > 0$, the following upper bound holds:*

$$\sup_{t\in[0,T]} \mathbf{E}\left(X_n(t) - X_0(t)\right)^2 \leq 2\left(R_n(T) + R_0(T)\right)^{\frac{1}{2}} e^{\frac{b_n T}{2}}$$
$$\times \Big(|a_n - a_0|\, T + |b_n - b_0|\, A_0^2(T) + |\sigma_n - \sigma_0|\, A_0(T)\Big)^{\frac{1}{2}}, \quad (22)$$

where $A_0(T)$ is defined in Theorem 12 and

$$R_n(T) = x_0^3 + \left(1 + \frac{3\sigma_n^2}{2a_n} + \frac{\sigma_n^4}{2a_n^2}\right)$$
$$\times \left(\frac{a_n^3}{b_n^3}\left(1 - e^{-b_n T}\right)^3 + \frac{3x_0 a_n^2}{b_n^2}\left(1 - e^{-b_n T}\right)^2\right)$$
$$+ \frac{3x_0^2 a_n}{b_n}\left(1 + \frac{\sigma_n^2}{a_n}\right)\left(1 - e^{-b_n T}\right).$$

(ii) Let $b_0 = 0$ and $b_n > 0$ for $n \geq 1$. Then for any $T > 0$, the following upper bound holds:

$$\sup_{t\in[0,T]} \mathbf{E}\left(X_n(t) - X_0(t)\right)^2 \leq 2\left(R_n(T) + \widetilde{R}_0(T)\right)^{\frac{1}{2}} e^{\frac{b_n T}{2}}$$

$$\times \left(|a_n - a_0|\, T + b_n B_0^2(T) + |\sigma_n - \sigma_0|\, B_0(T)\right)^{\frac{1}{2}},$$

where $B_0(T)$ is defined in Theorem 12 and

$$\widetilde{R}_0(T) = \left(\frac{a_0\sigma_0^4}{2} + \frac{3a_0^2\sigma_0^2}{2} + a_0^3\right) T^3$$

$$+ 3\left(\frac{x_0\sigma_0^4}{2} + \frac{3a_0x_0\sigma_0^2}{2} + a_0^2x_0\right) T^2 + 3x_0^2\left(\sigma_0^2 + a_0\right) T + x_0^3.$$

Theorem 19

(*i*) *Let* $b_n > 0$ *for all* $n \geq 0$. *Then for any* $T > 0$, *the following upper bound holds:*

$$\begin{aligned}
&\sup_{t\in[0,T]} \mathbf{E}(X_n(t) - X_0(t))^2 \\
&\leq e^{(b_n+b_0)T} \left(2\left|a_n - a_0\right| + \sigma_0^2 + 2\sigma_0\left|\sigma_n - \sigma_0\right|\right) \\
&\quad \times \left(\left|a_n - a_0\right| T + \left|b_n - b_0\right| A_0^2(T) + \left|\sigma_n - \sigma_0\right| A_0(T)\right) T \\
&\quad + e^{b_0 T} 2 e^{b_0 T} \left|b_n - b_0\right| T \left(D_n^2(T) + D_n(T) D_0(T)\right) \\
&\quad + e^{b_0 T} \left|\sigma_n - \sigma_0\right|^2 A_n^2(T),
\end{aligned} \tag{23}$$

where

$$A_n^2(T) = \int_0^T \mathbf{E}X_n(s)\,ds = \tfrac{1}{b_n}\left(x_n - \tfrac{a_n}{b_n}\right)\left(1 - e^{-b_n T}\right) + \tfrac{a_n}{b_n} T,$$

$$D_n^2(T) = \tfrac{x_0(\sigma_n^2 + 2a_n)}{b_n}\left(1 - e^{-bT}\right) + \tfrac{a(\sigma_n^2+2a_n)}{2b_n^2}\left(1 - e^{-bT}\right)^2 + x_0^2.$$

(ii) Let $b_0 = 0$ and $b_n > 0$ for $n \geq 1$. Then for any $T > 0$, the following upper bound holds:

$$\sup_{t\in[0,T]} \mathbf{E}(X_n(t) - X_0(t))^2$$
$$\leq \left(2\left|a_n - a_0\right| + \sigma_0^2 + 2\sigma_0\left|\sigma_n - \sigma_0\right|\right)$$
$$\times \left(\left|a_n - a_0\right| T + b_n B_0^2(T) + \left|\sigma_n - \sigma_0\right| B_0(T)\right) Te^{b_n T}$$
$$+ 2b_n T\left(D_n^2(T) + D_n(T)E_0(T)\right) + \left|\sigma_n - \sigma_0\right|^2 A_n^2(T).$$

where $B_0(T)$ is defined in Theorem 12 and

$$E_0^2(T) = x_0^2 + \left(\frac{\sigma_0^2}{2} + a_0\right)\left(2x_0 T + a_0 T^2\right).$$

Remark 20

Let us now discuss and compare the upper bounds established in Theorems 12, 18, and 19.

1. A key advantage of all three theorems is that they provide explicit rates of convergence in terms of the coefficients of the corresponding equations. This makes them particularly valuable for practical analysis.

2. Theorems 18 and 19 present upper bounds for the second moments, which are often crucial for practical applications. These bounds cannot be directly obtained from the results for the first moments, such as those provided by Theorem 12. Furthermore, it is worth noting that, in a similar manner, one can derive upper bounds for $\sup_{t\in[0,T]} \mathbf{E}(X_n(t) - X_0(t))^{2p}$ for any $p \geq 1$.

3. For a fixed $T > 0$, the convergence rates established in Theorems 12, 18, and 19 can be compared as follows. Assume that

$$|a_n - a_0| \leq \delta_n, \quad |b_n - b_0| \leq \delta_n, \quad |\sigma_n - \sigma_0| \leq \delta_n$$

for some sequence $\delta_n \downarrow 0$ as $n \to \infty$. Then for the quantity $\sup_{t\in[0,T]} \mathbf{E}\, |X_n(t) - X_0(t)|$ Theorems 12, 18, and 19 yield rates of convergence of orders $O(\delta_n)$, $O(\delta_n^{1/4})$, and $O(\delta_n^{1/2})$, respectively. Hence, from the perspective of convergence rates, Theorem 12 offers the fastest rate. Similarly, Theorem 19 demonstrates a superior rate of convergence compared to Theorem 18.

4. We can also compare Theorems 18 and 19 in the asymptotic case as $T \to \infty$. Note that the functions $R_n(T)$ are bounded, while $A_0^2(T)$ grows linearly with T. Consequently, the right-hand side of (22) behaves as $O(T^{1/2}e^{b_n T/2})$, as $T \to \infty$. In contrast, the right-hand side of (23) grows significantly faster, at a rate of $O(T^2 e^{(b_n+b_0)T})$. From this comparison, for large T, Theorem 18 provides a tighter bound than Theorem 19.

We now address the problem of identifying the squared Bessel process, i.e., the estimation of its parameters. Suppose we have continuous-time observations of a trajectory $\{Y_t, t \in [0, T]\}$ of the squared Bessel process (2) over some interval $[0, T]$. Note that these parameters are also defining for the CIR process, and the corresponding estimates can be easily modified for it. Moreover, we can assume that the parameter σ is known and focus on estimating the parameter a. For continuous-time observations, this assumption is natural, because σ can be determined almost surely from the observations on any fixed interval, as explained in the following remark.

Remark 21

(Estimation of σ) Let $T > 0$ be fixed, $\delta = \frac{T}{n}$, and $t_k = k\delta$, for $0 \leq k \leq n$. Then

$$\sum_{k=1}^{n} \left(Y_{t_k} - Y_{t_{k-1}}\right)^2 \to \sigma^2 \int_0^T Y_s \, ds \quad \text{a.s., as } n \to \infty.$$

Indeed, as it is clear, quadratic variation of the linear function tends to zero with the diameter of the partition of the interval. This implies that the parameter σ can be evaluated using the following identity:

$$\sigma^2 = \lim_{n\to\infty} \frac{\sum_{k=1}^{n} \left(Y_{t_k} - Y_{t_{k-1}}\right)^2}{\int_0^T Y_s \, ds} \quad \text{a.s.}$$

To estimate the parameter a, we apply the maximum likelihood method, see, e.g., [Mishura and Shevchenko(2017), Section 10.7]. First, we transform equation (2) using the Itô formula as follows:

$$\begin{aligned}\sqrt{Y_t} &= \sqrt{y_0} + \frac{1}{2}\int_0^t \frac{1}{\sqrt{Y_s}}\, dY_s - \frac{1}{8}\sigma^2 \int_0^t \frac{1}{\sqrt{Y_s}}\, ds \\ &= \sqrt{y_0} + \frac{a}{2}\int_0^t \frac{ds}{\sqrt{Y_s}} + \frac{\sigma}{2} W_t - \frac{1}{8}\sigma^2 \int_0^t \frac{1}{\sqrt{Y_s}}\, ds \\ &= \sqrt{y_0} + \frac{1}{2}\int_0^t \frac{a - \frac{\sigma^2}{4}}{\sqrt{Y_s}}\, ds + \frac{\sigma}{2} W_t.\end{aligned}$$

It follows from the Girsanov theorem that the likelihood function is then given by

$$\left.\frac{d\mathbf{Q}}{d\mathbf{P}}\right|_{[0,T]} = \exp\left\{ \frac{\frac{\sigma^2}{4} - a}{\sigma} \int_0^T \frac{1}{\sqrt{Y_s}}\, dW_s - \frac{1}{2}\left(\frac{\frac{\sigma^2}{4} - a}{\sigma}\right)^2 \int_0^T \frac{ds}{Y_s} \right\}.$$

We aim to find the value of a that maximizes this likelihood function.

To proceed, we set

$$-\theta = \frac{\sigma}{4} - \frac{a}{\sigma}$$

and minimize the likelihood function with respect to θ, replacing

$$-\theta \int_0^T \frac{1}{\sqrt{Y_s}} \, dW_s = -\frac{2\theta}{\sigma} \int_0^T \frac{1}{\sqrt{Y_s}} \, d\sqrt{Y_s} + \theta^2 \int_0^T \frac{ds}{Y_s}.$$

Letting $Z_t = \sqrt{Y_t}$, the expression simplifies to minimizing the following:

$$-\frac{2\theta}{\sigma} \int_0^T \frac{dZ_s}{Z_s} + \frac{1}{2}\theta^2 \int_0^T \frac{ds}{Z_s^2}.$$

This minimization leads to the following maximum likelihood estimator:

$$\hat{\theta}_T = \frac{2 \int_0^T \frac{dZ_s}{Z_s}}{\sigma \int_0^T \frac{ds}{Z_s^2}}.$$

Proposition 22

Let $2a > \sigma^2$. Then $\hat{\theta}_T$ is a strongly consistent estimator of θ, i.e.,

$$\hat{\theta}_T \to \theta \quad \text{a.s., when } T \to \infty.$$

Corollary 23

Let $2a > \sigma^2$. The maximum likelihood estimator of the parameter a of the squared Bessel process (2) is given by

$$\hat{a}_T = \sigma\hat{\theta}_T + \frac{\sigma^2}{4} = \frac{2\int_0^T \frac{d\sqrt{Y_s}}{\sqrt{Y_s}}}{\int_0^T \frac{ds}{Y_s}} + \frac{\sigma^2}{4}$$

and it is strongly consistent, as $T \to \infty$.

Instability and some functional limit theorems for the squared Bessel process

As we have already mentioned, the process X is ergodic, while Y is non-ergodic. Now our goal is to establish how these properties reflect in the notion of stochastic instability. There are several approaches to stochastic instability of the processes. We shall consider the following definition, see book [Kulinich et al.(2020)].

Definition 24

A stochastic process ξ is called stochastically unstable if for any constant $N > 0$

$$\lim_{t \to +\infty} \frac{1}{t} \int_0^t \mathbf{P}\left\{|\xi_s| < N\right\} ds = 0.$$

Proposition 25

Let $N > 0$ be an arbitrary constant.

(i) *The CIR process X has the following property:*

$$\lim_{t \to +\infty} \frac{1}{t} \int_0^t \mathbf{P}\left\{X_s < N\right\} ds = \frac{\gamma\left(\frac{2a}{\sigma^2}, \frac{2bN}{\sigma^2}\right)}{\Gamma\left(\frac{2a}{\sigma^2}\right)},$$

where $\gamma(a, x) = \int_0^x u^{a-1} e^{-u}\, du$ is the lower incomplete Gamma function.

(ii) *The squared Bessel process Y is stochastically unstable, i.e.,*

$$\lim_{t \to +\infty} \frac{1}{t} \int_0^t \mathbf{P}\left\{Y_s < N\right\} ds = 0.$$

Remark 26

Proposition 25 justifies why we say that we have phase transition when coefficient b of CIR process tends to zero and we get squared Bessel process. Indeed, squared Bessel process is stochastically unstable while, oppositely, CIR process is ergodic, and in this sense, stochastically stable. Therefore, we use here the term "phase transition" in order to describe principally different long-term behaviour of the respective dynamical system. In more detail the difference of the asymptotic behaviour is described for example in the paper [Cherstvy et al.(2013)] where the authors analyze the transition from an ergodic model to a non-ergodic one, for which standard analytical and statistical methods are insufficient for fitting the data, necessitating the development and application of alternative tools.

Now, note that in some sense, we were lucky because knowledge of the distribution allowed us to establish instability of the squared Bessel process directly, without application of the tools of stochastic analysis. However, we can establish instability for the approximations of the squared Bessel process whose distribution is unknown. So, again, consider the squared Bessel process determined as the unique solution of the equation (2). Let for simplicity $\sigma = 2$. General condition $a \geq \frac{\sigma^2}{2}$ leads in our case to $a \geq 2$. So, we assume that $a \geq 2$, then the trajectories of the solution are strictly positive with probability 1. Therefore, we can consider function $F(x) = \sqrt{x}$ and apply Itô formula to Y_t, obtaining equation

$$\begin{aligned}\sqrt{Y_t} &= \sqrt{y_0} + \frac{1}{2}\int_0^t \frac{a - \frac{\sigma^2}{4}}{\sqrt{Y_s}}\, ds + \frac{\sigma}{2} W_t \\ &= \sqrt{y_0} + \frac{1}{2}\int_0^t \frac{a-1}{\sqrt{Y_s}}\, ds + W_t,\end{aligned}$$

or, that is the same,

$$V_t = V_0 + \frac{1}{2}\int_0^t \frac{a-1}{V_s}\,ds + W_t,$$

where $V_t = \sqrt{Y_t}$, $V_0 = \sqrt{Y_0}$. Note that $a - 1 \geq 1$. Now our goal is to consider a smooth version of Bessel process, i.e., a solution of the stochastic differential equation

$$V_t^\varepsilon = \int_0^t \frac{c ds}{\sqrt{(V_s^\varepsilon)^2 + \varepsilon^2}} + W_t, \tag{24}$$

where $\varepsilon \neq 0$, $c > 0$, $V_0^\varepsilon = \sqrt{y}_0 > 0$. The coefficient $\frac{c}{\sqrt{x^2+\varepsilon^2}}$ is Lipschitz because it has a bounded derivative:

$$\left|\left(\frac{c}{\sqrt{x^2+\varepsilon^2}}\right)'\right| = \frac{c\,|x|}{(x^2+\varepsilon^2)^{3/2}} \leq c\,\frac{|x|}{(x^2+\varepsilon^2)^{1/2}} \cdot \frac{1}{x^2+\varepsilon^2} \leq \frac{c}{\varepsilon^2}.$$

Also it is bounded, therefore, due the standard existence-uniqueness theorem for stochastic differential equations, equation (24) has a unique strong solution.

We want to achieve three goals. First, we establish the convergence of V^ε to the (non-squared) Bessel process $V = \sqrt{Y}$, as $\varepsilon \to 0$, that is more or less the expected result.

Proposition 27

Let $c = \frac{1}{2}(a-1)$, $\varepsilon \to 0$. *Then for any* $t > 0$

$$V_t^\varepsilon \to V_t \quad a.s.$$

Second, we wish to establish stochastic instability of V^ε.

Proposition 28

Let $c > 0$. *Then the processes* V^ε *are stochastically unstable for any* $\varepsilon \neq 0$, *i.e.,*

$$\lim_{t\to+\infty} \frac{1}{t} \int_0^t \mathbf{P}\left\{V_s^\varepsilon < N\right\} ds = 0$$

for any constant $N > 0$.

Remark 29

Stochastic instability of V^ε does not follow from Proposition 25 because $V^\varepsilon \le V$. Oppositely, stochastic instability of V follows from Proposition 28. However, having explicit formulas for the distributions of X and $Y = (V)^2$, we preferred to give the direct proof to Proposition 25.

And finally, we establish a bit unexpected statement: if $\varepsilon \neq 0$ is fixed, then the properly normalised processess V^ε weakly converge to the (non-squared) Bessel process as time tends to infinity.

Theorem 30

Normalised stochastic process $Z_\varepsilon(t) = \frac{Y^\varepsilon_{tT}}{\sqrt{T}}$ converges weakly, as $T \to \infty$, to the Bessel process Y_t that is the solution of the equation

$$Y_t^2 = 3t + 2\int_0^t Y_s\, dW_s, \quad t \ge 0. \tag{25}$$

Milton Abramowitz and Irene A. Stegun, editors.
Handbook of mathematical functions with formulas, graphs, and mathematical tables.
Dover Publications, Inc., New York, 1992.

Aurélien Alfonsi.
Affine diffusions and related processes: simulation, theory and applications, volume 6 of *Bocconi & Springer Series*.
Springer, Cham; Bocconi University Press, Milan, 2015.

Mohamed Ben Alaya and Ahmed Kebaier.
Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions.
Stoch. Anal. Appl., 31(4):552–573, 2013.

Damiano Brigo and Fabio Mercurio.
A deterministic-shift extension of analytically-tractable and time-homogeneous short-rate models.
Finance Stoch., 5(3):369–387, 2001.

Alexander S. Cherny and Hans-Jürgen Engelbert.
Singular stochastic differential equations, volume 1858 of *Lecture Notes in Mathematics*.
Springer-Verlag, Berlin, 2005.

Andrey G. Cherstvy, Aleksei V. Chechkin, and Ralf Metzler.
Anomalous diffusion and ergodicity breaking in heterogeneous diffusion processes.
New J. Phys., 15:083039, 13, 2013.

John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross.
A theory of the term structure of interest rates.
Econometrica, 53(2):385–407, 1985.

Olena Dehtiar, Yuliya Mishura, and Kostiantyn Ralchenko.
Two methods of estimation of the drift parameters of the Cox–Ingersoll–Ross process: continuous observations.
Comm. Statist. Theory Methods, 51(19):6818–6833, 2022.

Marco Di Francesco and Kevin Kamm.
How to handle negative interest rates in a CIR framework.
SeMA J., 79(4):593–618, 2022.

Albert Einstein.
Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen.
Annalen der Physik, 322(8):549–560, 1905.

Albert Einstein.
Investigations on the Theory of the Brownian Movement.
Dover publications, New-York, 1956.
Edited with notes by R. Fürth.

Albert Einstein and Marian von Smoluchowski.
Untersuchungen über die Theorie der Brownschen Bewegung. Abhandlung ?ber die Brownsche Bewegung und verwandte Erscheinungen.
Ostwalds Klassiker der exakten Wissenschaften, Band 199. Verlag Harri Deutsch, 1997.

Anja Göing-Jaeschke and Marc Yor.
A survey and some generalizations of Bessel processes.
Bernoulli, 9(2):313–349, 2003.

Nobuyuki Ikeda and Shinzo Watanabe.
Stochastic differential equations and diffusion processes, volume 24 of *North-Holland Mathematical Library*.
North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, second edition, 1989.

Ioannis Karatzas and Steven E. Shreve.
Brownian motion and stochastic calculus, volume 113 of *Graduate Texts in Mathematics*.
Springer-Verlag, New York, second edition, 1991.

Cónall Kelly and Gabriel J. Lord.
An adaptive splitting method for the Cox-Ingersoll-Ross process.
Appl. Numer. Math., 186:252–273, 2023.

Grigorij Kulinich, Svitlana Kushnirenko, and Yuliya Mishura.
Asymptotic analysis of unstable solutions of stochastic differential equations, volume 9 of *Bocconi & Springer Series*.
Bocconi University Press; Springer, 2020.

Yoosef Maghsoodi.
Solution of the extended CIR term structure and bond option valuation.
Math. Finance, 6(1):89–109, 1996.

Y. Mishura, S. Posashkova, and G. Shevchenko.
Properties of solutions of stochastic differential equations with nonhomogeneous coefficients and non-Lipschitz diffusion.
Theory Probab. Math. Statist., 79:117–126, 2009.

Yuliya Mishura and Georgiy Shevchenko.
Theory and Statistical Applications of Stochastic Processes.
ISTE Ltd and John Wiley & Sons, 2017.

Yuliya Mishura and Anton Yurchenko-Tytarenko.
Standard and fractional reflected Ornstein-Uhlenbeck processes as the limits of square roots of Cox-Ingersoll-Ross processes.
Stochastics, 95(1):99–117, 2023.

Yuliya Mishura, Andrey Pilipenko, and Anton Yurchenko-Tytarenko.
Low-dimensional Cox-Ingersoll-Ross process.
Stochastics, 2024.

Ostap Okhrin, Michael Rockinger, and Manuel Schmid.
Distributional properties of continuous time processes: from CIR to Bates.
AStA Adv. Stat. Anal., 107(3):397–419, 2023.

K. Oldham, J. Myland, and J. Spanier.
An atlas of functions.
Springer, New York, second edition, 2009.

Giuseppe Orlando, Rosa Maria Mininni, and Michele Bufalo.
Interest rates calibration with a CIR model.
The Journal of Risk Finance, 20(4):370–387, 2019.

Daniel Revuz and Marc Yor.
Continuous martingales and Brownian motion, volume 293 of *Grundlehren der mathematischen Wissenschaften*.
Springer-Verlag, Berlin, third edition, 1999.

Marian Von Smoluchowski.
Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen.
Annalen der physik, 326(14):756–780, 1906.

THANK YOU VERY MUCH FOR YOUR ATTENTION!