Two methods of estimation of the drift parameters of the Cox–Ingersoll–Ross process: Continuous observations

Olena Dehtiar, Yuliya Mishura, and Kostiantyn Ralchenko

Department of Probability, Statistics and Actuarial Mathematics, Mechanics and Mathematics Faculty, Taras Shevchenko National University of Kyiv, Kiev, Ukraine

ABSTRACT

We consider a stochastic differential equation of the form $dr_t = (a - br_t)dt + \sigma \sqrt{r_t} dW_t$, where $a, b$ and $\sigma$ are positive constants. The solution corresponds to the Cox–Ingersoll–Ross process. We study the estimation of an unknown drift parameter $(a, b)$ by continuous observations of a sample path $\{r_t, t \in [0, T]\}$. First, we prove the strong consistency of the maximum likelihood estimator. Since this estimator is well-defined only in the case $2a > \sigma^2$, we propose another estimator that is defined and strongly consistent for all positive $a, b, \sigma$. The quality of the estimators is illustrated by simulation results.

ARTICLE HISTORY

Received 1 May 2020
Accepted 14 December 2020

KEYWORDS

Cox–Ingersoll–Ross process; continuous observations; parameter estimation; strong consistency

1. Introduction

The Cox–Ingersoll–Ross (CIR) process is a very famous object and is a unique solution of the following stochastic differential equation

$$dr_t = (a - br_t)dt + \sigma \sqrt{r_t} dW_t, \quad r_t|_{t=0} = r_0 > 0$$

(1)

where $W = \{W_t, t \geq 0\}$ is a Wiener process and $a, b, \sigma$ are positive constants.

There are many papers devoted to the construction and the asymptotic properties of the parameter estimators of the CIR process, based on conditional least squares and the maximum likelihood estimator see, e.g., Overbeck and Rydén (1997); Li and Ma (2015); De Rossi (2010); Barczy et al. (2019); Ben Alaya and Kebaier (2012, 2013). More precisely, Overbeck and Rydén (1997) used conditional least squares as a basic method and derived two estimators, which differ by the method of estimating of parameter $\sigma$: pseudo likelihood method and unweighted least squares method based on squared residuals. Moreover, strong consistency of these estimators was proved for both observations at equidistant time point and continuous observations. Nevertheless, the simulation study demonstrated that the maximum likelihood estimator outperforms the conditional least-squares estimators and the pseudo likelihood approach in most cases.

Ben Alaya and Kebaier (2012, 2013) presented a new approach to the investigation of the asymptotic behavior of the maximum likelihood estimator, depending on the values
of the parameters, both for continuous and discrete observations. Using an exact simu-
lation algorithm, the authors illustrated practical behavior of these estimators’ errors
covering ergodic and nonergodic situations. Asymptotic properties of maximum likeli-
hood estimator for the stable CIR process were studied in Li and Ma (2015). In Barczy
et al. (2019) the authors investigated the asymptotic behavior of the maximum likeli-
hood estimator for so-called jump-type CIR process, driven by a standard Wiener pro-
cess and a subordinator. They distinguish three cases: subcritical, critical and
supercritical.

The paper De Rossi (2010) contains the method that is based on the sequential
Monte Carlo techniques and shows how to construct a simulated maximum likeli-
hood procedure. This paper also describes two methods of computing the likelihood:
sampling and re-sampling algorithm which solves the problem of degeneracy for real-
istic sample sizes. In order to maximize the likelihood, the author applies the genetic
algorithm that relies on the survival of the fittest in determining the optimal param-
eter vector. Testing on simulated data confirmed that this approach allows not to
undermine the accuracy of the estimation procedure by the effect of simulation errors
and copes with larger parameter dimensions at a modest computational cost.
Avoiding the computational burden the MATLAB implementation of the estimation
routine is provided and tested in Kladivko (2007). Moreover, the simulation algo-
rum for the approximation of the CIR process trajectories was described in Milstein
and Schoenmakers (2015).

Concerning the discretization, due to the square-root diffusion coefficient the classical
Euler–Maruyama scheme does not preserve the non negativity of the process. Cozma
and Reisinger (2016) provided a brief discussion of the discretization schemes often
encountered in the finance literature. They also investigated the exponential integrability
properties which play an essential role in deriving strong convergence of Euler discret-
ization schemes. The high priority to the convergence results and its applications were
given in the paper Deelstra and Delbaen (1995). The Euler approach to constructing
discretization schemes was also introduced in Dereich, Neuenkirch, and Szpruch (2012).
Here the use of the drift-implicit square-root Euler method gives a strictly positive
approximation of the original CIR process and some global convergence results. The
paper Mishura and Munchak (2016) is the example of constructing additive and multi-
plicative discrete approximation schemes taking the Euler approximations of the CIR
process itself but replacing the increments of the Wiener process with i. i. d. bounded
vanishing symmetric random variables.

In this paper we investigate two estimators of the parameter \((a, b)\) by continuous
observations of a sample path of CIR process \(r = \{r_t, t \in [0, T]\}\) and prove their strong
consistency as \(T \to \infty\). The first one is the standard maximum likelihood estimator,
which was constructed and studied in Ben Alaya and Kebaier (2012, 2013). Compared
to the known results, we establish the strong consistency instead of weak one. Note that
the maximum likelihood estimator is well-defined only if \(2a \geq \sigma^2\), because, in particu-
ar, it contains the integral \(\int_0^T \frac{1}{r_t} dt\), which exists with probability one if and only if \(2a \geq \sigma^2\), see (Ben Alaya and Kebaier 2012, Prop. 4). For this reason, we decided to create
some statistics that converge regardless of whether \(2a \geq \sigma^2\) or not. On this way we cre-
ated a different estimator of the vector parameter \((a, b)\), which is strongly consistent for
all positive $a$, $\sigma$ and $b$. Another advantage of the new alternative estimator is that it has simpler form, therefore, it is computationally faster. It includes only two statistics of the process $r$, namely the Lebesgue integrals $\int_0^T r_t \, dt$ and $\int_0^T r_t^2 \, dt$, see Theorem 5 below. At the same time the maximum likelihood estimator depends on two Lebesgue integrals, $\int_0^T r_t \, dt$ and $\int_0^T r_t^2 \, dt$, on the stochastic integral $\int_0^T \frac{dr_t}{r_t}$ and on the process itself. Particular attention in the paper is paid to the a.s. asymptotic behavior of the integral $\int_0^T r_t^2 \, dt$, which is crucial for the construction of the alternative estimator. We would like to emphasize that in the present paper we restrict ourselves to the case $b > 0$. The boundary case $b = 0$ was investigated in Ben Alaya and Kebaier (2013), where the consistency and asymptotic distribution on the maximum likelihood estimator was derived, assuming that $2a \geq \sigma^2$. It worth mentioning that there is no consistent estimator for $a$ in the case $b < 0$, see (Overbeck 1998, Thm. 2 (v)).

The paper is organized as follows. In Section 2 we recall some basic properties of CIR process. Section 3 is devoted to the strong consistency of the maximum likelihood estimator. In Section 4 we introduce an alternative estimator and prove its strong consistency.

2. Preliminaries

In this section we consider the properties of the CIR process. Most of them will be useful for statistical parameter estimation, but some facts are of independent interest. These facts are well-known, but we combined them into one statement for the reader’s convenience.

**Proposition 1.** Assume that $2a \geq \sigma^2$. Then

1. The unique solution $r = \{r_t, t \geq 0\}$ of Equation (1) is positive with probability 1:
   \[
   \inf \{t \geq 0 : r_t = 0\} = +\infty \quad \text{a.s.} \quad (2)
   \]
   (with convention $\inf \emptyset = +\infty$). Moreover,
   \[
   \mathbb{P} \left\{ \lim_{t \to \infty} \sup_{t \geq 0} r_t = +\infty \right\} = \mathbb{P} \left\{ \lim_{t \to \infty} \inf_{t \geq 0} r_t = 0 \right\} = 1. \quad (3)
   \]

2. The process $r$ is ergodic and it has continuous stationary density that corresponds to gamma distribution and has the following form:
   \[
   p_\infty(x) = \frac{\beta^2}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} 1_{x > 0}, \quad \alpha = \frac{2a}{\sigma^2}, \beta = \frac{2b}{\sigma^2}. \quad (4)
   \]

3. For any function $f : \mathbb{R}^+ \to \mathbb{R}$ such that $\int_{\mathbb{R}^+} |f(x)| p_\infty(x) \, dx < \infty$ we have that
   \[
   \frac{1}{T} \int_0^T f(r_t) \, dt \to \int_{\mathbb{R}} f(x) p_\infty(x) \, dx, \quad \text{a.s., as } T \to \infty. \quad (5)
   \]

The results of Proposition 1 follow from the general theory of homogeneous diffusions. In particular, Equation (2) follows from the Feller’s test for explosions (see, e.g., Karatzas and Shreve 1991, Thm. 5.29, p. 348), and Equation (3) can be deduced from (Karatzas and Shreve 1991, Prop. 5.22, p. 345) (see also Mijatović and Urusov 2012,
Sec. 2.3). Statements 2) – 3) are based on the ergodic theory for homogeneous diffusions, see, e. g., (Kutoyants 2004, Thm. 1.16) or (Skorokhod 1989, Ch. 1, § 3). The conditions of these general results for the case of the CIR process can be easily verified, see, e. g., (Bel Hadj Khelifa et al. 2016, Example 4, p. 280). A direct proof of Equation (2) for the CIR process is given in (Alfonsi 2015, Sec. 1.2.4). The derivation of the stationary distribution (4) can be found, e. g., in (Alfonsi 2015, Eq. (1.24)).

Corollary 2. It follows immediately from Equation (5) (see also Remark after Proposition 4 in Ben Alaya and Kebaier (2012)) that in the case $2a > \sigma^2$ we have the following asymptotic relations:

$$\frac{1}{T} \int_0^T r_t \, dt \to \int \chi \rho_\infty(x) \ dx = \frac{\chi}{\beta} = \frac{a}{b}, \quad \text{a.s., as } T \to \infty,$$

(6)

$$\frac{1}{T} \int_0^T \frac{1}{r_t} \, dt \to \int \frac{1}{x} \rho_\infty(x) \ dx = \frac{\beta}{\chi - 1} = \frac{b}{a - \sigma^2/2}, \quad \text{a.s., as } T \to \infty,$$

(7)

$$\frac{1}{T} \int_0^T r_t^2 \, dt \to \int x^2 \rho_\infty(x) \ dx = \left( \frac{\chi}{\beta} \right)^2 + \frac{\chi}{\beta^2} = \frac{a^2}{b^2} + \frac{a \sigma^2}{2b^2}, \quad \text{a.s., as } T \to \infty.$$

(8)

Remark 1. It was proved in (Deelstra and Delbaen 1995, Thm. 1) that the convergence (6) holds also in the case $0 < 2a \leq \sigma^2$. The convergence (8) is valid for all positive $a$ and $\sigma$, this will be justified in Theorem 4 below.

Remark 2. The convergence (7) holds also in the case $2a = \sigma^2$, if the right-hand side in Equation (7) is understood as $+\infty$. For the proof it suffices to consider the function $\frac{1}{x}$ in place of $\frac{1}{x}$, and then let $N \to \infty$.

3. Maximum likelihood estimation

Let us recall the construction of the maximum likelihood estimator of the couple of unknown parameters $(a, b)$ by the continuous observations of $r$ over the interval $[0, T]$. We assume that $2a \geq \sigma^2$ throughout this section.

Dividing Equation (1) by $\sqrt{r_t}$ and integrating Equation (1) over the time interval $[0, s]$ we get the equality:

$$\int_0^s \frac{dr_t}{\sqrt{r_t}} = \int_0^s \left( \frac{a}{\sqrt{r_t}} - b \sqrt{r_t} \right) dt + \sigma W_s.$$

In order to construct likelihood function for the estimation of couple $(a, b)$ of parameters, we use the Girsanov theorem for the Wiener process with the drift that equals

$$\frac{1}{\sigma} \int_0^s \left( \frac{a}{\sqrt{r_t}} - b \sqrt{r_t} \right) dt.$$

The arguments given below have a somewhat formal character, since during the construction of the likelihood function we are not interested in whether we are really
dealing with probability measures, we are only interested in the form of the likelihood function. So, according to Girsanov theorem, under some additional technical assumptions which we do not take into account now, the process $\int_0^t \left( \frac{a}{\sqrt{r_t}} - b \sqrt{r_t} \right) dt + \sigma W_t$ will be a Wiener process on the interval $[0, T]$ w.r.t. to the new measure $Q_{(0,0)}$ such that

$$dQ_{(0,0)}/dQ_{(a,b)} = \mathcal{L} = \exp \left\{ - \int_0^T \frac{a - br_t}{\sigma r_t} dW_t - \frac{1}{2} \int_0^T \frac{(a - br_t)^2}{\sigma^2 r_t} dt \right\} = \exp \left\{ - \int_0^T \frac{a - br_t}{\sigma^2 r_t} dr_t + \frac{1}{2} \int_0^T \frac{(a - br_t)^2}{\sigma^2 r_t} dt \right\}.$$  

Since we are interested in maximization of the function that corresponds to the value of $dQ_{(a,b)}/dQ_{(0,0)}$, and taking into account that $r$ is an observable process, our likelihood function gets the form

$$\tilde{\mathcal{L}} = \exp \left\{ \int_0^T \frac{a - br_t}{\sigma^2 r_t} dr_t - \frac{1}{2} \int_0^T \frac{(a - br_t)^2}{\sigma^2 r_t} dt \right\}.$$  

So, our approach gives us the likelihood function nd in what follows we are dealing with it. The maximum likelihood estimator for the couple $(a, b)$ is constructed by maximizing of $\tilde{\mathcal{L}}$ with respect to $(a, b)$. It has the following form:

$$\hat{a}_T = \frac{\int_0^T r_t dt \int_0^T \frac{dt}{r_t} - T \cdot (r_T - r_0)}{\int_0^T r_t dt \cdot \int_0^T \frac{dt}{r_t} - T^2}; \quad (9)$$

$$\hat{b}_T = \frac{(r_0 - r_T) \int_0^T \frac{dt}{r_t} + T \int_0^T \frac{dt}{r_t}}{\int_0^T r_t dt \cdot \int_0^T \frac{dt}{r_t} - T^2}. \quad (10)$$

**Theorem 3.** Assume that $2a \geq \sigma^2$. Then the estimator $(\hat{a}_T, \hat{b}_T)$ is strongly consistent.

**Proof.** Taking into consideration that $\int_0^T \frac{dt}{r_t} = a \int_0^T \frac{dt}{\sqrt{r_t}} - b T + \sigma \int_0^T \frac{dW_t}{\sqrt{r_t}}$, we can represent Equation (9) in the following form:

$$\hat{a}_T = \frac{a \int_0^T \frac{dt}{r_t} \int_0^T r_t dt - b T \int_0^T r_t dt + \sigma \int_0^T \frac{dW_t}{\sqrt{r_t}} \int_0^T r_t dt - T(r_T - r_0)}{\int_0^T r_t dt \cdot \int_0^T \frac{dt}{r_t} - T^2}$$

$$= \frac{a \int_0^T \frac{dt}{r_t} \int_0^T r_t dt - b T \int_0^T r_t dt + \sigma \int_0^T \frac{dW_t}{\sqrt{r_t}} \int_0^T r_t dt - T \int_0^T (a - br_t) dt - \sigma T \int_0^T \sqrt{r_t} dW_t}{\int_0^T r_t dt \cdot \int_0^T \frac{dt}{r_t} - T^2}$$

$$= a + \frac{\sigma \int_0^T \frac{dW_t}{\sqrt{r_t}} \cdot \int_0^T r_t dt - \sigma T \int_0^T \sqrt{r_t} dW_t}{\int_0^T r_t dt \cdot \int_0^T \frac{dt}{r_t} - T^2}, \quad (11)$$

and similarly, estimator from Equation (10) can be presented as
\[
\hat{b}_T = \frac{(r_0 - r_T) \int_0^T \frac{dt}{r_t} + aT \int_0^T \frac{dt}{r_t} - bT^2 + \sigma T \int_0^T dW_t}{\int_0^T r_t dt \cdot \int_0^T \frac{dt}{r_t} - T^2 }
\]
\[
= - \int_0^T (a - br_t) dt \cdot \int_0^T \frac{dt}{r_t} - \sigma \int_0^T \sqrt{r_t} dW_t \cdot \int_0^T \frac{dt}{r_t} + aT \int_0^T \frac{dt}{r_t} - bT^2 + \sigma T \int_0^T \frac{dW_t}{\sqrt{r_t}}.
\]
\[
= b + \frac{-\sigma \int_0^T \sqrt{r_t} dW_t \cdot \int_0^T \frac{dt}{r_t} + \sigma T \int_0^T \frac{dW_t}{\sqrt{r_t}}}{\int_0^T r_t dt \cdot \int_0^T \frac{dt}{r_t} - T^2 }.
\]
(12)

Consider the remainder

\[
R^a_T = \frac{\sigma \int_0^T \frac{dW_t}{\sqrt{r_t}} \cdot \int_0^T r_t dt - \sigma T \int_0^T \frac{dW_t}{\sqrt{r_t}}}{\int_0^T r_t dt \cdot \int_0^T \frac{dt}{r_t} - T^2 }
\]

from Equation (11), and the remainder

\[
R^b_T = \frac{-\sigma \int_0^T \sqrt{r_t} dW_t \cdot \int_0^T \frac{dt}{r_t} + \sigma T \int_0^T \frac{dW_t}{\sqrt{r_t}}}{\int_0^T r_t dt \cdot \int_0^T \frac{dt}{r_t} - T^2 }
\]

from Equation (12) can be considered in the same lines. In the case \(2a > \sigma^2\) we rewrite \(R^b_T\) in the following form

\[
R^b_T = \frac{\sigma \left( \frac{1}{T} \int_0^T \frac{dW_t}{\sqrt{r_t}} \right) \cdot \left( \frac{1}{T} \int_0^T r_t dt \right)}{\frac{1}{T^2} \int_0^T r_t dt \cdot \int_0^T \frac{dt}{r_t} - 1}.
\]
(13)

According to Corollary 2, relations (5) and (6), the denominator in Equation (13) tends to \(\frac{a}{a - \sigma^2} - 1 > 0\). Moreover, both values in the numerator, \(\frac{1}{T} \int_0^T \frac{dW_t}{\sqrt{r_t}}\) and \(\frac{1}{T} \int_0^T \sqrt{r_t} dW_t\) tend to zero a.s. as \(T \to \infty\). Indeed,

\[
\frac{1}{T} \int_0^T \frac{dW_t}{\sqrt{r_t}} = \frac{\int_0^T \frac{dt}{r_t} \cdot \int_0^T \frac{dW_t}{\sqrt{r_t}}}{\int_0^T \frac{dt}{r_t}},
\]

and

\[
\frac{\int_0^T \frac{dt}{r_t}}{T} \to \frac{b}{a - \frac{\sigma^2}{2}}.
\]

This means that \(\int_0^T \frac{dt}{r_t} \to \infty\), \(T \to \infty\), but \(\int_0^T \frac{dt}{r_t}\) is a square characteristics of the locally square integrable martingale \(\int_0^T \frac{dW_t}{\sqrt{r_t}}\). According to the strong law of large numbers for locally square integrable martingales (Liptser and Shiryaev 1989), \((\int_0^T \frac{dt}{r_t})^{-1} \cdot \int_0^T \frac{dW_t}{\sqrt{r_t}} \to 0\) a.s. as \(T \to \infty\). Therefore, \(\frac{1}{T} \int_0^T \frac{dW_t}{\sqrt{r_t}} \to 0\) a.s. as \(T \to \infty\), and similar relation holds for \(\frac{1}{T} \int_0^T \sqrt{r_t} dW_t\). Together with Equation (13), this means that \(\hat{a}_T\) is strongly consistent. The case \(2a = \sigma^2\) is considered similarly, taking into account that in this case \(\frac{1}{T} \int_0^T \frac{1}{r_t} dt \to +\infty\) a.s. as \(T \to \infty\). \(\Box\)
Remark 3. The weak consistency and asymptotic normality of the maximum likelihood estimator \((\hat{a}, \hat{b})\) was obtained in Ben Alaya and Kebaier (2012, 2013). Mention that the weak consistency holds also in the boundary cases, when \(2a = \sigma^2\) and/or \(b = 0\) (however, joint asymptotic distribution is not normal in these cases). If \(2a < \sigma^2\), then the maximum likelihood estimator is not well-defined. If \(2a \geq \sigma^2\) and \(b < 0\), then it is not consistent.

Remark 4. An alternative parametrization of the CIR process is

\[
d r_t = \alpha (\mu - r_t) dt + \sigma \sqrt{r_t} dW_t, \quad r_t|_{t=0} = r_0 > 0,
\]

(14)

where \(W = \{W_t, t \geq 0\}\) is a Wiener process and \(\alpha, \mu, \sigma\) are positive constants. Assume that \(2\alpha\mu \geq \sigma^2\). Then the maximum likelihood estimator for the couple \((\alpha, \mu)\) constructed by the continuous observations of \(r_t\) over the interval \([0, T]\) has the following form

\[
\hat{\alpha}_T = \frac{T \int_0^T \frac{dr_t}{r_t} - \int_0^T \frac{dt}{r_t} \int_0^T dr_t}{\int_0^T r_t dt \int_0^T \frac{dt}{r_t} - T^2}, \quad \hat{\mu}_T = \frac{\int_0^T r_t dt \int_0^T \frac{dr_t}{r_t} - T(r_T - r_0)}{T \int_0^T \frac{dr_t}{r_t} - \int_0^T \frac{dt}{r_t} (r_T - r_0)}.
\]

The strong consistency of \((\hat{\alpha}_T, \hat{\mu}_T)\) follows from Theorem 3, if we take into account the relations

\[
\alpha = b, \quad \mu = \frac{a}{b}, \quad \hat{\alpha}_T = \hat{b}_T, \quad \hat{\mu}_T = \frac{\hat{a}_T}{\hat{b}_T}.
\]

(15)

4. An alternative approach to drift parameters estimation

The disadvantage of the maximum likelihood estimators is that they work only if \(a > \sigma^2/2\), however, a priori we do not know if this relation holds for the observed process. To avoid this circumstance, in this section we will introduce a new estimator for the parameter \((a, b)\) based on the statistics \(\int_0^T r_t \ dt\) and \(\int_0^T r_t^2 \ dt\). First, we will prove that the convergence (8) remains valid in the case \(0 < a \leq \sigma^2/2\).

Theorem 4. The following convergence holds:

\[
\frac{1}{T} \int_0^T r_t^2 \ dt \to \frac{a^2}{b^2} + \frac{a\sigma^2}{2b^2} \quad \text{a.s., as } T \to \infty.
\]

Proof. Using relation (27), we get the following equalities:

\[
\frac{1}{T} \int_0^T r_t^2 \ dt = \frac{1}{T} \int_0^T e^{-2bt} \left( r_0 - \frac{a}{b} e^{bt} + \sigma \int_0^t e^{bu} \sqrt{r_u} \ dW_u \right)^2 \ dt
\]

\[
= \frac{1}{T} \int_0^T e^{-2bt} \left( r_0 - \frac{a}{b} e^{bt} \right)^2 \ dt
\]

\[
+ \frac{2\sigma}{T} \int_0^T e^{-2bt} \left( r_0 - \frac{a}{b} e^{bt} \right) \left( \int_0^t e^{bu} \sqrt{r_u} \ dW_u \right) \ dt
\]

\[
+ \frac{\sigma^2}{T} \int_0^T e^{-2bt} \left( \int_0^t e^{bu} \sqrt{r_u} \ dW_u \right)^2 \ dt
\]

\[
=: I_1 + I_2 + I_3.
\]
The term $I_1$ is the subject of straightforward calculations:

$$I_1 = \frac{1}{T} \int_0^T e^{-2bt} \left( r_0 - \frac{a}{b} + \frac{a}{b} e^{bt} \right)^2 \, dt$$

$$= \frac{1}{T} \int_0^T \left( r_0 - \frac{a}{b} \right)^2 e^{-2bt} + \frac{2a}{b} \left( r_0 - \frac{a}{b} \right) e^{-bt} + \frac{a^2}{b^2} \, dt.$$  

$$= \left( r_0 - \frac{a}{b} \right)^2 \frac{1 - e^{-2bT}}{2bT} + \frac{2a}{b} \left( r_0 - \frac{a}{b} \right) \frac{1 - e^{-bT}}{bT} + \frac{a^2}{b^2}.$$  

Evidently, the following convergence holds:

$$I_1 \to \frac{a^2}{b^2}, \quad T \to \infty. \quad (17)$$

Let us consider $I_2$. The equality (27) implies that

$$\sigma \int_0^t e^{bu} \sqrt{r_u} \, dW_u = e^{bt} r_t - r_0 + \frac{a}{b} e^{bt}.$$  

Consequently, integral $I_2$ can be transformed as follows:

$$I_2 = 2\sigma \int_0^T e^{-2bt} \left( r_0 - \frac{a}{b} + \frac{a}{b} e^{bt} \right) \left( \int_0^t e^{bu} \sqrt{r_u} \, dW_u \right) \, dt$$

$$= 2 \int_0^T e^{-2bt} \left( r_0 - \frac{a}{b} + \frac{a}{b} e^{bt} \right) \left( e^{bt} r_t - r_0 + \frac{a}{b} e^{bt} \right) \, dt$$

$$= 2 \left( r_0 - \frac{a}{b} \right) \int_0^T e^{-bt} r_t \, dt + \frac{2a}{bT} \int_0^T r_t \, dt - \frac{2}{T} \int_0^T e^{-2bt} \left( r_0 - \frac{a}{b} + \frac{a}{b} e^{bt} \right)^2 \, dt. \quad (18)$$

Asymptotic relation (23) implies that the first term in the right-hand side of Equation (18) converges to zero a.s. as $T \to \infty$. By Equation (6), for the second term we have

$$\frac{2a}{bT} \int_0^T r_t \, dt \to \frac{2a^2}{b^2} \quad \text{a.s., as} \; T \to \infty.$$  

Finally, the third term equals $-2I_1$, therefore, by (17),

$$-\frac{2}{T} \int_0^T e^{-2bt} \left( r_0 - \frac{a}{b} + \frac{a}{b} e^{bt} \right)^2 \, dt = -2I_1 \to -\frac{2a^2}{b^2}, \quad \text{as} \; T \to \infty.$$  

Thus,

$$I_2 \to 0 \quad \text{a.s., as} \; T \to \infty. \quad (19)$$

It remains to study the asymptotic behavior of $I_3$. Note that

$$I_3 = \frac{\sigma^2}{T} \int_0^T e^{-2bt} Z_t^2 \, dt,$$

where $Z_t$ is defined in Equation (25). By Itô’s formula, from Equation (25) we get

$$Z_t^2 = \int_0^t e^{2bs} r_s \, ds + 2 \int_0^t Z_s e^{bs} \sqrt{r_s} \, dW_s.$$
Therefore we can present $I_3$ as the sum
\begin{equation}
I_3 = \frac{\sigma^2}{T} \int_0^T e^{-2bt} \int_0^t e^{2bs} r_s \, ds \, dt + \frac{2\sigma^2}{T} \int_0^T e^{-2bt} \int_0^t Z_s e^{bs} \sqrt{r_s} \, dW_s \, dt =: I_{3,1} + I_{3,2}.
\end{equation}

Using the Fubini theorem, we transform $I_{3,1}$ as follows:
\begin{align*}
I_{3,1} &= \frac{\sigma^2}{T} \int_0^T e^{2bs} r_s \int_s^T e^{-2bt} \, dt \, ds = \frac{\sigma^2}{T} \int_0^T e^{2bs} r_s \frac{e^{-2bs} - e^{-2bT}}{2b} \, ds \\
&= \frac{\sigma^2}{2bT} \int_0^T r_s \, ds - \frac{\sigma^2}{2bT} e^{-2bT} \int_0^T e^{2bs} r_s \, ds.
\end{align*}

Using asymptotic relations (6) and (24), we obtain
\begin{equation}
I_{3,1} \to \frac{aa^2}{2b^2}, \quad \text{a.s., as } T \to \infty.
\end{equation}

Changing the order of integration, we rewrite $I_{3,2}$ in the following way:
\begin{align*}
I_{3,2} &= \frac{2\sigma^2}{T} \int_0^T Z_s e^{bs} \sqrt{r_s} \int_s^T e^{-2bt} \, dt \, dW_s = \frac{2\sigma^2}{T} \int_0^T Z_s e^{bs} \sqrt{r_s} \frac{e^{-2bs} - e^{-2bT}}{2b} \, dt \, dW_s \\
&= \frac{\sigma^2}{bT} \int_0^T e^{-bs} Z_s \sqrt{r_s} \, dt \, dW_s - \frac{\sigma^2}{bT} e^{-2bT} \int_0^T e^{bs} Z_s \sqrt{r_s} \, dt \, dW_s.
\end{align*}

According to Lemma 8, both stochastic integrals in the right-hand side converge to zero a.s. Therefore, $I_{3,2} \to 0$ a.s., as $T \to \infty$. Combining this with Equations (20) and (21), we see that
\begin{equation}
I_3 \to \frac{aa^2}{2b^2} \quad \text{a.s., as } T \to \infty.
\end{equation}

Finally, inserting the convergences (17, 19, 22) into the equality (16), we conclude the proof.

Theorem 4 enables to construct a strongly consistent estimator for the parameter $(a, b)$.

**Theorem 5.** Define
\begin{align*}
\tilde{a}_T &= \frac{\sigma^2}{2} \cdot \frac{\left( \int_0^T r_t \, dt \right)^2}{T \int_0^T r_t^2 \, dt - \left( \int_0^T r_t \, dt \right)^2}, \\
\tilde{b}_T &= \frac{\sigma^2}{2} \cdot \frac{T \int_0^T r_t \, dt}{T \int_0^T r_t^2 \, dt - \left( \int_0^T r_t \, dt \right)^2}.
\end{align*}

Then vector $(\tilde{a}_T, \tilde{b}_T)$ is a strongly consistent estimator of $(a, b)$.

**Proof.** It follows from the convergence (6) and Theorem 4 that
\begin{align*}
\tilde{b}_T &= \frac{\sigma^2}{2} \cdot \frac{\frac{1}{T} \int_0^T r_t \, dt}{\frac{1}{T} \int_0^T r_t^2 \, dt - \left( \frac{1}{T} \int_0^T r_t \, dt \right)^2} \to \frac{\sigma^2}{2} \cdot \frac{a}{b^2} + \frac{aa^2}{2b^2} - \frac{a^2}{b^2} = b, \quad \text{a.s., as } T \to \infty.
\end{align*}
Further, from Equation (6) we have the convergence

\[ \tilde{a}_T = \bar{b}_T \cdot \frac{1}{T} \int_0^T r_t \, dt \to b \cdot \frac{a}{b} = a, \quad \text{a.s., as } T \to \infty. \]

Remark 5. For the model (14) the alternative estimator of \((x, \mu)\) can be defined as follows:

\[ \tilde{x}_T = \frac{\sigma^2}{2} \cdot \frac{T \int_0^T r_t \, dt}{T \int_0^T r_t^2 \, dt - \left( \int_0^T r_t \, dt \right)^2}, \quad \tilde{\mu}_T = \frac{1}{T} \int_0^T r_t \, dt. \]

The strong consistency of the estimator \((\tilde{x}_T, \tilde{\mu}_T)\) follows from Theorem 5 and the relations (15).

Table 1. Means and standard deviations of \(\tilde{a}_T\) and \(\tilde{b}_T\) for \(2a > \sigma^2\).

| \(a\) | \(b\) | \(\sigma\) | 10  | 50  | 100 | 150 | 200 |
|------|------|------|-----|-----|-----|-----|-----|
| 1    | 1    | 1    | 1.3520 | 1.0507 | 1.0187 | 1.0105 | 1.0073 |
|      |      |      | 0.5357 | 0.1628 | 0.0983 | 0.0818 | 0.0723 |
|      |      |      | 1.5328 | 1.1152 | 1.0535 | 1.0300 | 1.0226 |
|      |      |      | 0.5804 | 0.2435 | 0.1701 | 0.1306 | 0.1203 |
| 1    | 2    | 1    | 1.2013 | 1.0401 | 1.0144 | 1.0082 | 1.0046 |
|      |      |      | 0.3201 | 0.1077 | 0.0700 | 0.0557 | 0.0495 |
|      |      |      | 1.3028 | 1.0766 | 1.0270 | 1.0206 | 1.0140 |
|      |      |      | 0.4304 | 0.1618 | 0.1051 | 0.0974 | 0.0872 |
| 1    | 3    | 1    | 1.0915 | 1.0201 | 1.0192 | 1.0129 | 1.0141 |
|      |      |      | 0.2029 | 0.0821 | 0.0559 | 0.0470 | 0.0459 |
|      |      |      | 1.1142 | 1.0411 | 1.0241 | 1.0252 | 1.0239 |
|      |      |      | 0.2961 | 0.1282 | 0.0944 | 0.0810 | 0.0778 |

Table 1. Means and standard deviations of \(\tilde{a}_T\) and \(\tilde{b}_T\) for \(2a > \sigma^2\).
5. Simulations

In this section we illustrate the quality of the estimators, using Monte Carlo simulation. For each set of parameters $(a, b, \sigma)$, we generate 100 sample paths of the solution $r = \{r_t, t \in [0, T]\}$ to the Equation (1) using Euler’s approximation. We choose the initial value $r_0 = 1$ for all simulations. Then we compute means and standard deviations of the estimators at the times $T = 10, 50, 100, 150, 200$.

In the case $2a > \sigma^2$ we compare two estimators, $(\hat{a}_T, \hat{b}_T)$ and $(\tilde{a}_T, \tilde{b}_T)$. Tables 1 and 2 report the simulation results concerning the estimation of $a$ and $b$ respectively. We see that both estimators are consistent and behave similarly, however, the maximum likelihood estimator $(\hat{a}_T, \hat{b}_T)$ slightly outperforms the alternative estimator $(\tilde{a}_T, \tilde{b}_T)$.

For the case $2a < \sigma^2$ the maximum likelihood estimator does not make sense, so only the means and variances for $(\tilde{a}_T, \tilde{b}_T)$ are reported, see Tables 3 and 4. Clearly, the

| $a$ | $b$ | $\sigma$ | $T$ |
|-----|-----|---------|-----|
| 1   | 1   | 1       | 10  |
|     |     |         | 1.4514 | 1.0772 | 1.0244 |
|     |     |         | 0.5969 | 0.2184 | 0.1398 |
| 1   | 2   | 1       | 2.4399 | 2.1229 | 2.0677 |
|     |     |         | 0.6993 | 0.2658 | 0.1941 |
| 1   | 3   | 1       | 3.3578 | 3.0316 | 3.0287 |
|     |     |         | 0.8546 | 0.3701 | 0.2479 |
| 2   | 1   | 1       | 1.4767 | 1.0521 | 0.9980 |
|     |     |         | 0.6254 | 0.1944 | 0.1323 |
| 2   | 2   | 1       | 2.3087 | 2.0703 | 2.0437 |
|     |     |         | 0.6851 | 0.2865 | 0.1996 |
| 2   | 3   | 1       | 3.3561 | 3.0417 | 3.0420 |
|     |     |         | 0.8854 | 0.3546 | 0.2705 |
| 3   | 1   | 1       | 1.3899 | 1.0798 | 1.0337 |
|     |     |         | 0.5563 | 0.1882 | 0.1525 |
| 3   | 1   | 2       | 1.3616 | 1.0617 | 1.0209 |
|     |     |         | 0.4884 | 0.2102 | 0.1361 |
| 3   | 1   | 2       | 1.6055 | 1.1483 | 1.0447 |
|     |     |         | 0.6558 | 0.3057 | 0.2179 |

Table 2. Means and standard deviations of $\hat{b}_T$ and $\tilde{b}_T$ for $2a > \sigma^2$. 

| $a$ | $b$ | $\sigma$ | $T$ |
|-----|-----|---------|-----|
| 1   | 1   | 1       | 150 |
|     |     |         | 1.0176 | 0.1086 |
|     |     |         | 1.0363 | 0.0986 |
| 1   | 2   | 1       | 2.0335 | 2.0159 |
|     |     |         | 0.1693 | 0.1587 |
| 1   | 3   | 1       | 3.0094 | 3.0162 |
|     |     |         | 0.2071 | 0.1949 |
| 2   | 1   | 1       | 1.0056 | 1.0120 |
|     |     |         | 0.1214 | 0.1101 |
| 2   | 2   | 1       | 2.0104 | 2.0074 |
|     |     |         | 0.2007 | 0.1755 |
| 2   | 3   | 1       | 3.0354 | 3.0311 |
|     |     |         | 0.1811 | 0.1619 |
| 3   | 1   | 1       | 1.0264 | 1.0262 |
|     |     |         | 0.1283 | 0.1109 |
| 3   | 1   | 2       | 1.0038 | 0.9989 |
|     |     |         | 0.1011 | 0.0977 |
| 3   | 1   | 2       | 1.0220 | 1.0185 |
|     |     |         | 0.1451 | 0.1250 |

| $a$ | $b$ | $\sigma$ | $T$ |
|-----|-----|---------|-----|
| 1   | 1   | 1       | 200 |
|     |     |         | 1.0132 | 0.0986 |
|     |     |         | 1.0271 | 0.0986 |
| 1   | 2   | 1       | 2.0159 | 2.0074 |
|     |     |         | 0.1587 | 0.1498 |
| 1   | 3   | 1       | 3.0162 | 3.0094 |
|     |     |         | 0.1949 | 0.1949 |
| 2   | 1   | 1       | 1.0120 | 1.0074 |
|     |     |         | 0.1101 | 0.1004 |
| 2   | 2   | 1       | 2.0074 | 1.9755 |
|     |     |         | 0.1755 | 0.1324 |
| 2   | 3   | 1       | 3.0311 | 3.0354 |
|     |     |         | 0.1619 | 0.1375 |
| 3   | 1   | 1       | 1.0262 | 1.0262 |
|     |     |         | 0.1109 | 0.1004 |
| 3   | 1   | 2       | 1.0185 | 1.0185 |
|     |     |         | 0.1250 | 0.1004 |

Table 3. Means and standard deviations of $\hat{b}_T$ and $\tilde{b}_T$ for $2a < \sigma^2$.
numerical results confirm the consistency of this estimator. However, in many cases the rate of convergence is quite slow (compared to the case $2a > \sigma^2$), especially for $\hat{b}_T$.

Table 3. Means and standard deviations $\hat{a}_T$ for $2a < \sigma^2$.

| $a$ | $b$ | $\sigma$ | $T$ | $E[\hat{a}_T]$ | $\sigma[\hat{a}_T]$ |
|-----|-----|---------|----|----------------|-------------------|
| 1   | 1   | 2       | 10 | 1.6220        | 0.6134           |
|     |     |         | 50 | 1.1125        | 0.2720           |
|     |     |         | 100| 1.0583        | 0.2164           |
|     |     |         | 150| 1.0384        | 0.1826           |
|     |     |         | 200| 1.0348        | 0.1589           |
| 1   | 1   | 3       | 10 | 1.3276        | 0.3762           |
|     |     |         | 50 | 1.1253        | 0.2141           |
|     |     |         | 100| 1.1256        | 0.1442           |
|     |     |         | 150| 1.1407        | 0.1317           |
|     |     |         | 200| 1.1091        | 0.1144           |
| 1   | 2   | 2       | 10 | 1.3309        | 2.5947           |
|     |     |         | 50 | 1.0907        | 0.6717           |
|     |     |         | 100| 1.0570        | 0.5451           |
|     |     |         | 150| 1.0357        | 0.4372           |
|     |     |         | 200| 1.0212        | 0.4372           |
| 1   | 3   | 2       | 10 | 1.3113        | 1.3936           |
|     |     |         | 50 | 1.1099        | 0.2141           |
|     |     |         | 100| 1.0618        | 0.1442           |
|     |     |         | 150| 1.0432        | 0.1317           |
|     |     |         | 200| 1.0330        | 0.1144           |
| 1   | 3   | 3       | 10 | 1.3276        | 2.5786           |
|     |     |         | 50 | 1.1253        | 0.6717           |
|     |     |         | 100| 1.1256        | 0.5451           |
|     |     |         | 150| 1.1407        | 0.4372           |
|     |     |         | 200| 1.1091        | 0.3840           |

Table 4. Means and standard deviations $\hat{b}_T$ for $2a < \sigma^2$.

| $a$ | $b$ | $\sigma$ | $T$ | $E[\hat{b}_T]$ | $\sigma[\hat{b}_T]$ |
|-----|-----|---------|----|----------------|-------------------|
| 1   | 1   | 2       | 10 | 2.0929        | 2.0929           |
|     |     |         | 50 | 3.4455        | 5.5096           |
|     |     |         | 100| 3.2534        | 0.2816           |
|     |     |         | 150| 3.1348        | 0.2316           |
|     |     |         | 200| 3.0772        | 0.1994           |
| 1   | 1   | 3       | 10 | 4.5037        | 4.5037           |
|     |     |         | 50 | 3.4455        | 3.4455           |
|     |     |         | 100| 3.2534        | 3.2534           |
|     |     |         | 150| 3.1348        | 3.1348           |
|     |     |         | 200| 3.0772        | 3.0772           |
| 1   | 2   | 2       | 10 | 2.7734        | 1.2187           |
|     |     |         | 50 | 2.2172        | 0.4458           |
|     |     |         | 100| 2.1665        | 0.3275           |
|     |     |         | 150| 2.1346        | 0.2733           |
|     |     |         | 200| 2.1137        | 0.2507           |
| 2   | 3   | 3       | 10 | 2.5827        | 0.7142           |
|     |     |         | 50 | 2.1628        | 0.3657           |
|     |     |         | 100| 2.1024        | 0.2816           |
|     |     |         | 150| 2.0647        | 0.2316           |
|     |     |         | 200| 2.0350        | 0.1994           |
| 3   | 1   | 3       | 10 | 4.5037        | 4.5037           |
|     |     |         | 50 | 3.4455        | 3.4455           |
|     |     |         | 100| 3.2534        | 3.2534           |
|     |     |         | 150| 3.1348        | 3.1348           |
|     |     |         | 200| 3.0772        | 3.0772           |

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
Acknowledgments

We are grateful to the referees for their useful and valuable comments.

Funding

This work was supported by the National Research Fund of Ukraine under Grant 2020.02/0026.

ORCID

Yuliya Mishura http://orcid.org/0000-0002-6877-1800
Kostiantyn Ralchenko http://orcid.org/0000-0001-7208-3130

References

Alfonsi, A. 2015. Affine diffusions and related processes: Simulation, theory and applications, volume 6 of Bocconi & Springer Series. Cham: Springer

Barczy, M., M. Ben Alaya, A. Kebaier, and G. Pap. 2019. Asymptotic properties of maximum likelihood estimator for the growth rate of a stable CIR process based on continuous time observations. Statistics 53 (3):533–68. doi:10.1080/02331888.2019.1579216.

Bel Hadj Khelifa, M., Y. Mishura, K. Ralchenko, and M. Zili. 2016. Drift parameter estimation in stochastic differential equation with multiplicative stochastic volatility. Modern Stochastics: Theory and Applications 3 (4):269–85. doi:10.15559/16-VMSTA66.

Ben Alaya, M., and A. Kebaier. 2012. Parameter estimation for the square-root diffusions: Ergodic and nonergodic cases. Stochastic Models 28 (4):609–34. doi:10.1080/15326349.2012.726042.

Ben Alaya, M., and A. Kebaier. 2013. Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions. Stochastic Analysis and Applications 31 (4):552–73. doi:10.1080/07362994.2013.798175.

Cozma, A., and C. Reisinger. 2016. Exponential integrability properties of Euler discretization schemes for the Cox–Ingersoll–Ross process. Discrete and Continuous Dynamical Systems - Series B 21 (10):3359–77. doi:10.3934/dcdsb.2016101.

Deelstra, G., and F. Delbaen. 1995. Long-term returns in stochastic interest rate models. Insurance: Mathematics and Economics 17 (2):163–69. doi:10.1016/0167-6687(95)00018-N.

Dereich, S., A. Neuenkirch, and L. Szpruch. 2012. An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 468 (2140):1105–15. doi:10.1098/rspa.2011.0505.

Karatzas, I., and S. E. Shreve. 1991. Brownian motion and stochastic calculus, volume 113 of Graduate Texts in Mathematics, 2nd ed. New York: Springer-Verlag.

Kladívková, K. 2007. Maximum likelihood estimation of the Cox–Ingersoll–Ross process: The matlab implementation. Technical Computing Prague 7:8.

Kutoyants, Y. A. 2004. Statistical inference for ergodic diffusion processes. Springer Series in Statistics. London: Springer-Verlag London, Ltd.

Li, Z., and C. Ma. 2015. Asymptotic properties of estimators in a stable Cox–Ingersoll–Ross model. Stochastic Processes and Their Applications 125 (8):3196–233. doi:10.1016/j.spa.2015.03.002.

Liptser, R., and A. N. Shiryaev. 1989. Theory of martingales. Mathematics and its applications (Soviet Series), vol. 49. Dordrecht: Kluwer Academic Publishers Group.

Mijatović, A., and M. Urusov. 2012. Convergence of integral functionals of one-dimensional diffusions. Electron. Commun. Probab 17 (61):13.

Milstein, G. N., and J. Schoenmakers. 2015. Uniform approximation of the Cox–Ingersoll–Ross process. Advances in Applied Probability 47 (4):1132–56. doi:10.1239/aap/1449859803.
Appendix

In this appendix we prove several auxiliary results. The first result gives the asymptotics of two normalized Lebesgue integrals.

Lemma 6. Let $a > 0$, $b > 0$, $\sigma > 0$. Then the following normalized integrals asymptotically vanish as $T \to \infty$:

\[ \frac{1}{T} \int_0^T e^{-br_t} \, dt \to 0 \quad \text{a.s.,} \tag{23} \]

\[ \frac{1}{T^2} \int_0^T e^{-2br_t} \, dt \to 0 \quad \text{a.s.} \tag{24} \]

Proof. 1. In order to prove the asymptotic relation Equation (23), we rewrite the normalized integral as follows:

\[ \frac{1}{T} \int_0^T e^{-br_t} \, dt = \frac{1}{\sqrt{T}} \int_0^{\sqrt{T}} e^{-br_t} \, dt + \frac{1}{\sqrt{T}} \int_{\sqrt{T}}^T e^{-br_t} \, dt. \]

Then the first integral can be bounded in the following way:

\[ \frac{1}{\sqrt{T}} \int_0^{\sqrt{T}} e^{-br_t} \, dt \leq \frac{1}{\sqrt{T}} \int_0^\infty r_t \, dt = \frac{1}{\sqrt{T}} \int_0^{\sqrt{T}} r_t \, dt \to 0 \quad \text{a.s., as } T \to \infty, \]

since $\frac{1}{\sqrt{T}} \int_0^{\sqrt{T}} r_t \, dt \to \frac{\sigma}{T}$ a.s., as $T \to \infty$, by Equation (6), see Remark 1. Furthermore, the second integral can be bounded as

\[ \frac{1}{\sqrt{T}} \int_{\sqrt{T}}^T e^{-br_t} \, dt \leq e^{-b\sqrt{T}} \frac{1}{\sqrt{T}} \int_0^\infty r_t \, dt \leq e^{-b\sqrt{T}} \frac{1}{\sqrt{T}} \int_0^\infty r_t \, dt \to 0 \quad \text{a.s., as } T \to \infty, \]

where the convergence follows from Equation (6). Thus, relation Equation (23) is proved.

2. Note that $\int_0^T e^{2br_t} \, dt \geq \int_0^T r_t \, dt \to \infty$ a.s., as $T \to \infty$, by Equation (6). Therefore, applying L'Hôpital's rule, we conclude that

\[ \lim_{T \to \infty} \frac{\int_0^T e^{2br_t} \, dt}{Te^{2bT}} \quad \text{is } \lim_{T \to \infty} \frac{e^{2bT} r_T}{1 + 2bT} \quad \text{a. s.} \]

Now Equation (24) follows from the a. s. convergence $\frac{r_T}{T} \to 0$, $T \to \infty$. The latter convergence was established in the proof of Theorem 1 of Deelstra and Delbaen (1995). \qed

The next result presents the bounds for the moments of the related stochastic integral.
Lemma 7. Denote

\[ Z_t = \int_0^t e^{bu} \sqrt{r_u} \, dW_u, \quad t \geq 0. \] (25)

There exists a constant \( C > 0 \) such that for all \( t \geq 0 \),

\[ \mathbb{E}[Z_t^2] \leq Ce^{2bt}, \quad \mathbb{E}[Z_t^3] \leq Ce^{3bt}, \quad \mathbb{E}[Z_t^2 r_t] \leq Ce^{2bt}. \] (26)

The constant \( C \) may depend on \( a, b, \sigma \) and \( r_0 \).

Proof. 1. According to (Deelstra and Delbaen 1995, Eq. (1)), the process \( r \) satisfies the following relations

\[ r_t = e^{-bt} \left( r_0 + a \int_0^t e^{bu} \, du + \sigma \int_0^t e^{bu} \sqrt{r_u} \, dW_u \right) \]

\[ = e^{-bt} \left( r_0 - \frac{a}{b} + \frac{a}{b} e^{bt} + \sigma \int_0^t e^{bu} \sqrt{r_u} \, dW_u \right). \] (27)

Therefore, in particular, its expectation equals

\[ \mathbb{E}r_t = \left( r_0 - \frac{a}{b} \right) e^{-bt} + \frac{a}{b}. \]

Then for the 2nd moment we have the following representation:

\[ \mathbb{E}Z_t^2 = \mathbb{E} \left( \int_0^t e^{bu} \sqrt{r_u} \, dW_u \right)^2 = \int_0^t \left( \left( r_0 - \frac{a}{b} \right) e^{bu} + \frac{a}{b} e^{2bu} \right) \, du \]

\[ = \left( r_0 - \frac{a}{b} \right) \frac{e^{bt} - 1}{b} + \frac{a}{b} \cdot \frac{e^{2bt} - 1}{2b} = e^{2bt} \left( \frac{1}{2b^2} + \frac{1}{b^2} \left( r_0 - \frac{a}{b} \right) e^{-bt} - \frac{1}{b} \left( r_0 - \frac{a}{2b} \right) e^{-2bt} \right). \]

Consequently, \( \mathbb{E}Z_t^2 \leq Ce^{2bt} \) with \( C = \frac{a^2}{2b^2} + \frac{1}{b} |r_0 - \frac{a}{b}| + \frac{1}{b} |r_0 - \frac{a}{2b}|. \)

2. Let us consider \( \mathbb{E}Z_t^3 \). By Itô’s formula, from Equation (25) we have

\[ Z_t^3 = 3 \int_0^t Z_s^2 r_s \, ds + 3 \int_0^t Z_s^2 e^{bs} \sqrt{r_s} \, dW_s. \] (28)

Note that according to Karatzas and Shreve (1991, Problem 3.15, p. 306) (see also Mishura and Shevchenko 2017, Thm. 9.3), for any \( p \geq 1 \), \( \mathbb{E}\left[ \sup_{t \in [0, T]} |r_t|^{2p} \right] < \infty \). It follows from Equations (27) and (25) that

\[ r_s = e^{-bs} \left( r_0 - \frac{a}{b} + \frac{a}{b} e^{bs} + \sigma Z_s \right) \] (29)

Therefore, we have also that \( \mathbb{E}\left[ \sup_{t \in [0, T]} |Z_t|^{2p} \right] < \infty \) for any \( p \geq 1 \). Consequently, all the terms in the left-hand side and in the right-hand side of Equation (28) have bounded expectations. Hence,

\[ \mathbb{E}[Z_t^3] = 3\mathbb{E} \int_0^t Z_s e^{2bs} r_s \, ds. \] (30)

We insert Equation (29) into (30) and obtain
$$E[Z^3_t] = 3E\int_0^t Z e^{bs} \left( r_0 - \frac{a}{b} + \frac{a}{b} e^{bs} + \sigma Z_t \right) \, ds = 3\sigma \int_0^t e^{bs} E[Z^2_t] \, ds,$$

since $E[Z_t] = 0$. Applying the bound $E[Z^2_t] \leq C e^{2bs}$ from Equation (26), we get

$$E[Z^3_t] \leq 3\sigma C \int_0^t e^{3bs} \, ds = \frac{C\sigma}{b} \left( e^{3bt} - 1 \right) \leq \frac{C\sigma}{b} e^{3bt}.$$

3. We express $r_t$ through $Z_t$ by Equation (29), and get

$$E[Z^2_t r_t] = \left( \left( r_0 - \frac{a}{b} \right) e^{-bt} + \frac{a}{b} \right) E[Z^2_t] + \sigma e^{-bt} E[Z^3_t] \leq C e^{2bt},$$

where the inequality follows from the first two bounds in Equation (26).

**Lemma 8.** Let the process $Z$ be defined by Equation (25). Then the following normalized stochastic integrals vanish as $T \to \infty$:

$$\frac{1}{T} \int_0^T e^{-bt} Z_t \sqrt{r_t} \, dW_t \to 0 \quad \text{a.s.,} \quad (31)$$

$$\frac{1}{T} e^{-2bt} \int_0^T e^{bt} Z_t \sqrt{r_t} \, dW_t \to 0 \quad \text{a.s.} \quad (32)$$

**Proof.** 1. Obviously, the convergence (31) is equivalent to

$$\frac{1}{T+1} \int_0^T e^{-bt} Z_t \sqrt{r_t} \, dW_t \to 0 \quad \text{a.s., as } T \to \infty.$$

By Kronecker’s lemma (see, e.g., Deelstra and Delbaen 1995), in order to prove this convergence it suffices to show that

$$\int_0^\infty \frac{e^{-bt} Z_t \sqrt{r_t}}{t+1} \, dW_t$$

(33)

is well defined. Since the process $M_T = \int_0^T \frac{e^{-bt} Z_t \sqrt{r_t}}{t+1} \, dW_t$ is a martingale, it suffices to prove that

$$E(M)_\infty = E\left[ \int_0^\infty \frac{e^{-2bt} Z_t^2 r_t}{(t+1)^2} \, dt \right] < \infty. \quad (34)$$

But Equation (34) follows immediately from the third bound of Equation (26):

$$E\left[ \int_0^\infty \frac{e^{-2bt} Z_t^2 r_t}{(t+1)^2} \, dt \right] \leq C \int_0^\infty \frac{1}{(t+1)^2} \, dt = C.$$

Thus, Equation (31) is proved.

2. Similarly, by Kronecker’s lemma, the Equation (32) follows from the existence a. s. of the integral

$$\int_0^\infty \frac{e^{bt} Z_t \sqrt{r_t}}{t+1} \, dW_t = \int_0^\infty \frac{e^{-bt} Z_t \sqrt{r_t}}{t+1} \, dW_t.$$

This integral is well defined, because it coincides with the integral Equation (33) considered in the first part of the proof.