Modern Stochastics: Theory and Applications 6 (3) (2019) 377–395
https://doi.org/10.15559/19-VMSTA140

Maximum likelihood estimation in the non-ergodic fractional Vasicek model

Stanislav Lohvinenko, Kostiantyn Ralchenko*

Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, 64, Volodymyrs'ka St., 01601 Kyiv, Ukraine

stanislav.lohvinenko@gmail.com (S. Lohvinenko), k.ralchenko@gmail.com (K. Ralchenko)

Received: 27 May 2019, Revised: 7 August 2019, Accepted: 5 September 2019,
Published online: 23 September 2019

Abstract We investigate the fractional Vasicek model described by the stochastic differential equation $dX_t = (\alpha - \beta X_t)\,dt + \gamma\,dB_t^H$, $X_0 = x_0$, driven by the fractional Brownian motion B^H with the known Hurst parameter $H \in (1/2, 1)$. We study the maximum likelihood estimators for unknown parameters α and β in the non-ergodic case (when $\beta < 0$) for arbitrary $x_0 \in \mathbb{R}$, generalizing the result of Tanaka, Xiao and Yu (2019) for particular $x_0 = \alpha/\beta$, derive their asymptotic distributions and prove their asymptotic independence.

Keywords Fractional Brownian motion, fractional Vasicek model, maximum likelihood estimation, moment generating function, asymptotic distribution, non-ergodic process

2010 MSC 60G22, 62F10, 62F12

1 Introduction

The present paper deals with the fractional Vasicek model of the form

$$dX_t = (\alpha - \beta X_t)dt + \gamma dB_t^H, \quad X_0 = x_0 \in \mathbb{R}, \tag{1}$$

where B^H is the fractional Brownian motion with the Hurst index $H \in (1/2, 1)$. It is a generalization of the classical interest rate model proposed by O. Vasicek [34]

*Corresponding author.

in 1977. This generalization enables to study processes with long-range dependence, which arise in financial mathematics and several other areas such as telecommunication networks, investigation of turbulence and image processing. In recent years, many articles on various financial applications of the fractional Vasicek model (1) have appeared, see e.g. [8, 9, 12, 13, 30, 40]). In order to use this model in practice, a theory of parameter estimation is necessary.

Notice that in the particular case $\alpha = 0$, (1) is a so-called fractional Ornstein–Uhlenbeck process, introduced in [7]. The drift parameter estimation for it has been studied since 2002, see the paper [17], where the maximum likelihood estimation was considered. The asymptotic and exact distributions of the maximum likelihood estimator (MLE) were investigated later in [4, 31, 32]. Alternative approaches to the drift parameter estimation were proposed and studied in [3, 6, 14–16, 21]. We refer to the article [28] for a survey on this topic, and to the book [20] for its more detailed presentation.

In the general case, the least squares and ergodic-type estimators of unknown parameters α and β were studied in [27, 38, 39]. The corresponding MLEs of α and β were presented in [25]. Their consistency and asymptotic normality were proved there for the case $\beta > 0$. Slightly more general results were proved in [26], where joint asymptotic normality of MLE of the vector parameter (α, β) was established. Recently Tanaka et al. [33] investigated asymptotic behavior of MLEs in the cases $\beta = 0$ and $\beta < 0$. However, in the latter case the asymptotic distribution was obtained only under assumption that $x_0 = \frac{\alpha}{\beta}$. The study of the case $x_0 \neq \frac{\alpha}{\beta}$ requires a different technique and still remains an open problem. The goal of the present paper is to fill in the gap and to derive asymptotic distributions of the MLEs of α and β for arbitrary $x_0 \in \mathbb{R}$, $\alpha \in \mathbb{R}$ and $\beta < 0$. Moreover, we prove that the MLEs for α and β are asymptotically independent.

The asymptotic behavior of the process X and of the estimators substantially depends on the sign of the parameter β. If $\beta < 0$, then the process X behaves as $O_{\mathsf{P}}(e^{-\beta T})$ as $T \to \infty$, hence it is non-ergodic. If $\beta > 0$, then $X_T = O_{\mathsf{P}}(1)$, as $T \to \infty$, and the process has ergodic properties, see, e.g., [27]. The method for the hypothesis testing of the sign of β was developed in [22].

In this article we restrict ourselves to the case $\frac{1}{2} < H < 1$. Our proofs are based on the results of the papers [17] and [26], which are valid only for $H \in (\frac{1}{2}, 1)$ and cannot be immediately extended to the case $H \in (0, \frac{1}{2})$. In particular, the integral representation (7) below, which is the starting point for derivation of moment generating functions (MGFs) in Lemmas 1 and 2, holds for $H \in (0, \frac{1}{2})$ with different (and more complicated) kernel K_H. Therefore, the asymptotic behavior of the MLEs in this case requires a separate study.

The paper is organized as follows. In Section 2 we describe the model and the estimators, and introduce the notation. Section 3 contains the results on distributions and asymptotic behavior of stochastic processes involved into MLEs. In Section 4 we formulate and prove the main results on asymptotic distributions of MLEs. Some auxiliary facts and results concerning modified Bessel functions of the first kind and MGFs related to the normal distribution are collected in the appendices.

2 Preliminaries

Let $(\Omega, \mathfrak{F}, \{\mathfrak{F}_t\}, \mathsf{P})$ be a complete probability space with filtration. Let $B^H = \{B_t^H, t \ge 0\}$ be a fractional Brownian motion on this probability space, that is, a centered Gaussian process with the covariance function

$$\mathbb{E}B_t^H B_s^H = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right). \tag{2}$$

It follows from (2) that $\mathbb{E}(B_t^H - B_s^H)^2 = |t-s|^{2H}$. Hence, there exists a modification of B^H, which is δ-Hölder continuous for all $\delta \in (0, H)$.

We study the fractional Vasicek model, described by the stochastic differential equation

$$X_t = x_0 + \int_0^t (\alpha - \beta X_s)ds + \gamma B_t^H, \quad t \ge 0. \tag{3}$$

The main goal is to estimate parameters $\alpha \in \mathbb{R}$ and $\beta < 0$ by continuous observations of a trajectory of X on the interval $[0, T]$. We assume that the parameters $\gamma > 0$ and $H \in (1/2, 1)$ are known. This assumption is natural, because γ and H can be obtained explicitly from the observations by considering realized power variations, see Remark 1 below.

Equation (3) has a unique solution, which is given by

$$X_t = x_0 e^{-\beta t} + \frac{\alpha}{\beta}\left(1 - e^{-\beta t}\right) + \gamma \int_0^t e^{-\beta(t-s)} dB_s^H, \quad t \ge 0, \tag{4}$$

where $\int_0^t e^{-\beta(t-s)} dB_s^H$ is a path-wise Riemann–Stieltjes integral. It exists due to [7, Proposition A.1].

Following [18], for $0 < s < t \le T$ we define

$$\kappa_H = 2H\Gamma(3/2 - H)\Gamma(H + 1/2), \qquad \lambda_H = \frac{2H\Gamma(3-2H)\Gamma(H+1/2)}{\Gamma(3/2 - H)},$$
$$k_H(t,s) = \kappa_H^{-1} s^{1/2-H}(t-s)^{1/2-H}, \qquad w_t^H = \lambda_H^{-1} t^{2-2H}.$$

We introduce also three stochastic processes

$$P_H(t) = \frac{1}{\gamma}\frac{d}{dw_t^H}\int_0^t k_H(t,s) X_s\,ds,$$
$$Q_H(t) = \frac{1}{\gamma}\frac{d}{dw_t^H}\int_0^t k_H(t,s)(\alpha - \beta X_s)\,ds,$$
$$S_t = \frac{1}{\gamma}\int_0^t k_H(t,s)\,dX_s.$$

Note that by [25, Lemma 4.1]

$$Q_H(t) = \frac{\alpha}{\gamma} - \beta P_H(t). \tag{5}$$

According to [18, Theorem 1], the process S is an $(\mathfrak{F}_t)$-semimartingale with the decomposition

$$S_t = \int_0^t Q_H(s)\, dw_s^H + M_t^H, \tag{6}$$

where $M_t^H = \int_0^t k_H(t,s)\, dB_s^H$ is a Gaussian martingale with the variance function $\langle M^H \rangle = w^H$. The natural filtrations of processes S and X coincide. Moreover, the process X admits the following representation

$$X_t = \int_0^t K_H(t,s)\, dS_s, \tag{7}$$

where $K_H(t,s) = \gamma H(2H-1) \int_s^t r^{H-1/2}(r-s)^{H-3/2}\, dr$.

Remark 1. If we observe the whole path $\{X_t, t \in [0,T]\}$, then the parameters γ and H can be obtained from observations explicitly in the following way. Let $\{t_i^{(n)}\}$ be a partition of $[0,T]$, such that $\sup_i |t_{i+1}^{(n)} - t_i^{(n)}| \to 0$, as $n \to \infty$. Denote $Z_T = \int_0^T k_H(T,s)\, dX_s = \gamma S_T$. From (6) it follows that $\langle Z \rangle_T = \gamma^2 w_T^H$ a.s. Hence, the parameter γ is calculated as the limit

$$\gamma^2 = \left(w_T^H\right)^{-1} \lim_n \sum_i \left(Z_{t_{i+1}^{(n)}} - Z_{t_i^{(n)}} \right)^2 \quad \text{a.s.}$$

The Hurst index H can be evaluated as follows:

$$H = \frac{1}{2} - \frac{1}{2} \lim_n \log_2 \left(\frac{\sum_{i=1}^{2n-1} \left(X_{t_{i+1}^{(2n)}} - 2X_{t_i^{(2n)}} + X_{t_{i-1}^{(2n)}} \right)^2}{\sum_{i=1}^{n-1} \left(X_{t_{i+1}^{(n)}} - 2X_{t_i^{(n)}} + X_{t_{i-1}^{(n)}} \right)^2} \right) \quad \text{a.s.,}$$

see, e.g., [20, Sec. 3.1]. There exist several other methods of the Hurst index evaluation based on power variations of X. We refer to the books [5, 20] for further information on this subject.

Applying the analog of the Girsanov formula for a fractional Brownian motion ([18, Theorem 3], see also [19]) and (6), one can obtain the likelihood ratio $\frac{d\mathsf{P}_{\alpha,\beta}(T)}{d\mathsf{P}_{0,0}(T)}$ for the probability measure $\mathsf{P}_{\alpha,\beta}(T)$ corresponding to our model and the probability measure $\mathsf{P}_{0,0}(T)$ corresponding to the model with zero drift [25]:

$$\begin{aligned} \frac{d\mathsf{P}_{\alpha,\beta}(T)}{d\mathsf{P}_{0,0}(T)} &= \exp\left\{ \int_0^T Q_H(t)\, dS_t - \frac{1}{2} \int_0^T \left(Q_H(t)\right)^2 dw_t^H \right\} \\ &= \exp\left\{ \frac{\alpha}{\gamma} S_T - \beta \int_0^T P_H(t)\, dS_t - \frac{\alpha^2}{2\gamma^2} w_T^H \right. \\ &\quad \left. + \frac{\alpha\beta}{\gamma} \int_0^T P_H(t)\, dw_t^H - \frac{\beta^2}{2} \int_0^T \left(P_H(t)\right)^2 dw_t^H \right\}. \end{aligned} \tag{8}$$

MLEs of parameters α and β maximize (8) and have the following form [25]:

$$\widehat{\alpha}_T = \frac{S_T K_T - I_T J_T}{w_T^H K_T - J_T^2} \gamma, \qquad \widehat{\beta}_T = \frac{S_T J_T - w_T^H I_T}{w_T^H K_T - J_T^2}, \tag{9}$$

where

$$I_T = \int_0^T P_H(t)\,dS_t, \qquad J_T = \int_0^T P_H(t)\,dw_t^H, \qquad K_T = \int_0^T \bigl(P_H(t)\bigr)^2\,dw_t^H.$$

It is worth noting that using the definition of $P_H(t)$ one can easily represent J_T in the following way

$$J_T = \frac{1}{\gamma}\int_0^T k_H(T,s)X_s\,ds.$$

3 Auxiliary results

In this section we find exact and asymptotic distributions of the statistics S_T, I_T, J_T, K_T and related random variables and vectors.

We start with the bivariate MGF of the vector (S_T, I_T). For the case $\beta > 0$, it was derived in [26, Lemma 3.3]. However, the same proof is valid for the case $\beta < 0$. The following result is a reformulation of [26, Lemma 3.3], obtained by applying the formula (44) from Appendix A.

Lemma 1. *The moment generating function of* (S_T, I_T) *equals*

$$\begin{aligned} m_1^{(\alpha,\beta)}(\xi_1,\xi_2) &= \mathbb{E}\bigl[\exp\{\xi_1 S_T + \xi_2 I_T\}\bigr] \\ &= D^{(\alpha,\beta)}(\xi_2)^{-\frac{1}{2}} \exp\Biggl\{\frac{1}{8D^{(\alpha,\beta)}(\xi_2)}\sum_{i=1}^4 A_i^{(\alpha,\beta)}(\xi_1,\xi_2) - \frac{\xi_2 T}{2}\Biggr\}, \end{aligned}$$

where

$$\begin{aligned} D^{(\alpha,\beta)}(\xi_2) = {} & \left(1-\frac{\xi_2}{2\beta}\right)^2 + \frac{\xi_2^2}{4\beta^2}e^{-2\beta T} + \left(\frac{\xi_2}{\beta} - \frac{\xi_2^2}{2\beta^2}\right)\frac{(-\beta)\pi T}{4\sin\pi H}e^{-\beta T} \\ & \times \left[I_{-H}\left(-\frac{\beta T}{2}\right)I_{H-1}\left(-\frac{\beta T}{2}\right) + I_{1-H}\left(-\frac{\beta T}{2}\right)I_H\left(-\frac{\beta T}{2}\right)\right], \end{aligned} \tag{10}$$

$$A_1^{(\alpha,\beta)}(\xi_1,\xi_2) = \xi_2\bigl(c_1\bigl(\tfrac{\alpha}{\beta}\bigr)\xi_1 - c_2\bigl(\tfrac{\alpha}{\beta}\bigr)\xi_2\bigr)(-\beta)^{H-1}T^{1-H}e^{-\frac{3\beta T}{2}}I_{1-H}\bigl(-\tfrac{\beta T}{2}\bigr), \tag{11}$$

$$\begin{aligned} A_2^{(\alpha,\beta)}(\xi_1,\xi_2) = {} & \bigl(\xi_1^2 c_3 - \xi_1\xi_2 c_4\bigl(\tfrac{\alpha}{\beta}\bigr) + \xi_2^2 c_5\bigl(\tfrac{\alpha}{\beta}\bigr)\bigr)T^{2-2H}e^{-\beta T} \\ & \times I_{1-H}\bigl(-\tfrac{\beta T}{2}\bigr)I_{H-1}\bigl(-\tfrac{\beta T}{2}\bigr), \end{aligned} \tag{12}$$

$$A_3^{(\alpha,\beta)}(\xi_1,\xi_2) = \xi_2(\xi_2 - 2\beta)c_6\bigl(\tfrac{\alpha}{\beta}\bigr)(-\beta)^{2H-1}Te^{-\beta T}I_{1-H}\bigl(-\tfrac{\beta T}{2}\bigr)I_{-H}\bigl(-\tfrac{\beta T}{2}\bigr), \tag{13}$$

$$\begin{aligned} A_4^{(\alpha,\beta)}(\xi_1,\xi_2) = {} & \bigl(c_1\bigl(\tfrac{\alpha}{\beta}\bigr)\xi_1 - c_2\bigl(\tfrac{\alpha}{\beta}\bigr)\xi_2\bigr)(\xi_2 - 2\beta)(-\beta)^{H-1} \\ & \times T^{1-H}e^{-\frac{\beta T}{2}}I_{1-H}\bigl(-\tfrac{\beta T}{2}\bigr), \end{aligned} \tag{14}$$

$$c_1\bigl(\tfrac{\alpha}{\beta}\bigr) = \bigl(x_0 - \tfrac{\alpha}{\beta}\bigr)4\rho_H, \qquad c_4\bigl(\tfrac{\alpha}{\beta}\bigr) = \bigl(x_0 - \tfrac{\alpha}{\beta}\bigr)\rho_H 2^{2H+1}\Gamma(H),$$

$$c_2\bigl(\tfrac{\alpha}{\beta}\bigr) = \bigl(x_0 - \tfrac{\alpha}{\beta}\bigr)^2\frac{\lambda_H^* 2^{2H+1}\rho_H^2}{\Gamma(1-H)}, \qquad c_5\bigl(\tfrac{\alpha}{\beta}\bigr) = \bigl(x_0 - \tfrac{\alpha}{\beta}\bigr)^2\frac{\lambda_H^* 2^{4H-1}\rho_H^2\Gamma(H)}{\Gamma(1-H)},$$

$$c_3 = \frac{2\Gamma(H)\Gamma(1-H)}{\lambda_H^*}, \qquad c_6\left(\tfrac{\alpha}{\beta}\right) = \left(x_0 - \frac{\alpha}{\beta}\right)^2 2\lambda_H^* \rho_H^2,$$

$$\lambda_H^* = \frac{\lambda_H}{2-2H}, \qquad \rho_H = \frac{\sqrt{\pi}\,\Gamma(3/2-H)}{\gamma \kappa_H}.$$

The domain of the function $m_1^{(\alpha,\beta)}$ *equals* $\{(\xi_1,\xi_2) \in \mathbb{R}^2 : D^{(\alpha,\beta)}(\xi_2) > 0\}$.

The following lemma gives a joint MGF of (S_T, I_T, J_T, K_T).

Lemma 2. *The moment generating function of* (S_T, I_T, J_T, K_T) *equals*

$$\begin{aligned} m_2(\theta_1,\theta_2,\theta_3,\theta_4) &= \mathbb{E}\big[\exp\{\theta_1 S_T + \theta_2 I_T + \theta_3 J_T + \theta_4 K_T\}\big] \\ &= m_1^{(\alpha_1,\beta_1)}\left(\theta_1 + \frac{\alpha-\alpha_1}{\gamma}, \theta_2 - \beta + \beta_1\right) \exp\left\{\frac{\alpha_1^2 - \alpha^2}{2\gamma^2} w_T^H\right\}, \end{aligned}$$

where

$$\alpha_1 = -\frac{\gamma\theta_3 + \alpha\beta}{\sqrt{\beta^2 - 2\theta_4}}, \qquad \beta_1 = -\sqrt{\beta^2 - 2\theta_4}.$$

The domain of the function m_2 *equals*

$$\left\{(\theta_1,\theta_2,\theta_3,\theta_4) \in \mathbb{R}^4 : \theta_4 < \beta^2/2,\ D^{(\alpha,\beta)}\left(\theta_2 - \beta - \sqrt{\beta^2 - 2\theta_4}\right) > 0\right\},$$

where $D^{(\alpha,\beta)}$ *is defined in* (10).

Proof. The proof is the same as for [26, Theorem 3.4]. □

Lemma 3. *Under stated conditions the process* S_T *has the normal asymptotic distribution as* $T \to \infty$, *namely*

$$T^{H-1/2} e^{\beta T} S_T \xrightarrow{d} \mathcal{N}\left(\frac{(x_0 - \frac{\alpha}{\beta})\rho_H(-\beta)^{H-1/2}}{\sqrt{\pi}}, \frac{\Gamma(H)\Gamma(1-H)}{2\pi(-\beta)\lambda_H^*}\right). \tag{15}$$

Proof. We obtain the distribution via MGF. Using Lemma 1 we have

$$\mathbb{E}\big[\exp\{\theta T^{H-1/2} e^{\beta T} S_T\}\big] = m_1\big(\theta T^{H-1/2} e^{\beta T}, 0\big).$$

Taking each term of the function m_1 separately with $\xi_1 = \theta T^{H-1/2} e^{\beta T}$ and $\xi_2 = 0$ and applying (45) we obtain that $D(\xi_2) = 1$, $A_1(\xi_1,\xi_2) = A_3(\xi_1,\xi_2) = 0$,

$$\begin{aligned} A_2(\xi_1,\xi_2) &= \xi_1^2 c_3 T^{2-2H} e^{-\beta T} I_{1-H}\left(-\frac{\beta T}{2}\right) I_{H-1}\left(-\frac{\beta T}{2}\right) \\ &= \theta^2 c_3 T e^{\beta T} I_{1-H}\left(-\frac{\beta T}{2}\right) I_{H-1}\left(-\frac{\beta T}{2}\right) \\ &\to \frac{c_3}{\pi(-\beta)}\theta^2, \quad \text{as } T \to \infty, \end{aligned}$$

and

$$\begin{aligned}A_4(\xi_1,\xi_2) &= \xi_1 2c_1\big(\tfrac{\alpha}{\beta}\big)(-\beta)^H T^{1-H} e^{-\frac{\beta T}{2}} I_{1-H}\left(-\frac{\beta T}{2}\right)\\ &= \theta 2c_1\big(\tfrac{\alpha}{\beta}\big)(-\beta)^H T^{1/2} e^{\frac{\beta T}{2}} I_{1-H}\left(-\frac{\beta T}{2}\right)\\ &\to \frac{2c_1(\frac{\alpha}{\beta})(-\beta)^{H-1/2}}{\sqrt{\pi}}\theta, \quad \text{as } T\to\infty.\end{aligned}$$

Hence

$$\mathbb{E}\big[\exp\{\theta T^{H-1/2} e^{\beta T} S_T\}\big] \to \exp\left\{\frac{c_3}{8\pi(-\beta)}\theta^2 + \frac{c_1(\frac{\alpha}{\beta})(-\beta)^{H-1/2}}{4\sqrt{\pi}}\theta\right\},$$

as $T\to\infty$. This means that

$$T^{H-1/2} e^{\beta T} S_T \xrightarrow{d} \mathcal{N}\left(\frac{c_1(\frac{\alpha}{\beta})(-\beta)^{H-1/2}}{4\sqrt{\pi}}, \frac{c_3}{4\pi(-\beta)}\right), \quad \text{as } T\to\infty,$$

which is equivalent to (15). □

The following result is crucial for the derivation of the joint asymptotic distribution of MLE.

Lemma 4. *The vector of main components of the MLE has the following behavior*

$$\begin{pmatrix} T^{H-1}(S_T + \beta J_T - \frac{\alpha}{\gamma} w_T^H) \\ e^{\beta T}(I_T + \beta K_T) \\ e^{2\beta T} K_T \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \xi \\ \eta\zeta \\ \zeta^2 \end{pmatrix}, \quad \text{as } T\to\infty, \tag{16}$$

where ξ, η, ζ are independent and $\xi \overset{d}{=} \mathcal{N}(0,\lambda_H^{-1})$, $\eta \overset{d}{=} \mathcal{N}(0,1)$,

$$\zeta \overset{d}{=} \mathcal{N}\left(\frac{(x_0 - \frac{\alpha}{\beta})\rho_H \sqrt{\lambda_H^*}(-\beta)^{H-1}}{\sqrt{2\pi}}, \frac{1}{4\beta^2 \sin \pi H}\right). \tag{17}$$

Proof. We again obtain the stated asymptotic distribution via MGF of the presented vector. It could be easily reduced to already studied MGF. That said, using Lemma 2, we have

$$\begin{aligned}&\mathbb{E}\left[\exp\left\{\theta_1 T^{H-1}\left(S_T + \beta J_T - \frac{\alpha}{\gamma} w_T^H\right) + \theta_2 e^{\beta T}(I_T + \beta K_T) + \theta_3 e^{2\beta T} K_T\right\}\right]\\ &\quad = m_2\big(\theta_1 T^{H-1}, \theta_2 e^{\beta T}, \theta_1 \beta T^{H-1}, \theta_2 \beta e^{\beta T} + \theta_3 e^{2\beta T}\big) \exp\left\{-\frac{\theta_1\alpha}{\gamma} T^{H-1} w_T^H\right\}\\ &\quad = m_1^{(\alpha_1(T),\beta_1(T))}\left(\theta_1 T^{H-1} + \frac{\alpha-\alpha_1(T)}{\gamma}, \theta_2 e^{\beta T} - \beta + \beta_1(T)\right)\\ &\qquad \times \exp\left\{\frac{\alpha_1(T)^2 - \alpha^2}{2\gamma^2} w_T^H - \frac{\theta_1\alpha}{\gamma} T^{H-1} w_T^H\right\},\end{aligned} \tag{18}$$

where

$$\alpha_1(T) = \frac{\beta\gamma\theta_1 T^{H-1} + \alpha\beta}{-\sqrt{\beta^2 - 2(\theta_2\beta e^{\beta T} + \theta_3 e^{2\beta T})}},$$
$$\beta_1(T) = -\sqrt{\beta^2 - 2\big(\theta_2\beta e^{\beta T} + \theta_3 e^{2\beta T}\big)}.$$

Applying the Taylor series expansion, we get as $T \to \infty$

$$\begin{aligned}\alpha_1(T) &= \big(\gamma\theta_1 T^{H-1} + \alpha\big)\left[1 - 2\left(\frac{\theta_2}{\beta}e^{\beta T} + \frac{\theta_3}{\beta^2}e^{2\beta T}\right)\right]^{-1/2} \\ &= \big(\gamma\theta_1 T^{H-1} + \alpha\big)\left[1 + \left(\frac{\theta_2}{\beta}e^{\beta T} + \frac{\theta_3}{\beta^2}e^{2\beta T}\right) + O\big(e^{2\beta T}\big)\right] \\ &= \alpha + \gamma\theta_1 T^{H-1} + \frac{\theta_2\alpha}{\beta}e^{\beta T} + \frac{\theta_2\gamma\theta_1}{\beta}T^{H-1}e^{\beta T} + O\big(e^{2\beta T}\big) \qquad (19)\end{aligned}$$

and

$$\begin{aligned}\beta_1(T) &= \beta\left[1 - 2\left(\theta_2\frac{1}{\beta}e^{\beta T} + \theta_3\frac{1}{\beta^2}e^{2\beta T}\right)\right]^{1/2} \\ &= \beta\Bigg[1 - \left(\theta_2\frac{1}{\beta}e^{\beta T} + \theta_3\frac{1}{\beta^2}e^{2\beta T}\right) \\ &\qquad - \frac{1}{2}\left(\theta_2\frac{1}{\beta}e^{\beta T} + \theta_3\frac{1}{\beta^2}e^{2\beta T}\right)^2 + O\big(e^{3\beta T}\big)\Bigg] \\ &= \beta - \theta_2 e^{\beta T} + \frac{\theta_2^2 + 2\theta_3}{2(-\beta)}e^{2\beta T} + O\big(e^{3\beta T}\big). \qquad (20)\end{aligned}$$

Note that $\alpha_1(T) \to \alpha$ and $\beta_1(T) \to \beta$, as $T \to \infty$. Moreover, the arguments of the function $m_1^{(\alpha_1(T),\beta_1(T))}$ in (18) have the following asymptotic behavior:

$$\xi_1(T) := \theta_1 T^{H-1} + \frac{\alpha - \alpha_1(T)}{\gamma} = \frac{\theta_2\alpha}{-\beta\gamma}e^{\beta T} + \frac{\theta_1\theta_2}{-\beta}T^{H-1}e^{\beta T} + O\big(e^{2\beta T}\big), \quad (21)$$
$$\xi_2(T) := \theta_2 e^{\beta T} - \beta + \beta_1(T) = \frac{\theta_2^2 + 2\theta_3}{2(-\beta)}e^{2\beta T} + O\big(e^{3\beta T}\big), \quad (22)$$

as $T \to \infty$. Further, inserting (22) into (10), and applying the expansion (45) from Appendix A, we obtain

$$\begin{aligned}&D^{(\alpha_1(T),\beta_1(T))}\big(\xi_2(T)\big) \\ &\quad= \left(1 - \frac{(\theta_2^2 + 2\theta_3)e^{2\beta T}}{4\beta_1(T)(-\beta)}\right)^2 + \frac{(\theta_2^2 + 2\theta_3)^2 e^{4\beta T}}{16\beta_1(T)^2\beta^2}e^{-2\beta_1(T)T} \\ &\qquad+ \left(\frac{(\theta_2^2 + 2\theta_3)e^{2\beta T}}{2\beta_1(T)(-\beta)} - \frac{(\theta_2^2 + 2\theta_3)^2 e^{4\beta T}}{8\beta_1(T)^2\beta^2}\right)\frac{(-\beta_1(T))\pi T}{4\sin\pi H}e^{-\beta_1(T)T} \\ &\qquad\times\left[I_{-H}\left(-\frac{\beta_1(T)T}{2}\right)I_{H-1}\left(-\frac{\beta_1(T)T}{2}\right)\right.\end{aligned}$$

$$
\begin{aligned}
&\quad + I_{1-H}\Big(-\frac{\beta_1(T)T}{2}\Big)I_H\Big(-\frac{\beta_1(T)T}{2}\Big)\Big] + O\big(e^{\beta T}\big)\\
&\sim \bigg(1 - \frac{(\theta_2^2+2\theta_3)e^{2\beta T}}{4\beta_1(T)(-\beta)}\bigg)^2 + \frac{(\theta_2^2+2\theta_3)^2 e^{4\beta T}}{16\beta_1(T)^2\beta^2}e^{-2\beta_1(T)T}\\
&\quad + \bigg(\frac{(\theta_2^2+2\theta_3)e^{2\beta T}}{2\beta_1(T)(-\beta)} - \frac{(\theta_2^2+2\theta_3)^2 e^{4\beta T}}{8\beta_1(T)^2\beta^2}\bigg)\frac{1}{2\sin\pi H}e^{-2\beta_1(T)T} + O\big(e^{\beta T}\big)\\
&\to 1 - \frac{\theta_2^2+2\theta_3}{4\beta^2\sin\pi H}, \quad \text{as } T\to\infty. \qquad (23)
\end{aligned}
$$

It follows from (21), (22) that

$$
c_1\big(\tfrac{\alpha_1(T)}{\beta_1(T)}\big)\xi_1(T) - c_2\big(\tfrac{\alpha_1(T)}{\beta_1(T)}\big)\xi_2(T) \sim c_1\big(\tfrac{\alpha}{\beta}\big)\frac{\theta_2\alpha}{-\beta\gamma}e^{\beta T}, \quad \text{as } T\to\infty. \qquad (24)
$$

Using this relation, (22) and (45), we get from (11) that

$$
\begin{aligned}
&A_1^{(\alpha_1(T),\beta_1(T))}\big(\xi_1(T),\xi_2(T)\big)\\
&\quad = \xi_2(T)\big(c_1\big(\tfrac{\alpha_1(T)}{\beta_1(T)}\big)\xi_1(T) - c_2\big(\tfrac{\alpha_1(T)}{\beta_1(T)}\big)\xi_2(T)\big)\\
&\qquad \times \big(-\beta_1(T)\big)^{H-1}T^{1-H}e^{-\frac{3\beta_1(T)T}{2}}I_{1-H}\Big(-\frac{\beta_1(T)T}{2}\Big)\\
&\quad \sim \frac{\theta_2^2+2\theta_3}{2(-\beta)}e^{2\beta T}c_1\big(\tfrac{\alpha}{\beta}\big)\frac{\alpha}{(-\beta)\gamma}\theta_2 e^{\beta T}\big(-\beta_1(T)\big)^{H-1}T^{1-H}\\
&\qquad \times e^{-\frac{3\beta_1(T)T}{2}}I_{1-H}\Big(-\frac{\beta_1(T)T}{2}\Big)\\
&\quad \sim \frac{\alpha\theta_2(\theta_2^2+2\theta_3)}{2\beta^2\gamma}e^{3\beta T}c_1\big(\tfrac{\alpha}{\beta}\big)\big(-\beta_1(T)\big)^{H-3/2}\frac{1}{\sqrt{\pi}}T^{1/2-H}e^{-2\beta_1(T)T}\\
&\quad = O\big(T^{1/2-H}e^{\beta T}\big) \to 0, \quad \text{as } T\to\infty. \qquad (25)
\end{aligned}
$$

It follows from (21), (22) that

$$
\begin{aligned}
\xi_1(T)^2 &= \frac{\theta_2^2\alpha^2}{\beta^2\gamma^2}e^{2\beta T} + O\big(T^{H-1}e^{2\beta T}\big),\\
\xi_1(T)\xi_2(T) &= O\big(e^{3\beta T}\big), \qquad \xi_2(T)^2 = O\big(e^{4\beta T}\big),
\end{aligned}
$$

as $T\to\infty$. Therefore, by (12) and (45) we obtain

$$
\begin{aligned}
&A_2^{(\alpha_1(T),\beta_1(T))}\big(\xi_1(T),\xi_2(T)\big)\\
&\quad = \big(\xi_1(T)^2 c_3 - \xi_1(T)\xi_2(T)c_4\big(\tfrac{\alpha_1(T)}{\beta_1(T)}\big) + \xi_2(T)^2 c_5\big(\tfrac{\alpha_1(T)}{\beta_1(T)}\big)\big)\\
&\qquad \times T^{2-2H}e^{-\beta_1(T)T}I_{1-H}\Big(-\frac{\beta_1(T)T}{2}\Big)I_{H-1}\Big(-\frac{\beta_1(T)T}{2}\Big)\\
&\quad \sim c_3\frac{\alpha^2}{\beta^2\gamma^2}\theta_2^2 e^{2\beta T}T^{2-2H}e^{-\beta_1(T)T}I_{1-H}\Big(-\frac{\beta_1(T)T}{2}\Big)I_{H-1}\Big(-\frac{\beta_1(T)T}{2}\Big)
\end{aligned}
$$

$$
\begin{aligned}
&\sim e^{2\beta T} c_3 \frac{\alpha^2}{\beta^2\gamma^2}\theta_2^2 \frac{1}{(-\beta)\pi} T^{1-2H} e^{-2\beta_1(T)T} \\
&= O\big(T^{1-2H}\big) \to 0, \quad \text{as } T \to \infty.
\end{aligned} \tag{26}
$$

Note that $\xi_2(T) - 2\beta_1(T) \to -2\beta$, as $T \to \infty$, by (20) and (22). Hence, by (13) and (45),

$$
\begin{aligned}
&A_3^{(\alpha_1(T),\beta_1(T))}\big(\xi_1(T), \xi_2(T)\big) \\
&\quad = \xi_2(T)\big(\xi_2(T) - 2\beta_1(T)\big)c_6\big(\tfrac{\alpha_1(T)}{\beta_1(T)}\big) \\
&\qquad \times \big(-\beta_1(T)\big)^{2H-1} T e^{-\beta_1(T)T} I_{1-H}\bigg(-\frac{\beta_1(T)T}{2}\bigg) I_{-H}\bigg(-\frac{\beta_1(T)T}{2}\bigg) \\
&\quad \sim \frac{\theta_2^2 + 2\theta_3}{-2\beta} e^{2\beta T}(-2\beta)c_6\big(\tfrac{\alpha}{\beta}\big)\big(-\beta_1(T)\big)^{2H-2}\frac{1}{\pi}e^{-2\beta_1(T)T} \\
&\quad \to \frac{c_6(\frac{\alpha}{\beta})(-\beta)^{2H-2}}{\pi}\big(\theta_2^2 + 2\theta_3\big), \quad \text{as } T \to \infty.
\end{aligned} \tag{27}
$$

Similarly, using (14), (24) and (45), we get

$$
\begin{aligned}
&A_4^{(\alpha_1(T),\beta_1(T))}\big(\xi_1(T), \xi_2(T)\big) \\
&\quad = \big(c_1\big(\tfrac{\alpha_1(T)}{\beta_1(T)}\big)\xi_1(T) - c_2\big(\tfrac{\alpha_1(T)}{\beta_1(T)}\big)\xi_2(T)\big) \\
&\qquad \times \big(\xi_2(T) - 2\beta_1(T)\big)\big(-\beta_1(T)\big)^{H-1} T^{1-H} e^{-\frac{\beta_1(T)T}{2}} I_{1-H}\bigg(-\frac{\beta_1(T)T}{2}\bigg) \\
&\quad \sim c_1\big(\tfrac{\alpha}{\beta}\big)\frac{\theta_2\alpha}{-\beta\gamma} e^{\beta T}(-2\beta)\big(-\beta_1(T)\big)^{H-3/2}\frac{1}{\sqrt{\pi}} T^{1/2-H} e^{-\beta_1(T)T} \\
&\quad = O\big(T^{1/2-H}\big) \to 0, \quad \text{as } T \to \infty.
\end{aligned} \tag{28}
$$

Also, (19) implies

$$
\alpha_1(T)^2 = \alpha^2 + 2\alpha\gamma\theta_1 T^{H-1} + \gamma^2\theta_1^2 T^{2H-2} + O\big(e^{\beta T}\big), \quad \text{as } T \to \infty.
$$

Then, for the expression under the exponential function in (18) we have

$$
\begin{aligned}
\frac{\alpha_1(T)^2 - \alpha^2}{2\gamma^2} w_T^H - \frac{\theta_1\alpha}{\gamma} T^{H-1} w_T^H &= \frac{1}{2}\theta_1^2 T^{2H-2} w_T^H + O\big(w_T^H e^{\beta T}\big) \\
&= \frac{1}{2}\theta_1^2 \lambda_H^{-1} + O\big(w_T^H e^{\beta T}\big) \to \frac{1}{2}\theta_1^2\lambda_H^{-1}, \quad \text{as } T \to \infty,
\end{aligned} \tag{29}
$$

since $w_T^H = \lambda_H^{-1} T^{2-2H}$.

Thus, inserting (23) and (25)–(29) into (18), we arrive at

$$
\begin{aligned}
&\mathbb{E}\bigg[\exp\bigg\{\theta_1 T^{H-1}\bigg(S_T + \beta J_T - \frac{\alpha}{\gamma} w_T^H\bigg) + \theta_2 e^{\beta T}(I_T + \beta K_T) + \theta_3 e^{2\beta T} K_T\bigg\}\bigg] \\
&\quad \to \exp\bigg\{\frac{1}{2}\theta_1^2\lambda_H^{-1}\bigg\}\bigg(1 - \frac{\theta_2^2 + 2\theta_3}{4\beta^2 \sin\pi H}\bigg)^{-1/2}
\end{aligned}
$$

$$\times \exp\left\{ \frac{c_6(\frac{\alpha}{\beta})(-\beta)^{2H-2}(\theta_2^2+2\theta_3)}{8\pi\left(1-\frac{\theta_2^2+2\theta_3}{4\beta^2 \sin \pi H}\right)} \right\}, \quad \text{as } T \to \infty.$$

We see that the limit is a product of MGF of the normal random variable $\xi \overset{d}{=} \mathcal{N}(0, \lambda_H^{-1})$ and MGF of the random vector $\binom{\eta\zeta}{\zeta^2}$, where the random variables $\eta \overset{d}{=} \mathcal{N}(0,1)$ and $\zeta \overset{d}{=} \mathcal{N}\left(\frac{\sqrt{c_6(\frac{\alpha}{\beta})(-\beta)^{H-1}}}{2\sqrt{\pi}}, \frac{1}{4\beta^2 \sin \pi H}\right)$ are independent, see Lemma 5 in Appendix B. This concludes the proof, since $c_6(\frac{\alpha}{\beta}) = (x_0 - \frac{\alpha}{\beta})^2 2\lambda_H^* \rho_H^2$. □

Remark 2. In fact, $\mathcal{N}(0, \lambda_H^{-1})$ is an exact distribution of the random variable $T^{H-1}(S_T + \beta J_T - \frac{\alpha}{\gamma} w_T^H)$ for any T. It can be easily seen from the above proof by putting $\theta_2 = \theta_3 = 0$ (then $\alpha_1(T) = \alpha + \gamma\theta_1 T^{H-1}$, $\beta_1(T) = \beta$, $\xi_1(T) = \xi_2(T) = 0$).

The following series of corollaries will describe asymptotic distributions of minor components of the MLE.

First, by considering the convergence of the first component of the random vector in (16), we immediately get the following result.

Corollary 1. *For the process* $(S_T + \beta J_T)$ *we have*

$$\frac{1}{w_T^H}(S_T + \beta J_T) \xrightarrow{\mathsf{P}} \frac{\alpha}{\gamma}, \quad T \to \infty.$$

Next, we focus on the process I_T. In order to obtain its asymptotic behavior it suffices to write

$$e^{2\beta T} I_T = e^{2\beta T}(I_T + \beta K_T) - \beta e^{2\beta T} K_T$$

and then apply (16).

Corollary 2. *For the process* I_T *we have*

$$e^{2\beta T} I_T \xrightarrow{d} -\beta\zeta^2, \quad T \to \infty,$$

where ζ *has the normal distribution defined in* (17).

Finally, the asymptotic behavior of J_T can be easily derived using Lemma 3, Corollary 1 and the identity

$$-\beta T^{H-\frac{1}{2}} e^{\beta T} J_T = T^{H-\frac{1}{2}} e^{\beta T} S_T - T^{H-\frac{1}{2}} e^{\beta T}(S_T + \beta J_T).$$

Corollary 3. *For the process* J_T *we have*

$$T^{H-\frac{1}{2}} e^{\beta T} J_T \xrightarrow{d} \mathcal{N}\left(\frac{8(x_0 - \frac{\alpha}{\beta})\rho_H(-\beta)^{H-3/2}}{\sqrt{\pi}}, \frac{4\Gamma(H)\Gamma(1-H)}{\lambda_H^*(-\beta)^3\pi}\right), \quad T \to \infty.$$

4 Main results

Now we are ready to prove the main result of the article.

Theorem 1. *Let $\beta < 0$, $H \in (1/2, 1)$. Then*

$$\begin{pmatrix} T^{1-H}(\widehat{\alpha}_T - \alpha) \\ e^{-\beta T}(\widehat{\beta}_T - \beta) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \nu \\ \frac{\eta}{\zeta} \end{pmatrix}, \quad T \to \infty, \tag{30}$$

where $\nu \overset{d}{=} \mathcal{N}(0, \lambda_H \gamma^2)$, $\eta \overset{d}{=} \mathcal{N}(0, 1)$, *and*

$$\zeta \overset{d}{=} \mathcal{N}\left(\frac{(x_0 - \frac{\alpha}{\beta})\rho_H \sqrt{\lambda_H^*}(-\beta)^{H-1}}{\sqrt{2\pi}}, \frac{1}{4\beta^2 \sin \pi H} \right) \tag{31}$$

are independent random variables. In particular, the estimators $\widehat{\alpha}_T$ and $\widehat{\beta}_T$ are asymptotically independent.

Proof. Using (9) and the equality $T^{1-H} = \lambda_H w_T^H T^{H-1}$, we can write

$$\begin{aligned} T^{1-H}(\widehat{\alpha}_T - \alpha) &= \lambda_H w_T^H T^{H-1} \left(\frac{S_T K_T - I_T J_T}{w_T^H K_T - J_T^2} \gamma - \alpha \right) \\ &= \lambda_H w_T^H T^{H-1} \frac{\gamma S_T K_T - \gamma I_T J_T - \alpha w_T^H K_T + \alpha J_T^2}{w_T^H K_T - J_T^2} \\ &= \frac{e^{2\beta T} K_T \gamma \lambda_H T^{H-1}(S_T + \beta J_T - \frac{\alpha}{\gamma} w_T^H)}{e^{2\beta T} K_T - \frac{1}{w_T^H} e^{2\beta T} J_T^2} \\ &\quad + \frac{-\gamma \lambda_H T^{H-1} e^{2\beta T} J_T (I_T + \beta K_T) + \alpha \lambda_H T^{H-1} e^{2\beta T} J_T^2}{e^{2\beta T} K_T - \frac{1}{w_T^H} e^{2\beta T} J_T^2}. \end{aligned} \tag{32}$$

Note that by Corollary 3 and Lemma 4, we see that $J_T = O_{\mathsf{P}}(T^{\frac{1}{2}-H} e^{-\beta T})$, $I_T + \beta K_T = O_{\mathsf{P}}(e^{-\beta T})$, and $e^{2\beta T} K_T - \frac{1}{w_T^H} e^{2\beta T} J_T^2 \xrightarrow{d} \zeta^2$, as $T \to \infty$. Consequently, the second term in the right-hand side of (32) converges to zero in probability.

Further, by (9),

$$\begin{aligned} e^{-\beta T}(\widehat{\beta}_T - \beta) &= e^{-\beta T} \left(\frac{S_T J_T - w_T^H I_T}{w_T^H K_T - J_T^2} - \beta \right) \\ &= e^{-\beta T} \frac{S_T J_T - w_T^H I_T - \beta w_T^H K_T + \beta J_T^2}{w_T^H K_T - J_T^2} \\ &= \frac{-e^{\beta T}(I_T + \beta K_T) + e^{\beta T} J_T \frac{1}{w_T^H}(S_T + \beta J_T)}{e^{2\beta T} K_T - \frac{1}{w_T^H} e^{2\beta T} J_T^2}. \end{aligned} \tag{33}$$

Corollaries 1 and 3 imply that $e^{\beta T} J_T \frac{1}{w_T^H}(S_T + \beta J_T)$ converges to zero in probability. Then applying Lemma 4 and Slutsky's theorem, from (32), (33) we get the convergence (30). □

Remark 3. Unlike the ergodic case (studied in [26]), in the non-ergodic case the initial value x_0 affects the asymptotic bias of $\widehat{\beta}_T$. If $\beta < 0$, then the deterministic term $(x_0 - \frac{\alpha}{\beta})e^{-\beta t}$ in (4) does not converge to zero and, moreover, has the same asymptotic order $O(e^{-\beta T})$ as the stochastic term $\gamma \int_0^t e^{-\beta(t-s)} dB_s^H$. This implies that the asymptotic behavior of the statistics S_T, I_T, J_T, and K_T depends on x_0. A similar dependence on initial condition holds for the non-ergodic Ornstein–Uhlenbeck model driven by the Brownian motion (see [11] and [23, Prop. 3.46]) and some explosive autoregressive models [2, 35, 37].

Remark 4. The model (1) with $x_0 = \frac{\alpha}{\beta}$ was considered in [33, Th. 5.2]. In this particular case we have $\zeta \overset{d}{=} \mathcal{N}(0, \frac{1}{4\beta^2 \sin \pi H})$. Consequently,

$$\frac{e^{-\beta T}}{2\beta}(\widehat{\beta}_T - \beta) \xrightarrow{d} \frac{X\sqrt{\sin \pi H}}{Y}, \quad \text{as } T \to \infty,$$

where X and Y are two independent $\mathcal{N}(0, 1)$ random variables. This completely agrees with [33, Th. 5.2].

Remark 5 (Alternative parameterization)**.** An alternative specification of the fractional Vasicek model is

$$dX_t = \kappa(\mu - X_t)dt + \gamma dB_t^H, \quad t \in [0, T], \quad X_0 = x_0. \tag{34}$$

For the model (34), the MLEs of the parameters μ and κ have the following form [26]:

$$\widehat{\mu}_T = \frac{S_T K_T - I_T J_T}{S_T J_T - w_T^H I_T}\gamma, \qquad \widehat{\kappa}_T = \frac{S_T J_T - w_T^H I_T}{w_T^H K_T - J_T^2}.$$

One can establish the following result: if $\kappa < 0$ and $H \in (1/2, 1)$, then

$$\begin{pmatrix} T^{1-H}(\widehat{\mu}_T - \mu) \\ e^{-\kappa T}(\widehat{\kappa}_T - \kappa) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \widetilde{\nu} \\ \eta/\widetilde{\zeta} \end{pmatrix}, \quad \text{as } T \to \infty, \tag{35}$$

where $\widetilde{\nu} \overset{d}{=} \mathcal{N}(0, \frac{\lambda_H \gamma^2}{\kappa^2})$, $\eta \overset{d}{=} \mathcal{N}(0, 1)$, and

$$\widetilde{\zeta} \overset{d}{=} \mathcal{N}\left(\frac{(x_0 - \mu)\rho_H \sqrt{\lambda_H^*}(-\kappa)^{H-1}}{\sqrt{2\pi}}, \frac{1}{4\kappa^2 \sin \pi H}\right),$$

are independent random variables.

The proof of (35) is carried out by the delta-method. By Taylor's theorem, for the function $g(x, y) = \frac{x}{y}$, we have as $(x, y) \to (\alpha, \beta)$

$$g(x, y) - g(\alpha, \beta) = \frac{1}{\beta}(x - \alpha) - \frac{\alpha}{\beta^2}(y - \beta) + o\left(\sqrt{(x - \alpha)^2 + (y - \beta)^2}\right). \tag{36}$$

Multiplying both sides of (36) by T^{1-H}, and putting $x = \widehat{\alpha}_T$, $y = \widehat{\beta}_T$, we get

$$T^{1-H}\left(\frac{\widehat{\alpha}_T}{\widehat{\beta}_T} - \frac{\alpha}{\beta}\right) = \frac{1}{\beta}T^{1-H}(\widehat{\alpha}_T - \alpha) + R_T, \tag{37}$$

where

$$R_T = -\frac{\alpha}{\beta^2} T^{1-H}(\widehat{\beta}_T - \beta) + o_{\mathsf{P}}\Big(T^{1-H}\sqrt{(\widehat{\alpha}_T - \alpha)^2 + (\widehat{\beta}_T - \beta)^2}\Big) \xrightarrow{\mathsf{P}} 0, \qquad (38)$$

as $T \to \infty$, since $T^{1-H}(\widehat{\beta}_T - \beta) \xrightarrow{\mathsf{P}} 0$ and $T^{1-H}(\widehat{\alpha}_T - \alpha) \xrightarrow{d} \nu$ due to (30).

Finally, by Slutsky's theorem, from (30), (37) and (38) we obtain the convergence

$$\begin{pmatrix} T^{1-H}(\frac{\widehat{\alpha}_T}{\widehat{\beta}_T} - \frac{\alpha}{\beta}) \\ e^{-\beta T}(\widehat{\beta}_T - \beta) \end{pmatrix} = \begin{pmatrix} \frac{1}{\beta} T^{1-H}(\widehat{\alpha}_T - \alpha) \\ e^{-\beta T}(\widehat{\beta}_T - \beta) \end{pmatrix} + \begin{pmatrix} R_T \\ 0 \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \frac{\nu}{\beta} \\ \frac{\eta}{\zeta} \end{pmatrix}, \quad \text{as } T \to \infty,$$

which is equivalent to (35), since $\widehat{\mu}_T = \frac{\widehat{\alpha}_T}{\widehat{\beta}_T}$, $\widehat{\kappa}_T = \widehat{\beta}_T$, $\mu = \frac{\alpha}{\beta}$, $\kappa = \beta$.

Now let us consider the situation when one of the parameters is known. In this case we can obtain strong consistency of the corresponding MLEs (instead of weak one) by applying the strong law of large numbers for martingales, see, e.g., [24, Theorem 2.6.10].

Theorem 2. *Let $\beta < 0$ be known and $H \in (1/2, 1)$. The MLE for α is*

$$\widetilde{\alpha}_T = \frac{\gamma}{w_T^H}(S_T + \beta J_T). \qquad (39)$$

It is unbiased, strongly consistent and normal:

$$T^{1-H}(\widetilde{\alpha}_T - \alpha) \stackrel{d}{=} \mathcal{N}\big(0, \lambda_H \gamma^2\big). \qquad (40)$$

Proof. The form of the MLE (39) was established in [25, Eq. (3.2)]. The normality follows from Remark 2:

$$T^{1-H}(\widetilde{\alpha}_T - \alpha) = \lambda_H \gamma T^{H-1}\Big(S_T + \beta J_T - \frac{\alpha}{\gamma} w_T^H\Big) \stackrel{d}{=} \mathcal{N}\big(0, \lambda_H \gamma^2\big).$$

In order to obtain the strong consistency, we rewrite this equality using the relation

$$S_T + \beta J_T - \frac{\alpha}{\gamma} w_T^H = M_T^H, \qquad (41)$$

which follows from (6) and (5). Then

$$\widetilde{\alpha}_T - \alpha = \gamma \frac{M_T^H}{w_T^H}.$$

Recall that M_T^H is a martingale and $\langle M^H \rangle_T = w_T^H \to \infty$, as $T \to \infty$. Then $\frac{M_T^H}{w_T^H} \to 0$ a.s., as $T \to \infty$, by the strong law of large numbers for martingales [24, Th. 2.6.10, Cor. 1]. □

Remark 6. Actually, the statement of Theorem 2 is true regardless of the sign of β. For $\beta > 0$, (40) was proved in [25, Th. 3.1].

Theorem 3. *Let α be known, $H \in (1/2, 1)$ and $\beta < 0$. The MLE for β is*

$$\widetilde{\beta}_T = \frac{\frac{\alpha}{\gamma} J_T - I_T}{K_T}. \tag{42}$$

It is strongly consistent and

$$e^{-\beta T}(\widetilde{\beta}_T - \beta) \xrightarrow{d} \frac{\eta}{\zeta}, \quad \text{as } T \to \infty, \tag{43}$$

where η and ζ are the same as in Theorem 1.

Proof. The form of the MLE (42) is found in [25, Eq. (3.3)]. The strong consistency is established in the same way as in [25, Th. 3.2]. It follows from (41) that

$$\begin{aligned}\frac{\alpha}{\gamma} J_T - I_T - \beta K_T &= \frac{\alpha}{\gamma}\int_0^T P_H(t)\,dw_T^H - \int_0^T P_H(t)\,dS_t - \beta\int_0^T \big(P_H(t)\big)^2\,dw_t^H \\ &= -\int_0^T P_H(t)\,dM_T^H.\end{aligned}$$

Whence, (42) implies that

$$\widetilde{\beta}_T - \beta = \frac{\frac{\alpha}{\gamma} J_T - I_T - \beta K_T}{K_T} = -\frac{\int_0^T P_H(t)\,dM_T^H}{\int_0^T (P_H(t))^2\,dw_t^H}.$$

Since the process M^H is a martingale with the quadratic variation w^H, the process $\int_0^\cdot P_H(t)\,dM_t^H$ is a martingale with the quadratic variation $\int_0^\cdot (P_H(t))^2\,dw_t^H$. Note that $\int_0^T (P_H(t))^2\,dw_t^H = K_T \to \infty$ in probablility, by Lemma 4. This convergence holds almost surely, because $\int_0^T (P_H(t))^2\,dw_t^H$ is increasing in upper bound T. Therefore, by the strong law of large numbers for martingales [24, Th. 2.6.10, Cor. 1], we get that $\widetilde{\beta}_T \to \beta$ a.s., as $T \to \infty$.

Finally, the convergence (43) follows from the representation

$$e^{-\beta T}(\widetilde{\beta}_T - \beta) = \frac{\frac{\alpha}{\gamma} e^{\beta T} J_T - e^{\beta T}(I_T + \beta K_T)}{e^{2\beta T} K_T},$$

Lemma 4, Corollary 3, and Slutsky's theorem. □

Remark 7. The particular case when the parameter $\alpha = 0$ is known and $x_0 = 0$ was studied in [32]. Similarly to Remark 4, we see that in this case the convergence (43) takes the form

$$\frac{e^{-\beta T}}{2\beta}(\widetilde{\beta}_T - \beta) \xrightarrow{d} \frac{X\sqrt{\sin \pi H}}{Y}, \quad \text{as } T \to \infty,$$

where X and Y are two independent $\mathcal{N}(0, 1)$ random variables. This coincides with the result of [32, Th. 2].

A Modified Bessel function of the first kind

In this section we present some properties of the modified Bessel function of the first kind $I_\nu(x)$, which are helpful for our proofs. For more details on this topic we recommend the book [36]. Let $\nu > -1$, $x \in \mathbb{R}$. Then the function $I_\nu(x)$ could be defined by the following power series [29, Formula 50:6:1]:

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

Note, that if x is negative and ν is non-integer, then the function $I_\nu(x)$ is complex-valued. However, a function $I_\nu(x)/x^\nu$ is always real-valued. This function equals $2^{-\nu}/\Gamma(1+\nu)$ when $x = 0$ and it is even, i.e.

$$\frac{I_\nu(-x)}{(-x)^\nu} = \frac{I_\nu(x)}{x^\nu}, \tag{44}$$

see [29, Formula 50:2:1]. For large values of x the function $I_\nu(x)$ has the following asymptotic behavior [1, Formula 9.7.1]:

$$I_\nu(x) = \frac{e^x}{\sqrt{2\pi x}}\left(1 - \frac{4\nu^2 - 1}{8x} + O\big(x^{-2}\big)\right), \quad x \to \infty. \tag{45}$$

B MGFs related to the bivariate normal distribution

MGF of the product of two normal random variables $X \overset{d}{=} \mathcal{N}(m_1, \sigma_1^2)$, $Y \overset{d}{=} \mathcal{N}(m_2, \sigma_2^2)$ with the correlation coefficient $r = \frac{\operatorname{Cov}(X,Y)}{\sigma_1\sigma_2}$ equals (see e.g. [10])

$$\mathbb{E}\big[\exp\{tXY\}\big] = D^{-1/2} \exp\left\{\frac{(m_1^2\sigma_2^2 + m_2^2\sigma_1^2 - 2rm_1m_2\sigma_1\sigma_2)t^2 + 2m_1m_2t}{2D}\right\}, \tag{46}$$

where

$$D = \big(1 - (1+r)\sigma_1\sigma_2 t\big)\big(1 + (1-r)\sigma_1\sigma_2 t\big).$$

Lemma 5. *For two independent normal random variables $X \overset{d}{=} \mathcal{N}(m, \sigma^2)$ and $Y \overset{d}{=} \mathcal{N}(0, 1)$,*

$$\mathbb{E}\big[\exp\{\theta_1 XY + \theta_2 X^2\}\big] = \big[1 - \sigma^2(\theta_1^2 + 2\theta_2)\big]^{-\frac{1}{2}} \exp\left\{\frac{m^2(\theta_1^2 + 2\theta_2)}{2[1 - \sigma^2(\theta_1^2 + 2\theta_2)]}\right\}. \tag{47}$$

Proof. Evidently,

$$\theta_1 XY + \theta_2 X^2 = X(\theta_1 Y + \theta_2 X),$$

where $X \overset{d}{=} \mathcal{N}(m, \sigma^2)$, $\theta_1 Y + \theta_2 X \overset{d}{=} \mathcal{N}(\theta_2 m, \theta_1^2 + \theta_2^2\sigma^2)$, and

$$\operatorname{Cov}(X, \theta_1 Y + \theta_2 X) = \theta_2 \operatorname{Cov}(X, X) = \theta_2\sigma^2.$$

Applying (46) with $t = 1$, $m_1 = m$, $\sigma_1^2 = \sigma^2$, $m_2 = \theta_2 m$, $\sigma_2^2 = \theta_1^2 + \theta_2^2\sigma^2$, and $r = \frac{\theta_2\sigma}{\sqrt{\theta_1^2+\theta_2^2\sigma^2}}$, we get

$$\begin{aligned}
&\mathbb{E}\bigl[\exp\bigl\{X(\theta_1 Y + \theta_2 X)\bigr\}\bigr]\\
&\quad = D^{-1/2}\exp\left\{\frac{m^2(\theta_1^2+\theta_2^2\sigma^2)+\theta_2^2 m^2\sigma^2 - 2\theta_2^2 m^2\sigma^2 + 2\theta_2 m^2}{2D}\right\}\\
&\quad = D^{-1/2}\exp\left\{\frac{m^2(\theta_1^2+2\theta_2)}{2D}\right\},
\end{aligned}$$

where

$$\begin{aligned}
D &= \left(1-\left(1+\frac{\theta_2\sigma}{\sqrt{\theta_1^2+\theta_2^2\sigma^2}}\right)\sigma\sqrt{\theta_1^2+\theta_2^2\sigma^2}\right)\\
&\quad\times\left(1+\left(1-\frac{\theta_2\sigma}{\sqrt{\theta_1^2+\theta_2^2\sigma^2}}\right)\sigma\sqrt{\theta_1^2+\theta_2^2\sigma^2}\right)\\
&= \bigl(1-\sigma\sqrt{\theta_1^2+\theta_2^2\sigma^2}-\theta_2\sigma^2\bigr)\bigl(1+\sigma\sqrt{\theta_1^2+\theta_2^2\sigma^2}-\theta_2\sigma^2\bigr)\\
&= \bigl(1-\theta_2\sigma^2\bigr)^2-\sigma^2\bigl(\theta_1^2+\theta_2^2\sigma^2\bigr) = 1-\sigma^2\bigl(\theta_1^2+2\theta_2\bigr),
\end{aligned}$$

whence (47) follows. □

Acknowledgement

The authors are grateful to the referees for their valuable comments and suggestions.

Funding

The second author acknowledges that the present research is carried through within the frame and support of the ToppForsk project nr. 274410 of the Research Council of Norway with title STORM: Stochastics for Time-Space Risk Models.

References

[1] Abramowitz, M., Stegun, I.A. (eds.): Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover Publications, Inc., New York (1992). MR1225604

[2] Anderson, T.W.: On asymptotic distributions of estimates of parameters of stochastic difference equations. Ann. Math. Stat. **30**, 676–687 (1959). MR107347

[3] Belfadli, R., Es-Sebaiy, K., Ouknine, Y.: Parameter estimation for fractional Ornstein–Uhlenbeck processes: non-ergodic case. Front. Chem. Sci. Eng. **1**(1), 1–16 (2011)

[4] Bercu, B., Coutin, L., Savy, N.: Sharp large deviations for the fractional Ornstein–Uhlenbeck process. Theory Probab. Appl. **55**(4), 575–610 (2011). MR2859161

[5] Berzin, C., Latour, A., León, J.R.: Inference on the Hurst Parameter and the Variance of Diffusions Driven by Fractional Brownian Motion. Lecture Notes in Statistics, vol. 216. Springer, Cham (2014). MR3289986

[6] Bishwal, J.P.N.: Minimum contrast estimation in fractional Ornstein-Uhlenbeck process: Continuous and discrete sampling. Fract. Calc. Appl. Anal. **14**, 375–410 (2011). MR2837637

[7] Cheridito, P., Kawaguchi, H., Makoto, M.: Fractional Ornstein–Uhlenbeck processes. Electron. J. Probab. **8**, 1–14 (2003). MR1961165

[8] Chronopoulou, A., Viens, F.G.: Estimation and pricing under long-memory stochastic volatility. Ann. Finance **8**(2–3), 379–403 (2012). MR2922802

[9] Corlay, S., Lebovits, J., Véhel, J.L.: Multifractional stochastic volatility models. Math. Finance **24**(2), 364–402 (2014). MR3274947

[10] Craig, C.C.: The frequency of function of y/x. Ann. of Math. (2) **30**(1–4), 471–486 (1928/29). MR1502897

[11] Feigin, P.D.: Maximum likelihood estimation for continuous-time stochastic processes. Adv. Appl. Probab. **8**(4), 712–736 (1976). MR0426342

[12] Fink, H., Klüppelberg, C., Zähle, M.: Conditional distributions of processes related to fractional Brownian motion. J. Appl. Probab. **50**(1), 166–183 (2013). MR3076779

[13] Hao, R., Liu, Y., Wang, S.: Pricing credit default swap under fractional Vasicek interest rate model. J. Math. Finance **4**(1), 10–20 (2014)

[14] Hu, Y., Nualart, D.: Parameter estimation for fractional Ornstein–Uhlenbeck processes. Stat. Probab. Lett. **80**(11–12), 1030–1038 (2010). MR2638974

[15] Hu, Y., Song, J.: Parameter estimation for fractional Ornstein–Uhlenbeck processes with discrete observations. In: Malliavin Calculus and Stochastic Analysis. A Festschrift in Honor of David Nualart. Springer Proc. Math. Stat., vol. 34, pp. 427–442. Springer, New York (2013). MR3070455

[16] Hu, Y., Nualart, D., Zhou, H.: Parameter estimation for fractional Ornstein–Uhlenbeck processes of general Hurst parameter. Stat. Inference Stoch. Process. **22**(1), 111–142 (2019). MR3918739

[17] Kleptsyna, M.L., Le Breton, A.: Statistical analysis of the fractional Ornstein–Uhlenbeck type process. Stat. Inference Stoch. Process. **5**(3), 229–248 (2002). MR1943832

[18] Kleptsyna, M.L., Le Breton, A., Roubaud, M.-C.: Parameter estimation and optimal filtering for fractional type stochastic systems. Stat. Inference Stoch. Process. **3**(1–2), 173–182 (2000). MR1819294

[19] Kozachenko, Y., Melnikov, A., Mishura, Y.: On drift parameter estimation in models with fractional Brownian motion. Statistics **49**(1), 35–62 (2015). MR3304366

[20] Kubilius, K., Mishura, Y., Ralchenko, K.: Parameter Estimation in Fractional Diffusion Models. Bocconi & Springer Series, vol. 8. Springer (2017). MR3304366

[21] Kubilius, K., Mishura, Y., Ralchenko, K., Seleznjev, O.: Consistency of the drift parameter estimator for the discretized fractional Ornstein–Uhlenbeck process with Hurst index $H \in (0, \frac{1}{2})$. Electron. J. Stat. **9**(2), 1799–1825 (2015). MR3391120

[22] Kukush, A., Mishura, Y., Ralchenko, K.: Hypothesis testing of the drift parameter sign for fractional Ornstein–Uhlenbeck process. Electron. J. Stat. **11**(1), 385–400 (2017). MR3608678

[23] Kutoyants, Y.A.: Statistical Inference for Ergodic Diffusion Processes. Springer Series in Statistics, p. 481. Springer (2004). MR2144185

[24] Liptser, R.S., Shiryayev, A.N.: Theory of Martingales vol. 49. Kluwer Academic Publishers Group, Dordrecht (1989). MR1022664

[25] Lohvinenko, S., Ralchenko, K.: Maximum likelihood estimation in the fractional Vasicek model. Lith. J. Stat. **56**(1), 77–87 (2017)

[26] Lohvinenko, S., Ralchenko, K.: Asymptotic distribution of maximum likelihood estimator in fractional Vasicek model. Theory Probab. Math. Stat. **99**, 134–151 (2018). MR3908663

[27] Lohvinenko, S., Ralchenko, K., Zhuchenko, O.: Asymptotic properties of parameter estimators in fractional Vasicek model. Lith. J. Stat. **55**(1), 102–111 (2016)

[28] Mishura, Y., Ralchenko, K.: Drift parameter estimation in the models involving fractional Brownian motion. In: Modern Problems of Stochastic Analysis and Statistics. Springer Proc. Math. Stat., vol. 208, pp. 237–268. Springer, Cham (2017). MR3747669

[29] Oldham, K., Myland, J., Spanier, J.: An Atlas of Functions, 2nd edn. Springer (2009). MR2466333

[30] Song, L., Li, K.: Pricing option with stochastic interest rates and transaction costs in fractional Brownian markets. Discrete Dyn. Nat. Soc., 7056734-8 (2018). MR3842704

[31] Tanaka, K.: Distributions of the maximum likelihood and minimum contrast estimators associated with the fractional Ornstein–Uhlenbeck process. Stat. Inference Stoch. Process. **16**(3), 173–192 (2013). MR3123562

[32] Tanaka, K.: Maximum likelihood estimation for the non-ergodic fractional Ornstein–Uhlenbeck process. Stat. Inference Stoch. Process. **18**(3), 315–332 (2015). MR3395610

[33] Tanaka, K., Xiao, W., Yu, J.: Maximum likelihood estimation for the fractional Vasicek model. Research Collection School Of Economics, Paper No. 08-2019, 1–31 (2019)

[34] Vasicek, O.: An equilibrium characterization of the term structure. J. Financ. Econ. **5**(2), 177–188 (1977)

[35] Wang, X., Yu, J.: Limit theory for an explosive autoregressive process. Econ. Lett. **126**, 176–180 (2015). MR3298417

[36] Watson, G.N.: A Treatise on the Theory of Bessel Functions. Cambridge Mathematical Library. Cambridge University Press, Cambridge (1995). MR1349110

[37] White, J.S.: The limiting distribution of the serial correlation coefficient in the explosive case. Ann. Math. Stat. **29**, 1188–1197 (1958). MR100952

[38] Xiao, W., Yu, J.: Asymptotic theory for estimating drift parameters in the fractional Vasicek model. Econom. Theory **35**(1), 198–231 (2019). MR3904176

[39] Xiao, W., Yu, J.: Asymptotic theory for rough fractional Vasicek models. Econ. Lett. **177**, 26–29 (2019). MR3904332

[40] Xiao, W., Zhang, W., Zhang, X., Chen, X.: The valuation of equity warrants under the fractional Vasicek process of the short-term interest rate. Physica A **394**, 320–337 (2014). MR3123524