

RESEARCH

Open Access

Maximum likelihood estimators in linear regression models with Ornstein-Uhlenbeck process

Hongchang Hu¹, Xiong Pan^{2*} and Lifeng Xu¹

*Correspondence: pxjhlh@163.com
²Faculty of Information Engineering,
China University of Geosciences,
Wuhan, 430074, China
Full list of author information is
available at the end of the article

Abstract

The paper studies the linear regression model

$$y_t = x_t^T \beta + \varepsilon_t, \quad t = 1, 2, \dots, n,$$

where

$$d\varepsilon_t = \lambda(\mu - \varepsilon_t) dt + \sigma dB_t,$$

with parameters $\lambda, \sigma \in R^+$, $\mu \in R$ and $\{B_t, t \geq 0\}$ the standard Brownian motion. Firstly, the maximum likelihood (ML) estimators of β , λ and σ^2 are given. Secondly, under general conditions, the asymptotic properties of the ML estimators are investigated. And then, limiting distributions for likelihood ratio test statistics of the hypothesis are also given. Lastly, the validity of the method are illuminated by two real examples.

MSC: 62J05; 62M10; 60J60

Keywords: linear regression model; maximum likelihood estimator; Ornstein-Uhlenbeck process; asymptotic property; likelihood ratio test

1 Introduction

Consider the following linear regression model

$$y_t = x_t^T \beta + \varepsilon_t, \quad t = 1, 2, \dots, n, \quad (1.1)$$

where y_t 's are scalar response variables, x_t 's are explanatory variables, β is an m -dimensional unknown parameter, and $\{\varepsilon_t\}$ is an Ornstein-Uhlenbeck process, which satisfies the linear stochastic differential equation (SDE)

$$d\varepsilon_t = \lambda(\mu - \varepsilon_t) dt + \sigma dB_t \quad (1.2)$$

with parameters $\lambda, \sigma \in R^+$, $\mu \in R$ and $\{B_t, t \geq 0\}$ the standard Brownian motion.

It is well known that a linear regression model is the most important and popular model in the statistical literature, which attracts many people to investigate the model. For an

ordinary linear regression model (when the errors are independent and identically distributed (i.i.d.) random variables), Wang and Zhou [1], Anatolyev [2], Bai and Guo [3], Chen [4], Gil *et al.* [5], Hampel *et al.* [6], Cui [7], Durbin [8] and Li and Yang [9] used various estimation methods to obtain estimators of the unknown parameters in (1.1) and discussed some large or small sample properties of these estimators. Recently, linear regression with serially correlated errors has attracted increasing attention from statisticians and economists. One case of considerable interest is that the errors are autoregressive processes; Hu [10], Wu [11], and Fox and Taqqu [12] established its asymptotic normality with the usual p_n -normalization in the case of long memory stationary Gaussian observations errors. Giraitis and Surgailis [13] extended this result to non-Gaussian linear sequences. Koul and Surgailis [14] established the asymptotic normality of the Whittle estimator in linear regression models with non-Gaussian long memory moving average errors. Shiohama and Taniguchi [15] estimated the regression parameters in a linear regression model with autoregressive process. Fan [16] investigated moderate deviations for M -estimators in linear models with ϕ -mixing errors.

The Ornstein-Uhlenbeck process was originally introduced by Ornstein and Uhlenbeck [17] as a model for particle motion in a fluid. In physical sciences, the Ornstein-Uhlenbeck process is a prototype of a noisy relaxation process, whose probability density function $f(x, t)$ can be described by the Fokker-Planck equation (see Janczura *et al.* [18], Debbasch *et al.* [19], Gillespie [20], Ditlevsen and Lansky [21], Garbaczewski and Olkiewicz [22], Plastino and Plastino [23]):

$$\frac{\partial f(x, t)}{\partial t} = \frac{\partial}{\partial x} (\lambda(x - \mu)f(x, t)) + \frac{\sigma^2}{2} \frac{\partial^2 f(x, t)}{\partial x^2}.$$

This process is now widely used in many areas of application. The main characteristic of the Ornstein-Uhlenbeck process is the tendency to return towards the long-term equilibrium μ . This property, known as mean-reversion, is found in many real life processes, *e.g.*, in commodity and energy price processes (see Fasen [24], Yu [25], Geman [26]). There are a number of papers concerned with the Ornstein-Uhlenbeck process, for example, Janczura *et al.* [18], Zhang *et al.* [27], Rieder [28], Iacus [29], Bishwal [30], Shimizu [31], Zhang and Zhang [32], Chronopoulou and Viens [33], Lin and Wang [34] and Xiao *et al.* [35]. It is well known that the solution of model (1.2) is an autoregressive process. For a constant or functional or random coefficient autoregressive model, many people (for example, Magdalinos [36], Andrews and Guggenberger [37], Fan and Yao [38], Berk [39], Goldenshluger and Zeevi [40], Liebscher [41], Baran *et al.* [42], Distaso [43] and Harvill and Ray [44]) used various estimation methods to obtain estimators and discussed some asymptotic properties of these estimators, or investigated hypotheses testing.

By (1.1) and (1.2), we can obtain that the more general process satisfies the SDE

$$dy_t = \lambda(L(t, \lambda, \mu, \beta) - y_t) dt + \sigma dB_t, \quad (1.3)$$

where $L(t, \lambda, \mu, \beta)$ is a time-dependent mean reversion level with three parameters. Thus, model (1.3) is a general Ornstein-Uhlenbeck process. Its special cases have gained much attention and have been applied to many fields such as economics, physics, geography, geology, biology and agriculture. Dehling *et al.* [45] considered the model with maximum likelihood estimate, and proved strong consistency and asymptotic normality. Lin

and Wang [34] established the existence of a successful coupling for a class of stochastic differential equations given by (1.3). Bishwal [30] investigated the uniform rate of weak convergence of the minimum contrast estimator in the Ornstein-Uhlenbeck process (1.3).

The solution of model (1.2) is given by

$$\varepsilon_t = e^{-\lambda t} \varepsilon_0 + \mu(1 - e^{-\lambda t}) + \sigma \int_0^t e^{\lambda(s-t)} dB_s, \tag{1.4}$$

where $\int_0^t e^{\lambda(s-t)} dB_s \sim N(0, \frac{1 - \exp(-2\lambda t)}{2\lambda})$.

The process observed in discrete time is more relevant in statistics and economics. Therefore, by (1.4), the Ornstein-Uhlenbeck time series for $t = 1, 2, \dots, n$ is given by

$$\varepsilon_t = e^{-\lambda d} \varepsilon_{t-1} + \mu(1 - e^{-\lambda d}) + \sigma \sqrt{\frac{1 - e^{-2\lambda d}}{2\lambda}} \eta_t, \tag{1.5}$$

where $\eta_t \sim N(0, 1)$ i.i.d. random errors and with equidistant time lag d , fixed in advance. Models (1.1) and (1.5) include many special cases such as a linear regression model with constant coefficient autoregressive processes (when $\mu = 0$; see Hu [10], Wu [11], Maller [46], Pere [47] and Fuller [48]), Ornstein-Uhlenbeck time series or processes (when $\beta = 0$; see Rieder [28], Iacus [29], Bishwal [30], Shimizu [31] and Zhang and Zhang [32]), constant coefficient autoregressive processes (when $\mu = 0, \beta = 0$; see Chambers [49], Hamilton [50], Brockwell and Davis [51] and Abadir and Lucas [52], *etc.*).

The paper discusses models (1.1) and (1.5). The organization of the paper is as follows. In Section 2 some estimators of β, θ and σ^2 are given by the quasi-maximum likelihood method. Under general conditions, the existence and consistency of the quasi-maximum likelihood estimators as well as asymptotic normality are investigated in Section 3. The hypothesis testing is given in Section 4. Some preliminary lemmas are presented in Section 5. The main proofs of theorems are presented in Section 6, with two real examples in Section 7.

2 Estimation method

Without loss of generality, we assume that $\mu = 0, \varepsilon_0 = 0$ in the sequel. Write the ‘true’ model as

$$y_t = x_t^T \beta_0 + e_t, \quad t = 1, 2, \dots, n \tag{2.1}$$

and

$$e_t = \exp(-\lambda_0 d) e_{t-1} + \sigma_0 \sqrt{\frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0}} \eta_t, \tag{2.2}$$

where $\eta_t \sim N(0, 1)$ i.i.d.

By (2.2), we have

$$e_t = \sigma_0 \sqrt{\frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0}} \sum_{j=1}^t \exp\{-\lambda_0 d(t-j)\} \eta_j. \tag{2.3}$$

Thus e_t is measurable with respect to the σ -field H generated by $\eta_1, \eta_2, \dots, \eta_t$, and

$$Ee_t = 0, \quad \text{Var}(e_t) = \sigma_0^2 \frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0} \exp\{-\lambda_0 dt(t-1)\}. \tag{2.4}$$

Using similar arguments as those of Rieder [28] or Maller [46], we get the log-likelihood of y_2, y_3, \dots, y_n conditional on y_1 ,

$$\begin{aligned} \Psi_n(\beta, \lambda, \sigma^2) &= \log L_n \\ &= -\frac{1}{2}(n-1) \log\left(\frac{\pi\sigma^2}{\lambda}\right) - \frac{1}{2}(n-1) \log(1 - \exp(-2\lambda d)) \\ &\quad - \frac{\lambda}{\sigma^2(1 - \exp(-2\lambda d))} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})^2. \end{aligned} \tag{2.5}$$

We maximize (2.5) to obtain QML estimators denoted by $\hat{\sigma}_n^2, \hat{\beta}_n, \hat{\lambda}_n$ (when they exist). Then the first derivatives of Ψ_n may be written as

$$\frac{\partial \Psi_n}{\partial \sigma^2} = -\frac{n-1}{2\sigma^2} + \frac{\lambda}{\sigma^4(1 - \exp(-2\lambda d))} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})^2, \tag{2.6}$$

$$\begin{aligned} \frac{\partial \Psi_n}{\partial \lambda} &= \frac{n-1}{2\lambda} - \frac{(n-1)d \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} - \frac{2d\lambda \exp(-\lambda d)}{\sigma^2(1 - \exp(-2\lambda d))} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})\varepsilon_{t-1} \\ &\quad - \frac{1 - (1 + 2d\lambda) \exp(-2\lambda d)}{\sigma^2(1 - \exp(-2\lambda d))^2} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})^2 \end{aligned} \tag{2.7}$$

and

$$\frac{\partial \Psi_n}{\partial \beta} = \frac{2\lambda}{\sigma^2(1 - \exp(-2\lambda d))} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})(x_t - \exp(-\lambda d)x_{t-1}). \tag{2.8}$$

Thus $\hat{\sigma}_n^2, \hat{\beta}_n, \hat{\lambda}_n$ satisfy the following estimation equations:

$$\hat{\sigma}_n^2 = \frac{2\hat{\lambda}_n}{(n-1)(1 - \exp(-2\hat{\lambda}_n d))} \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d)\hat{\varepsilon}_{t-1})^2, \tag{2.9}$$

$$\begin{aligned} \frac{\hat{\sigma}_n^2(1 - (1 + 2d\hat{\lambda}_n) \exp(-2\hat{\lambda}_n d))}{2\hat{\lambda}_n} - \frac{2d\hat{\lambda}_n \exp(-\hat{\lambda}_n d)}{n-1} \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d)\hat{\varepsilon}_{t-1})\hat{\varepsilon}_{t-1} \\ - \frac{1 - (1 + 2d\hat{\lambda}_n) \exp(-2\hat{\lambda}_n d)}{(1 - \exp(-2\hat{\lambda}_n d))(n-1)} \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d)\hat{\varepsilon}_{t-1})^2 = 0 \end{aligned} \tag{2.10}$$

and

$$\sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d)\hat{\varepsilon}_{t-1})(x_t - \exp(-\hat{\lambda}_n d)x_{t-1}) = 0, \tag{2.11}$$

where

$$\hat{\varepsilon}_t = y_t - x_t^T \hat{\beta}_n. \tag{2.12}$$

To obtain our results, the following conditions are sufficient (see Maller [46]).

(A1) $X_n = \sum_{t=2}^n x_t x_t^T$ is positive definite for sufficiently large n and

$$\lim_{n \rightarrow \infty} \max_{1 \leq t \leq n} x_t^T X_n^{-1} x_t = 0. \tag{2.13}$$

(A2)

$$\limsup_{n \rightarrow \infty} |\tilde{\lambda}|_{\max} (X_n^{-\frac{1}{2}} Z_n X_n^{-\frac{T}{2}}) < 1, \tag{2.14}$$

where $Z_n = \frac{1}{2} \sum_{t=2}^n (x_t x_{t-1}^T + x_{t-1} x_t^T)$, $|\tilde{\lambda}|_{\max}(\cdot)$ denotes the maximum in absolute value of the eigenvalues of a symmetric matrix.

For ease of exposition, we shall introduce the following notations which will be used later in the paper.

Let $(m + 1)$ -vector $\theta = (\beta, \lambda)$. Define

$$S_n(\theta) = \sigma^2 \frac{\partial \Psi_n}{\partial \theta} = \sigma^2 \left(\frac{\partial \Psi_n}{\partial \beta}, \frac{\partial \Psi_n}{\partial \lambda} \right), \quad F_n(\theta) = -\sigma^2 \frac{\partial^2 \Psi_n}{\partial \theta \partial \theta^T}. \tag{2.15}$$

By (2.7) and (2.8), we get the components of $F_n(\theta)$

$$\begin{aligned} -\sigma^2 \frac{\partial^2 \Psi_n}{\partial \beta \partial \beta^T} &= \frac{2\lambda}{1 - \exp(-2\lambda d)} \sum_{t=2}^n (x_t - \exp(-\lambda d)x_{t-1})(x_t - \exp(-\lambda d)x_{t-1})^T \\ &= \frac{2\lambda}{1 - \exp(-2\lambda d)} X_n(\lambda), \end{aligned} \tag{2.16}$$

$$\begin{aligned} -\sigma^2 \frac{\partial^2 \Psi_n}{\partial \beta \partial \lambda} &= -\frac{2d\lambda \exp(-\lambda d)}{1 - \exp(-2\lambda d)} \sum_{t=2}^n (\varepsilon_{t-1}x_t + \varepsilon_t x_{t-1} - 2 \exp(-\lambda d)x_{t-1}\varepsilon_{t-1}) \\ &\quad - \frac{1 - (1 + 2d\lambda) \exp(-2\lambda d)}{(1 - \exp(-2\lambda d))^2} \\ &\quad \cdot \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})(x_t - \exp(-\lambda d)x_{t-1}) \end{aligned} \tag{2.17}$$

and

$$\begin{aligned} -\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} &= \frac{\sigma^2(n-1)}{2\lambda^2} - \frac{2\sigma^2(n-1)d^2 \exp(-2\lambda d)}{(1 - \exp(-2\lambda d))^2} + \frac{2d^2\lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} \sum_{t=2}^n \varepsilon_{t-1}^2 \\ &\quad + \frac{2d(1 - d\lambda - (1 + d\lambda) \exp(-2\lambda d)) \exp(-\lambda d)}{(1 - \exp(-2\lambda d))^2} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})\varepsilon_{t-1} \\ &\quad + \frac{2d \exp(-\lambda d)[1 - (1 + 2d\lambda) \exp(-2\lambda d)]}{(1 - \exp(-2\lambda d))^2} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})\varepsilon_{t-1} \\ &\quad + \frac{4d \exp(-2\lambda d)[d\lambda - 1 + (1 + d\lambda) \exp(-2\lambda d)]}{(1 - \exp(-2\lambda d))^3} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})^2 \\ &= \frac{\sigma^2(n-1)}{2\lambda^2} - \frac{2\sigma^2(n-1)d^2 \exp(-2\lambda d)}{(1 - \exp(-2\lambda d))^2} + \frac{2d^2\lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} \sum_{t=2}^n \varepsilon_{t-1}^2 \end{aligned}$$

$$\begin{aligned}
 & + \frac{2d \exp(-\lambda d)[(2-d\lambda) - d\lambda \exp(-2\lambda d)]}{(1 - \exp(-2\lambda d))^2} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})\varepsilon_{t-1} \\
 & + \frac{4d \exp(-2\lambda d)[d\lambda - 1 + (1+d\lambda)\exp(-2\lambda d)]}{(1 - \exp(-2\lambda d))^3} \\
 & \cdot \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})^2. \tag{2.18}
 \end{aligned}$$

Hence we have

$$F_n(\theta) = \begin{pmatrix} \frac{2\lambda}{1-\exp(-2\lambda d)} X_n(\lambda) & -\sigma^2 \frac{\partial^2 \Psi_n}{\partial \beta \partial \lambda} \\ * & -\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} \end{pmatrix}, \tag{2.19}$$

where the * indicates that the elements are filled in by symmetry. By (2.18), we have

$$\begin{aligned}
 E \left\{ -\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} \Big|_{\theta=\theta_0} \right\} & = (n-1)\sigma_0^2 \left\{ \frac{1}{2\lambda_0^2} + \frac{2d \exp(-2\lambda_0 d)[-1 + (1+d\lambda_0)\exp(-2\lambda_0 d)]}{\lambda_0(1 - \exp(-2\lambda_0 d))^2} \right\} \\
 & + \frac{2d^2 \lambda_0 \exp(-2\lambda_0 d)}{1 - \exp(-2\lambda_0 d)} \sum_{t=2}^n E e_{t-1}^2 \\
 & = (n-1)\sigma_0^2 \frac{[1 - (1+2d\lambda_0)\exp(-2\lambda_0 d)]^2}{2\lambda_0^2(1 - \exp(-2\lambda_0 d))^2} \\
 & + \frac{2d^2 \lambda_0 \exp(-2\lambda_0 d)}{1 - \exp(-2\lambda_0 d)} \sum_{t=2}^n E e_{t-1}^2 \\
 & = \Delta_n(\theta_0, \sigma_0) = O(n). \tag{2.20}
 \end{aligned}$$

Thus,

$$\begin{aligned}
 D_n & = E(F_n(\theta_0)) \\
 & = \begin{pmatrix} \frac{2\lambda_0}{1-\exp(-2\lambda_0 d)} X_n(\lambda_0) & 0 \\ 0 & \Delta_n(\theta_0, \sigma_0) \end{pmatrix} \tag{2.21}
 \end{aligned}$$

3 Large sample properties of the estimators

Theorem 3.1 *Suppose that conditions (A1)-(A2) hold. Then there is a sequence $A_n \downarrow 0$ such that, for each $A > 0$, as $n \rightarrow \infty$, the probability*

$$P\{ \text{there are estimators } \hat{\theta}_n, \hat{\sigma}_n^2 \text{ with } S_n(\hat{\theta}_n) = 0, \text{ and } (\hat{\theta}_n, \hat{\sigma}_n^2) \in N'_n(A) \} \rightarrow 1. \tag{3.1}$$

Furthermore,

$$(\hat{\theta}_n, \hat{\sigma}_n^2) \rightarrow_p (\theta_0, \sigma_0^2), \quad n \rightarrow \infty, \tag{3.2}$$

where, for each $n = 1, 2, \dots, A > 0$ and $A_n \in (0, \sigma_0^2)$, define neighborhoods

$$N_n(A) = \{ \theta \in R^{m+1} : (\theta - \theta_0)^T D_n(\theta - \theta_0) \leq A^2 \} \tag{3.3}$$

and

$$N'_n(A) = N_n(A) \cap \{\sigma^2 \in [\sigma_0^2 - A_n, \sigma_0^2 + A_n]\}. \tag{3.4}$$

Theorem 3.2 *Suppose that conditions (A1)-(A2) hold. Then*

$$\frac{1}{\hat{\sigma}_n} F_n^{\frac{T}{2}}(\hat{\theta}_n)(\hat{\theta}_n - \theta_0) \rightarrow_D N(0, I_{m+1}), \quad n \rightarrow \infty. \tag{3.5}$$

In the following, we will investigate some special cases in models (1.1) and (1.5). From Theorem 3.1 and Theorem 3.2, we obtain the following results. Here we omit their proofs.

Corollary 3.1 *If $\beta = 0$, then*

$$\frac{\sqrt{\Delta_n(\theta_0, \sigma_0)}}{\hat{\sigma}_n}(\hat{\lambda}_n - \lambda_0) \rightarrow_D N(0, 1), \quad n \rightarrow \infty. \tag{3.6}$$

Corollary 3.2 *If $\beta = 0$, then*

$$\sqrt{n}(\hat{\lambda}_n - \lambda_0) \rightarrow_D N(0, \sigma_0^2), \quad n \rightarrow \infty. \tag{3.7}$$

4 Hypothesis testing

In order to fit a data set $\{y_t, t = 1, 2, \dots, n\}$, we may use model (1.3) or an Ornstein-Uhlenbeck process with a constant mean level model

$$dy_t = \lambda(\mu - y_t) dt + \sigma dB_t. \tag{4.1}$$

If $\beta \neq 0$, then we use model (1.3), namely models (1.1) and (1.2). If $\beta = 0$, then we use model (1.4). How to know $\beta = 0$ or $\beta \neq 0$? In the section, we shall consider the question about hypothesis testing and obtain limiting distributions for likelihood ratio (LR) test statistics (see Fan and Jiang [53]).

Under the null hypothesis

$$H_0: \beta_0 = 0, \quad \lambda_0 > 0, \quad \sigma_0 > 0, \tag{4.2}$$

let $\hat{\beta}_{0n}, \hat{\lambda}_{0n}, \hat{\sigma}_{0n}^2$ be the corresponding ML estimators of β, λ, σ^2 . Also let

$$\hat{L}_n = -2\Psi_n(\hat{\beta}_n, \hat{\lambda}_n, \hat{\sigma}_n^2) \tag{4.3}$$

and

$$\hat{L}_{0n} = -2\Psi_n(\hat{\beta}_{0n}, \hat{\lambda}_{0n}, \hat{\sigma}_{0n}^2). \tag{4.4}$$

By (2.9) and (2.5), we have that

$$\begin{aligned} \hat{L}_n &= (n-1) \log\left(\frac{\pi \hat{\sigma}_n^2}{\hat{\lambda}_n}\right) + (n-1) \log(1 - \exp(-2\hat{\lambda}_n d)) \\ &\quad + \frac{2\hat{\lambda}_n}{\hat{\sigma}_n^2(1 - \exp(-2\hat{\lambda}_n d))} \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d)\hat{\varepsilon}_{t-1})^2 \end{aligned}$$

$$\begin{aligned}
 &= (n-1) \log\left(\frac{\pi \hat{\sigma}_n^2}{\hat{\lambda}_n}\right) + (n-1) \log(1 - \exp(-2\hat{\lambda}_n d)) + (n-1) \\
 &= (n-1) \log(\pi + 1) + (n-1) \log(\hat{\sigma}_n^2) \\
 &\quad + (n-1) (\log(1 - \exp(-2\hat{\lambda}_n d)) - \log(\hat{\lambda}_n)).
 \end{aligned} \tag{4.5}$$

And similarly,

$$\begin{aligned}
 \hat{L}_{0n} &= (n-1) \log(\pi + 1) + (n-1) \log(\hat{\sigma}_{0n}^2) \\
 &\quad + (n-1) (\log(1 - \exp(-2\hat{\lambda}_{0n} d)) - \log(\hat{\lambda}_{0n})).
 \end{aligned} \tag{4.6}$$

By (4.5) and (4.6), we have

$$\begin{aligned}
 \tilde{d}(n) &= \hat{L}_{0n} - \hat{L}_n \\
 &= (n-1) \log\left(\frac{\hat{\sigma}_{0n}^2}{\hat{\sigma}_n^2}\right) + (n-1) \left(\log\left(\frac{1 - \exp(-2\hat{\lambda}_{0n} d)}{1 - \exp(-2\hat{\lambda}_n d)}\right) - \log\left(\frac{\hat{\lambda}_{0n}}{\hat{\lambda}_n}\right)\right) \\
 &= (n-1) \left(\frac{\hat{\sigma}_{0n}^2}{\hat{\sigma}_n^2} - 1\right) + (n-1) \left(\frac{1 - \exp(-2\hat{\lambda}_{0n} d)}{1 - \exp(-2\hat{\lambda}_n d)} - \frac{\hat{\lambda}_{0n}}{\hat{\lambda}_n}\right) + o_p(1) \\
 &= (n-1) \frac{\hat{\sigma}_{0n}^2 - \hat{\sigma}_n^2}{\sigma_0^2} + (n-1) \left(\frac{1 - \exp(-2\hat{\lambda}_{0n} d)}{1 - \exp(-2\hat{\lambda}_n d)} - \frac{\hat{\lambda}_{0n}}{\hat{\lambda}_n}\right) + o_p(1) \\
 &= (n-1) \frac{\hat{\sigma}_{0n}^2 - \hat{\sigma}_n^2}{\sigma_0^2} + o_p(1).
 \end{aligned} \tag{4.7}$$

Large values of $\tilde{d}(n)$ suggest rejection of the null hypothesis.

Theorem 4.1 *Suppose that conditions (A1)-(A2) hold. If H_0 holds, then*

$$\tilde{d}(n) \rightarrow_D \chi^2(m), \quad n \rightarrow \infty. \tag{4.8}$$

5 Some lemmas

Throughout this paper, let C denote a generic positive constant which could take different value at each occurrence. To prove our main results, we first introduce the following lemmas.

Lemma 5.1 *If condition (A1) holds, then for any $\lambda \in R^+$ the matrix $X_n(\lambda)$ is positive definite for large enough n , and*

$$\lim_{n \rightarrow \infty} \max_{1 \leq t \leq n} x_t^T X_n^{-1}(\lambda) x_t = 0.$$

Proof Let $\tilde{\lambda}_1$ and $\tilde{\lambda}_m$ be the smallest and largest roots of $|Z_n - \tilde{\lambda} X_n| = 0$. Then from Ex. 22.1 of Rao [54],

$$\tilde{\lambda}_1 \leq \frac{u^T Z_n u}{u^T X_n u} \leq \tilde{\lambda}_m$$

for unit vectors u . Thus by (2.18) there are some $\delta \in (0, 1)$ and $n_0(\delta)$ such that $n \geq N_0$ implies

$$|u^T Z_n u| \leq (1 - \delta) u^T X_n u. \tag{5.1}$$

By (2.16) and (5.1), we have

$$\begin{aligned} u^T X_n(\lambda) u &= \sum_{t=2}^n (u^T (x_t - \exp(-\lambda d) x_{t-1}))^2 \\ &\geq \sum_{t=2}^n (u^T x_t)^2 + \min_{\lambda} \exp(-2\lambda d) \sum_{t=2}^n (u^T x_{t-1})^2 \\ &\quad - \max_{\lambda} \exp(-\lambda d) u^T Z_n u \\ &\geq u^T X_n u + \min_{\lambda} \exp(-2\lambda d) u^T X_n u - u^T Z_n u \\ &\geq \left(1 + \min_{\lambda} \exp(-2\lambda d) - (1 - \delta)\right) u^T X_n u \\ &= \left(\min_{\lambda} \exp(-2\lambda d) + \delta\right) u^T X_n u = C(\lambda, \delta) u^T X_n u. \end{aligned} \tag{5.2}$$

By Rao [54, p.60] and (2.17), we have

$$\frac{(u^T x_t)^2}{u^T X_n u} \rightarrow 0. \tag{5.3}$$

From (5.3) and $C(\lambda, \delta) > 0$,

$$x_t^T X_n^{-1}(\lambda) x_t = \sup_u \left(\frac{(u^T x_t)^2}{u^T X_n(\lambda) u} \right) \leq \sup_u \left(\frac{(u^T x_t)^2}{C(\lambda, \delta) u^T X_n u} \right) \rightarrow 0. \tag{5.4}$$

□

Lemma 5.2 *The matrix D_n is positive definite for large enough n , $E(S_n(\theta_0)) = 0$ and $\text{Var}(S_n(\theta_0)) = \sigma_0^2 D_n$.*

Proof Note that $X_n(\lambda_0)$ is positive definite and $\Delta_n(\theta_0, \sigma_0) > 0$. It is easy to show that the matrix D_n is positive definite for large enough n . By (2.8), we have

$$\begin{aligned} \sigma_0^2 E \left(\frac{\partial \Psi_n}{\partial \beta} \Big|_{\theta=\theta_0} \right) &= \frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)} \\ &\quad \cdot \sum_{t=2}^n E(e_t - \exp(-\lambda_0 d) e_{t-1}) (x_t - \exp(-\lambda_0 d) x_{t-1}) \\ &= \frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)} \sigma_0 \sqrt{\frac{1 - \exp(2d\lambda_0)}{2\lambda_0}} \\ &\quad \cdot \sum_{t=2}^n (x_t - \exp(-\lambda_0 d) x_{t-1}) E \eta_t \\ &= 0. \end{aligned} \tag{5.5}$$

Note that e_{t-1} and η_t are independent, so we have $E(\eta_t e_{t-1}) = 0$. Thus, by (2.7) and $E\eta_t = 0$, we have

$$\begin{aligned} E\left(\frac{\partial \Psi_n}{\partial \lambda} \Big|_{\theta=\theta_0}\right) &= \frac{n-1}{2\lambda_0} - \frac{(n-1)d \exp(-2\lambda_0 d)}{1 - \exp(-2\lambda_0 d)} - 0 \\ &\quad - \frac{1 - (1 + 2d\lambda_0) \exp(-2\lambda_0 d)}{\sigma_0^2 (1 - \exp(-2\lambda_0 d))^2} \sigma_0^2 \frac{1 - \exp(2d\lambda_0)}{2\lambda_0} \sum_{t=2}^n E\eta_t^2 \\ &= \frac{n-1}{2\lambda_0} - \frac{(n-1)d \exp(-2\lambda_0 d)}{1 - \exp(-2\lambda_0 d)} - \frac{1 - (1 + 2d\lambda_0) \exp(-2\lambda_0 d)}{2\lambda_0 (1 - \exp(-2\lambda_0 d))} (n-1) \\ &= 0. \end{aligned} \tag{5.6}$$

Hence, from (5.5) and (5.6),

$$E(S_n(\theta_0)) = \sigma_0^2 E\left(\frac{\partial \Psi_n}{\partial \beta} \Big|_{\theta=\theta_0}, \frac{\partial \Psi_n}{\partial \lambda} \Big|_{\theta=\theta_0}\right) = 0. \tag{5.7}$$

By (2.8) and (2.20), we have

$$\begin{aligned} \text{Var}\left(\sigma_0^2 \frac{\partial \Psi_n}{\partial \beta} \Big|_{\theta=\theta_0}\right) &= \text{Var}\left\{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)} \sum_{t=2}^n (e_t - \exp(-\lambda_0 d)e_{t-1})(x_t - \exp(-\lambda_0 d)x_{t-1})\right\} \\ &= \frac{2\sigma_0^2 \lambda_0}{1 - \exp(-2\lambda_0 d)} \text{Var}\left\{\sum_{t=2}^n (x_t - \exp(-\lambda_0 d)x_{t-1})\eta_t\right\} \\ &= \frac{2\sigma_0^2 \lambda_0}{1 - \exp(-2\lambda_0 d)} X_n(\lambda_0). \end{aligned} \tag{5.8}$$

Note that $\{\eta_t e_{t-1}, H_t\}$ is a martingale difference sequence with

$$\text{Var}(\eta_t e_{t-1}) = E\eta_t^2 Ee_{t-1}^2 = Ee_{t-1}^2,$$

so

$$\begin{aligned} \text{Var}\left(\sigma_0^2 \frac{\partial \Psi_n}{\partial \lambda} \Big|_{\theta=\theta_0}\right) &= E\left\{\sigma_0 d \exp(-\lambda_0 d) \sqrt{\frac{2\lambda_0}{1 - \exp(-\lambda_0 d)}} \sum_{t=2}^n \eta_t e_{t-1}\right\}^2 \\ &\quad + E\left\{\frac{\sigma_0^2 [1 - (1 + 2d\lambda_0) \exp(-2d\lambda_0)]}{2\lambda_0 (1 - \exp(-2\lambda_0 d))} \sum_{t=2}^n (\eta_t^2 - 1)\right\}^2 \\ &\quad + \frac{\sqrt{2}\sigma_0^3 d \exp(-\lambda_0 d) [1 - (1 + 2d\lambda_0) \exp(-2d\lambda_0)]}{\sqrt{\lambda_0} (1 - \exp(-2\lambda_0 d))^{\frac{3}{2}}} \\ &\quad \cdot E\left\{\sum_{t=2}^n \eta_t e_{t-1} \sum_{t=2}^n (\eta_t^2 - 1)\right\} \\ &= \frac{2\lambda_0 \sigma_0^2 d^2 \exp(-2\lambda_0 d)}{1 - \exp(-\lambda_0 d)} \sum_{t=2}^n Ee_{t-1}^2 \end{aligned}$$

$$\begin{aligned}
 & + \left\{ \frac{\sigma_0^2 [1 - (1 + 2d\lambda_0) \exp(-2d\lambda_0)]}{2\lambda_0(1 - \exp(-2\lambda_0d))} \right\}^2 (n-1)(E\eta_t^4 - 1) \\
 & + \frac{\sqrt{2}\sigma_0^3 d \exp(-\lambda_0d) [1 - (1 + 2d\lambda_0) \exp(-2d\lambda_0)]}{\sqrt{\lambda_0}(1 - \exp(-2\lambda_0d))^{\frac{3}{2}}} \\
 & \cdot \left(\sum_{t=2}^n E((\eta_t^2 - 1)\eta_t e_{t-1}) + \sum_{t \neq k} E(\eta_t e_{t-1}(\eta_k^2 - 1)) \right) \\
 & = \frac{2\lambda_0 \sigma_0^2 d^2 \exp(-2\lambda_0d)}{1 - \exp(-\lambda_0d)} \sum_{t=2}^n E e_{t-1}^2 \\
 & + 2(n-1) \left\{ \frac{\sigma_0^2 [1 - (1 + 2d\lambda_0) \exp(-2d\lambda_0)]}{2\lambda_0(1 - \exp(-2\lambda_0d))} \right\}^2 \\
 & = \sigma_0^2 \Delta_n(\theta_0, \sigma_0). \tag{5.9}
 \end{aligned}$$

By (2.7), (2.8), and noting that e_{t-1} and η_t are independent, we have

$$\begin{aligned}
 \text{Cov} \left(\sigma_0^2 \frac{\partial \Psi_n}{\partial \beta}, \sigma_0^2 \frac{\partial \Psi_n}{\partial \lambda} \right) \Big|_{\theta=\theta_0} & = -\sigma_0^3 \frac{1 - (1 + 2d\lambda) \exp(-2\lambda d)}{\sqrt{2\lambda_0}(1 - \exp(-2\lambda d))^{\frac{3}{2}}} \\
 & \cdot E \left(\sum_{t=2}^n \eta_t^2 \sum_{t=2}^n \eta_t (x_t - \exp(-\lambda d)x_{t-1}) \right) \\
 & = -\sigma_0^3 \frac{1 - (1 + 2d\lambda) \exp(-2\lambda d)}{\sqrt{2\lambda_0}(1 - \exp(-2\lambda d))^{\frac{3}{2}}} E\eta_t^3 \\
 & \cdot \sum_{t=2}^n (x_t - \exp(-\lambda d)x_{t-1}) = 0. \tag{5.10}
 \end{aligned}$$

From (5.8)-(5.10), it follows that $\text{Var}(S_n(\theta_0)) = \sigma_0^2 D_n$. The proof is completed. \square

Lemma 5.3 (Maller [55]) *Let W_n be a symmetric random matrix with eigenvalues $\tilde{\lambda}_j(n)$, $1 \leq j \leq d$. Then*

$$W_n \rightarrow_p I \iff \tilde{\lambda}_j(n) \rightarrow_p 1, \quad n \rightarrow \infty.$$

Lemma 5.4 *For each $A > 0$,*

$$\sup_{\theta \in N_n(A)} \| D_n^{-\frac{1}{2}} F_n(\theta) D_n^{-\frac{T}{2}} - \Phi_n \| \rightarrow_p 0, \quad n \rightarrow \infty \tag{5.11}$$

and also

$$\Phi_n \rightarrow_D \Phi, \tag{5.12}$$

$$\lim_{c \rightarrow 0} \limsup_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left\{ \inf_{\theta \in N_n(A)} \lambda_{\min} (D_n^{-\frac{1}{2}} F_n(\theta) D_n^{-\frac{T}{2}}) \leq c \right\} = 0, \tag{5.13}$$

where

$$\Phi_n = \begin{pmatrix} \frac{\lambda(1 - \exp(-2d\lambda_0))}{\lambda_0(1 - \exp(-2d\lambda))} I_m & 0 \\ 0 & \frac{-\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} |_{\theta=\theta_0}}{\Delta_n(\theta_0, \sigma_0)} \end{pmatrix}, \quad \Phi = I_{m+1}. \tag{5.14}$$

Proof Let $X_n(\lambda_0) = X_n^{\frac{1}{2}}(\lambda_0)X_n^{\frac{T}{2}}(\lambda_0)$ be a square root decomposition of $X_n(\lambda_0)$. Then

$$\begin{aligned}
 D_n &= \begin{pmatrix} \sqrt{\frac{2\lambda_0}{1-\exp(-2d\lambda_0)}}X_n^{\frac{1}{2}}(\lambda_0) & 0 \\ 0 & \sqrt{\Delta_n(\theta_0, \sigma_0)} \end{pmatrix} \\
 &\cdot \begin{pmatrix} \sqrt{\frac{2\lambda_0}{1-\exp(-2d\lambda_0)}}X_n^{\frac{T}{2}}(\lambda_0) & 0 \\ 0 & \sqrt{\Delta_n(\theta_0, \sigma_0)} \end{pmatrix} \\
 &= D_n^{\frac{1}{2}}D_n^{\frac{T}{2}}.
 \end{aligned} \tag{5.15}$$

Let $\theta \in N_n(A)$. Then

$$\begin{aligned}
 (\theta - \theta_0)^T D_n (\theta - \theta_0) &= \frac{2\lambda_0}{1 - \exp(-2d\lambda_0)} (\beta - \beta_0)^T X_n(\lambda_0) (\beta - \beta_0) \\
 &\quad + (\lambda - \lambda_0)^2 \Delta_n(\theta_0, \sigma_0) \leq A^2.
 \end{aligned} \tag{5.16}$$

From (2.20), (2.21) and (5.14),

$$D_n^{-\frac{1}{2}} F_n(\theta) D_n^{-\frac{T}{2}} - \Phi_n = \begin{pmatrix} W_{11} & W_{12} \\ & W_{22} \end{pmatrix}, \tag{5.17}$$

where

$$W_{11} = \frac{\lambda(1 - \exp(-2d\lambda_0))}{\lambda_0(1 - \exp(-2d\lambda))} \{X_n^{-\frac{1}{2}}(\lambda_0)X_n(\lambda)X_n^{-\frac{T}{2}}(\lambda_0) - I_m\}, \tag{5.18}$$

$$W_{12} = \frac{\sqrt{\frac{1-\exp(-2d\lambda_0)}{2\lambda_0}}X_n^{-\frac{1}{2}}(\lambda_0)(-\sigma^2 \frac{\partial^2 \Psi_n}{\partial \beta \partial \lambda})}{\sqrt{\Delta_n(\theta_0, \sigma_0)}} \tag{5.19}$$

and

$$W_{22} = \frac{-\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} - \sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} |_{\theta=\theta_0}}{\Delta_n(\theta_0, \sigma_0)}. \tag{5.20}$$

Let

$$N_n^\beta(A) = \left\{ \beta : \frac{2\lambda_0}{1 - \exp(-2d\lambda_0)} |(\beta - \beta_0)^T X_n^{\frac{1}{2}}(\lambda_0)|^2 \leq A^2 \right\} \tag{5.21}$$

and

$$N_n^\lambda(A) = \left\{ \theta : |\lambda - \lambda_0| \leq \frac{A}{\sqrt{\Delta_n(\theta_0, \sigma_0)}} \right\}. \tag{5.22}$$

As the first step, we will show that, for each $A > 0$,

$$\sup_{\theta \in N_n^\theta(A)} \|W_{11}\| \rightarrow 0, \quad n \rightarrow \infty. \tag{5.23}$$

In fact, note that

$$\begin{aligned} W_{11} &= \frac{\lambda(1 - \exp(-2d\lambda_0))}{\lambda_0(1 - \exp(-2d\lambda))} X_n^{-\frac{1}{2}}(\lambda_0)(X_n(\lambda) - X_n(\lambda_0))X_n^{-\frac{T}{2}}(\lambda_0) \\ &= \frac{\lambda(1 - \exp(-2d\lambda_0))}{\lambda_0(1 - \exp(-2d\lambda))} X_n^{-\frac{1}{2}}(\lambda_0)(T_1 + T_2 - T_3)X_n^{-\frac{T}{2}}(\lambda_0), \end{aligned} \tag{5.24}$$

where

$$\begin{aligned} T_1 &= \sum_{t=2}^n (\exp(-d\lambda_0) - \exp(-d\lambda))x_{t-1}(x_t - \exp(-d\lambda_0)x_{t-1})^T, \\ T_2 &= \sum_{t=2}^n (\exp(-d\lambda_0) - \exp(-d\lambda))(x_t - \exp(-d\lambda_0)x_{t-1})x_t^T \end{aligned}$$

and

$$T_3 = \sum_{t=2}^n (\exp(-d\lambda) - \exp(-d\lambda_0))^2 x_{t-1}x_{t-1}^T.$$

Let $u, v \in R^d$, $|u| = |v| = 1$, and let $u_n^T = u^T X_n^{-\frac{1}{2}}(\lambda_0)$, $v_n^T = X_n^{-\frac{T}{2}}(\lambda_0)v$. By the Cauchy-Schwarz inequality, Lemma 5.1 and noting $N_n^\lambda(A)$, we have

$$\begin{aligned} |u_n^T T_1 v_n| &= \left| (\exp(-d\lambda_0) - \exp(-d\lambda)) \sum_{t=2}^n u_n^T x_{t-1} (x_t - \exp(-d\lambda_0)x_{t-1})^T v_n \right| \\ &\leq \max |\exp(-d\lambda_0) - \exp(-d\lambda)| \left(\sum_{t=2}^n u_n^T x_t x_t^T u_n \right)^{\frac{1}{2}} \\ &\quad \cdot \left(\sum_{t=2}^n v_n^T (x_t - \exp(-d\lambda_0)x_{t-1})(x_t - \exp(-d\lambda_0)x_{t-1})^T v_n \right)^{\frac{1}{2}} \\ &\leq d|\lambda_0 - \lambda| \cdot \sqrt{n} \max_{1 \leq t \leq n} (x_t^T X_n^{-1}(\lambda_0)x_t) \cdot 1 \\ &\leq C \sqrt{\frac{n}{\Delta_n(\theta_0, \sigma_0)}} o(1) \rightarrow 0. \end{aligned} \tag{5.25}$$

Similar to the proof of T_1 , we easily obtain

$$|u_n^T T_2 v_n| \rightarrow 0. \tag{5.26}$$

By the Cauchy-Schwarz inequality, Lemma 5.1 and noting $N_n^\lambda(A)$, we have

$$\begin{aligned} |u_n^T T_3 v_n| &= \left| u_n^T \sum_{t=2}^n (\exp(-d\lambda_0) - \exp(-d\lambda))^2 x_{t-1}x_{t-1}^T v_n \right| \\ &\leq \max |\exp(-d\lambda_0) - \exp(-d\lambda)|^2 \\ &\quad \cdot \left(\sum_{t=2}^n u_n^T x_t x_t^T u_n \sum_{t=2}^n v_n^T x_t x_t^T v_n \right)^{\frac{1}{2}} \end{aligned}$$

$$\begin{aligned} &\leq n|\lambda_0 - \lambda|^2 \max_{1 \leq t \leq n} (x_t^T X_n^{-1}(\lambda_0) x_t) \\ &\leq \frac{nA^2}{\Delta_n(\theta_0, \sigma_0)} o(1) \rightarrow 0. \end{aligned} \tag{5.27}$$

Hence, (5.23) follows from (5.24)-(5.27).

For the second step, we will show that

$$W_{12} \rightarrow_p 0. \tag{5.28}$$

Note that

$$\varepsilon_t = y_t - x_t^T \beta = x_t^T (\beta_0 - \beta) + e_t \tag{5.29}$$

and

$$\varepsilon_t - \exp(-d\lambda_0)\varepsilon_{t-1} = (x_t - \exp(-d\lambda_0)x_{t-1})^T (\beta_0 - \beta) + \sigma_0 \sqrt{\frac{1 - \exp(-2d\lambda_0)}{2\lambda_0}} \eta_t. \tag{5.30}$$

Write

$$\begin{aligned} J &= \sqrt{\frac{1 - \exp(-2d\lambda_0)}{2\lambda_0}} X_n^{-\frac{1}{2}}(\lambda_0) \left(-\sigma^2 \frac{\partial^2 \Psi_n}{\partial \beta \partial \lambda} \right) \\ &= -\sqrt{\frac{1 - \exp(-2d\lambda_0)}{2\lambda_0}} X_n^{-\frac{1}{2}}(\lambda_0) \frac{2d\lambda \exp(-\lambda d)}{1 - \exp(-2\lambda d)} \\ &\quad \cdot \sum_{t=2}^n (\varepsilon_{t-1} x_t + \varepsilon_t x_{t-1} - 2 \exp(-\lambda d) x_{t-1} \varepsilon_{t-1}) \\ &\quad - \sqrt{\frac{1 - \exp(-2d\lambda_0)}{2\lambda_0}} X_n^{-\frac{1}{2}}(\lambda_0) \frac{1 - (1 + 2d\lambda) \exp(-2\lambda d)}{(1 - \exp(-2\lambda d))^2} \\ &\quad \cdot \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d) \varepsilon_{t-1}) (x_t - \exp(-\lambda d) x_{t-1}) \\ &= -\sqrt{\frac{1 - \exp(-2d\lambda_0)}{2\lambda_0}} \frac{2d\lambda \exp(-\lambda d)}{1 - \exp(-2\lambda d)} X_n^{-\frac{1}{2}}(\lambda_0) (T_1 + T_2 + 2T_3 + 2T_4 + 2T_5) \\ &\quad - \sqrt{\frac{1 - \exp(-2d\lambda_0)}{2\lambda_0}} \frac{1 - (1 + 2d\lambda) \exp(-2\lambda d)}{(1 - \exp(-2\lambda d))^2} X_n^{-\frac{1}{2}}(\lambda_0) T_6, \end{aligned} \tag{5.31}$$

where

$$\begin{aligned} T_1 &= \sum_{t=2}^n x_{t-1}^T (\beta_0 - \beta) (x_t - \exp(-\lambda_0 d) x_{t-1}), \\ T_2 &= \sum_{t=2}^n (x_t - \exp(-\lambda_0 d) x_{t-1})^T (\beta_0 - \beta) x_{t-1}, \\ T_3 &= \sum_{t=2}^n (\exp(-\lambda_0 d) - \exp(-\lambda d)) x_{t-1}^T (\beta_0 - \beta) x_{t-1}, \end{aligned}$$

$$T_4 = \sigma \sqrt{\frac{1 - \exp(-2\lambda d)}{2\lambda}} \sum_{t=2}^n \eta_t x_{t-1}, \quad T_5 = \sum_{t=2}^n e_{t-1} x_{t-1},$$

$$T_6 = \sigma \sqrt{\frac{1 - \exp(-2\lambda d)}{2\lambda}} \sum_{t=2}^n \eta_t (x_t - \exp(-\lambda d) x_{t-1}).$$

For $\beta \in N_n^\beta(A)$ and each $A > 0$, we have

$$\begin{aligned} |(\beta_0 - \beta)^T x_t|^2 &= (\beta_0 - \beta)^T X_n^{\frac{1}{2}} (\lambda_0) X_n^{-\frac{1}{2}} (\lambda_0) x_t x_t^T X_n^{-\frac{T}{2}} (\lambda_0) X_n^{\frac{T}{2}} (\lambda_0) (\beta_0 - \beta) \\ &\leq \max_{1 \leq t \leq n} (x_t^T X_n^{-1} (\lambda_0) x_t) (\beta_0 - \beta)^T X_n (\lambda_0) (\beta_0 - \beta) \\ &\leq A^2 \max_{1 \leq t \leq n} (x_t^T X_n^{-1} (\lambda_0) x_t). \end{aligned} \tag{5.32}$$

By (5.32) and Lemma 5.1, we have

$$\sup_{\beta \in N_n^\beta(A)} \max_{1 \leq t \leq n} |(\beta_0 - \beta)^T x_t| \rightarrow 0, \quad n \rightarrow \infty, A > 0. \tag{5.33}$$

Using the Cauchy-Schwarz inequality and (5.33), we obtain

$$\begin{aligned} u_n^T T_1 &= \sum_{t=2}^n u_n^T x_{t-1}^T (\beta_0 - \beta) (x_t - \exp(-\lambda_0 d) x_{t-1}) \\ &\leq \left\{ \sum_{t=2}^n (x_{t-1}^T (\beta_0 - \beta))^2 \right\}^{\frac{1}{2}} \\ &\quad \cdot \left\{ \sum_{t=2}^n u_n^T (x_t - \exp(-\lambda_0 d) x_{t-1}) (x_t - \exp(-\lambda_0 d) x_{t-1})^T u_n \right\}^{\frac{1}{2}} \\ &\leq \sqrt{n} \max_{1 \leq t \leq n} |(\beta_0 - \beta)^T x_t| = o(\sqrt{n}). \end{aligned} \tag{5.34}$$

Using a similar argument as T_1 , we obtain that

$$u_n^T T_2 = o_p(\sqrt{n}). \tag{5.35}$$

By the Cauchy-Schwarz inequality and (5.33), (5.25), we get

$$\begin{aligned} u_n^T T_3 &= \sum_{t=2}^n (\exp(-\lambda_0 d) - \exp(-\lambda d)) x_{t-1}^T (\beta_0 - \beta) u_n^T x_{t-1} \\ &\leq \left\{ \sum_{t=2}^n (\exp(-\lambda_0 d) - \exp(-\lambda d))^2 (x_{t-1}^T (\beta_0 - \beta))^2 \sum_{t=2}^n (u_n^T x_{t-1})^2 \right\}^{\frac{1}{2}} \\ &\leq C |\lambda_0 - \lambda| \left\{ \sum_{t=2}^n (x_{t-1}^T (\beta_0 - \beta))^2 \sum_{t=2}^n (u_n^T x_{t-1})^2 \right\}^{\frac{1}{2}} \\ &\leq C \frac{A}{\sqrt{\Delta_n(\theta_0, \sigma_0)}} \sqrt{n} o(1) o(\sqrt{n}) = o(\sqrt{n}). \end{aligned} \tag{5.36}$$

By (5.25), we have

$$\text{Var}(u_n^T T_4) = \sigma^2 \frac{1 - \exp(-2\lambda d)}{2\lambda} \sum_{t=2}^n (u_n^T x_{t-1})^2 = o(n). \tag{5.37}$$

Thus, by the Chebychev inequality and (5.37),

$$u_n^T T_4 = o_p(\sqrt{n}). \tag{5.38}$$

By Lemma 5.1 and (2.3), we have

$$\begin{aligned} \text{Var}(u_n^T T_5) &= \text{Var}\left(\sum_{t=2}^n u_n^T x_t e_{t-1}\right) \\ &= \sigma_0^2 \frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0} \\ &\quad \cdot \text{Var}\left\{\sum_{j=1}^{n-1} \left(\sum_{t=j+1}^n u_n^T x_t \exp\{-\lambda_0 d(t-1-j)\}\right) \eta_j\right\} \\ &= \sigma_0^2 \frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0} \sum_{j=1}^{n-1} \left(\sum_{t=j+1}^n u_n^T x_t \exp\{-\lambda_0 d(t-1-j)\}\right)^2 \\ &\leq \sigma_0^2 \frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0} \max_{2 \leq t \leq n} |u_n^T x_t| \\ &\quad \cdot \sum_{j=1}^{n-1} \left(\sum_{t=j+1}^n \exp\{-\lambda_0 d(t-1-j)\}\right)^2 \\ &\leq C \max_{2 \leq t \leq n} |u_n^T x_t| n = o(n). \end{aligned} \tag{5.39}$$

Thus, by the Chebychev inequality and (5.39),

$$u_n^T T_5 = o_p(\sqrt{n}). \tag{5.40}$$

Using a similar argument as T_4 , we obtain

$$u_n^T T_6 = o_p(\sqrt{n}). \tag{5.41}$$

Thus (5.28) follows immediately from (5.31), (5.34)-(5.36), (5.38), (5.40) and (5.41).

For the third step, we will show that

$$W_{22} \rightarrow_p \mathbf{0}. \tag{5.42}$$

Write that

$$\begin{aligned} J &= -\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} - \sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} \Big|_{\theta=\theta_0} \\ &= \frac{\sigma^2(n-1)}{2\lambda^2} - \frac{2\sigma^2(n-1)d^2 \exp(-2\lambda d)}{(1 - \exp(-2\lambda d))^2} + \frac{2d^2 \lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} \sum_{t=2}^n \varepsilon_{t-1}^2 \end{aligned}$$

$$\begin{aligned}
 & + \frac{2d \exp(-\lambda d)[(2-d\lambda) - d\lambda \exp(-2\lambda d)]}{(1 - \exp(-2\lambda d))^2} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})\varepsilon_{t-1} \\
 & + \frac{4d \exp(-2\lambda d)[d\lambda - 1 + (1+d\lambda) \exp(-2\lambda d)]}{(1 - \exp(-2\lambda d))^3} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})^2 \\
 & - \frac{\sigma_0^2(n-1)}{2\lambda_0^2} + \frac{2\sigma_0^2(n-1)d^2 \exp(-2\lambda_0 d)}{(1 - \exp(-2\lambda_0 d))^2} - \frac{2d^2\lambda_0 \exp(-2\lambda_0 d)}{1 - \exp(-2\lambda_0 d)} \sum_{t=2}^n e_{t-1}^2 \\
 & - \frac{2d \exp(-\lambda_0 d)[(2-d\lambda_0) - d\lambda_0 \exp(-2\lambda_0 d)]}{(1 - \exp(-2\lambda_0 d))^2} \sum_{t=2}^n (e_t - \exp(-\lambda_0 d)e_{t-1})e_{t-1} \\
 & - \frac{4d \exp(-2\lambda_0 d)[d\lambda_0 - 1 + (1+d\lambda_0) \exp(-2\lambda_0 d)]}{(1 - \exp(-2\lambda_0 d))^3} \\
 & \cdot \sum_{t=2}^n (e_t - \exp(-\lambda_0 d)e_{t-1})^2. \tag{5.43}
 \end{aligned}$$

By (3.3) and (3.4), we obtain that

$$\begin{aligned}
 T_1 &= \frac{\sigma^2(n-1)}{2\lambda^2} - \frac{\sigma_0^2(n-1)}{2\lambda_0^2} \\
 &= \frac{n-1}{2\lambda^2\lambda_0^2} (\sigma^2(\lambda_0^2 - \lambda^2) + \lambda^2(\sigma^2 - \sigma_0^2)) \\
 &= o(n) \tag{5.44}
 \end{aligned}$$

and

$$\begin{aligned}
 T_2 &= \frac{2\sigma_0^2(n-1)d^2 \exp(-2\lambda_0 d)}{(1 - \exp(-2\lambda_0 d))^2} - \frac{2\sigma^2(n-1)d^2 \exp(-2\lambda d)}{(1 - \exp(-2\lambda d))^2} \\
 &= \frac{2d^2(n-1)}{(1 - \exp(-2\lambda_0 d))^2(1 - \exp(-2\lambda d))^2} \{ \sigma_0(\exp(-\lambda_0 d) - \exp(-\lambda d)) \\
 & \quad + \exp(-\lambda d)(\sigma_0 - \sigma) + \exp(-\lambda d - \lambda_0 d) \\
 & \quad \cdot [\sigma(\exp(-\lambda_0 d) - \exp(-\lambda d)) + \exp(-\lambda d)(\sigma - \sigma_0)] \} \\
 & \quad \cdot (\sigma_0 \exp(-\lambda_0 d)(1 - \exp(-2\lambda d)) + \sigma \exp(-\lambda d)(1 - \exp(-2\lambda_0 d))) \\
 &= o(n). \tag{5.45}
 \end{aligned}$$

By (5.29), we have

$$\begin{aligned}
 T_3 &= \frac{2d^2\lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} \sum_{t=2}^n \varepsilon_{t-1}^2 - \frac{2d^2\lambda_0 \exp(-2\lambda_0 d)}{1 - \exp(-2\lambda_0 d)} \sum_{t=2}^n e_{t-1}^2 \\
 &= \frac{2d^2\lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} \sum_{t=2}^n \{ (x_t^T(\beta_0 - \beta))^2 + 2x_t^T(\beta_0 - \beta)e_t + e_t^2 \} \\
 & \quad - \frac{2d^2\lambda_0 \exp(-2\lambda_0 d)}{1 - \exp(-2\lambda_0 d)} \sum_{t=2}^n e_{t-1}^2 \\
 &= \frac{2d^2\lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} \sum_{t=2}^n (x_t^T(\beta_0 - \beta))^2 + \frac{2d^2\lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} \sum_{t=2}^n 2x_t^T(\beta_0 - \beta)e_t
 \end{aligned}$$

$$\begin{aligned}
 & + \left\{ \frac{2d^2\lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} - \frac{2d^2\lambda_0 \exp(-2\lambda_0 d)}{1 - \exp(-2\lambda_0 d)} \right\} \sum_{t=2}^n e_{t-1}^2 \\
 & = \frac{2d^2\lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} T_{31} + \frac{4d^2\lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} T_{32} + T_{33}.
 \end{aligned} \tag{5.46}$$

By (5.32), it is easy to show that

$$T_{31} = o(n). \tag{5.47}$$

By Lemma 5.1, (2.3) and (5.32), we have

$$\begin{aligned}
 \text{Var}(T_{32}) & = \text{Var} \left(\sum_{t=2}^n x_t^T (\beta_0 - \beta) e_t \right) \\
 & = \text{Var} \left\{ \sum_{j=1}^{n-1} \left(\sum_{t=j+1}^n x_t^T (\beta_0 - \beta) \exp\{-\lambda_0 d(t-1-j)\} \right) \eta_j \right\} \\
 & = \sum_{j=1}^{n-1} \left(\sum_{t=j+1}^n x_t^T (\beta_0 - \beta) \exp\{-\lambda_0 d(t-1-j)\} \right)^2 \\
 & \leq \max_{2 \leq t \leq n} |x_t^T (\beta_0 - \beta)| \sum_{j=1}^{n-1} \left(\sum_{t=j+1}^n \exp\{-\lambda_0 d(t-1-j)\} \right)^2 \\
 & \leq C \max_{2 \leq t \leq n} |x_t^T (\beta_0 - \beta)| n = o(n).
 \end{aligned} \tag{5.48}$$

Thus by the Chebychev inequality and (5.48),

$$T_{32} = o_p(\sqrt{n}). \tag{5.49}$$

Write

$$\begin{aligned}
 & \frac{2d^2\lambda \exp(-2\lambda d)}{1 - \exp(-2\lambda d)} - \frac{2d^2\lambda_0 \exp(-2\lambda_0 d)}{1 - \exp(-2\lambda_0 d)} \\
 & = \frac{2d^2}{(1 - \exp(-2\lambda d))(1 - \exp(-2\lambda_0 d))} U,
 \end{aligned} \tag{5.50}$$

where

$$U = \lambda \exp(-2\lambda d)(1 - \exp(-2\lambda_0 d)) - \lambda_0 \exp(-2\lambda_0 d)(1 - \exp(-2\lambda d)).$$

Note that

$$\begin{aligned}
 U & = \lambda \exp(-2\lambda d)(\exp(-2\lambda d) - \exp(-2\lambda_0 d)) \\
 & \quad + (\lambda(\exp(-2\lambda d) - \exp(-2\lambda_0 d)) + (\lambda - \lambda_0) \exp(-2\lambda_0 d))(1 - \exp(-2\lambda d)) \\
 & = o(1),
 \end{aligned} \tag{5.51}$$

so we have

$$T_{33} = o(n). \tag{5.52}$$

Thus, by (5.46), (5.47), (5.49) and (5.52), we have

$$T_3 = o(n). \tag{5.53}$$

By (5.29), we have

$$\begin{aligned} T_4 &= \frac{2d \exp(-\lambda d)[(2-d\lambda) - d\lambda \exp(-2\lambda d)]}{(1-\exp(-2\lambda d))^2} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})\varepsilon_{t-1} \\ &\quad - \frac{2d \exp(-\lambda_0 d)[(2-d\lambda_0) - d\lambda_0 \exp(-2\lambda_0 d)]}{(1-\exp(-2\lambda_0 d))^2} \sum_{t=2}^n (e_t - \exp(-\lambda_0 d)e_{t-1})e_{t-1} \\ &= \frac{2d \exp(-\lambda d)[(2-d\lambda) - d\lambda \exp(-2\lambda d)]}{(1-\exp(-2\lambda d))^2} \sigma \sqrt{\frac{1-\exp(-2\lambda d)}{2\lambda}} \sum_{t=2}^n x_{t-1}^T (\beta_0 - \beta) \eta_t \\ &\quad + \left\{ \frac{2d \exp(-\lambda d)[(2-d\lambda) - d\lambda \exp(-2\lambda d)]}{(1-\exp(-2\lambda d))^2} \sigma \sqrt{\frac{1-\exp(-2\lambda d)}{2\lambda}} \right. \\ &\quad \left. - \frac{2d \exp(-\lambda_0 d)[(2-d\lambda_0) - d\lambda_0 \exp(-2\lambda_0 d)]}{(1-\exp(-2\lambda_0 d))^2} \right. \\ &\quad \left. \cdot \sigma \sqrt{\frac{1-\exp(-2\lambda d)}{2\lambda}} \right\} \sum_{t=2}^n \eta_t e_{t-1} \\ &= T_{41} + T_{42}. \end{aligned} \tag{5.54}$$

It is easy to show that

$$T_{41} = o(n). \tag{5.55}$$

Note that $\{\eta_t e_{t-1}, H_t\}$ is a martingale difference sequence, so we have

$$\text{Var} \left(\sum_{t=2}^n \eta_t e_{t-1} \right) = \sum_{t=2}^n E e_{t-1}^2 = \Delta_n(\theta_0, \sigma_0).$$

Hence,

$$T_{42} = o(n). \tag{5.56}$$

By (5.54)-(5.56), we have

$$T_4 = o(n). \tag{5.57}$$

It is easily proved that

$$\begin{aligned} T_5 &= \frac{4d \exp(-2\lambda d)[d\lambda - 1 + (1+d\lambda) \exp(-2\lambda d)]}{(1-\exp(-2\lambda d))^3} \sum_{t=2}^n (\varepsilon_t - \exp(-\lambda d)\varepsilon_{t-1})^2 \\ &\quad - \frac{4d \exp(-2\lambda_0 d)[d\lambda_0 - 1 + (1+d\lambda_0) \exp(-2\lambda_0 d)]}{(1-\exp(-2\lambda_0 d))^3} \sum_{t=2}^n (e_t - \exp(-\lambda_0 d)e_{t-1})^2 \end{aligned}$$

$$\begin{aligned}
 &= \left\{ \frac{4d \exp(-2\lambda d)[d\lambda - 1 + (1 + d\lambda) \exp(-2\lambda d)]}{(1 - \exp(-2\lambda d))^3} \sigma \sqrt{\frac{1 - \exp(-2\lambda d)}{2\lambda}} \right. \\
 &\quad - \frac{4d \exp(-2\lambda_0 d)[d\lambda_0 - 1 + (1 + d\lambda_0) \exp(-2\lambda_0 d)]}{(1 - \exp(-2\lambda_0 d))^3} \\
 &\quad \left. \cdot \sigma_0 \sqrt{\frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0}} \right\} \sum_{t=2}^n \eta_t^2 = o(n). \tag{5.58}
 \end{aligned}$$

Hence, (5.42) follows immediately from (5.43)-(5.45), (5.53), (5.57) and (5.58). This completes the proof of (5.11) from (5.17), (5.23), (5.28) and (5.42).

It is well known that $\frac{\lambda(1-\exp(-2d\lambda_0))}{\lambda_0(1-\exp(-2d\lambda))} \rightarrow 1$ as $n \rightarrow \infty$. To prove (5.12), we need to show that

$$\frac{-\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} |_{\theta=\theta_0}}{\Delta_n(\theta_0, \sigma_0)} \rightarrow_p 1, \quad n \rightarrow \infty.$$

This follows immediately from (2.20) and the Markov inequality.

Finally, we will prove (5.13). By (5.11) and (5.12), we have

$$D_n^{-\frac{1}{2}} F(\theta) D_n^{-\frac{T}{2}} \rightarrow_p I_m, \quad n \rightarrow \infty \tag{5.59}$$

uniformly in $\theta \in N_n(A)$ for each $A > 0$. Thus, by Lemma 5.3,

$$\lambda_{\min}(D_n^{-\frac{1}{2}} F(\theta) D_n^{-\frac{T}{2}}) \rightarrow_p 1, \quad n \rightarrow \infty. \tag{5.60}$$

This implies (5.13). □

Lemma 5.5 (Hall and Heyde [56]) *Let $\{S_{ni}, F_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ be a zero-mean, square-integrable martingale array with differences X_{ni} , and let η^2 be an a.s. finite random variable. Suppose that $\sum_i E\{X_{ni}^2 I(|X_{ni}| > \varepsilon) | F_{n,i-1}\} \rightarrow_p 0$ for all $\varepsilon \rightarrow 0$, and $\sum_i E\{X_{ni}^2 | F_{n,i-1}\} \rightarrow_p \eta^2$. Then*

$$S_{nk_n} = \sum_i X_{ni} \rightarrow_D Z,$$

where the r.v. Z has the characteristic function $E\{\exp(-\frac{1}{2}\eta^2 t^2)\}$.

6 Proof of theorems

Proof of Theorem 3.1 Take $A > 0$, let

$$M_n(A) = \{\theta \in R^{m+1} : (\theta - \theta_0)^T D_n (\theta - \theta_0) = A^2\} \tag{6.1}$$

be the boundary of $N_n(A)$, and let $\theta \in M_n(A)$. Using (2.19) and the Taylor expansion, for each $\sigma^2 > 0$, we have

$$\begin{aligned}
 \Psi_n(\theta, \sigma^2) &= \Psi_n(\theta_0, \sigma^2) + (\theta - \theta_0)^T \frac{\partial \Psi_n(\theta_0, \sigma^2)}{\partial \theta} + \frac{1}{2} (\theta - \theta_0)^T \frac{\partial^2 \Psi_n(\theta_0, \sigma^2)}{\partial \theta \partial \theta^T} (\theta - \theta_0) \\
 &= \frac{1}{\sigma^2} \Psi_n(\theta_0, \sigma^2) + (\theta - \theta_0)^T S_n(\theta_0) - \frac{1}{2\sigma^2} (\theta - \theta_0)^T F_n(\tilde{\theta})(\theta - \theta_0), \tag{6.2}
 \end{aligned}$$

where $\tilde{\theta} = a\theta + (1 - a)\theta_0$ for some $0 \leq a \leq 1$.

Let $Q_n(\theta) = \frac{1}{2}(\theta - \theta_0)^T F_n(\tilde{\theta})(\theta - \theta_0)$ and $v_n(\theta) = \frac{1}{A} D_n^{-\frac{T}{2}}(\theta - \theta_0)$. Take $c > 0$ and $\theta \in M_n(A)$, and by (6.2), we obtain that

$$\begin{aligned}
 & P\{\Psi_n(\theta, \sigma^2) \geq \Psi_n(\theta_0, \sigma^2) \text{ for some } \theta \in M_n(A)\} \\
 & \leq P\{(\theta - \theta_0)^T S_n(\theta_0) \geq Q_n(\theta), Q_n(\theta) > cA^2 \text{ for some } \theta \in M_n(A)\} \\
 & \quad + P\{Q_n(\theta) \leq cA^2 \text{ for some } \theta \in M_n(A)\} \\
 & \leq P\{v_n^T(\theta) D_n^{-\frac{1}{2}} S_n(\theta_0) > cA \text{ for some } \theta \in M_n(A)\} \\
 & \quad + P\{v_n^T(\theta) D_n^{-\frac{1}{2}} F_n(\tilde{\theta}) D_n^{-\frac{T}{2}} v_n(\theta) \leq c \text{ for some } \theta \in M_n(A)\} \\
 & \leq P\{|D_n^{-\frac{1}{2}} S_n(\theta_0)| > cA\} \\
 & \quad + P\left\{\inf_{\theta \in N_n(A)} \lambda_{\min}(D_n^{-\frac{1}{2}} F_n(\tilde{\theta}) D_n^{-\frac{T}{2}}) \leq c\right\}. \tag{6.3}
 \end{aligned}$$

By Lemma 5.2 and the Chebychev inequality, we obtain

$$P\{|D_n^{-\frac{1}{2}} S_n(\theta_0)| > cA\} \leq \frac{\text{Var}(D_n^{-\frac{1}{2}} S_n(\theta_0))}{c^2 A^2} = \frac{\sigma_0^2}{c^2 A^2}. \tag{6.4}$$

Let $A \rightarrow \infty$, then $c \downarrow 0$, and using (5.13), we have

$$P\left\{\inf_{\varphi \in N_n(A)} \lambda_{\min}(D_n^{-\frac{1}{2}} F_n(\tilde{\theta}) D_n^{-\frac{T}{2}}) \leq c\right\} \rightarrow 0. \tag{6.5}$$

By (6.3)-(6.5), we have

$$\lim_{A \rightarrow \infty} \liminf_{n \rightarrow \infty} P\{\Psi_n(\theta, \sigma^2) < \Psi_n(\theta_0, \sigma^2) \text{ for all } \theta \in M_n(A)\} = 1. \tag{6.6}$$

By Lemma 5.3, $\lambda_{\min}(X_n(\theta_0)) \rightarrow \infty$ as $n \rightarrow \infty$. Hence $\lambda_{\min}(D_n) \rightarrow \infty$. Moreover, from (5.13), we have

$$\inf_{\theta \in N_n(A)} \lambda_{\min}(F_n(\theta)) \rightarrow_p \infty.$$

This implies that $\Psi_n(\theta, \sigma^2)$ is concave on $N_n(A)$. Noting this fact and (6.6), we get

$$\begin{aligned}
 & \lim_{A \rightarrow \infty} \liminf_{n \rightarrow \infty} P\left\{\sup_{\theta \in M_n(A)} \Psi_n(\theta, \sigma^2) < \Psi_n(\theta_0, \sigma^2), \Psi_n(\theta, \sigma^2) \text{ is concave on } N_n(A)\right\} \\
 & = 1. \tag{6.7}
 \end{aligned}$$

On the event in the brackets, the continuous function $\Psi_n(\theta, \sigma^2)$ has a unique maximum in θ over the compact neighborhood $N_n(A)$. Hence

$$\lim_{A \rightarrow \infty} \liminf_{n \rightarrow \infty} P\{S_n(\hat{\theta}_n(A)) = 0 \text{ for a unique } \hat{\theta}_n(A) \in N_n(A)\} = 1.$$

Moreover, there is a sequence $A_n \rightarrow \infty$ such that $\hat{\theta}_n = \hat{\theta}(A_n)$ satisfies

$$\liminf_{n \rightarrow \infty} P\{S_n(\hat{\theta}_n) = 0 \text{ and } \hat{\theta}_n \text{ maximizes } \Psi_n(\theta, \sigma^2) \text{ uniquely in } N_n(A)\} = 1.$$

This $\hat{\theta}_n = (\hat{\beta}_n, \hat{\lambda}_n)$ is a QML estimator for θ_0 . It is clearly consistent, and

$$\lim_{A \rightarrow \infty} \liminf_{n \rightarrow \infty} P\{\hat{\theta}_n \in N_n(A)\} = 1.$$

Since $\hat{\theta}_n = (\hat{\beta}_n, \hat{\lambda}_n)$ are ML estimators for θ_0 , $\hat{\sigma}_n^2$ is an ML estimator for σ_0^2 from (2.9).

To complete the proof, we will show that $\hat{\sigma}_n^2 \rightarrow \sigma_0^2$ as $n \rightarrow \infty$. If $\hat{\theta}_n \in N_n(A)$, then $\hat{\beta}_n \in N_n^\beta(A)$ and $\hat{\lambda}_n \in N_n^\lambda(A)$.

By (2.12) and (2.1), we have

$$\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d) \hat{\varepsilon}_{t-1} = (x_t - \exp(-\hat{\lambda}_n d) x_{t-1})^T (\beta_0 - \hat{\beta}_n) + (e_t - \exp(-\hat{\lambda}_n d) e_{t-1}). \quad (6.8)$$

By (2.9), (2.11) and (6.8), we have

$$\begin{aligned} (n-1)\hat{\sigma}_n^2 &= \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d) \hat{\varepsilon}_{t-1})^2 \\ &= \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d) \hat{\varepsilon}_{t-1}) \\ &\quad \cdot \{(x_t - \exp(-\hat{\lambda}_n d) x_{t-1})^T (\beta_0 - \hat{\beta}_n) + (e_t - \exp(-\hat{\lambda}_n d) e_{t-1})\} \\ &= \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d) \hat{\varepsilon}_{t-1}) (x_t - \exp(-\hat{\lambda}_n d) x_{t-1})^T (\beta_0 - \hat{\beta}_n) \\ &\quad + \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d) \hat{\varepsilon}_{t-1}) (e_t - \exp(-\hat{\lambda}_n d) e_{t-1}) \\ &= \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d) \hat{\varepsilon}_{t-1}) (e_t - \exp(-\hat{\lambda}_n d) e_{t-1}). \end{aligned} \quad (6.9)$$

From (6.8), it follows that

$$\begin{aligned} &\sum_{t=2}^n \{(x_t - \exp(-\hat{\lambda}_n d) x_{t-1})^T (\beta_0 - \hat{\beta}_n)\}^2 \\ &= \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d) \hat{\varepsilon}_{t-1})^2 \\ &\quad - 2 \sum_{t=2}^n (\hat{\varepsilon}_t - \exp(-\hat{\lambda}_n d) \hat{\varepsilon}_{t-1}) (e_t - \exp(-\hat{\lambda}_n d) e_{t-1}) \\ &\quad + \sum_{t=2}^n (e_t - \exp(-\hat{\lambda}_n d) e_{t-1})^2. \end{aligned} \quad (6.10)$$

From (2.2), we get

$$\begin{aligned} &\sum_{t=2}^n (e_t - \exp(-\hat{\lambda}_n d) e_{t-1})^2 \\ &= \sum_{t=2}^n \left(\exp(-\lambda_0 d) e_{t-1} + \sigma_0 \sqrt{\frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0}} \eta_t - \exp(-\hat{\lambda}_n d) e_{t-1} \right)^2 \end{aligned}$$

$$\begin{aligned}
 &= \sigma_0^2 \frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0} \sum_{t=2}^n \eta_t^2 \\
 &\quad + \sum_{t=2}^n (\exp(-\lambda_0 d) - \exp(-\hat{\lambda}_n d))^2 e_{t-1}^2 \\
 &\quad + 2\sigma_0 \sqrt{\frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0}} \sum_{t=2}^n (\exp(-\lambda_0 d) - \exp(-\hat{\lambda}_n d)) \eta_t e_{t-1}. \tag{6.11}
 \end{aligned}$$

By (6.9)-(6.11), we have

$$\begin{aligned}
 (n-1)\hat{\sigma}_n^2 &= \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \sum_{t=2}^n (e_t - \exp(-\hat{\lambda}_n d)e_{t-1})^2 \\
 &\quad - \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \sum_{t=2}^n ((x_t - \exp(-\hat{\lambda}_n d)x_{t-1})^T (\beta_0 - \hat{\beta}_n))^2 \\
 &= \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \cdot \sigma_0^2 \frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0} \sum_{t=2}^n \eta_t^2 \\
 &\quad + \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \cdot \sum_{t=2}^n (\exp(-\lambda_0 d) - \exp(-\hat{\lambda}_n d))^2 e_{t-1}^2 \\
 &\quad + 2 \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \cdot \sigma_0 \sqrt{\frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0}} \\
 &\quad \cdot \sum_{t=2}^n (\exp(-\lambda_0 d) - \exp(-\hat{\lambda}_n d)) \eta_t e_{t-1} \\
 &\quad - \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \sum_{t=2}^n ((x_t - \exp(-\hat{\lambda}_n d)x_{t-1})^T (\beta_0 - \hat{\beta}_n))^2 \\
 &= T_1 + T_2 + 2T_3 - T_4. \tag{6.12}
 \end{aligned}$$

By the law of large numbers and $\hat{\lambda}_n \rightarrow_p \lambda$, we have

$$\begin{aligned}
 \frac{1}{n-1} T_1 &= \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \cdot \frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0} \sigma_0^2 \cdot \frac{1}{n-1} \sum_{t=2}^n \eta_t^2 \\
 &\rightarrow_p \sigma_0^2 \frac{2\lambda_n}{1 - \exp(-2\lambda_n d)} \cdot \frac{1 - \exp(-2\lambda_0 d)}{2\lambda_0} \\
 &= \sigma_0^2 \quad (n \rightarrow \infty). \tag{6.13}
 \end{aligned}$$

By the Markov inequality, and noting that $ET_2 \leq CA^2$, we obtain

$$\frac{1}{n-1} T_2 \rightarrow_p 0 \quad (n \rightarrow \infty). \tag{6.14}$$

Since $\{(\exp(-\lambda_0 d) - \exp(-\hat{\lambda}_n d))\eta_t e_{t-1}, H_{t-1}\}$ is a martingale difference sequence with

$$\text{Var}\{(\exp(-\lambda_0 d) - \exp(-\hat{\lambda}_n d))\eta_t e_{t-1}\} = (\exp(-\lambda_0 d) - \exp(-\hat{\lambda}_n d))^2 Ee_{t-1}^2,$$

so we have

$$\begin{aligned} \text{Var}(T_3) &= \sum_{t=2}^n E((\exp(-\lambda_0 d) - \exp(-\hat{\lambda}_n d))\eta_t e_{t-1})^2 \\ &= \sum_{t=2}^n (\exp(-\lambda_0 d) - \exp(-\hat{\lambda}_n d))^2 Ee_{t-1}^2 \\ &\leq C(\lambda_0 - \hat{\lambda}_n)^2 \sum_{t=2}^n Ee_{t-1}^2 \leq CA^2. \end{aligned} \tag{6.15}$$

By the Chebychev inequality, we have

$$\frac{1}{n-1} T_3 \rightarrow_p 0 \quad (n \rightarrow \infty). \tag{6.16}$$

By (5.33), we have

$$\begin{aligned} T_4 &= \sum_{t=2}^n ((x_t^T(\beta_0 - \hat{\beta}_n) - \exp(-\hat{\lambda}_n d)x_{t-1})^T(\beta_0 - \hat{\beta}_n))^2 \\ &\leq 2 \sum_{t=2}^n (x_t^T(\beta_0 - \hat{\beta}_n))^2 + \sum_{t=2}^n (\exp(-\hat{\lambda}_n d)x_{t-1}^T(\beta_0 - \hat{\beta}_n))^2 \\ &= o(n). \end{aligned} \tag{6.17}$$

From (6.12)-(6.14), (6.16) and (6.17), we have $\hat{\sigma}_n^2 \rightarrow \sigma_0^2$.

We therefore complete the proof of Theorem 3.1. □

Proof of Theorem 3.2 It is easy to know that $S_n(\hat{\theta}_n) = 0$ and $F_n(\hat{\theta}_n)$ is nonsingular from Theorem 3.1. By the Taylor expansion, we have

$$0 = S_n(\hat{\theta}_n) = S_n(\theta_0) - F_n(\tilde{\theta}_n)(\hat{\theta}_n - \theta_0). \tag{6.18}$$

Since $\hat{\theta}_n \in N_n(A)$, also $\tilde{\theta}_n \in N_n(A)$. By (5.11), we have

$$F_n(\tilde{\theta}_n) = D_n^{\frac{1}{2}}(\Phi_n + \tilde{A}_n)D_n^{\frac{T}{2}}, \tag{6.19}$$

where \tilde{A}_n is a symmetric matrix with $\tilde{A}_n \rightarrow_p 0$. By (6.18) and (6.19), we have

$$D_n^{\frac{T}{2}}(\hat{\theta}_n - \theta_0) = D_n^{\frac{T}{2}}F_n^{-1}(\tilde{\theta}_n)S_n(\theta_0) = (\Phi_n + \tilde{A}_n)^{-1}D_n^{-\frac{1}{2}}S_n(\theta_0). \tag{6.20}$$

Similar to (6.20), we have

$$\begin{aligned} F_n(\hat{\theta}_n) &= D_n^{\frac{1}{2}}(\Phi_n + \hat{A}_n)D_n^{\frac{T}{2}} \\ &= (D_n^{\frac{1}{2}}(\Phi_n + \hat{A}_n)^{\frac{1}{2}})((\Phi_n + \hat{A}_n)^{\frac{T}{2}}D_n^{\frac{T}{2}}) \\ &= F_n^{\frac{1}{2}}(\hat{\theta}_n)F_n^{\frac{T}{2}}(\hat{\theta}_n). \end{aligned} \tag{6.21}$$

Here $\hat{A}_n \rightarrow_p 0$. By (6.20), (6.21), and noting that $\hat{\sigma}_n^2 \rightarrow_p \sigma_0^2$ and $D_n^{-\frac{1}{2}} S_n(\theta_0) = O_p(1)$, we obtain that

$$\begin{aligned} F_n^{\frac{T}{2}}(\hat{\theta}_n)(\hat{\theta}_n - \theta_0)/\hat{\sigma}_n &= (\Phi_n + \hat{A}_n)^{\frac{1}{2}} (\Phi_n + \tilde{A}_n)^{-1} D_n^{-\frac{1}{2}} S_n(\theta_0)/\hat{\sigma}_n \\ &= \Phi_n^{-\frac{1}{2}} D_n^{-\frac{1}{2}} S_n(\theta_0)/\sigma_0 + o_p(1). \end{aligned} \tag{6.22}$$

From (2.7) and (2.8), we have

$$\begin{aligned} \frac{S_n(\theta_0)}{\sigma_0} &= \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \left\{ \sum_{t=2}^n \eta_t (x_t - \exp(-\lambda_0 d)x_{t-1}), \right. \\ &\quad - d \exp(-\lambda_0 d) \sum_{t=2}^n \eta_t e_{t-1} \\ &\quad \left. - \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \frac{\sigma_0 [1 - (1 + 2d\lambda_0) \exp(-2\lambda_0 d)]}{4\lambda_0^2} \sum_{t=2}^n (\eta_t^2 - 1) \right\}. \end{aligned} \tag{6.23}$$

From (5.14) and (5.15), we have

$$\begin{aligned} \Phi_n^{-\frac{1}{2}} D_n^{-\frac{1}{2}} &= \begin{pmatrix} \left(\frac{\lambda(1 - \exp(-2d\lambda_0))}{\lambda_0(1 - \exp(-2d\lambda))}\right)^{-\frac{1}{2}} I_d & 0 \\ 0 & \sqrt{\frac{\Delta_n(\theta_0, \sigma_0)}{-\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} |_{\theta=\theta_0}}} \end{pmatrix} \\ &\cdot \begin{pmatrix} \left(\frac{2\lambda_0}{1 - \exp(-2d\lambda_0)}\right)^{-\frac{1}{2}} X_n^{-\frac{1}{2}}(\lambda_0) & 0 \\ 0 & \frac{1}{\sqrt{\Delta_n(\theta_0, \sigma_0)}} \end{pmatrix} \\ &= \begin{pmatrix} \left(\frac{2\lambda}{1 - \exp(-2d\lambda)}\right)^{-\frac{1}{2}} X_n^{-\frac{1}{2}}(\lambda_0) & 0 \\ 0 & \frac{1}{\sqrt{-\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} |_{\theta=\theta_0}}} \end{pmatrix}. \end{aligned} \tag{6.24}$$

By (6.23) and (6.24), we have

$$\begin{aligned} \Phi_n^{-\frac{1}{2}} D_n^{-\frac{1}{2}} S_n(\theta_0)/\sigma_0 &= \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \left\{ \left(\frac{2\lambda}{1 - \exp(-2d\lambda)}\right)^{-\frac{1}{2}} \sum_{t=2}^n \eta_t X_n^{-\frac{1}{2}}(\theta_0) (x_t - \exp(-\lambda_0 d)x_{t-1}), \right. \\ &\quad \frac{1}{\sqrt{-\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} |_{\theta=\theta_0}}} \cdot \left[-d \exp(-\lambda_0 d) \sum_{t=2}^n \eta_t e_{t-1} \right. \\ &\quad \left. \left. - \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \frac{\sigma_0 [1 - (1 + 2d\lambda_0) \exp(-2\lambda_0 d)]}{4\lambda_0^2} \sum_{t=2}^n (\eta_t^2 - 1) \right] \right\}. \end{aligned} \tag{6.25}$$

Let $u \in R^d$ with $|u| = 1$, and

$$a_{tn} = u \left(\frac{2\lambda}{1 - \exp(-2d\lambda)} \right)^{-\frac{1}{2}} X_n^{-\frac{1}{2}}(\lambda_0) (x_t - \exp(-\lambda_0 d)x_{t-1}).$$

Then $\max_{2 \leq t \leq n} a_{tn} = o(1)$, and we will consider the limiting distribution of the following 2-vector

$$\begin{aligned} & \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \left\{ \sum_{t=2}^n a_{tn} \eta_t, -\frac{1}{\sqrt{-\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} \Big|_{\theta=\theta_0}}} \left[d \exp(-\lambda_0 d) \sum_{t=2}^n \eta_t e_{t-1} \right. \right. \\ & \left. \left. + \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \frac{\sigma_0 [1 - (1 + 2d\lambda_0) \exp(-2\lambda_0 d)]}{4\lambda_0^2} \sum_{t=2}^n (\eta_t^2 - 1) \right] \right\}. \end{aligned} \quad (6.26)$$

Note that

$$-\sigma^2 \frac{\partial^2 \Psi_n}{\partial \lambda^2} \Big|_{\theta=\theta_0} = O_p(\Delta_n(\theta_0, \sigma_0)) = O_p(n).$$

Hence, by the Cramer-Wold device, it will suffice to find the asymptotic distribution of the following random

$$\begin{aligned} & \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \sum_{t=2}^n \left\{ u_1 a_{tn} \eta_t - \frac{u_2}{\sqrt{\Delta_n(\theta_0, \sigma_0)}} \left[d \exp(-\lambda_0 d) \eta_t e_{t-1} \right. \right. \\ & \left. \left. + \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \frac{\sigma_0 [1 - (1 + 2d\lambda_0) \exp(-2\lambda_0 d)]}{4\lambda_0^2} (\eta_t^2 - 1) \right] \right\} \\ & = \sum_{t=2}^n \zeta_t, \end{aligned} \quad (6.27)$$

where $(u_1, u_2) \in R^2$ with $u_1^2 + u_2^2 = 1$. Note that

$$\begin{aligned} E\{\zeta_t | H_{t-1}\} & = \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \left\{ u_1 a_{tn} E(\eta_t) - \frac{u_2}{\sqrt{\Delta_n(\theta_0, \sigma_0)}} \left[d \exp(-\lambda_0 d) E(\eta_t) e_{t-1} \right. \right. \\ & \left. \left. + \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \frac{\sigma_0 [1 - (1 + 2d\lambda_0) \exp(-2\lambda_0 d)]}{4\lambda_0^2} E(\eta_t^2 - 1) \right] \right\} \\ & = 0, \quad \text{a.s.,} \end{aligned} \quad (6.28)$$

so the sums in (6.27) are partial sums of a martingale triangular array to H_t , and we will verify the Lindeberg conditions for their convergence to normality.

By (6.27), and noting that $E\eta_t^3 = 0$, $E\eta_t^4 = 3$ and $\lambda \in N_n^\lambda(A)$, we have

$$\begin{aligned} \sum_{t=2}^n E(\zeta_t^2 | H_{t-1}) & = \frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)} \left\{ u_1^2 \sum_{t=2}^n a_{tn}^2 + u_2^2 \frac{1}{\Delta_n(\theta_0, \sigma_0)} \left[d^2 \exp(-2\lambda_0 d) \sum_{t=2}^n e_{t-1}^2 \right. \right. \\ & \left. \left. + \frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)} \frac{\sigma_0^2 [1 - (1 + 2d\lambda_0) \exp(-2\lambda_0 d)]^2}{16\lambda_0^2} \sum_{t=2}^n E(\eta_t^2 - 1)^2 \right] \right\} \\ & \quad - 2 \sum_{t=2}^n u_1 \frac{u_2}{\sqrt{n}} d \exp(-\lambda_0 d) a_{tn} e_{t-1} \\ & \quad + 2 \frac{u_2^2}{n} d \exp(-\lambda_0 d) \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0 d)}} \end{aligned}$$

$$\begin{aligned}
 & \left. \begin{aligned}
 & \frac{\sigma_0[1 - (1 + 2d\lambda_0)\exp(-2\lambda_0d)]}{4\lambda_0^2} E(\eta_t(\eta_t^2 - 1))e_{t-1} \\
 & - 2u_1a_{tn} \frac{u_2}{\sqrt{n}} \sqrt{\frac{2\lambda_0}{1 - \exp(-2\lambda_0d)}} \\
 & \frac{\sigma_0[1 - (1 + 2d\lambda_0)\exp(-2\lambda_0d)]}{4\lambda_0^2} E(\eta_t(\eta_t^2 - 1)) \left. \right\} \\
 & = u_1^2 u^2 \frac{2\lambda_0}{1 - \exp(-2\lambda_0d)} \left(\frac{2\lambda}{1 - \exp(-2d\lambda)} \right)^{-1} + u_2^2 + o_p(1) + 0 + 0 \\
 & = u_1^2 + u_2^2 + o_p(1) = 1 + o_p(1).
 \end{aligned} \tag{6.29}
 \end{aligned}$$

Let $\tilde{a}_{tn} = \min\{a_{tn}, \frac{1}{\sqrt{\Delta_n(\theta_0, \sigma_0)}}\}$ and $\tilde{\zeta}_t = \tilde{a}_{tn}\zeta_t$. Then $\tilde{a}_{tn} = o(1)$.
 For any $c > 0$,

$$\begin{aligned}
 \sum_{t=2}^n E\{\zeta_t^2 I(|\zeta_t| > c) | H_{t-1}\} &= \sum_{t=2}^n \int_c^\infty y^2 dP\{|\tilde{a}_{tn}\tilde{\zeta}_t| \leq y | H_{t-1}\} \\
 &= \sum_{t=2}^n \tilde{a}_{tn}^2 \int_{\frac{c}{\tilde{a}_{tn}}}^\infty y^2 dP\{|\tilde{\zeta}_t| \leq y | H_{t-1}\} \\
 &= o(1) \sum_{t=2}^n \tilde{a}_{tn}^2 = o(1)O_p(1) \rightarrow 0, \quad n \rightarrow \infty.
 \end{aligned} \tag{6.30}$$

This verifies the Lindeberg conditions, and by Lemma 5.5, we have

$$\sum_{t=2}^n \zeta_t \rightarrow_D N(0, 1).$$

Thus we complete the proof of Theorem 3.2. □

Proof of Theorem 4.1 Note that $\hat{\lambda}_{0n} \rightarrow \lambda_0$, $\hat{\lambda}_n \rightarrow \lambda_0$. Similarly to the proof of Theorem 4.1(3) in Maller [55], by (6.12) and Theorem 3.2, we have

$$\begin{aligned}
 \tilde{d}(n) &= \frac{1}{\sigma_0^2} \left\{ \left(\frac{2\hat{\lambda}_{0n}}{1 - \exp(-2\hat{\lambda}_{0n}d)} - \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \right) \sigma_0^2 \frac{1 - \exp(-2\lambda_0d)}{2\lambda_0} \sum_{t=2}^n \eta_t^2 \right. \\
 &+ \sum_{t=2}^n \left\{ \frac{2\hat{\lambda}_{0n}}{1 - \exp(-2\hat{\lambda}_{0n}d)} (\exp(-\lambda_0d) - \exp(-\hat{\lambda}_{0n}d))^2 \right. \\
 &- \left. \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} (\exp(-\lambda_0d) - \exp(-\hat{\lambda}_n d))^2 \right\} e_{t-1}^2 \\
 &+ 2\sigma_0 \sqrt{\frac{1 - \exp(-2\lambda_0d)}{2\lambda_0}} \sum_{t=2}^n \left\{ \frac{2\hat{\lambda}_{0n}}{1 - \exp(-2\hat{\lambda}_{0n}d)} (\exp(-\lambda_0d) - \exp(-\hat{\lambda}_{0n}d)) \right. \\
 &- \left. \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} (\exp(-\lambda_0d) - \exp(-\hat{\lambda}_n d)) \right\} \eta_t e_{t-1} \\
 &\left. + \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \sum_{t=2}^n ((x_t - \exp(-\hat{\lambda}_n d)x_{t-1})^T (\beta_0 - \hat{\beta}_n))^2 \right\}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{\sigma_0^2} \frac{2\hat{\lambda}_n}{1 - \exp(-2\hat{\lambda}_n d)} \sum_{t=2}^n ((x_t - \exp(-\hat{\lambda}_n d)x_{t-1})^T (\beta_0 - \hat{\beta}_n))^2 + o(1) \\
 &\rightarrow_D \chi^2(m).
 \end{aligned} \tag{6.31}$$

□

7 Empirical examples

In the section, we consider two empirical examples. The first one (β is a one-dimensional unknown parameter, namely $m = 1$) is water flowing in the Kootenay River in January, which is taken from Hampel *et al.* [6, p.310]. The second one (β is a 4-dimensional unknown parameter, namely $m = 4$) is the consumption of spirits in the United Kingdom, which is taken from Fuller [48].

7.1 Water flowing in the Kootenay river

By the ordinary least squares method, we obtain that

$$\hat{y}_t = 9.51371 + 0.47476x_t + \hat{\varepsilon}_t \tag{7.1}$$

and

$$\varepsilon_t = 0.2077\varepsilon_{t-1} + \eta_t, \quad t = 1, 2, \dots, 13, \tag{7.2}$$

where η_t is a sequence of uncorrelated $(0, 1.5013^2)$ random variables.

By the Huber-Dutter (HD) method, we obtain the following model (see Hu [10]):

$$\hat{y}_t = 9.51371 + 0.4745x_t + \hat{\varepsilon}_t \tag{7.3}$$

and

$$\varepsilon_t = 0.3024\varepsilon_{t-1} + \eta_t, \tag{7.4}$$

where η_t is a sequence of uncorrelated $(0, 1.0988^2)$ random variables.

By the ML method (take $d = 1$ and starting values for $\lambda^{(0)} = 1$, $(\sigma^2)^{(0)} = 1.5$, $\beta^{(0)} = 0.5$; here we use pattern search algorithms), we obtain the following model:

$$\hat{y}_t = 9.51371 + 0.48039x_t + \hat{\varepsilon}_t \tag{7.5}$$

and

$$\varepsilon_t = \exp(-1.80089)\varepsilon_{t-1} + 0.5184\eta_t, \tag{7.6}$$

where η_t is a sequence of uncorrelated $(0, 1)$ random variables.

By model (1.3), we obtain a general process $\{y_t\}$ satisfying the following SDE:

$$d(y_t - 9.51371) = (1.3455x_t + 9.51371 - y_t) dt + 0.9976 dB_t. \tag{7.7}$$

Since $1.5013^2 > 1.0988^2 > 0.5184^2$, our results excel the results of HD and the least squares method in mean squares error (MSE).

By (4.7), we obtain $\tilde{d}(13) = 362.4137 > 6.63 = \chi_{1-0.01}^2(1)$. It is shown that $\beta \neq 0$ at the significant level $\alpha = 0.01$. Thus we should apply the linear regression model (1.1) with Ornstein-Uhlenbeck process instead of only the Ornstein-Uhlenbeck process for the data.

It is shown that our estimation method and testing approach are valid in the case of $m = 1$. For a multidimensional parameter β , it is true in the following example.

7.2 Consumption of spirits in the UK

We will use the data studied by Fuller [48]. The data pertain to the consumption of spirits in the United Kingdom from 1870 to 1983. The dependent variable y_t is the annual per capita consumption of spirits in the United Kingdom. The explanatory variables x_{t1} and x_{t2} are per capita income and price of spirits, respectively, both deflated by a general price index. All data are in logarithms. The model suggested by Prest can be written as follows:

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \varepsilon_t, \quad (7.8)$$

where 1869 is the origin for t , $x_{t3} = \frac{t}{100}$, $x_{t4} = \frac{(t-35)^2}{10^4}$, and assume that ε_t is a stationary time series.

Fuller [48] obtained the estimated generalized least squares equation

$$\hat{y}_t = 2.36 + 0.72x_{t1} - 0.80x_{t2} - 0.81x_{t3} - 0.92x_{t4} \quad (7.9)$$

and

$$\varepsilon_t = 0.7633\varepsilon_{t-1} + \eta_t,$$

where η_t is a sequence of uncorrelated $(0, 0.000417)$ random variables.

Take $d = 1$ and starting values for

$$\lambda^{(0)} = 0.3, \quad (\sigma^2)^{(0)} = 0.0004, \quad \beta^{(0)} = (0.72, -0.80, -0.81, -0.92)^T.$$

Using our method, we obtain the following models:

$$\hat{y}_t = 2.36 + 0.73251x_{t1} - 0.80024x_{t2} - 0.86286x_{t3} - 0.60774x_{t4} \quad (7.10)$$

and

$$\varepsilon_t = \exp(-0.25319)\varepsilon_{t-1} + 0.0196\eta_t, \quad (7.11)$$

where η_t is a sequence of uncorrelated $(0, 1)$ random variables; or

$$d\varepsilon_t = -0.25319\varepsilon_t dt + 0.0221 dB_t. \quad (7.12)$$

Since $0.000417 > 0.00038461$, our results excel the results of Fuller [48] in MSE.

By (4.7), we obtain $\tilde{d}(69) = 100.2777 > 13.3 = \chi_{1-0.01}^2(4)$. It is shown that $\beta \neq 0$ at the significant level $\alpha = 0.01$.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Author details

¹School of Mathematics and Statistics, Hubei Normal University, Huangshi, 435002, China. ²Faculty of Information Engineering, China University of Geosciences, Wuhan, 430074, China.

Acknowledgements

This work was supported by the Natural Science Foundation of China (No. 41374017), and Science and Technology Research Projects of the Educational Department of Hubei Province (No. Q20142501).

Received: 5 February 2014 Accepted: 9 July 2014 Published: 19 Aug 2014

References

1. Wang, XM, Zhou, W: Bootstrap approximation to the distribution of M -estimates in a linear model. *Acta Math. Sin. Engl. Ser.* **20**(1), 93-104 (2004)
2. Anatolyev, S: Inference in regression models with many regressors. *J. Econom.* **170**, 368-382 (2012)
3. Bai, ZD, Guo, M: A paradox in least-squares estimation of linear regression models. *Stat. Probab. Lett.* **42**, 167-174 (1999)
4. Chen, X: Consistency of LS estimates of multiple regression under a lower order moment condition. *Sci. China Ser. A* **38**(12), 1420-1431 (1995)
5. Gil, GR, Engela, B, Norberto, C, Ana, C: Least squares estimation of linear regression models for convex compact random sets. *Adv. Data Anal. Classif.* **1**, 67-81 (2007)
6. Hampel, FR, Ronchetti, EM, Rousseeuw, PJ, Stahel, WA: *Robust Statistics*. Wiley, New York (1986)
7. Cui, H: On asymptotics of t -type regression estimation in multiple linear model. *Sci. China Ser. A* **47**(4), 628-639 (2004)
8. Durbin, L: A note on regression when there is extraneous information about one of the coefficients. *J. Am. Stat. Assoc.* **48**, 799-808 (1953)
9. Li, Y, Yang, H: A new stochastic mixed ridge estimator in linear regression model. *Stat. Pap.* **51**(2), 315-323 (2010)
10. Hu, HC: Asymptotic normality of Huber-Dutter estimators in a linear model with AR(1) processes. *J. Stat. Plan. Inference* **143**(3), 548-562 (2013)
11. Wu, WB: M -Estimation of linear models with dependent errors. *Ann. Stat.* **35**(2), 495-521 (2007)
12. Fox, R, Taqqu, MS: Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Stat.* **14**, 517-532 (1986)
13. Giraitis, L, Surgailis, D: A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. *Probab. Theory Relat. Fields* **86**, 87-104 (1990)
14. Koul, HL, Surgailis, D: Asymptotic normality of the Whittle estimator in linear regression models with long memory errors. *Stat. Inference Stoch. Process.* **3**, 129-147 (2000)
15. Shiohama, T, Taniguchi, M: Sequential estimation for time series regression models. *J. Stat. Plan. Inference* **123**, 295-312 (2004)
16. Fan, J: Moderate deviations for M -estimators in linear models with ϕ -mixing errors. *Acta Math. Sin. Engl. Ser.* **28**(6), 1275-1294 (2012)
17. Ornstein, LS, Uhlenbeck, GE: On the theory of Brownian motion. *Phys. Rev.* **36**, 823-841 (1930)
18. Janczura, J, Orzel, S, Wylomanska, A: Subordinated α -stable Ornstein-Uhlenbeck process as a tool for financial data description. *Physica A* **390**, 4379-4387 (2011)
19. Debbasch, F, Mallick, K, Rivet, JP: Relativistic Ornstein-Uhlenbeck process. *J. Stat. Phys.* **88**, 945-966 (1997)
20. Gillespie, D: Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. *Phys. Rev. E* **54**(2), 2084-2091 (1996)
21. Ditlevsen, S, Lansky, P: Estimation of the input parameters in the Ornstein-Uhlenbeck neuronal model. *Phys. Rev. E* **71**(1), 011907 (2005)
22. Garbaczewski, P, Olkiewicz, R: Ornstein-Uhlenbeck-Cauchy process. *J. Math. Phys.* **41**(10), 6843-6860 (2000)
23. Plastino, AR, Plastino, A: Non-extensive statistical mechanics and generalized Fokker-Planck equation. *Physica A* **222**, 347-354 (1995)
24. Fasen, V: Statistical estimation of multivariate Ornstein-Uhlenbeck processes and applications to co-integration reserved. *J. Econom.* (2012). doi:10.1016/j.jeconom.2012.08.019
25. Yu, J: Bias in the estimation of the mean reversion parameter in continuous time models. *J. Econom.* **169**, 114-122 (2012)
26. Geman, H: *Commodities and Commodity Derivatives*. Wiley, Chichester (2005)
27. Zhang, B, Grzelak, LA, Oosterlee, CM: Efficient pricing of commodity options with early-exercise under the Ornstein-Uhlenbeck process. *Appl. Numer. Math.* **62**, 91-111 (2012)
28. Rieder, S: Robust parameter estimation for the Ornstein-Uhlenbeck process. *Stat. Methods Appl.* (2012). doi:10.1007/s10260-012-0195-2
29. Iacus, S: *Simulation and Inference for Stochastic Differential Equations*. Springer, New York (2008)
30. Bishwal, JPN: Uniform rate of weak convergence of the minimum contrast estimator in the Ornstein-Uhlenbeck process. *Methodol. Comput. Appl. Probab.* **12**, 323-334 (2010)
31. Shimizu, Y: Local asymptotic mixed normality for discretely observed non-recurrent Ornstein-Uhlenbeck processes. *Ann. Inst. Stat. Math.* **64**, 193-211 (2012)
32. Zhang, S, Zhang, X: A least squares estimator for discretely observed Ornstein-Uhlenbeck processes driven by symmetric α -stable motions. *Ann. Inst. Stat. Math.* (2012). doi:10.1007/s10463-012-0362-0
33. Chronopoulou, A, Viens, FG: Estimation and pricing under long-memory stochastic volatility. *Ann. Finance* **8**, 379-403 (2012)
34. Lin, H, Wang, J: Successful couplings for a class of stochastic differential equations driven by Levy processes. *Sci. China Math.* **55**(8), 1735-1748 (2012)

35. Xiao, W, Zhang, W, Zhang, X: Minimum contrast estimator for fractional Ornstein-Uhlenbeck processes. *Sci. China Math.* **55**(7), 1497-1511 (2012)
36. Magdalinos, T: Mildly explosive autoregression under weak and strong dependence. *J. Econom.* **169**, 179-187 (2012)
37. Andrews, DWK, Guggenberger, P: Asymptotics for LS, GLS, and feasible GLS statistics in an AR(1) model with conditional heteroskedasticity. *J. Econom.* **169**, 196-210 (2012)
38. Fan, J, Yao, Q: *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer, New York (2005)
39. Berk, KN: Consistent autoregressive spectral estimates. *Ann. Stat.* **2**, 489-502 (1974)
40. Goldenshluger, A, Zeevi, A: Non-asymptotic bounds for autoregressive time-series modeling. *Ann. Stat.* **29**, 417-444 (2001)
41. Liebscher, E: Strong convergence of estimators in nonlinear autoregressive models. *J. Multivar. Anal.* **84**, 247-261 (2003)
42. Baran, S, Pap, G, Zuijlen, MV: Asymptotic inference for unit roots in spatial triangular autoregression. *Acta Appl. Math.* **96**, 17-42 (2007)
43. Distaso, W: Testing for unit root processes in random coefficient autoregressive models. *J. Econom.* **142**, 581-609 (2008)
44. Harvill, JL, Ray, BK: Functional coefficient autoregressive models for vector time series. *Comput. Stat. Data Anal.* **50**, 3547-3566 (2008)
45. Dehling, H, Franke, B, Kott, T: Drift estimation for a periodic mean reversion process. *Stat. Inference Stoch. Process.* **13**, 175-192 (2010)
46. Maller, RA: Asymptotics of regressions with stationary and nonstationary residuals. *Stoch. Process. Appl.* **105**, 33-67 (2003)
47. Pere, P: Adjusted estimates and Wald statistics for the AR(1) model with constant. *J. Econom.* **98**, 335-363 (2000)
48. Fuller, WA: *Introduction to Statistical Time Series*, 2nd edn. Wiley, New York (1996)
49. Chambers, MJ: Jackknife estimation of stationary autoregressive models. *J. Econom.* (2012). doi:10.1016/j.jeconom.2012.09.003
50. Hamilton, JD: *Time Series Analysis*. Princeton University Press, Princeton (1994)
51. Brockwell, PJ, Davis, RA: *Time Series: Theory and Methods*. Springer, New York (1987)
52. Abadir, KM, Lucas, A: A comparison of minimum MSE and maximum power for the nearly integrated non-Gaussian model. *J. Econom.* **119**, 45-71 (2004)
53. Fan, JQ, Jiang, JC: Nonparametric inference with generalized likelihood ratio tests. *Test* **16**, 409-444 (2007)
54. Rao, CR: *Linear Statistical Inference and Its Applications*. Wiley, New York (1973)
55. Maller, RA: Quadratic negligibility and the asymptotic normality of operator normed sums. *J. Multivar. Anal.* **44**, 191-219 (1993)
56. Hall, P, Heyde, CC: *Martingale Limit Theory and Its Application*. Academic Press, New York (1980)

10.1186/1029-242X-2014-301

Cite this article as: Hu et al.: Maximum likelihood estimators in linear regression models with Ornstein-Uhlenbeck process. *Journal of Inequalities and Applications* 2014, **2014**:301

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
