Electronic Journal of Statistics
Vol. 11 (2017) 2978–3007
ISSN: 1935-7524
DOI: 10.1214/17-EJS1298

Maximum likelihood estimation for a bivariate Gaussian process under fixed domain asymptotics

Daira Velandia, François Bachoc

Departamento de Ciencias Naturales y Exactas. Universidad de la Costa, Calle 58 Nº 55-66, (Barranquilla, Colombia) and Institut de Mathématiques de Toulouse. Université Toulouse III – Paul Sabatier (Toulouse, France)
e-mail: dvelandi@cuc.edu.co; francois.bachoc@math.univ-toulouse.fr

Moreno Bevilacqua*, **Xavier Gendre**

Instituto de Estadística. Universidad de Valparaíso (Valparaiso, Chile) and Institut de Mathématiques de Toulouse. Université Toulouse III – Paul Sabatier (Toulouse, France)
e-mail: moreno.bevilacqua@uv.cl; xavier.gendre@math.univ-toulouse.fr

and

Jean-Michel Loubes

Institut de Mathématiques de Toulouse. Université Toulouse III – Paul Sabatier (Toulouse, France)
e-mail: jean-michel.loubes@math.univ-toulouse.fr

Abstract: We consider maximum likelihood estimation with data from a bivariate Gaussian process with a separable exponential covariance model under fixed domain asymptotics. We first characterize the equivalence of Gaussian measures under this model. Then consistency and asymptotic normality for the maximum likelihood estimator of the microergodic parameters are established. A simulation study is presented in order to compare the finite sample behavior of the maximum likelihood estimator with the given asymptotic distribution.

Keywords and phrases: Bivariate exponential model, equivalent Gaussian measures, infill asymptotics, microergodic parameters.

Received July 2016.

Contents

1 Introduction 2979
2 Equivalence of Gaussian measures 2981
3 Consistency of the maximum likelihood estimator 2984
4 Asymptotic distribution 2988
5 Numerical experiments 2995

*The research work conducted by Moreno Bevilacqua was supported in part by FONDECYT grant 1160280 Chile.

6 Concluding remarks 2998
A Appendix section 2999
References 3005

1. Introduction

Gaussian processes are widely used in statistics to model spatial data. When fitting a Gaussian field, one has to deal with the issue of the estimation of its covariance. In many cases, a model is chosen for the covariance, which turns the problem into a parametric estimation problem. Within this framework, the maximum likelihood estimator (MLE) of the covariance parameters of a Gaussian stochastic process observed in $\mathbb{R}^d$, $d \geq 1$, has been deeply studied in the last years in the two following asymptotic frameworks.

The fixed domain asymptotic framework, sometimes called infill asymptotics [29, 7], corresponds to the case where more and more data are observed in some fixed bounded sampling domain (usually a region of $\mathbb{R}^d$). The increasing domain asymptotic framework corresponds to the case where the sampling domain increases with the number of observed data and where the distance between any two sampling locations is bounded away from 0. The asymptotic behavior of the MLE of the covariance parameters can be quite different in these two frameworks [37].

Consider first increasing-domain asymptotics. Then, generally speaking, for all (identifiable) covariance parameters, the MLE is consistent and asymptotically normal under some mild regularity conditions. The asymptotic covariance matrix is equal to the inverse of the (asymptotic) Fisher information matrix. This result was first shown by [22], and then extended in different directions by [8, 9, 24, 4].

The situation is significantly different under fixed domain asymptotics. Indeed, two types of covariance parameters can be distinguished: microergodic and non-microergodic parameters [14, 29]. A covariance parameter is microergodic if, for two different values of it, the two corresponding Gaussian measures are orthogonal, see [14, 29]. It is non-microergodic if, even for two different values of it, the two corresponding Gaussian measures are equivalent. Non-microergodic parameters cannot be estimated consistently, but misspecifying them asymptotically results in the same statistical inference as specifying them correctly [26, 27, 28, 37]. On the other hand, it is at least possible to consistently estimate microergodic covariance parameters, and misspecifying them can have a strong negative impact on inference.

Nevertheless, under fixed domain asymptotics, it has often proven to be challenging to establish the microergodicity or non-microergodicity of covariance parameters, and to provide asymptotic results for estimators of microergodic parameters. Most available results are specific to particular covariance models. When $d = 1$ and the covariance model is exponential, only a reparameterized quantity obtained from the variance and scale parameters is microergodic. It is shown in [33] that the MLE of this microergodic parameter is consistent and asymptotically normal. When $d > 1$ and for a separable exponential covariance function, all the covariance parameters are microergodic, and the asymptotic normality of the MLE is proved in [34]. Other results in this case are also given in [30, 1, 6]. Consistency of the MLE is shown as well in [21] for the scale covariance parameters of the Gaussian covariance function and in [20] for all the covariance parameters of the separable Matérn $3/2$ covariance function. Finally, for the entire isotropic Matérn class of covariance functions, all parameters are microergodic for $d > 4$ [2], and only reparameterized parameters obtained from the scale and variance are microergodic for $d \leq 3$ [35]. In [17], the asymptotic distribution of MLEs for these microergodic parameters is provided, generalizing previous results in [10] and [31].

All the results discussed above have been obtained when considering a univariate stochastic process. There are few results on maximum likelihood in the multivariate setting. Under increasing-domain asymptotics [5] extend the results of [22] to the bivariate case and consider the asymptotic distribution of the MLE for a large class of bivariate covariance models in order to test the independence between two Gaussian processes. In [11], asymptotic consistency of the tapered MLE for multivariate processes is established, also under increasing domain asymptotics. In [23], some results are given on the distribution of the MLE of the correlation parameter between the two components of a bivariate stochastic process with a separable structure, when the space covariance is known, regardless of the asymptotic framework. In [18], the fixed domain asymptotic results of [34] are extended to the multivariate case, for $d = 3$ and when the correlation parameters between the different Gaussian processes are known. Finally, under fixed domain asymptotics, in the bivariate case and when considering an isotropic Matérn model, [36] show which covariance parameters are microergodic.

In this paper, we extend the results of [33] (when $d = 1$ and the covariance function is exponential) to the bivariate case. Our main motivation to study this particular setting is that, on the one hand, its theoretical analysis remains tractable, as the likelihood function is available in close form, and as the number of covariance parameters to estimate remain moderate in the bivariate case. On the other hand, obtaining rigorous proofs in this bivariate setting is insightful, and is a first step toward the fixed-domain asymptotic analysis of more general multivariate settings. We discuss potential multivariate extensions in the conclusion section.

Note that bivariate processes observed along lines can occur in practical situations. As mentioned in [10], spatial data can be collected along linear flight paths, see for instance the international H20 project, [32, 19, 25]. Furthermore, comparing two spatial quantities, for visualization, or for detecting correlations is very standard, see for instance Figure 9 in [32].

In this paper, we first consider the equivalence of Gaussian measures, that is to say we characterize which covariance parameters are microergodic. In the univariate case, [1] characterize the equivalence of Gaussian measures with exponential covariance function using the entropy distance criteria. We extend their approach to the bivariate case. It turns out, similarly as in the univariate case, that not all covariance parameters are microergodic. Hence not all covariance parameters can be consistently estimated. Then we establish the consistency and the asymptotic normality of the MLE of the microergodic parameters. Some of our proof methods are natural extensions of those of [33] in the univariate case, while others are specific to the bivariate case. In particular, in the proof of Lemma 3, we provide asymptotic approximations and central limit theorems for terms involving the interaction of the two correlated Gaussian processes, which is a novelty compared to the univariate case. Also, in the Proof of Theorem 2, specific matrix manipulations are needed to isolate the microergodic parameter estimators in order to prove their asymptotic normality.

The paper falls into the following parts. In Section 2 we characterize the equivalence of Gaussian measures, and describe which covariance parameters are microergodic. In Section 3 we establish the strong consistency of the MLE of the microergodic parameters. Section 4 is devoted to its asymptotic normality. Some technical lemmas are needed in order to prove these results and, in particular, Lemma 3 is essential to prove the asymptotic normality results. The proofs of the technical lemmas are postponed to the appendix. Section 5 provides a simulation study that shows how well the given asymptotic distributions apply to finite sample cases. The final section provides a discussion and open problems for future research.

2. Equivalence of Gaussian measures

First we present some notations used in the whole paper.

If $A = (a_{ij})_{1\le i\le k, 1\le j\le n}$ is a $k \times n$ matrix and $B = (b_{ij})_{1\le i\le p, 1\le j\le q}$ is a $p \times q$ matrix, then the Kronecker product of the two matrices, denoted by $A \otimes B$, is the $kp \times nq$ block matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{k1}B & \dots & a_{kn}B \end{bmatrix}.$$

In the following, we will consider a stationary zero-mean bivariate Gaussian process observed on fixed compact subset T of $\mathbb{R}$, $Z(s) = \{(Z_1(s), Z_2(s))^\top, s \in T\}$ with covariance function indexed by a parameter $\psi = (\sigma_1^2, \sigma_2^2, \rho, \theta)^\top \in \mathbb{R}^4$, given by

$$\mathrm{Cov}_\psi(Z_i(s_l), Z_j(s_m)) = \sigma_i\sigma_j(\rho + (1-\rho)\mathbf{1}_{i=j})e^{-\theta|s_l - s_m|},\ i, j = 1, 2. \tag{2.1}$$

Note that $\sigma_1^2, \sigma_2^2 > 0$ are marginal variances parameters and $\theta > 0$ is a correlation decay parameter. The quantity ρ with $|\rho| < 1$ is the so-called colocated correlation parameter [13], that expresses the correlation between $Z_1(s)$ and $Z_2(s)$ for each s. For $i = 1, 2$, the covariance of the marginal process $Z_i(s)$ is $\mathrm{Cov}_\psi(Z_i(s_l), Z_i(s_m)) = \sigma_i^2 e^{-\theta|s_l - s_m|}$. Such process is known as the Ornstein-Uhlenbeck process and it has been widely used to model physical, biological, social, and many other phenomena. Denote by P_ψ the distribution of the bivariate process Z, under covariance parameter ψ. As we consider fixed domain asymptotic, the process $Z(s)$ is observed at an increasing number of points on a compact set T. Without loss of generality we consider $T = [0, 1]$ and denote by $0 \leq s_1 < \ldots < s_n \leq 1$ the observation points of the process. Let us notice that the points $s_1, \ldots, s_n$ are allowed to be permuted when new points are added and that these points are assumed to be dense in T when n tends towards infinity. The observations can thus be written as $Z_n = (Z_{1,n}^\top, Z_{2,n}^\top)^\top$ with $Z_{i,n} = (Z_i(s_1), \ldots, Z_i(s_n))^\top$ for $i = 1, 2$. Hence the observation vector Z_n follows a centered Gaussian distribution $Z_n \sim N(0, \Sigma(\psi))$ with covariance matrix $\Sigma(\psi) = A \otimes R$, given by

$$A = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}, \quad R = \left[e^{-\theta|s_m - s_l|}\right]_{1\leq m,l\leq n}, \tag{2.2}$$

and the associated likelihood function is given by

$$f_n(\psi) = (2\pi)^{-n}|\Sigma(\psi)|^{-1/2} e^{-\frac{1}{2}Z_n{}^\top \Sigma(\psi)^{-1} Z_n}. \tag{2.3}$$

The aim of this section is to provide a necessary and sufficient condition to warrant equivalence between two Gaussian measures P_{ψ_1} and P_{ψ_2} with $\psi_i = (\sigma_{i,1}^2, \sigma_{i,2}^2, \rho_i, \theta_i)^\top$, $i = 1, 2$.

Specifically let us define the symmetrized entropy

$$I_n(P_{\psi_1}, P_{\psi_2}) = E_{\psi_1} \log \frac{f_n(\psi_1)}{f_n(\psi_2)} + E_{\psi_2} \log \frac{f_n(\psi_2)}{f_n(\psi_1)}. \tag{2.4}$$

We assume in this section that the observation points are the terms of a growing sequence in the sense that, at each step, new points are added to the sampling scheme but none is deleted. This assumption ensures that $I_n(P_{\psi_1}, P_{\psi_2})$ is an increasing sequence.

Hence we may define the limit $I(P_{\psi_1}, P_{\psi_2}) = \lim_{n\to\infty} I_n(P_{\psi_1}, P_{\psi_2})$, possibly infinite. Then P_{ψ_1} and P_{ψ_2} are either equivalent or orthogonal if and only if $I(P_{\psi_1}, P_{\psi_2}) < \infty$ or $I(P_{\psi_1}, P_{\psi_2}) = \infty$ respectively (see Lemma 3 in page 77 of [14] whose arguments can be immediately extended to the multivariate case). Using this criterion, the following lemma characterizes the equivalence of the Gaussian measures P_{ψ_1} and P_{ψ_2}.

Lemma 1. *The two measures P_{ψ_1} and P_{ψ_2} are equivalent on the σ-algebra generated by $\{Z(s),\, s \in T\}$, if and only if $\sigma_{i,1}^2\theta_1 = \sigma_{i,2}^2\theta_2$, $i = 1, 2$ and $\rho_1 = \rho_2$ and orthogonal otherwise.*

Proof. Let us introduce $\Delta_i = s_i - s_{i-1}$ for $i = 2, \ldots, n$ and note that

$$\sum_{i=2}^{n} \Delta_i \leq 1 \quad \text{and} \quad \lim_{n\to\infty} \max_{2\leq i\leq n} \Delta_i = 0.$$

Let $R_j = \left[e^{-\theta_j|s_m - s_l|}\right]_{1\leq m,l\leq n}$, $j = 1, 2$. By expanding (2.4) we find that

$$I_n(P_{\psi_1}, P_{\psi_2}) \quad = \quad \frac{1}{2}\left\{\frac{1}{(1-\rho_2^2)}\left[\frac{\sigma_{1,1}^2}{\sigma_{1,2}^2} - 2\frac{\sigma_{1,1}\sigma_{2,1}\rho_1\rho_2}{\sigma_{1,2}\sigma_{2,2}} + \frac{\sigma_{2,1}^2}{\sigma_{2,2}^2}\right] tr(R_1 R_2^{-1})\right.$$

$$+ \frac{1}{(1-\rho_1^2)}\left[\frac{\sigma_{1,2}^2}{\sigma_{1,1}^2} - 2\frac{\sigma_{1,2}\sigma_{2,2}\rho_1\rho_2}{\sigma_{1,1}\sigma_{2,1}} + \frac{\sigma_{2,2}^2}{\sigma_{2,1}^2}\right] tr(R_2R_1^{-1})\Bigg\}$$
$$-2n.$$

If $\sigma_{i,1}^2\theta_1 = \sigma_{i,2}^2\theta_2$ and $\rho_1 = \rho_2$ for $i = 1, 2$ we obtain

$$I_n(P_{\psi_1}, P_{\psi_2}) = \frac{\theta_2}{\theta_1} tr(R_1R_2^{-1}) + \frac{\theta_1}{\theta_2} tr(R_2R_1^{-1}) - 2n.$$

In order to compute $tr(R_1R_2^{-1})$ and $tr(R_2R_1^{-1})$, we use some results in [3]. The matrix R_j can be written as follows,

$$R_j = \begin{pmatrix} 1 & e^{-\theta_j\Delta_2} & \cdots & e^{-\theta_j\sum_{i=2}^{n}\Delta_i} \\ e^{-\theta_j\Delta_2} & 1 & \cdots & e^{-\theta_j\sum_{i=3}^{n}\Delta_i} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-\theta_j\sum_{i=2}^{n}\Delta_i} & e^{-\theta_j\sum_{i=3}^{n}\Delta_i} & \cdots & 1 \end{pmatrix}$$

and R_j^{-1} can be written as

$$\begin{pmatrix} \frac{1}{1-e^{-2\theta_j\Delta_2}} & \frac{-e^{-\theta_j\Delta_2}}{1-e^{-2\theta_j\Delta_2}} & 0 & \cdots & 0 \\ \frac{-e^{-\theta_j\Delta_2}}{1-e^{-2\theta_j\Delta_2}} & \frac{1}{1-e^{-2\theta_j\Delta_2}} + \frac{e^{-2\theta_j\Delta_3}}{1-e^{-2\theta_j\Delta_3}} & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & & 0 \\ \vdots & \ddots & & \frac{1}{1-e^{-2\theta_j\Delta_{n-1}}} + \frac{e^{-2\theta_j\Delta_n}}{1-e^{-2\theta_j\Delta_n}} & \frac{-e^{-\theta_j\Delta_n}}{1-e^{-2\theta_j\Delta_n}} \\ 0 & \cdots & 0 & \frac{-e^{-\theta_j\Delta_n}}{1-e^{-2\theta_j\Delta_n}} & \frac{1}{1-e^{-2\theta_j\Delta_n}} \end{pmatrix}.$$

Since, $tr(R_jR_k^{-1}) = \sum_{i=1}^{n}\sum_{m=1}^{n}(R_j \otimes R_k^{-1})_{im},\ j,k = 1,2, j \neq k$, we have

$$\begin{aligned} tr(R_jR_k^{-1}) &= -2\sum_{i=2}^{n}\frac{e^{-(\theta_j+\theta_k)\Delta_i}}{1-e^{-2\theta_k\Delta_i}} + \sum_{i=2}^{n}\frac{1}{1-e^{-2\theta_k\Delta_i}} \\ &\quad + \sum_{i=3}^{n}\frac{e^{-2\theta_k\Delta_i}}{1-e^{-2\theta_k\Delta_i}} + \frac{1}{1-e^{-2\theta_k\Delta_2}} \\ &= \sum_{i=2}^{n}\frac{-2e^{-(\theta_j+\theta_k)\Delta_i} + 1 + e^{-2\theta_k\Delta_i}}{1-e^{-2\theta_k\Delta_i}} + 1 \\ &= \sum_{i=2}^{n}\frac{(e^{-\theta_k\Delta_i} - e^{-\theta_j\Delta_i})^2}{1-e^{-2\theta_k\Delta_i}} + \sum_{i=2}^{n}\frac{1-e^{-2\theta_j\Delta_i}}{1-e^{-2\theta_k\Delta_i}} + 1. \end{aligned}$$

Then, we can write $I_n(P_{\psi_1}, P_{\psi_2})$ as

$$I_n(P_{\psi_1}, P_{\psi_2}) = \frac{\theta_2}{\theta_1}\left(\sum_{i=2}^{n}\frac{(e^{-\theta_1\Delta_i} - e^{-\theta_2\Delta_i})^2}{1-e^{-2\theta_2\Delta_i}} + \sum_{i=2}^{n}\frac{1-e^{-2\theta_1\Delta_i}}{1-e^{-2\theta_2\Delta_i}} + 1\right)$$

$$+\frac{\theta_1}{\theta_2}\left(\sum_{i=2}^{n}\frac{(e^{-\theta_2\Delta_i}-e^{-\theta_1\Delta_i})^2}{1-e^{-2\theta_1\Delta_i}}+\sum_{i=2}^{n}\frac{1-e^{-2\theta_2\Delta_i}}{1-e^{-2\theta_1\Delta_i}}+1\right)-2n.$$

For $j,k=1,2$, $j\neq k$, as is obtained by Taylor expansion, since $\max_i \Delta_i$ tends to 0, we have

$$\max_{2\le i\le n}\left|\frac{1-e^{-2\theta_j\Delta_i}}{\Delta_i(1-e^{-2\theta_k\Delta_i})}-\frac{\theta_j}{\Delta_i\theta_k}\right|=O(1) \text{ and } \max_{2\le i\le n}\frac{(e^{-\theta_j\Delta_i}-e^{-\theta_k\Delta_i})^2}{\Delta_i(1-e^{-2\theta_k\Delta_i})}=O(1)$$

Since $\sum_i \Delta_i$ tends to 1,

$$I_n(P_{\psi_1},P_{\psi_2})=\frac{\theta_2}{\theta_1}(1+O(1))+\frac{\theta_1}{\theta_2}(1+O(1))$$

and $I(P_{\psi_1},P_{\psi_2})=\lim_{n\to\infty}I_n(P_{\psi_1},P_{\psi_2})<\infty$.

Then the two Gaussian measures P_{ψ_1} and P_{ψ_2} are equivalent on the $\sigma-$algebra generated by Z if and only if $\sigma_{i,1}^2\theta_1=\sigma_{i,2}^2\theta_2$, $i=1,2$, and $\rho_1=\rho_2$. □

Note that sufficient conditions for the equivalence of Gaussian measures using a generalization of the covariance model (2.1) are given in [36]. A consequence of the previous lemma is that it is not possible to estimate consistently all the parameters individually if the data are observed on a compact set T. However the microergodic parameters $\sigma_1^2\theta$, $\sigma_2^2\theta$ and ρ are consistently estimable. The following section is devoted to their estimation.

3. Consistency of the maximum likelihood estimator

Let $\widehat{\psi}=(\hat\theta,\hat\sigma_1^2,\hat\sigma_2^2,\hat\rho)$ be the MLE obtained by maximizing $f_n(\psi)$ with respect to ψ. In the rest of the paper, we will denote by θ_0, σ_{i0}^2, $i=1,2$ and ρ_0 the true but unknown parameters that have to be estimated. We let $var=var_{\psi_0}$, $cov=cov_{\psi_0}$ and $\mathbb{E}=\mathbb{E}_{\psi_0}$ denote the variance, covariance and expectation under P_{ψ_0}. In this section, we establish the strong consistency of the MLE of the microergodic parameters $\hat\rho$, $\hat\theta\hat\sigma_1^2$ and $\hat\theta\hat\sigma_2^2$.

We first consider an explicit expression for the negative log-likelihood function

$$l_n(\psi)=-2\log(f_n(\psi))=2n\log(2\pi)+\log|\Sigma(\psi)|+Z_n^\top\left[\Sigma(\psi)\right]^{-1}Z_n. \tag{3.1}$$

The explicit expression is given in the following lemma whose proof can be found in the appendix.

Lemma 2. *The negative log-likelihood function in Equation (3.1) can be written as*

$$l_n(\psi)=n\left[\log(2\pi)+\log(1-\rho^2)\right]+\sum_{k=1}^{2}\sum_{i=2}^{n}\log\left[\sigma_k^2\left(1-e^{-2\theta\Delta_i}\right)\right]$$
$$+\sum_{k=1}^{2}\log(\sigma_k^2)+\frac{1}{1-\rho^2}\left\{\sum_{k=1}^{2}\frac{1}{\sigma_k^2}\left(z_{k,1}^2+\sum_{i=2}^{n}\frac{\left(z_{k,i}-e^{-\theta\Delta_i}z_{k,i-1}\right)^2}{1-e^{-2\theta\Delta_i}}\right)\right.$$

$$- \frac{2\rho}{\sigma_1\sigma_2}\left(z_{1,1}z_{2,1} + \sum_{i=2}^{n}\frac{\left(z_{1,i} - e^{-\theta\Delta_i}z_{1,i-1}\right)\left(z_{2,i} - e^{-\theta\Delta_i}z_{2,i-1}\right)}{1 - e^{-2\theta\Delta_i}}\right)\Bigg\},$$

with $z_{k,i} = Z_k(s_i)$ *and* $\Delta_i = s_i - s_{i-1},\ i = 2, \ldots, n.$

The following theorem uses Lemma 2 in order to establish the strong consistency of MLE of the microergodic parameters ρ, $\theta\sigma_1^2$, $\theta\sigma_2^2$.

Theorem 1. *Let* $J = (a_\theta, b_\theta) \times (a_{\sigma_1}, b_{\sigma_1}) \times (a_{\sigma_2}, b_{\sigma_2}) \times (a_\rho, b_\rho)$, *with* $0 < a_\theta \leq \theta_0 \leq b_\theta < \infty$, $0 < a_{\sigma_1} \leq \sigma_{01}^2 \leq b_{\sigma_1} < \infty$, $0 < a_{\sigma_2} \leq \sigma_{02}^2 \leq b_{\sigma_2} < \infty$ *and* $-1 < a_\rho \leq \rho_0 \leq b_\rho < 1$. *Define* $\widehat{\psi} = (\hat{\theta}, \hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\rho})$ *as the minimum of the negative log-likelihood estimator, solution of*

$$l_n(\widehat{\psi}) = \min_{\psi \in J} l_n(\psi). \tag{3.2}$$

Then, with probability one, $\widehat{\psi}$ *exists for* n *large enough and when* $n \to +\infty$

$$\hat{\rho} \xrightarrow{a.s} \rho_0, \tag{3.3}$$

$$\hat{\theta}\hat{\sigma}_1^2 \xrightarrow{a.s} \theta_0\sigma_{01}^2, \tag{3.4}$$

$$\hat{\theta}\hat{\sigma}_2^2 \xrightarrow{a.s} \theta_0\sigma_{02}^2. \tag{3.5}$$

Proof. The proof follows the guideline of the consistency of the maximum likelihood estimation given in [33]. Hence consistency results given in (3.3), (3.4) and (3.5) hold as long as we can prove that there exists $0 < d < D < \infty$ such that for every $\epsilon > 0$, ψ and $\widetilde{\psi}$, with $\|\psi - \widetilde{\psi}\| > \epsilon$

$$\min_{\{\psi \in J,\ \|\psi - \widetilde{\psi}\| > \epsilon\}} \left\{ l_n(\psi) - l_n(\widetilde{\psi}) \right\} \to \infty \ \ a.s. \tag{3.6}$$

where $\widetilde{\psi} = (\widetilde{\theta}, \widetilde{\rho}^2, \widetilde{\sigma}_1^2, \widetilde{\sigma}_2^2)^\top \in J$ can be any nonrandom vector such that

$$\widetilde{\rho} = \rho_0, \quad \widetilde{\theta}\widetilde{\sigma}_1^2 = \theta_0\sigma_{01}^2, \quad \widetilde{\theta}\widetilde{\sigma}_2^2 = \theta_0\sigma_{02}^2.$$

In order to simplify our notation, let $W_{k,i,n} = \frac{z_{k,i} - e^{-\theta_0\Delta_i}z_{k,i-1}}{\left[\sigma_{0k}^2(1 - e^{-2\theta_0\Delta_i})\right]^{\frac{1}{2}}}$, $k = 1, 2$ and $i = 2, ..., n$. By the Markovian and Gaussian properties of Z_1 and Z_2, it follows that for each $i \geq 2$, $W_{k,i,n}$ is independent of $\{Z_{k,j}, j \leq i - 1\}$, $k = 1, 2$. Moreover $\{W_{k,i,n}, 2 \leq i \leq n\}$, $k = 1, 2$ are an i.i.d. sequences of standard Gaussian random variables. Using Lemma 2 we write,

$$\begin{aligned} l_n(\psi) \ = \ & \sum_{k=1}^{2}\sum_{i=2}^{n}\log\left[\sigma_k^2\left(1 - e^{-2\theta\Delta_i}\right)\right] + \frac{1}{1-\rho^2}\sum_{k=1}^{2}\sum_{i=2}^{n}\frac{\left(z_{k,i} - e^{-\theta\Delta_i}z_{k,i-1}\right)^2}{\sigma_k^2(1 - e^{-2\theta\Delta_i})} \\ & + \frac{2\rho}{(1-\rho^2)}\sum_{i=2}^{n}\frac{\left(z_{1,i} - e^{-\theta\Delta_i}z_{1,i-1}\right)\left(z_{2,i} - e^{-\theta\Delta_i}z_{2,i-1}\right)}{\sigma_1\sigma_2(1 - e^{-2\theta\Delta_i})} \\ & + n\log(2\pi) + c(\psi, n), \end{aligned}$$

with $c(\psi,n)=\sum_{k=1}^{2}\log\sigma_k^2+n\log\left(1-\rho^2\right)+\frac{1}{1-\rho^2}\left[\sum_{k=1}^{2}\frac{z_{k,1}^2}{\sigma_k^2}-2\rho\frac{z_{1,1}z_{2,1}}{\sigma_1\sigma_2}\right]$ and from the proof of Theorem 1 in [33], uniformly in $0<\theta\leq d_\theta$ and $\sigma_k^2\in[a_{\sigma_k},b_{\sigma_k}]$, $k=1,2$:

$$\sum_{i=2}^{n}\frac{\left(z_{k,i}-e^{-\theta\Delta_i}z_{k,i-1}\right)^2}{\sigma_k^2(1-e^{-2\theta\Delta_i})}=\sum_{i=2}^{n}\frac{\sigma_{0k}^2(1-e^{-2\theta_0\Delta_i})}{\sigma_k^2(1-e^{-2\theta\Delta_i})}W_{k,i,n}^2+O(n^{\frac{1}{2}}),\ k=1,2. \tag{3.7}$$

Moreover, from Cauchy-Schwarz inequality,

$$\sum_{i=2}^{n}\frac{\left(z_{1,i}-e^{-\theta\Delta_i}z_{1,i-1}\right)\left(z_{2,i}-e^{-\theta\Delta_i}z_{2,i-1}\right)}{\sigma_1\sigma_2(1-e^{-2\theta\Delta_i})}$$
$$\leq\prod_{k=1}^{2}\left(\sum_{i=2}^{n}\frac{\sigma_{0k}^2(1-e^{-2\theta_0\Delta_i})}{\sigma_k^2(1-e^{-2\theta\Delta_i})}W_{k,i,n}^2+O(n^{\frac{1}{2}})\right)^{\frac{1}{2}} \tag{3.8}$$

and from Lemma 2(ii) in [33] uniformly in $\theta\leq R$ and $\sigma_k^2\in[a_{\sigma_k},b_{\sigma_k}]$, for every $\alpha_k>0$, with $k=1,2$,

$$\sum_{i=2}^{n}\frac{\sigma_{0k}^2(1-e^{-2\theta_0\Delta_i})}{\sigma_k^2(1-e^{-2\theta\Delta_i})}W_{k,i,n}^2=\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2\theta}(n-1)+\frac{1}{\theta}O(n^{\frac{1}{2}+\alpha_k}),\ k=1,2. \tag{3.9}$$

Combining (3.7), (3.8) and (3.9), we can write,

$$\begin{aligned}
l_n(\psi)\quad\geq\quad & n\log(2\pi)+c(\psi,n)+\sum_{k=1}^{2}\sum_{i=2}^{n}\log\left[\sigma_k^2\left(1-e^{-2\theta\Delta_i}\right)\right]+\frac{1}{1-\rho^2}O(n^{\frac{1}{2}})\\
&+\frac{1}{1-\rho^2}\sum_{k=1}^{2}\sum_{i=2}^{n}\frac{\sigma_{0k}^2(1-e^{-2\theta_0\Delta_i})}{\sigma_k^2(1-e^{-2\theta\Delta_i})}W_{k,i,n}^2\\
&-\frac{2\rho}{1-\rho^2}n\prod_{k=1}^{2}\left(\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2\theta}+\frac{1}{n\theta}O(n^{\frac{1}{2}+\alpha_k})-\frac{1}{n}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2\theta}\right)^{\frac{1}{2}}.
\end{aligned}$$

Therefore

$$\begin{aligned}
l_n(\psi)-l_n(\widetilde{\psi})\geq\ & p(\psi,\widetilde{\psi},n)+\sum_{k=1}^{2}\sum_{i=2}^{n}log\left[\frac{\sigma_k^2(1-e^{-2\theta\Delta_i})}{\widetilde{\sigma}_k^2(1-e^{-2\widetilde{\theta}\Delta_i})}\right]\\
&+\sum_{k=1}^{2}\sum_{i=2}^{n}\left[\frac{1}{1-\rho^2}\frac{\sigma_{0k}^2(1-e^{-2\theta_0\Delta_i})}{\sigma_k^2(1-e^{-2\theta\Delta_i})}-\frac{1}{1-\widetilde{\rho}^2}\frac{\sigma_{0k}^2(1-e^{-2\theta_0\Delta_i})}{\widetilde{\sigma}_k^2(1-e^{-2\widetilde{\theta}\Delta_i})}\right]W_{k,i,n}^2\\
&-\frac{2\rho}{1-\rho^2}\frac{n}{\theta}\prod_{k=1}^{2}\left(\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}+\frac{1}{n}O(n^{\frac{1}{2}+\alpha_k})-\frac{1}{n}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}\right)^{\frac{1}{2}}\\
&+\frac{2\widetilde{\rho}}{1-\widetilde{\rho}^2}n\prod_{k=1}^{2}\left(1+\frac{1}{n\widetilde{\theta}}O(n^{\frac{1}{2}+\alpha_k})-\frac{1}{n}\right)^{\frac{1}{2}}
\end{aligned}$$

$$+\frac{1}{1-\rho^2}O(n^{\frac{1}{2}})+\frac{1}{1-\widetilde{\rho}^2}O(n^{\frac{1}{2}}), \tag{3.10}$$

where $p(\psi,\widetilde{\psi},n)=c(\psi,n)-c(\widetilde{\psi},n)$. From lemma 2 in [33], for some $M_k>0$ and uniformly in $\theta\le R$ and $\sigma_k^2\in[a_{\sigma_k},b_{\sigma_k}]$, $k=1,2$

$$\sum_{i=2}^{n}\log\left[\frac{\sigma_k^2(1-e^{-2\theta\Delta_i})}{\widetilde{\sigma}_k^2(1-e^{-2\widetilde{\theta}\Delta_i})}\right]\ge\sum_{i=2}^{n}\log\left(\frac{\theta}{M_k}\right)=(n-1)\log\left(\frac{\theta}{M_k}\right),\ k=1,2. \tag{3.11}$$

and

$$\sum_{i=2}^{n}\left[\frac{1}{1-\rho^2}\frac{\sigma_{0k}^2(1-e^{-2\theta_0\Delta_i})}{\sigma_k^2(1-e^{-2\theta\Delta_i})}-\frac{1}{1-\widetilde{\rho}^2}\frac{\sigma_{0k}^2(1-e^{-2\theta_0\Delta_i})}{\widetilde{\sigma}_k^2(1-e^{-2\widetilde{\theta}\Delta_i})}\right]W_{k,i,n}^2$$
$$=(n-1)\left(\frac{1}{1-\rho^2}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2\theta}-\frac{1}{1-\widetilde{\rho}^2}\right)+\theta^{-1}O(n^{\frac{1}{2}+\alpha_k}), \tag{3.12}$$

for $k=1,2$.

Let $\widetilde{\rho}=\min\{\widetilde{\rho},\rho\}$ and combining (3.11) and (3.12), we can write,

$$\begin{aligned}&l_n(\theta,\rho,\sigma_1^2,\sigma_2^2)-l_n(\widetilde{\theta},\widetilde{\rho},\widetilde{\sigma}_1^2,\widetilde{\sigma}_2^2)\\&\ge p(\psi,\widetilde{\psi},n)+\sum_{k=1}^{2}\frac{n}{\theta(1-\rho_0^2)}\left[\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}+O(n^{\gamma_k-1})\right]-\sum_{k=1}^{2}\frac{1}{1-\rho_0^2}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2\theta}\\&\quad+\sum_{k=1}^{2}(n-1)\left[log\left(\frac{M_k}{\theta}\right)-\frac{1}{1-\rho_0^2}\right]+\frac{2\rho_0}{1-\rho_0^2}n\prod_{k=1}^{2}\left(1+\frac{1}{n\widetilde{\theta}}O(n^{\frac{1}{2}+\alpha_k})-\frac{1}{n}\right)^{\frac{1}{2}}\\&\quad-\frac{2\rho_0}{1-\rho_0^2}\frac{n}{\theta}\prod_{k=1}^{2}\left(\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}+\frac{1}{n}O(n^{\frac{1}{2}+\alpha_k})-\frac{1}{n}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}\right)^{\frac{1}{2}}.\end{aligned}$$

for some $\gamma_k<1,\ k=1,2$, where the $O(n^{\gamma_k})$ term is uniform in $\theta\le R$.

Since some $log(\theta)^{-1}=o(\theta^{-1})$ as $\theta\downarrow 0$, we can choose δ_θ small enough so that for all $\theta\le\delta_\theta$, $\frac{\sigma_0^2\theta_0}{2b_\sigma\theta}-1-log(\frac{M}{\theta})\ge\eta$, which implies that with probability 1,

$$\begin{aligned}&l_n(\theta,\rho,\sigma_1^2,\sigma_2^2)-l_n(\widetilde{\theta},\widetilde{\rho},\widetilde{\sigma}_1^2,\widetilde{\sigma}_2^2)\\&\ge\frac{n}{(1-\rho_0^2)}\left\{\frac{1}{\theta}\sum_{k=1}^{2}\frac{\sigma_{0k}^2\theta_0}{2b_{\sigma_k}}+O(n^{\gamma_k-1})-2\rho_0\left[\prod_{k=1}^{2}\frac{1}{\theta}\left(\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}+O(n^{\alpha_k-\frac{1}{2}})-\frac{1}{n}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}\right)^{\frac{1}{2}}\right.\right.\\&\qquad\left.\left.-\prod_{k=1}^{2}\left(1+\frac{1}{\widetilde{\theta}}O(n^{\alpha_k-\frac{1}{2}})-\frac{1}{n}\right)^{\frac{1}{2}}\right]\right\}+p(\psi,\widetilde{\psi},n).\end{aligned}$$

Thus we get (3.6) by letting $d_\theta=\delta_\theta$.

Hence, since $p(\psi,\widetilde{\psi},n)\to 0$,

$$\frac{1}{\theta}\sum_{k=1}^{2}\frac{\sigma_{0k}^2\theta_0}{2b_{\sigma_k}}+O(n^{\gamma_k-1})\le\infty$$

and

$$-2\rho_0\left[\prod_{k=1}^{2}\frac{1}{\theta}\left(\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}+O(n^{\alpha_k-\frac{1}{2}})-\frac{1}{n}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}\right)^{\frac{1}{2}}-\prod_{k=1}^{2}\left(1+\frac{1}{\widetilde{\theta}}O(n^{\alpha_k-\frac{1}{2}})-\frac{1}{n}\right)^{\frac{1}{2}}\right]\leq\infty,$$

we prove that

$$\min_{\{\psi\in J,\,\|\psi-\widetilde{\psi}\|>\epsilon\}}\left\{l_n(\theta,\rho,\sigma_1^2,\sigma_2^2)-l_n(\widetilde{\theta},\widetilde{\rho},\widetilde{\sigma}_1^2,\widetilde{\sigma}_2^2)\right\}\to\infty$$

when $n\to\infty$, uniformly in $\theta\leq\delta_\theta$. □

4. Asymptotic distribution

Before we state the main result on the MLE asymptotic distribution, we need to introduce some notation that will be used throughout this paper. Because of Theorem 1, there exists a compact subset $\mathcal{S}$ of $(0,+\infty)\times(0,+\infty)\times(0,+\infty)\times(-1,1)$ of the form $\Theta\times\mathcal{V}\times\mathcal{V}\times\mathcal{R}$, such that a.s. $\hat{\psi}$ belongs to $\mathcal{S}$ for n large enough. We let $O_u(1)$ denote any real function $g_n(\theta,\rho,\sigma_1^2,\sigma_2^2)$ which satisfies $\sup_{(\theta,\rho,\sigma_1^2,\sigma_2^2)\in\mathcal{S}}|g_n(\theta,\rho,\sigma_1^2,\sigma_2^2)|=O(1)$. For example $\theta\sigma_1/(1-\rho^2)=O_u(1)$. We also let $O_{up}(1)$ denote any real function $g_n(\theta,\rho,\sigma_1^2,\sigma_2^2,Z_{1,n},Z_{2,n})$ which satisfies

$$\sup_{(\theta,\rho,\sigma_1^2,\sigma_2^2)\in\mathcal{S}}|g_n(\theta,\rho,\sigma_1^2,\sigma_2^2,Z_{1,n},Z_{2,n})|=O_p(1).$$

For example $z_{1,1}\sigma_1=O_{up}(1)$.

The following lemma is essential when establishing the asymptotic distribution of the microergodic parameters.

Lemma 3. *With the same notations and assumptions as in Theorem 1, let*

$$L(\theta)=\sum_{i=2}^{n}\frac{(z_{1,i}-e^{-\theta\Delta_i}z_{1,i-1})(z_{2,i}-e^{-\theta\Delta_i}z_{2,i-1})}{1-e^{-2\theta\Delta_i}},$$

and $G=[\partial/\partial\theta]L(\theta)$. Let for $n\in\mathbb{N}$ and $i=2,...,n$,

$$Y_{i,n}=\frac{\left(z_{1,i}-e^{-\theta_0\Delta_i}z_{1,i-1}\right)\left(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1}\right)}{\sigma_{01}\sigma_{02}\sqrt{1+\rho_0^2}(1-e^{-2\theta_0\Delta_i})}.\tag{4.1}$$

Then for all $n\in\mathbb{N}$, the $(Y_{i,n})_{i=2,\ldots,n}$ are independent with $\mathbb{E}(Y_{i,n})=\rho_0/(1+\rho_0^2)^{1/2}$ and $var(Y_{i,n})=1$. Furthermore we have

$$G=-\frac{\sigma_{01}\sigma_{02}\sqrt{1+\rho_0^2}\theta_0}{\theta^2}\sum_{i=2}^{n}Y_{i,n}+O_{up}(1).$$

Using the previous lemma, the following theorem establishes the asymptotic distribution of the MLE of the microergodic parameters. Specifically we consider three cases: first when both the colocated correlation and variance parameters are known, second when only the variance parameters are known and third when all the microergodic parameters are unknown.

Theorem 2. *With the same notation and assumptions as in Theorem 1, if* $a_{\sigma_k} = b_{\sigma_k} = \sigma_{0k}^2 = \hat{\sigma}_k^2$ *for* $k = 1, 2$, $a_\rho = b_\rho = \rho_0 = \hat{\rho}$ *and* $a_\theta < \theta_0 < b_\theta$ *then*

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \theta_0^2\right). \tag{4.2}$$

If $a_{\sigma_k} = b_{\sigma_k} = \sigma_{0k}^2 = \hat{\sigma}_k^2$ *for* $k = 1, 2$, $a_\rho < \rho_0 < b_\rho$ *and* $a_\theta < \theta_0 < b_\theta$, *then*

$$\sqrt{n}\begin{pmatrix}\hat{\theta} - \theta_0 \\ \hat{\rho} - \rho_0\end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \Sigma_{\theta\rho}\right), \tag{4.3}$$

where $\Sigma_{\theta\rho} = \begin{pmatrix}\theta_0^2(1+\rho_0^2) & \theta_0\rho_0(1-\rho_0^2) \\ \theta_0\rho_0(1-\rho_0^2) & (\rho_0^2-1)^2\end{pmatrix}$.

Finally, if $a_{\sigma_k} < \sigma_{0k}^2 < b_{\sigma_k}$ *for* $k = 1, 2$, $a_\rho < \rho_0 < b_\rho$ *and* $a_\theta < \theta_0 < b_\theta$, *then*

$$\sqrt{n}\begin{pmatrix}\hat{\sigma}_1^2\hat{\theta} - \sigma_{01}^2\theta_0 \\ \hat{\sigma}_2^2\hat{\theta} - \sigma_{02}^2\theta_0 \\ \hat{\rho} - \rho_0\end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \Sigma_f\right), \tag{4.4}$$

where $\Sigma_f = \begin{pmatrix}2(\theta_0\sigma_{01}^2)^2 & 2(\theta_0\rho_0\sigma_{01}\sigma_{02})^2 & \theta_0\rho_0\sigma_{01}^2(1-\rho_0^2) \\ 2(\theta_0\rho_0\sigma_{01}\sigma_{02})^2 & 2(\theta_0\sigma_{02}^2)^2 & \theta_0\rho_0\sigma_{02}^2(1-\rho_0^2) \\ \theta_0\rho_0\sigma_{01}^2(1-\rho_0^2) & \theta_0\rho_0\sigma_{02}^2(1-\rho_0^2) & (\rho_0^2-1)^2\end{pmatrix}$.

Proof. Let $s_x(\psi) = \frac{\partial}{\partial x} l_n(\psi)$ the derivative of the negative log-likelihood with respect to $x = \sigma_1^2, \sigma_2^2, \theta, \rho$. From Lemma 3 and from Equation (3.11) in [33] we can write, with $W_{k,i,n}$ as in the proof of Theorem 1,

$$s_\theta(\psi) = \frac{2n}{\theta} - \frac{1}{1-\rho^2}\left(\sum_{k=1}^{2}\sum_{i=2}^{n}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2\theta^2}W_{k,i,n}^2 - 2\rho(1+\rho_0^2)^{\frac{1}{2}}\frac{\sigma_{01}\sigma_{02}\theta_0}{\sigma_1\sigma_2\theta^2}\sum_{i=2}^{n}Y_{i,n}\right) + O_{up}(1). \tag{4.5}$$

Then from (4.5) we have

$$\begin{aligned}\theta^2(1-\rho^2)s_\theta(\psi) \quad = \quad & (n-1)\left[2\theta(1-\rho^2) - \theta_0\left(\frac{\sigma_{01}^2}{\sigma_1^2} - 2\rho\rho_0\frac{\sigma_{01}\sigma_{02}}{\sigma_1\sigma_2} + \frac{\sigma_{02}^2}{\sigma_2^2}\right)\right] \\ & - \sum_{k=1}^{2}\sum_{i=2}^{n}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}\xi_{k,i} + 2\rho(1+\rho_0^2)^{\frac{1}{2}}\theta_0\frac{\sigma_{01}\sigma_{02}}{\sigma_1\sigma_2}\sum_{i=2}^{n}\xi_{3,i} + O_{up}(1),\end{aligned} \tag{4.6}$$

with $\xi_{k,i} = W_{k,i,n}^2 - 1$, $k = 1, 2$ and $\xi_{3,i} = Y_{i,n} - \frac{\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}}$.

Then $\hat{\psi}$ satisfies $s_\theta(\hat{\psi}) = 0$ and in view of (4.6), we get

$$\begin{aligned}0 = \hat{\theta}^2(1-\hat{\rho}^2)s_\theta(\hat{\psi}) \quad = \quad & (n-1)\left[2\hat{\theta}(1-\hat{\rho}^2) - \theta_0\left(\frac{\sigma_{01}^2}{\hat{\sigma}_1^2} - 2\hat{\rho}\rho_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2} + \frac{\sigma_{02}^2}{\hat{\sigma}_2^2}\right)\right] \\ & - \sum_{k=1}^{2}\sum_{i=2}^{n}\frac{\sigma_{0k}^2\theta_0}{\hat{\sigma}_k^2}\xi_{k,i} + 2\hat{\rho}(1+\rho_0^2)^{\frac{1}{2}}\theta_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\sum_{i=2}^{n}\xi_{3,i} + O_p(1).\end{aligned} \tag{4.7}$$

If we set $a_{\sigma_k} = b_{\sigma_k} = \sigma_{0k}^2 = \hat{\sigma}_k^2$ for $k = 1, 2$ and $a_\rho = b_\rho = \rho_0 = \hat{\rho}$ in (4.7), we get

$$0 = 2(n-1)\left[(\hat{\theta} - \theta_0)(1-\rho_0^2)\right] - \theta_0\left[\sum_{k=1}^{2}\sum_{i=2}^{n}\xi_{k,i} - 2\rho_0(1+\rho_0^2)^{\frac{1}{2}}\sum_{i=2}^{n}\xi_{3,i}\right] + O_p(1). \tag{4.8}$$

Hence (4.8) implies

$$\sqrt{n}(\hat{\theta}-\theta_0)=\frac{\theta_0 n^{-\frac{1}{2}}}{2(1-\rho_0^2)}\left[\sum_{k=1}^{2}\sum_{i=2}^{n}\xi_{k,i}-2\rho_0(1+\rho_0^2)^{\frac{1}{2}}\sum_{i=2}^{n}\xi_{3,i}\right]+O_p\left(n^{-\frac{1}{2}}\right).$$

On the other hand, from the multivariate central limit theorem we get

$$n^{-\frac{1}{2}}\sum_{i=2}^{n}\begin{pmatrix}\xi_{1,i}\\ \xi_{2,i}\\ \xi_{3,i}\end{pmatrix}\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\Sigma_\xi\right), \tag{4.9}$$

where

$$\Sigma_\xi=\begin{pmatrix}2 & 2\rho_0^2 & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}}\\ 2\rho_0^2 & 2 & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}}\\ \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} & 1\end{pmatrix},$$

is obtained by calculating $Cov(\xi_{m,i},\xi_{l,i})$ for $m,l=1,2,3$ and $i=2,\ldots,n$.

Hence we have

$$\sqrt{n}(\hat{\theta}-\theta_0)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\Sigma_\theta\right),$$

where

$$\begin{aligned}\Sigma_\theta &= \frac{\theta_0^2}{4(1-\rho_0^2)^2}\begin{pmatrix}1 & 1 & -2\rho_0(1+\rho_0^2)^{\frac{1}{2}}\end{pmatrix}\begin{pmatrix}2 & 2\rho_0^2 & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}}\\ 2\rho_0^2 & 2 & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}}\\ \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} & 1\end{pmatrix}\\ &\quad\times\begin{pmatrix}1\\ 1\\ -2\rho_0(1+\rho_0^2)^{\frac{1}{2}}\end{pmatrix}.\end{aligned} \tag{4.10}$$

Then, computing the previous quadratic form, we get

$$\sqrt{n}(\hat{\theta}-\theta_0)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\theta_0^2\right),$$

so (4.2) is proved. Now, we first prove (4.3) and (4.4) for $\rho_0\in(-1,1)\diagdown\{0\}$ and discuss the case $\rho_0=0$ at the end of the proof. To show (4.3), take differentiation with respect to ρ. From the proof of Theorem 2 given in [33], and from arguments similar to those of the proof of Lemma 3, we get

$$\begin{aligned}s_\rho(\psi) &= \frac{2\rho}{(1-\rho^2)^2}\left(\sum_{k=1}^{2}\sum_{i=2}^{n}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2\theta}W_{k,i,n}^2-\frac{(1+\rho^2)(1+\rho_0^2)^{\frac{1}{2}}\sigma_{01}\sigma_{02}\theta_0}{\rho\sigma_1\sigma_2\theta}\sum_{i=2}^{n}Y_{i,n}\right)\\ &\quad-\frac{2n\rho}{(1-\rho^2)}+O_{up}(1).\end{aligned} \tag{4.11}$$

Then (4.11) implies

$$\begin{aligned}-\frac{(1-\rho^2)^2\theta}{2\rho}s_\rho(\psi) &= (n-1)\left[\theta(1-\rho^2)-\theta_0\left(\frac{\sigma_{01}^2}{\sigma_1^2}-\frac{(1+\rho^2)\rho_0}{\rho}\frac{\sigma_{01}\sigma_{02}}{\sigma_1\sigma_2}+\frac{\sigma_{02}^2}{\sigma_2^2}\right)\right]\\ &\quad-\sum_{k=1}^{2}\sum_{i=2}^{n}\frac{\sigma_{0k}^2\theta_0}{\sigma_k^2}\xi_{k,i}-\frac{(1+\rho^2)(1+\rho_0^2)^{\frac{1}{2}}\theta_0}{\rho}\frac{\sigma_{01}\sigma_{02}}{\sigma_1\sigma_2}\sum_{i=2}^{n}\xi_{3,i}\\ &\quad+O_{up}(1),\end{aligned} \tag{4.12}$$

Then $\hat{\psi}$ satisfies $s_\rho(\hat{\psi}) = 0$ and in view of (4.12), we get

$$\begin{aligned} 0 = -\frac{(1-\hat{\rho}^2)^2\hat{\theta}}{2\hat{\rho}} s_\rho(\hat{\psi}) &= (n-1)\left[\hat{\theta}(1-\hat{\rho}^2) - \theta_0\left(\frac{\sigma_{01}^2}{\hat{\sigma}_1^2} - \frac{(1+\hat{\rho}^2)\rho_0}{\hat{\rho}}\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2} + \frac{\sigma_{02}^2}{\hat{\sigma}_2^2}\right)\right] \\ &\quad - \sum_{k=1}^{2}\sum_{i=2}^{n}\frac{\sigma_{0k}^2\theta_0}{\hat{\sigma}_k^2}\xi_{k,i} - \frac{(1+\hat{\rho}^2)(1+\rho_0^2)^{\frac{1}{2}}\theta_0}{\hat{\rho}}\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\sum_{i=2}^{n}\xi_{3,i} \\ &\quad + O_p(1). \end{aligned} \tag{4.13}$$

Then we can write, from (4.7) and (4.13)

$$\begin{aligned} \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= (n-1)\begin{pmatrix} 2\hat{\theta}(1-\hat{\rho}^2) - \theta_0\left(\frac{\sigma_{01}^2}{\hat{\sigma}_1^2} - 2\hat{\rho}\rho_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2} + \frac{\sigma_{02}^2}{\hat{\sigma}_2^2}\right) \\ \hat{\theta}(1-\hat{\rho}^2) - \theta_0\left(\frac{\sigma_{01}^2}{\hat{\sigma}_1^2} - \frac{(1+\hat{\rho}^2)\rho_0}{\hat{\rho}}\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2} + \frac{\sigma_{02}^2}{\hat{\sigma}_2^2}\right) \end{pmatrix} \\ &\quad - \begin{pmatrix} \frac{\sigma_{01}^2\theta_0}{\hat{\sigma}_1^2} & \frac{\sigma_{02}^2\theta_0}{\hat{\sigma}_2^2} & -2\hat{\rho}(1+\rho_0^2)^{\frac{1}{2}}\theta_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2} \\ \frac{\sigma_{01}^2\theta_0}{\hat{\sigma}_1^2} & \frac{\sigma_{02}^2\theta_0}{\hat{\sigma}_2^2} & -\frac{(1+\hat{\rho}^2)(1+\rho_0^2)^{\frac{1}{2}}\theta_0}{\hat{\rho}}\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2} \end{pmatrix}\begin{pmatrix} \sum_{i=2}^n \xi_{1,i} \\ \sum_{i=2}^n \xi_{2,i} \\ \sum_{i=2}^n \xi_{3,i} \end{pmatrix} + O_p(1). \end{aligned} \tag{4.14}$$

If we set $a_{\sigma_k} = b_{\sigma_k} = \sigma_{0k}^2 = \hat{\sigma}_k^2$ for $k = 1, 2$ in (4.15), we get

$$\begin{aligned} \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= (n-1)\begin{pmatrix} 2\hat{\theta}(1-\hat{\rho}^2) - 2\theta_0(1-\hat{\rho}\rho_0) \\ \hat{\theta}(1-\hat{\rho}^2) - \theta_0\left(2 - \frac{(1+\hat{\rho}^2)\rho_0}{\hat{\rho}}\right) \end{pmatrix} \\ &\quad -\theta_0\begin{pmatrix} 1 & 1 & -2\hat{\rho}(1+\rho_0^2)^{\frac{1}{2}} \\ 1 & 1 & -\frac{(1+\hat{\rho}^2)(1+\rho_0^2)^{\frac{1}{2}}}{\hat{\rho}} \end{pmatrix}\begin{pmatrix} \sum_{i=2}^n \xi_{1,i} \\ \sum_{i=2}^n \xi_{2,i} \\ \sum_{i=2}^n \xi_{3,i} \end{pmatrix} + O_p(1) \\ &= (n-1)\begin{pmatrix} 2(\hat{\theta}-\theta_0) - 2\hat{\rho}\left(\hat{\theta}\hat{\rho} - \theta_0\rho_0\right) \\ (\hat{\theta}-\theta_0) - \hat{\rho}\left(\hat{\theta}\hat{\rho} - \theta_0\rho_0\right) - \frac{\theta_0}{\hat{\rho}}(\hat{\rho}-\rho_0) \end{pmatrix} \\ &\quad -\theta_0\left[\begin{pmatrix} 1 & 1 & -2\rho_0(1+\rho_0^2)^{\frac{1}{2}} \\ 1 & 1 & -\frac{(1+\rho_0^2)^{\frac{3}{2}}}{\rho_0} \end{pmatrix} + o_p(1)\right]\begin{pmatrix} \sum_{i=2}^n \xi_{1,i} \\ \sum_{i=2}^n \xi_{2,i} \\ \sum_{i=2}^n \xi_{3,i} \end{pmatrix} + O_p(1). \end{aligned} \tag{4.15}$$

Furthermore,

$$\begin{aligned} \begin{pmatrix} 2(\hat{\theta}-\theta_0) - 2\hat{\rho}\left(\hat{\theta}\hat{\rho} - \theta_0\rho_0\right) \\ (\hat{\theta}-\theta_0) - \hat{\rho}\left(\hat{\theta}\hat{\rho} - \theta_0\rho_0\right) - \frac{\theta_0}{\hat{\rho}}(\hat{\rho}-\rho_0) \end{pmatrix} &= \begin{pmatrix} 2 & -2\hat{\rho} & 0 \\ 1 & -\hat{\rho} & -\frac{\theta_0}{\hat{\rho}} \end{pmatrix}\begin{pmatrix} \hat{\theta}-\theta_0 \\ \hat{\theta}\hat{\rho} - \theta_0\rho_0 \\ \hat{\rho}-\rho_0 \end{pmatrix} \\ &= \left[\begin{pmatrix} 2 & -2\rho_0 & 0 \\ 1 & -\rho_0 & -\frac{\theta_0}{\rho_0} \end{pmatrix} + o_p(1)\right] \\ &\quad \times \begin{pmatrix} \hat{\theta}-\theta_0 \\ \rho_0(\hat{\theta}-\theta_0) + \hat{\theta}(\hat{\rho}-\rho_0) \\ \hat{\rho}-\rho_0 \end{pmatrix} \\ &= \left[\begin{pmatrix} 2 & -2\rho_0 & 0 \\ 1 & -\rho_0 & -\frac{\theta_0}{\rho_0} \end{pmatrix} + o_p(1)\right] \\ &\quad \times \left[\begin{pmatrix} 1 & 0 \\ \rho_0 & \hat{\theta} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \hat{\theta}-\theta_0 \\ \hat{\rho}-\rho_0 \end{pmatrix}\right] \\ &= \left[\begin{pmatrix} 2 & -2\rho_0 & 0 \\ 1 & -\rho_0 & -\frac{\theta_0}{\rho_0} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \rho_0 & \theta_0 \\ 0 & 1 \end{pmatrix} + o_p(1)\right] \\ &\quad \times \begin{pmatrix} \hat{\theta}-\theta_0 \\ \hat{\rho}-\rho_0 \end{pmatrix} \end{aligned}$$

$$= \left[\begin{pmatrix} 2-2\rho_0^2 & -2\rho_0\theta_0 \\ 1-\rho_0^2 & -\theta_0\rho_0 - \frac{\theta_0}{\rho_0} \end{pmatrix} + o_p(1)\right] \times \begin{pmatrix} \hat{\theta}-\theta_0 \\ \hat{\rho}-\rho_0 \end{pmatrix}. \tag{4.16}$$

By taking the inverse of the 2×2 matrix in (4.16), we get from (4.15):

$$\sqrt{n}\begin{pmatrix} \hat{\theta}-\theta_0 \\ \hat{\rho}-\rho_0 \end{pmatrix} = \theta_0 n^{-\frac{1}{2}} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{\rho_0}{2\theta_0} & -\frac{\rho_0}{2\theta_0} & \frac{\sqrt{1+\rho_0^2}}{\theta_0} \end{pmatrix} \begin{pmatrix} \sum_{i=2}^n \xi_{1,i} \\ \sum_{i=2}^n \xi_{2,i} \\ \sum_{i=2}^n \xi_{3,i} \end{pmatrix} + o_p(1).$$

From (4.9) we can get

$$\sqrt{n}\begin{pmatrix} \hat{\theta}-\theta_0 \\ \hat{\rho}-\rho_0 \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \Sigma_{\theta\rho}\right),$$

where

$$\Sigma_{\theta\rho} = \theta_0^2 \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{\rho_0}{2\theta_0} & -\frac{\rho_0}{2\theta_0} & \frac{\sqrt{1+\rho_0^2}}{\theta_0} \end{pmatrix} \begin{pmatrix} 2 & 2\rho_0^2 & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} \\ 2\rho_0^2 & 2 & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} \\ \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} & 1 \end{pmatrix} \times \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{\rho_0}{2\theta_0} & -\frac{\rho_0}{2\theta_0} & \frac{\sqrt{1+\rho_0^2}}{\theta_0} \end{pmatrix}^{\top}$$

Then, we get that

$$\sqrt{n}\begin{pmatrix} \hat{\theta}-\theta_0 \\ \hat{\rho}-\rho_0 \end{pmatrix} \underset{\mathcal{D}}{\longrightarrow} \mathcal{N}\left(0, \begin{pmatrix} \theta_0^2(1+\rho_0^2) & \theta_0\rho_0(1-\rho_0^2) \\ \theta_0\rho_0(1-\rho_0^2) & (-1+\rho_0^2)^2 \end{pmatrix}\right). \tag{4.17}$$

Let us now show (4.4). Similarly as for (4.11), we can show

$$s_{\sigma_1^2}(\psi) = \frac{n}{\sigma_1^2} - \frac{\sigma_{01}^2\theta_0}{(1-\rho^2)\sigma_1^4\theta}\sum_{i=2}^n W_{1,i,n}^2 + \rho(1+\rho_0^2)^{\frac{1}{2}}\frac{\sigma_{01}\sigma_{02}\theta_0}{(1-\rho^2)\sigma_1^3\sigma_2\theta}\sum_{i=2}^n Y_{i,n} + O_{up}(1). \tag{4.18}$$

Then (4.18) implies

$$\sigma_1^2(1-\rho^2)\theta s_{\sigma_1^2}(\psi) = (n-1)\left[\theta(1-\rho^2) - \theta_0\left(\frac{\sigma_{01}^2}{\sigma_1^2} - \rho\rho_0\frac{\sigma_{01}\sigma_{02}}{\sigma_1\sigma_2}\right)\right] - \frac{\sigma_{01}^2\theta_0}{\sigma_1^2}\sum_{i=2}^n \xi_{1,i} + \rho(1+\rho_0^2)^{\frac{1}{2}}\theta_0\frac{\sigma_{01}\sigma_{02}}{\sigma_1\sigma_2}\sum_{i=2}^n \xi_{3,i} + O_{up}(1). \tag{4.19}$$

Then $\hat{\psi}$ satisfies $s_{\sigma_1^2}(\hat{\psi}) = 0$ and in view of (4.19), we get

$$0 = \hat{\sigma}_1^2(1-\hat{\rho}^2)\hat{\theta}s_{\sigma_1^2}(\hat{\psi}) = (n-1)\left[\hat{\theta}(1-\hat{\rho}^2) - \theta_0\left(\frac{\sigma_{01}^2}{\hat{\sigma}_1^2} - \hat{\rho}\rho_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\right)\right] - \frac{\sigma_{01}^2\theta_0}{\hat{\sigma}_1^2}\sum_{i=2}^n \xi_{1,i} + \hat{\rho}(1+\rho_0^2)^{\frac{1}{2}}\theta_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\sum_{i=2}^n \xi_{3,i} + O_p(1). \tag{4.20}$$

Then we can write, from (4.7), (4.13) and (4.21)

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = (n-1)\begin{pmatrix} 2\hat{\theta}(1-\hat{\rho}^2)-\theta_0\left(\frac{\sigma_{01}^2}{\hat{\sigma}_1^2}-2\hat{\rho}\rho_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}+\frac{\sigma_{02}^2}{\hat{\sigma}_2^2}\right)\\ \hat{\theta}(1-\hat{\rho}^2)-\theta_0\left(\frac{\sigma_{01}^2}{\hat{\sigma}_1^2}-\frac{(1+\hat{\rho}^2)\rho_0}{\hat{\rho}}\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}+\frac{\sigma_{02}^2}{\hat{\sigma}_2^2}\right)\\ \hat{\theta}(1-\hat{\rho}^2)-\theta_0\left(\frac{\sigma_{01}^2}{\hat{\sigma}_1^2}-\hat{\rho}\rho_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\right)\end{pmatrix} \tag{4.21}$$

$$-\begin{pmatrix}\frac{\sigma_{01}^2\theta_0}{\hat{\sigma}_1^2} & \frac{\sigma_{02}^2\theta_0}{\hat{\sigma}_2^2} & -2\hat{\rho}(1+\rho_0^2)^{\frac{1}{2}}\theta_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\\ \frac{\sigma_{01}^2\theta_0}{\hat{\sigma}_1^2} & \frac{\sigma_{02}^2\theta_0}{\hat{\sigma}_2^2} & -\frac{(1+\hat{\rho}^2)(1+\rho_0^2)^{\frac{1}{2}}\theta_0}{\hat{\rho}}\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\\ \frac{\sigma_{01}^2\theta_0}{\hat{\sigma}_1^2} & 0 & -\hat{\rho}(1+\rho_0^2)^{\frac{1}{2}}\theta_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\end{pmatrix}\begin{pmatrix}\sum_{i=2}^n\xi_{1,i}\\ \sum_{i=2}^n\xi_{2,i}\\ \sum_{i=2}^n\xi_{3,i}\end{pmatrix}+O_p(1).$$

If all parameters are uknown, we get after some tedious algebra:

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = (n-1)\begin{pmatrix}\frac{1}{\hat{\sigma}_1^2} & \frac{1}{\hat{\sigma}_2^2} & -\frac{2\hat{\rho}}{\hat{\sigma}_1\hat{\sigma}_2}\\ \frac{1}{\hat{\sigma}_1^2} & \frac{1}{\hat{\sigma}_2^2} & -\frac{(\hat{\rho}^2+1)}{\hat{\rho}\hat{\sigma}_1\hat{\sigma}_2}\\ \frac{1}{\hat{\sigma}_1^2} & 0 & -\frac{\hat{\rho}}{\hat{\sigma}_1\hat{\sigma}_2}\end{pmatrix}\begin{pmatrix}\hat{\theta}\hat{\sigma}_1^2-\theta_0\sigma_{01}^2\\ \hat{\theta}\hat{\sigma}_2^2-\theta_0\sigma_{02}^2\\ \hat{\theta}\hat{\rho}\hat{\sigma}_1\hat{\sigma}_2-\theta_0\rho_0\sigma_{01}\sigma_{02}\end{pmatrix}$$

$$-\begin{pmatrix}\frac{\sigma_{01}^2\theta_0}{\hat{\sigma}_1^2} & \frac{\sigma_{02}^2\theta_0}{\hat{\sigma}_2^2} & -2\hat{\rho}(1+\rho_0^2)^{\frac{1}{2}}\theta_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\\ \frac{\sigma_{01}^2\theta_0}{\hat{\sigma}_1^2} & \frac{\sigma_{02}^2\theta_0}{\hat{\sigma}_2^2} & -\frac{(1+\hat{\rho}^2)(1+\rho_0^2)^{\frac{1}{2}}\theta_0}{\hat{\rho}}\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\\ \frac{\sigma_{01}^2\theta_0}{\hat{\sigma}_1^2} & 0 & -\hat{\rho}(1+\rho_0^2)^{\frac{1}{2}}\theta_0\frac{\sigma_{01}\sigma_{02}}{\hat{\sigma}_1\hat{\sigma}_2}\end{pmatrix}\begin{pmatrix}\sum_{i=2}^n\xi_{1,i}\\ \sum_{i=2}^n\xi_{2,i}\\ \sum_{i=2}^n\xi_{3,i}\end{pmatrix}+O_p(1).$$

Applying LU matrix factorization we get

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = (n-1)\begin{pmatrix}\frac{1}{\hat{\sigma}_1\hat{\sigma}_2} & 0 & 0\\ \frac{1}{\hat{\sigma}_1\hat{\sigma}_2} & \frac{1}{\hat{\sigma}_1\hat{\sigma}_2} & 0\\ \frac{1}{\hat{\sigma}_1\hat{\sigma}_2} & 0 & \frac{1}{\hat{\sigma}_1\hat{\sigma}_2}\end{pmatrix}\begin{pmatrix}\frac{\hat{\sigma}_2}{\hat{\sigma}_1} & \frac{\hat{\sigma}_1}{\hat{\sigma}_2} & -2\hat{\rho}\\ 0 & 0 & \frac{\hat{\rho}^2-1}{\hat{\rho}}\\ 0 & -\frac{\hat{\sigma}_1}{\hat{\sigma}_2} & \hat{\rho}\end{pmatrix}\begin{pmatrix}\hat{\theta}\hat{\sigma}_1^2-\theta_0\sigma_{01}^2\\ \hat{\theta}\hat{\sigma}_2^2-\theta_0\sigma_{02}^2\\ \hat{\theta}\hat{\rho}\hat{\sigma}_1\hat{\sigma}_2-\theta_0\rho_0\sigma_{01}\sigma_{02}\end{pmatrix}$$

$$-\theta_0\begin{pmatrix}\frac{1}{\hat{\sigma}_1\hat{\sigma}_2} & 0 & 0\\ \frac{1}{\hat{\sigma}_1\hat{\sigma}_2} & \frac{1}{\hat{\sigma}_1\hat{\sigma}_2} & 0\\ \frac{1}{\hat{\sigma}_1\hat{\sigma}_2} & 0 & \frac{1}{\hat{\sigma}_1\hat{\sigma}_2}\end{pmatrix}\begin{pmatrix}\sigma_{01}^2\frac{\hat{\sigma}_2}{\hat{\sigma}_1} & \sigma_{02}^2\frac{\hat{\sigma}_1}{\hat{\sigma}_2} & -2(1+\rho_0^2)^{\frac{1}{2}}\sigma_{01}\sigma_{02}\hat{\rho}\\ 0 & 0 & (1+\rho_0^2)^{\frac{1}{2}}\sigma_{01}\sigma_{02}\frac{\hat{\rho}^2-1}{\hat{\rho}}\\ 0 & -\sigma_{02}^2\frac{\hat{\sigma}_1}{\hat{\sigma}_2} & (1+\rho_0^2)^{\frac{1}{2}}\sigma_{01}\sigma_{02}\hat{\rho}\end{pmatrix}$$

$$\times\begin{pmatrix}\sum_{i=2}^n\xi_{1,i}\\ \sum_{i=2}^n\xi_{2,i}\\ \sum_{i=2}^n\xi_{3,i}\end{pmatrix}$$

$$+O_p(1).$$

Hence we get

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = (n-1)\left(\begin{pmatrix}\frac{\sigma_{02}}{\sigma_{01}} & \frac{\sigma_{01}}{\sigma_{02}} & -2\rho_0\\ 0 & 0 & \frac{\rho_0^2-1}{\rho_0}\\ 0 & -\frac{\sigma_{01}}{\sigma_{02}} & \rho_0\end{pmatrix}+o_p(1)\right)\begin{pmatrix}\hat{\theta}\hat{\sigma}_1^2-\theta_0\sigma_{01}^2\\ \hat{\theta}\hat{\sigma}_2^2-\theta_0\sigma_{02}^2\\ \hat{\theta}\hat{\rho}\hat{\sigma}_1\hat{\sigma}_2-\theta_0\rho_0\sigma_{01}\sigma_{02}\end{pmatrix}$$

$$-\theta_0\left(\begin{pmatrix}\sigma_{01}\sigma_{02} & \sigma_{01}\sigma_{02} & -2\rho_0(1+\rho_0^2)^{\frac{1}{2}}\sigma_{01}\sigma_{02}\\ 0 & 0 & \frac{(\rho_0^2-1)(1+\rho_0^2)^{\frac{1}{2}}}{\rho_0}\sigma_{01}\sigma_{02}\\ 0 & -\sigma_{01}\sigma_{02} & \rho_0(1+\rho_0^2)^{\frac{1}{2}}\sigma_{01}\sigma_{02}\end{pmatrix}+o_p(1)\right)$$

$$\times\begin{pmatrix}\sum_{i=2}^n\xi_{1,i}\\ \sum_{i=2}^n\xi_{2,i}\\ \sum_{i=2}^n\xi_{3,i}\end{pmatrix}$$

$$+O_p(1). \tag{4.22}$$

Furthermore, we have

$$\begin{pmatrix} \frac{\sigma_{02}}{\sigma_{01}} & \frac{\sigma_{01}}{\sigma_{02}} & -2\rho_0 \\ 0 & 0 & \frac{\rho_0^2-1}{\rho_0} \\ 0 & -\frac{\sigma_{01}}{\sigma_{02}} & \rho_0 \end{pmatrix}^{-1} \begin{pmatrix} -2\rho_0(1+\rho_0^2)^{\frac{1}{2}}\sigma_{01}\sigma_{02} & & \\ 0 & 0 & \frac{(\rho_0^2-1)(1+\rho_0^2)^{\frac{1}{2}}}{\rho_0}\sigma_{01}\sigma_{02} \\ 0 & -\sigma_{01}\sigma_{02} & \rho_0(1+\rho_0^2)^{\frac{1}{2}}\sigma_{01}\sigma_{02} \end{pmatrix}$$
$$= \begin{pmatrix} \sigma_{01}^2 & 0 & 0 \\ 0 & \sigma_{02}^2 & 0 \\ 0 & 0 & \sqrt{\rho_0^2+1}\sigma_{01}\sigma_{02} \end{pmatrix}. \tag{4.23}$$

Hence, from (4.23) and (4.22), we obtain

$$\sqrt{n}\begin{pmatrix} \hat{\theta}\hat{\sigma}_1^2 - \theta_0\sigma_{01}^2 \\ \hat{\theta}\hat{\sigma}_2^2 - \theta_0\sigma_{02}^2 \\ \hat{\theta}\hat{\rho}\hat{\sigma}_1\hat{\sigma}_2 - \theta_0\rho_0\sigma_{01}\sigma_{02} \end{pmatrix} = \theta_0 n^{-\frac{1}{2}} \begin{pmatrix} \sigma_{01}^2 & 0 & 0 \\ 0 & \sigma_{02}^2 & 0 \\ 0 & 0 & \sqrt{\rho_0^2+1}\sigma_{01}\sigma_{02} \end{pmatrix} \begin{pmatrix} \sum_{i=2}^n \xi_{1,i} \\ \sum_{i=2}^n \xi_{2,i} \\ \sum_{i=2}^n \xi_{3,i} \end{pmatrix} + o_p(1).$$

Hence from (4.9) we can get

$$\sqrt{n}\begin{pmatrix} \hat{\theta}\hat{\sigma}_1^2 - \theta_0\sigma_{01}^2 \\ \hat{\theta}\hat{\sigma}_2^2 - \theta_0\sigma_{02}^2 \\ \hat{\theta}\hat{\rho}\hat{\sigma}_1\hat{\sigma}_2 - \theta_0\rho_0\sigma_{01}\sigma_{02} \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \Sigma_{\theta\rho\sigma_1\sigma_2}\right),$$

where

$$\begin{aligned} \Sigma_{\theta\rho\sigma_1\sigma_2} &= \theta_0^2 \begin{pmatrix} \sigma_{01}^2 & 0 & 0 \\ 0 & \sigma_{02}^2 & 0 \\ 0 & 0 & \sqrt{\rho_0^2+1}\sigma_{01}\sigma_{02} \end{pmatrix} \begin{pmatrix} 2 & 2\rho_0^2 & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} \\ 2\rho_0^2 & 2 & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} \\ \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} & \frac{2\rho_0}{(1+\rho_0^2)^{\frac{1}{2}}} & 1 \end{pmatrix} \\ &\times \begin{pmatrix} \sigma_{01}^2 & 0 & 0 \\ 0 & \sigma_{02}^2 & 0 \\ 0 & 0 & \sqrt{\rho_0^2+1}\sigma_{01}\sigma_{02} \end{pmatrix}^{\top}. \end{aligned}$$

Then, we get:

$$\Sigma_{\theta\rho\sigma_1\sigma_2} = \begin{pmatrix} 2(\theta_0\sigma_{01}^2)^2 & 2(\theta_0\rho_0\sigma_{01}\sigma_{02})^2 & 2\theta_0^2\rho_0\sigma_{01}^3\sigma_{02} \\ 2(\theta_0\rho_0\sigma_{01}\sigma_{02})^2 & 2(\theta_0\sigma_{02}^2)^2 & 2\theta_0^2\rho_0\sigma_{02}^3\sigma_{01} \\ 2\theta_0^2\rho_0\sigma_{01}^3\sigma_{02} & 2\theta_0^2\rho_0\sigma_{02}^3\sigma_{01} & \theta_0^2(\rho_0^2+1)^2\sigma_{01}^2\sigma_{02}^2 \end{pmatrix}.$$

Let $f\begin{pmatrix} \hat{\theta}\hat{\sigma}_1^2 \\ \hat{\theta}\hat{\sigma}_2^2 \\ \hat{\theta}\hat{\rho}\hat{\sigma}_1\hat{\sigma}_2 \end{pmatrix} = \begin{pmatrix} \hat{\theta}\hat{\sigma}_1^2 \\ \hat{\theta}\hat{\sigma}_2^2 \\ \frac{\hat{\theta}\hat{\rho}\hat{\sigma}_1\hat{\sigma}_2}{\sqrt{\hat{\theta}\hat{\sigma}_1^2}\sqrt{\hat{\theta}\hat{\sigma}_2^2}} \end{pmatrix}$.

Then, using the multivariate Delta Method we get

$$\sqrt{n}\begin{pmatrix} \hat{\theta}\hat{\sigma}_1^2 - \theta_0\sigma_{01}^2 \\ \hat{\theta}\hat{\sigma}_2^2 - \theta_0\sigma_{02}^2 \\ \hat{\rho} - \rho_0 \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \Sigma_f\right),$$

where $\Sigma_f = H_f \Sigma_{\theta\rho\sigma_1\sigma_2} H_f^{\top}$ and $H_f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{\rho_0}{2\sigma_{01}^2\theta_0} & -\frac{\rho_0}{2\sigma_{02}^2\theta_0} & \frac{1}{\sigma_{01}\sigma_{02}\theta_0} \end{pmatrix}$.

Finally, we get

$$\sqrt{n}\begin{pmatrix}\hat{\sigma}_1^2\hat{\theta}-\sigma_{01}^2\theta_0\\ \hat{\sigma}_2^2\hat{\theta}-\sigma_{02}^2\theta_0\\ \hat{\rho}-\rho_0\end{pmatrix}\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\begin{pmatrix}2(\theta_0\sigma_{01}^2)^2 & 2(\theta_0\rho_0\sigma_{01}\sigma_{02})^2 & \theta_0\rho_0\sigma_{01}^2(1-\rho_0^2)\\ 2(\theta_0\rho_0\sigma_{01}\sigma_{02})^2 & 2(\theta_0\sigma_{02}^2)^2 & \theta_0\rho_0\sigma_{02}^2(1-\rho_0^2)\\ \theta_0\rho_0\sigma_{01}^2(1-\rho_0^2) & \theta_0\rho_0\sigma_{02}^2(1-\rho_0^2) & (\rho_0^2-1)^2\end{pmatrix}\right). \tag{4.24}$$

In the case $\rho_0 = 0$, we can show that (4.17) and (4.24) are still true with the same proof. The only difference is that we multiply the second line of (4.15) and (4.22) by $\hat{\rho}$. We skip the technical details. □

5. Numerical experiments

The main goal of this section is to compare the finite sample behavior of the MLE of the covariance parameters of model (2.1) with the asymptotic distribution given in Section 4. We consider two possible scenarios for our simulation study:

1. The variances parameters are known and we estimate jointly ρ_0 and θ_0.
2. We estimate jointly all the parameters σ_{01}^2, σ_{02}^2, ρ_0 and θ_0.

Under the first scenario we simulate, using the Cholesky decomposition, 1000 realizations from a bivariate zero mean stochastic process with covariance model (2.1) observed on $n = 200, 500$ points uniformly distributed in $[0, 1]$. We simulate fixing $\sigma_{01}^2 = \sigma_{02}^2 = 1$ and increasing values for the colocated correlation parameter and the scale parameter, that is $\rho_0 = 0, 0.2, 0.5$ and $\theta_0 = 3/x$ with $x = 0.2, 0.4, 0.6$. Note that θ_0 is parametrized in terms of practical range that is the correlation is lower than 0.05 when the distance between the points is greater than x. For each simulated realization, we compute $\hat{\rho}_i$ and $\hat{\theta}_i$, $i = 1, \ldots, 1000$, i.e. the MLE of the colocated correlation and scale parameters. Using the asymptotic distribution given in Equation (4.3), Tables 1, 2 compare the empirical quantiles of order $0.05, 0.25, 0.5, 0.75, 0.95$ of $[\sqrt{n}(\hat{\theta}_i-\theta_0)/\sqrt{\theta_0^2(1+\rho_0^2)}]_{i=1}^{1000}$ and $[\sqrt{n}(\hat{\rho}_i-\rho_0)/\sqrt{(\rho_0^2-1)^2}]_{i=1}^{1000}$ respectively, with the theoretical quantiles of the standard Gaussian distribution when $n = 200, 500$. The simulated variances of $\hat{\rho}_i$ and $\hat{\theta}_i$ for $i = 1, \ldots, 1000$ are also reported.

As a general comment, it can be noted that the asymptotic approximation given in Equation (4.3) improves and the variances of the MLE of ρ_0 and θ_0 decrease when increasing n from 200 to 500. When $n = 500$ the asymptotic approximation works very well.

Under the second scenario we set $\sigma_{01}^2 = \sigma_{02}^2 = 0.5$ and the other parameters as in scenario 1. In this case we simulate, using Cholesky decomposition, 1000 realizations from a bivariate zero mean stochastic process with covariance model (2.1) observed on $n = 500, 1000$ points uniformly distributed in $[0, 1]$. For each simulated realization, we obtain $\hat{\sigma}_{1i}^2$, $\hat{\sigma}_{2i}^2$ $\hat{\rho}_i$ and $\hat{\theta}_i$, $i = 1, \ldots, 1000$ the MLE of the two variances, the colocated correlation and scale parameters. Using the asymptotic distribution given in Equation (4.4), Tables 3, 4, 5 compare the empirical quantiles of order $0.05, 0.25, 0.5, 0.75, 0.95$ of $[\sqrt{n}(\hat{\sigma}_{1i}^2\hat{\theta}_i-\sigma_{01}^2\theta_0)/\sqrt{2(\sigma_{01}^2\theta_0)^2}]_{i=1}^{1000}$, $[\sqrt{n}(\hat{\sigma}_{2i}^2\hat{\theta}_i-\sigma_{02}^2\theta_0)/\sqrt{2(\sigma_{02}^2/\theta_0)^2}]_{i=1}^{1000}$ and $[\sqrt{n}(\hat{\rho}_i-\rho_0)/\sqrt{(\rho_0^2-1)^2}]_{i=1}^{1000}$ respectively, for $n = 500, 1000$ with the theoreti-

Table 1

For scenario 1: empirical quantiles, and variances of simulated MLE of ρ_0 for different values of ρ_0 and θ_0, when $n = 200, 500$.

n	θ_0	ρ_0	5%	25%	50%	75%	95%	Var
200	3/0.2	0	-1.6070	-0.6521	-0.0335	0.6812	1.7225	0.0051
500	3/0.2	0	-1.6416	-0.6255	0.0022	0.6675	1.6499	0.0019
200	3/0.2	0.2	-1.6755	-0.6749	-0.0161	0.7149	1.6455	0.0048
500	3/0.2	0.2	-1.6336	-0.6786	-0.0113	0.6712	1.6361	0.0018
200	3/0.2	0.5	-1.7768	-0.6809	-0.0232	0.6583	1.6119	0.0030
500	3/0.2	0.5	-1.6586	-0.6490	0.0146	0.6321	1.6709	0.0011
200	3/0.4	0	-1.6185	-0.6531	-0.0292	0.6852	1.7259	0.0051
500	3/0.4	0	-1.6454	-0.6248	-0.0029	0.6616	1.6457	0.0019
200	3/0.4	0.2	-1.6781	-0.6688	-0.0031	0.7142	1.6576	0.0048
500	3/0.4	0.2	-1.6291	-0.6750	-0.0059	0.6755	1.6629	0.0018
200	3/0.4	0.5	-1.7716	-0.6874	-0.0282	0.6580	1.6226	0.0030
500	3/0.4	0.5	-1.6436	-0.6534	0.0082	0.6270	1.6788	0.0011
200	3/0.6	0	-1.6179	-0.6554	-0.0288	0.6845	1.7200	0.0051
500	3/0.6	0	-1.6487	-0.6466	-0.0019	0.6645	1.6513	0.0019
200	3/0.6	0.2	-1.6908	-0.6694	-0.0088	0.7120	1.6681	0.0048
500	3/0.6	0.2	-1.6286	-0.6767	-0.0111	0.6704	1.6608	0.0018
200	3/0.6	0.5	-1.7810	-0.6950	-0.0354	0.6642	1.6121	0.0030
500	3/0.6	0.5	-1.6407	-0.6537	0.0073	0.6255	1.6686	0.0011
$\mathcal{N}(0,1)$			-1.6448	-0.6744	0	0.6744	1.6448	

Table 2

For scenario 1: empirical quantiles, and variances of simulated MLE of θ_0 for different values of ρ_0 and θ_0, when $n = 500, 1000$.

n	θ_0	ρ_0	5%	25%	50%	75%	95%	Var
200	3/0.2	0	-1.6567	-0.7382	-0.0978	0.6805	1.7761	2.50e-05
500	3/0.2	0	-1.6838	-0.7447	-0.0469	0.6684	1.6369	9.23e-06
200	3/0.2	0.2	-1.6176	-0.7432	-0.0651	0.6583	1.8583	2.61e-05
500	3/0.2	0.2	-1.6962	-0.7370	-0.0260	0.6533	1.6414	9.61e-06
200	3/0.2	0.5	-1.6032	-0.7028	-0.0725	0.6689	1.8607	3.12e-05
500	3/0.2	0.5	-1.6530	-0.7169	-0.0600	0.6758	1.6320	1.13e-05
200	3/0.4	0	-1.5910	-0.7551	-0.0907	0.6715	1.8092	9.68e-05
500	3/0.4	0	-1.6852	-0.7522	-0.0367	0.6661	1.6850	3.64e-05
200	3/0.4	0.2	-1.6073	-0.7242	-0.0731	0.6261	1.7977	1.01e-04
500	3/0.4	0.2	-1.6841	-0.7469	-0.0217	0.6649	1.6060	3.79e-05
200	3/0.4	0.5	-1.5561	-0.6992	-0.0599	0.6578	1.8200	1.02e-04
500	3/0.4	0.5	-1.6410	-0.7191	-0.0577	0.6772	1.6024	4.48e-05
200	3/0.6	0	-1.5563	-0.7307	-0.0847	0.6711	1.8093	2.15e-04
500	3/0.6	0	-1.6737	-0.7421	-0.0352	0.6635	1.6752	8.16e-05
200	3/0.6	0.2	-1.5693	-0.7187	-0.0694	0.6130	1.8244	2.01e-04
500	3/0.6	0.2	-1.6821	-0.7473	-0.0373	0.6579	1.6315	8.49e-05
200	3/0.6	0.5	-1.5666	-0.6765	-0.0638	0.6659	1.8175	2.05e-04
500	3/0.6	0.5	-1.6373	-0.7232	-0.0566	0.6669	1.6208	1.03e-04
$\mathcal{N}(0,1)$			-1.6448	-0.6744	0	0.6744	1.6448	

cal quantiles of the standard Gaussian distribution. The simulated variances of $\hat{\sigma}^2_{1i}\hat{\theta}_i$, $\hat{\sigma}^2_{2i}\hat{\theta}_i$ and $\hat{\rho}_i$ and for $i = 1, \ldots, 1000$ are also reported. As in the previous scenario, the asymptotic approximation given in Equation (4.4) improves and the variances of the MLE of ρ_0 and $\sigma^2_{0i}\theta_0$, $i = 1, 2$ reduce when increasing

TABLE 3
For scenario 2: empirical quantiles, and variances of simulated MLE of $\sigma_{01}^2\theta_0$ for different values of ρ_0 and θ_0, when $n = 500, 1000$.

n	θ_0	ρ_0	5%	25%	50%	75%	95%	Var
500	3/0.2	0	-1.4333	-0.5971	0.0547	0.7163	1.7152	0.2100
1000	3/0.2	0	-1.6085	-0.6291	0.0338	0.7331	1.65266	0.1102
500	3/0.2	0.2	-1.4331	-0.5964	0.0535	0.7160	1.7142	0.2106
1000	3/0.2	0.2	-1.6022	-0.6257	0.0356	0.7348	1.6526	0.1095
500	3/0.2	0.5	-1.4333	-0.5945	0.0520	0.7163	1.7151	0.2098
1000	3/0.2	0.5	-1.6115	-0.6327	0.0336	0.7339	1.6501	0.1110
500	3/0.4	0	-1.4277	-0.5827	0.0427	0.6999	1.6847	0.0519
1000	3/0.4	0	-1.6158	-0.6364	0.0370	0.7277	1.6263	0.0275
500	3/0.4	0.2	-1.4276	-0.5799	0.0427	0.6999	1.6844	0.0518
1000	3/0.4	0.2	-1.6109	-0.6299	0.0459	0.7412	1.6357	0.0276
500	3/0.4	0.5	-1.4276	-0.5827	0.0387	0.6938	1.6842	0.0517
1000	3/0.4	0.5	-1.6090	-0.6275	0.0380	0.7402	1.6346	0.0275
500	3/0.6	0	-1.4229	-0.5847	0.0406	0.6995	1.6997	0.0228
1000	3/0.6	0	-1.6241	-0.6314	0.0393	0.7411	1.6377	0.0123
500	3/0.6	0.2	-1.4235	-0.5833	0.0433	0.7090	1.6999	0.0228
1000	3/0.6	0.2	-1.6234	-0.6318	0.0343	0.7377	1.6365	0.0123
500	3/0.6	0.5	-1.4235	-0.5833	0.0433	0.7090	1.6999	0.0228
1000	3/0.6	0.5	-1.6234	-0.6318	0.0343	0.7377	1.6365	0.0123
$\mathcal{N}(0,1)$			-1.6448	-0.6744	0	0.6744	1.6448	

TABLE 4
For scenario 2: empirical quantiles, and variances of simulated MLE of $\sigma_{02}^2\theta_0$ for different values of ρ_0 and θ_0, when $n = 500, 1000$.

n	θ_0	ρ_0	5%	25%	50%	75%	95%	Var
500	3/0.2	0	-1.5318	-0.6282	0.0544	0.7382	1.8544	0.2336
1000	3/0.2	0	-1.5134	-0.6382	0.0628	0.7003	1.7527	0.1150
500	3/0.2	0.2	-1.5067	-0.6272	0.0411	0.7359	1.7854	0.2364
1000	3/0.2	0.2	-1.4653	-0.6415	0.0728	0.7239	1.7743	0.1155
500	3/0.2	0.5	-1.4734	-0.6078	0.0308	0.7732	1.8493	0.2336
1000	3/0.2	0.5	-1.4260	-0.6438	0.0192	0.7809	1.7520	0.1149
500	3/0.4	0	-1.5173	-0.6479	0.0598	0.7225	1.8452	0.0578
1000	3/0.4	0	-1.5014	-0.6395	0.0604	0.6989	1.7377	0.0287
500	3/0.4	0.2	-1.5164	-0.6275	0.0553	0.7537	1.7436	0.0580
1000	3/0.4	0.2	-1.4724	-0.6442	0.0494	0.7260	1.7822	0.0288
500	3/0.4	0.5	-1.4877	-0.6099	0.0252	0.7725	1.7729	0.0581
1000	3/0.4	0.5	-1.4488	-0.6495	0.0117	0.7565	1.7381	0.0287
500	3/0.6	0	-1.5448	-0.6447	0.0705	0.7226	1.8264	0.0257
1000	3/0.6	0	-1.4940	-0.6560	0.0548	0.7055	1.7365	0.0128
500	3/0.6	0.2	-1.5122	-0.6379	0.0668	0.7553	1.7310	0.0257
1000	3/0.6	0.2	-1.4466	-0.6450	0.0541	0.7316	1.7923	0.0128
500	3/0.6	0.5	-1.4768	-0.6128	0.0325	0.7605	1.7396	0.0258
1000	3/0.6	0.5	-1.4464	-0.6549	-0.0115	0.7551	1.7464	0.0128
$\mathcal{N}(0,1)$			-1.6448	-0.6744	0	0.6744	1.6448	

n from 500 to 1000. When $n = 1000$ the asymptotic approximation is quite satisfactory, with the exception of the case $\rho_0 = 0.5$ where some problems of convergence on the tails of the distributions can be noted, in particular when $\theta_0 = 3/0.4, 3/0.6$.

Table 5
For scenario 2: empirical quantiles, and variances of simulated MLE of ρ_0 for different values of ρ_0 and θ_0, when $n = 500, 1000$.

n	θ_0	ρ_0	5%	25%	50%	75%	95%	Var
500	3/0.2	0	-1.6477	-0.6271	0.0016	0.6795	1.6786	0.0019
1000	3/0.2	0	-1.7235	-0.6167	0.0516	0.6975	1.7051	0.0010
500	3/0.2	0.2	-1.6431	-0.6714	0.0037	0.6518	1.6418	0.0018
1000	3/0.2	0.2	-1.6620	-0.5992	0.0460	0.6906	1.6757	0.0009
500	3/0.2	0.5	-1.6193	-0.6434	0.0123	0.6220	1.6585	0.0011
1000	3/0.2	0.5	-1.6582	-0.6445	0.0563	0.6729	1.5996	0.0005
500	3/0.4	0	-1.6486	-0.6283	-0.0091	0.6684	1.6600	0.0019
1000	3/0.4	0	-1.7296	-0.6209	0.0365	0.6967	1.7151	0.0010
500	3/0.4	0.2	-1.6407	-0.6589	-0.0074	0.6509	1.6631	0.0018
1000	3/0.4	0.2	-1.6840	-0.6067	0.0253	0.6845	1.6823	0.0009
500	3/0.4	0.5	-1.6160	-0.6529	-0.0045	0.5987	1.6543	0.0010
1000	3/0.4	0.5	-1.6669	-0.6434	0.0577	0.6734	1.6171	0.0005
500	3/0.6	0	-1.6504	-0.6280	-0.0092	0.6890	1.6550	0.0019
1000	3/0.6	0	-1.7330	-0.6214	0.0370	0.6931	1.7297	0.0010
500	3/0.6	0.2	-1.6412	-0.6525	0.0050	0.6653	1.6603	0.0018
1000	3/0.6	0.2	-1.7102	-0.6111	0.0201	0.6738	1.6908	0.0009
500	3/0.6	0.5	-1.6536	-0.6510	0.0070	0.6169	1.6561	0.0011
1000	3/0.6	0.5	-1.6776	-0.6496	0.0617	0.6714	1.6175	0.0005
$\mathcal{N}(0,1)$			-1.6448	-0.6744	0	0.6744	1.6448	

6. Concluding remarks

In this paper we consider the fixed domain asymptotic properties of the MLE for a bivariate zero mean Gaussian process with a separable exponential covariance model. We characterize the equivalence of Gaussian measures under this model and we establish the consistency and the asymptotic normality of the MLE of the microergodic parameters. Analogue results under increasing domain asymptotics are obtained by [5]. It is interesting to note that the asymptotic distribution of the MLE of the colocated correlation parameter, between the two processes, does not depend on the asymptotic framework.

Our results can be extended in different directions. The most natural extension, is to consider a general number k of Gaussian processes $Z_1, ..., Z_k$, which covariance structure is of the form $\text{Cov}(Z_i(s_1), Z_j(s_2)) = e^{-\theta|s_1-s_2|}M_{i,j}$, where M is an unknown $k \times k$ covariance matrix. This corresponds to the model studied in this paper for $k = 2$. Rigorously studying the case of a general value of k is out of the scope of the present work, in particular because we think that identifiability and parametrization issues, for the estimation of the covariance matrix M, may arise. We plan to address this extension in future research.

Let us discuss a second potential extension. Let $\mathcal{M}(h, \nu, \theta) = \frac{2^{1-\nu}}{\Gamma(\nu)} (||h||\theta)^{\nu} \mathcal{K}_{\nu}(||h||\theta)$, $h \in \mathbb{R}^d$, $\nu, \theta > 0$, be the Matérn correlation model. A generalization of the bivariate covariance model (2.1) is then the following model:

$$Cov(Z_i(s), Z_j(s+h); \psi) = \sigma_i \sigma_j (\rho + (1-\rho)\mathbf{1}_{i=j})\mathcal{M}(h, \nu, \theta_{ij}), \quad i, j = 1, 2,$$

with $\theta_{12} = \theta_{21}$, $\sigma_1 > 0$, $\sigma_2 > 0$, where in this case $\psi = (\sigma_1^2, \sigma_2^2, \theta_{11}, \theta_{12}, \theta_{22}, \nu, \rho)^\top$. This is a special case of the bivariate Matérn model proposed in [13]. The authors give necessary and sufficient conditions in terms of ψ for the validity of this kind of model. Studying the asymptotic properties of the MLE of ψ would then be interesting. The main challenges in this case are the number of parameters involved and the fact that the covariance matrix cannot be factorized as a Kronecker product. Moreover for $\nu \neq 0.5$ the Markovian property of the process cannot be exploited.

Finally, another interesting extension is to consider the fixed domain asymptotic properties of the tapered maximum likelihood estimator in bivariate covariance models. This method of estimation has been proposed as a possible surrogate for the MLE when working with large data sets, see [12, 15]. Asymptotic properties of this estimator, under fixed domain asymptotics and in the univariate case, can be found in [16], [31] and [10]. Extensions of these results to the bivariate case would be interesting.

Appendix A: Appendix section

Proof of lemma 2

Let $\Sigma(\psi) = A \otimes R$, where the matrices A and R are defined in (2.2). First, using properties of the determinant of the Kroneker product, we have:

$$\log|\Sigma(\psi)| = log(|A|^n\,|R|^2) = n\log\left[\sigma_1^2\sigma_2^2(1-\rho^2)\right] + 2\log|R|\,.$$

From lemma 1 in [34], $|R| = \prod_{i=2}^n \left(1 - e^{-2\theta\Delta_i}\right)$. Then, we have

$$\log|\Sigma(\psi)| = n\log\left[\sigma_1^2\sigma_2^2(1-\rho^2)\right] + 2\sum_{i=2}^n \log\left(1 - e^{-2\theta\Delta_i}\right). \tag{A.1}$$

On the other hand, since $\Sigma(\psi)^{-1} = A^{-1} \otimes R^{-1}$, we obtain

$$\begin{aligned} Z_n^\top\left[\Sigma(\psi)\right]^{-1} Z_n &= \left[Z_{1,n}^\top, Z_{2,n}^\top\right]\left[\begin{pmatrix} \frac{1}{\sigma_1^2(1-\rho^2)}R^{-1} & -\frac{\rho}{\sigma_1\sigma_2(1-\rho^2)}R^{-1} \\ -\frac{\rho}{\sigma_1\sigma_2(1-\rho^2)}R^{-1} & \frac{1}{\sigma_2^2(1-\rho^2)}R^{-1}\end{pmatrix}\right]\begin{bmatrix} Z_{1,n} \\ Z_{2,n}\end{bmatrix} \\ &= \frac{1}{(1-\rho^2)}\left\{\frac{1}{\sigma_1^2}Z_{1,n}^\top R^{-1}Z_{1,n} + \frac{1}{\sigma_2^2}Z_{2,n}^\top R^{-1}Z_{2,n}\right. \\ &\qquad \left. -\frac{\rho}{\sigma_1\sigma_2}\left(Z_{2,n}^\top R^{-1}Z_{1,n} + Z_{1,n}^\top R^{-1}Z_{2,n}\right)\right\}. \end{aligned}$$

Then using Lemma 1 in [34] (Eq. 4.2) we obtain:

$$\begin{aligned} Z_n^\top\left[\Sigma(\psi)\right]^{-1} Z_n &= \frac{1}{(1-\rho^2)}\left\{\sum_{k=1}^2 \frac{1}{\sigma_k^2}\left(z_{k,1}^2 + \sum_{i=2}^n \frac{\left(z_{k,i} - e^{-\theta\Delta_i}z_{k,i-1}\right)^2}{1-e^{-2\theta\Delta_i}}\right)\right. \\ &\quad \left. - \frac{2\rho}{\sigma_1\sigma_2}\left(z_{1,1}z_{2,1} + \sum_{i=2}^n \frac{\left(z_{1,i} - e^{-\theta\Delta_i}z_{1,i-1}\right)\left(z_{2,i} - e^{-\theta\Delta_i}z_{2,i-1}\right)}{1-e^{-2\theta\Delta_i}}\right)\right\}. \end{aligned} \tag{A.2}$$

Combining (3.1), (A.1) and (A.2), we obtain

$$l_n(\psi) = n\left[\log(2\pi) + \log(1-\rho^2)\right] + \sum_{k=1}^2 \log(\sigma_k^2) + \sum_{k=1}^2\sum_{i=2}^n \log\left[\sigma_k^2\left(1 - e^{-2\theta\Delta_i}\right)\right]$$

$$+\frac{1}{1-\rho^2}\left\{\sum_{k=1}^{2}\frac{1}{\sigma_k^2}\left(z_{k,1}^2+\sum_{i=2}^{n}\frac{\left(z_{k,i}-e^{-\theta\Delta_i}z_{k,i-1}\right)^2}{1-e^{-2\theta\Delta_i}}\right)\right.$$
$$\left.-\frac{2\rho}{\sigma_1\sigma_2}\left(z_{1,1}z_{2,1}+\sum_{i=2}^{n}\frac{\left(z_{1,i}-e^{-\theta\Delta_i}z_{1,i-1}\right)\left(z_{2,i}-e^{-\theta\Delta_i}z_{2,i-1}\right)}{1-e^{-2\theta\Delta_i}}\right)\right\}.$$

Proof of lemma 3

By differentiation of $L(\theta)$ with respect to θ we obtain

$$\begin{aligned} G &= \sum_{i=2}^{n}\frac{\Delta_i e^{-\theta\Delta_i}z_{1,i-1}(z_{2,i}-e^{-\theta\Delta_i}z_{2,i-1})+(z_{1,i}-e^{-\theta\Delta_i}z_{1,i-1})\Delta_i e^{-\theta\Delta_i}z_{2,i-1}}{1-e^{-2\theta\Delta_i}} \\ &\quad -\sum_{i=2}^{n}\frac{(z_{1,i}-e^{-\theta\Delta_i}z_{1,i-1})(z_{2,i}-e^{-\theta\Delta_i}z_{2,i-1})2\Delta_i e^{-2\theta\Delta_i}}{(1-e^{-2\theta\Delta_i})^2.} = G_1-G_2, \end{aligned}$$

say. Let us first show that $G_1=O_{up}(1)$. Let for $i=2,...,n$, $A_{\theta,i}=\Delta_i e^{-\theta\Delta_i}/(1-e^{-2\theta\Delta_i})$. By symmetry of $Z_{1,n}$ and $Z_{2,n}$, in order to show $G_1=O_{up}(1)$, it is sufficient to show that

$$\sum_{i=2}^{n}A_{\theta,i}z_{1,i-1}(z_{2,i}-e^{-\theta\Delta_i}z_{2,i-1})=O_{up}(1). \qquad \text{(A.3)}$$

We have

$$\begin{aligned} \sum_{i=2}^{n}A_{\theta,i}z_{1,i-1}(z_{2,i}-e^{-\theta\Delta_i}z_{2,i-1}) &= \sum_{i=2}^{n}A_{\theta,i}z_{1,i-1}(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1}) \\ &\quad +\sum_{i=2}^{n}A_{\theta,i}z_{1,i-1}z_{2,i-1}(e^{-\theta_0\Delta_i}-e^{-\theta\Delta_i}) \\ &= T_1+T_2, \end{aligned}$$

say. Now, one can see from Taylor expansions, and since $\theta\in\Theta$ with Θ compact in $(0,\infty)$, that

$$S:=\sup_{\theta\in\Theta}\sup_{n\in\mathbb{N},i=2,\ldots,n}\left|\frac{A_{\theta,i}(e^{-\theta_0\Delta_i}-e^{-\theta\Delta_i})}{\Delta_i}\right|<\infty.$$

Hence

$$\begin{aligned} |T_2| &\leq \sup_{t\in[0,1]}|Z_1(t)Z_2(t)|S\sum_{i=1}^{n}\Delta_i \\ &= O_{up}(1). \end{aligned}$$

Let us now consider T_1. We have, for any $k<i$

$$\mathbb{E}\left(\left\{z_{1,i-1}(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1})z_{1,k-1}(z_{2,k}-e^{-\theta_0\Delta_k}z_{2,k-1})\right\}\right.$$
$$=\mathbb{E}\left\{\mathbb{E}\left[z_{1,i-1}(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1})z_{1,k-1}(z_{2,k}-e^{-\theta_0\Delta_k}z_{2,k-1})\right.\right. \qquad \text{(A.4)}$$

$$
\begin{aligned}
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad |z_{1,1},...,z_{1,i-1},z_{2,1},...,z_{2,i-1}]\} \\
&= \mathbb{E}\left\{z_{1,i-1}z_{1,k-1}(z_{2,k}-e^{-\theta_0\Delta_k}z_{2,k-1})\mathbb{E}\left[(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1})\right.\right. \qquad (A.5)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \left.\left.|z_{1,1},...,z_{1,i-1},z_{2,1},...,z_{2,i-1}\right]\right\}. \qquad (A.6)
\end{aligned}
$$

Let us show that $\mathbb{E}\left[z_{2,i}|z_{1,1},...,z_{1,i-1},z_{2,1},...,z_{2,i-1}\right] = e^{-\theta_0\Delta_i}z_{2,i-1}$. Let r be the $1\times(i-1)$ vector $(e^{-(s_i-s_1)\theta_0}, e^{-(s_i-s_2)\theta_0},...,e^{-(s_i-s_{i-1})\theta_0})^\top$, $R = [e^{-|s_a-s_b|\theta_0})]_{a,b=1}^{i-1}$ and let $V_k = (z_{k,1},...,z_{k,i-1})^\top$ for $k=1,2$. Then

$$
\begin{aligned}
&\mathbb{E}\left[z_{2,i}|z_{1,1},...,z_{1,i-1},z_{2,1},...,z_{2,i-1}\right] \\
&= \mathbb{E}\left[z_{2,i}|z_{1,1}/\sigma_{01},...,z_{1,i-1}/\sigma_{01},z_{2,1}/\sigma_{02},...,z_{2,i-1}/\sigma_{02}\right] \\
&= [\rho_0\sigma_{02}r^\top, \sigma_{02}r^\top]\left[\begin{pmatrix}1 & \rho_0\\ \rho_0 & 1\end{pmatrix}^{-1}\otimes R^{-1}\right]\begin{bmatrix}(1/\sigma_{01})V_1\\(1/\sigma_{02})V_2\end{bmatrix} \\
&= [\rho_0\sigma_{02}r^\top, \sigma_{02}r^\top]\begin{bmatrix}\frac{1}{1-\rho_0^2}R^{-1} & \frac{-\rho_0}{1-\rho_0^2}R^{-1}\\ \frac{-\rho_0}{1-\rho_0^2}R^{-1} & \frac{1}{1-\rho_0^2}R^{-1}\end{bmatrix}\begin{bmatrix}(1/\sigma_{01})V_1\\(1/\sigma_{02})V_2\end{bmatrix} \\
&= \frac{1}{1-\rho_0^2}\left(\rho_0\sigma_{02}r^\top R^{-1}V_1/\sigma_{01} - \rho_0^2\sigma_{02}r^\top R^{-1}V_2/\sigma_{02} - \rho_0\sigma_{02}r^\top R^{-1}V_1/\sigma_{01}\right. \\
&\qquad\qquad \left. +\sigma_{02}r^\top R^{-1}V_2/\sigma_{02}\right) \\
&= r^\top R^{-1}V_2.
\end{aligned}
$$

Now, it is well known from the Markovian property of Z_2 that $r^\top R^{-1}V_2 = e^{-\theta_0\Delta_i}z_{2,i-1}$. Hence, we have $\mathbb{E}\left[z_{2,i}|z_{1,1},...,z_{1,i-1},z_{2,1},...,z_{2,i-1}\right] = e^{-\theta_0\Delta_i}z_{2,i-1}$, which together with (A.4) gives

$$
\mathbb{E}\left(\{z_{1,i-1}(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1})z_{1,k-1}(z_{2,k}-e^{-\theta_0\Delta_k}z_{2,k-1})\right\} = 0
$$

for $k < i$. Hence

$$
\begin{aligned}
\mathbb{E}\left(T_1^2\right) &= \mathbb{E}\left(\left[\sum_{i=2}^n A_{\theta,i}z_{1,i-1}(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1})\right]^2\right) \\
&= \sum_{i=2}^n \mathbb{E}\left(A_{\theta,i}^2 z_{1,i-1}^2(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1})^2\right). \qquad (A.7)
\end{aligned}
$$

Now, one can see from Taylor expansions that

$$
S' := \sup_{\theta\in\Theta}\sup_{n\in\mathbb{N},i=2,\dots,n}|A_{\theta,i}| < \infty.
$$

Hence

$$
\mathbb{E}\left(T_1^2\right) \leq S'\sum_{i=2}^n\sqrt{\mathbb{E}\left(z_{1,i-1}^4\right)}\sqrt{\mathbb{E}\left((z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1})^4\right)}
$$

$$= \quad S' \sum_{i=2}^{n} \sqrt{3}\sigma_{01}^2 \sqrt{3}\sigma_{02}^2 (1 - e^{-2\theta_0 \Delta_i}).$$

One can see that

$$S'' := \sup_{\theta\in\Theta} \sup_{n\in\mathbb{N}, i=2,\dots,n} \left| \frac{(1 - e^{-2\theta_0\Delta_i})}{\Delta_i} \right| < \infty.$$

Hence,

$$\mathbb{E}\left(T_1^2\right) \leq S'S''3\sigma_{01}^2\sigma_{02}^2 \sum_{i=2}^{n} \Delta_i = O_p(1). \tag{A.8}$$

Hence $T_1 = 0_{up}(1)$ and (A.3) is proved. Hence, we have, with

$$B_{\theta,i} = \frac{2\Delta_i e^{-2\theta\Delta_i}}{(1 - e^{-2\theta\Delta_i})^2},$$

$$G = O_{up}(1) - \sum_{i=2}^{n} (z_{1,i} - e^{-\theta\Delta_i} z_{1,i-1})(z_{2,i} - e^{-\theta\Delta_i} z_{2,i-1}) B_{\theta,i}. \tag{A.9}$$

Furthermore, using $a_\theta b_\theta c_\theta = a_{\theta_0} b_{\theta_0} c_\theta - a_{\theta_0} b_{\theta_0} c_\theta + a_{\theta_0} b_\theta c_\theta - a_{\theta_0} b_\theta c_\theta + a_\theta b_\theta c_\theta$, we have

$$\begin{aligned}
&\sum_{i=2}^{n} (z_{1,i} - e^{-\theta\Delta_i} z_{1,i-1})(z_{2,i} - e^{-\theta\Delta_i} z_{2,i-1}) B_{\theta,i} \\
&= \sum_{i=2}^{n} (z_{1,i} - e^{-\theta_0\Delta_i} z_{1,i-1})(z_{2,i} - e^{-\theta_0\Delta_i} z_{2,i-1}) B_{\theta,i} \\
&\quad + \sum_{i=2}^{n} (z_{1,i} - e^{-\theta_0\Delta_i} z_{1,i-1})(e^{-\theta_0\Delta_i} z_{2,i-1} - e^{-\theta\Delta_i} z_{2,i-1}) B_{\theta,i} \\
&\quad + \sum_{i=2}^{n} (e^{-\theta_0\Delta_i} z_{1,i-1} - e^{-\theta\Delta_i} z_{1,i-1})(z_{2,i} - e^{-\theta\Delta_i} z_{2,i-1}) B_{\theta,i} \\
&= \sum_{i=2}^{n} (z_{1,i} - e^{-\theta_0\Delta_i} z_{1,i-1})(z_{2,i} - e^{-\theta_0\Delta_i} z_{2,i-1}) B_{\theta,i} + R_1 + R_2,
\end{aligned} \tag{A.10}$$

say. We now show that $R_1, R_2 = O_{up}(1)$. For R_1, we have

$$\begin{aligned}
R_1 \quad &= \quad \sum_{i=2}^{n} (z_{1,i} - e^{-\theta_0\Delta_i} z_{1,i-1})(e^{-\theta_0\Delta_i} z_{2,i-1} - e^{-\theta\Delta_i} z_{2,i-1}) B_{\theta,i} \\
&= \quad \sum_{i=2}^{n} z_{2,i-1}(z_{1,i} - e^{-\theta_0\Delta_i} z_{1,i-1})(e^{-\theta_0\Delta_i} - e^{-\theta\Delta_i}) B_{\theta,i}.
\end{aligned}$$

As for T_1 in (A.7),

$$\mathbb{E}(R_1^2) = \sum_{i=2}^{n} \mathbb{E}\left(z_{2,i-1}^2 (z_{1,i} - e^{-\theta_0\Delta_i} z_{1,i-1})^2\right) (e^{-\theta_0\Delta_i} - e^{-\theta\Delta_i})^2 B_{\theta,i}^2.$$

One can show using Taylor expansions that

$$S^{(3)} := \sup_{\theta \in \Theta} \sup_{n \in \mathbb{N}, i=2,\ldots,n} \left| B_{\theta,i}^2 (e^{-\theta_0 \Delta_i} - e^{-\theta \Delta_i})^2 \right| < \infty.$$

Hence

$$\mathbb{E}(R_1^2) \leq S^{(3)} \sum_{i=2}^{n} \mathbb{E}\left(z_{2,i-1}^2 (z_{1,i} - e^{-\theta_0 \Delta_i} z_{1,i-1})^2 \right) = O_u(1)$$

as for (A.8). Hence $R_1 = O_{up}(1)$. For R_2, we have

$$\begin{aligned} R_2 &= \sum_{i=2}^{n} (e^{-\theta_0 \Delta_i} z_{1,i-1} - e^{-\theta \Delta_i} z_{1,i-1})(z_{2,i} - e^{-\theta \Delta_i} z_{2,i-1}) B_{\theta,i} \\ &= \sum_{i=2}^{n} B_{\theta,i} (e^{-\theta_0 \Delta_i} - e^{-\theta \Delta_i}) z_{1,i-1} (z_{2,i} - e^{-\theta \Delta_i} z_{2,i-1}) \\ &= \sum_{i=2}^{n} C_{\theta,i} z_{1,i-1} (z_{2,i} - e^{-\theta \Delta_i} z_{2,i-1}), \end{aligned} \quad \text{(A.11)}$$

say. We can thus show that $R_2 = O_{up}(1)$ as for (A.3). Indeed, the only difference between (A.11) and (A.3) is that $A_{\theta,i}$ is replaced by $C_{\theta,i}$. To show (A.3) we only use that

$$\sup_{\theta \in \Theta} \sup_{n \in \mathbb{N}, i=2,\ldots,n} |A_{\theta,i}| < \infty.$$

We can see from Taylor expansions that

$$\sup_{\theta \in \Theta} \sup_{n \in \mathbb{N}, i=2,\ldots,n} |C_{\theta,i}| < \infty.$$

Hence, as for (A.3), we can show that $R_2 = O_{up}(1)$. Hence, from (A.9) and (A.10), we have,

$$G = O_{up}(1) - \sum_{i=2}^{n} (z_{1,i} - e^{-\theta_0 \Delta_i} z_{1,i-1})(z_{2,i} - e^{-\theta_0 \Delta_i} z_{2,i-1}) B_{\theta,i}. \quad \text{(A.12)}$$

Let, for $i = 2, ..., n$,

$$X_i = (z_{1,i} - e^{-\theta_0 \Delta_i} z_{1,i-1})(z_{2,i} - e^{-\theta_0 \Delta_i} z_{2,i-1}) B_{\theta,i}.$$

For $k < i$ we have

$$\begin{aligned} &\mathbb{E}\left((z_{1,i} - e^{-\theta_0 \Delta_i} z_{1,i-1})(z_{2,k} - e^{-\theta_0 \Delta_k} z_{2,k-1}) \right) \\ &= \rho_0 \sigma_{01} \sigma_{02} \Big(e^{-(s_i - s_k)\theta_0} - e^{-\theta_0 \Delta_i} e^{-(s_{i-1} - s_k)\theta_0} - e^{-\theta_0 \Delta_k} e^{-(s_i - s_{k-1})\theta_0} \\ &\qquad + e^{-\theta_0 (\Delta_i + \Delta_k)} e^{-(s_{i-1} - s_{k-1})\theta_0} \Big) = 0. \end{aligned}$$

Hence, for $k < i$ (and for $k \neq i$ by symmetry), the random variables $(z_{1,i} - e^{-\theta_0 \Delta_i} z_{1,i-1})$ and $(z_{2,k} - e^{-\theta_0 \Delta_k} z_{2,k-1})$ are independent. In addition, the random variables $(z_{j,i} - e^{-\theta_0 \Delta_i} z_{j,i-1})$ and $(z_{j,k} - e^{-\theta_0 \Delta_k} z_{j,k-1})$ are also independent for $j = 1, 2$ and $k \neq i$.

Hence, the $n-1$ Gaussian vectors $\left\{\left[(z_{1,i}-e^{-\theta_0\Delta_i}z_{1,i-1}),(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1})\right]\right\}_{i=2,\dots,n}$ are mutually independent. Thus, the $\{X_i\}_{i=2,\dots,n}$ are independent random variables.

We also have

$$\begin{aligned}\sum_{i=2}^{n}X_i &= \sum_{i=2}^{n}\left(z_{1,i}-e^{-\theta_0\Delta_i}z_{1,i-1}\right)\left(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1}\right)\frac{2\Delta_i e^{-2\theta\Delta_i}}{(1-e^{-2\theta\Delta_i})^2}\\ &= \sum_{i=2}^{n}\frac{\left(z_{1,i}-e^{-\theta_0\Delta_i}z_{1,i-1}\right)\left(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1}\right)}{\sigma_{01}\sigma_{02}\sqrt{1+\rho_0^2}(1-e^{-2\theta_0\Delta_i})}\\ &\quad\times\frac{\sigma_{01}\sigma_{02}\sqrt{1+\rho_0^2}(1-e^{-2\theta_0\Delta_i})2\Delta_i e^{-2\theta\Delta_i}}{(1-e^{-2\theta\Delta_i})^2}.\end{aligned}$$

Let

$$D_{\theta,i}=\frac{\sigma_{01}\sigma_{02}\sqrt{1+\rho_0^2}(1-e^{-2\theta_0\Delta_i})2\Delta_i e^{-2\theta\Delta_i}}{(1-e^{-2\theta\Delta_i})^2}$$

and let $Y_{i,n}$ be as in (4.1). Then, let

$$\begin{aligned}T &= \left|\sum_{i=2}^{n}X_i-\left(\frac{\sigma_{01}\sigma_{02}\sqrt{1+\rho_0^2}\theta_0}{\theta^2}\right)\sum_{i=2}^{n}Y_{i,n}\right|\\ &= \left|\sum_{i=2}^{n}Y_{i,n}\left(D_{\theta,i}-\frac{\sigma_{01}\sigma_{02}\sqrt{1+\rho_0^2}\theta_0}{\theta^2}\right)\right|.\end{aligned}$$

On the other hand

$$\begin{aligned}\mathbb{E}(Y_{i,n}) &= \frac{\mathbb{E}(z_{1,i}z_{2,i})-e^{-\theta_0\Delta_i}\mathbb{E}(z_{1,i}z_{2,i-1})-e^{-\theta_0\Delta_i}\mathbb{E}(z_{1,i-1}z_{2,i})+e^{-2\theta_0\Delta_i}\mathbb{E}(z_{1,i-1}z_{2,i-1})}{\sigma_{01}\sigma_{02}(1+\rho_0^2)^{1/2}(1-e^{-2\theta_0\Delta_i})}\\ &= \frac{\sigma_{01}\sigma_{02}\rho_0\left[1-e^{-2\theta_0\Delta_i}-e^{-2\theta_0\Delta_i}+e^{-2\theta_0\Delta_i}\right]}{\sigma_{01}\sigma_{02}(1+\rho_0^2)^{1/2}(1-e^{-2\theta_0\Delta_i})}\\ &= \frac{\sigma_{01}\sigma_{02}\rho_0(1-e^{-2\theta_0\Delta_i})}{\sigma_{01}\sigma_{02}(1+\rho_0^2)^{1/2}(1-e^{-2\theta_0\Delta_i})}\\ &= \frac{\rho_0}{(1+\rho_0^2)^{1/2}},\end{aligned}$$

Furthermore,

$$\begin{aligned}\mathbb{E}(Y_{i,n}^2) &= \frac{\mathbb{E}\left[(z_{1,i}-e^{-\theta_0\Delta_i}z_{1,i-1})^2(z_{2,i}-e^{-\theta_0\Delta_i}z_{2,i-1})^2\right]}{\left[\sigma_{01}\sigma_{02}(1+\rho_0^2)^{1/2}(1-e^{-2\theta_0\Delta_i})\right]^2}\\ &= \frac{\sigma_{01}^2\sigma_{02}^2\left[1+2\rho_0^2+e^{-2\theta_0\Delta_i}(e^{-2\theta_0\Delta_i}-2)+2\rho_0^2e^{-2\theta_0\Delta_i}(e^{-2\theta_0\Delta_i}-2)\right]}{\left[\sigma_{01}\sigma_{02}(1+\rho_0^2)^{1/2}(1-e^{-2\theta_0\Delta_i})\right]^2}\\ &= \frac{1+2\rho_0^2}{1+\rho_0^2},\end{aligned}$$

as is obtained by using Isserlis' theorem for correlated Gaussian random variables. Furtermore

$$Var(Y_{i,n}) = \mathbb{E}(Y_{i,n}^2)-[\mathbb{E}(Y_{i,n})]^2$$

$$
\begin{aligned}
&= \frac{1+2\rho_0^2}{1+\rho_0^2} - \left(\frac{\rho_0}{(1+\rho_0^2)^{1/2}} \right)^2 \\
&= 1.
\end{aligned}
$$

Hence $\mathbb{E}(|Y_{i,n}| \leq \sqrt{2})$ and so

$$
\begin{aligned}
\mathbb{E}(T) &\leq \sqrt{2} \sum_{i=2}^{n} \left| D_{\theta,i} - \frac{\sigma_{01}\sigma_{02}\sqrt{1+\rho_0^2}\theta_0}{\theta^2} \right| \\
&= \sqrt{2}\sigma_{01}\sigma_{02}\sqrt{1+\rho_0^2} \sum_{i=2}^{n} \left| \frac{(1-e^{-2\theta_0\Delta_i})2\Delta_i e^{-2\theta\Delta_i}}{(1-e^{-2\theta\Delta_i})^2} - \frac{\theta_0}{\theta^2} \right| .
\end{aligned}
$$

One can show, from a Taylor expansion and since $\theta \in \Theta$ with Θ compact in $(0, +\infty)$, that

$$
\sup_{n\in\mathbb{N}, i=2,\dots,n} \sup_{\theta\in\Theta} \frac{1}{\Delta_i} \left| \frac{(1-e^{-2\theta_0\Delta_i})2\Delta_i e^{-2\theta\Delta_i}}{(1-e^{-2\theta\Delta_i})^2} - \frac{\theta_0}{\theta^2} \right| < \infty.
$$

Hence $\mathbb{E}(T) = O_u(1)$ and $T = 0_{up}(1)$. Hence, finally

$$
G = -\frac{\sigma_{01}\sigma_{02}\sqrt{1+\rho_0^2}\theta_0}{\theta^2} \sum_{i=2}^{n} Y_{i,n} + O_{up}(1).
$$

References

[1] Abt, M. and Welch, W. (1998). Fisher information and maximum-likelihood estimation of covariance parameters in Gaussian stochastic processes. *The Canadian Journal of Statistics*, 26:127–137. MR1624393

[2] Anderes, E. (2010). On the consistent separation of scale and variance for Gaussian random fields. *The Annals of Statistics*, 38:870–893. MR2604700

[3] Antognini, A. B. and Zagoraiou, M. (2010). Exact optimal designs for computer experiments via Kriging metamodelling. *Journal of Statistical Planning and Inference*, 140:2607–2617. MR2644082

[4] Bachoc, F. (2014). Asymptotic analysis of the role of spatial sampling for covariance parameter estimation of Gaussian processes. *Journal of Multivariate Analysis*, 125:1–35. MR3163828

[5] Bevilacqua, M., Vallejos, R., and Velandia, D. (2015). Assessing the significance of the correlation between the components of a bivariate Gaussian random field. *Environmetrics*, 26:545–556. MR3431929

[6] Chen, H., Simpson, D., and Ying, Z. (2000). Infill asymptotics for a stochastic process model with measurement error. *Statistica Sinica*, 10:141–156. MR1742105

[7] Cressie, N. (1993). *Statistics for spatial data*. J. Wiley. MR1239641

[8] Cressie, N. and Lahiri, S. (1993). The asymptotic distribution of REML estimators. *Journal of Multivariate Analysis*, 45:217–233. MR1221918

[9] Cressie, N. and Lahiri, S. (1996). Asymptotics for REML estimation of spatial covariance parameters. *Journal of Statistical Planning and Inference*, 50:327–341. MR1394135
[10] Du, J., Zhang, H., and Mandrekar, V. S. (2009). Fixed-domain asymptotic properties of tapered maximum likelihood estimators. *The Annals of Statistics*, 37:3330–3361. MR2549562
[11] Furrer, R., Bachoc, F., and Du, J. (2016). Asymptotic properties of multivariate tapering for estimation and prediction. *Journal of Multivariate Analysis*, 149:177–191. MR3507322
[12] Furrer, R., Genton, M. G., and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 15(3):502–523. MR2291261
[13] Gneiting, T., Kleiber, W., and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105:1167–1177. MR2752612
[14] Ibragimov, I. A. and Rozanov, Y. A. (1978). *Gaussian Random Processes*. Springer-Verlag New York. MR0543837
[15] Kaufman, C. G., Schervish, M. J., and Nychka, D. W. (2008a). Covariance Tapering for Likelihood-Based Estimation in Large Spatial Data Sets. *Journal of the American Statistical Association*, 103(484):1545–1555.
[16] Kaufman, C. G., Schervish, M. J., and Nychka, D. W. (2008b). Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103:1545–1555. MR2504203
[17] Kaufman, C. G. and Shaby, B. A. (2013). The role of the range parameter for estimation and prediction in geostatistics. *Biometrika*, 100:473–484. MR3068447
[18] Lehrke, S. G. and Ghorai, J. K. (2010). Large sample properties of ML estimator of the parameters of multivariate O–U random fields. *Communications in Statistics – Theory and Methods*, 39(4):738–752. MR2745317
[19] LeMone, M. A., Chen, F., Alfieri, J. G., Cuenca, R., Hagimoto, Y., Blanken, P. D., Niyogi, D., Kang, S., Davis, K., and Grossman, R. (2007). NCAR/CU surface vegetation observation network during the international H2O project 2002 field campaign. *Bull. Amer. Meteor. Soc.*, 88: 65–81.
[20] Loh, W. L. (2005). Fixed-domain asymptotics for a subclass of Matérn-type Gaussian random fields. *The Annals of Statistics*, 33:2344–2394. MR2211089
[21] Loh, W. L. and Lam, T. K. (2000). Estimating structured correlation matrices in smooth Gaussian random fields models. *The Annals of Statistics*, 28:880–904. MR1792792
[22] Mardia, K. V. and Marshall, R. J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, 71:135–146. MR0738334
[23] Pascual, F. G. and Zhang, H. (2006). Estimation of linear correlation coefficient of two correlated spatial processes. *Sankhyā: The Indian Journal of Statistics*, 68:307–325. MR2303086

[24] Shaby, B. A. and Ruppert, D. (2012). Tapered covariance: Bayesian estimation and asymptotics. *Journal of Computational and Graphical Statistics*, 21(2):433–452. MR2945475

[25] Stassberg, D., LeMone, M. A., Warner, T., and Alfieri, J. G. (2008). Comparison of observed 10 m wind speeds to those based on Monin–Obukhov similarity theory using aircraft and surface data from the international H2O project. *Mon. Wea. Rev.*, 136:964–972.

[26] Stein, M. (1988). Asymptotically efficient prediction of a random field with a misspecified covariance function. *The Annals of Statistics*, 16:55–63. MR0924856

[27] Stein, M. (1990a). Bounds on the efficiency of linear predictions using an incorrect covariance function. *The Annals of Statistics*, 18:1116–1138. MR1062701

[28] Stein, M. (1990b). Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure. *The Annals of Statistics*, 18:850–872. MR1056340

[29] Stein, M. L. (1999). *Interpolation of Spatial Data*. Springer Series in Statistics. Springer-Verlag New York. MR1697409

[30] van der Vaart, A. (1996). Maximum likelihood estimation under a spatial sampling scheme. *The Annals of Statistics*, 5:2049–2057. MR1421160

[31] Wang, D. and Loh, W. L. (2011). On fixed-domain asymptotics and covariance tapering in Gaussian random field models. *Electronic Journal of Statistics*, 5:238–269. MR2792553

[32] Weckwerth, T. M., Parsons, D. B., Koch, S. E., Moore, J. A., LeMone, M. A., Demoz, B. B., Flamant, C., Geerts, B., Wang, J., and Feltz, W. F. (2004). An overview of the international H2O project (IHOP 2002) and some preliminary highlights. *Bull. Amer. Meteor. Soc.*, 85:253–277.

[33] Ying, Z. (1991). Asymptotic properties of a maximum likelihood estimator with data from a Gaussian process. *Journal of Multivariate Analysis*, 36:280–296. MR1096671

[34] Ying, Z. (1993). Maximum likelihood estimation of parameters under a spatial sampling scheme. *The Annals of Statistics*, 21:1567–1590. MR1241279

[35] Zhang, H. (2004). Inconsistent estimation and asymptotically equivalent interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99:250–261. MR2054303

[36] Zhang, H. and Cai, W. (2015). When doesn't cokriging outperform Kriging? *Statistical Science*, 30:176–180. MR3353100

[37] Zhang, H. and Zimmerman, D. L. (2005). Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika*, 92:921–936. MR2234195