PARAMETER ESTIMATION FOR GAUSSIAN RANDOM FIELDS AND MULTIVARIATE GAUSSIAN RANDOM PROCESSES UNDER FIXED-DOMAIN ASYMPTOTICS

By

Haoxiang Feng

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Statistics—Doctor of Philosophy

2024

ABSTRACT

This dissertation explores parameter estimation for Gaussian random fields and multivariate Gaussian random processes under fixed-domain asymptotics, a crucial framework for modeling spatial and temporal data. Unlike increasing-domain asymptotics, fixed-domain asymptotics involve a growing number of observations within a fixed, bounded region, leading to denser data. This scenario is common in applications such as image processing, where the spatial domain is constrained by the finite size of the sensor array.

First, we study the parameter estimation for a Gaussian field with a multiplicative covariance function, which is particularly relevant in computer experiments. We propose an increment-based estimator for estimating variance and scale parameters, and subsequent analysis shows that the estimator is both strongly consistent and asymptotically normal.

Next, we extend the analysis to the bivariate Ornstein-Uhlenbeck process, constructing an explicit estimator that is strongly consistent and asymptotically normal. This estimator, requiring no prior parameter information, is shown to have the same asymptotic covariance matrix as that of the maximum likelihood estimator (MLE).

Finally, we investigate asymptotic properties of MLE for the isotropic powered exponential field. Unlike the Matérn model, the spectral density of the powered exponential model poses analytical challenges. We also establish conditions for the equivalence of Gaussian measures, providing a contrast to the orthogonality conditions found in earlier studies.

This dissertation is dedicated to my parents.

ACKNOWLEDGEMENTS

The research in this dissertation was partially supported by the NSF grant DMS-2153846.

I would like to express my deepest gratitude to my advisor, Dr. Yimin Xiao, for his guidance, support, and patience throughout my PhD research. My sincere thanks also go to Dr. Haolei Weng, Dr. Chih-Li Sung, and Dr. Antoine Ayache for serving on my guidance committee and for their invaluable feedback and encouragement. I would also like to thank my friends for their support.

I am grateful to the Department of Statistics and Probability and Michigan State University for their support during my time in East Lansing.

Lastly, I extend my heartfelt thanks to my beloved Yaqian for her enduring love and understanding.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Multivariate Gaussian random fields defined on $\mathbb{R}^d$ are widely used to model spatial and temporal data. When fitting a random field to the data, it is common to assume that its covariance function belongs to a family of functions parameterized by a set of parameters which are then estimated based on the data. This turns the modeling problem into a parameter estimation one (see, e.g., Chilès and Delfiner 2012, for an introduction to spatial statistical techniques).

Three prevalent asymptotic frameworks exist in spatial statistics: increasing-domain asymptotics, fixed-domain asymptotics (see, e.g., M. L. Stein 1999a, chap. 3.3) and mixed-domain asymptotics (see, e.g., Lahiri 2003). In increasing-domain asymptotics, the distance between neighboring observation points remains above a positive threshold, causing the sampling region to expand as the number of observations increases. In fixed-domain asymptotics which is also called infill asymptotics in Cressie (1993), the number of observation points rises within a bounded sampling domain, resulting in increasingly dense observation points. In mixed-domain asymptotics, the sampling region expands, and at the same time, increasingly dense observation points fill in any given subregion of the sampling region.

Fixed-domain asymptotics can occur in the process of spatial data. For example, we can model the analog signal a camera receives as

$$X_a(s) = f(s) + X(s), \quad s \in T,$$

where $T \subset \mathbb{R}^2$ is a fixed and bounded set because the charge-coupled device (CCD) of a camera has a finite surface. We use $f(s)$ to represent the original signal, and $X(s)$ is the intrinsic physical noise which is modeled by a random field. Subsequently, an analog-to-digital converter converts $X_a(s)$ to a digital signal, which can be modeled as

$$X_d(\mathbf{n}) = \langle X_a(s), \phi_{\mathbf{n}}(s) \rangle, \quad \mathbf{0} \leq \mathbf{n} \leq \mathbf{N},$$

where $\{\phi_{\mathbf{n}}(s)\}$ are the sensor responses. Here the two-dimensional vector $\mathbf{n}$ represents the pixel and $\mathbf{N}$ denotes the image size. If we plan to study the performance of a denoising algorithm based on measurements $\{X_d(\mathbf{n})\}$ as $\mathbf{N}$ increases, it seems reasonable to adopt the fixed-domain asymptotic framework (see, e.g., Mallat 2008, chap. 11 for denoising).

In the theory of fixed-domain asymptotics for multivariate Gaussian fields, there is a key concept called the equivalence of Gaussian measures. Specifically, for two probability measures P_0 and P_1 on a measurable space $(\Omega, \mathcal{F})$, say that P_0 is absolutely continuous with respect to P_1 if for all $A \in \mathcal{F}$, $P_1(A) = 0$ implies $P_0(A) = 0$. Define P_0 and P_1 to be equivalent, written $P_0 \equiv P_1$, if they are mutually absolutely continuous. Define P_0 and P_1 to be orthogonal, written $P_0 \perp P_1$, if there exists $A \in \mathcal{F}$ such that $P_1(A) = 0$ implies $P_0(A) = 1$. Feldman (1958) proved that two Gaussian measures are either equivalent or orthogonal. The concept plays an important role in both prediction and estimation of multivariate Gaussian fields under fixed-domain asymptotics (see, e.g., M. L. Stein 1988; M. Stein 1990; M. L. Stein 1993; M. L. Stein 1999b; M. L. Stein 2004).

Since this work focuses mainly on parameter estimation problems of multivariate Gaussian processes and univariate Gaussian fields under fixed-domain asymptotics, we first review recent results and relevant methods in the literature. On the one hand, Mason and Xiao (2002) studied sample path properties of operator fractional Brownian motions. Amblard and Coeurjolly (2011) studied parameter estimation of multivariate fractional Brownian motion with an increment-based method and proved the strong consistency and asymptotic normality under increasing-domain asymptotics. Didier and Pipiras (2011) provided the integral representations of operator fractional Brownian motions in the spectral and time domains, respectively. Subsequently, Abry and Didier (2018) constructed estimators for operator fractional Brownian motion with wavelets, and showed the estimators are consistent and asymptotically normal under increasing-domain asymptotics. It is of interest to note that the corresponding high-pass filter of the wavelet used by Abry and Didier (2018) is a type of increment. On the other hand, Zhou and Xiao (2018) estimated the fractal indices of bivariate stationary Gaussian processes with the increment-based method under fixed-domain asymptotics and proved the consistency and asymptotic normality. Gneiting, Kleiber, and Schlather (2010) introduced a flexible parametric family of matrix-valued covariance functions for multivariate stationary Gaussian random fields, where each constituent component is a Matérn field. Zhang and Cai (2015) gave a sufficient condition for two probability measures corresponding to matrix-valued Matérn covariance functions to be equivalent and displayed an explicit example where cokriging is identical to kriging (best linear unbiased prediction). Velandia et al. (2017) showed that the MLE for a bivariate Ornstein-Uhlenbeck process is strongly consistent and asymptotically normal under fixed-domain asymptotics given some prior information of the parameter. Bachoc et al. (2022) provided conditions for equivalence of measures associated with multivariate Gaussian random fields and studied misspecified cokriging prediction for multivariate Matérn and generalized Wendland fields under fixed-domain asymptotics.

Compared with MLE, the increment-based estimator can achieve linear computational complexity, and it is easily computed with no maximization required in many cases. Unsurprisingly, there has been a long history of using increment-based methods to study properties of random fields. Actually, it can be traced back to the quadratic variation theorem in Lévy (1940). Since then, there has been a growing number of papers focused on parameter estimation for univariate random fields with increments under fixed-domain asymptotics. Kent and Wood (1997) estimated the fractal index of Gaussian processes using increments. Chan and Wood (2000) consistently estimated the fractal dimension with asymptotic normality using increments of the random field observed on a regular grid in $\mathbb{R}^2$. Anderes (2010) proved both variance and scale parameters from the Matérn model can be consistently estimated when $d > 4$ based on the increment-based method, but asymptotic distributions were not given. The increment-based method was then generalized by Loh, Sun, and Wen (2021) and Loh (2015) to estimate the smoothness parameter of the Matérn field irregularly sampled on $\mathbb{R}^d$. However, both papers did not study the asymptotic distribution of the estimators.

In Chapter 2, we propose a method based on findings from Chan and Wood (2000) and Loh (2015) for estimating the variance and scale parameters in the model considered in Ying (1993) and study both its strong consistency and asymptotic normality with irregularly spaced data under fixed-domain asymptotics. Models with a multiplicative covariance function like the one considered in Ying (1993) are popular in computer experiments; (see, e.g., Sacks, Welch, et al. 1989; Sacks, Schiller, and Welch 1989; Paulo 2005; Bayarr et al. 2009; Peng and Wu 2014).

In contrast to the increasing-domain asymptotics where the MLE of all (identifiable) parameters is consistent and asymptotically normal under some mild regularity conditions (Mardia and Marshall 1984), there is no general result for the asymptotic properties of MLE under fixed-domain asymptotics. However, there is quite a bit of literature on fixed-domain asymptotics of MLE when assuming that the covariance belongs to a parametric family. For the univariate Ornstein-Uhlenbeck process, Ying (1991) showed the product of its variance and scale parameters can be consistently estimated by MLE and the estimator is asymptotically normal. Regarding the isotropic Matérn model, with known smoothness parameter ν and free variance σ^2 and scale parameter α, Zhang (2004) showed the MLE of $\sigma^2\alpha^{2\nu}$ is strongly consistent for dimension $d \leq 3$, and Du, Zhang, and Mandrekar (2009) showed the asymptotic normality of the estimator when $d = 1$. Wang and Loh (2011) extended the asymptotic result to dimension $d \leq 3$. Bevilacqua et al. (2019) studied the MLE for the generalized Wendland model and derived similar results to those for the Matérn model.

In Chapter 3, we construct an estimator for the bivariate Ornstein-Uhlenbeck process considered in Velandia et al. (2017) and study its strong consistency and asymptotic normality. The construction is inspired by the results in Ying (1991) and Zhang (2004). Compared to the MLE from Velandia et al. (2017), our estimator has an explicit form and does not require any prior information of the parameter. Meanwhile, it turns out that the estimator has the same asymptotic covariance matrix as that of the MLE, echoing the fact that cokriging is identical to kriging for this model as shown in Zhang and Cai (2015).

In Chapter 4, we analyze the asymptotic properties of MLE for the isotropic powered exponential field using tools from M. L. Stein (2004), Zhang (2004), and Wang and Loh (2011). Compared with the Matérn model, the spectral density of the powered exponential model cannot be analytically expressed except in some special cases, which brings new challenges to the analysis. Furthermore, we also establish the parameter condition for the equivalence of Gaussian measures, which contrasts with the orthogonality condition in Theorem 5 from Anderes (2010).

CHAPTER 2

A CLASS OF ORNSTEIN-UHLENBECK FIELDS

2.1 Introduction

Stationary Gaussian fields with multiplicative covariance functions have been successfully applied to the modeling of computer experiments. Sacks, Welch, et al. (1989) and Sacks, Schiller, and Welch (1989) proposed the use of a Gaussian field with the multiplicative powered exponential covariance function in their modeling of computer experiments. Observing the undesirable properties of the corresponding Gaussian field for modeling computer experiments, M. L. Stein (1989) proposed using a stationary Gaussian field model, $X(\mathbf{u}), \mathbf{u} \in [0,1]^d$, with mean 0 and the multiplicative Matérn covariance function,

$$\begin{aligned}&\operatorname{cov}\left(X\left(\mathbf{u}\right), X\left(\mathbf{v}\right)\right)\\ =&\sigma^2 \prod_{i=1}^{d} \frac{2^{1-\nu}}{\Gamma(\nu)} (\theta_i |\mathbf{u}[i] - \mathbf{v}[i]|)^{\nu} K_{\nu}(\theta_i |\mathbf{u}[i] - \mathbf{v}[i]|), \quad \forall\, \mathbf{u}, \mathbf{v} \in [0,1]^d,\end{aligned} \tag{2.1}$$

where $\theta_1, \cdots, \theta_d$ and σ^2 are strictly positive parameters, and K_ν is the modified Bessel function of the second kind with order $\nu > 0$. By Equation (32) of M. L. Stein (1999a), the spectral density corresponding to Eq. (2.1) is

$$f(\boldsymbol{\omega}) = \sigma^2 \prod_{i=1}^{d} \frac{\Gamma(\nu + \frac{1}{2})\theta_i^{2\nu}}{\Gamma(\nu)\pi^{1/2}} \frac{1}{(\theta_i^2 + (\boldsymbol{\omega}[i])^2)^{\nu+(1/2)}}. \tag{2.2}$$

The parameter ν controls the smoothness of the random field X. Specifically, X is m times mean square differentiable if and only if $\nu > m$.

If $\nu = n + \frac{1}{2}$, Eq. (2.1) reduces to the product of an exponential function and a polynomial, in that

$$\begin{aligned}&\operatorname{cov}\left(X\left(\mathbf{u}\right), X\left(\mathbf{v}\right)\right)\\ =&\sigma^2 \prod_{i=1}^{d} \exp\left(-\theta_i |\mathbf{u}[i] - \mathbf{v}[i]|\right) \sum_{k=0}^{n} \frac{(n+k)!}{(2n)!} \begin{pmatrix} n \\ k \end{pmatrix} (2\theta_i |\mathbf{u}[i] - \mathbf{v}[i]|)^{n-k},\end{aligned} \tag{2.3}$$

for $n = 0, 1, \cdots$ by Eq.(8.468) of Gradshteyn et al. (1981). Therefore, when $\nu = 0.5$, Eq. (2.1) reduces to

$$\mathrm{cov}\left(X\left(\mathbf{u}\right), X\left(\mathbf{v}\right)\right) = \sigma^2 \prod_{i=1}^{d} \exp\left(-\theta_i |\mathbf{u}[i] - \mathbf{v}[i]|\right). \tag{2.4}$$

In this regard, Ying (1993) showed both the strong consistency and the asymptotic normality of the maximum likelihood estimator (MLE) of $(\sigma^2, \theta_1, \cdots, \theta_d)$ when $d \geq 2$. Subsequently, van der Vaart (1996) proved the MLE is also asymptotically efficient for dimension $d = 2$. When $\nu = 1.5$, Eq. (2.1) reduces to

$$\mathrm{cov}\left(X\left(\mathbf{u}\right), X\left(\mathbf{v}\right)\right) = \sigma^2 \prod_{i=1}^{d} \left(1 + \theta_i |\mathbf{u}[i] - \mathbf{v}[i]|\right) \exp\left(-\theta_i |\mathbf{u}[i] - \mathbf{v}[i]|\right), \tag{2.5}$$

and Loh (2005) constructed a consistent estimator of $(\sigma^2, \theta_1, \cdots, \theta_d)$ based on the maximum likelihood method when $d \geq 3$.

In this chapter, we use the increment-based method from Chan and Wood (2000) and Loh (2015) to estimate the variance and scale parameters in Eq. (2.4) based on irregularly sampled data and study both its consistency and asymptotic normality under infill asymptotics. For simplicity, we first study the dimension $d = 2$ in detail; therefore, there are only three parameters $(\lambda, \mu, \sigma^2)'$ being defined in our model below. Then we generalize the method to an arbitrary d in Section 2.5. Our main motivation to study this particular case is that, on the one hand, properties of the MLE in this case have been well studied so that a comprehensive comparison can be made between the MLE and the increment-based method in terms of the consistency and the asymptotic normality. On the other hand, one obvious advantage of the increment-based method is that the construction of the estimator does not involve the inverse of the covariance matrix compared to the MLE, which greatly reduces the computational complexity of the estimation. Actually, in our case, the computational complexity is $O(n)$ where n is the sample size. However, to our best knowledge, the applicability of the increment-based method on a Gaussian field with a multiplicative covariance function under an irregular sampling scheme had not yet been studied. The proofs in this case lay the foundation for the asymptotic analysis of more general cases.

The rest of this chapter is organized as follows. In Section 2.2, we study the Ornstein-Uhlenbeck (OU) process with a covariance function $\Gamma_\mu(t) = \sigma^2 e^{-\mu|t|}$ by the increment-based method, which is the first step towards the asymptotic analysis of the OU field with the covariance function as Eq. (2.4). In Section 2.3, we establish the strong consistency and the asymptotic normality of the estimator of $(\lambda, \mu, \sigma^2)'$. In Section 2.4, we present a simulation study on the efficiency of the estimator in finite-sample cases. In Section 2.5, we generalize the increment-based method to any dimension d and derive the corresponding strong consistency and asymptotic normality results.

We end the introduction with some notation. For any real-valued sequences $a_n, b_n, a_n \sim b_n$ means $\lim_{n\to\infty} b_n/a_n = 1$. We write $\mathbf{j} = (\mathbf{j}[1], \cdots, \mathbf{j}[d])$, with brackets used to denote components of $\mathbf{j}$, and $\mathbf{0} \le \mathbf{j} \le \mathbf{n}$ is equivalent to $\mathbf{0} \le \mathbf{j}[\ell] \le \mathbf{n}[\ell]$ for $\ell = 1, \cdots, d$. If $\mathbf{j} \in \mathbf{Z}^d$ is a multi-index, we write $|\mathbf{j}| = \sum_{\ell=1}^{d} |\mathbf{j}[\ell]|$.

2.2 Sampling over $\mathbb{R}$

In this section, we explore the performance of the quadratic variation built on the increment introduced in Section 2 in Loh (2015). For the Ornstein-Uhlenbeck process X with mean 0 and the covariance function $\Gamma_\mu(t) = \sigma^2 e^{-\mu|t|}$, we aim to estimate $\sigma^2\mu$ with observations $X(t_{n,1}), \cdots, X(t_{n,n})$, where $0 = t_{n,1} < t_{n,2} < \cdots < t_{n,n-1} < t_{n,n} = 1$. For brevity, we write $t_{n,i} = t_i$ and $X(t_i) = X_i, i = 1, \cdots, n$.

For $\ell \in \mathbb{N}^+$, and $i = 1, \cdots, n-\ell$, we plan to design the increment such that

$$\sum_{k=0}^{\ell} a_{i,k} t_{(i+k)}^{q} = \begin{cases} 0, & \forall\, q = 0, \cdots, (\ell-1), \\ \ell! & \text{if } q = \ell, \end{cases} \tag{2.6}$$

where we use the convention $0^0 = 1$. The system of equations can be expressed in the following matrix form

$$\begin{pmatrix} t^0_{(i+0)} & t^0_{(i+1)} & \cdots & t^0_{(i+\ell)} \\ t^1_{(i+0)} & t^1_{(i+1)} & \cdots & t^1_{(i+\ell)} \\ \vdots & \vdots & \ddots & \vdots \\ t^\ell_{(i+0)} & t^\ell_{(i+1)} & \cdots & t^\ell_{(i+\ell)} \end{pmatrix} \begin{pmatrix} a_{i,0} \\ a_{i,1} \\ \vdots \\ a_{i,\ell} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \ell! \end{pmatrix}. \tag{2.7}$$

The matrix on the left hand side is a Vandermonde matrix and the last column of the inverse of the Vandermonde matrix is

$$\left(\frac{1}{\prod_{0\le s\le \ell, s\neq 0}(t_{i+0}-t_{i+s})}, \frac{1}{\prod_{0\le s\le \ell, s\neq 1}(t_{i+1}-t_{i+s})}, \cdots, \frac{1}{\prod_{0\le s\le \ell, s\neq \ell}(t_{i+\ell}-t_{i+s})} \right)'. \tag{2.8}$$

Therefore,

$$a_{i,k} = \frac{\ell!}{\prod_{0\le s\le \ell, s\neq k}(t_{i+k}-t_{i+s})}, \quad \forall\, k = 0, \cdots, \ell. \tag{2.9}$$

Over here, we only consider the case $\ell = 1$; therefore, the coefficients of the increment can be simplified as

$$a_{i,0} = -1/\Delta_i, \quad a_{i,1} = 1/\Delta_i, \tag{2.10}$$

where $\Delta_i = t_{i+1} - t_i$; (see Kent and Wood 1997, for a full exposition on the effect of ℓ). Then define the quadratic variation as

$$V_n = \sum_{i=1}^{n-1} (\nabla X_i)^2, \tag{2.11}$$

where

$$\nabla X_i = a_{i,0} X_i + a_{i,1} X_{i+1}. \tag{2.12}$$

To study the limiting moments of V_n, we impose a regularity condition on sampling points $\{t_i\}_{i=1}^n$ as follows.

Condition 1. *For $n \ge 2$, define $t_i = \varphi((i-1)/(n-1)), i = 1, \cdots n$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is a twice continuously differentiable function satisfying $\varphi(0) = 0, \varphi(1) = 1$ and*

$$\min_{0\le s\le 1} \varphi^{(1)}(s) > 0.$$

It follows from Condition 1 that there exist positive constants C_0 and C_1 such that

$$0 < C_0/n \le \min_{1\le i\le n-1}(t_{i+1}-t_i) \le \max_{1\le i\le n-1}(t_{i+1}-t_i) \le C_1/n. \tag{2.13}$$

With the Taylor expansion, we have

$$e^x = 1 + x + \frac{e^{x\theta_x}}{2!}x^2, \quad 0<\theta_x<1. \tag{2.14}$$

So

$$\Gamma_\mu(t) = \sigma^2 - \sigma^2\mu|t| + \frac{e^{-\mu|t|\theta}}{2!}\sigma^2\mu^2|t|^2, \tag{2.15}$$

where $\quad 0<\theta<1$ depends on $(-\mu|t|)$. Then

$$\begin{aligned}
\mathrm{E}((\nabla X_i)^2) &= \sum_{k_1=0}^{1}\sum_{k_2=0}^{1} a_{i,k_1}a_{i,k_2}\Gamma_\mu(t_{(i+k_2)}-t_{(i+k_1)}) \\
&= -\sigma^2\mu\sum_{k_1=0}^{1}\sum_{k_2=0}^{1} a_{i,k_1}a_{i,k_2}|t_{(i+k_2)}-t_{(i+k_1)}| \\
&+ \sum_{k_1=0}^{1}\sum_{k_2=0}^{1} a_{i,k_1}a_{i,k_2}\frac{e^{-\mu|t_{(i+k_2)}-t_{(i+k_1)}|\theta}}{2!}\sigma^2\mu^2|t_{(i+k_2)}-t_{(i+k_1)}|^2 \\
&= -2(\sigma^2\mu)a_{i,0}a_{i,1}(|t_{i+1}-t_i|) + O(1).
\end{aligned} \tag{2.16}$$

And the $O(1)$ in the above equation comes from the fact that the last term in the second last equation can be uniformly bounded over $i = 1, \cdots, (n-1)$, namely,

$$\begin{aligned}
&\left|\sum_{k_1=0}^{1}\sum_{k_2=0}^{1} a_{i,k_1}a_{i,k_2}\frac{e^{-\mu|t_{(i+k_2)}-t_{(i+k_1)}|\theta}}{2!}\sigma^2\mu^2|t_{(i+k_2)}-t_{(i+k_1)}|^2\right| \\
\le&\sum_{k_1=0}^{1}\sum_{k_2=0}^{1}\left|a_{i,k_1}a_{i,k_2}\frac{e^{-\mu|t_{(i+k_2)}-t_{(i+k_1)}|\theta}}{2!}\sigma^2\mu^2|t_{(i+k_2)}-t_{(i+k_1)}|^2\right| \\
\le&\sum_{k_1=0}^{1}\sum_{k_2=0}^{1}\left|a_{i,k_1}a_{i,k_2}\frac{1}{2}\sigma^2\mu^2|t_{(i+k_2)}-t_{(i+k_1)}|^2\right| \\
=&\left|a_{i,k_0}a_{i,k_1}\Delta_i^2\right|\sigma^2\mu^2 \\
=&\sigma^2\mu^2.
\end{aligned} \tag{2.17}$$

Meanwhile, by the Taylor expansion again, we have

$$\begin{aligned}
&-2(\sigma^2\mu)\sum_{i=1}^{n-1}(a_{i,0}a_{i,1}|t_{i+1}-t_i|)\\
=&2(\sigma^2\mu)\sum_{i=1}^{n-1}1/\left(\varphi(\frac{i}{n-1})-\varphi(\frac{i-1}{n-1})\right)\\
=&2(\sigma^2\mu)\sum_{i=1}^{n-1}1/\left(\varphi^{(1)}(\frac{i-1}{n-1})\frac{1}{(n-1)}+\varphi^{(2)}(\frac{i-1}{n-1}+\frac{\theta_i}{n-1})\frac{1}{2(n-1)^2}\right) \qquad (2.18)\\
=&2(\sigma^2\mu)\sum_{i=1}^{n-1}\frac{(n-1)}{\varphi^{(1)}(\frac{i-1}{n-1})}\left(1-\left(1+\alpha_i\frac{\varphi^{(2)}(\frac{i-1}{n-1}+\frac{\theta_i}{n-1})}{2(n-1)\varphi^{(1)}(\frac{i-1}{n-1})}\right)^{-2}\frac{\varphi^{(2)}(\frac{i-1}{n-1}+\frac{\theta_i}{n-1})}{2(n-1)\varphi^{(1)}(\frac{i-1}{n-1})}\right)\\
=&2(\sigma^2\mu)n^2\int_0^1\{\varphi^{(1)}(s)\}^{-1}ds+O(n), \quad \text{as } n\to\infty,
\end{aligned}$$

where $\theta_i, \alpha_i \in (0,1)$ depend on i. Notice that the second last equation above holds when n is big enough since

$$\frac{\varphi^{(2)}(\frac{i-1}{n-1}+\frac{\theta_i}{n-1})}{2(n-1)\varphi^{(1)}(\frac{i-1}{n-1})}>-1,$$

uniformly over $i=1,\cdots,(n-1)$ under this case. As for the $O(n)$ term in Eq. (2.18), we only need to notice that for n sufficiently big,

$$\left(1+\alpha_i\frac{\varphi^{(2)}(\frac{i-1}{n-1}+\frac{\theta_i}{n-1})}{2(n-1)\varphi^{(1)}(\frac{i-1}{n-1})}\right)^{-2}<2,$$

uniformly over $i=1,\cdots,(n-1)$. Therefore,

$$\left|\frac{(n-1)}{\varphi^{(1)}(\frac{i-1}{n-1})}\left(1+\alpha_i\frac{\varphi^{(2)}(\frac{i-1}{n-1}+\frac{\theta_i}{n-1})}{2(n-1)\varphi^{(1)}(\frac{i-1}{n-1})}\right)^{-2}\frac{\varphi^{(2)}(\frac{i-1}{n-1}+\frac{\theta_i}{n-1})}{2(n-1)\varphi^{(1)}(\frac{i-1}{n-1})}\right|\leq\frac{\max_{x\in[0,1]}|\varphi^{(2)}(x)|}{\min_{x\in[0,1]}|\varphi^{(1)}(x)|^2}, \qquad (2.19)$$

uniformly over $i=1,\cdots,(n-1)$ when n big enough. Therefore, with Eqs. (2.16) and (2.18), we have

$$\begin{aligned}
\mathrm{E}(V_n)&=\sum_{i=1}^{n-1}\mathrm{E}((\nabla X_i)^2)\\
&=-2(\sigma^2\mu)\sum_{i=1}^{n-1}a_{i,0}a_{i,1}(|t_{i+1}-t_i|)+O(n) \qquad (2.20)\\
&=2(\sigma^2\mu)n^2\int_0^1\{\varphi^{(1)}(s)\}^{-1}ds+O(n),
\end{aligned}$$

as $n \to \infty$. From Eq. (2.20), we expect

$$V_n \Big/ \left(2n^2 \int_0^1 \{\varphi^{(1)}(s)\}^{-1} ds \right) \xrightarrow{\text{a.s.}} \sigma^2 \mu, \qquad \text{as } n \to \infty.$$

So we first study $V_n / \operatorname{E}(V_n)$ and then replace $\operatorname{E}(V_n)$ by $2n^2 \int_0^1 \{\varphi^{(1)}(s)\}^{-1} ds$ to get the desired estimator of $\sigma^2\mu$.

Lemma 2.1. *If the sampling function φ meets Condition 1, then*

$$\lim_{n\to\infty} n \operatorname{var}(V_n / \operatorname{E}(V_n)) = 2 \frac{\int_0^1 \{\varphi^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi^{(1)}(s)\}^{-1} ds \right)^2} \triangleq \phi_0. \tag{2.21}$$

Proof. By Eq. (2.11), we have

$$\begin{aligned}
\operatorname{var}(V_n) &= \operatorname{var}\left(\sum_{i=1}^{n-1} (\nabla X_i)^2 \right) \\
&= \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \operatorname{cov}((\nabla X_i)^2, (\nabla X_j)^2) \\
&= \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} 2[\operatorname{E}(\nabla X_i \nabla X_j)]^2 \\
&= \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} 2 \left[\sum_{k_1=0}^{1} \sum_{k_2=0}^{1} a_{i,k_1} a_{j,k_2} \Gamma_\mu (t_{i+k_1} - t_{j+k_2}) \right]^2.
\end{aligned} \tag{2.22}$$

If $i = j$, similar to the derivation of $\operatorname{E}(V_n)$, we have

$$\begin{aligned}
P_n \triangleq & \sum_{i=1}^{n-1} 2 \left[\sum_{k_1=0}^{1} \sum_{k_2=0}^{1} a_{i,k_1} a_{i,k_2} \Gamma_\mu (t_{i+k_1} - t_{i+k_2}) \right]^2 \\
=& 2 \sum_{i=1}^{n-1} \left[\frac{2(\sigma^2\mu)(n-1)}{\varphi^{(1)}(\frac{i-1}{n-1})} \right. \\
& \left. \times \left(1 - \left(1 + \alpha_i \frac{\varphi^{(2)}(\frac{i-1}{n-1} + \frac{\theta_i}{n-1})}{2(n-1)\varphi^{(1)}(\frac{i-1}{n-1})} \right)^{-2} \frac{\varphi^{(2)}(\frac{i-1}{n-1} + \frac{\theta_i}{n-1})}{2(n-1)\varphi^{(1)}(\frac{i-1}{n-1})} \right) + O(1) \right]^2 \\
=& 2 \sum_{i=1}^{n-1} \left[2(\sigma^2\mu) \frac{(n-1)}{\varphi^{(1)}(\frac{i-1}{n-1})} + O(1) \right]^2 \\
=& 2 \left[(2\sigma^2\mu n)^2 n \int_0^1 \{\varphi^{(1)}(s)\}^{-2} ds + O(n^2) \right].
\end{aligned} \tag{2.23}$$

If $i \neq j$, because of the symmetry, we only need to study the case $j > i$, then $t_{j+k_2} - t_{i+k_1} \geq 0$ for all $k_1, k_2 = 0, 1$. By the Taylor expansion (see, e.g., Loh 2015, Lemma 4),

$$\Gamma_\mu(t_{j+k_2} - t_{i+k_1}) = \sum_{s=0}^{1} \frac{\Gamma_\mu^{(s)}(t_j - t_i)}{s!}[(t_{j+k_2} - t_{i+k_1}) - (t_j - t_i)]^s + \int_{t_j - t_i}^{t_{j+k_2} - t_{i+k_1}} [(t_{j+k_2} - t_{i+k_1}) - t]\Gamma_\mu^{(2)}(t)dt.$$

Meanwhile, regarding $\Gamma_\mu(t) = \sigma^2 e^{-\mu|t|}$, $\Gamma_\mu^{(2)}(t) = \sigma^2\mu^2 e^{-\mu t}$ for $t \geq 0$. So

$$\begin{aligned} &\left|\sum_{k_1=0}^{1}\sum_{k_2=0}^{1} a_{i,k_1}a_{j,k_2}\Gamma_\mu(t_{j+k_2} - t_{i+k_1})\right| \\ =&\left|\sum_{k_1=0}^{1}\sum_{k_2=0}^{1} a_{i,k_1}a_{j,k_2}\left(\int_{t_j-t_i}^{t_{j+k_2}-t_{i+k_1}} [(t_{j+k_2} - t_{i+k_1}) - t]\,\Gamma_\mu^{(2)}(t)dt\right)\right| \\ \leq&\sum_{k_1=0}^{1}\sum_{k_2=0}^{1} |a_{i,k_1}a_{j,k_2}|\left(\int_{\mathcal{I}} |(t_{j+k_2} - t_{i+k_1}) - t|\,dt\right)\max_{t\in\mathcal{I}}\left|\Gamma_\mu^{(2)}(t)\right| \\ \leq&4\frac{C_1^2}{C_0^2}\sigma^2\mu^2, \end{aligned} \tag{2.24}$$

where $\mathcal{I} \triangleq ((t_j - t_i) \wedge (t_{j+k_2} - t_{i+k_1}), (t_j - t_i) \vee (t_{j+k_2} - t_{i+k_1}))$ and the last inequality holds by noticing that

$$\max_{k_1,k_2} |(t_{j+k_2} - t_{i+k_1}) - (t_j - t_i)| \leq \max\{\Delta_i, \Delta_j\} \leq \frac{C_1}{n}. \tag{2.25}$$

Therefore,

$$Q_n \triangleq \sum_{i\neq j} 2\left[\sum_{k_1=0}^{1}\sum_{k_2=0}^{1} a_{i,k_1}a_{j,k_2}\Gamma_\mu(t_{i+k_1} - t_{j+k_2})\right]^2 = O(n^2). \tag{2.26}$$

Combining Eqs. (2.22), (2.23) and (2.26), we have

$$\begin{aligned} &\lim_{n\to\infty} n\,\mathrm{var}(V_n/\,\mathrm{E}(V_n)) \\ =&\lim_{n\to\infty} n\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} 2\left[\sum_{k_1=0}^{1}\sum_{k_2=0}^{1} a_{i,k_1}a_{j,k_2}\Gamma_\mu(t_{i+k_1} - t_{j+k_2})\right]^2 /(\mathrm{E}(V_n))^2 \\ =&\lim_{n\to\infty} n(P_n + Q_n)/f_n^2 \\ =&2\frac{\int_0^1(\varphi^{(1)}(s))^{-2}ds}{\left[\int_0^1(\varphi^{(1)}(s))^{-1}ds\right]^2}. \end{aligned} \tag{2.27}$$

□

Remark 2.1. *By Jensen's inequality,*

$$\frac{\int_0^1 (\varphi^{(1)}(s))^{-2} ds}{[\int_0^1 (\varphi^{(1)}(s))^{-1} ds]^2} \geq 1,$$

and the equality holds if and only if $\varphi^{(1)}(s) = const$ on $[0,1]$. This implies the estimator has the same asymptotic variance as that of MLE only when the sampling scheme is regular. If we consider a family of sampling functions

$$\left\{ \varphi_i(x) = \frac{i}{i+2} \left((x + \frac{1}{i})^2 - \frac{1}{i^2} \right)^2 : \quad i \in \mathbb{N}_+ \right\},$$

then

$$\frac{\int_0^1 \{\varphi_i^{(1)}(s)\}^{-2} ds}{\left[\int_0^1 \{\varphi_i^{(1)}(s)\}^{-1} ds\right]^2} \to +\infty, \quad \text{as } i \to \infty.$$

Loh (2015) showed that $V_n / \mathrm{E}(V_n) \xrightarrow{\text{a.s.}} 1$. In the following, we derive a weaker upper bound for the convergence rate based on Lemma 2.1, which will be used in the subsequent multidimensional analysis.

Lemma 2.2. $\forall \xi > 0$, *there exist a finite constant $C_5 > 0$ such that when n is sufficiently large,*

$$\mathrm{P}\left(\left| \frac{V_n}{\mathrm{E}(V_n)} - 1 \right| \geq \xi \right) \leq 2e^{-C_5 \xi \sqrt{n} / \sqrt{\phi_0}}. \tag{2.28}$$

Proof. The proof follows the argument of Lemma 1 given in Zhou and Xiao (2018). Writing

$$\begin{aligned} \mathbf{Y}_n &= \left(\frac{\nabla X_1}{\sqrt{\mathrm{E}(V_n)}}, \cdots, \frac{\nabla X_{n-1}}{\sqrt{\mathrm{E}(V_n)}} \right)', \\ \Sigma_n &= (\Sigma_{i,j})_{(n-1)\times(n-1)} = \mathrm{cov}(\mathbf{Y}_n). \end{aligned} \tag{2.29}$$

Let $\Lambda_n = \mathrm{diag}(\lambda_1, \cdots, \lambda_{n-1})$ be the diagonal matrix whose diagonal entries are the eigenvalues of Σ_n, and let $\mathbf{U} = (U_1, \cdots, U_{n-1})' \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{n-1})$. Then, we have

$$V_n / \mathrm{E}(V_n) = \mathbf{Y}_n' \mathbf{Y}_n \stackrel{d}{=} \mathbf{U}' \Lambda_n \mathbf{U}.$$

We apply the Hanson and Wright inequality (Hanson and Wright 1971) to bound the tail probability of the quadratic forms and obtain

$$\begin{aligned} &\mathrm{P}(|V_n/\mathrm{E}(V_n) - \mathrm{E}(V_n/\mathrm{E}(V_n))| > \xi) \\ =&\mathrm{P}(|\mathbf{U}'\Lambda_n\mathbf{U} - \mathrm{tr}(\Lambda_n)| > \xi) \\ \leq& 2\exp\left\{-\min\left(C_3\frac{\xi}{\|\Lambda_n\|_2}, C_4\frac{\xi^2}{\|\Lambda_n\|_F^2}\right)\right\}, \end{aligned} \tag{2.30}$$

where $\|\Lambda_n\|_2$ and $\|\Lambda_n\|_F$ are the ℓ_2 norm and Frobenius norm of Λ_n, respectively; C_3 and C_4 are positive constants independent of Λ_n and ξ. Note that

$$\phi_n \triangleq \mathrm{var}(V_n/\mathrm{E}(V_n)) = \mathrm{var}(\mathbf{U}'\Lambda_n\mathbf{U}) = 2\mathrm{tr}(\Lambda_n^2) = 2\|\Lambda_n\|_F^2.$$

By Lemma 2.1, we have $\|\Lambda_n\|_F^2 = \frac{1}{2}\phi_n \sim \frac{1}{2n}\phi_0$. Combining the above with the fact that $\|\Lambda_n\|_2 \leq \|\Lambda_n\|_F$, we see that

$$\frac{1}{\|\Lambda_n\|_2} \geq \sqrt{\frac{n}{\phi_0}},$$

when n is sufficiently large. Meanwhile $\frac{1}{\|\Lambda_n\|_F^2} \sim \frac{2n}{\phi_0}$ as $n \to \infty$. Hence, Eq. (2.30) decays exponentially with rate $\sqrt{n}$ when $n \to \infty$. Consequently, when n is sufficiently large,

$$\mathrm{P}(|V_n/\mathrm{E}(V_n) - 1| > \xi) \leq 2e^{-C_5\xi\sqrt{n}/\sqrt{\phi_0}},$$

where $C_5 \triangleq C_3$. □

Theorem 2.1. *Let V_n be as in Eq.* (2.11)*, and suppose Condition 1 holds. Then*

$$V_n/\mathrm{E}(V_n) \xrightarrow{a.s.} 1, \qquad \text{as } n \to \infty. \tag{2.31}$$

Proof. Combining Lemma 2.2 and the Borel-Cantelli lemma, we have the conclusion. □

2.3 Sampling over $\mathbb{R}^2$

Let $X(\mathbf{u}), \mathbf{u} \in [0,1]^2$ be a zero-mean univariate Gaussian field with covariance function

$$\mathrm{cov}\left(X\left(\mathbf{u}\right), X\left(\mathbf{v}\right)\right) = \sigma^2 \exp\left(-\lambda|\mathbf{u}[1] - \mathbf{v}[1]| - \mu|\mathbf{u}[2] - \mathbf{v}[2]|\right), \tag{2.32}$$

where $(\sigma^2, \lambda, \mu)' \in \mathbb{R}_+^3$. By Ying (1993), we know that σ^2, λ and μ can be consistently estimated.

Let φ_h and φ_v satisfy Condition 1. We consider the observation locations over $\mathbb{R}^2$ which are of the form

$$\mathbf{u}(\mathbf{i}) \triangleq \left(\varphi_h\left(\frac{\mathbf{i}[1]-1}{m-1}\right), \varphi_v\left(\frac{\mathbf{i}[2]-1}{n-1}\right)\right)',$$

where $1 \leq \mathbf{i}[1] \leq m, 1 \leq \mathbf{i}[2] \leq n$, and $n/m \to \rho \in (0, \infty)$. For simplicity, we write $X_{\mathbf{i}} = X(\mathbf{u}(\mathbf{i}))$. Since the observation region is in $\mathbb{R}^2$, we will use letters v and h to emphasize operations in the sense of the vertical and horizontal directions respectively. To simplify the notation, let

$$\varphi_{(h,j)} = \varphi_h\left(\frac{j-1}{m-1}\right), \qquad \varphi_{(v,l)} = \varphi_v\left(\frac{l-1}{n-1}\right),$$

for $j = 1, \cdots, m$ and $l = 1, \cdots, n$. Let

$$\Delta_{h,j} = \varphi_{(h,j+1)} - \varphi_{(h,j)},$$
$$\Delta_{v,k} = \varphi_{(v,k+1)} - \varphi_{(v,k)},$$

for $j = 1, \cdots, (m-1)$ and $k = 1, \cdots, (n-1)$. Define the index set

$$\mathcal{I}_{(m,n)} = \{\mathbf{i} : (1,1)' \leq \mathbf{i} \leq (m-1, n-1)'\}, \tag{2.33}$$

If $\mathbf{j} \in \mathcal{I}_{(m,n)}$ is a vector, $\Delta_{h,\mathbf{j}}$ and $\Delta_{v,\mathbf{j}}$ are defined as

$$\Delta_{h,\mathbf{j}} = \varphi_{(h,\mathbf{j}[1]+1)} - \varphi_{(h,\mathbf{j}[1])},$$
$$\Delta_{v,\mathbf{j}} = \varphi_{(v,\mathbf{j}[2]+1)} - \varphi_{(v,\mathbf{j}[2])}.$$

Similar to the increment in Eq. (2.10), we define 4 sets of increments

$$\{a_{h;\mathbf{i},0}, a_{h;\mathbf{i},1}\}, \qquad \{a_{h;\mathbf{i},(0,0)}, a_{h;\mathbf{i},(1,0)}\},$$
$$\{a_{v;\mathbf{i},0}, a_{v;\mathbf{i},1}\}, \qquad \{a_{v;\mathbf{i},(0,0)}, a_{v;\mathbf{i},(0,1)}\},$$

as

$$a_{h;\mathbf{i},0} = a_{h;\mathbf{i},(0,0)} = -1/\Delta_{h,\mathbf{i}}, \qquad a_{h;\mathbf{i},1} = a_{h;\mathbf{i},(1,0)} = 1/\Delta_{h,\mathbf{i}}, \tag{2.34}$$

$$a_{v;\mathbf{i},0} = a_{v;\mathbf{i},(0,0)} = -1/\Delta_{v,\mathbf{i}}, \qquad a_{v;\mathbf{i},1} = a_{v;\mathbf{i},(0,1)} = 1/\Delta_{v,\mathbf{i}}. \tag{2.35}$$

Based on Condition 1, we also have positive constants $C_{h,0}$ and $C_{h,1}$ such that

$$\begin{aligned} 0 < C_{h,0}/m \le \min_{1\le i\le (m-1)} \left(\varphi_h\left(\frac{i}{m-1}\right) - \varphi_h\left(\frac{i-1}{m-1}\right)\right) \\ \le \max_{1\le i\le (m-1)} \left(\varphi_h\left(\frac{i}{m-1}\right) - \varphi_h\left(\frac{i-1}{m-1}\right)\right) \le C_{h,1}/m. \end{aligned}$$

$C_{v,0}$ and $C_{v,1}$ can be similarly defined.

2.3.1 The vertical and horizontal increments

Similar to the definition of V_n in Eq. (2.11) for the univariate case, for each fixed $j = 1, \cdots, (m-1)$, let

$$V_{v,j} = \sum_{k=1}^{n-1} \left(\nabla_v X_{(j,k)}\right)^2, \tag{2.36}$$

where

$$\nabla_v X_{(j,k)} = a_{v;(j,k),(0,0)} X_{(j,k)} + a_{v;(j,k),(0,1)} X_{(j,k+1)}. \tag{2.37}$$

From the structure of the covariance function Eq. (2.32), the Gaussian process $X_{(i,\cdot)}$ has the same distribution across $1 \le i \le m$. Therefore, to estimate $\sigma^2\mu$, we first construct the weighted quantity

$$\bar{Z}_{v,(m,n)} = \sum_{j=1}^{m-1} \Delta_{h,j} V_{v,j} / \operatorname{E}(V_{v,j}) = \sum_{j=1}^{m-1} \Delta_{h,j} V_{v,j} / \operatorname{E}(V_{v,1}). \tag{2.38}$$

Switching the direction from vertical to horizontal, for each fixed $j = 1, \cdots, (n-1)$, we can also define

$$\begin{aligned} \nabla_h X_{(k,j)} &= a_{h;(k,j),(0,0)} X_{(k,j)} + a_{h;(k,j),(1,0)} X_{(k+1,j)}, \\ V_{h,j} &= \sum_{k=1}^{m-1} \left(\nabla_h X_{(k,j)}\right)^2. \end{aligned}$$

Therefore, to estimate $\sigma^2\lambda$, we construct the weighted quantity

$$\bar{Z}_{h,(m,n)} = \sum_{j=1}^{n-1} \Delta_{v,j} V_{h,j} / \operatorname{E}(V_{h,j}) = \sum_{j=1}^{n-1} \Delta_{v,j} V_{h,j} / \operatorname{E}(V_{h,1}). \tag{2.39}$$

From the definitions of $\bar{Z}_{v,(m,n)}$ and $\bar{Z}_{h,(m,n)}$, it is reasonable to just study statistical properties of $\bar{Z}_{v,(m,n)}$. To analyze the variance of $\bar{Z}_{v,(m,n)}$, we first note that for any $j, k = 1, \cdots, (m-1)$,

$$\operatorname{cov}(V_{v,j}, V_{v,k}) = e^{-2\lambda\left|\varphi_{(h,j)} - \varphi_{(h,k)}\right|} \operatorname{var}(V_{v,1}). \tag{2.40}$$

Actually, denote

$$P(\lambda) = \begin{pmatrix} 1 & e^{-\lambda\left|\varphi_{(h,j)} - \varphi_{(h,k)}\right|} \\ e^{-\lambda\left|\varphi_{(h,j)} - \varphi_{(h,k)}\right|} & 1 \end{pmatrix}, \tag{2.41}$$

and

$$B(\mu) = \left(e^{-\mu\left|\varphi_{(v,p)} - \varphi_{(v,q)}\right|}\right)_{n \times n}, \tag{2.42}$$

where $p, q = 1, \cdots, n$, then the $2n$-vector

$$Y_{2n} \triangleq \Big(\big(X_{(j,l)}, l \in \{1, \cdots, n\}\big), \big(X_{(k,l)}, l \in \{1, \cdots, n\}\big)\Big)'$$

has the distribution

$$Y_{2n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 P(\lambda) \otimes B(\mu)), \tag{2.43}$$

where $\otimes$ is the Kronecker product. Let

$$R_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad R_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Since $V_{v,j}$ is a quadratic form, it can be expressed as

$$V_{v,j} = Y_{2n}'(R_1 \otimes F) Y_{2n}, \tag{2.44}$$

where F is a $n \times n$ semi-positive definite matrix. Similarly,

$$V_{v,k} = Y_{2n}'(R_2 \otimes F) Y_{2n}. \tag{2.45}$$

Therefore,

$$\begin{aligned}
&\operatorname{cov}(V_{v,j}, V_{v,k})\\
=&\operatorname{cov}\left(Y_{2n}'(R_1 \otimes F)Y_{2n}, Y_{2n}'(R_2 \otimes F)Y_{2n}\right)\\
=&2\operatorname{tr}\big((R_1 \otimes F)\sigma^2(P(\lambda) \otimes B(\mu))(R_2 \otimes F)\sigma^2(P(\lambda) \otimes B(\mu))\big)\\
=&2\sigma^4 \operatorname{tr}((R_1 P(\lambda) R_2 P(\lambda)) \otimes (FB(\mu)FB(\mu)))\\
=&e^{-2\lambda\left|\varphi_{(h,j)}-\varphi_{(h,k)}\right|} 2\sigma^4 \operatorname{tr}((FB(\mu)FB(\mu)))\\
=&e^{-2\lambda\left|\varphi_{(h,j)}-\varphi_{(h,k)}\right|} \operatorname{var}(V_{v,1}).
\end{aligned} \tag{2.46}$$

Based on Eqs. (2.40) and (2.46), we have

$$\begin{aligned}
\operatorname{var}\left(\bar{Z}_{v,(m,n)}\right) &= \operatorname{var}\left(\sum_{j=1}^{m-1} \Delta_{h,j} V_{v,j} / \operatorname{E}(V_{v,1})\right)\\
&= \operatorname{var}\left(\frac{V_{v,1}}{\operatorname{E}(V_{v,1})}\right) \sum_{j=1}^{m-1}\sum_{k=1}^{m-1} e^{-2\lambda\left|\varphi_{(h,j)}-\varphi_{(h,k)}\right|}(\Delta_{h,j}\Delta_{h,k}).
\end{aligned} \tag{2.47}$$

Based on the definition of $\{\Delta_{h,j}\}_{j=1}^{m-1}$ and Condition 1, we have

$$\sum_{j=1}^{m-1}\sum_{k=1}^{m-1} e^{-2\lambda\left|\varphi_{(h,j)}-\varphi_{(h,k)}\right|}(\Delta_{h,j}\Delta_{h,k}) \xrightarrow{m\to\infty} \iint_{[0,1]^2} e^{-2\lambda|x-y|}\, dx\, dy. \tag{2.48}$$

A direct calculation shows that

$$C_\lambda \triangleq \iint_{[0,1]^2} e^{-2\lambda|x-y|}\, dx\, dy = \frac{1}{\lambda} - \frac{1-e^{-2\lambda}}{2\lambda^2}. \tag{2.49}$$

We are ready to prove the following theorem on asymptotic properties of $\bar{Z}_{v,(m,n)}$.

Theorem 2.2. *Let $\bar{Z}_{v,(m,n)}$ be as in Eq. (2.38). Suppose the two sampling functions $\varphi_h(\cdot)$ and $\varphi_v(\cdot)$ satisfy Condition 1. Then*

$$\lim_{n\to\infty} n \operatorname{var}\left(\bar{Z}_{v,(m,n)}\right) = 2C_\lambda \frac{\int_0^1 \{\varphi_v^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_v^{(1)}(s)\}^{-1} ds\right)^2}, \tag{2.50}$$

and

$$\bar{Z}_{v,(m,n)} \xrightarrow[n\to\infty]{a.s.} 1. \tag{2.51}$$

Proof. Combining Lemma 2.1 and Eqs. (2.47) to (2.49), we have Eq. (2.50). As for the almost sure convergence, let $\xi > 0$, then by Lemma 2.2,

$$\begin{aligned} &\mathrm{P}\left(|\bar{Z}_{v,(m,n)} - \mathrm{E}(\bar{Z}_{v,(m,n)})| > \xi\right) \\ \leq &\mathrm{P}\left(\sum_{i=1}^{m-1} \Delta_{h,i}\left|\frac{V_{v,i}}{\mathrm{E}(V_{v,i})} - 1\right| > \xi\right) \\ \leq &\sum_{i=1}^{m-1} \mathrm{P}\left(\Delta_{h,i}\left|\frac{V_{v,i}}{\mathrm{E}(V_{v,i})} - 1\right| > \xi/m\right) \\ \leq &m\,\mathrm{P}\left(\left|\frac{V_{v,1}}{\mathrm{E}(V_{v,1})} - 1\right| > \xi/C_{h,1}\right) \\ \leq &2me^{-C_5\sqrt{n}\xi/C_{h,1}/\sqrt{\phi_0}}, \end{aligned} \tag{2.52}$$

when n is big enough. So applying the Borel-Cantelli lemma, we have the almost sure convergence. □

2.3.2 The square increment

In this part, we focus on estimating $\sigma^2\mu\lambda$. Define the increment $\{b_{\mathbf{k}}\}_{\mathbf{k}=\mathbf{0}}^{\mathbf{1}}$ as

$$b_{(0,0)} = b_{(1,1)} = 1, \qquad\qquad b_{(1,0)} = b_{(0,1)} = -1,$$

and we call it the square increment since its support is the four vertices of the square; (see Chan and Wood 2000, for a fuller description of it) . In the following, we reserve the symbol s to emphasize the square increment. Based on the structure of the covariance function, for each fixed $1 \leq j \leq m$, $\{X_{(j,l)}\}_{l=1}^{n}$ contains information of $\sigma^2\mu$. And for each fixed $1 \leq l \leq n$, $\{X_{(j,l)}\}_{j=1}^{m}$ contains information of $\sigma^2\lambda$. So naturally we define the quantity as

$$\nabla_s X_{\mathbf{i}} = \frac{\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{1}} b_{\mathbf{k}} X_{\mathbf{i}+\mathbf{k}}}{\Delta_{h,\mathbf{i}}\Delta_{v,\mathbf{i}}}. \tag{2.53}$$

Define the quadratic variation

$$V_{s,(m,n)} = \sum_{\mathbf{i}\in\mathcal{I}_{(m,n)}} (\nabla_s X_{\mathbf{i}})^2, \tag{2.54}$$

and its transformation

$$\bar{Z}_{s,(m,n)} = V_{s,(m,n)}/\,\mathrm{E}\left(V_{s,(m,n)}\right), \tag{2.55}$$

then for fixed index $\mathbf{i}, \mathbf{j} \in \mathcal{I}_{(m,n)}$,

$$
\begin{aligned}
& \mathrm{E}\left(\nabla_s X_{\mathbf{i}} \nabla_s X_{\mathbf{j}}\right) \\
=& \frac{1}{\Delta_{h,\mathbf{i}}\Delta_{v,\mathbf{i}}\Delta_{h,\mathbf{j}}\Delta_{v,\mathbf{j}}} \sum_{\mathbf{k},\mathbf{l}=\mathbf{0}}^{\mathbf{1}} b_{\mathbf{k}} b_{\mathbf{l}} \, \mathrm{E}\left(X_{\mathbf{i}+\mathbf{k}} X_{\mathbf{j}+\mathbf{l}}\right) \\
=& \frac{\sigma^2}{\Delta_{h,\mathbf{i}}\Delta_{v,\mathbf{i}}\Delta_{h,\mathbf{j}}\Delta_{v,\mathbf{j}}} \sum_{\mathbf{k},\mathbf{l}=\mathbf{0}}^{\mathbf{1}} b_{\mathbf{k}} b_{\mathbf{l}} e^{-\lambda\left|\varphi_{(h,(\mathbf{i}+\mathbf{k})[1])}-\varphi_{(h,(\mathbf{j}+\mathbf{l})[1])}\right|-\mu\left|\varphi_{(v,(\mathbf{i}+\mathbf{k})[2])}-\varphi_{(v,(\mathbf{j}+\mathbf{l})[2])}\right|} \\
=& \frac{\sigma^2}{\Delta_{h,\mathbf{i}}\Delta_{h,\mathbf{j}}} \left(\sum_{l_1,k_1=0}^{1} \mathtt{b}_{k_1}\mathtt{b}_{l_1} e^{-\lambda\left|\varphi_{(h,(\mathbf{i}[1]+k_1))}-\varphi_{(h,(\mathbf{j}[1]+l_1))}\right|} \right) \\
& \times \frac{1}{\Delta_{v,\mathbf{i}}\Delta_{v,\mathbf{j}}} \left(\sum_{l_2,k_2=0}^{1} \mathtt{b}_{k_2}\mathtt{b}_{l_2} e^{-\mu\left|\varphi_{(v,(\mathbf{i}[2]+k_2))}-\varphi_{(v,(\mathbf{j}[2]+l_2))}\right|} \right) \\
=& \sigma^2 \left(\sum_{l,k=0}^{1} a_{h;\mathbf{i},k} a_{h;\mathbf{j},l} e^{-\lambda\left|\varphi_{(h,(\mathbf{i}[1]+k))}-\varphi_{(h,(\mathbf{j}[1]+l))}\right|} \right) \\
& \times \left(\sum_{l,k=0}^{1} a_{v;\mathbf{i},k} a_{v;\mathbf{j},l} e^{-\mu\left|\varphi_{(v,(\mathbf{i}[2]+k))}-\varphi_{(v,(\mathbf{j}[2]+l))}\right|} \right),
\end{aligned}
\tag{2.56}
$$

where $\mathtt{b}_0 = -1, \mathtt{b}_1 = 1$. So we have decomposed the covariance between $\nabla_s X_{\mathbf{i}}$ and $\nabla_s X_{\mathbf{j}}$ into the product of the covariance in the univariate case. Based on the derivation of Eqs. (2.18) and (2.20), we have

$$\begin{aligned}
\mathrm{E}\left(V_{s,(m,n)}\right) &= \sum_{\mathbf{i}\in\mathcal{I}_{(m,n)}} \mathrm{E}\left[(\nabla_s X_{\mathbf{i}})^2\right] \\
&= \sum_{\mathbf{i}\in\mathcal{I}_{(m,n)}} \sigma^2 \left(\sum_{l,k=0}^{1} a_{h;\mathbf{i},k} a_{h;\mathbf{i},l} e^{-\lambda\left|\varphi_{(h,(\mathbf{i}[1]+k))}-\varphi_{(h,(\mathbf{i}[1]+l))}\right|}\right) \\
&\quad \times \left(\sum_{l,k=0}^{1} a_{v;\mathbf{i},k} a_{v;\mathbf{i},l} e^{-\mu\left|\varphi_{(v,(\mathbf{i}[2]+k))}-\varphi_{(v,(\mathbf{i}[2]+l))}\right|}\right) \\
&= \sigma^2 \left(\sum_{i_1=1}^{m-1}\left(\sum_{l,k=0}^{1} a_{h;i_1,k} a_{h;i_1,l} e^{-\lambda\left|\varphi_{(h,(i_1+k))}-\varphi_{(h,(i_1+l))}\right|}\right)\right) \\
&\quad \times \left(\sum_{i_2=1}^{n-1}\left(\sum_{l,k=0}^{1} a_{v;i_2,k} a_{v;i_2,l} e^{-\mu\left|\varphi_{(v,(i_2+k))}-\varphi_{(v,(i_2+l))}\right|}\right)\right) \\
&= \sigma^2 \left(2\lambda m^2 \int_0^1 \{\varphi_h^{(1)}(s)\}^{-1} ds + O(m)\right) \\
&\quad \times \left(2\mu n^2 \int_0^1 \{\varphi_v^{(1)}(s)\}^{-1} ds + O(n)\right) \\
&= 4\sigma^2\lambda\mu(mn)^2 \int_0^1 \{\varphi_h^{(1)}(s)\}^{-1} ds \int_0^1 \{\varphi_v^{(1)}(s)\}^{-1} ds + O(n^3),
\end{aligned} \tag{2.57}$$

as $n \to \infty$.

Theorem 2.3. *Let $\bar{Z}_{s,(m,n)}$ be as in Eq. (2.55). Suppose the two sampling functions $\varphi_h(\cdot)$ and $\varphi_v(\cdot)$ satisfy Condition 1. Then*

$$\lim_{n\to\infty} mn \,\mathrm{var}\left(\bar{Z}_{s,(m,n)}\right) = 2\frac{\int_0^1 \{\varphi_v^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_v^{(1)}(s)\}^{-1} ds\right)^2} \frac{\int_0^1 \{\varphi_h^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_h^{(1)}(s)\}^{-1} ds\right)^2}, \tag{2.58}$$

and

$$\bar{Z}_{s,(m,n)} \xrightarrow[n\to\infty]{a.s.} 1. \tag{2.59}$$

Proof. As with the univariate case, the variance of $V_{s,(m,n)}$ can be written as

$$\begin{aligned}
\mathrm{var}\left(V_{s,(m,n)}\right) &= \mathrm{var}\left(\sum_{\mathbf{i}\in\mathcal{I}_{(m,n)}} (\nabla_s X_{\mathbf{i}})^2\right) \\
&= \sum_{\mathbf{i}\in\mathcal{I}_{(m,n)}} \sum_{\mathbf{j}\in\mathcal{I}_{(m,n)}} 2[\mathrm{E}\left(\nabla_s X_{\mathbf{i}} \nabla_s X_{\mathbf{j}}\right)]^2.
\end{aligned} \tag{2.60}$$

As with the derivation of Eq. (2.23), based on Eqs. (2.17) and (2.19), we have that

$$\sum_{l,k=0}^{1} a_{v;\mathbf{i},k} a_{v;\mathbf{i},l} e^{-\mu\left|\varphi_{(v,(\mathbf{i}[2]+k))}-\varphi_{(v,(\mathbf{i}[2]+l))}\right|} = 2\mu\frac{(n-1)}{\varphi_v^{(1)}\left(\frac{\mathbf{i}[2]-1}{n-1}\right)} + O(1), \tag{2.61}$$

as $n \to \infty$ uniformly over $\mathbf{i} \in \mathcal{I}_{(m,n)}$.

We first deal with the case $\mathbf{i} = \mathbf{j}$,

$$\begin{aligned}
P_{(m,n)} &\triangleq \sum_{\mathbf{i}\in\mathcal{I}_{(m,n)}} 2\left\{\mathrm{E}\left[(\nabla_s X_{\mathbf{i}})^2\right]\right\}^2 \\
&= \sum_{\mathbf{i}\in\mathcal{I}_{(m,n)}} 2\left\{\sigma^2\left(\sum_{l,k=0}^{1} a_{h;\mathbf{i},k} a_{h;\mathbf{i},l} e^{-\lambda\left|\varphi_{(h,(\mathbf{i}[1]+k))}-\varphi_{(h,(\mathbf{i}[1]+l))}\right|}\right)\right. \\
&\quad \times \left.\left(\sum_{l,k=0}^{1} a_{v;\mathbf{i},k} a_{v;\mathbf{i},l} e^{-\mu\left|\varphi_{(v,(\mathbf{i}[2]+k))}-\varphi_{(v,(\mathbf{i}[2]+l))}\right|}\right)\right\}^2 \\
&= 2\sigma^4 \sum_{\mathbf{i}\in\mathcal{I}_{(m,n)}} \left[2\lambda\frac{(m-1)}{\varphi_h^{(1)}(\frac{\mathbf{i}[1]-1}{m-1})} + O(1)\right]^2 \left[2\mu\frac{(n-1)}{\varphi_v^{(1)}(\frac{\mathbf{i}[2]-1}{n-1})} + O(1)\right]^2 \\
&= 2\sigma^4 \left\{\sum_{i_1=1}^{m-1}\left[2\lambda\frac{(m-1)}{\varphi_h^{(1)}(\frac{i_1-1}{m-1})} + O(1)\right]^2\right\} \\
&\quad \times \left\{\sum_{i_2=1}^{n-1}\left[2\mu\frac{(n-1)}{\varphi_v^{(1)}(\frac{i_2-1}{n-1})} + O(1)\right]^2\right\} \\
&= 2\sigma^4\left[(2\lambda m)^2 m\int_0^1 \{\varphi_h^{(1)}(s)\}^{-2}ds + O(m^2)\right] \\
&\quad \times \left[(2\mu n)^2 n\int_0^1 \{\varphi_v^{(1)}(s)\}^{-2}ds + O(n^2)\right] \\
&= 2(4\sigma^2\lambda\mu)^2(mn)^3\int_0^1 \{\varphi_h^{(1)}(s)\}^{-2}ds\int_0^1 \{\varphi_v^{(1)}(s)\}^{-2}ds + O(n^5),
\end{aligned} \tag{2.62}$$

as $n \to \infty$.

Secondly, we consider the case where $\mathbf{i}$ and $\mathbf{j}$ are different but in the same row or column. Because of the similarity between the row and the column case, we only show the case

$\mathbf{j}[1] = \mathbf{i}[1]$ and $\mathbf{j}[2] \neq \mathbf{i}[2]$. Based on Eq. (2.24), we have

$$\begin{aligned}
Q_{(m,n)} &\triangleq \sum_{\substack{\mathbf{i}\in\mathcal{I}_{(m,n)}\\ \mathbf{j}[1]=\mathbf{i}[1]\\ \mathbf{j}[2]\neq\mathbf{i}[2]}} 2\{\mathrm{E}\left[\nabla_s X_{\mathbf{i}} \nabla_s X_{\mathbf{j}}\right]\}^2 \\
&= \sum_{\substack{\mathbf{i}\in\mathcal{I}_{(m,n)}\\ \mathbf{j}[1]=\mathbf{i}[1]\\ \mathbf{j}[2]\neq\mathbf{i}[2]}} 2\Bigg\{\sigma^2 \left(\sum_{l,k=0}^{1} a_{h;\mathbf{i},k} a_{h;\mathbf{i},l} e^{-\lambda\left|\varphi_{(h,(\mathbf{i}[1]+k))}-\varphi_{(h,(\mathbf{i}[1]+l))}\right|}\right) \\
&\qquad \times \left(\sum_{l,k=0}^{1} a_{v;\mathbf{i},k} a_{v;\mathbf{j},l} e^{-\mu\left|\varphi_{(v,(\mathbf{i}[2]+k))}-\varphi_{(v,(\mathbf{j}[2]+l))}\right|}\right)\Bigg\}^2 \\
&\leq n \sum_{\mathbf{i}\in\mathcal{I}_{(m,n)}} 2\Bigg\{\sigma^2 \left(\sum_{l,k=0}^{1} a_{h;\mathbf{i},k} a_{h;\mathbf{i},l} e^{-\lambda\left|\varphi_{(h,(\mathbf{i}[1]+k))}-\varphi_{(h,(\mathbf{i}[1]+l))}\right|}\right) \qquad (2.63)\\
&\qquad \times \left(\left[4\frac{C_{v,1}^2}{C_{v,0}^2}\mu^2\right]\right)\Bigg\}^2 \\
&= 2\sigma^4 n \sum_{\mathbf{i}\in\mathcal{I}_{(m,n)}} \left[2\lambda \frac{(m-1)}{\varphi_h^{(1)}(\frac{\mathbf{i}[1]-1}{m-1})} + O(1)\right]^2 \left[4\frac{C_{v,1}^2}{C_{v,0}^2}\mu^2\right]^2 \\
&= 2\sigma^4 n \left\{\sum_{i_1=1}^{m-1} \left[2\lambda \frac{(m-1)}{\varphi_h^{(1)}(\frac{i_1-1}{m-1})} + O(1)\right]^2\right\} \left\{\sum_{i_2=1}^{n-1} \left[4\frac{C_{v,1}^2}{C_{v,0}^2}\mu^2\right]^2\right\} \\
&= 8\sigma^4 C_q \lambda^2 m^3 n^2 \int_0^1 \{\varphi_h^{(1)}(s)\}^{-2} ds + O(n^4),
\end{aligned}$$

as $n \to \infty$, where

$$C_q \triangleq 4\frac{C_{v,1}^2}{C_{v,0}^2}\mu^2.$$

Finally, we consider the case $\mathbf{j}[1] \neq \mathbf{i}[1]$ and $\mathbf{j}[2] \neq \mathbf{i}[2]$. Based on Eq. (2.24), it should be easy to see that

$$R_{(m,n)} \triangleq \sum_{\substack{\mathbf{i}\in\mathcal{I}_{(m,n)}\\ \mathbf{j}[1]\neq\mathbf{i}[1]\\ \mathbf{j}[2]\neq\mathbf{i}[2]}} 2\{\mathrm{E}\left[\nabla_s X_{\mathbf{i}} \nabla_s X_{\mathbf{j}}\right]\}^2 = O(n^4), \qquad (2.64)$$

as $n \to \infty$.

Combining Eqs. (2.62) to (2.64), we see that

$$\mathrm{var}\left(V_{s,(m,n)}\right) =$$

$$2(4\sigma^2\lambda\mu)^2(mn)^3 \int_0^1 \{\varphi_h^{(1)}(s)\}^{-2} ds \int_0^1 \{\varphi_v^{(1)}(s)\}^{-2} ds + O(n^5), \quad (2.65)$$

as $n \to \infty$. Consequently, it follows from Eqs. (2.57) and (2.65) that Eq. (2.58) holds.

With Eq. (2.58) on hand, following the proof of Theorem 2.1, we have

$$\bar{Z}_{s,(m,n)} \xrightarrow[n\to\infty]{\text{a.s.}} 1. \tag{2.66}$$

This proves Theorem 2.3. □

2.3.3 The asymptotic distribution of the estimator

In summary, we have those three quantities

$$\bar{Z}_{v,(m,n)} = \sum_{i=1}^{m-1} \Delta_{h,i} V_{v,i} / \operatorname{E}(V_{v,i}), \tag{2.67}$$

$$\bar{Z}_{h,(m,n)} = \sum_{i=1}^{n-1} \Delta_{v,i} V_{h,i} / \operatorname{E}(V_{h,i}), \tag{2.68}$$

$$\bar{Z}_{s,(m,n)} = V_{s,(m,n)} / \operatorname{E}\left(V_{s,(m,n)}\right). \tag{2.69}$$

Denote $\bar{Z}_{(m,n)} \triangleq \left(\bar{Z}_{v,(m,n)}, \bar{Z}_{h,(m,n)}, \bar{Z}_{s,(m,n)}\right)'$, and $\Phi_{(m,n)} \triangleq \operatorname{cov}\left(\bar{Z}_{(m,n)}\right)$.

Lemma 2.3. *Suppose the two sampling functions $\varphi_h(\cdot), \varphi_v(\cdot)$ satisfy Condition 1 and*

$$\lim_{m\to\infty} \frac{n}{m} = \rho.$$

Then

$$n\Phi_{(m,n)} \to \begin{pmatrix} 2C_\lambda \frac{\int_0^1 \{\varphi_v^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_v^{(1)}(s)\}^{-1} ds\right)^2} & 0 & 0 \\ 0 & 2\rho C_\mu \frac{\int_0^1 \{\varphi_h^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_h^{(1)}(s)\}^{-1} ds\right)^2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \triangleq \Phi_0. \tag{2.70}$$

Proof. Elements on the main diagonal of Φ_0 are implied by Theorems 2.2 and 2.3. Moreover, by the Cauchy-Schwarz inequality

$$\left|\operatorname{cov}\left(\bar{Z}_{v,(m,n)}, \bar{Z}_{s,(m,n)}\right)\right| \le \sqrt{\left(\operatorname{var}\left(\bar{Z}_{v,(m,n)}\right)\right)\left(\operatorname{var}\left(\bar{Z}_{s,(m,n)}\right)\right)} = O(n^{\frac{-3}{2}}).$$

So $\Phi_0(1,3) = 0$. It remains to prove $\Phi_0(1,2) = 0$. Let $f_\lambda(t) = e^{-\lambda|t|}$, then

$$f_\lambda^{(1)}(t) = \begin{cases} -\lambda e^{-\lambda t}, & \text{if } t > 0, \\ \lambda e^{\lambda t}, & \text{if } t < 0. \end{cases} \tag{2.71}$$

The key step in our proof is to show that for any $\mathbf{i}, \mathbf{j} \in \mathcal{I}_{(m,n)}$,

$$\left| \mathrm{E} \left(X_{\mathbf{i}+(0,1)} - X_{\mathbf{i}} \right) \left(X_{\mathbf{j}+(1,0)} - X_{\mathbf{j}} \right) \right| \leq \sigma^2 \lambda \mu \Delta_{h,\mathbf{j}} \Delta_{v,\mathbf{i}}. \tag{2.72}$$

Actually,

$$\begin{aligned}
& \left| \mathrm{E} \left(X_{\mathbf{i}+(0,1)} - X_{\mathbf{i}} \right) \left(X_{\mathbf{j}+(1,0)} - X_{\mathbf{j}} \right) \right| \\
=& \sigma^2 \left| e^{-\lambda|\varphi_{(h,\mathbf{i}[1])} - \varphi_{(h,\mathbf{j}[1])}| - \mu|\varphi_{(v,\mathbf{i}[2])} - \varphi_{(v,\mathbf{j}[2])}|} \right. \\
& + e^{-\lambda|\varphi_{(h,\mathbf{i}[1])} - \varphi_{(h,\mathbf{j}[1]+1)}| - \mu|\varphi_{(v,\mathbf{i}[2]+1)} - \varphi_{(v,\mathbf{j}[2])}|} \\
& - e^{-\lambda|\varphi_{(h,\mathbf{i}[1])} - \varphi_{(h,\mathbf{j}[1]+1)}| - \mu|\varphi_{(v,\mathbf{i}[2])} - \varphi_{(v,\mathbf{j}[2])}|} \\
& \left. - e^{-\lambda|\varphi_{(h,\mathbf{i}[1])} - \varphi_{(h,\mathbf{j}[1])}| - \mu|\varphi_{(v,\mathbf{i}[2]+1)} - \varphi_{(v,\mathbf{j}[2])}|} \right| \\
=& \sigma^2 \left| \left(e^{-\lambda|\varphi_{(h,\mathbf{i}[1])} - \varphi_{(h,\mathbf{j}[1])}|} - e^{-\lambda|\varphi_{(h,\mathbf{i}[1])} - \varphi_{(h,\mathbf{j}[1]+1)}|} \right) \right. \\
& \left. \left(e^{-\mu|\varphi_{(v,\mathbf{i}[2])} - \varphi_{(v,\mathbf{j}[2])}|} - e^{-\mu|\varphi_{(v,\mathbf{i}[2]+1)} - \varphi_{(v,\mathbf{j}[2])}|} \right) \right|.
\end{aligned}$$

Since

$$\left(\varphi_{(h,\mathbf{i}[1])} - \varphi_{(h,\mathbf{j}[1])} \right) \left(\varphi_{(h,\mathbf{i}[1])} - \varphi_{(h,\mathbf{j}[1]+1)} \right) \geq 0,$$

it follows from Eq. (2.71) and the mean value theorem that

$$\left| \left(e^{-\lambda|\varphi_{(h,\mathbf{i}[1])} - \varphi_{(h,\mathbf{j}[1])}|} - e^{-\lambda|\varphi_{(h,\mathbf{i}[1])} - \varphi_{(h,\mathbf{j}[1]+1)}|} \right) \right| \leq \lambda \Delta_{h,\mathbf{j}}. \tag{2.73}$$

The same method shows

$$\left| \left(e^{-\mu|\varphi_{(v,\mathbf{i}[2])} - \varphi_{(v,\mathbf{j}[2])}|} - e^{-\mu|\varphi_{(v,\mathbf{i}[2]+1)} - \varphi_{(v,\mathbf{j}[2])}|} \right) \right| \leq \mu \Delta_{v,\mathbf{i}}. \tag{2.74}$$

Therefore, Eq. (2.72) holds. Based on the definition of $\nabla_v X_{\mathbf{i}}$ in Eq. (2.37), we have that

$$\left| \mathrm{E} \left(\nabla_v X_{\mathbf{i}} \nabla_h X_{\mathbf{j}} \right) \right| = \frac{\left| \mathrm{E} \left(X_{\mathbf{i}+(0,1)} - X_{\mathbf{i}} \right) \left(X_{\mathbf{j}+(1,0)} - X_{\mathbf{j}} \right) \right|}{\Delta_{h,\mathbf{j}} \Delta_{v,\mathbf{i}}} \leq \sigma^2 \lambda \mu, \tag{2.75}$$

which implies that for any $(i,j)' \in \mathcal{I}_{(m,n)}$,

$$\begin{aligned}\operatorname{cov}(V_{v,i}, V_{h,j}) &= \sum_{k=1}^{n-1}\sum_{s=1}^{m-1} \operatorname{cov}\big((\nabla_v X_{(i,k)})^2, (\nabla_h X_{(s,j)})^2\big) \\ &= \sum_{k=1}^{n-1}\sum_{s=1}^{m-1} 2\big(\mathrm{E}\,\nabla_v X_{(i,k)} \nabla_h X_{(s,j)}\big)^2 \\ &\le 2mn(\sigma^2\lambda\mu)^2.\end{aligned} \tag{2.76}$$

Finally,

$$\begin{aligned}\Phi_N(1,2) &= \operatorname{cov}\left(\sum_{i=1}^{m-1} \Delta_{h,i} V_{v,i}/\,\mathrm{E}(V_{v,i}), \sum_{j=1}^{n-1} \Delta_{v,j} V_{h,j}/\,\mathrm{E}(V_{h,j})\right) \\ &= \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} \Delta_{h,i}\Delta_{v,j} \operatorname{cov}(V_{v,i}, V_{h,j})/\,\mathrm{E}(V_{v,i})/\,\mathrm{E}(V_{h,j}) \\ &\le 2mn(\sigma^2\lambda\mu)^2/\,\mathrm{E}(V_{v,1})/\,\mathrm{E}(V_{h,1}) \\ &= O((mn)^{-1}).\end{aligned} \tag{2.77}$$

This proves $\Phi_0(1,2) = 0$. □

Theorem 2.4.

$$\sqrt{n}(\bar{Z}_{(m,n)} - \mathbf{1}) \to \mathcal{N}(\mathbf{0}, \Phi_0). \tag{2.78}$$

Proof. The following argument imitates the method of Chan and Wood. Let $L = (m-1)(n-1)$. Fix an 3-vector $\mathbf{f} \in \mathbb{R}_+^3$ and define the $3L \times 3L$ diagonal matrix F_L by $F_L = \operatorname{diag}\{\mathbf{f}', \cdots, \mathbf{f}'\}$ for $L \ge 1$. Define the $3L$-vector $\mathbf{W}_L = (\mathbf{Y}_L'(\mathbf{j}), \mathbf{j} \in \mathcal{I}_{(m,n)})'$, where

$$\mathbf{Y}_L(\mathbf{j}) = \left(\sqrt{\Delta_{h,\mathbf{j}}}\frac{\nabla_v X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{v,1})}}, \sqrt{\Delta_{v,\mathbf{j}}}\frac{\nabla_h X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{h,1})}}, \frac{\nabla_s X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{s,(m,n)})}}\right)'.$$

By construction, we have

$$\begin{aligned}S_L &\triangleq \sqrt{n}\mathbf{f}'\big(\bar{Z}_{(m,n)} - \mathbf{1}\big) \\ &= \sqrt{n}(\mathbf{W}_L' F_L \mathbf{W}_L - \mathrm{E}(\mathbf{W}_L' F_L \mathbf{W}_L)).\end{aligned} \tag{2.79}$$

Let V_L denote the covariance matrix of $\mathbf{W}_L$. Note that each entry of V_L is of the form $\sigma_L^{ab}(\mathbf{i},\mathbf{j})$ with $a, b \in \{v, h, s\}$. For example,

$$\sigma_L^{vh}(\mathbf{i},\mathbf{j}) = \operatorname{cov}\left(\sqrt{\Delta_{h,\mathbf{i}}}\frac{\nabla_v X_{\mathbf{i}}}{\sqrt{\mathrm{E}(V_{v,1})}}, \sqrt{\Delta_{v,\mathbf{j}}}\frac{\nabla_h X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{h,1})}}\right). \tag{2.80}$$

For convenience, $\sigma_L^{ab}(\cdot)$ will be denoted as $\sigma^{ab}(\cdot)$ below. $V_L^{\frac{1}{2}}$ is defined as the symmetric positive definite square root of V_L. Denote by $\Lambda_L = \mathrm{diag}(\lambda_{1,L}, \cdots, \lambda_{3L,L})$ the diagonal matrix whose diagonal entries are eigenvalues of $2n^{\frac{1}{2}} V_L^{\frac{1}{2}} F_L V_L^{\frac{1}{2}}$. Then for $\epsilon_L \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{3L\times 3L})$, we have

$$n^{\frac{1}{2}} \mathbf{W}_L' F_L \mathbf{W}_L \overset{\mathrm{d}}{=} n^{\frac{1}{2}} \epsilon_L' V_L^{\frac{1}{2}} F_L V_L^{\frac{1}{2}} \epsilon_L \overset{\mathrm{d}}{=} \frac{1}{2} \epsilon_L' \Lambda_L \epsilon_L. \tag{2.81}$$

Therefore, for all $|\theta| < \frac{1}{\max(\lambda_{1,L},\cdots,\lambda_{3L,L})}$, the cumulant generating function of S_L is given by

$$\kappa_L(\theta) \triangleq \ln\big(\mathrm{E}(e^{\theta S_L})\big) = -\frac{1}{2} \sum_{q=1}^{3L} \{\ln(1-\theta\lambda_{q,L}) + \theta\lambda_{q,L}\}, \tag{2.82}$$

(see Khuri 2009, chap. 5). To obtain the limit of $\kappa_L(\theta)$ as $n \to \infty$, we first prove

$$\mathrm{tr}\big(\Lambda_L^4\big) = \sum_{q=1}^{3L} \lambda_{q,L}^4 \to 0, \quad \text{as } n \to \infty. \tag{2.83}$$

Direct calculation shows that

$$\begin{aligned} \mathrm{tr}\big(\Lambda_L^4\big) &= 16n^2\, \mathrm{tr}\left((V_L^{\frac{1}{2}} F_L V_L^{\frac{1}{2}})^4 \right) \\ &= 16n^2\, \mathrm{tr}\left((V_L F_L)^4 \right) \\ &= 16n^2 \sum_{u_1\in\{v,h,s\}} \cdots \sum_{u_4\in\{v,h,s\}} f_{u_1}\cdots f_{u_4} \Delta_L(u_1,\cdots,u_4), \end{aligned} \tag{2.84}$$

where

$$\begin{aligned} &\Delta_L(u_1,\cdots,u_4) \\ = &\sum_{\mathbf{i}_1\in\mathcal{I}_{(m,n)}} \cdots \sum_{\mathbf{i}_4\in\mathcal{I}_{(m,n)}} \sigma^{u_1u_2}(\mathbf{i}_1,\mathbf{i}_2)\sigma^{u_2u_3}(\mathbf{i}_2,\mathbf{i}_3)\sigma^{u_3u_4}(\mathbf{i}_3,\mathbf{i}_4)\sigma^{u_4u_1}(\mathbf{i}_1,\mathbf{i}_4). \end{aligned} \tag{2.85}$$

For $a, b \in \{v,h,s\}$ and $\mathbf{i},\mathbf{j} \in \mathcal{I}_{(m,n)}$, if we are able to find an upper bound $\hat{\sigma}^{ab}(\mathbf{i}-\mathbf{j})$ for each $\big|\sigma^{ab}(\mathbf{i},\mathbf{j})\big|$, the proof can be completed by imitating the stationary case; (see Chan and Wood 2000, (7.15)).

- For the vertical case, we have

$$
\begin{aligned}
&|\sigma^{vv}(\mathbf{i},\mathbf{j})| \\
=&\left|\operatorname{cov}\left(\sqrt{\Delta_{h,\mathbf{i}}}\frac{\nabla_v X_{\mathbf{i}}}{\sqrt{\mathrm{E}(V_{v,1})}}, \sqrt{\Delta_{h,\mathbf{j}}}\frac{\nabla_v X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{v,1})}}\right)\right| \\
=&\frac{\sqrt{\Delta_{h,\mathbf{i}}\Delta_{h,\mathbf{j}}}}{\mathrm{E}(V_{v,1})}\left|\operatorname{cov}\left(\nabla_v X_{\mathbf{i}}, \nabla_v X_{\mathbf{j}}\right)\right| \\
\leq&\frac{C_{h,1}}{\mathrm{E}(V_{v,1})m}e^{-\lambda\left|\varphi_{(h,\mathbf{i}[1])}-\varphi_{(h,\mathbf{j}[1])}\right|}\left|\sum_{s,k=0}^{1} a_{v;\mathbf{i},k}a_{v;\mathbf{j},s}\Gamma_\mu\big(\varphi_{(v,(\mathbf{i}[2]+k))}-\varphi_{(v,(\mathbf{j}[2]+s))}\big)\right| \\
\leq&\frac{C_{h,1}}{\mathrm{E}(V_{v,1})m}\left|\sum_{s,k=0}^{1} a_{v;\mathbf{i},k}a_{v;\mathbf{j},s}\Gamma_\mu\big(\varphi_{(v,(\mathbf{i}[2]+k))}-\varphi_{(v,(\mathbf{j}[2]+s))}\big)\right|.
\end{aligned}
$$

  Then imitating the proof of Lemma 2.1, we get that there exist some constants $C_1, C_2 > 0$ such that

$$
|\sigma^{vv}(\mathbf{i},\mathbf{j})| \leq \begin{cases} \frac{C_1}{mn}, & \text{if } \mathbf{i}[2]-\mathbf{j}[2]=0 \ , \\ \frac{C_2}{mn^2}, & \text{if } \mathbf{i}[2]-\mathbf{j}[2]\neq 0 \ , \end{cases} \tag{2.86}
$$

  uniformly over $\mathbf{i},\mathbf{j} \in \mathcal{I}_{(m,n)}$ when n is big enough. Define

$$
\hat{\sigma}^{vv}(\mathbf{x}) = \frac{C_1}{mn}\mathbf{1}_0(\mathbf{x}[2]) + \frac{C_2}{mn^2}\mathbf{1}_{\mathbb{R}\backslash 0}(\mathbf{x}[2]),
$$

  then

$$
|\sigma^{vv}(\mathbf{i},\mathbf{j})| \leq \hat{\sigma}^{vv}(\mathbf{i}-\mathbf{j}),
$$

  when n is big enough.

- For the square case, noticing Eq. (2.56), we have that there exist some constants

$$
C_3, C_4, C_5, C_6 > 0,
$$

  such that

$$
\begin{aligned}
&|\sigma^{ss}(\mathbf{i},\mathbf{j})| \\
=&\left|\mathrm{E}\left(\nabla_s X_{\mathbf{i}}\nabla_s X_{\mathbf{j}}\right)\right| / \,\mathrm{E}(V_{s,(m,n)}) \\
\leq&\frac{1}{(mn)^2}\big(C_3 m\mathbf{1}_0((\mathbf{i}-\mathbf{j})[1]) + C_4\mathbf{1}_{\mathbb{R}\backslash 0}((\mathbf{i}-\mathbf{j})[1])\big) \\
&\times\big(C_5 n\mathbf{1}_0((\mathbf{i}-\mathbf{j})[2]) + C_6\mathbf{1}_{\mathbb{R}\backslash 0}((\mathbf{i}-\mathbf{j})[2])\big),
\end{aligned} \tag{2.87}
$$

uniformly over $\mathbf{i},\mathbf{j}\in\mathcal{I}_{(m,n)}$ when n is big enough. So we define $\hat{\sigma}^{ss}(\cdot)$ as

$$\hat{\sigma}^{ss}(\mathbf{x}) = \frac{1}{(mn)^2}\big(C_3 m \mathbf{1}_0(\mathbf{x}[1]) + C_4\mathbf{1}_{\mathbb{R}\setminus 0}(\mathbf{x}[1])\big) \times \big(C_5 n\mathbf{1}_0(\mathbf{x}[2]) + C_6\mathbf{1}_{\mathbb{R}\setminus 0}(\mathbf{x}[2])\big).$$

- For the cross term between the vertical and the horizontal,

$$\begin{aligned}
&\left|\sigma^{vh}(\mathbf{i},\mathbf{j})\right| \\
=&\left|\operatorname{cov}\left(\sqrt{\Delta_{h,\mathbf{i}}}\frac{\nabla_v X_\mathbf{i}}{\sqrt{\mathrm{E}(V_{v,1})}}, \sqrt{\Delta_{v,\mathbf{j}}}\frac{\nabla_h X_\mathbf{j}}{\sqrt{\mathrm{E}(V_{h,1})}}\right)\right| \\
=&\frac{\sqrt{\Delta_{h,\mathbf{i}}}\sqrt{\Delta_{v,\mathbf{j}}}}{\sqrt{\mathrm{E}(V_{v,1})}\sqrt{\mathrm{E}(V_{h,1})}}\left|\mathrm{E}\left(\nabla_v X_\mathbf{i}\nabla_h X_\mathbf{j}\right)\right|.
\end{aligned} \tag{2.88}$$

Then by Eq. (2.75), there exists some constant C_7 such that

$$\left|\sigma^{vh}(\mathbf{i},\mathbf{j})\right| \le \frac{C_7}{(mn)^{\frac{3}{2}}},$$

uniformly over $\mathbf{i},\mathbf{j}\in\mathcal{I}_{(m,n)}$ when n is big enough. In this case,

$$\hat{\sigma}^{vh}(\mathbf{x}) = \frac{C_7}{(mn)^{\frac{3}{2}}}$$

is a constant function with respect to $\mathbf{x}$.

- Lastly, we study the cross term between the vertical and the square,

$$\begin{aligned}
&|\sigma^{vs}(\mathbf{i},\mathbf{j})| \\
=&\left|\operatorname{cov}\left(\sqrt{\Delta_{h,\mathbf{i}}}\frac{\nabla_v X_\mathbf{i}}{\sqrt{\mathrm{E}(V_{v,1})}}, \frac{\nabla_s X_\mathbf{j}}{\sqrt{\mathrm{E}(V_{s,(m,n)})}}\right)\right| \\
=&\frac{\sqrt{\Delta_{h,\mathbf{i}}}}{\sqrt{\mathrm{E}(V_{v,1})\,\mathrm{E}(V_{s,(m,n)})}}\left|\operatorname{cov}\left(\nabla_v X_\mathbf{i}, (\nabla_v X_{\mathbf{j}+(1,0)} - \nabla_v X_\mathbf{j})/\Delta_{h,\mathbf{j}}\right)\right| \\
=&\frac{\sqrt{\Delta_{h,\mathbf{i}}}}{\sqrt{\mathrm{E}(V_{v,1})\,\mathrm{E}(V_{s,(m,n)})}}\left|e^{-\lambda\left|\varphi_{(h,\mathbf{i}[1])}-\varphi_{(h,\mathbf{j}[1]+1)}\right|} - e^{-\lambda\left|\varphi_{(h,\mathbf{i}[1])}-\varphi_{(h,\mathbf{j}[1])}\right|}\right|/\Delta_{h,\mathbf{j}} \\
&\times\left|\sum_{s,k=0}^{1} a_{v;\mathbf{i},k}a_{v;\mathbf{j},s}\Gamma_\mu\big(\varphi_{(v,(\mathbf{i}[2]+k))} - \varphi_{(v,(\mathbf{j}[2]+s))}\big)\right| \\
\le&\lambda\frac{\sqrt{\Delta_{h,\mathbf{i}}}}{\sqrt{\mathrm{E}(V_{v,1})\,\mathrm{E}(V_{s,(m,n)})}}\left|\sum_{s,k=0}^{1} a_{v;\mathbf{i},k}a_{v;\mathbf{j},s}\Gamma_\mu\big(\varphi_{(v,(\mathbf{i}[2]+k))} - \varphi_{(v,(\mathbf{j}[2]+s))}\big)\right|,
\end{aligned} \tag{2.89}$$

where the last inequality is based on Eq. (2.73). Therefore, we conclude that there exist some constants $C_8, C_9 > 0$ such that

$$|\sigma^{vs}(\mathbf{i},\mathbf{j})| \leq \begin{cases} \frac{C_8}{m^{1.5}n}, & \text{if } \mathbf{i}[2]-\mathbf{j}[2]=0 \ , \\ \frac{C_9}{m^{1.5}n^2}, & \text{if } \mathbf{i}[2]-\mathbf{j}[2]\neq 0 \ , \end{cases} \tag{2.90}$$

uniformly over $\mathbf{i},\mathbf{j} \in \mathcal{I}_{(m,n)}$ when n is big enough. Define

$$\hat{\sigma}^{vs}(\mathbf{x}) = \frac{C_8}{m^{1.5}n}\mathbf{1}_0(\mathbf{x}[2]) + \frac{C_9}{m^{1.5}n^2}\mathbf{1}_{\mathbb{R}\setminus 0}(\mathbf{x}[2]),$$

then

$$|\sigma^{vs}(\mathbf{i},\mathbf{j})| \leq \hat{\sigma}^{vs}(\mathbf{i}-\mathbf{j}),$$

when n is big enough.

So combining all the cases above, we have

$$\begin{aligned} &|\Delta_L(u_1,\cdots,u_4)| \\ \leq &\sum_{\mathbf{i}_1\in\mathcal{I}_{(m,n)}}\cdots\sum_{\mathbf{i}_4\in\mathcal{I}_{(m,n)}} |\sigma^{u_1u_2}(\mathbf{i}_1,\mathbf{i}_2)\sigma^{u_2u_3}(\mathbf{i}_2,\mathbf{i}_3)\sigma^{u_3u_4}(\mathbf{i}_3,\mathbf{i}_4)\sigma^{u_4u_1}(\mathbf{i}_1,\mathbf{i}_4)| \\ \leq &\sum_{\mathbf{i}_1\in\mathcal{I}_{(m,n)}}\cdots\sum_{\mathbf{i}_4\in\mathcal{I}_{(m,n)}} \hat{\sigma}^{u_1u_2}(\mathbf{i}_1-\mathbf{i}_2)\hat{\sigma}^{u_2u_3}(\mathbf{i}_2-\mathbf{i}_3)\hat{\sigma}^{u_3u_4}(\mathbf{i}_3-\mathbf{i}_4)\hat{\sigma}^{u_4u_1}(\mathbf{i}_1-\mathbf{i}_4), \end{aligned} \tag{2.91}$$

when n is big enough.

Define the index set

$$\mathcal{D}_{(m,n)} \triangleq \{\mathbf{i}-\mathbf{j} : \mathbf{i},\mathbf{j}\in\mathcal{I}_{(m,n)}\}. \tag{2.92}$$

For each triple ($\mathbf{h}_1$, $\mathbf{h}_2$, $\mathbf{h}_3$) which satisfies $\mathbf{h}_a \in \mathcal{D}_{(m,n)}, 1 \leq a \leq 3$, the cardinality of the set

$$\{(\mathbf{i}_1,\mathbf{i}_2,\mathbf{i}_3,\mathbf{i}_4) : \mathbf{i}_a \in \mathcal{I}_n, a=1,\cdots,4; \mathbf{h}_a = \mathbf{i}_a - \mathbf{i}_{a+1}, 1\leq a\leq 3\}$$

is bounded by L. It follows that when n is big enough

$$\begin{aligned} &|\Delta_L(u_1,\cdots,u_4)| \\ \leq &L\sum_{\mathbf{h}_1\in\mathcal{D}_{(m,n)}}\cdots\sum_{\mathbf{h}_3\in\mathcal{D}_{(m,n)}} \hat{\sigma}^{u_1u_2}(\mathbf{h}_1)\hat{\sigma}^{u_2u_3}(\mathbf{h}_2)\hat{\sigma}^{u_3u_4}(\mathbf{h}_3)\hat{\sigma}^{u_4u_1}(\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3) \\ \leq &C_{10}\prod_{a=1}^{3}\left(\sum_{\mathbf{h}_a\in\mathcal{D}_{(m,n)}}\hat{\sigma}^{u_au_{a+1}}(\mathbf{h})\right). \end{aligned} \tag{2.93}$$

The last inequality holds since there exists some constant $C_{10} > 0$ such that

$$\hat{\sigma}^{ab}(h) \le \frac{C_{10}}{mn},$$

for all possible $a, b \in \{v, h, s\}$ and $\mathbf{h} \in \mathcal{D}_{(m,n)}$ when n is big enough.

Direct calculation shows that

$$\sum_{\mathbf{h}\in\mathcal{D}_{(m,n)}} \hat{\sigma}^{vv}(\mathbf{h}) = O(\frac{1}{n}), \qquad \sum_{\mathbf{h}\in\mathcal{D}_{(m,n)}} \hat{\sigma}^{ss}(\mathbf{h}) = O(\frac{1}{mn}),$$

$$\sum_{\mathbf{h}\in\mathcal{D}_{(m,n)}} \hat{\sigma}^{vh}(\mathbf{h}) = O(\frac{1}{\sqrt{mn}}), \qquad \sum_{\mathbf{h}\in\mathcal{D}_{(m,n)}} \hat{\sigma}^{vs}(\mathbf{h}) = O(\frac{1}{\sqrt{mn}}).$$

Combining Eqs. (2.84) and (2.93), we conclude that

$$\operatorname{tr}\left(\Lambda_L^4\right) = n^2 O(n^{-3}) = O(n^{-1}), \tag{2.94}$$

as $n \to \infty$. Namely, Eq. (2.83) holds. Meanwhile, note that Eq. (2.83) implies that as $n \to \infty$,

$$\max_{1\le q\le 3L}\{\lambda_{q,L}\} \le \left(\sum_{1\le q\le 3L} \lambda_{q,L}^4\right)^{\frac{1}{4}} \to 0. \tag{2.95}$$

Expanding Eq. (2.82) about $\theta = 0$ using Taylor's theorem, we obtain

$$\kappa_L(\theta) = \frac{1}{2}\sum_{q=1}^{3L}\left\{\frac{1}{2}(\theta\lambda_{q,L})^2 + \frac{1}{3}(\theta\lambda_{q,L})^3 + \frac{1}{4}(\theta\lambda_{q,L})^4(1-\theta_{q,L}^*\lambda_{q,L})^{-4}\right\}, \tag{2.96}$$

for some $\theta_{q,L}^*$ which satisfies $0 \le \left|\theta_{q,L}^*\right| \le |\theta|$.

Let us first consider the term $\sum_{q=1}^{3L} \frac{1}{2}(\theta\lambda_{q,L})^2$,

$$\begin{aligned}\sum_{q=1}^{3L} \frac{1}{2}\lambda_{q,L}^2 &= \frac{1}{2}\operatorname{tr}\left(\Lambda_L^2\right) = 2n \operatorname{tr}\left((V_L F_L)^2\right) \\ &= \operatorname{var}(\sqrt{n} W_L' F_L W_L) = n\mathbf{f}'\Phi_{(m,n)}\mathbf{f}.\end{aligned} \tag{2.97}$$

It follows from Lemma 2.3 that, as $n \to \infty$,

$$\sum_{q=1}^{3L} \frac{1}{2}\lambda_{q,L}^2 = n\mathbf{f}'\Phi_{(m,n)}\mathbf{f} \to \mathbf{f}'\Phi_0\mathbf{f}. \tag{2.98}$$

Second, Eq. (2.83) implies that, as $n \to \infty$,

$$\left| \sum_{q=1}^{3L} \lambda_{q,L}^3 \right| \leq \max_{1 \leq q \leq 3L} \{\lambda_{q,L}\} \sum_{q=1}^{3L} \lambda_{q,L}^2 \to 0. \tag{2.99}$$

Third, note that $\delta = \sup_{L \geq 1} \max_{1 \leq q \leq 3L} \{\lambda_{q,L}\}$ is positive and finite. If we restrict attention to $|\theta| \leq \frac{1}{2\delta}$, then

$$(1 - \theta_{q,L}^* \lambda_{q,L}) > 1/2, \tag{2.100}$$

for all q and L. Hence, combining Eq. (2.98), Eq. (2.99) and Eq. (2.100), we have that for $\theta \in (-\frac{1}{2\delta}, \frac{1}{2\delta})$,

$$\kappa_L(\theta) \to \theta^2 \mathbf{f}' \Phi_0 \mathbf{f}/2, \quad \text{as } n \to \infty, \tag{2.101}$$

which leads to the conclusion that $S_L \to \mathcal{N}(\mathbf{0}, \mathbf{f}'\Phi_0\mathbf{f})$ in distribution. This completes the proof. □

Let

$$R_v \triangleq \mathrm{E}(V_{v,1}) \Big/ \left(2n^2 \int_0^1 \{\varphi_v^{(1)}(s)\}^{-1} ds \right), \tag{2.102}$$

$$R_h \triangleq \mathrm{E}(V_{h,1}) \Big/ \left(2m^2 \int_0^1 \{\varphi_h^{(1)}(s)\}^{-1} ds \right), \tag{2.103}$$

$$R_s \triangleq \mathrm{E}(V_{s,(m,n)}) \Big/ \left(4(mn)^2 \int_0^1 \{\varphi_h^{(1)}(s)\}^{-1} ds \int_0^1 \{\varphi_v^{(1)}(s)\}^{-1} ds \right). \tag{2.104}$$

Based on Eqs. (2.18), (2.20) and (2.57), we have that

$$\mathbf{r}_n \triangleq (R_v, R_h, R_s) = \big(\sigma^2\mu + O(n^{-1}), \sigma^2\lambda + O(n^{-1}), \sigma^2\lambda\mu + O(n^{-1})\big), \tag{2.105}$$

as $n \to \infty$. Define the estimator of

$$\mathbf{p} \triangleq \big(\sigma^2\mu, \sigma^2\lambda, \sigma^2\lambda\mu\big)'$$

as

$$\hat{Z}_{(m,n)} = (\bar{Z}_{v,(m,n)} R_v, \bar{Z}_{h,(m,n)} R_h, \bar{Z}_{s,(m,n)} R_s)',$$

and

$$\mathbf{P} \triangleq \begin{pmatrix} \sigma^2\mu & & \\ & \sigma^2\lambda & \\ & & \sigma^2\lambda\mu \end{pmatrix}.$$

Consequently, the estimator of (λ, μ, σ^2) is defined as

$$\begin{aligned} \hat{\lambda} &= \hat{Z}_{(m,n)}[3]/\hat{Z}_{(m,n)}[1], \\ \hat{\mu} &= \hat{Z}_{(m,n)}[3]/\hat{Z}_{(m,n)}[2], \\ \hat{\sigma}^2 &= \hat{Z}_{(m,n)}[1]\hat{Z}_{(m,n)}[2]/\hat{Z}_{(m,n)}[3]. \end{aligned}$$

We end up with the corollary below:

Corollary 2.1. *Suppose the two sampling functions $\varphi_h(\cdot)$ and $\varphi_v(\cdot)$ satisfy Condition 1 and $\lim_{m\to\infty} \frac{n}{m} = \rho$, then*

$$\sqrt{n}\left(\begin{pmatrix} \hat{\lambda} \\ \hat{\mu} \\ \hat{\sigma}^2 \end{pmatrix} - \begin{pmatrix} \lambda \\ \mu \\ \sigma^2 \end{pmatrix}\right) \xrightarrow[n\to\infty]{d} \mathcal{N}(\mathbf{0}, \Sigma), \tag{2.106}$$

where Σ is defined as

$$\begin{pmatrix} 2C_\lambda \frac{\int_0^1 \{\varphi_v^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_v^{(1)}(s)\}^{-1} ds\right)^2} \lambda^2 & 0 & -2C_\lambda \frac{\int_0^1 \{\varphi_v^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_v^{(1)}(s)\}^{-1} ds\right)^2} \sigma^2\lambda \\ & 2\rho C_\mu \frac{\int_0^1 \{\varphi_h^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_h^{(1)}(s)\}^{-1} ds\right)^2} \mu^2 & -2\rho C_\mu \frac{\int_0^1 \{\varphi_h^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_h^{(1)}(s)\}^{-1} ds\right)^2} \sigma^2\mu \\ & & 2\sigma^4 \left(\rho C_\mu \frac{\int_0^1 \{\varphi_h^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_h^{(1)}(s)\}^{-1} ds\right)^2} + C_\lambda \frac{\int_0^1 \{\varphi_v^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_v^{(1)}(s)\}^{-1} ds\right)^2}\right) \end{pmatrix},$$

and

$$\begin{pmatrix} \hat{\lambda} \\ \hat{\mu} \\ \hat{\sigma}^2 \end{pmatrix} \xrightarrow[n\to\infty]{a.s.} \begin{pmatrix} \lambda \\ \mu \\ \sigma^2 \end{pmatrix}. \tag{2.107}$$

Proof. Let

$$\mathbf{R}_n \triangleq \begin{pmatrix} R_v & & \\ & R_h & \\ & & R_s \end{pmatrix},$$

then we have

$$\sqrt{n}\Big(\hat{Z}_{(m,n)} - \mathbf{p}\Big) = \sqrt{n}\big[\mathbf{R}_n\big(\bar{Z}_{(m,n)} - \mathbf{1}\big) + (\mathbf{r}_n - \mathbf{p})\big]. \tag{2.108}$$

Since

$$\mathbf{R}_n \xrightarrow{n\to\infty} \mathbf{P},$$

we have

$$\sqrt{n}\mathbf{R}_n\big(\bar{Z}_{(m,n)} - \mathbf{1}\big) \xrightarrow{\mathrm{d}} \mathcal{N}(\mathbf{0}, \mathbf{P}\Phi_0\mathbf{P}'),$$

based on Theorem 2.4. Moreover, as $n \to \infty$,

$$\sqrt{n}(\mathbf{r}_n - \mathbf{p}) = \Big(O(n^{-\frac{1}{2}}), O(n^{-\frac{1}{2}}), O(n^{-\frac{1}{2}})\Big)'.$$

Therefore, we have

$$\sqrt{n}\Big(\hat{Z}_{(m,n)} - \mathbf{p}\Big) \xrightarrow{\mathrm{d}} \mathcal{N}(\mathbf{0}, \mathbf{P}\Phi_0\mathbf{P}'). \tag{2.109}$$

Applying the delta method, we get Eq. (2.106).

The almost sure convergence is the direct consequence of Theorems 2.2 and 2.3. □

2.4 Simulations

Consider the 2-dimensional OU field with $\lambda = 0.5, \mu = 10, \sigma^2 = 4, \frac{n}{m} = 2$ and two sampling functions as

$$\varphi_h(x) = \frac{100}{102}\left(\left(x + \frac{1}{100}\right)^2 - \frac{1}{100^2}\right),$$
$$\varphi_v(x) = \frac{20}{22}\left(\left(x + \frac{1}{20}\right)^2 - \frac{1}{20^2}\right).$$

With $m = 100, 150, 200, \cdots, 500$, we run 1000 realizations for each case and estimate λ, μ and σ^2.

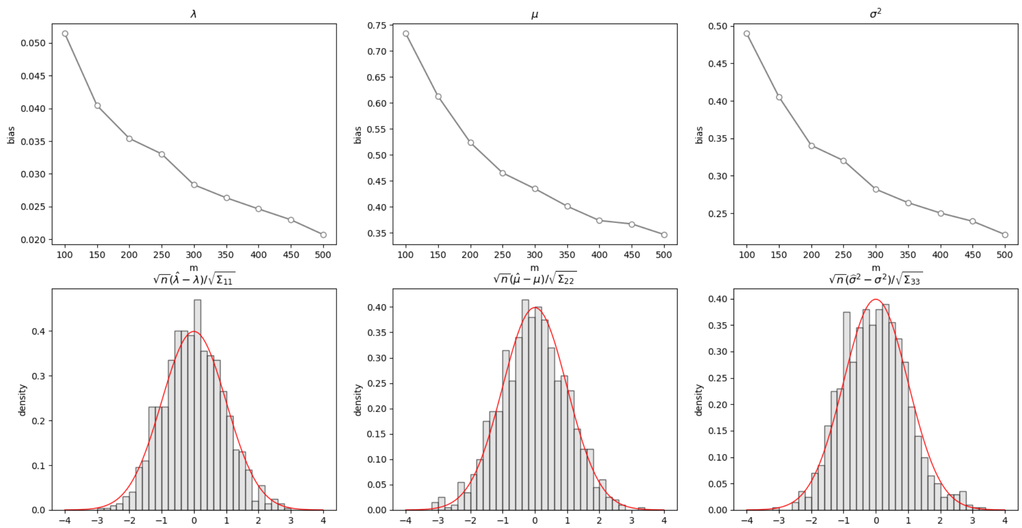

Figure 2.1 Simulated estimation for the 2-dimensional OU field . Plots in the first row present the averaged absolute value of bias for each sample size and each parameter; plots in the second row present the distribution of normalized bias when $n = 1000$, where the red curve is the density of $N(0, 1)$.

2.5 Extensions to higher-dimensional spaces

In this section, we generalize the results for dimension $d = 2$ to higher-dimensional spatial processes. Let $X(\mathbf{u}), \mathbf{u} \in [0, 1]^d$, denote a spatial Gaussian process with the covariance function Eq. (2.4). Like the case $d = 2$, for $\ell = 1, \cdots, d$, φ_ℓ satisfies Condition 1. We consider the $N = \prod_{\ell=1}^{d} n_\ell$ observation locations over $\mathbb{R}^d$ which are of the form

$$\mathbf{u}(\mathbf{i}) \triangleq \left(\varphi_1\left(\frac{\mathbf{i}[1]-1}{n_1-1}\right), \varphi_2\left(\frac{\mathbf{i}[2]-1}{n_2-1}\right), \cdots, \varphi_d\left(\frac{\mathbf{i}[d]-1}{n_d-1}\right)\right)',$$

where $1 \leq \mathbf{i}[\ell] \leq n_\ell$ and $N^{1/d}/n_\ell \to \rho_\ell \in (0, \infty), \ell = 1, \cdots, d$, and we write $X_\mathbf{i} = X(\mathbf{u}(\mathbf{i}))$. To simplify the notation, let

$$\varphi_j^{(\ell)} = \varphi_\ell\left(\frac{j-1}{n_\ell-1}\right),$$

for $j = 1, \cdots, n_\ell$ and $\ell = 1, \cdots, d$. Furthermore, define

$$\Delta_\mathbf{k}^{(\ell)} = \varphi_{\mathbf{k}[\ell]+1}^{(\ell)} - \varphi_{\mathbf{k}[\ell]}^{(\ell)},$$

for $\mathbf{k} \in \mathcal{I}_N \triangleq \{\mathbf{i} : \mathbf{1} \leq \mathbf{i} \leq ((n_1 - 1), \cdots, (n_d - 1))'\}$ and $\ell = 1, \cdots, d$. If $k = 1, \cdots, (n_\ell - 1)$ is a scalar, then

$$\Delta_k^{(\ell)} = \varphi_{k+1}^{(\ell)} - \varphi_k^{(\ell)}.$$

To estimate $\sigma^2 \prod_{\ell=1}^d \theta_\ell$, we first introduce the increment $\{b_\mathbf{k}\}_{\mathbf{k}=\mathbf{0}}^{\mathbf{1}}$ which is defined as

$$b_\mathbf{k} = (-1)^{(d-|\mathbf{k}|)}. \tag{2.110}$$

Accordingly, define

$$\nabla_s X_\mathbf{i} = \frac{\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{1}} b_\mathbf{k} X_{\mathbf{i}+\mathbf{k}}}{\prod_{\ell=1}^d \Delta_\mathbf{i}^{(\ell)}}, \tag{2.111}$$

$$V_{s,N} = \sum_{\mathbf{i}\in\mathcal{I}_N} (\nabla_s X_\mathbf{i})^2, \tag{2.112}$$

$$\bar{Z}_{s,N} = V_{s,N} / \operatorname{E}(V_{s,N}). \tag{2.113}$$

Similarly to Eq. (2.57), it is not hard to see

$$\operatorname{E}(V_{s,N}) = \sigma^2 \prod_{\ell=1}^d \left(2\theta_\ell n_\ell^2 \int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds \right) + O(N^{2-1/d}). \tag{2.114}$$

As for $\sigma^2 \prod_{\ell\neq j} \theta_\ell, j = 1, \cdots, d$, we first define $\{b_\mathbf{k}^{(j)}\}$ as follows:

$$b_\mathbf{k}^{(j)} = (-1)^{(d-1-|\mathbf{k}|)}, \tag{2.115}$$

where $\mathbf{k} = \mathbf{0}, \cdots, (\mathbf{1} - \mathbf{e}_j)$ and $\{\mathbf{e}_j\}$ is the standard basis of $\mathbb{R}^d$; moreover, let

$$\nabla_j X_\mathbf{i} = \frac{\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{1}-\mathbf{e}_j} b_\mathbf{k}^{(j)} X_{\mathbf{i}+\mathbf{k}}}{\prod_{\ell\neq j} \Delta_\mathbf{i}^{(\ell)}}, \tag{2.116}$$

$$V_{j,m} = \sum_{\mathbf{i}\in\mathcal{I}_{j,m,N}} (\nabla_j X_\mathbf{i})^2, \tag{2.117}$$

$$\bar{Z}_{j,N} = \sum_{m=1}^{n_j-1} \Delta_m^{(j)} V_{j,m} / \operatorname{E}(V_{j,m}) = \sum_{m=1}^{n_j-1} \Delta_m^{(j)} V_{j,m} / \operatorname{E}(V_{j,1}), \tag{2.118}$$

where $\mathcal{I}_{j,m,N} \triangleq \{\mathbf{i} \in \mathcal{I}_N : \mathbf{i}[j] = m\}$. Consequently,

$$\operatorname{E}(V_{j,m}) = \sigma^2 \prod_{\ell\neq j} \left(2\theta_\ell n_\ell^2 \int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds \right) + O(N^{2-3/d}). \tag{2.119}$$

Finally, define the $(d+1)$-vector $\bar{Z}_N \triangleq (\bar{Z}_{1,N}, \cdots, \bar{Z}_{d,N}, \bar{Z}_{s,N})'$, and $\Phi_N \triangleq \operatorname{cov}\left(\bar{Z}_N\right)$. We study the limiting covariance of $\bar{Z}_N$ in the lemma below.

Lemma 2.4.

$$N\,\mathrm{var}\left(\bar{Z}_{s,N}\right) \to 2\prod_{\ell=1}^{d} \frac{\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2}ds}{\left(\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1}ds\right)^2}, \tag{2.120}$$

and

$$N^{1-1/d}\Phi_N \to \begin{pmatrix} \Sigma_d & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix} \triangleq \Phi_0, \tag{2.121}$$

where Σ_d is a diagonal matrix with the j-th diagonal element as

$$2C_{\theta_j}\prod_{\ell\neq j}\rho_\ell \frac{\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2}ds}{\left(\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1}ds\right)^2}. \tag{2.122}$$

Proof. We follow the proof of Theorem 2.3 to establish the former part of Lemma 2.4. To begin with,

$$\begin{aligned}\mathrm{var}\left(V_{s,N}\right) &= \mathrm{var}\left(\sum_{\mathbf{i}\in\mathcal{I}_N}(\nabla_s X_{\mathbf{i}})^2\right) \\ &= \sum_{\mathbf{i}\in\mathcal{I}_N}\sum_{\mathbf{j}\in\mathcal{I}_N} 2[\mathrm{E}\left(\nabla_s X_{\mathbf{i}}\nabla_s X_{\mathbf{j}}\right)]^2.\end{aligned} \tag{2.123}$$

Imitating Eq. (2.62), we claim that

$$\begin{aligned}P_N &\triangleq \sum_{\mathbf{i}\in\mathcal{I}_N} 2\big\{\mathrm{E}\left[(\nabla_s X_{\mathbf{i}})^2\right]\big\}^2 \\ &= 2\sigma^4\prod_{\ell=1}^{d}\left[(2\theta_\ell n_\ell)^2 n_\ell \int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2}ds + O(n_\ell^2)\right].\end{aligned} \tag{2.124}$$

As for the cross terms, without loss of generality, we assume $\mathbf{i}$ and $\mathbf{j}$ are the same only in terms of the first k coordinates where $k = 0, \cdots, (d-1)$. Imitating Eq. (2.63), we have

$$\begin{aligned}Q_{N,k} &\triangleq \sum_{\mathbf{i},\mathbf{j}} 2\{\mathrm{E}\left[\nabla_s X_{\mathbf{i}}\nabla_s X_{\mathbf{j}}\right]\}^2 \\ &= O\left(2\sigma^4\prod_{\ell=1}^{k}\left\{(2\theta_\ell n_\ell)^2 n_\ell\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2}ds\right\}\prod_{j=k+1}^{d} n_j^2\right).\end{aligned} \tag{2.125}$$

By Eqs. (2.124) and (2.125), we see that

$$\mathrm{var}\left(V_{s,N}\right) = 2\sigma^4\prod_{\ell=1}^{d}\left[(2\theta_\ell n_\ell)^2 n_\ell\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2}ds\right] + O(N^{3-1/d}).$$

Combining Eq. (2.114) with the above equation, we establish Eq. (2.120).

Next, we deal with Σ_d. Based on Theorem 2.2, we claim that

$$\prod_{\ell\neq j} n_\ell \operatorname{var}\left(\bar{Z}_{j,N}\right) \to 2C_{\theta_j} \prod_{\ell\neq j} \frac{\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds\right)^2}, \tag{2.126}$$

where C_{θ_j} is a coefficient defined in Eq. (2.49). Moreover, since $N^{1/d}/n_\ell \to \rho_\ell$, the diagonal elements of Σ_d have the form Eq. (2.122). In order to show that the off-diagonal elements of Σ_d are 0, for example, $(\Sigma_d)_{1,2} = 0$, we define $\{b_{\mathbf{k}}^{(12)}\}$ as follows:

$$b_{\mathbf{k}}^{(12)} = (-1)^{(d-2-|\mathbf{k}|)}, \tag{2.127}$$

where $\mathbf{k} = \mathbf{0}, \cdots, (\mathbf{1} - \mathbf{e}_1 - \mathbf{e}_2)$. Correspondingly, define

$$\nabla_{12} X_{\mathbf{i}} = \frac{\sum_{\mathbf{k}} b_{\mathbf{k}}^{(12)} X_{\mathbf{i}+\mathbf{k}}}{\prod_{\ell\notin\{1,2\}} \Delta_{\mathbf{i}}^{(\ell)}}. \tag{2.128}$$

Similar to the derivation of Eq. (2.56), we claim that

$$\begin{aligned} &\mathrm{E}\left(\nabla_{12} X_{\mathbf{i}} \nabla_{12} X_{\mathbf{j}}\right) \\ =&\sigma^2 e^{-\theta_1\left|\varphi_{\mathbf{i}[1]}^{(1)} - \varphi_{\mathbf{j}[1]}^{(1)}\right| - \theta_2\left|\varphi_{\mathbf{i}[2]}^{(1)} - \varphi_{\mathbf{j}[2]}^{(1)}\right|} \prod_{\ell=3}^{d} \left(\sum_{l,k=0}^{1} a_{\mathbf{i},k}^{(\ell)} a_{\mathbf{j},l}^{(\ell)} e^{-\theta_\ell\left|\varphi_{(\mathbf{i}[\ell]+k)}^{(\ell)} - \varphi_{(\mathbf{j}[\ell]+l)}^{(\ell)}\right|}\right), \end{aligned} \tag{2.129}$$

where $a_{\mathbf{i},0}^{(\ell)} = -1/\Delta_{\mathbf{i}}^{(\ell)}$ and $a_{\mathbf{i},1}^{(\ell)} = 1/\Delta_{\mathbf{i}}^{(\ell)}$. Meanwhile, notice that

$$\nabla_1 X_{\mathbf{i}} = (\nabla_{12} X_{\mathbf{i}+e_2} - \nabla_{12} X_{\mathbf{i}})/\Delta_{\mathbf{i}}^{(2)},$$
$$\nabla_2 X_{\mathbf{j}} = (\nabla_{12} X_{\mathbf{j}+e_1} - \nabla_{12} X_{\mathbf{j}})/\Delta_{\mathbf{j}}^{(1)},$$

same as the deduction of Eq. (2.75), we get

$$\left|\mathrm{E}\left(\nabla_1 X_{\mathbf{i}} \nabla_2 X_{\mathbf{j}}\right)\right| \leq \sigma^2 \theta_1 \theta_2 \prod_{\ell=3}^{d} \left|\sum_{l,k=0}^{1} a_{\mathbf{i},k}^{(\ell)} a_{\mathbf{j},l}^{(\ell)} e^{-\theta_\ell\left|\varphi_{(\mathbf{i}[\ell]+k)}^{(\ell)} - \varphi_{(\mathbf{j}[\ell]+l)}^{(\ell)}\right|}\right|. \tag{2.130}$$

And

$$\begin{aligned}
&\operatorname{cov}\left(V_{1,m}, V_{2,n}\right) \\
=&\sum_{\mathbf{i}\in\mathcal{I}_{1,m,N}}\sum_{\mathbf{j}\in\mathcal{I}_{2,n,N}} 2|\mathrm{E}\left(\nabla_1 X_{\mathbf{i}}\nabla_2 X_{\mathbf{j}}\right)|^2 \\
\leq&\sum_{\mathbf{i}\in\mathcal{I}_{1,m,N}}\sum_{\mathbf{j}\in\mathcal{I}_{2,n,N}} 2\big(\sigma^2\theta_1\theta_2\big)^2\prod_{\ell=3}^{d}\left(\sum_{l,k=0}^{1} a_{\mathbf{i},k}^{(\ell)} a_{\mathbf{j},l}^{(\ell)} e^{-\theta_\ell\left|\varphi^{(\ell)}_{(\mathbf{i}[\ell]+k)}-\varphi^{(\ell)}_{(\mathbf{j}[\ell]+l)}\right|}\right)^2 \\
=&n_1 n_2\prod_{\ell=3}^{d} O(n_\ell^3);
\end{aligned} \tag{2.131}$$

see Lemma 2.1 for the derivation of $O(n_\ell^3)$ in the last equation. By Eq. (2.119) and the definitions of $\bar{Z}_{1,N}$ as well as $\bar{Z}_{2,N}$, we see that

$$\left(\Phi_N\right)_{1,2} = O\left(\frac{n_1 n_2\prod_{\ell=3}^{d} n_\ell^3}{\prod_{\ell\neq 1} n_\ell^2\prod_{\ell\neq 2} n_\ell^2}\right) = O(N^{-1});$$

therefore, $\left(\Sigma_d\right)_{1,2} = 0$. Similarly, we can prove $\left(\Sigma_d\right)_{i,j} = 0$, where $i\neq j$ and $i,j = 1,\cdots,d$.

Lastly, by the Cauchy-Schwarz inequality, we have that

$$\left(\Phi_N\right)_{d+1,i} = O(N^{1/(2d)-1}), \tag{2.132}$$

which leads to $\left(\Phi_0\right)_{d+1,i} = 0$ for $i = 1,\cdots,d$. Therefore, we have proven Lemma 2.4. □

Theorem 2.5.

$$\sqrt{N^{1-1/d}}(\bar{Z}_N - \mathbf{1}) \to \mathcal{N}(\mathbf{0}, \Phi_0). \tag{2.133}$$

Proof. The following is just a reiteration of the proof of Theorem 2.4 for higher-dimensional spaces. Let $L = \prod_{\ell=1}^{d}(n_\ell - 1)$. Fix an $(d+1)$-vector $\mathbf{f}\in\mathbb{R}_+^{(d+1)}$ and define the $(d+1)L\times(d+1)L$ diagonal matrix F_L by $F_L = \operatorname{diag}\{\mathbf{f}',\cdots,\mathbf{f}'\}$ for $L\geq 1$. Define the $(d+1)L$-vector $\mathbf{W}_L = (\mathbf{Y}_L'(\mathbf{j}), \mathbf{j}\in\mathcal{I}_N)'$, where

$$\mathbf{Y}_L(\mathbf{j}) = \left(\sqrt{\Delta_{\mathbf{j}}^{(1)}}\frac{\nabla_1 X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{1,1})}},\cdots,\sqrt{\Delta_{\mathbf{j}}^{(d)}}\frac{\nabla_d X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{d,1})}},\frac{\nabla_s X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{s,N})}}\right)'.$$

By construction, we have

$$\begin{aligned}
S_L &\triangleq \sqrt{N^{1-1/d}}\mathbf{f}'(\bar{Z}_N - \mathbf{1}) \\
&= \sqrt{N^{1-1/d}}(\mathbf{W}_L' F_L\mathbf{W}_L - \mathrm{E}(\mathbf{W}_L' F_L\mathbf{W}_L)).
\end{aligned} \tag{2.134}$$

Let V_L denote the covariance matrix of $\mathbf{W}_L$. Note that each entry of V_L is of the form $\sigma_L^{ab}(\mathbf{i},\mathbf{j})$ with $a,b=1,\cdots,(d+1)$. For example,

$$\sigma_L^{12}(\mathbf{i},\mathbf{j}) = \operatorname{cov}\left(\sqrt{\Delta_{\mathbf{j}}^{(1)}}\frac{\nabla_1 X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{1,1})}}, \sqrt{\Delta_{\mathbf{j}}^{(2)}}\frac{\nabla_2 X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{2,1})}}\right), \tag{2.135}$$

and

$$\sigma_L^{1(d+1)}(\mathbf{i},\mathbf{j}) = \operatorname{cov}\left(\sqrt{\Delta_{\mathbf{j}}^{(1)}}\frac{\nabla_1 X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{1,1})}}, \frac{\nabla_s X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{s,N})}}\right). \tag{2.136}$$

For convenience, $\sigma_L^{ab}(\cdot)$ will be denoted as $\sigma^{ab}(\cdot)$ below. $V_L^{\frac{1}{2}}$ is defined as the symmetric positive definite square root of V_L. Denote by

$$\Lambda_L = \operatorname{diag}(\lambda_{1,L},\cdots,\lambda_{(d+1)L,L}),$$

the diagonal matrix whose diagonal entries are eigenvalues of $2\sqrt{N^{1-1/d}}V_L^{\frac{1}{2}}F_L V_L^{\frac{1}{2}}$. Then for $\epsilon_L \sim \mathcal{N}(\mathbf{0},\mathbf{I}_{(d+1)L\times(d+1)L})$, we have

$$\sqrt{N^{1-1/d}}\mathbf{W}_L' F_L \mathbf{W}_L \stackrel{\mathrm{d}}{=} \sqrt{N^{1-1/d}}\epsilon_L' V_L^{\frac{1}{2}} F_L V_L^{\frac{1}{2}}\epsilon_L \stackrel{\mathrm{d}}{=} \frac{1}{2}\epsilon_L'\Lambda_L\epsilon_L. \tag{2.137}$$

Therefore, for all $|\theta| < \frac{1}{\max(\lambda_{1,L},\cdots,\lambda_{(d+1)L,L})}$, the cumulant generating function of S_L is given by

$$\kappa_L(\theta) \triangleq \ln\big(\mathrm{E}(e^{\theta S_L})\big) = -\frac{1}{2}\sum_{q=1}^{(d+1)L}\{\ln(1-\theta\lambda_{q,L})+\theta\lambda_{q,L}\}, \tag{2.138}$$

(see Khuri 2009, chap. 5). To obtain the limit of $\kappa_L(\theta)$ as $N\to\infty$, we first prove

$$\operatorname{tr}\big(\Lambda_L^4\big) = \sum_{q=1}^{(d+1)L}\lambda_{q,L}^4 \to 0, \quad \text{as } N\to\infty. \tag{2.139}$$

Direct calculation shows that

$$\begin{aligned}\operatorname{tr}\big(\Lambda_L^4\big) &= 16N^{2-2/d}\operatorname{tr}\left((V_L^{\frac{1}{2}}F_L V_L^{\frac{1}{2}})^4\right)\\ &= 16N^{2-2/d}\operatorname{tr}\left((V_L F_L)^4\right)\\ &= 16N^{2-2/d}\sum_{u_1=1}^{d+1}\cdots\sum_{u_4=1}^{d+1} f_{u_1}\cdots f_{u_4}\Delta_L(u_1,\cdots,u_4),\end{aligned} \tag{2.140}$$

where

$$\begin{aligned}&\Delta_L(u_1,\cdots,u_4)\\&=\sum_{\mathbf{i}_1\in\mathcal{I}_N}\cdots\sum_{\mathbf{i}_4\in\mathcal{I}_N}\sigma^{u_1u_2}(\mathbf{i}_1,\mathbf{i}_2)\sigma^{u_2u_3}(\mathbf{i}_2,\mathbf{i}_3)\sigma^{u_3u_4}(\mathbf{i}_3,\mathbf{i}_4)\sigma^{u_4u_1}(\mathbf{i}_1,\mathbf{i}_4).\end{aligned}\tag{2.141}$$

For $a,b=1,\cdots,(d+1)$ and $\mathbf{i},\mathbf{j}\in\mathcal{I}_N$, if we are able to find an upper bound $\hat{\sigma}^{ab}(\mathbf{i}-\mathbf{j})$ for each $\left|\sigma^{ab}(\mathbf{i},\mathbf{j})\right|$, the proof can be completed by imitating the stationary case; (see Chan and Wood 2000, (7.15)). Without loss of generality, we only consider $a,b\in\{1,2,(d+1)\}$:

- For the 1-1 case, we have

$$\begin{aligned}&\left|\sigma^{11}(\mathbf{i},\mathbf{j})\right|\\=&\left|\operatorname{cov}\left(\sqrt{\Delta_{\mathbf{i}}^{(1)}}\frac{\nabla_1X_{\mathbf{i}}}{\sqrt{\mathrm{E}(V_{1,1})}},\sqrt{\Delta_{\mathbf{j}}^{(1)}}\frac{\nabla_1X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{1,1})}}\right)\right|\\=&\frac{\sqrt{\Delta_{\mathbf{i}}^{(1)}\Delta_{\mathbf{j}}^{(1)}}}{\mathrm{E}(V_{1,1})}\left|\operatorname{cov}\left(\nabla_1X_{\mathbf{i}},\nabla_1X_{\mathbf{j}}\right)\right|\\\leq&\frac{\sigma^2C_{1,1}}{\mathrm{E}(V_{1,1})n_1}e^{-\theta_1\left|\varphi^{(1)}_{\mathbf{i}[1]}-\varphi^{(1)}_{\mathbf{j}[1]}\right|}\prod_{\ell=2}^{d}\left|\sum_{l,k=0}^{1}a^{(\ell)}_{\mathbf{i},k}a^{(\ell)}_{\mathbf{j},l}e^{-\theta_\ell\left|\varphi^{(\ell)}_{(\mathbf{i}[\ell]+k)}-\varphi^{(\ell)}_{(\mathbf{j}[\ell]+l)}\right|}\right|\\\leq&\frac{\sigma^2C_{1,1}}{\mathrm{E}(V_{1,1})n_1}\prod_{\ell=2}^{d}\left|\sum_{l,k=0}^{1}a^{(\ell)}_{\mathbf{i},k}a^{(\ell)}_{\mathbf{j},l}e^{-\theta_\ell\left|\varphi^{(\ell)}_{(\mathbf{i}[\ell]+k)}-\varphi^{(\ell)}_{(\mathbf{j}[\ell]+l)}\right|}\right|,\end{aligned}$$

  where $C_{1,1}$ is defined based on Eq. (2.13). Define

$$\hat{\sigma}^{11}(\mathbf{x})=\frac{C_5}{n_1\prod_{\ell=2}^{d}n_\ell^2}\prod_{\ell=2}^{d}\left(C_1^{(\ell)}n_\ell\mathbf{1}_0(\mathbf{x}[\ell])+C_2^{(\ell)}\mathbf{1}_{\mathbb{R}\backslash0}(\mathbf{x}[\ell])\right),$$

  where $C_5,C_1^{(\ell)}$, and $C_2^{(\ell)}$ are positive constants. we claim that

$$\left|\sigma^{11}(\mathbf{i},\mathbf{j})\right|\leq\hat{\sigma}^{11}(\mathbf{i}-\mathbf{j}),$$

  uniformly over $\mathbf{i},\mathbf{j}\in\mathcal{I}_N$ when N is big enough.

- For the square case, we have that

$$\begin{aligned}&\left|\sigma^{ss}(\mathbf{i},\mathbf{j})\right|\\=&\left|\mathrm{E}\left(\nabla_sX_{\mathbf{i}}\nabla_sX_{\mathbf{j}}\right)\right|/\,\mathrm{E}(V_{s,N})\\\leq&\frac{C_6}{N^2}\prod_{\ell=1}^{d}\left(C_1^{(\ell)}n_\ell\mathbf{1}_0(\mathbf{x}[\ell])+C_2^{(\ell)}\mathbf{1}_{\mathbb{R}\backslash0}(\mathbf{x}[\ell])\right),\end{aligned}\tag{2.142}$$

uniformly over $\mathbf{i}, \mathbf{j} \in \mathcal{I}_N$ when N is big enough. So we define $\hat{\sigma}^{ss}(\cdot)$ as

$$\hat{\sigma}^{ss}(\mathbf{x}) = \frac{C_6}{N^2} \prod_{\ell=1}^{d} \left(C_1^{(\ell)} n_\ell \mathbf{1}_0(\mathbf{x}[\ell]) + C_2^{(\ell)} \mathbf{1}_{\mathbb{R}\setminus 0}(\mathbf{x}[\ell]) \right),$$

where C_6 is a positive constant.

- For the 1-2 cross term,

$$\begin{aligned} &\left|\sigma^{12}(\mathbf{i}, \mathbf{j})\right| \\ =&\left|\operatorname{cov}\left(\sqrt{\Delta_{\mathbf{i}}^{(1)}} \frac{\nabla_1 X_{\mathbf{i}}}{\sqrt{\mathrm{E}(V_{1,1})}}, \sqrt{\Delta_{\mathbf{j}}^{(2)}} \frac{\nabla_2 X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{2,1})}}\right)\right| \\ =&\frac{\sqrt{\Delta_{\mathbf{i}}^{(1)}}\sqrt{\Delta_{\mathbf{j}}^{(2)}}}{\sqrt{\mathrm{E}(V_{1,1})}\sqrt{\mathrm{E}(V_{2,1})}} \left|\mathrm{E}\left(\nabla_1 X_{\mathbf{i}} \nabla_2 X_{\mathbf{j}}\right)\right|. \end{aligned} \tag{2.143}$$

According to Eq. (2.130), there exists some constant C_7 such that

$$\left|\sigma^{12}(\mathbf{i}, \mathbf{j})\right| \leq \frac{C_7}{(n_1 n_2)^{\frac{3}{2}} \prod_{\ell=3} n_\ell^2} \prod_{\ell=3}^{d} \left(C_1^{(\ell)} n_\ell \mathbf{1}_0(\mathbf{x}[\ell]) + C_2^{(\ell)} \mathbf{1}_{\mathbb{R}\setminus 0}(\mathbf{x}[\ell]) \right),$$

uniformly over $\mathbf{i}, \mathbf{j} \in \mathcal{I}_N$ when n is big enough. In this case,

$$\hat{\sigma}^{12}(\mathbf{x}) = \frac{C_7}{(n_1 n_2)^{\frac{3}{2}} \prod_{\ell=3}^{d} n_\ell^2} \prod_{\ell=3}^{d} \left(C_1^{(\ell)} n_\ell \mathbf{1}_0(\mathbf{x}[\ell]) + C_2^{(\ell)} \mathbf{1}_{\mathbb{R}\setminus 0}(\mathbf{x}[\ell]) \right),$$

where C_7 is a positive constant.

- Lastly, we study the 1-s cross term ,

$$\begin{aligned} &\left|\sigma^{1s}(\mathbf{i}, \mathbf{j})\right| \\ =&\left|\operatorname{cov}\left(\sqrt{\Delta_{\mathbf{i}}^{(1)}} \frac{\nabla_1 X_{\mathbf{i}}}{\sqrt{\mathrm{E}(V_{1,1})}}, \frac{\nabla_s X_{\mathbf{j}}}{\sqrt{\mathrm{E}(V_{s,N})}}\right)\right| \\ =&\frac{\sqrt{\Delta_{\mathbf{i}}^{(1)}}}{\sqrt{\mathrm{E}(V_{1,1})\,\mathrm{E}(V_{s,N})}} \left|\operatorname{cov}\left(\nabla_1 X_{\mathbf{i}}, (\nabla_1 X_{\mathbf{j}+(1,0)} - \nabla_1 X_{\mathbf{j}})/\Delta_{\mathbf{j}}^{(1)}\right)\right| \\ =&\frac{\sigma^2\sqrt{\Delta_{\mathbf{i}}^{(1)}}}{\sqrt{\mathrm{E}(V_{1,1})\,\mathrm{E}(V_{s,N})}} \left| e^{-\theta_1\left|\varphi_{\mathbf{i}[1]}^{(1)} - \varphi_{\mathbf{j}[1]+1}^{(1)}\right|} - e^{-\theta_1\left|\varphi_{\mathbf{i}[1]}^{(1)} - \varphi_{\mathbf{j}[1]}^{(1)}\right|}\right| / \Delta_{\mathbf{j}}^{(1)} \\ &\times \prod_{\ell=2}^{d} \left| \sum_{l,k=0}^{1} a_{\mathbf{i},k}^{(\ell)} a_{\mathbf{j},l}^{(\ell)} e^{-\theta_\ell\left|\varphi_{(\mathbf{i}[\ell]+k)}^{(\ell)} - \varphi_{(\mathbf{j}[\ell]+l)}^{(\ell)}\right|} \right| \\ \leq&\frac{\theta_1\sigma^2\sqrt{\Delta_{\mathbf{i}}^{(1)}}}{\sqrt{\mathrm{E}(V_{1,1})\,\mathrm{E}(V_{s,N})}} \prod_{\ell=2}^{d} \left| \sum_{l,k=0}^{1} a_{\mathbf{i},k}^{(\ell)} a_{\mathbf{j},l}^{(\ell)} e^{-\theta_\ell\left|\varphi_{(\mathbf{i}[\ell]+k)}^{(\ell)} - \varphi_{(\mathbf{j}[\ell]+l)}^{(\ell)}\right|} \right|, \end{aligned} \tag{2.144}$$

where the last inequality is based on Eq. (2.73). Define

$$\hat{\sigma}^{1s}(\mathbf{x}) = \frac{C_8}{(n_1)^{\frac{3}{2}} \prod_{\ell=2}^{d} n_\ell^2} \prod_{\ell=2}^{d} \Big(C_1^{(\ell)} n_\ell \mathbf{1}_0(\mathbf{x}[\ell]) + C_2^{(\ell)} \mathbf{1}_{\mathbb{R}\backslash 0}(\mathbf{x}[\ell]) \Big),$$

where C_8 is a positive constant, then

$$\Big| \sigma^{1s}(\mathbf{i}, \mathbf{j}) \Big| \leq \hat{\sigma}^{1s}(\mathbf{i} - \mathbf{j}),$$

when N is big enough.

So combining all the cases above, we have

$$\begin{aligned} &|\Delta_L(u_1, \cdots, u_4)| \\ \leq &\sum_{\mathbf{i}_1 \in \mathcal{I}_N} \cdots \sum_{\mathbf{i}_4 \in \mathcal{I}_N} |\sigma^{u_1 u_2}(\mathbf{i}_1, \mathbf{i}_2) \sigma^{u_2 u_3}(\mathbf{i}_2, \mathbf{i}_3) \sigma^{u_3 u_4}(\mathbf{i}_3, \mathbf{i}_4) \sigma^{u_4 u_1}(\mathbf{i}_1, \mathbf{i}_4)| \\ \leq &\sum_{\mathbf{i}_1 \in \mathcal{I}_N} \cdots \sum_{\mathbf{i}_4 \in \mathcal{I}_N} \hat{\sigma}^{u_1 u_2}(\mathbf{i}_1 - \mathbf{i}_2) \hat{\sigma}^{u_2 u_3}(\mathbf{i}_2 - \mathbf{i}_3) \hat{\sigma}^{u_3 u_4}(\mathbf{i}_3 - \mathbf{i}_4) \hat{\sigma}^{u_4 u_1}(\mathbf{i}_1 - \mathbf{i}_4), \end{aligned} \tag{2.145}$$

when N is big enough.

Define the index set

$$\mathcal{D}_N \triangleq \{\mathbf{i} - \mathbf{j} : \mathbf{i}, \mathbf{j} \in \mathcal{I}_N\}. \tag{2.146}$$

For each triple ($\mathbf{h}_1$, $\mathbf{h}_2$, $\mathbf{h}_3$) which satisfies $\mathbf{h}_a \in \mathcal{D}_N, 1 \leq a \leq 3$, the cardinality of the set

$$\{(\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{i}_4) : \mathbf{i}_a \in \mathcal{I}_N, a = 1, \cdots, 4; \mathbf{h}_a = \mathbf{i}_a - \mathbf{i}_{a+1}, 1 \leq a \leq 3\}$$

is bounded by L. It follows that when N is big enough

$$\begin{aligned} &|\Delta_L(u_1, \cdots, u_4)| \\ \leq &L \sum_{\mathbf{h}_1 \in \mathcal{D}_N} \cdots \sum_{\mathbf{h}_3 \in \mathcal{D}_N} \hat{\sigma}^{u_1 u_2}(\mathbf{h}_1) \hat{\sigma}^{u_2 u_3}(\mathbf{h}_2) \hat{\sigma}^{u_3 u_4}(\mathbf{h}_3) \hat{\sigma}^{u_4 u_1}(\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3) \\ \leq &C_9 \prod_{a=1}^{3} \left(\sum_{\mathbf{h}_a \in \mathcal{D}_N} \hat{\sigma}^{u_a u_{a+1}}(\mathbf{h}) \right). \end{aligned} \tag{2.147}$$

The last inequality holds since there exists some constant $C_9 > 0$ such that

$$\hat{\sigma}^{ab}(h) \leq \frac{C_9}{N},$$

for all possible $a, b \in \{1, 2, (d+1)\}$ and $\mathbf{h} \in \mathcal{D}_N$ when N is big enough.

Since the integrand $\hat{\sigma}^{ab}(h)$ is separable with respect to the counting measure, direct calculation shows that

$$\sum_{\mathbf{h}\in\mathcal{D}_N} \hat{\sigma}^{11}(\mathbf{h}) = O\left(\frac{1}{\prod_{\ell=2}^d n_\ell}\right), \qquad \sum_{\mathbf{h}\in\mathcal{D}_N} \hat{\sigma}^{ss}(\mathbf{h}) = O\left(\frac{1}{N}\right),$$
$$\sum_{\mathbf{h}\in\mathcal{D}_N} \hat{\sigma}^{12}(\mathbf{h}) = O\left(\frac{1}{\sqrt{n_1 n_2}\prod_{\ell=3}^d n_\ell}\right), \qquad \sum_{\mathbf{h}\in\mathcal{D}_N} \hat{\sigma}^{1s}(\mathbf{h}) = O\left(\frac{1}{\sqrt{n_1}\prod_{\ell=2}^d n_\ell}\right).$$

Combining Eqs. (2.140) and (2.147), we conclude that

$$\operatorname{tr}\big(\Lambda_L^4\big) = N^{2(1-1/d)} O(N^{-3(1-1/d)}) = O(N^{1/d-1}), \tag{2.148}$$

as $N \to \infty$. Namely, Eq. (2.139) holds. Meanwhile, note that Eq. (2.139) implies that as $N \to \infty$,

$$\max_{1\le q\le (d+1)L} \{\lambda_{q,L}\} \le \left(\sum_{1\le q\le (d+1)L} \lambda_{q,L}^4\right)^{\frac{1}{4}} \to 0. \tag{2.149}$$

Expanding Eq. (2.138) about $\theta = 0$ using Taylor's theorem, we obtain

$$\kappa_L(\theta) = \frac{1}{2}\sum_{q=1}^{(d+1)L} \{\frac{1}{2}(\theta\lambda_{q,L})^2 + \frac{1}{3}(\theta\lambda_{q,L})^3 + \frac{1}{4}(\theta\lambda_{q,L})^4(1-\theta_{q,L}^*\lambda_{q,L})^{-4}\}, \tag{2.150}$$

for some $\theta_{q,L}^*$ which satisfies $0 \le |\theta_{q,L}^*| \le |\theta|$.

Let us first consider the term $\sum_{q=1}^{(d+1)L} \frac{1}{2}(\theta\lambda_{q,L})^2$:

$$\begin{aligned}\sum_{q=1}^{(d+1)L} \frac{1}{2}\lambda_{q,L}^2 &= \frac{1}{2}\operatorname{tr}\big(\Lambda_L^2\big) = 2N^{1-1/d}\operatorname{tr}\big((V_L F_L)^2\big)\\ &= \operatorname{var}(\sqrt{N^{1-1/d}}W_L' F_L W_L) = N^{1-1/d}\mathbf{f}'\Phi_N\mathbf{f}.\end{aligned} \tag{2.151}$$

It follows from Lemma 2.4 that, as $N \to \infty$,

$$\sum_{q=1}^{(d+1)L} \frac{1}{2}\lambda_{q,L}^2 = N^{1-1/d}\mathbf{f}'\Phi_N\mathbf{f} \to \mathbf{f}'\Phi_0\mathbf{f}. \tag{2.152}$$

Second, Eq. (2.139) implies that, as $N \to \infty$,

$$\left| \sum_{q=1}^{(d+1)L} \lambda_{q,L}^3 \right| \le \max_{1 \le q \le (d+1)L} \{\lambda_{q,L}\} \sum_{q=1}^{(d+1)L} \lambda_{q,L}^2 \to 0. \tag{2.153}$$

Third, note that $\delta = \sup_{L \ge 1} \max_{1 \le q \le (d+1)L} \{\lambda_{q,L}\}$ is positive and finite. If we restrict attention to $|\theta| \le \frac{1}{2\delta}$, then

$$(1 - \theta_{q,L}^* \lambda_{q,L}) > 1/2, \tag{2.154}$$

for all q and L. Hence, combining Eq. (2.152), Eq. (2.153) and Eq. (2.154), we have that for $\theta \in (-\frac{1}{2\delta}, \frac{1}{2\delta})$,

$$\kappa_L(\theta) \to \theta^2 \mathbf{f}' \Phi_0 \mathbf{f}/2, \quad \text{as } n \to \infty, \tag{2.155}$$

which leads to the conclusion that $S_L \to \mathcal{N}(\mathbf{0}, \mathbf{f}'\Phi_0 \mathbf{f})$ in distribution. This completes the proof. $\square$

In this final part, we propose the estimator of $(\theta_1, \cdots, \theta_d, \sigma^2)'$ which is

$$\begin{aligned}
\hat{\theta}_1 &\triangleq \frac{\prod_{\ell \ne 1} \left(2n_\ell^2 \int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds \right) \mathrm{E}\left(V_{s,N}\right)}{\prod_{\ell=1}^d \left(2n_\ell^2 \int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds \right) \mathrm{E}\left(V_{1,1}\right)} \frac{\bar{Z}_{s,N}}{\bar{Z}_{1,N}}, \\
&\vdots \\
\hat{\theta}_d &\triangleq \frac{\prod_{\ell \ne d} \left(2n_\ell^2 \int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds \right) \mathrm{E}\left(V_{s,N}\right)}{\prod_{\ell=1}^d \left(2n_\ell^2 \int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds \right) \mathrm{E}\left(V_{d,1}\right)} \frac{\bar{Z}_{s,N}}{\bar{Z}_{d,N}}, \\
\hat{\sigma}^2 &\triangleq \frac{\left(\prod_{\ell=1}^d \left(2n_\ell^2 \int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds \right) \right)^{(d-1)} \prod_{\ell=1}^d \mathrm{E}\left(V_{\ell,1}\right)}{\prod_{\ell=1}^d \left(\prod_{j \ne \ell} \left(2n_j^2 \int_0^1 \{\varphi_j^{(1)}(s)\}^{-1} ds \right) \right) \left(\mathrm{E}\left(V_{s,N}\right)\right)^{(d-1)}} \frac{\prod_{\ell=1}^d \bar{Z}_{\ell,N}}{\left(\bar{Z}_{s,N}\right)^{(d-1)}}.
\end{aligned} \tag{2.156}$$

In a matrix form, Eq. (2.156) can be represented as

$$\begin{pmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_d \\ \hat{\sigma}^2 \end{pmatrix} \triangleq \begin{pmatrix} k_1 & & & \\ & \ddots & & \\ & & k_d & \\ & & & k_{(d+1)} \end{pmatrix} \begin{pmatrix} r_1 \\ \vdots \\ r_d \\ r_{(d+1)} \end{pmatrix}, \tag{2.157}$$

where

$$\left(r_1, \cdots, r_d, r_{(d+1)}\right) \triangleq \left(\frac{\bar{Z}_{s,N}}{\bar{Z}_{1,N}}, \cdots, \frac{\bar{Z}_{s,N}}{\bar{Z}_{d,N}}, \frac{\prod_{\ell=1}^{d} \bar{Z}_{\ell,N}}{\left(\bar{Z}_{s,N}\right)^{(d-1)}}\right), \tag{2.158}$$

and $\left(k_1, \cdots, k_d, k_{(d+1)}\right)$ is the coefficients of $\left(r_1, \cdots, r_d, r_{(d+1)}\right)$ in Eq. (2.156). Based on Eqs. (2.114) and (2.119), it is easy to see that

$$\begin{aligned} k_\ell &= \theta_\ell + O(N^{-1/d}), \\ k_{(d+1)} &= \sigma^2 + O(N^{-1/d}), \end{aligned} \tag{2.159}$$

where $\ell = 1, \cdots, d$. In the following corollary, We state all the results in high-dimensional spaces :

Corollary 2.2.

$$\begin{pmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_d \\ \hat{\sigma}^2 \end{pmatrix} \xrightarrow{a.s.} \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_d \\ \sigma^2 \end{pmatrix}, \tag{2.160}$$

and

$$N^{(d-1)/(2d)} \left(\begin{pmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_d \\ \hat{\sigma}^2 \end{pmatrix} - \begin{pmatrix} k_1 \\ \vdots \\ k_d \\ k_{(d+1)} \end{pmatrix} \right) \xrightarrow[N\to\infty]{d} \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \tilde{\Sigma}_d & \mathbf{b} \\ \mathbf{b}' & c \end{pmatrix}\right), \tag{2.161}$$

where $\tilde{\Sigma}_d$ is a diagonal matrix with the diagonal elements as

$$\left(\tilde{\Sigma}_d\right)_{j,j} = 2C_{\theta_j} \prod_{\ell \neq j} \rho_\ell \frac{\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds\right)^2} \theta_j^2,$$

and

$$\mathbf{b}_j = -2C_{\theta_j} \prod_{\ell \neq j} \rho_\ell \frac{\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds\right)^2} \theta_j \sigma^2,$$

$$c = 2\left(\sum_{j=1}^{d} C_{\theta_j} \prod_{\ell \neq j} \rho_\ell \frac{\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds\right)^2}\right) \sigma^4.$$

Proof. The almost sure convergence is implied by Eq. (2.159) and Lemma 2.4 and the Hanson-Wright inequality; see, e.g., Theorem 2.1.

Applying the delta method to Theorem 2.5, we get that

$$N^{(d-1)/(2d)}\left(\begin{pmatrix} r_1 \\ \vdots \\ r_d \\ r_{(d+1)} \end{pmatrix} - \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}\right) \xrightarrow[N\to\infty]{\mathrm{d}} \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \Gamma_d & \mathbf{e} \\ \mathbf{e}' & f \end{pmatrix}\right), \tag{2.162}$$

where Γ_d is a diagonal matrix with the diagonal elements as

$$(\Gamma_d)_{j,j} = 2C_{\theta_j} \prod_{\ell \neq j} \rho_\ell \frac{\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds\right)^2},$$

and

$$\mathbf{e}_j = -2C_{\theta_j} \prod_{\ell \neq j} \rho_\ell \frac{\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds\right)^2},$$

$$f = 2\left(\sum_{j=1}^d C_{\theta_j} \prod_{\ell \neq j} \rho_\ell \frac{\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-2} ds}{\left(\int_0^1 \{\varphi_\ell^{(1)}(s)\}^{-1} ds\right)^2}\right);$$

moreover,

$$\left(\begin{pmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_d \\ \hat{\sigma}^2 \end{pmatrix} - \begin{pmatrix} k_1 \\ \vdots \\ k_d \\ k_{(d+1)} \end{pmatrix}\right) = \begin{pmatrix} k_1 & & & \\ & \ddots & & \\ & & k_d & \\ & & & k_{(d+1)} \end{pmatrix} \left(\begin{pmatrix} r_1 \\ \vdots \\ r_d \\ r_{(d+1)} \end{pmatrix} - \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}\right). \tag{2.163}$$

Direct calculation shows the form of the right hand side of Eq. (2.161). This proves Corollary 2.2. □

CHAPTER 3

THE MULTIVARIATE ORNSTEIN-UHLENBECK PROCESS

3.1 Introduction

In this section, we review the definition and some basic properties of the multivariate Ornstein-Uhlenbeck (OU) process; (see Gardiner 1985, for a fuller exposition). The multivariate Ornstein-Uhlenbeck (OU) process $Z(t) \in \mathbb{R}^n, t \in \mathbb{R}$ is defined by the stochastic differential equation (SDE)

$$dZ(t) + AZ(t)dt = BdW(t), \tag{3.1}$$

where $A \in \mathbb{R}^{n\times n}$ is a matrix having eigenvalues with strictly positive real part, $B \in \mathbb{R}^{n\times n}$, and $W(t)$ is a n-dimensional Wiener process. Note that

$$e^{At}dZ(t) + Ae^{At}Z(t)dt = d\big(e^{At}Z(t)\big), \tag{3.2}$$

and $\lim\limits_{t\to-\infty} e^{At}Z(t) = 0$, so we have

$$e^{At}Z(t) = \int_{-\infty}^{t} e^{Au}BdW(u), \tag{3.3}$$

that is,

$$Z(t) = \int_{-\infty}^{t} e^{-A(t-u)}BdW(u). \tag{3.4}$$

Based on the form of $Z(t)$, we see that

$$\mathrm{E}(Z(t)) = \mathbf{0}, \tag{3.5}$$

$$\mathrm{cov}(Z(t), Z(s)) = \int_{-\infty}^{\min(t,s)} e^{-A(t-u)}BB'e^{-A'(s-u)}du. \tag{3.6}$$

Let us define the stationary covariance matrix Σ as

$$\Sigma \triangleq \mathrm{cov}(Z(t), Z(t)). \tag{3.7}$$

Then

$$\begin{aligned} A\Sigma + \Sigma A' &= \int_{-\infty}^{t} Ae^{-A(t-u)}BB'e^{-A'(t-u)}du \\ &\quad + \int_{-\infty}^{t} e^{-A(t-u)}BB'e^{-A'(t-u)}A'du \\ &= \int_{-\infty}^{t} \frac{d}{du}e^{-A(t-u)}BB'e^{-A'(t-u)}du \\ &= BB'. \end{aligned} \tag{3.8}$$

From Eq. (3.6), we see that if $t > s$,

$$\begin{aligned} \operatorname{cov}(Z(t), Z(s)) &= e^{-A(t-s)} \int_{-\infty}^{s} e^{-A(s-u)}BB'e^{-A'(s-u)}du \\ &= e^{-A(t-s)}\Sigma, \quad t > s, \end{aligned} \tag{3.9}$$

and similarly,

$$\operatorname{cov}(Z(t), Z(s)) = \Sigma e^{-A'(s-t)}, \quad t < s. \tag{3.10}$$

Velandia et al. (2017) considered the parameter estimation of the multivariate OU under fixed-domain asymptotics when

$$A = \begin{pmatrix} \theta & 0 \\ 0 & \theta \end{pmatrix}, \quad BB' = 2\theta \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}. \tag{3.11}$$

Combining Eqs. (3.8) to (3.10), we see that

$$\operatorname{cov}(Z(t), Z(s)) = e^{-\theta|t-s|} \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}, \tag{3.12}$$

where σ_1^2, σ_2^2 are marginal variance parameters, $\theta > 0$ is called a correlation decay parameter, and the quantity ρ with $|\rho| < 1$ is called the colocated correlation parameter (Gneiting, Kleiber, and Schlather 2010). For simplicity, let $\psi = (\sigma_1^2, \sigma_2^2, \rho, \theta)'$. They showed the maximum likelihood estimator (MLE) of $(\theta\sigma_1^2, \theta\sigma_2^2, \rho)'$ is strongly consistent and asymptotically normal given some prior information of ψ in Velandia et al. (2017). However, the MLE in this case must be found numerically. Therefore, we plan to construct an explicit estimator of $(\theta\sigma_1^2, \theta\sigma_2^2, \rho)'$, which is computationally efficient in practice and asymptotically identical to the MLE under fixed-domain asymptotics. The asymptotic results are formally stated in Theorem 3.3.

3.2 Asymptotic theory for estimation

Under fixed-domain asymptotics, the random process $Z(t) \triangleq (Z_1(t), Z_2(t))'$ is sampled increasingly densely over a fixed and bounded set T. Without loss of generality, we consider $T = [0, 1]$, and the sampling design $T_n = \{t_1, \cdots, t_n\}$ with

$$0 \le t_1 < t_2 < \cdots < t_n \le 1. \tag{3.13}$$

The corresponding observation is denoted as

$$Z_n = (Z'_{1,n}, Z'_{2,n})', \tag{3.14}$$

where $Z_{i,n} = (Z_i(t_1), \cdots, Z_i(t_n))'$ for $i = 1, 2$. Let

$$\Sigma(\psi) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix} \otimes R(\theta), \tag{3.15}$$

where $\otimes$ is the Kronecker product, and

$$R(\theta) = \left(e^{-\theta|t_i - t_j|}\right)_{1 \le i,j \le n}$$

is a matrix-valued function. By Eq. (3.12), we see that $Z_n \sim \mathcal{N}(\mathbf{0}, \Sigma(\psi))$. The associated likelihood function is given by

$$L_n(\psi) = (2\pi)^{-n} |\Sigma(\psi)|^{-1/2} \exp\left\{ -\frac{1}{2} Z'_n \Sigma(\psi)^{-1} Z_n \right\}. \tag{3.16}$$

In the rest of this section, we will denote by $\theta_0, \sigma_{01}^2, \sigma_{02}^2$ and ρ_0 the true but unknown parameter. The following theorem defines the asymptotic behavior of the MLE with respect to $L_n(\psi)$.

Theorem 3.1 (Velandia et al. (2017))**.** *Let T_n be dense in $[0, 1]$ as n goes to infinity. 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 $0 < a_\theta \le \theta_0 \le b_\theta < \infty$, $0 < a_{\sigma_1} \le \sigma_{01}^2 \le b_{\sigma_1} < \infty$, $0 < a_{\sigma_2} \le \sigma_{02}^2 \le b_{\sigma_2} < \infty$, and $-1 < a_\rho \le \rho_0 \le b_\rho < 1$. Define $\hat{\psi} = (\hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\rho}, \hat{\theta})'$ as*

$$\hat{\psi} = \underset{\psi \in J}{\operatorname{argmax}}\, L_n(\psi). \tag{3.17}$$

Then

$$\hat{\theta}\hat{\sigma}_1^2 \xrightarrow[n\to\infty]{a.s.} \theta_0\sigma_{01}^2, \tag{3.18}$$

$$\hat{\theta}\hat{\sigma}_2^2 \xrightarrow[n\to\infty]{a.s.} \theta_0\sigma_{02}^2, \tag{3.19}$$

$$\hat{\rho} \xrightarrow[n\to\infty]{a.s.} \rho_0, \tag{3.20}$$

and

$$\sqrt{n}\left(\begin{pmatrix}\hat{\theta}\hat{\sigma}_1^2\\ \hat{\theta}\hat{\sigma}_2^2\\ \hat{\rho}\end{pmatrix} - \begin{pmatrix}\theta_0\sigma_{01}^2\\ \theta_0\sigma_{02}^2\\ \rho_0\end{pmatrix}\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} 2\theta_0^2\sigma_{01}^4 & 2\theta_0^2\sigma_{01}^2\sigma_{02}^2\rho_0^2 & \theta_0\rho_0\sigma_{01}^2(1-\rho_0^2) \\ 2\theta_0^2\sigma_{01}^2\sigma_{02}^2\rho_0^2 & 2\theta_0^2\sigma_{02}^4 & \theta_0\rho_0\sigma_{02}^2(1-\rho_0^2) \\ \theta_0\rho_0\sigma_{01}^2(1-\rho_0^2) & \theta_0\rho_0\sigma_{02}^2(1-\rho_0^2) & (1-\rho_0^2)^2 \end{pmatrix}\right). \tag{3.21}$$

Our estimator for $(\theta_0\sigma_{01}^2, \theta_0\sigma_{02}^2, \rho_0)'$ is inspired by the results in Ying (1991) where the parameter estimation for univariate OU under fixed-domain asymptotics were studied. With the observation $Z_{1,n} = (Z_1(t_1), \cdots, Z_1(t_n))'$, we have the likelihood function

$$L_{1,n}(\sigma_1^2, \theta) = (2\pi)^{-n/2}\left|\sigma_1^2 R(\theta)\right|^{-1/2} \exp\left\{-\frac{1}{2\sigma_1^2} Z_{1,n}' R(\theta)^{-1} Z_{1,n}\right\}. \tag{3.22}$$

The following result can be found in Theorem 1 of Ying (1991).

Theorem 3.2 (Ying (1991))**.** *For any fixed* $\theta_f > 0$, *let*

$$\hat{\sigma}_1^2 = \underset{\sigma_1^2\in(0,\infty)}{\operatorname{argmax}}\, L_{1,n}(\sigma_1^2, \theta_f). \tag{3.23}$$

Then

$$\hat{\sigma}_1^2\theta_f \xrightarrow[n\to\infty]{a.s.} \sigma_{01}^2\theta_0, \tag{3.24}$$

under the sampling design Eq. (3.13)*.*

Note that for the fixed $\theta_f > 0$, $\hat{\sigma}_1^2$ has an explicit form, that is,

$$\hat{\sigma}_1^2 = \frac{1}{n} Z_{1,n}' R^{-1}(\theta_f) Z_{1,n}. \tag{3.25}$$

Therefore, we define the estimator of $\sigma_{01}^2\theta_0$ as

$$\widehat{\sigma_1^2\theta} \triangleq \frac{\theta_f}{n} Z_{1,n}' R^{-1}(\theta_f) Z_{1,n}. \tag{3.26}$$

Similarly, the estimator of $\sigma_{02}^2\theta_0$ is defined as

$$\widehat{\sigma_2^2\theta} \triangleq \frac{\theta_f}{n} Z_{2,n}' R^{-1}(\theta_f) Z_{2,n}. \tag{3.27}$$

To estimate ρ_0, we notice that $Z_3(t) \triangleq Z_1(t) + Z_2(t)$ is also a mean 0 univariate OU process with covariance function

$$\text{cov}(Z_3(t), Z_3(s)) = (\sigma_{01}^2 + \sigma_{02}^2 + 2\sigma_{01}\sigma_{02}\rho_0)e^{-\theta_0|s-t|}. \tag{3.28}$$

As a result, we define

$$\hat{\rho} \triangleq \frac{Z_{1,n}' R^{-1}(\theta_f) Z_{2,n}}{\sqrt{Z_{1,n}' R^{-1}(\theta_f) Z_{1,n} \cdot Z_{2,n}' R^{-1}(\theta_f) Z_{2,n}}}. \tag{3.29}$$

The following theorem establishes the strong consistency and asymptotic normality of our estimator for $(\sigma_{01}^2\theta_0, \sigma_{02}^2\theta_0, \rho_0)'$.

Theorem 3.3. *Under the sampling design Eq.* (3.13)*,*

$$\begin{pmatrix} \widehat{\sigma_1^2\theta} \\ \widehat{\sigma_2^2\theta} \\ \hat{\rho} \end{pmatrix} \xrightarrow[n\to\infty]{a.s.} \begin{pmatrix} \sigma_{01}^2\theta_0 \\ \sigma_{02}^2\theta_0 \\ \rho_0 \end{pmatrix}, \tag{3.30}$$

and

$$\sqrt{n}\left(\begin{pmatrix} \widehat{\sigma_1^2\theta} \\ \widehat{\sigma_2^2\theta} \\ \hat{\rho} \end{pmatrix} - \begin{pmatrix} \sigma_{01}^2\theta_0 \\ \sigma_{02}^2\theta_0 \\ \rho_0 \end{pmatrix}\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} 2\theta_0^2\sigma_{01}^4 & 2\theta_0^2\sigma_{01}^2\sigma_{02}^2\rho_0^2 & \theta_0\rho_0\sigma_{01}^2(1-\rho_0^2) \\ 2\theta_0^2\sigma_{01}^2\sigma_{02}^2\rho_0^2 & 2\theta_0^2\sigma_{02}^4 & \theta_0\rho_0\sigma_{02}^2(1-\rho_0^2) \\ \theta_0\rho_0\sigma_{01}^2(1-\rho_0^2) & \theta_0\rho_0\sigma_{02}^2(1-\rho_0^2) & (1-\rho_0^2)^2 \end{pmatrix}\right). \tag{3.31}$$

Proof. The strong consistency in Eq. (3.30) directly follows Theorem 3.2.

In the following, we deal with the asymptotic normality part. Let σ_{03}^2 be the true but unknown marginal variance of the OU process $Z_3(t)$ which is the sum of $Z_1(t)$ and $Z_1(t)$, that is,

$$\sigma_{03}^2 \triangleq \sigma_{01}^2 + \sigma_{02}^2 + 2\sigma_{01}\sigma_{02}\rho_0.$$

Meanwhile, define

$$Z_{3,n} = (Z_3(t_1), \cdots, Z_3(t_n))'.$$

To simplify notations, for $i = 1, 2, 3$, we write $Z_{i,n}$ as Z_i for the rest of the proof. For $i \in \{1, 2, 3\}$, and $k \in \{1, \dots, n\}$, let

$$W_{i,k} = \frac{z_{i,k} - e^{-\theta_0 \Delta_k} z_{i,k-1}}{(\sigma_{0i}^2(1 - e^{-2\theta_0 \Delta_k}))^{1/2}},$$

where $\Delta_k \triangleq t_k - t_{k-1}$, and $z_{i,k} \triangleq Z_i(t_k)$. From the proof of Theorem 2 in Ying (1991), we see that for $i = 1, 2, 3$, as $n \to \infty$,

$$\sqrt{n}\left(\frac{Z_i' R^{-1}(\theta_f) Z_i}{n}\theta_f - \sigma_{0i}^2\theta_0\right) = n^{-1/2}\sigma_{0i}^2\theta_0 \sum_{k=2}^{n}(W_{i,k}^2 - 1) + O_p(n^{-1/2}); \tag{3.32}$$

(see also, Du, Zhang, and Mandrekar 2009, (B.36)). Since

$$2Z_1^T R^{-1}(\theta_f) Z_2 = Z_3^T R^{-1}(\theta_f) Z_3 - Z_1^T R^{-1}(\theta_f) Z_1 - Z_2^T R^{-1}(\theta_f) Z_2,$$

there is

$$\begin{aligned}
&\sqrt{n}\left(\frac{Z_1^T R^{-1}(\theta_f) Z_2}{n}\theta_f - \sigma_{01}\sigma_{02}\rho_0\theta_0\right) \\
=&\frac{\sqrt{n}}{2}\Bigg[\left(\frac{Z_3' R^{-1}(\theta_f) Z_3}{n}\theta_f - \sigma_{03}^2\theta_0\right) - \left(\frac{Z_1' R^{-1}(\theta_f) Z_1}{n}\theta_f - \sigma_{01}^2\theta_0\right) \\
&- \left(\frac{Z_2' R^{-1}(\theta_f) Z_2}{n}\theta_f - \sigma_{02}^2\theta_0\right)\Bigg] \\
=&\frac{\theta_0}{2\sqrt{n}}\sum_{k=2}^{n}\left[\sigma_{03}^2\big(W_{3,k}^2 - 1\big) - \sigma_{01}^2\big(W_{1,k}^2 - 1\big) - \sigma_{02}^2\big(W_{2,k}^2 - 1\big)\right] + O_p(n^{-1/2}) \\
=&n^{-1/2}\sum_{k=2}^{n}\frac{\tilde{W}_{3,k} - \tilde{W}_{1,k} - \tilde{W}_{2,k}}{2} + O_p(n^{-1/2}),
\end{aligned} \tag{3.33}$$

where $\tilde{W}_{i,k} = \sigma_{0i}^2\theta_0\big(W_{i,k}^2 - 1\big)$ for $i = 1, 2, 3$. Denote

$$\tilde{W}_{nk} = \left(\tilde{W}_{1,k}, \tilde{W}_{2,k}, \frac{\tilde{W}_{3,k} - \tilde{W}_{1,k} - \tilde{W}_{2,k}}{2}\right)',$$

then for any fixed n, it is not hard to verify that $\left\{\tilde{W}_{nk}\right\}_{k=2}^{n}$ are independent random vectors with zero mean and finite variance. Since for any n and i, $W_{i,k} \sim \mathcal{N}(0, 1)$, we have

$$\operatorname{var}\big(W_{i,k}^2 - 1\big) = 2, \qquad \operatorname{cov}\big(W_{i,k}^2 - 1, W_{j,k}^2 - 1\big) = 2[\operatorname{cov}(W_{i,k}, W_{j,k})]^2 \triangleq 2c_{i,j}^2.$$

From the definition, it is straightforward that $c_{1,2} = \rho_0$. Moreover,

$$\begin{aligned} c_{1,3} &= \operatorname{cov}\left(W_{1,k}, W_{3,k}\right) \\ &= \operatorname{cov}\left[W_{1,k}, \frac{\sigma_{01}}{\sigma_{03}}\left(W_{1,k} + \frac{\sigma_{02}}{\sigma_{01}} W_{2,k}\right)\right] \\ &= \frac{\sigma_{01}}{\sigma_{03}} \operatorname{var}(W_{1,k}) + \frac{\sigma_{02}}{\sigma_{03}} c_{1,2} \\ &= \frac{\sigma_{01} + \sigma_{02}\rho_0}{\sqrt{\sigma_{01}^2 + \sigma_{02}^2 + 2\sigma_{01}\sigma_{02}\rho_0}}. \end{aligned}$$

Similarly,

$$c_{2,3} = \frac{\sigma_{01}\rho_0 + \sigma_{02}}{\sqrt{\sigma_{01}^2 + \sigma_{02}^2 + 2\sigma_{01}\sigma_{02}\rho_0}}.$$

Thus,

$$\operatorname{cov}(\tilde{W}_{nk}) = \theta_0^2 \begin{pmatrix} 2\sigma_{01}^4 & 2\sigma_{01}^2\sigma_{02}^2\rho_0^2 & \alpha \\ 2\sigma_{01}^2\sigma_{02}^2\rho_0^2 & 2\sigma_{02}^4 & \beta \\ \alpha & \beta & \gamma \end{pmatrix}, \tag{3.34}$$

where

$$\begin{aligned} \alpha &= (2\sigma_{01}^2\sigma_{03}^2 c_{1,3}^2 - 2\sigma_{01}^4 - 2\sigma_{01}^2\sigma_{02}^2 c_{1,2}^2)/2 = 2\sigma_{01}^3\sigma_{02}\rho_0, \\ \beta &= (2\sigma_{02}^2\sigma_{03}^2 c_{2,3}^2 - 2\sigma_{02}^4 - 2\sigma_{01}^2\sigma_{02}^2 c_{1,2}^2)/2 = 2\sigma_{01}\sigma_{02}^3\rho_0, \\ \gamma &= \frac{1}{4}(2\sigma_{03}^4 + 2\sigma_{01}^4 + 2\sigma_{02}^4 - 4\sigma_{01}^2\sigma_{03}^2 c_{1,3}^2 - 4\sigma_{02}^2\sigma_{03}^2 c_{2,3}^2 + 4\sigma_{01}^2\sigma_{02}^2 c_{1,2}^2) = \sigma_{01}^2\sigma_{02}^2(1+\rho_0^2). \end{aligned}$$

By the multivariate Lindeberg-Feller CLT, as $n \to \infty$,

$$n^{-1/2} \sum_{k=2}^{n} \tilde{W}_{nk} \xrightarrow{\mathrm{d}} \mathcal{N}(0, V), \tag{3.35}$$

where

$$V = \operatorname{cov}(\tilde{W}_{nk}) = \theta_0^2 \begin{pmatrix} 2\sigma_{01}^4 & 2\sigma_{01}^2\sigma_{02}^2\rho_0^2 & 2\sigma_{01}^3\sigma_{02}\rho_0 \\ 2\sigma_{01}^2\sigma_{02}^2\rho_0^2 & 2\sigma_{02}^4 & 2\sigma_{01}\sigma_{02}^3\rho_0 \\ 2\sigma_{01}^3\sigma_{02}\rho_0 & 2\sigma_{01}\sigma_{02}^3\rho_0 & \sigma_{01}^2\sigma_{02}^2(1+\rho_0^2) \end{pmatrix}, \tag{3.36}$$

By Eqs. (3.32), (3.33) and (3.35), we have that as $n \to \infty$,

$$\sqrt{n}\left(\theta_f \begin{pmatrix} Z_1' R^{-1}(\theta_f) Z_1/n \\ Z_2' R^{-1}(\theta_f) Z_2/n \\ Z_1' R^{-1}(\theta_f) Z_2/n \end{pmatrix} - \begin{pmatrix} \sigma_{01}^2\theta_0 \\ \sigma_{02}^2\theta_0 \\ \sigma_{01}\sigma_{02}\rho_0\theta_0 \end{pmatrix}\right) \xrightarrow{\mathrm{d}} \mathcal{N}(0, V). \tag{3.37}$$

Define function $g(x_1, x_2, x_3) = (x_1, x_2, x_3/\sqrt{x_1 x_2})' : \mathbb{R}^3 \to \mathbb{R}^3$. Then

$$\left(\widehat{\sigma_1^2\theta}, \widehat{\sigma_2^2\theta}, \hat{\rho}\right)' = g\big(\theta_f Z_1' R^{-1}(\theta_f) Z_1/n, \theta_f Z_2' R^{-1}(\theta_f) Z_2/n, \theta_f Z_1' R^{-1}(\theta_f) Z_2/n\big).$$

Using the multivariate Delta method,

$$\sqrt{n}\left(\begin{pmatrix}\widehat{\sigma_1^2\theta} \\ \widehat{\sigma_2^2\theta} \\ \hat{\rho}\end{pmatrix} - \begin{pmatrix}\sigma_{01}^2\theta_0 \\ \sigma_{02}^2\theta_0 \\ \rho_0\end{pmatrix}\right) \xrightarrow{\text{d}} \mathcal{N}(0, (\nabla g)V(\nabla g)'), \tag{3.38}$$

where

$$(\nabla g)_{i,j} = \frac{\partial g_i}{\partial x_j},$$

and

$$(\nabla g)V(\nabla g)' = \begin{pmatrix} 2\theta_0^2\sigma_{01}^4 & 2\theta_0^2\sigma_{01}^2\sigma_{02}^2\rho_0^2 & \theta_0\rho_0\sigma_{01}^2(1-\rho_0^2) \\ 2\theta_0^2\sigma_{01}^2\sigma_{02}^2\rho_0^2 & 2\theta_0^2\sigma_{02}^4 & \theta_0\rho_0\sigma_{02}^2(1-\rho_0^2) \\ \theta_0\rho_0\sigma_{01}^2(1-\rho_0^2) & \theta_0\rho_0\sigma_{02}^2(1-\rho_0^2) & (1-\rho_0^2)^2 \end{pmatrix}.$$

□

Remark 3.1. *Note that the construction of our estimator is based on each component of the bivariate process $Z(s)$, and it has the same asymptotic matrix as that in Theorem 3.1. This echos the fact that cokriging is identical to kriging under the bivariate OU model in Zhang and Cai (2015).*

3.3 Simulations

In this section, we investigate the finite sample performance of the estimator $\left(\widehat{\sigma_1^2\theta}, \widehat{\sigma_2^2\theta}, \hat{\rho}\right)'$ with $\theta_f = 1$ in Eqs. (3.26), (3.27) and (3.29). The simulations are conducted over an irregular grid within $[0, 1]$ with $n = 200, 250, \ldots, 1000$. Specifically, we have considered a regular grid with increment $1/n$ over $[0, 1]$. Then the grid points have been perturbed, adding a uniform random value on $[-1/(5n), 1/(5n)]$. We simulate, using the R package RandomFields developed by Schlather et al. (2017), and then we estimate with 1000 realizations from a zero mean bivariate OU process where $\sigma_1^2 = \sigma_2^2 = 1$, $\rho = 0.2$ and $\theta = 0.2, 0.5$.

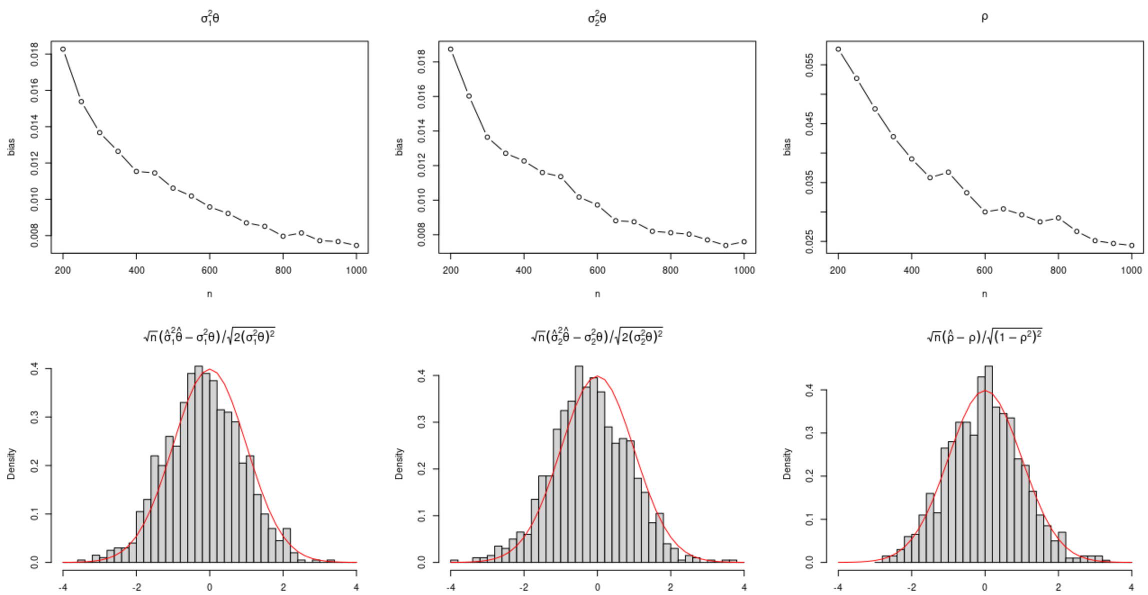

Figure 3.1 Simulated estimation for the bivariate OU process with $\theta = 0.2$. Plots in the first row present the averaged absolute value of bias for each sample size and each parameter; plots in the second row present the distribution of normalized bias when $n = 1000$, where the red curve is the density of $N(0, 1)$.

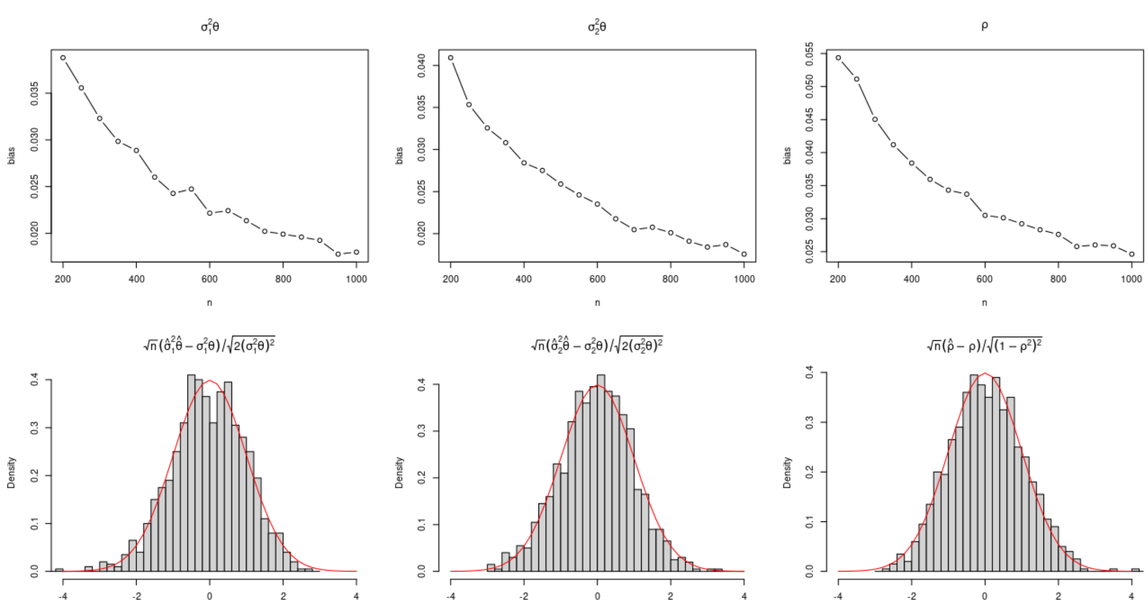

Figure 3.2 The setting is the same as Fig. 3.1 except $\theta = 0.5$.

CHAPTER 4

THE POWERED EXPONENTIAL FIELD

4.1 Introduction

In this chapter, we consider a stationary, isotropic Gaussian random field $\{X(\mathbf{t}) : \mathbf{t} \in \mathbb{R}^d\}$ with mean 0 and covariance function

$$C(\mathbf{u}-\mathbf{v}) \triangleq \operatorname{cov}\{X(\mathbf{u}), X(\mathbf{v})\} = \sigma^2 e^{-\theta\|\mathbf{u}-\mathbf{v}\|^{\alpha}}, \quad \forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^d, \tag{4.1}$$

where $\alpha \in (0, 2)$, $\theta > 0$ and $\sigma > 0$. The corresponding spectral density is defined as the Fourier transform of Eq. (4.1):

$$f_{\sigma,\theta}(\boldsymbol{\omega}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\boldsymbol{\omega}'\mathbf{t}} C(\mathbf{t}) d\mathbf{t}. \tag{4.2}$$

Note that $f_{\sigma,\theta} : \mathbb{R}^d \to \mathbb{R}$ is actually the probability density function of a sub-Gaussian random vector (see, e.g., Samorodnitsky and Taqqu 2017, Proposition 2.5.5). By the Bochner's theorem (see, e.g., Adler and Taylor 2007, Theorem 5.4.1), Eq. (4.1) is a covariance in $\mathbb{R}^d$ for any d.

Another popular covariance function in spatial statistics is the Matérn covariance function

$$C_M(\mathbf{t}) = \sigma^2 \frac{(\beta\|\mathbf{t}\|)^{\nu}}{2^{\nu-1}\Gamma(\nu)} \mathcal{K}_{\nu}(\beta\|\mathbf{t}\|), \quad \forall\, \mathbf{t} \in \mathbb{R}^d, \tag{4.3}$$

where ν, β and σ^2 are strictly positive, and $\mathcal{K}_\nu$ is the modified Bessel function of the second kind. Two covariance functions $C(\mathbf{t})$ and $C_M(\mathbf{t})$ exhibit the same behavior at the origin if $\alpha = 2\nu$. Specifically,

$$\sigma^2 - C(\mathbf{t}) \asymp \sigma^2 - C_M(\mathbf{t}) \asymp \|\mathbf{t}\|^{\alpha}, \tag{4.4}$$

as $\|\mathbf{t}\| \to 0$ given $\alpha = 2\nu$. As a result, the realizations of $X(\mathbf{t})$ have fractal dimension $(d + 1 - \alpha/2)$ with probability 1 (Gneiting and Schlather 2004; Adler 2010). However, Contrary to the fact that for $d \leq 3$, σ^2 and β in Eq. (4.3) cannot be estimated consistently under fixed-domain asymptotics proven by Zhang (2004), Anderes (2010) showed that both σ^2 and θ in Eq. (4.1) can be consistently estimated given $\alpha \in (0, d/2) \setminus \{1\}$ under fixed-domain asymptotics. This shows that the powered exponential model and the Matérn model have different statistical properties when $d \le 3$. As noted in Zhang (2004), the spectral densities corresponding to the powered exponential model do not have a closed form except in some special cases, which brings challenges when deriving results on the equivalence of Gaussian measures.

The remainder of this chapter is organized as follows. In Section 4.2, we characterize the equivalence of Gaussian measure under the powered exponential model when $\alpha \in (d/2, 2)$. In Section 4.3, we establish the strong consistency and asymptotic normality of the MLE for $\sigma^2\theta$. In Section 4.4, we provide simulations of the MLE under finite sample cases.

4.2 The equivalence of Gaussian measures

To study sufficient conditions for the equivalence of two Gaussian measures for the powered exponential class, we use Theorem A.1 in M. L. Stein (2004):

Let $P_i,\ i = 1, 2$, be two probability measures such that under P_i, the process $\mathbf{X}(\mathbf{s})$, $\mathbf{s} \in \mathbb{R}^d$, is stationary Gaussian with mean 0 and a second-order spectral density $f_i(\mathbf{v})$, $\mathbf{v} \in \mathbb{R}^d$. If, for some $\alpha > d$, $f_1(\mathbf{v})|\mathbf{v}|^\alpha$ is bounded away from 0 and ∞ as $|\mathbf{v}| \to \infty$, and for some finite c,

$$\int_{|\mathbf{v}|>c} \left\{ \frac{f_1(\mathbf{v}) - f_2(\mathbf{v})}{f_1(\mathbf{v})} \right\}^2 d\mathbf{v} < \infty, \tag{4.5}$$

then for any bounded subset $D \subset \mathbb{R}^d$, $P_1 \equiv P_2$ on the paths of $X(\mathbf{s})$, $\mathbf{s} \in D$.

Since Eq. (4.1) and the characteristic function of a stable distribution have the same form, we will take advantage of the knowledge from stable distributions to verify conditions in Theorem A.1 in M. L. Stein (2004) for the powered exponential family, which is inspired by Formula (3.4) in Wang (2010). In Samorodnitsky and Taqqu (2017), a stable distribution is defined as follows:

Definition 4.1. A random variable X is said to have a stable distribution $S_\alpha(\sigma, \beta, \mu)$ if there are parameters $0 < \alpha \le 2$, $\sigma \ge 0$, $-1 \le \beta \le 1$, and μ real such that its characteristic function has the follow form:

$$\mathrm{E}\exp\{i\omega X\} = \begin{cases} \exp\big\{-\sigma^\alpha|\omega|^\alpha(1 - i\beta(\text{sign } \omega)\tan\frac{\pi\alpha}{2}) + i\mu\omega\big\}, & \text{if } \alpha \ne 1, \\ \exp\big\{-\sigma|\omega|(1 + i\beta\frac{2}{\pi}(\text{sign } \omega)\ln|\omega|) + i\mu\omega\big\}, & \text{if } \alpha = 1, \end{cases} \tag{4.6}$$

where

$$\text{sign }\omega = \begin{cases} 1, & \text{if } \omega > 0, \\ 0, & \text{if } \omega = 0, \\ -1, & \text{if } \omega < 0. \end{cases}$$

Now consider $X \sim S_\alpha(1, 0, 0)$, $0 < \alpha < 1$. By Eq. (4.6), it has ch.f.

$$\psi(\omega) = \mathrm{E}\exp\{i\omega X\} = \exp\{-|\omega|^\alpha\}.$$

And its density function can be expanded into a convergent series as follows (see, e.g., Bergström 1952, Equation (4); Feller 1991, Lemma 1, p. 583)

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\psi(\omega)\exp\{-i\omega x\}d\omega = -\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\frac{\Gamma(\alpha k+1)}{x^{\alpha k+1}}\sin\left(k\frac{\alpha\pi}{2}\right), \quad x > 0.$$

Obviously, for $\theta > 0$, $\theta^{\frac{1}{\alpha}}X$ has ch.f.

$$\psi_\theta(\omega) = \exp\{-\theta|\omega|^\alpha\}.$$

Meanwhile

$$\begin{aligned} f_\theta(x) &= \theta^{-\frac{1}{\alpha}}f(\theta^{-\frac{1}{\alpha}}x) \\ &= -\theta^{-\frac{1}{\alpha}}\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\frac{\Gamma(\alpha k+1)}{x^{\alpha k+1}}\theta^{\frac{\alpha k+1}{\alpha}}\sin\left(k\frac{\alpha\pi}{2}\right) \\ &= \theta\frac{\Gamma(\alpha+1)}{\pi}\sin\left(\frac{\alpha\pi}{2}\right)x^{-(\alpha+1)} - \theta^2\frac{\Gamma(2\alpha+1)}{2\pi}\sin(\alpha\pi)x^{-(2\alpha+1)} + O(x^{-(3\alpha+1)}), \end{aligned} \tag{4.7}$$

as $x \to \infty$. And for $1 < \alpha < 2$, we have the same asymptotic expansion Eq. (4.7) according to Eq. (5) in Bergström (1952).

As for the multidimensional case, let $\mathbf{Y}$ be a centered d-dimensional isotropic stable random vector with characteristic function $\exp\{-\theta|\mathbf{u}|^\alpha\}$. The amplitude of $\mathbf{Y}$ is defined by

$$R = \|\mathbf{Y}\| = \sqrt{\mathbf{Y}_1^2 + \cdots + \mathbf{Y}_d^2}.$$

Then the density function of R has the following asymptotic series expansion:

$$\begin{aligned} f_R(r) = &\frac{\theta}{\pi\Gamma(d/2)}\Gamma\left(\frac{\alpha+2}{2}\right)\Gamma\left(\frac{\alpha+d}{2}\right)\sin\left(\frac{\alpha\pi}{2}\right)\left(\frac{2}{r}\right)^{\alpha+1} \\ &- \frac{\theta^2}{2\pi\Gamma(d/2)}\Gamma(\alpha+1)\Gamma\left(\frac{2\alpha+d}{2}\right)\sin(\alpha\pi)\left(\frac{2}{r}\right)^{2\alpha+1} + O(r^{-3\alpha-1}), \end{aligned} \tag{4.8}$$

as $r \to \infty$ for $\alpha \in (0,1) \cup (1,2)$. Meanwhile for $\mathbf{y} \neq \mathbf{0}$,

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\Gamma(d/2)}{2\pi^{d/2}} |\mathbf{y}|^{1-d} f_R(|\mathbf{y}|); \tag{4.9}$$

(see Nolan 2005, Section 3.1 and 3.2 for details). And direct calculation shows that the asymptotic expression $f_{\mathbf{Y}}(\mathbf{y})$ is the same as Eq. (4.7) when $d = 1$.

Lemma 4.1. *Let P_i, $i = 1, 2$, be two probability measures such that under P_i, the process $\mathbf{X}(\mathbf{s})$, $\mathbf{s} \in \Omega \subset \mathbb{R}^d$, is stationary Gaussian with mean 0 and a covariance function $C_i(\mathbf{t}) = \sigma_i^2 e^{-\theta_i ||\mathbf{t}||^\alpha}$ with parameters $(\sigma_i, \theta_i, \alpha)$ where $\alpha \in (0,2)$. If $\alpha \in \left(\frac{d}{2}, 2\right)$, then for any bounded infinite set $\Omega \subset \mathbb{R}^d$, the Gaussian measures P_1 and P_2 are equivalent if and only if $\sigma_1^2\theta_1 = \sigma_2^2\theta_2$.*

Proof. Note that when $\alpha = 1$ and $d = 1$, Theorem 2 in Zhang 2004 implies the theorem holds. Below, we only consider the case $\alpha \in \left(\frac{d}{2}, 2\right) \setminus \{1\}$.

" $\Leftarrow$ " Here we follow the proof of Theorem 2 in Zhang 2004. For powered exponential covariance function $C(\mathbf{t}) = \sigma^2 e^{-\theta ||\mathbf{t}||^\alpha}$ in $\mathbb{R}^d$, the corresponding isotropic spectral density has the following asymptotic expansion:

$$\begin{aligned} f(u) &\triangleq \sigma^2 f_{\mathbf{Y}}((u, 0, \cdots, 0)) \\ &= \sigma^2 \frac{\Gamma(d/2)}{2\pi^{d/2}} u^{1-d} f_R(u) \\ &= \sigma^2 \frac{\Gamma(d/2)}{2\pi^{d/2}} u^{1-d} \Bigg(\frac{\theta}{\pi\Gamma(d/2)} \Gamma\left(\frac{\alpha+2}{2}\right) \Gamma\left(\frac{\alpha+d}{2}\right) \sin\left(\frac{\alpha\pi}{2}\right) \left(\frac{2}{u}\right)^{\alpha+1} \\ &\quad - \frac{\theta^2}{2\pi\Gamma(d/2)} \Gamma(\alpha+1) \Gamma\left(\frac{2\alpha+d}{2}\right) \sin(\alpha\pi) \left(\frac{2}{u}\right)^{2\alpha+1} + O(u^{-3\alpha-1}) \Bigg) \\ &= \sigma_1^2\theta_1 \frac{2^\alpha}{\pi^{d/2+1}} \Gamma(\frac{\alpha+2}{2}) \Gamma(\frac{\alpha+d}{2}) \sin\left(\frac{\alpha\pi}{2}\right) u^{-(\alpha+d)} \\ &\quad + \sigma_1^2\theta_1^2 \frac{2^{2\alpha-1}}{\pi^{d/2+1}} \Gamma(\alpha+1) \Gamma(\frac{2\alpha+d}{2}) \sin(\alpha\pi) u^{-(2\alpha+d)} + O\big(u^{-(3\alpha+d)}\big), \end{aligned} \tag{4.10}$$

as $u \to \infty$. Moreover, Eq. (4.5) can be expressed as

$$\int_c^\infty u^{d-1} \left\{ \frac{f_1(u) - f_2(u)}{f_1(u)} \right\}^2 du < \infty, \tag{4.11}$$

where f_1 and f_2 are isotropic spectral densities corresponding to P_1 and P_2. Assuming that $\sigma_1^2\theta_1 = \sigma_2^2\theta_2$, then

$$\begin{aligned}
&\lim_{u\to\infty} u^{\alpha+d} f_1(u) \\
=&\lim_{u\to\infty} u^{\alpha+d}\sigma_1^2 \frac{\Gamma(d/2)}{2\pi^{d/2}} u^{1-d} \frac{\theta_1}{\pi\Gamma(d/2)} \Gamma(\frac{\alpha+2}{2})\Gamma(\frac{\alpha+d}{2}) \sin\left(\frac{\alpha\pi}{2}\right)(\frac{2}{u})^{\alpha+1} \\
=&\sigma_1^2\theta_1 \frac{2^\alpha}{\pi^{d/2+1}} \Gamma(\frac{\alpha+2}{2})\Gamma(\frac{\alpha+d}{2}) \sin\left(\frac{\alpha\pi}{2}\right).
\end{aligned}$$

And

$$\begin{aligned}
&\left\{\frac{f_1(u)-f_2(u)}{f_1(u)}\right\}^2 \\
=&\left\{\frac{\sigma_1^2 f_{R,1}(u) - \sigma_2^2 f_{R,2}(u)}{\sigma_1^2 f_{R,1}(u)}\right\}^2 \\
\lesssim&\left(2^{\alpha-1}\frac{(\sigma_2^2\theta_2^2 - \sigma_1^2\theta_1^2)}{\sigma_1^2\theta_1} \frac{\Gamma(\alpha+1)\Gamma(\frac{2\alpha+d}{2})\sin(\alpha\pi)}{\Gamma(\frac{\alpha+2}{2})\Gamma(\frac{\alpha+d}{2})\sin\left(\frac{\alpha\pi}{2}\right)} u^{-\alpha}\right)^2,
\end{aligned} \tag{4.12}$$

for all u big enough. So the integral in Eq. (4.11) is finite when $\alpha \in \left(\frac{d}{2}, 2\right) \setminus \{1\}$. Therefore, the two measures are equivalent.

" $\Rightarrow$ " If $\sigma_1^2\theta_1 \neq \sigma_2^2\theta_2$, let $\sigma_0^2 = \sigma_2^2\theta_2/\theta_1$. Then $\sigma_0^2\theta_1 = \sigma_2^2\theta_2$ and the two powered exponential covariance functions $C(\mathbf{t}; \sigma_0^2, \theta_1, \alpha)$ and $C(\mathbf{t}; \sigma_2^2, \theta_2, \alpha)$ define two equivalent measures. We only need to show that for any countable infinite subset $T \subset \Omega$, $C(\mathbf{t}; \sigma_0^2, \theta_1, \alpha)$ and $C(\mathbf{t}; \sigma_1^2, \theta_1, \alpha)$ define two orthogonal Gaussian measures on $\sigma\{X(t), t \in T\}$.

For Gaussian random variables $\{X(t), t \in T\}$, we re-index them as $\{X(n), n \in \mathbb{N}\}$. Obviously, any finite many of them has a positive definite covariance matrix with respect to both Gaussian measures.

Applying Gram-Schmidt orthogonalization to $\{X(n), n \in \mathbb{N}\}$ with the inner product induced by $C(\mathbf{t}; \sigma_0^2, \theta_1, \alpha)$, we get $\{\xi_i\}_{i=1}^\infty$ which is a sequence of independent standard Gaussian random variables with respect to the Gaussian measure induced by $C(\mathbf{t}; \sigma_0^2, \theta_1, \alpha)$.

And $\{\xi_i\}_{i=1}^\infty$ is a sequence of independent Gaussian random variables with mean zero and variance σ_1^2/σ_0^2 with respect to the Gaussian measure induced by $C(\mathbf{t}; \sigma_1^2, \theta_1, \alpha)$.

Let $\epsilon \triangleq \frac{|\sigma_1^2/\sigma_0^2-1|}{2}$, then by law of large numbers, we have

$$\lim_{n\to\infty} P_{\sigma_0^2,\theta_1,\alpha}\left(\left|\sum_{k=1}^{n}\xi_k^2/n-1\right|>\epsilon\right)=0. \tag{4.13}$$

Meanwhile,

$$\begin{aligned}
1 &= \lim_{n\to\infty} P_{\sigma_1^2,\theta_1,\alpha}\left(\left|\sum_{k=1}^{n}\xi_k^2/n-\frac{\sigma_1^2}{\sigma_0^2}\right|<\left|\sigma_1^2/\sigma_0^2-1\right|-\epsilon\right)\\
&= \lim_{n\to\infty} P_{\sigma_1^2,\theta_1,\alpha}\left(\left|\sigma_1^2/\sigma_0^2-1\right|-\left|\sum_{k=1}^{n}\xi_k^2/n-\frac{\sigma_1^2}{\sigma_0^2}\right|>\epsilon\right)\\
&\leq \lim_{n\to\infty} P_{\sigma_1^2,\theta_1,\alpha}\left(\left|\sum_{k=1}^{n}\xi_k^2/n-1\right|>\epsilon\right).
\end{aligned}$$

Then

$$\lim_{n\to\infty} P_{\sigma_1^2,\theta_1,\alpha}\left(\left|\sum_{k=1}^{n}\xi_k^2/n-1\right|>\epsilon\right)=1. \tag{4.14}$$

Finally, combining Eq. (4.13) and Eq. (4.14), we have $P_{\sigma_0^2,\theta_1,\alpha}\perp P_{\sigma_1^2,\theta_1,\alpha}$ on $\sigma\{X(t),t\in T\}$, thus on $\sigma\{X(t),t\in\Omega\}$. □

Remark 4.1. *For d-dimensional isotropic Ornstein-Uhlenbeck process with covariance function*

$$V_d(t,s;\sigma^2,\theta)=\sigma^2 e^{-\theta\left[\sum_{i=1}^{d}(t_i-s_i)^2\right]^{\frac{1}{2}}},$$

setting the smoothness parameter ν as $\frac{1}{2}$ in the Matérn model, we have that for $d\leq 3$, $V_d(\cdot;\sigma_1^2,\theta_1)$ and $V_d(\cdot;\sigma_2^2,\theta_2)$ induces two Gaussian measures that are equivalent if and only if $\sigma_1^2\theta_1=\sigma_2^2\theta_2$ according to Theorem 2 in Zhang (2004).

With Lemma 4.1, we see that σ and θ cannot be consistently estimated when $\alpha\in(d/2,2)$ under fixed-domain asymptotics (see Zhang 2004, for a detailed explanation). However, as we will show in the next section, the MLE of $\sigma^2\theta$ is consistent and asymptotic normal.

4.3 Asymptotic properties of the MLE

Let the observation of Gaussian field $X(t)$ be

$$\mathbf{X}_n=\{X(\mathbf{t}_1),X(\mathbf{t}_2),\cdots,X(\mathbf{t}_n)\}', \tag{4.15}$$

where $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_n \in [0, T]^d$ are distinct locations, and $[0, T]^d \subset \mathbb{R}^d$ is the fixed cube with side T.

Theorem 4.1. *Let $\{X(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d\}$, $d \in \{1, 2, 3\}$, be stationary isotropic Gaussian field with mean 0 and covariance function $C(\mathbf{t}) = \sigma_0^2 e^{-\theta_0 ||\mathbf{t}||^\alpha}$ where σ_0^2 and θ_0 are unknown and α is known. Let D_n, $n = 1, 2, \cdots$, be an increasing (nested) sequence of finite subsets of $[0, T]^d \subset \mathbb{R}^d$, and $L_n(\sigma^2, \theta)$ be the likelihood function when the process is observed at locations in D_n. For any fixed $\theta_1 > 0$, let $\hat{\sigma}_n^2$ maximize $L_n(\sigma^2, \theta_1)$. Then when $\alpha \in (\frac{d}{2}, 2)$, $\hat{\sigma}_n^2 \theta_1 \to \sigma_0^2 \theta_0$, with P_0 probability 1, where P_0 is the Gaussian measure defined by the powered exponential covariance function corresponding to parameter values σ_0^2, θ_0 and α. Namely,*

$$\frac{\mathbf{X}_n' \Gamma_{(n,\theta_1)}^{-1} \mathbf{X}_n}{n} \theta_1 \xrightarrow[n \to \infty]{P_0 - a.s.} \sigma_0^2 \theta_0, \tag{4.16}$$

where $\Gamma_{(n,\theta_0)}$ is the true correlation matrix.

Proof. The proof follows the same arguments of the proof of Theorem 3 in Zhang (2004). □

To derive the asymptotic normality, the key is to control the error between quadratic forms with the misspecified correlation matrix in Eq. (4.16) and the true one respectively. Before stating the main result on the asymptotic normality, we need to introduce several preliminaries that will be used in the main proof.

Let $\sigma_1^2 = \sigma_0^2 \theta_0 / \theta_1$, and P_1 is the Gaussian measure corresponding to mean 0 and the covariance function $C(\mathbf{t}; \sigma_1^2, \theta_1, \alpha)$ in the following. First, we observe that by the eigendecomposition

$$\big(\sigma_0^2 \Gamma_{(n,\theta_0)}\big)^{-1/2} \big(\sigma_1^2 \Gamma_{(n,\theta_1)}\big) \big(\sigma_0^2 \Gamma_{(n,\theta_0)}\big)^{-1/2} = U_n \Lambda_n U_n', \tag{4.17}$$

where U_n is an orthogonal matrix consisting of the eigenvectors and

$$\Lambda_n = \operatorname{diag}(\lambda_{1,n}, \cdots, \lambda_{n,n})$$

is a diagonal matrix with the eigenvalues $\{\lambda_{k,n}\}_{k=1}^n$. Define

$$\mathbf{Y}_n = U_n' \big(\sigma_0^2 \Gamma_{(n,\theta_0)}\big)^{-1/2} \mathbf{X}_n, \tag{4.18}$$

then

$$\frac{1}{\sqrt{n}}\Big(\mathbf{X}_n'\big(\sigma_1^2\Gamma_{(n,\theta_1)}\big)^{-1}\mathbf{X}_n - \mathbf{X}_n'\big(\sigma_0^2\Gamma_{(n,\theta_0)}\big)^{-1}\mathbf{X}_n\Big) = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\big(\lambda_{k,n}^{-1}-1\big)Y_{k,n}^2,$$

where $(Y_{1,n},\cdots,Y_{n,n})' = \mathbf{Y}_n \sim \mathcal{N}(\mathbf{0}, I_n)$ under P_0, and $\mathbf{Y}_n \sim \mathcal{N}(\mathbf{0}, \Lambda_n)$ under P_1.

Second, we plan to represent $\mathbf{Y}_n$ in the frequency domain. The spectral representation theorem (see, e.g., Adler and Taylor 2007, Theorem 2.4.2) says $X(\mathbf{t})$ has the following spectral representation

$$X(\mathbf{t}) \overset{\text{fdd}}{=} \int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}'\mathbf{t}} M(d\boldsymbol{\omega}), \tag{4.19}$$

where M is a Gaussian random measure on $\mathbb{R}^d$ whose control measure has the Radon-Nikodym derivative $f_{\sigma,\theta}$ with respect to the Lebesgue measure. The $\overset{\text{fdd}}{=}$ means that both sides in Eq. (4.19) have the same finite-dimensional distributions. Subsequently, we introduce two isomorphic Hilbert spaces. Define

$$H^0 \triangleq \left\{\sum_{i=1}^{n} r_i X(\mathbf{t}_i) : r_1,\cdots,r_n \in \mathbb{R}, \mathbf{t}_1,\cdots,\mathbf{t}_n \in [0,T]^d\right\}, \tag{4.20}$$

and

$$L^0 \triangleq \left\{\sum_{i=1}^{n} r_i e^{i\boldsymbol{\omega}'\mathbf{t}_i} : r_1,\cdots,r_n \in \mathbb{R}, \mathbf{t}_1,\cdots,\mathbf{t}_n \in [0,T]^d\right\}. \tag{4.21}$$

Eq. (4.19) implies that

$$\operatorname{cov}\left\{\sum_{j=1}^{n} a_j X(\mathbf{t}_j), \sum_{k=1}^{m} b_k X(\mathbf{t}_k)\right\} = \left(\sum_{j=1}^{n} a_j \exp\{i\boldsymbol{\omega}'\mathbf{t}_j\}, \sum_{k=1}^{m} b_k \exp\{i\boldsymbol{\omega}'\mathbf{t}_k\}\right)_{f_{\sigma,\theta}}, \tag{4.22}$$

where the inner product $(\cdot,\cdot)_{f_{\sigma,\theta}}$ is defined as $(\phi,\psi)_{f_{\sigma,\theta}} \triangleq \int_{\mathbb{R}^d} \phi(\boldsymbol{\omega})\overline{\psi(\boldsymbol{\omega})}d\boldsymbol{\omega}$. Therefore, two Hilbert spaces $H(f_{\sigma,\theta})$, as the closure of H^0 with respect to the inner product $\operatorname{cov}\{\cdot,\cdot\}$, and $L(f_{\sigma,\theta})$, as the closure of L^0 with respect to the inner product $(\cdot,\cdot)_{f_{\sigma,\theta}}$, are isomorphic; (see, e.g., M. L. Stein 1999a, chap. 2.6; Ibragimov and Rozanov 1978, chap. 3.1.3). Define

$$(\psi_1,\cdots,\psi_n)' = U_n'\big(\sigma_0^2\Gamma_{(n,\theta_0)}\big)^{-1/2}(\exp\{i\boldsymbol{\omega}'\mathbf{t}_1\},\cdots,\exp\{i\boldsymbol{\omega}'\mathbf{t}_n\})'. \tag{4.23}$$

Combining Eqs. (4.18) and (4.22), we see that for $j,k = 1,\cdots,n$,

$$\begin{aligned}(\psi_j,\psi_k)_{f_{\sigma_0,\theta_0}} &= \delta_{j,k},\\ (\psi_j,\psi_k)_{f_{\sigma_1,\theta_1}} &= \lambda_{j,n}\delta_{j,k},\end{aligned} \tag{4.24}$$

where $\delta_{j,k} = 1$ if $j = k$ and is 0 otherwise.

Then, to bound $\{\lambda_{k,n}\}_{k=1}^{n}$, we first introduce the entropy distance between two equivalent Gaussian measures (see M. L. Stein (1999a) and Ibragimov and Rozanov (1978) for a fuller exposition about the entropy distance). Lemma 4.1 implies $P_0 \equiv P_1$, and the entropy distance r between P_1 and P_0 is defined as

$$r \triangleq -\left[\mathrm{E}_0 \ln\left(\frac{P_1}{P_0}\right) + \mathrm{E}_1 \ln\left(\frac{P_0}{P_1}\right)\right], \tag{4.25}$$

where $\frac{P_0}{P_1}$ and $\frac{P_1}{P_0}$ are the Radon-Nikodym derivative on the σ-algebra generated by the Gaussian process X. Moreover, the conditional expectation of $\frac{P_1}{P_0}$ given $\mathbf{X}_n$ is

$$\begin{aligned} p_n &\triangleq \mathrm{E}_0\left(\frac{P_1}{P_0}\middle| X_n\right) \\ &= \frac{f_1(X_n)}{f_0(X_n)} \\ &= \left(\frac{\left|\left(\sigma_0^2\Gamma_{(n,\theta_0)}\right)\right|}{\left|\left(\sigma_1^2\Gamma_{(n,\theta_1)}\right)\right|}\right)^{1/2} e^{-\frac{1}{2}\mathbf{X}_n'\left(\sigma_1^2\Gamma_{(n,\theta_1)}\right)^{-1}\mathbf{X}_n + \frac{1}{2}\mathbf{X}_n'\left(\sigma_0^2\Gamma_{(n,\theta_0)}\right)^{-1}\mathbf{X}_n}, \end{aligned} \tag{4.26}$$

where $f_1(x)$ and $f_0(x)$ are the probability density functions of $\mathbf{X}_n$ with respect to P_1 and P_0 respectively. Define r_n as

$$\begin{aligned} r_n &\triangleq -\mathrm{E}_0 \ln(p_n) + \mathrm{E}_1 \ln(p_n) \\ &= \frac{1}{2}\mathrm{E}_0\, \mathbf{X}_n'\Big(\left(\sigma_1^2\Gamma_{(n,\theta_1)}\right)^{-1} - \left(\sigma_0^2\Gamma_{(n,\theta_0)}\right)^{-1}\Big)\mathbf{X}_n \\ &\quad + \frac{1}{2}\mathrm{E}_1\, \mathbf{X}_n'\Big(\left(\sigma_0^2\Gamma_{(n,\theta_0)}\right)^{-1} - \left(\sigma_1^2\Gamma_{(n,\theta_1)}\right)^{-1}\Big)\mathbf{X}_n \\ &= \frac{1}{2}\left[\mathrm{tr}\left(\left(\sigma_1^2\Gamma_{(n,\theta_1)}\right)^{-1}\left(\sigma_0^2\Gamma_{(n,\theta_0)}\right)\right) + \mathrm{tr}\left(\left(\sigma_0^2\Gamma_{(n,\theta_0)}\right)^{-1}\left(\sigma_1^2\Gamma_{(n,\theta_1)}\right)\right) - 2n\right] \\ &= \frac{1}{2}\sum_{k=1}^{n}\left(\frac{1}{\lambda_{k,n}} + \lambda_{k,n} - 2\right), \end{aligned} \tag{4.27}$$

where $\{\lambda_{k,n}\}_{k=1}^{n}$ are defined in Eq. (4.17).

Lemma 4.2. *There exist two positive constants $0 < C_i \leq C_s < \infty$ such that $C_i \leq \lambda_{k,n} \leq C_s$ for all n and $1 \leq k \leq n$.*

Proof. Theorem 4 in M. L. Stein (1999a, p.117) implies $0 \le r < +\infty$ in Eq. (4.25). And by Jensen's inequality, $r_n \le r$ for all n, which implies that

$$\frac{1}{2}\left(\frac{1}{\lambda_{k,n}} + \lambda_{k,n} - 2\right) \le r,$$

for all n and $1 \le k \le n$. Let

$$C_i = \frac{(2+2r) - \sqrt{(2+2r)^2 - 4}}{2}, \qquad C_s = \frac{(2+2r) + \sqrt{(2+2r)^2 - 4}}{2}, \tag{4.28}$$

then we have

$$C_i \le \lambda_{k,n} \le C_s.$$

This proves Lemma 4.2. □

Finally, we describe properties of the spectral density function $f_{\sigma,\theta}(\boldsymbol{\omega})$ in the following lemma.

Lemma 4.3. *The spectral density $f_{\sigma,\theta}(\boldsymbol{\omega})$ defined in Eq. (4.2) is bounded, uniformly continuous and strictly positive on $\mathbb{R}^d$. Furthermore, $f_{\sigma,\theta}(\boldsymbol{\omega}) \asymp \|\boldsymbol{\omega}\|^{-(\alpha+d)}$ as $\|\boldsymbol{\omega}\| \to \infty$.*

Proof. Since $\int_{\mathbb{R}^d} C(\mathbf{t})d\mathbf{t} < \infty$, $f_{\sigma,\theta}(\boldsymbol{\omega})$ is bounded and uniformly continuous by the properties of Fourier transform. Furthermore, based on Formula (7.5) in Khoshnevisan, Xiao, and Zhong (2003), for $\boldsymbol{\omega} \ne 0$,

$$f_{1,1}(\boldsymbol{\omega}) = \|\boldsymbol{\omega}\|^{-d} \int_0^\infty \nu\left(\frac{s}{\|\boldsymbol{\omega}\|^2}\right) g^{(\alpha/2)}(s)ds, \tag{4.29}$$

where the function ν is defined as

$$\nu(s) = (4\pi s)^{-d/2} e^{-\frac{1}{4s}},$$

$g^{(\alpha/2)}(s)$ is the density function of the random variable $\tau(1)$, and where $\tau = \tau(t), t \ge 0$ is a stable subordinator of index $\frac{\alpha}{2}$. Finally,

$$f_{1,1}(\mathbf{0}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} C(\mathbf{t})d\mathbf{t} > 0.$$

So we have $f_{1,1}(\boldsymbol{\omega}) > 0$. The fact that $f_{\sigma,\theta}(\boldsymbol{\omega}) \asymp \|\boldsymbol{\omega}\|^{-(\alpha+d)}$ as $\|\boldsymbol{\omega}\| \to \infty$ has been verified in the proof of Lemma 4.1. This completes the proof. □

Theorem 4.2. *With the same notation and assumptions as in Theorem 4.1, we have*

$$\sqrt{n}\big(\hat{\sigma}_n^2\theta_1 - \sigma_0^2\theta_0\big) \xrightarrow{d} \mathcal{N}\big(0, 2(\sigma_0^2\theta_0)^2\big), \tag{4.30}$$

with respect to P_0.

Proof. Note that Ying (1991) has proven the theorem when $\alpha = 1$ and $d = 1$. Therefore, we only consider the case $\alpha \in \big(\frac{d}{2}, 2\big) \setminus \{1\}$, and we follow the arguments from Wang (2010) in the following. Splitting $(\hat{\sigma}_n^2\theta_1 - \sigma_0^2\theta_0)$ into two parts, we have

$$\begin{aligned}\sqrt{n}\big(\hat{\sigma}_n^2\theta_1 - \sigma_0^2\theta_0\big) = \frac{\sigma_0^2\theta_0}{\sqrt{n}}\Big(\mathbf{X}_n'\big(\sigma_1^2\Gamma_{(n,\theta_1)}\big)^{-1}\mathbf{X}_n - \mathbf{X}_n'\big(\sigma_0^2\Gamma_{(n,\theta_0)}\big)^{-1}\mathbf{X}_n\Big) \\ + \frac{\sigma_0^2\theta_0}{\sqrt{n}}\Big(\mathbf{X}_n'\big(\sigma_0^2\Gamma_{(n,\theta_0)}\big)^{-1}\mathbf{X}_n - n\Big).\end{aligned} \tag{4.31}$$

Since $\big(\sigma_0^2\Gamma_{(n,\theta_0)}\big)^{-1/2}\mathbf{X}_n$ consists of i.i.d. standard normal variables with respect to P_0, $\mathbf{X}_n'\big(\sigma_0^2\Gamma_{(n,\theta_0)}\big)^{-1}\mathbf{X}_n$ is the sum of i.i.d. variables having χ_1^2 distribution. Hence, the Lindeberg-Feller theorem implies

$$\frac{\sigma_0^2\theta_0}{\sqrt{n}}\Big(\mathbf{X}_n'\big(\sigma_0^2\Gamma_{(n,\theta_0)}\big)^{-1}\mathbf{X}_n - n\Big) \xrightarrow{\mathrm{d}} \mathcal{N}\big(0, 2(\sigma_0^2\theta_0)^2\big). \tag{4.32}$$

Thus to prove Theorem 4.2, it suffices to show that

$$\frac{\sigma_0^2\theta_0}{\sqrt{n}}\Big(\mathbf{X}_n'\big(\sigma_1^2\Gamma_{(n,\theta_1)}\big)^{-1}\mathbf{X}_n - \mathbf{X}_n'\big(\sigma_0^2\Gamma_{(n,\theta_0)}\big)^{-1}\mathbf{X}_n\Big) \xrightarrow{P_0} 0. \tag{4.33}$$

Let $m = \big\lfloor\frac{\alpha+d}{2}\big\rfloor + 1$ and $\kappa = \frac{\alpha+d}{4m} \in (0, \frac{1}{2})$, where $\lfloor\cdot\rfloor$ denotes the greatest integer function. Define the integrable function c_0 and its Fourier transform ξ_0 as

$$c_0(\mathbf{x}) = \|\mathbf{x}\|^{\kappa-d}\mathcal{I}(\|\mathbf{x}\| \le 1), \quad \mathbf{x} \in \mathbb{R}^d, \tag{4.34}$$

$$\xi_0(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} e^{-i\mathbf{x}'\boldsymbol{\omega}}c_0(\mathbf{x})d\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^d. \tag{4.35}$$

It follows from Lemma 4.5 that $\xi_0 : \mathbb{R}^d \to \mathbb{R}$ is a continuous, isotropic and strictly positive function and $\xi_0(\boldsymbol{\omega}) \asymp \|\boldsymbol{\omega}\|^{-\kappa}$ as $\|\boldsymbol{\omega}\| \to \infty$.

Let $c_1 = c_0 * \cdots * c_0$ denote the $2m$-fold convolution of the function c_0 with itself. Then $\mathrm{supp}(c_1) \subset \{\mathbf{x} : \|\mathbf{x}\| \le 2m\}$ and

$$\xi_1(\boldsymbol{\omega}) = \xi_0(\boldsymbol{\omega})^{2m}. \tag{4.36}$$

It follows from Lemma 4.3 that there exist positive constants $0 < c_{\xi_1} \le C_{\xi_1} < \infty$ such that for all $\boldsymbol{\omega} \in \mathbb{R}^d$,

$$c_{\xi_1} \le \frac{f_{\sigma_0,\theta_0}(\boldsymbol{\omega})}{\xi_1(\boldsymbol{\omega})^2} \le C_{\xi_1}. \tag{4.37}$$

Define

$$\eta(\boldsymbol{\omega}) = \frac{f_{\sigma_1,\theta_1}(\boldsymbol{\omega}) - f_{\sigma_0,\theta_0}(\boldsymbol{\omega})}{\xi_1(\boldsymbol{\omega})^2}. \tag{4.38}$$

It follows from the proof of Lemma 4.1 that there exists a constant C_η such that for all $\boldsymbol{\omega} \in \mathbb{R}^d$,

$$|\eta(\boldsymbol{\omega})| \le \frac{C_\eta}{1 + \|\boldsymbol{\omega}\|^\alpha}, \tag{4.39}$$

and $\eta : \mathbb{R}^d \to \mathbb{R}$ is square integrable when $\alpha \in \left(\frac{d}{2}, 2\right) \setminus \{1\}$. From the theory of the Fourier transform of $L^2(\mathbb{R}^d)$ functions, there exists a square integrable function $g : \mathbb{R}^d \to \mathbb{R}$ such that

$$\int_{\mathbb{R}^d} (\eta(\boldsymbol{\omega}) - \hat{g}_k(\boldsymbol{\omega}))^2 d\boldsymbol{\omega} \xrightarrow{k \to \infty} 0, \tag{4.40}$$

where

$$\hat{g}_k(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} e^{-i\boldsymbol{\omega}'\mathbf{x}} g(\mathbf{x}) \mathcal{I}\{\|\mathbf{x}\|_{\max} \le k\} d\mathbf{x}, \quad k \in \mathbb{N}. \tag{4.41}$$

In order to illustrate the following Lemma, we need to introduce the approximate identity $e_n(\mathbf{x})$ defined by Formula (35) of Wang and Loh (2011). Let $\{\epsilon_n\}_{n=1}^\infty$ be a positive sequence such that $\epsilon_n \to 0$ as $n \to \infty$. Define

$$e_n(\mathbf{x}) = \frac{1}{C_e \epsilon_n^d} \tilde{c}_1\left(\frac{\mathbf{x}}{\epsilon_n}\right), \quad \mathbf{x} \in \mathbb{R}^d, \tag{4.42}$$

and

$$\tilde{\xi}_1(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} e^{-i\mathbf{x}'\boldsymbol{\omega}} \tilde{c}_1(\mathbf{x}) d\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^d, \tag{4.43}$$

where $C_e = \int_{\mathbb{R}^d} \tilde{c}_1(\mathbf{x}) d\mathbf{x}$ and $\tilde{c}_1 = \tilde{c}_0 * \cdots * \tilde{c}_0$, the $2m_a$-fold convolution of $\tilde{c}_0$, with $\tilde{c}_0(\mathbf{x}) = \|\mathbf{x}\|^{\frac{a+d}{4m_a} - d} \mathcal{I}\{\|\mathbf{x}\| \le 1\}$ and $m_a = \left\lfloor \frac{a+d}{2} \right\rfloor + 1$. Here a is an arbitrary positive constant. For the Fourier transform of e_n, we have

$$\hat{e}_n(\boldsymbol{\omega}) = \frac{\tilde{\xi}_1(\epsilon_n \boldsymbol{\omega})}{C_e}. \tag{4.44}$$

Lemma 4.5 implies that there exists a constant $C_{\hat{e}}$ (independent of $\boldsymbol{\omega}$ and n) such that for all $\boldsymbol{\omega} \in \mathbb{R}^d$,

$$|\hat{e}_n(\boldsymbol{\omega})| \leq \frac{C_{\hat{e}}}{(1+\epsilon_n\|\boldsymbol{\omega}\|)^{(a+d)/2}}. \tag{4.45}$$

Lemma 4.4. *Let e_n and g be as in Eqs. (4.41) and (4.42) respectively and β_0 is any constant satisfying $0 < \beta_0 < \min\{2, 2\alpha - d\}$. Then there exists a constant C_{β_0} such that*

$$\int_{\mathbb{R}^d} |e_n * g(\mathbf{x}) - g(\mathbf{x})|^2 d\mathbf{x} \leq C_{\beta_0}\epsilon_n^{\beta_0}. \tag{4.46}$$

Proof. The Plancherel theorem implies that for $\mathbf{y} \in \mathbb{R}^d$,

$$\begin{aligned}\int_{\mathbb{R}^d} |g(\mathbf{x}-\mathbf{y}) - g(\mathbf{x})|^2 d\mathbf{x} &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left|e^{-i\boldsymbol{\omega}'\mathbf{y}}\eta(\boldsymbol{\omega}) - \eta(\boldsymbol{\omega})\right|^2 d\boldsymbol{\omega} \\ &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left|(e^{-i\boldsymbol{\omega}'\mathbf{y}} - 1)\eta(\boldsymbol{\omega})\right|^2 d\boldsymbol{\omega} \\ &\leq \frac{2^{2-\beta_0}\|\mathbf{y}\|^{\beta_0}}{(2\pi)^d}\int_{\mathbb{R}^d}\|\boldsymbol{\omega}\|^{\beta_0}|\eta(\boldsymbol{\omega})|^2 d\boldsymbol{\omega}.\end{aligned} \tag{4.47}$$

Hence

$$\begin{aligned}&\left[\int_{\mathbb{R}^d}|e_n * g(\mathbf{x}) - g(\mathbf{x})|^2 d\mathbf{x}\right]^{1/2} \\ =&\left[\int_{\mathbb{R}^d}\left|\int_{\|\mathbf{y}\|\leq 2m_a\epsilon_n}[g(\mathbf{x}-\mathbf{y}) - g(\mathbf{x})]e_n(\mathbf{y})d\mathbf{y}\right|^2 d\mathbf{x}\right]^{1/2} \\ \leq&\int_{\|\mathbf{y}\|\leq 2m_a\epsilon_n}\left[\int_{\mathbb{R}^d}|g(\mathbf{x}-\mathbf{y}) - g(\mathbf{x})|^2 d\mathbf{x}\right]^{1/2} e_n(\mathbf{y})d\mathbf{y} \\ \leq&\frac{2^{(2-\beta_0)/2}(2m_a\epsilon_n)^{\beta_0/2}}{(2\pi)^{d/2}}\left[\int_{\mathbb{R}^d}\|\boldsymbol{\omega}\|^{\beta_0}|\eta(\boldsymbol{\omega})|^2 d\boldsymbol{\omega}\right]^{1/2} \\ \leq&\frac{C_\eta 2^{(2-\beta_0)/2}(2m_a\epsilon_n)^{\beta_0/2}}{(2\pi)^{d/2}}\left[\int_{\mathbb{R}^d}\frac{\|\boldsymbol{\omega}\|^{\beta_0}}{(1+\|\boldsymbol{\omega}\|^{\alpha})^2}d\boldsymbol{\omega}\right]^{1/2}.\end{aligned} \tag{4.48}$$

This proves Lemma 4.4. □

Let $\mathbf{x}, \mathbf{y} \in [0, T]^d$, and observing that $\mathrm{supp}(c_1) \subset [-2m, 2m]^d$, we obtain

$$
\begin{aligned}
b(\mathbf{x}, \mathbf{y}) &\triangleq \mathrm{E}_{f_{\sigma_1,\theta_1}}[X(\mathbf{x})X(\mathbf{y})] - \mathrm{E}_{f_{\sigma_0,\theta_0}}[X(\mathbf{x})X(\mathbf{y})] \\
&= \int_{\mathbb{R}^d} e^{i(\mathbf{x}-\mathbf{y})'\boldsymbol{\omega}}[f_{\sigma_1,\theta_1}(\boldsymbol{\omega}) - f_{\sigma_0,\theta_0}(\boldsymbol{\omega})]d\boldsymbol{\omega} \\
&= \int_{\mathbb{R}^d} e^{i(\mathbf{x}-\mathbf{y})'\boldsymbol{\omega}}\eta(\boldsymbol{\omega})\xi_1(\boldsymbol{\omega})^2 d\boldsymbol{\omega} \\
&= (2\pi)^d \int_{\mathbb{R}^{2d}} g(\mathbf{s}-\mathbf{t})c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})d\mathbf{s}d\mathbf{t} \\
&= (2\pi)^d \int_{\mathbb{R}^{2d}} e_n * g(\mathbf{s}-\mathbf{t})c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})d\mathbf{s}d\mathbf{t} \\
&\quad + (2\pi)^d \int_{\mathbb{R}^{2d}} [g(\mathbf{s}-\mathbf{t}) - e_n * g(\mathbf{s}-\mathbf{t})]c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})d\mathbf{s}d\mathbf{t} \\
&= (2\pi)^d \int_{\mathbb{R}^{2d}} e_n * g(\mathbf{s}-\mathbf{t})c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})d\mathbf{s}d\mathbf{t} \\
&\quad + (2\pi)^d \int_{\mathbb{R}^{2d}} h_n^*(\mathbf{s}, \mathbf{t})c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})d\mathbf{s}d\mathbf{t},
\end{aligned}
\tag{4.49}
$$

where

$$
h_n^*(\mathbf{s}, \mathbf{t}) = [g(\mathbf{s}-\mathbf{t}) - e_n * g(\mathbf{s}-\mathbf{t})]\mathcal{I}\{\|\mathbf{s}+\mathbf{t}\|_{\max} \leq 4m + 2T\}, \quad \forall\, \mathbf{s}, \mathbf{t} \in \mathbb{R}^d.
$$

Let $\eta_n^* : \mathbb{R}^d \to \mathbb{C}$ denote the $L^2(\mathbb{R}^d)$ Fourier transform of $g - e_n * g$. This implies that

$$
\int_{\mathbb{R}^d} \left(\eta_n^*(\boldsymbol{\omega}) - \hat{g}_{n,k}^*(\boldsymbol{\omega})\right)^2 d\boldsymbol{\omega} \xrightarrow{k\to\infty} 0, \tag{4.50}
$$

where

$$
\hat{g}_{n,k}^*(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} e^{-i\boldsymbol{\omega}'\mathbf{x}}[g(\mathbf{x}) - e_n * g(\mathbf{x})]\mathcal{I}\{|\mathbf{x}|_{\max} \leq k\}d\mathbf{x}, \quad k \in \mathbb{N}. \tag{4.51}
$$

Thus as in (24) in Wang and Loh (2011), we have

$$
\begin{aligned}
&(2\pi)^d \int_{\mathbb{R}^{2d}} h_n^*(\mathbf{s}, \mathbf{t})c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})d\mathbf{s}d\mathbf{t} \\
=&(2\pi)^{-d} \int_{\mathbb{R}^{2d}} e^{i(\boldsymbol{\omega}'\mathbf{x}-\boldsymbol{\nu}'\mathbf{y})}\eta_n^*\left(\frac{\boldsymbol{\omega}+\boldsymbol{\nu}}{2}\right)\kappa\left(\frac{\boldsymbol{\omega}-\boldsymbol{\nu}}{2}\right)\xi_1(\boldsymbol{\omega})\xi_1(\boldsymbol{\nu})d\boldsymbol{\omega}d\boldsymbol{\nu},
\end{aligned}
\tag{4.52}
$$

where $\kappa(\mathbf{x}) = 2^{-d}\int_{\mathbb{R}^d} e^{-i\mathbf{x}'\mathbf{t}}\mathcal{I}\{\|\mathbf{t}\|_{\max} \leq 4m + 2T\}d\mathbf{t}, \quad \mathbf{x} \in \mathbb{R}^d$. We observe that since $\mathcal{I}\{\|\mathbf{t}\|_{\max} \leq 4m + 2T\} \in L_1 \cap L_2$, κ is continuous and $\kappa \in L_2(\mathbb{R}^d)$. Now we define

$$
h_n^{**}(\mathbf{s}, \mathbf{t}) = \int_{\|\mathbf{u}\|_{\max} \leq 2m+2m_a+T} e_n(\mathbf{s}-\mathbf{u})g(\mathbf{u}-\mathbf{t})d\mathbf{u}, \quad \forall\, \mathbf{s}, \mathbf{t} \in \mathbb{R}^d.
$$

Then the function $h_n^{**}:\mathbb{R}^d\to\mathbb{C}$ is square-integrable and

$$\begin{aligned}
&(2\pi)^d\int_{\mathbb{R}^{2d}} e_n * g(\mathbf{s}-\mathbf{t})c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})d\mathbf{s}d\mathbf{t}\\
=&(2\pi)^d\int_{\mathbb{R}^{2d}} h_n^{**}(\mathbf{s},\mathbf{t})c_1(\mathbf{x}-\mathbf{s})c_1(\mathbf{y}-\mathbf{t})d\mathbf{s}d\mathbf{t}\\
=&(2\pi)^{-d}\int_{\mathbb{R}^{2d}} e^{i(\boldsymbol{\omega}'\mathbf{x}-\boldsymbol{\nu}'\mathbf{y})}\xi_1(\boldsymbol{\omega})\xi_1(\boldsymbol{\nu})\\
&\quad\times\left(\int_{\|\mathbf{u}\|_{\max}\le 2m+2m_a+T} e^{-i(\boldsymbol{\omega}'\mathbf{u}-\boldsymbol{\nu}'\mathbf{u})}\hat{e}_n(\boldsymbol{\omega})\eta(\boldsymbol{\nu})d\mathbf{u}\right)d\boldsymbol{\omega}d\boldsymbol{\nu}.
\end{aligned}\tag{4.53}$$

It follows from Eqs. (4.52) and (4.53), that for $\mathbf{x},\mathbf{y}\in[0,T]^d$,

$$\begin{aligned}
b(\mathbf{x},\mathbf{y}) =& (2\pi)^{-d}\int_{\mathbb{R}^{2d}} e^{i(\boldsymbol{\omega}'\mathbf{x}-\boldsymbol{\nu}'\mathbf{y})}\eta_n^*\left(\frac{\boldsymbol{\omega}+\boldsymbol{\nu}}{2}\right)\kappa\left(\frac{\boldsymbol{\omega}-\boldsymbol{\nu}}{2}\right)\xi_1(\boldsymbol{\omega})\xi_1(\boldsymbol{\nu})d\boldsymbol{\omega}d\boldsymbol{\nu}\\
&+(2\pi)^{-d}\int_{\mathbb{R}^{2d}} e^{i(\boldsymbol{\omega}'\mathbf{x}-\boldsymbol{\nu}'\mathbf{y})}\xi_1(\boldsymbol{\omega})\xi_1(\boldsymbol{\nu})\\
&\times\left(\int_{\|\mathbf{u}\|_{\max}\le 2m+2m_a+T} e^{-i(\boldsymbol{\omega}'\mathbf{u}-\boldsymbol{\nu}'\mathbf{u})}\hat{e}_n(\boldsymbol{\omega})\eta(\boldsymbol{\nu})d\mathbf{u}\right)d\boldsymbol{\omega}d\boldsymbol{\nu}.
\end{aligned}\tag{4.54}$$

Let $\{\psi_1,\cdots,\psi_n\}$ be as in Eq. (4.23), and then for $k=1,\cdots,n$,

$$(\psi_k,\psi_k)_{f_{\sigma_1,\theta_1}}-(\psi_k,\psi_k)_{f_{\sigma_0,\theta_0}}=\lambda_{k,n}-1=\nu_{k,n}^*+\nu_{k,n}^{**},\tag{4.55}$$

where

$$\nu_{k,n}^* = (2\pi)^{-d}\int_{\mathbb{R}^{2d}}\psi_k(\boldsymbol{\omega})\overline{\psi_k(\boldsymbol{\nu})}\eta_n^*\left(\frac{\boldsymbol{\omega}+\boldsymbol{\nu}}{2}\right)\kappa\left(\frac{\boldsymbol{\omega}-\boldsymbol{\nu}}{2}\right)\xi_1(\boldsymbol{\omega})\xi_1(\boldsymbol{\nu})d\boldsymbol{\omega}d\boldsymbol{\nu},$$

and

$$\begin{aligned}
\nu_{k,n}^{**} =& (2\pi)^{-d}\int_{\mathbb{R}^{2d}}\psi_k(\boldsymbol{\omega})\overline{\psi_k(\boldsymbol{\nu})}\xi_1(\boldsymbol{\omega})\xi_1(\boldsymbol{\nu})\\
&\times\left(\int_{\|\mathbf{u}\|_{\max}\le 2m+2m_a+T} e^{-i(\boldsymbol{\omega}'\mathbf{u}-\boldsymbol{\nu}'\mathbf{u})}\hat{e}_n(\boldsymbol{\omega})\eta(\boldsymbol{\nu})d\mathbf{u}\right)d\boldsymbol{\omega}d\boldsymbol{\nu}.
\end{aligned}$$

Using Bessel's inequality, we have

$$
\begin{aligned}
\sum_{k=1}^{n}\left|\nu_{k,n}^{**}\right| &\leq (2\pi)^{-d}\sum_{k=1}^{n}\int_{\|\mathbf{u}\|_{\max}\leq 2m+2m_a+T}\left|\int_{\mathbb{R}^d} e^{-i\boldsymbol{\omega}'\mathbf{u}}\psi_k(\boldsymbol{\omega})\xi_1(\boldsymbol{\omega})\hat{e}_n(\boldsymbol{\omega})d\boldsymbol{\omega}\right| \\
&\quad\times\left|\int_{\mathbb{R}^d} e^{i\boldsymbol{\nu}'\mathbf{u}}\overline{\psi_k(\boldsymbol{\nu})}\xi_1(\boldsymbol{\nu})\eta(\boldsymbol{\nu})d\boldsymbol{\nu}\right|d\mathbf{u} \\
&\leq \frac{1}{2(2\pi)^d}\int_{\|\mathbf{u}\|_{\max}\leq 2m+2m_a+T}\sum_{k=1}^{n}\left\{\left|\int_{\mathbb{R}^d} e^{-i\boldsymbol{\omega}'\mathbf{u}}\psi_k(\boldsymbol{\omega})\frac{\xi_1(\boldsymbol{\omega})}{f_{\sigma_0,\theta_0}^{1/2}(\boldsymbol{\omega})}\hat{e}_n(\boldsymbol{\omega})\right.\right. \\
&\quad\left.\times f_{\sigma_0,\theta_0}^{1/2}(\boldsymbol{\omega})d\boldsymbol{\omega}\right|^2+\left.\left|\int_{\mathbb{R}^d} e^{i\boldsymbol{\nu}'\mathbf{u}}\overline{\psi_k(\boldsymbol{\nu})}\frac{\xi_1(\boldsymbol{\nu})}{f_{\sigma_0,\theta_0}^{1/2}(\boldsymbol{\nu})}\eta(\boldsymbol{\nu})f_{\sigma_0,\theta_0}^{1/2}(\boldsymbol{\nu})d\boldsymbol{\nu}\right|^2\right\}d\mathbf{u} \\
&\leq \frac{1}{2(2\pi)^d}\int_{\|\mathbf{u}\|_{\max}\leq 2m+2m_a+T}\int_{\mathbb{R}^d}\frac{\xi_1^2(\boldsymbol{\omega})}{f_{\sigma_0,\theta_0}(\boldsymbol{\omega})}|\hat{e}_n(\boldsymbol{\omega})|^2 d\boldsymbol{\omega}d\mathbf{u} \\
&\quad+\frac{1}{2(2\pi)^d}\int_{\|\mathbf{u}\|_{\max}\leq 2m+2m_a+T}\int_{\mathbb{R}^d}\frac{\xi_1^2(\boldsymbol{\nu})}{f_{\sigma_0,\theta_0}(\boldsymbol{\nu})}|\eta(\boldsymbol{\nu})|^2 d\boldsymbol{\nu}d\mathbf{u} \\
&\leq \frac{(2m+2m_a+T)^d}{2\pi^d}\left\{\sup_{\mathbf{s}\in\mathbb{R}^d}\frac{\xi_1^2(\mathbf{s})}{f_{\sigma_0,\theta_0}(\mathbf{s})}\right\}\int_{\mathbb{R}^d}|\hat{e}_n(\mathbf{x})|^2+|\eta(\mathbf{x})|^2 d\mathbf{x},
\end{aligned}
\tag{4.56}
$$

and

$$
\begin{aligned}
\sum_{k=1}^{n}\left|\nu_{k,n}^{*}\right|^2 &\leq (2\pi)^{-d}\left\{\sup_{\mathbf{s}\in\mathbb{R}^d}\frac{\xi_1^2(\mathbf{s})}{f_{\sigma_0,\theta_0}(\mathbf{s})}\right\}^2\int_{\mathbb{R}^{2d}}\left|\eta_n^*\left(\frac{\boldsymbol{\omega}+\boldsymbol{\nu}}{2}\right)\kappa\left(\frac{\boldsymbol{\omega}-\boldsymbol{\nu}}{2}\right)\right|^2 d\boldsymbol{\omega}d\boldsymbol{\nu} \\
&= \pi^{-d}\left\{\sup_{\mathbf{s}\in\mathbb{R}^d}\frac{\xi_1^2(\mathbf{s})}{f_{\sigma_0,\theta_0}(\mathbf{s})}\right\}^2\int_{\mathbb{R}^d}|\eta_n^*(\boldsymbol{\omega})|^2 d\boldsymbol{\omega}\int_{\mathbb{R}^d}|\kappa(\boldsymbol{\nu})|^2 d\boldsymbol{\nu}.
\end{aligned}
\tag{4.57}
$$

From Eqs. (4.37) and (4.45), there exists constant C_2 independent of n such that

$$
\sum_{k=1}^{n}\left|\nu_{k,n}^{**}\right| \leq \frac{C_2}{\epsilon_n^{a+d}}. \tag{4.58}
$$

Meanwhile, Combining Lemma 4.4 and Eq. (4.37), we observe that there exists constant C_1 independent of n such that

$$
\sum_{k=1}^{n}\left|\nu_{k,n}^{*}\right|^2 \leq C_1\epsilon_n^{\beta_0}. \tag{4.59}
$$

Moreover, Jensen's inequality implies that

$$
\sum_{k=1}^{n}\left|\nu_{k,n}^{*}\right| \leq \left(n\sum_{k=1}^{n}\left|\nu_{k,n}^{*}\right|^2\right)^{1/2} \leq \sqrt{C_1 n\epsilon_n^{\beta_0}}. \tag{4.60}
$$

So we conclude that

$$
\begin{aligned}
\sum_{k=1}^{n}|\lambda_{k,n}-1| &\leq \sum_{k=1}^{n}\left(\left|\nu_{k,n}^{*}\right|+\left|\nu_{k,n}^{**}\right|\right) \\
&\leq \sqrt{C_1 n\epsilon_n^{\beta_0}}+\frac{C_2}{\epsilon_n^{a+d}}.
\end{aligned}
\tag{4.61}
$$

Finally for any constant $\vartheta > 0$, using Markov's inequality, Lemma 4.2 and Eq. (4.61) we obtain

$$
\begin{aligned}
&P_0\left(\frac{1}{\sqrt{n}}\left|\mathbf{X}_n'\big(\sigma_1^2\Gamma_{(n,\theta_1)}\big)^{-1}\mathbf{X}_n - \mathbf{X}_n'\big(\sigma_0^2\Gamma_{(n,\theta_0)}\big)^{-1}\mathbf{X}_n\right| > \vartheta\right) \\
=&P_0\left(\frac{1}{\sqrt{n}}\left|\sum_{k=1}^{n}\big(\lambda_{k,n}^{-1}-1\big)Y_{k,n}^2\right| > \vartheta\right) \\
\leq&\frac{1}{\vartheta\sqrt{n}}\sum_{k=1}^{n}\left|\lambda_{k,n}^{-1}-1\right| \\
\leq&\frac{1}{\vartheta\sqrt{n}}\left\{\max_{1\leq j\leq n}\lambda_{j,n}^{-1}\right\}\sum_{k=1}^{n}\left|\lambda_{k,n}-1\right| \\
\leq&\frac{1}{C_i\vartheta\sqrt{n}}\left(\sqrt{C_1 n\epsilon_n^{\beta_0}}+\frac{C_2}{\epsilon_n^{a+d}}\right) \\
=&\frac{C_1^{1/2}\epsilon_n^{\beta_0/2}}{C_i\vartheta}+\frac{C_2}{C_i\vartheta n^{1/2}\epsilon_n^{a+d}}.
\end{aligned}
\tag{4.62}
$$

Choose ϵ_n such that $\epsilon_n \to 0$ and $n^{1/2}\epsilon_n^{a+d} \to \infty$ as $n \to \infty$. It follows from Eq. (4.62) that

$$
\frac{\sigma_0^2\theta_0}{\sqrt{n}}\left(\mathbf{X}_n'\big(\sigma_1^2\Gamma_{(n,\theta_1)}\big)^{-1}\mathbf{X}_n - \mathbf{X}_n'\big(\sigma_0^2\Gamma_{(n,\theta_0)}\big)^{-1}\mathbf{X}_n\right) \xrightarrow{P_0} 0. \tag{4.63}
$$

This proves Theorem 4.2. □

Lemma 4.5. *Let* ξ_0 *be as in Eq.* (4.35). *Then* $\xi_0 : \mathbb{R}^d \to \mathbb{R}$ *is a continuous, isotropic and strictly positive function and* $\xi_0(\boldsymbol{\omega}) \asymp \|\boldsymbol{\omega}\|^{-\kappa}$ *as* $\|\boldsymbol{\omega}\| \to \infty$.

Proof. We split the proof into two parts. When $d = 1$, we follow the proof of Lemma 5 in Du, Zhang, and Mandrekar (2009). Notice that ξ_0 is continuous and real symmetric since $c_0 \in L^1(\mathbb{R})$ and c_0 is real symmetric. Meanwhile,

$$
\xi_0(\omega) = \int_{-1}^{1} e^{-i\omega x}|x|^{\kappa-1}dx = 2\int_0^1 \cos(\omega x)x^{\kappa-1}dx. \tag{4.64}
$$

When $\omega > 0$, let $u = \omega x$. We have

$$
\xi_0(\omega) = 2\omega^{-\kappa}\int_0^{\omega}\cos(u)u^{\kappa-1}du. \tag{4.65}
$$

Notice that $\int_0^\infty \cos(u)u^{\kappa-1}du = \Gamma(\kappa)\cos\big(\frac{\pi\kappa}{2}\big)$. So

$$
\xi_0(\omega) \asymp |\omega|^{-\kappa}, \quad \text{as } |\omega| \to \infty.
$$

As for the strict positiveness, first notice that $\xi_0(0) = \int_{-1}^{1} |x|^{\kappa-1}dx > 0$. For any $\omega > 0$, it suffices to show

$$y(\omega) = \int_0^\omega \cos(u)u^{-\delta}du > 0, \tag{4.66}$$

where $\delta = 1 - \kappa \in [1/2, 1)$. Note that $y'(\omega) = \cos(\omega)\omega^{-\delta}$ and $y''(\omega) = -\sin(\omega)\omega^{-\delta} - \delta\cos(\omega)\omega^{-\delta-1}$. Therefore, the minimum points over $(0, \infty)$ are $\{2k\pi + 3\pi/2, k = 0, 1, \cdots\}$. We first claim $y(3\pi/2)$ is the global minimum. Actually

$$\begin{aligned}
&\int_{2k\pi+3\pi/2}^{2(k+1)\pi+3\pi/2} \cos(u)u^{-\delta}du \\
=&\int_{2k\pi+3\pi/2}^{2k\pi+5\pi/2} \cos(u)u^{-\delta}du + \int_{2k\pi+5\pi/2}^{2(k+1)\pi+3\pi/2} \cos(u)u^{-\delta}du \\
=&\int_{3\pi/2}^{5\pi/2} \cos(s)(2k\pi + s)^{-\delta}ds - \int_{3\pi/2}^{5\pi/2} \cos(s)(2k\pi + \pi + s)^{-\delta}du > 0.
\end{aligned} \tag{4.67}$$

And

$$\begin{aligned}
y(3\pi/2) =& \int_0^{\pi/4} \cos(u)u^{-\delta}du + \int_{\pi/4}^{\pi/2} \cos(u)u^{-\delta}du \\
&+ \int_{\pi/2}^{\pi} \cos(u)u^{-\delta}du + \int_{\pi}^{3\pi/2} \cos(u)u^{-\delta}du \\
&\geq \cos(\pi/4)\frac{(\pi/4)^{1-\delta}}{1-\delta} + \frac{1}{(\pi/2)^\delta}\left(1 - \frac{\sqrt{2}}{2}\right) - \frac{1}{(\pi/2)^\delta} - \frac{1}{\pi^\delta} \\
&\geq \frac{\sqrt{2}}{2}\sqrt{\pi} + \frac{2}{\pi}\left(1 - \frac{\sqrt{2}}{2}\right) - \sqrt{\frac{2}{\pi}} - \frac{1}{\sqrt{\pi}} > 0.
\end{aligned} \tag{4.68}$$

This completes the strict positiveness.

When $d \geq 2$, we follow the proof of Lemma 2 in the Supplement of Bevilacqua et al. (2019). Obviously, $\xi_0(\mathbf{0}) > 0$. Let U_d be the uniform probability measure on $S^{d-1} = \{\mathbf{u} \in$

$\mathbb{R}^d : \|\mathbf{u}\| = 1\}$. By isotropy, we have for all $\boldsymbol{\omega} \in \mathbb{R}^d \setminus \{0\}$,

$$\begin{aligned}
\xi_0(\boldsymbol{\omega}) &= \int_{S^{d-1}} \int_{\|\mathbf{x}\|\le 1} e^{-i\|\boldsymbol{\omega}\|\mathbf{x}'\mathbf{u}} \|\mathbf{x}\|^{\kappa-d} d\mathbf{x} U_d(d\mathbf{u}) \\
&= \int_{\|\mathbf{x}\|\le 1} \Gamma(d/2) \left(\frac{2}{\|\boldsymbol{\omega}\|\|\mathbf{x}\|} \right)^{(d-2)/2} J_{(d-2)/2}(\|\boldsymbol{\omega}\|\|\mathbf{x}\|) \|\mathbf{x}\|^{\kappa-d} d\mathbf{x} \\
&= (2\pi)^{d/2} \|\boldsymbol{\omega}\|^{\frac{2-d}{2}} \int_0^1 r^{\kappa-\frac{d}{2}} J_{(d-2)/2}(\|\boldsymbol{\omega}\| r) dr \qquad (4.69) \\
&= (2\pi)^{d/2} \|\boldsymbol{\omega}\|^{-\kappa} \int_0^{\|\boldsymbol{\omega}\|} r^{\kappa-\frac{d}{2}} J_{(d-2)/2}(r) dr \\
&= 2\kappa^{-1} \pi^{d/2} \Gamma(d/2)^{-1} {}_1F_2(\kappa/2; \kappa/2+1, d/2; -(\|\boldsymbol{\omega}\|/2)^2),
\end{aligned}$$

where $J_{(d-2)/2}$ is the Bessel function of the first kind, and the generalized hypergeometric function ${}_1F_2$ is defined as

$$ {}_1F_2(a; b, c; z) = \sum_{k=0}^{\infty} \frac{(a)_k z^k}{(b)_k (c)_k k!}, \qquad (4.70)$$

with $(q)_k = \Gamma(q+k)/\Gamma(q)$ for $k \in \mathbb{N} \cup \{0\}$. See M. L. Stein (1999a, p.43) for the first and second equation in Eq. (4.69); the third equation is derived by the spherical coordinates transform; and see Prudnikov, Brychkov, and Marichev (1986, 1.8.1.1, p.37) for the last equation. From Theorem 2 in Fields and Ismail (1975), we get that $\xi_0 > 0$ for $d \ge 2$. Then ξ_0 is a continuous, isotropic and strictly positive on $\mathbb{R}^d$.

Next, we prove the integral $\int_0^{\infty} r^{\kappa-\frac{d}{2}} J_{(d-2)/2}(r) dr$ exists. Based on the series expansion (see, e.g., Abramowitz and Stegun 1968, p. 9.1.10)

$$J_\nu(x) = \left(\frac{x}{2}\right)^{\nu} \sum_{k=0}^{\infty} \frac{\left(-\frac{1}{4}x^2\right)^k}{k!\Gamma(\nu+k+1)},$$

we have

$$J_{(d-2)/2}(r) = O(r^{(d-2)/2}),$$

as $r \to 0$. So the integrand is integrable around the origin. Meanwhile, Based on the asymptotic expansion (see, e.g., Abramowitz and Stegun 1968, p. 9.2.1)

$$J_\nu(x) = \sqrt{2/(\pi x)} \left\{ \cos\left(x - \frac{1}{2}\nu\pi - \frac{1}{4}\pi \right) + O(|x|^{-1}) \right\}, \quad \text{as } |x| \to \infty,$$

we have

$$J_{(d-2)/2}(r) = O(r^{-1/2}),$$

as $r \to \infty$. So the integrand is integrable around the infinity. Overall, $\int_0^\infty r^{\kappa-\frac{d}{2}} J_{(d-2)/2}(r)dr$ exists. Therefore, $\xi_0(\boldsymbol{\omega}) \asymp \|\boldsymbol{\omega}\|^{-\kappa}$ as $\|\boldsymbol{\omega}\| \to \infty$. This completes the proof. □

4.4 Simulations

In this section, we investigate the finite sample performance of the estimator $\hat{\sigma}_n^2\theta_1$ of $\sigma_0^2\theta_0$. We simulate, using the R package RandomFields developed by Schlather et al. (2017), and then we estimate with 1000 realizations from a zero mean Gaussian field with the covariance function Eq. (4.1) where

- $\sigma_0 = 1$ and $\theta_1 = 1$ are fixed.
- $\theta_0 = 2, 5$ or 10.
- For $d = 2$ and $\alpha = 1.2$, we generated $X(\mathbf{t})$ on a regular grid of $[0,1]^2$ with sample size $n = 50m$, that is,
$$\left\{X\left(\frac{i_1}{40}, \frac{i_2}{50}\right) : 1 \le i_1 \le m, 1 \le i_2 \le 50\right\},$$
where $m = 20, 30$ or 40.
- For $d = 3$ and $\alpha = 1.8$, we generated $X(\mathbf{t})$ on a regular grid of $[0,1]^3$ with sample size $n = 100m$, that is,
$$\left\{X\left(\frac{i_1}{20}, \frac{i_2}{10}, \frac{i_3}{10}\right) : 1 \le i_1 \le m, 1 \le i_2, i_3 \le 10\right\},$$
where $m = 10, 15$ or 20.

Tables 4.1 and 4.2 summarize the percentiles of order $0.05, 0.25, 0.5, 0.75, 0.95$, bias and sample standard deviation (SD) of $\hat{\sigma}_n^2\theta_1$. Fig. 4.1 shows histograms of $\sqrt{n/2}(\hat{\sigma}_n^2\theta_1/(\sigma_0^2\theta_0) - 1)$ when $d = 2$ and $n = 2000$.

Remark 4.2. *As we see from Tables 4.1 and 4.2, the estimator $\hat{\sigma}_n^2\theta_1$ has a larger bias when the fixed scale parameter θ_1 is further away from the true parameter θ_0. This phenomenon not only occurs under the powered exponential model but also other models. See, for example,*

Table S-1 in the supplement of Kaufman and Shaby (2013) for the Matérn model and Table 2 in Bevilacqua et al. (2019) for the generalized Wendland model.

θ_0	n	5%	25%	50%	75%	95%	Bias	SD
2	1000	1.853	1.939	2.006	2.067	2.157	0.004	0.093
	1500	1.887	1.953	2.007	2.056	2.126	0.006	0.073
	2000	1.903	1.961	2.007	2.049	2.112	0.006	0.065
5	1000	4.677	4.891	5.056	5.216	5.458	0.058	0.237
	1500	4.761	4.928	5.056	5.176	5.368	0.055	0.184
	2000	4.797	4.948	5.056	5.159	5.320	0.054	0.156
10	1000	9.533	9.924	10.243	10.541	10.983	0.240	0.448
	1500	9.635	9.978	10.245	10.481	10.850	0.239	0.372
	2000	9.718	10.009	10.236	10.457	10.795	0.240	0.324

Table 4.1 Percentiles, Bias, and SD of $\hat{\sigma}_n^2\theta_1$ when $d = 2$

θ_0	n	5%	25%	50%	75%	95%	Bias	SD
2	1000	1.933	2.017	2.085	2.147	2.250	0.085	0.096
	1500	1.956	2.027	2.075	2.133	2.217	0.079	0.079
	2000	1.965	2.026	2.076	2.122	2.191	0.076	0.070
5	1000	5.453	5.738	5.925	6.120	6.392	0.927	0.286
	1500	5.506	5.708	5.875	6.028	6.246	0.873	0.231
	2000	5.506	5.718	5.857	5.991	6.166	0.852	0.196
10	1000	13.152	13.817	14.297	14.788	15.498	4.315	0.715
	1500	13.233	13.736	14.118	14.501	15.083	4.138	0.566
	2000	13.298	13.727	14.039	14.369	14.875	4.055	0.483

Table 4.2 Percentiles, Bias, and SD of $\hat{\sigma}_n^2\theta_1$ when $d = 3$

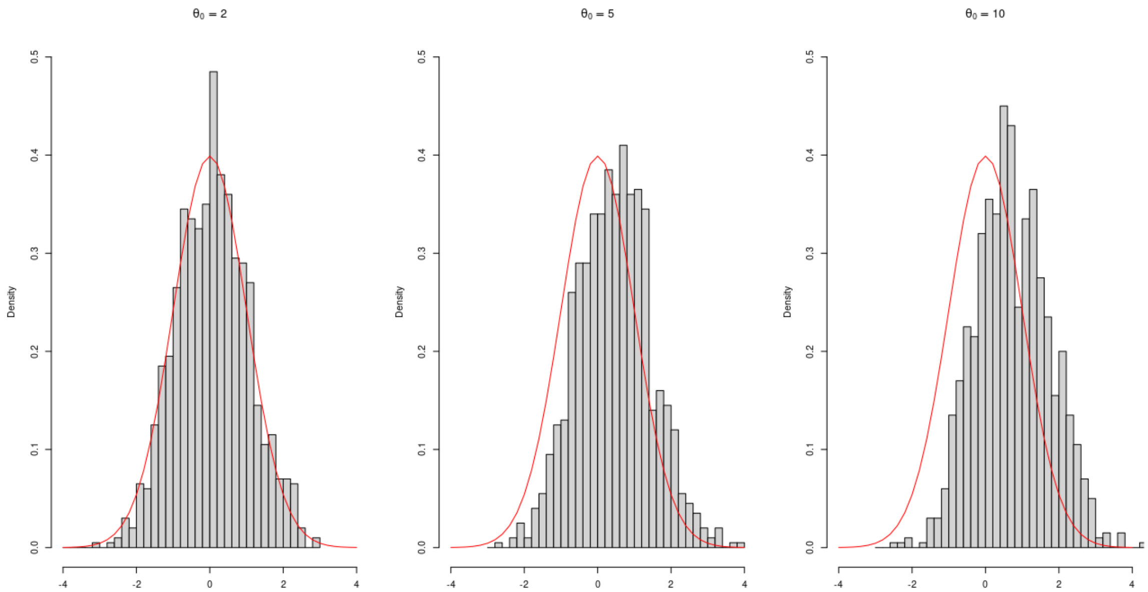

Figure 4.1 Histograms of $\sqrt{n/2}(\hat{\sigma}_n^2\theta_1/(\sigma_0^2\theta_0) - 1)$ with $\alpha = 1.2$ when $d = 2$ and $n = 2000$. The parameter θ_0 is $2, 5$ and 10 from left to right. The red curve is the density of $N(0, 1)$.

BIBLIOGRAPHY

Abramowitz, Milton and Irene A Stegun (1968). Handbook of mathematical functions with formulas, graphs, and mathematical tables. Vol. 55. US Government printing office.

Abry, Patrice and Gustavo Didier (2018). Wavelet estimation for operator fractional Brownian motion. In: *Bernoulli* 24 (2).

Adler, Robert J (2010). The Geometry of Random Fields.

Adler, Robert J and Jonathan E Taylor (2007). Random Fields and Geometry.

Amblard, Pierre Olivier and Jean François Coeurjolly (2011). Identification of the multivariate fractional brownian motion. In: *IEEE Transactions on Signal Processing* 59 (11).

Anderes, Ethan (2010). On the consistent separation of scale and variance for Gaussian random fields. In: *Annals of Statistics* 38 (2).

Bachoc, François, Emilio Porcu, Moreno Bevilacqua, Reinhard Furrer, and Tarik Faouzi (2022). Asymptotically equivalent prediction in multivariate geostatistics. In: *Bernoulli* 28 (4).

Bayarr, M. J., James O. Berger, Eliza S. Calder, Keith Dalbey, Simon Lunagomez, Abani K. Patra, E. Bruce Pitman, Elaine T. Spiller, and Robert L. Wolpert (2009). Using statistical and computer models to quantify volcanic hazards. In: *Technometrics* 51 (4).

Bergström, Harald (1952). On some expansions of stable distribution functions. In: *Arkiv för matematik* 2 (4).

Bevilacqua, Moreno, Tarik Faouzi, Reinhard Furrer, and Emilio Porcu (Apr. 2019). Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics. In: *The Annals of Statistics* 47 (2).

Chan, Grace and Andrew T.A. Wood (2000). Increment-based estimators of fractal dimension for two-dimensional surface data. In: *Statistica Sinica* 10 (2).

Chilès, Jean Paul and Pierre Delfiner (2012). Geostatistics: Modeling Spatial Uncertainty: Second Edition.

Cressie, Noel A. C. (Sept. 1993). Statistics for Spatial Data. Wiley. ISBN: 9780471002550.

Didier, Gustavo and Vladas Pipiras (Feb. 2011). Integral representations and properties of operator fractional Brownian motions. In: *Bernoulli* 17 (1).

Du, Juan, Hao Zhang, and V. S. Mandrekar (Dec. 2009). Fixed-domain asymptotic properties of tapered maximum likelihood estimators. In: *The Annals of Statistics* 37 (6A).

Feldman, Jacob (1958). Equivalence and perpendicularity of Gaussian processes. In: *Pacific Journal of Mathematics* 8 (4).

Feller, William (1991). An introduction to probability theory and its applications, Volume 2. Vol. 81. John Wiley & Sons.

Fields, Jerry L. and Mourad El-Houssieny Ismail (May 1975). On the Positivity of some ${}_1F_2$'s. In: *SIAM Journal on Mathematical Analysis* 6 (3), pp. 551–559.

Gardiner, C.W. (1985). Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences. Proceedings in Life Sciences. Springer-Verlag. ISBN: 9783540113577.

Gneiting, Tilmann, William Kleiber, and Martin Schlather (2010). Matérn cross-covariance functions for multivariate random fields. In: *Journal of the American Statistical Association* 105 (491).

Gneiting, Tilmann and Martin Schlather (2004). Stochastic models that separate fractal dimension and the hurst effect. In: *SIAM Review* 46 (2).

Gradshteyn, I. S., I. M. Ryzhik, Alan Jeffrey, Y. V. Geronimus, M. Y. Tseytlin, and Y. C. Fung (1981). Table of Integrals, Series, and Products. In: *Journal of Biomechanical Engineering* 103 (1).

Hanson, D. L. and F. T. Wright (June 1971). A Bound on Tail Probabilities for Quadratic Forms in Independent Random Variables. In: *The Annals of Mathematical Statistics* 42 (3), pp. 1079–1083.

Ibragimov, I. A. and Y. A. Rozanov (1978). Gaussian Random Processes. Springer New York. ISBN: 978-1-4612-6277-0.

Kaufman, C. G. and B. A. Shaby (June 2013). The role of the range parameter for estimation and prediction in geostatistics. In: *Biometrika* 100 (2), pp. 473–484.

Kent, John T. and Andrew T.A. Wood (1997). Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. In: *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 679–699.

Khoshnevisan, Davar, Yimin Xiao, and Yuquan Zhong (2003). Measuring the range of an additive Lévy process. In: *Annals of Probability* 31 (2).

Khuri, Andre I. (Oct. 2009). Linear Model Methodology. Chapman and Hall/CRC. ISBN: 9781420010442.

Lahiri, S N (2003). Central Limit Theorems for Weighted Sums of a Spatial Process under a Class of Stochastic and Fixed Designs.

Lévy, M. Paul (1940). Le Mouvement Brownien Plan. In: *American Journal of Mathematics* 62 (1/4).

Loh, Wei-Liem (Oct. 2005). Fixed-domain asymptotics for a subclass of Matérn-type Gaussian random fields. In: *The Annals of Statistics* 33 (5).

Loh, Wei-Liem (2015). Estimating the smoothness of a Gaussian random field from irregularly spaced data via higher-order quadratic variations. In: *Annals of Statistics* 43 (6).

Loh, Wei-Liem, Saifei Sun, and Jun Wen (Dec. 2021). On fixed-domain asymptotics, parameter estimation and isotropic Gaussian random fields with Matérn covariance functions. In: *The Annals of Statistics* 49 (6).

Mallat, Stephane (2008). A Wavelet Tour of Signal Processing: The Sparse Way.

Mardia, K. V. and R. J. Marshall (Apr. 1984). Maximum Likelihood Estimation of Models for Residual Covariance in Spatial Regression. In: *Biometrika* 71 (1), p. 135.

Mason, J. D. and Yimin Xiao (Jan. 2002). Sample Path Properties of Operator-Self-Similar Gaussian Random Fields. In: *Theory of Probability & Its Applications* 46 (1), pp. 58–78.

Nolan, John (2005). Multivariate stable densities and distribution functions: general and elliptical case Deutsche Bundesbank's 2005 Annual Fall Conference Multivariate stable densities and distribution functions: general and elliptical case.

Paulo, Rui (2005). Default priors for Gaussian processes.

Peng, Chien Yu and C. F. Jeff Wu (2014). On the choice of nugget in kriging modeling for deterministic computer experiments. In: *Journal of Computational and Graphical Statistics* 23 (1).

Prudnikov, A.P., Yury Brychkov, and O.I. Marichev (1986). Integrals and series: special functions. Vol. 2. CRC press.

Sacks, Jerome, Susannah B. Schiller, and William J. Welch (Feb. 1989). Designs for Computer Experiments. In: *Technometrics* 31 (1), pp. 41–47.

Sacks, Jerome, William J. Welch, Toby J. Mitchell, and Henry P. Wynn (1989). Design and analysis of computer experiments. In: *Statistical Science* 4 (4).

Samorodnitsky, Gennady and Murad S. Taqqu (Nov. 2017). Stable Non-Gaussian Random

Processes. Routledge. ISBN: 9780203738818.

Schlather, Martin, Alexander Malinowski, Marco Oesting, Daphne Boecker, Kirstin Strokorb, Sebastian Engelke, Johannes Martini, Felix Ballani, Olga Moreva, Jonas Auel, Peter J Menck, Sebastian Gross, Ulrike Ober, Christoph Berreth, Katharina Burmeister, Juliane Manitz, Paulo Ribeiro, Richard Singleton, Ben Pfaff, and R Core Team (2017). *RandomFields: Simulation and Analysis of Random Fields*. R package version 3.1.50.

Stein, Michael (June 1990). Uniform Asymptotic Optimality of Linear Predictions of a Random Field Using an Incorrect Second-Order Structure. In: *The Annals of Statistics* 18 (2).

Stein, Michael L. (Mar. 1988). Asymptotically Efficient Prediction of a Random Field with a Misspecified Covariance Function. In: *The Annals of Statistics* 16 (1).

Stein, Michael L. (Nov. 1989). [Design and Analysis of Computer Experiments]: Comment. In: *Statistical Science* 4 (4).

Stein, Michael L. (Aug. 1993). A simple condition for asymptotic optimality of linear predictions of random fields. In: *Statistics & Probability Letters* 17 (5), pp. 399–404.

Stein, Michael L. (1999a). Interpolation of Spatial Data. Springer New York. ISBN: 978-1-4612-7166-6.

Stein, Michael L. (1999b). Predicting random fields with increasing dense observations. In: *Annals of Applied Probability* 9 (1).

Stein, Michael L. (June 2004). Equivalence of Gaussian measures for some nonstationary random fields. In: *Journal of Statistical Planning and Inference* 123 (1), pp. 1–11.

van der Vaart, Aad (Oct. 1996). Maximum likelihood estimation under a spatial sampling scheme. In: *The Annals of Statistics* 24 (5).

Velandia, Daira, François Bachoc, Moreno Bevilacqua, Xavier Gendre, and Jean Michel Loubes (2017). Maximum likelihood estimation for a bivariate Gaussian process under fixed domain asymptotics. In: *Electronic Journal of Statistics* 11 (2).

Wang, Daqing (2010). Fixed domain asymptotics and consistent estimation for Gaussian random field models in spatial statistics and computer experiments. In: *Nat. Univ. Singapore PhD thesis, Singapore.*

Wang, Daqing and Wei-Liem Loh (Jan. 2011). On fixed-domain asymptotics and covariance tapering in Gaussian random field models. In: *Electronic Journal of Statistics* 5 (none).

Ying, Zhiliang (1991). Asymptotic properties of a maximum likelihood estimator with data from a Gaussian process. In: *Journal of Multivariate Analysis* 36 (2).

Ying, Zhiliang (Sept. 1993). Maximum Likelihood Estimation of Parameters under a Spatial Sampling Scheme. In: *The Annals of Statistics* 21 (3).

Zhang, Hao (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. In: *Journal of the American Statistical Association* 99 (465).

Zhang, Hao and Wenxiang Cai (May 2015). When Doesn't Cokriging Outperform Kriging? In: *Statistical Science* 30 (2).

Zhou, Yuzhen and Yimin Xiao (2018). Joint asymptotics for estimating the fractal indices of bivariate Gaussian processes. In: *Journal of Multivariate Analysis* 165.