PARAMETER ESTIMATION FOR UNIVARIATE AND BIVARIATE GAUSSIAN PROCESSES AND FIELDS

By

Nian Liu

A DISSERTATION

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

Statistics—Doctor of Philosophy

2024

ABSTRACT

Gaussian random fields are widely studied in various subject areas. This dissertation focuses on estimating covariance parameters of stationary Gaussian random fields based on both regularly and irregularly spaced sampling points, as well as investigating the infill asymptotic properties of the estimators.

We first consider a bivariate Gaussian random process and propose an increment-based estimator for the smoothness parameter in the cross-covariance function, for which the strong consistency and asymptotic normality hold under the infill asymptotic framework. We further study the joint asymptotic distribution of estimators for smoothness parameters in the cross-covariance and autocovariance functions. Subsequently, we estimate the scale parameter and range parameters of a univariate anisotropic Ornstein-Uhlenbeck field based on quadratic forms of vectors of observations. The estimators we propose are computationally more efficient than the maximum likelihood estimators but have similar infill asymptotic performances with MLEs. Another computational complexity reduction method we use is the Vecchia approximation. We estimate the scale parameter in the Matérn covariance function using the maximizer of the likelihood approximated by the standard Vecchia approach. We study the bias resulting from a misspecified range parameter and the conditioning variables of the Vecchia approximation. The theoretical results in this work are illustrated by simulations.

ACKNOWLEDGEMENTS

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

I would like to express my genuine gratitude to my advisor, Dr. Yimin Xiao, for his support, encouragement, and guidance in my research and career development. I would also like to express my appreciation to Dr. Andrew Finley, Dr. Shlomo Levental, Dr. Haolei Weng, and Dr. Dongsheng Wu for serving on my guidance committee and providing me with valuable suggestions. In addition, I appreciate the faculty and staff in the Department of Statistics and Probability for their help during my PhD program.

I would also like to thank my family and friends for their care and support. I am more than fortunate to be surrounded by such warm and kind people.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION 1

CHAPTER 2 ESTIMATION OF SMOOTHNESS PARAMETERS 5
- 2.1 Introduction 5
- 2.2 Estimating the Cross Smoothness Parameter 6
- 2.3 Irregular Sampling 23

CHAPTER 3 ANISOTROPIC ORNSTEIN-UHLENBECK FIELD 37
- 3.1 Introduction 37
- 3.2 Product Estimation 38
- 3.3 Separable Estimation 48
- 3.4 Simulation 54
- 3.5 Discussion 55

CHAPTER 4 VECCHIA APPROXIMATION 58
- 4.1 Introduction 58
- 4.2 Maximum Likelihood Estimator for σ^2 59
- 4.3 Simulation 64

BIBLIOGRAPHY 68

APPENDIX A QUADRATIC VARIATIONS FROM IRREGULAR SAMPLING 76

APPENDIX B HIGH EXCURSION PROBABILITY 84

APPENDIX C STOCHASTIC PARTIAL DIFFERENTIAL EQUATION 86

CHAPTER 1

INTRODUCTION

Gaussian random fields (GRFs) are essential tools in spatial statistics, physics, finance, image processing, and other various areas. A random field, as a generalization of a stochastic process, is a collection of random variables indexed by elements in a topological space, which could be taken as $\mathbb{R}^d$ ($d \geq 1$). This work focuses on estimating covariance parameters of stationary GRFs and investigating infill asymptotic properties of the estimators.

The covariance function of a univariate stationary isotropic GRF $\{X(\mathrm{t}), \mathrm{t} \in \mathbb{R}^d\}$ considered by Anderes and Stein (2008) and Loh (2015) is written as

$$\mathrm{Cov}(X(\mathrm{s}), X(\mathrm{t}+\mathrm{s})) = \sum_{k=0}^{\lfloor \nu \rfloor} \beta_k ||\mathrm{t}||^{2k} + \beta_\nu^* G_\nu(||\mathrm{t}||) + O(||\mathrm{t}||^{2\nu+\tau}) \quad \text{as } ||\mathrm{t}|| \to 0, \quad \forall \mathrm{s}, \mathrm{t} \in \mathbb{R}^d, \tag{1.1}$$

where $|| \cdot ||$ denotes the Euclidean distance, $\beta_0 > 0$, $\beta_\nu^* \neq 0$, and $\tau > 0$ are constants, $\lfloor \nu \rfloor = \max\{\nu_0 \in \mathbb{Z} : \nu_0 < \nu\}$, and $G_\nu : [0, \infty) \mapsto \mathbb{R}$ is defined by

$$G_\nu(x) = \begin{cases} x^{2\nu} + x^{2\nu}(\log x - 1) 1_{\mathbb{Z}}(\nu), & x > 0, \\ 0, & x = 0. \end{cases}$$

This model includes the Matérn and exponential classes of covariance functions, which are widely used in spatial interpolation (Stein, 1999; Gramacy, 2020).

The isotropic exponential class covariance function is defined as

$$\sigma^2 \exp\left(-\theta ||s||^{2\nu}\right), \quad \mathrm{s} \in \mathbb{R}^d, \tag{1.2}$$

where $\sigma^2 > 0$, $\theta > 0$, $0 < \nu \leq 1$. The case when $0 < \nu < 1$ is contained in model (1.1) with $\beta_0 = \sigma^2$. When $\nu = 1/2$, the function (1.2) is called the Ornstein-Uhlenbeck covariance function, which is also a special case of the Matérn class of covariance functions. The Matérn covariance model

$$(\theta ||\mathrm{t}||)^\nu K_\nu(\theta ||\mathrm{t}||), \quad \mathrm{t} \in \mathbb{R}^d, \tag{1.3}$$

where K_ν is the modified Bessel function of the second kind with order ν, was proposed by von Kármán (1948) with $\nu = 1/3$ and $d = 3$. Some properties of the Matérn model were demonstrated in

Matérn (1986), Kent (1989), and Stein (1999). The stochastic partial differential equation (SPDE) that generates a Gaussian process on $\mathbb{R}^d$ with the Matérn covariance function is presented in Whittle (1954) and Whittle (1963) as

$$\left(\nabla^2 - \theta^2\right)^p \xi(\mathrm{x}) = \epsilon(\mathrm{x}), \quad \mathrm{x} \in \mathbb{R}^d, \tag{1.4}$$

where ∇^2 is the Laplace operator, $\theta > 0$ and $p > d/4$ are constants, ϵ is the Gaussian white noise with unit variance. The covariance function of ξ as a solution to (1.4) is

$$E(\xi(\mathrm{s})\xi(\mathrm{t}+\mathrm{s})) = \frac{(||\mathrm{t}||/\theta)^{2p-d/2} K_{2p-d/2}(\theta||\mathrm{t}||)}{2^{2p-1}\Gamma(2p)}, \quad \mathrm{t}, \mathrm{s} \in \mathbb{R}^d. \tag{1.5}$$

A more general class of stationary GRFs on $\mathbb{R}^2$ derived from second-order SPDEs was discussed by Heine (1955). Later, Vecchia (1985) introduced the derivation of covariance functions from spectral densities of stationary GRFs on $\mathbb{R}^2$, and showed the corresponding SPDEs. One generalization of model (1.3) is the spatio-temporal covariance function (Cressie and Huang, 1999; Gneiting, 2002; De Iaco et al., 2002; Ma, 2005, 2008). Jones and Zhang (1997) considered the spatio-temporal random field defined by the SPDE

$$\left(\left(\sum_{i=1}^{d} \frac{\partial^2}{\partial s_i^2}\right)^p - c\frac{\partial}{\partial t}\right) Z(\mathrm{s};t) = \epsilon(\mathrm{s};t), \quad \mathrm{s} = (s_1, s_2, \ldots, s_d)' \in \mathbb{R}^d, t \in \mathbb{R},$$

where $p > d/2$ and $c > 0$ are constants, $\epsilon(s;t)$ is the Gaussian white noise.

For the multivariate GRF $\{X(\mathrm{t}), \mathrm{t} \in \mathbb{R}^d\}$, where $X \in \mathbb{R}^p$ and $p \geq 1$, Gneiting et al. (2010) introduced a multivariate Matérn model, where the marginal and cross-covariance functions of a multivariate spatial random field are all of the Matérn type. Hu et al. (2013) introduced an approach to construct multivariate Gaussian random fields (GRFs) using systems of SPDEs. Based on systems of SPDEs with additive type G noise whose marginal covariance functions are of Matérn type, Bolin and Wallin (2020) formulated a new class of multivariate non-Gaussian models. SPDE models for GRFs are also researched by Hu and Steinsland (2016), Leonenko et al. (2011), Carrizo Vergara (2018), and Lindgren et al. (2011, 2022).

The Matérn and exponential classes of covariance functions both have mainly three types of parameters: the scale parameter σ^2, which equals the variance of $X(\mathrm{t})$ at any $\mathrm{t} \in \mathbb{R}^d$; the range parameter θ, which measures how fast the correlation decays with the distance; and the smoothness parameter ν, which controls the smoothness such as mean square differentiability of the random field. More specifically, X is n times mean square differentiable if and only if $n < \nu$ (Stein, 1999; Anderes and Stein, 2008).

The increasing-domain asymptotics and infill (fixed-domain) asymptotics are two frameworks under which the covariance parameter estimations for GRFs have been studied (Cressie, 1993; Stein, 1999). Under the increasing-domain asymptotic framework, the minimum distance between sampling locations is bounded away from zero, and the sampling region grows as the sample size N increases. Under infill asymptotics, the sampling region is fixed and bounded, and the mesh of the sampling points decreases as the sample size N tends to infinity. Besides, there is another asymptotic framework called hybrid asymptotics or mixed domain asymptotics, under which the sampling locations increasingly densely fill in any given subregion of the unbounded sampling region (Stein, 1999; Lahiri, 2003; Lahiri and Mukherjee, 2004; Chang et al., 2017).

This work focuses on the infill asymptotic framework, which plays an important role in spatial sampling design and kriging (Stein, 1999; Zhu and Zhang, 2006). Assuming the smoothness parameter ν is known, Zhang (2004), Du et al. (2009), Wang and Loh (2011), and Kaufman and Shaby (2013) provided infill asymptotic results for the MLE and tapered MLE of the microergodic parameter of the GRF with the Matérn covariance function; while Bevilacqua et al. (2019) studied infill asymptotics for MLE of the microergodic parameter in the generalized Wendland covariance function, which exhibits the same behavior as of the Matérn function at the origin according to Gneiting (2002). Using quadratic variations defined based on irregularly spaced sampling designs (more details described in Appendices A.2-A.3), Loh et al. (2021) also estimated the microergodic parameter of the Matérn covariance function under the infill asymptotic framework.

The estimation of the smoothness parameter has also been widely studied. Regarding the fractal dimension, which is a measure of the smoothness of sample paths of a stochastic process, existing approaches of estimation include the box-counting method (Hall and Wood, 1993), variogram estimator (Constantine and Hall, 1994), periodogram-based estimator (Chan et al., 1995), variation method (Dubuc et al., 1989), etc. The infill asymptotic behavior of increment-based estimators for the smoothness parameter of a stationary GRF was studied by Kent and Wood (1997), Chan and Wood (2000), Loh (2015), and Loh et al. (2021). For time series or spatial data, Gneiting et al. (2012) discussed various types of estimators of its fractal dimension under the infill asymptotic framework, considering both stationary and nonstationary univariate GRF models. Zhou and Xiao (2018) studied the joint infill asymptotic properties of increment-based estimators for smoothness parameters in the autocovariance functions of two coordinates of $\{X(t) = (X_1(t), X_2(t))^T, t \in \mathbb{R}\}$, which extended the work of Kent and Wood (1997) to the bivariate case.

The subsequential chapters are organized as follows. In Chapter 2, we consider the bivariate model $\{X(t) = (X_1(t), X_2(t))^T, t \in \mathbb{R}\}$ studied by Zhou and Xiao (2018) and propose an increment-based estimator for the smoothness parameter in the cross-covariance function of $X(t)$, based on both regularly and irregularly spaced sampling points. The strong consistency and asymptotic normality of the estimator are demonstrated under the infill asymptotic framework. In Chapter 3, we estimate the scale parameter and range parameters of a univariate anisotropic Ornstein-Uhlenbeck field on $\mathbb{R}^2$. The estimators we propose have similar asymptotic behaviors with MLEs, but with less computational cost. In Chapter 4, we estimate the scale parameter in the Matérn covariance function using MLE, whose computational complexity is reduced by the Vecchia approximation. We study the bias resulting from a misspecified range parameter and the conditioning variables of the Vecchia approximation. Simulation results are presented in each chapter to illustrate the theoretical results.

CHAPTER 2

ESTIMATION OF SMOOTHNESS PARAMETERS

2.1 Introduction

Based on the infill asymptotic behaviors of quadratic variations (Lévy, 1940; Baxter, 1956; Grenander, 1981), the increment-based methods have been used by several authors to consistently estimate the smoothness parameter of a univariate stationary Gaussian random field under the infill asymptotic framework (Istas and Lang, 1997; Kent and Wood, 1997; Chan and Wood, 2000; Loh, 2015; Loh et al., 2021). Consider a Gaussian process X observed on $0 = t_0 < t_1 < \cdots < t_n = 1$, Istas and Lang (1997) and Kent and Wood (1997) independently generalized the quadratic variation defined as $\sum_{j=1}^{n}(X(t_j) - X(t_{j-1})^2$ using vectors of increment. The empirical mean of squared process defined by Kent and Wood (1997) is equivalent to the empirical quadratic variation studied by Istas and Lang (1997). An increment of order p is vector $a = (a_{-J}, a_{1-J}, \ldots, a_J)^T \in \mathbb{R}^{2J+1}$ ($J > 0$) satisfying

$$\sum_{j=-J}^{J} j^q a_j \begin{cases} = 0, & 0 \le q \le p, \\ \neq 0, & q = p + 1. \end{cases}$$

The increment-based estimators could also be used for estimating the fractal dimension of nonstationary GRFs (Zhu and Stein, 2002; Begyn, 2005; Kubilius and Melichov, 2010).

Denote by $X = \{(X_1(t), X_2(t))^T, t \in \mathbb{R}\}$ a bivariate stationary Gaussian process with zero mean and covariance function

$$C(t) = \begin{pmatrix} C_{11}(t) & C_{12}(t) \\ C_{21}(t) & C_{22}(t) \end{pmatrix}. \tag{2.1}$$

Assume that as $|t| \to 0$,

$$C_{ii}(t) = \sigma_i^2 - c_{ii}|t|^{\alpha_{ii}} + o(|t|^{\alpha_{ii}}), \tag{2.2}$$

$$C_{ij}(t) = \rho\sigma_1\sigma_2(1 - c_{12}|t|^{\alpha_{12}} + o(|t|^{\alpha_{12}})), \tag{2.3}$$

where $\sigma_i, c_{ii}, c_{ij} > 0$, $\alpha_{ii} \in (0, 2)$, $|\rho| \in (0, 1)$, $i, j \in \{1, 2\}$, $i \neq j$. Following the framework of Gneiting et al. (2010), Zhou and Xiao (2018) imposed the following assumptions to make the covariance function (2.1) valid:

$$\alpha_{12} > (\alpha_{11} + \alpha_{22})/2$$
$$\text{or } \alpha_{12} = (\alpha_{11} + \alpha_{22})/2 \text{ and } c_{12}^2\rho^2\sigma_1^2\sigma_2^2 < c_{11}c_{22}.$$

2.2 Estimating the Cross Smoothness Parameter

Consider the Gaussian process X modeled by (2.1-2.3). When $\alpha_{12} = (\alpha_{11} + \alpha_{22})/2$, the cross smoothness parameter α_{12} could be estimated using estimators for α_{11} and α_{22}. This case can be treated by using the results in Zhou and Xiao (2018). In the following, we focus on the case when $\alpha_{12} > (\alpha_{11} + \alpha_{22})/2$ and construct an increment-based estimator for α_{12}.

The regularity conditions below are introduced for the convenience of subsequent analysis. Consider the condition (A_q) in Kent and Wood (1997) for the qth derivative of covariance function C_{ij}, that is,

$$C_{ij}^{(q)}(t) = -A_{ij}\frac{\alpha_{ij}!}{q!}|t|^{\alpha_{ij}-q} + o(|t|^{\alpha_{ij}-q}) \tag{2.4}$$

as $|t| \to 0$, where $q \geq 1$, $i, j \in \{1, 2\}$, $A_{ii} = c_{ii}$, $A_{12} = A_{21} = \rho\sigma_1\sigma_2 c_{12}$, and $\alpha_{ij}!/q! = \alpha_{ij}(\alpha_{ij} - 1)\dots(\alpha_{ij} - q + 1)$.

Under the infill asymptotics framework, Section 2.2.1 discusses the covariation of X, and Section 2.2.2 further studies asymptotic properties of the increment-based estimator for α_{12}. Some simulation results are presented in Section 2.2.3.

2.2.1 Covariation

Let $a = (a_{-J}, a_{1-J}, \dots, a_J)^T$ be an increment of order p. Denote by $X_{n,i}^u \in \mathbb{R}^{n(2J+1)}$ the vector of observations of component X_i, where $i = 1, 2$, $u = 1, 2, \dots, m$ and $n \in \mathbb{Z}^+$. For $j = 1, 2, \dots, 2J + 1$ and $k = 1, 2, \dots, n$, let

$$(X_{n,i}^u)_{j+(k-1)(2J+1)} = X_i\left(\frac{k + u(j - J - 1)}{n}\right).$$

In other words, for $k = 1, 2, \dots, n(2J + 1)$,

$$(X_{n,i}^u)_k = X_i\left(\frac{k_J + 1 + u(k - k_J(2J + 1) - J - 1)}{n}\right),$$

where $k_J = \max\{j \in \mathbb{Z} : j < k/(2J+1)\}$. Define

$$Y_n^u := \begin{pmatrix} Y_{n,1}^u \\ Y_{n,2}^u \end{pmatrix} = \begin{pmatrix} n^{\alpha_{11}/2}(I_n \otimes a^T) & 0 \\ 0 & n^{\alpha_{22}/2}(I_n \otimes a^T) \end{pmatrix} \begin{pmatrix} X_{n,1}^u \\ X_{n,2}^u \end{pmatrix},$$

where $\otimes$ denotes the Kronecker product. More specifically, for $k = 1, \dots, n$,

$$(Y_{n,i}^u)_k = n^{\alpha_{ii}/2} \sum_{j=1}^{2J+1} a_{j-J-1} (X_{n,i}^u)_{j+(k-1)(2J+1)}.$$

Denote by

$$Z_{n,12}^u(k) = n^{\alpha_{12}-(\alpha_{11}+\alpha_{22})/2} (Y_{n,1}^u)_k (Y_{n,2}^u)_k, \quad k = 1, \dots, n$$

and define the covariation as

$$\begin{aligned} \bar{Z}_{n,12}^u &= \frac{1}{n} \sum_{j=1}^{n} Z_{n,12}^u(j) \\ &= \frac{1}{2} n^{\alpha_{12}-(\alpha_{11}+\alpha_{22})/2-1} (Y_n^u)^T \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix} Y_n^u. \end{aligned} \tag{2.5}$$

We first discuss the infill asymptotic properties of covariations $\bar{Z}_{n,12}^u$, based on which the estimator for α_{12} will be constructed (see (2.27) below).

Theorem 1. *Assume (2.4) holds for $q = 2p + 3$ and $i, j \in \{1, 2\}$, then $\forall u = 1, \dots, m$,*

$$\bar{Z}_{n,12}^u \xrightarrow{P} A u^{\alpha_{12}}$$

as $n \to \infty$ if $\alpha_{11} + \alpha_{22} < 2\alpha_{12} < \alpha_{11} + \alpha_{22} + 1 < 4p + 4$ or $4p + 3 < \alpha_{11} + \alpha_{22} < 2\alpha_{12} < 4p + 4$, where $A = -\rho\sigma_1\sigma_2 c_{12} \sum_{k,l=-J}^{J} a_k a_l |k - l|^{\alpha_{12}}$.

Proof. Based on (2.2) and (2.3), for any $j, k = 1, \ldots, n$ and any $u, v = 1, \ldots, m$,

$$\begin{aligned}
\sigma^{uv}_{n,ir}(k-j) &:= E[(Y^u_{n,i})_j (Y^v_{n,r})_k] = n^{(\alpha_{ii}+\alpha_{rr})/2} \sum_{s,t=-J}^{J} a_s a_t E\left[X_i\left(\frac{j+su}{n}\right) X_r\left(\frac{k+tu}{n}\right)\right] \\
&= n^{(\alpha_{ii}+\alpha_{rr})/2} \sum_{s,t} a_s a_t C_{ir}\left(\frac{j-k+su-tv}{n}\right) \\
&= -A_{ir} n^{(\alpha_{ii}+\alpha_{rr})/2-\alpha_{ir}} \sum_{s,t} a_s a_t |j-k+su-tv|^{\alpha_{ir}} + o(n^{(\alpha_{ii}+\alpha_{rr})/2-\alpha_{ir}}) \\
&\to \begin{cases} -A_{ii} \sum_{s,t} a_s a_t |j-k+su-tv|^{\alpha_{ii}}, & i = r \\ 0, & i \neq r \end{cases}
\end{aligned} \tag{2.6}$$

as $n \to \infty$, where $i, r \in \{1, 2\}$. Thus,

$$\begin{aligned}
E[Z^u_{n,12}(j)] &= n^{\alpha_{12}-(\alpha_{11}+\alpha_{22})/2} E[(Y^u_{n,1})_j (Y^u_{n,2})_j] \\
&= -\rho\sigma_1\sigma_2 c_{12} \sum_{k,l} a_k a_l |k-l|^{\alpha_{12}} u^{\alpha_{12}} + o(1) \\
&\to A u^{\alpha_{12}} \quad \text{as} \quad n \to \infty,
\end{aligned} \tag{2.7}$$

where $A = 0$ if $\alpha_{12}/2 \in \mathbb{Z}$ and $p \geq \alpha_{12}/2$, due to the fact that $\sum_{k,l} a_k a_l (k-l)^r = 0$ for $r \leq 2p+1$.

If (2.4) holds for $q = 2p+3$, then $\forall -n < h < n$, there exists h^* between h and $h + su - tv$ such that

$$\begin{aligned}
&\sum_{s,t} a_s a_t C_{ir}\left(\frac{h+su-tv}{n}\right) \\
=&\frac{2(uv)^{p+1}}{(2p+2)!n^{2p+2}} \left(D_1^2 C_{ir}^{(2p+2)}\left(\frac{h}{n}\right) + \frac{u+v}{n(2p+3)} D_1 D_2 C_{ir}^{(2p+3)}\left(\frac{h^*}{n}\right)\right),
\end{aligned} \tag{2.8}$$

where $i, r \in \{1, 2\}$, $D_1 = \sum_s a_s s^{p+1}$, $D_2 = \sum_s a_s s^{p+2}$. As a result, when $j - k = h$,

$$\begin{aligned}
Cov(Z^u_{n,12}(j), Z^v_{n,12}(k)) &= E[Z^u_{n,12}(j) Z^v_{n,12}(k)] - E[Z^u_{n,12}(j)] E[Z^v_{n,12}(k)] \\
&= n^{2\alpha_{12}-(\alpha_{11}+\alpha_{22})} \Big(E[(Y^u_{n,1})_j (Y^v_{n,1})_k] E[(Y^u_{n,2})_j (Y^v_{n,2})_k] \\
&\quad + E[(Y^u_{n,1})_j (Y^v_{n,2})_k] E[(Y^v_{n,1})_k (Y^u_{n,2})_j] \Big) \\
&= n^{2\alpha_{12}} \left(\frac{2(uv)^{p+1} D_1}{(2p+2)! n^{2p+2}}\right)^2 (F^{uv}_{n,12}(h)^2 + F^{uv}_{n,11}(h) F^{uv}_{n,22}(h)),
\end{aligned} \tag{2.9}$$

where for $i, r \in \{1, 2\}$,

$$F^{uv}_{n,ir}(h) = D_1 C_{ir}^{(2p+2)}\left(\frac{h}{n}\right) + \frac{u+v}{n(2p+3)} D_2 C_{ir}^{(2p+3)}\left(\frac{h^*}{n}\right).$$

As $h/n \to 0$,

$$F^{uv}_{n,12}(h)^2 = \left(\frac{h}{n}\right)^{2\alpha_{12}-(4p+4)} \left(A_{12}\frac{\alpha_{12}!}{(2p+2)!}\right)^2 \left(D_1 D_2 \frac{u+v}{2p+3} 2(\alpha_{12}-2p-2)|h|^{-1} + D_1^2 + D_2^2 \frac{(u+v)^2}{(2p+3)^2}(\alpha_{12}-2p-2)^2|h|^{-2}\right)(1+o(1)),$$

$$F^{uv}_{n,11}(h)F^{uv}_{n,22}(h) = \left(\frac{h}{n}\right)^{\alpha_{11}+\alpha_{22}-(4p+4)} A_{11}A_{22}\frac{\alpha_{11}!}{(2p+2)!}\frac{\alpha_{22}!}{(2p+2)!}\left(D_1 D_2 \frac{u+v}{2p+3}(\alpha_{11} + \alpha_{22} - 4p - 4)|h|^{-1} + D_1^2 + D_2^2\frac{(u+v)^2}{(2p+3)^2}(\alpha_{11}-2p-2)(\alpha_{22} - 2p - 2)|h|^{-2}\right)(1+o(1)).$$

It was shown in the proof of Theorem 1 in Kent and Wood (1997) that as $n \to \infty$,

$$\sum_{h=-n+1}^{n-1}\left(1-\frac{|h|}{n}\right)|h|^a = \begin{cases} O(1), & \text{if } a < -1; \\ O(n^{a+1}), & \text{if } a > -1. \end{cases}$$

Hence, as $n \to \infty$,

$$\begin{aligned} Cov(\bar{Z}^u_{n,12}, \bar{Z}^v_{n,12}) &= \frac{1}{n}\sum_{h=-n+1}^{n-1}\left(1-\frac{|h|}{n}\right) Cov(Z^u_{n,12}(0), Z^v_{n,12}(h)) \\ &= n^{2\alpha_{12}-(4p+4)-1}\left(\frac{2(uv)^{p+1}D_1}{(2p+2)!}\right)^2 \\ &\quad \sum_{h=-n+1}^{n-1}\left(1-\frac{|h|}{n}\right)\left(F^{uv}_{n,12}(h)^2 + F^{uv}_{n,11}(h)F^{uv}_{n,22}(h)\right) \\ &= \begin{cases} O(n^{2\alpha_{12}-(\alpha_{11}+\alpha_{22})-1}), & \text{if } \alpha_{11}+\alpha_{22} < 4p+3; \\ O(n^{2\alpha_{12}-(4p+4)}), & \text{if } \alpha_{11}+\alpha_{22} > 4p+3. \end{cases} \end{aligned} \tag{2.10}$$

It is induced from (2.7) and (2.10) that, when $\alpha_{11}+\alpha_{22} < 2\alpha_{12} < \alpha_{11}+\alpha_{22}+1 < 4p+4$ or $4p+3 < \alpha_{11}+\alpha_{22} < 2\alpha_{12} < 4p+4$, $\bar{Z}^u_{n,12} \xrightarrow{P} Au^{\alpha_{12}}$ as $n \to \infty$. □

Remark. Under the conditions of Theorem 1, we have natural consequences as follows.

(i) Take $p = 0$, then for $\alpha_{11}+\alpha_{22} < 3$, $\bar{Z}^u_{n,12} \xrightarrow{P} Au^{\alpha_{12}}$ as $n \to \infty$ if $\alpha_{11}+\alpha_{22} < 2\alpha_{12} < \alpha_{11}+\alpha_{22}+1$; for $\alpha_{11}+\alpha_{22} > 3$, the convergence holds if $\alpha_{11}+\alpha_{22} < 2\alpha_{12} < 4$.

(ii) Take $p \geq 1$, then for any $\alpha_1, \alpha_2 \in (0,2)$, $\bar{Z}^u_{n,12} \xrightarrow{P} Au^{\alpha_{12}}$ as $n \to \infty$ if $\alpha_{11} + \alpha_{22} < 2\alpha_{12} < \alpha_{11} + \alpha_{22} + 1$.

The convergence in probability in Theorem 1 can be strengthened to almost sure convergence by applying the following lemma and the Borel–Cantelli Lemma.

Lemma 1. *Under conditions in Theorem 1, $\forall u = 1, \ldots, m$, there exists a constant $C \in (0, \infty)$ independent of n such that for all large enough n and $\forall 0 < \xi < 1$,*

$$P\left(\left|\frac{(\bar{Z}^u_{n,12})^2 - E(\bar{Z}^u_{n,12})^2}{E(\bar{Z}^u_{n,12})^2}\right| > \xi\right) \leq C \exp\left(-n^{\min\{\alpha_{11}+\alpha_{22}+1, 4p+4\}/2-\alpha_{12}} \frac{\xi}{4-\xi}\right). \tag{2.11}$$

Proof. For $n \geq 1$ and $u = 1, \ldots, m$, denote

$$M^u_n = \frac{1}{2} n^{\alpha_{12}-(\alpha_{11}+\alpha_{22})/2-1} (\Sigma_Y^{1/2})^T \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix} \Sigma_Y^{1/2},$$

then according to (2.5), $\bar{Z}^u_{n,12} \stackrel{\mathrm{d}}{=} U^T M^u_n U$, where $U \sim N(0, I_{2n})$. By the Hanson-Wright inequality, there exists constants C_1, C_2 that do not depend on n or u such that $\forall 0 < \xi < 1$,

$$P\left(\left|\frac{\bar{Z}^u_{n,12} - E\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}}\right| > \xi\right) \leq 2 \exp\left(-\min\left\{\frac{C_1 \xi |E\bar{Z}^u_{n,12}|}{||M^u_n||_2}, \frac{C_2 \xi^2 |E\bar{Z}^u_{n,12}|^2}{||M^u_n||_F^2}\right\}\right).$$

Under the conditions in Theorem 1, as $n \to \infty$ there is

$$||M^u_n||_F^2 = tr((M^u_n)^2) = var(\bar{Z}^u_{n,12})/2 = \begin{cases} O(n^{2\alpha_{12}-(\alpha_{11}+\alpha_{22})-1}), & \text{if } \alpha_{11}+\alpha_{22} < 4p+3; \\ O(n^{2\alpha_{12}-(4p+4)}), & \text{if } \alpha_{11}+\alpha_{22} > 4p+3. \end{cases} \tag{2.12}$$

Since $E\bar{Z}^u_{n,12} \to Au^{\alpha_{12}}$ as $n \to \infty$ and $||M^u_n||_2 \leq ||M^u_n||_F$, there exists a constant $C_0 \in (0, \infty)$ that does not depend on n but may depend on u such that

$$P\left(\left|\frac{\bar{Z}^u_{n,12} - E\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}}\right| > \xi\right) \leq C_0 \exp\left(-n^{\min\{\alpha_{11}+\alpha_{22}+1, 4p+4\}/2-\alpha_{12}} \xi\right). \tag{2.13}$$

Under the conditions in Theorem 1,

$$\frac{(E\bar{Z}^u_{n,12})^2}{E(\bar{Z}^u_{n,12})^2} = \frac{E(\bar{Z}^u_{n,12})^2 - var(\bar{Z}^u_{n,12})}{E(\bar{Z}^u_{n,12})^2} \to 1 \quad \text{as } n \to \infty.$$

Thus, $\forall 0 < \xi < 1$, $1 - \xi/2 < (E\bar{Z}^u_{n,12})^2/E(\bar{Z}^u_{n,12})^2 < 1 + \xi/2$ when n is large enough. Together with (2.13) it implies

$$
\begin{aligned}
&P\left(\left|\frac{(\bar{Z}^u_{n,12})^2 - E(\bar{Z}^u_{n,12})^2}{E(\bar{Z}^u_{n,12})^2}\right| > \xi\right)\\
\le& P\left(\frac{(E\bar{Z}^u_{n,12})^2}{E(\bar{Z}^u_{n,12})^2}\left|\left(\frac{\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}}\right)^2 - 1\right| + \left|\frac{(E\bar{Z}^u_{n,12})^2}{E(\bar{Z}^u_{n,12})^2} - 1\right| > \xi\right)\\
=& P\left(\left|\left(\frac{\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}}\right)^2 - 1\right| > \frac{\xi + (E\bar{Z}^u_{n,12})^2/E(\bar{Z}^u_{n,12})^2 - 1}{(E\bar{Z}^u_{n,12})^2/E(\bar{Z}^u_{n,12})^2}\right)\\
\le& P\left(\left|\left(\frac{\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}}\right)^2 - 1\right| > \frac{\xi - \xi/2}{1 - \xi/2}\right) \quad \text{for large } n\\
=& P\left(\left|\left(\frac{\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}}\right)^2 - 1\right| > \frac{\xi}{2 - \xi}\right)\\
\le& P\left(\left|\frac{\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}} - 1\right| \cdot \left|\frac{\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}} + 1\right| > \frac{\xi}{2 - \xi}, \left|\frac{\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}} - 1\right| \le \frac{\xi}{2 - \xi}\right)\\
&+ P\left(\left|\frac{\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}} - 1\right| > \frac{\xi}{2 - \xi}\right)\\
\le& P\left(\left|\frac{\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}} - 1\right| > \frac{\xi/(2 - \xi)}{2 + \xi/(2 - \xi)}\right) + P\left(\left|\frac{\bar{Z}^u_{n,12}}{E\bar{Z}^u_{n,12}} - 1\right| > \frac{\xi}{2 - \xi}\right)\\
\le& C \exp\left(-n^{\min\{\alpha_{11}+\alpha_{22}+1, 4p+4\}/2 - \alpha_{12}} \frac{\xi}{4 - \xi}\right)
\end{aligned}
$$

for some constant $C \in (0, \infty)$ that is independent of n and ξ but may depend on u. □

The joint asymptotic distribution of the covariations is presented in the following theorem.

Theorem 2. *Denote by* $\bar{Z}_{n,12} = (\bar{Z}^1_{n,12}, \ldots, \bar{Z}^m_{n,12})^T$ *and take* $p \ge 1$. *When* $\alpha_{11} + \alpha_{22} < 2\alpha_{12}$ *and (2.4) holds for* $q = 2p + 2$,

$$n^{1/2 + (\alpha_{11}+\alpha_{22})/2 - \alpha_{12}} (\bar{Z}_{n,12} - E\bar{Z}_{n,12}) \xrightarrow{d} N(0, \Phi) \tag{2.14}$$

as $n \to \infty$, *where the matrix* $\Phi \in \mathbb{R}^{m\times m}$ *has entries*

$$\Phi_{u,v} = A_{11}A_{22} \sum_{h=-\infty}^{\infty} \sum_{s,t,j,l=-J}^{J} a_s a_t a_j a_l |h + su - tv|^{\alpha_{11}} |h + ju - lv|^{\alpha_{22}}, \quad 1 \le u, v \le m. \tag{2.15}$$

Proof. By the Cramér-Wold theorem, to prove the asymptotic normality of $\bar{Z}_{n,12}$, it suffices to show that $\forall \boldsymbol{\gamma} \in \mathbb{R}^m$,

$$n^{1/2+(\alpha_{11}+\alpha_{22})/2-\alpha_{12}} \boldsymbol{\gamma}^T (\bar{Z}_{n,12} - E\bar{Z}_{n,12}) \xrightarrow{d} N(0, \boldsymbol{\gamma}^T \Phi \boldsymbol{\gamma}) \tag{2.16}$$

as $n \to \infty$.

Denote by

$$W_n = (Y_{n,1}^1(1), \ldots, Y_{n,1}^m(1), Y_{n,1}^1(2), \ldots, Y_{n,1}^m(n), Y_{n,2}^1(1), \ldots, Y_{n,2}^m(n))^T \in \mathbb{R}^{2mn}, \tag{2.17}$$

then

$$n^{1/2+(\alpha_{11}+\alpha_{22})/2-\alpha_{12}} \boldsymbol{\gamma}^T \bar{Z}_{n,12} = \frac{1}{2} n^{-1/2} W_n^T \begin{pmatrix} 0 & diag(1_n \otimes \boldsymbol{\gamma}) \\ diag(1_n \otimes \boldsymbol{\gamma}) & 0 \end{pmatrix} W_n,$$

where $diag(x)$ maps a vector x to a diagonal matrix whose diagonal is x, $1_n \in \mathbb{R}^n$ is a vector with all its entries equals 1. Let $V_n = Cov(W_n)$ and

$$G_n = \frac{1}{2} n^{-1/2} (V_n^{1/2})^T \begin{pmatrix} 0 & diag(1_n \otimes \boldsymbol{\gamma}) \\ diag(1_n \otimes \boldsymbol{\gamma}) & 0 \end{pmatrix} V_n^{1/2}, \tag{2.18}$$

then $n^{1/2+(\alpha_{11}+\alpha_{22})/2-\alpha_{12}} \boldsymbol{\gamma}^T \bar{Z}_{n,12} \stackrel{d}{=} \epsilon_n^T G_n \epsilon_n \stackrel{d}{=} \epsilon_n^T diag(\text{eig}(G_n)) \epsilon_n$ for $\epsilon_n \sim N(0, I_{2mn})$.

It follows from the proof of Theorem 2 in Zhou and Xiao (2018) that (2.16) holds if $\text{Tr}(G_n^4) \to 0$ and $2\text{Tr}(G_n^2) \to \boldsymbol{\gamma}^T \Phi \boldsymbol{\gamma}$ as $n \to \infty$.

Let

$$H_n = V_n \begin{pmatrix} 0 & diag(1_n \otimes \boldsymbol{\gamma}) \\ diag(1_n \otimes \boldsymbol{\gamma}) & 0 \end{pmatrix},$$

then for $i_1, i_2 \in \{1, 2\}$, $j_1, j_2 \in \{1, \ldots, n\}$ and $k_1, k_2 \in \{1, \ldots, m\}$,

$$H_n\left((i_1 - 1)mn + (j_1 - 1)m + k_1, (i_2 - 1)mn + (j_2 - 1)m + k_2\right) = \gamma_{k_2} \sigma_{n, i_1(3-i_2)}^{k_1 k_2}(j_2 - j_1).$$

Thus,

$$\begin{aligned}
\mathrm{Tr}(H_n^4) = & \sum_{k_1,\ldots,k_4=1}^{m} \gamma_{k_1}\gamma_{k_2}\gamma_{k_3}\gamma_{k_4} \sum_{i_1,\ldots,i_4=1}^{2} \sum_{j_1,\ldots,j_4=1}^{n} \Big(\sigma_{n,i_1(3-i_2)}^{k_1k_2}(j_2-j_1) \\
& \sigma_{n,i_2(3-i_3)}^{k_2k_3}(j_3-j_2)\sigma_{n,i_3(3-i_4)}^{k_3k_4}(j_4-j_3)\sigma_{n,i_4(3-i_1)}^{k_4k_1}(j_1-j_4)\Big) \\
\le & \sum_{k_1,\ldots,k_4=1}^{m} |\gamma_{k_1}\gamma_{k_2}\gamma_{k_3}\gamma_{k_4}| \sum_{i_1,\ldots,i_4=1}^{2} n \sum_{|h_1|,|h_2|,|h_3|<n} \Big| \sigma_{n,i_1(3-i_2)}^{k_1k_2}(h_1) \\
& \sigma_{n,i_2(3-i_3)}^{k_2k_3}(h_2)\sigma_{n,i_3(3-i_4)}^{k_3k_4}(h_3)\sigma_{n,i_4(3-i_1)}^{k_4k_1}(h_1+h_2+h_3)\Big|,
\end{aligned}$$

$$\begin{aligned}
\mathrm{Tr}(H_n^2) &= 2\sum_{k_1,k_2=1}^{m}\gamma_{k_1}\gamma_{k_2}\sum_{j_1,j_2=1}^{n}\left(\left(\sigma_{n,12}^{k_1k_2}(j_2-j_1)\right)^2 + \sigma_{n,11}^{k_1k_2}(j_2-j_1)\sigma_{n,22}^{k_1k_2}(j_2-j_1)\right) \\
&= 2n\sum_{k_1,k_2=1}^{m}\gamma_{k_1}\gamma_{k_2}\sum_{|h|<n}\left(1-\frac{|h|}{n}\right)\left(\left(\sigma_{n,12}^{k_1k_2}(h)\right)^2 + \sigma_{n,11}^{k_1k_2}(h)\sigma_{n,22}^{k_1k_2}(h)\right).
\end{aligned}$$

For any fixed h, the convergence of $\sigma_{n,ir}^{uv}(h)$ as $n \to \infty$ is presented in (2.6). By Theorem 1 in Kent and Wood (1997) and Lemma 2 in Zhou and Xiao (2018), when $\alpha_{11}+\alpha_{22} < 2\alpha_{12}$ and (2.4) holds for $q = 2p+2$,

$$\sigma_{n,ii}^{uv}(h) = O(|h|^{\alpha_{ii}-2p-2}) \quad \text{and} \quad \sigma_{n,12}^{uv}(h) = O(|h|^{(\alpha_{11}+\alpha_{22})/2-2p-2}) \tag{2.19}$$

uniformly for $n > |h|$. If $p \ge 1$, then $\alpha_{ii} - 2p - 2 < -2$ and $(\alpha_{11}+\alpha_{22})/2 - 2p - 2 < -2$ hold for any $\alpha_{11}, \alpha_{22} \in (0,2)$. Hence there exists a constant $c_0 > 0$ such that

$$\begin{aligned}
& \sum_{h_1,h_2,h_3=1-n}^{n-1} \left| \sigma_{n,i_1(3-i_2)}^{k_1k_2}(h_1)\sigma_{n,i_2(3-i_3)}^{k_2k_3}(h_2)\sigma_{n,i_3(3-i_4)}^{k_3k_4}(h_3)\sigma_{n,i_4(3-i_1)}^{k_4k_1}(h_1+h_2+h_3)\right| \\
& \le c_0 \sum_{h_1,h_2,h_3=1-n}^{n-1} \Bigg(|h_1|^{\frac{\alpha_{i_1i_1}+\alpha_{(3-i_2)(3-i_2)}}{2}-2p-2} |h_2|^{\frac{\alpha_{i_2i_2}+\alpha_{(3-i_3)(3-i_3)}}{2}-2p-2} \\
& \qquad |h_3|^{\frac{\alpha_{i_3i_3}+\alpha_{(3-i_4)(3-i_4)}}{2}-2p-2}\Bigg) \\
& = O(1)
\end{aligned}$$

as $n \to \infty$, $\forall i_1, i_2, i_3, i_4 \in \{1,2\}$. Consequently, $\mathrm{Tr}(H_n^4) = O(n)$ and

$$\mathrm{Tr}(G_n^4) = \left(\frac{1}{2}n^{-1/2}\right)^4 \mathrm{Tr}(H_n^4) = O(n^{-1}) \to 0$$

as $n \to \infty$.

For $u, v \in \{1, \dots, m\}$ and $h \in \mathbb{Z}$, define

$$d_n^{uv}(h) := 1_{|h|<n} \left(1 - \frac{|h|}{n}\right) \left(\left(\sigma_{n,12}^{uv}(h)\right)^2 + \sigma_{n,11}^{uv}(h)\sigma_{n,22}^{uv}(h)\right).$$

Then for any fixed h,

$$d_n^{uv}(h) \to A_{11}A_{22} \sum_{s,t,j,l=-J}^{J} a_s a_t a_j a_l |h + su - tv|^{\alpha_{11}} |h + ju - lv|^{\alpha_{22}}$$

as $n \to \infty$. Moreover,

$$d_n^{uv}(h) \le \left(\sigma_{n,12}^{uv}(h)\right)^2 + \sigma_{n,11}^{uv}(h)\sigma_{n,22}^{uv}(h) = O(|h|^{\alpha_{11}+\alpha_{22}-4p-4})$$

uniformly for $n > |h|$. If $p \ge 1$, then $\alpha_{11} + \alpha_{22} - 4p - 4 < -4$ and $\sum_{h=-\infty}^{\infty} |h|^{\alpha_{11}+\alpha_{22}-4p-4} < \infty$. Thus for any $u, v \in \{1, \dots, m\}$, $\{d_n^{uv}(h), h \in \mathbb{Z}\}$ is dominated by a summable sequence. It therefore follows from the dominated convergence theorem that

$$\begin{aligned}
\mathrm{Tr}(G_n^2) &= \frac{1}{4n}\mathrm{Tr}(H_n^2) \\
&= \frac{1}{2} \sum_{k_1,k_2=1}^{m} \gamma_{k_1}\gamma_{k_2} \sum_{h=-\infty}^{\infty} d_n^{k_1k_2}(h) \\
&\to \frac{A_{11}A_{22}}{2} \sum_{k_1,k_2=1}^{m} \gamma_{k_1}\gamma_{k_2} \sum_{h=-\infty}^{\infty} \sum_{s,t,j,l=-J}^{J} a_s a_t a_j a_l |h + sk_1 - tk_2|^{\alpha_{11}} |h + jk_1 - lk_2|^{\alpha_{22}} \\
&:= \frac{1}{2}\boldsymbol{\gamma}^T \Phi \boldsymbol{\gamma} \quad \text{as } n \to \infty,
\end{aligned}$$

where $\Phi \in \mathbb{R}^{m \times m}$ is a constant matrix with entries defined in (2.15).

This proves Theorem 2. □

Take $p = 1$, $J = 1$, and $a = (1, -2, 1)^T$, we further discuss the joint asymptotic distribution of covariations defined in this chapter and the quadratic variations $\bar{Z}_{n,1}$, $\bar{Z}_{n,2}$ studied by Zhou and Xiao (2018), where $\bar{Z}_{n,i} = (\bar{Z}_{n,i}^1, \dots, \bar{Z}_{n,i}^m)^T$ and

$$\bar{Z}_{n,i}^u = \frac{1}{n}(Y_{n,i}^u)^T Y_{n,i}^u, \quad u = 1, \dots, m, \quad i = 1, 2. \tag{2.20}$$

Theorem 3. *When $\alpha_{11} + \alpha_{22} < 2\alpha_{12}$ and (2.4) holds for $q = 4$,*

$$n^{D_\alpha} \begin{pmatrix} \bar{Z}_{n,1} - E\bar{Z}_{n,1} \\ \bar{Z}_{n,2} - E\bar{Z}_{n,1} \\ \bar{Z}_{n,12} - E\bar{Z}_{n,12} \end{pmatrix} \xrightarrow{d} N\left(0, \begin{pmatrix} \Phi_1 & & \\ & \Phi_2 & \\ & & \Phi \end{pmatrix}\right) \tag{2.21}$$

as $n \to \infty$, where

$$D_\alpha = \begin{pmatrix} \frac{1}{2} & & \\ & \frac{1}{2} & \\ & & \frac{1+\alpha_{11}+\alpha_{22}}{2} - \alpha_{12} \end{pmatrix},$$

the matrix $\Phi \in \mathbb{R}^{m\times m}$ is as defined in Theorem 2, and matrices $\Phi_i \in \mathbb{R}^{m\times m}$ have entries as

$$(\Phi_i)_{u,v} = 2A_{ii}^2 \sum_{h=-\infty}^{\infty} \left(\sum_{s,t=-1}^{1} a_s a_t |h + su - tv|^{\alpha_{ii}} \right)^2, \quad i = 1, 2. \tag{2.22}$$

Proof. By the Cramér-Wold theorem, it suffices to prove that $\forall \boldsymbol{\gamma}_1 = (\gamma_{1,1}, \ldots, \gamma_{1,m})^T$, $\boldsymbol{\gamma}_2 = (\gamma_{2,1}, \ldots, \gamma_{2,m})^T$, and $\boldsymbol{\gamma}_{12} = (\gamma_{12,1}, \ldots, \gamma_{12,m})^T \in \mathbb{R}^m$,

$$\begin{aligned} &\sqrt{n}\left(\boldsymbol{\gamma}_1^T(\bar{Z}_{n,1} - E\bar{Z}_{n,1}) + \boldsymbol{\gamma}_2^T(\bar{Z}_{n,2} - E\bar{Z}_{n,2}) + n^{\frac{\alpha_{11}+\alpha_{22}}{2} - \alpha_{12}} \boldsymbol{\gamma}_{12}^T(\bar{Z}_{n,12} - E\bar{Z}_{n,12})\right) \\ \xrightarrow{d} &N(0, \boldsymbol{\gamma}_1^T \Phi_1 \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2^T \Phi_2 \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_{12}^T \Phi \boldsymbol{\gamma}_{12}) \end{aligned} \tag{2.23}$$

as $n \to \infty$.

Recall the notation W_n defined in (2.17) and $V_n = \mathrm{Cov}(W_n)$, let

$$\Lambda_n = \frac{2}{\sqrt{n}} (V_n^{1/2})^T \Gamma_n V_n^{1/2}, \tag{2.24}$$

where

$$\Gamma_n = \begin{pmatrix} diag(1_n \otimes \boldsymbol{\gamma}_1) & 0 \\ 0 & diag(1_n \otimes \boldsymbol{\gamma}_2) \end{pmatrix}. \tag{2.25}$$

It follows from definitions of $\bar{Z}_{n,1}, \bar{Z}_{n,2}$, and $\bar{Z}_{n,12}$ that

$$\begin{aligned} n^{D_\alpha} \begin{pmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \\ \boldsymbol{\gamma}_{12} \end{pmatrix}^T \begin{pmatrix} \bar{Z}_{n,1} \\ \bar{Z}_{n,2} \\ \bar{Z}_{n,12} \end{pmatrix} &= \frac{1}{\sqrt{n}} W_n^T \begin{pmatrix} diag(1_n \otimes \boldsymbol{\gamma}_1) & \frac{1}{2} diag(1_n \otimes \boldsymbol{\gamma}_{12}) \\ \frac{1}{2} diag(1_n \otimes \boldsymbol{\gamma}_{12}) & diag(1_n \otimes \boldsymbol{\gamma}_2) \end{pmatrix} W_n \\ &\overset{d}{=} \epsilon_n^T \left(G_n + \frac{1}{2}\Lambda_n \right) \epsilon_n, \end{aligned}$$

where $\epsilon_n \sim N(0, I_{3mn})$ and G_n is defined in (2.18). Therefore, it remains to prove

$$\mathrm{Tr}\left((G_n + \frac{1}{2}\Lambda_n)^2\right) \to \frac{1}{2}\left(\boldsymbol{\gamma}_1^T\Phi_1\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2^T\Phi_2\boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_{12}^T\Phi\boldsymbol{\gamma}_{12}\right)$$

and

$$\mathrm{Tr}\left((G_n + \frac{1}{2}\Lambda_n)^4\right) \to 0$$

as $n \to \infty$.

It has been proved by Zhou and Xiao (2018) that as $n \to \infty$,

$$\mathrm{Tr}(\Lambda_n^2) \to 2\left(\boldsymbol{\gamma}_1^T\Phi_1\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2^T\Phi_2\boldsymbol{\gamma}_2\right) \quad \text{and} \quad \mathrm{Tr}(\Lambda_n^4) \to 0$$

when $\alpha_{11} + \alpha_{22} < 2\alpha_{12}$ and (2.4) holds for $q = 4$. Since conditions in Theorem 2 are satisfied, we also have

$$\mathrm{Tr}(G_n^2) \to \frac{1}{2}\boldsymbol{\gamma}_{12}^T\Phi\boldsymbol{\gamma}_{12} \quad \text{and} \quad \mathrm{Tr}(G_n^4) \to 0$$

as $n \to \infty$. Moreover,

$$\begin{aligned}
\mathrm{Tr}(G_n\Lambda_n) &= \frac{1}{n}\mathrm{Tr}\left(V_n\begin{pmatrix} 0 & diag(1_n \otimes \boldsymbol{\gamma}_{12}) \\ diag(1_n \otimes \boldsymbol{\gamma}_{12}) & 0 \end{pmatrix}V_n\Gamma_n\right) \\
&= \frac{1}{n}\sum_{\ell_1,\ell_2=1}^{2mn}(H_n)_{\ell_1,\ell_2}(V_n\Gamma_n)_{\ell_2,\ell_1} \\
&= \frac{1}{n}\sum_{k_1,k_2=1}^{m}\sum_{i_1,i_2=1}^{2}\gamma_{i_1,k_1}\gamma_{12,k_2}\sum_{j_1,j_2=1}^{n}\sigma_{n,i_1(3-i_2)}^{k_1k_2}(j_2-j_1)\sigma_{n,i_2i_1}^{k_2k_1}(j_2-j_1) \\
&= \sum_{k_1,k_2=1}^{m}\sum_{i_1,i_2=1}^{2}\gamma_{i_1,k_1}\gamma_{12,k_2}\sum_{|h|<n}\left(1-\frac{|h|}{n}\right)\sigma_{n,i_1(3-i_2)}^{k_1k_2}(h)\sigma_{n,i_2i_1}^{k_2k_1}(h) \\
&\to 0 \quad \text{as } n \to \infty
\end{aligned} \tag{2.26}$$

by the dominated convergence theorem, since $\sigma_{n,12}^{uv}(h) \to 0$ as $n \to \infty$ for any $u, v = 1, \ldots, m$ and any fixed h. Due to the fact that

$$\mathrm{Card}\{(j_1, \ldots, j_4) : 1 \le j_1, \ldots, j_4 \le n, j_{i+1} - j_i = h_i \ (i = 1, 2, 3)\} \le n,$$

we have

$$
\begin{aligned}
\operatorname{Tr}\left((G_n\Lambda_n)^2\right) &= \frac{1}{n^2}\operatorname{Tr}\left((H_nV_n\Gamma_n)^2\right) \\
&= \frac{1}{n^2}\sum_{\ell_1,\ldots,\ell_4=1}^{2mn}(H_n)_{\ell_1,\ell_2}(V_n\Gamma_n)_{\ell_2,\ell_3}(H_n)_{\ell_3,\ell_4}(V_n\Gamma_n)_{\ell_4,\ell_1} \\
&= \frac{1}{n^2}\sum_{k_1,\ldots,k_4=1}^{m}\left(\sum_{i_1,\ldots,i_4=1}^{2}\gamma_{i_1,k_1}\gamma_{12,k_2}\gamma_{i_3,k_3}\gamma_{12,k_4}\right. \\
&\quad \left.\sum_{j_1,\ldots,j_4=1}^{n}\sigma_{n,i_1(3-i_2)}^{k_1k_2}(j_2-j_1)\sigma_{n,i_2i_3}^{k_2k_3}(j_2-j_3)\sigma_{n,i_3(3-i_4)}^{k_3k_4}(j_4-j_3)\sigma_{n,i_4i_1}^{k_4k_1}(j_4-j_1)\right) \\
&\le \frac{1}{n}\sum_{k_1,\ldots,k_4=1}^{m}\left(\sum_{i_1,\ldots,i_4=1}^{2}\gamma_{i_1,k_1}\gamma_{12,k_2}\gamma_{i_3,k_3}\gamma_{12,k_4}\right. \\
&\quad \left.\sum_{h_1,h_2,h_3=1-n}^{n-1}\sigma_{n,i_1(3-i_2)}^{k_1k_2}(h_1)\sigma_{n,i_2i_3}^{k_2k_3}(h_2)\sigma_{n,i_3(3-i_4)}^{k_3k_4}(h_3)\sigma_{n,i_4i_1}^{k_4k_1}(h_1+h_2+h_3)\right).
\end{aligned}
$$

Follow similar steps in the proof of Theorem 2, there exists a constant $c_0 > 0$ such that

$$
\begin{aligned}
\operatorname{Tr}\left((G_n\Lambda_n)^2\right) &\le \frac{c_0}{n}\sum_{k_1,\ldots,k_4=1}^{m}\left(\sum_{i_1,\ldots,i_4=1}^{2}|\gamma_{12,k_2}\gamma_{i_3,k_3}\gamma_{12,k_4}|\right. \\
&\quad \left.\sum_{h_1,h_2,h_3=1-n}^{n-1}|h_1|^{\frac{\alpha_{i_1i_1}+\alpha_{(3-i_2)(3-i_2)}}{2}-4}|h_2|^{\frac{\alpha_{i_2i_2}+\alpha_{i_3i_3}}{2}-4}|h_3|^{\frac{\alpha_{i_3i_3}+\alpha_{(3-i_4)(3-i_4)}}{2}-4}\right) \\
&= O(n^{-1}) \quad \text{as } n\to\infty,
\end{aligned}
$$

since $\forall i, j = 1, 2,\ \frac{1}{2}(\alpha_{ii}+\alpha_{jj}) - 4 < -2$.

Consequently, as $n \to \infty$,

$$
\begin{aligned}
\operatorname{Tr}\left((G_n+\frac{1}{2}\Lambda_n)^2\right) &= \operatorname{Tr}(G_n^2) + \frac{1}{4}\operatorname{Tr}(\Lambda_n^2) + \operatorname{Tr}(G_n\Lambda_n) \\
&\to \frac{1}{2}\left(\boldsymbol{\gamma}_1^T\Phi_1\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2^T\Phi_2\boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_{12}^T\Phi\boldsymbol{\gamma}_{12}\right),
\end{aligned}
$$

where entries of Φ_1, Φ_2, and Φ are defined in (2.22) and (2.15). The Cauchy–Schwarz inequality implies that

$$\begin{aligned}\mathrm{Tr}\left((G_n+\frac{1}{2}\Lambda_n)^4\right) &= \mathrm{Tr}(G_n^4)+2\mathrm{Tr}(G_n^3\Lambda_n)+\frac{1}{2}\mathrm{Tr}((G_n\Lambda_n)^2)+\mathrm{Tr}(G_n^2\Lambda_n^2)\\ &\quad+\frac{1}{2}\mathrm{Tr}(G_n\Lambda_n^3)+\frac{1}{2^4}\mathrm{Tr}(\Lambda_n^4)\\ &\le \mathrm{Tr}(G_n^4)+2\sqrt{\mathrm{Tr}(G_n^6)\mathrm{Tr}(\Lambda_n^2)}+\frac{1}{2}\mathrm{Tr}((G_n\Lambda_n)^2)+\sqrt{\mathrm{Tr}(G_n^4)\mathrm{Tr}(\Lambda_n^4)}\\ &\quad+\frac{1}{2}\sqrt{\mathrm{Tr}(G_n^2)\mathrm{Tr}(\Lambda_n^6)}+\frac{1}{2^4}\mathrm{Tr}(\Lambda_n^4)\\ &\to 0\end{aligned}$$

as $n\to\infty$. This finishes the proof using the convergence of the moment generating function. □

2.2.2 Convergence of Estimator

Define the estimator of α_{12} as

$$\hat{\alpha}_{12}=\frac{1}{2}\sum_{u=1}^{m}L_u\log(\bar{Z}_{n,12}^u)^2, \tag{2.27}$$

where $\{L_u, u=1,\dots,m\}$ is a list of constants satisfying $\sum_{u=1}^m L_u=0$ and $\sum_{u=1}^m L_u\log u=1$. Plug in the definition of $\bar{Z}_{n,12}^u$ given in (2.5), then $\hat{\alpha}_{12}$ is a function of the observed process X_n^u and increment a only, written as

$$\begin{aligned}\hat{\alpha}_{12}&=\frac{1}{2}\sum_{u=1}^{m}L_u\log\left(\frac{1}{2}n^{\alpha_{12}-1}X_n^{uT}\begin{pmatrix}0 & I_n\otimes(aa^T)\\ I_n\otimes(aa^T) & 0\end{pmatrix}X_n^u\right)^2\\ &=\frac{1}{2}\sum_{u=1}^{m}L_u\log\left(X_n^{uT}\begin{pmatrix}0 & I_n\otimes(aa^T)\\ I_n\otimes(aa^T) & 0\end{pmatrix}X_n^u\right)^2,\end{aligned} \tag{2.28}$$

where $X_n^u=((X_{n,1}^u)^T,(X_{n,2}^u)^T)^T$.

Theorem 4. *Assume the increment $a=(a_{-J},a_{1-J},\dots,a_J)^T$ of order p satisfies*

$$\sum_{k,l=-J}^{J}a_ka_l|k-l|^{\alpha_{12}}\neq 0,$$

and (2.4) holds for $q=2p+3$ and $i,j\in\{1,2\}$. If $\alpha_{11}+\alpha_{22}<2\alpha_{12}<\alpha_{11}+\alpha_{22}+1<4p+4$ or $4p+3<\alpha_{11}+\alpha_{22}<2\alpha_{12}<4p+4$, then $\hat{\alpha}_{12}\xrightarrow{a.s.}\alpha_{12}$ as $n\to\infty$.

Proof. It follows from Lemma 1 and the Borel–Cantelli Lemma that $\forall u = 1, \ldots, m$,

$$\frac{(\bar{Z}^u_{n,12})^2}{E(\bar{Z}^u_{n,12})^2} \xrightarrow{a.s.} 1 \quad \text{as } n \to \infty.$$

When $\alpha_{11} + \alpha_{22} < 2\alpha_{12} < \alpha_{11} + \alpha_{22} + 1 < 4p + 4$ or $4p + 3 < \alpha_{11} + \alpha_{22} < 2\alpha_{12} < 4p + 4$, (2.7) and (2.10) imply that

$$E(\bar{Z}^u_{n,12})^2 = Cov(\bar{Z}^u_{n,12}) + (E\bar{Z}^u_{n,12})^2 \to A^2 u^{2\alpha_{12}},$$

where $A = -\rho\sigma_1\sigma_2 c_{12} \sum_{k,l} a_k a_l |k - l|^{\alpha_{12}}$. When $\sum_{k,l} a_k a_l |k - l|^{\alpha_{12}} \neq 0$, $\hat{\alpha}_{12}$ defined in (2.27) can be written as

$$\begin{aligned}
\hat{\alpha}_{12} &= \frac{1}{2}\sum_{u=1}^{m} L_u \left(\log \frac{(\bar{Z}^u_{n,12})^2}{E(\bar{Z}^u_{n,12})^2} + \log E(\bar{Z}^u_{n,12})^2\right) \\
&= \frac{1}{2}\sum_{u=1}^{m} L_u \log \frac{(\bar{Z}^u_{n,12})^2}{E(\bar{Z}^u_{n,12})^2} + \frac{1}{2}\sum_{u=1}^{m} L_u \log E(\bar{Z}^u_{n,12})^2 \\
&\xrightarrow{a.s.} \frac{1}{2}\sum_{u=1}^{m} L_u \log 1 + \frac{1}{2}\sum_{u=1}^{m} L_u \log(A^2 u^{2\alpha_{12}}) = \alpha_{12}
\end{aligned}$$

as $n \to \infty$ by the continuous mapping theorem. □

To derive the asymptotic normality of $\hat{\alpha}_{12}$, we further assume that as $t \to 0$,

$$C_{12}(t) = C_{21}(t) = \rho\sigma_1\sigma_2(1 - c_{12}|t|^{\alpha_{12}} + O(|t|^{\alpha_{12}+\beta_{12}})), \tag{2.29}$$

for some $\beta_{12} > 0$. It follows from (2.7) that $E[Z^u_{n,12}(j)] = Au^{\alpha_{12}} + O(n^{-\beta_{12}})$. The following corollary is straightforward when a further assumption is made on β_{12}.

Corollary 1. *Under conditions in Theorem 2, if* $\alpha_{12} + \beta_{12} > (\alpha_{11} + \alpha_{22} + 1)/2$*, then*

$$n^{1/2+(\alpha_{11}+\alpha_{22})/2-\alpha_{12}}(\bar{Z}_{n,12} - A\phi) \xrightarrow{d} N(0, \Phi) \tag{2.30}$$

as $n \to \infty$*, where* $\phi \in \mathbb{R}^m$ *and* $\phi_j = j^{\alpha_{12}}$, $j = 1, \ldots, m$.

The asymptotic normality of $\hat{\alpha}_{12}$ is then induced by the multivariate delta method.

Theorem 5. *Take $p \geq 1$ and assume (2.4) holds for $q = 2p+2$. When $A \neq 0$, if $\alpha_{11} + \alpha_{22} < 2\alpha_{12}$ and $\alpha_{12} + \beta_{12} > (\alpha_{11} + \alpha_{22} + 1)/2$, then*

$$n^{1/2+(\alpha_{11}+\alpha_{22})/2-\alpha_{12}}(\hat{\alpha}_{12} - \alpha_{12}) \xrightarrow{d} N(0, A^{-2}\tilde{L}^T\Phi\tilde{L}) \tag{2.31}$$

as $n \to \infty$, where $\tilde{L} = (L_1, L_2/2^{\alpha_{12}}, \ldots, L_m/m^{\alpha_{12}})^T \in \mathbb{R}^m$.

Proof. Define a mapping $f : \mathbb{R}^m \to \mathbb{R}$ by

$$f(x) = \frac{1}{2}\sum_{u=1}^{m} L_u \log x_u^2, \quad \forall x = (x_1, \ldots, x_m) \in \mathbb{R}^m.$$

Then $f(\bar{Z}_{n,12}) = \hat{\alpha}_{12}$, $f(A\phi) = \alpha_{12}$. When $A \neq 0$, f is continuously differentiable in a neighborhood of $A\phi$ and $\nabla f(A\phi) = A^{-1}\tilde{L}$.

Use the multivariate Taylor's theorem,

$$n^{1/2+(\alpha_{11}+\alpha_{22})/2-\alpha_{12}}(\hat{\alpha}_{12} - \alpha_{12}) = n^{1/2+(\alpha_{11}+\alpha_{22})/2-\alpha_{12}}\nabla f(A_n)(\bar{Z}_{n,12} - A\phi),$$

where $|A_n - A\phi| < |\bar{Z}_{n,12} - A\phi|$. As $n \to \infty$, Theorem 1 implies $\bar{Z}_{n,12} \xrightarrow{P} A\phi$, so we also have $A_n \xrightarrow{P} A\phi$. Applying the continuous mapping theorem, $\nabla f(A_n) \xrightarrow{P} \nabla f(A\phi)$. It follows from Corollary 1 and Slutsky's theorem that as $n \to \infty$,

$$n^{1/2+(\alpha_{11}+\alpha_{22})/2-\alpha_{12}}\nabla f(A_n)(\bar{Z}_{n,12} - A\phi) \xrightarrow{d} \nabla f(A\phi)N(0, \Phi) \stackrel{d}{=} N(0, A^{-2}\tilde{L}^T\Phi L).$$

This finishes the proof. □

Take $p = 1$, $J = 1$, and $a = (1, -2, 1)^T$. As was studied by Kent and Wood (1997) and Zhou and Xiao (2018), the estimators

$$\hat{\alpha}_{ii} = \sum_{u=1}^{m} L_{i,u} \log \bar{Z}_{n,i}^u, \quad i = 1, 2 \tag{2.32}$$

are strongly consistent and jointly converge in distribution to a multivariate Gaussian distribution, where $\bar{Z}_{n,i}^u$'s are defined in (2.20), $L_{i,u}$'s are constants such that $\sum_{u=1}^m L_{i,u} = 0$ and $\sum_{u=1}^m L_{i,u} \log u = 1$. The following theorem presents the joint asymptotic distribution of $\hat{\alpha}_{11}$, $\hat{\alpha}_{22}$, and $\hat{\alpha}_{12}$ as $n \to \infty$.

Theorem 6. *Assume that as* $|t| \to 0$, *(2.29) holds with* $\alpha_{12} + \beta_{12} > (\alpha_{11} + \alpha_{22} + 1)/2$, *and*

$$C_{ii}(t) = \sigma_i^2 - c_{ii}|t|^{\alpha_{ii}} + O(|t|^{\alpha_{ii}+\beta_{ii}}), \quad i = 1, 2$$

for some constants $\beta_{11}, \beta_{22} > 1/2$. *If* $2\alpha_{12} > \alpha_{11} + \alpha_{22}$, $\alpha_{12} \neq 2$, *and (2.4) holds for* $q = 4$, *then as* $n \to \infty$,

$$n^{D_\alpha} \begin{pmatrix} \hat{\alpha}_{11} - \alpha_{11} \\ \hat{\alpha}_{22} - \alpha_{22} \\ \hat{\alpha}_{12} - \alpha_{12} \end{pmatrix} \xrightarrow{d} N \left(0, \begin{pmatrix} A_1^{-2} \tilde{L}_1^T \Phi_1 \tilde{L}_1 & & \\ & A_2^{-2} \tilde{L}_2^T \Phi_2 \tilde{L}_2 & \\ & & A^{-2} \tilde{L}_3^T \Phi \tilde{L}_3 \end{pmatrix} \right), \tag{2.33}$$

where $A_i = c_{ii}(8 - 2^{\alpha_{ii}+1})$ *and* $\tilde{L}_i = (L_{i,1}, L_{i,2}/2^{\alpha_{ii}}, \ldots, L_{i,m}/m^{\alpha_{ii}})^T \in \mathbb{R}^m$ *for* $i = 1, 2$, $A = \rho\sigma_1\sigma_2 c_{12}(8 - 2^{\alpha_{12}+1})$, $\tilde{L}_3 = (L_{3,1}, L_{3,2}/2^{\alpha_{12}}, \ldots, L_{3,m}/m^{\alpha_{12}})^T \in \mathbb{R}^m$, *the matrices* $\Phi_1, \Phi_2, \Phi \in \mathbb{R}^{m\times m}$ *and* D_α *are as defined in Theorem 3.*

Proof. When $a = (1, -2, 1)^T$, we have

$$A = -\rho\sigma_1\sigma_2 c_{12} \sum_{k,l=-J}^{J} a_k a_l |k - l|^{\alpha_{12}} = \rho\sigma_1\sigma_2 c_{12}(8 - 2^{\alpha_{12}+1}).$$

It follows from (2.7) and Equation (14) in Zhou and Xiao (2018) that as $n \to \infty$,

$$n^{D_\alpha} \begin{pmatrix} E\bar{Z}_{n,1} - A_1\phi^1 \\ E\bar{Z}_{n,2} - A_2\phi^2 \\ E\bar{Z}_{n,12} - A\phi \end{pmatrix} = \begin{pmatrix} O\left(n^{1/2-\beta_{11}}\right) \\ O\left(n^{1/2-\beta_{22}}\right) \\ O\left(n^{(1+\alpha_{11}+\alpha_{22})/2-\alpha_{12}-\beta_{12}}\right) \end{pmatrix} \to 0 \tag{2.34}$$

if $\beta_{11}, \beta_{22} > 1/2$ and $\alpha_{12} + \beta_{12} > (\alpha_{11} + \alpha_{22} + 1)/2$, where $\phi^i = (1, 2^{\alpha_{ii}}, \ldots, m^{\alpha_{ii}})^T$ for $i = 1, 2$, and $\phi = (1, 2^{\alpha_{12}}, \ldots, m^{\alpha_{12}})^T$. Together with Theorem 3 this implies that

$$n^{D_\alpha} \begin{pmatrix} \bar{Z}_{n,1} - A_1\phi^1 \\ \bar{Z}_{n,2} - A_2\phi^2 \\ \bar{Z}_{n,12} - A\phi \end{pmatrix} \xrightarrow{d} N \left(0, \begin{pmatrix} \Phi_1 & & \\ & \Phi_2 & \\ & & \Phi \end{pmatrix} \right) \tag{2.35}$$

as $n \to \infty$.

Define a mapping $\mathrm{f} : \mathbb{R}_{>0}^{2m} \times \mathbb{R} \mapsto \mathbb{R}^3$ as

$$\mathrm{f(x)} = \begin{pmatrix} \sum_{u=1}^m L_{1,u} \log x_{1,u} \\ \sum_{u=1}^m L_{2,u} \log x_{2,u} \\ \frac{1}{2} \sum_{u=1}^m L_{3,u} \log x_{3,u}^2 \end{pmatrix}$$

for any x $= (x_{1,1}, \ldots, x_{1,m}, x_{2,1}, \ldots, x_{2,m}, x_{3,1}, \ldots, x_{3,m}) \in \mathbb{R}^{2m}_{>0} \times \mathbb{R}$, where $L_{i,u}$'s are constants such that $\sum_{u=1}^{m} L_{i,u} = 0$ and $\sum_{u=1}^{m} L_{i,u} \log u = 1$, $\forall i \in \{1,2,3\}$. Denote by $\bar{Z}_n = (\bar{Z}_{n,1}^T, \bar{Z}_{n,2}^T, \bar{Z}_{n,12}^T)^T$ and $\boldsymbol{\phi} = (A_1(\phi^1)^T, A_2(\phi^2)^T, A\phi^T)^T$, then

$$\mathrm{f}(\bar{Z}_n) = (\hat{\alpha}_{11}, \hat{\alpha}_{22}, \hat{\alpha}_{12})^T, \quad \mathrm{f}(\boldsymbol{\phi}) = (\alpha_{11}, \alpha_{22}, \alpha_{12})^T.$$

When $\alpha_{12} \neq 2$, $A = \rho\sigma_1\sigma_2 c_{12}(8 - 2^{\alpha_{12}+1}) \neq 0$ and f is thus continuously differentiable in a neighborhood of $\boldsymbol{\phi}$. Moreover, $\nabla \mathrm{f}(\boldsymbol{\phi}) = (A_1^{-1}\tilde{L}_1^T, A_2^{-1}\tilde{L}_2^T, A^{-1}\tilde{L}_3^T)^T$.

In a similar manner as in the proof of Theorem 5, it could be proved that as $n \to \infty$,

$$\begin{aligned} n^{D_\alpha}\begin{pmatrix} \hat{\alpha}_{11} - \alpha_{11} \\ \hat{\alpha}_{22} - \alpha_{22} \\ \hat{\alpha}_{12} - \alpha_{12} \end{pmatrix} &\xrightarrow{d} \nabla \mathrm{f}(\boldsymbol{\phi}) N\left(0, \begin{pmatrix} \Phi_1 & & \\ & \Phi_2 & \\ & & \Phi \end{pmatrix}\right) \\ &\stackrel{d}{=} N\left(0, \begin{pmatrix} A_1^{-2}\tilde{L}_1^T\Phi_1\tilde{L}_1 & & \\ & A_2^{-2}\tilde{L}_2^T\Phi_2\tilde{L}_2 & \\ & & A^{-2}\tilde{L}_3^T\Phi\tilde{L}_3 \end{pmatrix}\right). \end{aligned}$$

This finishes the proof. □

2.2.3 Simulation

Denote by M_ν the Matérn covariance function with parameter ν. Namely,

$$\begin{aligned} M_\nu(t) &= 2^{1-\nu}\Gamma(\nu)^{-1}|t|^\nu K_\nu(|t|) \\ &= 1 - \frac{\Gamma(1-\nu)}{4^\nu\Gamma(1+\nu)}|t|^{2\nu} + \frac{1}{4(1-\nu)}|t|^2 + O(|t|^{2\nu+2}) + O(|t|^4) \quad \text{as } t \to 0. \end{aligned}$$

Take $C_{11} = C_{22} = M_{0.5}$ and $C_{12} = C_{21} = 0.5M_{0.55}$. Let $m = 50$, $p = 1$, $a = (1, -2, 1)^T$ and $n \in \{200, 250, \ldots, 1500\}$. For each value of n, generate 3000 independent realizations of the process X. In this case, $\sigma_1 = \sigma_2 = 1$, $\alpha_{11} = \alpha_{22} = 1$, $\rho = 0.5$, $\alpha_{12} = 1.1 > (\alpha_{11} + \alpha_{22})/2$, $\beta_{12} = 0.9$, $c_{12} = 0.5^{1.1}\Gamma(1 - 0.55)/\Gamma(1 + 0.55)$, $c_{11} = c_{22} = 0.5\Gamma(0.5)/\Gamma(1.5)$,

$$A = -\rho\sigma_1\sigma_2 c_{12} \sum_{k,l} a_k a_l |k - l|^{\alpha_{12}} = c_{12}(4 - 2^{1.1}) \approx 1.9177 \neq 0,$$

$$\alpha_{12} + \beta_{12} = 2 > 3/2 = (\alpha_{11} + \alpha_{22} + 1)/2.$$

It follows from Theorem 2 that $\forall u = 1, \ldots, m$,

$$\begin{aligned}
\Phi_{u,u} &= A_{11}A_{22} \sum_{h=-\infty}^{\infty} \sum_{s,t,j,l=-J}^{J} a_s a_t a_j a_l |h + su - tv|^{\alpha_{11}} |h + ju - lv|^{\alpha_{22}} \\
&= (A_{11})^2 \sum_{h=-\infty}^{\infty} (6|h| - 4|h+u| + |h+2u| - 4|h-u| + |h+2u|)^2 \\
&= \left(\frac{\Gamma(0.5)}{2\Gamma(1.5)}\right)^2 \left(16u^2 + 2\sum_{h=1}^{u} (6h - 4(h+u) + 4u - 4(u-h))^2 \right. \\
&\quad \left. +2\sum_{h=u+1}^{2u} (6h - 4(h+u) + 4u - 4(h-u))^2 + 2\sum_{h=2u+1}^{\infty} (6h - 4(h+u) + 2h - 4(h-u))^2 \right) \\
&= \frac{8}{3}(4u^3 + 5u)
\end{aligned}$$

is the asymptotic marginal variance of $n^{1/2+(\alpha_{11}+\alpha_{22})/2-\alpha_{12}} \bar{Z}^u_{n,12}$ as (2.15) presented. The empirical marginal distributions of $\bar{Z}^u_{n,12}$ ($u = 1, 10, 20, 30, 40, 50$) when $n = 1500$ are shown in Figure 2.1, where 3000 realizations are presented in the histogram.

Take $\hat{\alpha}_{12}$ as the ordinary least squares estimator for β_1 in the linear regression model

$$\frac{1}{2}\log(\bar{Z}_{n,12})^2 = \begin{pmatrix} 1 & \log 1 \\ 1 & \log 2 \\ \vdots & \vdots \\ 1 & \log m \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix},$$

then as was simplified by Kent and Wood (1997),

$$\hat{\alpha}_{12} = \frac{1}{2}\sum_{u=1}^{m} \frac{\log u - \frac{1}{m}\sum_{v=1}^{m} \log v}{\sum_{u=1}^{m}\left(\log u - \frac{1}{m}\sum_{v=1}^{m}\log v\right)^2} \log(\bar{Z}^u_{n,12})^2,$$

which is an example of the estimator defined in (2.27). Since conditions in Theorem 4 are satisfied, $\hat{\alpha}_{12}$ is a strongly consistent estimator for α_{12}. The asymptotic normality follows from Theorem 5. Figure 2.3 and 2.2 confirm these claims.

2.3 Irregular Sampling

Since regularly spaced data is not always available, it is of practical importance to study estimators of the smoothness parameter based on irregular sampling designs. Given observations of

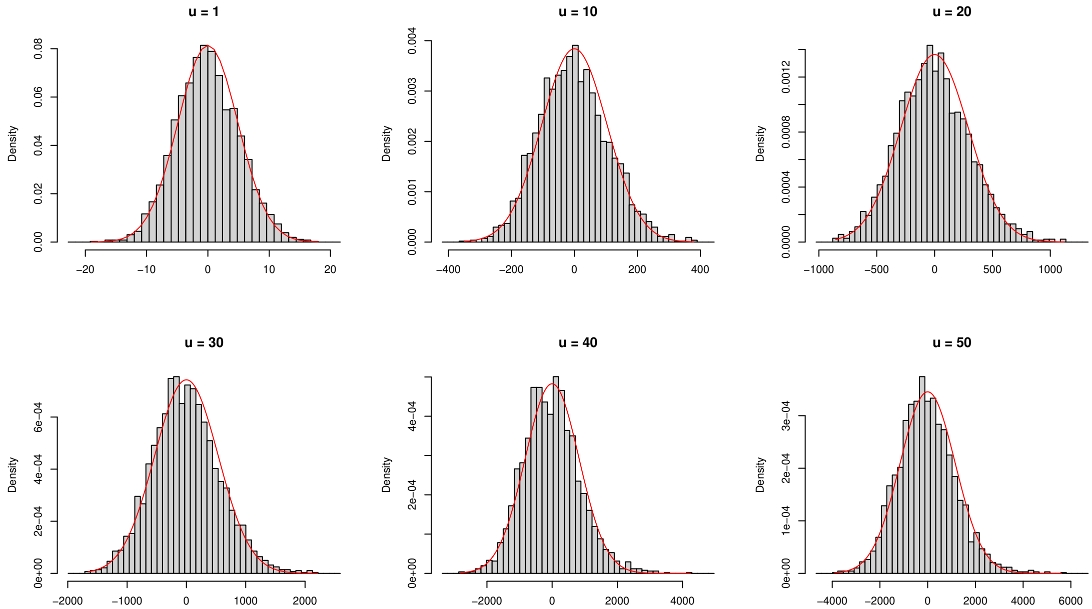

Figure 2.1The empirical distribution of $\sqrt{n^{1-2\alpha_{12}+\alpha_{11}+\alpha_{22}}}(\bar{Z}_n^u - Au^{\alpha_{12}})$ when $n = 1500$ with 3000 realizations. The red curve is the density function of $N(0, 8(4u^3 + 5u)/3)$.

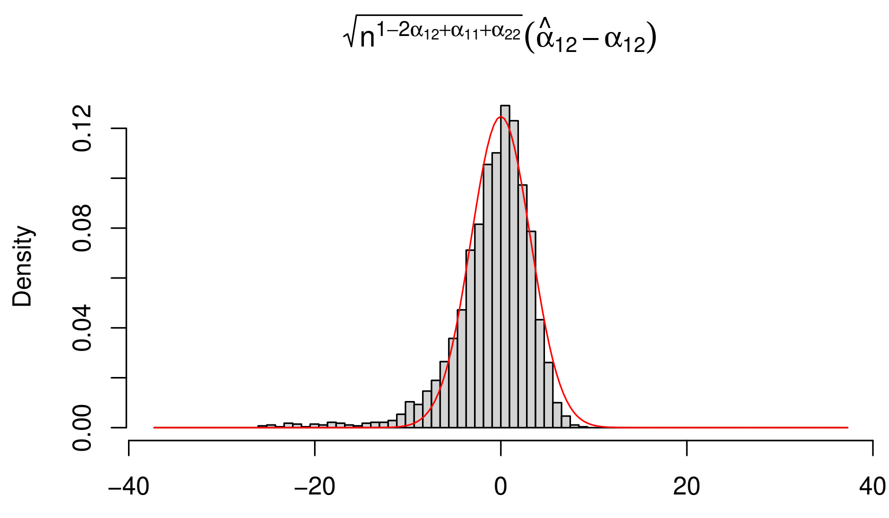

Figure 2.2The empirical distribution of $\sqrt{n^{1-2\alpha_{12}+\alpha_{11}+\alpha_{22}}}(\hat{\alpha}_{12} - \alpha_{12})$ when $n = 1500$ with 3000 realizations. The red curve is the density function of $N(0, A^{-2}\tilde{L}^T\Phi_n\tilde{L})$, where Φ_n is the empirical covariance matrix of $\bar{Z}_{n,12}$ with 3000 realizations when $n = 1500$.

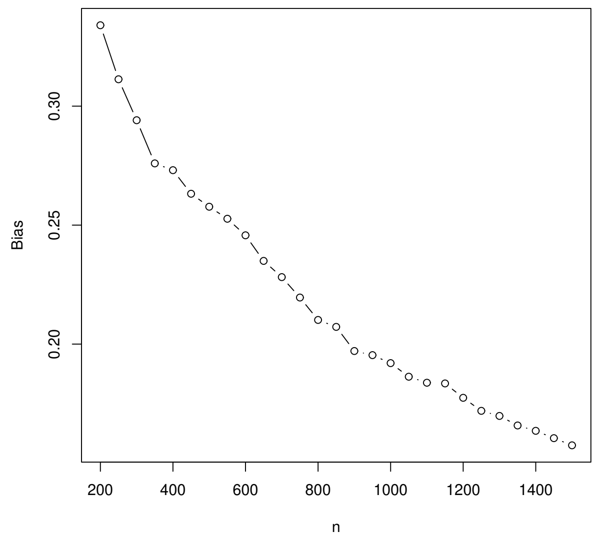

Figure 2.3The average absolute value of bias among 3000 realizations when $n = 200, 250, \ldots, 1500$.

a Gaussian process, constructing quadratic variations of a certain order is an essential step when defining increment-based estimators of the smoothness parameter. When the observation locations are not evenly spaced, coefficients of the increment discussed in Section 2.2 will be related to distances between sampling points. Begyn (2005), Loh (2015), and Loh et al. (2021) proposed several irregular sampling designs, based on which the infill asymptotic properties of quadratic variations are studied. Details of the irregular sampling designs are included in Appendix A.

In Section 2.3.1, we discuss the joint behaviors of quadratic variations for two coordinates in the bivariate model based on the deformed sampling design. In Section 2.3.2, we define a strong consistent estimator for the cross smoothness parameter and present the rate of almost sure convergence for estimators based on the stratified sampling design.

2.3.1 Quadratic Variations

Consider a special case of the bivariate stationary Gaussian process $X(t) = (X_1(t), X_2(t))$ defined in (2.1-2.3). Let the autocovariance function for each coordinate of X and the cross-covariance function of X all take the following form such that $\forall t, s \in \mathbb{R}$ and $\forall i, j \in \{1, 2\}$,

$$C_{ij}(t) = \sum_{k=0}^{\lfloor \alpha_{ij}/2 \rfloor} \beta_k (\theta_{ij}|t|)^{2k} + \beta^*_{\alpha_{ij}} G_{\alpha_{ij}}(\theta_{ij}|t|) + O(|t|^{\alpha_{ij}+\tau}) \tag{2.36}$$

as $|t| \to 0$ for some constant $\tau > 0$, where $\beta_0 = \sigma_i\sigma_j(\rho + (1-\rho)1_{i=j})$, $\lfloor x \rfloor = \max\{x_0 \in \mathbb{Z} : x_0 < x\}$, $\beta^*_{\alpha_{ij}} \neq 0$, and $G_{\alpha_{ij}} : [0, \infty) \mapsto \mathbb{R}$ is defined by

$$G_{\alpha_{ij}}(x) = x^{\alpha_{ij}} + x^{\alpha_{ij}}(\log x - 1)1_{\mathbb{Z}}(\alpha_{ij}/2)$$

when $x > 0$ and $G_{\alpha_{ij}}(0) = 0$.

Under the setting of deformed sampling design defined in (A.3), we study the cross-covariance of quadratic variations defined in (A.6) for coordinates X_1 and X_2.

Proposition 1. *For dilation $\theta \in \{1, 2\}$ and the order of increment $\ell \in \{1, 2, \ldots, \lfloor (n-1)/\theta \rfloor\}$,*

$$\frac{E(V^1_{\theta,\ell}V^2_{\theta,\ell})}{EV^1_{\theta,\ell}EV^2_{\theta,\ell}} = \begin{cases} O(n^{\alpha_{11}+\alpha_{22}-2\alpha_{12}-1}) & \text{if } \alpha_{12} < 2\ell - 1/2, \\ O(n^{\alpha_{11}+\alpha_{22}-2\alpha_{12}-1}\log n) & \text{if } \alpha_{12} = 2\ell - 1/2, \\ O(n^{\alpha_{11}+\alpha_{22}-4\ell}) & \text{if } \alpha_{12} > 2\ell - 1/2, \end{cases}$$

where $V^i_{\theta,\ell}$ is the quadratic variation of X_i $(i = 1, 2)$ as defined in (A.6).

Proof. For the brevity of symbols, denote by $a_i = (a_{\theta,\ell;i,k})_{k=0}^{\ell}$ the vector of increment defined in (A.4). Write $X^j_i = (X_j(t_{i+\theta k}))_{k=0}^{\ell}$ and $\nabla_{\theta,\ell}X^j_i = a_i^T X^j_i$. Then

$$\begin{aligned} E(V^1_{\theta,\ell}V^2_{\theta,\ell}) &= E\sum_{i,j=1}^{n-\theta\ell}\left(\left(\sum_{k=0}^{\ell} a_{\theta,\ell;i,k}X_1(t_{i+\theta k})\right)^2\left(\sum_{k=0}^{\ell} a_{\theta,\ell;j,k}X_2(t_{j+\theta k})\right)^2\right) \\ &= E\sum_{i,j=1}^{n-\theta\ell}\left(\left(a_i^T X_i^1\right)^2\left(a_j^T X_j^2\right)^2\right) \\ &= \sum_{i,j=1}^{n-\theta\ell}\left(E\left[(X_i^1)^T(a_i a_i^T)X_i^1\right]E\left[(X_j^2)^T(a_j a_j^T)X_j^2\right] + 2\left(E\left[(X_i^1)^T(a_i a_j^T)X_j^2\right]\right)^2\right) \\ &= \sum_{i,j=1}^{n-\theta\ell} E(\nabla_{\theta,\ell}X_i^1)^2 E(\nabla_{\theta,\ell}X_j^2)^2 + 2\sum_{i,j=1}^{n-\theta\ell}\left(E\left[(X_i^1)^T(a_i a_j^T)X_j^2\right]\right)^2. \end{aligned}$$

By Theorem 1 (a) in Loh (2015),

$$\sum_{i,j=1}^{n-\theta\ell} E(\nabla_{\theta,\ell}X_i^1)^2 E(\nabla_{\theta,\ell}X_j^2)^2 = EV_{\theta,\ell}^1 EV_{\theta,\ell}^2 = O(n^{2\ell+1-\alpha_{11}}) \cdot O(n^{2\ell+1-\alpha_{22}}) \tag{2.37}$$

as $n \to \infty$.

With the cross-covariance function defined in (2.36),

$$\sum_{i,j=1}^{n-\theta\ell} \left(E\left[(X_i^1)^T (a_i a_j^T) X_j^2 \right] \right)^2 = O\left(\sum_{i,j=1}^{n-\theta\ell} \left(\sum_{p,q=0}^{\ell} a_{\theta,\ell;i,p} a_{\theta,\ell;j,q} |t_{i+\theta p} - t_{j+\theta q}|^{\alpha_{12}} \right)^2 \right)$$

as $n \to \infty$. The properties of ℓth order increment imply that as $n \to \infty$,

$$\begin{aligned}
&\sum_{i,j=1}^{n-\theta\ell} \left(\sum_{p,q=0}^{\ell} a_{\theta,\ell;i,p} a_{\theta,\ell;j,q} |t_{i+\theta p} - t_{j+\theta q}|^{\alpha_{12}} \right)^2 \\
&= \sum_{|i-j|\le\theta\ell+1} \left(\sum_{p,q=0}^{\ell} O(n^{2\ell}) \left(\frac{i-j+\theta(p-q)}{n-1} \varphi^{(1)}(0) + O(n^{-2}) \right)^{\alpha_{12}} \right)^2 \\
&\quad + \sum_{|i-j|>\theta\ell+1} \left(\sum_{p,q=0}^{\ell} a_{\theta,\ell;i,p} a_{\theta,\ell;j,q} |t_{i+\theta p} - t_{j+\theta q}|^{\alpha_{12}} \right)^2 \\
&:= A_n + B_n,
\end{aligned}$$

where $A_n = O(n^{1+4\ell-2\alpha_{12}})$ and

$$\begin{aligned}
B_n &\le \sum_{|i-j|>\theta\ell+1} \left(\sum_{p,q=0}^{\ell} |a_{\theta,\ell;i,p} a_{\theta,\ell;j,q}| \cdot |t_{i+\theta p} - t_{j+\theta q}|^{\alpha_{12}+2\ell-2\ell} \right)^2 \\
&\le \sum_{|i-j|>\theta\ell+1} \left(\max_{0\le p,q\le\ell} |t_{i+\theta p} - t_{j+\theta q}|^{\alpha_{12}-2\ell} \sum_{p,q=0}^{\ell} |a_{\theta,\ell;i,p} a_{\theta,\ell;j,q} (t_{i+\theta p} - t_{j+\theta q})^{2\ell}| \right)^2 \\
&= O(1) \sum_{|i-j|>\theta\ell+1} \max_{0\le p,q\le\ell} |t_{i+\theta p} - t_{j+\theta q}|^{2\alpha_{12}-4\ell} \\
&= O(n^2) \int_{1/n}^{1} s^{2\alpha_{12}-4\ell} ds.
\end{aligned}$$

Thus,

$$\sum_{i,j=1}^{n-\theta\ell} \left(\sum_{p,q=0}^{\ell} a_{\theta,\ell;i,p} a_{\theta,\ell;j,q} |t_{i+\theta p} - t_{j+\theta q}|^{\alpha_{12}} \right)^2 = \begin{cases} O(n^{1+4\ell-2\alpha_{12}}) & \text{if } \alpha_{12} < 2\ell - 1/2, \\ O(n^2 \log n) & \text{if } \alpha_{12} = 2\ell - 1/2, \\ O(n^2) & \text{if } \alpha_{12} > 2\ell - 1/2. \end{cases}$$

This finishes the proof together with (2.37). □

For a stationary GRF X on $\mathbb{R}^d$ with zero mean and the isotropic Matérn covariance function

$$C(\mathrm{t}) = \frac{\sigma^2(\eta||\mathrm{t}||)^\nu}{2^{\nu-1}\Gamma(\nu)}\kappa_\nu(\eta||\mathrm{t}||), \quad \forall \mathrm{t} \in \mathbb{R}^d, \tag{2.38}$$

where $\sigma, \eta, \nu > 0$ are constants, we discuss the finite sample joint distribution of $V_{1,1,\ell}$ and $V_{2,1,\ell}$ in the remaining of this section. The quadratic variations $V_{\theta,d,\ell}$ are defined in (A.21). Consider the case when $d = 1$ and $0 < \nu < \ell$, $\nu \notin \mathbb{Z}$. Write

$$\nabla_{\theta,\ell}X = \left(\nabla_{\theta,1,\ell}X_i\right)_{i=1}^{n-2\ell\omega_n}$$

and denote by $V_{uv}(n,\ell) = (\nabla_{u,\ell}X)^T\nabla_{v,\ell}X$, $W_{uv}(n,\ell) = Cov(\nabla_{u,\ell}X, \nabla_{v,\ell}X)$ for $u, v \in \{1,2\}$. For the brevity, write $V_{uv}(n,\ell)$ as V_{uv} and $W_{uv}(n,\ell)$ as W_{uv} in the following text.

It follows from Eq.(15) in Loh et al. (2021) that as $n \to \infty$,

$$\begin{aligned}
&(W_{uv})_{i,i+h}\\
=&\beta_\nu^* \sum_{j,k=0}^{\ell} c_{\mathbf{i},u,1,\ell}(j)c_{\mathbf{i+h},v,1,\ell}(k)\left|\frac{h+(vk-uj)\omega_n+\delta_{i+h,k}-\delta_{i,j}}{n}\right|^{2\nu} + O\left(\left(\frac{\omega_n}{n}\right)^{2\ell}\right)\\
&+O\left(\left(\frac{\omega_n}{n}\right)^{2\nu+2}\right)\\
=&\beta_\nu^* \sum_{j,k=0}^{\ell}\left(c_\ell(j)+O(\omega_n^{-1})\right)\left(c_\ell(k)+O(\omega_n^{-1})\right)\left|\frac{h+(vk-uj)\omega_n+\delta_{i+h,k}-\delta_{i,j}}{n}\right|^{2\nu} + o\left(\left(\frac{\omega_n}{n}\right)^{2\nu}\right)\\
=&\beta_\nu^* \sum_{j,k=0}^{\ell} c_\ell(j)c_\ell(k)\left|\frac{h+(vk-uj)\omega_n+\delta_{i+h,k}-\delta_{i,j}}{n}\right|^{2\nu} + o\left(\left(\frac{\omega_n}{n}\right)^{2\nu}\right)\\
=&\left(\frac{\omega_n}{n}\right)^{2\nu}\beta_\nu^* \sum_{j,k=0}^{\ell} c_\ell(j)c_\ell(k)\left|vk-uj+\frac{h+\delta_{i+h,k}-\delta_{i,j}}{\omega_n}\right|^{2\nu} + o\left(\left(\frac{\omega_n}{n}\right)^{2\nu}\right)
\end{aligned} \tag{2.39}$$

for any $1 \le i \le i+h \le n-2\ell\omega_n$. Denote by $a_{uv}(\nu,\ell) = \beta_\nu^*\sum_{j,k=0}^{\ell} c_\ell(j)c_\ell(k)\,|vk-uj|^{2\nu}$, then $\forall 1 \le i \le i+h \le n-2\ell\omega_n$,

$$(n/\omega_n)^{2\nu}(W_{uv})_{i,i+h} \to a_{uv}(\nu,\ell) \tag{2.40}$$

as $n \to \infty$.

Take $\epsilon \sim N(0, I_{n-2\ell\omega_n})$, then for $\theta = 1, 2$,

$$\frac{(n/\omega_n)^{2\nu}}{n-2\ell\omega_n} V_{\theta,1,\ell} \stackrel{d}{=} \frac{(n/\omega_n)^{2\nu}}{n-2\ell\omega_n} \epsilon^T W_{\theta\theta} \epsilon \stackrel{d}{=} \epsilon^T \left(\frac{(n/\omega_n)^{2\nu}}{n-2\ell\omega_n} diag(\mathrm{eig}(W_{\theta\theta})) \right) \epsilon := \epsilon^T \Lambda_n^\theta \epsilon,$$

the cumulant generating function of which is

$$\begin{aligned} \log E e^{t\epsilon^T \Lambda_n^\theta \epsilon} &= \sum_{k=1}^{n-2\ell\omega_n} \log(1-2t\lambda_k)^{-1/2} \\ &= \frac{1}{2} \sum_{m=1}^{\infty} \frac{(2t)^m}{m} \sum_{k=1}^{n-2\ell\omega_n} \lambda_k^m, \end{aligned}$$

where $t < \min(\lambda_k^{-1})$ and λ_k, $k = 1, \ldots, n-2\ell\omega_n$ are diagonal elements of Λ_n^θ.

Denote by $r_n = \frac{(n/\omega_n)^{2\nu}}{n-2\ell\omega_n}$ and recall the notation $W_\theta = V_{\theta,1,\ell}/EV_{\theta,1,\ell}$ for $\theta = 1, 2$. Write $H_n = W_{22} - W_{21} W_{11}^{-1} W_{12}$, then $\nabla_{2,\ell} X | \nabla_{1,\ell} X \sim N(W_{21} W_{11}^{-1} \nabla_{1,\ell} X, H_n)$ and the moment generating function of $V_{2,1,\ell} | \nabla_{1,\ell} X$ is

$$\begin{aligned} & M_{V_{2,1,\ell}|\nabla_{1,\ell}X}(t) \\ =& |I - 2tH_n|^{-1/2} \exp\left(-\frac{1}{2} (\nabla_{1,\ell} X)^T W_{11}^{-1} W_{12} \left(I - (I - 2tH_n)^{-1} \right) H_n^{-1} W_{21} W_{11}^{-1} \nabla_{1,\ell} X \right), \end{aligned} \tag{2.41}$$

where I is the $(n-2\ell\omega_n)$-dimensional identity matrix. Moreover, the moment generating function of the vector $\tilde{V} := r_n (V_{1,1,\ell}, V_{2,1,\ell})^T$ is

$$M_{\tilde{V}}(s,t) = \left| I_{2(n-2\ell\omega_n)} - 2 \begin{pmatrix} r_n t H_n & 0 \\ 0 & H_n^{st} W_{11} \end{pmatrix} \right|^{-1/2}, \tag{2.42}$$

where

$$H_n^{st} = r_n s I - \frac{1}{2} W_{11}^{-1} W_{12} \left(I - (I - 2r_n t H_n)^{-1} \right) H_n^{-1} W_{21} W_{11}^{-1}.$$

This is due to the fact that

$$\begin{aligned}
M_{\tilde{V}}(s,t) &= E\left[e^{r_n(sV_{1,1,\ell}+tV_{2,1,\ell})}\right] \\
&= E\left[e^{r_n sV_{1,1,\ell}} E\left[e^{r_n tV_{2,1,\ell}} | \nabla_{1,\ell}X\right]\right] \\
&= E\left[e^{r_n sV_{1,1,\ell}} M_{V_{2,1,\ell}|\nabla_{1,\ell}X}(r_n t)\right] \\
&= |I - 2r_n t H_n|^{-1/2} E\left[\exp\left((\nabla_{1,\ell}X)^T H_n^{st} \nabla_{1,\ell}X\right)\right] \\
&= |I - 2r_n t H_n|^{-1/2} M_{(\nabla_{1,\ell}X)^T H_n^{st}\nabla_{1,\ell}X}(1) \\
&= |I - 2r_n t H_n|^{-1/2} |I - 2H_n^{st} W_{11}|^{-1/2} \\
&= \left| I_{2(n-2\ell\omega_n)} - 2\begin{pmatrix} r_n t H_n & 0 \\ 0 & H_n^{st} W_{11} \end{pmatrix} \right|^{-1/2}.
\end{aligned}$$

2.3.2 Estimating Smoothness Parameters

We first consider a univariate stationary GRF X on $\mathbb{R}^d$ with zero mean and the isotropic Matérn covariance function (2.38). Based on the stratified design introduced in Appendix A.2.3, the following results on the rate of convergence hold for $\hat{\nu}_{n,\ell}$ defined in (A.26).

Proposition 2. *When $d \in \{1,2,3\}$ and $\ell \in \mathbb{Z}^+$,*

1. *if $0 < \nu \leq \ell - 1$, then*

$$n^{d(1-\gamma_0)/2-k}(\hat{\nu}_{n,\ell} - \nu) \xrightarrow{a.s.} 0 \quad \text{as } n \to \infty$$

for any $(d(1-\gamma_0)/2 - \gamma_0) \vee (d/2 - 2)(1-\gamma_0) < k < d(1-\gamma_0)/2$;

2. *if $\ell - 1 < \nu < \ell - d/4$, then*

$$n^{d(1-\gamma_0)/2-k}(\hat{\nu}_{n,\ell} - \nu) \xrightarrow{a.s.} 0 \quad \text{as } n \to \infty$$

for any $(d(1-\gamma_0)/2 - \gamma_0) \vee (d/2 - 2\ell + 2\nu)(1-\gamma_0) < k < d(1-\gamma_0)/2$;

3. *if $\nu = \ell - d/4$, then*

$$n^{d(1-\gamma_0)/2-k}(\log n)^{-1/2}(\hat{\nu}_{n,\ell} - \nu) \xrightarrow{a.s.} 0 \quad \text{as } n \to \infty$$

for any $(d(1-\gamma_0)/2-\gamma_0) \vee (d/2-2\ell+2\nu)(1-\gamma_0) < k < d(1-\gamma_0)/2$*;*

4. *if* $\ell - d/4 < \nu < \ell$*, then*

$$n^{(2\ell-2\nu)(1-\gamma_0)-k}(\hat{\nu}_{n,\ell}-\nu) \overset{a.s.}{\to} 0 \quad \text{as } n \to \infty$$

for any $(2\ell-2\nu)(1-\gamma_0)-\gamma_0 < k < (2\ell-2\nu)(1-\gamma_0)$.

Proof. Theorem 1(a) in Loh et al. (2021) implies that as $n \to \infty$

$$\begin{aligned}
\hat{\nu}_{n,\ell}-\nu &= \frac{\log(V_{2,d,\ell}/V_{1,d,\ell})-\log(2^{2\nu})}{2\log 2} \\
&= \frac{1}{2\log 2}\log\left(\frac{\frac{V_{2,d,\ell}/EV_{2,d,\ell}}{V_{1,d,\ell}/EV_{1,d,\ell}}\cdot\frac{EV_{2,d,\ell}}{EV_{1,d,\ell}}}{2^{2\nu}}\right) \\
&= \frac{1}{2\log 2}\log\left(\frac{\frac{V_{2,d,\ell}/EV_{2,d,\ell}}{V_{1,d,\ell}/EV_{1,d,\ell}}\left(2^{2\nu}+O(h(n))\right)}{2^{2\nu}}\right) \\
&= \frac{1}{2\log 2}\log\left(\frac{V_{2,d,\ell}/EV_{2,d,\ell}}{V_{1,d,\ell}/EV_{1,d,\ell}}\left(1+O(h(n))\right)\right),
\end{aligned} \tag{2.43}$$

where

$$h(n) = \begin{cases} n^{-\gamma_0}+n^{(\gamma_0-1)((2\ell-2\nu)\wedge 2)} & \text{if } \nu \notin \mathbb{Z}, \\ n^{-\gamma_0}+n^{2(\gamma_0-1)}\log n & \text{if } \nu = \ell-1, \\ n^{-\gamma_0}+n^{2(\gamma_0-1)} & \text{if } 0 < \nu \le \ell-2, \nu \in \mathbb{Z}. \end{cases}$$

Denote by $W_\theta = V_{\theta,d,\ell}/EV_{\theta,d,\ell}$ for $\theta = 1, 2$, then it suffices to find the convergence rate of $W_2/W_1 - 1$. It was proved in Loh et al. (2021) (P21-25) that

$$P(|W_\theta - 1| \ge \epsilon) \le 2\exp\left(-C\min\left\{\frac{\epsilon}{a_n}, \frac{\epsilon^2}{b_n}\right\}\right), \quad \forall \epsilon > 0,$$

where as $n \to \infty$,

$$a_n = \begin{cases} O(n^{d(\gamma_0-1)}) & \text{if } \nu < \ell - d/2, \\ O(n^{d(\gamma_0-1)}\log n) & \text{if } \nu = \ell - d/2, \\ O(n^{(2\ell-2\nu)(\gamma_0-1)}) & \text{if } \ell - d/2 < \nu < \ell, \end{cases}$$

$$b_n = \begin{cases} O(n^{d(\gamma_0-1)}) & \text{if } \nu < \ell - d/4, \\ O(n^{d(\gamma_0-1)}\log n) & \text{if } \nu = \ell - d/4, \\ O(n^{(4\ell-4\nu)(\gamma_0-1)}) & \text{if } \ell - d/4 < \nu < \ell. \end{cases}$$

Then for any positive constant c_0,

$$P(c_0|W_\theta - 1| \geq \epsilon) \leq 2\exp\left(-C\min\left\{\frac{\epsilon}{c_0 a_n}, \frac{\epsilon^2}{c_0^2 b_n}\right\}\right), \quad \forall \epsilon > 0.$$

By the Borel-Cantelli lemma, for $\theta = 1, 2$, $f(n,k)(W_\theta - 1) \to 0$ a.s. as $n \to \infty$ for any $k > 0$, where

$$f(n,k) = \begin{cases} n^{d(1-\gamma_0)/2-k} & \text{if } \nu < \ell - d/4, \\ n^{d(1-\gamma_0)/2-k}(\log n)^{-1/2} & \text{if } \nu = \ell - d/4, \\ n^{(2\ell-2\nu)(1-\gamma_0)-k} & \text{if } \ell - d/4 < \nu < \ell. \end{cases}$$

Thus, $f(n,k)(W_2/W_1 - 1) = f(n,k)((W_2-1)-(W_1-1))/W_1 \to 0$ a.s. as $n \to \infty$ for any $k > 0$. It follows from (2.43) that as $n \to \infty$,

$$\begin{aligned} f(n,k)(\hat{\nu}_{n,\ell} - \nu) &= \frac{f(n,k)}{2\log 2}\log\left(\frac{W_2}{W_1}(1 + O(h(n)))\right) \\ &\sim f(n,k)\left(\frac{W_2}{W_1}(1+O(h(n))) - 1\right) \\ &= f(n,k)\,(W_2/W_1 - 1) + f(n,k)O(h(n)). \end{aligned}$$

When $d \in \{1,2,3\}$, it always holds that $\ell - 1 < \ell - d/4 < \ell$ and $d/4 \notin \mathbb{Z}$, so

$$f(n,k)h(n) = \begin{cases} n^{d(1-\gamma_0)/2-\gamma_0-k} + n^{(d/2-2)(1-\gamma_0)-k} & \text{if } 0 < \nu < \ell - 1, \\ n^{d(1-\gamma_0)/2-\gamma_0-k} + n^{(d/2-2)(1-\gamma_0)-k}\log n & \text{if } \nu = \ell - 1, \\ n^{d(1-\gamma_0)/2-\gamma_0-k} + n^{(d/2-2\ell+2\nu)(1-\gamma_0)-k} & \text{if } \ell - 1 < \nu < \ell - d/4, \\ (n^{d(1-\gamma_0)/2-\gamma_0-k} + n^{(d/2-2\ell+2\nu)(1-\gamma_0)-k})(\log n)^{-1/2} & \text{if } \nu = \ell - d/4, \\ n^{(2\ell-2\nu)(1-\gamma_0)-\gamma_0-k} + n^{-k} & \text{if } \ell - d/4 < \nu < \ell. \end{cases}$$

This finishes the proof. □

Remark 1. *Briefly speaking, as $n \to \infty$, it holds that*

$$n^{(1-\gamma_0)(d/2\wedge(2\ell-2\nu))-k}(\hat{\nu}-\nu) \overset{a.s.}{\to} 0 \quad \text{if } \nu \neq \ell - d/4, \tag{2.44}$$

$$n^{d(1-\gamma_0)/2-k}(\log n)^{-1/2}(\hat{\nu}-\nu) \overset{a.s.}{\to} 0 \quad \text{if } \nu = \ell - d/4, \tag{2.45}$$

where k is a constant whose range depends on d, γ_0, and $\ell - \nu$.

In the remaining of this section, we consider a bivariate Gaussian process $X(t) = (X_1(t), X_2(t))$ with zero mean and covariance function

$$C(t) = \begin{pmatrix} C_{11}(t) & C_{12}(t) \\ C_{21}(t) & C_{22}(t) \end{pmatrix},$$

where C_{ij} is the Matérn covariance function

$$C_{ij}(t) = \frac{\sigma_{ij}^2(\eta_{ij}|t|)^{\nu_{ij}}}{2^{\nu_{ij}-1}\Gamma(\nu_{ij})}\kappa_{\nu_{ij}}(\eta_{ij}|t|), \quad \forall t \in \mathbb{R}, \tag{2.46}$$

where $i, j \in \{1, 2\}$, $\sigma_{12} = \sigma_{21} = \rho\sigma_{11}\sigma_{22}$, $\nu_{ij}, \eta_{ij}, \sigma_{11}, \sigma_{22} > 0$, $|\rho| \in (0, 1)$.

Under the stratified sampling design introduced in Appendix A.2.3, write

$$Y_{n,1}^{\theta} = (\nabla_{\theta,1,\ell}^1 X_1, \nabla_{\theta,1,\ell}^1 X_2, \ldots, \nabla_{\theta,1,\ell}^1 X_{n-2\ell\omega_n})^T,$$

$$Y_{n,2}^{\theta} = (\nabla_{\theta,1,\ell}^2 X_1, \nabla_{\theta,1,\ell}^2 X_2, \ldots, \nabla_{\theta,1,\ell}^2 X_{n-2\ell\omega_n})^T,$$

$$Y_n^{\theta} = \begin{pmatrix} Y_{n,1}^{\theta} \\ Y_{n,2}^{\theta} \end{pmatrix} \in \mathbb{R}^{2(n-2\ell\omega_n)},$$

and define the covariation as

$$\begin{aligned} Z_{n,12}^{\theta} &= \sum_{1\leq i\leq n-2\ell\omega_n} \left(\nabla_{\theta,1,\ell}^1 X_i\right)\left(\nabla_{\theta,1,\ell}^2 X_i\right) \\ &= \frac{1}{2}(Y_n^{\theta})^T \begin{pmatrix} 0 & I_{n-2\ell\omega_n} \\ I_{n-2\ell\omega_n} & 0 \end{pmatrix} Y_n^{\theta}, \end{aligned} \tag{2.47}$$

where $\theta \in \{1, 2\}$, $\ell \in \mathbb{Z}^+$, and

$$\nabla_{\theta,1,\ell}^k X_i = \sum_{j=0}^{\bar{\ell}} c_{\mathbf{i},\theta,1,\ell}(j) X_k(\mathbf{x}_{\mathbf{i},j}), \quad i \in \{1, \ldots, n-2\ell\omega_n\}, \quad k = 1, 2. \tag{2.48}$$

Proposition 3. *When* $2(\nu_{11}+\nu_{22}) < 4\nu_{12} < \{(2(\nu_{11}+\nu_{22})+1) \wedge 4\ell\}$ *and* $\nu_{11} \vee \nu_{22} < \ell$,

$$\frac{Z^{\theta}_{n,12}}{EZ^{\theta}_{n,12}} \xrightarrow{a.s.} 1 \quad as\ n \to \infty, \tag{2.49}$$

where $\theta \in \{1,2\}$ *and* $\ell \in \mathbb{Z}^+$.

Proof. It follows from Theorem 1 (a) in Loh et al. (2021) that as $n \to \infty$,

$$EZ^{\theta}_{n,12} = \left(\frac{\omega_n \theta}{n}\right)^{2\nu_{12}} (n - 2\ell\omega_n)\left(\beta^* \sum_{1\le j,k\le \ell} c_{j,\theta,1,\ell}c_{k,\theta,1,\ell}G_{\nu_{12}}(|j-k|) + o(1)\right), \tag{2.50}$$

where $\theta \in \{1,2\}$ and $\ell \in \mathbb{Z}^+$. For $k = 1,2$, let

$$\nabla^k_{\theta,\ell}X = \left(\nabla^k_{\theta,1,\ell}X_i\right)_{i=1}^{n-2\ell\omega_n}$$

and write $W^k_{\theta\theta}(n,\ell) = Cov(\nabla^k_{\theta,\ell}X, \nabla^k_{\theta,\ell}X)$, $W^{12}_{\theta\theta}(n,\ell) = Cov(\nabla^1_{\theta,\ell}X, \nabla^2_{\theta,\ell}X)$. Then the variance of the covariation follows

$$\begin{aligned}
var\left(\frac{Z^{\theta}_{n,12}}{EZ^{\theta}_{n,12}}\right) &= \frac{E(Z^{\theta}_{n,12})^2 - (EZ^{\theta}_{n,12})^2}{(EZ^{\theta}_{n,12})^2} \\
&= \frac{\sum_{1\le i,j\le n-2\ell\omega_n} E\left(\nabla^1_{\theta,1,\ell}X_i \nabla^1_{\theta,1,\ell}X_j \nabla^2_{\theta,1,\ell}X_i \nabla^2_{\theta,1,\ell}X_j\right) - (EZ^{\theta}_{n,12})^2}{(EZ^{\theta}_{n,12})^2} \\
&= \frac{(EZ^{\theta}_{n,12})^2 + \sum_{1\le i,j\le n-2\ell\omega_n}\left((W^1_{\theta\theta})_{i,j}(W^2_{\theta\theta})_{i,j} + (W^{12}_{\theta\theta})^2_{i,j}\right) - (EZ^{\theta}_{n,12})^2}{(EZ^{\theta}_{n,12})^2} \\
&= \frac{1}{(EZ^{\theta}_{n,12})^2} \sum_{1\le i,j\le n-2\ell\omega_n} (W^1_{\theta\theta})_{i,j}(W^2_{\theta\theta})_{i,j} + (W^{12}_{\theta\theta})^2_{i,j}.
\end{aligned}$$

It follows from the same manner as in (3.18-3.19) of Loh et al. (2021) that, based on the definition of $c_{i,\theta,1,\ell}$ in (A.20) and the Taylor expansion of the function C_{12},

$$\frac{1}{(EZ^{\theta}_{n,12})^2} \sum_{1\le i,j\le n-2\ell\omega_n} (W^{12}_{\theta\theta})^2_{i,j} = \begin{cases} O\left(\frac{\omega_n}{n}\right), & 0 < \nu_{12} < \ell - 1/4, \\ O\left(\frac{\omega_n}{n}\log\frac{n}{\omega_n}\right), & \nu_{12} = \ell - 1/4, \\ O\left(\left(\frac{\omega_n}{n}\right)^{4\ell-4\nu_{12}}\right), & \ell - 1/4 < \nu_{12} < \ell \end{cases} \tag{2.51}$$

as $n \to \infty$. Similarly, when $\nu_{11} \vee \nu_{22} < \ell$,

$$\frac{1}{(EZ^{\theta}_{n,12})^2} \sum_{1 \le i,j \le n-2\ell\omega_n} (W^1_{\theta\theta})_{i,j}(W^2_{\theta\theta})_{i,j}$$

$$= \begin{cases} O\left(\left(\frac{\omega_n}{n}\right)^{2\nu_{11}+2\nu_{22}-4\nu_{12}+1}\right), & 0 < 2(\nu_{11}+\nu_{22}) < 4\ell - 1, \\ O\left(\left(\frac{\omega_n}{n}\right)^{2\nu_{11}+2\nu_{22}-4\nu_{12}+1} \log \frac{n}{\omega_n}\right), & 2(\nu_{11}+\nu_{22}) = 4\ell - 1, \\ O\left(\left(\frac{\omega_n}{n}\right)^{4\ell-4\nu_{12}}\right), & 4\ell - 1 < 2(\nu_{11}+\nu_{22}) < 4\ell \end{cases} \tag{2.52}$$

as $n \to \infty$. Thus,

$$var\left(\frac{Z^{\theta}_{n,12}}{EZ^{\theta}_{n,12}}\right) = \begin{cases} O\left(\left(\frac{\omega_n}{n}\right)^{2\nu_{11}+2\nu_{22}-4\nu_{12}+1}\right), & 0 < 2(\nu_{11}+\nu_{22}) < 4\nu_{12} \le 4\ell - 1, \\ O\left(\left(\frac{\omega_n}{n}\right)^{2\nu_{11}+2\nu_{22}-4\nu_{12}+1} \log \frac{n}{\omega_n}\right), & 4\ell - 1 = 2(\nu_{11}+\nu_{22}) < 4\nu_{12} < 4\ell, \\ O\left(\left(\frac{\omega_n}{n}\right)^{4\ell-4\nu_{12}}\right), & 4\ell - 1 < 2(\nu_{11}+\nu_{22}) < 4\nu_{12} < 4\ell \end{cases} \tag{2.53}$$

as $n \to \infty$. Consequently, when $2(\nu_{11}+\nu_{22}) < 4\nu_{12} < \{(2(\nu_{11}+\nu_{22})+1) \wedge 4\ell\}$ and $\nu_{11} \vee \nu_{22} < \ell$,

$$\frac{Z^{\theta}_{n,12}}{EZ^{\theta}_{n,12}} \xrightarrow{P} 1 \quad \text{as } n \to \infty.$$

According to the definition in (2.47),

$$\frac{Z^{\theta}_{n,12}}{EZ^{\theta}_{n,12}} \stackrel{\mathrm{d}}{=} U^T \Sigma^{\theta}_n U,$$

where $U \sim N(0, I_{2(n-2\ell\omega_n)})$ and

$$\Sigma^{\theta}_n = \frac{1}{2EZ^{\theta}_{n,12}} \mathrm{Cov}(Y^{\theta}_n)^{1/2} \begin{pmatrix} 0 & I_{n-2\ell\omega_n} \\ I_{n-2\ell\omega_n} & 0 \end{pmatrix} \mathrm{Cov}(Y^{\theta}_n)^{1/2}.$$

The Hanson-Wright inequality implies that there exists an absolute constant $C > 0$ such that $\forall \epsilon > 0$,

$$P\left(\left|\frac{Z^{\theta}_{n,12}}{EZ^{\theta}_{n,12}} - 1\right| \ge \epsilon\right) = P\left(\left|U^T\Sigma^{\theta}_n U - E[U^T\Sigma^{\theta}_n U]\right| \ge \epsilon\right)$$
$$\le 2\exp\left(-C \min\left\{\frac{\epsilon}{||\Sigma^{\theta}_n||_2}, \frac{\epsilon^2}{||\Sigma^{\theta}_n||_F^2}\right\}\right). \tag{2.54}$$

Since $||\Sigma^{\theta}_n||_2 \le ||\Sigma^{\theta}_n||_F$ and

$$||\Sigma^{\theta}_n||_F^2 = \frac{1}{2} var\left(\frac{Z^{\theta}_{n,12}}{EZ^{\theta}_{n,12}}\right),$$

the Borel-Cantelli lemma together with (2.53) and (2.54) induces that if $2(\nu_{11} + \nu_{22}) < 4\nu_{12} < \{(2(\nu_{11} + \nu_{22}) + 1) \wedge 4\ell\}$ and $\nu_{11} \vee \nu_{22} < \ell$, then

$$\frac{Z^{\theta}_{n,12}}{EZ^{\theta}_{n,12}} \stackrel{a.s.}{\to} 1 \quad \text{as } n \to \infty. \tag{2.55}$$

This finishes the proof. □

Consequently, the estimator defined as

$$\hat{\nu}_{12} = \frac{\log(Z^2_{n,12}/Z^1_{n,12})^2}{4 \log 2} \tag{2.56}$$

is a strongly consistent estimator for ν_{12} based on irregularly spaced data.

Theorem 7. *Under the conditions of Proposition 3,*

$$\hat{\nu}_{12} \stackrel{a.s.}{\to} \nu_{12} \quad \textit{as } n \to \infty. \tag{2.57}$$

Proof. It follows from (2.50) that

$$\frac{EZ^2_{n,12}}{EZ^1_{n,12}} \to 2^{2\nu_{12}} \quad \text{as } n \to \infty.$$

By the result of Proposition 3,

$$\begin{aligned} \frac{Z^2_{n,12}}{Z^1_{n,12}} &= \frac{Z^2_{n,12}/EZ^2_{n,12}}{Z^1_{n,12}/EZ^1_{n,12}} \cdot \frac{EZ^2_{n,12}}{EZ^1_{n,12}} \\ &\stackrel{a.s.}{\to} 2^{2\nu_{12}} \quad \text{as } n \to \infty. \end{aligned} \tag{2.58}$$

The proof is completed by applying the continuous mapping theorem. □

CHAPTER 3

ANISOTROPIC ORNSTEIN-UHLENBECK FIELD

3.1 Introduction

Proposed by Uhlenbeck and Ornstein (1930), the Ornstein-Uhlenbeck process is widely used in spatial statistics and finance. Denote by $\{W(u,t); u,t \in \mathbb{R}_+\}$ a standard Wiener field, then the random field

$$X(u,t) = \sigma \exp(-\lambda u - \mu t) W\left(e^{2\lambda u}, e^{2\mu t}\right), \quad u,t \in \mathbb{R} \tag{3.1}$$

is a zero-mean stationary Ornstein-Uhlenbeck field on $\mathbb{R}^2$ with covariance function

$$\operatorname{Cov}\left(X\left(u,t\right), X\left(v,s\right)\right) = \sigma^2 \exp\left(-\lambda|u-v| - \mu|t-s|\right), \quad \forall u,t,v,s \in \mathbb{R}, \tag{3.2}$$

where $(\sigma^2, \lambda, \mu) \in \mathbb{R}^3_{>0}$. As indicated by Theorem 7.2 in Piterbarg (1995), the parameters σ^2, λ, and μ characterize the high excursion probability of X on a closed Jordan set (the details are provided in Appendix B). Estimating their values is thus of significance in extreme value theory and has applications in risk assessment for rare events.

Ying (1993) proves the strong consistency and asymptotic normality of the maximum likelihood estimators (MLEs) for σ^2, λ, and μ in (3.2), thus has presented the identifiability of the parameters. The MLEs are asymptotically efficient as shown by van der Vaart (1996). The MLE is also commonly used to estimate covariance parameter under other models. For Gaussian random fields on $\mathbb{R}^d$ ($d = 1,2,3$) with the isotropic Matérn covariance function, Bachoc et al. (2019) studied the asymptotic distributions of MLE and constrained MLE for the variance and correlation length parameters. Bevilacqua et al. (2019) investigated strong consistency and asymptotic distribution of the MLE for the microergodic parameters in generalized Wendland covariance functions. However, the calculation of precision matrices and numerical optimizations usually make it computationally expensive to get MLEs. To reduce the computational cost, approaches aiming at sparse covariance matrices or sparse precision matrices have been widely studied, such as covariance tapering (Furrer et al., 2006; Kaufman et al., 2008; Du et al., 2009) and Vecchia approximations (Vecchia, 1988; Pardo-Igúzquiza and Dowd, 1997; Katzfuss and Guinness, 2021).

For Gaussian random fields with the covariance function

$$\mathrm{Cov}\left(X\left(\mathbf{u}\right), X\left(\mathbf{v}\right)\right) = \sigma^2 \prod_{i=1}^{d} \exp\left(-\theta_i |u_i - v_i|^{\gamma}\right), \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^d, \tag{3.3}$$

Lam and Loh (2000) proved the strong consistency of MLEs for $\theta_1, \ldots, \theta_d$ when $\gamma = 2$, based on observations on a regular lattice. Later, Wang (2010) provided consistent estimators for σ^2 and $\theta_1, \ldots, \theta_d$ using quadratic variations and spectral analysis when $d \geq 2$ and $0 < \gamma < 1$. The covariance function of the Ornstein-Uhlenbeck field X we consider in this chapter is a special case of (3.3) with $d = 2$ and $\gamma = 1$. Since X is Markovian, its precision matrix has sparse closed-form expression (Baldi Antognini and Zagoraiou, 2010), which reduces the computational complexity and the memory storage requirement of MLEs. The estimators we propose in this chapter are computationally more efficient than MLEs, while their strong consistency and asymptotic normality still hold.

This chapter is organized as follows. We formulate estimations for $\sigma^2\mu$, $\sigma^2\lambda$, and $\sigma^2\lambda\mu$ in Section 3.2 based on MLEs. Section 3.3 includes estimations for σ^2, λ, and μ, as well as the asymptotic behaviors of the estimators. Some simulation results are presented in Section 3.4. In Section 3.5, conclusions and our future research plans are provided.

3.2 Product Estimation

Denote by $x_{ij} = X\left(u_i, t_j\right)$, $x_i = (x_{i1}, \ldots, x_{in})^T$, $x = (x_1^T, x_2^T, \ldots, x_m^T)^T \in \mathbb{R}^{mn}$ and

$$A\left(\lambda\right) = \left(e^{-\lambda|u_i - u_j|}\right)_{m\times m}, B\left(\mu\right) = \left(e^{-\mu|t_i - t_j|}\right)_{n\times n}, \tag{3.4}$$

where $0 = u_0 < u_1 < \cdots < u_m = 1$, $0 = t_0 < t_1 < \cdots < t_n = 1$. Then $x \sim N\left(0, \sigma^2 A\left(\lambda\right) \otimes B\left(\mu\right)\right)$. For notational convenience, write $\Delta u_i = u_i - u_{i-1}$ $(i = 1, \cdots, m)$ and $\Delta t_i = t_i - t_{i-1}$ $(i = 1, \cdots, n)$. Suppose $\max_i \Delta u_i \to 0$ as $m \to \infty$ and $\max_i \Delta t_i \to 0$ as $n \to \infty$. Define estimators for $\sigma^2\mu$, $\sigma^2\lambda$, and $\sigma^2\lambda\mu$ as

$$\widehat{\sigma^2\mu} = \frac{1}{n} \sum_{i=1}^{m} x_{i\cdot}^T B^{-1}(1) x_{i\cdot} \Delta u_i, \tag{3.5}$$

$$\widehat{\sigma^2\lambda} = \frac{1}{m} \sum_{j=1}^{n} x_{\cdot j}^T A^{-1}(1) x_{\cdot j} \Delta t_j, \tag{3.6}$$

$$\widehat{\sigma^2\lambda\mu} = \frac{1}{mn} x^T \left(A^{-1}(1) \otimes B^{-1}(1)\right) x. \tag{3.7}$$

In what follows, we discuss the asymptotic behaviors of the estimators in (3.5-3.7) as $n \to \infty$ and $m \to \infty$.

Proposition 4. *Under model (3.2), as $n \to \infty$ and $m \to \infty$,*

$$E\widehat{\sigma^2\mu} = \sigma^2\mu - \sigma^2\frac{(\mu+1)^2-4}{2n} + o(n^{-1}),$$

$$E\widehat{\sigma^2\lambda} = \sigma^2\lambda - \sigma^2\frac{(\lambda+1)^2-4}{2m} + o(m^{-1}),$$

$$E\widehat{\sigma^2\lambda\mu} = \sigma^2\lambda\mu - \sigma^2\frac{m\lambda((\mu+1)^2-4) + n\mu((\lambda+1)^2-4)}{2mn} + o(n^{-1}) + o(m^{-1}).$$

Proof. For any $1 \le i \le m$, since $x_{i\cdot} \sim N\left(0, \sigma^2 B(\mu)\right)$, we have

$$E\left(\frac{1}{n} x_{i\cdot}^T B^{-1}(1) x_{i\cdot}\right) = \frac{\sigma^2}{n}\mathrm{Tr}\left(M_\mu^B\right),$$

where $M_\mu^B = B^{-1}(1)B(\mu)$. As a result,

$$E\widehat{\sigma^2\mu} = \sum_{i=1}^m E\left(\frac{1}{n} x_{i\cdot}^T B^{-1}(1) x_{i\cdot}\right)\Delta u_i = \frac{\sigma^2}{n}\mathrm{Tr}\left(M_\mu^B\right)$$

because $\sum_{i=1}^m \Delta u_i = 1$.

It is well known that the $n \times n$ precision matrix $B^{-1}(1)$ has entries as

$$\left(B^{-1}(1)\right)_{i,j} = \begin{cases} \frac{1}{1-\exp(-2|t_1-t_2|)}, & \text{if } i=j=1, \\ \frac{1}{1-\exp(-2|t_{n-1}-t_n|)}, & \text{if } i=j=n, \\ \frac{1}{1-\exp(-2|t_{i-1}-t_i|)} + \frac{\exp(-2|t_i-t_{i+1}|)}{1-\exp(-2|t_i-t_{i+1}|)}, & \text{if } 1<i=j<n, \\ -\frac{\exp(-|t_i-t_j|)}{1-\exp(-2|t_i-t_j|)}, & \text{if } |i-j|=1, \\ 0, & \text{if } |i-j|>1. \end{cases}$$

Thus, the entries of M_μ^B are

$$\left(M_\mu^B\right)_{i,j} = \begin{cases} \tilde{B}_2 b_{1j} - q_{1j}, & \text{if } i=1, \\ (\tilde{B}_i + B_i) b_{ij} - p_{ij} - q_{ij}, & \text{if } 1<i<n, \\ \tilde{B}_n b_{nj} - p_{nj}, & \text{if } i=n, \end{cases} \tag{3.8}$$

where $1 \le j \le n$,

$$B_i = \frac{\exp(-2|t_{i+1} - t_i|)}{1 - \exp(-2|t_{i+1} - t_i|)}, \quad \tilde{B}_i = 1 + B_{i-1}, \quad b_{ij} = \exp(-\mu|t_i - t_j|),$$

$$p_{ij} = \frac{\exp(-|t_{i-1} - t_i|)}{1 - \exp(-2|t_{i-1} - t_i|)} b_{(i-1)j}, \quad q_{ij} = \frac{\exp(-|t_{i+1} - t_i|)}{1 - \exp(-2|t_{i+1} - t_i|)} b_{(i+1)j}.$$

Since $\max_i \Delta t_i \to 0$ as $n \to \infty$, it further holds that

$$\begin{aligned}
\mathrm{Tr}\left(M_\mu^B\right) &= n + 2\sum_{i=2}^{n} \frac{e^{-2\Delta t_i}(1 - e^{-(\mu-1)\Delta t_i})}{1 - e^{-2\Delta t_i}} \\
&= n + (\mu - 1)\sum_{i=2}^{n}\left(1 - \Delta t_i + O\left((\Delta t_i)^2\right)\right) \\
&= n + (\mu - 1)(n - 1 - (1 - t_1) + o(1))
\end{aligned}$$

and $E\widehat{\sigma^2\mu} = \frac{\sigma^2}{n}\mathrm{Tr}\left(M_\mu^B\right) = \sigma^2\mu - \sigma^2\frac{(\mu+1)^2-4}{2n} + o(n^{-1})$ as $n \to \infty$. In a similar manner, there is

$$E\widehat{\sigma^2\lambda} = \sigma^2\lambda - \sigma^2\frac{(\lambda+1)^2 - 4}{2m} + o(m^{-1})$$

as $m \to \infty$.

Moreover,

$$\begin{aligned}
E\widehat{\sigma^2\lambda\mu} &= \frac{1}{mn}Ex^T\left(A^{-1}(1) \otimes B^{-1}(1)\right)x \\
&= \frac{\sigma^2}{mn}\mathrm{Tr}\left(\left(A^{-1}(1) \otimes B^{-1}(1)\right)(A(\lambda) \otimes B(\mu))\right) \\
&= \frac{\sigma^2}{mn}\mathrm{Tr}\left(A^{-1}(1)A(\lambda)\right)\mathrm{Tr}\left(B^{-1}(1)B(\mu)\right) \\
&= \frac{1}{\sigma^2}E\widehat{\sigma^2\lambda}E\widehat{\sigma^2\mu} \\
&= \sigma^2\lambda\mu - \sigma^2\frac{m\lambda((\mu+1)^2 - 4) + n\mu((\lambda+1)^2 - 4)}{2mn} + o(n^{-1}) + o(m^{-1})
\end{aligned}$$

as $n \to \infty$ and $m \to \infty$. This finishes the proof. □

Proposition 4 indicates that $\widehat{\sigma^2\lambda}$, $\widehat{\sigma^2\mu}$, and $\widehat{\sigma^2\lambda\mu}$ are asymptotically unbiased estimators for $\sigma^2\lambda$, $\sigma^2\mu$, and $\sigma^2\lambda\mu$. To further study the convergence of variances of the estimators, we first introduce the following lemma regarding variances of quadratic forms.

Lemma 2. *Under model (3.2), as $n \to \infty$ and $m \to \infty$,*

$$\mathrm{Var}\left(\frac{1}{n} x_{i\cdot}^T B^{-1}(1) x_{i\cdot}\right) = \frac{2}{n}(\sigma^2\mu)^2 + O(n^{-2}), \quad \forall 1 \le i \le m, \tag{3.9}$$

$$\mathrm{Var}\left(\frac{1}{m} x_{\cdot j}^T A^{-1}(1) x_{\cdot j}\right) = \frac{2}{m}(\sigma^2\lambda)^2 + O(m^{-2}), \quad \forall 1 \le j \le n. \tag{3.10}$$

Proof. Since $x_{i\cdot} \sim N\left(0, \sigma^2 B(\mu)\right)$ for any $1 \le i \le m$, we have

$$\mathrm{Var}\left(\frac{1}{n} x_{i\cdot}^T B^{-1}(1) x_{i\cdot}\right) = 2\left(\frac{\sigma^2}{n}\right)^2 \mathrm{Tr}\left((M_\mu^B)^2\right),$$

where $M_\mu^B = B^{-1}(1)B(\mu)$. Recall the entries of M_μ^B in (3.8), we thus have

$$\begin{aligned}
\mathrm{Tr}\left((M_\mu^B)^2\right) =& \left(\tilde{B}_2 b_{11} - q_{11}\right)^2 + 2\left(\tilde{B}_2 b_{1n} - q_{1n}\right)\left(\tilde{B}_n b_{n1} - p_{n1}\right) + \left(\tilde{B}_n b_{nn} - p_{nn}\right)^2 \\
&+ 2\sum_{i=2}^{n-1}\left(\tilde{B}_2 b_{1i} - q_{1i}\right)\left(\left(\tilde{B}_i + B_i\right) b_{i1} - p_{i1} - q_{i1}\right) \\
&+ 2\sum_{i=2}^{n-1}\left(\tilde{B}_n b_{ni} - p_{ni}\right)\left(\left(\tilde{B}_i + B_i\right) b_{in} - p_{in} - q_{in}\right) \\
&+ \sum_{k=2}^{n-1}\sum_{i=2}^{n-1}\left(\left(\tilde{B}_k + B_k\right) b_{ki} - p_{ki} - q_{ki}\right)\left(\left(\tilde{B}_i + B_i\right) b_{ik} - p_{ik} - q_{ik}\right).
\end{aligned}$$

For the convenience of expression, we introduce a few more notations as below. Denote by

$$T_1 = (\tilde{B}_2 - q_{11})^2 + (\tilde{B}_n - p_{nn})^2 + \sum_{k=2}^{n-1}(\tilde{B}_k + B_k - p_{kk} - q_{kk})^2,$$

$$T_2 = \sum_{i=2}^{n-1}\left((\tilde{B}_2 b_{i1} - q_{1i})\left((\tilde{B}_i + B_i)b_{i1} - p_{i1} - q_{i1}\right) + (\tilde{B}_n b_{in} - p_{ni})\left((\tilde{B}_i + B_i)b_{in} - p_{in} - q_{in}\right)\right),$$

$$T_3 = \sum_{\substack{i,k=2 \\ k \ne i}}^{n-1}\left((\tilde{B}_k + B_k)b_{ki} - p_{ki} - q_{ki}\right)\left((\tilde{B}_i + B_i)b_{ik} - p_{ik} - q_{ik}\right),$$

$$T_4 = (\tilde{B}_2 b_{1n} - q_{1n})(\tilde{B}_n b_{1n} - p_{n1}),$$

then $\mathrm{Tr}\left((M_\mu^B)^2\right) = T_1 + 2T_2 + T_3 + 2T_4$. As $n \to \infty$,

$$\begin{aligned}
T_1 =& \frac{1}{2}(\mu+1)^2 - \frac{1}{4}(\mu+1)(\mu^2-1)(\Delta t_2 + \Delta t_n) + (n-2)\mu^2 - \frac{1}{2}\mu(\mu^2-1)(t_n - t_2 + t_{n-1} - t_1) \\
&+ \sum_{k=2}^{n} O((\Delta t_k)^2) + \sum_{k=2}^{n-1} O(\Delta t_k \Delta t_{k+1}) \\
=& n\mu^2 + O(1).
\end{aligned}$$

It thus remains to prove $2T_2 + T_3 + 2T_4 = O(1)$ as $n \to \infty$.

As was previously defined,

$$\tilde{B}_2 b_{1k} - q_{1k} = e^{-\mu(t_k - t_1)} \frac{1 - e^{-(1-\mu)\Delta t_2}}{1 - e^{-2\Delta t_2}}, \quad \forall k \geq 2,$$

$$\tilde{B}_n b_{nk} - p_{nk} = e^{-\mu(t_n - t_k)} \frac{1 - e^{-(1-\mu)\Delta t_n}}{1 - e^{-2\Delta t_n}}, \quad \forall k \leq n-1.$$

For any $2 \leq i, k \leq n-1$ and $i \neq k$,

$$\begin{aligned}
&(\tilde{B}_k + B_k) b_{ki} - p_{ki} - q_{ki} \\
=& \frac{e^{-\mu|t_k - t_i|} - e^{-\Delta t_k - \mu|t_{k-1} - t_i|}}{1 - e^{-2\Delta t_k}} + \frac{e^{-2\Delta t_{k+1} - \mu|t_k - t_i|} - e^{-\Delta t_{k+1} - \mu|t_{k+1} - t_i|}}{1 - e^{-2\Delta t_{k+1}}} \\
=& \begin{cases} e^{-\mu(t_k - t_i)} \left(\frac{1 - e^{-(1-\mu)\Delta t_k}}{1 - e^{-2\Delta t_k}} + \frac{e^{-2\Delta t_{k+1}} - e^{-(1+\mu)\Delta t_{k+1}}}{1 - e^{-2\Delta t_{k+1}}} \right), & \text{if } i \leq k-1, \\ e^{-\mu(t_i - t_k)} \left(\frac{1 - e^{-(1+\mu)\Delta t_k}}{1 - e^{-2\Delta t_k}} + \frac{e^{-2\Delta t_{k+1}} - e^{-(1-\mu)\Delta t_{k+1}}}{1 - e^{-2\Delta t_{k+1}}} \right), & \text{if } i \geq k+1. \end{cases}
\end{aligned}$$

Thus as $n \to \infty$,

$$\begin{aligned}
T_3 =& \frac{1}{8} \left(1 - \mu^2\right)^2 \sum_{\substack{i,k=2 \\ i<k}}^{n-1} e^{-2\mu(t_k - t_i)} \left(t_{k+1} - t_{k-1} + O((\Delta t_k)^2) + O((\Delta t_{k+1})^2) \right) (t_{i+1} - t_{i-1} \\
& + O((\Delta t_i)^2) + O((\Delta t_{i+1})^2)\Big) \\
\propto& \sum_{\substack{i,k=2 \\ i<k}}^{n-1} e^{-2\mu(t_k - t_i)} (t_{k+1} - t_{k-1})(t_{i+1} - t_{i-1}) + o(1) \\
\leq& 2 \sum_{k=2}^{n-1} (\Delta t_k + \Delta t_{k+1}) + o(1) \\
=& O(1).
\end{aligned}$$

Similarly, $T_2 = O(1)$ and $T_4 = O(1)$ as $n \to \infty$. This finishes the proof of (3.9). The proof of (3.10) follows the same manner and is thus omitted. □

Based on Lemma 2, the rates of convergence for $\widehat{\sigma^2 \lambda}$, $\widehat{\sigma^2 \mu}$, and $\widehat{\sigma^2 \lambda \mu}$ are derived as follows.

Proposition 5. *Under model (3.2), as $n \to \infty$ and $m \to \infty$,*

$$\operatorname{Var}(\widehat{\sigma^2 \mu}) = \frac{1}{n\lambda^2} \left(2\lambda - 1 + e^{-2\lambda}\right) (\sigma^2 \mu)^2 + O(n^{-2}),$$

$$\mathrm{Var}(\widehat{\sigma^2\lambda}) = \frac{1}{m\mu^2}\left(2\mu - 1 + e^{-2\mu}\right)(\sigma^2\lambda)^2 + O(m^{-2}),$$

$$\mathrm{Var}(\widehat{\sigma^2\lambda\mu}) = \frac{2}{mn}\left(\sigma^2\lambda\mu\right)^2 + O\left(m^{-2}n^{-1}\right) + O\left(m^{-1}n^{-2}\right).$$

Proof. Under model (3.2),

$$\begin{aligned}
\mathrm{Var}(\widehat{\sigma^2\mu}) &= \mathrm{Var}\left(\frac{1}{n}x^T\left(D_m \otimes B^{-1}(1)\right)x\right) \\
&= 2\left(\frac{\sigma^2}{n}\right)^2 \mathrm{Tr}\left(\left((D_m A(\lambda)) \otimes \left(B^{-1}(1)B(\mu)\right)\right)^2\right) \\
&= 2\left(\frac{\sigma^2}{n}\right)^2 \mathrm{Tr}\left((D_m A(\lambda))^2\right)\mathrm{Tr}\left((M_\mu^B)^2\right),
\end{aligned}$$

where D_m denotes the $m \times m$ diagonal matrix with $(D_m)_{ii} = \Delta u_i$, $i = 1, 2, \ldots, m$.

As $m \to \infty$,

$$\begin{aligned}
\mathrm{Tr}\left((D_m A(\lambda))^2\right) &= \sum_{i,j=1}^{m} \Delta u_i \Delta u_j e^{-2\lambda|u_i - u_j|} \\
&\to \int_0^1 \int_0^1 e^{-2\lambda|x-y|}dxdy \\
&= \frac{2\lambda - 1 + e^{-2\lambda}}{2\lambda^2}.
\end{aligned} \tag{3.11}$$

It follows from the proof of Lemma 2 that $\mathrm{Tr}\left((M_\mu^B)^2\right) = n\mu^2 + O(1)$ as $n \to \infty$. Thus,

$$\mathrm{Var}(\widehat{\sigma^2\mu}) = \frac{1}{n\lambda^2}\left(2\lambda - 1 + e^{-2\lambda}\right)(\sigma^2\mu)^2 + O(n^{-2})$$

as $n \to \infty$ and $m \to \infty$. The proof for the variance of $\widehat{\sigma^2\lambda}$ follows the same manner.

Moreover, as $n \to \infty$ and $m \to \infty$,

$$\begin{aligned}
\mathrm{Var}(\widehat{\sigma^2\lambda\mu}) &= 2\left(\frac{\sigma^2}{mn}\right)^2 \mathrm{Tr}\left(\left(\left(A^{-1}(1)A(\lambda)\right) \otimes \left(B^{-1}(1)B(\mu)\right)\right)^2\right) \\
&= \frac{1}{2\sigma^4}\mathrm{Var}\left(\frac{1}{n}x_{1\cdot}^T B^{-1}(1)x_{1\cdot}\right)\mathrm{Var}\left(\frac{1}{m}x_{\cdot j}^T A^{-1}(1)x_{\cdot j}\right) \\
&= \frac{2}{mn}\left(\sigma^2\lambda\mu\right)^2 + O\left(m^{-2}n^{-1}\right) + O\left(m^{-1}n^{-2}\right)
\end{aligned}$$

by the results of Lemma 2. □

For each of the estimators formulated in (3.5-3.7), its asymptotic distribution is shown in the following theorem.

Theorem 8. *Under model (3.2), as $n \to \infty$ and $m \to \infty$,*

$$\sqrt{n}(\widehat{\sigma^2\mu} - \sigma^2\mu) \xrightarrow{d} N\left(0, \left(\frac{2}{\lambda} - \frac{1-e^{-2\lambda}}{\lambda^2}\right)(\sigma^2\mu)^2\right),$$

$$\sqrt{m}(\widehat{\sigma^2\lambda} - \sigma^2\lambda) \xrightarrow{d} N\left(0, \left(\frac{2}{\mu} - \frac{1-e^{-2\mu}}{\mu^2}\right)(\sigma^2\lambda)^2\right).$$

Furthermore, when $m = rn$ and $n \to \infty$,

$$\sqrt{mn}(\widehat{\sigma^2\lambda\mu} - \sigma^2\lambda\mu) \xrightarrow{d} N\left(-\sigma^2\frac{r\lambda((\mu+1)^2-4)+\mu((\lambda+1)^2-4)}{2\sqrt{r}}, 2(\sigma^2\lambda\mu)^2\right).$$

Proof. Under model (3.2), the joint density of x is

$$p^J_{mn}(\sigma^2,\lambda,\mu) := (2\pi\sigma^2)^{-mn/2}|\,(A(\lambda)\otimes B(\mu))\,|^{-1/2}\exp\left(-\frac{1}{2\sigma^2}x^T\,(A(\lambda)\otimes B(\mu))^{-1}\,x\right). \quad (3.12)$$

For any $m, n \in \mathbb{Z}^+$,

$$\begin{aligned}
\sqrt{mn}(\widehat{\sigma^2\lambda\mu} - \sigma^2\lambda\mu) &= \frac{1}{\sqrt{mn}}x^T\left(\left(A^{-1}(1)\otimes B^{-1}(1)\right) - \lambda\mu\left(A^{-1}(\lambda)\otimes B^{-1}(\mu)\right)\right)x \\
&\quad + \sqrt{mn}\lambda\mu\left(x^T\frac{A^{-1}(\lambda)\otimes B^{-1}(\mu)}{mn}x - \sigma^2\right) \\
&= \frac{2\sigma^2}{\sqrt{mn}}\left(E\log\frac{p^J_{mn}(\sigma^2\lambda\mu,1,1)}{p^J_{mn}(\sigma^2,\lambda,\mu)} - \log\frac{p^J_{mn}(\sigma^2\lambda\mu,1,1)}{p^J_{mn}(\sigma^2,\lambda,\mu)}\right) \\
&\quad + \frac{1}{\sqrt{mn}}Ex^T\left(\left(A^{-1}(1)\otimes B^{-1}(1)\right) - \lambda\mu\left(A^{-1}(\lambda)\otimes B^{-1}(\mu)\right)\right)x \\
&\quad + \sqrt{mn}\lambda\mu\left(x^T\frac{A^{-1}(\lambda)\otimes B^{-1}(\mu)}{mn}x - \sigma^2\right) \\
&= \frac{2\sigma^2}{\sqrt{mn}}\left(E\log\frac{p^J_{mn}(\sigma^2\lambda\mu,1,1)}{p^J_{mn}(\sigma^2,\lambda,\mu)} - \log\frac{p^J_{mn}(\sigma^2\lambda\mu,1,1)}{p^J_{mn}(\sigma^2,\lambda,\mu)}\right) \\
&\quad + \sqrt{mn}\left(E\widehat{\sigma^2\lambda\mu} - \sigma^2\lambda\mu\right) \\
&\quad + \frac{\lambda\mu}{\sqrt{mn}}\left(x^T\left(A^{-1}(\lambda)\otimes B^{-1}(\mu)\right)x - Ex^T\left(A^{-1}(\lambda)\otimes B^{-1}(\mu)\right)x\right).
\end{aligned}$$

Since the probability measure corresponding to $p^J_{mn}(\sigma^2,\lambda,\mu)$ and the probability measure corresponding to $p^J_{mn}(\sigma^2\lambda\mu,1,1)$ are equivalent (Ying, 1993), the Radon-Nikodym derivative satisfies

(Ibragimov and Rozanov, 1978)

$$P\left(0 < \lim_{mn\to\infty} \frac{p^J_{mn}(\sigma^2\lambda\mu, 1, 1)}{p^J_{mn}(\sigma^2, \lambda, \mu)} < \infty\right) = 1 \quad \text{and} \; -\infty < E\log\left(\lim_{mn\to\infty} \frac{p^J_{mn}(\sigma^2\lambda\mu, 1, 1)}{p^J_{mn}(\sigma^2, \lambda, \mu)}\right) < \infty.$$

Thus as $mn \to \infty$,

$$\frac{2\sigma^2}{\sqrt{mn}}\left(E\log\frac{p^J_{mn}(\sigma^2\lambda\mu, 1, 1)}{p^J_{mn}(\sigma^2, \lambda, \mu)} - \log\frac{p^J_{mn}(\sigma^2\lambda\mu, 1, 1)}{p^J_{mn}(\sigma^2, \lambda, \mu)}\right) = \frac{2\sigma^2}{\sqrt{mn}}\left(O(1) - O_p(1)\right) = o_p(1).$$

By the Central Limit Theorem, as $mn \to \infty$

$$\frac{\lambda\mu}{\sqrt{mn}}\left(x^T\left(A^{-1}(\lambda)\otimes B^{-1}(\mu)\right)x - Ex^T\left(A^{-1}(\lambda)\otimes B^{-1}(\mu)\right)x\right) \xrightarrow{d} N(0, 2(\sigma^2\lambda\mu)^2).$$

By Proposition 4, when $m = rn$ and $n \to \infty$,

$$\sqrt{mn}\left(E\widehat{\sigma^2\lambda\mu} - \sigma^2\lambda\mu\right) = -\sigma^2\frac{r\lambda((\mu+1)^2-4) + \mu((\lambda+1)^2-4)}{2\sqrt{r}} + o(1).$$

As a result, when $m = rn$ and $n \to \infty$,

$$\sqrt{mn}(\widehat{\sigma^2\lambda\mu} - \sigma^2\lambda\mu) \xrightarrow{d} N\left(-\sigma^2\frac{r\lambda((\mu+1)^2-4) + \mu((\lambda+1)^2-4)}{2\sqrt{r}}, 2(\sigma^2\lambda\mu)^2\right). \tag{3.13}$$

For any $0 \le u \le 1$, the joint density of $y_{u\cdot} := (X(u,t_1), X(u,t_2), \ldots, X(u,t_n))$ is

$$p^B_n(\sigma^2, \mu; u) := (2\pi\sigma^2)^{-n/2}|B(\mu)|^{-1/2}\exp\left(-\frac{1}{2\sigma^2}y^T_{u\cdot}B^{-1}(\mu)y_{u\cdot}\right). \tag{3.14}$$

Recall that D_m is the $m\times m$ diagonal matrix with $(D_m)_{ii} = \Delta u_i$, $i = 1, 2, \ldots, m$. For any $m, n \in \mathbb{Z}^+$,

$$\begin{aligned}
\sqrt{n}(\widehat{\sigma^2\mu} - \sigma^2\mu) &= \frac{1}{\sqrt{n}}x^T\left(\left(D_m\otimes B^{-1}(1)\right) - \mu\left(D_m\otimes B^{-1}(\mu)\right)\right)x \\
&\quad + \sqrt{n}\sigma^2\mu\left(x^T\frac{D_m\otimes B^{-1}(\mu)}{\sigma^2 n}x - 1\right) \\
&= \frac{2\sigma^2}{\sqrt{n}}\sum_{i=1}^m \Delta u_i\left(E\log\frac{p^B_n(\sigma^2\mu, 1; u_i)}{p^B_n(\sigma^2, \mu; u_i)} - \log\frac{p^B_n(\sigma^2\mu, 1; u_i)}{p^B_n(\sigma^2, \mu; u_i)}\right) \\
&\quad + \sqrt{n}\sigma^2\mu\left(x^T\frac{D_m\otimes B^{-1}(\mu)}{\sigma^2 n}x - 1\right) \qquad (3.15) \\
&\quad + \frac{1}{\sqrt{n}}Ex^T\left(\left(D_m\otimes B^{-1}(1)\right) - \mu\left(D_m\otimes B^{-1}(\mu)\right)\right)x \\
&= \frac{2\sigma^2}{\sqrt{n}}\left(E\log\frac{p^B_n(\sigma^2\mu, 1; u_1)}{p^B_n(\sigma^2, \mu; u_1)} - \sum_{i=1}^m \Delta u_i\log\frac{p^B_n(\sigma^2\mu, 1; u_i)}{p^B_n(\sigma^2, \mu; u_i)}\right) \\
&\quad + \sqrt{n}\left(E\widehat{\sigma^2\mu} - \sigma^2\mu\right) \\
&\quad + \frac{\mu}{\sqrt{n}}\left(x^T\left(D_m\otimes B^{-1}(\mu)\right)x - Ex^T\left(D_m\otimes B^{-1}(\mu)\right)x\right). \qquad (3.16)
\end{aligned}$$

By Proposition 4, as $m \to \infty$ and $n \to \infty$,

$$\sqrt{n}\left(E\widehat{\sigma^2\mu} - \sigma^2\mu\right) = o(1). \tag{3.17}$$

Denote by

$$H_n := \frac{1}{\sqrt{n}}\left(D_m \otimes B^{-1}(\mu)\right)(A(\lambda) \otimes B(\mu)),$$

then as $n \to \infty$,

$$\mathrm{Tr}(H_n^2) = \sum_{k,j=1}^{m} \Delta u_k \Delta u_j e^{-2\lambda|u_k - u_j|} \to \frac{1}{\lambda} - \frac{1 - e^{-2\lambda}}{2\lambda^2},$$

$$\mathrm{Tr}(H_n^4) < \frac{n}{n^2} \to 0.$$

The convergence of moment generating function implies that as $m \to \infty$ and $n \to \infty$,

$$\frac{\mu}{\sqrt{n}}\left(x^T\left(D_m \otimes B^{-1}(\mu)\right)x - Ex^T\left(D_m \otimes B^{-1}(\mu)\right)x\right) \xrightarrow{d} N\left(0, \left(\frac{2}{\lambda} - \frac{1 - e^{-2\lambda}}{\lambda^2}\right)(\sigma^2\mu)^2\right). \tag{3.18}$$

Since $\forall 0 \leq u \leq 1$, the probability measure corresponding to $p_n^B(\sigma^2, \mu; u)$ and the probability measure corresponding to $p_n^B(\sigma^2\mu, 1; u)$ are equivalent (Ying, 1991), the Radon-Nikodym derivative satisfies (Ibragimov and Rozanov, 1978)

$$P\left(0 < \rho_u^B < \infty\right) = 1, \tag{3.19}$$

$$-\infty < E\log\rho_u^B < \infty, \tag{3.20}$$

where $\rho_u^B = \lim_{n\to\infty} \frac{p_n^B(\sigma^2\mu, 1; u)}{p_n^B(\sigma^2, \mu; u)}$.

Moreover, since the probability measure corresponding to $p_{mn}^J(\sigma^2, \lambda, \mu)$ and the probability measure corresponding to $p_{mn}^J(\sigma^2\mu, \lambda, 1)$ are equivalent (Ying, 1993), the Radon-Nikodym derivative satisfies (Ibragimov and Rozanov, 1978)

$$P\left(0 < \lim_{mn\to\infty} \frac{p_{mn}^J(\sigma^2\mu, \lambda, 1)}{p_{mn}^J(\sigma^2, \lambda, \mu)} < \infty\right) = 1.$$

Thus as $m, n \to \infty$,

$$\begin{aligned}
&\log\frac{p_{mn}^J(\sigma^2\mu, \lambda, 1)}{p_{mn}^J(\sigma^2, \lambda, \mu)} \\
=& -\frac{m}{2}\log\frac{|\sigma^2\mu B(1)|}{|\sigma^2 B(\mu)|} - \frac{1}{2}x^T\left(A^{-1}(\lambda) \otimes \left(\frac{1}{\sigma^2\mu}B^{-1}(1) - \frac{1}{\sigma^2}B^{-1}(\mu)\right)\right)x \\
=& O_p(1).
\end{aligned}$$

For any $m, n \geq 1$, denote by

$$J_{mn} = x^T \left(\left(\frac{1}{m} A^{-1}(\lambda) - D_m \right) \otimes \left(\frac{1}{\sigma^2 \mu} B^{-1}(1) - \frac{1}{\sigma^2} B^{-1}(\mu) \right) \right) x.$$

Since $\mathrm{Tr}\,(D_m A(\lambda)) = \sum_{i=1}^m \Delta u_i = 1$ and $\mathrm{Tr}\left((D_m A(\lambda))^2\right) = \sum_{i,j=1}^m \Delta u_i \Delta u_j e^{-2\lambda |u_i - u_k|} = O(1)$, there are

$$\begin{aligned}
EJ_{mn} &= \mathrm{Tr} \left(\left(\left(\frac{1}{m} A^{-1}(\lambda) - D_m \right) \otimes \left(\frac{1}{\sigma^2 \mu} B^{-1}(1) - \frac{1}{\sigma^2} B^{-1}(\mu) \right) \right) \left(A(\lambda) \otimes \sigma^2 B(\mu) \right) \right) \\
&= \mathrm{Tr} \left(\left(\frac{1}{m} I_m - D_m A(\lambda) \right) \otimes \left(\frac{1}{\mu} B^{-1}(1) B(\mu) - I_n \right) \right) \\
&= \mathrm{Tr} \left(\frac{1}{m} I_m - D_m A(\lambda) \right) \mathrm{Tr} \left(\frac{1}{\mu} B^{-1}(1) B(\mu) - I_n \right) \\
&= 0, \quad \forall m, n \geq 1,
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}(J_{mn}) &= 2\mathrm{Tr} \left(\left(\left(\frac{1}{m} I_m - D_m A(\lambda) \right) \otimes \left(\frac{1}{\mu} B^{-1}(1) B(\mu) - I_n \right) \right)^2 \right) \\
&= \frac{1}{m} \left(1 + m\mathrm{Tr} \left((D_m A(\lambda))^2 \right) - 2\mathrm{Tr}\,(D_m A(\lambda)) \right) \left(\frac{1}{\mu^2} \mathrm{Tr} \left((M_\mu^B)^2 \right) + \mathrm{Tr}(I_n) - \frac{2}{\mu} \mathrm{Tr} \left(M_\mu^B \right) \right) \\
&= \frac{1}{m} O(m) \left(n + O(1) + n - 2(n + O(1)) \right) \\
&= O(1) \quad \text{as } m, n \to \infty,
\end{aligned}$$

where $M_\mu^B = B^{-1}(1) B(\mu)$. Thus, $J_{mn} = O_p(1)$ as $m, n \to \infty$. Hence,

$$\begin{aligned}
\sum_{i=1}^m \Delta u_i \log \frac{p_n^B(\sigma^2 \mu, 1; u_i)}{p_n^B(\sigma^2, \mu; u_i)} &= -\frac{1}{2} \log \frac{|\sigma^2 \mu B(1)|}{|\sigma^2 B(\mu)|} - \frac{1}{2} x^T \left(D_m \otimes \left(\frac{1}{\sigma^2 \mu} B^{-1}(1) - \frac{1}{\sigma^2} B^{-1}(\mu) \right) \right) x \\
&= \frac{1}{m} \log \frac{p_{mn}^J(\sigma^2 \mu, \lambda, 1)}{p_{mn}^J(\sigma^2, \lambda, \mu)} + \frac{1}{2} J_{mn} \\
&= O_p(1) \qquad (3.21)
\end{aligned}$$

as $m, n \to \infty$. Moreover, it is implied by (3.20) that

$$E \log \frac{p_n^B(\sigma^2 \mu, 1; u_1)}{p_n^B(\sigma^2, \mu; u_1)} = O(1). \qquad (3.22)$$

As a result of (3.17-3.22), as $m \to \infty$ and $n \to \infty$,

$$\sqrt{n}(\widehat{\sigma^2\mu} - \sigma^2\mu) \xrightarrow{d} N\left(0, \left(\frac{2}{\lambda} - \frac{1-e^{-2\lambda}}{\lambda^2}\right)(\sigma^2\mu)^2\right). \tag{3.23}$$

Similarly, for any $0 \le t \le 1$, the joint density of $y_{\cdot t} := (X(u_1,t), X(u_2,t), \dots, X(u_m,t))$ is

$$p_m^A(\sigma^2, \lambda; t) := (2\pi\sigma^2)^{-m/2}|A(\lambda)|^{-1/2} \exp\left(-\frac{1}{2\sigma^2} y_{\cdot t}^T A^{-1}(\lambda) y_{\cdot t}\right). \tag{3.24}$$

Denote by $\tilde{D}_n$ the $n \times n$ diagonal matrix with $(\tilde{D}_n)_{ii} = \Delta t_i$, $i = 1, 2, \dots, n$. Then for any $m, n \in \mathbb{Z}^+$,

$$\begin{aligned}
\sqrt{m}(\widehat{\sigma^2\lambda} - \sigma^2\lambda) &= \frac{2\sigma^2}{\sqrt{m}}\left(E \log \frac{p_m^A(\sigma^2\lambda, 1; t_1)}{p_m^A(\sigma^2, \lambda; t_1)} - \sum_{i=1}^n \Delta t_i \log \frac{p_m^A(\sigma^2\lambda, 1; t_i)}{p_m^A(\sigma^2, \lambda; t_i)}\right) + \sqrt{m}\left(E\widehat{\sigma^2\lambda} - \sigma^2\lambda\right) \\
&\quad + \frac{\lambda}{\sqrt{m}}\left(x^T\left(A^{-1} \otimes \tilde{D}_n\right)x - Ex^T\left(A^{-1} \otimes \tilde{D}_n\right)x\right) \\
&= o_p(1) + o(1) + \sigma^2\lambda\left(x^T\tilde{H}_m x - Ex^T\tilde{H}_m x\right) \quad \text{as } m, n \to \infty,
\end{aligned} \tag{3.25}$$

where $\tilde{H}_m := \frac{1}{\sqrt{m}}\left(A^{-1}(\lambda) \otimes \tilde{D}_n\right)(A(\lambda) \otimes B(\mu))$. Thus,

$$\sqrt{m}(\widehat{\sigma^2\lambda} - \sigma^2\lambda) \xrightarrow{d} N\left(0, \left(\frac{2}{\mu} - \frac{1-e^{-2\mu}}{\mu^2}\right)(\sigma^2\lambda)^2\right) \tag{3.26}$$

as $m, n \to \infty$. □

3.3 Separable Estimation

Based on the results presented in Section 3.2, define estimators

$$\hat{\lambda} = \frac{\widehat{\sigma^2\lambda\mu}}{\widehat{\sigma^2\mu}} = \frac{x^T\left(A^{-1}(1) \otimes B^{-1}(1)\right)x}{m\sum_{i=1}^m x_{i\cdot}^T B^{-1}(1) x_{i\cdot}\Delta u_i}, \tag{3.27}$$

$$\hat{\mu} = \frac{\widehat{\sigma^2\lambda\mu}}{\widehat{\sigma^2\lambda}} = \frac{x^T\left(A^{-1}(1) \otimes B^{-1}(1)\right)x}{n\sum_{j=1}^n x_{\cdot j}^T A^{-1}(1) x_{\cdot j}\Delta t_j}, \tag{3.28}$$

and

$$\hat{\sigma}^2 = \frac{\widehat{\sigma^2\lambda}\widehat{\sigma^2\mu}}{\widehat{\sigma^2\lambda\mu}} = \frac{\left(\sum_{j=1}^n x_{\cdot j}^T A^{-1}(1) x_{\cdot j}\Delta t_j\right)\left(\sum_{i=1}^m x_{i\cdot}^T B^{-1}(1) x_{i\cdot}\Delta u_i\right)}{x^T\left(A^{-1}(1) \otimes B^{-1}(1)\right)x}, \tag{3.29}$$

where $\widehat{\sigma^2\mu}$, $\widehat{\sigma^2\lambda}$, and $\widehat{\sigma^2\lambda\mu}$ are defined in (3.5-3.7), matrices A and B are defined in (3.4). The main results of this chapter are regarding the joint asymptotic normality and the strong consistency of $\hat{\lambda}$, $\hat{\mu}$, and $\hat{\sigma}^2$.

Theorem 9. *Under model (3.2), if $m/n \to r$ as $n \to \infty$, then*

$$\sqrt{m}\begin{pmatrix} \hat{\lambda}-\lambda \\ \hat{\mu}-\mu \\ \hat{\sigma}^2-\sigma^2 \end{pmatrix} \xrightarrow{d} N\left(0, \begin{pmatrix} rC_\lambda & 0 & -r\sigma^2\frac{C_\lambda}{\lambda} \\ 0 & C_\mu & -\sigma^2\frac{C_\mu}{\mu} \\ -r\sigma^2\frac{C_\lambda}{\lambda} & -\sigma^2\frac{C_\mu}{\mu} & \sigma^4\left(\frac{C_\mu}{\mu^2}+r\frac{C_\lambda}{\lambda^2}\right) \end{pmatrix}\right) \quad as\ n \to \infty, \tag{3.30}$$

where $C_\lambda = 2\lambda - 1 + e^{-2\lambda}$ and $C_\mu = 2\mu - 1 + e^{-2\mu}$.

Proof. It was shown in the proof of Theorem 8 that when $m/n \to r$ and $n \to \infty$,

$$\begin{aligned}\sqrt{m}\begin{pmatrix} \widehat{\sigma^2\mu}-\sigma^2\mu \\ \widehat{\sigma^2\lambda}-\sigma^2\lambda \\ \widehat{\sigma^2\lambda\mu}-\sigma^2\lambda\mu \end{pmatrix} &= \begin{pmatrix} \sqrt{r}\sigma^2\mu\left(x^T\frac{D_m\otimes B^{-1}(\mu)}{\sigma^2\sqrt{n}}x - Ex^T\frac{D_m\otimes B^{-1}(\mu)}{\sigma^2\sqrt{n}}x\right)+o_p(1) \\ \sigma^2\lambda\left(x^T\frac{A^{-1}(\lambda)\otimes\tilde{D}_n}{\sigma^2\sqrt{m}}x - Ex^T\frac{A^{-1}(\lambda)\otimes\tilde{D}_n}{\sigma^2\sqrt{m}}x\right)+o_p(1) \\ \frac{\sigma^2\lambda\mu}{\sqrt{n}}\left(x^T\frac{A^{-1}(\lambda)\otimes B^{-1}(\mu)}{\sigma^2\sqrt{mn}}x - Ex^T\frac{A^{-1}(\lambda)\otimes B^{-1}(\mu)}{\sigma^2\sqrt{mn}}x\right)+O\left(\frac{1}{\sqrt{n}}\right)+o_p(1) \end{pmatrix} \\ &= V - EV + o_p(1),\end{aligned}$$

where

$$V = \left(\sqrt{r}\sigma^2\mu\left(x^T\frac{D_m\otimes B^{-1}(\mu)}{\sigma^2\sqrt{n}}x\right),\quad \sigma^2\lambda\left(x^T\frac{A^{-1}(\lambda)\otimes\tilde{D}_n}{\sigma^2\sqrt{m}}x\right),\quad \frac{\sigma^2\lambda\mu}{\sqrt{n}}\left(x^T\frac{A^{-1}(\lambda)\otimes B^{-1}(\mu)}{\sigma^2\sqrt{mn}}x\right)\right)^T.$$

For any $\gamma = (\gamma_1, \gamma_2, \gamma_3)^T \in \mathbb{R}^3_{>0}$,

$$\begin{aligned}\gamma^T V &= x^T\left(\gamma_1\sqrt{r}\sigma^2\mu\frac{D_m\otimes B^{-1}(\mu)}{\sigma^2\sqrt{n}} + \gamma_2\sigma^2\lambda\frac{A^{-1}(\lambda)\otimes\tilde{D}_n}{\sigma^2\sqrt{m}} + \gamma_3\frac{\sigma^2\lambda\mu}{\sqrt{n}}\frac{A^{-1}(\lambda)\otimes B^{-1}(\mu)}{\sigma^2\sqrt{mn}}\right)x \\ &:= x^T M_{mn} x.\end{aligned}$$

It was revealed in the proof of Theorem 8 that

$$\tilde{M}_{mn} := M_{mn}\sigma^2 A(\lambda)\otimes B(\mu)) \tag{3.31}$$

$$= \gamma_1\sqrt{r}\sigma^2\mu H_n + \gamma_2\sigma^2\lambda\tilde{H}_m + \gamma_3\frac{\sigma^2\lambda\mu}{n\sqrt{m}}I_{mn}, \tag{3.32}$$

where matrices H_n and $\tilde{H}_m$ satisfy that as $m, n \to \infty$,

$$\mathrm{Tr}(H_n^2) \to \frac{1}{\lambda} - \frac{1-e^{-2\lambda}}{2\lambda^2}, \quad \mathrm{Tr}(\tilde{H}_m^2) \to \frac{1}{\mu} - \frac{1-e^{-2\mu}}{2\mu^2};$$

$$\mathrm{Tr}(H_n^k) = o(1), \quad \mathrm{Tr}(\tilde{H}_m^k) = o(1), \quad \forall k \geq 3.$$

Moreover, $\forall m, n \in Z^+$,

$$\mathrm{Tr}(H_n) = \frac{1}{\sqrt{n}}\mathrm{Tr}(D_m A(\lambda) \otimes I_n) = \sqrt{n},$$
$$\mathrm{Tr}(\tilde{H}_m) = \frac{1}{\sqrt{m}}\mathrm{Tr}(I_m \otimes \tilde{D}_n B(\mu)) = \sqrt{m};$$
$$\mathrm{Tr}(H_n\tilde{H}_m) = \frac{1}{\sqrt{mn}}\mathrm{Tr}(D_m A(\lambda) \otimes \tilde{D}_n B(\mu)) = \frac{\mathrm{Tr}(D_m A(\lambda))\mathrm{Tr}(\tilde{D}_n B(\mu))}{\sqrt{mn}} = \frac{1}{\sqrt{mn}};$$
$$\mathrm{Tr}(H_n^k\tilde{H}_m) = \frac{\mathrm{Tr}\left((D_m A(\lambda))^k\right)\mathrm{Tr}(\tilde{D}_n B(\mu))}{\sqrt{n^k m}} = \frac{\mathrm{Tr}(H_n^k)}{n\sqrt{m}}, \quad \mathrm{Tr}(H_n\tilde{H}_m^k) = \frac{\mathrm{Tr}(\tilde{H}_m^k)}{m\sqrt{n}}, \quad \forall k \geq 2;$$
$$\mathrm{Tr}\left((H_n\tilde{H}_m)^2\right) = \frac{\mathrm{Tr}\left((D_m A(\lambda))^2\right)\mathrm{Tr}\left((\tilde{D}_n B(\mu))^2\right)}{mn} = \frac{\mathrm{Tr}(H_n^2)\mathrm{Tr}(\tilde{H}_m^2)}{mn}.$$

Thus when $m/n \to r$ and $n \to \infty$,

$$\begin{aligned}\mathrm{Tr}(\tilde{M}_{mn}^2) &= (\gamma_1\sqrt{r}\sigma^2\mu)^2\mathrm{Tr}(H_n^2) + (\gamma_2\sigma^2\lambda)^2\mathrm{Tr}(\tilde{H}_m^2) + O(\mathrm{Tr}(H_n\tilde{H}_m)) + O\left(\frac{\mathrm{Tr}(H_n)}{n\sqrt{m}}\right)\\ &\quad + O\left(\frac{\mathrm{Tr}(\tilde{H}_m)}{n\sqrt{m}}\right) + O\left(\frac{1}{n}\right)\\ &\to (\gamma_1\sqrt{r}\sigma^2\mu)^2\frac{2\lambda - 1 + e^{-2\lambda}}{2\lambda^2} + (\gamma_2\sigma^2\lambda)^2\frac{2\mu - 1 + e^{-2\mu}}{2\mu^2},\end{aligned} \tag{3.33}$$

$$\begin{aligned}\mathrm{Tr}(\tilde{M}_{mn}^4) &= O\left(\mathrm{Tr}(H_n^4)\right) + O\left(\mathrm{Tr}(\tilde{H}_m^4)\right) + O\left(\mathrm{Tr}((H_n\tilde{H}_m)^2)\right) + O\left(\mathrm{Tr}(H_n^3\tilde{H}_m)\right) + O\left(\mathrm{Tr}(H_n\tilde{H}_m^3)\right)\\ &\quad + \frac{1}{n\sqrt{m}}\left(O\left(\mathrm{Tr}(H_n^3)\right) + O\left(\mathrm{Tr}(\tilde{H}_m^3)\right) + O\left(\mathrm{Tr}(H_n^2\tilde{H}_m)\right) + O\left(\mathrm{Tr}(H_n\tilde{H}_m^2)\right)\right)\\ &\quad + \frac{1}{n^2 m}\left(O\left(\mathrm{Tr}(H_n^2)\right) + O\left(\mathrm{Tr}(H_n\tilde{H}_m)\right) + O\left(\mathrm{Tr}(\tilde{H}_m^2)\right)\right)\\ &\quad + \frac{1}{n^3\sqrt{m^3}}\left(O\left(\mathrm{Tr}(H_n)\right) + O\left(\mathrm{Tr}(\tilde{H}_m)\right)\right)\\ &\to 0,\end{aligned} \tag{3.34}$$

Hence, the convergence of the moment generating function for $\gamma^T(V - EV)$ implies that it is asymptotically Gaussian with zero mean and the variance equals

$$2\lim_{m/n\to r, n\to\infty}\mathrm{Tr}(\tilde{M}_{mn}^2) = 2\left(r(\gamma_1\sigma^2\mu)^2\frac{2\lambda - 1 + e^{-2\lambda}}{2\lambda^2} + (\gamma_2\sigma^2\lambda)^2\frac{2\mu - 1 + e^{-2\mu}}{2\mu^2}\right).$$

By the Cramér–Wold theorem, when $m/n \to r$ as $n \to \infty$,

$$\sqrt{m}\begin{pmatrix}\widehat{\sigma^2\mu} - \sigma^2\mu\\ \widehat{\sigma^2\lambda} - \sigma^2\lambda\\ \widehat{\sigma^2\lambda\mu} - \sigma^2\lambda\mu\end{pmatrix} \xrightarrow{d} N\left(0, \begin{pmatrix}2r(\sigma^2\mu)^2\frac{2\lambda-1+e^{-2\lambda}}{2\lambda^2} & 0 & 0\\ 0 & 2(\sigma^2\lambda)^2\frac{2\mu-1+e^{-2\mu}}{2\mu^2} & 0\\ 0 & 0 & 0\end{pmatrix}\right). \tag{3.35}$$

Define function $g : \mathbb{R}^3_{>0} \mapsto \mathbb{R}^3_{>0}$ as

$$g(x, y, z) = (z/x, z/y, xy/z), \quad \forall (x, y, z) \in \mathbb{R}^3_{>0}. \tag{3.36}$$

Then the Jacobian matrix of g is

$$J_g(x, y, z) = \begin{pmatrix} -z/x^2 & 0 & 1/x \\ 0 & -z/y^2 & 1/y \\ y/z & x/z & -xy/z^2 \end{pmatrix}.$$

It follows from the definition that

$$g\left(\widehat{\sigma^2\mu}, \widehat{\sigma^2\lambda}, \widehat{\sigma^2\lambda\mu}\right) = \left(\hat{\lambda}, \hat{\mu}, \hat{\sigma}^2\right),$$

$$g\left(\sigma^2\mu, \sigma^2\lambda, \sigma^2\lambda\mu\right) = \left(\lambda, \mu, \sigma^2\right),$$

$$J_g\left(\sigma^2\mu, \sigma^2\lambda, \sigma^2\lambda\mu\right) \begin{pmatrix} 2r(\sigma^2\mu)^2 \frac{2\lambda-1+e^{-2\lambda}}{2\lambda^2} & 0 & 0 \\ 0 & 2(\sigma^2\lambda)^2 \frac{2\mu-1+e^{-2\mu}}{2\mu^2} & 0 \\ 0 & 0 & 0 \end{pmatrix} J_g\left(\sigma^2\mu, \sigma^2\lambda, \sigma^2\lambda\mu\right)^T$$

$$= \begin{pmatrix} r(2\lambda - 1 + e^{-2\lambda}) & 0 & -\frac{r\sigma^2}{\lambda}(2\lambda - 1 + e^{-2\lambda}) \\ 0 & 2\mu - 1 + e^{-2\mu} & -\frac{\sigma^2}{\mu}(2\mu - 1 + e^{-2\mu}) \\ -\frac{r\sigma^2}{\lambda}(2\lambda - 1 + e^{-2\lambda}) & -\frac{\sigma^2}{\mu}(2\mu - 1 + e^{-2\mu}) & \sigma^4\left(\frac{2\mu-1+e^{-2\mu}}{\mu^2} + r\frac{2\lambda-1+e^{-2\lambda}}{\lambda^2}\right) \end{pmatrix}.$$

Thus when $m/n \to r$ as $n \to \infty$,

$$\sqrt{m}\begin{pmatrix} \hat{\lambda} - \lambda \\ \hat{\mu} - \mu \\ \hat{\sigma}^2 - \sigma^2 \end{pmatrix}$$

$$\xrightarrow{d} N\left(0, \begin{pmatrix} r(2\lambda - 1 + e^{-2\lambda}) & 0 & -\frac{r\sigma^2}{\lambda}(2\lambda - 1 + e^{-2\lambda}) \\ 0 & 2\mu - 1 + e^{-2\mu} & -\frac{\sigma^2}{\mu}(2\mu - 1 + e^{-2\mu}) \\ -\frac{r\sigma^2}{\lambda}(2\lambda - 1 + e^{-2\lambda}) & -\frac{\sigma^2}{\mu}(2\mu - 1 + e^{-2\mu}) & \sigma^4\left(\frac{2\mu-1+e^{-2\mu}}{\mu^2} + r\frac{2\lambda-1+e^{-2\lambda}}{\lambda^2}\right) \end{pmatrix}\right).$$

The proof is finished using the multivariate delta method. □

Remark 2. *The estimators* $\hat{\lambda}$ *and* $\hat{\mu}$ *are asymptotically independent. This is due to the zero entries of* J_g *as well as the asymptotic independence of* $\widehat{\sigma^2\mu}$ *and* $\widehat{\sigma^2\lambda}$*, which is based on the fact that* $Tr(D_m A(\lambda)) = Tr(\tilde{D}_n B(\mu)) = 1, \forall m, n.$

Besides the asymptotic normality, estimators $\hat{\lambda}$, $\hat{\mu}$, and $\hat{\sigma}^2$ are also strongly consistent.

Theorem 10. *Under model (3.2), as* $m, n \to \infty$*,*

$$\left(\hat{\lambda}, \hat{\mu}, \hat{\sigma}^2\right) \stackrel{a.s.}{\to} \left(\lambda, \mu, \sigma^2\right). \tag{3.37}$$

Proof. Since the function g defined in (3.36) is a continuous function, the continuous mapping theorem makes it suffice to prove

$$\left(\widehat{\sigma^2\mu}, \widehat{\sigma^2\lambda}, \widehat{\sigma^2\lambda\mu}\right) \stackrel{a.s.}{\to} \left(\sigma^2\mu, \sigma^2\lambda, \sigma^2\lambda\mu\right)$$

as $m, n \to \infty$.

For any $(\lambda_0, \mu_0) \in \mathbb{R}^2_{>0}$, there always exists a compact region $\mathbb{C}$ in $\mathbb{R}^2_{>0}$ that contains (λ_0, μ_0) and $(1, 1)$ as its interior points. Therefore (4.13) and (4.14) in the proof of Theorem 1 in Ying (1993) both hold. Namely, as $n \to \infty$,

$$\begin{aligned}
&x_{1\cdot}^T B^{-1}(1) x_{1\cdot} + \sum_{i=2}^{m} \frac{(x_{i\cdot} - e^{-\varepsilon_i} x_{(i-1)\cdot})^T B^{-1}(1)(x_{i\cdot} - e^{-\varepsilon_i} x_{(i-1)\cdot})}{1 - e^{-2\varepsilon_i}} \\
&\stackrel{a.s.}{=} \lambda_0\mu_0\sigma_0^2 \sum_{i=2}^{m}\sum_{k=2}^{n} \omega_{ik}^2 + [\lambda_0\sigma_0^2 + \lambda_0\mu_0\sigma_0^2(1-\mu_0) + \frac{\lambda_0(1-\mu_0)^2\sigma_0^2}{2}]m \\
&\quad + [\mu_0\sigma_0^2 + \lambda_0\mu_0\sigma_0^2(1-\lambda_0) + \frac{\mu_0(1-\lambda_0)^2\sigma_0^2}{2}]n + o(n).
\end{aligned}$$

As a result, as $m, n \to \infty$,

$$\begin{aligned}
&l_{m,n}(1,1,\sigma^2) - l_{m,n}(1,1,\lambda_0\mu_0\sigma_0^2)\\
=&(1+m-1+n-1+(m-1)(n-1))\log(\frac{\sigma^2}{\lambda_0\mu_0\sigma_0^2})\\
&+(\frac{1}{\sigma^2}-\frac{1}{\lambda_0\mu_0\sigma_0^2})[x_1^T B^{-1}(1)x_1 + \sum_{i=2}^{m}\frac{(x_i - e^{-\varepsilon_i}x_{i-1})^T B^{-1}(1)(x_i - e^{-\varepsilon_i}x_{i-1})}{1-e^{-2\varepsilon_i}}]\\
=&(\frac{\lambda_0\mu_0\sigma_0^2}{\sigma^2}-1)\sum_{i=2}^{m}\sum_{k=2}^{n}\omega_{ik}^2 - (m-1)(n-1)\log(\frac{\lambda_0\mu_0\sigma_0^2}{\sigma^2})\\
&+(\frac{1}{\sigma^2}-\frac{1}{\lambda_0\mu_0\sigma_0^2})[\lambda_0\sigma_0^2 + \lambda_0\mu_0\sigma_0^2(1-\mu_0) + \frac{\lambda_0(1-\mu_0)^2\sigma_0^2}{2}]m\\
&+(\frac{1}{\sigma^2}-\frac{1}{\lambda_0\mu_0\sigma_0^2})[\mu_0\sigma_0^2 + \lambda_0\mu_0\sigma_0^2(1-\lambda_0) + \frac{\mu_0(1-\lambda_0)^2\sigma_0^2}{2}]n\\
&+(m+n-1)\log(\frac{\sigma^2}{\lambda_0\mu_0\sigma_0^2}) + o(n)\\
\overset{a.s.}{=}&(m-1)(n-1)(\frac{\lambda_0\mu_0\sigma_0^2}{\sigma^2} - 1 - \log(\frac{\lambda_0\mu_0\sigma_0^2}{\sigma^2})) + o(mn),
\end{aligned} \tag{3.38}$$

where the last equality holds since $\sum_{i=2}^{m}\sum_{k=2}^{n}(\omega_{ik}^2 - 1) = o(mn)$ almost surely. Thus,

$$l_{m,n}(1,1,\sigma^2) - l_{m,n}(1,1,\lambda_0\mu_0\sigma_0^2) \to \infty \quad \text{a.s.}$$

as $m, n \to \infty$ if $\sigma^2 \neq \lambda_0\mu_0\sigma_0^2$. Together with Lemma 4 in Ying (1991), the result above entails

$$\operatorname*{argmin}_{\sigma^2} l_{m,n}(1,1,\sigma^2) \overset{a.s.}{\to} \lambda_0\mu_0\sigma_0^2 \tag{3.39}$$

as $m, n \to \infty$. Hence as $m, n \to \infty$, $\widehat{\sigma^2\lambda\mu} \overset{a.s.}{\to} \sigma^2\lambda\mu$.

It remains to prove that as $m, n \to \infty$, $\widehat{\sigma^2\mu} \overset{a.s.}{\to} \sigma^2\mu$ and $\widehat{\sigma^2\lambda} \overset{a.s.}{\to} \sigma^2\lambda$. It follows from the definition that under model (3.2),

$$\widehat{\sigma^2\mu} = \frac{1}{n}x^T\left(D_m \otimes B^{-1}(1)\right)x \overset{d}{=} \epsilon^T\Lambda_{mn}\epsilon,$$

where $\epsilon \sim N(0, I_{mn})$ and Λ_{mn} is a diagonal matrix whose diagonal entries are eigenvalues of the matrix

$$\frac{\sigma^2}{n}\left((A^{1/2}(\lambda))^T D_m A^{1/2}(\lambda)\right) \otimes \left((B^{1/2}(\mu))^T B^{-1}(1)B^{1/2}(\mu)\right).$$

By the result of Proposition 5,

$$\begin{aligned}||\Lambda_{mn}||_F^2 =&\mathrm{Tr}\left(\left(\frac{\sigma^2}{n}\left((A^{1/2}(\lambda))^T D_m A^{1/2}(\lambda)\right)\otimes\left((B^{1/2}(\mu))^T B^{-1}(1) B^{1/2}(\mu)\right)\right)^2\right)\\ =&\frac{1}{2}\operatorname{Var}\left(\widehat{\sigma^2\mu}\right)\\ =&O(n^{-1}) \quad \text{as } m,n\to\infty.\end{aligned}$$

Moreover, $||\Lambda_{mn}||_2 \leq ||\Lambda_{mn}||_F = O(n^{-1/2})$ as $m,n\to\infty$. Thus, the Hanson-Wright inequality implies that for sufficiently large n, $\exists C_0 > 0$ such that

$$\begin{aligned}P\left(\left|\widehat{\sigma^2\mu} - E\widehat{\sigma^2\mu}\right| \geq \xi\right) &\leq 2\exp\left(-C\min\left\{\frac{\xi}{||\Lambda_{mn}||_2}, \frac{\xi^2}{||\Lambda_{mn}||_F^2}\right\}\right)\\ &\leq 2\exp(-C_0\sqrt{n}\xi), \quad \forall\xi > 0,\end{aligned} \tag{3.40}$$

where $C > 0$ is an absolute constant. It hence follows from the Borel–Cantelli lemma that $\widehat{\sigma^2\mu} - E\widehat{\sigma^2\mu} \overset{a.s.}{\to} 0$ as $m,n\to\infty$. By the results of Proposition 4,

$$\widehat{\sigma^2\mu} - \sigma^2\mu = \widehat{\sigma^2\mu} - E\widehat{\sigma^2\mu} + E\widehat{\sigma^2\mu} - \sigma^2\mu \overset{a.s.}{\to} 0 \tag{3.41}$$

as $m,n\to\infty$.

In a similar manner, it can be proved that $\widehat{\sigma^2\lambda} \overset{a.s.}{\to} \sigma^2\lambda$ as $m,n\to\infty$. This finishes the proof. □

3.4 Simulation

Let $\lambda = 0.5$, $\mu = 10$, $\sigma^2 = 4$. For each value of the sample size $n = 500, 600, \ldots, 2000$ and $m = 0.5n$, we set irregular sampling locations as $u_0 = t_0 = 0$, $u_m = t_n = 1$, and

$$(u_i, t_j) = \left(\frac{i}{m} + U_u^i, \frac{j}{n} + U_t^j\right), \quad \forall 0 < i < m, 0 < j < n,$$

where $U_u^i \overset{i.i.d.}{\sim} U\left(-\frac{1}{2m}, \frac{1}{2m}\right)$ and $U_t^j \overset{i.i.d.}{\sim} U\left(-\frac{1}{2n}, \frac{1}{2n}\right)$ are independent uniformly distributed random variables. Given sampling locations, we run 1000 realizations and calculate $\hat{\lambda}$, $\hat{\mu}$, and $\hat{\sigma}^2$ as defined in Section 3.3. One realization when $n = 500$ is shown in Figure 3.1. The averaged absolute value

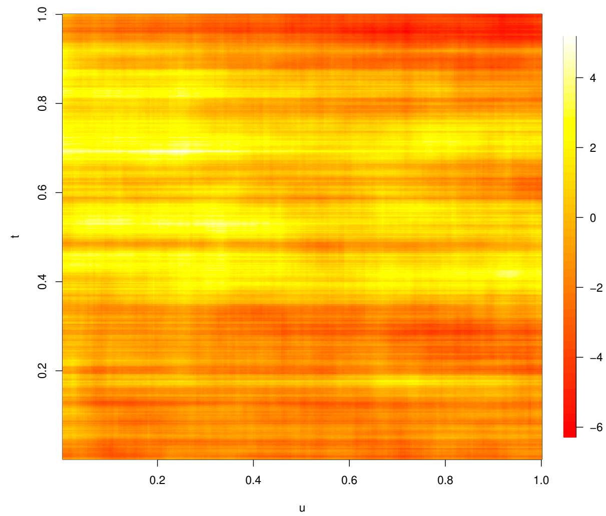

Figure 3.1A simulated OU field with $m = 250$ and $n = 500$.

Table 3.1Empirical quantiles of standardized bias when estimating λ.

	λ			$N(0,1)$
n	500	1000	2000	
5%	-1.4462	-1.5188	-1.4547	-1.6448
25%	-0.6030	-0.5308	-0.5681	-0.6744
50%	0.1559	0.0893	0.0763	0
75%	0.9224	0.7342	0.7039	0.6744
95%	1.9886	1.8533	1.7377	1.6448

of bias for each sample size and the histogram of bias when $n = 2000$ are shown in Figure 3.2. For $n = 500, 1000, 2000$, some empirical quantiles of

$$\frac{\sqrt{m}(\hat{\lambda} - \lambda)}{\sqrt{r(2\lambda - 1 + e^{-2\lambda})}}, \quad \frac{\sqrt{m}(\hat{\mu} - \mu)}{\sqrt{2\mu - 1 + e^{-2\mu}}}, \quad \text{and} \quad \frac{\sqrt{m}(\hat{\sigma}^2 - \sigma^2)}{\sqrt{\sigma^4\left(\frac{2\mu - 1 + e^{-2\mu}}{\mu^2} + r\frac{2\lambda - 1 + e^{-2\lambda}}{\lambda^2}\right)}}$$

are shown in Tables 3.1-3.3.

3.5 Discussion

We proposed estimators for covariance parameters of an anisotropic Ornstein-Uhlenbeck field observed on $[0, 1]^2$. The estimators $\hat{\lambda}$, $\hat{\mu}$, and $\hat{\sigma}^2$ formulated in Section 3.3 are strongly consistent and have lower computational complexity than the MLEs of λ, μ, and σ^2. As the sample size goes to infinity, the estimators we proposed asymptotically follow normal distribution, but have higher

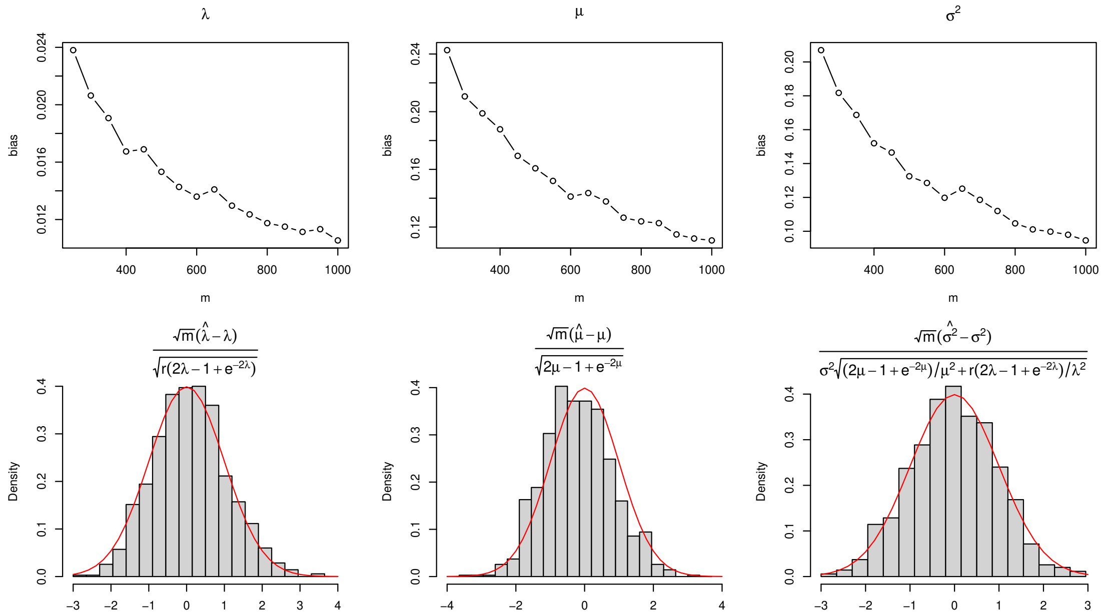

Figure 3.2The plots in the first row present averaged absolute values of bias for $n = 500, 600, \ldots, 2000$ and $m = n/2$ among 1000 realizations. The second row of plots present the empirical distributions of bias with 1000 realizations when $n = 2000$ and $m = 1000$, where the red curve indicates the density function of $N(0, 1)$.

Table 3.2Empirical quantiles of standardized bias when estimating μ.

	μ			$N(0,1)$
n	500	1000	2000	
5%	-1.9248	-1.8978	-1.6819	-1.6448
25%	-1.0806	-0.9460	-0.8577	-0.6744
50%	-0.3960	-0.3193	-0.2047	0
75%	0.3473	0.3897	0.4762	0.6744
95%	1.4002	1.3349	1.5174	1.6448

Table 3.3Empirical quantiles of standardized bias when estimating σ^2.

	σ^2			$N(0,1)$
n	500	1000	2000	
5%	-1.6784	-1.5597	-1.6782	-1.6448
25%	-0.7708	-0.6583	-0.6788	-0.6744
50%	-0.0558	-0.0277	0	0
75%	0.7103	0.6323	0.6719	0.6744
95%	1.7328	1.5499	1.4902	1.6448

variance compared with the MLEs studied by Ying (1993). This presents a trade-off between the computational cost and the estimation accuracy.

The sampling grid based on which $\hat{\lambda}$, $\hat{\mu}$, and $\hat{\sigma}^2$ are formulated is defined by lines parallel to the coordinate axes. For a significantly anisotropic OU field such as the one shown in Figure 3.1, the coordinate axes are distinguishable. When values of λ and μ are close, however, it could be difficult to determine directions along which observations should be taken. It is thus of interest to study the properties of estimators when sampling directions are not parallel to the coordinate axes.

The main results presented in this chapter focus on the asymptotic behaviors of the estimators. It would also be interesting to study their finite-sample distributions and measure the distance between a finite-sample distribution and the asymptotic distribution. The statistical inference for parameters λ, μ, and σ^2 is also worth analyzing. The exploration of these topics is reserved for future research work.

CHAPTER 4

VECCHIA APPROXIMATION

4.1 Introduction

Consider a zero-mean Gaussian process X with the Matérn covariance function

$$Cov(X(t), X(t+d)) = K(d) = \sigma^2 \frac{(\theta d)^\nu}{\Gamma(\nu) 2^{\nu-1}} \mathcal{K}_\nu(\theta d), \tag{4.1}$$

where $\theta > 0$, $\nu > 0$, $\sigma^2 > 0$, Γ is the gamma function, and $\mathcal{K}_\nu$ is the modified Bessel function of the second kind. Denote by $X_n = (X(t_1^n), X(t_2^n), \ldots, X(t_n^n))$ the observations of X with sample size n. When $\nu \neq \frac{1}{2}$, X is not Markovian and the sparse precision matrix of X_n discussed in Chapter 3 is not valid. It is thus necessary to study other approaches to reduce the computational cost of the MLE. The existing approaches to achieve computational efficiency include covariance tapering (Furrer et al., 2006; Kaufman et al., 2008; Du et al., 2009), Gaussian Markov random fields representation (Rue and Held, 2005; Lindgren et al., 2011), multiresolution approximation (Nychka et al., 2015; Katzfuss, 2017), etc.

The Vecchia approximation is a method to reduce the computational burden through sparse precision matrices. Write the joint density function of $X(t_1^n), X(t_2^n), \ldots, X(t_n^n)$ as

$$f_n = f_{X(t_1^n)} \prod_{i=2}^{n} f_{X(t_i^n)|X(t_{i-1}^n)\ldots X(t_1^n)}.$$

The Vecchia's method (Vecchia, 1988) approximates f_n by

$$\hat{f}_n = f_{X(t_1^n)} \prod_{i=2}^{n} f_{X(t_i^n)|X(t_{i-1}^n)\ldots X(t_{1\vee(i-k)}^n)} \tag{4.2}$$

for some $k \ll n$, which makes the precision matrix of X_n a band matrix and could thus significantly reduce the computational complexity. The accuracy of Vecchia approximation has been discussed in both theoretical and practical aspects (Stein et al., 2004; Datta et al., 2016; Guinness, 2018; Finley et al., 2019; Zhang et al., 2021; Cao et al., 2022). Under a more general framework proposed by Katzfuss and Guinness (2021), where the conditioning vector contains both observed data and latent variables, the nearest-neighbor Gaussian process, latent autoregressive process, multiresolution approximation, and many other popular Gaussian process approximation methods are special cases of the Vecchia approach.

In the remainder of this chapter, we focus on the standard Vecchia approximation and estimate the scale parameter in the Matérn covariance function by MLE solved from the approximated likelihood. The effects of the misspecified range parameter and the conditioning variables on the bias are discussed in Section 4.2, and simulation results are presented in Section 4.3.

4.2 Maximum Likelihood Estimator for σ^2

Under a regular sampling design on fixed domain, we have $t_i^n = i/n$ for $i = 1, 2, \ldots, n$. When ν is known, the expectation of MLE for σ^2 from Vecchia approximation satisfies the following results.

Proposition 6. *Denote by $\hat{\sigma}^2$ the MLE for σ^2 from Vecchia approximation with ν known and θ replaced by some fixed $\theta_0 > 0$. When $k = 1$ in (4.2), $E\hat{\sigma}^2 = \sigma^2$ for any $n \geq 2$ if $\theta_0 = \theta$, and*

$$E\hat{\sigma}^2 = \begin{cases} \sigma^2 \left(\frac{\theta}{\theta_0}\right)^{2\nu} + O(n^{2\nu-2}) + O(n^{-1}) + O(n^{-2\nu}), & \nu < 1, \\ \sigma^2 \left(\frac{\theta}{\theta_0}\right)^{2} + O((\log n)^{-1}), & \nu = 1, \\ \sigma^2 \left(\frac{\theta}{\theta_0}\right)^{2} + O(n^{-1}) + O(n^{2-2\nu}), & \nu > 1, \nu \notin \mathbb{Z} \end{cases}$$

as $n \to \infty$ if $\theta_0 \neq \theta$. When $k = 2$ in (4.2) and $\theta_0 \neq \theta$,

$$E\hat{\sigma}^2 = \begin{cases} \sigma^2 \left(\frac{\theta}{\theta_0}\right)^{2\nu} + O(n^{2\nu-2}) + O(n^{-1}) + O(n^{-2\nu}), & \nu < 1, \\ \sigma^2 \left(\frac{\theta}{\theta_0}\right)^{2} + O((\log n)^{-1}), & \nu = 1, \\ \sigma^2 \left(\frac{\theta}{\theta_0}\right)^{2\nu} + O(n^{-1}) + O(n^{2-2\nu}) + O(n^{2\nu-4}), & 1 < \nu < 2, \\ \sigma^2 \left(\frac{\theta}{\theta_0}\right)^{4} + O((\log n)^{-1}), & \nu = 2, \\ \sigma^2 \left(\frac{\theta}{\theta_0}\right)^{4} + \frac{\sigma^2\beta^2}{6\tau-\beta^2}\left(\left(\frac{\theta}{\theta_0}\right)^{2} - 1\right)^2 + O(n^{-1}) + O(n^{4-2\nu}), & \nu > 2, \nu \notin \mathbb{Z} \end{cases}$$

as $n \to \infty$, where $\tau = \frac{\Gamma(1-\nu)}{2^5\Gamma(3-\nu)}$ and $\beta = \frac{1}{4(1-\nu)}$.

Proof. Denote for $1 \leq i \leq n$ that

$$K_{n,i}^0 = \frac{(\theta_0 i/n)^\nu}{\Gamma(\nu)2^{\nu-1}}\mathcal{K}_\nu(\theta_0 i/n)$$

for some fixed $\theta_0 > 0$, and write $K_{n,i} = \sigma^{-2}K(i/n)$. It follows from (9.6.2) and (9.6.10) in Abramowitz and Stegun (1948) that for $\nu \notin \mathbb{Z}$,

$$\frac{x^\nu}{\Gamma(\nu)2^{\nu-1}}\mathcal{K}_\nu(x) = 1 - \alpha x^{2\nu} + \beta x^2 + \tau x^4 + O(x^{2\nu+2}) + O(x^6) + O(x^{2\nu+4}) \quad \text{as } x \to 0, \tag{4.3}$$

where $\alpha = \frac{\Gamma(1-\nu)}{4^\nu\Gamma(1+\nu)}$, $\tau = \frac{\Gamma(1-\nu)}{2^5\Gamma(3-\nu)}$, and $\beta = \frac{1}{4(1-\nu)}$. The gamma function Γ on $\mathbb{R}$ is defined as

$$\Gamma(x) = \begin{cases} \int_0^\infty t^{x-1}e^{-t}\mathrm{d}t, & x > 0, \\ \frac{\Gamma(x+n+1)}{x(x+1)\cdots(x+n)}, & x < 0, x \notin \mathbb{Z}, \end{cases} \tag{4.4}$$

where n is chosen such that $x+n > 0$. For $\nu \in \mathbb{Z}$, it follows from (9.6.10) and (9.6.11) in Abramowitz and Stegun (1948) that

$$\begin{aligned}\frac{x^\nu}{\Gamma(\nu)2^{\nu-1}}\mathcal{K}_\nu(x) &= \sum_{k=0}^{\nu-1}(-1)^k\frac{(\nu-k-1)!}{k!(\nu-1)!}\left(\frac{x}{2}\right)^{2k} + \frac{2(-1)^{\nu+1}}{(\nu-1)!}\log\left(\frac{x}{2}\right)\sum_{k=0}^{\infty}\frac{1}{k!(\nu+k)!}\left(\frac{x}{2}\right)^{2\nu+2k} \\ &\quad + (-1)^\nu\sum_{k=1}^{\infty}\frac{\sum_{h=1}^{k}\frac{2}{h} + \sum_{h=k+1}^{k+\nu}\frac{1}{h} - 2\gamma}{k!(\nu+k)!(\nu-1)!}\left(\frac{x}{2}\right)^{2\nu+2k} \\ &\quad + \frac{(-1)^\nu}{(\nu-1)!\nu!}\left(\sum_{h=1}^{\nu}\frac{1}{h} - 2\gamma\right)\left(\frac{x}{2}\right)^{2\nu} \\ &= \sum_{k=0}^{\infty}\left(c_{\nu,k}x^{2k} + \tilde{c}_{\nu,k}x^{2\nu+2k}\log x\right),\end{aligned} \tag{4.5}$$

where γ is the Euler's constant, $c_{\nu,k}, \tilde{c}_{\nu,k}$ are constants depending only on ν and k.

Case 1. When $k = 1$, the approximated joint density is

$$\hat{f}_n(x_1, \ldots, x_n) = (2\pi\sigma^2)^{-\frac{n}{2}}(1 - K_{n,1}^2)^{-\frac{n-1}{2}}\exp\left(-\frac{1}{2\sigma^2}\left(x_1^2 + \frac{1}{1-K_{n,1}^2}\sum_{i=2}^{n}(x_i - x_{i-1}K_{n,1})^2\right)\right) \tag{4.6}$$

since $X(t_i^n)|X(t_{i-1}^n) \sim N\left(X(t_{i-1}^n)K_{n,1}, \sigma^2(1-K_{n,1}^2)\right)$. Hence,

$$\log\hat{f}_n(x_1, \ldots, x_n)\ |_{\theta=\theta_0} = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\left(x_1^2 + \frac{1}{1-(K_n^0)^2}\sum_{i=2}^{n}(x_i - x_{i-1}K_n^0)^2\right) + C, \tag{4.7}$$

where $K_n^0 = K_{n,1}^0$, C is a constant not depending on σ^2. The MLE of σ^2 calculated from (4.7) is thus

$$\hat{\sigma}^2 = \frac{1}{n}\left(x_1^2 + \frac{1}{1-(K_n^0)^2}\sum_{i=2}^{n}(x_i - x_{i-1}K_n^0)^2\right), \tag{4.8}$$

where $x_i = X(i/n)$, $i = 1, \ldots, n$. Under model (4.1), there is

$$E\hat{\sigma}^2 = \frac{\sigma^2}{n}\left(1 + (n-1)\frac{1 + (K_n^0)^2 - 2K_n^0 K_{n,1}}{1 - (K_n^0)^2}\right)$$

for any $n \geq 2$. Consequently, $E\hat{\sigma}^2 = \sigma^2$ always holds when $\theta_0 = \theta$. Cases when $\theta_0 \neq \theta$ are discussed below.

When $0 < \nu < 1$, (4.3) implies that as $n \to \infty$,

$$\begin{aligned}\frac{1 + (K_n^0)^2 - 2K_n^0 K_{n,1}}{1 - (K_n^0)^2} &= \frac{\theta^{2\nu} + \alpha n^{-2\nu}\theta^{2\nu}(\theta^{2\nu} - \theta_0^{2\nu}/2) - n^{2\nu-2}\theta^2\beta/\alpha + O(n^{-2})}{\theta_0^{2\nu} + \alpha n^{-2\nu}(\theta_0^{4\nu}/2) - n^{2\nu-2}\theta_0^2\beta/\alpha + O(n^{-2})} \\ &= \left(\frac{\theta}{\theta_0}\right)^{2\nu} + O(n^{2\nu-2}) + O(n^{-2\nu}) + O(n^{-2}).\end{aligned}$$

Hence,

$$E\hat{\sigma}^2 = \sigma^2\left(\frac{\theta}{\theta_0}\right)^{2\nu} + O(n^{2\nu-2}) + O(n^{-1}) + O(n^{-2\nu}).$$

When $\nu > 1$ and $\nu \notin \mathbb{Z}$, (4.3) implies that as $n \to \infty$,

$$\begin{aligned}\frac{1 + (K_n^0)^2 - 2K_n^0 K_{n,1}}{1 - (K_n^0)^2} &= \frac{-2\beta\theta^2/n^2 + 2\alpha\theta^{2\nu}/n^{2\nu} + \beta^2(\theta_0^4 - 2\theta^2\theta_0^2)/n^4 - 2\tau(\theta/n)^4 + O(n^{-2-2\nu})}{-2\beta\theta_0^2/n^2 + 2\alpha\theta_0^{2\nu}/n^{2\nu} - \beta^2(\theta_0/n)^4 - 2\tau(\theta_0/n)^4 + O(n^{-2-2\nu})} \\ &= \left(\frac{\theta}{\theta_0}\right)^{2\nu} + O(n^{2-2\nu}) + O(n^{-2\nu}) + O(n^{-2})\end{aligned}$$

and

$$E\hat{\sigma}^2 = \sigma^2\left(\frac{\theta}{\theta_0}\right)^{2\nu} + O(n^{2-2\nu-2}) + O(n^{-1}).$$

When $\nu = 1$, it follows from (4.5) that

$$\frac{x^\nu}{\Gamma(\nu)2^{\nu-1}}\mathcal{K}_\nu(x) = 1 + c_1x^2\log(1/x) + c_2x^2 + c_3x^4\log(1/x) + c_4x^4 + O(x^6\log x) \tag{4.9}$$

as $x \to 0$, where c_1, c_2, c_3, c_4 are constants only depending on ν. Thus,

$$\begin{aligned}&\frac{1 + (K_n^0)^2 - 2K_n^0 K_{n,1}}{1 - (K_n^0)^2} \\ =&\frac{r^2 2c_1n^{-2}\log n + 2c_2n^{-2} - 2c_1n^{-2}(r^2\log r + (1 + r^2)c_2/c_1) + O(n^{-4}(\log n)^2)}{2c_1n^{-2}\log n - 2c_2n^{-2} + O(n^{-4}(\log n)^2)} \\ =&r^2 + O((\log n)^{-1}) + O((\log n)^{-2}),\end{aligned}$$

where $r = \theta/\theta_0$. Hence,

$$E\hat{\sigma}^2 = \sigma^2 \left(\frac{\theta}{\theta_0}\right)^2 + O((\log n)^{-1}).$$

Case 2. When $k = 2$, the approximated joint density is

$$\begin{aligned}&\hat{f}_n(x_1, \dots, x_n)\\ =&\frac{(2\pi\sigma^2 b)^{-\frac{n}{2}} b}{\sqrt{1 - K_{n,1}^2}} \exp\left(-\frac{1}{2\sigma^2}\left(x_1^2 + \frac{(x_2 - K_{n,1}x_1)^2}{1 - K_{n,1}^2} + \frac{1}{b}\sum_{i=3}^{n}(x_i - a_1 x_{i-1} - a_2 x_{i-2})^2\right)\right),\end{aligned}$$

where $a_1 = \frac{K_{n,1} - K_{n,1}K_{n,2}}{1-(K_{n,1})^2}$, $a_2 = \frac{K_{n,2} - (K_{n,1})^2}{1-(K_{n,1})^2}$, and $b = 1 - \frac{(K_{n,1})^2 + (K_{n,2})^2 - 2(K_{n,1})^2 K_{n,2}}{1-(K_{n,1})^2}$. This is due to

$$\begin{pmatrix} X(t_i^n) \\ X(t_{i-1}^n) \\ X(t_{i-2}^n) \end{pmatrix} \sim N\left(0, \sigma^2 \begin{pmatrix} 1 & K(|t_i^n - t_{i-1}^n|) & K(|t_i^n - t_{i-2}^n|) \\ K(|t_i^n - t_{i-1}^n|) & 1 & K(|t_{i-1}^n - t_{i-2}^n|) \\ K(|t_i^n - t_{i-2}^n|) & K(|t_{i-1}^n - t_{i-2}^n|) & 1 \end{pmatrix}\right)$$

and the regular sampling design, which implies that $\forall 3 \le i \le n$,

$$X(t_i^n)|(X(t_{i-1}^n), X(t_{i-2}^n)) \sim N\left(a_1 X(t_{i-1}^n) + a_2 X(t_{i-2}^n), \sigma^2 b\right).$$

Take $\arg\max_{\sigma^2} \log \hat{f}_n$ and plug in $\theta = \theta_0$, then

$$\hat{\sigma}^2 = \frac{1}{n}\left(x_1^2 + \frac{(x_2 - K_{n,1}^0 x_1)^2}{1 - (K_{n,1}^0)^2} + \frac{1}{b^0}\sum_{i=3}^{n}(x_i - a_1^0 x_{i-1} - a_2^0 x_{i-2})^2\right), \tag{4.10}$$

where $a_1^0 = \frac{K_{n,1}^0 - K_{n,1}^0 K_{n,2}^0}{1-(K_{n,1}^0)^2}$, $a_2^0 = \frac{K_{n,2}^0 - (K_{n,1}^0)^2}{1-(K_{n,1}^0)^2}$, and $b^0 = \frac{(K_{n,1}^0)^2 + (K_{n,2}^0)^2 - 2(K_{n,1}^0)^2 K_{n,2}^0}{1-(K_{n,1}^0)^2}$. This estimator can also be written as a quadratic form

$$\hat{\sigma}^2 = \frac{1}{n} X_n^T M^{-1} X_n,$$

where $X_n = (X(t_1^n), X(t_2^n), \dots, X(t_n^n))$ and

$$M^{-1} = \begin{pmatrix} \frac{1}{1-(K_{n,1}^0)^2} + \frac{(a_2^0)^2}{b^0} & \frac{a_1^0 a_2^0}{b^0} - \frac{K_{n,1}^0}{1-(K_{n,1}^0)^2} & -\frac{a_2^0}{b_0} & & & & \\ \frac{a_1^0 a_2^0}{b^0} - \frac{K_{n,1}^0}{1-(K_{n,1}^0)^2} & \frac{1}{1-(K_{n,1}^0)^2} + \frac{a_{12}^{02}}{b^0} & \frac{a_{12}^0}{b_0} & -\frac{a_2^0}{b_0} & & & \\ -\frac{a_2^0}{b_0} & \frac{a_{12}^0}{b_0} & \frac{1+a_{12}^{02}}{b_0} & \frac{a_{12}^0}{b_0} & \ddots & & \\ & -\frac{a_2^0}{b_0} & \frac{a_{12}^0}{b_0} & \ddots & \ddots & \ddots & \\ & & & \ddots & \frac{1+a_{12}^{02}}{b_0} & \frac{a_{12}^0}{b_0} & -\frac{a_2^0}{b_0} \\ & & & \ddots & \frac{a_{12}^0}{b_0} & \frac{1+(a_1^0)^2}{b_0} & -\frac{a_1^0}{b_0} \\ & & & & -\frac{a_2^0}{b_0} & -\frac{a_1^0}{b_0} & \frac{1}{b_0} \end{pmatrix}$$

is an n-dimensional pentadiagonal matrix, where $a_{12}^0 = a_1^0 a_2^0 - a_1^0$, $a_{12}^{02} = (a_1^0)^2 + (a_2^0)^2$.

Denote by $\sigma^2\Sigma$ the covariance matrix of X_n, then $\Sigma_{ij} = K_{n,|i-j|}$ and

$$\begin{aligned}
E\hat{\sigma}^2 &= \frac{\sigma^2}{n}\mathrm{Tr}(M^{-1}\Sigma) \\
&= \frac{\sigma^2}{n}\left(2 + (n-2)\left(\frac{(1+(K_{n,1}^0)^2 - 2K_{n,1}^0 K_{n,1})(1-K_{n,2}^0)}{(1+K_{n,2}^0 - 2(K_{n,1}^0)^2)(1-(K_{n,1}^0)^2)}\right.\right. \\
&\qquad \left.\left. + \frac{2(K_{n,2}^0 - (K_{n,1}^0)^2)(1-K_{n,2})}{(1+K_{n,2}^0 - 2(K_{n,1}^0)^2)(1-K_{n,2}^0)}\right)\right) \\
&:= \frac{\sigma^2}{n}\left(2 + (n-2)A_n\right). \qquad (4.11)
\end{aligned}$$

Consequently, $E\hat{\sigma}^2 = \sigma^2$ always holds when $\theta_0 = \theta$. Cases when $\theta_0 \neq \theta$ are discussed below.

After similar steps as did in Case 1, it follows from (4.3) that when $\nu \notin \mathbb{Z}$,

$$A_n = \begin{cases} \left(\frac{\theta}{\theta_0}\right)^{2\nu} + O(n^{2\nu-2}) + O(n^{-2\nu}), & \text{if } \nu < 1, \\ \left(\frac{\theta}{\theta_0}\right)^{2\nu} + O(n^{2-2\nu}) + O(n^{2\nu-4}), & \text{if } 1 < \nu < 2, \\ \left(\frac{\theta}{\theta_0}\right)^{4} + \frac{\beta^2}{6\tau-\beta^2}\left(\left(\frac{\theta}{\theta_0}\right)^2 - 1\right)^2 + O(n^{4-2\nu}) + O(n^{-2}), & \text{if } \nu > 2. \end{cases}$$

When $\nu = 1$, it follows from (4.5) and (4.9) that as $n \to \infty$,

$$\frac{1+(K_{n,1}^0)^2 - 2K_{n,1}^0 K_{n,1}}{1-(K_{n,1}^0)^2} = r^2 + \frac{r^2\log r}{\log(\theta_0/n)} + \frac{c_2 r^2 \log r}{c_1(\log(\theta_0/n))^2} + O((\log n)^{-3}),$$

$$\frac{1-K_{n,2}^0}{1+K_{n,2}^0 - 2(K_{n,1}^0)^2} = -\frac{\log(\theta_0/n)}{\log 2} - \frac{\log 2 - c_2/c_1}{\log 2} + O(n^{-2}(\log n)^3),$$

$$\frac{K_{n,2}^0 - (K_{n,1}^0)^2}{1+K_{n,2}^0 - 2(K_{n,1}^0)^2} = \frac{\log(\theta_0/n)}{2\log 2} + \frac{2\log 2 - c_2/c_1}{2\log 2} + O(n^{-2}(\log n)^3),$$

$$\frac{1-K_{n,2}}{1-K_{n,2}^0} = r^2 + \frac{r^2\log r}{\log(2\theta_0/n)} + \frac{c_2 r^2\log r}{c_1(\log(2\theta_0/n))^2} + O((\log n)^{-3}),$$

where $r = \theta/\theta_0$. Hence,

$$A_n = \left(\frac{\theta}{\theta_0}\right)^2 + O((\log n)^{-1}) + O((\log n)^{-2}).$$

Similarly, when $\nu = 2$, it follows from (4.5) that

$$\frac{x^\nu}{\Gamma(\nu)2^{\nu-1}}\mathcal{K}_\nu(x) = 1 + c_2' x^2 + c_3' x^4 \log(1/x) + c_4' x^4 + O(x^6 \log x) \qquad (4.12)$$

as $x \to 0$, where c_2', c_3', c_4' are constants only depending on ν. Thus, as $n \to \infty$,

$$\frac{1 + (K_{n,1}^0)^2 - 2K_{n,1}^0 K_{n,1}}{1 - (K_{n,1}^0)^2} = r^2 + (r^2 - r^4)\frac{c_3'}{c_2'}\left(\frac{\theta_0}{n}\right)^2 \log\left(\frac{\theta_0}{n}\right) + O(n^{-2}),$$

$$\begin{aligned}\frac{1 - K_{n,2}^0}{1 + K_{n,2}^0 - 2(K_{n,1}^0)^2} = &-\frac{4}{3} + \frac{c_2' n^2 (16c_4' - 16c_3' \log 2 - 2(c_2')^2 - 4c_4')}{36(c_3'\theta_0 \log(\theta_0/n))^2} \\ &+ \frac{c_2' n^2}{3c_3'\theta_0^2 \log(\theta_0/n)} + O(n^2(\log n)^{-3}),\end{aligned}$$

$$\begin{aligned}\frac{K_{n,2}^0 - (K_{n,1}^0)^2}{1 + K_{n,2}^0 - 2(K_{n,1}^0)^2} = &\frac{7}{6} - \frac{c_2' n^2 (16c_4' - 16c_3' \log 2 - 2(c_2')^2 - 4c_4')}{72(c_3'\theta_0 \log(\theta_0/n))^2} \\ &- \frac{c_2' n^2}{6c_3'\theta_0^2 \log(\theta_0/n)} + O(n^2(\log n)^{-3}),\end{aligned}$$

$$\frac{1 - K_{n,2}}{1 - K_{n,2}^0} = r^2 + 4(r^2 - r^4)\frac{c_3'}{c_2'}\left(\frac{\theta_0}{n}\right)^2 \log\left(\frac{2\theta_0}{n}\right) + O(n^{-2}),$$

$$A_n = \left(\frac{\theta}{\theta_0}\right)^4 + O((\log n)^{-1}) + O((\log n)^{-2}).$$

This together with (4.11) finishes the proof. □

Remark. Only $k = 1, 2$ are considered in Proposition 6 since the corresponding Vecchia approximation is computationally efficient. If θ is known, then taking $\theta_0 = \theta$ when construct $\hat{\sigma}^2$ will result in unbiased estimator for σ^2.

4.3 Simulation

Let $\sigma^2 = 1$ and $\theta = 5$ in (4.1). For each value of $n \in \{200, 250, \ldots, 1000\}$, generate 15000 independent realizations of X. In the following text, denote by $\sigma_{\nu,k}^2 = \lim_{n\to\infty} E\hat{\sigma}^2$, whose value is proved in Proposition 6.

Fix $\theta_0 = 1$ when solving for MLE of σ^2 using the Vecchia approximation (4.2). For $(\nu, k) \in \{(0.3, 1), (1.3, 1), (1.3, 2)\}$, the first row of plots in Figure 4.1 presents the boxplot of $\hat{\sigma}^2 - \sigma_{\nu,k}^2$ among 15000 realizations at each sample size n. The second row of plots in Figure 4.1 presents the empirical distribution of $\hat{\sigma}^2 - \sigma_{\nu,k}^2$ when $n = 1000$, where the red curve indicates the density

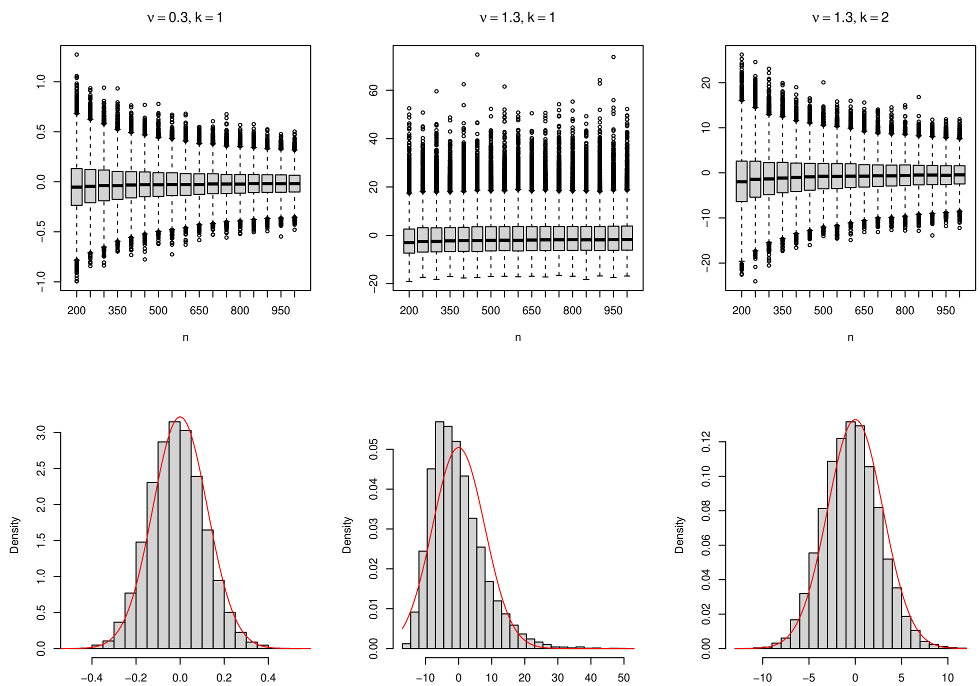

Figure 4.1Empirical distributions of bias with 15000 realizations. ($\sigma^2 = 1, \theta = 5, \theta_0 = 1.$)

function of normal distribution with zero mean and standard deviation equals the empirical standard deviation of $\hat{\sigma}^2 - \sigma^2_{\nu,k}$ among 15000 realizations. For the same three pairs of values of (ν, k), Figure 4.2 presents the average and standard deviation of absolute values of $\hat{\sigma}^2 - \sigma^2_{\nu,k}$ at each sample size n among 15000 realizations when $(\nu, k) = (0.3, 1)$ and $(\nu, k) = (1.3, 2)$. For the case when $(\nu, k) = (1.3, 1)$, 50000 realizations are generated since the estimator $\hat{\sigma}^2$ has a larger variance.

Fix $\theta_0 = \theta = 5$ when solving for MLE of σ^2 using the Vecchia approximation (4.2), then $\sigma^2_{\nu,k} = \sigma^2 = 1$. For the same dataset of realizations, plots in Figure 4.3 include the boxplot of $\hat{\sigma}^2 - \sigma^2$ among 15000 realizations at each sample size n, as well as the empirical distribution of $\hat{\sigma}^2 - \sigma^2$ when $n = 1000$, where the red curve indicates the density function of normal distribution with zero mean and standard deviation equals the empirical standard deviation of $\hat{\sigma}^2 - \sigma^2$ among 15000 realizations. Figure 4.4 presents the average and standard deviation of absolute values of $\hat{\sigma}^2 - \sigma^2$ at each sample size n among 15000 realizations when $(\nu, k) = (0.3, 1)$ and $(\nu, k) = (1.3, 2)$. For the case when $(\nu, k) = (1.3, 1)$, since the variance of $\hat{\sigma}^2$ is larger, 50000 realizations are generated.

The first row of plots in Figure 4.2 and Figure 4.4 illustrate Proposition 6. Furthermore, it is

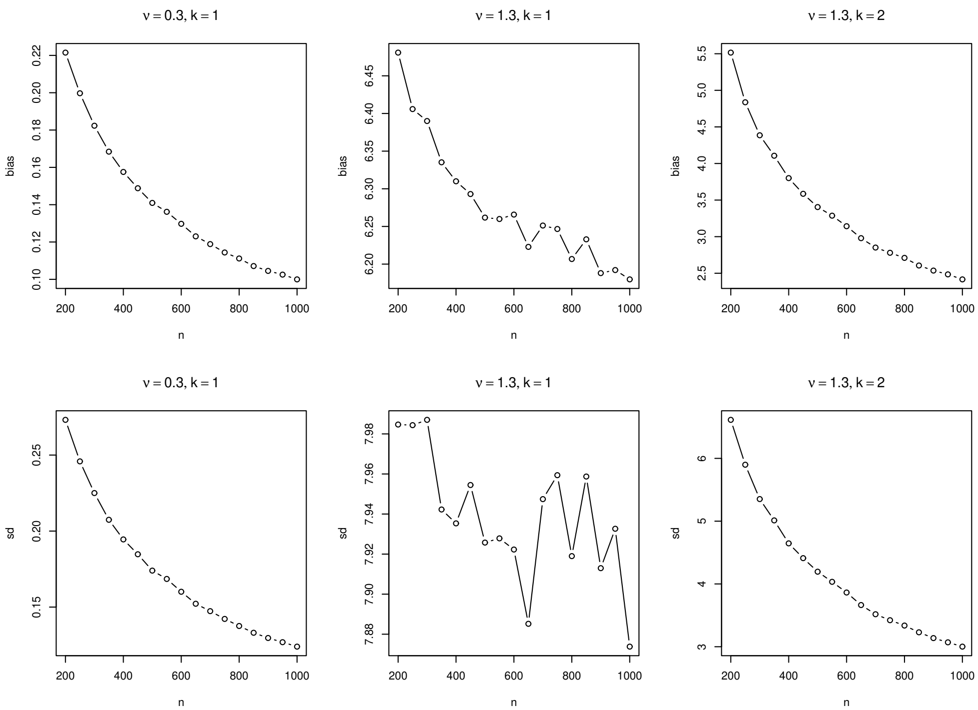

Figure 4.2The average and standard deviation for absolute value of bias when $n = 200, 250, \ldots, 1000$. ($\sigma^2 = 1, \theta = 5, \theta_0 = 1$.)

indicated by the simulation results that when $k < \nu$, the standard deviation of $\hat{\sigma}^2$ is not significantly reduced as the sample size increases, and the empirical distribution of $\hat{\sigma}^2 - \sigma^2_{\nu,k}$ appears to be right-skewed. When $k > \nu$, however, the standard deviation of $\hat{\sigma}^2$ decreases as the sample size increases, and the empirical distribution of $\hat{\sigma}^2 - \sigma^2_{\nu,k}$ when $n = 1000$ is close to normal distribution. As is observed from Figure 4.2, the standard deviation of $\hat{\sigma}^2$ when $(\nu, k) = (0.3, 1)$ is smaller compared with the case when $(\nu, k) = (1.3, 2)$. Let $\theta_0 = \theta$, then $(\nu, k) = (0.3, 1)$ and $(\nu, k) = (1.3, 2)$ result in similar values of the standard deviation of $\hat{\sigma}^2$, as is shown in Figure 4.4.

For future research, it is interesting to perform theoretical analysis for more asymptotic properties of $\hat{\sigma}^2$, including the convergence rate of its variance and its asymptotic distribution. The sampling design considered in this chapter is limited to a regular grid on the line, which is also the sampling design studied in Section III of Zhang et al. (2021). It is challenging but interesting to extend the existing results to irregular sampling designs on $\mathbb{R}$ ($d \geq 1$).

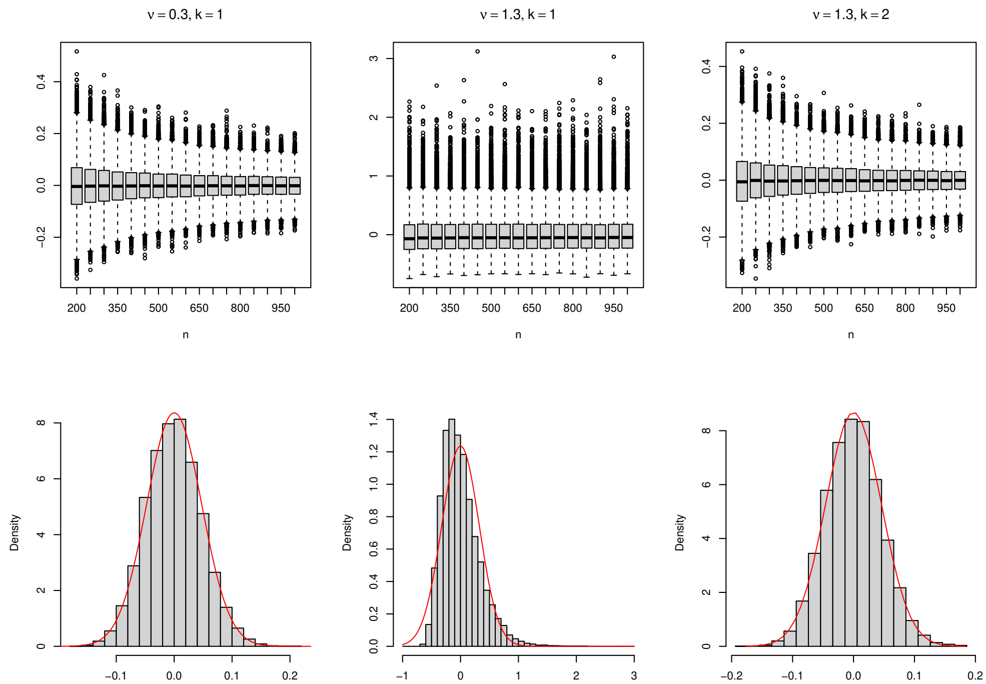

Figure 4.3Empirical distributions of bias with 15000 realizations. ($\sigma^2 = 1, \theta = 5, \theta_0 = 5$.)

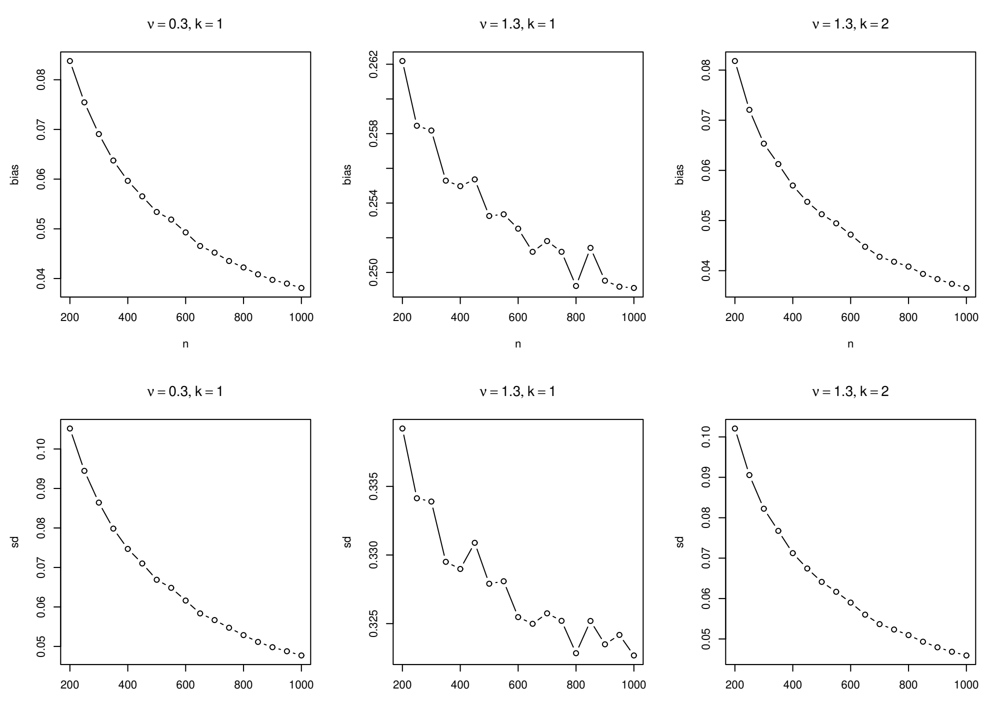

Figure 4.4The average and standard deviation for absolute value of bias when $n = 200, 250, \ldots, 1000$. ($\sigma^2 = 1, \theta = 5, \theta_0 = 5$.)

BIBLIOGRAPHY

[1] Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55. US Government printing office, 1948.

[2] Ethan B. Anderes and Michael L. Stein. Estimating deformations of isotropic Gaussian random fields on the plane. *The Annals of Statistics*, 36(2):719 – 741, 2008. doi: 10.1214/009053607000000893. URL https://doi.org/10.1214/009053607000000893.

[3] François Bachoc, Agnès Lagnoux, and Andrés F. López-Lopera. Maximum likelihood estimation for Gaussian processes under inequality constraints. *Electronic Journal of Statistics*, 13 (2):2921 – 2969, 2019. doi: 10.1214/19-EJS1587. URL https://doi.org/10.1214/19-EJS1587.

[4] Alessandro Baldi Antognini and Maroussa Zagoraiou. Exact optimal designs for computer experiments via kriging metamodelling. *Journal of Statistical Planning and Inference*, 140 (9):2607–2617, 2010. ISSN 0378-3758. doi: https://doi.org/10.1016/j.jspi.2010.03.027. URL https://www.sciencedirect.com/science/article/pii/S0378375810001400.

[5] Glen Earl Baxter. A strong limit theorem for gaussian processes. *Proceedings of the American Mathematical Society*, 7(3):522–527, 1956. URL https://doi.org/10.1090/S0002-9939-1956-0090920-6.

[6] Arnaud Begyn. Quadratic Variations along Irregular Subdivisions for Gaussian Processes. *Electronic Journal of Probability*, 10:691 – 717, 2005. doi: 10.1214/EJP.v10-245. URL https://doi.org/10.1214/EJP.v10-245.

[7] Moreno Bevilacqua, Tarik Faouzi, Reinhard Furrer, and Emilio Porcu. Estimation and prediction using generalized wendland covariance functions under fixed domain asymptotics. *The Annals of Statistics*, 47(2):pp. 828–856, 2019. ISSN 00905364, 21688966. URL https://www.jstor.org/stable/26581883.

[8] David Bolin and Jonas Wallin. Multivariate type g matérn stochastic partial differential equation random fields. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82(1):215–239, 2020.

[9] Jian Cao, Joseph Guinness, Marc G Genton, and Matthias Katzfuss. Scalable gaussian-process regression and variable selection using vecchia approximations. *Journal of Machine Learning Research*, 23(348):1–30, 2022.

[10] Ricardo Carrizo Vergara. *Development of geostatistical models using stochastic partial differential equations*. Theses, Université Paris sciences et lettres, December 2018. URL https://pastel.hal.science/tel-02188146.

[11] Grace Chan and Andrew T.A. Wood. Increment-based estimators of fractal dimension for two-dimensional surface data. *Statistica Sinica*, 10, 2000. ISSN 10170405.

[12] Grace Chan, Peter Hall, and D. S. Poskitt. Periodogram-Based Estimators of Fractal Properties. *The Annals of Statistics*, 23(5):1684 – 1711, 1995. doi: 10.1214/aos/1176324319. URL https://doi.org/10.1214/aos/1176324319.

[13] Chih-Hao Chang, Hsin-Cheng Huang, and Ching-Kang Ing. Mixed domain asymptotics for a stochastic process model with time trend and measurement error. *Bernoulli*, 23(1):159 – 190, 2017. doi: 10.3150/15-BEJ740. URL https://doi.org/10.3150/15-BEJ740.

[14] A. G. Constantine and Peter Hall. Characterizing Surface Smoothness Via Estimation of Effective Fractal Dimension. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56(1):97–113, 12 1994. ISSN 0035-9246. doi: 10.1111/j.2517-6161.1994.tb01963.x. URL https://doi.org/10.1111/j.2517-6161.1994.tb01963.x.

[15] Noel Cressie and Hsin-Cheng Huang. Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, 94(448):1330–1340, 1999. ISSN 01621459. URL http://www.jstor.org/stable/2669946.

[16] Noel A. C. Cressie. *Statistics for Spatial Data*. J. Wiley, 1993.

[17] Abhirup Datta, Sudipto Banerjee, Andrew O Finley, and Alan E Gelfand. Hierarchical nearest-neighbor gaussian process models for large geostatistical datasets. *Journal of the American Statistical Association*, 111(514):800–812, 2016.

[18] Sandra De Iaco, Donald Myers, and Donato Posa. Nonseparable space-time covariance models: Some parametric families. *Mathematical Geology*, 34:23–42, 01 2002. ISSN 1573-8868. doi: 10.1023/A:1014075310344.

[19] Juan Du, Hao Zhang, and V. S. Mandrekar. Fixed-domain asymptotic properties of tapered maximum likelihood estimators. *The Annals of Statistics*, 37(6A):3330–3361, 2009. ISSN 00905364, 21688966. URL http://www.jstor.org/stable/25662196.

[20] B. Dubuc, J. F. Quiniou, C. Roques-Carmes, C. Tricot, and S. W. Zucker. Evaluating the fractal dimension of profiles. *Phys. Rev. A*, 39:1500–1512, Feb 1989. doi: 10.1103/PhysRevA.39.1500. URL https://link.aps.org/doi/10.1103/PhysRevA.39.1500.

[21] Andrew O. Finley, Abhirup Datta, Bruce D. Cook, Douglas C. Morton, Hans E. Andersen, and Sudipto Banerjee. Efficient algorithms for bayesian nearest neighbor gaussian processes. *Journal of Computational and Graphical Statistics*, 28(2):401–414, 2019. doi: 10.1080/10618600.2018.1537924. URL https://doi.org/10.1080/10618600.2018.1537924. PMID: 31543693.

[22] Reinhard Furrer, Marc G Genton, and Douglas Nychka. Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 15(3):502–523, 2006. doi: 10.1198/106186006X132178. URL https://doi.org/10.1198/106186006X132178.

[23] Reinhard Furrer, Marc G. Genton, and Douglas Nychka. Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 15(3):502–523, 2006. ISSN 10618600. URL http://www.jstor.org/stable/27594195.

[24] R. Gay and C. C. Heyde. On a class of random field models which allows long range dependence. *Biometrika*, 77(2):401–403, 1990. ISSN 00063444. URL http://www.jstor.org/stable/2336820.

[25] Tilmann Gneiting. Nonseparable, stationary covariance functions for space-time data. *Journal of the American Statistical Association*, 97(458):590–600, 2002. ISSN 01621459. URL http://www.jstor.org/stable/3085674.

[26] Tilmann Gneiting, William Kleiber, and Martin Schlather. Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105(491):1167–1177, 2010. doi: 10.1198/jasa.2010.tm09420. URL https://doi.org/10.1198/jasa.2010.tm09420.

[27] Tilmann Gneiting, Hana Ševčíková, and Donald B. Percival. Estimators of Fractal Dimension: Assessing the Roughness of Time Series and Spatial Data. *Statistical Science*, 27(2):247 – 277, 2012. doi: 10.1214/11-STS370. URL https://doi.org/10.1214/11-STS370.

[28] Robert B Gramacy. *Surrogates: Gaussian process modeling, design, and optimization for the applied sciences*. Chapman and Hall/CRC, 2020.

[29] Ulf Grenander. *Abstract Inference*. Wiley, 1981.

[30] Joseph Guinness. Permutation and grouping methods for sharpening gaussian process approximations. *Technometrics*, 60(4):415–429, 2018.

[31] Peter Hall and Andrew Wood. On the performance of box-counting estimators of fractal dimension. *Biometrika*, 80(1):246–251, 03 1993. ISSN 0006-3444. doi: 10.1093/biomet/80.1.246. URL https://doi.org/10.1093/biomet/80.1.246.

[32] V. Heine. Models for two-dimensional stationary stochastic processes. *Biometrika*, 42(1/2):170–178, 1955. ISSN 00063444. URL http://www.jstor.org/stable/2333434.

[33] Xiangping Hu and Ingelin Steinsland. Spatial modeling with system of stochastic partial differential equations. *Wiley Interdisciplinary Reviews: Computational Statistics*, 8(2):112–125, 2016.

[34] Xiangping Hu, Daniel Simpson, Finn Lindgren, and Håvard Rue. Multivariate gaussian random fields using systems of stochastic partial differential equations. *arXiv preprint arXiv:1307.1379*, 2013.

[35] I.A. Ibragimov and Y.A. Rozanov. *Gaussian random processes*. Springer-Verlag, 1978.

[36] Jacques Istas and Gabriel Lang. Quadratic variations and estimation of the local hölder index of a gaussian process. *Annales de l'I.H.P. Probabilités et statistiques*, 33(4):407–436, 1997. URL http://www.numdam.org/item/AIHPB_1997__33_4_407_0/.

[37] Richard H. Jones and Yiming Zhang. Models for continuous stationary space-time processes. In Timothy G. Gregoire, David R. Brillinger, Peter J. Diggle, Estelle Russek-Cohen, William G. Warren, and Russell D. Wolfinger, editors, *Modelling Longitudinal and Spatially Correlated Data*, pages 289–298, New York, NY, 1997. Springer New York. ISBN 978-1-4612-0699-6.

[38] Matthias Katzfuss. A multi-resolution approximation for massive spatial datasets. *Journal of the American Statistical Association*, 112(517):201–214, 2017. doi: 10.1080/01621459.2015.1123632. URL https://doi.org/10.1080/01621459.2015.1123632.

[39] Matthias Katzfuss and Joseph Guinness. A general framework for vecchia approximations of gaussian processes. *Statistical Science*, 36(1):124–141, 2021.

[40] C. G. Kaufman and B. A. Shaby. The role of the range parameter for estimation and prediction in geostatistics. *Biometrika*, 100(2):473–484, 2013. ISSN 00063444. URL http://www.jstor.org/stable/43304571.

[41] Cari G. Kaufman, Mark J. Schervish, and Douglas W. Nychka. Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103(484):1545–1555, 2008. doi: 10.1198/016214508000000959. URL https://doi.org/10.1198/016214508000000959.

[42] Cari G. Kaufman, Mark J. Schervish, and Douglas W. Nychka. Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103(484):1545–1555, 2008. ISSN 01621459. URL http://www.jstor.org/stable/27640203.

[43] John T. Kent. Continuity Properties for Random Fields. *The Annals of Probability*, 17(4):1432 – 1440, 1989. doi: 10.1214/aop/1176991163. URL https://doi.org/10.1214/aop/1176991163.

[44] John T. Kent and Andrew T. A. Wood. Estimating the fractal dimension of a locally self-similar gaussian process by using increments. *Journal of the Royal Statistical Society. Series B (Methodological)*, 59(3):679–699, 1997. ISSN 00359246. URL http://www.jstor.org/stable/2346018.

[45] K Kubilius and D Melichov. Quadratic variations and estimation of the hurst index of the solution of sde driven by a fractional brownian motion. *Lithuanian Mathematical Journal*, 50 (4):401–417, Nov 2010.

[46] S. N. Lahiri. Central limit theorems for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhyā: The Indian Journal of Statistics (2003-2007)*, 65(2):

356–388, 2003. ISSN 09727671. URL http://www.jstor.org/stable/25053269.

[47] S. N. Lahiri and Kanchan Mukherjee. Asymptotic distributions of m-estimators in a spatial regression model under some fixed and stochastic spatial sampling designs. *Annals of the Institute of Statistical Mathematics*, 56(2):225–250, June 2004.

[48] Tao-Kai Lam and Wei-Liem Loh. Estimating structured correlation matrices in smooth Gaussian random field models. *The Annals of Statistics*, 28(3):880 – 904, 2000. doi: 10.1214/aos/1015952003. URL https://doi.org/10.1214/aos/1015952003.

[49] Nikolai Leonenko, Maria D. Ruiz-Medina, and Murad S. Taqqu. Fractional Elliptic, Hyperbolic and Parabolic Random Fields. *Electronic Journal of Probability*, 16(none):1134 – 1172, 2011. doi: 10.1214/EJP.v16-891. URL https://doi.org/10.1214/EJP.v16-891.

[50] Paul Lévy. Le mouvement brownien plan. *American Journal of Mathematics*, 62:487–550, 1940. URL https://mathscinet.ams.org/mathscinet-getitem?mr=2734.

[51] Finn Lindgren, Håvard Rue, and Johan Lindström. An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011. doi: https://doi.org/10.1111/j.1467-9868.2011.00777.x. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9868.2011.00777.x.

[52] Finn Lindgren, David Bolin, and Håvard Rue. The spde approach for gaussian and non-gaussian fields: 10 years and still running. *Spatial Statistics*, 50:100599, 2022. ISSN 2211-6753. doi: https://doi.org/10.1016/j.spasta.2022.100599. URL https://www.sciencedirect.com/science/article/pii/S2211675322000057. Special Issue: The Impact of Spatial Statistics.

[53] Wei-Liem Loh. Estimating the smoothness of a Gaussian random field from irregularly spaced data via higher-order quadratic variations. *The Annals of Statistics*, 43(6):2766 – 2794, 2015. doi: 10.1214/15-AOS1365. URL https://doi.org/10.1214/15-AOS1365.

[54] Wei-Liem Loh, Saifei Sun, and Jun Wen. On fixed-domain asymptotics, parameter estimation and isotropic Gaussian random fields with Matérn covariance functions. *The Annals of Statistics*, 49(6):3127 – 3152, 2021. doi: 10.1214/21-AOS2077. URL https://doi.org/10.1214/21-AOS2077.

[55] Wei-Liem Loh, Saifei Sun, and Jun Wen. Supplement to "On fixed-domain asymptotics, parameter estimation and isotropic Gaussian random fields with Matérn covariance functions", 2021.

[56] Chunsheng Ma. Spatio-temporal variograms and covariance models. *Advances in Applied Probability*, 37(3):706–725, 2005. ISSN 00018678. URL http://www.jstor.org/stable/30037351.

[57] Chunsheng Ma. Recent developments on the construction of spatio-temporal covariance models. *Stochastic Environmental Research and Risk Assessment*, 22:39–47, 03 2008. doi: 10.1007/s00477-007-0154-x.

[58] Bertil Matérn. *Spatial Variation*. Springer New York, NY, 1986.

[59] M.M. Meerschaert and A. Sikorskii. *Stochastic Models for Fractional Calculus*. De Gruyter Studies in Mathematics. De Gruyter, 2011. ISBN 9783110258165.

[60] Douglas Nychka, Soutir Bandyopadhyay, Dorit Hammerling, Finn Lindgren, and Stephan Sain. A multiresolution gaussian process model for the analysis of large spatial datasets. *Journal of Computational and Graphical Statistics*, 24(2):579–599, 2015. doi: 10.1080/10618600.2014.914946. URL https://doi.org/10.1080/10618600.2014.914946.

[61] Eulogio Pardo-Igúzquiza and Peter A Dowd. Amle3d: A computer program for the inference of spatial covariance parameters by approximate maximum likelihood estimation. *Computers & Geosciences*, 23(7):793–805, 1997.

[62] Vladimir I. Piterbarg. *Asymptotic Methods in the Theory of Gaussian Processes and Fields*. American Mathematical Society, 1995.

[63] Havard Rue and Leonhard Held. *Gaussian Markov random fields: theory and applications*. Chapman and Hall/CRC, 2005. URL https://doi.org/10.1201/9780203492024.

[64] Michael L. Stein, Zhiyi Chi, and Leah J. Welty. Approximating Likelihoods for Large Spatial Data Sets. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 66(2): 275–296, 04 2004. ISSN 1369-7412. doi: 10.1046/j.1369-7412.2003.05512.x. URL https://doi.org/10.1046/j.1369-7412.2003.05512.x.

[65] Michael Leonard Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, 1999.

[66] George E Uhlenbeck and Leonard S Ornstein. On the theory of the brownian motion. *Physical review*, 36(5):823, 1930.

[67] Balth. van der Pol and H. Bremmer. *Operational calculus based on the two-sided Laplace integral*. Cambridge University Press, 1950.

[68] Aad van der Vaart. Maximum likelihood estimation under a spatial sampling scheme. *The Annals of Statistics*, 24(5):2049 – 2057, 1996. doi: 10.1214/aos/1069362309. URL https://doi.org/10.1214/aos/1069362309.

[69] A. V. Vecchia. A general class of models for stationary two-dimensional random processes. *Biometrika*, 72(2):281–291, 1985. ISSN 00063444. URL http://www.jstor.org/stable/2336080.

[70] A. V. Vecchia. Estimation and model identification for continuous spatial processes. *J. Roy. Statist. Soc. Ser. B*, 50(2):297–312, 1988. ISSN 0035-9246. URL http://links.jstor.org/sici?sici=0035-9246(1988)50:2<297:EAMIFC>2.0.CO;2-D&origin=MSN.

[71] Theodore von Kármán. Progress in the statistical theory of turbulence. *Proceedings of the National Academy of Sciences of the United States of America*, 34(11):530–539, 1948. doi: https://doi.org/10.1073/pnas.34.11.530.

[72] Daqing Wang. Fixed domain asymptotics and consistent estimation for gaussian random field models in spatial statistics and computer experiments, 2010. URL https://scholarbank.nus.edu.sg/handle/10635/19060.

[73] Daqing Wang and Wei-Liem Loh. On fixed-domain asymptotics and covariance tapering in Gaussian random field models. *Electronic Journal of Statistics*, 5(none):238 – 269, 2011. doi: 10.1214/11-EJS607. URL https://doi.org/10.1214/11-EJS607.

[74] P. Whittle. On stationary processes in the plane. *Biometrika*, 41(3/4):434–449, 1954. ISSN 00063444. URL http://www.jstor.org/stable/2332724.

[75] P. Whittle. Stochastic processes in several dimensions. *Bulletin of the International Statistical Institute*, 40:974–994, 1963.

[76] Zhiliang Ying. Asymptotic properties of a maximum likelihood estimator with data from a gaussian process. *Journal of Multivariate Analysis*, 36(2):280 – 296, 1991. doi: 10.1016/0047-259X(91)90062-7.

[77] Zhiliang Ying. Maximum likelihood estimation of parameters under a spatial sampling scheme. *The Annals of Statistics*, 21, 9 1993. ISSN 0090-5364. doi: 10.1214/aos/1176349272.

[78] Hao Zhang. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99(465):250–261, 2004. doi: 10.1198/016214504000000241. URL https://doi.org/10.1198/016214504000000241.

[79] Lu Zhang, Wenpin Tang, and Sudipto Banerjee. Fixed-domain asymptotics under vecchia's approximation of spatial process likelihoods. *arXiv preprint arXiv:2101.08861*, 2021. URL https://doi.org/10.48550/arXiv.2101.08861.

[80] Yuzhen Zhou and Yimin Xiao. Joint asymptotics for estimating the fractal indices of bivariate gaussian processes. *Journal of Multivariate Analysis*, 165:56–72, 2018. ISSN 0047-259X. doi: https://doi.org/10.1016/j.jmva.2017.12.001. URL https://www.sciencedirect.com/science/article/pii/S0047259X17307509.

[81] Zhengyuan Zhu and Michael L. Stein. Parameter estimation for fractional brownian surfaces. *Statistica Sinica*, 12(3):863–883, 2002. ISSN 10170405, 19968507. URL http://www.jstor.org/stable/24306999.

[82] Zhengyuan Zhu and Hao Zhang. Spatial sampling design under the infill asymptotic framework. *Environmetrics*, 17(4):323–337, 2006. doi: https://doi.org/10.1002/env.772. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/env.772.

APPENDIX A

QUADRATIC VARIATIONS FROM IRREGULAR SAMPLING

A.1 $d = 1$

(6) studied quadratic variations defined using irregular observations of process $(X_t)_{t\in[0,1]}$ with Gaussian increments. Suppose (X_t) is observed at

$$0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_{N_n}^{(n)} = 1, \quad n \in \mathbb{N}$$

and denote by $\Delta t_k^{(n)} = t_{k+1}^{(n)} - t_k^{(n)}$, $k = 0, \ldots, N_n - 1$. Write $\Delta t_k^{(n)}$ as Δt_k for brevity. Let

$$\Delta X_k = \Delta t_{k-1} X_{t_{k+1}} + \Delta t_k X_{t_{k-1}} - (\Delta t_{k-1} + \Delta t_k) X_{t_k}. \tag{A.1}$$

It is straightforward that

$$t_{k+1}^q \Delta t_{k-1} + t_{k-1}^q \Delta t_k - t_k^q (\Delta t_{k-1} + \Delta t_k) = 0, \quad q = 0, 1;$$

$$t_{k+1}^2 \Delta t_{k-1} + t_{k-1}^2 \Delta t_k - t_k^2 (\Delta t_{k-1} + \Delta t_k) \neq 0.$$

The second order quadratic variation is then defined as

$$\mathcal{V}_n(X) = 2 \sum_{k=1}^{N_n - 1} \frac{\Delta t_k (\Delta X_k)^2}{(\Delta t_{k-1})^{\frac{3-\gamma}{2}} (\Delta t_k)^{\frac{3-\gamma}{2}} (\Delta t_{k-1} + \Delta t_k)}, \tag{A.2}$$

where $\gamma > 0$ is related to the smoothness of (X_t). For example, if (X_t) is a fractional Brownian motion with Hurst's index H, then $\gamma = 2 - 2H$.

Denote by $m_n = \max\{\Delta t_k^{(n)}; 0 \leq k \leq N_n - 1\}$ and $p_n = \min\{\Delta t_k^{(n)}; 0 \leq k \leq N_n - 1\}$. It is assumed in (6) that

(i) For a sequence of positive real numbers $(l_k)_{k\geq 1}$,

$$\lim_{n\to\infty} \sup_{1\leq k\leq N_n - 1} \left| \frac{\Delta t_{k-1}^{(n)}}{\Delta t_k^{(n)}} - l_k \right| = 0;$$

(ii) $m_n = O(p_n)$ as $n \to \infty$;

(iii) $p_n = o(\frac{1}{\log n})$ as $n \to \infty$.

With irregular observations satisfying the assumptions above, the almost sure convergence of $\mathcal{V}_n(X)$ is proved under some regularity conditions on (X_t).

Although (6) considered a general class of irregular observations, the quadratic variation defined in (A.2) could not be evaluated when γ is unknown. Also, γ could not be estimated when $m_n = p_n = \frac{1}{N_n}$ does not hold. The quadratic variations defined by (53), however, do not depend on unknown parameters.

(53) considered a stationary, isotropic Gaussian random field X on $\mathbb{R}^d$, $d = 1, 2$. When $d = 1$, define irregular lattice points

$$t_i = \varphi\left(\frac{i-1}{n-1}\right), \quad i = 1, \ldots, n \tag{A.3}$$

for $n \geq 2$, where $\varphi : \mathbb{R} \mapsto \mathbb{R}$ is a twice continuously differentiable function with $\varphi(0) = 0$, $\varphi(1) = 1$ and $\min_{0 \leq s \leq 1} \varphi'(s) > 0$.

For $\theta \in \{1, 2\}$ and $\ell \in \{1, 2, \ldots, \lfloor (n-1)/\theta \rfloor\}$, define

$$a_{\theta,\ell;i,k} = \frac{\ell!}{\prod_{0 \leq j \leq \ell, j \neq k} (t_{i+\theta k} - t_{i+\theta j})}, \quad k = 0, \ldots, \ell, \tag{A.4}$$

$$\nabla_{\theta,\ell} X_i = \sum_{k=0}^{\ell} a_{\theta,\ell;i,k} X(t_{i+\theta k}), \quad i = 1, \ldots, n - \theta\ell. \tag{A.5}$$

Lemma 1 in (53) shows that

$$\sum_{k=0}^{\ell} a_{\theta,\ell;i,k} t_{i+\theta k}^q = \begin{cases} 0, & q = 0, \ldots, \ell - 1 \\ \ell!, & q = \ell. \end{cases}$$

The ℓth order quadratic variations are defined as

$$V_{\theta,\ell} = \sum_{i=1}^{n-\theta\ell} \left(\nabla_{\theta,\ell} X_i\right)^2, \quad \theta \in \{1, 2\}, \ell \in \{1, 2, \ldots, \lfloor (n-1)/\theta \rfloor\}. \tag{A.6}$$

A.2 $d > 1$

A.2.1 Observations along a curve

(53) studied the case when $d = 2$ and X is observed along a fixed curve in $\mathbb{R}^2$. Assume that

(i) $\exists \epsilon > 0$, $L > 0$ s.t. $\gamma : (-\epsilon, L + \epsilon) \mapsto \mathbb{R}^d$ is a C^2-curve parameterized by arc length;

(ii) $\exists C > 0$ s.t. $||\gamma(t^*) - \gamma(t)|| \geq C|t^* - t|, \forall t^*, t \in [0, L]$.

Denote by $X_i = X(\gamma(t_i))$ and $d_{i,j} = ||\gamma(t_i) - \gamma(t_j)||$ for $1 \leq i, j \leq n$, where t_i is defined in (A.3). For $\theta, \ell \in \{1, 2\}$, define

$$b_{\theta,\ell;i,k} = \frac{\ell}{\prod_{0\leq j\leq \ell, j\neq k}(d_{i,i+\theta k} - d_{i,i+\theta j})}, \quad k = 0, \ldots, \ell, \tag{A.7}$$

$$\tilde{\nabla}_{\theta,\ell} X_i = \sum_{k=0}^{\ell} b_{\theta,\ell;i,k} X_{i+\theta k}, \quad i = 1, \ldots, n - \theta\ell. \tag{A.8}$$

Lemma 1 in (53) shows that

$$\sum_{k=0}^{\ell} b_{\theta,\ell;i,k} d_{i,i+\theta k}^{q} = \begin{cases} 0, & q = 0, \ldots, \ell - 1 \\ \ell, & q = \ell. \end{cases}$$

The ℓth order quadratic variations are constructed as

$$\tilde{V}_{\theta,\ell} = \sum_{i=1}^{n-\theta\ell} \left(\tilde{\nabla}_{\theta,\ell} X_i\right)^2, \quad \theta, \ell \in \{1, 2\}. \tag{A.9}$$

A.2.2 Observations on deformed lattice

When $d = 2$ and X is observed on deformed lattice points in $\mathbb{R}^2$, (53) also defined corresponding second order quadratic variations.

Consider an open set Ω in $\mathbb{R}^2$ with $[0, 1]^2 \subset \Omega$, and a $C^2(\Omega)$ diffeomorphism $\tilde{\varphi} : \Omega \mapsto \mathbb{R}^2$. Let $\tilde{\varphi} = (\varphi_1, \varphi_2)$. Write $X_{i_1,i_2} = X(\mathbf{x}^{i_1,i_2})$, where $\mathbf{x}^{i_1,i_2} = (x_1^{i_1,i_2}, x_2^{i_1,i_2})' = (\varphi_1(i_1/n, i_2/n), \varphi_2(i_1/n, i_2/n))'$ for $1 \leq i_1, i_2 \leq n$.

For $\theta \in \{1, 2\}$ and $1 \leq i_1, i_2 \leq n - \theta$, let

$$A_{\theta;i_1,i_2} = \begin{pmatrix} x_1^{i_1+\theta,i_2} - x_1^{i_1,i_2} & x_2^{i_1+\theta,i_2} - x_2^{i_1,i_2} \\ x_1^{i_1,i_2+\theta} - x_1^{i_1,i_2} & x_2^{i_1,i_2+\theta} - x_2^{i_1,i_2} \end{pmatrix},$$

$$B_{\theta;i_1,i_2} = \begin{pmatrix} x_1^{i_1+\theta,i_2} - x_1^{i_1+\theta,i_2+\theta} & x_2^{i_1+\theta,i_2} - x_2^{i_1+\theta,i_2+\theta} \\ x_1^{i_1,i_2+\theta} - x_1^{i_1+\theta,i_2+\theta} & x_2^{i_1,i_2+\theta} - x_2^{i_1+\theta,i_2+\theta} \end{pmatrix}.$$

Then define

$$\begin{pmatrix} \tilde{\nabla}_{\theta,1} X_{i_1,i_2} \\ \tilde{\nabla}_{\theta,2} X_{i_1,i_2} \end{pmatrix} = B_{\theta;i_1,i_2}^{-1} \begin{pmatrix} X_{i_1+\theta,i_2} - X_{i_1+\theta,i_2+\theta} \\ X_{i_1,i_2+\theta} - X_{i_1+\theta,i_2+\theta} \end{pmatrix} - A_{\theta;i_1,i_2}^{-1} \begin{pmatrix} X_{i_1+\theta,i_2} - X_{i_1,i_2} \\ X_{i_1,i_2+\theta} - X_{i_1,i_2} \end{pmatrix} \tag{A.10}$$

$$= \sum_{0 \le k_1,k_2 \le 1} \begin{pmatrix} c_{\theta,1;i_1,i_2}^{k_1,k_2} X_{i_1+\theta k_1, i_2+\theta k_2} \\ c_{\theta,2;i_1,i_2}^{k_1,k_2} X_{i_1+\theta k_1, i_2+\theta k_2} \end{pmatrix}, \tag{A.11}$$

where $B_{\theta;i_1,i_2}^{-1}$ and $A_{\theta;i_1,i_2}^{-1}$ exist for large enough n since $\tilde{\varphi}$ is a diffeomorphism. Lemma 2 in (53) shows that for $j, \ell \in \{1, 2\}$,

$$\sum_{0 \le k_1,k_2 \le 1} c_{\theta,\ell;i_1,i_2}^{k_1,k_2} \left(x_j^{i_1+\theta k_1, i_2+\theta k_2} \right)^q = 0, \quad q = 0, 1.$$

The second order quadratic variations are defined as

$$\tilde{V}_{\theta,\ell} = \sum_{1 \le i_1,i_2 \le n-\theta} \left(\tilde{\nabla}_{\theta,\ell} X_{i_1,i_2} \right)^2, \quad \theta, \ell \in \{1, 2\}. \tag{A.12}$$

For quadratic variations defined in (A.6), (A.9) and (A.12), the rates of their expectations and variances as $n \longrightarrow \infty$ are proved by (53) under some regularity conditions on X.

(54) focused on the stationary GRF X on $\mathbb{R}^d$ with isotropic Matérn covariance function, and studied quadratic variations constructed from irregular observations of X when $d > 2$.

The definition in (A.12) is extended to the case where X is observed on $[0, 1]^d$ and $d \in \mathbb{Z}^+$. Consider an open set Ω in $\mathbb{R}^d$ with $[0, 1]^d \subset \Omega$, and a $C^2(\Omega)$ diffeomorphism $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_d) : \Omega \mapsto \mathbb{R}^d$. Write

$$\mathbf{x}(\mathbf{i}) = (x_1(\mathbf{i}), \dots, x_d(\mathbf{i}))' = \left(\varphi_1 \left(\frac{\mathbf{i}}{n} \right), \dots, \varphi_d \left(\frac{\mathbf{i}}{n} \right) \right)'$$

and $X_{i_1,\dots,i_d} = X(\mathbf{x}(\mathbf{i}))$, where $\mathbf{i} = (i_1, \dots, i_d)'$ and $1 \le i_1, \dots, i_d \le n$. The sample size is thus n^d.

For $\theta \in \{1, 2\}$ and $\ell \in \mathbb{Z}^+$, let

$$\bar{\ell} = \sum_{l=1}^{\ell} \begin{pmatrix} l + d - 1 \\ d - 1 \end{pmatrix}, \tag{A.13}$$

$$\mathbf{x}_{\mathbf{i},j} = (x_{\mathbf{i},j;1}, \dots, x_{\mathbf{i},j;d})' = \mathbf{x}(i_1 + k_1\theta, \dots, i_d + k_d\theta), \quad j = 0, \dots, \bar{\ell},$$

$$\tilde{\mathbf{y}}_{\mathbf{i},j} = \frac{n}{\theta}(\mathbf{x}_{\mathbf{i},j} - \mathbf{x}_{\mathbf{i},0}), \quad j = 1, \dots, \bar{\ell},$$

where $i_1, \ldots, i_d \in \{1, \ldots, n-\ell\theta\}$, $k_1, \ldots, k_d \in \{0, 1, \ldots, \ell\}$ and $\sum_{i=1}^{d} k_i \in \{0, 1, \ldots, \ell\}$, j denotes the lexicographical order of combinations $(k_1, \ldots, k_d)$, $\mathbf{x}_{\mathbf{i},0} = \mathbf{x}(\mathbf{i})$. The detailed rule of ordering is described in Section 5.1 of (54).

For $l = 1, \ldots, \ell$ and $\mathbf{s} = (s_1, \ldots, s_d)' \in \mathbb{R}^d$, define

$$\mathbf{a}^{\langle d,l\rangle}(\mathbf{s}) = \left(\prod_{k=1}^{d} \frac{s_k^{l_k}}{l_k!}\right) \in \mathbb{R}^{\binom{l+d-1}{d-1}}, \tag{A.14}$$

where $l_1, \ldots, l_d \in \{0, 1, \ldots, \ell\}$ and $\sum_{i=1}^{d} l_i = l$. The elements of $\mathbf{a}^{\langle d,l\rangle}(\mathbf{s})$ are arranged in lexicographic ordering with respect to $(l_1, \ldots, l_d)$. Define a $\bar{\ell} \times \bar{\ell}$ matrix

$$\tilde{A}_{\mathbf{i},\theta,d,\ell} = \begin{pmatrix} \mathbf{a}^{\langle d,1\rangle}(\tilde{\mathbf{y}}_{\mathbf{i},1}) & \mathbf{a}^{\langle d,2\rangle}(\tilde{\mathbf{y}}_{\mathbf{i},1}) & \cdots & \mathbf{a}^{\langle d,\ell\rangle}(\tilde{\mathbf{y}}_{\mathbf{i},1}) \\ \mathbf{a}^{\langle d,1\rangle}(\tilde{\mathbf{y}}_{\mathbf{i},2}) & \mathbf{a}^{\langle d,2\rangle}(\tilde{\mathbf{y}}_{\mathbf{i},2}) & \cdots & \mathbf{a}^{\langle d,\ell\rangle}(\tilde{\mathbf{y}}_{\mathbf{i},\bar{\ell}}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}^{\langle d,1\rangle}(\tilde{\mathbf{y}}_{\mathbf{i},\bar{\ell}}) & \mathbf{a}^{\langle d,2\rangle}(\tilde{\mathbf{y}}_{\mathbf{i},1}) & \cdots & \mathbf{a}^{\langle d,\ell\rangle}(\tilde{\mathbf{y}}_{\mathbf{i},\bar{\ell}}) \end{pmatrix} \tag{A.15}$$

and assume $|\tilde{A}_{\mathbf{i},\theta,d,\ell}| \neq 0$ for all $i_1, \ldots, i_d \in \{1, \ldots, n-\ell\theta\}$.

Denote by $\tilde{A}^{-1}_{\mathbf{i},\theta,d,\ell} = \left(\tilde{\alpha}^{j,k}_{\mathbf{i},\theta,d,\ell}\right)_{1\le j,k\le\bar{\ell}}$ and let

$$\tilde{c}_{\mathbf{i},\theta,d,\ell}(j) = \begin{cases} \tilde{\alpha}^{\bar{\ell},j}_{\mathbf{i},\theta,d,\ell}, & \forall j = 1, \ldots, \bar{\ell}, \\ -\sum_{k=1}^{\bar{\ell}} \tilde{\alpha}^{\bar{\ell},k}_{\mathbf{i},\theta,d,\ell}, & \text{if } j = 0. \end{cases} \tag{A.16}$$

For $\theta \in \{1, 2\}$ and $\ell \in \mathbb{Z}^+$, define

$$\tilde{\nabla}_{\theta,d,\ell} X_{i_1,\ldots,i_d} = \sum_{j=0}^{\bar{\ell}} \tilde{c}_{\mathbf{i},\theta,d,\ell}(j) X(\mathbf{x}_{\mathbf{i},j}), \quad i_1, \ldots, i_d \in \{1, \ldots, n-2\ell\}. \tag{A.17}$$

The ℓth order quadratic variation is then defined as

$$\tilde{V}_{\theta,d,\ell} = \sum_{1\le i_1,\ldots,i_d\le n-2\ell} \left(\tilde{\nabla}_{\theta,d,\ell} X_{i_1,\ldots,i_d}\right)^2. \tag{A.18}$$

A.2.3 Stratified sampling

Let

$$\mathbf{x}(\mathbf{i}) = (x_1(\mathbf{i}), \ldots, x_d(\mathbf{i}))' = \left(\frac{i_1 - 1 + \delta_{\mathbf{i};1}}{n}, \ldots, \frac{i_d - 1 + \delta_{\mathbf{i};d}}{n}\right)' \in [0,1)^d,$$

where $\mathbf{i} = (i_1, \ldots, i_d)'$ and $1 \le i_1, \ldots, i_d \le n$; $0 \le \delta_{\mathbf{i};k} < 1$ ($k = 1, \ldots, d$) are constants that can vary with n. Let ω_n be an integer depending only on n such that $\omega_n = O(n^{\gamma_0})$ as $n \to \infty$, where $\gamma_0 \in (0, 1)$ is a constant.

For $\theta \in \{1, 2\}$ and $\ell \in \mathbb{Z}^+$, let

$$\mathbf{x}_{\mathbf{i},j} = (x_{\mathbf{i},j;1}, \ldots, x_{\mathbf{i},j;d})' = \mathbf{x}(i_1 + k_1\omega_n\theta, \ldots, i_d + k_d\omega_n\theta), \quad j = 0, \ldots, \bar{\ell},$$

$$\mathbf{y}_{\mathbf{i},j} = \frac{n}{\omega_n\theta}(\mathbf{x}_{\mathbf{i},j} - \mathbf{x}_{\mathbf{i},0}), \quad j = 1, \ldots, \bar{\ell},$$

where $i_1, \ldots, i_d \in \{1, \ldots, n - \ell\omega_n\theta\}$, other notations are as defined in Section A.2.2. Define a $\bar{\ell} \times \bar{\ell}$ matrix

$$A_{\mathbf{i},\theta,d,\ell} = \begin{pmatrix} \mathbf{a}^{\langle d,1\rangle}(\mathbf{y}_{\mathbf{i},1}) & \mathbf{a}^{\langle d,2\rangle}(\mathbf{y}_{\mathbf{i},1}) & \cdots & \mathbf{a}^{\langle d,\ell\rangle}(\mathbf{y}_{\mathbf{i},1}) \\ \mathbf{a}^{\langle d,1\rangle}(\mathbf{y}_{\mathbf{i},2}) & \mathbf{a}^{\langle d,2\rangle}(\mathbf{y}_{\mathbf{i},2}) & \cdots & \mathbf{a}^{\langle d,\ell\rangle}(\mathbf{y}_{\mathbf{i},\bar{\ell}}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}^{\langle d,1\rangle}(\mathbf{y}_{\mathbf{i},\bar{\ell}}) & \mathbf{a}^{\langle d,2\rangle}(\mathbf{y}_{\mathbf{i},1}) & \cdots & \mathbf{a}^{\langle d,\ell\rangle}(\mathbf{y}_{\mathbf{i},\bar{\ell}}) \end{pmatrix}, \tag{A.19}$$

where $\mathbf{a}^{\langle d,l\rangle}(\cdot)$ is defined in (A.14). Assume $|A_{\mathbf{i},\theta,d,\ell}| \ne 0$ for all $i_1, \ldots, i_d \in \{1, \ldots, n - \ell\omega_n\theta\}$. Then denote by $A^{-1}_{\mathbf{i},\theta,d,\ell} = \left(\alpha^{j,k}_{\mathbf{i},\theta,d,\ell}\right)_{1 \le j,k \le \bar{\ell}}$. Let

$$c_{\mathbf{i},\theta,d,\ell}(j) = \begin{cases} \alpha^{\bar{\ell},j}_{\mathbf{i},\theta,d,\ell}, & \forall j = 1, \ldots, \bar{\ell}, \\ -\sum_{k=1}^{\bar{\ell}} \alpha^{\bar{\ell},k}_{\mathbf{i},\theta,d,\ell}, & \text{if } j = 0. \end{cases} \tag{A.20}$$

The ℓth order quadratic variation is then defined as

$$V_{\theta,d,\ell} = \sum_{1 \le i_1, \ldots, i_d \le n - 2\ell\omega_n} \left(\nabla_{\theta,d,\ell} X_{i_1,\ldots,i_d}\right)^2, \tag{A.21}$$

where $\theta \in \{1, 2\}$, $\ell \in \mathbb{Z}^+$ and

$$\nabla_{\theta,d,\ell} X_{i_1,\ldots,i_d} = \sum_{j=0}^{\bar{\ell}} c_{\mathbf{i},\theta,d,\ell}(j) X(\mathbf{x}_{\mathbf{i},j}), \quad i_1, \ldots, i_d \in \{1, \ldots, n - 2\ell\omega_n\}. \tag{A.22}$$

A.3 Randomized Sampling Design

Section 4 in (54) considered random sampling on $[0, 1)^d$, where $d \in \{1, 2, 3\}$. It is an extension of the stratified sampling discussed in Section A.2.3.

Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be a sequence of i.i.d. random vectors in $\mathbb{R}^d$ that are independent of the GRF X. Assume the probability density function $p(\mathbf{x})$ of $\mathbf{x}_1$ satisfies

$$\int_{[0,1)^d} p(\mathbf{x})\mathrm{d}\mathbf{x} = 1 \quad \text{and} \quad \inf_{[0,1)^d} p(\mathbf{x}) \geq p_0 > 0. \tag{A.23}$$

When p_0 in (A.23) is unknown, let

$$n_\tau = \left\lfloor \left(\frac{N}{\tau \log^2(N)} \right)^{1/d} \right\rfloor, \quad \forall \tau > 0.$$

Let $\hat{\tau}$ be the smallest real number greater than or equal to 1 such that

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \cap \prod_{j=1}^{d} \left[\frac{i_j - 1}{n_{\hat{\tau}}}, \frac{i_j}{n_{\hat{\tau}}} \right) \neq \varnothing, \quad \forall i_1, \ldots, i_d \in \{1, \ldots, n_{\hat{\tau}}\}.$$

Consider the effective sample only:

$$\left\{ \{\mathbf{x}_j, X(\mathbf{x}_j)\} : \mathbf{x}_j \in \prod_{j=1}^{d} \left[\frac{i_j - 1}{n_{\hat{\tau}}}, \frac{i_j}{n_{\hat{\tau}}} \right), i_1, \ldots, i_d \in \{1, \ldots, n_{\hat{\tau}}\}, j \in \{1, \ldots, N\} \right\}. \tag{A.24}$$

Take a subset of $\mathbf{x}_j$'s in (A.24) such that for each $\mathbf{i} = (i_1, \ldots, i_d)'$ with $1 \leq i_1, \ldots, i_d \leq n_{\hat{\tau}}$, there is strictly one j satisfying $\mathbf{x}_j \in \prod_{j=1}^{d} \left[\frac{i_j-1}{n_{\hat{\tau}}}, \frac{i_j}{n_{\hat{\tau}}} \right)$. Write the selected $\mathbf{x}_j$ as $\mathbf{x}(\mathbf{i})$. The randomized sampling design is then reduced to the stratified sampling design with a sample size of $n_{\hat{\tau}}^d$. Thus, the ℓth order quadratic variations could be defined as in (A.21), where $\theta \in \{1, 2\}$, $\ell \in \mathbb{Z}^+$ and n is replaced by $n_{\hat{\tau}}$.

When p_0 in (A.23) is known, let $\tau_0 = 3/p_0$ and

$$\bar{n}_\tau = \left\lfloor \left(\frac{N}{\tau \log(N)} \right)^{1/d} \right\rfloor,$$

where $\tau \geq \tau_0$. Let $\bar{\tau}$ be the smallest real number greater than or equal to τ_0 such that

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \cap \prod_{j=1}^{d} \left[\frac{i_j - 1}{\bar{n}_{\bar{\tau}}}, \frac{i_j}{\bar{n}_{\bar{\tau}}} \right) \neq \varnothing, \quad \forall i_1, \ldots, i_d \in \{1, \ldots, \bar{n}_{\bar{\tau}}\}.$$

The effective sample is defined as in (A.24) by replacing $n_{\hat{\tau}}$ with $\bar{n}_{\bar{\tau}}$. Similarly, the ℓth order quadratic variations are defined as in (A.21), where $\theta \in \{1, 2\}$, $\ell \in \mathbb{Z}^+$ and n is replaced by $\bar{n}_{\bar{\tau}}$.

A.4 Estimating Smoothness Parameters

Based on the a.s. convergence of the quadratic variation defined in (A.2), when a fractional Ornstein-Uhlenbeck process O^H is observed from regular sampling, its fractional parameter $H \in (0, 1)$ has a strongly consistent estimator as

$$\hat{H}_n = \frac{1}{2} - \frac{\log\left(\sum_{k=1}^{N_n-1}\left(O^H_{\frac{k+1}{N_n}} + O^H_{\frac{k-1}{N_n}} - 2O^H_{\frac{k}{N_n}}\right)^2\right)}{2\log N_n}, \tag{A.25}$$

where $1/N_n = o(1/\log n)$.

Quadratic variations constructed in (A.6), (A.9) and (A.12) are used to estimate the smoothness parameter ν in covariance function (1.1).

The estimators of ν defined by (53) are minimizers of functions that depend on sampling locations and quadratic variations. Although with no closed form expressions, the estimators are proved to be strongly consistent when $\ell > \nu$ and observations are on $[0, 1]$ or along a curve. When X is observed on deformed lattice and $\nu \in (0, 2)$, $\ell \in \{1, 2\}$, the estimator defined using (A.12) is proved to be strongly consistent as well.

The Matérn covariance function belongs to the class of functions defined in (1.1). To estimate its smoothness parameter ν, define

$$\hat{\nu}_{n,\ell} = \frac{\log(V_{2,d,\ell}/V_{1,d,\ell})}{2\log 2}, \tag{A.26}$$

where $V_{\theta,d,\ell}$, $\theta = 1, 2$ are quadratic variations defined in (A.18), (A.21) and Section A.3, corresponding to different kinds of sampling design. When $\ell > \nu$, it is proved by (54) that $\hat{\nu}_{n,\ell} \to \nu$ a.s. as $n \to \infty$.

APPENDIX B

HIGH EXCURSION PROBABILITY

We first introduce some notations and definitions presented in (62).

The *structural modulus* of vector $\mathbf{t} \in \mathbb{R}^n$ is defined as

$$|\mathbf{t}|_{E,\alpha} = \sum_{i=1}^{k} \left(\sum_{j=E(i-1)+1}^{E(i)} t_j^2 \right)^{\alpha_i/2},$$

where $E = \{e_1, e_2, \ldots, e_k\}$, $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_k\}$, $e_i, \alpha_i \in \mathbb{Z}^+$ $(i = 1, 2, \ldots, k)$, $\sum_{i=1}^{k} e_i = n$, $E(i) = \sum_{j=0}^{i} e_j$, $e_0 = 0$. A *structure* (E, α) defines a partition of the space $\mathbb{R}^n$ into a direct product of orthogonal subspaces ($\mathbb{R}^n = \times_{i=1}^{k} \mathbb{R}^{e_i}$) such that the restrictions of the structural modulus $|\mathbf{t}|_{E,\alpha}$ on either of them is a Euclidean norm taken to the degree α_i, $i = 1, 2, \ldots, k$, respectively.

Example 1. *Let $n = k = 2$ and $E = \{1, 1\}$, then $E(0) = 0$, $E(1) = 1$, $E(2) = 2$, and*

$$|\mathbf{t}|_{E,\alpha} = |t_1|^{\alpha_1} + |t_2|^{\alpha_2}, \quad \forall \mathbf{t} = (t_1, t_2) \in \mathbb{R}^2,$$

where $\alpha_1, \alpha_2 \in \mathbb{Z}^+$.

Let $\chi(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^n$ be a Gaussian field with continuous trajectories, and

$$E\chi(\mathbf{t}) = -|\mathbf{t}|_{E,\alpha},$$

$$\mathrm{Cov}\,(\chi(\mathbf{t}), \chi(\mathbf{s})) = |\mathbf{t}|_{E,\alpha} + |\mathbf{s}|_{E,\alpha} - |\mathbf{t} - \mathbf{s}|_{E,\alpha},$$

where $\alpha_i \leq 2$ makes the covariance function valid. For any compact set $T \subset \mathbb{R}^n$ and matrix $M \in \mathbb{R}^{n \times n}$, denote by

$$H^{M}_{(E,\alpha),(E',\alpha')}(T) = E \exp\left(\max_{T} \left\{ \chi(\mathbf{t}) - |M\mathbf{t}|_{E',\alpha'} \right\} \right).$$

Write $H_{E,\alpha}(T) = H^{\mathbf{0}}_{(E,\alpha),(E',\alpha')}(T)$, where $\mathbf{0}$ is the zero matrix.

A set $A \subset \mathbb{R}^n$ is called *Jordan measurable* if its interior and closure have the same Lebesgue measure, i.e. its boundary has Lebesgure measure zero. The system $\{A_u, u > 0\}$ is said to *blow up slowly with the rate* $\kappa > 0$ if each of these sets contains a unit cube and $\mathrm{mes}(A_u) = O(e^{\kappa u^2/2})$ as $u \to \infty$.

Theorem 7.2 in (62) is presented as below, where the subscript $\cdot_{E,\alpha}$ is written as $\cdot_\alpha$ for short.

Theorem 11. *(62) Let* $\{X(\mathbf{t}), \mathbf{t} \in \mathbb{R}^n\}$ *be a Gaussian homogeneous field with zero mean and the covariance function* $r(\mathbf{t})$ *satisfies that there exists a non-degenerate matrix C and a structure* (E, α) *such that*

$$\begin{aligned} &r(C\mathbf{t}) = 1 - |\mathbf{t}|_\alpha + o(|\mathbf{t}|_\alpha) \quad \text{as } t \to 0, \\ &r(\mathbf{t}) \to 0 \quad \text{as } t \to \infty. \end{aligned} \tag{B.1}$$

Then there exists a number $\kappa > 0$ *such that for any system of closed Jordan sets, blowing up slowly with the rate* κ*,*

$$P\left(\max_{\mathbf{t}\in A_u} X(\mathbf{t}) > u\right) = H_\alpha \mathrm{mes}(A_u)|\det C^{-1}| \prod_{i=1}^{k} u^{2e_i/\alpha_i} \Psi(u)(1 + o(1)) \quad \text{as } u \to \infty, \tag{B.2}$$

where

$$H_\alpha = \lim_{t\to\infty} \frac{H_\alpha([0,t]^n)}{t^n}$$

and $\Psi(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty \exp(-x^2/2)\mathrm{d}x$.

Remark 3. *The zero-mean stationary Ornstein-Uhlenbeck field X with covariance function defined in (3.2) taking* $\sigma^2 = 1$ *satisfies conditions in Theorem 11 with* $n = 2$*,* $E = \{1, 1\}$*,* $\alpha = \{1, 1\}$*, and* $C = \begin{pmatrix} 1/\lambda & 0 \\ 0 & 1/\mu \end{pmatrix}$.

APPENDIX C

STOCHASTIC PARTIAL DIFFERENTIAL EQUATION

Write the two-sided Laplace transform of a function h as

$$\mathcal{L}_h(p) = \int_{-\infty}^{\infty} e^{-px} h(x)\mathrm{d}x, \tag{C.1}$$

and denote by D^n the differential operator of order n, i.e. $D^n h(x) = \frac{\mathrm{d}^n}{\mathrm{d}x^n} h(x)$. It follows from the differentiation rule presented on Page 48-50 of (67) that

$$\mathcal{L}_{D^n h}(p) = p^n \mathcal{L}_h(p), \quad \forall n \in \mathbb{Z}^+ \tag{C.2}$$

when

$$\lim_{x\to\infty} e^{-px} h(x) - \lim_{x\to-\infty} e^{-px} h(x) = 0.$$

The case when $n \notin \mathbb{Z}^+$ is discussed in (59). We first introduce the definition of fractional derivatives below. For any $\alpha > 0$, define the fractional difference operator Δ^α as

$$\Delta^\alpha f(x) = \sum_{j=0}^{\infty} \frac{\Gamma(\alpha+1)}{j!\Gamma(\alpha-j+1)} (-1)^j f(x-jh)$$

and write the fractional derivative in the Grünwald-Letnikov finite difference form as

$$D^\alpha f(x) := \frac{\mathrm{d}^\alpha f(x)}{\mathrm{d}x^\alpha} = \lim_{h\to 0} \frac{\Delta^\alpha f(x)}{h^\alpha}. \tag{C.3}$$

Alternative integral forms for the fractional derivative are also presented in (59), as shown in Tables C.1-C.2. Consider the Riemann-Liouville fractional derivative of order $0 < \alpha < 1$, of which the Laplace transform is written as

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{-px} D^\alpha f(x)\mathrm{d}x &= \int_{-\infty}^{\infty} e^{-px} \frac{\mathrm{d}}{\mathrm{d}x} \int_0^\infty f(x-y) \frac{y^{-\alpha}}{\Gamma(1-\alpha)} \mathrm{d}y\mathrm{d}x \\
&= \frac{1}{\Gamma(1-\alpha)} \left(\left[e^{-px} \int_0^\infty f(x-y) y^{-\alpha} \mathrm{d}y \right]_{x=-\infty}^{\infty} \right. \\
&\quad \left. - \int_{-\infty}^{\infty} \int_0^\infty f(x-y) y^{-\alpha} \mathrm{d}y \mathrm{d}e^{-px} \right) \\
&:= \frac{1}{\Gamma(1-\alpha)} (I_1 - I_2),
\end{aligned}$$

where

$$I_2 = -p \int_0^\infty e^{-py} y^{-\alpha} \int_{-\infty}^\infty f(z) e^{-pz} \mathrm{d}z \mathrm{d}y$$
$$= -p^\alpha \mathcal{L}_f(p)$$

when $e^{-px} y^{-\alpha} f(x-y)$ is integrable. If it further holds that

$$\lim_{x\to\infty} e^{-px} \int_0^\infty f(x-y) y^{-\alpha} \mathrm{d}y - \lim_{x\to-\infty} e^{-px} \int_0^\infty f(x-y) y^{-\alpha} \mathrm{d}y = 0,$$

then

$$\mathcal{L}_{D^\alpha f}(p) = p^\alpha \mathcal{L}_f(p).$$

Generator form	$\int_0^\infty (f(x) - f(x-y)) \frac{\alpha y^{-\alpha-1}}{\Gamma(1-\alpha)} \mathrm{d}y$
Caputo form	$\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}x} f(x-y) \frac{y^{-\alpha}}{\Gamma(1-\alpha)} \mathrm{d}y$
Riemann-Liouville form	$\frac{\mathrm{d}}{\mathrm{d}x} \int_0^\infty f(x-y) \frac{y^{-\alpha}}{\Gamma(1-\alpha)} \mathrm{d}y$

Table C.1Alternative integral forms for the fractional derivative when $0 < \alpha < 1$.

Generator form	$\int_0^\infty (f(x-y) - f(x) + y \frac{\mathrm{d}}{\mathrm{d}x} f(x)) \frac{\alpha(\alpha-1) y^{-\alpha-1}}{\Gamma(2-\alpha)} \mathrm{d}y$
Caputo form	$\int_0^\infty \frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x-y) \frac{y^{1-\alpha}}{\Gamma(2-\alpha)} \mathrm{d}y$
Riemann-Liouville form	$\frac{\mathrm{d}^2}{\mathrm{d}x^2} \int_0^\infty f(x-y) \frac{y^{1-\alpha}}{\Gamma(2-\alpha)} \mathrm{d}y$

Table C.2Alternative integral forms for the fractional derivative when $1 < \alpha < 2$.

Consider the stochastic partial differential equation (SPDE)

$$L\left(\frac{\partial}{\partial t_1}, \frac{\partial}{\partial t_2}\right) X(t_1, t_2) = \epsilon(t_1, t_2), \quad t_1, t_2 \in \mathbb{R}, \tag{C.4}$$

where L is a linear differential operator. The Green's function of L satisfies

$$L\left(\frac{\partial}{\partial t_1}, \frac{\partial}{\partial t_2}\right) G(t_1, t_2) = \delta_0(t_1)\delta_0(t_2), \quad t_1, t_2 \in \mathbb{R}, \tag{C.5}$$

where δ_0 is the Dirac measure at 0.

When ϵ is the Gaussian white noise, it holds that

$$E[\epsilon(s_1, s_2)\epsilon(s_1 + t_1, s_2 + t_2)] = \delta_0(t_1)\delta_0(t_2), \quad \forall s_1, s_2, t_1, t_2 \in \mathbb{R}. \tag{C.6}$$

The covariance function of X is thus

$$\begin{aligned}
C(t_1, t_2) &:= E[X(s_1, s_2)X(s_1 + t_1, s_2 + t_2)], \quad \forall s_1, s_2, t_1, t_2 \in \mathbb{R} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(s_1, s_2)G(s_1 + t_1, s_2 + t_2)\mathrm{d}s_1\mathrm{d}s_2, \quad \forall t_1, t_2 \in \mathbb{R}.
\end{aligned} \tag{C.7}$$

As presented in (32), when the operator L takes the form of

$$L\left(\frac{\partial}{\partial t_1}, \frac{\partial}{\partial t_2}\right) = c_1\frac{\partial^2}{\partial t_1^2} + c_2\frac{\partial^2}{\partial t_2^2} + c_3\frac{\partial^2}{\partial t_1 \partial t_2} + c_4\frac{\partial}{\partial t_1} + c_5\frac{\partial}{\partial t_2} + c_6, \tag{C.8}$$

the Laplace transforms of the Green's function and the covariance function derived from (C.4) satisfy

$$\mathcal{L}_G(p, q) = \frac{1}{L(p, q)}, \tag{C.9}$$

$$\mathcal{L}_C(p, q) = \frac{1}{L(p, q)L(-p, -q)}. \tag{C.10}$$

As a special case of (C.8), the elliptic form of the operator L is discussed in (74), where the corresponding SPDE is

$$\left(\frac{\partial^2}{\partial t_1^2} + \frac{\partial^2}{\partial t_2^2} - \gamma^2\right)X(t_1, t_2) = \epsilon(t_1, t_2). \tag{C.11}$$

Denote by K_ℓ the modified Bessel functions of the second kind. The Green's function for (C.11) is thus

$$G(t_1, t_2) = \mathcal{L}^{-1}\frac{1}{p^2 + q^2 - \gamma^2} = \frac{1}{2\pi}K_0\left(\gamma\sqrt{t_1^2 + t_2^2}\right).$$

The spectral density function of X as the Fourier transform of the covariance function C is derived as

$$\begin{aligned}
f_X(\xi, \eta) &= \frac{1}{(2\pi)^2}L_C(i\xi, i\eta) \\
&= \frac{1}{(2\pi)^2\left(-\xi^2 - \eta^2 - \gamma^2\right)^2} \\
&\propto \frac{1}{\left(\xi^2 + \eta^2 + \gamma^2\right)^2}.
\end{aligned}$$

(24) considered the SPDE

$$(\nabla^2 - \beta^2)^\nu X(t_1, t_2) = \epsilon(t_1, t_2), \tag{C.12}$$

where $\nabla^2 = \partial^2/\partial t_1^2 + \partial^2/\partial t_2^2$, ϵ is a white noise field, $\beta \in \mathbb{R}$, $\nu > 0$, and

$$(\nabla^2 - \beta^2)^\nu = (-1)^\nu \sum_{j=0}^{\infty} \binom{\nu}{j} (-\nabla^2)^j \beta^{2(\nu-j)}. \tag{C.13}$$

The Green's function of $(\nabla^2 - \beta^2)^\nu$ satisfies

$$(-1)^\nu \sum_{j=0}^{\infty} \binom{\nu}{j} (-\nabla^2)^j \beta^{2(\nu-j)} G(t_1, t_2) = \delta_0(t_1)\delta_0(t_2). \tag{C.14}$$

Taking Laplace transform on both sides of equation (C.14) yields

$$(-1)^\nu \sum_{j=0}^{\infty} \binom{\nu}{j} (-p^2 - q^2)^j \beta^{2(\nu-j)} \mathcal{L}_G(p, q) = 1.$$

Thus,

$$\begin{aligned} \mathcal{L}_G(p, q) &= \left(\sum_{j=0}^{\infty} \binom{\nu}{j} (p^2 + q^2)^j (-\beta^2)^{\nu-j} \right)^{-1} \\ &= \frac{1}{\left(p^2 + q^2 - \beta^2\right)^\nu}. \end{aligned}$$

The spectral density function of X is

$$\begin{aligned} f_X(\xi, \eta) &= \frac{1}{(2\pi)^2} \left((-1)^\nu \sum_{j=0}^{\infty} \binom{\nu}{j} \left(-(i\xi)^2 - (i\eta)^2 \right)^j \beta^{2(\nu-j)} \right)^{-2} \\ &\propto \frac{1}{\left(\xi^2 + \eta^2 + \beta^2\right)^{2\nu}}, \end{aligned}$$

which is also presented in (75).