TAIL ESTIMATION OF THE SPECTRAL DENSITY UNDER FIXED-DOMAIN ASYMPTOTICS

By

Wei-Ying Wu

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Statistics

2011

ABSTRACT

TAIL ESTIMATION OF THE SPECTRAL DENSITY UNDER FIXED-DOMAIN ASYMPTOTICS

By

Wei-Ying Wu

For spatial statistics, two asymptotic approaches are usually considered: increasing domain asymptotics and fixed-domain asymptotics (or infill asymptotics). For increasing domain asymptotics, sampled data increase with the increasing spatial domain, while under infill asymptotics, data are observed on a fixed region with the distance between neighboring observations tending to zero. The consistency and asymptotic results under these two asymptotic frameworks can be quite different. For example, not all parameters are consistently estimated under infill asymptotics while consistency holds for those parameters under increasing asymptotics (Zhang 2004).

For a stationary Gaussian random field on $\mathbb{R}^d$ with the spectral density $f(\boldsymbol{\lambda})$ that satisfies $f(\boldsymbol{\lambda}) \sim c\,|\boldsymbol{\lambda}|^{-\theta}$ as $|\boldsymbol{\lambda}| \to \infty$, the parameters c and θ control the tail behavior of the spectral density where θ is related to the smoothness of random field and c can be used to determine the orthogonality of probability measures for a fixed θ. Specially, c corresponds to the microergodic parameter mentioned in Du et al. (2009) when Matérn covariance is assumed. Additionally, under infill asymptotics, the tail behavior of the spectral density dominates the performance of the prediction, and the equivalence of the probability measures. Based on those reasons, it is significant in statistics to estimate c and θ.

When the explicit form of f is known, its corresponding covariance structure can be computed through the Fourier transformation. Therefore, spatial domain methodologies like Maximum Likelihood Estimator (MLE) or Tapering MLE can be used for the estimation

of c and θ. Unfortunately, the exact form of f should be unknown in practice. Under this situation, spatial domain methods will not be applied without the covariance information. In my work, for data observed on grid points, two methods which utilize tail frequency information are proposed to estimate c and θ. One of them can be viewed as a weighted local Whittle type estimator. Under proposed approaches, the explicit form of f and the restriction of the dimension are not necessary. The asymptotic properties of the proposed estimators under infill asymptotics (or fixed-domain asymptotics) are investigated in this dissertation together with simulation studies.

ACKNOWLEDGMENT

I would like to give my deepest thank to my supervisors, Professor Yimin Xiao and Chae Young Lim for their invaluable support and encouragement in my PhD studying at Michigan State University. They are always patiently sharing and explaining their knowledge for me. Without their help and guidance, it would have been impossible for me to finish this work. Also, I thank Professor Chae Young Lim for providing me financial support from her MSU grant (MSU 08-IRGP-1532) for my studies in the year 2009-2010.

I also appreciate the help from my other committee members, Professor Mark M. Meerschaert and Professor Zhengfang Zhou. Thank you for your detailed suggestion and time on my dissertation.

My next thank to my best friends and classmates - Wei-Wen Hsu, Gengxin Li, Hsiu-Ching Chang and Sumit Sinha for their help and suggestions in the past years.

I am indebted to Professor Lijian Yang, Ms. Suzanne Watson and Mr. Eric Segur for their tremendous help and constant support.

I want to thank all the people who have helped and inspired me during my doctoral studying.

Most importantly, to give my thanks to my wife Wei-Ling, my parents and other families, whose patient love enabled me to complete this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

With recent advances in technology, we are facing enormous amount of data sets. When those data sets are observed on a regular grid, spectral analysis is popularly used due to fast computation using the Fast Fourier Transform. For example, parameters of the spectral density of a stationary lattice process can be estimated using a Whittle likelihood [Whittle, (1954)], which is more efficient in terms of computation compared to the maximum likelihood method on a spatial domain.

In my dissertation, I propose new methodologies developed from the perspective of spectral analysis to estimate parameters that control the tail behavior of the spectral density for a stationary Gaussian random field under fixed-domain asymptotics, which is one of two famous sampling schemes in spatial statistics. The second sampling scheme is the increasing domain asymptotics. Before explaining my research problem, I first introduce these two sampling schemes and their differences.

1.1 Increasing domain and fixed-domain asymptotics

Spatial data on a grid often can be regarded as a realization of a random field on a lattice. That is, for a random field, $Z(\boldsymbol{s})$ on $\mathbb{R}^d$, data is observed at $\phi \boldsymbol{J}$ for $\boldsymbol{J} \in \prod_{j=1}^{d}\{1, \cdots, m_j\}$, where ϕ is a grid length. When ϕ is fixed and the sample size is increasing (increasing domain asymptotics), asymptotic properties of parameter estimates on a spectral domain have been studied by many authors [see, e.g., Whittle (1954), Guyon (1982, 1995), Boissy et al. (2005) and Guo et al. (2009)]. For example, Guyon (1982) studied asymptotic properties of estimators using a Whittle likelihood or its variants when a parametric model is assumed for the spectral density of a stationary process on a lattice. Guo et al. (2009) studied asymptotic properties of estimators of long-range dependence parameters for anisotropic spatial linear processes using a local Whittle likelihood method in which a parametric form near zero frequency is only assumed. This is an extension of Robinson's research (1995) on time series.

For spatial data, it is often natural to assume that the data is observed on a bounded domain of interest, therefore more observations on the bounded domain means that the distance between observations, ϕ, decreases as the number of observations increases. This sampling scheme requires a different asymptotic framework, called fixed-domain asymptotics [Stein (1999)] (or infill asymptotics [Cressie (1993)]).

It has been shown that the asymptotic results under fixed-domain asymptotics can be different from the results under increasing-domain asymptotics [see, e.g., Mardia and Marshall (1984), Ying (1991, 1993), and Zhang (2004)]. For example, Zhang (2004) showed not all parameters in the Matérn covariance model of a stationary Gaussian random field on $\mathbb{R}^d$ are consistently estimable when d is smaller than or equal to 3. He also showed that a reparameterized quantity which is a function of variance and scale parameters can be estimated consistently by the maximum likelihood method. On the other hand, under increasing-domain asymptotics, the maximum likelihood estimators (MLEs) of variance and scale parameters for a stationary Gaussian process are consistent and asymptotically normal [Mardia and Marshall (1984)]. Although not all parameters can be estimated consistently under fixed-domain asymptotics, a microergodic parameter can be estimated consistently [see, e.g., Ying (1991, 1993), Zhang (2004), Zhang and Zimmerman (2005), Du et al. (2009), and Anderes (2010)]. The microergodicity of functions of parameters determines the equivalence of probability measures, whereby a microergodic parameter is the quantity that affects asymptotic mean squared prediction error under fixed-domain asymptotics. [Stein (1990, 1999)].

Although there have been more asymptotic results available recently under fixed-domain asymptotics, it is still very few in contrast with vast literature on increasing-domain asymptotics. Also, most results are for specific models of covariance functions. For example, Ying (1991, 1993) and Chen et al. (2000) studied asymptotic properties of estimators for a microergodic parameter in the exponential covariance function, while Zhang (2004), Loh (2005), Kaufman et al. (2008), Du et al. (2009) and Anderes (2010) investigated asymptotic properties of estimators for the Matérn covariance function. For the estimation of the fractal dimension in the spatial domain under the fixed-domain asymptotics, Constantine and Hall (1994) estimated effective fractal dimension using variogram for a non-Gaussian stationary process on $\mathbb{R}$. Chan and Wood (2004) introduced an increment-based estimator of the fractal dimension of a function of a stationary Gaussian random field on $\mathbb{R}^d$ when $d = 1$ or 2. These asymptotic results are established in the spatial domain.

Asymptotic results in the spectral domain are even less under fixed-domain asymptotics.

Stein (1995) studied asymptotic properties of a spatial periodogram of a filtered version of a stationary Gaussian random field. Lim and Stein (2008) extended results of Stein (1995) and showed asymptotic normality of a smoothed spatial cross-periodogram under fixed-domain asymptotics. Regarding the parameter estimation in the spectral domain under fixed-domain asymptotics, Chan et al. (1995) proposed a periodogram-based estimator of the fractal dimension of a stationary Gaussian random field when $d = 1$.

In the above discussions, it follows that the properties under increasing domain and fixed domain are quite different and more research works are required for fixed-domain asymptotics. In the next Section, I will begin to introduce my research problem under fixed-domain asymptotics.

1.2 The tail behavior of the spectral density

In this dissertation, I propose estimators of parameters that control the tail behavior of the spectral density for a stationary Gaussian random field when the data is observed on a grid within a bounded domain and study their asymptotic properties under fixed-domain asymptotics. Let $f(\boldsymbol{\lambda})$ be the spectral density of a stationary Gaussian random field, $Z(\boldsymbol{s})$ on $\mathbb{R}^d$ and we assume that

$$f(\boldsymbol{\lambda}) \;\sim\; c\,|\boldsymbol{\lambda}|^{-\theta} \quad \text{as} \;\; |\boldsymbol{\lambda}| \;\to\; \infty, \boldsymbol{\lambda} \in R^d \tag{1.1}$$

where $|\cdot|$ is a usual Euclidean norm and $\theta > d$ to ensure integrability of f. That is, we only assume power law for the tail behavior of the spectral density and do not assume any specific parametric form of the spectral density. In the following subsection, the reasons for interest in the tail behavior will be introduced from two perspectives; the equivalence of probability measures and the prediction.

1.2.1 Equivalence of probability measures

The equivalence between two probability measures P_1 and P_2 on a measurable space $\{\Omega, \mathcal{F})$ is that $P_1(A) = 0$ for any $A \in \mathcal{F}$ implies $P_2(A) = 0$ and denoted by $P_1 \equiv P_2$. We usually assume $\mathcal{F}$ is generated by the paths of the process $\{Z(\boldsymbol{s}), \boldsymbol{s} \in D\}$. When the stationarity is considered for the process, many criteria based on the spectral densities have been developed to classify the equivalence of probability measures [see, e.g.,Ibragimov (1978), Yadrenko (1983) and Du (2009a)].

Theorem 1. *(Yadrenko (1983)) Let $P_i, i = 1, 2$ be two probability measures such that under P_i, the process $\{Z(\boldsymbol{s}), \boldsymbol{s} \in R^d\}$ is stationary Gaussian with mean 0 and a second-order spectral density $f_i(\boldsymbol{\lambda}), \boldsymbol{\lambda} \in R^d$. If, for some $\theta > d$, $f_1(\boldsymbol{\lambda})\, |\boldsymbol{\lambda}|^\theta$ is bounded away from 0 and ∞ as $|\boldsymbol{\lambda}| \to \infty$, and for some finite c,*

$$\int_{|\boldsymbol{\lambda}|>c} \left\{ \frac{f_2(\boldsymbol{\lambda}) - f_1(\boldsymbol{\lambda})}{f_1(\boldsymbol{\lambda})} \right\}^2 d\boldsymbol{\lambda} < \infty. \tag{1.2}$$

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

The integrability of (1.2) is determined by the tail of spectral densities. For example, if $f_i(\boldsymbol{\lambda})$'s are isotopic, i.e., depend only on $|\boldsymbol{\lambda}|$, (1.2) will hold when there exists some $\epsilon > 0$ such that

$$\frac{f_1(\boldsymbol{\lambda})}{f_2(\boldsymbol{\lambda})} - 1 = O(|\boldsymbol{\lambda}|^{-(d/2+\epsilon)}) \ \text{ as} |\boldsymbol{\lambda}| \to \infty. \tag{1.3}$$

This implies the equivalence of probability measures can be verified by the decay degree of their spectral densities.

Many applications of the equivalence of measures have been explored to reduce the computational burden like a tapering method. Let $l_n(\theta)$ be the log likelihood of data observed:

$$l_n(\theta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log[\det V_n] - \frac{1}{2}X_n^{\prime}V_n^{-1}X_n. \tag{1.4}$$

where n is sample size, X_n is a data vector and V_n is the covariance matrix. The computation cost to obtain Maximum Likelihood Estimator (MLE) can be expensive.

To reduce computational burden, a tapering method on the covariance function can be used:

$$\tilde{V}(l,\theta) = V(l,\theta) \circ V_{tap}(l).$$

where $V(l,\theta)$ is the covariance function of the underlying process that depends on parameter θ (possibly a set of parameters), $V_{tap}(l)$ is the taper, a known positive function, that is 0 after a threshold distance and "$\circ$" is Schur or Hadamard product. By replacing $V(l,\theta)$ with $\tilde{V}(l,\theta)$, tapered likelihood is attained as

$$l_{n,tap}(\theta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log[\det \tilde{V_n}(l,\theta)] - \frac{1}{2}X_n^{\prime}\tilde{V_n}(l,\theta)_n^{-1}X_n. \tag{1.5}$$

The consistency of the estimator based on $l_{n,tap}(\theta)$ holds if the probability measure under $\tilde{V}(l,\theta)$ is equivalent to the one under $V(l,\theta)$ [see. Zhang (2004)]. More theoretical discussion about a tapered method is found in the Chapter 3 [Du, (2009a)].

1.2.2 Prediction under fixed-domain asymptotics

The another motivation to study tail behavior of the spectral density comes from its role in prediction. In spatial statistics, the best linear unbiased prediction is called kriging. Let process $Z(\boldsymbol{s})$ be a mean zero stationary process and data is sampled at locations $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3...., \}$ which are dense in a bounded region $D \subseteq R^d$, which implies that the infill sampling is used. Further, assume $\boldsymbol{s}^*$ be a new location that we would like to explore. Let $\hat{Z}(\boldsymbol{s}^*, n)$ be the best linear unbiased prediction of $Z(\boldsymbol{s}^*)$ based on the data $Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2), ..., Z(\boldsymbol{s}_n)$ and $e(\boldsymbol{s}^*, n)$ be the error between $Z(\boldsymbol{s}^*)$ and $\hat{Z}(\boldsymbol{s}^*, n)$. The following theorem [Stein (1998), p. 136] compares the prediction performance between a correct measure P_1 and a misspecified measure P_2.

Theorem 2. *(Stein 1999, p.252) Let $Z(\boldsymbol{s})$ be a mean zero stationary Gaussian random field under probability measure P_i with spectral density f_i, for $i = 1, 2$. If there exist some $\rho > 0$ such that $f_1(\lambda)|\lambda|^\rho$ is bounded away from 0 and ∞, and $\frac{f_2(\lambda)}{f_1(\lambda)} \to 1$ as $|\lambda| \to \infty$,*

$$\begin{aligned} &\lim_{n\to\infty} \frac{E_1(e_2(\boldsymbol{s}^*, n) - e_1(\boldsymbol{s}^*, n))^2}{E_1(e_1(\boldsymbol{s}^*, n))^2} = 0 \\ &\lim_{n\to\infty} \frac{E_2(e_2(\boldsymbol{s}^*, n))^2}{E_1(e_2(\boldsymbol{s}^*, n))^2} = 0 \end{aligned} \tag{1.6}$$

where $E_i(\cdot)$ and $e_i(\cdot)$ is the expectation and prediction error under probability measure P_i, for $i = 1, 2$.

The above result means no matter which probability measures we used for prediction performance is asymptotically equivalent under the fixed-domain sampling if the tail behavior of f_2 is as that of f_1. Thus, understanding the tail behavior of the spectral density is of great importance in spatial statistics.

In my dissertation, we introduce two approaches to estimate parameters that control the tail behavior of the spectral density. That is c and θ in (1.1). One of the proposed estimators is obtained by minimizing an objective function that can be viewed as a weighted Whittle likelihood, in which Fourier frequencies near a pre-specified non-zero frequency are considered. This approach is similar to the local Whittle likelihood method introduced by Robinson (1995) for estimating a long-range dependence parameter in time series analysis. For a stationary lattice process, Robinson (1995) proposed to estimate a long-range dependence parameter by minimizing the Whittle likelihood over Fourier frequencies near zero since the long-range parameter dependence is controlled by the behavior of the spectral density near zero. Meanwhile, we are interested in estimating parameters that govern the spectral density of a random field when the frequency is very large so that we need to focus on Fourier frequencies that are away from zero.

In our work, we establish consistency and asymptotic normality of estimators of c and estimators of θ, respectively, when the other parameter is known. Some properties are also discussed when both parameters are unknown. Specially, if the Matérn covariance model is considered, c is related to a microergodic parameter. Consider the Matérn spectral density given as

$$f(\boldsymbol{\lambda}) = \frac{\sigma^2\alpha^{2\nu}}{\pi^{d/2}(\alpha^2 + |\boldsymbol{\lambda}|^2)^{\nu+d/2}}, \ \lambda \in R^d. \tag{1.7}$$

Matérn spectral density has three parameters (σ^2, α, ν), where σ^2 is the variance parameter, α is the scale parameter and ν is the smoothness parameter. Since the Matérn spectral density satisfies

$$f(\boldsymbol{\lambda}) \ \sim \ \frac{\sigma^2\alpha^{2\nu}}{\pi^{\frac{d}{2}}}|\boldsymbol{\lambda}|^{-(2\nu+d)}$$

as $|\boldsymbol{\lambda}| \to \infty$, we have $c \equiv \sigma^2\alpha^{2\nu}/\pi^{d/2}$ and $\theta \equiv 2\nu + d$, and $\sigma^2\alpha^{2\nu}$ is a microergodic parameter. Thus, estimating $\sigma^2\alpha^{2\nu}$ when ν is known is equivalent to estimate c when θ is known. There are several references that investigate estimation of $\sigma^2\alpha^{2\nu}$ in the spatial domain. Zhang (2004) showed that σ^2 and α can be estimated only in the form of $\sigma^2\alpha^{2\nu}$ under fixed-domain asymptotics when ν is known and $d \leq 3$. Du et al. (2009) investigated asymptotic properties of the MLE and the tapered MLE of $\sigma^2\alpha^{2\nu}$ when ν is known, α is fixed and $d = 1$ for a stationary Gaussian process. Anderes (2010) proposed an increment-based estimator of $\sigma^2\alpha^{2\nu}$ for a geometric anisotropic Matérn covariance function and showed that α can be estimated separately when $d > 4$.

The parameter θ is related to the fractal index (or fractal dimension) when the process $\{Z(\boldsymbol{s}), \boldsymbol{s} \in R^d\}$ is a stationary isotropic Gaussian process. For example, for a stationary Gaussian random field on $\mathbb{R}^d$, suppose that its covariance function $C(t)$ satisfies

$$C(t) \sim C(0) - k|t|^{\alpha} \quad \text{as} \quad |t| \to 0 \tag{1.8}$$

for some k and $0 < \alpha \leq 2$. In this case, α is the fractal index that governs the roughness of sample paths of the process and the fractal dimension D becomes $D = d + (1 - \alpha/2)$. This follows from Theorem 5.1 in Xue and Xiao (2010). When $\alpha = 2$ in (1.8), it is possible that the sample function may be differentiable. This can be determined by the smoothness of $C(t)$ in items of the spectral measure of $\{Z(s), s \in R^d\}$. Further information is in Adler and Taylor (2007) and Xue and Xiao (2010).

On Abelian type theorem, (1.8) holding the corresponding spectral density satisfies

$$f(\boldsymbol{\lambda}) \sim k'|\boldsymbol{\lambda}|^{-(\alpha+d)} \quad \text{as} \quad |\boldsymbol{\lambda}| \to \infty$$

so that $\theta \equiv \alpha + d$ in our settings.

The rest of this dissertation is organized in the following manner. First, in Chapter 2, we explain our settings and assumptions. We extend the results in Stein (1995) and Lim (2008) to more relaxed condition and then introduce our estimators and state theorems for the asymptotic properties of the proposed estimators. Simulation study will be presented in Chapter 3. In Chapter 4, we will discuss some issues related to our approach and possible extension of the current work. In the final chapter, we give proofs of our theoretical results.

Chapter 2

Main Results

2.1 Preliminary

In this work, we consider a stationary Gaussian random field, $Z(\boldsymbol{s})$ on $\mathbb{R}^d$ with the spectral density $f(\boldsymbol{\lambda})$ that satisfies (1.1). Define a lattice process $Y_\phi(\boldsymbol{J})$ by $Y_\phi(\boldsymbol{J}) \equiv Z(\phi \boldsymbol{J})$, where $\boldsymbol{J} \in \mathbb{Z}^d$, the set of d-dimensional integer-valued vectors. The corresponding spectral density of $Y_\phi(\boldsymbol{J})$ is

$$\bar{f}_\phi(\boldsymbol{\lambda}) = \phi^{-d} \sum_{\boldsymbol{Q} \in \mathbb{Z}^d} f\left(\frac{\boldsymbol{\lambda} + 2\pi \boldsymbol{Q}}{\phi}\right),$$

for $\boldsymbol{\lambda} \in (-\pi, \pi]^d$. Typically, $\bar{f}_\phi(\boldsymbol{\lambda})$ may have a peak near the origin which is getting higher as $\phi \to 0$. This causes a problem to estimate the spectral density using the periodogram [Stein (1995)]. To alleviate the problem, we consider a discrete Laplacian operator to difference the data, which is proposed by Stein (1995). The Laplacian operator is defined by

$$\Delta_\phi Z(\boldsymbol{s}) = \sum_{j=1}^{d} \left\{ Z(\boldsymbol{s} + \phi\, e_j) - 2Z(\boldsymbol{s}) + Z(\boldsymbol{s} - \phi\, e_j) \right\},$$

where e_j is the unit vector whose jth entry is 1. Depending on the behavior of the spectral density at high frequencies, we can apply the Laplacian operator iteratively to control the peak near the origin. Define $Y_\phi^\tau(\boldsymbol{J}) \equiv \left(\Delta_\phi\right)^\tau Z(\boldsymbol{s})$ as the lattice process obtained by applying the Laplacian operator τ times. Then as shown by Stein (1995), its corresponding spectral density becomes

$$\bar{f}_\phi^\tau(\boldsymbol{\lambda}) = \left\{ \sum_{j=1}^{d} 4\sin^2(\lambda_j/2) \right\}^{2\tau} \bar{f}_\phi(\boldsymbol{\lambda}). \tag{2.1}$$

Under the condition of (1.1), the limit of $\bar{f}_\phi^\tau(\boldsymbol{\lambda})$ as $\phi \to 0$ after scaling by $\phi^{d-\theta}$ is

$$\phi^{d-\theta} \bar{f}_\phi^\tau(\boldsymbol{\lambda}) \;\to\; c \left\{ \sum_{j=1}^{d} 4\sin^2(\lambda_j/2) \right\}^{2\tau} \sum_{\boldsymbol{Q} \in \mathbb{Z}^d} |\boldsymbol{\lambda} + 2\pi \boldsymbol{Q}|^{-\theta}$$

for $\boldsymbol{\lambda} \neq \boldsymbol{0}$. Define

$$g_{c,\theta}(\boldsymbol{\lambda}) = \begin{cases} c \left\{ \sum_{j=1}^{d} 4\sin^2(\lambda_j/2) \right\}^{2\tau} \sum_{\boldsymbol{Q} \in \mathbb{Z}^d} |\boldsymbol{\lambda} + 2\pi \boldsymbol{Q}|^{-\theta}, & \boldsymbol{\lambda} \in (-\pi, \pi]^d \backslash \{\boldsymbol{0}\}, \\ 0, & \boldsymbol{\lambda} = \boldsymbol{0}. \end{cases} \tag{2.2}$$

The limit function, $g_{c,\theta}(\boldsymbol{\lambda})$ is integrable by choosing τ such that $4\tau - \theta > -d$. When $d = 1$, simple differencing is preferred as discussed in Stein (1995). Then, 4τ will be replaced with 2τ in our results.

Now suppose that $Z(\boldsymbol{s})$ is observed on the lattice $\phi\boldsymbol{J}$. More specifically, we assume that we observe $Y_\phi^\tau(\boldsymbol{J})$ at $\boldsymbol{J} \in T_m = \{1, ..., m\}^d$ after differencing $Z(\boldsymbol{s})$ using the Laplacian operator τ times. We further assume that $\phi = m^{-1}$ so that the number of observations increases within a bounded observation domain. The spectral density of $Y_\phi^\tau(\boldsymbol{J})$ can be estimated by a periodogram which is defined using a discrete Fourier transform of the data. That is, the periodogram is defined by

$$I_m^\tau(\boldsymbol{\lambda}) = (2\pi m)^{-d} \left|D(\boldsymbol{\lambda})\right|^2,$$

where $D(\boldsymbol{\lambda})$ is the discrete Fourier transform of the data given by

$$D(\boldsymbol{\lambda}) = \sum_{\boldsymbol{J}\in T_m} Y_\delta^\tau(\boldsymbol{J}) \exp\{-i\,\boldsymbol{\lambda}^T\boldsymbol{J}\}.$$

We consider the periodogram only at Fourier frequencies, $2\pi m^{-1}\boldsymbol{J}$ for $\boldsymbol{J} \in \mathcal{T}_m = \{-\lfloor(m-1)/2\rfloor, \cdots, m - \lfloor m/2\rfloor\}^d$, where $\lfloor\, x\,\rfloor$ is the largest integer not greater than x. A smoothed periodogram at Fourier frequencies is defined by

$$\hat{I}_m^\tau\left(\frac{2\pi\boldsymbol{J}}{m}\right) = \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) I_m^\tau\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{K})}{m}\right),$$

with weights $W_h(\boldsymbol{K})$ given by

$$W_h(\boldsymbol{K}) = \frac{\Lambda_h\left(2\pi\boldsymbol{K}/m\right)}{\sum_{\boldsymbol{L}\in\mathcal{T}_m}\Lambda_h\left(2\pi\boldsymbol{L}/m\right)}, \tag{2.3}$$

where

$$\Lambda_h(\boldsymbol{s}) = \frac{1}{h}\Lambda\left(\frac{\boldsymbol{s}}{h}\right) \mathrm{I}_{\{||\,\boldsymbol{s}\,||\leq h\}}$$

for a symmetric continuous function Λ on $\mathbb{R}^d$ that satisfies $\Lambda(\boldsymbol{s}) \geq 0$ and $\Lambda(\mathbf{0}) > 0$ and $\mathrm{I}_{\mathcal{A}}$ is the indicator function of the set $\mathcal{A}$. The norm $||\cdot||$ is defined by $||\,\boldsymbol{s}\,|| = \max\{|s_1|, |s_2|, ..., |s_d|\}$.

For positive functions a and b, $a(\boldsymbol{\lambda}) \asymp b(\boldsymbol{\lambda})$ for $\boldsymbol{\lambda} \in \mathcal{A}$ means that there exist constants C_1 and C_2 such that $0 < C_1 \leq a(\boldsymbol{\lambda})/b(\boldsymbol{\lambda}) \leq C_2 < \infty$ for all possible $\boldsymbol{\lambda} \in \mathcal{A}$. For asymptotic results in this paper, we consider the following assumption on the spectral density $f(\boldsymbol{\lambda})$.

Assumption 1. *The spectral density* $f(\boldsymbol{\lambda})$ *of a stationary Gaussian random field* $\{Z(\boldsymbol{s}), \boldsymbol{s} \in \mathbb{R}^d\}$,

A. $f(\boldsymbol{\lambda}) \sim c\,|\boldsymbol{\lambda}|^{-\theta}$ *as* $|\boldsymbol{\lambda}| \to \infty$,

B. $f(\boldsymbol{\lambda})$ *is twice differentiable and there exists a positive constant* C *such that for* $|\boldsymbol{\lambda}| > C$,

$$f(\boldsymbol{\lambda}) \asymp (1+|\boldsymbol{\lambda}|)^{-\theta}, \left|\frac{\partial}{\partial \boldsymbol{\lambda}_j} f(\boldsymbol{\lambda})\right| \asymp (1+|\boldsymbol{\lambda}|)^{-(\theta+1)} \quad \textit{and}$$
$$\left|\frac{\partial^2}{\partial \boldsymbol{\lambda}_j \partial \boldsymbol{\lambda}_k} f(\boldsymbol{\lambda})\right| \asymp (1+|\boldsymbol{\lambda}|)^{-(\theta+2)} \tag{2.4}$$

for $j, k = 1, ..., d$.

2.2 Asymptotic properties of a smoothed periodogram

Asymptotic properties of a spatial periodogram and a smoothed spatial periodogram under fixed-domain asymptotics were investigated by Stein (1995) and Lim and Stein (2008). They assume that spectral density f is twice differentiable and satisfies (2.4) for all $\boldsymbol{\lambda} \in R^d$.

This assumption tells us that the spectral density $f(\boldsymbol{\lambda})$ behaves like $(1+|\boldsymbol{\lambda}|)^{-\theta}$ for all $\boldsymbol{\lambda}$, which is much stronger condition than (1.1). However this condition allows to find asymptotic bounds of expectation, variance and covariance of a spatial periodogram at Fourier frequency $2\pi\boldsymbol{J}/m$ for each $m \neq 0$ and $\boldsymbol{J}$ such that $\|\boldsymbol{J}\| \neq 0$. Consistency and asymptotic normality of a smoothed spatial periodogram at Fourier frequency $2\pi\boldsymbol{J}/m$, however, are shown when $\lim_{m\to\infty} 2\pi\boldsymbol{J}/m = \boldsymbol{\mu} \neq \mathbf{0}$, that is, $\boldsymbol{J}$ should not be closed to zero asymptotically. Since we make use of asymptotic properties of a smoothed spatial periodogram at such Fourier frequency under more general assumption (Assumption 1), we extend some of the results in Stein (1995) and Lim and Stein (2008) under Assumption 1. We focus only on a smoothed spatial periodogram in the following theorem, but results for a smoothed spatial cross-periodogram can be shown similarly. Throughout the dissertation, denote

$$\xrightarrow{p} \text{ by convergence in probability;}$$

$$\xrightarrow{d} \text{ by convergence in distribution.}$$

Theorem 3. *Suppose that the spectral density f of a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbb{R}^d$ satisfies Assumption 1. Also suppose that $4\tau > \theta - 1$ and $h = \mathfrak{C}m^{-\gamma}$ for some $\mathfrak{C} > 0$ where γ satisfies $\max\{(d-2)/d, 0\} < \gamma < 1$. Further, assume that $\lim_{m\to\infty} 2\pi\boldsymbol{J}/m = \boldsymbol{\mu}$ and $0 < \|\boldsymbol{\mu}\| < \pi$. Then, we have*

$$\frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{\bar{f}_{\phi}^{\tau}(2\pi\boldsymbol{J}/m)} \xrightarrow{p} 1 \tag{2.5}$$

and

$$m^{\eta}\left(m^{-(d-\theta)}\hat{I}_m^{\tau}\left(2\pi\boldsymbol{J}/m\right)-g_{c,\theta}(\boldsymbol{\mu})\right) \ \xrightarrow{d}\ \mathcal{N}\left(0,\frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d g_{c,\theta}^2(\boldsymbol{\mu})\right), \tag{2.6}$$

where $\eta = d(1-\gamma)/2$ *and* $\Lambda_r = \int_{[-1,1]^d}\Lambda^r(\boldsymbol{s})d\boldsymbol{s}$.

Remark 1. *The function* $g_{c,\theta}$ *is integrable under* $4\tau > \theta - d$ *which is satisfied by the condition* $4\tau > \theta - 1$*. The condition* $4\tau > \theta - 1$ *is necessary to show*

$$\mathrm{E}\left(\hat{I}_m^{\tau}\left(2\pi\boldsymbol{J}/m\right)/\bar{f}_{\phi}^{\,\tau}\left(2\pi\boldsymbol{J}/m\right)\right) \ \rightarrow\ 1$$

and the condition $\max\{(d-2)/d, 0\} < \gamma < 1$ *is needed to show*

$$\mathrm{Var}\left(\hat{I}_m^{\tau}\left(2\pi\boldsymbol{J}/m\right)/\bar{f}_{\phi}^{\,\tau}\left(2\pi\boldsymbol{J}/m\right)\right) \ \rightarrow\ 0$$

so that (2.5) can be shown.

2.3 Approach I

To estimate parameters, c and θ, we consider the following objective function to be minimized.

$$L(c,\theta) = \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\Big\{ \log\Big(m^{d-\theta} g_{c,\theta}\left(2\pi(\boldsymbol{J}+\boldsymbol{K})/m\right)\Big) + \frac{1}{m^{d-\theta}} \frac{I_m^{\mathcal{T}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \Big\}, \tag{2.7}$$

where $W_h(\boldsymbol{K})$ is given in (2.3). In $L(c,\theta)$, $2\pi\boldsymbol{J}/m$ is any given Fourier frequency that satisfies $\|\boldsymbol{J}\| \asymp m$ so that $2\pi\boldsymbol{J}/m$ is away from $\mathbf{0}$.

$L(c,\theta)$ can be viewed as a weighted Whittle likelihood function. When Λ is a nonzero constant function, $W_h(\boldsymbol{K}) \equiv 1/|\mathcal{K}|$ for $\boldsymbol{K} \in \mathcal{K}$, where $\mathcal{K} = \{\boldsymbol{K} \in \mathcal{T}_m : ||2\pi\boldsymbol{K}/m|| \leq h\}$ and $|\mathcal{K}|$ is the number of elements in the set $\mathcal{K}$. Then, $L(c,\theta)$ is the form of a local Whittle likelihood for the lattice data $\{Y_\delta^{\mathcal{T}}(\boldsymbol{J}), \boldsymbol{J} \in T_m\}$ in which the true spectral density is replaced with $m^{d-\theta} g_{c,\theta}$. Note that $g_{c,\theta}(\boldsymbol{\lambda})$ is the limit of the spectral density of $Y_\delta^{\mathcal{T}}(\boldsymbol{J})$ after being scaled by $m^{-(d-\theta)}$ for non-zero $\boldsymbol{\lambda}$ when $\phi = m^{-1}$. The summation in $L(c,\theta)$ is over the Fourier frequencies near $2\pi\boldsymbol{J}/m$ by letting $h \to 0$ as $m \to \infty$. While a local Whittle likelihood method to estimate a long-range dependence parameter for time series considers Fourier frequencies near zero, we consider Fourier frequencies near a pre-specified non-zero frequency. For example, by choosing $\boldsymbol{J}$ such that $\lfloor 2\pi\boldsymbol{J}/m \rfloor = (\pi/2)\,\mathbf{1}_d$, where $\mathbf{1}_d$ is the d-dimensional vector of ones, $L(c,\theta)$ considers frequencies only near $(\pi/2)\mathbf{1}_d$.

2.3.1 Estimation of c under the known θ

We consider the estimator of c by minimizing $L(c, \theta)$ when θ is known. Thus, the proposed estimator of c when θ is known as θ_0 is given by

$$\hat{c} = \arg\min_{c \in \mathcal{C}} L(c, \theta_0),$$

where $\mathcal{C}$ is the parameter space of c. $\hat{c}$ has the explicit expression obtained by $\partial L(c, \theta_0)/\partial c = 0$:

$$\hat{c} = \sum_{\boldsymbol{K} \in \mathcal{T}_m} W_h(\boldsymbol{K}) \frac{1}{m^{d-\theta_0}} \frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_0(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}, \tag{2.8}$$

where $g_0 \equiv g_{1,\theta_0}$. The following theorem establishes the consistency and asymptotic normality of the estimator $\hat{c}$.

Theorem 4. *Suppose that the spectral density f of a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbb{R}^d$ satisfies Assumption 1. Also suppose that $4\tau > \theta_0 - 1$ for a known θ_0 and $h = \mathfrak{C} m^{-\gamma}$ for some $\mathfrak{C} > 0$ where γ satisfies $d/(d+2) < \gamma < 1$. Further, assume that $\boldsymbol{J}$ satisfies $\lfloor 2\pi \boldsymbol{J}/m \rfloor = (\pi/2)\,\mathbf{1}_d$ and the true parameter c is in the interior of the parameter space $\mathcal{C}$ which is a closed interval. Then, for $\hat{c}$ given in (2.8), we have*

$$\hat{c} \ \xrightarrow{p} \ c, \tag{2.9}$$

and

$$m^{\eta}(\hat{c} - c) \ \xrightarrow{d} \ \mathcal{N}\left(0\,,\, c^2 \frac{\Lambda_2}{\Lambda_1^2} \left(\frac{2\pi}{\mathfrak{C}}\right)^d\right), \tag{2.10}$$

where $\Lambda_r = \int_{[-1,1]^d} \Lambda^r(\boldsymbol{s}) d\boldsymbol{s}$ *and* $\eta = d(1-\gamma)/2$.

Remark 2. *Theorem 4 can also be proved when we replace* θ_0 *in (2.8) with a consistent estimator* $\hat{\theta}$ *as long as the estimator* $\hat{\theta}$ *satisfies* $\hat{\theta} - \theta_0 = o_p((\log(m))^{-1})$.

Remark 3. *We can prove Theorem 4 for* $\boldsymbol{J}$ *such that* $\lim_{m\to\infty} 2\pi\boldsymbol{J}/m = \boldsymbol{\mu}$ *and* $0 < \|\boldsymbol{\mu}\| < \pi$ *instead of the specific choice of* $\lfloor 2\pi\boldsymbol{J}/m \rfloor = (\pi/2)\mathbf{1}_d$, *which we choose for simplicity in the proof.*

When we choose Λ as a constant function and $\mathfrak{C} = (1/2)\pi^2$, we have

$$m^{\eta}(\hat{c} - c) \;\xrightarrow{d}\; \mathcal{N}\left(0\,,\, 2^d c^2 \pi^{-d}\right).$$

For the Matérn spectral density given in (3.1) with $d = 1$, Du et al. (2009) showed that for any fixed α_1 with known ν, maximum likelihood estimator of σ^2 satisfies

$$n^{1/2}(\hat{\sigma}^2\alpha_1^{2\nu} - \sigma_0^2\alpha_0^{2\nu}) \;\xrightarrow{d}\; \mathcal{N}\left(0\,,\, 2(\sigma_0^2\alpha_0^{2\nu})^2\right), \tag{2.11}$$

where n is the sample size, and σ_0^2 and α_0 are true parameters. Note that m is the sample size of Y_δ^τ which is the τ times differenced lattice process of $Z(\boldsymbol{s})$. Since $\pi^{1/2}c = \sigma^2\alpha^{2\nu}$ for $d = 1$, we have the same asymptotic variance as in (2.11). However, our approach has a slower convergence rate since $\eta < 1/3$ when $d = 1$ as we used partial information. This is also the case for a local Whittle likelihood method in Robinson (1995).

2.3.2 Estimation of θ under the known c

To estimate θ, we assume that c is known as c_0. The proposed estimator of θ is then given by

$$\hat{\theta} = \arg\min_{\theta \in \Theta} L(c_0, \theta), \tag{2.12}$$

where Θ is the parameter space of θ. The consistency and the convergence rate of the proposed estimator $\hat{\theta}$ are given in the following Theorem.

Theorem 5. *Suppose that the spectral density f of a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbb{R}^d$ satisfies Assumption 1. Also suppose that $4\tau > \theta - 1$ and $h = \mathfrak{C}m^{-\gamma}$ for some $\mathfrak{C} > 0$ where γ satisfies $d/(d+2) < \gamma < 1$. Further, assume that $\boldsymbol{J}$ satisfies $\lfloor 2\pi \boldsymbol{J}/m \rfloor = (\pi/2)\, \mathbf{1}_d$ and the true parameter θ is in the interior of the parameter space Θ which is a closed interval. Then, for $\hat{\theta}$ given in (2.12), we have*

$$\hat{\theta} \xrightarrow{p} \theta. \tag{2.13}$$

In addition,

$$\hat{\theta} - \theta = o_p((\log m)^{-1}). \tag{2.14}$$

Remark 4. *The consistency of $\hat{\theta}$ is not enough to determine the asymptotic distribution of $\hat{\theta}$ since we have θ in the exponent of m in the expression of $L(c, \theta)$. For the proof of the asymptotic distribution, we need the rate of convergence given in (2.14).*

From Theorem 5, we can now show the following Theorem for the asymptotic distribution of $\hat{\theta}$.

Theorem 6. *Under the conditions of Theorem 5, we have*

$$\log(m)\, m^{\eta}(\hat{\theta}-\theta) \ \xrightarrow{d}\ \mathcal{N}\left(0, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right),$$

where $\eta = d(1-\gamma)/2$.

Remark 5. *Note that we have a different convergence rate for* $\hat{\theta}$ *compared to the convergence rate for* $\hat{c}$ *given in Theorem 4. The additional term* $\log(m)$ *is from the fact that* θ *is in the exponent of* m *in the expression of* $L(c,\theta)$.

2.3.3 Estimation under unknown c and θ

In the previous discussion, we consider estimation of one parameter when the other parameter is known. But in practice, both may be unknown. In order to handle this situation, c is assigned as any fixed value c^*. The estimator of θ is then defined by

$$\hat{\theta} = \arg\min_{\theta\in\Theta} L(c^*,\theta). \tag{2.15}$$

Theorem 7. *Suppose that the spectral density* f *of a stationary Gaussian random field* $Z(\boldsymbol{s})$ *on* $\mathbb{R}^d$ *satisfies Assumption 1. Also suppose that* $4\tau > \theta - 1$ *and* $h = \mathfrak{C}m^{-\gamma}$ *for some* $\mathfrak{C} > 0$ *where* γ *satisfies* $d/(d+2) < \gamma < 1$. *Further, assume that* $\boldsymbol{J}$ *satisfies* $\lfloor 2\pi\boldsymbol{J}/m\rfloor = (\pi/2)\,\mathbf{1}_d$ *and the true parameter* θ *is in the interior of the parameter space* Θ *which is a closed interval. Then, for* $\hat{\theta}$ *given in (2.15), we have*

$$\hat{\theta} \ \xrightarrow{p}\ \theta. \tag{2.16}$$

Furthermore,

$$\hat{\theta} - \theta \ = O_p((\log m)^{-1}). \tag{2.17}$$

In contrast to Theorem 5, The convergence rate of $\hat{\theta}$ is slower. With this convergence rate, we can not prove asymptotic distribution of $\hat{\theta}$. Also, we could consider the estimator of c_0 by minimizing $L(\hat{\theta}, c)$, where $\hat{\theta}$ defined in (2.15), that is,

$$\hat{c} = \sum_{\boldsymbol{K} \in \mathcal{T}_m} W_h(\boldsymbol{K}) \frac{1}{m^{d-\hat{\theta}}} \frac{I_m^{\mathcal{T}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{\hat{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}, \tag{2.18}$$

where $\hat{\theta}$ is the estimate of θ given in (2.15) with the fixed c^*. But, the consistency of $\hat{c}$ is not guaranteed. Instead, we obtain the following results which can be easily derived from

Corollary 1. $\hat{c} - c_0 = O_p(1)$.

2.4 Approach II

In Section 2.3, we developed a local Whittle type estimator which utilizes Fourier frequency information around $2\pi\boldsymbol{J}/m = (\pi/2)\mathbf{1}_d$. However, as the sample size increases, the Fourier frequencies used in the estimator will be very closed to $2\pi\boldsymbol{J}/m = (\pi/2)\mathbf{1}_d$. Thus, we could use $g_{c,\theta}(\cdot)$ only at $[2\pi\boldsymbol{J}/m]$. In this Section, we provide another estimation methodology which uses directly the smoothed periodogram with a fixed frequency. Alternative estimator is obtained by minimizing

$$R(c,\theta) \quad = \quad \log\left(m^{d-\theta} g_{c,\theta}\left(2\pi\boldsymbol{J}/m\right)\right) + \frac{1}{m^{d-\theta}} \frac{\hat{I}^{\tau}_m(2\pi\boldsymbol{J}/m)}{g_{c,\theta}(2\pi\boldsymbol{J}/m)}. \tag{2.19}$$

Asymptotic properties will be discussed in the rest of this Section, and the organization is same as the Section 2.3. Most theoretical results of the new estimators are identical with those obtained in Section 2.3 but require some changes in proof.

2.4.1 Estimation of c under known θ

The estimator of c is established by minimizing $R(c,\theta)$ when θ is known. Thus, when θ is known as θ_0, the proposed estimator of c is given by

$$\hat{c} = \arg\min_{c\in\mathcal{C}} R(c,\theta_0),$$

where $\mathcal{C}$ is the parameter space of c. By the similar way in Section 2.3, the exact form of $\hat{c}$ is obtained by solving the equation $\partial R(c,\theta_0)/\partial c = 0$:

$$\hat{c} = \frac{1}{m^{d-\theta_0}} \frac{\hat{I}^{\tau}_m(2\pi\boldsymbol{J}/m)}{g_0(2\pi\boldsymbol{J}/m)}, \tag{2.20}$$

where $g_0 \equiv g_{1,\theta_0}$. The same consistency and asymptotic results as in Section 2.3 hold for this estimator.

Theorem 8. *Suppose that the spectral density f of a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbb{R}^d$ satisfies Assumption 1. Also suppose that $4\tau > \theta_0 - 1$ for a known θ_0 and $h = \mathfrak{C} m^{-\gamma}$ for some $\mathfrak{C} > 0$ where γ satisfies $d/(d+2) < \gamma < 1$. Further, assume that $\boldsymbol{J}$ satisfies $\lfloor 2\pi \boldsymbol{J}/m \rfloor = (\pi/2)\,\mathbf{1}_d$ and the true parameter c is in the interior of the parameter space $\mathcal{C}$ which is a closed interval. Then, for $\hat{c}$ given in (2.20), we have*

$$\hat{c} \ \xrightarrow{p} \ c, \tag{2.21}$$

and

$$m^{\eta}(\hat{c} - c) \ \xrightarrow{d} \ \mathcal{N}\left(0\,,\, c^2 \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right), \tag{2.22}$$

where $\Lambda_r = \int_{[-1,1]^d} \Lambda^r(\boldsymbol{s}) d\boldsymbol{s}$ and $\eta = d(1-\gamma)/2$.

2.4.2 Estimation of θ under known c

Using (2.19), we can consider

$$\hat{\theta} = \arg\min_{\theta \in \Theta} R(c_0, \theta), \tag{2.23}$$

where Θ is the parameter space of θ when we assume that c is known as c_0. In the following Theorem, the consistency and the convergence rate of the new estimator $\hat{\theta}$ defined in (2.23) are provided.

Theorem 9. *Suppose that the spectral density f of a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbb{R}^d$ satisfies Assumption 1. Also suppose that $4\tau > \theta - 1$ and $h = \mathfrak{C}m^{-\gamma}$ for some $\mathfrak{C} > 0$ where γ satisfies $d/(d+2) < \gamma < 1$. Further, assume that $\boldsymbol{J}$ satisfies $\lfloor 2\pi \boldsymbol{J}/m \rfloor = (\pi/2)\,\mathbf{1}_d$ and the true parameter θ is in the interior of the parameter space Θ which is a closed interval. Then, for $\hat{\theta}$ given in (2.23), we have*

$$\hat{\theta} \xrightarrow{p} \theta. \tag{2.24}$$

In addition,

$$\hat{\theta} - \theta = o_p((\log m)^{-1}). \tag{2.25}$$

Remark 6. *With the similar way in the Section 2.3, the rate of convergence given in (2.25) is also useful for studying the asymptotic properties of $\hat{\theta}$. The same result as in the Section 2.3 will be shown.*

Theorem 10. *Under the conditions of Theorem 9, we have*

$$\log(m)\, m^{\eta}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}\left(0, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right),$$

where $\eta = d(1-\gamma)/2$.

2.4.3 Estimation under unknown θ and c

In this subsection, we also consider the situation when both parameters are unknown. With a given c^* which may be different from the true value c_0, the estimator of θ is established by

$$\hat{\theta} = \arg\min_{\theta \in \Theta} R(c^*, \theta). \tag{2.26}$$

Then, we have the following results which are similar to see 2.3.3.

Theorem 11. *Suppose that the spectral density f of a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbb{R}^d$ satisfies Assumption 1. Also suppose that $4\tau > \theta_0 - 1$ for a known θ_0 and $h = \mathfrak{C} m^{-\gamma}$ for some $\mathfrak{C} > 0$ where γ satisfies $d/(d+2) < \gamma < 1$. Further, assume that $\boldsymbol{J}$ satisfies $\lfloor 2\pi \boldsymbol{J}/m \rfloor = (\pi/2)\, \mathbf{1}_d$ and the true parameter θ is in the interior of the parameter space Θ which is a closed interval. Then, for $\hat{\theta}$ given in (2.26), we have*

$$\hat{\theta} \xrightarrow{p} \theta \tag{2.27}$$

Moreover,

$$\hat{\theta} - \theta = O_p((\log m)^{-1}). \tag{2.28}$$

If

$$\hat{c} = \frac{1}{m^{d-\hat{\theta}}} \frac{\hat{I}_m^{\tau}(2\pi \boldsymbol{J}/m)}{g_{\hat{\theta}}(2\pi \boldsymbol{J}/m)}, \tag{2.29}$$

is viewed as the estimator of true value c_0, we can show $\hat{c} - c_0 = O_p(1)$.

Based on the quantity of c^*, the overestimation and underestimation of $\hat{\theta}$ for true value θ_0 can be found in the following result.

Theorem 12. *(i) When* $c^* < c_0$*, there exists* M *such that* $P\left(\theta_0 \geq \hat{\theta}\right) = 0$ *for* $m > M$*.*

(ii) When $c^* > c_0$*, there exists* M *such that* $P\left(\theta_0 \geq \hat{\theta}\right) = 0$ *for* $m > M$*.*

Remark 7. *The properties of overestimation and underestimation for the first approach are also found from simulation study. However, theoretical results will be more complicated than second approach because the effect from* B_m *and* C_m *should be pored.*

Chapter 3

Simulation Study

In this chapter, simulation studies with various many models are introduced to validate the asymptotical results obtained in Chapter 2. Although estimators constructed in Chapter 2 work for high dimensional situation, one dimensional Matérn covariance model with various parameter values are considered here.

Let $Z(s)$ is a stationary Gaussian process on $\mathbb{R}$ with a Matérn covariance function whose spectral density follows (see e.g., Stein 1999, pp. 31)

$$f(\boldsymbol{\lambda}) = \sigma^2(\alpha^2 + \boldsymbol{\lambda})^{-\nu-1/2}. \tag{3.1}$$

Data are generated from the subroutine "mnrnd" in Matlab with covariances following (3.1). We consider the region $D = [0, 10]$ with different grid size $\phi = 0.1, 0.05$ and 0.025 which corresponds to $m = 100, 200,$ and 400. 500 data sets are simulated for each case. So that we have 500 parameter estimates.

For the sake of simplifying computation, function Λ is a constant function so that $W_h(\boldsymbol{K})$ is same for each $\boldsymbol{K} \in \mathcal{K}$. The four times finite difference operator ($\tau = 4$) is applied on the simulated data, and $C = 1$ and $\gamma = 1/3$ is chosen for the bandwidth . The notations used in the tables are defined as follows: m is sample size, $|\mathcal{K}|$ is the number of non-zero weights $W_h(\boldsymbol{K})$, Bias is the average of the bias obtained by estimations, and STD is the standard deviation of estimates.

In the first example, we consider $(\alpha, \sigma^2, \nu) = (1, 1/\pi, 1/2)$. In this case, true parameters of (c, θ) are $(1/\pi, 2)$. Table 3.1 and Table 3.2 are results of estimates θ and c, respectively.

Bias of Table 3.1 and Table 3.2 shows errors between our estimations and true value are less than 10^{-2} and STD means the estimations are very concentrated. Compared with the sample size (m), under the present bandwidth setting, the number of non-zero weights, $|\mathcal{K}|$, seems to be small for each $\boldsymbol{K} \in \mathcal{K}$, that is, small number of frequencies are used. The wider bandwidth setting is also considered by replacing $C = 1$ with $C = 5$, and simulation output is shown in Table 3.3. The Bias and STD are slightly improved in the new bandwidth setting.

The second simulation example comes from (3.1) with $(\alpha, \sigma^2, \nu) = (1, 1/\pi, 3/2)$ which implies $(c, \theta) = (1/\pi, 4)$. Under the same setting in the previous example with $C = 1$, The Bias and STD in Table 3.4 and 3.5 show similar results. Further, $C = 5$ is again applied to have wider bandwidth and the output is shown in Table 3.4. Although STD is improved, Bias in Table 3.6 did not be improved. From Table 3.3 and 3.6, the accuracy of estimation seems to be affected by which bandwidth we select. Therefore, it is important to find an optimal bandwidth. We will investigate this as a future research.

Under the same simulation setting as Table 3.6, the second approach is also applied and the output are shown in Table 3.7. Compared with Table 3.6, the performance of the second approach seems to be similar with the first one. This matches those theoretical results we found before.

We consider the estimating θ when c is also unknown. In previous examples whose true value are $(\theta, c) = (2, 1/\pi)$ and $(\theta, c) = (4, 1/\pi)$. θ is estimated when c is assumed as 2, 1, 0.2 and 0.1. The simulation output of previous two examples under different c are shown in Table 3.8 and Table 3.9, and their histograms are placed in the Figure 3.1 and 3.2. When c is bigger than true value, the Bias is positive and grows as c increase. In the Figure 3.1 and 3.2, if the selected c is $1/\pi$ (true value of c), the estimates distributed around the both sides of the true value of θ ($\theta = 4$). Meanwhile, when c is not equal to $1/\pi$, most of estimates is left or right of the true value. Moreover, in the Figure 3.3, trend of estimations is gradually moving to true value as the increase of sample size.

Table 3.1: Estimation of θ under known c

m	$\lvert\mathcal{K}\rvert$	$W_h(\boldsymbol{K})$	Bias	STD
100	7	1/7	0.039	0.129
200	10	1/10	0.009	0.088
400	17	1/17	0.009	0.05

Table 3.2: Estimation of c under known θ

m	$\lvert\mathcal{K}\rvert$	$W_h(\boldsymbol{K})$	Bias	STD
100	7	1/7	0.00072	0.12
200	10	1/10	0.0039	0.0945
400	17	1/17	0.0024	0.078

Table 3.3: Estimation of θ under known c

m	$\lvert\mathcal{K}\rvert$	$W_h(\boldsymbol{K})$	Bias	STD
100	33	1/33	-0.0024	0.0618
200	52	1/52	0.004	0.038
400	83	1/83	0.002	0.0256

Table 3.4: Estimation of θ under known c

m	$\|\mathcal{K}\|$	$W_h(\boldsymbol{K})$	Bias	STD
100	7	1/7	0.032	0.138
200	10	1/10	0.02	0.094
400	17	1/17	0.011	0.058

Table 3.5: Estimation of c under known θ

m	$\|\mathcal{K}\|$	$W_h(\boldsymbol{K})$	Bias	STD
100	7	1/7	0.014	0.132
200	10	1/10	0.003	0.094
400	17	1/17	-0.003	0.077

Table 3.6: Estimation of θ under known c

m	$\|\mathcal{K}\|$	$W_h(\boldsymbol{K})$	Bias	STD
100	33	1/33	0.04	0.066
200	52	1/52	0.031	0.042
400	83	1/83	-0.027	0.027

Table 3.7: Estimation of θ under known c (Second approach)

m	$\lvert\mathcal{K}\rvert$	$W_h(\boldsymbol{K})$	Bias	STD
100	33	1/33	0.004	0.077
200	52	1/52	0.026	0.047
400	83	1/83	-0.027	0.03

Table 3.8: Estimation of θ under unknown c for Example 1

c	$\lvert\mathcal{K}\rvert$	$W_h(\boldsymbol{K})$	Bias	STD
2	52	1/52	0.4907	0.0421
1	52	1/52	0.2996	0.0419
$1/\pi$	52	1/52	0.004	0.0378
0.2	52	1/52	-0.1364	0.0417
0.1	52	1/52	-0.3180	0.0378

Table 3.9: Estimation of θ under unknown c for Example 2

c	$\lvert\mathcal{K}\rvert$	$W_h(\boldsymbol{K})$	Bias	STD
2	52	1/52	0.5332	0.0418
1	52	1/52	0.2245	0.0415
$1/\pi$	52	1/52	0.031	0.042
0.2	52	1/52	-0.1309	0.0413
0.1	52	1/52	-0.3331	0.0415

Figure 3.1: Histogram of Example 1 on different c.

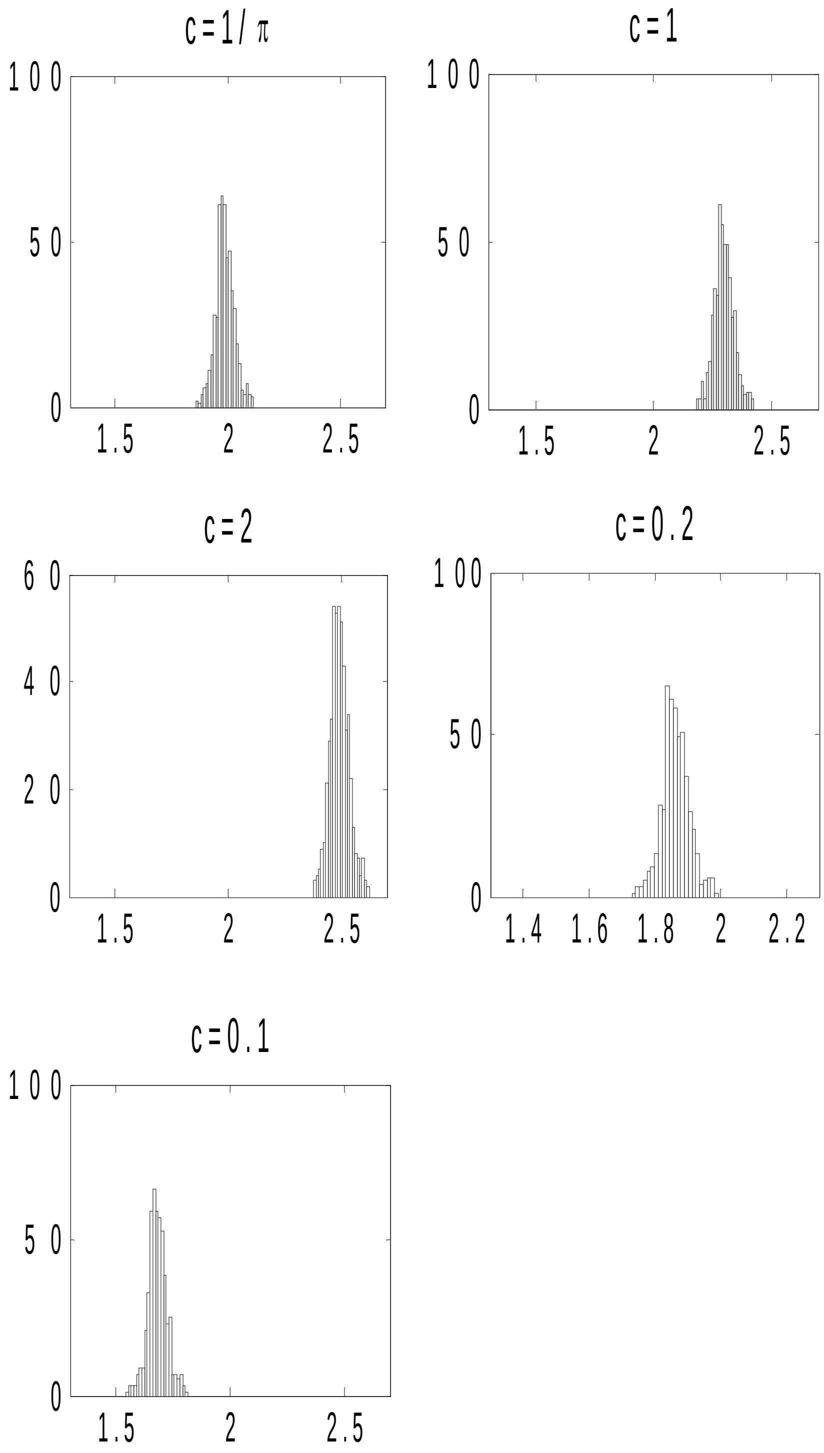


Figure 3.2: Histogram of Example 2 on different c.

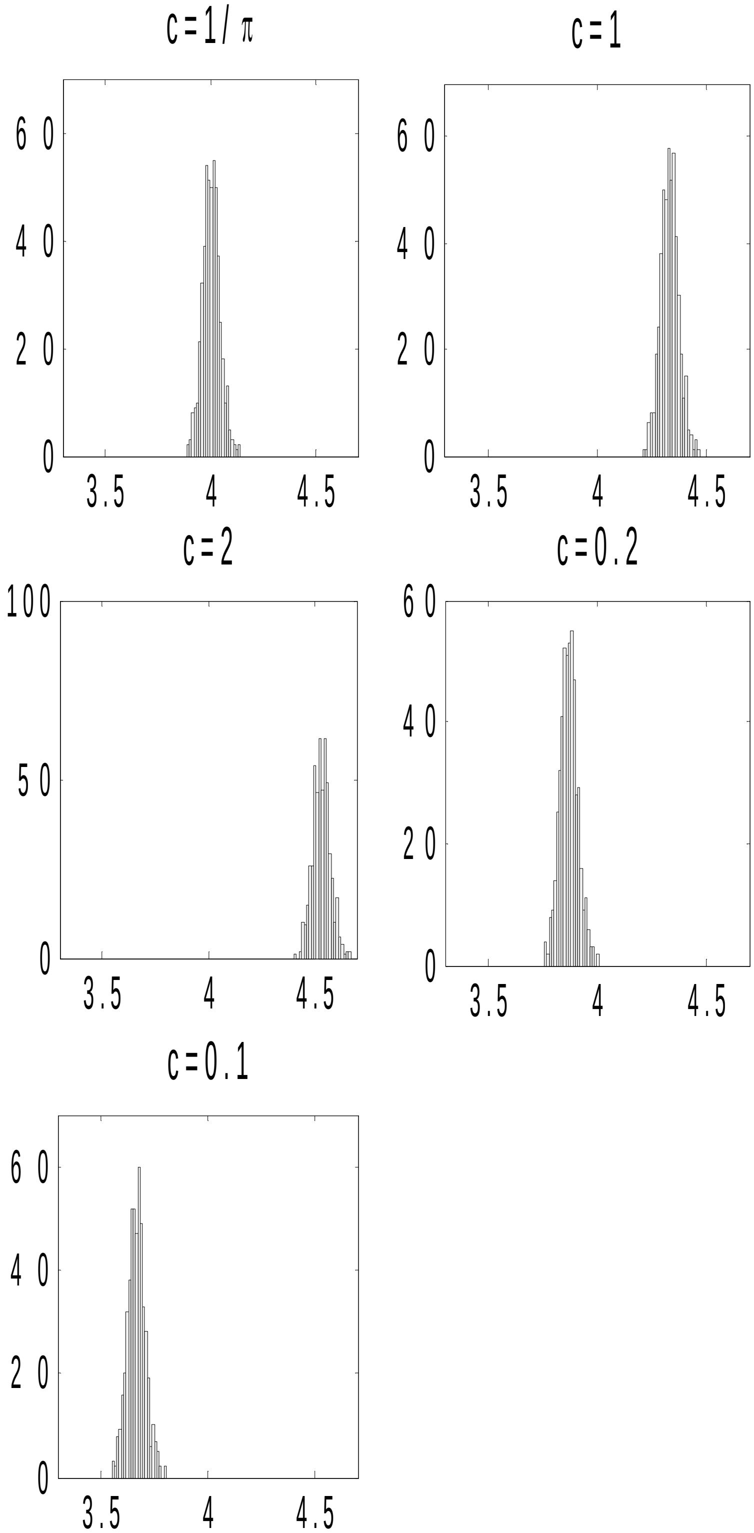
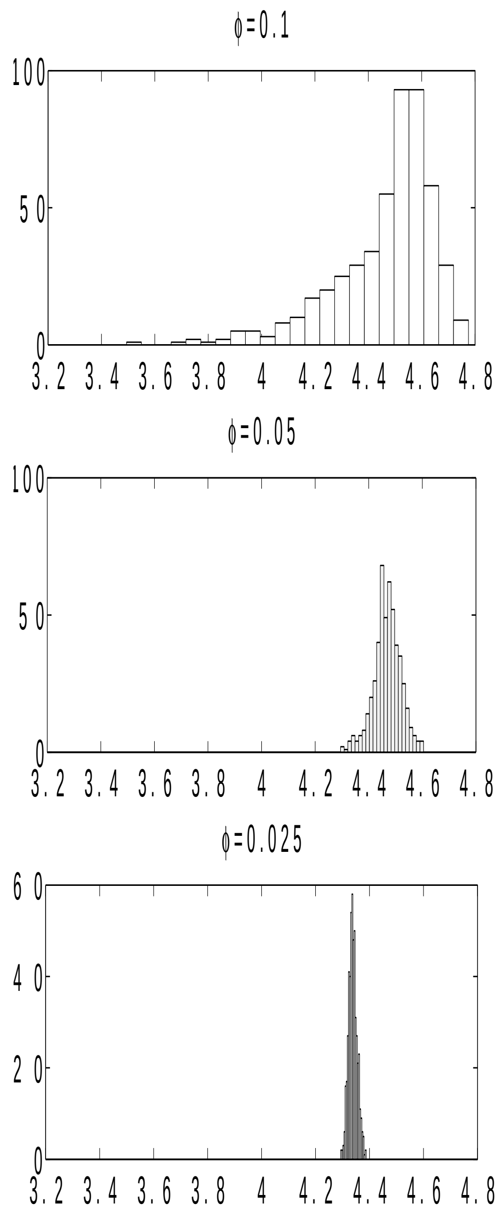


Figure 3.3: Histogrm of Example 2 with different grid sizes on wrong c.



Chapter 4

Discussion

In this dissertation, we first extended the result of Stein and Lim (2008) on weaker assumptions. Then, we proposed two approaches to estimate c and θ that govern the tail behavior of the spectral density of a stationary Gaussian random field on $\mathbb{R}^d$. The proposed estimators are obtained by minimizing the objective function given in (2.7) and (2.19). The first approach makes use of frequency information around $2\pi\boldsymbol{J}/m$. The second approach only employ the information from $[2\pi\boldsymbol{J}/m] = (\pi/2)1_d$. Regarding proofs of asymptotic results and simulation comparison, there is not much difference between these two approaches.

As mentioned in Chapter 2, the objective function given in (2.7) is similar to the one used in the local Whittle likelihood method when a kernel function Λ in $W_h(\boldsymbol{K})$ is constant. When we replace $m^{d-\theta}g_{c,\theta}$ with $\bar{f}^{\tau}_{\phi}(\boldsymbol{\lambda})$ and remove $W_h(\boldsymbol{K})$ in (2.7), it can be thought of an approximation to the likelihood of $\boldsymbol{Y}^{\tau}_{\phi}(\boldsymbol{J})$. This approximation, however, has not been verified under fixed-domain asymptotics. One might think that we can apply a similar technique to prove the validity of Whittle approximation to the likelihood since $\boldsymbol{Y}^{\tau}_{\phi}(\boldsymbol{J})$ is a lattice process. However, the spectral density $\bar{f}^{\tau}_{\phi}(\boldsymbol{\lambda})$ of $\boldsymbol{Y}^{\tau}_{\phi}(\boldsymbol{J})$ converges to zero, which require a different approach and further investigation is needed.

The weights in (2.7) is controlled by h, a bandwidth, which can be interpreted as a proportion of Fourier frequencies to be considered in the objective function. In our theorems, we assume $h = \mathfrak{C}m^{-\gamma}$ for some constant $\mathfrak{C}$. In proofs, we make use of the properties of a smoothed spatial periodogram $\hat{I}_m^{\mathcal{T}}$. Simulation results are also changing with different bandwidth. Thus, we could find the optimal bandwidth that minimizes the mean squared error of $\hat{I}_m^{\mathcal{T}}$. However, finding the mean squared error of $\hat{I}_m^{\mathcal{T}}$ needs explicit expressions of the bias and variance of $\hat{I}_m^{\mathcal{T}}(\boldsymbol{\lambda})$ and this requires further investigation. It will be more useful when we can estimate c and θ together or estimate θ when c is unknown. Due to the form of $g_{c,\theta}$, proving their asymptotic properties under fixed-domain asymptotics is challenging and needs different mathematics.

Although some contributions including theoretical results are made for the case which both parameters are unknown, more efforts are still need. In the current method, to estimate θ, c was pretended to be a fixed number c^* but convergence rate of $\hat{\theta}$ may be slower. To handle this problem, we believe updating c^* through $\hat{\theta}$ could be more reasonable, but how to update both estimators by an iterative way is still open.

The approaches of the fractal index could be another alternative way to research the tail behavior of the spectral density. By Abelian type theorem, some relationships between the tail of the spectral density and the origin of the covariance function have been existed. In this situation, the methodologies for the fractal index may be useful but the detail have to be carefully considered. Also, we believe our approaches should be available for the stationary increment process.

Finally, in our work, data are sampled from on the regular grid points. But in practice, irregular situation is more interesting. Several works or ideas discussed for increasing domain asymptotics may be also valid for fixed-domain asymptotics. Meanwhile, we are also interested in extending our univariate approaches univariate to multivariate situation.

Chapter 5

Appendix

5.1 The properties of $g_{c,\theta}(\boldsymbol{\lambda})$

Some properties of the function $g_{c,\theta}(\boldsymbol{\lambda})$ are discussed in this Appendix. These properties will be used in the proofs given in Appendix 5.2.1. Recall that

$$g_{c,\theta}(\boldsymbol{\lambda}) = c\left\{\sum_{j=1}^{d} 4\sin^2(\lambda_j/2)\right\}^{2\tau} \sum_{\boldsymbol{Q}\in\mathbb{Z}^d} |\boldsymbol{\lambda} + 2\pi\boldsymbol{Q}|^{-\theta}.$$

For a function $g_{c,\theta}(\boldsymbol{\lambda})$, let ∇g be the gradient of g with respect to $\boldsymbol{\lambda}$ and let $\dot{g}$ and $\ddot{g}$ denote the first and second derivatives of $g_{c,\theta}(\boldsymbol{\lambda})$ with respect to θ, respectively. That is,$\nabla g = (\partial g/\partial\lambda_1, \cdots, \partial g/\partial\lambda_d)$, $\dot{g} = \partial g_{c,\theta}(\boldsymbol{\lambda})/\partial\theta$ and $\ddot{g} = \partial^2 g_{c,\theta}(\boldsymbol{\lambda})/\partial\theta^2$.

We denote $\mathcal{A}_\rho = [-\pi,\pi]^d \setminus (-\rho,\rho)^d$ for a fixed ρ that satisfies $0 < \rho < 1$. Since we assume that the parameter space Θ is a closed interval in Chapter 2, let $\Theta = [\theta_L, \theta_U]$ and $\theta_L > d$. Although Lemma 1 can be shown for any fixed ρ with $0 < \rho < 1$, we further assume that ρ is small enough so that all Fourier frequencies near $(\pi/2)\mathbf{1}_d$ considered in $R(c,\theta)$ are contained in $\mathcal{A}_\rho$.

Lemma 1. *The following properties hold for* $g_{c,\theta}(\boldsymbol{\lambda})$*. Let* $c > 0$ *be a fixed constant.*

(a) There exist constants K_L *and* K_U *such that for all* $(\theta, \boldsymbol{\lambda}) \in \Theta \times \mathcal{A}_\rho$,

$$0 \; < \; K_L \; \leq \; g_{c,\theta}(\boldsymbol{\lambda}) \; \leq \; K_U \; < \; \infty. \tag{5.1}$$

(b) For any $\theta_1, \theta_2 \in \Theta$*, there exist constants* K_L *and* K_U *such that for all* $\boldsymbol{\lambda} \in \mathcal{A}_\rho$,

$$0 \; < \; K_L \; \leq \; g_{c,\theta_1}(\boldsymbol{\lambda})/g_{c,\theta_2}(\boldsymbol{\lambda}) \; \leq \; K_U \; < \; \infty. \tag{5.2}$$

(c) ∇g, $\dot{g}$, $\ddot{g}$, $\dot{g}/g$ *and* $\nabla(\dot{g}/g)$ *are uniformly bounded on* $\Theta \times \mathcal{A}_\rho$.

(d) $g_{c,\theta}(\boldsymbol{\lambda})$ *is continuous on* $\Theta \times \mathcal{A}_\rho$.

Proof. Since $g_{c,\theta}(\boldsymbol{\lambda})$ is linear in c, it will be enough just consider $g_{1,\theta}(\boldsymbol{\lambda})$. First, we find the upper and lower bounds of $\sum_{\boldsymbol{Q} \in \mathbb{Z}^d} |\boldsymbol{\lambda} + 2\pi \boldsymbol{Q}|^{-\theta}$. For all $(\theta, \boldsymbol{\lambda}) \in \Theta \times \mathcal{A}_\rho$, we have

$$\sum_{\boldsymbol{Q} \in \mathbb{Z}^d} |\boldsymbol{\lambda} + 2\pi \boldsymbol{Q}|^{-\theta} \geq \pi^{-\theta_U} > 0$$

and

$$\begin{aligned} \sum_{\boldsymbol{Q} \in \mathbb{Z}^d} |\boldsymbol{\lambda} + 2\pi \boldsymbol{Q}|^{-\theta} \;\; &\leq \;\; \sum_{\boldsymbol{Q} \in \mathbb{Z}^d \setminus \{\boldsymbol{0}\}} |\boldsymbol{\lambda} + 2\pi \boldsymbol{Q}|^{-\theta_L} + \epsilon^{-\theta_U} \\ &\leq \;\; (2\pi)^d \epsilon^{d-\theta_L}/(d-\theta_L) + \epsilon^{-\theta_U}, \end{aligned}$$

where the last inequality follows from

$$\begin{aligned}\sum_{\boldsymbol{Q}\in\mathbb{Z}^d\setminus\boldsymbol{0}}|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta_L} &\leq \int_{|\boldsymbol{y}|\geq 1}|\boldsymbol{\lambda}+2\pi\boldsymbol{y}|^{-\theta_L}d\boldsymbol{y}\\ &\leq \int_{|\boldsymbol{z}|\geq\epsilon}(2\pi)^d|\boldsymbol{z}|^{-\theta_L}d\boldsymbol{z}\\ &= \int_{x\,\geq\,\epsilon}(2\pi)^d x^{d-1}x^{-\theta_L}dx\\ &= (2\pi)^d\epsilon^{d-\theta_L}/(\theta_L-d), \end{aligned} \tag{5.3}$$

since $\theta_L > d$. Thus, we have

$$0 \;<\; k_L \;\leq\; \sum_{\boldsymbol{Q}\in\mathbb{Z}^d}|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta} \;\leq\; k_U < \infty, \tag{5.4}$$

where $k_L = \pi^{-\theta_U}$ and $k_U = (2\pi)^d\epsilon^{d-\theta_L}/(\theta_L-d)+\epsilon^{-\theta_U}$.

Then, (a) follows from (5.4),

$$(4d\sin^2(\epsilon/2))^{2\tau} \;\leq\; \left\{\sum_{j=1}^{d}4\sin^2(\lambda_j/2)\right\}^{2\tau} \;\leq\; (4d)^{2\tau},$$

and by setting $K_L \equiv c\,(4d\sin^2(\epsilon/2))^{2\tau}k_L$ and $K_U \equiv c\,(4d)^{2\tau}k_U$.

(b) follows from observing that $\sum_{\boldsymbol{Q}\in\mathbb{Z}^d}|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}$ has lower and upper bounds that are uniform on $\Theta\times\mathcal{A}_\rho$ as given in (5.4).

For (c), we have

$$\begin{aligned}
\left|\frac{\partial g}{\partial \lambda_i}\right| &= c\left|4\tau\left\{\sum_{j=1}^{d} 4\sin^2(\lambda_j/2)\right\}^{2\tau-1}\sin(\lambda_i)\sum_{\boldsymbol{Q}\in\mathbb{Z}^d}|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}\right. \\
&\quad \left. -\,\theta\left\{\sum_{j=1}^{d} 4\sin^2(\lambda_j/2)\right\}^{2\tau-1}\sum_{\boldsymbol{Q}\in\mathbb{Z}^d}(\lambda_i+2\pi Q_i)\,|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta-2}\right| \\
&\le K\sum_{\boldsymbol{Q}\in\mathbb{Z}^d}|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta} \\
&\le K\,k_U
\end{aligned}$$

for some constant $K > 0$ and k_U given in (5.4), which implies uniform boundedness of ∇g on $\Theta \times \mathcal{A}_\rho$. For the uniform bound of $\dot{g}$ and $\ddot{g}$, we first compute $\dot{g}$ and $\ddot{g}$:

$$\begin{aligned}
\dot{g} &= -c\left\{\sum_{j=1}^{d} 4\sin^2(\lambda_j/2)\right\}^{2\tau}\sum_{\boldsymbol{Q}\in\mathbb{Z}^d}|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}\log|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|\,, \\
\ddot{g} &= c\left\{\sum_{j=1}^{d} 4\sin^2(\lambda_j/2)\right\}^{2\tau}\sum_{\boldsymbol{Q}\in\mathbb{Z}^d}|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}\left(\log|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|\right)^2.
\end{aligned}$$

Since we can find x_0 and K such that for a given $\beta > 0$, $|\log x| \le Kx^\beta$ for all $x > x_0$, we can show that there exist n_0, K_1 and K_2 that satisfy

$$|\dot{g}| \le K_1 + K_2 \sum_{\boldsymbol{Q}\in\mathbb{Z}^d, ||\boldsymbol{Q}||\ge n_0} |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta+\beta}$$

for some fixed $\beta > 0$. When we choose $\beta = (\theta_L - \theta)/2$, we can show that

$$\sum_{\boldsymbol{Q}\in\mathbb{Z}^d, ||\boldsymbol{Q}||\ge n_0} |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta+\beta} < \infty$$

using a similar argument to show (5.3), which leads to uniform boundedness of $\dot{g}$. Similarly, we can show uniform boundedness of $\ddot{g}$.

The uniform boundedness of $\dot{g}/g$ follows from uniform boundedness of $\dot{g}$ and (a). To show uniform boundedness of $\nabla(\dot{g}/g)$, consider

$$\frac{\partial}{\partial \lambda_i}(\dot{g}/g) = -\frac{\sum_{\boldsymbol{Q}\in\mathbb{Z}^d} |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta-2}(\lambda_i+2\pi Q_i)(1-\theta\log|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|)}{\sum_{\boldsymbol{Q}\in\mathbb{Z}^d} |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}}$$
$$+\frac{\left(\sum_{\boldsymbol{Q}\in\mathbb{Z}^d} |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}\log|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|\right)\left(-\theta\sum_{\boldsymbol{Q}\in\mathbb{Z}^d} |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta-2}(\lambda_i+2\pi Q_i)\right)}{\left(\sum_{\boldsymbol{Q}\in\mathbb{Z}^d} |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}\right)^2}.$$

Since denominators in the expression of $\partial(\dot{g}/g)/\partial\lambda_i$ have uniform lower bounds as shown in (5.4), it is enough to find uniform bounds of numerators to show uniform boundedness of $\partial(\dot{g}/g)/\partial\lambda_i$. By observing that $|\lambda_i+2\pi Q_i| \le |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|$ and $|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-1} \le K$ for some $K>0$ on $\mathcal{A}_\rho$, we can show that each numerator in the expression of $\partial(\dot{g}/g)/\partial\lambda_i$ is uniformly bounded on $\Theta\times\mathcal{A}_\rho$ using a similar argument to show uniform boundedness of $\dot{g}$.

To show (d), it is enough to show the continuity of $\sum_{\boldsymbol{Q}\in\mathbb{Z}^d} |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}$ on $\Theta\times\mathcal{A}_\rho$ since $\left\{\sum_{j=1}^d 4\sin^2(\lambda_j/2)\right\}^{2\tau}$ is continuous on $\mathcal{A}_\rho$. It can be easily shown that

$$\sum_{\boldsymbol{Q}\in\mathbb{Z}^d, ||\boldsymbol{Q}||>n} |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}$$

converges to zero uniformly on $\Theta\times\mathcal{A}_\rho$ as $n\to\infty$, which implies the uniform convergence of $\sum_{\boldsymbol{Q}\in\mathbb{Z}^d, ||\boldsymbol{Q}||\le n} |\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}$ to $g(\theta,\boldsymbol{\lambda})$. Thus, the continuity of $g_{c,\theta}(\boldsymbol{\lambda})$ in $\boldsymbol{\lambda}$ follows from the continuity of $|\boldsymbol{\lambda}+2\pi\boldsymbol{Q}|^{-\theta}$.

□

5.2 Proofs of Theorems in Section 2

5.2.1 Proofs of Theorems in Section 2.2

Proof of Theorem 3. If $f(\boldsymbol{\lambda})$ satisfies (2.4) for all $\boldsymbol{\lambda}$, (2.5) and (2.6) hold by results in Stein (1995) and Lim and Stein (2008). To prove (2.5) and (2.6) when (2.4) holds only for large $\boldsymbol{\lambda}$, we need to show that the effect of $f(\boldsymbol{\lambda})$ on $|\boldsymbol{\lambda}| \le C$ is negligible.

Consider a spectral density $k(\boldsymbol{\lambda})$ which satisfies $k(\boldsymbol{\lambda}) \sim c|\boldsymbol{\lambda}|^{-\theta}$ as $|\boldsymbol{\lambda}| \to \infty$ and $k(\boldsymbol{\lambda})$ is twice differentiable and satisfies (2.4) for all $\boldsymbol{\lambda}$. Also assume that $k(\boldsymbol{\lambda}) \equiv f(\boldsymbol{\lambda})$ for $|\boldsymbol{\lambda}| > C$. Let $I_m^{f,\tau}(\boldsymbol{\lambda})$ be the periodogram at $\boldsymbol{\lambda}$ from the observations under $f(\boldsymbol{\lambda})$ and

$$\begin{aligned} & a_{m,\phi}^{f,\tau}(\boldsymbol{J},\boldsymbol{K}) \\ & \quad = (2\pi m)^{-d} \int_{\mathbb{R}^d} \left\{ \sum_{j=1}^{d} 4\sin^2\left(\frac{\phi\lambda_j}{2}\right) \right\}^{2\tau} f(\boldsymbol{\lambda})\Phi(\boldsymbol{\lambda},\boldsymbol{J},\boldsymbol{K})d\boldsymbol{\lambda}. \end{aligned}$$

where

$$\Phi(\boldsymbol{\lambda},\boldsymbol{J},\boldsymbol{k}) = \prod_{j=1}^{d} \frac{\sin^2\left(\frac{m\phi\lambda_j}{2}\right)}{\sin\left(\frac{\phi\lambda_j}{2}+\frac{\pi J_j}{m}\right)\sin\left(\frac{\phi\lambda_j}{2}+\frac{\pi K_j}{m}\right)}$$

Note that

$$\begin{aligned} E\left(I_m^{f,\tau}(2\pi\boldsymbol{J}/m)\right) &= a_{m,\phi}^{f,\tau}(\boldsymbol{J},\boldsymbol{J}), \\ \mathrm{Var}\left(I_m^{f,\tau}(2\pi\boldsymbol{J}/m)\right) &= a_{m,\phi}^{f,\tau}(\boldsymbol{J},\boldsymbol{J})^2 + a_{m,\phi}^{f,\tau}(\boldsymbol{J},-\boldsymbol{J})^2. \end{aligned}$$

(2.5) and (2.6) follow from Theorems 3, 6 and 12 in Lim and Stein (2008) when these Theorems hold for f under Assumption 1. The key part of proofs of these Theorems under

Assumption 1 is to show

$$\frac{E\left(I_m^{f,\tau}(2\pi\boldsymbol{J}/m)\right)}{\bar{f}_\phi^\tau(2\pi\boldsymbol{J}/m)} = 1+O(m^{-\beta_1}) \tag{5.5}$$

$$\frac{\mathrm{Var}\left(I_m^{f,\tau}(2\pi\boldsymbol{J}/m)\right)}{\bar{f}_\phi^\tau(2\pi\boldsymbol{J}/m)^2} = 1+O(m^{-\beta_2}), \tag{5.6}$$

for some $\beta_1, \beta_2 > 0$. Once (5.5) and (5.6) are shown, the other parts of proofs are similar to the proofs in Lim and Stein (2008).

Since results in Stein (1995) and Lim and Stein (2008) hold for $k(\boldsymbol{\lambda})$, we have (5.5) and (5.6) for $k(\boldsymbol{\lambda})$. Then, (5.5) and (5.6) for $f(\boldsymbol{\lambda})$ follow from

$$\left|a_{m,\phi}^{f,\tau}(\boldsymbol{J},\pm\boldsymbol{J}) - a_{m,\phi}^{k,\tau}(\boldsymbol{J},\pm\boldsymbol{J})\right| = O(m^{-d-4\tau}), \tag{5.7}$$

for $\boldsymbol{J}$ that satisfies $\|\boldsymbol{J}\| \asymp m$ and $2\boldsymbol{J}/m \notin \mathbb{Z}^d$. (5.7) holds since

$$\begin{aligned}
&\left|a_{m,\phi}^{f,\tau}(\boldsymbol{J},\pm\boldsymbol{J}) - a_{m,\phi}^{k,\tau}(\boldsymbol{J},\pm\boldsymbol{J})\right| \\
&= \left|(2\pi m)^{-d}\int_{|\boldsymbol{\lambda}|\le C}\left\{\sum_{j=1}^d 4\sin^2\left(\frac{\phi\lambda_j}{2}\right)\right\}^{2\tau}(f(\boldsymbol{\lambda})-k(\boldsymbol{\lambda}))\Phi(\boldsymbol{\lambda},\boldsymbol{J},\boldsymbol{k})d\boldsymbol{\lambda}\right| \\
&\le (2\pi m)^{-d}\int_{|\boldsymbol{\lambda}|\le C}\left\{\sum_{j=1}^d 4\sin^2\left(\frac{\phi\lambda_j}{2}\right)\right\}^{2\tau}|f(\boldsymbol{\lambda})-k(\boldsymbol{\lambda})|\,\Phi(\boldsymbol{\lambda},\boldsymbol{J},\boldsymbol{k})d\boldsymbol{\lambda} \\
&\le v\, m^{-d-4\tau}
\end{aligned}$$

for some positive constant v since $k(\boldsymbol{\lambda}) \equiv f(\boldsymbol{\lambda})$ for $|\boldsymbol{\lambda}| > C$ and $\left\|\phi\lambda_j/2 \pm \pi J_j/m\right\|$ stays away from zero and π when m is large.

□

5.2.2 Proofs of Theorems in Section 2.3

Proof of Theorem 4. To show weak consistency of $\hat{c}$, we consider upper and lower bounds of $\hat{c}$. Let

$$\boldsymbol{K}^{\mathcal{U}} = \mathrm{argmax}_{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0}\, g_0\left(2\pi(\boldsymbol{J}+\boldsymbol{K})/m\right)$$

and

$$\boldsymbol{K}^{\mathcal{L}} = \mathrm{argmin}_{\boldsymbol{K}\in\mathcal{T}_m,, W_h(\boldsymbol{K})\neq 0}\, g_0\left(2\pi(\boldsymbol{J}+\boldsymbol{K})/m\right).$$

Recall that $g_0 = g_{1,\theta_0}$. Then, we have

$$\begin{aligned}\frac{\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0} g_0(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{U}})/m)} &\leq \hat{c} \\ &\leq \frac{\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0} g_0(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{L}})/m)}\end{aligned}$$

which can be rewritten as

$$\begin{aligned}\frac{c\,\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{U}})/m)} &\leq \hat{c} \\ &\leq \frac{c\,\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{L}})/m)} \qquad (5.8)\end{aligned}$$

with probability one. Note that both $g_{c,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{U}})/m)$ and $g_{c,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{L}})/m)$ converge to $g_{c,\theta_0}((\pi/2)\mathbf{1}_d)$ by continuity of $g_{c,\theta}(\boldsymbol{\lambda})$ and $m^{-(d-\theta_0)}\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)$ converges to $g_{c,\theta_0}((\pi/2)\mathbf{1}_d)$ in probability by Theorem 3. Thus, it follows that $\hat{c}$ converges to c in probability.

For the asymptotic distribution of $\hat{c}$, note that we have

$$m^{\eta}\left(\frac{\hat{I}_m^{\tau}\left(2\pi\boldsymbol{J}/m\right)}{m^{d-\theta_0}}-g_{c,\theta_0}\left((\pi/2)\mathbf{1}_d\right)\right) \xrightarrow{d} \mathcal{N}\left(0\,,\,\frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d g_{c,\theta_0}^2((\pi/2)\mathbf{1}_d)\right) \tag{5.9}$$

from Proposition 12 in Lim and Stein (2008) and

$$m^{\eta}\left(g_{c,\theta_0}\left(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{E}})/m\right)-g_{c,\theta_0}\left((\pi/2)\mathbf{1}_d\right)\right) \longrightarrow 0, \tag{5.10}$$

for $\mathcal{E}=\mathcal{U}$ or $\mathcal{L}$, since $4\tau>\theta_0-1, h=\mathfrak{C}m^{-\gamma}$ and $\frac{d}{d+2}<\gamma<1$. Then, (2.10) follows from (5.50) and (5.10).

□

To prove Theorem 5, we consider following lemmas.

Lemma 2. *Consider a function $h_m(x)=-\log(x)+d_m(x-1)$, where d_m is positive and a function of a positive integer m. Also assume that $d_m\to 1$ as $m\to\infty$. Then, for a given r with $0<r<1$, there exist $\delta_r>0$ and M_r such that for all $m\geq M_r$,*

$$h_m(x)>\delta_r,$$

for any $x\in\mathfrak{Z}_r$, where $\mathfrak{Z}_r=\{z:|z-1|>r, z>0\}$.

Proof. It can be easily shown that for any positive integer m, $h_m(x)$ is a convex function on $(0,\infty)$ and minimized at $x=1/d_m$ with $h_m(1/d_m)\leq 0$. Let $h_\infty(x)=-\log(x)+x-1$. Since $d_m\to 1$, for any $r\in(0,1)$, there exists $M_r>0$ such that for all $m\geq M_r$, we have $|1/d_m-1|\leq r$ and $\min\{h_m(1-r),h_m(1+r)\}>(1/2)\min\{h_\infty(1-r),h_\infty(1+r)\}>0$.

Hence for all $x \in \mathfrak{Z}_r$, we have

$$h_r(x) \geq \min\{h_m(1-r), h_m(1+r)\} > (1/2)\min\{h_\infty(1-r), h_\infty(1+r)\} \equiv \delta_r.$$

□

The following lemma shows that $L(c_0, \theta_1) - L(c_0, \theta_0)$ can be bounded from below by three terms and two of them can be neglected.

Lemma 3. *For a positive integer m and $\theta_1 \in \Theta$, we have*

$$L(c_0, \theta_1) - L(c_0, \theta_0) \geq A_m + B_m + C_m,$$

where

$$\begin{aligned} A_m = &-\log\left(m^{\theta_1-\theta_0}\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}\right) \\ &+ \frac{\hat{I}_m^\delta(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}\left(m^{\theta_1-\theta_0}\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}-1\right), \end{aligned} \tag{5.11}$$

$$B_m = \log\left(\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}\frac{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}\right) \tag{5.12}$$

$$C_m = \frac{\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}\left(1-\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)}\right). \tag{5.13}$$

In (5.11)-(5.13), $\boldsymbol{K}_M$, $\boldsymbol{K}_m$, $\boldsymbol{S}_M$ and $\boldsymbol{S}_m$ are defined as

$$\begin{aligned}
\boldsymbol{K}_M &= \arg\max_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m), \\
\boldsymbol{K}_m &= \arg\min_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m), \\
\boldsymbol{S}_M &= \arg\max_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} \log\left(\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\right), \\
\boldsymbol{S}_m &= \arg\min_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}.
\end{aligned}$$

Furthermore,

$$\sup_{\theta\in\Theta} |B_m| = o(1), \tag{5.14}$$

$$C_m = o_p(1), \tag{5.15}$$

where (5.15) is under the conditions of Theorem 5.

Proof. From the expression of $L(c,\theta)$ given in (2.7), we have

$$\begin{aligned}
&L(c_0,\theta_1) - L(c_0,\theta_0) \\
&= -\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \log\left(m^{\theta_1-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \right) \\
&\quad + \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} m^{\theta_1-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \\
&\quad - \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \\
&\geq \log\left(m^{\theta_1-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)} \right) \\
&\quad + \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} m^{\theta_1-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} \\
&\quad - \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)} \\
&= -\log\left(m^{\theta_1-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}{g_{c_0,\theta_m}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)} \right) + \frac{\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} \\
&\times \left(m^{\theta_1-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} - \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)} \right) \\
&=: H_m.
\end{aligned}$$

H_m is further decomposed as

$$\begin{aligned}
H_m &= -\log\left(m^{\theta^1-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} \right) \\
&\quad + \frac{\hat{I}_m^\delta(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} \left(m^{\theta_1-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta^1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} - 1 \right) \\
&\quad + \log\left(\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)} \frac{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} \right) \\
&\quad + \frac{\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} \left(1 - \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)} \right),
\end{aligned}$$

which is $A_m + B_m + C_m$ given in (5.11)-(5.13).

Note that $2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m$, $2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m$, $2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m$ and $2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m$ converge to $(\pi/2)\mathbf{1}_d$ as $m \to \infty$. Note also that the convergence of $2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m$ and $2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m$ holds for θ_1 uniformly on Θ, because $h \to 0$.

The continuity of $g_{c_0,\theta}$ in Lemma 1 implies that as $m \to \infty$,

$$\log\left(\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}\frac{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}\right) \quad \longrightarrow \quad 0 \tag{5.16}$$

holds for θ_1 uniformly on Θ, therefore, $\sup_\Theta |B_m| = o(1)$. Also, we have

$$m^{-(d-\theta_0)}\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)/g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m) \ \xrightarrow{p}\ 1,$$

since $m^{-(d-\theta_0)}\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)/g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)$ converges to one in probability by Theorem 3 and $g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)$ converges to $g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)$. Thus, together with

$$1 - \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)} \quad \to \quad 0,$$

C_m converges to one in probability.

□

Theorem 13 (Egorov theorem (Folland 1999))**.** *Suppose that $\nu(X) < \infty$, and $f_1, f_2, \ldots$ and f are measurable complex-valued functions on X such that $f_n \to f$ a.e. Then for every $\epsilon > 0$ there exists $E \subseteq X$ such that $\nu(E) < \epsilon$ and $f_n \to f$ uniformly on E^c.*

Proof of Theorem 5. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be the probability space where a stationary Gaussian random field $Z(\boldsymbol{s})$ is defined. To emphasize dependence on m, we use $\hat{\theta}_m$ instead of $\hat{\theta}$ in this proof. Note that we have

$$P(L(c_0, \hat{\theta}_m) - L(c_0, \theta_0) \le 0) = 1 \tag{5.17}$$

for any positive integer m, due to the definition of $\hat{\theta}_m$. We are going to prove the theorem by deriving a contradiction to (5.17) when $\hat{\theta}_m$ does not converge to θ_0 in probability.

Suppose that $\hat{\theta}_m$ does not converge to θ_0 in probability. Then, there exist $\epsilon > 0$, $\delta > 0$ and M_1 such that for $m \ge M_1$,

$$P(|\hat{\theta}_m - \theta_0| > \epsilon) > \delta.$$

We define $\mathcal{D}_m = \{\omega \in \Omega : |\hat{\theta}_m - \theta_0| > \epsilon\}$. By Lemma 3, we have

$$L(c_0, \hat{\theta}_m) - L(c_0, \theta_0) \ge A_m + B_m + C_m,$$

where A_m, B_m and C_m are given in (5.11)-(5.13) with $\theta_1 = \hat{\theta}$. Also, note that

$$A_m = h_m \left(m^{\hat{\theta}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_m)/m)}{g_{c_0,\hat{\theta}}(2\pi(\boldsymbol{J} + \boldsymbol{S}_m)/m)} \right),$$

where $h_m(\cdot)$ is defined in Lemma 2 with

$$d_m = \frac{\hat{I}_m^{\delta}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{K}_M)/m)}, \tag{5.18}$$

where $\boldsymbol{K}_M$ is defined in Lemma 3.

We are going to show that there exist $\{m_k\}$, a subsequence of $\{m\}$ and a subset of $\mathcal{D}_{m_k}$

such that for large enough m_k, $A_{m_k} + B_{m_k} + C_{m_k}$ is bounded away from zero.

By Theorem 3 and the convergence of $g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{K}_M)/m)$ to $g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)$, we have $d_m \xrightarrow{p} 1$. Then, there exists $\{m_k\}$, a subsequence of $\{m\}$ such that d_{m_k} converges to one almost surely. By (5.17) in Lemma 3, almost sure convergence of d_{m_k} implies that C_{m_k} given in (5.13) converges to zero almost surely. To use Lemma 2, we need uniform convergence of d_{m_k} which is obtained by Egorov's Thoerem (Folland, 1999). By Egorov's Theorem, there exists $\mathcal{G}_\delta \subset \Omega$ such that d_{m_k} and C_{m_k} converge uniformly on $\mathcal{G}_\delta$ and $P(\mathcal{G}_\delta) > 1 - \delta/2$.

On the other hand, there exists a M_2, which does not depend on ω, such that for $m_k \geq M_2$,

$$\left| m_k^{\hat{\theta}_{m_k} - \theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)}{g_{c_0,\hat{\theta}_{m_k}}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)} - 1 \right| > \frac{1}{2} \tag{5.19}$$

for all $\omega \in \mathcal{D}_{m_k}$, because of the uniform boundedness of $g_{c_0,\theta_0}/g_{c_0,\theta_1}$.

Let $\mathcal{H}_{m_k} = \mathcal{D}_{m_k} \cap \mathcal{G}_{m_k}$. Note that $P(\mathcal{H}_{m_k}) > \delta/2 > 0$ for $m_k \geq M_1$. Then, by Lemma 2 with $r = 1/2$, there exist $\delta_r > 0$ and M_r such that for $m_k \geq M_r$,

$$\begin{aligned}
A_{m_k} &= -\log\left(m_k^{\hat{\theta}_{m_k} - \theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)} \right) \\
&\quad + \frac{\hat{I}^{\delta}_{m_k}(2\pi\boldsymbol{J}/m_k)}{m_k^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{K}_M)/m_k)} \left(m_k^{\hat{\theta}_{m_k} - \theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)}{g_{c_0,\theta_1}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)} - 1 \right) \\
&> \delta_r
\end{aligned} \tag{5.20}$$

uniformly on $\mathcal{H}_{m_k}$. Note here that $M_r \geq \max\{M_1, M_2\}$.

By the uniform convergence of $|B_m|$ on Θ shown in Lemma 3, there exists a M_3 such that for $m_k \geq M_3$,

$$\left| B_{m_k} \right| \;<\; \frac{\delta_r}{4} \tag{5.21}$$

with $\theta_1 = \hat{\theta}_{m_k}(\omega)$ uniformly for $\omega \in \Omega$. The uniform convergence of C_{m_k} on $\mathcal{G}_\delta$ allows us to find M_4 such that for $m_k \geq M_4$,

$$\left| C_{m_k} \right| \;<\; \frac{\delta_r}{4} \tag{5.22}$$

uniformly on $\mathcal{H}_{m_k}$.

Therefore, for $m_k \geq \max\{M_r, M_3, M_4\}$, we have $A_{m_k} + B_{m_k} + C_{m_k} \geq A_{m_k} - |B_{m_k}| - |C_{m_k}| > \delta_r/2$ on $\mathcal{H}_{m_k}$ which leads

$$L(c_0, \hat{\theta}_{m_k}) - L(c_0, \theta_0) > \frac{\delta_r}{2} \tag{5.23}$$

on $\mathcal{H}_{m_k}$. Since $P(\mathcal{H}_{m_k}) > \delta/2 > 0$, it contradicts to (5.17) which completes the proof. Here, we do not need $P(\cap_k \mathcal{H}_{m_k}) \;>\; 0$ since (5.17) should holds for any $m > 0$.

To show (2.14), it is enough to show that $m^{\hat{\theta}-\theta_0} \xrightarrow{p} 1$ which is equivalent to show that

$$\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_m)/m)}{g_{c_0,\hat{\theta}}(2\pi(\boldsymbol{J} + \boldsymbol{S}_m)/m)} \;\xrightarrow{p}\; 1, \tag{5.24}$$

$$m^{\hat{\theta}-\theta_0}\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_m)/m)}{g_{c_0,\hat{\theta}}(2\pi(\boldsymbol{J} + \boldsymbol{S}_m)/m)} \;\xrightarrow{p}\; 1. \tag{5.25}$$

(5.24) follows from the consistency of $\hat{\theta}$ and the continuity of $g_{c_0,\theta}$ shown in Lemma 5.1.

To show (5.25), notice that we have

$$P\left(L(c_0,\hat{\theta}) - L(c_0,\theta_0) \ \leq \ 0\right) = 1 \tag{5.26}$$

for each $m > 0$ by the definition of $\hat{\theta}$ and we have

$$P\left(L(c_0,\hat{\theta}) - L(c_0,\theta_0) \geq A_m + B_m + C_m\right) = 1$$

by Lemma 3.

Suppose that (5.25) does not hold. Then, there exists $r > 0$, $\delta > 0$ and M_1 such that

$$P\left(\left| m^{\hat{\theta}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\hat{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} - 1\right| \ > \ r\right) > \delta$$

for all $m \geq M_1$. On the other hand, there exists $\{m_k\}$, a subsequence of $\{m\}$, such that $d_{m_k} \to 1$, $B_m \to 0$ and $C_m \to 0$ almost surely, where d_m is given in (5.18), B_m and C_m are given in (5.12) and (5.13) with $\theta_1 = \hat{\theta}$. Then, by Egorov's Thoerem, there exists $\Omega_\delta \subset \Omega$ such that $P(\Omega_\delta) > 1 - \delta/2$ and d_{m_k}, B_m and C_m are uniformly convergent on Ω_δ. As in Lemma 2, for a_{m_k}, a nonzero solution of $h_{m_k}(b_{m_k}) = 0$, where

$$b_m = m^{\hat{\theta}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\hat{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)},$$

there exists M_2 such that $|a_{m_k} - 1| \leq r$ uniformly on Ω_δ for all $m_k \geq M_2$. Now, define

$$\mathcal{D}_m = \left\{\omega \ : \ \left| m^{\hat{\theta}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\hat{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} - 1\right| \ > \ r\right\}. \tag{5.27}$$

Note that $P(\mathcal{D}_{m_k} \cap \Omega_\delta) \geq \delta/2 > 0$ for all $m_k \geq \max\{M_1, M_2\}$. Similarly to the proof of Lemma 2, for each $m_k \geq \max\{M_1, M_2\}$, there exists $\delta_r > 0$ such that $A_{m_k} > \delta_r$ for all $\omega \in \mathcal{D}_{m_k} \cap \Omega_\delta$. This implies that

$$P(A_{m_k} > \delta_r) \geq \delta/2$$

for each $m_k \geq \max\{M_1, M_2\}$. Note that δ_r does not depend on m_k which can be seen in Lemma 2. Meanwhile, there exists M_3 such that for $m_k \geq M_3$,

$$|B_{m_k}| \leq \delta_r/4, \quad |C_{m_k}| \leq \delta_r/4$$

for all $\omega \in \Omega_\delta$. Hence we have

$$P\left(L(c_0, \hat{\theta}) - L(c_0, \theta_0) \;>\; \delta_r/2\right) \geq \delta/2$$

for $m_k \geq \max\{M_1, M_2, M_3\}$, which contradicts to (5.26). Thus, (5.25) is proved. □

Alternative Proof of Theorem 5. To show the consistency of $\hat{\theta}$, for a given $\epsilon > 0$ such that $0 < \epsilon < \min\{\theta_U - \theta_0, \theta_L - \theta_0\}/2$, define $\Theta_\epsilon = \{\theta : |\theta - \theta_0| \leq \epsilon\}$ and Θ_ϵ^c is the complement of Θ_ϵ. Then, we have

$$\begin{aligned} P\left(\hat{\theta} \in \Theta_\epsilon^c \cap \Theta\right) &= P\left(\inf_{\Theta_\epsilon^c \cap \Theta} L(c_0, \theta) \leq \inf_{\Theta_\epsilon \cap \Theta} L(c_0, \theta)\right) \\ &\leq P\left(\inf_{\Theta_\epsilon^c \cap \Theta} (L(c_0, \theta) - L(c_0, \theta_0)) \;\leq\; 0\right). \end{aligned}$$

By Lemma 3, we also have

$$\begin{aligned}\inf_{\Theta_\epsilon^c \cap \Theta} (L(c_0,\theta) - L(c_0,\theta_0)) &\geq \inf_{\Theta_\epsilon^c \cap \Theta} (A_m + B_m + C_m) \\ &\geq \inf_{\Theta_\epsilon^c \cap \Theta} (A_m - |B_m|) + C_m \\ &\geq \inf_{\Theta_\epsilon^c \cap \Theta} A_m - \sup_{\Theta} |B_m| + C_m,\end{aligned}$$

where A_m, B_m and C_m are given in (5.11)-(5.13). Thus, to show the consistency of $\hat{\theta}$, it is enough to show that there exists $\delta > 0$ such that

$$P\left(\inf_{\Theta_\epsilon^c \cap \Theta} A_m + C_m \ > \ \delta \right) \ \longrightarrow \ 1.$$

since B_m is deterministic with $\sup_\Theta |B_m| \ \to \ 0$ as $m \ \to \ \infty$. We can consider A_m as

$$A_m = h_m \left(m^{\theta-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} \right),$$

where $h_m(\cdot)$ is defined in Lemma 2 with

$$d_m = \frac{\hat{I}_m^\delta(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}, \tag{5.28}$$

where $\boldsymbol{K}_M$ is defined in Lemma 3. For $\theta \in \Theta_\epsilon^c \ \cap \ \Theta$, if $\theta > \theta_0 + \epsilon$,

$$m^{\theta-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} \ \longrightarrow \ \infty$$

as $m \to \infty$, because of the uniform boundedness of $g_{c_0,\theta_0}/g_{c_0,\theta}$ shown in Lemma 1. Similarly, if $\theta < \theta_0 - \epsilon$,

$$m^{\theta-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} \;\longrightarrow\; 0$$

as $m \to \infty$. Thus, there exists M_1 such that for $m \geq M_1$,

$$\left| m^{\theta-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} - 1 \right| \;>\; \frac{1}{2}, \tag{5.29}$$

for all $\theta \in \Theta_\epsilon^c \,\cap\, \Theta$, because of the uniform boundedness of $g_{c_0,\theta_0}/g_{c_0,\theta}$.

By Theorem 12 in Lim and Stein (2008) and the convergence of $g_{c_0,\theta_0}\,(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)$ to $g_{c_0,\theta_0}\,((\pi/2)\mathbf{1}_d)$, $d_m \xrightarrow{p} 1$. Similarly, we can show that $C_m \;\xrightarrow{p}\; 0$. Then, there exists a $\delta > 0$ such that

$$P\left(\inf_{\Theta_\epsilon^c \,\cap\, \Theta} A_m + C_m \;>\; \delta \right) \;\longrightarrow\; 1 \tag{5.30}$$

by Lemma 2 with $r = 1/2$ and the fact that randomness of A_m and C_m comes from the same quantity d_m. This completes the proof of (2.13).

□

To proof Theorem 6, we consider the following Lemma.

Lemma 4. *Under the conditions of Theorem 5, let $\eta = d(1-\gamma)/2$, we have*

(a)

$$m^{\eta}\left(\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \;-\; 1\right)$$
$$\xrightarrow{d} \;\; \mathcal{N}\left(0, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right), \qquad (5.31)$$

(b)

$$\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\left(1 \;-\; \frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\right)\frac{\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}$$
$$= O_p(m^{-\eta}) \qquad (5.32)$$

Proof. To prove (5.31), we find the asymptotic distribution of its lower and upper bounds. It can be easily shown that

$$LB_m \;\le\; m^{\eta}\left(\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \;-\; 1\right) \;\le\; UB_m,$$

where

$$LB_m \;=\; m^{\eta}\left(\frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} \;-\; 1\right), \qquad (5.33)$$

$$UB_m \;=\; m^{\eta}\left(\frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)} \;-\; 1\right) \qquad (5.34)$$

with $\boldsymbol{K}_M$ and $\boldsymbol{K}_m$ as given in Lemma 3. We rewrite LB_m as

$$LB_m = m^{\eta}\Bigg(\left(\frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)} - 1\right)\frac{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} + \frac{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} - 1\Bigg).$$

By Lemma 1 and $\gamma > d/(d+2)$, we have

$$\frac{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} \longrightarrow 1,$$

$$m^{\eta}\left(\frac{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} - 1\right) \longrightarrow 0.$$

Thus, by Theorem 3,

$$LB_m \xrightarrow{d} \mathcal{N}\left(0, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{c}}\right)^d\right).$$

Similarly, we can show

$$UB_m \xrightarrow{d} \mathcal{N}\left(0, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{c}}\right)^d\right).$$

The lower and upper bounds converge to the same distribution which implies (5.31).

To show (5.32), we rewrite the LHS of (5.32) as

$$\begin{aligned}
&\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\left(1 \;-\; \frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\right)\frac{\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\\
&\quad= \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\frac{\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} - \frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\\
&\quad- \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\frac{\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\\
&\quad+ \frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}.
\end{aligned}$$

By Lemma 1 and $\gamma > d/(d+2)$, we can show that

$$m^\eta\left(\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\frac{\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} - \frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\right) \;\longrightarrow\; 0.$$

Also, it can be easily shown that

$$LB_m \;\le\; \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\frac{\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \;\le\; UB_m,$$

where

$$\begin{aligned}
LB_m &= \frac{\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\frac{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{P}_m)/m)}{g^2_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{P}_m)/m)},\\
UB_m &= \frac{\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\frac{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{P}_M)/m)}{g^2_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{P}_M)/m)},
\end{aligned}$$

with

$$
\begin{aligned}
\boldsymbol{P}_M &= \arg\max_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} \frac{\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g^2_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}, \\
\boldsymbol{P}_m &= \arg\min_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} \frac{\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g^2_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}.
\end{aligned}
$$

By Lemma 1, $\gamma > d/(d+2)$ and Theorem 3, we can show that

$$
\begin{aligned}
m^\eta\left(LB_m - \frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\right) &\xrightarrow{d} \mathcal{N}\left(0, \left(\frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\right)^2 \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right), \\
m^\eta\left(UB_m - \frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\right) &\xrightarrow{d} \mathcal{N}\left(0, \left(\frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\right)^2 \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right).
\end{aligned}
$$

This completes the proof of (5.32). □

Proof of Theorem 6. Let $\dot{L} = \partial L/\partial\theta$ and $\ddot{L} = \partial^2 L/\partial\theta^2$. To show the asymptotic distribution of $\hat{\theta}$, we consider the Taylor expansion of $\dot{L}(c_0, \hat{\theta})$ around θ_0,

$$
\dot{L}(c_0, \hat{\theta}) = \dot{L}(c_0, \theta_0) + \ddot{L}(c_0, \bar{\theta})(\hat{\theta} - \theta_0),
$$

where $\bar{\theta}$ lies on the line segment between $\hat{\theta}$ and θ_0. Since $\dot{L}(c_0, \hat{\theta}) = 0$, we have

$$
\log(m) m^\eta(\hat{\theta} - \theta_0) = -log(m) m^\eta \left(\ddot{L}(c_0, \bar{\theta})\right)^{-1} \dot{L}(c_0, \theta_0).
$$

Thus, it is enough to show

$$
(\log(m))^{-1} m^\eta \dot{L}(c_0, \theta_0) \xrightarrow{d} \mathcal{N}\left(0, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right), \tag{5.35}
$$

$$
(\log(m))^{-2} \ddot{L}(c_0, \bar{\theta}) \xrightarrow{p} 1. \tag{5.36}
$$

Since

$$\begin{aligned}
\dot{L}(c_0,\theta_0) &= -\log(m) + \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \\
&\quad - \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m) \\
&\quad \times \frac{\Big(-\log(m) m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m) + m^{d-\theta_0}\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)\Big)}{\Big(m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)\Big)^2} \\
&= \log(m)\left(\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} - 1\right) \\
&\quad + \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \left(1 - \frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\right) \frac{\dot{g}_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)},
\end{aligned}$$

we see that (5.35) follows from Lemma 4.

Next we prove (5.36). After some simplification, we have

$$\begin{aligned}
\ddot{L}(c_0,\bar{\theta}) &= (\log(m))^2 \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\bar{\theta}} g_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \\
&\quad - 2\log(m) \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)\dot{g}_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\bar{\theta}} g^2_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \\
&\quad + 2 \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)\dot{g}^2_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\bar{\theta}} g^3_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \\
&\quad + \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \left(1 - \frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\bar{\theta}} g_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\right) \frac{\ddot{g}_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \\
&\quad - \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{\dot{g}^2_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g^2_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \\
&=: E_1 + E_2,
\end{aligned}$$

where E_1 is the first term with $(\log(m))^2$ and E_2 is the last four terms in the expression of

$\ddot{L}(c_0, \bar{\theta})$.

First, we want to show that

$$(\log(m))^{-2} E_1 \xrightarrow{p} 1. \tag{5.37}$$

It can be easily shown that

$$LB_m \le (\log(m))^{-2} E_1 \le UB_m,$$

where

$$\begin{aligned} LB_m &= \frac{\hat{I}^{\tau}_m(2\pi \boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)} \frac{m^{\bar{\theta}-\theta_0} g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{P}_M)/m)}, \\ UB_M &= \frac{\hat{I}^{\tau}_m(2\pi \boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)} \frac{m^{\bar{\theta}-\theta_0} g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{P}_m)/m)} \end{aligned}$$

with

$$\begin{aligned} \boldsymbol{P}_M &= \arg\max_{\{\boldsymbol{K} \in \mathcal{T}_m, W_h(\boldsymbol{K}) \neq 0\}} g_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m), \\ \boldsymbol{P}_m &= \arg\min_{\{\boldsymbol{K} \in \mathcal{T}_m, W_h(\boldsymbol{K}) \neq 0\}} g_{c_0,\bar{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m). \end{aligned}$$

By Theorem 3, (2.14) in Theorem 5 and Lemma 1, we can show that both LB_m and UB_m converge to one in probability, which in turn implies (5.37). In a similar way, we can show that $(\log(m))^{-1} E_2 = O_p(1)$. Thus, together with (5.37), we can show (5.36), which completes the proof.

□

In order to prove Theorem 7, we will extend Lemma 2 to more generalized situation.

Lemma 5. *Consider a function* $h_m(x) = -\log(x) + d_m(x-1)$, $(x > 0)$, *where* $\{d_m\}$ *is a sequence of positive numbers such that* $d_m \to d > 0$ *as* $m \to \infty$. *Then, there exists some* $r_l \in (0,1)$ *and* $r_u \in (1,\infty)$, $\delta_r > 0$ *and* M *such that* $\forall m \geq M$, *we have*

$$h_m(x) > 1$$

$\forall x \in (0, r_l] \cup [r_u, \infty)$.

Proof. Since $d_m \to d > 0$, then $\forall \epsilon \in (0,d), \exists M$ s.t. $\forall m \geq M$,

$$|d_m - d| < \epsilon$$

or $d - \epsilon < d_m < d + \epsilon$.

$\forall c > 0$ fixed, note that the function $f_c(x) = -\log(x) + c(x-1), (x > 0)$ has the following properties:

(i)

$$f_c(x) \to \infty \text{ as } x \to 0^+ \text{ or } x \to \infty.$$

(ii) $f_c'(x) = -\frac{1}{x} + c$. So, $f_c'(x) = 0 \Leftrightarrow x = \frac{1}{c}$.

$$f_c'(x) < 0 \text{ if } x < \frac{1}{c} \text{ and } f_c'(x) > 0 \text{ if } x > \frac{1}{c}.$$

(iii) $f_c(x)$ attains its minimum at $x = \frac{1}{c}$ and $f_c(x) \leq 0$. ($f_c(\frac{1}{c}) < 0$ if $c \neq 1$. Otherwise,

$f_1(1) = 0.$) Hence we can find $x_1 < \frac{1}{c} < x_2$ such that

$$f_c(x) \geq 1$$

if $0 < x \leq x_1$ or $x \geq x_2$. Now we apply the above facts to $c = d - \epsilon$ or $c = d + \epsilon$ to get the following:

(a) If $0 < x \leq x_1$, then

$$h_m(x) = -\log(x) + d_m(x-1) \geq -\log(x) + (d+\epsilon)(x-1) \geq 1.$$

(b) If $x \geq x_2$, then

$$h_m(x) = -\log(x) + d_m(x-1) \geq -\log(x) + (d-\epsilon)(x-1) \geq 1.$$

Therefore, we have proved the Lemma. □

To prove Theorem 11, we first find the lower bound for $L(c^*, \theta_1) - L(c^*, \theta_0)$. The construction of this lower bound follows by replacing c_0 in (5.11),(5.12) and (5.13) in Lemma 3 with c^*. The lower bounded is also established by three terms and two of them are dominated by the other.

Lemma 6. *For a positive integer m and any $\theta_1 \in \Theta$, we have*

$$L(c^*, \theta_1) - L(c^*, \theta_0) \geq A_m + B_m + C_m,$$

where

$$A_m = -\log\left(m^{\theta_1-\theta_0}\frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c^*,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}\right) + \frac{\hat{I}^{\delta}_m(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}\left(m^{\theta_1-\theta_0}\frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c^*,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}-1\right), \tag{5.38}$$

$$B_m = \log\left(\frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}\frac{g_{c^*,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}{g_{c^*,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}\right) \tag{5.39}$$

$$C_m = \frac{\hat{I}^{\tau}_m(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}\left(1-\frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}{g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)}\right). \tag{5.40}$$

In (5.38)-(5.40), $\boldsymbol{K}_M$*,* $\boldsymbol{K}_m$*,* $\boldsymbol{S}_M$ *and* $\boldsymbol{S}_m$ *are defined as*

$$\begin{aligned}
\boldsymbol{K}_M &= \arg\max_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m),\\
\boldsymbol{K}_m &= \arg\min_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m),\\
\boldsymbol{S}_M &= \arg\max_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} \log\left(\frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c^*,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\right),\\
\boldsymbol{S}_m &= \arg\min_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} \frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c^*,\theta_1}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}.
\end{aligned}$$

Furthermore,

$$\sup_{\theta\in\Theta}|B_m| = o(1), \tag{5.41}$$

$$C_m = o_p(1), \tag{5.42}$$

where (5.42) is under the conditions of Theorem 5.

Proof of Lemma 6. The procedure of the proof for this Lemma is the same as Lemma 6. Therefore, we will not introduce the details.

□

Proof of Theorem 7. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be the probability space where a stationary Gaussian random field $Z(\boldsymbol{s})$ is defined. To emphasize dependence on m, we use $\hat{\theta}_m$ instead of $\hat{\theta}$ in this proof. Note that we have

$$P(L(c^*, \hat{\theta}_m) - L(c^*, \theta_0) \leq 0) = 1 \tag{5.43}$$

for any positive integer m, due to the definition of $\hat{\theta}_m$. We are going to prove the theorem by deriving a contradiction to (5.43) when $\hat{\theta}_m$ does not converge to θ_0 in probability.

Suppose that $\hat{\theta}_m$ does not converge to θ_0 in probability. Then, there exist $\epsilon > 0$, $\delta > 0$ and M_1 such that for $m \geq M_1$,

$$P(|\hat{\theta}_m - \theta_0| > \epsilon) > \delta.$$

We define $\mathcal{D}_m = \{\omega \in \Omega : |\hat{\theta}_m - \theta_0| > \epsilon\}$. By Lemma 3, we have

$$L(c^*, \hat{\theta}_m) - L(c^*, \theta_0) \geq A_m + B_m + C_m,$$

where A_m, B_m and C_m are given in (5.38)-(5.40) with $\theta_1 = \hat{\theta}$. Also, note that

$$A_m = h_m \left(m^{\hat{\theta} - \theta_0} \frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_m)/m)}{g_{c^*,\hat{\theta}}(2\pi(\boldsymbol{J} + \boldsymbol{S}_m)/m)} \right),$$

where $h_m(\cdot)$ is defined in Lemma 5 with

$$d_m = \frac{\hat{I}_m^{\delta}(2\pi \boldsymbol{J}/m)}{m^{d-\theta_0} g_{c^*,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{K}_M)/m)}, \tag{5.44}$$

where $\boldsymbol{K}_M$ is defined in Lemma 3.

We are going to show that there exist $\{m_k\}$, a subsequence of $\{m\}$ and a subset of $\mathcal{D}_{m_k}$ such that for large enough m_k, $A_{m_k} + B_{m_k} + C_{m_k}$ is bounded away from zero.

By Theorem 3 and the convergence of $g_{c^*,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{K}_M)/m)$ to $g_{c^*,\theta_0}((\pi/2)\mathbf{1}_d)$, we have $d_m \xrightarrow{p} d = c_0/c^*$. Then, there exists $\{m_k\}$, a subsequence of $\{m\}$ such that d_{m_k} converges to one almost surely. By (5.17) in Lemma 3, almost sure convergence of d_{m_k} implies that C_{m_k} given in (5.40) converges to zero almost surely. To use Lemma 5, we need uniform convergence of d_{m_k} which is obtained by Egorov's Thoerem (Folland, 1999). By Egorov's Theorem, there exists $\mathcal{G}_\delta \subset \Omega$ such that d_{m_k} and C_{m_k} converge uniformly on $\mathcal{G}_\delta$ and $P(\mathcal{G}_\delta) > 1 - \delta/2$.

On the other hand, there exists a M_2, which does not depend on ω, such that for $m_k \geq M_2$,

$$m_k^{\hat{\theta}_{m_k} - \theta_0} \frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)}{g_{c^*,\hat{\theta}_{m_k}}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)} \tag{5.45}$$

falls on the outside of (r_l, r_u) for all $\omega \in \mathcal{D}_{m_k}$, because of the uniform boundedness of $g_{c^*,\theta_0}/g_{c^*,\theta_1}$.

Let $\mathcal{H}_{m_k} = \mathcal{D}_{m_k} \cap \mathcal{G}_{m_k}$. Note that $P(\mathcal{H}_{m_k}) > \delta/2 > 0$ for $m_k \geq M_1$. Then, by Lemma 5, there exist $\delta_r > 0$ and M_r such that for $m_k \geq M_r$,

$$\begin{aligned} A_{m_k} &= -\log\left(m_k^{\hat{\theta}_{m_k} - \theta_0} \frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)}{g_{c^*,\theta_1}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)} \right) \\ &\quad + \frac{\hat{I}^\delta_{m_k}(2\pi \boldsymbol{J}/m_k)}{m_k^{d-\theta_0} g_{c^*,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{K}_M)/m_k)} \left(m_k^{\hat{\theta}_{m_k} - \theta_0} \frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)}{g_{c^*,\theta_1}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)} - 1 \right) \\ &> \delta_r \end{aligned} \tag{5.46}$$

uniformly on $\mathcal{H}_{m_k}$. Note here that $M_r \geq \max\{M_1, M_2\}$.

By the uniform convergence of $|B_m|$ on Θ shown in Lemma 6, there exists a M_3 such that for $m_k \geq M_3$,

$$\left| B_{m_k} \right| < \frac{\delta_r}{4} \tag{5.47}$$

with $\theta_1 = \hat{\theta}_{m_k}(\omega)$ uniformly for $\omega \in \Omega$. The uniform convergence of C_{m_k} on $\mathcal{G}_\delta$ allows us to find M_4 such that for $m_k \geq M_4$,

$$\left| C_{m_k} \right| < \frac{\delta_r}{4} \tag{5.48}$$

uniformly on $\mathcal{H}_{m_k}$.

Therefore, for $m_k \geq \max\{M_r, M_3, M_4\}$, we have $A_{m_k} + B_{m_k} + C_{m_k} \geq A_{m_k} - |B_{m_k}| - |C_{m_k}| > \delta_r/2$ on $\mathcal{H}_{m_k}$ which leads

$$L(c^*, \hat{\theta}_{m_k}) - L(c^*, \theta_0) > \frac{\delta_r}{2} \tag{5.49}$$

on $\mathcal{H}_{m_k}$. Since $P(\mathcal{H}_{m_k}) > \delta/2 > 0$, it contradicts to (5.43) which completes the proof. Here, we do not need $P(\cap_k \mathcal{H}_{m_k}) > 0$ since (5.43) should holds for any $m > 0$.

(2.17) comes from

$$\lim_{m\to\infty} P\left(m^{\hat{\theta}-\theta_0} \frac{g_{c^*,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_m)/m)}{g_{c^*,\hat{\theta}}(2\pi(\boldsymbol{J} + \boldsymbol{S}_m)/m)} \in (r_l, r_u) \right) = 1.$$

Otherwise, the same contradiction to (5.43) will be found.

To prove (2.29), let

$$\begin{aligned}\boldsymbol{K}_M &= \arg\max_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} g_{\hat{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m),\\ \boldsymbol{K}_m &= \arg\min_{\{\boldsymbol{K}\in\mathcal{T}_m, W_h(\boldsymbol{K})\neq 0\}} g_{\hat{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m).\end{aligned}$$

Assume

$$\hat{c} = \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{1}{m^{d-\hat{\theta}}} \frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{\hat{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}.$$

$$\begin{aligned}& c\, m^{\hat{\theta}-\theta} \frac{1}{m^{d-\theta}} \frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{g_{\theta,c}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} \frac{g_{\theta}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}{g_{\hat{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}\\ &\leq \hat{c} \leq c\, m^{\hat{\theta}-\theta} \frac{1}{m^{d-\theta}} \frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{g_{\theta,c}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)} \frac{g_{\theta}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)}{g_{\hat{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)}\end{aligned}$$

By Theorem 3 and the convergence of $g_{c,\theta}\left(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m\right)$ and $g_{c,\theta}\left(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m\right)$ to $g_{c,\theta}\left((\pi/2)\mathbf{1}_d\right)$,

$$\frac{1}{m^{d-\theta}} \frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{g_{\theta,c}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} \to^p 1$$

and

$$\frac{1}{m^{d-\theta}} \frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{g_{\theta,c}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)} \to^p 1.$$

Corollary 1 is verified because $\hat{\theta}-\theta = O_p(\log(m)^{-1})$ and the boundedness of $g_{c,\theta}$.

□

5.2.3 Proofs of Theorems in Section 2.4

The idea to verify the theoretical results of the second estimator defined in Section 2.4 are similar with Section 2.3. The procedures of proofs will be simpler and worked out by Theorem 3 and Lemma 2.

Proof of Theorem 8. Compared with the first estimator in Section 2.3, the consistency of $\hat{c}$ will be directly attained because $m^{-(d-\theta_0)}\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)$ converges to $g_{c,\theta_0}((\pi/2)\mathbf{1}_d)$ in probability by Theorem 3.

$$\hat{c} = \frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_0(2\pi\boldsymbol{J}/m)} \to^p c.$$

The asymptotic distribution of $\hat{c}$ comes from Theorem 3

$$\begin{aligned} m^{\eta}\left(\frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}} - g_{c,\theta_0}((\pi/2)\mathbf{1}_d)\right) & \\ \xrightarrow{d} \mathcal{N}\left(0\,,\, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d g_{c,\theta_0}^2((\pi/2)\mathbf{1}_d)\right) & \qquad (5.50) \end{aligned}$$

□

Proof of Theorem 9. For all θ_1 and θ_2 in Θ,

$$\begin{aligned} R(c_0,\theta_1) - R(c_0,\theta_2) = -\log\left(m^{\theta_1-\theta_2}\frac{g_{c_0,\theta_2}(2\pi\boldsymbol{J}/m)}{g_{c_0,\theta_1}(2\pi\boldsymbol{J}/m)}\right) \\ + \frac{\hat{I}_m^{\delta}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_2}g_{c_0,\theta_2}(2\pi\boldsymbol{J}/m)}\left(m^{\theta_1-\theta_2}\frac{g_{c_0,\theta_2}(2\pi\boldsymbol{J}/m)}{g_{c_0,\theta^1}(2\pi\boldsymbol{J}/m)} - 1\right) \end{aligned}$$

We also suppose that $Z(\boldsymbol{s})$ is a stationary Gaussian random field defined on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and replace $\hat{\theta}_m$ with $\hat{\theta}$ in this proof. The main idea of proving the theorem is looking for a contradiction to

$$P(R(c_0, \hat{\theta}_m) - R(c_0, \theta_0) \leq 0) = 1 \tag{5.51}$$

for any positive integer m, due to the definition of $\hat{\theta}_m$ when $\hat{\theta}_m$ does not converge to θ_0 in probability.

Suppose that $\hat{\theta}_m$ does not converge to θ_0 in probability. Then, there exist $\epsilon > 0$, $\delta > 0$ and M_1 such that for $m \geq M_1$,

$$P(|\hat{\theta}_m - \theta_0| > \epsilon) > \delta.$$

We define $\mathcal{D}_m = \{\omega \in \Omega : |\hat{\theta}_m - \theta_0| > \epsilon\}$.

Assume

$$d_m = \frac{\hat{I}_m^{\delta}(2\pi \boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)}, \tag{5.52}$$

By Theorem 3, we know $d_m \xrightarrow{p} 1$. Then, there exists $\{m_k\}$, a subsequence of $\{m\}$ such that d_{m_k} converges to one almost surely. To use Lemma 2, we need uniform convergence of d_{m_k} which is obtained by Egorov's Thoerem (Folland, 1999). By Egorov's Theorem, there exists $\mathcal{G}_\delta \subset \Omega$ such that d_{m_k} converge uniformly on $\mathcal{G}_\delta$ and $P(\mathcal{G}_\delta) > 1 - \delta/2$.

On the other hand, there exists a M_2, which does not depend on ω, such that for $m_k \geq$

M_2,

$$\left| m_k^{\hat\theta_{m_k}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m_k)}{g_{c_0,\hat\theta_{m_k}}(2\pi \boldsymbol{J}/m_k)} - 1 \right| \;>\; \frac{1}{2} \tag{5.53}$$

for all $\omega \in \mathcal{D}_{m_k}$, because of the uniform boundedness of $g_{c_0,\theta_0}/g_{c_0,\theta_1}$.

Let $\mathcal{H}_{m_k} = \mathcal{D}_{m_k} \cap\, \mathcal{G}_{m_k}$. Note that $P(\mathcal{H}_{m_k}) > \delta/2 > 0$ for $m_k \geq M_1$. Then, by Lemma 2 with $r = 1/2$, there exist $\delta_r > 0$ and M_r such that for $m_k \geq M_r$,

$$\begin{aligned}
& -\log\left(m_k^{\hat\theta_{m_k}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m_k)}{g_{c_0,\theta_1}(2\pi \boldsymbol{J}/m_k)} \right) \\
& \quad + \frac{\hat I^{\delta}_{m_k}(2\pi \boldsymbol{J}/m_k)}{m_k^{d-\theta_0} g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m_k)} \left(m_k^{\hat\theta_{m_k}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m_k)}{g_{c_0,\theta_1}(2\pi \boldsymbol{J}/m_k)} - 1 \right) \\
> \;& \delta_r
\end{aligned} \tag{5.54}$$

uniformly on $\mathcal{H}_{m_k}$. Note here that $M_r \geq \max\{M_1, M_2\}$.

Since $P(\mathcal{H}_{m_k}) > \delta/2 > 0$, it contradicts to (5.51) which completes the proof because (5.51) holds for any $m > 0$.

To show (2.14), it is enough to show that $m^{\hat\theta-\theta_0} \stackrel{p}{\longrightarrow} 1$ which is equivalent to show that

$$m^{\hat\theta-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\hat\theta}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} \quad \stackrel{p}{\longrightarrow} \quad 1 \tag{5.55}$$

$$\text{because} \quad \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\hat\theta}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} \quad \stackrel{p}{\longrightarrow} \quad 1. \tag{5.56}$$

(5.56) follows from the consistency of $\hat\theta$ and the continuity of $g_{c_0,\theta}$ shown in Lemma 5.1. To show (5.55), notice that we have

$$P\Big(R(c_0,\hat\theta) - R(c_0,\theta_0) \;\leq\; 0\Big) = 1 \tag{5.57}$$

for each $m > 0$ by the definition of $\hat{\theta}$.

Suppose that (5.55) does not hold. Then, there exists $r > 0$, $\delta > 0$ and M_1 such that

$$P\left(\left| m^{\hat{\theta}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)}{g_{c_0,\hat{\theta}}(2\pi \boldsymbol{J}/m)} - 1 \right| \;>\; r\right) > \delta$$

for all $m \geq M_1$. On the other hand, there exists $\{m_k\}$, a subsequence of $\{m\}$, such that $d_{m_k} \to 1$. Then, by Egorov's Thoerem, there exists $\Omega_\delta \subset \Omega$ such that $P(\Omega_\delta) > 1 - \delta/2$ and d_{m_k}. Now, define

$$\mathcal{D}_m = \left\{ \omega \;:\; \left| m^{\hat{\theta}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)}{g_{c_0,\hat{\theta}}(2\pi \boldsymbol{J}/m)} - 1 \right| \;>\; r \right\}. \tag{5.58}$$

Note that $P(\mathcal{D}_{m_k} \cap \Omega_\delta) \geq \delta/2 > 0$ for all $m_k \geq \max\{M_1, M_r\}$. Similarly to the proof of Lemma 2, for each $m_k \geq \max\{M_1, M_r\}$, there exists $\delta_r > 0$ such that $R(c_0, \hat{\theta}) - R(c_0, \theta_0) > \delta_r$ for all $\omega \in \mathcal{D}_{m_k} \cap \Omega_\delta$. This implies that

$$P(R(c_0, \hat{\theta}) - R(c_0, \theta_0) > \delta_r) \geq \delta/2$$

for each $m_k \geq \max\{M_1, M_r\}$. Note that δ_r does not depend on m_k which can be seen in Lemma 2. □

Proof of Theorem 10. Let $\dot{R} = \partial L/\partial\theta$ and $\ddot{R} = \partial^2 R/\partial\theta^2$. To show the asymptotic distribution of $\hat{\theta}$, we consider the Taylor expansion of $\dot{R}(c_0, \hat{\theta})$ around θ_0,

$$\dot{R}(c_0, \hat{\theta}) \;=\; \dot{R}(c_0, \theta_0) + \ddot{R}(c_0, \bar{\theta})(\hat{\theta} - \theta_0),$$

where $\bar{\theta}$ lies on the line segment between $\hat{\theta}$ and θ_0. Since $\dot{R}(c_0, \hat{\theta}) = 0$, we have

$$\log(m) m^{\eta}(\hat{\theta} - \theta_0) = -\log(m) m^{\eta} \left(\ddot{R}(c_0, \bar{\theta})\right)^{-1} \dot{R}(c_0, \theta_0).$$

Thus, it is enough to show

$$(\log(m))^{-1} m^{\eta} \dot{R}(c_0, \theta_0) \xrightarrow{d} \mathcal{N}\left(0, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right), \tag{5.59}$$

$$(\log(m))^{-2} \ddot{R}(c_0, \bar{\theta}) \xrightarrow{p} 1. \tag{5.60}$$

Since

$$\begin{aligned}
\dot{R}(c_0, \theta_0) &= -\log(m) + \frac{\dot{g}_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)}{g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)} - \hat{I}_m^{\tau}(2\pi(\boldsymbol{J})/m) \\
&\quad \times \frac{\left(-\log(m) m^{d-\theta_0} g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m) + m^{d-\theta_0} \dot{g}_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)\right)}{\left(m^{d-\theta_0} g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)\right)^2} \\
&= \log(m)\left(\frac{\hat{I}_m^{\tau}(2\pi \boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)} - 1\right) \\
&\quad + \left(1 - \frac{\hat{I}_m^{\tau}(2\pi \boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)}\right) \frac{\dot{g}_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)}{g_{c_0,\theta_0}(2\pi \boldsymbol{J}/m)},
\end{aligned}$$

we see that (5.59) follows from Lemma 4.

Next we prove (5.60). After some simplification, we have

$$\begin{aligned}
\ddot{R}(c_0,\bar{\theta}) &= (\log(m))^2 \frac{\hat{I}_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\bar{\theta}} g_{c_0,\bar{\theta}}(2\pi\boldsymbol{J}/m)} - 2\log(m) \frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)\dot{g}_{c_0,\bar{\theta}}(2\pi\boldsymbol{J}/m)}{m^{d-\bar{\theta}} g^2_{c_0,\bar{\theta}}(2\pi\boldsymbol{J}/m)} \\
&\quad + 2\frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)\dot{g}^2_{c_0,\bar{\theta}}(2\pi\boldsymbol{J}/m)}{m^{d-\bar{\theta}} g^3_{c_0,\bar{\theta}}(2\pi\boldsymbol{J}/m)} + \left(1 - \frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\bar{\theta}} g_{c_0,\bar{\theta}}(2\pi\boldsymbol{J}/m)}\right) \frac{\ddot{g}_{c_0,\bar{\theta}}(2\pi\boldsymbol{J}/m)}{g_{c_0,\bar{\theta}}(2\pi\boldsymbol{J}/m)} \\
&\quad - \frac{\dot{g}^2_{c_0,\bar{\theta}}(2\pi\boldsymbol{J}/m)}{g^2_{c_0,\bar{\theta}}(2\pi\boldsymbol{J}/m)} \\
&=: E_1 + E_2,
\end{aligned}$$

where E_1 is the first term with $(\log(m))^2$ and E_2 is the last four terms in the expression of $\ddot{R}(c_0,\bar{\theta})$.

First, we know that

$$(\log(m))^{-2} E_1 \xrightarrow{p} 1.$$

and

$$(\log(m))^{-1} E_2 = O_p(1)$$

from Theorem 3, (2.14) in Theorem 5 and Lemma 1.

□

Proof of Theorem 11. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be the probability space where a stationary Gaussian random field $Z(\boldsymbol{s})$ is defined. To emphasize dependence on m, we use $\hat{\theta}_m$ instead of $\hat{\theta}$ in this proof. Note that we have From the previous discussion,

$$R(c^*, \hat{\theta}_m) - R(c^*, \theta_0) = -\log\left(m^{\hat{\theta}_m - \theta_0} \frac{g_{*,\theta_0}(2\pi \boldsymbol{J}/m)}{g_{c^*,\hat{\theta}_m}(2\pi \boldsymbol{J}/m)} \right)$$
$$+ \frac{\hat{I}_m^{\delta}(2\pi \boldsymbol{J}/m)}{m^{d-\theta_0} g_{*,\theta_0}(2\pi \boldsymbol{J}/m)} \left(m^{\hat{\theta}_m - \theta_0} \frac{g_{c^*,\theta_0}(2\pi \boldsymbol{J}/m)}{g_{c^*,\hat{\theta}_m}(2\pi \boldsymbol{J}/m)} - 1 \right)$$

and

$$P(R(c^*, \hat{\theta}_m) - R(c^*, \theta_0) \leq 0) = 1, \forall m. \tag{5.61}$$

for any positive integer m, due to the definition of $\hat{\theta}_m$. We are going to prove the theorem by deriving a contradiction to (5.61) when $\hat{\theta}_m$ does not converge to θ_0 in probability.

Suppose that $\hat{\theta}_m$ does not converge to θ_0 in probability. Then, there exist $\epsilon > 0$, $\delta > 0$ and M_1 such that for $m \geq M_1$,

$$P(|\hat{\theta}_m - \theta_0| > \epsilon) > \delta.$$

We define $\mathcal{D}_m = \{\omega \in \Omega : |\hat{\theta}_m - \theta_0| > \epsilon\}$ and

$$d_m = \frac{\hat{I}_m^{\delta}(2\pi \boldsymbol{J}/m)}{m^{d-\theta_0} g_{c^*,\theta_0}(2\pi \boldsymbol{J}/m)}. \tag{5.62}$$

By Theorem 3, we have $d_m \xrightarrow{p} d = c_0/c^*$. Then, there exists $\{m_k\}$, a subsequence of $\{m\}$ such that d_{m_k} converges to one almost surely. To use Lemma 5, we need uniform convergence of d_{m_k} which is obtained by Egorov's Thoerem (Folland, 1999). By Egorov's

Theorem, there exists $\mathcal{G}_\delta \subset \Omega$ such that d_{m_k} converges uniformly on $\mathcal{G}_\delta$ and $P(\mathcal{G}_\delta) > 1-\delta/2$.

On the other hand, there exists a M_2, which does not depend on ω, such that for $m_k \geq M_2$,

$$m_k^{\hat{\theta}_{m_k}-\theta_0} \frac{g_{c^*,\theta_0}(2\pi \boldsymbol{J}/m_k)}{g_{c^*,\hat{\theta}_{m_k}}(2\pi \boldsymbol{J}/m_k)} \tag{5.63}$$

falls on the outside of (r_l, r_u) for all $\omega \in \mathcal{D}_{m_k}$, because of the uniform boundedness of $g_{c^*,\theta_0}/g_{c^*,\theta_1}$.

Let $\mathcal{H}_{m_k} = \mathcal{D}_{m_k} \cap \mathcal{G}_{m_k}$. Note that $P(\mathcal{H}_{m_k}) > \delta/2 > 0$ for $m_k \geq M_1$. Then, by Lemma 5, there exist $\delta_r > 0$ and M_r such that for $m_k \geq M_r$,

$$\begin{aligned} & R(c^*, \hat{\theta}_m) - R(c^*, \theta_0) && (5.64) \\ &= -\log\left(m_k^{\hat{\theta}_{m_k}-\theta_0} \frac{g_{c^*,\theta_0}(2\pi \boldsymbol{J}/m_k)}{g_{c^*,\theta_1}(2\pi \boldsymbol{J}/m_k)} \right) \\ &\quad + \frac{\hat{I}^{\delta}_{m_k}(2\pi \boldsymbol{J}/m_k)}{m_k^{d-\theta_0} g_{c^*,\theta_0}(2\pi \boldsymbol{J}/m_k)} \left(m_k^{\hat{\theta}_{m_k}-\theta_0} \frac{g_{c^*,\theta_0}(2\pi \boldsymbol{J}/m_k)}{g_{c^*,\theta_1}(2\pi \boldsymbol{J}/m_k)} - 1 \right) \\ &> \delta_r && (5.65) \end{aligned}$$

uniformly on $\mathcal{H}_{m_k}$. Note here that $M_r \geq \max\{M_1, M_2\}$. It contradicts to (5.61) which completes the proof. Here, we do not need $P(\cap_k \mathcal{H}_{m_k}) > 0$ since (5.61) should holds for any $m > 0$.

Assume

$$\hat{c} = \frac{1}{m^{d-\hat{\theta}}} \frac{\hat{I}^{\tau}_m(2\pi \boldsymbol{J}/m)}{g_{\hat{\theta}}(2\pi \boldsymbol{J}/m)}.$$

Then,

$$\hat{c} = c\, m^{\hat{\theta}-\theta} \frac{1}{m^{d-\theta}} \frac{\hat{I}_m^{\mathcal{T}}(2\pi \boldsymbol{J}/m)}{g_{\theta,c}(2\pi \boldsymbol{J}/m)} \frac{g_{\theta}(2\pi \boldsymbol{J}/m)}{g_{\hat{\theta}}(2\pi \boldsymbol{J}/m)}.$$

(2.29) is proven because of $\hat{\theta} - \theta = O_p(\log(m)^{-1})$, the boundedness of $g_{c,\theta}$ and Theorem 3. □

Lemma 7. *Consider a function $h_m(x) = -\log(x) + d_m(x-1)$, where d_m is positive and a function of a positive integer m. Also assume that $d_m \to d > 1$ (or < 1) as $m \to \infty$. Then, there exists some M such that for all $m \geq M$,*

$$h_m(x) > 0$$

for any $x > 1$ (or $x < 1$).

Proof. (For $d > 1$) As the previous discussion, we have known $h_m(x)$ is a convex function on $(0, \infty)$ for any positive integer m and minimized at $x = 1/d_m$ with $h_m(1/d_m) \leq 0$. Because $d_m \to d > 1$, there exists M such that $d_m > 1$ if $m \geq M$. There exists two intersection points between x-axis of $h_m(x)$ will be 1 and $u_m < 1$. Since the convexity of $h_m(x)$, when $m \geq M$, $h_m(x) < 0$ if $x > 1$ (or $x < 1$).

□

Proof of Theorem 12. Suppose that the result (i) of Theorem 12 does not hold, then there exists δ and M_1 such that

$$P\left(\theta_0 < \hat{\theta}_m\right) > \delta$$

for $m > M_1$.

By Theorem 3, we have $d_m \xrightarrow{p} d = c_0/c^*$. Then, there exists $\{m_k\}$, a subsequence of $\{m\}$ such that d_{m_k} converges to one almost surely. To use Lemma 7, we need uniform convergence of d_{m_k} which is obtained by Egorov's Thoerem (Folland, 1999) . By Egorov's Theorem, there exists $\mathcal{G}_\delta \subset \Omega$ such that d_{m_k} converges to uniformly on $\mathcal{G}_\delta$ and $P(\mathcal{G}_\delta) > 1 - \delta/2$. Assume $c^* < c_0$. Then, d_m converge $c_0/c^* > 1$. By the uniform convergence, there exists M such that $d_{m_k} > 1$ when $m_k > M$.

Assume that $\omega \in \Omega_{mk} = \{\omega : \theta_0 < \hat{\theta}_{m_k}\}$.

$$g_{c^*,\theta_0}(2\pi \boldsymbol{J}/m) > g_{c^*,\hat{\theta}_{m_k}}(2\pi \boldsymbol{J}/m)$$

because of the monotonicity of $g_{c^*,\theta}$ about θ.

$$m_k^{\hat{\theta}_{m_k} - \theta_0} \frac{g_{c^*,\theta_0}(2\pi \boldsymbol{J}/m_k)}{g_{c^*,\hat{\theta}_{m_k}}(2\pi \boldsymbol{J}/m_k)} > 1 \tag{5.66}$$

for all $\omega \in \Omega_{m_k} \bigcap \mathcal{G}_\delta$. Because $R(c^*, \hat{\theta}_{m_k}) - R(c^*, \theta_0) > 0$ on $\Omega_{m_k} \bigcap \mathcal{G}_\delta$ and $P(\Omega_{m_k} \bigcap \mathcal{G}_\delta) > 0$ when $m_k > M$, this contradicts to (5.61) will be found. The result (ii) will also be proven in a similar way.

□

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Adler, R. J. and Taylor, J. E.(2007) Random fields and geometry. *Springer Monographs in Mathematics Springer, New York.*

[2] Boissy, Y., Bhattacharyya, B.B., Li, X. and Richardson, G.D. (2005) Parameter estimates for fractional autoregressive spatial processes. *Ann. Statist.* **33**, 2553-2567.

[3] Chan, G., Hall, P., and Poskitt, D.S. (1995). Periodogram-based estimators of fractal properties. *Ann. Statist.* **23**, 1684-1711.

[4] Chan, G. and Wood, A. T. A. (2000). Increment-based estimators of fractal dimension for two-dimensional surface data. *Statist. Sinica* **10**, 343-376.

[5] Chan, G. and Wood, A. T. A. (2004). Estimation of fractal dimension for a class of non-Gaussian stationary processes and fields. *Ann. Statist.* **32**, 1222-1260.

[6] Chen, H.-S., Simpson, D.G. and Ying, Z. (2000). Infill asymptotics for a stochastic process model with measurement error. *Statist. Sinica* **10**, 141-156.

[7] Constantine, A. G. , Hall, P. (1994). Characterizing surface smoothness via estimation of effective fractal dimension. *J. R. Statist. Soc. B.* **56**, 97-113.

[8] Cressie, N. A. C (1993). Statistics for spatial data (rev. ed.). John Wiley, New York.

[9] Du, J. (2009a). Asymptotic and computational methods in spatial statistics. (Ph. D. Thesis). *Michigan State University*, 1-111.

[10] Du, J., Zhang, H. and Mandrekar, V. S. (2009). Fixed-domain asymptotic properties of tapered maximum likelihood estimators. *Ann. Statist.* **100**, 993-1028.

[11] Folland, Gerald B. Real Analysis. *Wiley-Interscience Publication., New York, 1999.*

[12] Furrer, R., Genton, M. G. and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets. Journal of Computational and Graphical Statistics. *Journal of Computational and Graphical Statistics.* **15**, 502-523.

[13] Guo, H., Lim, C. and Meerschaert, M. M (2009). Local Whittle estimator for anisotropic random fields. *J. Multivariate Anal.* **100**, 993-1028.

[14] Guyon, X. (1982). Parameter estimation for a stationary process on a d-dimensional lattice. *Biometrika.* **69**, 95-105.

[15] Guyon, X. (1995). Random fields on a network : modeling, statistics, and applications. Springer-Verlag, New York.

[16] Ibragimov, I. A. and Rozanov, Y. A. (1978). Gaussian random processes. Springer, New York. MR0543837

[17] Kaufman, C., Schervish, M. and Nychka, D. (2008). Covariance tapering for likelihood-based estimation in large spatial datasets. *J. Amer. Statist. Assoc.* **103**, 1545-1555.

[18] Loh, W.-L. (2005). Fixed-domain asymptotics for a subclass of Matern-type Gaussian random fields. *Ann. Statist.* **33**, 2344-2394.

[19] Lim, C. and Stein, M. L. (2008). Properties of spatial cross-periodograms using fixed-domain asymptotics. *J. Multivariate Anal.* **99**, 1962-1984.

[20] Mardia, K. V. and Marshall, R. J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika.* **71**, 135-146.

[21] Robinson, P. M. (1995). Gaussian semiparametric estimation of long range dependence. *Ann. Statist.* **23**, 1630-1661.

[22] Stein, M. L. (1988). Asymptotically efficient prediction of a random field with a misspec- ified covariance function. *Ann. Statist.* **16**, 55-63.

[23] Stein, M. L. (1990). Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure. *Ann. Statist.* **18**, 850-872.

[24] STEIN, M. L. (1990). Bounds on the efficiency of linear predictions using an incorrect covariance function. *Ann. Statist.* **18**, 1116-1138.

[25] STEIN, M. L. (1993). A simple condition for asymptotic optimality of linear predictions of random fields. *Statistics and Probability Letters.* **17**, 399-404.

[26] STEIN, M. L. (1995). Fixed-domain asymptotics for spatial periodograms. *J. Amer. Statist. Assoc.* **432**, 1962-1984.

[27] STEIN, M. L. (1999). Interpolation of spatial Data. *Springer, New York.*

[28] WHITTLE, P. (1954). On stationary processes in the plane. *Biometrika.* **41**, 434-449.

[29] XUE, Y. AND XIAO, Y. (2010). Fractal and smoothness properties of space-time Gaussian models. To appear in Frontiers Math.

[30] YADRENKO, M. (1983). Spectral Theory of Random Fields, New York: Optimization Software.

[31] YING, Z. (1991). Asymptotic properties of a maximum likelihood estimator with data from a Gaussian process. *J. Multivariate Anal.* **36**, 280-296.

[32] YING, Z. (1993). Maximum likelihood estimation of parameters under a spatial sampling scheme. *Ann. Statist.* **21**, 1567-1590.

[33] ZHANG, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Statist. Assoc.* **465**, 250-261.

[34] ZHANG, H. AND ZIMMERMAN, D. L. (2005). Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika.* **92**, 921-936.