Technical Report RM 692, Department of Statistics and Probability, Michigan State University

Tail Estimation of the Spectral Density under Fixed-Domain Asymptotics

Wei-Ying Wu, Chae Young Lim [*] and Yimin Xiao[†]

Wei-Ying Wu,
Department of Statistics & Probability
Michigan State University,
East Lansing, MI 48824
e-mail: wu@stt.msu.edu

Chae Young Lim,
Department of Statistics & Probability
Michigan State University,
East Lansing, MI 48824
e-mail: lim@stt.msu.edu

Yimin Xiao,
Department of Statistics & Probability,
Michigan State University,
East Lansing, MI 48824
e-mail: xiaoyimi@stt.msu.edu

Abstract: Consider a stationary Gaussian random field on $\mathbb{R}^d$ with the spectral density $f(\boldsymbol{\lambda})$ that satisfies $f(\boldsymbol{\lambda}) \sim c\,|H\boldsymbol{\lambda}|^{-\theta}$ as $|\boldsymbol{\lambda}| \to \infty$ for some nonsingular matrix H. The parameters c and θ control the tail behavior of the spectral density. c is related to a microergodic parameter and θ is related to a fractal index. For data observed on a grid, we propose estimators of c and θ by minimizing an objective function, which can be viewed as a weighted local Whittle likelihood and study their asymptotic properties under fixed-domain asymptotics.

AMS 2000 subject classifications: Primary 62M30; secondary 62H11.
Keywords and phrases: fixed-domain asymptotics, fractal index, fractal dimension, Gaussian random fields, infill asymptotics, microergodic, Whittle likelihood .

1. Introduction

With recent advances in technology, we are facing enormous amount of data sets. When data sets are observed on a regular grid, spectral analysis is popular

*Supported by MSU grant 08-IRGP-1532
†Supported in part by NSF grant DMS-1006903

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.

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 are 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, 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 process using a local Whittle likelihood method in which a parametric form near zero frequency is only assumed. This is an extension of Robinson (1995) for time series.

For spatial data, however, it is often natural to assume that the data are observed on a bounded domain of interest. More observations on the bounded domain implies that the distance between observations, ϕ, is decreasing 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)]). The classical asymptotic framework when the sampling distance is fixed (i.e. ϕ is fixed) is called increasing-domain asymptotics to differentiate from fixed-domain asymptotics.

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), Zhang (2004)]. For example, Zhang (2004) showed not all parameters in the Matérn covariance function of a stationary Gaussian random field on $\mathbb{R}^d$ are consistently estimable when d is smaller than or equal to 3, while a reparameterized quantity of variance and scale parameters can be estimated consistently by the maximum likelihood method. On the other hand, the maximum likelihood estimators (MLEs) of variance and scale parameters for a stationary Gaussian process under increasing-domain asymptotics 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), Anderes (2010)]. The microergodicity of functions of parameters determines the equivalence of probability measures and a microergodic parameter is the quantity that affects asymptotic mean squared prediction error under fixed-domain asymptotics. [Stein (1990a, 1990b, 1999)].

Although there have been more asymptotic results available recently under fixed-domain asymptotics, it is still a 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. Moreover, these asymptotic results are established in the spatial domain. Asymptotic work 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 this paper, we propose estimators of parameters that control the tail behavior of the spectral density for a stationary Gaussian random field when the data are 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

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

where $|\cdot|$ is a usual Euclidean norm, H is a nonsingular matrix and $\theta > d$ to ensure integrability of f. That is, we assume a power law for the tail behavior of the spectral density and do not assume any specific parametric form of the spectral density. Also, this assumption allows a wide range of anisotropic spectral densities by introducing H. The proposed estimators are obtained by minimizing an objective function that can be viewed as a weighted local 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 a Whittle likelihood over Fourier frequencies near zero since the long-range dependence parameter controls 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.

We establish consistency and asymptotic normality of an estimator of c and an estimator of θ, respectively, when the other parameter is known. The parameter c is related to a microergodic parameter. For example, consider the Matérn spectral density given as

$$f(\boldsymbol{\lambda}) = \frac{\sigma^2 \alpha^{2\nu}}{\pi^{d/2}(\alpha^2 + |\boldsymbol{\lambda}|^2)^{\nu+d/2}}. \tag{1.2}$$

The 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 a tapered MLE of $\sigma^2\alpha^{2\nu}$ when ν is known, α is fixed and $d = 1$ for a stationary Gaussian random field. 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). For example, for a stationary Gaussian random field, $Z(\boldsymbol{s})$, $\boldsymbol{s} \in \mathbb{R}^d$, suppose that its covariance function $C(\boldsymbol{t})$ satisfies

$$C(\boldsymbol{t}) \ \sim \ C(\boldsymbol{0}) - k|\boldsymbol{t}|^{\alpha} \quad \text{as} \ \ |\boldsymbol{t}| \ \to \ 0 \tag{1.3}$$

for some k and $0 < \alpha \leq 2$. In this case, α is the fractal index that governs the roughness of sample paths of a random field and the fractal dimension D becomes $D = d + (1 - \alpha/2)$. This follows from Adler (1981, Chapter 8) or Theorem 5.1 in Xue and Xiao (2010) where more general results are proven for anisotropic Gaussian random fields. When $\alpha = 2$, it is possible that the sample function may be differentiable. This can be determined by the smoothness of $C(\boldsymbol{t})$ or in terms of the spectral measure of $Z(\boldsymbol{s})$ (see, e.g., Adler and Taylor (2007) or Xue and Xiao (2010) for further information). By an Abelian type theorem, (1.3) holds if the corresponding spectral density satisfies

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

so that $\theta \equiv \alpha + d$ in our settings. There is a number of references that construct estimators based on fractal properties of processes. For example, 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 for the fractal dimension of a stationary Gaussian random field on $\mathbb{R}^d$ with $d = 1$ or 2. Compared to these works in the spatial domain, the work by Chan et al. (1995) to estimate the fractal dimension is done in the spectral domain.

In Section 2, we explain our settings and assumptions, and in Section 3 introduce our estimators and state the main theorems for the asymptotic properties of the proposed estimators. Section 4 discusses some issues related to our approach and possible extension of the current work. All proofs are given in Appendices.

2. Preliminaries

In this paper, 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 $\boldsymbol{\lambda} \in (-\pi, \pi]^d$. $\bar{f}_\phi(\boldsymbol{\lambda})$ has 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

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

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

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

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

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

for $\boldsymbol{\lambda} \neq \boldsymbol{0}$. Define for $\boldsymbol{\lambda} \in (-\pi, \pi]^d$,

$$g_{c,\theta}(\boldsymbol{\lambda}) = c \left\{ \sum_{j=1}^{d} 4\sin^2\left(\frac{\lambda_j}{2}\right) \right\}^{2\tau} \times \sum_{\boldsymbol{Q} \in \mathbb{Z}^d} |H(\boldsymbol{\lambda} + 2\pi\boldsymbol{Q})|^{-\theta} \, \mathbb{I}_{(\boldsymbol{\lambda} \neq \boldsymbol{0})}, \tag{2.2}$$

where $\mathbb{I}_A = 1$ if A is true and zero, otherwise. 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 in Section 3.

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 fixed 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, periodogram is defined by

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

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

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

We consider the periodogram only at Fourier frequencies, $2\pi m^{-1}\boldsymbol{J}$ for $\boldsymbol{J} \in \mathcal{T}_m \equiv \{-\lfloor (m-1)/2 \rfloor, \cdots, m - \lfloor m/2 \rfloor\}^d$, 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

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

for a symmetric continuous function Λ on $\mathbb{R}^d$ that satisfies $\Lambda(\boldsymbol{s}) \geq 0$ and $\Lambda(\boldsymbol{0}) > 0$ 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. *For a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbb{R}^d$, the spectral density $f(\boldsymbol{\lambda})$ satisfies*

(A) $f(\boldsymbol{\lambda}) \sim c\,|H\boldsymbol{\lambda}|^{-\theta}$ *as* $|\boldsymbol{\lambda}| \to \infty$, *for some* $c > 0$, $\theta > d$ *and a nonsingular matrix* H,

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

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

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

3. Main Results

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 the spectral density f is twice differentiable and satisfies (2.4) for all $\boldsymbol{\lambda} \in \mathbb{R}^d$, which implies that the spectral density $f(\boldsymbol{\lambda})$ behaves like $(1+|\boldsymbol{\lambda}|)^{-\theta}$ for all $\boldsymbol{\lambda}$. This is much stronger condition than Assumption 1. This stronger condition allows to find asymptotic bounds for the expectation, variance and covariance of a spatial periodogram at Fourier frequency $2\pi\boldsymbol{J}/m$ by m and $\boldsymbol{J}$ for $\|\boldsymbol{J}\| \neq 0$. Consistency and asymptotic normality of a smoothed spatial periodogram at Fourier frequency $2\pi\boldsymbol{J}/m$, however, are only available when $\lim 2\pi\boldsymbol{J}/m = \boldsymbol{\mu} \neq \mathbf{0}$, that is, $\boldsymbol{J}$ should not be close to zero asymptotically. Since we make use of asymptotic properties of a smoothed spatial periodogram at such Fourier frequency, we extend those results in Stein (1995) and Lim and Stein (2008) under Assumption 1. We focus on only a smoothed spatial periodogram in the following theorem, but results for a smoothed spatial cross-periodogram can be shown similarly. Throughout the paper, let $\xrightarrow{p}$ denote the convergence in probability and $\xrightarrow{d}$ denote the convergence in distribution.

Theorem 3.1. *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 $0 < \|\boldsymbol{\mu}\| < \pi$. Let $\eta = d(1-\gamma)/2$. Then, we have*

$$\frac{\hat{I}_m^\tau\,(2\pi\boldsymbol{J}/m)}{\bar{f}_\phi^{\,\tau}\,(2\pi\boldsymbol{J}/m)} \xrightarrow{p} 1 \tag{3.1}$$

and

$$m^\eta\left(m^{-(d-\theta)}\hat{I}_m^\tau\,(2\pi\boldsymbol{J}/m) - g_{c,\theta}(\boldsymbol{\mu})\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d g_{c,\theta}^2(\boldsymbol{\mu})\right), \tag{3.2}$$

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

Remark 3.1. *$g_{c,\theta}$ is integrable under $4\tau > \theta - d$ which is satisfied by the condition $4\tau > \theta - 1$. $4\tau > \theta - 1$ is necessary to show $E\left(\hat{I}_m^\tau\,(2\pi\boldsymbol{J}/m)\,/\,\bar{f}_\phi^{\,\tau}\,(2\pi\boldsymbol{J}/m)\right) \to 1$ and the condition $\max\{(d-2)/d, 0\} < \gamma < 1$ is needed to show $Var\left(\hat{I}_m^\tau\,(2\pi\boldsymbol{J}/m)\,/\,\bar{f}_\phi^{\,\tau}\,(2\pi\boldsymbol{J}/m)\right) \to 0$ so that (3.1) can be shown.*

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

$$L(c,\theta) = \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\left\{\log\left(m^{d-\theta}g_{c,\theta}\left(2\pi(\boldsymbol{J}+\boldsymbol{K})/m\right)\right) + \frac{1}{m^{d-\theta}}\frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\right\}, \tag{3.3}$$

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 local Whittle likelihood. If 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 $\{Y_\delta^\tau(\boldsymbol{J}), \boldsymbol{J} \in T_m\}$ 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 $Y_\delta^\tau(\boldsymbol{J})$ after scaling 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 $(\pi/2)\mathbf{1}_d$.

For the estimation of c, we minimize $L(c, \theta)$ with a known θ. 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 $\partial L(c, \theta_0)/\partial c = 0$:

$$\hat{c} = \sum_{\boldsymbol{K} \in \mathcal{T}_m} W_h(\boldsymbol{K}) \frac{1}{m^{d-\theta_0}} \frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_0(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}, \tag{3.4}$$

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

Theorem 3.2. *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. Let $\eta = d(1-\gamma)/2$. Then, for $\hat{c}$ given in (3.4), we have*

$$\hat{c} \xrightarrow{p} c, \tag{3.5}$$

and

$$m^\eta(\hat{c} - c) \xrightarrow{d} \mathcal{N}\left(0\,,\, c^2 \frac{\Lambda_2}{\Lambda_1^2} \left(\frac{2\pi}{\mathfrak{C}}\right)^d\right), \tag{3.6}$$

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

Remark 3.2. *We can prove Theorem 3.2 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 have chosen for simplicity of the proof.*

Remark 3.3. *Without much difficulty, Theorem 3.2 can be also proved when we replace θ_0 with a consistent estimator $\hat{\theta}$ as long as the estimator $\hat{\theta}$ satisfies $\hat{\theta} - \theta_0 = o_p((\log(m))^{-1})$ which implies $m^{\hat{\theta}-\theta_0} \xrightarrow{p} 1$.*

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 (1.2) with $d = 1$, Du et al. (2009) showed that for any fixed α_1 with known ν, the MLE of σ^2 satisfies

$$n^{1/2}(\hat{\sigma}^2\alpha_1^{2\nu} - \sigma_0^2\alpha_0^{2\nu}) \ \xrightarrow{d}\ \mathcal{N}\left(0\,,\, 2(\sigma_0^2\alpha_0^{2\nu})^2\right), \tag{3.7}$$

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_ϕ^τ which is the τ times differenced lattice process of $Z(\boldsymbol{s})$ so that $m = n - 2\tau$ for the simple differencing and $m = n - 4\tau$ for the Laplace differencing. Since $\pi^{1/2}c = \sigma^2\alpha^{2\nu}$ for $d = 1$, we have the same asymptotic variance as in (3.7). 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).

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

$$\hat{\theta} = \arg\min_{\theta\in\Theta} L(c_0, \theta), \tag{3.8}$$

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 3.3. *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 $d/(d+2) < \gamma < 1$. 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 (3.8), we have*

$$\hat{\theta} \ \xrightarrow{p}\ \theta. \tag{3.9}$$

In addition,

$$\hat{\theta} - \theta \ = o_p((\log m)^{-1}). \tag{3.10}$$

Remark 3.4. *The consistency of $\hat{\theta}$ is not enough to prove 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 (3.10).*

From Theorem 3.3, we can now show the following Theorem for the asymptotic distribution of $\hat{\theta}$.

Theorem 3.4. *Under the conditions of Theorem 3.3, 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 3.5. *Note that we have a different convergence rate for $\hat{\theta}$ compared to the convergence rate for $\hat{c}$ given in Theorem 3.2. The additional term $\log(m)$ is from the fact that θ is in the exponent of m in the expression of $L(c, \theta)$.*

4. Discussion

We proposed estimators of 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 (3.3). As mentioned in Section 3, this objective function 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 $\bar{f}_{\phi}^{\tau}(\boldsymbol{\lambda})$ and remove $W_h(\boldsymbol{K})$ in the expression given in (3.3), it can be thought of a Whittle approximation to the likelihood of $\boldsymbol{Y}_{\phi}^{\tau}(\boldsymbol{J})$. 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 a Whittle approximation to the likelihood under increasing-domain asymptotics since $\boldsymbol{Y}_{\phi}^{\tau}(\boldsymbol{J})$ is a lattice process. However, the spectral density $\bar{f}_{\phi}^{\tau}(\boldsymbol{\lambda})$ of $\boldsymbol{Y}_{\phi}^{\tau}(\boldsymbol{J})$ converges to zero as $\phi \to 0$, which requires a different approach and further investigation is needed.

The weights in (3.3) 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 $\hat{I}_m^{\tau}$. Thus, we could find the optimal bandwidth that minimizes the mean squared error of $\hat{I}_m^{\tau}$. However, finding the mean squared error of $\hat{I}_m^{\tau}$ needs explicit first order asymptotic expressions of the bias and variance of $\hat{I}_m^{\tau}(\boldsymbol{\lambda})$, which are not yet available. 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 but worthwhile investigating further.

The Assumption 1 (A) implies that the spectral density is regularly varying at infinity with exponent $-\theta$. Together with a smoothness condition (Assumption 1 (B)), it satisfies assumptions (A1) and (A2) in Stein (2002), which includes slowly varying tail behavior. Assumptions (A1) and (A2) in Stein (2002) guarantee that there is a screening effect, that is, one can get a nearly optimal predictor at a location $\boldsymbol{s}$ based on the observations nearest to $\boldsymbol{s}$ (see Stein (2002) for further details). Then, our result that one can estimate tail behavior using only local information can be seen as a kind of analogue to a screening effect.

Appendix A: The properties of $g_{c,\theta}(\boldsymbol{\lambda})$

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

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

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

For a function $g_{c,\theta}(\boldsymbol{\lambda})$, let ∇g denote 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 $\ddot{g} = \partial^2 g_{c,\theta}(\boldsymbol{\lambda})/\partial\theta^2$.

Let$\mathcal{A}_\rho = [-\pi,\pi]^d \setminus (-\rho,\rho)^d$ for a fixed ρ that satisfies $0 < \rho < 1$. Since we assume that the parameter space Θ is a closed interval in Section 3, let $\Theta = [\theta_L, \theta_U]$ and $\theta_L > d$. Although Lemma A.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 $L(c,\theta)$ are contained in $\mathcal{A}_\rho$.

Lemma A.1. *The following properties hold for $g_{c,\theta}(\boldsymbol{\lambda})$.*

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

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

$$0 < K_L \leq g_{c,\theta}(\boldsymbol{\lambda}) \leq K_U < \infty. \tag{A.1}$$

(c) There exist K_L and K_U such that for all $\boldsymbol{\lambda} \in \mathcal{A}_\rho$ and all $\theta_1, \theta_2 \in \Theta$,

$$0 < K_L \leq g_{c,\theta_1}(\boldsymbol{\lambda})/g_{c,\theta_2}(\boldsymbol{\lambda}) \leq K_U < \infty. \tag{A.2}$$

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

Proof. We prove the Lemma when H is an identity matrix for simplicity. The results with a general nonsingular matrix H can be followed without much difficulty.

To show (a), it is enough to show 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 $g_{1,\theta}(\boldsymbol{\lambda})$. Thus, continuity of $g_{c,\theta}(\boldsymbol{\lambda})$ follows from the continuity of $|\boldsymbol{\lambda} + 2\pi\boldsymbol{Q}|^{-\theta}$.

To prove the remaining parts, 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

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

and

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

where the last inequality follows from

$$\begin{aligned}\sum_{\boldsymbol{Q}\in\mathbb{Z}^d\setminus\boldsymbol{0}} |\boldsymbol{\lambda} + 2\pi\boldsymbol{Q}|^{-\theta_L} &\le \int_{|\boldsymbol{y}|\ge 1} |\boldsymbol{\lambda} + 2\pi\boldsymbol{y}|^{-\theta_L} d\boldsymbol{y} \\ &\le \int_{|\boldsymbol{z}|\ge\epsilon} (2\pi)^d |\boldsymbol{z}|^{-\theta_L} d\boldsymbol{z} \\ &= \int_{x\,\ge\,\epsilon} (2\pi)^d x^{d-1} x^{-\theta_L} dx \\ &= (2\pi)^d \epsilon^{d-\theta_L}/(\theta_L - d),\end{aligned} \tag{A.3}$$

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

$$0 \ < \ k_L \ \le \sum_{\boldsymbol{Q}\in\mathbb{Z}^d} |\boldsymbol{\lambda} + 2\pi\boldsymbol{Q}|^{-\theta} \ \le \ k_U < \infty, \tag{A.4}$$

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

Then, (b) follows from (A.4),

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

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

(c) 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 (A.4).

For (d), we have

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

for some constant $K > 0$ and k_U given in (A.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 $\beta > 0$, $|\log x| \le K x^{\beta}$ 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 (A.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 (b). To show uniform boundedness of $\nabla(\dot{g}/g)$, consider

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

Since denominators in the expression of $\partial(\dot{g}/g)/\partial\lambda_i$ have uniform lower bounds as shown in (A.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 $\dot{g}$.

□

Appendix B: Proofs of Theorems in Section 3

Proof of Theorem 3.1. If $f(\boldsymbol{\lambda})$ satisfies (2.4) for all $\boldsymbol{\lambda}$, (3.1) and (3.2) hold by results in Stein (1995) and Lim and Stein (2008). To prove (3.1) and (3.2) 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 $k(\boldsymbol{\lambda})\sim c|H\boldsymbol{\lambda}|^{-\theta}$ 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

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

Note that

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

(3.1) and (3.2) 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

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

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

for some $\beta_1, \beta_2 > 0$. Once (B.1) and (B.2) 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 (B.1) and (B.2) for $k(\boldsymbol{\lambda})$. Then, (B.1) and (B.2) for $f(\boldsymbol{\lambda})$ follow from

$$\left|a_{m,\phi}^{f,\tau}(\boldsymbol{J}, \pm\boldsymbol{J}) - a_{m,\phi}^{k,\tau}(\boldsymbol{J}, \pm\boldsymbol{J})\right| = O(m^{-d-4\tau}), \tag{B.3}$$

for $\boldsymbol{J}$ that satisfies $\|\boldsymbol{J}\| \asymp m$ and $2\boldsymbol{J}/m \notin \mathbb{Z}^d$. (B.3) holds since

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

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

□

Proof of Theorem 3.2. 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

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

which can be rewritten as

$$\frac{c\,\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{U}})/m)} \le \hat{c} \le \frac{c\,\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{L}})/m)} \tag{B.4}$$

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 3.1. Thus, it follows that $\hat{c}$ converges to c in probability.

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

$$m^\eta\left(\frac{\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}} - g_{c,\theta_0}((\pi/2)\mathbf{1}_d)\right) \xrightarrow{d} \mathcal{N}\left(0\,,\ \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d g_{c,\theta_0}^2((\pi/2)\mathbf{1}_d)\right) \tag{B.5}$$

from Theorem 3.1 and

$$m^\eta\left(g_{c,\theta_0}\left(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{E}})/m\right) - g_{c,\theta_0}((\pi/2)\mathbf{1}_d)\right) \longrightarrow 0, \tag{B.6}$$

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, (3.6) follows from (B.5) and (B.6).

□

To prove Theorem 3.3, we consider following lemmas.

Lemma B.1. *Consider a function $h_m(x) = -\log(x) + d_m(x-1)$, where d_m is a positive 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 $M_r>0$ such that for all $m \ge M_r$,*

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

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

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 $h_m(1/d_m) \le 0$. 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 $m \ge M_r$, we have $|1/d_m - 1| \le r$ 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

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

□

Lemma B.2. *For a positive integer m and $\theta \in \Theta$, we have*

$$L(c_0,\theta) - L(c_0,\theta_0) \ge A_m + B_m + C_m,$$

where

$$A_m = -\log\left(m^{\theta-\theta_0}\frac{g_{c_0,\theta_0}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{S}_m)}{m}\right)}{g_{c_0,\theta}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{S}_m)}{m}\right)}\right) + \frac{\hat{I}_m^\tau\left(\frac{2\pi\boldsymbol{J}}{m}\right)}{m^{d-\theta_0}g_{c_0,\theta_0}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{K}_M)}{m}\right)} \times\left(m^{\theta-\theta_0}\frac{g_{c_0,\theta_0}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{S}_m)}{m}\right)}{g_{c_0,\theta}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{S}_m)}{m}\right)} - 1\right), \tag{B.7}$$

$$B_m = \log\left(\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}\frac{g_{c_0,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}{g_{c_0,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}\right), \tag{B.8}$$

$$C_m = \frac{\hat{I}_m^\tau(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}\left(1 - \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)}\right) \tag{B.9}$$

In (B.7)-(B.9), $\boldsymbol{K}_M$, $\boldsymbol{K}_m$, $\boldsymbol{S}_M$ and $\boldsymbol{S}_m$ are defined as

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

Furthermore,

$$\sup_{\theta\in\Theta}|B_m| = o(1), \tag{B.10}$$

$$C_m = o_p(1), \tag{B.11}$$

where (B.11) holds under the conditions of Theorem 3.3.

Proof. From the expression of $L(c,\theta)$ given in (3.3), we have

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

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

which is $A_m + B_m + C_m$ given in (B.7)-(B.9).

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 uniformly on Θ, because $h\to 0$.

The continuity of $g_{c_0,\theta}$ in Lemma A.1 implies that

$$\log\left(\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}\frac{g_{c_0,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{S}_M)/m)}{g_{c_0,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}\right) \longrightarrow 0 \quad \text{(B.12)}$$

holds for θ 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 3.1 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

$$1-\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)} \quad \to \quad 0, \qquad \text{(B.13)}$$

C_m converges to zero in probability.

□

Proof of Theorem 3.3. Let $(\Omega, \mathcal{F}, P)$ 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

$$P(L(c_0,\hat{\theta}_m) - L(c_0,\theta_0) \le 0) = 1 \qquad \text{(B.14)}$$

for each positive integer m by the definition of $\hat{\theta}_m$. We are going to prove consistency of $\hat{\theta}_m$ by deriving a contradiction to (B.14) 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 $\mathcal{D}_m = \{\omega \in \Omega : |\hat{\theta}_m(\omega) - \theta_0| > \epsilon\}$. By Lemma B.2, we have

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

where A_m, B_m and C_m are given in (B.7)-(B.9) with $\theta = \hat{\theta}_m(\omega)$, $\omega \in \mathcal{D}_m$. We are going to show that there exist $\{m_k\}$, a subsequence of $\{m\}$ and a subset of $\mathcal{D}_{m_k}$ on which $A_{m_k} + B_{m_k} + C_{m_k}$ is bounded away from zero for large enough m_k.

Note that

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

where $h_m(\cdot)$ is defined in Lemma B.1 with

$$d_m = \frac{\hat{I}_m^{\tau}(2\pi \boldsymbol{J}/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{K}_M)/m)}, \tag{B.15}$$

where $\boldsymbol{K}_M$ is defined in Lemma B.2.

By Theorem 3.1 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 (B.13) in the proof of Lemma B.2, almost sure convergence of d_{m_k} implies that C_{m_k} defined in (B.9) converges to zero almost surely. To use Lemma B.1, we need uniform convergence of d_{m_k}. By Egorov's Theorem (Folland, 1999), there exists $\mathcal{G}_\delta \subset \Omega$ such that d_{m_k} and C_{m_k} converge uniformly on $\mathcal{G}_\delta$ and $P(\mathcal{G}_\delta) > 1 - \delta/2$. Let $\mathcal{H}_{m_k} = \mathcal{D}_{m_k} \cap \mathcal{G}_\delta$. Note that $P(\mathcal{H}_{m_k}) > \delta/2 > 0$ for $m_k \geq M_1$.

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

$$\left| m_k^{\hat{\theta}_{m_k} - \theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)}{g_{c_0,\hat{\theta}_{m_k}}(2\pi(\boldsymbol{J} + \boldsymbol{S}_{m_k})/m_k)} - 1 \right| > \frac{1}{2} \tag{B.16}$$

for all $\omega \in \mathcal{D}_{m_k}$, because of the uniform boundedness of $g_{c_0,\theta_0}/g_{c_0,\theta}$. Then, by Lemma B.1 with $r = 1/2$, there exist $\delta_r > 0$ and $M_r \geq \max\{M_1, M_2\}$ such that for $m_k \geq M_r$,

$$
\begin{aligned}
A_{m_k} &= -\log\left(m_k^{\hat{\theta}_{m_k}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_{m_k})/m_k)}{g_{c_0,\hat{\theta}_{m_k}}(2\pi(\boldsymbol{J}+\boldsymbol{S}_{m_k})/m_k)}\right) \\
&\quad + \frac{\hat{I}^{\tau}_{m_k}(2\pi\boldsymbol{J}/m_k)}{m_k^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m_k)} \\
&\qquad\qquad \times \left(m_k^{\hat{\theta}_{m_k}-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_{m_k})/m_k)}{g_{c_0,\hat{\theta}_{m_k}}(2\pi(\boldsymbol{J}+\boldsymbol{S}_{m_k})/m_k)} - 1\right) \\
&> \delta_r
\end{aligned}
\tag{B.17}
$$

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

By the uniform convergence of $|B_m|$ on Θ shown in Lemma B.2, there exists a M_3 such that for $m_k \geq M_3$,

$$
|\, B_{m_k} \,| \quad < \quad \frac{\delta_r}{4} \tag{B.18}
$$

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

$$
|\, C_{m_k} \,| \quad < \quad \frac{\delta_r}{4} \tag{B.19}
$$

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

$$
L(c_0, \hat{\theta}_{m_k}) - L(c_0, \theta_0) > \frac{\delta_r}{2} \tag{B.20}
$$

on $\mathcal{H}_{m_k}$. Since $P(\mathcal{H}_{m_k}) > \delta/2 > 0$, it contradicts to (B.14) which completes the proof. Here, we do not need $P(\cap_k \mathcal{H}_{m_k}) > 0$ since (B.14) should holds for any $m > 0$.

To show the convergence rate of $\hat{\theta}_m$ given in (3.10), it is enough to show that $m^{\hat{\theta}_m - \theta_0} \xrightarrow{p} 1$ which is equivalent to show that

$$
\frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\hat{\theta}_m}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} \quad \xrightarrow{p} \quad 1, \tag{B.21}
$$

$$
m^{\hat{\theta}_m-\theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\hat{\theta}_m}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} \quad \xrightarrow{p} \quad 1. \tag{B.22}
$$

(B.21) follows from the consistency of $\hat{\theta}_m$ and the continuity of $g_{c_0,\theta}$ shown in Lemma A.1. To show (B.22), we consider a similar argument to show consistency of $\hat{\theta}$. For simplicity, we reset notations such as r, δ, δ_r, M and $\mathcal{D}_m$, etc. used in the proof of consistency.

Suppose that (B.22) does not hold. Then, there exists $r > 0$, $\delta > 0$ and M_1 such that

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

for all $m \geq M_1$. On the other hand, there exists $\{m_k\}$, a subsequence of $\{m\}$, such that $d_{m_k} \ \to \ 1$, $B_{m_k} \ \to \ 0$ and $C_{m_k} \ \to \ 0$ almost surely, where d_m is given in (B.15), B_m and C_m are given in (B.8) and (B.9) with $\theta = \hat{\theta}_m(\omega)$. 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_k} and C_{m_k} are uniformly convergent on Ω_δ. As in Lemma B.1, for a_{m_k}, a nonzero solution of $h_{m_k}(b_{m_k}) = 0$, where

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

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

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

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 B.1, 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 B.1.

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

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

for $m_k \geq \max\{M_1, M_2, M_3\}$, which contradicts to (B.14). Thus, (B.22) is proved. □

Note that an alternative proof of the consistency of $\hat{\theta}$ is available [Wu (2011)].

To proof Theorem 3.4, we consider the following Lemma.

Lemma B.3. *Under the conditions of Theorem 3.3, let $\eta = d(1-\gamma)/2$, we have*

(a)

$$m^{\eta}\left(\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right), \tag{B.23}$$

(b)

$$\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\left(1 - \frac{I_m^{\tau}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{K})}{m}\right)}{m^{d-\theta_0}g_{c_0,\theta_0}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{K})}{m}\right)}\right) \times \frac{\dot{g}_{c_0,\theta_0}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{K})}{m}\right)}{g_{c_0,\theta_0}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{K})}{m}\right)} = O_p(m^{-\eta}) \tag{B.24}$$

Proof. To prove (B.23), we find the asymptotic distribution of its lower and upper bounds. It can be easily shown that

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

where

$$L_m = m^{\eta}\left(\frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_M)/m)} - 1\right), \tag{B.25}$$

$$U_m = m^{\eta}\left(\frac{\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}_m)/m)} - 1\right) \tag{B.26}$$

with $\boldsymbol{K}_M$ and $\boldsymbol{K}_m$ as defined in Lemma B.2. We rewrite L_m as

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

By Lemma A.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 3.1,

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

Similarly, we can show

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

The convergence of lower and upper bounds to the same distribution implies (B.23).

To show (B.24), we rewrite the LHS of (B.24) as

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

By Lemma A.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

$$L'_m \le \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0}g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\frac{\dot g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} \le U'_m,$$

where

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

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 A.1, $\gamma > d/(d+2)$ and Theorem 3.1, we can show that

$$m^{\eta}\left(L'_m - \frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\right) \xrightarrow{d} \mathcal{N}\left(0, \left(\frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\right)^2 \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right),$$
$$m^{\eta}\left(U'_m - \frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\right) \xrightarrow{d} \mathcal{N}\left(0, \left(\frac{\dot{g}_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}{g_{c_0,\theta_0}((\pi/2)\mathbf{1}_d)}\right)^2 \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right).$$

This completes the proof of (B.24).

□

Proof of Theorem 3.4. 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

$$(\log(m))^{-1}m^{\eta}\dot{L}(c_0, \theta_0) \xrightarrow{d} \mathcal{N}\left(0, \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right), \tag{B.27}$$
$$(\log(m))^{-2}\ddot{L}(c_0, \bar{\theta}) \xrightarrow{p} 1. \tag{B.28}$$

Since $\dot{L}(c_0, \theta_0)$

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

we see that (B.27) follows from Lemma B.3.

Next we prove (B.28). 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

$$(\log(m))^{-2} E_1 \stackrel{p}{\longrightarrow} 1. \tag{B.29}$$

It can be easily shown that

$$LB''_m \le (\log(m))^{-2} E_1 \le U''_m,$$

where

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

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 3.1, (3.10) in Theorem 3.3 and Lemma A.1, we can show that both L''_m and U''_m converge to one in probability, which in turn implies (B.29). In a similar way, we can show that $(\log(m))^{-1}E_2 = O_p(1)$. Thus, together with (B.29), we can show (B.28), which completes the proof.

□

References

[1] Adler, R. J. (1981). *The geometry of random fields.* John Wiley & Sons Ltd., New York.

[2] Adler, R. J. and Taylor, J. E. (2007). *Random fields and geometry.* Springer.

[3] Anderes, E. (2010). On the consistent separation of scale and variance for Gaussian random fields. *Ann. Statist.* **38** 870-893.

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

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

[6] 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.

[7] 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.

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

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

[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, G. B. (1999). *Real analysis.* Wiley-Interscience Publication, New York.

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

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

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

[15] 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.

[16] Loh, W.-L. (2005). Fixed-domain asymptotics for a subclass of Matern-

type Gaussian random fields. *Ann. Statist.* **33** 2344-2394.

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

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

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

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

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

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

[23] Stein, M. L. (1999). *Interpolation of spatial data*. Springer, New York.

[24] Stein, M. L. (2002). The screening effect in kriging. *Ann. Statist.* **30** 298-323.

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

[26] Wu, W. (2011). Tail estimation of the spectral density under fixed-domain asymptotics. Ph.D. Dissertation, Department of Statistics and Probability, Michigan State University.

[27] Xue, Y. and Xiao, Y. (2010). Fractal and smoothness properties of space-time Gaussian models. *Frontiers Math.* to appear.

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

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

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

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