Tail Estimation of the Spectral Density for a Stationary Gaussian Random Field

Wei-Ying Wu[a], Chae Young Lim[b,1], Yimin Xiao[c,2]

[a] *Institute of Statistical Science, Academia Sinica, Taipei, Taiwan 115*
[b] *Department of Statistics & Probability, Michigan State University, East Lansing, MI 48824*
[c] *Department of Statistics & Probability, Michigan State University, East Lansing, MI 48824*

Abstract

Consider a stationary Gaussian random field on $\mathbb{R}^d$ with 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. 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, study their properties under the fixed-domain asymptotics and provide simulation results.

Keywords: fixed-domain asymptotics, fractal index, fractal dimension, Gaussian random fields, infill asymptotics, microergodic parameter, Whittle likelihood

1. Introduction

For a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbb{R}^d$, we have the spectral representation

$$Z(\boldsymbol{s}) = \int_{\mathbb{R}^d} \exp(i\boldsymbol{\lambda}^T \boldsymbol{s}) M(d\boldsymbol{\lambda}),$$

where M is a complex valued Gaussian random measure. The spectral measure F is defined as $F(d\boldsymbol{\lambda}) = \mathrm{E}(|M(d\boldsymbol{\lambda})|^2)$ which yields the covariance function of Z,

$$K(\boldsymbol{x}) = \int_{\mathbb{R}^d} \exp(i\boldsymbol{\lambda}^T \boldsymbol{x}) F(d\boldsymbol{\lambda}),$$

that is, $K(\boldsymbol{x}) = \mathrm{cov}(Z(\boldsymbol{s}+\boldsymbol{x}), Z(\boldsymbol{s}))$. When F has a density f, we call it the spectral density of Z. Moreover, if $K \in L^1(\mathbb{R}^d)$, then we have the inversion formula

$$f(\boldsymbol{\lambda}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp(-i\boldsymbol{\lambda}^T \boldsymbol{x}) K(\boldsymbol{x}) d\boldsymbol{x}.$$

Email addresses: `wuweiying1011@stat.sinica.edu.tw` (Wei-Ying Wu), `lim@stt.msu.edu` (Chae Young Lim), `xiaoyimi@stt.msu.edu` (Yimin Xiao)

[1] Supported by MSU grant 08-IRGP-1532
[2] Supported in part by NSF grant DMS-1006903

The behavior of the spectral density for large values of $|\boldsymbol{\lambda}|$ plays a central role in characterizing the local properties of the Gaussian random field Z, where $|\boldsymbol{\lambda}|$ is a usual Euclidean norm. In this paper, we assume that the spectral density f of Z satisfies

$$f(\boldsymbol{\lambda}) \;\sim\; c\,|\boldsymbol{\lambda}|^{-\theta} \quad \text{as} \;\; |\boldsymbol{\lambda}| \;\to\; \infty, \tag{1}$$

where $\theta > d$ 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.

We propose estimators of c and θ in (1) which 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 the fixed-domain asymptotics (or infill asymptotics) [see, e.g. Cressie (1993) and Stein (1999)]. Some information on the fixed-domain asymptotics is given in Section 2. 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. We establish consistency and asymptotic normality of the proposed estimators of c and θ, respectively, when the other parameter is known. Also, we obtain the consistency result of the estimator for θ with any fixed $c > 0$.

In the following two subsections, we describe relationships between tail behavior of the spectral density and statistical properties of a random field, and we recall some related methods on estimating tail parameters in the literature. In Section 2, we explain our settings and assumptions, and in Section 3, we introduce our estimators and state the main theorems for the asymptotic properties of the proposed estimators. Simulation studies are given in Section 4. Section 5 discusses some issues related to our approach and possible extension of the current work. All proofs are given in Appendices.

1.1. Tail behavior of the spectral density

The tail behavior of the spectral densities is a main criterion for verifying the equivalence of probability measures generated by stationary Gaussian random fields. The latter is useful for determining the optimality of prediction errors of linear predictors. For example, let P be the probability measure corresponding to the spectral density f which is assumed to be true and P_1 be another probability measure with the spectral density f_1 for a mean zero stationary Gaussian random field Z. If for some $\theta > d$,

$$0 < f(\boldsymbol{\lambda})|\boldsymbol{\lambda}|^{\theta} < \infty \;\; \text{as} \;\; |\boldsymbol{\lambda}| \to \infty \;\; \text{and} \tag{2}$$

$$\int_{|\boldsymbol{\lambda}|>R} \left(\frac{f(\boldsymbol{\lambda}) - f_1(\boldsymbol{\lambda})}{f(\boldsymbol{\lambda})}\right)^2 d\boldsymbol{\lambda} \;<\; \infty \tag{3}$$

for any $R > 0$, then two probability measures P and P_1 are equivalent. Further details of equivalence of Gaussian probability measures and the conditions for equivalence can be found in Ibragimov and Rozanov (1978), Yadrenko (1983) and Stein (1999). The condition (2) together with $f_1(\boldsymbol{\lambda})/f(\boldsymbol{\lambda}) \to 1$ implies that the best linear predictor (BLP) under an incorrect probability measure P_1 is asymptotically equivalent to the BLP under the correct P. We refer to Stein (1999) for further details.

The decay rate of the spectral density as $|\boldsymbol{\lambda}| \to \infty$ is related to the smoothness of a random field. For example, for a stationary Gaussian random field Z, suppose that its covariance function $K(\boldsymbol{x})$ satisfies

$$K(\boldsymbol{x}) \;\sim\; K(\mathbf{0}) - k|\boldsymbol{x}|^{\alpha} \quad \text{as} \;\; |\boldsymbol{x}| \;\to\; 0 \tag{4}$$

for some k and $0 < \alpha \leq 2$. In this case, α is the fractal index that governs the roughness of sample paths of Z and the fractal dimension D of the trajectories of Z is given by $D = d + (1 - \alpha/2)$. This follows from Adler (1981, Chapter 8) or Theorem 5.1 in Xue and Xiao (2011) where more general results are proven for anisotropic (not necessarily stationary) Gaussian random fields. When $\alpha = 2$, it is possible that the sample function may be differentiable. This can be determined by the smoothness of $K(\boldsymbol{x})$ or in terms of the spectral measure of Z [see, e.g., Adler and Taylor (2007) or Xue and Xiao (2011) for further information]. By an Abelian type theorem, (4) holds if the corresponding spectral density satisfies

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

By assuming (1), the fractal index α under our assumption becomes $\alpha = \theta - d$. Throughout the paper, we assume that $Z(\boldsymbol{s})$ has zero mean for simplicity.

The parameter c in (1) also has some interpretation. For example, consider the Matérn spectral density which is widely used for modeling the second-order structure of a random field,

$$f(\boldsymbol{\lambda}) = \frac{\sigma^2 \rho^{2\nu}}{\pi^{d/2}(\rho^2 + |\boldsymbol{\lambda}|^2)^{\nu + d/2}}. \tag{5}$$

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

$$f(\boldsymbol{\lambda}) \;\sim\; \frac{\sigma^2 \rho^{2\nu}}{\pi^{\frac{d}{2}}} |\boldsymbol{\lambda}|^{-(2\nu+d)} \tag{6}$$

as $|\boldsymbol{\lambda}| \to \infty$, we have $c \equiv \sigma^2\rho^{2\nu}/\pi^{d/2}$ and $\theta \equiv 2\nu + d$. Here $\sigma^2\rho^{2\nu}$, which is a function of c, is known as a microergodic parameter (Zhang 2004). The microergodicity of functions of parameters determines the equivalence of probability measures so that the microergodic parameter is the quantity that affects asymptotic mean squared prediction error under fixed-domain settings [see, e.g., Stein (1990a, 1990b, 1999)]. It is well known that not all parameters can be estimated consistently under the fixed-domain asymptotics, but the microergodic parameter can be estimated consistently [see, e.g., Ying (1991, 1993), Zhang (2004), Zhang and Zimmerman (2005), Kaufman et al. (2008), Du et al. (2009), Anderes (2010) and Wang and Loh (2011)].

1.2. Related works

There are a number of references that construct estimators for θ based on fractal properties of processes when $d < \theta \leq d + 2$. For example, Kent and Wood (1997) estimated fractal dimension using increments of a stationary Gaussian process on $\mathbb{R}$ and Chan and Wood (2000) proposed increment-based estimators for both Gaussian and non-Gaussian processes on $\mathbb{R}^2$. Constantine and Hall (1994) estimated effective fractal dimension using variogram for a non-Gaussian stationary process on $\mathbb{R}$. Davies and Hall (1999) studied estimation of fractal dimension using variogram for both isotropic and anisotropic Gaussian processes on $\mathbb{R}$ and $\mathbb{R}^2$. Chan and Wood (2004) considered variogram of increments of processes to estimate fractal dimension for both Gaussian and non-Gaussian processes on $\mathbb{R}$ and $\mathbb{R}^2$. Compared to those works in the spatial domain, the work by Chan et al. (1995) used periodogram in the spectral domain to estimate the fractal dimension.

There is also a number of references regarding the estimation of a microergodic parameter. Ying (1991, 1993) studied asymptotic properties of a microergodic parameter in the exponential covariance model, while Zhang (2004), Loh (2005), Kaufman et al. (2008), Du et al. (2009), Anderes (2010) and Wang and Loh (2011) investigated asymptotic properties of estimators for the Matérn covariance model. Zhang (2004) showed that σ^2 and ρ in (5) can be estimated only in the form of $\sigma^2\rho^{2\nu}$ under the fixed-domain asymptotics when ν is known and $d \leq 3$. Kaufman et al. (2008) showed the strong consistency of the tapered MLE of $\sigma^2\rho^{2\nu}$ when $d \leq 3$ and Du et al. (2009) showed consistency and asymptotic normality of the MLE and the tapered MLE of $\sigma^2\rho^{2\nu}$ when $d = 1$. Both articles assumed that the parameters ν and ρ are known. Wang and Loh (2011) extended asymptotic results of the tapered MLE in Kaufman et al. (2008) and Du et al. (2009). Anderes (2010) proposed an increment-based estimator of $\sigma^2\rho^{2\nu}$ for a geometric anisotropic Matérn covariance model and showed that ρ can be estimated separately when $d > 4$.

The aforementioned references on the estimation of the microergodic parameter in the the Matérn covariance model assume that ν is known (i.e. θ in our setting). Thus, estimation of c with known θ is equivalent to estimate the microergodic parameter with known ν when the spectral density is the Matérn model. Also, the Gaussian assumption was made on the processes, which is crucial for their proofs as well as ours.

2. Preliminaries

We consider a stationary Gaussian random field, $Z(\boldsymbol{s})$ on $\mathbb{R}^d$ with the spectral density $f(\boldsymbol{\lambda})$ that satisfies (1) and suppose that $Z(\boldsymbol{s})$ is observed on the lattice $\phi\boldsymbol{J}$ with spacing ϕ and $\boldsymbol{J} \in T_m = \{1, ..., m\}^d$. Then, we consider a corresponding lattice process $Y_\phi(\boldsymbol{J})$ defined as $Y_\phi(\boldsymbol{J}) \equiv Z(\phi\boldsymbol{J})$ for $\boldsymbol{J} \in \mathbb{Z}^d$, 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. 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 $Y_\phi^\tau(\boldsymbol{J}) \equiv (\Delta_\phi)^\tau Z(\boldsymbol{J})$ as the lattice process obtained by applying the Laplacian operator τ times. Then the corresponding spectral density of Y_ϕ^τ becomes

$$\bar{f}_\phi^\tau(\boldsymbol{\lambda}) = \left\{ \sum_{j=1}^{d} 4\sin^2\left(\frac{\lambda_j}{2}\right) \right\}^{2\tau} \bar{f}_\phi(\boldsymbol{\lambda}). \tag{7}$$

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

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

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

$$g_{c,\theta}(\boldsymbol{\lambda}) = c \left\{ \sum_{j=1}^{d} 4\sin^2\left(\frac{\lambda_j}{2}\right) \right\}^{2\tau} \times \sum_{\boldsymbol{Q} \in \mathbb{Z}^d} |(\boldsymbol{\lambda} + 2\pi\boldsymbol{Q})|^{-\theta}\, \mathbb{I}_{(\boldsymbol{\lambda} \neq \boldsymbol{0})}, \tag{8}$$

where $\mathrm{I}_{\mathcal{A}}$ is the indicator function of the set $\mathcal{A}$.

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. Other differencing method can be considered as long as it can alleviate peakness near origin of the spectral density.

For simplicity of notation, in the subsequent analysis, 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, which is called a fixed-domain setting. For spatial data, it is more 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 is decreasing as the number of observations increases. This sampling scheme requires a different asymptotic framework, called the fixed-domain asymptotics and the classical asymptotic framework with a fixed sampling distance is called the increasing-domain asymptotics to differentiate from the fixed-domain asymptotics.

The spectral density of $Y_\phi^\tau(\boldsymbol{J})$ can be estimated by a periodogram,

$$I_m^\tau(\boldsymbol{\lambda}) = (2\pi m)^{-d} \left| \sum_{\boldsymbol{J} \in \mathcal{T}_m} Y_\phi^\tau(\boldsymbol{J}) \exp\{-i\, \boldsymbol{\lambda}^T \boldsymbol{J}\} \right|^2 .$$

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

$$W_h(\boldsymbol{K}) = \frac{\Lambda_h(2\pi \boldsymbol{K}/m)}{\sum_{\boldsymbol{L} \in \mathcal{T}_m} \Lambda_h(2\pi \boldsymbol{L}/m)}, \tag{9}$$

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 $\Lambda(\boldsymbol{s}) \ge 0$ and $\Lambda(\boldsymbol{0}) > 0$. 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 $0 < C_1 \le a(\boldsymbol{\lambda})/b(\boldsymbol{\lambda}) \le C_2 < \infty$ for all possible $\boldsymbol{\lambda} \in \mathcal{A}$. For asymptotic results in this paper, we consider the following assumption.

Assumption 1. *For a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbb{R}^d$, let $f(\boldsymbol{\lambda})$ be the corresponding spectral density.*

(A) $f(\boldsymbol{\lambda})$ *satisfies* $f(\boldsymbol{\lambda}) \sim c\,|\boldsymbol{\lambda}|^{-\theta}$ *as* $|\boldsymbol{\lambda}| \to \infty$, *for some constants* $c > 0$, $\theta > d$.

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

$$\left|\frac{\partial}{\partial \lambda_j} f(\boldsymbol{\lambda})\right| (1+|\boldsymbol{\lambda}|)^{\theta+1} \quad \text{and} \qquad (10)$$
$$\left|\frac{\partial^2}{\partial \lambda_j \partial \lambda_k} f(\boldsymbol{\lambda})\right| (1+|\boldsymbol{\lambda}|)^{\theta+2}$$

are uniformly bounded for $j, k = 1, ..., d$.

(C) *Furthermore,* $f(\boldsymbol{\lambda})$ *satisfies* $f(\boldsymbol{\lambda}) \asymp (1+|\boldsymbol{\lambda}|)^{-\theta}$ *for all possible* $\boldsymbol{\lambda}$.

3. Main Results

Asymptotic properties of a spatial periodogram and a smoothed spatial periodogram under the fixed-domain asymptotics were investigated by Stein (1995) and Lim and Stein (2008). They assume that the spectral density f satisfies Assumption 1. There is some overlap between Assumption 1 (A) and (C), and the latter says that the spectral density $f(\boldsymbol{\lambda})$ is bounded near the origin. This stronger condition is used to find asymptotic bounds of the expectation, variance and covariance of a spatial periodogram at Fourier frequency $2\pi \boldsymbol{J}/m$ by m and $\boldsymbol{J}$ for $\|\boldsymbol{J}\| \neq 0$. However, consistency and asymptotic normality of the smoothed spatial periodogram at Fourier frequency $2\pi \boldsymbol{J}/m$ are 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 the smoothed spatial periodogram at such Fourier frequencies in this paper, Assumption 1 (C) may not be necessary. Indeed, we are able to remove Assumption 1 (C) to prove the consistency and asymptotic normality of the smoothed spatial periodogram with an additional condition on θ in Theorem 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 paper, let $\xrightarrow{p}$ denote the convergence in probability and $\xrightarrow{d}$ denote the convergence in distribution.

Theorem 1. *Suppose that the spectral density* f *of a stationary Gaussian random field* $Z(\boldsymbol{s})$ *on* $\mathbb{R}^d$ *satisfies Assumption 1 (A) and (B). Also suppose that* $4\tau > \theta - 1$, $d < \theta < 2d$ *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*

$$\frac{\hat{I}_m^{\tau}(2\pi \boldsymbol{J}/m)}{\bar{f}_{\phi}^{\tau}(2\pi \boldsymbol{J}/m)} \xrightarrow{p} 1 \qquad (11)$$

and

$$m^{\eta}\left(m^{-(d-\theta)}\hat{I}_m^{\tau}\left(2\pi\boldsymbol{J}/m\right)-g_{c,\theta}(\boldsymbol{\mu})\right) \tag{12}$$

$$\xrightarrow{d} \quad \mathcal{N}\left(0,\frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d g_{c,\theta}^2(\boldsymbol{\mu})\right),$$

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

Remark 1. *The function* $g_{c,\theta}$ *is integrable under* $4\tau>\theta-d$ *which is satisfied if* $4\tau>\theta-1$. *The latter is necessary to show* $E\left(\hat{I}_m^{\tau}\left(2\pi\boldsymbol{J}/m\right)/\bar{f}_{\phi}^{\tau}\left(2\pi\boldsymbol{J}/m\right)\right)\to 1$ *and the condition* $\max\{(d-2)/d,0\}<\gamma<1$ *is needed to show* $Var\left(\hat{I}_m^{\tau}\left(2\pi\boldsymbol{J}/m\right)/\bar{f}_{\phi}^{\tau}\left(2\pi\boldsymbol{J}/m\right)\right)\to 0$ *so that (11) follows.*

Remark 2. *The condition* $d<\theta<2d$ *is not very restrictive.* $\theta<d+2$ *makes the random field* $Z(\boldsymbol{s})$ *to have fractal sample path. As we mentioned in Section 1.2, many authors only consider this case. The condition* $\theta<2d$ *is more relaxed than* $\theta<d+2$ *when* $d>2$.

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

$$L(c,\theta) \;=\; \sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K})\left\{\log\left(m^{d-\theta}g_{c,\theta}\left(2\pi(\boldsymbol{J}+\boldsymbol{K})/m\right)\right)+\frac{1}{m^{d-\theta}}\frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{c,\theta}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}\right\}, \tag{13}$$

where $W_h(\boldsymbol{K})$ is given in (9). 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 $\boldsymbol{0}$.

The function $L(c,\theta)$ can be viewed as a weighted local Whittle likelihood. If Λ is a nonzero constant function, then $W_h(\boldsymbol{K})\equiv 1/|\mathcal{K}|$ for $\boldsymbol{K}\in\mathcal{K}$, where $\mathcal{K}=\{\boldsymbol{K}\in\mathcal{T}_m:\|2\pi\boldsymbol{K}/m\|\le h\}$ and $|\mathcal{K}|$ is the number of elements in the set $\mathcal{K}$. In this case, $L(c,\theta)$ is in the form of a local Whittle likelihood for the lattice data $\{Y_{\phi}^{\tau}(\boldsymbol{J}),\boldsymbol{J}\in T_m\}$ 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 $Y_{\phi}^{\tau}(\boldsymbol{J})$ 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, $L(c,\theta)$ is determined by frequencies 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. The estimator $\hat{c}$ has an explicit expression, which is obtained by solving $\partial L(c, \theta_0)/\partial c = 0$:

$$\hat{c} = \sum_{\boldsymbol{K} \in \mathcal{T}_m} W_h(\boldsymbol{K}) \frac{1}{m^{d-\theta_0}} \frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_0(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}, \tag{14}$$

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

Theorem 2. *Suppose that the spectral density f of a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbb{R}^d$ satisfies either Assumption 1 (A) and (B) with $d < \theta < 2d$ or Assumption 1 (A)- (C). 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 (14), we have*

$$\hat{c} \xrightarrow{p} c, \tag{15}$$

and

$$m^\eta(\hat{c} - c) \xrightarrow{d} \mathcal{N}\left(0\,,\, c^2 \frac{\Lambda_2}{\Lambda_1^2}\left(\frac{2\pi}{\mathfrak{C}}\right)^d\right), \tag{16}$$

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

Remark 3. *We can prove Theorem 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.*

When we choose a constant function Λ 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 (5) with $d = 1$, Du et al. (2009) showed that for any fixed α_1 with known ν, the MLE of σ^2 satisfies

$$n^{1/2}(\hat{\sigma}^2\alpha_1^{2\nu} - \sigma_0^2\alpha_0^{2\nu}) \xrightarrow{d} \mathcal{N}\left(0\,,\, 2(\sigma_0^2\alpha_0^{2\nu})^2\right), \tag{17}$$

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 (17). 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 $c > 0$ is fixed. The proposed estimator of θ is then given by

$$\hat{\theta} = \arg\min_{\theta \in \Theta} L(c, \theta), \tag{18}$$

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. *Suppose that the spectral density f of Z satisfies either Assumption 1 (A) and (B) with $d < \theta < 2d$ or Assumption 1 (A)- (C). 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 (18), we have*

$$\hat{\theta} \ \xrightarrow{p} \ \theta. \tag{19}$$

In addition, if c is known as the true value c_0,

$$\hat{\theta} - \theta \ = o_p\big((\log m)^{-1}\big), \tag{20}$$

and if c is different from the true value,

$$\hat{\theta} - \theta \ = O_p\big((\log m)^{-1}\big). \tag{21}$$

Remark 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 determining the asymptotic distribution, we will make use of the rate of convergence given in (20). However, we were not able to prove asymptotic distribution when c is a fixed positive constant other than the true value due to the slower convergence rate given in (21). The rate given in (20) for a known c is not optimal and the optimal rate is obtained in Theorem 4.*

Remark 5. *We could consider the estimator of c by minimizing $L(c, \hat{\theta})$, where $\hat{\theta}$ is obtained with any fixed c. That is,*

$$\tilde{c} = \sum_{\boldsymbol{K} \in \mathcal{T}_m} W_h(\boldsymbol{K}) \frac{1}{m^{d-\hat{\theta}}} \frac{I_m^{\tau}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{g_{\hat{\theta}}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}. \tag{22}$$

However, consistency of $\tilde{c}$ is not guaranteed. Instead, $\tilde{c} - c_0 = O_p(1)$ can be easily derived.

From Theorem 3, we show the following result on the asymptotic distribution of $\hat{\theta}$.

Theorem 4. *Under the conditions of Theorem 3, if c is known as the true value c_0, we have*

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

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

Remark 6. *Note that we have a different convergence rate for $\hat{\theta}$ compared to the convergence rate for $\hat{c}$ given in Theorem 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. Numerical Results

In this section, we study performance of our proposed estimators based on simulated data. A real data example can be found in Wu (2011).

4.1. Simulation studies

We first give an example on $\mathbb{R}$. Consider the spectral density that corresponds to a damped oscillation covariance function on $\mathbb{R}$ introduced in Yaglom (1987). The damped oscillation covariance function is given by

$$K(x) = \sigma^2 \exp(-\alpha|x|) \cos(\omega_0 x), \tag{23}$$

for $x \in \mathbb{R}$, where α and ω_0 are positive. The corresponding spectral density is

$$f(\lambda) = \frac{A(\lambda^2 + b^2)}{\lambda^4 + 2a\lambda^2 + b^4}, \tag{24}$$

for $\lambda \in \mathbb{R}$, where $A = \sigma^2\alpha/\pi$, $a = \alpha^2 - \omega_0^2$ and $b = \sqrt{\alpha^2 + \omega_0^2}$. The spectral density in (24) satisfies $f(\lambda) \sim c|\lambda|^{-\theta}$ as $|\lambda| \to \infty$, where $c = A$ and $\theta = 2$. We consider two sets of parameters; (1) $(\sigma^2, \alpha, \omega_0) = (1, 1, 1)$ and (2) $(\sigma^2, \alpha, \omega_0) = (4, 1, 1)$. The corresponding true values of (c, θ) are (1) $(c_0, \theta_0) = (1/\pi, 2)$ and (2) $(c_0, \theta_0) = (4/\pi, 2)$.

For each set of parameters, we generate 100 realizations from the Gaussian process with covariance function given in (23) for grid sizes, $m = 100, 200$ on $[0, 10]$. To estimate parameters, we choose a constant function Λ for $W_h(\cdot)$ and consider various bandwidths h. Since the value of h corresponds to a number of frequencies, $|\mathcal{K}|$, used in the objective function (13), we give different values of $|\mathcal{K}|$ by setting $|\mathcal{K}|/m = 0.1, 0.2, 0.3$. Also, we consider $\tau = 1, 2, 3$ to see the effect of different choice of τ. Table 1,2 and 3 are results of the simulation study. Each table shows the bias (Bias) and the standard deviation (STD) scaled by 10^2.

First of all, the effect of different τ is marginal on the estimates with respect to the Bias and the STD. Table 1 gives the Bias and the STD of $\hat{c}$ given $\theta = \theta_0$. Table 2 gives the Bias and the STD of $\hat{\theta}$ given $c = c_0$. The Bias and the STD decrease overall as m increases which is consistent with our theoretical findings. When $|\mathcal{K}|/m$ increases (i.e. more frequencies are used), the STD decreases as one can expect. On the other hand, mostly the Bias for $|\mathcal{K}|/m = 0.3$ increased compared to the Bias for $|\mathcal{K}|/m = 0.2$. This shows that we need to investigate choice of bandwidth h for better performance of the proposed method. We will discuss more on this issue in Section 5. Table 3 shows the

Bias and STD of the estimate $\hat{\theta}$ given fixed c. We have larger Bias when c is not the true value while the STD does not change much. The Bias decreases as m increases, but the rate is slow. These results are expected from our theoretical finding (21).

The second example is the Matérn spectral density given in (5) on $\mathbb{R}^2$. We consider two sets of parameters; (1) $(\sigma^2, \rho, \nu) = (1, \pi, 1)$ and (2) $(\sigma^2, \rho, \nu) = (1, 1, 0.5)$. The corresponding true values of (c, θ) are (1) $(c_0, \theta_0) = (\pi, 4)$ and (2) $(c_0, \theta_0) = (1/\pi, 3)$. For each set of parameters, we generate 100 realizations from the Gaussian random field with the Matérn covariance function for grid sizes, $m = 50, 100$ on $[0, 10]^2$. Simulation settings are similar to the first example. Table 4,5 and 6 are results of the simulation study. Each table shows the bias (Bias) and the standard error (STD) scaled by 10^2. The results are overall comparable to the results in the first example, which illustrates the theoretical findings on $\mathbb{R}^2$.

5. Discussion

We have 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 (13). 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.

The weights in (13) 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 $\mathfrak{C}$ and appropriate γ. The simulation study shows that the choice of h affects the estimation result. In the proofs, we use a smoothed spatial periodogram $\hat{I}_m^{\tau}$ and its properties. Thus, if we could find the optimal bandwidth that minimizes the mean squared error of $\hat{I}_m^{\tau}$, it can be used as a guideline for h. However, finding the mean squared error of $\hat{I}_m^{\tau}$ needs explicit first order asymptotic expressions of the bias and the variance of $\hat{I}_m^{\tau}(\boldsymbol{\lambda})$, which need further investigation.

Another quantity that affects the performance of the estimator is τ, the number of differencing. From the asymptotic result, any τ that satisfies $4\tau > \theta - 1$ when we use a Laplace differencing ($2\tau > \theta - 1$ when we use a simple differencing for $d = 1$) should produce a consistent estimator. However, there is no practical guideline on the choice of τ when we estimate θ. On the other hand, the simulation results show that variability (STD) does not depend much on the choice of τ while there are some variation in bias. Thus, one simple recommendation is to try several values of τ and consider an averaged quantity. One could investigate a leading term in bias and see the possibility to reduce bias due to the choice of τ.

It will be more useful if we can estimate the parameters c and θ together. One approach is already considered by estimating θ with unknown c but we have only the

Table 1: Estimation of c with known θ for the spectral density given in (24) on $\mathbb{R}$. The values of Bias and STD are scaled by 10^2.

$(\sigma^2, \alpha, \omega_0) = (1, 1, 1)$, $(c_0, \theta_0) = (1/\pi, 2)$						
	$\tau = 1$		$\tau = 2$		$\tau = 3$	
$m = 100$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m = 0.1$	1.23	8.98	1.28	9.08	2.08	10.11
0.2	1.13	7.72	1.03	7.69	1.23	7.44
0.3	0.82	5.90	0.80	5.81	2.10	6.37
$m = 200$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m = 0.1$	0.53	7.65	0.40	7.64	0.72	7.68
0.2	0.47	5.23	0.43	5.32	0.78	5.34
0.3	0.64	4.23	0.61	4.26	1.39	4.22

$(\sigma^2, \alpha, \omega_0) = (4, 1, 1)$, $(c_0, \theta_0) = (4/\pi, 2)$						
	$\tau = 1$		$\tau = 2$		$\tau = 3$	
$m = 100$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m = 0.1$	4.90	45.22	4.90	44.32	10.78	43.10
0.2	6.23	29.62	7.25	29.02	10.57	30.38
0.3	6.67	23.16	8.03	23.43	12.47	27.56
$m = 200$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m = 0.1$	3.48	31.74	4.10	31.56	4.30	31.05
0.2	1.33	20.35	1.74	20.42	2.32	21.73
0.3	2.22	16.53	2.93	16.55	6.38	19.96

Table 2: Estimation of θ with known c for the spectral density given in (24) on $\mathbb{R}$. The values of Bias and STD are scaled by 10^2.

$(\sigma^2, \alpha, \omega_0) = (1, 1, 1)$, $(c_0, \theta_0) = (1/\pi, 2)$						
	$\tau = 1$		$\tau = 2$		$\tau = 3$	
$m = 100$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m = 0.1$	0.03	9.13	0.01	9.23	-0.43	10.21
0.2	-0.26	7.92	-0.14	8.02	-0.35	7.60
0.3	-0.33	6.13	-0.27	6.20	-1.18	6.28
$m = 200$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m = 0.1$	0.40	6.22	0.50	6.20	0.20	6.25
0.2	-0.05	4.39	0.01	4.45	-0.25	4.45
0.3	-0.24	3.49	-0.20	3.52	-0.74	3.43

$(\sigma^2, \alpha, \omega_0) = (4, 1, 1)$, $(c_0, \theta_0) = (4/\pi, 2)$						
	$\tau = 1$		$\tau = 2$		$\tau = 3$	
$m = 100$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m = 0.1$	0.61	10.77	0.58	10.68	-0.99	9.88
0.2	-0.69	7.12	-0.94	7.02	-1.68	7.14
0.3	-1.22	5.42	-1.49	5.41	-2.23	6.05
$m = 200$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m = 0.1$	0.09	6.47	-0.06	6.37	-0.12	6.19
0.2	0.07	4.20	-0.01	4.18	-0.06	4.42
0.3	-0.22	3.93	-0.35	3.36	-0.86	3.78

Table 3: Estimation of θ with unknown c for the spectral density given in (24) on $\mathbb{R}$. The values of Bias and STD are scaled by 10^2. The values of c used are $1/5$, $1/\pi$ and $1/2$.

$(\sigma^2, \alpha, \omega_0) = (1, 1, 1)$, $(c_0, \theta_0) = (1/\pi, 2)$						
$\lvert\mathcal{K}\rvert/m = 0.2$	$\tau = 1$		$\tau = 2$		$\tau = 3$	
$m = 100$	Bias	STD	Bias	STD	Bias	STD
$c = 1/5$	−15.10	7.70	−14.98	7.80	−15.36	7.38
$1/\pi$	−0.26	7.92	−0.14	8.01	−0.35	7.59
$1/2$	14.49	8.11	14.60	8.20	14.56	7.77
$m = 200$	Bias	STD	Bias	STD	Bias	STD
$c = 1/5$	−12.29	4.33	−12.27	4.40	−12.50	4.38
$1/\pi$	−0.05	4.39	0.01	4.45	−0.25	4.45
$1/2$	12.03	4.43	12.09	4.49	11.84	4.50

consistency result. We conjecture that the asymptotic distribution is normal after removing the bias term but it is challenging to show the asymptotic distribution due to the term m^θ in the objective function.

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). Our results show that one can estimate tail behavior using only local information. This can be seen as an analogue to a screening effect.

As introduced in Section 1.1, c is related to a microergodic parameter. The previous references on the estimation of microergodic parameters under the fixed-domain asymptotics are limited in a sense that they assume the parametric covariance model such as the Matérn model or its variants with known θ. The consistency and asymptotic distribution of MLE and Tapered MLE for the isotropic Matérn model were established in Du et al. (2009) for $d = 1$. Kaufman et al. (2009) studied the tapered MLEs for the isotropic Matérn model with $d \leq 3$ but showed only consistency. Wang and Loh (2011) have extended the results of Du, et al. (2009) from $d = 1$ to $d \leq 3$. Anderes (2010) considered an increment-based estimator for a geometric anisotropic Matérn model with $d \leq 3$ but showed only consistency. On the other hand, our method does not require a

Table 4: Estimation of c with known θ for the Matérn spectral density. The values of Bias and STD are scaled by 10^2.

$(\sigma^2, \rho, \nu) = (1, \pi, 1)$, $(c_0, \theta_0) = (\pi, 4)$						
	$\tau = 1$		$\tau = 2$		$\tau = 3$	
$m = 50$	Bias	STD	Bias	STD	Bias	STD
$\|\mathcal{K}\|/m^2 = 0.1$	-26.06	19.33	-26.57	19.55	-22.47	20.20
0.2	-19.30	14.91	-20.66	15.02	-16.79	15.59
0.3	-14.61	12.67	-16.94	12.75	-12.39	13.52
$m = 100$	Bias	STD	Bias	STD	Bias	STD
$\|\mathcal{K}\|/m^2 = 0.1$	-3.44	11.59	-4.61	11.46	-2.44	11.44
0.2	7.51	7.55	8.30	7.79	16.20	8.57
0.3	6.60	6.52	7.70	6.72	16.05	7.55

$(\sigma^2, \rho, \nu) = (1, 1, 0.5)$, $(c_0, \theta_0) = (1/\pi, 3)$						
	$\tau = 1$		$\tau = 2$		$\tau = 3$	
$m = 50$	Bias	STD	Bias	STD	Bias	STD
$\|\mathcal{K}\|/m^2 = 0.1$	-0.10	2.04	0.04	2.05	0.72	2.10
0.2	0.11	1.70	0.10	1.69	0.73	1.74
0.3	0.90	1.53	0.74	1.49	1.59	1.55
$m = 100$	Bias	STD	Bias	STD	Bias	STD
$\|\mathcal{K}\|/m^2 = 0.1$	0.20	1.04	0.25	1.03	0.60	1.04
0.2	1.09	0.75	0.87	0.71	1.20	0.71
0.3	1.07	0.62	0.86	0.58	1.23	0.60

Table 5: Estimation of θ with known c for the Matérn spectral density. The values of Bias and STD are scaled by 10^2.

$(\sigma^2, \rho, \nu) = (1, \pi, 1)$, $(c_0, \theta_0) = (\pi, 4)$						
	$\tau = 1$		$\tau = 2$		$\tau = 3$	
$m = 50$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m^2 = 0.1$	4.71	2.19	4.06	1.60	3.46	1.34
0.2	4.41	2.23	3.63	1.61	3.06	1.36
0.3	2.79	2.31	0.87	1.76	0.03	1.53
$m = 100$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m^2 = 0.1$	0.33	1.07	0.44	1.06	0.26	1.05
0.2	-0.62	0.66	-0.67	0.68	-1.32	0.73
0.3	-0.57	0.58	-0.65	0.60	-1.35	0.65

$(\sigma^2, \rho, \nu) = (1, 1, 0.5)$, $(c_0, \theta_0) = (1/\pi, 3)$						
	$\tau = 1$		$\tau = 2$		$\tau = 3$	
$m = 50$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m^2 = 0.1$	0.15	2.06	-.0.014	2.06	-0.78	2.07
0.2	0.03	1.69	0.01	1.69	-0.67	1.70
0.3	-0.60	1.47	-0.55	1.45	-1.38	1.47
$m = 100$	Bias	STD	Bias	STD	Bias	STD
$\lvert\mathcal{K}\rvert/m^2 = 0.1$	-0.14	0.85	-0.18	0.84	-0.44	0.84
0.2	-0.81	0.58	-0.65	0.55	-0.89	0.55
0.3	-0.82	0.48	-0.66	0.46	-0.94	0.47

Table 6: Estimation of θ with unknown c for the Matérn spectral density. The values of Bias and STD are scaled by 10^2. The values of c used are 2, π and 4.

$(\sigma^2, \rho, \nu) = (1, \pi, 1)$, $(c_0, \theta_0) = (\pi, 4)$						
$\lvert\mathcal{K}\rvert/m^2 = 0.2$	$\tau = 1$		$\tau = 2$		$\tau = 3$	
$m = 50$	Bias	STD	Bias	STD	Bias	STD
$c = 2$	-12.81	1.58	-13.33	1.59	-16.40	1.78
π	4.41	2.23	3.63	1.61	3.06	1.36
4	13.10	1.61	12.70	1.63	10.11	1.75
$m = 100$	Bias	STD	Bias	STD	Bias	STD
$c = 2$	-14.06	0.68	-13.43	0.64	-13.49	0.66
π	-0.62	0.66	-0.67	0.68	-1.32	0.73
4	5.6	0.68	6.19	0.65	6.15	0.66

particular parametric model for the spectral density and we establish both consistency and asymptotic distribution of the estimator. The restriction on d is also less limited since we only need $4\tau > \theta - 1$ and $\theta > d$. The condition $\theta > d$ is required to guarantee the integrability of f, which is necessary for all spectral density functions.

The parameter θ is related to a fractal index when $d < \theta \leq d + 2$ so that our method can be used to estimate the fractal index. One advantage of our method compared to other methods for estimating the fractal index is that we can handle the data on $\mathbb{R}^d$ with $d \geq 3$. A possible application of our work is surface metrology in which surface roughness is of interest. For example, our method can be used to measure the uniformness of gray scale levels on 3-D image of materials where each 3-D pixel has a gray scale value.

[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, New York.

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

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

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

[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] Constantine, A. G. and 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] Davies, S. and Hall, P. (1999). Fractal Analysis of surface roughness by using spatial data. *J. R. Statist. Soc. B* **61**, 3-37.

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

[13] 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.

[14] Kent, J. T and Wood, A. T. A. (1997). Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. *J. R. Statist. Soc. B* **59**, 679-699.

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

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

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

[18] 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.

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

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

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

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

[23] Wang, D. and Loh, W. (2011). On fixed-domain asymptotics and covariance tapering in Gaussian random field models. *Electron. J. Stat.* **5** 238-269.

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

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

[26] Xue, Y. and Xiao, Y. (2011). Fractal and smoothness properties of space-time Gaussian models. *Frontiers Math.* **6** 1217-1246.

[27] Yadrenko, M. (1983). *Spectral theory of random fields.* New York: Optimization Software.

[28] Yaglom, A. M. (1987). *Correlation theory of stationary and related random functions I: Basic Results.* Springer, New York.

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

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

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

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. The proof is given in Wu (2011). 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 $\boldsymbol{\lambda} \neq \mathbf{0}$.

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 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 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 $\Theta \times \mathcal{A}_\rho$.

Appendix B. Proofs of Theorems in Section 3

Proof of Theorem 1. If $f(\boldsymbol{\lambda})$ satisfies Assumption 1 (A)-(C), (11) and (12) hold by results in Stein (1995) and Lim and Stein (2008). To prove (11) and (12) when only Assumption 1 (A) and (B) hold, we need to show that the effect of $f(\boldsymbol{\lambda})$ on $|\boldsymbol{\lambda}| \leq C$ is negligible.

Let $I_m^{f,\tau}(\boldsymbol{\lambda})$ be the periodogram at $\boldsymbol{\lambda}$ from the observations under $f(\boldsymbol{\lambda})$ and

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

Note that

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

(11) and (12) follow from Theorems 3, 6 and 12 in Lim and Stein (2008) when these Theorems hold for f under Assumption 1 (A) and (B). The key part of proofs of these Theorems under Assumption 1 (A) and (B) is to show

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

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

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).

Consider a spectral density $k(\boldsymbol{\lambda})$ that satisfies Assumption 1 (A)-(C) and $k(\boldsymbol{\lambda}) \sim \kappa f(\boldsymbol{\lambda})$ for some $\kappa > 0$. Such $k(\boldsymbol{\lambda})$ can be easily constructed. 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

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

for $\boldsymbol{J}$ that satisfies $\|\boldsymbol{J}\| \asymp m$ and $2\boldsymbol{J}/m \notin \mathbb{Z}^d$. Note that for an $\epsilon > 0$, there exists $C > 0$ such that for $|\boldsymbol{\lambda}| > C$, $|k(\boldsymbol{\lambda})/\kappa f(\boldsymbol{\lambda}) - 1| < \epsilon$. Then, (B.3) holds since

$$\begin{aligned} &\left| a_{m,\phi}^{f,\tau}(\boldsymbol{J},\pm\boldsymbol{J}) - a_{m,\phi}^{k,\tau}(\boldsymbol{J},\pm\boldsymbol{J}) \right| \\ &\leq (2\pi m)^{-d} \int_{|\boldsymbol{\lambda}|\leq C} \left\{ \sum_{j=1}^d 4\sin^2\left(\phi\lambda_j/2\right) \right\}^{2\tau} |f(\boldsymbol{\lambda}) - k(\boldsymbol{\lambda})| \\ &\quad \times \prod_{j=1}^d \frac{\sin^2\left(m\phi\lambda_j/2\right)}{|\sin\left(\phi\lambda_j/2 + \pi J_j/m\right)\sin\left(\phi\lambda_j/2 \pm \pi J_j/m\right)|} d\boldsymbol{\lambda} \\ &\quad + (2\pi m)^{-d} \int_{|\boldsymbol{\lambda}|> C} \left\{ \sum_{j=1}^d 4\sin^2\left(\phi\lambda_j/2\right) \right\}^{2\tau} |k(\boldsymbol{\lambda})|\,|f(\boldsymbol{\lambda})/k(\boldsymbol{\lambda}) - 1| \\ &\quad \times \prod_{j=1}^d \frac{\sin^2\left(m\phi\lambda_j/2\right)}{|\sin\left(\phi\lambda_j/2 + \pi J_j/m\right)\sin\left(\phi\lambda_j/2 \pm \pi J_j/m\right)|} d\boldsymbol{\lambda} \\ &\leq v_1\, m^{-d-4\tau} + v_2\, m^{-d} \end{aligned}$$

for some positive constants v_1 and v_2 since $\|\phi\lambda_j/2 \pm \pi J_j/m\|$ stays away from zero and π when m is large. Note that $\phi = m^{-1}$.

$\square$

We prove the remaining theorems under Assumption 1 (A) and (B) with $d < \theta < 2d$ for simplicity. When we consider Assumption 1 (A)-(C) without $\theta < 2d$, we only need to replace the reference to Theorem 1 with the reference to Lim and Stein (2008).

Proof of Theorem 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

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

which can be rewritten as

$$\begin{aligned}\frac{c\,\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{U}})/m)} &\leq \hat{c} \\ &\leq \frac{c\,\hat{I}_m^{\tau}(2\pi\boldsymbol{J}/m)}{m^{d-\theta_0} g_{c,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K}^{\mathcal{L}})/m)} \qquad \text{(B.4)}\end{aligned}$$

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

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

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

from Theorem 1 and

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

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

□

To prove Theorems 3, we consider following lemmas.

Lemma 2. *Consider a function $h_z(x) = -\log(x) + z(x-1)$ and a positive function of m, d_m, for a positive integer m. Assume that $d_m \to d > 0$ as $m \to \infty$. Then, for any* *given* $r_l \in (0, \min\{1, 1/d\})$ *and* $r_u \in (\max\{1, 1/d\}, \infty)$ *such that* $\min\{h_d(r_l), h_d(r_u)\} > 0$, *there exist* $\delta_r > 0$ *and* $M_r > 0$ *such that for all* $m \geq M_r$, $h_{d_m}(x) > \delta_r$, *for any* $x \in \mathfrak{Z}_r \equiv (0, r_l] \cup [r_u, \infty)$.

Proof. It can be easily shown that for any positive z, $h_z(x)$ is a convex function on $(0, \infty)$ and minimized at $x = 1/z$ with $h_z(1/z) \leq 0$. Since $d_m \to d$, for given $r_l \in (0, \min\{1, 1/d\})$ and $r_u \in (\max\{1, 1/d\}, \infty)$ such that $\min\{h_d(r_l), h_d(r_u)\} > 0$, there exists $M_r > 0$ such that for all $m \geq M_r$, we have $|1/d_m - 1/d| < \min\{1/d - r_l, r_u - 1/d\}$ and $\min\{h_{d_m}(r_l), h_{d_m}(r_u)\} > (1/2)\min\{h_d(r_l), h_d(r_u)\} > 0$. Hence for all $x \in \mathfrak{Z}_r$, we have $h_r(x) \geq \min\{h_{d_m}(r_l), h_{d_m}(r_u)\} > (1/2)\min\{h_d(r_l), h_d(r_u)\} \equiv \delta_r$.

□

Lemma 3. *For a positive integer* m, $\theta \in \Theta$ *and* $c > 0$, *we have*

$$L(c, \theta) - L(c, \theta_0) \geq A_m + B_m + C_m,$$

where

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

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

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

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

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

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.

Proof. From the expression of $L(c,\theta)$ given in (13), we have

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

H_m is further decomposed as $H_m = A_m + B_m + C_m$ where A_m, B_m and C_m are 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 $g_{c,\theta}$ in Lemma 1 implies that

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

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

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

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

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

C_m converges to zero in probability.

□

Proof of Theorem 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

$$
P(L(c,\hat{\theta}_m) - L(c,\theta_0) \leq 0) = 1 \tag{B.14}
$$

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 3, we have $L(c, \hat{\theta}_m(\omega)) - L(c, \theta_0) \geq A_m + B_m + C_m$, 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

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

where $h_{d_m}(\cdot)$ is defined in Lemma 2 with

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

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

By Theorem 1 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 $d_m \xrightarrow{p} c_0/c$. Let $d = c_0/c$. Then, there exists $\{m_k\}$, a subsequence of $\{m\}$ such that d_{m_k} converges to d almost surely. By (B.13) in the proof of Lemma 3, almost sure convergence of d_{m_k} implies that C_{m_k} defined in (B.9) converges to zero almost surely. To use Lemma 2, 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, because of the uniform boundedness of $g_{c,\theta_0}/g_{c,\theta}$, there exists a M_2, which does not depend on ω, such that for $m_k \geq M_2$,

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

for all $\omega \in \mathcal{D}_{m_k}$, where a is a solution of $h_d(x) = 0$ for $d \neq 1$ and $a \neq 1$. When $d = 1$, set $a = 1$. Then, by Lemma 2 with $r_l = 1 - \max\{|1-a|, 1/2\}$ and $r_u = 1 + \max\{|1-a|, 1/2\}$, there exist $\delta_r > 0$ and $M_r \geq \max\{M_1, M_2\}$ such that for $m_k \geq M_r$,

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

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

By the uniform convergence of $|B_m|$ on Θ shown in Lemma 3, 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

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

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 (20) when $c = c_0$, 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 1. To show (B.22), we consider a similar argument that was used to show consistency of $\hat{\theta}_m$. For simplicity, we reset notations such as ϵ, δ, δ_r, M and $\mathcal{D}_m$, etc. that were used in the proof of consistency.

Suppose that (B.22) does not hold. Then, there exists $1 > \epsilon > 0$, $\delta > 0$ and M_1 such that $P\left(\left| m^{\hat{\theta}_m - \theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\hat{\theta}_m}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} - 1 \right| \ > \ \epsilon\right) > \delta$ for all $m \geq M_1$. Let $\mathcal{D}_m = \left\{\omega : \left| m^{\hat{\theta}_m - \theta_0} \frac{g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c_0,\hat{\theta}_m}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} - 1 \right| > \epsilon\right\}$. On the other hand, there exists $\{m_k\}$, a subsequence of $\{m\}$, such that $d_{m_k} \ \rightarrow \ 1$, $B_{m_k} \ \rightarrow \ 0$ and $C_{m_k} \ \rightarrow \ 0$ almost surely, where d_m is given in (B.15), B_m and C_m are given in (B.8) and (B.9) with $c = c_0$. 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 converged uniformly on Ω_δ.

Then, by Lemma 2 with $r_l = 1 - \epsilon$ and $r_u = 1 + \epsilon$ for $d = 1$, there exist δ_r and M_2 such that for $m_k \geq M_2$, $A_{m_k} > \delta_r$ 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 $m_k \geq \max\{M_1, M_2\}$ since $P(\mathcal{D}_{m_k} \cap \Omega_\delta) \geq \delta/2 > 0$ for all $m_k \geq \max\{M_1, M_r\}$. Note that δ_r does not depend on m_k which can be seen in Lemma 2.

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 $P\Big(L(c_0, \hat{\theta}_m) - L(c_0, \theta_0) > \delta_r/2\Big) \geq \delta/2$ for $m_k \geq \max\{M_1, M_2, M_3\}$, which contradicts to (B.14) with $c = c_0$. Thus, (B.22) is proved.

The convergence rate of $\hat{\theta}_m$ when $c \neq c_0$ can be shown by showing that there exist r_l and r_u such that $\lim_{m\to\infty} P(r_l < m^{\hat{\theta}_m - \theta_0} \frac{g_{c,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)}{g_{c,\hat{\theta}_m}(2\pi(\boldsymbol{J}+\boldsymbol{S}_m)/m)} < r_u) = 1$ using a similar contradiction argument.

□

To prove Theorem 4, we consider the following Lemma.

Lemma 4. *Under the conditions of Theorem 3, let $\eta = d(1-\gamma)/2$, we have*

$$(a) \qquad m^\eta \left(\sum_{\boldsymbol{K}\in\mathcal{T}_m} W_h(\boldsymbol{K}) \frac{I_m^\tau(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)}{m^{d-\theta_0} g_{c_0,\theta_0}(2\pi(\boldsymbol{J}+\boldsymbol{K})/m)} - 1 \right) \tag{B.23}$$

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

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

$$\times \frac{\dot{g}_{c_0,\theta_0}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{K})}{m}\right)}{g_{c_0,\theta_0}\left(\frac{2\pi(\boldsymbol{J}+\boldsymbol{K})}{m}\right)} = O_p(m^{-\eta})$$

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

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

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 3. We rewrite L_m as

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

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

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

Thus, by Theorem 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

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

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

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

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

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

This completes the proof of (B.24).

□

Proof of Theorem 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 $\dot{L}(c_0,\theta_0)$ equals

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

we see that (B.27) follows from Lemma 4.

Next we prove (B.28). After some simplification, we write $\ddot{L}(c_0, \bar{\theta})$ as

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

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 show that

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

It can be easily shown that $L''_m \leq (\log(m))^{-2} E_1 \leq U''_m$, 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 1, (20) in Theorem 3 and Lemma 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 $(\log(m))^{-1} E_2 = O_p(1)$. This, together with (B.29), yields (B.28). The proof is complete. □