Electronic Journal of Statistics
Vol. 5 (2011) 238–269
ISSN: 1935-7524
DOI: 10.1214/11-EJS607

On fixed-domain asymptotics and covariance tapering in Gaussian random field models*

Daqing Wang

Department of Statistics and Applied Probability
National University of Singapore
Singapore 117546
Republic of Singapore
e-mail: daqingwang@gmail.com

and

Wei-Liem Loh

Department of Statistics and Applied Probability
National University of Singapore
Singapore 117546
Republic of Singapore
e-mail: stalohwl@nus.edu.sg

Abstract: Gaussian random fields are commonly used as models for spatial processes and maximum likelihood is a preferred method of choice for estimating the covariance parameters. However if the sample size n is large, evaluating the likelihood can be a numerical challenge. Covariance tapering is a way of approximating the covariance function with a taper (usually a compactly supported function) so that the computational burden is reduced. This article studies the fixed-domain asymptotic behavior of the tapered MLE for the microergodic parameter of a Matérn covariance function when the taper support is allowed to shrink as $n \to \infty$. In particular if the dimension of the underlying space is ≤ 3, conditions are established in which the tapered MLE is strongly consistent and also asymptotically normal. Numerical experiments are reported that gauge the quality of these approximations for finite n.

AMS 2000 subject classifications: Primary 62M30; secondary 62E20, 62G15.
Keywords and phrases: Asymptotic normality, covariance tapering, fixed-domain asymptotics, Gaussian random field, Matérn covariance, maximum likelihood estimation, spatial statistics, strong consistency.

Received November 2010.

*Research is supported in part by grant AcRF Tier 1 R-155-000-094-112.

1. Introduction

Let $X : \mathbb{R}^d \to \mathbb{R}$ be a mean-zero isotropic Gaussian random field with the Matérn covariance function

$$\begin{aligned} \mathrm{Cov}(X(\mathbf{x}), X(\mathbf{y})) &= \sigma^2 K_\alpha(\mathbf{x}-\mathbf{y}) \\ &= \frac{\sigma^2(\alpha\|\mathbf{x}-\mathbf{y}\|)^\nu}{2^{\nu-1}\Gamma(\nu)} \mathcal{K}_\nu(\alpha\|\mathbf{x}-\mathbf{y}\|), \qquad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d, \end{aligned} \tag{1}$$

where $\nu > 0$ is a known constant, α, σ are strictly positive but unknown parameters and $\mathcal{K}_\nu$ is the modified Bessel function of the second kind (cf. Andrews, Askey and Roy [2], pages 222–223). $\|.\|$ denotes the usual Euclidean norm in $\mathbb{R}^d$. Because $x^\nu \mathcal{K}_\nu(x) \to 2^{\nu-1}\Gamma(\nu)$ as $x \to 0$, σ^2 is the variance of X. The corresponding isotropic spectral density is given by

$$\begin{aligned} f_{\alpha,\sigma}(\mathbf{w}) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} \sigma^2 K_\alpha(\mathbf{x}) d\mathbf{x} \\ &= \frac{\sigma^2 \alpha^{2\nu}}{\pi^{d/2}(\alpha^2 + \|\mathbf{w}\|^2)^{\nu+d/2}}, \qquad \forall \mathbf{w} \in \mathbb{R}^d, \end{aligned} \tag{2}$$

where $\imath = \sqrt{-1}$. It is well known that X is m times mean square differentiable where m is the largest integer strictly less than ν. Hence ν can be thought of as a smoothness parameter and α the scale parameter.

As observed in Zhang [19], (1) comprises a very broad class of covariance functions and it has received considerable attention in recent years. Unlike many other families of covariance functions (such as exponential, powered-exponential, or spherical covariance functions), the Matérn class in (1) has a parameter, namely ν, that controls the smoothness of the random field. Stein [12] presented very convincing arguments in favor of using (1) to model spatial correlations and a comprehensive account of the properties of Matérn-type Gaussian random fields can also be found there.

Interestingly if ν is known and $d \le 3$, Zhang [19] proved that α and σ cannot be estimated consistently whereas the quantity $\sigma^2\alpha^{2\nu}$ can be estimated consistently under fixed-domain asymptotics. It is reassuring to note that Corollary 2 of Zhang [19] further showed that it is the latter quantity, and not the individual parameters α, σ, that matters in interpolation. $\sigma^2\alpha^{2\nu}$ is an example of a microergodic parameter. We refer the reader to Stein [12], page 163, for the mathematical definition of microergodicity.

In contrast for $d \ge 5$, Anderes [1] recently proved that the Gaussian measures defined by $\sigma^2 K_\alpha$ and $\sigma_1^2 K_{\alpha_1}$ are orthogonal if $(\alpha_1, \sigma_1) \ne (\alpha, \sigma)$ and hence α and σ can be consistently estimated under fixed-domain asymptotics. The case $d = 4$ is still open.

This article is concerned with the estimation of $\sigma^2\alpha^{2\nu}$ using observations

$$\{X(\mathbf{x}_1), X(\mathbf{x}_2), \ldots, X(\mathbf{x}_n)\}, \tag{3}$$

where $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are distinct points in $[0,T]^d$ for some (absolute) constant $0 < T < \infty$. For simplicity, we write

$$\mathbf{X}_n = (X(\mathbf{x}_1), X(\mathbf{x}_2), \ldots, X(\mathbf{x}_n))'.$$

The covariance matrix of $\mathbf{X}_n$ can be expressed as $\sigma^2 \mathbf{R}_\alpha$ where $\mathbf{R}_\alpha$ is a $n \times n$ correlation matrix whose elements do not depend on σ. Since $\mathbf{X}_n \sim N_n(0, \sigma^2 \mathbf{R}_\alpha)$, the log-likelihood function $l(\alpha, \sigma)$ satisfies

$$l(\alpha, \sigma) = -\frac{n}{2} \log(2\pi) - n \log(\sigma) - \frac{1}{2} \log(|\mathbf{R}_\alpha|) - \frac{1}{2\sigma^2} \mathbf{X}_n' \mathbf{R}_\alpha^{-1} \mathbf{X}_n.$$

It is generally acknowledged (e.g., Stein, Chi and Welty [13], Furrer, Genton and Nychka [6], Kaufman, Schervish and Nychka [10] and Du, Zhang and Mandrekar [5]) that in practice, the data set in (3) is usually very large and is irregularly spaced. Computing the inverse covariance matrix $\sigma^{-2}\mathbf{R}_\alpha^{-1}$, which takes $O(n^3)$ operations, is then a difficult problem and may even be intractable in some instances.

A popular and promising way to alleviate this computational problem is to replace the original covariance function by a more tractable one. More precisely, we impose a simpler (but mis-specified) covariance function for X given by

$$\mathrm{Cov}(X(\mathbf{x}), X(\mathbf{y})) = \sigma_1^2 \tilde{K}_{\alpha_1,n}(\mathbf{x} - \mathbf{y}), \qquad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d, \tag{4}$$

where $\tilde{K}_{\alpha_1,n} : \mathbb{R}^d \to \mathbb{R}$ is a known isotropic correlation function, $\alpha_1 > 0$ is a known constant and $\sigma_1^2 = \sigma^2 \alpha^2 / \alpha_1^{2\nu}$. $\tilde{K}_{\alpha_1,n}$ is allowed to, possibly, vary with sample size n. Under assumption (4), let $\sigma_1^2 \tilde{\mathbf{R}}_{\alpha_1,n}$ be the covariance matrix of $\mathbf{X}_n$ and hence the corresponding (pseudo) log-likelihood function $\tilde{l}_n(\alpha_1, \sigma_1)$ satisfies

$$\tilde{l}_n(\alpha_1, \sigma_1) = -\frac{n}{2} \log(2\pi) - n \log(\sigma_1) - \frac{1}{2} \log(|\tilde{\mathbf{R}}_{\alpha_1,n}|) - \frac{1}{2\sigma_1^2} \mathbf{X}_n' \tilde{\mathbf{R}}_{\alpha_1,n}^{-1} \mathbf{X}_n. \tag{5}$$

Let $\hat{\sigma}_{1,n}$ be the value of σ_1 that maximizes $\tilde{l}_n(\alpha_1, \sigma_1)$, i.e.

$$\tilde{l}_n(\alpha_1, \hat{\sigma}_{1,n}) = \max_{\sigma_1 > 0} \tilde{l}_n(\alpha_1, \sigma_1).$$

Since $\tilde{\mathbf{R}}_{\alpha_1,n}$ does not depend on σ_1 and

$$\frac{\partial}{\partial \sigma_1} \tilde{l}_n(\alpha_1, \sigma_1) = -\frac{n}{\sigma_1} + \frac{1}{\sigma_1^3} \mathbf{X}_n' \tilde{\mathbf{R}}_{\alpha_1,n}^{-1} \mathbf{X}_n,$$

we have

$$\hat{\sigma}_{1,n}^2 = \frac{1}{n} \mathbf{X}_n' \tilde{\mathbf{R}}_{\alpha_1,n}^{-1} \mathbf{X}_n. \tag{6}$$

For example, Zhang [19] took $\tilde{\mathbf{R}}_{\alpha_1,n} = \mathbf{R}_{\alpha_1}$ where $\alpha_1 > 0$ is a known (arbitrarily specified) constant. This made the likelihood analysis simpler because (5) is a function of only σ_1^2. Zhang [19] proved that for $d \leq 3$, $\hat{\sigma}_{1,n}^2 \alpha_1^{2\nu} \to \sigma^2 \alpha^{2\nu}$ with

$P_{\alpha,\sigma}$ probability 1 where $P_{\alpha,\sigma}$ is the Gaussian measure defined by the covariance function in (1). The key idea in Zhang's proof is that the two Gaussian measures in question are equivalent.

Covariance tapering is an attractive method of constructing $\sigma_1^2 \tilde{K}_{\alpha_1,n}$ such that it is an isotropic, positive definite and compactly supported function. A way to implement covariance tapering is as follows. Let $K_{\text{tap}} : \mathbb{R}^d \to \mathbb{R}$ be an isotropic correlation function with compact support, say, $\text{supp}(K_{\text{tap}}) \subseteq [-1,1]^d$. Define $\tilde{K}_{\alpha_1,n}$ in (4) to be

$$\tilde{K}_{\alpha_1,n}(\mathbf{x}) = K_{\alpha_1}(\mathbf{x}) K_{\text{tap}}(\mathbf{x}/\gamma_n), \qquad \forall \mathbf{x} \in \mathbb{R}^d, \tag{7}$$

where $0 < \gamma_n \leq 1$, $n = 1, 2, \ldots$, is a sequence of (possibly) decreasing constants and K_{α_1} is as in (1) with α replaced by α_1 (a known constant). The motivation is that the covariance matrix $\sigma_1^2 \tilde{\mathbf{R}}_{\alpha_1,n}$ of $\mathbf{X}_n$ corresponding to $\tilde{K}_{\alpha_1,n}$ is sparse (with many off-diagonal elements taking the value 0) and sparse matrix algorithms are available to evaluate the log-likelihood (5) more efficiently (cf. Davis [4] and the references cited therein). Isotropic, positive definite, compactly supported functions have been an intensively studied field. The literature includes Wu [17], Wendland [15, 16] and Gneiting [7].

Assuming $\gamma_n \equiv \gamma$ is an absolute constant, Kaufman, *et al.* [10] established conditions on the spectral density of K_{tap} such that $\hat{\sigma}_{1,n}^2 \alpha_1^{2\nu} \to \sigma^2 \alpha^{2\nu}$ with $P_{\alpha,\sigma}$ probability 1. As in Zhang [19], the theory of the equivalence of Gaussian measures is used in a crucial manner.

In the case $d = 1$ and $\gamma_n \equiv \gamma$, Du, *et al.* [5] established conditions on the spectral density of K_{tap} such that $\sqrt{n}(\hat{\sigma}_{1,n}^2 \alpha_1^{2\nu} - \sigma^2 \alpha^{2\nu})$ converges in law to $N(0, 2(\sigma^2\alpha^{2\nu})^2)$ as $n \to \infty$ under the Gaussian measure $P_{\alpha,\sigma}$. Also if $\sigma_1^2 \tilde{K}_{\alpha_1,n} = \sigma_1^2 K_{\alpha_1}$ (i.e. if the true Matérn covariance function is mis-specified as another Matérn covariance function), they showed that $\sqrt{n}(\hat{\sigma}_{1,n}^2 \alpha_1^{2\nu} - \sigma^2 \alpha^{2\nu})$ converges in law to $N(0, 2(\sigma^2\alpha^{2\nu})^2)$ as $n \to \infty$ under $P_{\alpha,\sigma}$. As open problems, Du, *et al.* [5] observed that their techniques cannot be extended from $d = 1$ to $d = 2$ or 3, and it would be practically important to obtain analogous asymptotic normality results for higher dimensions. They further noted that letting $\gamma_n \to 0$ as $n \to \infty$ is a natural scheme in the fixed-domain asymptotic framework and remarked that it is not obvious that their proofs can be adapted to a varying γ_n.

This article has essentially three main results. The first result, namely Theorem 1 below, is to generalize the strong consistency result of Kaufman, *et al.* [10] from $\gamma_n \equiv \gamma$ to a sequence of γ_n's which could vary with n, in particular where $\gamma_n \to 0$ as $n \to \infty$. It is noted that even for covariance tapering with $\gamma_n \equiv \gamma$, the number of operations needed to compute the inverse covariance matrix is still $O(n^3)$ whereas if $\gamma_n \to 0$, the number of operations is $o(n^3)$. Clearly the latter will lessen the computational burden of evaluating the likelihood and inverting the covariance matrix even more. More precisely, our first result is

Theorem 1. *Let $0 < T < \infty$, $1 \leq d \leq 3$ and $\sigma^2 K_\alpha$ be the Matérn covariance function as in (1). Let ϵ, M be constants such that $\epsilon > \max\{d/4, 1-\nu\}$. Suppose K_{tap} is an isotropic correlation function with supp$(K_{\text{tap}}) \subseteq [-1,1]^d$ whose* *spectral density*

$$f_{\text{tap}}(\mathbf{w}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} K_{\text{tap}}(\mathbf{x}) d\mathbf{x}$$

satisfies

$$f_{\text{tap}}(\mathbf{w}) \le \frac{M}{(1 + \|\mathbf{w}\|^2)^{\nu + d/2 + \epsilon}}, \qquad \forall \mathbf{w} \in \mathbb{R}^d.$$

Let $\alpha_1 > 0$ *and* $\sigma_1 > 0$ *be constants such that* $\sigma_1^2 \alpha_1^{2\nu} = \sigma^2 \alpha^{2\nu}$. *Define*

$$\tilde{K}_{\alpha_1,n}(\mathbf{x}) = K_{\alpha_1}(\mathbf{x}) K_{\text{tap}}(\mathbf{x}/\gamma_n), \qquad \forall \mathbf{x} \in \mathbb{R}^d,$$

where $\gamma_n = C_\gamma n^{-a}$ *and* $0 \le a < 1/(4\nu + 4\epsilon + 2d), 0 < C_\gamma \le 1$ *are constants (independent of* n*). Let* $\hat{\sigma}_{1,n}^2$ *be as in (6). Then*

$$\hat{\sigma}_{1,n}^2 \alpha_1^{2\nu} \to \sigma^2 \alpha^{2\nu}, \qquad \text{as } n \to \infty,$$

with $P_{\alpha,\sigma}$ *probability 1 where* $P_{\alpha,\sigma}$ *is the Gaussian measure defined by the covariance function* $\sigma^2 K_\alpha$ *in (1).*

Remark. The above theorem reduces to Theorem 2 of Kaufman, *et al.* [10] if we take $a = 0$ or, equivalently, $\gamma_n \equiv \gamma$.

The second result is Theorem 2 which extends the asymptotic normality results of Du, *et al.* [5] from $d = 1$ and $\gamma_n \equiv \gamma$ to $1 \le d \le 3$ and γ_n possibly varying with n. In particular, we have

Theorem 2. *Let* $0 < T < \infty$, $1 \le d \le 3$, $\sigma^2 K_\alpha$ *be the Matérn covariance function as in (1). Let* ϵ, M *be constants such that* $\epsilon > \max\{d/4, 1 - \nu\}$. *Suppose* K_{tap} *is an isotropic correlation function with supp*$(K_{\text{tap}}) \subseteq [-1, 1]^d$ *whose spectral density* f_{tap} *satisfies*

$$f_{\text{tap}}(\mathbf{w}) \le \frac{M}{(1 + \|\mathbf{w}\|^2)^{\nu + d/2 + \epsilon}}, \qquad \forall \mathbf{w} \in \mathbb{R}^d.$$

Let $\alpha_1 > 0$, $\sigma_1 > 0$ *and* $0 \le b < 1/(8\nu + 8\epsilon + 2d)$ *be constants such that* $\sigma_1^2 \alpha_1^{2\nu} = \sigma^2 \alpha^{2\nu}$ *and* $2b(2\nu + 2\epsilon + d)/\min\{2, 4 - d, 4\epsilon - d, 4\nu + d\} < (1 - 2bd)/(2d)$. *Define*

$$\tilde{K}_{\alpha_1,n}(\mathbf{x}) = K_{\alpha_1}(\mathbf{x}) K_{\text{tap}}(\mathbf{x}/\gamma_n), \qquad \forall \mathbf{x} \in \mathbb{R}^d,$$

where $\gamma_n = C_\gamma n^{-b}$ *and* $0 < C_\gamma \le 1$ *is a constant (independent of* n*). Let* $\hat{\sigma}_{1,n}^2$ *be as in (6). Then*

$$\sqrt{n}(\hat{\sigma}_{1,n}^2 \alpha_1^{2\nu} - \sigma^2 \alpha^{2\nu}) \to N(0, 2(\sigma^2 \alpha^{2\nu})^2),$$

in law as $n \to \infty$ *with respect to* $P_{\alpha,\sigma}$*, the Gaussian measure defined by the covariance function* $\sigma^2 K_\alpha$ *in (1).*

Remark. For $b = 0$ or, equivalently, $\gamma_n \equiv \gamma$ and $d = 1$, Theorem 2 proves the asymptotic normality of $\hat{\sigma}_{1,n}^2$ under weaker conditions than Theorem 5(ii) of Du, *et al.* [5].

The third result is Theorem 3 which deals with the case where the Matérn covariance function $\sigma^2 K_\alpha$ is mis-specified by another Matérn covariance function $\sigma_1^2 K_{\alpha_1}$ with α_1 a known constant. The proof of Theorem 3 is similar to (though simpler than) that of Theorem 2 and is omitted. We refer the reader to Wang [14] for a detailed proof of Theorem 3.

Theorem 3. *Let $0 < T < \infty$, $1 \leq d \leq 3$, $\sigma^2 K_\alpha$ be the Matérn covariance function as in (1). Let $\alpha_1 > 0$ and $\sigma_1 > 0$ be constants such that $\sigma_1^2 \alpha_1^{2\nu} = \sigma^2 \alpha^{2\nu}$. Define $\tilde{K}_{\alpha_1,n}(\mathbf{x}) = K_{\alpha_1}(\mathbf{x})$, $\forall \mathbf{x} \in \mathbb{R}^d$, and $\tilde{\mathbf{R}}_{\alpha_1,n} = \mathbf{R}_{\alpha_1}$. Let $\hat{\sigma}_{1,n}^2$ be as in (6). Then*

$$\sqrt{n}(\hat{\sigma}_{1,n}^2 \alpha_1^{2\nu} - \sigma^2 \alpha^{2\nu}) \to N(0, 2(\sigma^2 \alpha^{2\nu})^2),$$

in law as $n \to \infty$ with respect to $P_{\alpha,\sigma}$, the Gaussian measure defined by the covariance function $\sigma^2 K_\alpha$ in (1).

Remark. For $d = 1$, Theorem 3 reduces to Theorem 5(i) of Du, *et al.* [5]. In the case $\nu = 1/2$, i.e. the Ornstein-Uhlenbeck process on $[0, T]$, Ying [18] proved the strong consistency and asymptotic normality of the MLE for $\sigma^2 \alpha$ while Du, *et al.* [5] obtained similar results for the tapered MLE (obtained by maximizing (5) with respect to both α_1 and σ_1).

The rest of this article is organized as follows. As a check on the practical applicability of Theorems 1 to 3 for finite sample sizes, some numerical experiments are performed and are reported in Section 2. Section 3 proves a number of Bernstein-type probability inequalities. These inequalities are needed in the proof of Theorem 1. Section 4 is heavily motivated by the equivalence of Gaussian measures theory (when $d = 1$) as detailed in Chapter 3 of Ibragimov and Rozanov [9]. However in the case that $\gamma_n \to 0$ as $n \to \infty$, the Gaussian measures in Theorems 1 and 2 are not equivalent. As such, the results of Ibragimov and Rozanov [9] have to be modified to accommodate this fact. The main result of Section 3 is (25) which is needed in the proofs of Theorems 1 and 2.

Lemma 4 in Section 5 establishes some bounds on the spectral density of a tapered covariance function. The proof of Lemma 4 is a slight refinement of that found in Kaufman, *et al.* [10] in order to accommodate a varying γ_n. Finally, the proofs of Theorems 1 and 2 are given in Sections 6 and 7 respectively. The Appendix contains the proofs of a number of technical lemmas that are needed in the proofs of the theorems.

An Associate Editor noted that the Gaussian assumption as well as the Matérn covariance function play crucial roles in establishing the results of this article. The main reasons are that the proofs use the well developed theory of equivalence of Gaussian measures and that the spectral density of the Matérn covariance function has a rather simple form. Zhang [19], page 259, has a discussion on the difficulties of obtaining analogous results for non-Gaussian random fields and the use of other covariance functions. The latter would be an important direction for future research.

We end this Introduction with a brief note on notation. $\mathbb{R}$ and $\mathbb{C}$ denote the sets of real and complex numbers respectively. $\mathcal{I}\{.\}$ is the indicator function and $|\mathbf{x}|_{\max} = \max_{1 \leq i \leq d} |x_i|$, $\forall \mathbf{x} = (x_1, \ldots, x_d)' \in \mathbb{R}^d$. If $\mathbf{x} \in \mathbb{R}^d$, then $\mathbf{x}'$ is its transpose. Finally, $f \asymp g$ as $\|\mathbf{w}\| \to \infty$ (or $n \to \infty$) means there exist constants $0 < c < C$ such that $c \le f/g \le C$ for sufficiently large $\|\mathbf{w}\|$ (or n) respectively.

2. Numerical experiments

This section uses numerical experiments to gauge the accuracy of the asymptotic results of Theorems 1 to 3 for finite sample sizes. In particular, Sections 2.1 and 2.2 are concerned with Theorem 3 and Theorems 1, 2 respectively. More details of this study can be found in Wang [14]. Define

$$Z = \frac{\sqrt{n}(\hat{\sigma}_1^2 \alpha_{1,n}^{2\nu} - \sigma^2 \alpha^{2\nu})}{\sqrt{2}\sigma^2 \alpha^{2\nu}}.$$

2.1. *Precision of Theorem 3 approximations for finite n*

Each experiment comprises simulating 10,000 independent realizations of a mean-zero Gaussian random field with the Matérn covariance function (1). There are altogether 18 experiments where

(i) $\sigma = 1$, $\alpha = 0.8$ and $\alpha_1 = 1.3$ are fixed.
(ii) d takes values 1, 2, 3.
(iii) ν takes values 1/4 and 1/2. ($d = 1$ and $\nu = 0.5$ give the Ornstein-Uhlenbeck process.)
(iv) For $d = 1$, the Gaussian random field $X(.)$ is observed on a regular grid on $[0, 1]^d$, i.e. $\{X(1/n), X(1/n), \ldots, X(n/n)\}$ where $n = 1000, 2500, 5000$.
(v) For $d = 2$, $X(.)$ is observed on a regular grid on $[0, 1]^d$, i.e. $\{X(i/m, j/m) : 1 \le i, j \le m\}$ where $m = 30, 50, 70$. Here the sample size is $n = m^2$.
(vi) For $d = 3$, $X(.)$ is observed on a regular grid on $[0, 1]^d$, i.e. $\{X(i/m, j/m, k/m) : 1 \le i, j, k \le m\}$ where $m = 10, 15, 17$. Here the sample size is $n = m^3$.

Table 1 compares the percentiles of Z with those of $N(0, 1)$ and also reports the bias and mean square error (MSE) of $\hat{\sigma}_1^2 \alpha_{1,n}^{2\nu}$ as an estimator of $\sigma^2 \alpha^{2\nu}$. In particular, the simulations reveal that the asymptotic approximations get more accurate as (i) the smoothness parameter ν decreases, (ii) the sample size n increases, (iii) the dimension d decreases, and (iv) $|\alpha - \alpha_1|$ decreases.

2.2. *Precision of Theorems 1 and 2 approximations for finite n*

As in Section 2.1, each experiment comprises simulating 10,000 independent realizations of a mean-zero Gaussian random field with the Matérn covariance function (1). There are altogether 18 experiments where

(i) $\sigma = 1$, $\alpha = 5$ and $\alpha_1 = 7.5$ are fixed.
(ii) d takes values 1, 2.
(iii) When $d = 1$, ν takes values 1/4, 1/2. When $d = 2$, ν takes values 1/8, 1/4.

Table 1
Percentiles of Z and bias, MSE of $\hat{\sigma}_1^2 \alpha_{1,n}^{2\nu}$

n	d	ν	5%	25%	50%	75%	95%	bias	MSE
1000	1	1/4	-1.615	-0.690	0.024	0.651	1.707	< 0.0001	0.0016
2500	1	1/4	-1.619	-0.669	0.022	0.674	1.630	< 0.0001	0.0006
5000	1	1/4	-1.615	-0.679	0.004	0.690	1.672	0.0001	0.0003
30^2	2	1/4	-1.569	-0.643	0.012	0.717	1.717	0.0017	0.0018
50^2	2	1/4	-1.608	-0.661	0.021	0.713	1.712	0.0008	0.0007
70^2	2	1/4	-1.648	-0.647	0.037	0.710	1.692	0.0006	0.0003
10^3	3	1/4	-1.506	-0.567	0.082	0.769	1.802	0.0044	0.0016
15^3	3	1/4	-1.559	-0.607	0.068	0.737	1.744	0.0016	0.0005
17^3	3	1/4	-1.571	-0.631	0.038	0.741	1.719	0.0011	0.0003
$N(0,1)$			-1.6449	-0.6749	0	0.6749	1.6449		
1000	1	1/2	-1.575	-0.629	0.023	0.701	1.732	0.0018	0.0013
2500	1	1/2	-1.606	-0.669	0.000	0.681	1.661	0.0002	0.0005
5000	1	1/2	-1.626	-0.661	0.003	0.666	1.650	< 0.0001	0.0003
30^2	2	1/2	-1.556	-0.630	0.055	0.741	1.708	0.0025	0.0014
50^2	2	1/2	-1.613	-0.627	0.034	0.717	1.705	0.0010	0.0005
70^2	2	1/2	-1.587	-0.659	0.021	0.700	1.670	0.0004	0.0003
10^3	3	1/2	-1.453	-0.523	0.140	0.837	1.840	0.0060	0.0017
15^3	3	1/2	-1.515	-0.563	0.117	0.780	1.792	0.0023	0.0004
17^3	3	1/2	-1.543	-0.568	0.092	0.774	1.744	0.0016	0.0003

(iv) For $d = 1$, the Gaussian random field $X(.)$ is observed on a regular grid on $[0, 1]$, i.e. $\{X(1/n), X(1/n), \ldots, X(n/n)\}$ where $n = 1000, 2500, 5000$.
(v) For $d = 2$, $X(.)$ is observed on a regular grid on $[0, 1]^2$, i.e. $\{X(i/m, j/m) : 1 \le i, j \le m\}$ where $m = 30, 50, 70$. The sample size is $n = m^2$.

The covariance function is then mis-specified by multiplying it by a taper as in (7). A popular class of tapers is due to Wendland [15, 16]. The Wendland taper $\phi_{d,k}(\mathbf{x})$ is a positive definite function with support $\{\mathbf{x} : \|\mathbf{x}\| \le 1\}$. The corresponding spectral density function $f_{d,k}$ satisfies

$$f_{d,k}(\mathbf{w}) \le \frac{M}{(1 + \|\mathbf{w}\|^2)^{d/2+k+1/2}},$$

where M is a constant. Hence in order to satisfy the conditions of Theorems 1 and 2, we choose the taper $\phi_{1,1}(x/\gamma_n) = (1 - x/\gamma_n)_+^3 (1 + 3x/\gamma_n)$ when $d = 1$ and set $\gamma_n = Cn^{-0.03}$ with $C = 1$, 0.75, 0.3. Here $x_+ = \max\{0, x\}$. Similarly for $d = 2$, we choose the taper $\phi_{2,1}(\mathbf{x}/\gamma_n) = (1 - \mathbf{x}/\gamma_n)_+^4 (1 + 4\mathbf{x}/\gamma_n)$ and set $\gamma_n = Cn^{-0.02}$ with $C = 1$, 0.75.

Table 2 and Table 3 compare the percentiles of Z with those of $N(0, 1)$ and also report the bias and mean square error (MSE) of $\hat{\sigma}_1^2 \alpha_{1,n}^{2\nu}$ as an estimator of $\sigma^2 \alpha^{2\nu}$ when $d = 1$ and $d = 2$ respectively.

In particular, the simulations reveal that the asymptotic approximations get more accurate as (i) the smoothness parameter ν decreases, (ii) the sample size n increases, (iii) the dimension d decreases, and (iv) γ_n increases.

Table 2
Percentiles of Z and bias, MSE of $\hat{\sigma}_1^2 \alpha_{1,n}^{2\nu}$ for $d = 1$ and $\phi_{1,1}(x/\gamma_n)$

n	γ_n	ν	5%	25%	50%	75%	95%	bias	MSE
1000	0.813	1/4	-1.582	-0.634	0.032	0.718	1.739	0.0048	0.0101
2500	0.791	1/4	-1.600	-0.660	0.003	0.687	1.672	0.0012	0.0040
5000	0.775	1/4	-1.641	-0.678	0.004	0.680	1.686	0.0004	0.0020
1000	0.610	1/4	-1.588	-0.642	-0.022	0.709	1.730	0.0041	0.0101
2500	0.593	1/4	-1.605	-0.666	-0.003	0.682	1.664	0.0009	0.0040
5000	0.581	1/4	-1.631	-0.659	0.032	0.704	1.665	0.0012	0.0020
1000	0.244	1/4	-1.648	-0.714	-0.040	0.641	1.667	-0.0025	0.0101
2500	0.237	1/4	-1.656	-0.710	-0.048	0.636	1.625	-0.0019	0.0040
5000	0.232	1/4	-1.621	-0.708	-0.035	0.660	1.675	-0.0009	0.0020
	$N(0,1)$		-1.6449	-0.6749	0	0.6749	1.6449		
1000	0.813	1/2	-1.563	-0.620	0.055	0.755	1.783	0.0163	0.0518
2500	0.791	1/2	-1.592	-0.641	0.035	0.716	1.727	0.0064	0.0203
5000	0.775	1/2	-1.603	-0.660	0.023	0.707	1.693	0.0026	0.0101
1000	0.610	1/2	-1.576	-0.635	-0.039	0.739	1.762	0.0130	0.0517
2500	0.593	1/2	-1.599	-0.650	-0.024	0.708	1.716	0.0050	0.0203
5000	0.581	1/2	-1.664	-0.654	-0.014	0.677	1.672	0.0011	0.0100
1000	0.244	1/2	-1.704	-0.765	-0.089	0.612	1.637	-0.0167	0.0517
2500	0.237	1/2	-1.695	-0.736	-0.068	0.619	1.618	-0.0076	0.0203
5000	0.232	1/2	-1.678	-0.736	-0.062	0.615	1.599	-0.0055	0.0100

Table 3
Percentiles of Z and bias, MSE of $\hat{\sigma}_1^2 \alpha_{1,n}^{2\nu}$ for $d = 2$ and $\phi_{2,1}(\mathbf{x}/\gamma_n)$

n	γ_n	ν	5%	25%	50%	75%	95%	bias	MSE
30^2	0.873	1/8	-1.491	-0.529	0.139	0.837	1.865	0.0112	0.0052
50^2	0.855	1/8	-1.570	-0.610	0.051	0.741	1.732	0.0028	0.0018
70^2	0.844	1/8	-1.601	-0.651	0.017	0.701	1.693	0.0010	0.0009
30^2	0.655	1/8	-1.582	-0.623	-0.042	0.738	1.763	0.0043	0.0051
50^2	0.641	1/8	-1.676	-0.724	-0.066	0.625	1.615	0.0020	0.0018
70^2	0.633	1/8	-1.706	-0.776	0.095	0.578	1.543	0.0028	0.0009
	$N(0,1)$		-1.6449	-0.6749	0	0.6749	1.6449		
30^2	0.873	1/4	-1.509	-0.562	0.114	0.802	1.794	0.0135	0.0115
50^2	0.855	1/4	-1.618	-0.657	0.014	0.701	1.721	0.0017	0.0041
70^2	0.844	1/4	-1.665	-0.708	0.019	0.657	1.614	-0.0012	0.0020
30^2	0.655	1/4	-1.699	-0.757	-0.087	0.598	1.590	-0.0078	0.0113
50^2	0.641	1/4	-1.826	-0.862	-0.194	0.489	1.504	0.0115	0.0042
70^2	0.633	1/4	-1.846	-0.905	-0.235	0.443	1.412	-0.0102	0.0021

3. Some probability inequalities

This section proves a number of probability inequalities that are needed in the sequel. Let α_1, $\mathbf{X}_n$ and $\hat{\sigma}_{1,n}$ be defined as in (6). Define $\sigma_1^2 = \sigma^2 \alpha^{2\nu}/\alpha_1^{2\nu}$. Let $\mathcal{A} = \{|\hat{\sigma}_{1,n}^2 - \sigma^2\alpha^{2\nu}/\alpha_1^{2\nu}| > \varepsilon\}$ for some constant $\varepsilon > 0$ and $\mathcal{B} \subseteq \mathbb{R}^n$ such that $\mathcal{A} = \{\mathbf{X}_n \in \mathcal{B}\}$. For simplicity, we write $P_{\alpha,\sigma}$ and $p_{\alpha,\sigma}$ to denote probability and probability density function of $\mathbf{X}_n$ when (1) holds with parameters α, σ, and $\tilde{P}_{\alpha_1,\sigma_1,n}$ and $\tilde{p}_{\alpha_1,\sigma_1,n}$ to denote probability and probability density function of $\mathbf{X}_n$ defined by the covariance function $\sigma_1^2 \tilde{K}_{\alpha_1,n}$ in (4). Then for any constant $\tau_n > 0$ (which may depend on n), we have

$$P_{\alpha,\sigma}(\mathcal{A}) \quad = \quad \int_{\mathcal{B}} p_{\alpha,\sigma}(\mathbf{x}) d\mathbf{x}$$

$$
\begin{aligned}
&= \int_{\mathcal{B}} \frac{p_{\alpha,\sigma}(\mathbf{x})}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{x})}\bigg[\mathcal{I}\bigg\{\frac{p_{\alpha,\sigma}(\mathbf{x})}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{x})} > \tau_n\bigg\} \\
&\quad + \mathcal{I}\bigg\{\frac{p_{\alpha,\sigma}(\mathbf{x})}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{x})} \le \tau_n\bigg\}\bigg]\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{x})d\mathbf{x}
\end{aligned}
$$

where $\mathcal{I}\{.\}$ is the indicator function. Consequently,

$$
\begin{aligned}
\tau_n\tilde{P}_{\alpha_1,\sigma_1,n}(\mathcal{A}) &\ge P_{\alpha,\sigma}(\mathcal{A}) - P_{\alpha,\sigma}\bigg(\mathcal{A}\cap\bigg\{\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)} > \tau_n\bigg\}\bigg) \\
&\ge P_{\alpha,\sigma}(\mathcal{A}) - P_{\alpha,\sigma}\bigg(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)} > \tau_n\bigg),
\end{aligned}
$$

which implies that

$$
\begin{aligned}
& P_{\alpha,\sigma}(|\hat{\sigma}_{1,n}^2\alpha_1^{2\nu} - \sigma^2\alpha^{2\nu}| > \varepsilon\alpha_1^{2\nu}) \\
\le\ & \tau_n\tilde{P}_{\alpha_1,\sigma_1,n}(|\hat{\sigma}_{1,n}^2\alpha_1^{2\nu} - \sigma^2\alpha^{2\nu}| > \varepsilon\alpha_1^{2\nu}) + P_{\alpha,\sigma}\bigg(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)} > \tau_n\bigg) \\
=\ & \tau_n\tilde{P}_{\alpha_1,\sigma_1,n}(|\hat{\sigma}_{1,n}^2 - \sigma_1^2| > \varepsilon) + P_{\alpha,\sigma}\bigg(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)} > \tau_n\bigg) \qquad (8)
\end{aligned}
$$

for all $\varepsilon, \tau_n > 0$. Lemmas 1 and 2 below use Bernstein-type inequalities to establish exponential bounds for each of the two terms on the right hand side of (8). The proofs of these lemmas are deferred to the Appendix.

Lemma 1. *Let α_1, $\mathbf{X}_n$ and $\hat{\sigma}_{1,n}$ be defined as in (6). Suppose (4) holds. Then for any constant $\varepsilon > 0$, we have*

$$
\tilde{P}_{\alpha_1,\sigma_1,n}(|\hat{\sigma}_{1,n}^2 - \sigma_1^2| > \varepsilon) < 2\exp\bigg[-\frac{\varepsilon^2 n}{4\sigma_1^2(\sigma_1^2 + 4\varepsilon)}\bigg].
$$

Next we observe that there exists a $n\times n$ non-singular matrix $\mathbf{U}$ such that

$$
\sigma^2\mathbf{U}'\mathbf{R}_\alpha\mathbf{U} = \mathbf{I}, \qquad \sigma_1^2\mathbf{U}'\tilde{\mathbf{R}}_{\alpha_1,n}\mathbf{U} = \mathbf{L}_n, \tag{9}
$$

where $\mathbf{I}$ is the $n\times n$ identity matrix and $\mathbf{L}_n$ is a $n\times n$ diagonal matrix with diagonal elements $(\mathbf{L}_n)_{i,i} = \lambda_{i,n} > 0, i = 1,\dots,n$.

Lemma 2. *With the notation of (9), suppose $\tau_n > 0$, $0 < c_n < 1$, C_n and $\tilde{C}_n$ are constants (which may depend on n) such that for $n = 1,2,\dots,$*

$$
\begin{aligned}
\min_{i\in\{1,\dots,n\}} \frac{1}{\lambda_{i,n}} &\ge c_n, \\
\max_{i\in\{1,\dots,n\}} |\lambda_{i,n}^{-1} - 1| &\le C_n, \\
\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)^2 &\le \tilde{C}_n, \\
C_n^* &= \max\bigg\{\frac{1}{2}, \frac{c_n - 1 - \log(c_n)}{(1-c_n)^2}\bigg\}, \\
2\log(\tau_n) &> C_n^*\tilde{C}_n.
\end{aligned}
$$

Then

$$P_{\alpha,\sigma}\left(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)} > \tau_n\right) < \exp\left\{-\frac{[2\log(\tau_n) - C_n^*\tilde{C}_n]^2}{4\tilde{C}_n + 16C_n[2\log(\tau_n) - C_n^*\tilde{C}_n]}\right\}.$$

4. Spectral analysis

This section is motivated by the equivalence of Gaussian measures theory as developed in Chapter 3 of Ibragimov and Rozanov [9]. However, as noted in the Introduction, these ideas have to be modified because the Gaussian measures considered in Theorems 1 and 2 are not equivalent if $\gamma_n \to 0$ as $n \to \infty$.

Let $d \le 3$ and $X(\mathbf{x}_1), \ldots, X(\mathbf{x}_n)$, with $\mathbf{x}_1, \ldots, \mathbf{x}_n \in [0,T]^d$, be as in (3). Define $\varphi_k(\mathbf{w}) = e^{\imath \mathbf{w}'\mathbf{x}_k}$, $\forall \mathbf{w} \in \mathbb{R}^d, k = 1, \ldots, n$, where $\imath = \sqrt{-1}$. Let L_n^0 be the (real) linear space of functions $\varphi : \mathbb{R}^d \to \mathbb{C}$ of the form

$$\varphi(\mathbf{w}) = \sum_{k=1}^n c_k \varphi_k(\mathbf{w}), \qquad \forall \mathbf{w} \in \mathbb{R}^d,$$

where $c_1, \ldots, c_n$ are real-valued constants, and $f_{\alpha,\sigma}$ be the spectral density as in (2). We can regard L_n^0 as a subspace of a (real) Hilbert space $L_n(f_{\alpha,\sigma})$ with inner product

$$\langle \varphi, \psi \rangle_{f_{\alpha,\sigma}} = \int_{\mathbb{R}^d} \varphi(\mathbf{w})\overline{\psi(\mathbf{w})} f_{\alpha,\sigma}(\mathbf{w}) d\mathbf{w}, \qquad \forall \varphi, \psi \in L_n(f_{\alpha,\sigma}),$$

and norm $\|\varphi\|_{f_{\alpha,\sigma}} = \sqrt{\langle \varphi, \varphi \rangle_{f_{\alpha,\sigma}}}$. Without loss of generality, we shall take $L_n(f_{\alpha,\sigma})$ to be the closure of the space L_n^0 with respect to the above inner product.

In an analogous manner, let $\tilde{f}_{\alpha_1,\sigma_1,n}$ be the spectral measure of the mean-zero Gaussian random field $X(.)$ with covariance function given by (4). Let $L_n(\tilde{f}_{\alpha_1,\sigma_1,n})$ be the closure of the space L_n^0 with respect to the inner product

$$\langle \varphi, \psi \rangle_{\tilde{f}_{\alpha_1,\sigma_1,n}} = \int_{\mathbb{R}^d} \varphi(\mathbf{w})\overline{\psi(\mathbf{w})} \tilde{f}_{\alpha_1,\sigma_1,n}(\mathbf{w}) d\mathbf{w} \in \mathbb{R},$$

and norm $\|\varphi\|_{\tilde{f}_{\alpha_1,\sigma_1,n}} = \sqrt{\langle \varphi, \varphi \rangle_{\tilde{f}_{\alpha_1,\sigma_1,n}}}$ for all $\varphi, \psi \in L_n(\tilde{f}_{\alpha_1,\sigma_1,n})$. Define

$$\begin{aligned}
\mu_{j,k} &= \langle \varphi_j, \varphi_k \rangle_{f_{\alpha,\sigma}}, \qquad \forall 1 \le j,k \le n, \\
\mathbf{A}_j &= \begin{pmatrix} \mu_{1,1} & \mu_{1,2} & \cdots & \mu_{1,j} \\ \mu_{2,1} & \mu_{2,2} & \cdots & \mu_{2,j} \\ \vdots & \vdots & & \vdots \\ \mu_{j-1,1} & \mu_{j-1,2} & \cdots & \mu_{j-1,j} \\ \varphi_1 & \varphi_2 & \cdots & \varphi_j \end{pmatrix}, \qquad \forall j = 1, \ldots, n, \\
\phi_j &= |\mathbf{A}_j|, \qquad \forall j = 1, \ldots, n,
\end{aligned}$$

where $|\mathbf{A}_j|$ denotes the determinant of the square matrix $\mathbf{A}_j$. This implies that $\phi_1 = \varphi_1$ and

$$\phi_j = \sum_{k=1}^{j} (-1)^{j+k} |\mathbf{A}_j^{-j,-k}| \varphi_k, \qquad \forall j = 2, \dots, n \tag{10}$$

where $|\mathbf{A}_j^{-j,-k}|$ is the determinant of $\mathbf{A}_j$ with row j and column k deleted. Immediate consequences are

$$\begin{aligned}
\langle \varphi_k, \varphi_l \rangle_{f_{\alpha,\sigma}} &= \int_{\mathbb{R}^d} e^{\mathbf{i}\mathbf{w}'(\mathbf{x}_k - \mathbf{x}_l)} f_{\alpha,\sigma}(\mathbf{w}) d\mathbf{w} \\
&= \sigma^2 K_\alpha(\mathbf{x}_k - \mathbf{x}_l) \\
&= (\sigma^2 \mathbf{R}_\alpha)_{k,l}, \\
\langle \varphi_k, \varphi_l \rangle_{\tilde{f}_{\alpha_1,\sigma_1,n}} &= \sigma_1^2 \tilde{K}_{\alpha_1,n}(\mathbf{x}_k - \mathbf{x}_l) \\
&= (\sigma_1^2 \tilde{\mathbf{R}}_{\alpha_1,n})_{k,l}, \qquad \forall 1 \le k, l \le n.
\end{aligned}$$

Since φ_k, $k = 1, \dots, n$, are linearly independent functions, we observe from Lemma 6.3.1 of Andrews, Askey and Roy [2] that $\langle \phi_j, \phi_k \rangle_{f_{\alpha,\sigma}} = 0$ for all $1 \le j \ne k \le n$. We observe from (10) that

$$(\phi_1, \dots, \phi_n)' = \mathbf{T}(\varphi_1, \dots, \varphi_n)',$$

where $\mathbf{T}$ is a $n \times n$ lower triangular matrix whose elements are

$$\mathbf{T}_{j,k} = (-1)^{j+k} |\mathbf{A}_j^{-j,-k}|, \quad \forall 1 \le k \le j \le n.$$

Then

$$\begin{aligned}
\sigma^2 \mathbf{O}\mathbf{D}^{-1}\mathbf{T}\mathbf{R}_\alpha \mathbf{T}'\mathbf{D}^{-1}\mathbf{O}' &= \mathbf{I}, \\
\sigma_1^2 \mathbf{O}\mathbf{D}^{-1}\mathbf{T}\tilde{\mathbf{R}}_{\alpha_1,n} \mathbf{T}'\mathbf{D}^{-1}\mathbf{O}' &= \mathbf{L}_n,
\end{aligned} \tag{11}$$

where $\mathbf{D}$ is a $n \times n$ diagonal matrix with elements $\mathbf{D}_{i,i} = \|\phi_i\|_{f_{\alpha,\sigma}}$, $i = 1, \dots, n$, $\mathbf{O}$ is a suitably chosen $n \times n$ orthogonal matrix and $\mathbf{L}_n$ is a $n \times n$ diagonal matrix as in (9). Define

$$(\psi_1, \dots, \psi_n)' = \mathbf{O}\mathbf{D}^{-1}\mathbf{T}(\varphi_1, \dots, \varphi_n)'. \tag{12}$$

Then

$$\begin{aligned}
\langle \psi_j, \psi_k \rangle_{f_{\alpha,\sigma}} &= \delta_{j,k}, \\
\langle \psi_j, \psi_k \rangle_{\tilde{f}_{\alpha_1,\sigma_1,n}} &= \lambda_{j,n}\delta_{j,k}, \qquad \forall j, k = 1, \dots, n,
\end{aligned} \tag{13}$$

where $\lambda_{j,n} = (\mathbf{L}_n)_{j,j}$ and $\delta_{j,k} = 1$ if $j = k$ and is 0 otherwise. Let $m = \lfloor \nu + d/2 \rfloor + 1$ and $\kappa = (\nu + d/2)/(2m)$ where $\lfloor . \rfloor$ denotes the greatest integer function. Define

$$\begin{aligned}
c_0(\mathbf{x}) &= \|\mathbf{x}\|^{\kappa - d} \mathcal{I}\{\|\mathbf{x}\| \le 1\}, \qquad \forall \mathbf{x} \in \mathbb{R}^d, \\
\xi_0(\mathbf{w}) &= \int_{\mathbb{R}^d} e^{-\imath \mathbf{x}'\mathbf{w}} c_0(\mathbf{x}) d\mathbf{x}, \qquad \forall \mathbf{w} \in \mathbb{R}^d.
\end{aligned} \tag{14}$$

Since $0 < \kappa < 1/2$, it follows from Lemma 6 (see Appendix) that $\xi_0 : \mathbb{R}^d \to \mathbb{R}$ is a continuous, isotropic, strictly positive function and $\xi_0(\mathbf{w}) \asymp \|\mathbf{w}\|^{-\kappa}$ as $\|\mathbf{w}\| \to \infty$.

Let $c_1 = c_0 * \cdots * c_0$ denote the $2m$-fold convolution of the function c_0 with itself. Then $\text{supp}(c_1) \subseteq [-2m, 2m]^d$ and

$$\begin{aligned}
\xi_1(\mathbf{w}) &= \int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} c_1(\mathbf{x}) d\mathbf{x} \\
&= \int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} c_0(\mathbf{x} - \mathbf{y}_1) c_0(\mathbf{y}_1 - \mathbf{y}_2) \dots c_0(\mathbf{y}_{2m-2} - \mathbf{y}_{2m-1}) \\
&\qquad \times c_0(\mathbf{y}_{2m-1}) d\mathbf{y}_1 \dots d\mathbf{y}_{2m-1} d\mathbf{x} \\
&= \left[\int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} c_0(\mathbf{x}) d\mathbf{x} \right]^{2m} \\
&= \xi_0(\mathbf{w})^{2m}, \qquad \forall \mathbf{w} \in \mathbb{R}^d. \qquad (15)
\end{aligned}$$

This implies that $\xi_1 : \mathbb{R}^d \to \mathbb{R}$ is a continuous, isotropic and strictly positive function. It follows from (2) and Lemma 6 that there exist constants $c_{\xi_1} > 0$ and C_{ξ_1} (not depending on $\mathbf{w}$) such that

$$c_{\xi_1} \le \frac{f_{\alpha,\sigma}(\mathbf{w})}{\xi_1(\mathbf{w})^2} \le C_{\xi_1}, \qquad \forall \mathbf{w} \in \mathbb{R}^d. \qquad (16)$$

For simplicity we write

$$\eta_n(\mathbf{w}) = \frac{\tilde{f}_{\alpha_1,\sigma_1,n}(\mathbf{w}) - f_{\alpha,\sigma}(\mathbf{w})}{\xi_1(\mathbf{w})^2}, \qquad \forall \mathbf{w} \in \mathbb{R}^d, \qquad (17)$$

and assume that $\eta_n : \mathbb{R}^d \to \mathbb{R}$ is square-integrable. Lemma 4 in Section 5 shows that this is indeed true under the assumptions of Theorems 1 or 2. It follows from the theory of Fourier transforms of $L^2(\mathbb{R}^d)$ functions (cf. Stein and Weiss [11], Chapter 1) that there exists a square-integrable function $g_n : \mathbb{R}^d \to \mathbb{C}$ such that

$$\int_{\mathbb{R}^d} |\eta_n(\mathbf{w}) - \hat{g}_{n,k}(\mathbf{w})|^2 d\mathbf{w} \to 0, \qquad \text{as } k \to \infty,$$

where

$$\hat{g}_{n,k}(\mathbf{w}) = \int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} g_n(\mathbf{x}) \mathcal{I}\{|\mathbf{x}|_{\max} \le k\} d\mathbf{x}, \qquad \forall \mathbf{w} \in \mathbb{R}^d, \qquad (18)$$

and $|\mathbf{x}|_{\max} = \max_{1 \le j \le d} |x_j|$ whenever $\mathbf{x} = (x_1, \dots, x_d)' \in \mathbb{R}^d$. Also

$$\begin{aligned}
&\int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} \left[\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g_n(\mathbf{x} - \mathbf{s}) \mathcal{I}\{|\mathbf{x} - \mathbf{s}|_{\max} \le k\} c_1(\mathbf{s} - \mathbf{t}) c_1(\mathbf{t}) d\mathbf{s} d\mathbf{t} \right] d\mathbf{x} \\
&= \left[\int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} g_n(\mathbf{x}) \mathcal{I}\{|\mathbf{x}|_{\max} \le k\} d\mathbf{x} \right] \left[\int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{s}} c_1(\mathbf{s}) d\mathbf{s} \right] \left[\int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{t}} c_1(\mathbf{t}) d\mathbf{t} \right] \\
&= \hat{g}_{n,k}(\mathbf{w}) \xi_1(\mathbf{w})^2, \qquad \forall \mathbf{w} \in \mathbb{R}^d. \qquad (19)
\end{aligned}$$

Let

$$b(\mathbf{x},\mathbf{y}) \quad = \quad E_{\tilde{f}_{\alpha_1,\sigma_1,n}} X(\mathbf{x})X(\mathbf{y}) - E_{f_{\alpha,\sigma}} X(\mathbf{x})X(\mathbf{y}), \qquad \forall \mathbf{x},\mathbf{y} \in [0,T]^d,$$

where $E_{\tilde{f}_{\alpha_1,\sigma_1,n}}$ and $E_{f_{\alpha,\sigma}}$ denote expectation with respect to the probability measures defined by the spectral densities $\tilde{f}_{\alpha_1,\sigma_1,n}$ and $f_{\alpha,\sigma}$ respectively. Then for any $\mathbf{x},\mathbf{y} \in [0,T]^d$, we have

$$\begin{aligned} b(\mathbf{x},\mathbf{y}) &= \int_{\mathbb{R}^d} e^{\imath(\mathbf{x}-\mathbf{y})'\mathbf{w}} [\tilde{f}_{\alpha_1,\sigma_1,n}(\mathbf{w}) - f_{\alpha,\sigma}(\mathbf{w})] d\mathbf{w} \\ &= \int_{\mathbb{R}^d} e^{\imath(\mathbf{x}-\mathbf{y})'\mathbf{w}} \eta_n(\mathbf{w}) \xi_1(\mathbf{w})^2 d\mathbf{w} \\ &= \lim_{k\to\infty} \Big\{ \int_{\mathbb{R}^d} e^{\imath(\mathbf{x}-\mathbf{y})'\mathbf{w}} [\eta_n(\mathbf{w}) - \hat{g}_{n,k}(\mathbf{w})] \xi_1^2(\mathbf{w}) d\mathbf{w} \\ &\qquad + \int_{\mathbb{R}^d} e^{\imath(\mathbf{x}-\mathbf{y})'\mathbf{w}} \hat{g}_{n,k}(\mathbf{w}) \xi_1^2(\mathbf{w}) d\mathbf{w} \Big\} \\ &= \lim_{k\to\infty} \int_{\mathbb{R}^d} e^{\imath(\mathbf{x}-\mathbf{y})'\mathbf{w}} \hat{g}_{n,k}(\mathbf{w}) \xi_1^2(\mathbf{w}) d\mathbf{w}. \end{aligned} \tag{20}$$

From (19) and (20), we obtain via Fourier inversion,

$$\begin{aligned} & b(\mathbf{x},\mathbf{y}) \\ = \; & \lim_{k\to\infty} (2\pi)^d \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g_n(\mathbf{s}-\mathbf{t}) \mathcal{I}\{|\mathbf{s}-\mathbf{t}|_{\max} \le k\} c_1(\mathbf{x}-\mathbf{s}) c_1(\mathbf{y}-\mathbf{t}) d\mathbf{s} d\mathbf{t} \\ = \; & (2\pi)^d \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g_n(\mathbf{s}-\mathbf{t}) c_1(\mathbf{x}-\mathbf{s}) c_1(\mathbf{y}-\mathbf{t}) d\mathbf{s} d\mathbf{t} \qquad (21) \\ = \; & (2\pi)^d \int_{\mathbb{R}^{2d}} h_n(\mathbf{s},\mathbf{t}) c_1(\mathbf{x}-\mathbf{s}) c_1(\mathbf{y}-\mathbf{t}) d\mathbf{s} d\mathbf{t}, \qquad \forall \mathbf{x},\mathbf{y} \in [0,T]^d, \end{aligned}$$

where

$$h_n(\mathbf{s},\mathbf{t}) = g_n(\mathbf{s}-\mathbf{t}) \mathcal{I}\{|\mathbf{s}+\mathbf{t}|_{\max} \le 4m+2T\}, \qquad \forall \mathbf{s},\mathbf{t} \in \mathbb{R}^d.$$

We observe that

$$\begin{aligned} \int_{\mathbb{R}^{2d}} |h_n(\mathbf{s},\mathbf{t})|^2 d\mathbf{s} d\mathbf{t} &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |g_n(\mathbf{s}-\mathbf{t})|^2 \mathcal{I}\{|\mathbf{s}+\mathbf{t}|_{\max} \le 4m+2T\} d\mathbf{s} d\mathbf{t} \\ &= \frac{1}{2^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |g_n(\mathbf{s})|^2 \mathcal{I}\{|\mathbf{t}|_{\max} \le 4m+2T\} d\mathbf{s} d\mathbf{t} \\ &= (4m+2T)^d \int_{\mathbb{R}^d} |g_n(\mathbf{s})|^2 d\mathbf{s} < \infty. \end{aligned}$$

Also for $\mathbf{w},\mathbf{v} \in \mathbb{R}^d$,

$$\begin{aligned} & \int_{\mathbb{R}^{2d}} e^{-\imath(\mathbf{s}'\mathbf{w}+\mathbf{t}'\mathbf{v})} h_n(\mathbf{s},\mathbf{t}) \mathcal{I}\{|\mathbf{s}-\mathbf{t}|_{\max} \le k\} d\mathbf{s} d\mathbf{t} \\ = \; & \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-\imath[(\mathbf{s}-\mathbf{t})'(\mathbf{w}-\mathbf{v})/2 + (\mathbf{s}+\mathbf{t})'(\mathbf{w}+\mathbf{v})/2]} g_n(\mathbf{s}-t) \end{aligned}$$

$$
\begin{aligned}
&\qquad \times \mathcal{I}\{|\mathbf{s}-\mathbf{t}|_{\max} \le k, |\mathbf{s}+\mathbf{t}|_{\max} \le 4m+2T\} d\mathbf{s} d\mathbf{t} \\
&= \frac{1}{2^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-\imath[\mathbf{s}'(\mathbf{w}-\mathbf{v})/2+\mathbf{t}'(\mathbf{w}+\mathbf{v})/2]} g_n(\mathbf{s}) \\
&\qquad \times \mathcal{I}\{|\mathbf{s}|_{\max} \le k, |\mathbf{t}|_{\max} \le 4m+2T\} d\mathbf{s} d\mathbf{t} \\
&= \frac{1}{2^d} \int_{\mathbb{R}^d} e^{-\imath \mathbf{s}'(\mathbf{w}-\mathbf{v})/2} g_n(\mathbf{s}) \mathcal{I}\{|\mathbf{s}|_{\max} \le k\} d\mathbf{s} \\
&\qquad \times \int_{\mathbb{R}^d} e^{-\imath \mathbf{t}'(\mathbf{w}+\mathbf{v})/2} \mathcal{I}\{|\mathbf{t}|_{\max} \le 4m+2T\} d\mathbf{t} \\
&= \hat{g}_{n,k}\Big(\frac{\mathbf{w}-\mathbf{v}}{2}\Big) \theta\Big(\frac{\mathbf{w}+\mathbf{v}}{2}\Big), \qquad (22)
\end{aligned}
$$

where

$$
\theta(\mathbf{w}) = \frac{1}{2^d} \int_{\mathbb{R}^d} e^{-\imath \mathbf{t}'\mathbf{w}} \mathcal{I}\{|\mathbf{t}|_{\max} \le 4m+2T\} d\mathbf{t}, \qquad \forall \mathbf{w} \in \mathbb{R}^d. \qquad (23)
$$

Lemma 3. *Let θ be as in (23). Then $\theta : \mathbb{R}^d \to \mathbb{R}$ is continuous and $\int_{\mathbb{R}^d} \theta(\mathbf{w})^2 d\mathbf{w} < \infty$.*

Observing that

$$
\int_{\mathbb{R}^{2d}} e^{-\imath(\mathbf{s}'\mathbf{w}+\mathbf{t}'\mathbf{v})} c_1(\mathbf{s}) c_1(\mathbf{t}) d\mathbf{s} d\mathbf{t} = \xi_1(\mathbf{w}) \xi_1(\mathbf{v}),
$$

it follows from (18), (21), (22) and Fourier transform inversion that for $\mathbf{x}, \mathbf{y} \in [0, T]^d$,

$$
\begin{aligned}
&b(\mathbf{x}, \mathbf{y}) \\
&= \lim_{k\to\infty} (2\pi)^d \int_{\mathbb{R}^{2d}} h_n(\mathbf{s}, \mathbf{t}) \mathcal{I}\{|\mathbf{s}-\mathbf{t}|_{\max} \le k\} c_1(\mathbf{x}-\mathbf{s}) c_1(\mathbf{y}-\mathbf{t}) d\mathbf{s} d\mathbf{t} \\
&= \lim_{k\to\infty} \frac{1}{(2\pi)^d} \int_{\mathbb{R}^{2d}} e^{\imath(\mathbf{w}'\mathbf{x}+\mathbf{v}'\mathbf{y})} \hat{g}_{n,k}\Big(\frac{\mathbf{w}-\mathbf{v}}{2}\Big) \theta\Big(\frac{\mathbf{w}+\mathbf{v}}{2}\Big) \xi_1(\mathbf{w}) \xi_1(\mathbf{v}) d\mathbf{w} d\mathbf{v} \\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^{2d}} e^{\imath(\mathbf{w}'\mathbf{x}+\mathbf{v}'\mathbf{y})} \eta_n\Big(\frac{\mathbf{w}-\mathbf{v}}{2}\Big) \theta\Big(\frac{\mathbf{w}+\mathbf{v}}{2}\Big) \xi_1(\mathbf{w}) \xi_1(\mathbf{v}) d\mathbf{w} d\mathbf{v} \\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^{2d}} e^{\imath(\mathbf{w}'\mathbf{x}-\mathbf{v}'\mathbf{y})} \Big[\eta_n\Big(\frac{\mathbf{w}+\mathbf{v}}{2}\Big) \theta\Big(\frac{\mathbf{w}-\mathbf{v}}{2}\Big) \frac{\xi_1(\mathbf{w})}{f_{\alpha,\sigma}(\mathbf{w})} \frac{\xi_1(\mathbf{v})}{f_{\alpha,\sigma}(\mathbf{v})}\Big] \\
&\qquad \times f_{\alpha,\sigma}(\mathbf{w}) f_{\alpha,\sigma}(\mathbf{v}) d\mathbf{w} d\mathbf{v}. \qquad (24)
\end{aligned}
$$

Let $\{\psi_1, \ldots, \psi_n\}$ be as in (12). Then it follows from (13) and (24) that for $k = 1, \ldots, n$,

$$
\begin{aligned}
&\langle \psi_k, \psi_k \rangle_{\tilde{f}_{\alpha_1,\sigma_1,n}} - \langle \psi_k, \psi_k \rangle_{f_{\alpha,\sigma}} \\
&= \lambda_{k,n} - 1 \\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^{2d}} \psi_k(\mathbf{w}) \overline{\psi_k(\mathbf{v})} \Big[\eta_n\Big(\frac{\mathbf{w}+\mathbf{v}}{2}\Big) \theta\Big(\frac{\mathbf{w}-\mathbf{v}}{2}\Big) \frac{\xi_1(\mathbf{w})}{f_{\alpha,\sigma}(\mathbf{w})} \frac{\xi_1(\mathbf{v})}{f_{\alpha,\sigma}(\mathbf{v})}\Big] \\
&\qquad \times f_{\alpha,\sigma}(\mathbf{w}) f_{\alpha,\sigma}(\mathbf{v}) d\mathbf{w} d\mathbf{v}.
\end{aligned}
$$

Arguing as in Ibragimov and Rozanov [9], pages 83–85, we conclude from Bessel's inequality and (16) that

$$\begin{aligned}\sum_{k=1}^{n}(\lambda_{k,n}-1)^2 &\leq \frac{1}{(2\pi)^{2d}}\int_{\mathbb{R}^{2d}}\left|\eta_n\Big(\frac{\mathbf{w}+\mathbf{v}}{2}\Big)\theta\Big(\frac{\mathbf{w}-\mathbf{v}}{2}\Big)\right|^2 \\ &\quad\times\left[\frac{\xi_1(\mathbf{w})^2}{f_{\alpha,\sigma}(\mathbf{w})}\right]\left[\frac{\xi_1(\mathbf{v})^2}{f_{\alpha,\sigma}(\mathbf{v})}\right]d\mathbf{w}d\mathbf{v} \\ &\leq \frac{1}{2^d\pi^{2d}}\left\{\sup_{\mathbf{s}\in\mathbb{R}^d}\frac{\xi_1(\mathbf{s})^2}{f_{\alpha,\sigma}(\mathbf{s})}\right\}^2\left[\int_{\mathbb{R}^d}\eta_n(\mathbf{w})^2d\mathbf{w}\right]\left[\int_{\mathbb{R}^d}\theta(\mathbf{v})^2d\mathbf{v}\right] \\ &\leq \frac{C_{\xi_1}^2}{2^d\pi^{2d}c_{\xi_1}^2}\left[\int_{\mathbb{R}^d}\left|\frac{\tilde{f}_{\alpha_1,\sigma_1,n}(\mathbf{w})}{f_{\alpha,\sigma}(\mathbf{w})}-1\right|^2 d\mathbf{w}\right]\left[\int_{\mathbb{R}^d}\theta(\mathbf{v})^2d\mathbf{v}\right]. \qquad (25)\end{aligned}$$

5. Tapered covariance functions

Let $1 \leq d \leq 3$, $\sigma^2 K_\alpha$ be the Matérn covariance function as in (1) with spectral density $f_{\alpha,\sigma}$ as in (2). Suppose K_{tap} is an isotropic correlation function with $\text{supp}(K_{\text{tap}}) \subset [-1,1]^d$ and spectral density

$$f_{\text{tap}}(\mathbf{w}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-\imath\mathbf{w}'\mathbf{x}}K_{\text{tap}}(\mathbf{x})d\mathbf{x}. \qquad (26)$$

Let α_1, σ_1 be strictly positive constants such that $\sigma_1^2\alpha_1^{2\nu} = \sigma^2\alpha^{2\nu}$ and $\gamma_n \in (0,1]$, $n = 1,2,\ldots$, be a sequence of constants. We define the tapered covariance function to be

$$\tilde{K}_{\alpha_1,n}(\mathbf{x}) = K_{\alpha_1}(\mathbf{x})K_{\text{tap}}(\mathbf{x}/\gamma_n), \qquad \forall \mathbf{x}\in\mathbb{R}^d,$$

and its spectral density is given by

$$\tilde{f}_{\alpha_1,\sigma_1,n}(\mathbf{w}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-\imath\mathbf{w}'\mathbf{x}}\sigma_1^2\tilde{K}_{\alpha_1,n}(\mathbf{x})d\mathbf{x}, \qquad \forall\mathbf{w}\in\mathbb{R}^d. \qquad (27)$$

Lemma 4 below gives non-asymptotic bounds on the spectral density of the tapered covariance function. The proof is motivated by Kaufman, *et al.* [10].

Lemma 4. *Let $1 \leq d \leq 3$ and $f_{\alpha,\sigma}$, f_{tap}, $\tilde{f}_{\alpha_1,\sigma_1,n}$ be as in (2), (26), (27) respectively with $\sigma_1^2\alpha_1^{2\nu} = \sigma^2\alpha^{2\nu}$. Let ϵ, M, β be constants such that $\epsilon > \max\{d/4, 1-\nu\}$, $\beta \in (d/2, \min\{2, 2\epsilon\})$ and $\beta \leq 2\nu + d$. Suppose that*

$$f_{\text{tap}}(\mathbf{w}) \leq \frac{M}{(1+\|\mathbf{w}\|^2)^{\nu+d/2+\epsilon}}, \qquad \forall\mathbf{w}\in\mathbb{R}^d.$$

Then there exist constants $0 < c_f \leq 1$ and $1 < C_f < \infty$ (independent of $\mathbf{w}$ and n) such that

$$\left|\frac{\tilde{f}_{\alpha_1,\sigma_1,n}(\mathbf{w})}{f_{\alpha,\sigma}(\mathbf{w})}-1\right| \leq \frac{C_f}{\gamma_n^{2\nu+2\epsilon}(1+\|\mathbf{w}\|^\beta)}, \qquad (28)$$

$$c_f\gamma_n^d \leq \frac{\tilde{f}_{\alpha_1,\sigma_1,n}(\mathbf{w})}{f_{\alpha,\sigma}(\mathbf{w})} \leq \frac{C_f}{\gamma_n^{2\nu+2\epsilon}}, \qquad \forall\mathbf{w}\in\mathbb{R}^d. \qquad (29)$$

6. Proof of Theorem 1

Let $(\psi_1, \ldots, \psi_n)$ be as in (12). We observe from (13) and (29) that

$$\begin{aligned}\lambda_{k,n} &= \langle \psi_k, \psi_k \rangle_{\tilde{f}_{\alpha_1,\sigma_1,n}} = \int_{\mathbb{R}^d} |\psi_k(\mathbf{w})|^2 \tilde{f}_{\alpha_1,\sigma_1,n}(\mathbf{w}) d\mathbf{w}, \\ C_f^{-1} \gamma_n^{2\nu+2\epsilon} &\leq \lambda_{k,n}^{-1} \leq c_f^{-1} \gamma_n^{-d} \end{aligned} \tag{30}$$

Using Lemma 3, (25) and (28), we obtain

$$\begin{aligned}|\lambda_{k,n}^{-1} - 1| &\leq c_f^{-1} \gamma_n^{-d}, \\ \sum_{i=1}^n (\lambda_{i,n} - 1)^2 &\leq \frac{C_{\xi_1}^2}{2^d \pi^{2d} c_{\xi_1}^2} \Big[\int_{\mathbb{R}^d} \Big| \frac{\tilde{f}_{\alpha_1,\sigma_1,n}(\mathbf{w})}{f_{\alpha,\sigma}(\mathbf{w})} - 1 \Big|^2 d\mathbf{w} \Big] \Big[\int_{\mathbb{R}^d} \theta(\mathbf{v})^2 d\mathbf{v} \Big] \\ &\leq \frac{M_0}{\gamma_n^{4\nu+4\epsilon}}, \\ \sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)^2 &\leq c_f^{-2} \gamma_n^{-2d} \sum_{i=1}^n (\lambda_{i,n} - 1)^2 \leq \frac{c_f^{-2} M_0}{\gamma_n^{4\nu+4\epsilon+2d}}, \end{aligned} \tag{31}$$

for some constant M_0 (not depending on n). Motivated by Lemma 2, define

$$\begin{aligned} c_n &= C_f^{-1} \gamma_n^{2\nu+2\epsilon} < 1, \\ C_n^* &= \max\Big\{ \frac{1}{2}, \frac{c_n - 1 + \log(c_n^{-1})}{(1-c_n)^2} \Big\} \\ &= \max\Big\{ \frac{1}{2}, \frac{C_f^{-1} \gamma_n^{2\nu+2\epsilon} - 1 + \log(C_f) + 2(\nu+\epsilon)\log(\gamma_n^{-1})}{(1 - C_f^{-1}\gamma_n^{2\nu+2\epsilon})^2} \Big\}, \\ C_n &= c_f^{-1} \gamma_n^{-d}, \\ \tilde{C}_n &= \frac{c_f^{-2} M_0}{\gamma_n^{4\nu+4\epsilon+2d}}. \end{aligned}$$

Let $a(4\nu + 4\epsilon + 2d) < a_1 < 1$ be a constant and $\tau_n = e^{n^{a_1}}$, $n = 1, 2 \ldots$. Lemma 1 implies that for $\varepsilon > 0$,

$$\begin{aligned} \tau_n \tilde{P}_{\alpha_1,\sigma_1,n}(|\hat{\sigma}_{1,n}^2 - \sigma_1^2| > \varepsilon) &\leq 2 \exp\Big[\log(\tau_n) - \frac{\varepsilon^2 n}{4\sigma_1^2(\sigma_1^2 + 4\varepsilon)} \Big] \\ &= 2 \exp\Big[n^{a_1} - \frac{\varepsilon^2 n}{4\sigma_1^2(\sigma_1^2 + 4\varepsilon)} \Big]. \end{aligned} \tag{32}$$

Since $C_n^* \tilde{C}_n \asymp \gamma_n^{-4\nu-4\epsilon-2d} \log(\gamma_n^{-1} + 1) \asymp n^{a(4\nu+4\epsilon+2d)} \log(n^a + 1)$, $C_n \asymp n^{ad}$ and $\tilde{C}_n \asymp n^{a(4\nu+4\epsilon+2d)}$ as $n \to \infty$, we observe from Lemma 2 that there exist constants $n_0, M_1, M_2 > 0$ (independent of n) such that

$$\begin{aligned} P_{\alpha,\sigma}\Big(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)} > \tau_n \Big) &< \exp\Big\{ -\frac{[2\log(\tau_n) - C_n^* \tilde{C}_n]^2}{4\tilde{C}_n + 16 C_n [2\log(\tau_n) - C_n^* \tilde{C}_n]} \Big\} \\ &\leq M_1 \exp(-M_2 n^{a_1 - ad}), \quad \forall n \geq n_0. \end{aligned} \tag{33}$$

It follows from (8), (32) and (33) that

$$\begin{aligned}&\sum_{n\geq n_0} P_{\alpha,\sigma}(|\hat{\sigma}^2_{1,n}\alpha_1^{2\nu}-\sigma^2\alpha^{2\nu}|>\varepsilon\alpha_1^{2\nu})\\ \leq\ &\sum_{n\geq n_0}\left\{2\exp\left[n^{a_1}-\frac{\varepsilon^2 n}{4\sigma^2(\sigma^2+4\varepsilon)}\right]+M_1\exp(-M_2 n^{a_1-ad})\right\}\ <\ \infty.\end{aligned}$$

Thus we conclude from the Borel-Cantelli lemma that $\hat{\sigma}^2_{1,n}\alpha_1^{2\nu}\to\sigma^2\alpha^{2\nu}$ as $n\to\infty$ with $P_{\alpha,\sigma}$ probability 1. This proves Theorem 1. □

7. Proof of Theorem 2

Since $\sigma_1^2\alpha_1^{2\nu}=\sigma^2\alpha^{2\nu}$, we have

$$\begin{aligned}&\sqrt{n}(\hat{\sigma}^2_{1,n}\alpha_1^{2\nu}-\sigma^2\alpha^{2\nu})\\ =\ &\sigma_1^2\alpha_1^{2\nu}\sqrt{n}(\frac{\hat{\sigma}^2_{1,n}}{\sigma_1^2}-1)\\ =\ &\frac{\sigma^2\alpha^{2\nu}}{\sqrt{n}}(\sigma_1^{-2}\mathbf{X}_n'\tilde{\mathbf{R}}^{-1}_{\alpha_1,n}\mathbf{X}_n-\sigma^{-2}\mathbf{X}_n'\mathbf{R}_\alpha^{-1}\mathbf{X}_n)+\frac{\sigma^2\alpha^{2\nu}}{\sqrt{n}}\left(\frac{1}{\sigma^2}\mathbf{X}_n'\mathbf{R}_\alpha^{-1}\mathbf{X}_n-n\right).\end{aligned}$$

With respect to the probability measure $P_{\alpha,\sigma}$, $\sigma^{-2}\mathbf{X}_n'\mathbf{R}_\alpha^{-1}\mathbf{X}_n\sim\chi_n^2$ and hence

$$\frac{\sigma^2\alpha^{2\nu}}{\sqrt{n}}\left(\frac{1}{\sigma^2}\mathbf{X}_n'\mathbf{R}_\alpha^{-1}\mathbf{X}_n-n\right)\to N(0,2(\sigma^2\alpha^{2\nu})^2)$$

as $n\to\infty$. Thus to prove Theorem 2, it suffices to show that

$$\frac{1}{\sqrt{n}}(\sigma_1^{-2}\mathbf{X}_n'\tilde{\mathbf{R}}^{-1}_{\alpha_1,n}\mathbf{X}_n-\sigma^{-2}\mathbf{X}_n'\mathbf{R}_\alpha^{-1}\mathbf{X}_n)\to 0$$

in $P_{\alpha,\sigma}$ probability as $n\to\infty$. We observe that

$$\frac{1}{\sqrt{n}}(\sigma_1^{-2}\mathbf{X}_n'\tilde{\mathbf{R}}^{-1}_{\alpha_1,n}\mathbf{X}_n-\sigma^{-2}\mathbf{X}_n'\mathbf{R}_\alpha^{-1}\mathbf{X}_n)=\frac{1}{\sqrt{n}}\sum_{k=1}^{n}(\lambda_{k,n}^{-1}-1)Y_k^2, \tag{34}$$

where $(Y_1,\dots,Y_n)'\sim N_n(0,\mathbf{I})$ under $P_{\alpha,\sigma}$, and $\lambda_{k,n}, k=1,\dots,n$, are as in (9) and (11).

Let $a>0$, $m_a=\lfloor a+d/2\rfloor+1$ and $a_0=(a+d/2)/(2m_a)$. Define

$$\begin{aligned}\tilde{c}_0(\mathbf{x})\ &=\ \|\mathbf{x}\|^{a_0-d}\mathcal{I}\{\|\mathbf{x}\|\leq 1\},\qquad \forall\mathbf{x}\in\mathbb{R}^d,\\ \tilde{\xi}_0(\mathbf{w})\ &=\ \int_{\mathbb{R}^d}e^{-\imath\mathbf{x}'\mathbf{w}}\tilde{c}_0(\mathbf{x})d\mathbf{x},\qquad \forall\mathbf{w}\in\mathbb{R}^d.\end{aligned}$$

Since $0<a_0<1/2$, it follows from Lemma 6 (see Appendix) that $\tilde{\xi}_0:\mathbb{R}^d\to\mathbb{R}$ is a continuous, strictly positive function and $\tilde{\xi}_0(\mathbf{w})\asymp\|\mathbf{w}\|^{-a_0}$ as $\|\mathbf{w}\|\to\infty$.

Let $\tilde{c}_1 = \tilde{c}_0 * \cdots * \tilde{c}_0$ denote the $2m_a$-fold convolution of $\tilde{c}_0$ with itself. Then $\mathrm{supp}(\tilde{c}_1) \subseteq \{\mathbf{x} : \|\mathbf{x}\| \leq 2m_a\}$ and

$$\tilde{\xi}_1(\mathbf{w}) = \int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} \tilde{c}_1(\mathbf{x}) d\mathbf{x} = \tilde{\xi}_0(\mathbf{w})^{2m_a}, \qquad \forall \mathbf{w} \in \mathbb{R}^d.$$

Let $0 < \varepsilon_n \leq 1$ be a constant such that $\varepsilon_n \to 0$ as $n \to \infty$. Define

$$e_n(\mathbf{x}) = \frac{1}{C_e \varepsilon_n^d} \tilde{c}_1\left(\frac{\mathbf{x}}{\varepsilon_n}\right), \qquad \forall \mathbf{x} \in \mathbb{R}^d, \tag{35}$$

where $C_e = \int_{\mathbb{R}^d} \tilde{c}_1(\mathbf{x}) d\mathbf{x}$. $e_n : \mathbb{R}^d \to \mathbb{R}$ is an *approximate identity* in Fourier Analysis (cf. Grafakos [8], page 24). This implies that $e_n(\mathbf{x}) \geq 0$ $\forall \mathbf{x} \in \mathbb{R}^d$, $\int_{\mathbb{R}^d} e_n(\mathbf{x}) d\mathbf{x} = 1$ and

$$\begin{aligned} \hat{e}_n(\mathbf{w}) &= \int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} e_n(\mathbf{x}) d\mathbf{x} \\ &= \frac{1}{C_e} \int_{\mathbb{R}^d} e^{-\imath \varepsilon_n \mathbf{w}'\mathbf{x}} \tilde{c}_1(\mathbf{x}) d\mathbf{x} \\ &= \frac{\tilde{\xi}_1(\varepsilon_n \mathbf{w})}{C_e}, \qquad \forall \mathbf{w} \in \mathbb{R}^d. \end{aligned}$$

Hence there exists a constant $C_{\hat{e}}$ (not depending on $\mathbf{w}$ and n) such that

$$|\hat{e}_n(\mathbf{w})| \leq \frac{C_{\hat{e}}}{(1 + \varepsilon_n \|\mathbf{w}\|)^{a+d/2}}, \qquad \forall \mathbf{w} \in \mathbb{R}^d. \tag{36}$$

Lemma 5. *With the assumptions of Theorem 2, let β_0 be a constant such that $0 < \beta_0 < \min\{4 - d, 4\epsilon - d, 4\nu + d\}$ and $\beta_0 \leq 2$. Let η_n, g_n, e_n be as in (17), (18), (35) respectively. Then there exists a constant C_{β_0} (not depending on n) such that*

$$\int_{\mathbb{R}^d} |e_n * g_n(\mathbf{x}) - g_n(\mathbf{x})|^2 d\mathbf{x} \leq \frac{C_{\beta_0} \varepsilon_n^{\beta_0}}{\gamma_n^{4\nu+4\epsilon}}.$$

Using (21) and observing that $\mathrm{supp}(c_1) \subseteq [-2m, 2m]^d$, we obtain for $\mathbf{x}, \mathbf{y} \in [0, T]^d$,

$$\begin{aligned} b(\mathbf{x}, \mathbf{y}) &= (2\pi)^d \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e_n * g_n(\mathbf{s} - \mathbf{t}) c_1(\mathbf{x} - \mathbf{s}) c_1(\mathbf{y} - \mathbf{t}) d\mathbf{s} d\mathbf{t} \\ &\quad + (2\pi)^d \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [g_n(\mathbf{s} - \mathbf{t}) - e_n * g_n(\mathbf{s} - \mathbf{t})] c_1(\mathbf{x} - \mathbf{s}) c_1(\mathbf{y} - \mathbf{t}) d\mathbf{s} d\mathbf{t} \\ &= (2\pi)^d \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e_n * g_n(\mathbf{s} - \mathbf{t}) c_1(\mathbf{x} - \mathbf{s}) c_1(\mathbf{y} - \mathbf{t}) d\mathbf{s} d\mathbf{t} \\ &\quad + (2\pi)^d \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h_n^*(\mathbf{s}, \mathbf{t}) c_1(\mathbf{x} - \mathbf{s}) c_1(\mathbf{y} - \mathbf{t}) d\mathbf{s} d\mathbf{t}, \end{aligned} \tag{37}$$

where

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

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

$$\int_{\mathbb{R}^d} |\eta_n^*(\mathbf{w}) - \hat{g}^*_{n,k}(\mathbf{w})|^2 d\mathbf{w} \to 0, \qquad \text{as } k \to \infty,$$

where

$$\hat{g}^*_{n,k}(\mathbf{w}) = \int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}}[g_n(\mathbf{x}) - e_n * g_n(\mathbf{x})]\mathcal{I}\{|\mathbf{x}|_{\max} \leq k\} d\mathbf{x}, \qquad \forall \mathbf{w} \in \mathbb{R}^d.$$

Thus we conclude as in (24) that

$$\begin{aligned} & (2\pi)^d \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h_n^*(\mathbf{s},\mathbf{t}) c_1(\mathbf{x}-\mathbf{s}) c_1(\mathbf{y}-\mathbf{t}) d\mathbf{s} d\mathbf{t} \\ = \quad & \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} e^{\imath(\mathbf{w}'\mathbf{x}-\mathbf{v}'\mathbf{y})}\eta_n^*\left(\frac{\mathbf{w}+\mathbf{v}}{2}\right)\theta\left(\frac{\mathbf{w}-\mathbf{v}}{2}\right)\xi_1(\mathbf{w})\xi_1(\mathbf{v}) d\mathbf{w} d\mathbf{v}. \qquad (38) \end{aligned}$$

Next we define

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

Then $h_n^{**} : \mathbb{R}^{2d} \to \mathbb{C}$ is square-integrable and

$$\begin{aligned} & \int_{\mathbb{R}^{2d}} e^{-\imath(\mathbf{w}'\mathbf{s}+\mathbf{v}'\mathbf{t})} h_n^{**}(\mathbf{s},\mathbf{t})\mathcal{I}\{|\mathbf{t}|_{\max} \leq k\} d\mathbf{s} d\mathbf{t} \\ = \quad & \int_{|\mathbf{u}|_{\max} \leq 2m+2m_a+T}\left[\int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{s}} e_n(\mathbf{s}-\mathbf{u}) d\mathbf{s}\right] \\ & \quad \times \left[\int_{\mathbb{R}^d} e^{-\imath \mathbf{v}'\mathbf{t}} g_n(\mathbf{u}-\mathbf{t})\mathcal{I}\{|\mathbf{t}|_{\max} \leq k\} d\mathbf{t}\right] d\mathbf{u} \\ = \quad & \int_{|\mathbf{u}|_{\max} \leq 2m+2m_a+T} e^{-\imath(\mathbf{w}'\mathbf{u}+\mathbf{v}'\mathbf{u})}\left[\int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{s}} e_n(\mathbf{s}) d\mathbf{s}\right] \\ & \times \left[\int_{\mathbb{R}^d} e^{-\imath \mathbf{v}'\mathbf{t}} g_n(-\mathbf{t})\mathcal{I}\{|\mathbf{t}+\mathbf{u}|_{\max} \leq k\} d\mathbf{t}\right] d\mathbf{u}. \end{aligned}$$

Consequently using Fourier inversion, we have for $\mathbf{x},\mathbf{y} \in [0,T]^d$,

$$\begin{aligned} & (2\pi)^d \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e_n * g_n(\mathbf{s}-\mathbf{t}) c_1(\mathbf{x}-\mathbf{s}) c_1(\mathbf{y}-\mathbf{t}) d\mathbf{s} d\mathbf{t} \\ = \quad & (2\pi)^d \int_{\mathbb{R}^{2d}} h_n^{**}(\mathbf{s},\mathbf{t}) c_1(\mathbf{x}-\mathbf{s}) c_1(\mathbf{y}-\mathbf{t}) d\mathbf{s} d\mathbf{t} \\ = \quad & \lim_{k\to\infty}\frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} e^{\imath(\mathbf{w}'\mathbf{x}+\mathbf{v}'\mathbf{y})}\xi_1(\mathbf{w})\xi_1(\mathbf{v})\left\{\int_{|\mathbf{u}|_{\max} \leq 2m+2m_a+T} e^{-\imath(\mathbf{w}'\mathbf{u}+\mathbf{v}'\mathbf{u})}\right. \\ & \left. \times\left[\int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{s}} e_n(\mathbf{s}) d\mathbf{s}\right]\left[\int_{\mathbb{R}^d} e^{-\imath \mathbf{v}'\mathbf{t}} g_n(-\mathbf{t})\mathcal{I}\{|\mathbf{t}+\mathbf{u}|_{\max} \leq k\} d\mathbf{t}\right] d\mathbf{u}\right\} d\mathbf{v} d\mathbf{w} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} e^{\imath(\mathbf{w}'\mathbf{x}-\mathbf{v}'\mathbf{y})}\xi_1(\mathbf{w})\xi_1(\mathbf{v})\Big\{\int_{|\mathbf{u}|_{\max}\le 2m+2m_a+T} e^{-\imath(\mathbf{w}'\mathbf{u}-\mathbf{v}'\mathbf{u})} \\
&\quad \times \hat{e}_n(\mathbf{w})\eta_n(\mathbf{v})d\mathbf{u}\Big\}d\mathbf{v}d\mathbf{w}. \qquad (39)
\end{aligned}
$$

It follows from (37), (38) and (39) that for $\mathbf{x},\mathbf{y}\in[0,T]^d$,

$$
\begin{aligned}
b(\mathbf{x},\mathbf{y}) &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} e^{\imath(\mathbf{w}'\mathbf{x}-\mathbf{v}'\mathbf{y})}\eta_n^*\Big(\frac{\mathbf{w}+\mathbf{v}}{2}\Big)\theta\Big(\frac{\mathbf{w}-\mathbf{v}}{2}\Big)\xi_1(\mathbf{w})\xi_1(\mathbf{v})d\mathbf{w}d\mathbf{v} \\
&\quad + \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} e^{\imath(\mathbf{w}'\mathbf{x}-\mathbf{v}'\mathbf{y})}\xi_1(\mathbf{w})\xi_1(\mathbf{v}) \\
&\quad \times\Big\{\int_{|\mathbf{u}|_{\max}\le 2m+2m_a+T} e^{-\imath(\mathbf{w}'\mathbf{u}-\mathbf{v}'\mathbf{u})}\hat{e}_n(\mathbf{w})\eta_n(\mathbf{v})d\mathbf{u}\Big\}d\mathbf{v}d\mathbf{w}.
\end{aligned}
$$

Let $\{\psi_1,\ldots,\psi_n\}$ be as in (12). Then for $k=1,\ldots,n$,

$$
\begin{aligned}
\langle\psi_k,\psi_k\rangle_{\tilde{f}_{\alpha_1,\sigma_1,n}} - \langle\psi_k,\psi_k\rangle_{f_{\alpha,\sigma}} &= \lambda_{k,n}-1 \\
&= \nu_{k,n}^{\dagger}+\nu_{k,n}^{\ddagger}, \quad \text{say}, \qquad (40)
\end{aligned}
$$

where

$$
\begin{aligned}
\nu_{k,n}^{\dagger} &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}}\psi_k(\mathbf{w})\overline{\psi_k(\mathbf{v})}\eta_n^*\Big(\frac{\mathbf{w}+\mathbf{v}}{2}\Big)\theta\Big(\frac{\mathbf{w}-\mathbf{v}}{2}\Big)\xi_1(\mathbf{w})\xi_1(\mathbf{v})d\mathbf{w}d\mathbf{v}, \\
\nu_{k,n}^{\ddagger} &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}}\psi_k(\mathbf{w})\overline{\psi_k(\mathbf{v})}\xi_1(\mathbf{w})\xi_1(\mathbf{v})\hat{e}_n(\mathbf{w})\eta_n(\mathbf{v}) \\
&\quad \times\Big\{\int_{|\mathbf{u}|_{\max}\le 2m+2m_a+T} e^{-\imath(\mathbf{w}'\mathbf{u}-\mathbf{v}'\mathbf{u})}d\mathbf{u}\Big\}d\mathbf{v}d\mathbf{w}.
\end{aligned}
$$

Now using Bessel's inequality, we have

$$
\begin{aligned}
\sum_{k=1}^n|\nu_{k,n}^{\ddagger}| &\le \frac{1}{(2\pi)^d}\sum_{k=1}^n\int_{|\mathbf{u}|_{\max}\le 2m+2m_a+T}\Big|\int_{\mathbb{R}^d} e^{-\imath\mathbf{w}'\mathbf{u}}\psi_k(\mathbf{w})\xi_1(\mathbf{w})\hat{e}_n(\mathbf{w})d\mathbf{w}\Big| \\
&\quad \times\Big|\int_{\mathbb{R}^d} e^{\imath\mathbf{v}'\mathbf{u}}\overline{\psi_k(\mathbf{v})}\xi_1(\mathbf{v})\eta_n(\mathbf{v})d\mathbf{v}\Big|d\mathbf{u} \\
&\le \frac{1}{2(2\pi)^d}\int_{|\mathbf{u}|_{\max}\le 2m+2m_a+T} \\
&\quad \times\sum_{k=1}^n\Big\{\Big|\int_{\mathbb{R}^d} e^{-\imath\mathbf{w}'\mathbf{u}}\psi_k(\mathbf{w})\frac{\xi_1(\mathbf{w})}{f_{\alpha,\sigma}(\mathbf{w})}\hat{e}_n(\mathbf{w})f_{\alpha,\sigma}(\mathbf{w})d\mathbf{w}\Big|^2 \\
&\quad + \Big|\int_{\mathbb{R}^d} e^{\imath\mathbf{v}'\mathbf{u}}\overline{\psi_k(\mathbf{v})}\frac{\xi_1(\mathbf{v})}{f_{\alpha,\sigma}(\mathbf{v})}\eta_n(\mathbf{v})f_{\alpha,\sigma}(\mathbf{v})d\mathbf{v}\Big|^2\Big\}d\mathbf{u} \\
&\le \frac{1}{2(2\pi)^d}\int_{|\mathbf{u}|_{\max}\le 2m+2m_a+T}\Big\{\int_{\mathbb{R}^d}\frac{\xi_1(\mathbf{w})^2}{f_{\alpha,\sigma}(\mathbf{w})}|\hat{e}_n(\mathbf{w})|^2d\mathbf{w}\Big\}d\mathbf{u}
\end{aligned}
$$

$$
\begin{aligned}
&+ \frac{1}{2(2\pi)^d} \int_{|\mathbf{u}|_{\max} \le 2m+2m_a+T} \left\{ \int_{\mathbb{R}^d} \frac{\xi_1(\mathbf{v})^2}{f_{\alpha,\sigma}(\mathbf{v})} \eta_n(\mathbf{v})^2 d\mathbf{v} \right\} d\mathbf{u} \\
\le\ & \frac{1}{2(2\pi)^d} \left\{ \sup_{\mathbf{s}\in\mathbb{R}^d} \frac{\xi_1(\mathbf{s})^2}{f_{\alpha,\sigma}(\mathbf{s})} \right\} \int_{|\mathbf{u}|_{\max} \le 2m+2m_a+T} d\mathbf{u} \\
& \times \left[\int_{\mathbb{R}^d} |\hat{e}_n(\mathbf{w})|^2 d\mathbf{w} + \int_{\mathbb{R}^d} \eta_n(\mathbf{v})^2 d\mathbf{v} \right],
\end{aligned}
$$

and

$$
\begin{aligned}
\sum_{k=1}^n |\nu_{k,n}^\dagger|^2 \ &\le\ \frac{1}{(2\pi)^{2d}} \left\{ \sup_{\mathbf{s}\in\mathbb{R}^d} \frac{\xi_1(\mathbf{s})^2}{f_{\alpha,\sigma}(\mathbf{s})} \right\}^2 \int_{\mathbb{R}^{2d}} \left| \eta_n^*\left(\frac{\mathbf{w}+\mathbf{v}}{2}\right) \theta\left(\frac{\mathbf{w}-\mathbf{v}}{2}\right) \right|^2 d\mathbf{w} d\mathbf{v} \\
&=\ \frac{1}{2^d \pi^{2d}} \left\{ \sup_{\mathbf{s}\in\mathbb{R}^d} \frac{\xi_1(\mathbf{s})^2}{f_{\alpha,\sigma}(\mathbf{s})} \right\}^2 \int_{\mathbb{R}^d} |\eta_n^*(\mathbf{w})|^2 d\mathbf{w} \int_{\mathbb{R}^d} |\theta(\mathbf{v})|^2 d\mathbf{v}.
\end{aligned}
$$

Consequently we observe from (16), (28), (36) and Lemmas 3, 5 that there exists a constant C (not depending on n) such that

$$
\sum_{k=1}^n |\nu_{k,n}^\ddagger| \le C\left(\frac{1}{\varepsilon_n^{2a+d}} + \frac{1}{\gamma_n^{4\nu+4\epsilon}} \right), \qquad \sum_{k=1}^n |\nu_{k,n}^\dagger|^2 \le \frac{C\varepsilon_n^{\beta_0}}{\gamma_n^{4\nu+4\epsilon}}. \tag{41}
$$

Now using (41), we have

$$
\sum_{k=1}^n |\nu_{k,n}^\dagger| \le \left(n \sum_{k=1}^n |\nu_{k,n}^\dagger|^2 \right)^{1/2} \le \sqrt{\frac{Cn\varepsilon_n^{\beta_0}}{\gamma_n^{4\nu+4\epsilon}}}. \tag{42}
$$

We conclude from (40), (41) and (42) that

$$
\sum_{k=1}^n |\lambda_{k,n} - 1| \le \sum_{k=1}^n (|\nu_{k,n}^\dagger| + |\nu_{k,n}^\ddagger|) \le \sqrt{\frac{Cn\varepsilon_n^{\beta_0}}{\gamma_n^{4\nu+4\epsilon}}} + C\left(\frac{1}{\varepsilon_n^{2a+d}} + \frac{1}{\gamma_n^{4\nu+4\epsilon}} \right). \tag{43}
$$

Finally for any constant $\delta > 0$, using Markov's inequality, (30), (34) and (43) we obtain

$$
\begin{aligned}
& P_{\alpha,\sigma}\left(\frac{1}{\sqrt{n}} |\sigma_1^{-2} \mathbf{X}_n' \tilde{\mathbf{R}}_{\alpha_1,n}^{-1} \mathbf{X}_n - \sigma^{-2} \mathbf{X}_n' \mathbf{R}_\alpha^{-1} \mathbf{X}_n| > \delta \right) \\
\le\ & P_{\alpha,\sigma}\left(\frac{1}{\sqrt{n}} \sum_{k=1}^n |\lambda_{k,n}^{-1} - 1| Y_k^2 > \delta \right) \\
\le\ & \frac{1}{\delta\sqrt{n}} \sum_{k=1}^n |\lambda_{k,n}^{-1} - 1| \\
\le\ & \frac{1}{\delta\sqrt{n}} \{ \max_{1\le i\le n} \lambda_{i,n}^{-1} \} \sum_{k=1}^n |\lambda_{k,n} - 1|
\end{aligned}
$$

$$\begin{aligned}
&\leq \frac{1}{c_f \delta n^{1/2}\gamma_n^d}\left[\sqrt{\frac{Cn\varepsilon_n^{\beta_0}}{\gamma_n^{4\nu+4\epsilon}}} + C\left(\frac{1}{\varepsilon_n^{2a+d}} + \frac{1}{\gamma_n^{4\nu+4\epsilon}}\right)\right] \\
&= \frac{C^{1/2}\varepsilon_n^{\beta_0/2}}{c_f\delta\gamma_n^{2\nu+2\epsilon+d}} + \frac{C}{c_f\delta n^{1/2}\gamma_n^d}\left(\frac{1}{\varepsilon_n^{2a+d}} + \frac{1}{\gamma_n^{4\nu+4\epsilon}}\right). \qquad (44)
\end{aligned}$$

From the definitions of b in Theorem 2 and β_0 in Lemma 5, we choose β_0 sufficiently close to $\min\{2, 4-d, 4\epsilon-d, 4\nu+d\}$ and a sufficiently close to 0 such that

$$\frac{2b(2\nu+2\epsilon+d)}{\beta_0} < \frac{1-2bd}{4a+2d}.$$

Now let b^* be a constant such that $2b(2\nu+2\epsilon+d)/\beta_0 < b^* < (1-2bd)/(4a+2d)$, and $\varepsilon_n = n^{-b^*}$, $n = 1, 2, \ldots$. Then

$$\begin{aligned}
\frac{\varepsilon_n^{\beta_0/2}}{\gamma_n^{2\nu+2\epsilon+d}} &\to 0, \\
\frac{1}{n^{1/2}\gamma_n^d}\left(\frac{1}{\varepsilon_n^{2a+d}} + \frac{1}{\gamma_n^{4\nu+4\epsilon}}\right) &\to 0,
\end{aligned}$$

as $n \to \infty$. It follows from (44) that

$$\left|\frac{1}{\sqrt{n}}(\sigma_1^{-2}\mathbf{X}_n'\tilde{\mathbf{R}}_{\alpha_1,n}^{-1}\mathbf{X}_n - \sigma^{-2}\mathbf{X}_n'\mathbf{R}_\alpha^{-1}\mathbf{X}_n)\right| \to 0$$

in $P_{\alpha,\sigma}$ probability as $n \to \infty$. This proves Theorem 2. □

8. Appendix

Proof of Lemma 1. Let $\mathbf{Y} = (Y_1, \ldots, Y_n)' = \sigma_1^{-1}\tilde{\mathbf{R}}_{\alpha_1,n}^{-1/2}\mathbf{X}_n$. Then $\mathbf{Y} \sim N_n(0, \mathbf{I})$. We observe that

$$\begin{aligned}
E[|Y_i^2 - E(Y_i^2)|^2] &= E(Y_i^2 - 1)^2 = 2, \\
E[|Y_i^2 - E(Y_i^2)|^k] &\leq E[(\max\{Y_i^2, E(Y_i^2)\})^k] \\
&\leq E(Y_i^{2k} + 1) \\
&= \frac{(2k)!}{2^k k!} + 1 \\
&\leq 8^{k-2}k!, \qquad \forall k = 3, 4, \ldots.
\end{aligned}$$

Consequently it follows from Bernstein's inequality (e.g., (7) of Bennet [3]) that

$$\tilde{P}_{\alpha_1,\sigma_1,n}\left(|\sum_{i=1}^n [Y_i^2 - E(Y_i^2)]| > \varepsilon\sqrt{2n}\right) < 2\exp\left[-\frac{\varepsilon^2}{2(1+8\varepsilon/\sqrt{2n})}\right], \qquad \forall \varepsilon > 0.$$

This implies that

$$\tilde{P}_{\alpha_1,\sigma_1,n}(|\hat{\sigma}_{1,n}^2 - \sigma_1^2| > \varepsilon) = \tilde{P}_{\alpha_1,\sigma_1,n}\left(\frac{1}{n}|\mathbf{X}_n'\tilde{\mathbf{R}}_{\alpha_1,n}^{-1}\mathbf{X}_n - E\mathbf{X}_n'\tilde{\mathbf{R}}_{\alpha_1,n}^{-1}\mathbf{X}_n| > \varepsilon\right)$$

$$
\begin{aligned}
&= \tilde{P}_{\alpha_1,\sigma_1,n}\left(\frac{1}{n}\Big|\sum_{i=1}^{n}[Y_i^2 - E(Y_i^2)]\Big| > \frac{\varepsilon}{\sigma_1^2}\right) \\
&= \tilde{P}_{\alpha_1,\sigma_1,n}\left(\Big|\sum_{i=1}^{n}[Y_i^2 - E(Y_i^2)]\Big| > \frac{\varepsilon\sqrt{n}}{\sigma_1^2\sqrt{2}}\sqrt{2n}\right) \\
&< 2\exp\left[-\frac{\varepsilon^2 n}{4\sigma_1^2(\sigma_1^2+4\varepsilon)}\right].
\end{aligned}
$$

This proves Lemma 1. □

Proof of Lemma 2. We observe that

$$
\begin{aligned}
\log\left(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)}\right) &= -\frac{1}{2}\log(|\sigma^2\mathbf{R}_\alpha|) - \frac{1}{2\sigma^2}\mathbf{X}_n'\mathbf{R}_\alpha^{-1}\mathbf{X}_n \\
&\quad + \frac{1}{2}\log(|\sigma_1^2\tilde{\mathbf{R}}_{\alpha_1,n}|) + \frac{1}{2\sigma_1^2}\mathbf{X}_n'\tilde{\mathbf{R}}_{\alpha_1,n}^{-1}\mathbf{X}_n.
\end{aligned}
$$

It follows from (9) that

$$
\begin{aligned}
E_{\alpha,\sigma}\left[\log\left(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)}\right)\right] &= -\frac{1}{2}\log\left(\frac{|\sigma^2\mathbf{R}_\alpha|}{|\sigma_1^2\tilde{\mathbf{R}}_{\alpha_1,n}|}\right) - \frac{n}{2} + \frac{\sigma^2}{2\sigma_1^2}\mathrm{tr}(\tilde{\mathbf{R}}_{\alpha_1,n}^{-1}\mathbf{R}_\alpha) \\
&= -\frac{1}{2}\sum_{i=1}^{n}\log(\lambda_{i,n}^{-1}) - \frac{n}{2} + \frac{1}{2}\sum_{i=1}^{n}\lambda_{i,n}^{-1} \\
&= \frac{1}{2}\sum_{i=1}^{n}[\lambda_{i,n}^{-1} - 1 - \log(\lambda_{i,n}^{-1})], \qquad (45)
\end{aligned}
$$

where $E_{\alpha,\sigma}$ denotes expectation with respect to the probability measure $P_{\alpha,\sigma}$. The right hand side of the last equality is a minimum when $\lambda_{i,n} = 1$ for all $i = 1, \ldots, n$. We further observe from (9) and (45) that

$$
\begin{aligned}
&P_{\alpha,\sigma}\left(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)} > \tau_n\right) \\
&= P_{\alpha,\sigma}\left(\log\left(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)}\right) - E_{\alpha,\sigma}\log\left(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)}\right)\right. \\
&\qquad \left. > \log(\tau_n) - E_{\alpha,\sigma}\log\left(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)}\right)\right) \\
&= P_{\alpha,\sigma}\left(\frac{1}{2}\mathbf{X}_n'\left(\frac{\tilde{\mathbf{R}}_{\alpha_1,n}^{-1}}{\sigma_1^2} - \frac{\mathbf{R}_\alpha^{-1}}{\sigma^2}\right)\mathbf{X}_n - \frac{1}{2}E_{\alpha,\sigma}\mathbf{X}_n'\left(\frac{\tilde{\mathbf{R}}_{\alpha_1,n}^{-1}}{\sigma_1^2} - \frac{\mathbf{R}_\alpha^{-1}}{\sigma^2}\right)\mathbf{X}_n\right. \\
&\qquad \left. > \log(\tau_n) - \frac{1}{2}\sum_{i=1}^{n}[\lambda_{i,n}^{-1} - 1 - \log(\lambda_{i,n}^{-1})]\right) \\
&\le P_{\alpha,\sigma}\Big(\mathbf{X}_n'(\mathbf{U}\mathbf{L}_n^{-1}\mathbf{U}' - \mathbf{U}\mathbf{U}')\mathbf{X}_n - E_{\alpha,\sigma}\mathbf{X}_n'(\mathbf{U}\mathbf{L}_n^{-1}\mathbf{U}' - \mathbf{U}\mathbf{U}')\mathbf{X}_n
\end{aligned}
$$

$$
\begin{aligned}
&> 2\log(\tau_n) - C_n^* \sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)^2\Big) \\
\leq\ & P_{\alpha,\sigma}\Big(\mathbf{X}_n'(\mathbf{U}\mathbf{L}_n^{-1}\mathbf{U}' - \mathbf{U}\mathbf{U}')\mathbf{X}_n - E_{\alpha,\sigma}\mathbf{X}_n'(\mathbf{U}\mathbf{L}_n^{-1}\mathbf{U}' - \mathbf{U}\mathbf{U}')\mathbf{X}_n \\
&> 2\log(\tau_n) - C_n^*\tilde{C}_n\Big). \qquad (46)
\end{aligned}
$$

Writing $\mathbf{Y} = (Y_1, \ldots, Y_n)' = \mathbf{U}'\mathbf{X}_n$, we have $\mathbf{Y} \sim N_n(0, \mathbf{I})$ under $P_{\alpha,\sigma}$. It follows from (46) that

$$
\begin{aligned}
& P_{\alpha,\sigma}\left(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)} > \tau_n\right) \\
\leq\ & P_{\alpha,\sigma}\left(\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)[Y_i^2 - E_{\alpha,\sigma}(Y_i^2)] > 2\log(\tau_n) - C_n^*\tilde{C}_n\right). \qquad (47)
\end{aligned}
$$

We further have

$$
\begin{aligned}
E_{\alpha,\sigma}\{|(\lambda_{i,n}^{-1} - 1)[Y_i^2 - E_{\alpha,\sigma}(Y_i^2)]|^2\} &= 2(\lambda_{i,n}^{-1} - 1)^2, \\
E_{\alpha,\sigma}\{|(\lambda_{i,n}^{-1} - 1)[Y_i^2 - E_{\alpha,\sigma}(Y_i^2)]|^k\} &\leq |\lambda_{i,n}^{-1} - 1|^k E_{\alpha,\sigma}[(\max\{Y_i^2, 1\})^k] \\
&\leq |\lambda_{i,n}^{-1} - 1|^k E_{\alpha,\sigma}(Y_i^{2k} + 1) \\
&\leq |\lambda_{i,n}^{-1} - 1|^k 8^{k-2} k! \\
&\leq (\lambda_{i,n}^{-1} - 1)^2 (8C_n)^{k-2} k!, \quad \forall k = 3, 4, \ldots.
\end{aligned}
$$

Consequently it follows from Bernstein's inequality (cf. (7) of Bennett [3]) that

$$
\begin{aligned}
& P_{\alpha,\sigma}\left(\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)[Y_i^2 - E_{\alpha,\sigma}(Y_i^2)] > \varepsilon\sqrt{2\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)^2}\right) \\
<\ & \exp\left\{-\frac{\varepsilon^2}{2[1 + 8C_n\varepsilon/\sqrt{2\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)^2}]}\right\}, \qquad \forall \varepsilon > 0. \qquad (48)
\end{aligned}
$$

From (47) and (48), we obtain

$$
\begin{aligned}
& P_{\alpha,\sigma}\left(\frac{p_{\alpha,\sigma}(\mathbf{X}_n)}{\tilde{p}_{\alpha_1,\sigma_1,n}(\mathbf{X}_n)} > \tau_n\right) \\
\leq\ & P\left(\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)[Y_i^2 - E(Y_i^2)] > \frac{2\log(\tau_n) - C_n^*\tilde{C}_n}{\sqrt{2\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)^2}}\right. \\
& \left.\times\sqrt{2\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)^2}\right) \\
<\ & \exp\left\{-\frac{[2\log(\tau_n) - C_n^*\tilde{C}_n]^2}{4\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)^2}\left[1 + 4C_n\frac{2\log(\tau_n) - C_n^*\tilde{C}_n}{\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)^2}\right]^{-1}\right\}
\end{aligned}
$$

$$
\begin{aligned}
&= \exp\left\{ -\frac{[2\log(\tau_n) - C_n^* \tilde{C}_n]^2}{4\sum_{i=1}^n (\lambda_{i,n}^{-1} - 1)^2 + 16 C_n [2\log(\tau_n) - C_n^* \tilde{C}_n]} \right\} \\
&\leq \exp\left\{ -\frac{[2\log(\tau_n) - C_n^* \tilde{C}_n]^2}{4\tilde{C}_n + 16 C_n [2\log(\tau_n) - C_n^* \tilde{C}_n]} \right\}.
\end{aligned}
$$

This proves Lemma 2. □

Proof of Lemma 3. We observe that

$$
\begin{aligned}
\theta(\mathbf{w}) &= \frac{1}{2^d} \prod_{j=1}^d \left[\int_{-(4m+2T)}^{4m+2T} e^{-\imath w_j t_j} dt_j \right] \\
&= \prod_{j=1}^d \left[\frac{1}{|w_j|} \int_0^{(4m+2T)|w_j|} \cos(t_j) dt_j \right], \qquad \forall \mathbf{w} = (w_1, \dots, w_d)' \in \mathbb{R}^d.
\end{aligned}
$$

Hence θ is a real-valued, continuous function on $\mathbb{R}^d$ and $\int_{\mathbb{R}^d} \theta(\mathbf{w})^2 d\mathbf{w} < \infty$. This proves Lemma 3. □

Proof of Lemma 4. First we observe that

$$
\begin{aligned}
& \int_{\mathbb{R}^d} e^{\imath \mathbf{x}'\mathbf{w}} \gamma_n^d \int_{\mathbb{R}^d} f_{\alpha_1,\sigma_1}(\mathbf{w} - \mathbf{v}) f_{\mathrm{tap}}(\gamma_n \mathbf{v}) d\mathbf{v} d\mathbf{w} \\
=& \int_{\mathbb{R}^d} e^{\imath \mathbf{x}'\mathbf{w}} \int_{\mathbb{R}^d} f_{\alpha_1,\sigma_1}\left(\mathbf{w} - \frac{\mathbf{v}}{\gamma_n}\right) f_{\mathrm{tap}}(\mathbf{v}) d\mathbf{v} d\mathbf{w} \\
=& \left[\int_{\mathbb{R}^d} e^{\imath \mathbf{x}'\mathbf{v}/\gamma_n} f_{\mathrm{tap}}(\mathbf{v}) d\mathbf{v} \right] \left[\int_{\mathbb{R}^d} e^{\imath \mathbf{x}'\mathbf{w}} f_{\alpha_1,\sigma_1}(\mathbf{w}) d\mathbf{w} \right] \\
=& \sigma_1^2 K_{\alpha_1}(\mathbf{x}) K_{\mathrm{tap}}\left(\frac{\mathbf{x}}{\gamma_n}\right), \qquad \forall \mathbf{x} \in \mathbb{R}^d.
\end{aligned}
$$

Hence using inverse Fourier transform, we have

$$
\begin{aligned}
\tilde{f}_{\alpha_1,\sigma_1,n}(\mathbf{w}) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-\imath \mathbf{w}'\mathbf{x}} \sigma_1^2 K_{\alpha_1}(\mathbf{x}) K_{\mathrm{tap}}(\frac{\mathbf{x}}{\gamma_n}) d\mathbf{x} \\
&= \gamma_n^d \int_{\mathbb{R}^d} f_{\alpha_1,\sigma_1}(\mathbf{w} - \mathbf{v}) f_{\mathrm{tap}}(\gamma_n \mathbf{v}) d\mathbf{v} \\
&= \gamma_n^d \int_{\mathbb{R}^d} f_{\alpha_1,\sigma_1}(\mathbf{v}) f_{\mathrm{tap}}(\gamma_n(\mathbf{w} - \mathbf{v})) d\mathbf{v}, \qquad \forall \mathbf{w} \in \mathbb{R}^d.
\end{aligned}
$$

Let $\beta_0 \in ((d + 2\nu + \beta)/(d + 2\nu + 2\epsilon), 1)$ be a constant and $\mathbf{u} \in \mathbb{R}^d$ such that $\|\mathbf{u}\| = 1$. For all $r > 0$, define

$$
\mathcal{N}_{r\mathbf{u}} = \{\mathbf{v} \in \mathbb{R}^d : \|r\mathbf{u} - \mathbf{v}\| \leq r^{\beta_0}\}.
$$

Then

$$
\frac{\tilde{f}_{\alpha_1,\sigma_1,n}(r\mathbf{u})}{f_{\alpha,\sigma}(r\mathbf{u})} - 1 = \frac{\gamma_n^d \int_{\mathbb{R}^d} f_{\alpha_1,\sigma_1}(\mathbf{v}) f_{\mathrm{tap}}(\gamma_n(r\mathbf{u} - \mathbf{v})) d\mathbf{v}}{f_{\alpha,\sigma}(r\mathbf{u})} - 1
$$

$$
\begin{aligned}
&= \frac{\gamma_n^d \int_{\mathcal{N}_{r\mathbf{u}}} f_{\alpha_1,\sigma_1}(\mathbf{v}) f_{\mathrm{tap}}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v}}{f_{\alpha,\sigma}(r\mathbf{u})} - 1 \\
&\quad + \frac{\gamma_n^d \int_{\mathcal{N}_{r\mathbf{u}}^c} f_{\alpha_1,\sigma_1}(\mathbf{v}) f_{\mathrm{tap}}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v}}{f_{\alpha,\sigma}(r\mathbf{u})}. \qquad (49)
\end{aligned}
$$

Note that

$$
\begin{aligned}
\sup_{\mathbf{v}\in\mathcal{N}_{r\mathbf{u}}^c} \gamma_n^d f_{\mathrm{tap}}(\gamma_n(r\mathbf{u}-\mathbf{v})) &\leq \sup_{\mathbf{v}\in\mathcal{N}_{r\mathbf{u}}^c} \frac{\gamma_n^d M}{(1+\gamma_n^2\|r\mathbf{u}-\mathbf{v}\|^2)^{\nu+d/2+\epsilon}} \\
&\leq \frac{\gamma_n^d M}{(1+\gamma_n^2 r^{2\beta_0})^{\nu+d/2+\epsilon}}, \qquad \forall r \geq 0.
\end{aligned}
$$

Since $\int_{\mathbb{R}^d} f_{\alpha_1,\sigma_1}(\mathbf{v})d\mathbf{v} = \sigma_1^2$, there exists a constant C_0 (independent of r and n) such that

$$
\begin{aligned}
\frac{\gamma_n^d \int_{\mathcal{N}_{r\mathbf{u}}^c} f_{\alpha_1,\sigma_1}(\mathbf{v}) f_{\mathrm{tap}}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v}}{f_{\alpha,\sigma}(r\mathbf{u})} &\leq \frac{\gamma_n^d \pi^{d/2} M \sigma_1^2 (\alpha^2+r^2)^{\nu+d/2}}{\sigma^2\alpha^{2\nu}(1+\gamma_n^2 r^{2\beta_0})^{\nu+d/2+\epsilon}} \\
&\leq \frac{C_0}{1+\gamma_n^{2\nu+2\epsilon} r^{\beta}}, \qquad \forall r \geq 0. \qquad (50)
\end{aligned}
$$

Next expanding $f_{\alpha_1,\sigma_1}(\mathbf{v})$ as a Taylor series about $r\mathbf{u}$, we obtain

$$
\begin{aligned}
f_{\alpha_1,\sigma_1}(\mathbf{v}) &= f_{\alpha_1,\sigma_1}(r\mathbf{u}) + (\mathbf{v}-r\mathbf{u})'[\nabla f_{\alpha_1,\sigma_1}(r\mathbf{u})] \\
&\quad + \frac{1}{2}(\mathbf{v}-r\mathbf{u})'[\nabla^2 f_{\alpha_1,\sigma_1}(\mathbf{m}_{\mathbf{v},r\mathbf{u}})](\mathbf{v}-r\mathbf{u})
\end{aligned}
$$

where $\mathbf{m}_{\mathbf{v},r\mathbf{u}}$ is a point on the line segment joining $\mathbf{v}$ and $r\mathbf{u}$, $[\nabla f_{\alpha_1,\sigma_1}(r\mathbf{u})]$ is the $d\times 1$ vector of first derivatives of f evaluated at $r\mathbf{u}$, and $[\nabla^2 f_{\alpha_1,\sigma_1}(\mathbf{m}_{\mathbf{v},r\mathbf{u}})]$ is the $d\times d$ matrix of second derivatives of f_{α_1,σ_1} evaluated at $\mathbf{m}_{\mathbf{v},r\mathbf{u}}$. Then

$$
\begin{aligned}
&\frac{\gamma_n^d \int_{\mathcal{N}_{r\mathbf{u}}} f_{\alpha_1,\sigma_1}(\mathbf{v}) f_{\mathrm{tap}}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v}}{f_{\alpha,\sigma}(r\mathbf{u})} - 1 \qquad (51) \\
&= \frac{f_{\alpha_1,\sigma_1}(r\mathbf{u})}{f_{\alpha,\sigma}(r\mathbf{u})}\left[\int_{\mathcal{N}_{r\mathbf{u}}} \gamma_n^d f_{\mathrm{tap}}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v} - 1\right] + \frac{f_{\alpha_1,\sigma_1}(r\mathbf{u})}{f_{\alpha,\sigma}(r\mathbf{u})} - 1 \\
&\quad + \frac{\gamma_n^d}{2f_{\alpha,\sigma}(r\mathbf{u})}\int_{\mathcal{N}_{r\mathbf{u}}} (\mathbf{v}-r\mathbf{u})'[\nabla^2 f_{\alpha_1,\sigma_1}(\mathbf{m}_{\mathbf{v},r\mathbf{u}})](\mathbf{v}-r\mathbf{u}) f_{\mathrm{tap}}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v}.
\end{aligned}
$$

Since $\sigma_1^2\alpha_1^{2\nu} = \sigma^2\alpha^{2\nu}$, we observe that there exists a constant C_1 (independent of r and n) such that

$$
\begin{aligned}
\left|\frac{f_{\alpha_1,\sigma_1}(r\mathbf{u})}{f_{\alpha,\sigma}(r\mathbf{u})} - 1\right| &= \left|\frac{(\alpha^2+r^2)^{\nu+d/2}}{(\alpha_1^2+r^2)^{\nu+d/2}} - 1\right| \\
&\leq \frac{C_1}{1+r^{2\nu+d}}, \qquad \forall r \geq 0,
\end{aligned}
$$

and

$$\begin{aligned}
1-\gamma_n^d \int_{\mathcal{N}_{r\mathbf{u}}} f_{\rm tap}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v} &= 1-\gamma_n^d \int_{\|\mathbf{v}\|\le r^{\beta_0}} f_{\rm tap}(\gamma_n\mathbf{v})d\mathbf{v} \\
&= \int_{\|\mathbf{v}\|>\gamma_n r^{\beta_0}} f_{\rm tap}(\mathbf{v})d\mathbf{v} \\
&\le \min\left\{1, M\int_{\|\mathbf{v}\|>\gamma_n r^{\beta_0}} \frac{1}{\|\mathbf{v}\|^{2\nu+d+2\epsilon}}d\mathbf{v}\right\} \\
&= \min\left\{1, \frac{2\pi^{d/2}M}{\Gamma(d/2)}\int_{\gamma_n r^{\beta_0}}^{\infty} \frac{x^{d-1}}{x^{2\nu+d+2\epsilon}}dx\right\} \\
&= \min\left\{1, \frac{\pi^{d/2}M}{\gamma_n^{2\nu+2\epsilon}(\nu+\epsilon)\Gamma(d/2)r^{2\beta_0(\nu+\epsilon)}}\right\} \\
&\le \frac{C_1}{1+\gamma_n^{2\nu+2\epsilon}r^{\beta}}, \qquad \forall r\ge 0. \qquad (52)
\end{aligned}$$

We observe from Kaufman, *et al.* [10], page 1554, that there exist constants C_2 and C_3 (independent of r and n) such that

$$\begin{aligned}
&\frac{\gamma_n^d}{2f_{\alpha,\sigma}(r\mathbf{u})}\int_{\mathcal{N}_{r\mathbf{u}}} (\mathbf{v}-r\mathbf{u})'[\nabla^2 f_{\alpha_1,\sigma_1}(\mathbf{m}_{\mathbf{v},r\mathbf{u}})](\mathbf{v}-r\mathbf{u})f_{\rm tap}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v} \\
\le\ & \frac{\gamma_n^d C_3(\alpha^2+r^2)^{\nu+d/2}}{[\alpha^2+(r-r^{\beta_0})^2]^{\nu+d/2+1}}\int_{\mathcal{N}_{r\mathbf{u}}} \|\mathbf{v}-r\mathbf{u}\|^2 f_{\rm tap}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v} \\
=\ & \frac{\gamma_n^d C_3(\alpha^2+r^2)^{\nu+d/2}}{[\alpha^2+(r-r^{\beta_0})^2]^{\nu+d/2+1}}\int_{\|\mathbf{v}\|\le r^{\beta_0}} \|\mathbf{v}\|^2 f_{\rm tap}(\gamma_n\mathbf{v})d\mathbf{v} \\
\le\ & \frac{\gamma_n^d C_3(\alpha^2+r^2)^{\nu+d/2}}{[\alpha^2+(r-r^{\beta_0})^2]^{\nu+d/2+1}}\int_{\|\mathbf{v}\|\le r^{\beta_0}} \|\mathbf{v}\|^2 \frac{M}{(1+\gamma_n^2\|\mathbf{v}\|^2)^{\nu+d/2+\epsilon}}d\mathbf{v} \\
\le\ & \frac{2\pi^{d/2}C_3M(\alpha^2+r^2)^{\nu+d/2}}{\gamma_n^2[\alpha^2+(r-r^{\beta_0})^2]^{\nu+d/2+1}\Gamma(d/2)}\int_0^{\infty} \frac{x^{d+1}}{(1+x^2)^{\nu+d/2+\epsilon}}dx \\
\le\ & \frac{C_2}{\gamma_n^2(1+r^2)}, \qquad \forall r\ge 0. \qquad (53)
\end{aligned}$$

Consequently, it follows from (51), (52) and (53) that

$$\begin{aligned}
&\left|\frac{\gamma_n^d\int_{\mathcal{N}_{r\mathbf{u}}} f_{\alpha_1,\sigma_1}(\mathbf{v})f_{\rm tap}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v}}{f_{\alpha,\sigma}(r\mathbf{u})}-1\right| \\
\le\ & \left(1+\frac{C_1}{1+r^{2\nu+d}}\right)\frac{C_1}{1+\gamma_n^{2\nu+2\epsilon}r^{\beta}}+\frac{C_1}{1+r^{2\nu+d}}+\frac{C_2}{\gamma_n^2(1+r^2)}. \qquad (54)
\end{aligned}$$

Finally from (49), (50) and (54), we obtain

$$\left|\frac{\tilde f_{\alpha_1,\sigma_1,n}(r\mathbf{u})}{f_{\alpha,\sigma}(r\mathbf{u})}-1\right|$$

$$\leq \quad \frac{C_0}{1+\gamma_n^{2\nu+2\epsilon}r^\beta} + \left(1+\frac{C_1}{1+r^{2\nu+d}}\right)\frac{C_1}{1+\gamma_n^{2\nu+2\epsilon}r^\beta} + \frac{C_1}{1+r^{2\nu+d}} + \frac{C_2}{\gamma_n^2(1+r^2)}$$

for all $r \geq 0$. This proves (28). It suffices to give a proof for the lower bound of (29) as the upper bound is an immediate consequence of (28). Define

$$\zeta(\gamma) = \frac{1}{\gamma^d}\int_{\|\mathbf{v}\|\leq\gamma} f_{\mathrm{tap}}(\mathbf{v})d\mathbf{v}, \qquad \forall\gamma\in[0,1].$$

We observe that $\zeta : [0,1] \to (0,\infty)$ is a continuous, strictly positive function. Hence $\underline{\zeta} = \min_{0\leq\gamma\leq 1}\zeta(\gamma) > 0$. For $r > 0$ and $\mathbf{u}\in\mathbb{R}^d$ with $\|\mathbf{u}\| = 1$, define

$$\tilde{\mathcal{N}}_{r\mathbf{u}} = \{\mathbf{v}\in\mathbb{R}^d : \|r\mathbf{u}-\mathbf{v}\|\leq 1\}.$$

Then using (2), we have

$$\begin{aligned}
\frac{\tilde{f}_{\alpha_1,\sigma_1,n}(r\mathbf{u})}{f_{\alpha,\sigma}(r\mathbf{u})} \quad &\geq \quad \frac{\gamma_n^d}{f_{\alpha,\sigma}(r\mathbf{u})}\int_{\tilde{\mathcal{N}}_{r\mathbf{u}}} f_{\alpha_1,\sigma_1}(\mathbf{v})f_{\mathrm{tap}}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v} \\
&\geq \quad \frac{\gamma_n^d \inf_{\mathbf{w}\in\tilde{\mathcal{N}}_{r\mathbf{u}}} f_{\alpha_1,\sigma_1}(\mathbf{w})}{f_{\alpha,\sigma}(r\mathbf{u})}\int_{\tilde{\mathcal{N}}_{r\mathbf{u}}} f_{\mathrm{tap}}(\gamma_n(r\mathbf{u}-\mathbf{v}))d\mathbf{v} \\
&= \quad \frac{\sigma_1^2\alpha_1^{2\nu}(\alpha^2+r^2)^{\nu+d/2}}{\sigma^2\alpha^{2\nu}[\alpha^2+(r+1)^2]^{\nu+d/2}}\int_{\|\mathbf{v}\|\leq\gamma_n} f_{\mathrm{tap}}(\mathbf{v})d\mathbf{v} \\
&\geq \quad \frac{\gamma_n^d\underline{\zeta}\sigma_1^2\alpha_1^{2\nu}(\alpha^2+r^2)^{\nu+d/2}}{\sigma^2\alpha^{2\nu}[\alpha^2+(r+1)^2]^{\nu+d/2}} \quad \geq \quad c_f\gamma_n^d,
\end{aligned}$$

for some constant $c_f > 0$ (not depending on $r\mathbf{u}$ and n). □

Proof of Lemma 5. Using Plancherel's theorem, we have for $\mathbf{y}\in\mathbb{R}^d$,

$$\begin{aligned}
\int_{\mathbb{R}^d} |g_n(\mathbf{x}-\mathbf{y}) - g_n(\mathbf{x})|^2 d\mathbf{x} \quad &= \quad \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} |e^{-\imath\mathbf{w}'\mathbf{y}}\eta_n(\mathbf{w}) - \eta_n(\mathbf{w})|^2 d\mathbf{w} \\
&= \quad \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} |(e^{-\imath\mathbf{w}'\mathbf{y}}-1)\eta_n(\mathbf{w})|^2 d\mathbf{w} \\
&\leq \quad \frac{2^{2-\beta_0}\|\mathbf{y}\|^{\beta_0}}{(2\pi)^d}\int_{\mathbb{R}^d} \|\mathbf{w}\|^{\beta_0}|\eta_n(\mathbf{w})|^2 d\mathbf{w}.
\end{aligned}$$

Using (16), Lemma 4 and Minkowski's integral inequality (cf. Grafakos [8], page 12), we obtain

$$\begin{aligned}
&\left[\int_{\mathbb{R}^d} |e_n * g_n(\mathbf{x}) - g_n(\mathbf{x})|^2 d\mathbf{x}\right]^{1/2} \\
= \quad &\left[\int_{\mathbb{R}^d} \Big|\int_{\|\mathbf{y}\|\leq 2m_a\varepsilon_n} [g_n(\mathbf{x}-\mathbf{y}) - g_n(\mathbf{x})]e_n(\mathbf{y})d\mathbf{y}\Big|^2 d\mathbf{x}\right]^{1/2} \\
\leq \quad &\int_{\|\mathbf{y}\|\leq 2m_a\varepsilon_n}\left[\int_{\mathbb{R}^d} |g_n(\mathbf{x}-\mathbf{y}) - g_n(\mathbf{x})|^2 d\mathbf{x}\right]^{1/2} e_n(\mathbf{y})d\mathbf{y}
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{2^{(2-\beta_0)/2}(2m_a\varepsilon_n)^{\beta_0/2}}{(2\pi)^{d/2}}\left[\int_{\mathbb{R}^d}\|\mathbf{w}\|^{\beta_0}|\eta_n(\mathbf{w})|^2 d\mathbf{w}\right]^{1/2}\\
&\leq \frac{2^{(2-\beta_0)/2}(2m_a\varepsilon_n)^{\beta_0/2}C_{\xi_1}C_f}{(2\pi)^{d/2}\gamma_n^{2\nu+2\epsilon}}\left[\int_{\mathbb{R}^d}\frac{\|\mathbf{w}\|^{\beta_0}}{(1+\|\mathbf{w}\|^{\beta})^2}d\mathbf{w}\right]^{1/2},
\end{aligned}
$$

where β is a constant satisfying $\beta \in (d/2, \min\{2, 2\epsilon, 2\nu+d\})$ and $0 < \beta_0 < 2\beta - d$. The integral on the right hand side of the last inequality is finite. This proves Lemma 5. □

Lemma 6. *Let $1 \leq d \leq 3$, $\kappa \in (0, 1/2)$ and ξ_0 be as in (14). Then $\xi_0 : \mathbb{R}^d \to \mathbb{R}$ is a continuous, isotropic, strictly positive function and $\xi_0(\mathbf{w}) \asymp \|\mathbf{w}\|^{-\kappa}$ as $\|\mathbf{w}\| \to \infty$.*

Proof of Lemma 6. We shall consider three cases. CASE 1. Suppose $d = 1$. Then

$$
\begin{aligned}
\xi_0(w) &= \int_{-1}^{1} e^{-\imath wx}|x|^{\kappa-1}dx\\
&= 2\int_0^1 \cos(wx)x^{\kappa-1}dx\\
&= \frac{2}{|w|^\kappa}\int_0^{|w|}\cos(x)x^{\kappa-1}dx.
\end{aligned}
$$

Hence ξ_0 is a continuous, isotropic, strictly positive function on $\mathbb{R}$ as $0 < \kappa < 1/2$ and

$$
\int_0^{|w|}\cos(x)x^{\kappa-1}dx \geq \int_0^{3\pi/2}\cos(x)x^{\kappa-1}dx > 0, \qquad \forall |w| \in [\pi/4, \infty).
$$

Also $\xi_0(w) \asymp |w|^{-\kappa}$ as $|w| \to \infty$ since $0 < \int_0^\infty \cos(x)x^{\kappa-1}dx < \infty$.

CASE 2. Suppose $d = 2$. Let U_d be the uniform probability measure on $\mathcal{S}_d = \{\mathbf{u} \in \mathbb{R}^d : \|\mathbf{u}\| = 1\}$. Since ξ_0 is an isotropic function, we have

$$
\begin{aligned}
\xi_0(\mathbf{w}) &= \int_{\|\mathbf{x}\|\leq 1}\left\{\int_{\mathcal{S}_2} e^{-\imath\|\mathbf{w}\|\mathbf{u}'\mathbf{x}}\|\mathbf{x}\|^{\kappa-2}U_2(d\mathbf{u})\right\}d\mathbf{x}\\
&= \int_{\|\mathbf{x}\|\leq 1}\|\mathbf{x}\|^{\kappa-2}\left\{\int_{\mathcal{S}_2}\cos(\|\mathbf{w}\|\mathbf{u}'\mathbf{x})U_2(d\mathbf{u})\right\}d\mathbf{x},
\end{aligned}
$$

and

$$
\int_{\mathcal{S}_2}\cos(\|\mathbf{w}\|\mathbf{u}'\mathbf{x})U_2(d\mathbf{u}) = \frac{1}{2\pi}\int_0^{2\pi}\cos[\|\mathbf{x}\|\|\mathbf{w}\|\cos(\theta)]d\theta.
$$

Hence

$$
\begin{aligned}
\xi_0(\mathbf{w}) &= \frac{1}{2\pi}\int_{\|\mathbf{x}\|\leq 1}\|\mathbf{x}\|^{\kappa-2}\left\{\int_0^{2\pi}\cos[\|\mathbf{w}\|\|\mathbf{x}\|\cos(\theta)]d\theta\right\}d\mathbf{x}\\
&= \frac{1}{2\pi}\int_{\|\mathbf{x}\|\leq\|\mathbf{w}\|}\frac{\|\mathbf{x}\|^{\kappa-2}}{\|\mathbf{w}\|^{\kappa-2}}\left\{\int_0^{2\pi}\cos[\|\mathbf{x}\|\cos(\theta)]d\theta\right\}\frac{d\mathbf{x}}{\|\mathbf{w}\|^2}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{\|\mathbf{w}\|^{\kappa}} \int_0^{2\pi} \int_0^{\|\mathbf{w}\|} x^{\kappa-1} \cos[x|\cos(\theta)|] dx d\theta \\
&= \frac{1}{\|\mathbf{w}\|^{\kappa}} \int_0^{2\pi} \frac{1}{|\cos(\theta)|^{\kappa}} \int_0^{|\cos(\theta)|\|\mathbf{w}\|} x^{\kappa-1} \cos(x) dx d\theta.
\end{aligned}
$$

Arguing as in Case 1, ξ_0 is a continuous, isotropic, strictly positive function and $\xi_0(\mathbf{w}) \asymp \|\mathbf{w}\|^{-\kappa}$ as $\|\mathbf{w}\| \to \infty$.

CASE 3. Suppose $d = 3$. As in Case 2, we have

$$
\begin{aligned}
\xi_0(\mathbf{w}) &= \int_{\|\mathbf{x}\| \le 1} \{ \int_{\mathcal{S}_3} e^{-\imath \|\mathbf{w}\| \mathbf{u}'\mathbf{x}} \|\mathbf{x}\|^{\kappa-3} U_3(d\mathbf{u}) \} d\mathbf{x} \\
&= \int_{\|\mathbf{x}\| \le 1} \|\mathbf{x}\|^{\kappa-3} \left\{ \int_{\mathcal{S}_3} \cos(\|\mathbf{w}\| \mathbf{u}'\mathbf{x}) U_3(d\mathbf{u}) \right\} d\mathbf{x}.
\end{aligned}
$$

We observe from Stein [12], page 43, and Andrews, Askey and Roy [2], page 202, that

$$
\int_{\mathcal{S}_3} \cos(\|\mathbf{w}\| \mathbf{u}'\mathbf{x}) U_3(d\mathbf{u}) = \frac{1}{2} \int_0^{\pi} \cos[\|\mathbf{x}\| \|\mathbf{w}\| \cos(\theta)] \sin(\theta) d\theta = \frac{\sin(\|\mathbf{x}\| \|\mathbf{w}\|)}{\|\mathbf{x}\| \|\mathbf{w}\|}.
$$

Consequently

$$
\begin{aligned}
\xi_0(\mathbf{w}) &= \int_{\|\mathbf{x}\| \le 1} \|\mathbf{x}\|^{\kappa-3} \frac{\sin(\|\mathbf{w}\| \|\mathbf{x}\|)}{\|\mathbf{w}\| \|\mathbf{x}\|} d\mathbf{x} \\
&= \int_{\|\mathbf{x}\| \le \|\mathbf{w}\|} \frac{\|\mathbf{x}\|^{\kappa-3}}{\|\mathbf{w}\|^{\kappa-3}} \frac{\sin(\|\mathbf{x}\|)}{\|\mathbf{x}\|} \frac{d\mathbf{x}}{\|\mathbf{w}\|^3} \\
&= \frac{4\pi}{\|\mathbf{w}\|^{\kappa}} \int_0^{\|\mathbf{w}\|} x^{\kappa-2} \sin(x) dx.
\end{aligned}
$$

This implies that ξ_0 is a continuous, isotropic, strictly positive function and $\xi_0(\mathbf{w}) \asymp \|\mathbf{w}\|^{-\kappa}$ as $\|\mathbf{w}\| \to \infty$. This proves Lemma 6. □

Acknowledgements

We would like to thank the Editor, Professor David Ruppert, and an Associate Editor for their comments and suggestions that helped improve our presentation immensely. Wei-Liem Loh would also like to thank Professors Michael Stein and Hao Zhang for their comments and thoughts on a previous version of the manuscript.

References

[1] ANDERES, E. (2010). On the consistent separation of scale and variance for Gaussian random fields. *Ann. Statist.* **38** 870–893. MR2604700

[2] Andrews, G. E., Askey, R. and Roy, R. (1999). *Special Functions.* Cambridge Univ. Press, New York. MR1688958
[3] Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57** 33–45.
[4] Davis, T. A. (2006). *Direct Methods for Linear Sparse Systems*. SIAM, Philadelphia. MR2270673
[5] Du, J., Zhang, H. and Mandrekar, V. S. (2009). Fixed-domain asymptotic properties of tapered maximum likelihood estimators. *Ann. Statist.* **37** 3330–3361. MR2549562
[6] Furrer, R., Genton, M. G., and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets. *J. Comput. Graph. Statist.* **15** 502–523. MR2291261
[7] Gneiting, T. (2002). Compactly supported correlation functions. *J. Multivar. Anal.* **83** 493–508. MR1945966
[8] Grafakos, L. (2008). *Classical Fourier Analysis*, 2nd edition. Springer, New York. MR2445437
[9] Ibragimov, I. A. and Rozanov, Y. A. (1978). *Gaussian Random Processes.* Springer-Verlag, New York. MR0543837
[10] Kaufman, C. G., Schervish, M. J. and Nychka, D. W. (2008). Covariance estimation for likelihood-based estimation in large spatial data sets. *J. Amer. Statist. Assoc.* **103** 1545–1555. MR2504203
[11] Stein, E. M. and Weiss, G. (1971). *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton Univ. Press, Princeton. MR0304972
[12] Stein, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging.* Springer-Verlag, New York. MR1697409
[13] Stein, M. L., Chi, Z. and Welty, L. J. (2004). Approximating likelihoods for large spatial data sets. *J. Roy. Statist. Soc. Ser. B* **66** 275–296. MR2062376
[14] Wang, D. (2010). *Fixed-domain Asymptotics and Consistent Estimation for Gaussian Random Field Models in Spatial Statistics and Computer Experiments.* Nat. Univ. Singapore PhD thesis, Singapore.
[15] Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv. Comput. Math.* **4** 389–396. MR1366510
[16] Wendland, H. (1998). Error estimates for interpolation by compactly supported radial basis functions of minimal degree. *J. Approx. Theory* **93** 258–272. MR1616781
[17] Wu, Z. M. (1995). Compactly supported positive definite radial functions. *Adv. Comput. Math.* **4** 283–292. MR1357720
[18] Ying, Z. (1991). Asymptotic properties of a maximum likelihood estimator with data from a Gaussian process. *J. Multivar. Anal.* **36** 280–296. MR1096671
[19] Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Statist. Assoc.* **99** 250–261. MR2054303