Bernoulli **29**(3), 2023, 2519–2543
https://doi.org/10.3150/22-BEJ1551

Estimating the parameters of some common Gaussian random fields with nugget under fixed-domain asymptotics

WEI-LIEM LOH[a] and SAIFEI SUN[b]

Department of Statistics and Data Science, National University of Singapore, Singapore, [a]*stalohwl@nus.edu.sg,* [b]*saifeisun@u.nus.edu*

This article considers parameter estimation for a class of Gaussian random fields on $[0,1)^d$ that are observed with measurement error and irregularly spaced design sites. This class comprises Gaussian random fields with suitably smooth mean functions and isotropic powered exponential, Matérn or generalized Wendland covariance functions. Under fixed-domain asymptotics, consistent estimators are proposed for three microergodic parameters, namely the nugget, the smoothness parameter and a parameter related to the coefficient of the principal irregular term of the isotropic covariance function. Upper bounds for the convergence rate of these estimators are established. Simulations are conducted to study the finite sample accuracy of the proposed estimators.

Keywords: Convergence rate; fixed-domain asymptotics; Gaussian random field; generalized Wendland; multivariate discrete differentiation; nugget; powered exponential; quadratic variation; Matérn; space-filling design

1. Introduction

This article focuses on the following well known model in spatial statistics (see, for example, Chang, Huang and Ing (2017), Chen, Simpson and Ying (2000), Cressie (1991), Stein (1999), Tang, Zhang and Banerjee (2021) and the references cited therein):

$$X(\mathbf{x}_i) = m_X(\mathbf{x}_i) + Z(\mathbf{x}_i) + \epsilon(\mathbf{x}_i), \qquad \forall i = 1,\dots,n, \tag{1}$$

where

(i) the design sites are $\mathbf{x}_i \in [0,1)^d$, $i = 1,\dots,n$,
(ii) Z is a mean-zero isotropic Gaussian random field on $\mathbb{R}^d$ with covariance function K_Z,
(iii) the measurement errors $\epsilon(\mathbf{x}_1),\dots,\epsilon(\mathbf{x}_n)$ are i.i.d. $N(0,\tau)$ random variables, with $\tau \geq 0$, that are also independent of Z. The parameter τ is commonly known as the nugget.
(iv) m_X is a suitably smooth mean function of X.

As Tang, Zhang and Banerjee (2021) noted, spatial process models popular in geostatistics often represent the observed data as the sum of a smooth underlying process and white noise. The variation in the white noise is attributed to measurement error, or micro-scale variability, and is called the nugget.

For definiteness, we assume that K_Z belongs to one of three popular classes of covariance functions, namely $K_Z \in \{K_P, K_M, K_W\}$ where K_P is the powered exponential covariance function, K_M is the Matérn covariance function and K_W is the generalized Wendland covariance function (defined below).

1350-7265

Powered exponential. The isotropic powered exponential covariance function is given by (e.g. Sacks et al. (1989))

$$K_P(\mathbf{x},\mathbf{y}) = \sigma^2 \exp\{-(\alpha\|\mathbf{x}-\mathbf{y}\|)^{2\nu}\} = \sigma^2 \sum_{j=0}^{\infty} (-1)^j \frac{\alpha^{2\nu j}\|\mathbf{x}-\mathbf{y}\|^{2\nu j}}{j!}, \tag{2}$$

for all $\mathbf{x},\mathbf{y} \in \mathbb{R}^d$ where $\nu \in (0,1), \alpha, \sigma^2$ are unknown positive constants and $\|.\|$ denotes the Euclidean norm in $\mathbb{R}^d$.

Matérn. The isotropic Matérn covariance function is given by (e.g. Loh (2015), Stein (1999))

$$K_M(\mathbf{x},\mathbf{y}) = \frac{\sigma^2(\alpha\|\mathbf{x}-\mathbf{y}\|)^\nu}{2^{\nu-1}\Gamma(\nu)} \mathcal{K}_\nu(\alpha\|\mathbf{x}-\mathbf{y}\|), \qquad \forall \mathbf{x},\mathbf{y} \in \mathbb{R}^d, \tag{3}$$

where ν, α, σ^2 are unknown positive constants and $\mathcal{K}_\nu(.)$ is the modified Bessel function of the second kind (cf. Andrews, Askey and Roy (1999)).

Generalized Wendland. The isotropic generalized Wendland covariance function is given by (e.g. Bevilacqua et al. (2019), Sun (2020))

$$K_W(\mathbf{x},\mathbf{y}) = \frac{\sigma^2 \mathcal{I}\{\|\mathbf{x}-\mathbf{y}\| \in [0,\xi)\}}{B(2\nu,\mu)} \int_{\|\mathbf{x}-\mathbf{y}\|/\xi}^{1} (u^2 - \frac{\|\mathbf{x}-\mathbf{y}\|^2}{\xi^2})^{\nu-1/2}(1-u)^{\mu-1}\,du, \tag{4}$$

for all $\mathbf{x},\mathbf{y} \in \mathbb{R}^d$ where $B(.,.)$ is the beta function, $\mathcal{I}\{.\}$ is the indicator function and ν, ξ, μ, σ^2 are unknown positive constants such that $\nu \geq 1/2$ and $\mu \geq \nu + d/2$.

It follows from (1) that X is a Gaussian random field with mean function m_X and covariance function

$$K_X(\mathbf{x},\mathbf{y}) = \tau \mathcal{I}\{\mathbf{x}=\mathbf{y}\} + K_Z(\mathbf{x},\mathbf{y}), \qquad \forall \mathbf{x},\mathbf{y} \in [0,1)^d. \tag{5}$$

Given the observed sample

$$\{(\mathbf{x}_1, X(\mathbf{x}_1)), \ldots, (\mathbf{x}_n, X(\mathbf{x}_n))\} \tag{6}$$

and $K_Z \in \{K_P, K_M, K_W\}$, our aim is to estimate consistently (as $n \to \infty$) the nugget τ, the smoothness parameter ν and the parameter η (defined in Section 7). We are motivated by the following basic question:

Suppose the design sites $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are i.i.d. random vectors where $\mathbf{x}_1$ has the uniform distribution on $[0,1)^d$, $d \in \{1,2,3\}$, and the observed sample is $\{\mathbf{x}_1, X(\mathbf{x}_1)\}, \ldots, \{\mathbf{x}_n, X(\mathbf{x}_n)\}$ with $K_Z = K_M$ (Matérn covariance). How can one construct consistent estimators for τ, ν and $\eta = \sigma^2\alpha^{2\nu}$ as $n \to \infty$?

As far as we know, this question is open. Even though the likelihood is Gaussian, likelihood methods (such as MLE) appear to be analytically intractable under fixed-domain asymptotics. The latter asymptotics imply that as sample size $n \to \infty$, the n sites get to be increasingly dense in a (fixed) compact set in $\mathbb{R}^d$. The dependence among all the observations $X(\mathbf{x}_1), \ldots, X(\mathbf{x}_n)$ remains strong as $n \to \infty$. The dimension of the $n \times n$ covariance matrix of the observations increases to infinity as $n \to \infty$. This is exacerbated by the fact that $\mathbf{x}_i$'s are irregularly spaced. All these indicate that any theoretical analysis of the MLEs for τ or ν is a formidable (or even intractable) task with respect to fixed-domain asymptotics. Indeed as far as we know, the consistency of the MLEs for τ or ν are still open problems in the setting of the above question.

In the case where it is known that $\tau = 0$ (i.e. no nugget) and X is an isotropic Gaussian random field with Matérn covariance function as in (3), Loh, Sun and Wen (2021) recently constructed consistent estimators for ν and $\eta = \sigma^2\alpha^{2\nu}$ when the design sites $\mathbf{x}_i$ are i.i.d. on $[0,1)^d, d \in \{1,2,3\}$. However the results in Loh, Sun and Wen (2021) do not apply when $\tau \geq 0$ is unknown.

This article uses the crucial Lemma 1 which first appeared in Loh, Sun and Wen (2021). While Loh, Sun and Wen (2021) and this article use higher-order quadratic variations, the higher-order quadratic variations are significantly different as the nugget model (1) is much more complex than the corresponding isotropic model with $\tau = 0$, $K_Z = K_M$ and m_X is a constant. The theoretical results are also more complicated, e.g. the upper bounds for the convergence rate of the proposed estimators depend on the value of the smoothness parameter ν which is not the case in Loh, Sun and Wen (2021). Perhaps more importantly, this article shows the versatility of Lemma 1 in tackling many fixed-domain asymptotic problems that are otherwise intractable thus far using any other method.

We shall now review related literature on the nugget model (1). Chen, Simpson and Ying (2000) investigated the asymptotic properties of the maximum likelihood estimators in model (1) where $d = 1$, $m_X = 0$ and $K_Z = K_M$ with $\nu = 1/2$ (Ornstein-Uhlenbeck process). Chang, Huang and Ing (2017) assumes the mean function m_X is a linear regression term in model (1) with $d = 1$ and $K_Z = K_M$ with $\nu = 1/2$ and proves the consistency of the maximum likelihood estimators with respect to mixed-domain asymptotics (but not fixed-domain asymptotics). While fixed-domain asymptotics constrain the domain to be bounded and fixed, mixed-domain asymptotics allow the domain to expand without bound as sample size increases to infinity. Recently, Tang, Zhang and Banerjee (2021) studied the consistency properties of the maximum likelihood estimators in model (1) where $K_Z = K_M$ with the smoothness parameter ν known. A crucial assumption behind the papers Chang, Huang and Ing (2017), Chen, Simpson and Ying (2000), Tang, Zhang and Banerjee (2021) is that the smoothness parameter ν is known.

The following are related literature on model (1) without the nugget. Assuming that $\tau = 0$, $m_X = 0$ and $K_Z = K_M$ with ν known, consistent estimators for $\eta = \sigma^2\alpha^{2\nu}$ have been proposed by Du, Zhang and Mandrekar (2009), Kaufman and Shaby (2013), Keshavarz, Nguyen and Scott (2019), Wang and Loh (2011), Zhang (2004) with respect to fixed-domain asymptotics. Assuming that $\tau = 0$, $m_X = 0$ and $K_Z = K_W$ with ν and μ known, consistent estimators for $\sigma^2\xi^{-2\nu}$ have been proposed by Bevilacqua et al. (2019) with respect to fixed-domain asymptotics. In contrast, this article assumes all four parameters (namely ν, σ^2, μ, ξ) of K_W, the mean function m_X and the nugget τ are unknown.

Finally, Wen, Sun and Loh (2021) estimates the smoothness parameter of a class of nonstationary Gaussian random fields without nugget where the observations are taken on a smooth curve with respect to fixed-domain asymptotics. The curve being essentially 1 dimensional, the construction of the higher-order quadratic variations in Wen, Sun and Loh (2021) does not require Lemma 1. Using Lemma 1, this article constructs consistent estimators for ν, τ and η in model (1) where the design sites $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are randomly selected in $[0,1)^d$, $\eta = \sigma^2\alpha^{2\nu}$ if $K_Z \in \{K_P, K_M\}$ and $\eta = \sigma^2\xi^{-2\nu}/B(2\nu,\mu)$ if $K_Z = K_W$. This method is applicable to other space-filling designs too (see Section 8).

We shall end this section with an outline of the rest of the paper. Section 2 presents series expansions of the covariance functions K_M and K_W that are needed in the sequel. Sections 3 and 4 develop a new class of higher-order quadratic variations $V_{u,\theta,d,\ell}$ based on $X(\mathbf{x}(\mathbf{i})), 1 \leq i_1, \ldots, i_d \leq n$, where $\mathbf{x}(\mathbf{i})$'s are chosen via stratified sampling in $[0,1)^d$. The latter is assumed in Sections 3 to 7. Theorems 1 to 4 establish useful asymptotic properties of $V_{u,\theta,d,\ell}$ that are needed in Sections 5 to 7. Section 5 proposes the estimator $\hat{\tau}_\ell$ for the nugget parameter τ. This is accomplished by using $V_{0,\theta,d,\ell}$ to filter out the asymptotic contributions of τ from those of other parameters. Theorem 5 establishes an upper bound for the convergence rate of $\hat{\tau}_\ell$ when $K_Z \in \{K_P, K_M, K_W\}$.

Section 6 proposes the estimator $\hat{\nu}_P$ for the smoothness parameter ν if $K_Z = K_P$ and the estimator $\hat{\nu}$ for ν when $K_Z \in \{K_M, K_W\}$. This is done by using $V_{1,\theta,d,\ell}$ to filter out the asymptotic effects of ν from those of other parameters. Theorems 6, 7 establish upper bounds for the convergence rate of $\hat{\nu}_P, \hat{\nu}$ respectively. Section 7 proposes the estimator $\hat{\eta}_P$ for η if $K_Z = K_P$ and the estimator $\hat{\eta}$ for η when

$K_Z \in \{K_M, K_W\}$. Theorems 8, 9 establish upper bounds for the convergence rate of $\hat{\eta}_P, \hat{\eta}$ respectively. Section 8 adapts the results on stratified sampling in Sections 3 to 7 to space-filling designs in $[0,1)^d$. In particular, the resulting estimators $\hat{\tau}^{\mathrm{S}}_{\ell}$, $\hat{\nu}^{\mathrm{S}}$, $\hat{\eta}^{\mathrm{S}}$ are the first provably consistent estimators for the parameters τ, ν and η when $K_Z \in \{K_M, K_W\}$ with respect to i.i.d. sampling designs. This answers the question posed earlier. Section 9 reports a simulation study to gauge the accuracy of the proposed estimators for τ, ν and η when $d \in \{1,2\}$ and $K_Z \in \{K_P, K_M\}$. The online supplement contains the proofs of all the results in this article as well as the continuation of the simulation study of Section 9.

2. Some preliminary results

For $N \in \mathbb{Z}_+ = \{1,2,\ldots\}$, let $C^N(S)$ be the set of functions $f : S \to \mathbb{R}$ that are N times continuously differentiable (i.e. all Nth-order partial derivatives of f exist and are continuous). Let $\lfloor . \rfloor$ and $\lceil . \rceil$ denote the greatest integer function and least integer function respectively. Writing $a_n \asymp b_n$ means $0 < \liminf_{n\to\infty} a_n/b_n \le \limsup_{n\to\infty} a_n/b_n < \infty$ and $a_n \sim b_n$ means $\lim_{n\to\infty} a_n/b_n = 1$. If A is a matrix, then A' is its transpose. For $x, y \in \mathbb{R}$, $x \wedge y = \min\{x,y\}$ and $x \vee y = \max\{x,y\}$. Condition 1 below will be needed in the sequel.

Condition 1. *Suppose $X(\mathbf{x})$, $\mathbf{x} \in [0,1)^d$, is a Gaussian random field as in (1) with unknown mean function $m_X(\mathbf{x}) = \mathbb{E}X(\mathbf{x})$ where $m_X \in C^N(\mathbb{R}^d)$ for some integer N such that $N \ge \nu + d/2$ if $K_Z \in \{K_M, K_W\}$ and $N = 2$ if $K_Z = K_P$.*

The mean function m_X needs to be sufficiently smooth in order to avoid confounding with the smoothness parameter ν of K_Z. Define $G_s : [0,\infty) \to \mathbb{R}$, $s > 0$, by $G_s(0) = 0$ and for all $t > 0$,

$$G_s(t) = \begin{cases} t^{2s}\log(t), & \forall s \in \mathbb{Z}_+, \\ t^{2s}, & \text{otherwise.} \end{cases}$$

We observe from Loh (2015) that K_M in (3) has the following infinite series representation:

$$K_M(\mathbf{x},\mathbf{y}) = \sum_{j=0}^{\infty} \Big\{ \beta_j \|\mathbf{x}-\mathbf{y}\|^{2j} + \beta^*_{\nu+j} G_{\nu+j}(\|\mathbf{x}-\mathbf{y}\|) \Big\}, \qquad \forall \mathbf{x},\mathbf{y} \in \mathbb{R}^d, \tag{7}$$

where $\beta_j, \beta^*_{\nu+j}, j = 0,1,\ldots,$ are constants such that $\beta^*_\nu \neq 0$ and $(-1)^k \beta_k > 0$, whenever k is an integer satisfying $0 \le k < \nu$. Sun (2020), pages 10 to 15, proves that K_W in (4) has the following infinite series representation:

$$\begin{aligned} K_W(\mathbf{x},\mathbf{y}) = \mathcal{I}\{\|\mathbf{x}-\mathbf{y}\| \in [0,\xi)\} \sum_{j=0}^{\infty} \Big\{ &\beta_j \|\mathbf{x}-\mathbf{y}\|^{2j} \\ &+ \beta_{\nu,j+1} \|\mathbf{x}-\mathbf{y}\|^{2\nu+2j+1} + \beta^*_{\nu+j} G_{\nu+j}(\|\mathbf{x}-\mathbf{y}\|) \Big\}, \qquad \forall \mathbf{x},\mathbf{y} \in \mathbb{R}^d, \end{aligned} \tag{8}$$

where $\beta_j, \beta_{\nu,j+1}, \beta^*_{\nu+j}, j = 0,1,\ldots,$ are constants such that $\beta^*_\nu \neq 0$, $(-1)^k \beta_k > 0$, whenever k is an integer satisfying $0 \le k < \nu$ and $\beta_{\nu,j+1} = 0$, $j = 0,1,\ldots,$ whenever $\nu \in \{(2k-1)/2 : k \in \mathbb{Z}_+\}$. (7) and (8) imply that if $K_Z \in \{K_M, K_W\}$, then K_Z can be expressed as

$$K_Z(\mathbf{x},\mathbf{y}) = \mathcal{I}\{\|\mathbf{x}-\mathbf{y}\| \in [0,\xi)\} \sum_{j=0}^{\infty} \Big\{ \beta_j \|\mathbf{x}-\mathbf{y}\|^{2j}$$

$$+\beta_{\nu,j+1}\|\mathbf{x}-\mathbf{y}\|^{2\nu+2j+1}+\beta^*_{\nu+j}G_{\nu+j}(\|\mathbf{x}-\mathbf{y}\|)\Big\}, \qquad \forall \mathbf{x},\mathbf{y}\in\mathbb{R}^d, \tag{9}$$

where

(i) $0^0=1$, $\beta_j,\beta_{\nu,j+1},\beta^*_{\nu+j}$ are constants where $\beta^*_\nu\neq 0$ and $(-1)^k\beta_k>0$, whenever k is an integer satisfying $0\le k<\nu$,
(ii) $\xi\in(0,\infty]$,
(iii) $\beta_{\nu,j+1}=0$, $j=0,1,\ldots$, whenever $\nu\in\{(2k-1)/2:k\in\mathbb{Z}_+\}$.

By using (9), unified proofs are obtained when K_Z is either K_M or K_W for all the theorems of this article. The principal irregular term of $K_Z(\mathbf{x},\mathbf{y})$ in (9) is $\beta^*_\nu G_\nu(\|\mathbf{x}-\mathbf{y}\|)\mathcal{I}\{\|\mathbf{x}-\mathbf{y}\|\in[0,\xi)\}$, cf. (Stein, 1999). The method proposed in this article fails if the covariance function K_Z does not possess a principal irregular term. An example of such a covariance function is the infinitely differentiable isotropic squared exponential function (i.e. K_P in (2) with $\nu=1$). While this is the case, some parameter estimation fixed-domain consistency results for a mean-zero Gaussian random field with squared exponential function can be found in Loh and Lam (2000).

3. Stratified sampling design

For brevity, we write $\mathbf{i}=(i_1,\ldots,i_d)'\in\mathbb{Z}^d_+$. Suppose $\mathbf{x}(\mathbf{i})=(x_1(\mathbf{i}),\ldots,x_d(\mathbf{i}))'$, $1\le i_1,\ldots,i_d\le n$, are design sites in $[0,1)^d$ such that

$$x_k(\mathbf{i})=\frac{i_k-1}{n}+\frac{\delta_{\mathbf{i};k}}{n}, \tag{10}$$

where $\delta_{\mathbf{i};k}\in[0,1)$, $k=1,\ldots,d$. The $\delta_{\mathbf{i};k}$'s can vary with n. Then

$$\mathbf{x}(\mathbf{i})\in[\frac{i_1-1}{n},\frac{i_1}{n})\times\ldots\times[\frac{i_d-1}{n},\frac{i_d}{n}), \qquad \forall 1\le i_1,\ldots,i_d\le n.$$

For simplicity, we shall assume that the $\delta_{\mathbf{i};k}$'s are (non-random) constants in Sections 3 to 7 and are random, independent of the Gaussian random field X, in Section 8.

Let $\omega_n\in\mathbb{Z}_+$ such that $\omega_n\to\infty$ and $\omega_n=O(n^\gamma)$ as $n\to\infty$ where $\gamma\in(0,1)$ is a constant. We observe that Corollary 1 of Loh, Sun and Wen (2021) has constructed constants $c^{k_1,\ldots,k_d}_{\mathbf{i},\theta,d,\ell}$ such that Lemma 1 holds.

Lemma 1. *Let $d,\ell\in\mathbb{Z}_+$ and $\theta\in\{1,2\}$. Given $\mathbf{x}(\mathbf{i})$, $1\le i_1,\ldots,i_d\le n$, as in (10), there exist constants $c^{k_1,\ldots,k_d}_{\mathbf{i},\theta,d,\ell}$, $1\le i_1,\ldots,i_d\le n-\ell\omega_n\theta$, such that*

$$\sum_{0\le k_1,\ldots,k_d\le\ell} c^{k_1,\ldots,k_d}_{\mathbf{i},\theta,d,\ell}\, x_d(i_1+k_1\omega_n\theta,\ldots,i_d+k_d\omega_n\theta)^\ell=\ell!(\frac{\omega_n\theta}{n})^\ell,$$

and

$$\sum_{0\le k_1,\ldots,k_d\le\ell} c^{k_1,\ldots,k_d}_{\mathbf{i},\theta,d,\ell}\prod_{j=1}^d x_j(i_1+k_1\omega_n\theta,\ldots,i_d+k_d\omega_n\theta)^{l_j}=0,$$

for all integers $l_1,\ldots,l_d$ satisfying $0\le l_1,\ldots,l_{d-1}\le\ell$, $0\le l_d\le\ell-1$ and $0\le l_1+\ldots+l_d\le\ell$.

The results of Lemma 1 are exact, not approximations. Lemma 1 is crucial in the construction of the higher-order quadratic variations in Section 4. It should also be noted that the choice of coordinate d in Lemma 1 is arbitrary. By symmetry, coordinate d can be replaced by any other coordinate with appropriate changes in the construction of the constants $c^{k_1,\ldots,k_d}_{\mathbf{i},\theta,d,\ell}$. It is further proved in Corollary 2 of Loh, Sun and Wen (2021) that

$$c^{k_1,\ldots,k_d}_{\mathbf{i},\theta,d,\ell} = c^{k_1,\ldots,k_d}_{d,\ell} + O(\omega_n^{-1}), \qquad \forall 0 \le k_1,\ldots,k_d \le \ell, \tag{11}$$

as $n \longrightarrow \infty$ uniformly over $1 \le i_1,\ldots,i_d \le n - 2\ell\omega_n$ and $\delta_{\mathbf{j};k} \in [0,1), 1 \le j_1,\ldots,j_d \le n, 1 \le k \le d$,

$$\sum_{0\le k_1,\ldots,k_d \le \ell} c^{k_1,\ldots,k_d}_{d,\ell} k_d^{\ell} = \ell!,$$

and

$$\sum_{0\le k_1,\ldots,k_d \le \ell} c^{k_1,\ldots,k_d}_{d,\ell} \prod_{j=1}^{d} k_j^{l_j} = 0,$$

for all integers $l_1,\ldots,l_d$ satisfying $0 \le l_1,\ldots,l_{d-1} \le \ell, 0 \le l_d \le \ell - 1$ and $0 \le l_1 + \ldots + l_d \le \ell$. It is noted that the $c^{k_1,\ldots,k_d}_{d,\ell}$'s do not depend on θ, n and $\mathbf{i} = (i_1,\ldots,i_d)'$. The explicit computations of $c^{k_1,\ldots,k_d}_{\mathbf{i},\theta,d,\ell}$'s and $c^{k_1,\ldots,k_d}_{d,\ell}$'s are given in Loh, Sun and Wen (2021).

Finally, the results of Lemma 1 can be regarded as a multivariate discrete differentiation operation since successive differentiation would also eliminate low-order multivariate polynomials. While multivariate discrete differentiation is known in the literature for lattice data (e.g. Sadhanala et al. (2021)), Lemma 1 appears to be the first to apply to scattered data. Thus at least in principle, Lemma 1 may be useful to problems beyond those considered here.

4. A new class of quadratic variations

Suppose Condition 1 holds. Let $d,\ell \in \mathbb{Z}_+$, $\theta \in \{1,2\}$, n, ω_n be as in Section 3 and $\mathbf{i} = (i_1,\ldots,i_d)'$ such that $1 \le i_1,\ldots,i_d \le n - \ell\omega_n\theta$. Define

$$\nabla_{\theta,d,\ell} X\big(\mathbf{x}(\mathbf{i})\big) = \sum_{0\le k_1,\ldots,k_d \le \ell} c^{k_1,\ldots,k_d}_{\mathbf{i},\theta,d,\ell} X\big(\mathbf{x}(i_1 + k_1\omega_n\theta,\ldots,i_d + k_d\omega_n\theta)\big),$$

and the ℓth-order quadratic variation $V_{u,\theta,d,\ell} = V_{u,\theta,d,\ell}(n)$ by

$$V_{u,\theta,d,\ell} = \sum_{\mathbf{i}\in\Xi_{u,n}} \big\{\nabla_{\theta,d,\ell} X\big(\mathbf{x}(\mathbf{i})\big)\big\}\big\{\nabla_{\theta,d,\ell} X\big(\mathbf{x}(\mathbf{i} + u\mathbf{e}_1)\big)\big\}, \tag{12}$$

where $\mathbf{e}_1 = (1,0,\ldots,0)' \in \mathbb{R}^d$, $u \in \{0,1\}$ and $\Xi_{u,n} = \{\mathbf{i} : 1 \le i_1 + u, i_1,\ldots,i_d \le n - 2\ell\omega_n\}$. We observe that $V_{0,\theta,d,\ell} \ge 0$ but $V_{1,\theta,d,\ell}$ can be negative. The quadratic variation $V_{0,\theta,d,\ell}$ is used to estimate τ in Section 5 while $V_{1,\theta,d,\ell}$ is used to filter out the effects of τ and hence is used to estimate ν in Section 6.

Condition 2. *Suppose ω_n is an even integer such that $\omega_n \asymp n^{\gamma}$ where $\gamma \in (0,1)$ is a constant.*

Under Condition 2, ω_n is an even integer and hence

$$\begin{aligned}&\left\{\mathbf{x}(i_1+k_1\omega_n\theta,\ldots,i_d+k_d\omega_n\theta):0\le k_1,\ldots,k_d\le\ell\right\}\\ \bigcap&\left\{\mathbf{x}(i_1+1+k_1\omega_n\theta,i_2+k_2\omega_n\theta,\ldots,i_d+k_d\omega_n\theta):0\le k_1,\ldots,k_d\le\ell\right\}=\varnothing.\end{aligned}\tag{13}$$

We observe from (5) and (13) that

$$\begin{aligned}\mathbb{E}(V_{u,\theta,d,\ell})=&\sum_{\mathbf{i}\in\Xi_{u,n}}\sum_{0\le k_1,\ldots,k_{2d}\le\ell}c_{\mathbf{i},\theta,d,\ell}^{k_1,\ldots,k_d}c_{\mathbf{i}+u\mathbf{e}_1,\theta,d,\ell}^{k_{d+1},\ldots,k_{2d}}m_X\left(\mathbf{x}(i_1+k_1\omega_n\theta,\ldots,i_d+k_d\omega_n\theta)\right)\\&\times m_X\left(\mathbf{x}(i_1+u+k_{d+1}\omega_n\theta,i_2+k_{d+2}\omega_n\theta,\ldots,i_d+k_{2d}\omega_n\theta)\right)\\&+\tau(1-u)\sum_{\mathbf{i}\in\Xi_{0,n}}\sum_{0\le k_1,\ldots,k_d\le\ell}(c_{\mathbf{i},\theta,d,\ell}^{k_1,\ldots,k_d})^2+\sum_{\mathbf{i}\in\Xi_{u,n}}\sum_{0\le k_1,\ldots,k_{2d}\le\ell}\\&\times c_{\mathbf{i},\theta,d,\ell}^{k_1,\ldots,k_d}c_{\mathbf{i}+u\mathbf{e}_1,\theta,d,\ell}^{k_{d+1},\ldots,k_{2d}}K_Z\Big(\mathbf{x}(i_1+k_1\omega_n\theta,\ldots,i_d+k_d\omega_n\theta),\\&\qquad\mathbf{x}(i_1+u+k_{d+1}\omega_n\theta,i_2+k_{d+2}\omega_n\theta,\ldots,i_d+k_{2d}\omega_n\theta)\Big).\end{aligned}$$

Define

$$\begin{aligned}\tilde{K}_Z(\mathbf{x}-\mathbf{y})&=K_Z(\mathbf{x},\mathbf{y}),\\ \tilde{K}_Z^{(a_1,\ldots,a_d)}(\mathbf{x})&=\frac{\partial^{a_1+\ldots+a_d}}{\partial x_1^{a_1}\ldots\partial x_d^{a_d}}\tilde{K}_Z(\mathbf{x}),\qquad\forall\mathbf{x},\mathbf{y}\in\mathbb{R}^d,\end{aligned}$$

where $a_1,\ldots,a_d$ are nonnegative integers if the partial derivative exists.

Lemma 2. *Suppose $K_Z\in\{K_P,K_M,K_W\}$, $d\in\mathbb{Z}_+$ and $\xi=\infty$ if $K_Z\in\{K_P,K_M\}$. Then $\tilde{K}_Z^{(a_1,\ldots,a_d)}(.)$ is a continuous function on an open set containing $\left\{\mathbf{x}-\mathbf{y}:\mathbf{x},\mathbf{y}\in[0,1)^d,\|\mathbf{x}-\mathbf{y}\|\in(0,\xi)\right\}$ and there exists a positive constant C_ℓ such that*

$$|\tilde{K}_Z^{(a_1,\ldots,a_d)}(\mathbf{x}-\mathbf{y})|\le\begin{cases}C_\ell\|\mathbf{x}-\mathbf{y}\|^{2\nu-2\ell}, & \text{if } \nu<\ell,\\ C_\ell\{|\log(\|\mathbf{x}-\mathbf{y}\|)|+1\}, & \text{if } \nu=\ell,\\ C_\ell, & \text{if } \nu>\ell,\end{cases}$$

for all $a_1+\ldots+a_d=2\ell$ and $\mathbf{x},\mathbf{y}\in[0,1)^d$ satisfying $\|\mathbf{x}-\mathbf{y}\|\in(0,\xi)$. In addition there exists a constant C such that

$$\tilde{K}_Z(\mathbf{x}-\mathbf{y})\le C(1-\frac{\|\mathbf{x}-\mathbf{y}\|}{\xi})^{2\nu+(d-1)/2}\mathcal{I}\{\|\mathbf{x}-\mathbf{y}\|\le\xi\},\qquad\forall\mathbf{x},\mathbf{y}\in\mathbb{R}^d.$$

Theorems 1 to 4 and Proposition 1 below establish some useful asymptotic properties of $V_{u,\theta,d,\ell}$ that will be needed in Sections 5 to 7. Define

$$H_\ell(s)=\sum_{0\le k_1,\ldots,k_{2d}\le\ell}c_{d,\ell}^{k_1,\ldots,k_d}c_{d,\ell}^{k_{d+1},\ldots,k_{2d}}G_s\left(\|(k_1,\ldots,k_d)'-(k_{d+1},\ldots,k_{2d})'\|\right)$$

for all $s\in(0,\ell)$, $\ell\in\mathbb{Z}_+$.

Theorem 1. *Suppose $K_Z = K_P$ and Conditions 1, 2 hold. Let $d,\ell \in \mathbb{Z}_+$, $\tau \ge 0$, $\theta \in \{1,2\}$, $u \in \{0,1\}$ and $V_{u,\theta,d,\ell}$ be as in (12). Then as $n \to \infty$ uniformly over $\delta_{\mathbf{i};k} \in [0,1), 1 \le i_1,\dots,i_d \le n, 1 \le k \le d$, we have*

$$\mathbb{E}(V_{u,\theta,d,\ell}) = \tau(1-u)\sum_{\mathbf{i}\in\Xi_{0,n}}\sum_{0\le k_1,\dots,k_d\le\ell}(c_{\mathbf{i},\theta,d,\ell}^{k_1,\dots,k_d})^2 - \sigma^2\alpha^{2\nu}(\frac{\omega_n\theta}{n})^{2\nu}|\Xi_{u,n}|H_\ell(\nu)$$

$$+n^d(\frac{\omega_n}{n})^{2\nu}O\Big\{\frac{1}{\omega_n}+\frac{u}{\omega_n^{2\nu}}+(\frac{\omega_n}{n})^{(2\nu)\wedge(2\ell-2\nu)}\Big\} \asymp n^d\{\tau(1-u)+(\frac{\omega_n}{n})^{2\nu}\}.$$

Theorem 2. *Suppose $K_Z \in \{K_M, K_W\}$ as in (9) and Conditions 1, 2 hold. Let $d,\ell \in \mathbb{Z}_+$, $\tau \ge 0$, $\theta \in \{1,2\}$, $u \in \{0,1\}$ and $V_{u,\theta,d,\ell}$ be as in (12).*

(a) Suppose $\nu < \ell$. Then as $n \to \infty$ uniformly over $\delta_{\mathbf{i};k} \in [0,1), 1 \le i_1,\dots,i_d \le n, 1 \le k \le d$, we have

$$\mathbb{E}(V_{u,\theta,d,\ell}) = \tau(1-u)\sum_{\mathbf{i}\in\Xi_{0,n}}\sum_{0\le k_1,\dots,k_d\le\ell}(c_{\mathbf{i},\theta,d,\ell}^{k_1,\dots,k_d})^2 + \beta_\nu^*(\frac{\omega_n\theta}{n})^{2\nu}|\Xi_{u,n}|H_\ell(\nu)$$

$$+n^d(\frac{\omega_n}{n})^{2\nu}O\Big\{(\frac{\omega_n}{n})^{[2(\ell\wedge N)-2\nu]\wedge 2}+\frac{1}{\omega_n}+\frac{u}{\omega_n^{2\nu}}$$

$$+(\frac{\omega_n}{n})^2\log(\frac{n}{\omega_n})\mathcal{I}\{\ell=\nu+1\}+\beta_{\nu,1}\frac{\omega_n}{n}\Big\} \asymp n^d\{\tau(1-u)+(\frac{\omega_n}{n})^{2\nu}\}.$$

(b) Suppose $\nu = \ell$. Then as $n \to \infty$ uniformly over the $\delta_{\mathbf{i};k}$'s, we have

$$\mathbb{E}(V_{u,\theta,d,\ell}) = \tau(1-u)\sum_{\mathbf{i}\in\Xi_{0,n}}\sum_{0\le k_1,\dots,k_d\le\ell}(c_{\mathbf{i},\theta,d,\ell}^{k_1,\dots,k_d})^2$$

$$+(-1)^{\ell+1}(2\ell)!\beta_\ell^*(\frac{\omega_n\theta}{n})^{2\ell}|\Xi_{u,n}|\log(\frac{n}{\omega_n})+O\{n^d(\frac{\omega_n}{n})^{2\ell}\}$$

$$\asymp n^d\{\tau(1-u)+(\frac{\omega_n}{n})^{2\ell}\log(\frac{n}{\omega_n})\}.$$

(c) Suppose $\nu > \ell$. Then as $n \to \infty$ uniformly over the $\delta_{\mathbf{i};k}$'s, we have

$$\mathbb{E}(V_{u,\theta,d,\ell}) = \tau(1-u)\sum_{\mathbf{i}\in\Xi_{0,n}}\sum_{0\le k_1,\dots,k_d\le\ell}(c_{\mathbf{i},\theta,d,\ell}^{k_1,\dots,k_d})^2 + (-1)^\ell\beta_\ell(2\ell)!(\frac{\omega_n\theta}{n})^{2\ell}|\Xi_{u,n}|$$

$$+(\frac{\omega_n\theta}{n})^{2\ell}\sum_{\mathbf{i}\in\Xi_{u,n}}[m_X^{(0,\dots,0,\ell)}(\mathbf{x}(\mathbf{i}))]^2$$

$$+n^dO\Big\{(\frac{\omega_n}{n})^{2\nu}+(\frac{\omega_n}{n})^{2\nu}\log(\frac{n}{\omega_n})\mathcal{I}\{\nu=\ell+1\}+(\frac{\omega_n}{n})^{2\ell+1}\Big\}$$

$$\asymp n^d\{\tau(1-u)+(\frac{\omega_n}{n})^{2\ell}\}.$$

Theorem 3. *Suppose $K_Z \in \{K_P, K_M, K_W\}$ and Conditions 1, 2 hold. Let $d,\ell \in \mathbb{Z}_+$, $\tau \ge 0$, $\theta \in \{1,2\}$ and $V_{0,\theta,d,\ell}$ be as in (12). Then as $n \to \infty$ uniformly over the $\delta_{\mathbf{i};k}$'s, we have*

$$\mathrm{Var}(V_{0,\theta,d,\ell}) = O\big\{n^d\mathcal{I}\{\tau>0\}\big\} + \begin{cases} O\{n^{2d}(\omega_n/n)^{4\nu+d}\}, & \text{if } \nu < \ell - d/4,\\ O\{n^{2d}(\omega_n/n)^{4\nu+d}\log(n/\omega_n)\}, & \text{if } \nu = \ell - d/4,\\ O\{n^{2d}(\omega_n/n)^{4\ell}\}, & \text{if } \nu > \ell - d/4.\end{cases}$$

Next we write $\mathbf{Y} = (Y_1, \ldots, Y_{2|\Xi_{1,n}|})'$ where

$$X_{i_1,\ldots,i_d} = X(\mathbf{x}(\mathbf{i})), \qquad \forall \mathbf{i} = (i_1, \ldots, i_d)' \in \Xi_{1,n},$$

$$\{Y_1, \ldots, Y_{|\Xi_{1,n}|}\} = \Big\{ \frac{\nabla_{\theta,d,\ell} X_{i_1,\ldots,i_d}}{\sqrt{\mathbb{E}(V_{1,\theta,d,\ell})}} : (i_1, \ldots, i_d)' \in \Xi_{1,n} \Big\},$$

$$\{Y_{|\Xi_{1,n}|+1}, \ldots, Y_{2|\Xi_{1,n}|}\} = \Big\{ \frac{\nabla_{\theta,d,\ell} X_{i_1+1,i_2,\ldots,i_d}}{\sqrt{\mathbb{E}(V_{1,\theta,d,\ell})}} : (i_1, \ldots, i_d)' \in \Xi_{1,n} \Big\}.$$

For $1 \le j, j^* \le |\Xi_{1,n}|$ such that

$$Y_j = \frac{\nabla_{\theta,d,\ell} X_{i_1,\ldots,i_d}}{\sqrt{\mathbb{E}(V_{1,\theta,d,\ell})}}, \qquad Y_{j^*} = \frac{\nabla_{\theta,d,\ell} X_{i_1^*,\ldots,i_d^*}}{\sqrt{\mathbb{E}(V_{1,\theta,d,\ell})}},$$

$j < j^*$ if and only if there exists an $1 \le m \le d$ such that $i_a = i_a^*$, $1 \le a \le m-1$, and $i_m < i_m^*$. If

$$Y_j = \frac{\nabla_{\theta,d,\ell} X_{i_1,\ldots,i_d}}{\sqrt{\mathbb{E}(V_{1,\theta,d,\ell})}}, \qquad \forall j = 1, \ldots, |\Xi_{1,n}|,$$

then

$$Y_{|\Xi_{1,n}|+j} = \frac{\nabla_{\theta,d,\ell} X_{i_1+1,i_2,\ldots,i_d}}{\sqrt{\mathbb{E}(V_{1,\theta,d,\ell})}}.$$

Define the $2|\Xi_{1,n}| \times 2|\Xi_{1,n}|$ symmetric matrix $A = (a_{i,k})_{1 \le i,k \le 2|\Xi_{1,n}|}$ by

$$a_{j,|\Xi_{1,n}|+j} = a_{|\Xi_{1,n}|+j,j} = 1/2, \qquad \forall j = 1, \ldots, |\Xi_{1,n}|,$$
$$a_{i,k} = 0, \qquad \text{otherwise.}$$

Then

$$\frac{V_{1,\theta,d,\ell}}{\mathbb{E}(V_{1,\theta,d,\ell})} = \mathbf{Y}' A \mathbf{Y}.$$

Define the $2|\Xi_{1,n}| \times 2|\Xi_{1,n}|$ matrix $\Sigma_{\mathbf{Y}} = \mathbb{E}\{(\mathbf{Y} - \mu_{\mathbf{Y}})(\mathbf{Y} - \mu_{\mathbf{Y}})'\}$ where $\mu_{\mathbf{Y}} = \mathbb{E}\mathbf{Y}$. We observe that

$$\frac{V_{1,\theta,d,\ell}}{\mathbb{E}(V_{1,\theta,d,\ell})} = \mathbf{Z}'\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2} \mathbf{Z} + \mathbf{Z}'\Sigma_{\mathbf{Y}}^{1/2} A \mu_{\mathbf{Y}} + \mu_{\mathbf{Y}}' A \Sigma_{\mathbf{Y}}^{1/2} \mathbf{Z} + \mu_{\mathbf{Y}}' A \mu_{\mathbf{Y}},$$

where $\mathbf{Z} = (Z_1, \ldots, Z_{2|\Xi_{1,n}|})' \sim N_{2|\Xi_{1,n}|}(\mathbf{0}, I)$. Also it follows from the Hanson-Wright inequality (see Hanson and Wright (1971)) that for all $s > 0$,

$$\begin{aligned}
&\mathbb{P}\Big(|\mathbf{Z}'\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2} \mathbf{Z} - \mathbb{E}(\mathbf{Z}'\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2} \mathbf{Z})| \ge s \Big) \\
&= \mathbb{P}\Big(|\mathbf{Z}'\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2} \mathbf{Z} + \mu_{\mathbf{Y}}' A \mu_{\mathbf{Y}} - 1| \ge s \Big) \\
&\le 2 \exp\Big(- C \min\Big\{ \frac{s}{\|(\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2})_{\text{abs}}\|_2}, \frac{s^2}{\|\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2}\|_F^2} \Big\} \Big),
\end{aligned}$$

where $C > 0$ is an absolute constant and $\|.\|_2$, $\|.\|_F$ are the spectral, Frobenius norms respectively. Here $(\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2})_{\text{abs}}$ is the $2|\Xi_{1,n}| \times 2|\Xi_{1,n}|$ matrix whose elements are absolute values of the corresponding elements of $\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2}$. Since $\|(\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2})_{\text{abs}}\|_2 \leq \|\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2}\|_F$ (e.g. see Golub and Van Loan (1983) or Horn and Johnson (1985)), we have for all $s > 0$,

$$\begin{aligned} &\mathbb{P}\Big(|\mathbf{Z}'\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2}\mathbf{Z} + \mu_{\mathbf{Y}}' A \mu_{\mathbf{Y}} - 1| \geq s\Big) \\ &\leq 2\exp\Big(- C\min\Big\{\frac{s}{\|\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2}\|_F}, \frac{s^2}{\|\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2}\|_F^2}\Big\}\Big). \end{aligned} \tag{14}$$

Condition 3. *Suppose ω_n is an even integer such that $\omega_n \asymp n^\gamma$ where γ is a positive constant satisfying $1 - d/(4\nu) < \gamma < 1$.*

Define

$$q_{1,\ell,\omega_n} = \begin{cases} \big[n^{-d/2}(n/\omega_n)^{2\nu}\mathcal{I}\{\tau > 0\} + (\omega_n/n)^{d/2}\big]^{-1}, & \text{if } \nu < \ell - d/4, \\ \big[n^{-d/2}(n/\omega_n)^{2\nu}\mathcal{I}\{\tau > 0\} + (\omega_n/n)^{d/2}\log^{1/2}(n/\omega_n)\big]^{-1}, & \text{if } \nu = \ell - d/4, \\ \big[n^{-d/2}(n/\omega_n)^{2\nu}\mathcal{I}\{\tau > 0\} + (\omega_n/n)^{2\ell-2\nu}\big]^{-1}, & \text{if } \nu \in (\ell - d/4, \ell), \\ \big[n^{-d/2}(n/\omega_n)^{2\ell}\log^{-1}(n/\omega_n)\mathcal{I}\{\tau > 0\} + \log^{-1}(n/\omega_n)\big]^{-1}, & \text{if } \nu = \ell, \\ \big[n^{-d/2}(n/\omega_n)^{2\ell}\mathcal{I}\{\tau > 0\} + 1\big]^{-1}, & \text{if } \nu > \ell, \end{cases}$$

and

$$q_{2,\ell,\omega_n} = q_{1,\ell,\omega_n} \times \begin{cases} (n/\omega_n)^{2(\ell\wedge N)-2\nu}, & \text{if } \nu < \ell, \\ \log(n/\omega_n), & \text{if } \nu = \ell, \\ 1, & \text{if } \nu > \ell. \end{cases}$$

Under Condition 3, $q_{1,\ell,\omega_n} \longrightarrow \infty$ as $n \longrightarrow \infty$ if $\nu < \ell$.

Proposition 1. *Suppose $K_Z \in \{K_P, K_M, K_W\}$. Let $d, \ell \in \mathbb{Z}_+$, $\tau \geq 0$ and Conditions 1, 3 hold. Then as $n \longrightarrow \infty$ uniformly over the $\delta_{\mathbf{i};k}$'s, we have*

$$\|\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2}\|_F = O(q_{1,\ell,\omega_n}^{-1}).$$

Using Proposition 1 and (14), we obtain Theorem 4.

Theorem 4. *Suppose $K_Z \in \{K_P, K_M, K_W\}$. Let $d, \ell \in \mathbb{Z}_+, \tau \geq 0, \theta \in \{1, 2\}, V_{1,\theta,d,\ell}$ be as in (12) and Conditions 1, 3 hold. Then there exist positive constants C, C_0 such that for sufficiently large n,*

$$\begin{aligned} \mathbb{P}\Big(|\frac{V_{1,\theta,d,\ell}}{\mathbb{E}(V_{1,\theta,d,\ell})} - 1| \geq s\Big) &\leq 2\exp\big(- C\min\{q_{1,\ell,\omega_n}s, q_{1,\ell,\omega_n}^2 s^2\}\big) \\ &\quad + \min\Big\{1, \frac{C_0}{q_{2,\ell,\omega_n}^{1/2} s}\exp\big(- Cq_{2,\ell,\omega_n}s^2\big)\Big\}, \qquad \forall s > 0, \end{aligned}$$

uniformly over the $\delta_{\mathbf{i};k}$'s.

5. Estimating the nugget τ

We propose the following estimator $\hat{\tau}_\ell = \hat{\tau}_\ell(n)$ for τ based on the sample given by (6). Define

$$\hat{\tau}_\ell = \frac{V_{0,1,d,\ell}}{\sum_{\mathbf{i}\in\Xi_{0,n}} \sum_{0\le k_1,\dots,k_d\le\ell} (c_{\mathbf{i},1,d,\ell}^{k_1,\dots,k_d})^2} \tag{15}$$

where $V_{0,1,d,\ell}$ is as in (12). Theorem 5 below establishes upper bounds to the convergence rate of $\hat{\tau}_\ell$.

Theorem 5. *Suppose $K_Z \in \{K_P, K_M, K_W\}$. Let $\tau \ge 0, d, \ell \in \mathbb{Z}_+$, $\hat{\tau}_\ell$ be as in (15) and Conditions 1, 2 hold. Then*

$$\mathbb{E}\{(\hat{\tau}_\ell - \tau)^2\} = \begin{cases} O\{n^{-d}\mathcal{I}\{\tau > 0\} + (\omega_n/n)^{4\nu}\}, & \text{if } \nu < \ell, \\ O\{n^{-d}\mathcal{I}\{\tau > 0\} + (\omega_n/n)^{4\ell}\log^2(n/\omega_n)\}, & \text{if } \nu = \ell, \\ O\{n^{-d}\mathcal{I}\{\tau > 0\} + (\omega_n/n)^{4\ell}\}, & \text{if } \nu > \ell, \end{cases}$$

as $n \to \infty$ uniformly over $\delta_{\mathbf{i};k} \in [0,1), 1 \le i_1, \dots, i_d \le n, 1 \le k \le d$.

The upper bounds in Theorem 5 indicate that the convergence rate of $\hat{\tau}_\ell$ is likely to be arbitrarily slow when $\nu \to 0+$. The latter is also seen in the simulations on $\hat{\tau}_1$ in Section 9.

6. Estimating the smoothness parameter ν

For $\ell \in \mathbb{Z}_+$, define

$$\hat{\nu}_\ell = \frac{1}{2\log(2)} \log\left(\frac{V_{1,2,d,\ell} \vee \varepsilon_{n,\ell}}{V_{1,1,d,\ell} \vee \varepsilon_{n,\ell}}\right), \tag{16}$$

where $\varepsilon_{n,\ell} = n^d(\omega_n/n)^{2\ell}$. If $K_Z = K_P$, we have $\nu < 1 \le \ell$ for all $\ell \in \mathbb{Z}_+$. Then we have the following results.

Proposition 2. *Suppose $K_Z = K_P$. Let $d, \ell \in \mathbb{Z}_+$, $\tau \ge 0$, $\hat{\nu}_\ell$ be as in (16) and Conditions 1, 3 hold. Then as $n \to \infty$ uniformly over the $\delta_{\mathbf{i};k}$'s, we have*

$$\mathbb{E}\{(\hat{\nu}_\ell - \nu)^2\} = O\Big\{q_{1,\ell,\omega_n}^{-2} + q_{2,\ell,\omega_n}^{-1} + \frac{1}{\omega_n^{2\wedge(4\nu)}} + \Big(\frac{\omega_n}{n}\Big)^{(4\nu)\wedge(4\ell-4\nu)}\Big\}.$$

Proposition 3. *Suppose $K_Z \in \{K_M, K_W\}$ as in (9). Let $d, \ell \in \mathbb{Z}_+$, $\tau \ge 0$, $\nu < \ell$, $\hat{\nu}_\ell$ be as in (16) and Conditions 1, 3 hold. Then as $n \to \infty$ uniformly over the $\delta_{\mathbf{i};k}$'s, we have*

$$\begin{aligned}\mathbb{E}\{(\hat{\nu}_\ell - \nu)^2\} = O\Big\{&q_{1,\ell,\omega_n}^{-2} + q_{2,\ell,\omega_n}^{-1} + \frac{1}{\omega_n^{2\wedge(4\nu)}} + \Big(\frac{\omega_n}{n}\Big)^{(4\ell-4\nu)\wedge 4} \\ &+\Big(\frac{\omega_n}{n}\Big)^4 \log^2(n)\mathcal{I}\{\ell = \nu + 1\} + \Big(\frac{\omega_n}{n}\Big)^2 \mathcal{I}\{\beta_{\nu,1} \ne 0\}\Big\}.\end{aligned}$$

Condition 4. *Suppose ω_n is an even integer such that $\omega_n \asymp n^\gamma$ where γ is a positive constant satisfying $1 - d/4 \le \gamma < 1$.*

If $K_Z = K_P$ and assuming Condition 4 holds, we propose an alternative estimator $\hat{\nu}_P$ for ν which adaptively selects the ℓ in Proposition 2. Set ℓ be the smallest positive integer satisfying

$$\hat{\nu}_\ell \leq \ell - \frac{d}{4}, \qquad n^{-d}(\frac{n}{\omega_n})^{2\ell} V_{1,1,d,\ell} \geq (\frac{n}{\omega_n})^{d/2} \log(\frac{n}{\omega_n}).$$

If $\ell \in \{1,2\}$, define $\hat{\nu}_P = \hat{\nu}_\ell$. If $\ell \notin \{1,2\}$, define $\hat{\nu}_P = \hat{\nu}_0 = 0.99$. Hence we write $\hat{\nu}_P = \hat{\nu}_P(n) = \hat{\nu}_\ell$ where $\ell \in \{0,1,2\}$ is a random integer.

Theorem 6. *Suppose $K_Z = K_P$. Let $d \in \{1,2,3\}$, $\tau \geq 0$, $\hat{\nu}_P$ be defined as above and Conditions 1, 4 hold. Then as $n \to \infty$ uniformly over $\delta_{\mathbf{i};k} \in [0,1), 1 \leq i_1,\dots,i_d \leq n, 1 \leq k \leq d$, we have*

$$\mathbb{E}\{(\hat{\nu}_P - \nu)^2\} = O\Big\{n^{-d}(\frac{n}{\omega_n})^{4\nu} \mathcal{I}\{\tau > 0\} + (\frac{\omega_n}{n})^{d\wedge(4\nu)} + \frac{1}{\omega_n^{2\wedge(4\nu)}}\Big\}.$$

Condition 5. *Let $\ell \in \mathbb{Z}_+$ and $\omega_n = \omega_{n,\ell}$ be an even integer such that $c_{1,\ell} < \omega_{n,\ell}/n^{\gamma_\ell} < c_{2,\ell}$, $\forall n \in \mathbb{Z}_+$, where $c_{1,\ell}, c_{2,\ell}, \gamma_\ell$ are positive constants (independent of n) satisfying $1 - d/(4\ell) \leq \gamma_\ell < 1$.*

Condition 5 is verifiable since the lower bound for γ_ℓ does not depend on the unknown smoothness parameter ν (unlike Condition 3).

Remark. Under Condition 5, we shall replace ω_n by $\omega_{n,\ell}$ in the definition of $V_{u,\theta,d,\ell}$ in (12).

If $K_Z \in \{K_M, K_W\}$, we propose the following estimator $\hat{\nu}$ for ν in which a known upper bound for ν is not required as a condition for its consistency. Let M_n be an integer such that $M_n \to \infty$ and $M_n = O(n^{\gamma_0})$ as $n \to \infty$ for some positive constant γ_0. Let ℓ be the smallest positive integer satisfying

$$\hat{\nu}_\ell \leq \ell - \frac{d}{4}, \qquad n^{-d}(\frac{n}{\omega_{n,\ell}})^{2\ell} V_{1,1,d,\ell} \geq (\frac{n}{\omega_{n,\ell}})^{d/2} \log(\frac{n}{\omega_{n,\ell}}).$$

If $\ell \in \{1,\dots,M_n\}$, define $\hat{\nu} = \hat{\nu}_\ell$. If $\ell \notin \{1,\dots,M_n\}$, define $\hat{\nu} = \hat{\nu}_0 = M_n$. Hence we write $\hat{\nu} = \hat{\nu}(n) = \hat{\nu}_\ell$ where $\ell \in \{0,\dots,M_n\}$ is a random integer.

Theorem 7. *Suppose $K_Z \in \{K_M, K_W\}$ as in (9). Let $d \in \{1,2,3\}$, $\tau \geq 0$, $\hat{\nu}$ be defined as above and Conditions 1, 5 hold. Then as $n \to \infty$ uniformly over $\delta_{\mathbf{i};k} \in [0,1), 1 \leq i_1,\dots,i_d \leq n, 1 \leq k \leq d$, we have*

$$\mathbb{E}\{(\hat{\nu} - \nu)^2\} = O\Big\{n^{-d}(\frac{n}{\omega_{n,\ell_0}})^{4\nu} \mathcal{I}\{\tau > 0\} + (\frac{\omega_{n,\ell_0}}{n})^{d} + \frac{1}{\omega_{n,\ell_0}^{2\wedge(4\nu)}} + (\frac{\omega_{n,\ell_0}}{n})^{2} \mathcal{I}\{\beta_{\nu,1} \neq 0\}\Big\},$$

where ℓ_0 is the integer satisfying $\ell_0 - 1 - d/4 \leq \nu < \ell_0 - d/4$.

The convergence rate upper bounds of Theorems 6 and 7 are intriguing in that they depend on the value of the smoothness parameter ν. This is in contrast to when it is known that $\tau = 0$, i.e. no nugget (cf. Loh, Sun and Wen (2021) for $K_Z = K_M$ and Sun (2020) for $K_Z = K_W$).

Theorems 6 and 7 can be extended to larger values of d but the proofs and convergence rate upper bounds will be more complicated. In spatial statistical applications, $d \in \{1,2,3\}$ will usually suffice.

7. Estimating the parameter η

Suppose $K_Z = K_P$ as in (2). Define

$$\eta = \sigma^2 \alpha^{2\nu},$$
$$g(\nu) = -(\frac{\omega_n}{n})^{2\nu} |\Xi_{1,n}| H_1(\nu),$$

where $H_1(\nu), 0 < \nu < 1$, is as in Theorem 1. We observe from (2) that the principal irregular term of K_P is $-\sigma^2\alpha^{2\nu}\|\mathbf{x}-\mathbf{y}\|^{2\nu}$, cf. Stein (1999). Hence $-\eta = -\sigma^2\alpha^{2\nu}$ is the coefficient of the principal irregular term of K_P. Let $\hat{\nu}_1$ be as in (16) and

$$\tilde{\nu} = \begin{cases} 0, & \text{if } \hat{\nu}_1 < 0, \\ \hat{\nu}_1, & \text{if } 0 \le \hat{\nu}_1 \le 1, \\ 1, & \text{if } \hat{\nu}_1 > 1, \end{cases}$$

$$\hat{\eta}_P = \frac{V_{1,1,d,1} \vee \varepsilon_{n,1}}{g(\tilde{\nu})}, \tag{17}$$

where $\varepsilon_{n,1} = n^d(\omega_n/n)^2$. Theorem 8 computes upper bounds for the convergence rate of $\log(\hat{\eta}_P)$ as an estimator for $\log(\eta)$.

Theorem 8. *Suppose $K_Z = K_P$. Let $d \in \mathbb{Z}_+$, $\tau \ge 0$, $\hat{\eta}_P$ be as in (17) and Conditions 1, 3 hold. Then as $n \to \infty$ uniformly over the $\delta_{\mathbf{i};k}$'s, we have*

$$\mathbb{E}|\log(\hat{\eta}_P) - \log(\eta)| = O\Big\{\big[q_{1,1,\omega_n}^{-1} + q_{2,1,\omega_n}^{-1/2} + \frac{1}{\omega_n^{1\wedge(2\nu)}} + (\frac{\omega_n}{n})^{(2\nu)\wedge(2-2\nu)}\big]\log(\frac{n}{\omega_n})\Big\}.$$

Next we consider $K_Z \in \{K_M, K_W\}$. If $K_Z = K_M$, define $L = 0$, $\eta = \sigma^2\alpha^{2\nu}$ and for $n_0 \in \mathbb{Z}_+$,

$$\zeta_{n_0}(\nu) = \begin{cases} -\pi/\{2^{2\nu}\sin(\nu\pi)\Gamma(\nu)\Gamma(\nu+1)\}, & \text{if } \nu \in (0,\infty)\backslash\mathbb{Z}_+, \\ (-1)^{\nu+1}/\{2^{2\nu-1}\Gamma(\nu)\Gamma(\nu+1)\}, & \text{if } \nu \in \mathbb{Z}_+, \end{cases}$$
$$\zeta_{n_0}^c(\nu) = 0,$$
$$h_\ell(\nu) = \zeta_{n_0}(\nu) H_\ell(\nu),$$
$$g_\ell(\nu) = (\frac{\omega_{n,\ell}}{n})^{2\nu} |\Xi_{1,n}| h_\ell(\nu), \qquad \forall \nu \in (0,\ell), \ell \in \mathbb{Z}_+,$$

where $H_\ell(\nu)$ is as in Theorem 2. There has been a lot of work on estimating $\eta = \sigma^2\alpha^{2\nu}$ when $K_Z = K_M$ under the assumption that ν is known. (e.g. see Du, Zhang and Mandrekar (2009), Kaufman and Shaby (2013), Keshavarz, Nguyen and Scott (2019), Tang, Zhang and Banerjee (2021), Wang and Loh (2011), Zhang (2004)). Here the novelty is that $\tau, \nu, \alpha, \sigma^2$ and the mean function m_X are all unknown. Define $h_\ell(0) = \lim_{s\to 0+} h_\ell(s)$ and $h_\ell^{(1)}(0) = \lim_{s\to 0+} h_\ell^{(1)}(s)$ where $h_\ell^{(1)}(.)$ is the first derivative of $h_\ell(.)$. We note that ζ_{n_0} is independent of n_0 if $K_Z = K_M$. If $K_Z = K_W$, define $L = 1/2$, $\eta = \sigma^2\xi^{-2\nu}/B(2\nu,\mu)$ and

$$\zeta_{n_0}(\nu) = \begin{cases} \sum_{j=0}^{n_0}(-1)^{j+1}\binom{\nu-1/2}{j}(2\nu-2j)^{-1}, & \text{if } \nu \in [1/2,\infty)\backslash\mathbb{Z}_+, \\ (-1)^{\nu-1}\binom{\nu-1/2}{\nu}, & \text{if } \nu \in \mathbb{Z}_+, \end{cases}$$

$$\zeta^c_{n_0}(\nu) = \begin{cases} \sum_{j=n_0+1}^{\infty}(-1)^{j+1}\binom{\nu-1/2}{j}(2\nu-2j)^{-1}, & \text{if } \nu \in [1/2,\infty)\backslash\mathbb{Z}_+, \\ 0, & \text{if } \nu \in \mathbb{Z}_+, \end{cases}$$

$$h_\ell(\nu) = \zeta_{n_0}(\nu)H_\ell(\nu),$$

$$g_\ell(\nu) = (\frac{\omega_{n,\ell}}{n})^{2\nu}|\Xi_{1,n}|h_\ell(\nu), \qquad \forall \nu \in [1/2,\ell), \ell \in \mathbb{Z}_+,$$

where $n_0 \in \mathbb{Z}_+$ such that $n_0 \to \infty$ as $n \to \infty$ and for $x \geq 0$,

$$\binom{x}{0} = 1, \qquad \binom{x}{j} = \frac{\prod_{i=1}^{j}(x-i+1)}{j!}, \qquad \forall j = 1,\dots.$$

Consistent estimators for $\sigma^2\xi^{-2\nu}$ have been proposed by Bevilacqua et al. (2019) with respect to fixed-domain asymptotics under the assumptions that $\tau = 0$ and ν, μ are known. In contrast, here all four parameters (namely ν, σ^2, μ, ξ) of K_W, the mean function m_X and the nugget τ are unknown.

Lemma 3. *Suppose $K_Z \in \{K_M, K_W\}$. For $\ell - d/4 > L$,*

$$\sup_{s\in[L,\ell-d/4]} |h^{(1)}_\ell(s)/h_\ell(s)| = O(1)$$

as $n_0 \to \infty$.

If $K_Z = K_M$, pages 39 to 43 of the online supplement to Loh, Sun and Wen (2021) or pages 193 to 199 of Sun (2020) prove that $h_\ell(.)$, $h^{(1)}_\ell(.)$ are continuous functions on $[0,\ell)$ and $h_\ell(\nu) > 0$ for all $\nu \in [0,\ell)$. Hence Lemma 3 holds since $h_\ell(.)$ is independent of n_0. If $K_Z = K_W$, Sun (2020), page 124, proves that Lemma 3 holds by showing on pages 59 to 60 that $\sup_{s\in[1/2,\ell-d/4]} |\zeta^c_{n_0}(s)| = O(n_0^{-1})$ as $n_0 \to \infty$. Now suppose $K_Z \in \{K_M, K_W\}$. Let $\hat{\nu}_\ell$ be as in (16). Define for $\ell \in \mathbb{Z}_+$,

$$\tilde{\nu}_\ell = \begin{cases} L, & \text{if } \hat{\nu}_\ell < L, \\ \hat{\nu}_\ell, & \text{if } L \leq \hat{\nu}_\ell \leq \ell - d/4, \\ \ell - d/4, & \text{if } \hat{\nu}_\ell > \ell - d/4, \end{cases}$$

$$\hat{\eta}_\ell = \frac{V_{1,1,d,\ell} \vee \varepsilon_{n,\ell}}{g_\ell(\tilde{\nu}_\ell)}, \tag{18}$$

where $\varepsilon_{n,\ell} = n^d(\omega_n/n)^{2\ell}$.

Proposition 4. *Suppose $K_Z \in \{K_M, K_W\}$. Let $d, \ell \in \mathbb{Z}_+$, $\tau \geq 0$, $\nu < \ell - d/4$, $\hat{\eta}_\ell$ be as in (18) and Conditions 1, 3 hold. Then as $n \to \infty$ uniformly over the $\delta_{\mathbf{i};k}$'s, we have*

$$\begin{aligned} &\mathbb{E}|\log(\hat{\eta}_\ell) - \log(\eta)| \\ &\leq \log(1 + |\frac{\zeta^c_{n_0}(\nu)}{\zeta_{n_0}(\nu)}|) + O\Big\{\Big[q^{-1}_{1,\ell,\omega_n} + q^{-1/2}_{2,\ell,\omega_n} + (\frac{\omega_n}{n})^2 \\ &\quad + (\frac{\omega_n}{n})^2\log(n)\mathcal{I}\{\ell = \nu + 1\} + \frac{1}{\omega_n^{1\wedge(2\nu)}} + (\frac{\omega_n}{n})\mathcal{I}\{\beta_{\nu,1} \neq 0\}\Big]\log(\frac{n}{\omega_n})\Big\}. \end{aligned}$$

Next let $\hat{\nu}$ be as in Theorem 7 and $\tilde{M}_n$ be a positive constant such that $\tilde{M}_n \asymp n^{\gamma_1}$ as $n \to \infty$ for some positive constant γ_1. Define

$$\hat{\eta} = \begin{cases} 1/\tilde{M}_n, \text{ if } \hat{\nu} = \hat{\nu}_\ell \text{ for some } \ell \in \{1,\dots,M_n\} \text{ and } \hat{\eta}_\ell < 1/\tilde{M}_n, \\ \hat{\eta}_\ell, \quad \text{if } \hat{\nu} = \hat{\nu}_\ell \text{ for some } \ell \in \{1,\dots,M_n\} \text{ and } 1/\tilde{M}_n \le \hat{\eta}_\ell \le \tilde{M}_n, \\ \tilde{M}_n, \quad \text{if } \hat{\nu} = \hat{\nu}_\ell \text{ for some } \ell \in \{1,\dots,M_n\} \text{ and } \hat{\eta}_\ell > \tilde{M}_n, \\ \eta_0, \quad \text{if } \hat{\nu} = \hat{\nu}_0, \end{cases} \tag{19}$$

where η_0 is a positive constant. Theorem 9 below proves that $\hat{\eta} = \hat{\eta}(n)$ is a consistent estimator for η without the assumption of a known upper bound for ν.

Theorem 9. *Suppose $K_Z \in \{K_M, K_W\}$. Let $d \in \{1,2,3\}$, $\tau \ge 0$, $\hat{\eta}$ be as in (19) and Conditions 1, 5 hold. Then as $n \to \infty$ uniformly over $\delta_{\mathbf{i};k} \in [0,1), 1 \le i_1,\dots,i_d \le n, 1 \le k \le d$, we have*

$$\mathbb{E}|\log(\hat{\eta}) - \log(\eta)| \le \log(1 + |\frac{\zeta^c_{n_0}(\nu)}{\zeta_{n_0}(\nu)}|) + O\Big\{\Big[n^{-d/2}(\frac{n}{\omega_{n,\ell_0}})^{2\nu}\mathcal{I}\{\tau > 0\} + (\frac{\omega_{n,\ell_0}}{n})^{d/2} + \frac{1}{\omega_{n,\ell_0}^{1\wedge(2\nu)}} + (\frac{\omega_{n,\ell_0}}{n})\mathcal{I}\{\beta_{\nu,1} \ne 0\}\Big] \log(\frac{n}{\omega_{n,\ell_0}})\Big\},$$

where ℓ_0 is the integer satisfying $\ell_0 - 1 - d/4 \le \nu < \ell_0 - d/4$.

The following is an immediate corollary of Theorem 9.

Corollary 1. *Suppose the assumptions of Theorem 9 hold.*
(i) If $K_Z = K_M$,

$$\mathbb{E}|\log(\hat{\eta}) - \log(\eta)| = O\Big\{\Big[n^{-d/2}(\frac{n}{\omega_{n,\ell_0}})^{2\nu}\mathcal{I}\{\tau > 0\} + (\frac{\omega_{n,\ell_0}}{n})^{d/2} + \frac{1}{\omega_{n,\ell_0}^{1\wedge(2\nu)}}\Big] \log(\frac{n}{\omega_{n,\ell_0}})\Big\},$$

as $n \to \infty$ uniformly over $\delta_{\mathbf{i};k} \in [0,1), 1 \le i_1,\dots,i_d \le n, 1 \le k \le d$.
(ii) If $K_Z = K_W$,

$$\mathbb{E}|\log(\hat{\eta}) - \log(\eta)| = O\Big\{\frac{1}{n_0} + \Big[n^{-d/2}(\frac{n}{\omega_{n,\ell_0}})^{2\nu}\mathcal{I}\{\tau > 0\} + (\frac{\omega_{n,\ell_0}}{n})^{(d/2)\wedge 1} + \frac{1}{\omega_{n,\ell_0}^{1\wedge(2\nu)}}\Big] \log(\frac{n}{\omega_{n,\ell_0}})\Big\},$$

as $n \to \infty$ uniformly over $\delta_{\mathbf{i};k} \in [0,1), 1 \le i_1,\dots,i_d \le n, 1 \le k \le d$.

Finally, the $O(.)$'s in Sections 4 to 7 are uniform over $\delta_{\mathbf{i};k} \in [0,1), 1 \le i_1,\dots,i_d \le n, 1 \le k \le d$. Hence the lemmas, propositions and theorems in these sections hold when the $\delta_{\mathbf{i};k}$'s are random and are independent of the Gaussian random field X.

8. Space-filling designs

This section extends the results for stratified sampling in Sections 5 to 7 to more general space-filling designs, e.g. see Chapter 5 of Santner, Williams and Notz (2003). A key feature of such designs on

$[0,1)^d$ is that, with high probability, the n design sites get to be dense in $[0,1)^d$ as $n \to \infty$. The space-filling designs include i.i.d. sampling, Latin hypercube sampling Loh (1996), McKay, Beckman and Conover (1979), Stein (1987) and orthogonal array sampling Owen (1992).

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be a sequence of random vectors in $[0,1)^d$ that are independent of the Gaussian random field X in (1). Let $n_1 = n_1(n) \in \mathbb{Z}_+$ such that $n_1 \le \lfloor n^{1/d} \rfloor$ and $n_1 \to \infty$ as $n \to \infty$. For brevity, we write $\mathbf{i} = (i_1, \ldots, i_d)' \in \mathbb{Z}_+^d$ and

$$\Omega_{\mathbf{i};n_1} = [\frac{i_1 - 1}{n_1}, \frac{i_1}{n_1}) \times \ldots \times [\frac{i_d - 1}{n_1}, \frac{i_d}{n_1}), \qquad \forall 1 \le i_1, \ldots, i_d \le n_1,$$

$$p_n = \mathbb{P}\Big(\bigcup_{1 \le i_1, \ldots, i_d \le n_1} \bigcap_{j=1}^{n} \{\mathbf{x}_j \notin \Omega_{\mathbf{i};n_1}\} \Big). \tag{20}$$

If $\sum_{n=1}^{\infty} p_n < \infty$, it follows from the Borel-Cantelli lemma that

$$\mathbb{P}\Big(\bigcup_{1 \le i_1, \ldots, i_d \le n_1} \bigcap_{j=1}^{n} \{\mathbf{x}_j \notin \Omega_{\mathbf{i};n_1}\} \ \text{ i.o.} \Big) = 0. \tag{21}$$

Let $\hat{n}$ be the largest integer less than or equal to n_1 such that

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \cap \Omega_{\mathbf{i};\hat{n}} \neq \emptyset, \qquad \forall 1 \le i_1, \ldots, i_d \le \hat{n}.$$

Consequently it follows from (20) that $p_n = \mathbb{P}(\hat{n} \neq n_1)$ and from (21) that $\hat{n} \to n_1$ as $n \to \infty$ almost surely. Denote the $\mathbf{x}_j \in \Omega_{\mathbf{i};\hat{n}}$ as

$$\mathbf{x}(\mathbf{i}) = \big(x_1(\mathbf{i}), \ldots, x_d(\mathbf{i})\big)', \qquad \forall 1 \le i_1, \ldots, i_d \le \hat{n},$$

such that

$$x_k(\mathbf{i}) = \frac{i_k - 1}{\hat{n}} + \frac{\delta_{\mathbf{i};k}}{\hat{n}},$$

where $0 \le \delta_{\mathbf{i};k} < 1$, $k = 1, \ldots, d$. If there are more than one $\mathbf{x}_j$'s in $\Omega_{\mathbf{i};\hat{n}}$, we can choose any of these $\mathbf{x}_j$'s to be $\mathbf{x}(\mathbf{i})$. Now given $\mathbf{x}(\mathbf{i}) \in \Omega_{\mathbf{i};\hat{n}}, 1 \le i_1, \ldots, i_d \le \hat{n}$, this randomized sampling design reduces to the stratified sampling design of Section 3 where the observed sample is

$$\Big\{ \big\{\mathbf{x}(\mathbf{i}), X(\mathbf{x}(\mathbf{i}))\big\} : \mathbf{x}(\mathbf{i}) \in \Omega_{\mathbf{i};\hat{n}}, 1 \le i_1, \ldots, i_d \le \hat{n} \Big\}.$$

We note that the effective sample size is correspondingly reduced from n to $\hat{n}^d$. For $\ell \in \mathbb{Z}_+$, define the ℓth-order quadratic variation $V_{u,\theta,d,\ell} = V_{u,\theta,d,\ell}(\hat{n})$ to be as in Section 4 where n is replaced by $\hat{n}$.

Let $\hat{\tau}_\ell = \hat{\tau}_\ell(n)$ be as in Section 5 and $\hat{\tau}_\ell(\hat{n})$ be the former with n replaced by $\hat{n}$. We propose the following estimator $\hat{\tau}_\ell^{\mathrm{S}}$ for τ where the superscript S denotes "space-filling". Define

$$\hat{\tau}_\ell^{\mathrm{S}} = \begin{cases} \hat{\tau}_\ell(\hat{n}), & \text{if } \hat{\tau}_\ell(\hat{n}) \le M_n^{\mathrm{S}}, \\ M_n^{\mathrm{S}}, & \text{if } \hat{\tau}_\ell(\hat{n}) > M_n^{\mathrm{S}}, \end{cases} \tag{22}$$

for some positive constant M_n^{S} such that $M_n^{\mathrm{S}} \to \infty$ as $n \to \infty$. Now Theorem 10 follows from Theorem 5.

Theorem 10. *Suppose $K_Z \in \{K_P, K_M, K_W\}$. Let $\tau \geq 0, d, \ell \in \mathbb{Z}_+$, p_n, $\hat{\tau}_\ell^{\mathrm{S}}$ be as in (20), (22) respectively and Conditions 1, 2 hold. Then as $n \to \infty$,*

$$\mathbb{E}\{(\hat{\tau}_\ell^{\mathrm{S}} - \tau)^2\} = \begin{cases} O\big\{n_1^{-d}\mathcal{I}\{\tau > 0\} + (\omega_{n_1}/n_1)^{4\nu} + p_n (M_n^{\mathrm{S}})^2\big\}, & \text{if } \nu < \ell, \\ O\big\{n_1^{-d}\mathcal{I}\{\tau > 0\} + (\omega_{n_1}/n_1)^{4\ell} \log^2(n_1/\omega_{n_1}) + p_n (M_n^{\mathrm{S}})^2\big\}, & \text{if } \nu = \ell, \\ O\big\{n_1^{-d}\mathcal{I}\{\tau > 0\} + (\omega_{n_1}/n_1)^{4\ell} + p_n (M_n^{\mathrm{S}})^2\big\}, & \text{if } \nu > \ell. \end{cases}$$

Consequently $\hat{\tau}_\ell^{\mathrm{S}} \to \tau$ in probability and $\hat{\tau}_\ell^{\mathrm{S}}$ is a consistent estimator for τ if $p_n (M_n^{\mathrm{S}})^2 \to 0$ as $n \to \infty$.

Let $\hat{\nu}_P = \hat{\nu}_P(n)$ and $\hat{\nu} = \hat{\nu}(n)$ be as in Section 6. Define

$$\hat{\nu}_P^{\mathrm{S}} = \begin{cases} 0 & \text{if } \hat{\nu}_P(\hat{n}) \leq 0, \\ \hat{\nu}_P(\hat{n}) & \text{if } 0 < \hat{\nu}_P(\hat{n}) < 1, \\ 1 & \text{if } \hat{\nu}_P(\hat{n}) \geq 1, \end{cases}$$

$$\hat{\nu}^{\mathrm{S}} = \begin{cases} \hat{\nu}(\hat{n}) & \text{if } \hat{\nu}(\hat{n}) \geq 0, \\ 0 & \text{if } \hat{\nu}(\hat{n}) < 0. \end{cases}$$

Theorems 11 and 12 follows from Theorems 6 and 7 respectively.

Theorem 11. *Suppose $K_Z = K_P$. Let $d \in \{1,2,3\}$, $\tau \geq 0$, $\hat{\nu}_P^{\mathrm{S}}$ be as above and Conditions 1, 4 hold. Then as $n \to \infty$,*

$$\mathbb{E}\{(\hat{\nu}_P^S - \nu)^2\} = O\Big\{n_1^{-d}\big(\frac{n_1}{\omega_{n_1}}\big)^{4\nu}\mathcal{I}\{\tau > 0\} + \big(\frac{\omega_{n_1}}{n_1}\big)^{d \wedge (4\nu)} + \frac{1}{\omega_{n_1}^{2\wedge(4\nu)}} + p_n\Big\}.$$

Consequently $\hat{\nu}_P^S \to \nu$ in probability and $\hat{\nu}_P^S$ is a consistent estimator for ν if $p_n \to 0$ as $n \to \infty$.

Theorem 12. *Suppose $K_Z \in \{K_M, K_W\}$ as in (9). Let $d \in \{1,2,3\}$, $\tau \geq 0$, $\hat{\nu}^{\mathrm{S}}$ be as above and Conditions 1, 5 hold. Then as $n \to \infty$,*

$$\mathbb{E}\{(\hat{\nu}^{\mathrm{S}} - \nu)^2\} = O\Big\{n_1^{-d}\big(\frac{n_1}{\omega_{n_1,\ell_0}}\big)^{4\nu}\mathcal{I}\{\tau > 0\} + \big(\frac{\omega_{n_1,\ell_0}}{n_1}\big)^{d} + \frac{1}{\omega_{n_1,\ell_0}^{2\wedge(4\nu)}} + \big(\frac{\omega_{n_1,\ell_0}}{n_1}\big)^{2}\mathcal{I}\{\beta_{\nu,1} \neq 0\} + p_n M_n^2\Big\},$$

where ℓ_0 is the integer satisfying $\ell_0 - 1 - d/4 \leq \nu < \ell_0 - d/4$. Consequently $\hat{\nu}^{\mathrm{S}} \to \nu$ in probability and $\hat{\nu}^{\mathrm{S}}$ is a consistent estimator for ν if $p_n M_n^2 \to 0$ as $n \to \infty$.

Next suppose $\tilde{M}_n$ is a positive constant such that $\tilde{M}_n \to \infty$ as $n \to \infty$. Let $\hat{\eta}_P$ be as in (17) and $\hat{\eta}$ be as in (19). Define

$$\tilde{\eta}_P^S(n) = \begin{cases} 1/\tilde{M}_n & \text{if } \hat{\eta}_P < 1/\tilde{M}_n, \\ \hat{\eta}_P & \text{if } 1/\tilde{M}_n \leq \hat{\eta}_P \leq \tilde{M}_n, \\ \tilde{M}_n & \text{if } \hat{\eta}_P > \tilde{M}_n, \end{cases}$$

$\hat{\eta}_P^{\mathrm{S}} = \tilde{\eta}_P^S(\hat{n})$ and $\hat{\eta}^{\mathrm{S}} = \hat{\eta}(\hat{n})$. Theorems 13 and 14 follow from Theorems 8 and 9 respectively.

Theorem 13. *Suppose $K_Z = K_P$. Let $d \in \mathbb{Z}_+$, $\tau \geq 0$, $\hat{\eta}_P^S$ be as above and Conditions 1, 3 hold. Then as $n \to \infty$,*

$$\begin{aligned}
&\mathbb{E}|\log(\hat{\eta}_P^{\mathrm{S}}) - \log(\eta)| \\
&= O\Big\{\big[q_{1,1,\omega_{n_1}}^{-1} + q_{2,1,\omega_{n_1}}^{-1/2} + \frac{1}{\omega_{n_1}^{1\wedge(2\nu)}} + (\frac{\omega_{n_1}}{n_1})^{(2\nu)\wedge(2-2\nu)}\big] \log(\frac{n_1}{\omega_{n_1}}) + p_n \log(\tilde{M}_n)\Big\}.
\end{aligned}$$

Consequently $\hat{\eta}_P^{\mathrm{S}} \to \eta$ in probability and $\hat{\eta}_P^{\mathrm{S}}$ is a consistent estimator for η if $p_n \log(\tilde{M}_n) \to 0$ as $n \to \infty$.

Theorem 14. *Suppose $K_Z \in \{K_M, K_W\}$. Let $d \in \{1,2,3\}$, $\tau \geq 0$, $\hat{\eta}^{\mathrm{S}}$ be as above and Conditions 1, 5 hold. Then as $n \to \infty$,*

$$\begin{aligned}
&\mathbb{E}|\log(\hat{\eta}^{\mathrm{S}}) - \log(\eta)| \\
&\leq \log(1 + |\frac{\zeta_{n_0}^c(\nu)}{\zeta_{n_0}(\nu)}|) + O\Big\{\Big[n_1^{-d/2}(\frac{n_1}{\omega_{n_1,\ell_0}})^{2\nu} \mathcal{I}\{\tau > 0\} + (\frac{\omega_{n_1,\ell_0}}{n_1})^{d/2} + \frac{1}{\omega_{n_1,\ell_0}^{1\wedge(2\nu)}} \\
&\quad + (\frac{\omega_{n_1,\ell_0}}{n_1}) \mathcal{I}\{\beta_{\nu,1} \neq 0\}\Big] \log(\frac{n_1}{\omega_{n_1,\ell_0}}) + p_n \log(\tilde{M}_n)\Big\},
\end{aligned}$$

where ℓ_0 is the integer satisfying $\ell_0 - 1 - d/4 \leq \nu < \ell_0 - d/4$. Consequently $\hat{\eta}^{\mathrm{S}} \to \eta$ in probability and $\hat{\eta}^{\mathrm{S}}$ is a consistent estimator for η if $p_n \log(\tilde{M}_n) \to 0$ as $n \to \infty$.

As illustration, we shall present upper bounds for p_n for the following two designs: i.i.d. sampling and Latin hypercube sampling.

Example 1 (i.i.d. design). Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be a sequence of i.i.d. random vectors where $\mathbf{x}_1$ has probability density function $p(\mathbf{x})$, $\mathbf{x} \in [0,1)^d$, satisfying $\int_{[0,1)^d} p(\mathbf{x}) d\mathbf{x} = 1$ and $\inf_{[0,1)^d} p(\mathbf{x}) = p_0 > 0$. Set $n_1 = \lfloor (\frac{n}{\log^2(n)})^{1/d} \rfloor$. Then it follows from Loh, Sun and Wen (2021) that

$$p_n \leq \frac{n^{1-p_0 \log(n)}}{\log^2(n)}.$$

Example 2 (Latin hypercube design). Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be a Latin hypercube sample in $[0,1)^d$. The main feature of Latin hypercube sampling is that, in contrast to simple random sampling, it stratifies on each input dimension. More precisely, for positive integers d and n, let

(1) $\pi_1, \ldots, \pi_d$ be random permutations of $\{1, \ldots, n\}$ each uniformly distributed over all the $n!$ possible permutations,
(2) $U_{i_1,\ldots,i_d,j}, 1 \leq i_1, \ldots, i_d \leq n, 1 \leq j \leq d$, be $[0,1)$ uniform random variables,
(3) the $U_{i_1,\ldots,i_d,j}$'s, π_k's and the Gaussian random field X all be stochastically independent.

A Latin hypercube sample of size n (taken from the d-dimensional hypercube $[0,1)^d$) is defined to be $\{\mathbf{x}_i = \mathbf{x}(\pi_1(i), \ldots, \pi_d(i)) : i = 1, \ldots, n\}$ where for all $1 \leq i_1, \ldots, i_d \leq n$,

$$\begin{aligned}
\mathbf{x}(i_1, \ldots, i_d) &= (x_1(i_1, \ldots, i_d), \ldots, x_d(i_1, \ldots, i_d)), \\
x_j(i_1, \ldots, i_d) &= (i_j - 1 + U_{i_1,\ldots,i_d,j})/n, \qquad \forall j = 1, \ldots, d.
\end{aligned}$$

If $d = 1$, the Latin hypercube sample $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is just a stratified sample as in (10) where the $\delta_{\mathbf{i},k}$'s are i.i.d. $U[0,1)$ random variables.

Hence we assume hereafter that $d \in \{2,3\}$. Let $n_1 = \lfloor \{n/\log^c(n)\}^{1/d} \rfloor$ where $c > 1$ is a constant and $a_n = \lfloor n/n_1 \rfloor \sim n^{1-1/d} \log^{c/d}(n)$. Then $a_n/n \leq 1/n_1$. We observe that for $1 \leq i_1, \ldots, i_d \leq n_1$,

$$\mathbb{P}\Big(\bigcap_{j=1}^{n} \{\mathbf{x}_j \notin \Omega_{i_1,\ldots,i_d;n_1}\}\Big) \leq \mathbb{P}\Big(\bigcap_{j=1}^{n} \{\mathbf{x}_j \notin [0, \frac{a_n - 2}{n})^d\}\Big).$$

Now,

$$\begin{aligned}
&\mathbb{P}\Big(\big|\{(x_{j1}, x_{j2})' \in [0, \frac{a_n - 2}{n})^2 : j \in \{1, \ldots, n\}\}\big| = k\Big) \\
&= \binom{a_n - 2}{k} \Big\{ \prod_{s=1}^{k} (\frac{a_n - 2 - s + 1}{n}) \Big\} \Big\{ \prod_{t=1}^{a_n - 2 - k} \frac{n - (a_n - 2) - t + 1}{n} \Big\}
\end{aligned}$$

for $k = 0, \ldots, a_n - 2$. Hence for sufficiently large n,

$$\begin{aligned}
p_n &= \mathbb{P}\Big(\bigcup_{1 \leq i_1,\ldots,i_d \leq n_1} \bigcap_{j=1}^{n} \{\mathbf{x}_j \notin \Omega_{i_1,\ldots,i_d;n_1}\}\Big) \\
&\leq \sum_{1 \leq i_1,\ldots,i_d \leq n_1} \mathbb{P}\Big(\bigcap_{j=1}^{n} \{\mathbf{x}_j \notin [0, \frac{a_n - 2}{n})^d\}\Big) \\
&\leq n_1^d \Big\{ \mathcal{I}\{d = 2\} \mathbb{P}\Big(\big|\{(x_{j1}, x_{j2})' \in [0, \frac{a_n - 2}{n})^2 : j \in \{1, \ldots, n\}\}\big| = 0\Big) \\
&\quad + \mathcal{I}\{d = 3\} \sum_{k=0}^{a_n - 2} \Big\{ \prod_{u=1}^{k} \frac{n - (a_n - 2) - u + 1}{n} \Big\} \\
&\quad \times \mathbb{P}\Big(|\{(x_{j1}, x_{j2})' \in [0, \frac{a_n - 2}{n})^2 : j \in \{1, \ldots, n\}\}| = k\Big) \Big\} \\
&\leq n_1^d \Big\{ \mathcal{I}\{d = 2\} \Big\{ \prod_{i=1}^{a_n - 2} 1 - \frac{(a_n - 2) + i - 1}{n} \Big\} + \mathcal{I}\{d = 3\} \sum_{k=0}^{a_n - 2} \binom{a_n - 2}{k} \\
&\quad \times \Big\{ \prod_{u=1}^{k} \frac{n - (a_n - 2) - u + 1}{n} \Big\} \Big\{ \prod_{s=1}^{k} \frac{(a_n - 2) - s + 1}{n} \Big\} \\
&\quad \times \Big\{ \prod_{t=1}^{a_n - 2 - k} \frac{n - (a_n - 2) - t + 1}{n} \Big\} \Big\} \\
&\leq n_1^d \Big\{ \mathcal{I}\{d = 2\} (1 - \frac{a_n - 2}{n})^{a_n - 2} \\
&\quad + \mathcal{I}\{d = 3\} \sum_{k=0}^{a_n - 2} \binom{a_n - 2}{k} \Big\{ \frac{(a_n - 2)(n - a_n + 2)}{n^2} \Big\}^k (\frac{n - a_n + 2}{n})^{a_n - 2 - k} \Big\} \\
&= n_1^d \{1 - (\frac{a_n - 2}{n})^{d-1}\}^{a_n - 2} \leq n_1^d e^{-(a_n - 2)^d / n^{d-1}}
\end{aligned}$$

$$\sim \frac{n}{\log^c(n)} e^{-\log^c(n)},$$

as $n \longrightarrow \infty$.

9. Simulation study

Let X be a Gaussian random field as in (1) where $d \in \{1,2\}$, m_X is a constant and $K_Z \in \{K_P, K_M\}$. For simulations when $K_Z = K_W$, we refer the reader to Sun (2020). This section presents a simulation study to gauge the finite sample accuracy of the estimators for τ, ν and η where the design sites $\mathbf{x}(\mathbf{i}), 1 \le i_1, \ldots, i_d \le n$, are chosen via stratified sampling as in Section 3 such that the $\delta_{\mathbf{i};k}$'s are i.i.d. $U[0,1)$ and are independent of X. Consequently, the sample size is n^d.

The computation of the estimators is relatively quick without any high-dimensional nonconvex optimization. The difficulty lies in the accurate simulation of the $X(\mathbf{x}(\mathbf{i}))$'s when sample size n^d is large and smoothness parameter is not small; e.g. see Lindgren, Rue and Lindström (2011), Stein (2012), Wood and Chan (1994) and the references cited therein.

NUGGET PARAMETER τ

Powered exponential. Suppose X is as in (1) with $m_X = 1$ and $K_Z = K_P$ in (2). We perform two experiments below on $\hat{\tau}_1$. In each experiment, set $\sigma^2 = \alpha = 1$, $\omega_n = 2\lfloor n^{1/4}/2 \rfloor, \nu \in \{0.1, 0.5, 0.9\}$ and $\tau \in \{0, 0.5, 1, 1.5, 2\}$.

Experiment 1. Set $d = 1$, $n \in \{500, 1000\}$, see Table 1.

Experiment 2. Set $d = 2$, $n \in \{40, 60\}$, see Table 2.

Table 1. (Powered exponential) Estimated mean absolute errors of $\hat{\tau}_1$, i.e. $\mathbb{E}|\hat{\tau}_1 - \tau|$, with stratified sampling design when $d = 1$ replicated independently 100 times (standard errors within parentheses)

		$n = 500$			$n = 1000$	
τ	$\nu = 0.1$	$\nu = 0.5$	$\nu = 0.9$	$\nu = 0.1$	$\nu = 0.5$	$\nu = 0.9$
0	0.311(0.002)	0.008(0.000)	0.000(0.000)	0.284(0.001)	0.004(0.000)	0.000(0.000)
0.5	0.302(0.006)	0.031(0.003)	0.029(0.002)	0.286(0.004)	0.023(0.002)	0.022(0.002)
1.0	0.324(0.010)	0.057(0.004)	0.063(0.004)	0.273(0.007)	0.050(0.004)	0.039(0.003)
1.5	0.299(0.014)	0.092(0.007)	0.093(0.007)	0.291(0.009)	0.066(0.005)	0.070(0.005)
2.0	0.319(0.015)	0.110(0.008)	0.127(0.010)	0.294(0.013)	0.087(0.006)	0.095(0.006)

Table 2. (Powered exponential) Estimated mean absolute errors of $\hat{\tau}_1$, i.e. $\mathbb{E}|\hat{\tau}_1 - \tau|$, with stratified sampling design when $d = 2$ replicated independently 100 times (standard errors within parentheses)

		$n = 40$			$n = 60$	
τ	$\nu = 0.1$	$\nu = 0.5$	$\nu = 0.9$	$\nu = 0.1$	$\nu = 0.5$	$\nu = 0.9$
0	0.414(0.002)	0.043(0.000)	0.003(0.000)	0.389(0.001)	0.029(0.000)	0.002(0.000)
0.5	0.415(0.005)	0.042(0.003)	0.022(0.002)	0.389(0.003)	0.028(0.002)	0.013(0.001)
1.0	0.416(0.008)	0.062(0.004)	0.040(0.003)	0.390(0.005)	0.042(0.003)	0.027(0.002)
1.5	0.424(0.010)	0.070(0.005)	0.065(0.005)	0.390(0.006)	0.050(0.004)	0.045(0.004)
2.0	0.411(0.013)	0.088(0.007)	0.090(0.007)	0.368(0.008)	0.060(0.005)	0.053(0.004)

Table 3. (Matérn) Estimated mean absolute errors of $\hat{\tau}_1$ with stratified sampling design when $d = 1$ replicated independently 100 times (standard errors within parentheses)

τ	$n = 500$ $\nu = 0.5$	$\nu = 1$	$\nu = 1.5$	$n = 1000$ $\nu = 0.5$	$\nu = 1$	$\nu = 1.5$
0	0.008(0.000)	0.000(0.000)	0.000(0.000)	0.004(0.000)	0.000(0.000)	0.000(0.000)
0.5	0.031(0.002)	0.027(0.002)	0.028(0.002)	0.021(0.002)	0.022(0.002)	0.022(0.002)
1.0	0.072(0.005)	0.063(0.005)	0.056(0.004)	0.049(0.004)	0.039(0.003)	0.044(0.003)
1.5	0.094(0.008)	0.118(0.009)	0.096(0.006)	0.071(0.005)	0.068(0.005)	0.066(0.005)
2.0	0.116(0.008)	0.141(0.010)	0.126(0.010)	0.080(0.007)	0.088(0.006)	0.085(0.006)

Table 4. (Matérn) Estimated mean absolute errors of $\hat{\tau}_1$ with stratified sampling design when $d = 2$ replicated independently 100 times (standard errors within parentheses)

τ	$n = 40$ $\nu = 0.5$	$\nu = 1$	$\nu = 1.5$	$n = 60$ $\nu = 0.5$	$\nu = 1$	$\nu = 1.5$
0	0.043(0.000)	0.004(0.000)	0.001(0.000)	0.029(0.000)	0.002(0.000)	0.000(0.000)
0.5	0.049(0.003)	0.021(0.002)	0.022(0.002)	0.027(0.002)	0.013(0.001)	0.014(0.001)
1.0	0.067(0.005)	0.041(0.003)	0.043(0.003)	0.033(0.002)	0.030(0.002)	0.029(0.002)
1.5	0.069(0.006)	0.074(0.006)	0.064(0.005)	0.061(0.004)	0.037(0.003)	0.040(0.003)
2.0	0.089(0.007)	0.084(0.006)	0.090(0.007)	0.072(0.005)	0.067(0.004)	0.065(0.005)

Matérn. Suppose X is as in (1) with $m_X = 1$ and $K_Z = K_M$ in (3). We perform two experiments below on $\hat{\tau}_1$. In each experiment, set $\sigma^2 = \alpha = 1$, $\omega_n = 2\lfloor n^{1/4}/2\rfloor, \nu \in \{0.5, 1, 1.5\}$ and $\tau \in \{0, 0.5, 1, 1.5, 2\}$.

Experiment 3. Set $d = 1$, $n \in \{500, 1000\}$, see Table 3.

Experiment 4. Set $d = 2$, $n \in \{40, 60\}$, see Table 4.

Discussion. Tables 1 to 4 show that $\hat{\nu}_1$ performs well except when $\nu = 0.1$. The latter is in line with the upper bounds in Theorem 5 which indicate that the convergence rate of $\hat{\tau}_1$ is likely to be arbitrarily slow when $\nu \to 0+$.

SMOOTHNESS PARAMETER ν AND MICROERGODIC PARAMETER η

Powered exponential. Suppose X is as in (1) with $m_X = 1$ and $K_Z = K_P$ in (2) with $\sigma^2 = 5, \alpha = 3, \tau = 0.5$. We perform two experiments below. In each experiment, $\hat{\nu}_1$ denotes the estimator of ν as in (16) and $\hat{\eta}_P$ denotes the estimator of $\eta = \sigma^2\alpha^{2\nu}$ as in (17).

The loss functions used for $\hat{\nu}_1$ and $\hat{\eta}_P$ are $|\hat{\nu}_1 - \nu|$ and $|\log(\hat{\eta}_P)/\log(\eta) - 1|$ respectively.

Experiment 5. Set $d = 1$, $n \in \{1000, 2000\}$, $\omega_n = 2\lfloor n^{1-d/4}/20\rfloor$, $\nu \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, see Table 5.

Experiment 6. Set $d = 2$, $n \in \{40, 80\}$, $\omega_n = 2\lfloor n^{1-d/4}/2\rfloor$, $\nu \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, see Table 6.

Matérn. Suppose X is as in (1) with $m_X = 1$ and $K_Z = K_M$ in (3) with $\sigma^2 = 5, \alpha = 3, \tau = 0.5$. We perform two experiments below. In each experiment,

- $\hat{\nu}$ denotes the estimator of ν in Theorem 7 where $M_n = 4$.
- $\hat{\eta}$ denotes the estimator of $\eta = \sigma^2\alpha^{2\nu}$ in (19) with $\tilde{M}_n = 100$, $\eta_0 = 1$.

Table 5. (Powered exponential) Estimated mean absolute errors of $\hat{\nu}_1$ and $\log(\hat{\eta}_P)/\log(\eta)$ with stratified sampling design when $d = 1$ replicated independently 100 times (standard errors within parentheses)

ν	$\lvert\hat{\nu}_1 - \nu\rvert$ $n = 1000$	$n = 2000$	$\lvert\log(\hat{\eta}_P)/\log(\eta) - 1\rvert$ $n = 1000$	$n = 2000$
0.1	0.109 (0.008)	0.079 (0.006)	0.434(0.031)	0.403(0.027)
0.3	0.067 (0.005)	0.058 (0.004)	0.252(0.020)	0.210(0.015)
0.5	0.089 (0.007)	0.079 (0.006)	0.273(0.021)	0.251(0.020)
0.7	0.153 (0.010)	0.113 (0.009)	0.340(0.023)	0.286(0.020)
0.9	0.203 (0.019)	0.189 (0.017)	0.579(0.056)	0.511(0.053)

Table 6. (Powered exponential) Estimated mean absolute errors of $\hat{\nu}_1$ and $\log(\hat{\eta}_P)/\log(\eta)$ with stratified sampling design when $d = 2$ replicated independently 100 times (standard errors within parentheses)

ν	$\lvert\hat{\nu}_1 - \nu\rvert$ $n = 40$	$n = 80$	$\lvert\log(\hat{\eta}_P)/\log(\eta) - 1\rvert$ $n = 40$	$n = 80$
0.1	0.199 (0.017)	0.118(0.008)	0.747(0.038)	0.676(0.026)
0.3	0.116 (0.009)	0.076(0.006)	0.393(0.024)	0.332(0.018)
0.5	0.144 (0.010)	0.104(0.009)	0.313(0.021)	0.272(0.020)
0.7	0.204 (0.015)	0.107(0.009)	0.354(0.023)	0.225(0.017)
0.9	0.204 (0.013)	0.115(0.008)	0.284(0.019)	0.204(0.015)

Table 7. (Matérn) Estimated mean absolute errors of $\hat{\nu}$ and $\log(\hat{\eta})/\log(\eta)$ with stratified sampling design when $d = 1$ replicated independently 100 times (standard errors within parentheses)

ν	$\lvert\hat{\nu} - \nu\rvert$ $n = 1000$	$n = 2000$	$n = 4000$	$\lvert\log(\hat{\eta})/\log(\eta) - 1\rvert$ $n = 1000$	$n = 2000$	$n = 4000$
0.1	0.114 (0.009)	0.090 (0.006)	0.081 (0.006)	0.359(0.025)	0.288(0.021)	0.281 (0.021)
0.3	0.084 (0.006)	0.059 (0.005)	0.055 (0.004)	0.278(0.021)	0.226(0.016)	0.217 (0.017)
0.5	0.085 (0.006)	0.073 (0.006)	0.054 (0.004)	0.237(0.017)	0.222(0.020)	0.165 (0.013)
0.7	0.191 (0.019)	0.134 (0.013)	0.102 (0.009)	0.419(0.034)	0.323(0.029)	0.250 (0.022)
0.9	0.370 (0.020)	0.346 (0.035)	0.282 (0.020)	0.631(0.036)	0.563(0.040)	0.541 (0.041)

The loss functions used for $\hat{\nu}$ and $\hat{\eta}$ are $|\hat{\nu} - \nu|$ and $|\log(\hat{\eta})/\log(\eta) - 1|$ respectively.

Experiment 7. Set $d = 1$, $n \in \{1000, 2000, 4000\}$, $\omega_{n,\ell} = 2\lfloor n^{1-d/(4\ell)}/20\rfloor$, $\nu \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, see Table 7.

Experiment 8. Set $d = 2$, $n \in \{40, 60, 80\}$, $\omega_{n,\ell} = 2\lfloor n^{1-d/(4\ell)}/(4\ell - 2)\rfloor$, $\nu \in \{0.1, 0.3, 0.5, 0.7, 0.9, 1.1, 1.3\}$, see Table 8.

Discussion. Tables 5 to 8 indicate that $\hat{\nu}_1$ and $\hat{\nu}$ perform reasonably well for relatively small values of ν. For larger values of ν, larger sample sizes are needed for the asymptotics of Proposition 2 and Theorem 7 to apply. On the other hand, Tables 5 to 8 indicate that much larger sample sizes are needed for the estimators $\hat{\eta}_P$ and $\hat{\eta}$ to perform well. For the reproducibility of the simulation studies, we have uploaded as supplementary material a *Mathematica* program that was used to compute Table 7 of Experiment 7. This program sets $n = 4000$, $\tau = 0.5$, $\sigma^2 = 5$, $\alpha = 3$ and $\nu = 0.5$.

For a given n, selecting an appropriate ω_n is a challenging problem. We do not have a satisfactory answer and much more work is needed in this direction. As it stands, a naive way to choose ω_n is to have

Table 8. (Matérn) Estimated mean absolute errors of $\hat{\nu}$ and $\log(\hat{\eta})/\log(\eta)$ with stratified sampling design when $d = 2$ replicated independently 100 times (standard errors within parentheses)

	$\lvert\hat{\nu} - \nu\rvert$			$\lvert\log(\hat{\eta})/\log(\eta) - 1\rvert$		
ν	$n = 40$	$n = 60$	$n = 80$	$n = 40$	$n = 60$	$n = 80$
0.1	0.174 (0.039)	0.165 (0.009)	0.141(0.008)	0.569(0.032)	0.320(0.024)	0.313(0.019)
0.3	0.128 (0.009)	0.097 (0.009)	0.070(0.006)	0.342(0.026)	0.208(0.019)	0.188(0.016)
0.5	0.191 (0.013)	0.142 (0.011)	0.120(0.009)	0.425(0.025)	0.328(0.016)	0.292(0.016)
0.7	0.266 (0.014)	0.194 (0.012)	0.183(0.011)	0.441(0.024)	0.351(0.021)	0.345(0.022)
0.9	0.348 (0.035)	0.216 (0.015)	0.211(0.016)	0.466(0.030)	0.340(0.024)	0.328(0.024)
1.1	0.495 (0.061)	0.316 (0.042)	0.219(0.15)	0.490(0.035)	0.350(0.028)	0.303(0.022)
1.3	0.558 (0.062)	0.464 (0.054)	0.361(0.052)	0.505(0.035)	0.482(0.039)	0.349(0.028)

a prior guess of the true values of the unknown model parameters and use simulations to determine a value of ω_n such that the proposed estimators perform well in a neighborhood of the true parameters. It would be reassuring if the resulting estimates occur in that neighborhood of the true parameters. This is the way we chose the values of ω_n in our simulation study in Section 9.

This section will be continued in Section 22 of the online supplement Loh and Sun (2023). In Section 22.1, a simulation study is conducted to gauge the accuracy of the proposed estimators where the design sites are i.i.d. random vectors each uniformly distributed on $[0,1)^d$, $d \in \{1,2\}$. Section 22.2 of Loh and Sun (2023) presents a simulation study for the maximum likelihood estimators (MLEs) of τ, ν and $\eta = \sigma^2\alpha^{2\nu}$ for a mean-zero isotropic Matérn Gaussian random field when $d = 1$. Qualitatively, the accuracy of the MLEs are similar to those estimators proposed here. In particular, the error for estimating τ is relatively small while the error for estimating ν is also high for the sample size $n = 500$ used in the simulations. Maximum likelihood estimation simulations are very time-consuming and hence we chose $n = 500$ and not any larger n. These simulations further indicate that estimating the parameter $\eta = \sigma^2\alpha^{\nu}$ accurately is challenging for reasonable sample sizes. One possible explanation is that the unknown ν occurs in the exponent of α. Hence any error in estimating ν results in a possible exponential increase in estimation error for η.

On the complexity or running time between the MLEs and the proposed estimators, there is really no comparison. The proposed estimators can be calculated easily and quickly without any high dimensional nonconvex optimization. In contrast, as Section 22.2 indicates, the computation of the MLEs involves maximization of the log-profile likelihood over 3 unknown parameters. The latter involves high dimensional nonconvex optimization where the log-profile likelihood has many local maxima. Thus it is unclear how to locate the global maximum of the log-profile likelihood except to perform a comprehensive search over the whole parameter space.

Acknowledgements

The authors would like to thank Professor Davy Paindaveine, the Associate Editor and two anonymous referees for their constructive comments that improved the quality of this paper. The second author would also like to thank Professors Cheng Li, Michael Stein and Jin-Ting Zhang for their insightful comments on her PhD thesis Sun (2020).

Funding

The first author was supported by AcRF Tier 1 Grant R-155-000-209-114.

Supplementary Material

Supplementary material for "Estimating the parameters of some common Gaussian random fields with nugget under fixed-domain asymptotics" (DOI: 10.3150/22-BEJ1551SUPP; .pdf). The supplementary material Loh and Sun (2023), available online, contains proofs of the results in this article as well as the continuation of the simulation study in Section 9.

References

Andrews, G.E., Askey, R. and Roy, R. (1999). *Special Functions*. *Encyclopedia of Mathematics and Its Applications* **71**. Cambridge: Cambridge Univ. Press. MR1688958 https://doi.org/10.1017/CBO9781107325937

Bevilacqua, M., Faouzi, T., Furrer, R. and Porcu, E. (2019). Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics. *Ann. Statist.* **47** 828–856. MR3909952 https://doi.org/10.1214/17-AOS1652

Chang, C.-H., Huang, H.-C. and Ing, C.-K. (2017). Mixed domain asymptotics for a stochastic process model with time trend and measurement error. *Bernoulli* **23** 159–190. MR3556770 https://doi.org/10.3150/15-BEJ740

Chen, H.-S., Simpson, D.G. and Ying, Z. (2000). Infill asymptotics for a stochastic process model with measurement error. *Statist. Sinica* **10** 141–156. MR1742105

Cressie, N.A.C. (1991). *Statistics for Spatial Data*. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. New York: Wiley. MR1127423

Du, J., Zhang, H. and Mandrekar, V.S. (2009). Fixed-domain asymptotic properties of tapered maximum likelihood estimators. *Ann. Statist.* **37** 3330–3361. MR2549562 https://doi.org/10.1214/08-AOS676

Golub, G.H. and Van Loan, C.F. (1983). *Matrix Computations*. *Johns Hopkins Series in the Mathematical Sciences* **3**. Baltimore, MD: Johns Hopkins Univ. Press. MR0733103

Hanson, D.L. and Wright, F.T. (1971). A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Stat.* **42** 1079–1083. MR0279864 https://doi.org/10.1214/aoms/1177693335

Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis*. Cambridge: Cambridge Univ. Press. MR0832183 https://doi.org/10.1017/CBO9780511810817

Kaufman, C.G. and Shaby, B.A. (2013). The role of the range parameter for estimation and prediction in geostatistics. *Biometrika* **100** 473–484. MR3068447 https://doi.org/10.1093/biomet/ass079

Keshavarz, H., Nguyen, X. and Scott, C. (2019). Local inversion-free estimation of spatial Gaussian processes. *Electron. J. Stat.* **13** 4224–4279. MR4021265 https://doi.org/10.1214/19-EJS1592

Lindgren, F., Rue, H. and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 423–498. MR2853727 https://doi.org/10.1111/j.1467-9868.2011.00777.x

Loh, W.-L. (1996). On Latin hypercube sampling. *Ann. Statist.* **24** 2058–2080. MR1421161 https://doi.org/10.1214/aos/1069362310

Loh, W.-L. (2015). Estimating the smoothness of a Gaussian random field from irregularly spaced data via higher-order quadratic variations. *Ann. Statist.* **43** 2766–2794. MR3405611 https://doi.org/10.1214/15-AOS1365

Loh, W.-L. and Lam, T.-K. (2000). Estimating structured correlation matrices in smooth Gaussian random field models. *Ann. Statist.* **28** 880–904. MR1792792 https://doi.org/10.1214/aos/1015952003

Loh, W.-L. and Sun, S. (2023). Supplement to "Estimating the parameters of some common Gaussian random fields with nugget under fixed-domain asymptotics." https://doi.org/10.3150/22-BEJ1551SUPP

Loh, W.-L., Sun, S. and Wen, J. (2021). On fixed-domain asymptotics, parameter estimation and isotropic Gaussian random fields with Matérn covariance functions. *Ann. Statist.* **49** 3127–3152. MR4352525 https://doi.org/10.1214/21-aos2077

McKay, M.D., Beckman, R.J. and Conover, W.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21** 239–245. MR0533252 https://doi.org/10.2307/1268522

Owen, A.B. (1992). Orthogonal arrays for computer experiments, integration and visualization. *Statist. Sinica* **2** 439–452. MR1187952

Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn, H.P. (1989). Design and analysis of computer experiments. *Statist. Sci.* **4** 409–435. MR1041765

Sadhanala, V., Wang, Y.-X., Hu, A.J. and Tibshirani, R.J. (2021). Multivariate trend filtering for lattice data. Available at arXiv:2112.14758v1.

Santner, T.J., Williams, B.J. and Notz, W.I. (2003). *The Design and Analysis of Computer Experiments*. *Springer Series in Statistics*. New York: Springer. MR2160708 https://doi.org/10.1007/978-1-4757-3799-8

Stein, M. (1987). Large sample properties of simulations using Latin hypercube sampling. *Technometrics* **29** 143–151. MR0887702 https://doi.org/10.2307/1269769

Stein, M.L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. *Springer Series in Statistics*. New York: Springer. MR1697409 https://doi.org/10.1007/978-1-4612-1494-6

Stein, M.L. (2012). Simulation of Gaussian random fields with one derivative. *J. Comput. Graph. Statist.* **21** 155–173. MR2913361 https://doi.org/10.1198/jcgs.2010.10069

Sun, S. (2020). *Parameter Estimation for Isotropic Gaussian Random Fields with the Generalized Wendland Covariance Function or the Matérn Covariance Function with Nugget.* PhD thesis, Department of Statistics and Applied Probability, National University of Singapore.

Tang, W., Zhang, L. and Banerjee, S. (2021). On identifiability and consistency of the nugget in Gaussian spatial process models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **83** 1044–1070. MR4349127 https://doi.org/10.1111/rssb.12472

Wang, D. and Loh, W.-L. (2011). On fixed-domain asymptotics and covariance tapering in Gaussian random field models. *Electron. J. Stat.* **5** 238–269. MR2792553 https://doi.org/10.1214/11-EJS607

Wen, J., Sun, S. and Loh, W.-L. (2021). Smoothness estimation of nonstationary Gaussian random fields from irregularly spaced data observed along a curve. *Electron. J. Stat.* **15** 6071–6150. MR4355704 https://doi.org/10.1214/21-ejs1941

Wood, A.T.A. and Chan, G. (1994). Simulation of stationary Gaussian processes in $[0,1]^d$. *J. Comput. Graph. Statist.* **3** 409–432. MR1323050 https://doi.org/10.2307/1390903

Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Statist. Assoc.* **99** 250–261. MR2054303 https://doi.org/10.1198/016214504000000241

Received February 2022 and revised July 2022