Maximum likelihood estimation and prediction error for a Matérn model on the circle

Sébastien J. Petit*[1, 2]

[1] Université Paris-Saclay, CNRS, CentraleSupélec,
Laboratoire des signaux et systèmes, 91190, Gif-sur-Yvette, France

[2]Laboratoire National de Métrologie et d'Essais, 78197 Trappes Cedex, France

December 13, 2022

Abstract

This work considers Gaussian process interpolation with a periodized version of the Matérn covariance function (Stein, 1999, Section 6.7) with Fourier coefficients $\phi\left(\alpha^2+j^2\right)^{-\nu-1/2}$. Convergence rates are studied for the joint maximum likelihood estimation of ν and ϕ when the data is sampled according to the model. The mean integrated squared error is also analyzed with fixed and estimated parameters, showing that maximum likelihood estimation yields asymptotically the same error as if the ground truth was known. Finally, the case where the observed function is a "deterministic" element of a continuous Sobolev space is also considered, suggesting that bounding assumptions on some parameters can lead to different estimates.

1 Introduction

Gaussian process regression or kriging is a common technique for inferring an unknown function from data, which has applications in geostatistics (Stein, 1999), computer experiments (Santner et al., 2003), and machine learning (Rasmussen and Williams, 2006). Stein (1999) stresses the importance of choosing a covariance function that fits the problem, promotes the use of the Matérn (1986) family of covariance functions, and advocates using maximum likelihood to estimate its parameters.

A distinction is generally made between increasing and fixed-domain asymptotic frameworks for parameter estimation of Gaussian processes (see, e.g., Bachoc, 2021, for a review). While several increasing-domain asymptotic frameworks have been exhaustively studied (see, e.g., Mardia and Marshall, 1984; Bachoc, 2014), fixed-domain frameworks are studied only with simplifications and for a restrained number of parameters to our knowledge (Ying, 1991, 1993; van der Vaart, 1996; Zhang, 2004; Loh, 2006; Anderes, 2010).

*sebastien.petit@lne.fr

Considering fixed-domain asymptotics, the regularity parameter of the Matérn covariance function seems to have been little studied, although Stein (1999) presents numerous results suggesting it as the most impactful from the point of view of prediction error. To study parameter estimation, Stein (1999, Section 6.7) proposes an asymptotic framework[1] with uniformly distributed observations on the torus and makes a conjecture about the asymptotic behavior of maximum likelihood based on the Fisher information matrix. This topic has only recently regained popularity. Indeed, Chen et al. (2021) used the previous framework to show that the estimation of the regularity parameter is consistent if the other parameters remain fixed. Moreover, Karvonen (2022) has recently shown an asymptotic lower bound in the general case of a "nice" bounded domain of $\mathbb{R}^d$, also covering the case of a deterministic function from a continuous Sobolev space. Other results were obtained in similar frameworks (Szabó et al., 2015; Knapik et al., 2016; Karvonen et al., 2020).

This article presents three main contributions. First, we focus on the one-dimensional case of the framework proposed by Stein (1999, Section 6.7) to give an asymptotic normality result for maximum likelihood parameter estimation. Then, we leverage these convergence rates to analyze the expected integrated error, taking constant factors into account and showing that estimating the parameters yields the same error asymptotically as if the ground truth was known. Finally, we investigate the deterministic case by deriving the large sample limit of the likelihood criterion in a particular case. This suggests that bounding assumptions on some parameters can lead to different estimates.

The article is organized as follows. First, Section 2.1 details the asymptotic framework and notations. Then, Section 3 gives the main results. Finally, Section 4 provides a few results on the deterministic case.

2 Gaussian process interpolation on the circle

2.1 Framework

Let $f : [0, 1] \to \mathbb{R}$ be a continuous periodic function observed on a regular grid:

$$\{j/n,\ 0 \le j \le n-1\}.$$

Consider the periodic version of the stationary Matérn covariance function $\left\{k_\theta,\ \theta \in (0,+\infty)^3\right\}$ proposed by Stein (1999, Section 6.7) and defined, for $\theta = (\nu, \phi, \alpha) \in (0,+\infty)^3$, by the uniformly absolutely convergent Fourier serie with coefficients:

$$\underline{c}_j(\theta) = \frac{\phi}{(\alpha^2 + j^2)^{\nu+1/2}},\ j \in \mathbb{Z}. \tag{1}$$

The parameter α is not identifiable as different values yield equivalent probability measures. However, ν and ϕ are identifiable (see, e.g., Stein, 1999, Chapter 4 and Section 6.7). Stein (1999) advocates the *regularity* parameter ν as the key quantity governing the asymptotics of the kriging error. The amplitude parameter ϕ does not impact the kriging predictor but is known to be critical for uncertainty quantification (see, e.g., Stein, 1993b) and to be consistently estimated by maximum likelihood if ν and α are known (see, e.g., Zhang, 2004, who use a different parametrization).

The usual task in kriging is to infer the function f from the data

$$Z = (f(0), \ldots, f(1-1/n))^\mathsf{T}. \tag{2}$$

[1] This framework is similar to that proposed by Stein (1993a) in a slightly different setting, where a different conjecture is made about the error on noisy training data.

2.2 Best linear prediction

The function f is usually predicted using the posterior mean function given by the kriging equations (Matheron, 1971). This predictor can be written simply in the framework presented above.

Proposition 1. *Let $n \geq 1$ and $f\colon [0, 1] \to \mathbb{R}$ a continuous periodic function. Suppose also that f is the pointwise absolute limit of its Fourier serie. Writing $\widehat{f}_n$ for the posterior mean function given Z and the parameter $\theta \in \Theta$, we have:*

$$\widehat{f}_n(x) = \sum_{j\in\mathbb{Z}} \left(\frac{\sum_{j_1 \in j+n\mathbb{Z}} c_{j_1}(f)}{\sum_{j_1 \in j+n\mathbb{Z}} \underline{c}_{j_1}(\theta)} \right) \underline{c}_j(\theta) e^{2\pi i x j}, \ x \in [0, 1], \tag{3}$$

where the $c_j(f)$s are the Fourier coefficients of f.

The expression (3) shows how the posterior mean function approximates f: it transforms the Fourier coefficients of k_θ into those of f using the ratio of the discrete Fourier transforms. Finally, we also define the integrated squared error:

$$\mathrm{ISE}_n(\nu, \alpha; f) = \int_0^1 \left(f - \widehat{f}_n \right)^2. \tag{4}$$

Note that it does not depend on ϕ.

2.3 Maximum likelihood estimation

Given the observations Z and $\Theta \subset (0, +\infty)^3$, the maximum likelihood estimate is defined by $\widehat{\theta}_n = \left(\widehat{\nu}_n, \widehat{\phi}_n, \widehat{\alpha}_n \right)$ minimizing (a linear transform of) the negative log-likelihood:

$$\mathbb{L}_n(\theta) = n^{-1} \left(\ln(\det(K_\theta)) + Z^\mathsf{T} K_\theta^{-1} Z \right), \quad \theta = (\nu, \phi, \alpha) \in \Theta, \tag{5}$$

with ties broken arbitrarily and K_θ the covariance matrix of Z under θ.

The parameters ν and α will be assumed to be bounded in this work, i.e., we take $\Theta = N \times (0, +\infty) \times A$ with N and A bounded away from zero and infinity. Most of our results will be stated with A being a singleton, with α enforced to a value (which will not necessarily be the ground truth). This type of assumption is more or less standard in the field, as it leads to simplifications (see, e.g., Ying, 1991; Loh, 2006; Chen et al., 2021). However, keeping ϕ unbounded is key to our results and for discussing the deterministic case in Section 4.

2.4 A symmetrized version of the Hurwitz zeta function

As Stein (1999, Section 6.7) points out, a great simplification emerges when we restrict ourselves to the framework presented in Section 2.1. This will allow us to derive a sharp analysis, taking into account constant factors in the evaluation of the prediction accuracy with estimated parameters. Similarly to Stein (1999), the function

$$\gamma\colon \ (\alpha; x) \in (1, +\infty) \times (0, 1) \mapsto \sum_{j\in\mathbb{Z}} \frac{1}{|j + x|^\alpha},$$

will play a major role in deriving the limiting constants. It is smooth and related to the Hurwitz zeta function ζ_H by:

$$\gamma(\alpha; x) = \zeta_H(\alpha; x) + \zeta_H(\alpha; 1 - x), \quad x \in (0, 1). \tag{6}$$

Moreover, the function $\gamma(\alpha;\cdot)$ is symmetric with respect to $1/2$ for $\alpha>1$.

Integrals involving $\gamma(\alpha;\cdot)$, for some $\alpha>0$, will appear in the main body of the article. All of them are well-defined, and integrability statements are postponed to the proofs of Appendix A.

3 Main results

3.1 Standing assumptions

Consider the framework presented in Section 2.1 and suppose that the function is sampled according to the (real-valued) centered Gaussian process:

$$\xi : x \in [0,\, 1] \mapsto \frac{1}{\sqrt{2}} \sum_{j\in\mathbb{Z}} \sqrt{c_j(\theta_0)}\left(U_{1,|j|} + iU_{2,|j|}\operatorname{sign}(j)\right) e^{2\pi i x j}, \tag{7}$$

with $\theta_0=(\nu_0,\phi_0,\alpha_0)\in(0,+\infty)^3$ and $(U_{q,j})_{q\in\{1,2\},j\geq 0}$ independent random variables such that $U_{2,0}=0$, $U_{1,0}\sim\mathcal{N}(0,2)$, and $U_{q,j}\sim\mathcal{N}(0,1)$ for $q\in\{1,2\}$ and $j\geq 1$. The convergence of the expansion (7) is meant pointwise both in $L^2(P)$ and almost surely. It is easy to check that $\xi\sim\mathrm{GP}\left(0,k_{\theta_0}\right)$.

Let $\widehat{\theta}_n=\left(\widehat{\nu}_n,\widehat{\phi}_n,\widehat{\alpha}_n\right)$ be a maximum likelihood estimate defined in Section 2.3 for some $\Theta=N\times(0,+\infty)\times A$ with $A\subset(0,+\infty)$ and $N=[a_\nu,b_\nu]\ni\nu_0$. The following sections give convergence rates in terms of parameter estimation and predictions.

3.2 Convergence rate of maximum likelihood estimation

The following result states the consistency for the identifiable parameters and an upper bound of the corresponding rates.

Theorem 2. *Let $0<\beta<1/4$ and $A=[a_\alpha,b_\alpha]$ be bounded away from zero and infinity. The bounds $\widehat{\nu}_n-\nu_0=o_P\left(n^{-\beta}\right)$ and $\widehat{\phi}_n-\phi_0=o_P\left(\ln(n)n^{-\beta}\right)$ hold in probability.*

Informally, Theorem 2 is proven by showing that a linear transformation of

$$\inf_{\phi>0}\mathbb{L}_n(\nu,\,\phi,\,\alpha)$$

converges to

$$\ln\left(\gamma(2\nu+1;\cdot)\right)+\ln\left(\int_0^1\frac{\gamma(2\nu_0+1;\cdot)}{\gamma(2\nu+1;\cdot)}\right), \tag{8}$$

for $\nu>\nu_0-1/2$. Jensen inequality then shows that (8) is maximized by taking $\nu=\nu_0$.

Furthermore, similarly to Stein (1999, Section 6.7), let us define

$$\psi_\nu : x\in(0,\,1)\mapsto\frac{\sum_{j\in\mathbb{Z}}|x+j|^{-2\nu-1}\ln|x+j|}{\sum_{j\in\mathbb{Z}}|x+j|^{-2\nu-1}},\quad\text{for }\nu>0,$$

which is square integrable on $(0,\,1)$ and $e(\nu)=\mathsf{E}\left(\psi_\nu(U)\right)$ and $V(\nu)=\operatorname{Var}\left(\psi_\nu(U)\right)$, with $U\sim\mathcal{U}(0,\,1)$. The following result proves the conjecture made by Stein (1999, Section 6.7) when $\widehat{\nu}_n$ is bounded and α_0 is known.

Theorem 3. *Suppose that $A=\{\alpha_0\}$. Then*

$$\sqrt{2n}\begin{pmatrix}\frac{\widehat{\phi}_n-\phi_0}{2\phi_0}-\left(\ln(n)+e(\nu_0)\right)\left(\widehat{\nu}_n-\nu_0\right)\\ \sqrt{V(\nu_0)}\left(\widehat{\nu}_n-\nu_0\right)\end{pmatrix}\rightsquigarrow\mathcal{N}\left(0,\,I_2\right). \tag{9}$$

3.3 Convergence rate of the integrated prediction error

This section states our results about the expectation of (4) with estimated parameters. To avoid technical difficulties, we suppose that $\nu_0 > 1/2$ in this section. We begin with the case of fixed parameters.

For $\nu, \nu_0 > 0$, define

$$\vartheta_{\nu;\nu_0} : x \in (0,\, 1) \mapsto \frac{\gamma(4\nu+2;x)\,\gamma(2\nu_0+1;x)}{\gamma^2\,(2\nu+1;x)} + \gamma(2\nu_0+1;x) - 2\frac{\gamma(2\nu+2\nu_0+2;x)}{\gamma(2\nu+1;x)}.$$

The function $\vartheta_{\nu;\nu_0}$ is smooth on $(0,\,1)$, symmetric with respect to $1/2$, and a Taylor expansion around $x = 0$ shows that it is integrable if $\nu > (\nu_0 - 1)/2$.

The following result states the asymptotics of the prediction error with fixed parameters.

Theorem 4. *Let $\theta = (\nu,\, \phi,\, \alpha) \in N \times \mathbb{R} \times A$. Then,*

$$\mathsf{E}\left(\mathrm{ISE}_n\left(\nu,\, \alpha;\, \xi\right)\right) \lesssim \frac{1}{n^{4\nu+2}},\ \text{for } \nu < (\nu_0 - 1)/2 \tag{10}$$

$$\mathsf{E}\left(\mathrm{ISE}_n\left(\nu,\, \alpha;\, \xi\right)\right) \lesssim \frac{\ln(n)}{n^{2\nu_0}},\ \text{for } \nu = (\nu_0 - 1)/2 \tag{11}$$

and

$$\mathsf{E}\left(\mathrm{ISE}_n\left(\nu,\, \alpha;\, \xi\right)\right) \sim \frac{\phi_0 \int_0^1 \vartheta_{\nu;\nu_0}}{n^{2\nu_0}},\ \text{otherwise.} \tag{12}$$

The symbol $\lesssim$ denotes an inequality up to a universal constant.

This result shows that half of the smoothness is sufficient for optimal convergence rates. However, the constant $\int_0^1 \vartheta_{\nu;\nu_0}$ is minimized by taking $\nu = \nu_0$. This is in line with the result of Stein (1999, Theorem 3) obtained in a different framework.

Then, our last result gives the asymptotic behavior of the prediction error with estimated ν and ϕ and fixed—but not necessarily known—α.

Theorem 5. *Let $A = \{\alpha\}$ with $0 < \alpha < +\infty$. Then,*

$$\mathsf{E}\left(\mathrm{ISE}_n\left(\widehat{\nu}_n,\, \alpha;\, \xi\right)\right) \sim \frac{\phi_0 \int_0^1 \vartheta_{\nu_0;\nu_0}}{n^{2\nu_0}}. \tag{13}$$

This last result shows that estimating the parameters is optimal both in terms of the rate and the constant.

4 The deterministic case

This section studies the case of an unknown fixed deterministic function from a continuous Sobolev space. Let $\beta > 0$ and define the Sobolev space

$$H^{\beta+1/2}\left[0,\,1\right] = \left\{ g \in L^2\left[0,\,1\right],\ \|g\|^2_{H^{\beta+1/2}[0,\,1]} = \sum_{j\in\mathbb{Z}} (1+j^2)^{\beta+1/2} |c_j(f)|^2 < +\infty \right\} \tag{14}$$

of (continuous) periodic functions.

Furthermore, let $f \in H^{\beta+1/2}[0,1]$ for some $\beta > 0$ and define the smoothness

$$\nu_0(f) = \inf\left\{\beta > 0,\ f \notin H^{\beta+1/2}[0,1]\right\}$$

of f as Wang and Jing (2022) and Karvonen (2022). We will assume the existence of some $\beta \in (0,+\infty)$ such that $f \notin H^{\beta+1/2}[0,1]$ implying thus that $\nu_0(f)$ is finite.

Suppose that $\widehat{\theta}_n = \left(\widehat{\nu}_n, \widehat{\phi}_n, \widehat{\alpha}_n\right)$ is estimated by maximizing the likelihood as in Section 2.3 for $\Theta = N \times F \times A$ with N and A bounded away from zero and infinity. This section discusses the behavior of $\widehat{\nu}_n$ under three assumptions on F: 1) a singleton; 2) a range bounded away from zero and infinity; and 3) the whole $(0,+\infty)$. For the last case, the definition

$$\mathbb{M}_n^f\colon\ (\nu,\alpha) \in N \times A \mapsto \inf_{\phi>0} \mathbb{L}_n(\nu,\phi,\alpha) + (2\nu_0(f)+1)\ln(n) - 1,$$

will be used.

On "nice" bounded regions of $\mathbb{R}^d$, Karvonen (2022) shows that $\liminf \widehat{\nu}_n \geq \nu_0(f)$ if α and ϕ are fixed[2]. The following result shows that it holds on the circle no matter the assumption on F.

Proposition 6. *If N and A are bounded away from zero and infinity, then*

$$\liminf \widehat{\nu}_n \geq \nu_0(f)$$

holds for the three previous assumptions on F.

Regarding the precise behavior of $\widehat{\nu}_n$ above $\nu_0(f)$, Karvonen (2022) conjectures that $\widehat{\nu}_n$ converges to $\nu_0(f)+1/2$ if ϕ and α are fixed but deems that a joint estimation may give a different behavior. The rest of this section is devoted to supporting this idea.

Write $\approx$ for a two-way inequality up to universal constants and use the more stringent assumption that

$$\left|c_j(f)\right| \approx |j|^{-\nu_0(f)-1}$$

for simplicity. The following result is a minor adaptation of the reasoning used by Chen et al. (2021) and Karvonen (2022), showing that the conjecture is verified more generally if ϕ is bounded. The proof is omitted for brevity.

Proposition 7. *Suppose that $\left|c_j(f)\right| \approx |j|^{-\nu_0(f)-1}$ and that N, A, and F are bounded away from zero and infinity. Then, the convergence $\widehat{\nu}_n \to \nu_0(f)+1/2$ holds.*

However, our last results suggest that it does not hold when $F = (0,+\infty)$.

Proposition 8. *Suppose that $c_j(f) = |j|^{-\nu_0(f)-1}$, for $j \neq 0$. Then, for $\nu > \nu_0(f)$ and $\alpha > 0$, we have*

$$\mathbb{M}_n^f(\nu,\alpha) \to \mathbb{M}_\infty^f(\nu) = \int_0^1 \ln(\gamma(2\nu+1;\cdot)) + \ln\left(\int_0^1 \frac{\gamma^2(\nu_0(f)+1;\cdot)}{\gamma(2\nu+1;\cdot)}\right).$$

It is possible to obtain a uniform version of the previous convergence. However, it is omitted for brevity and because we could not identify the minimizer(s) of the limit analytically. Figure 1 shows a numerical approximation of $\mathbb{M}_\infty^f$.

[2]Note that Karvonen (2022) does not assume that ν is bounded.

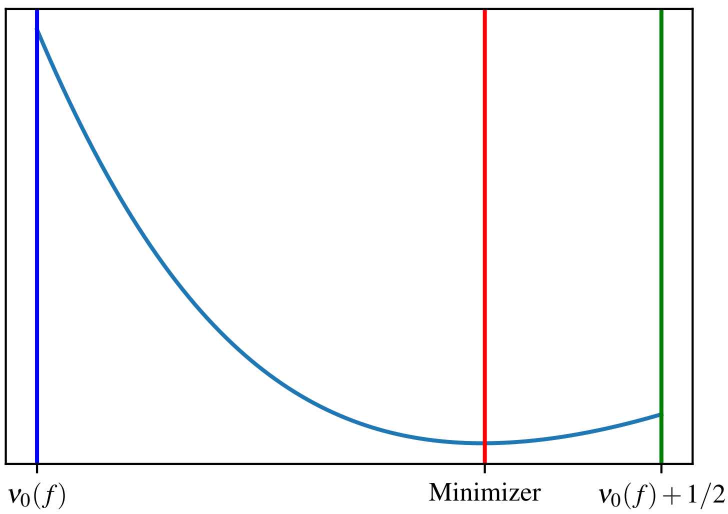

Figure 1: The function $\mathbb{M}_{\infty}^{f}$, for $v_0(f) = 1$. A numerical approximation of the minimizer is about 1.359. Note that $\mathbb{M}_{\infty}^{f}$ is approximated numerically using finite sums for γ and discretizations for the integrals.

A Proofs

A.1 Additional notations

The symbol $\lesssim$ denotes an inequality up to a universal constant. For compactness, the symbol $\approx$ is used when the two-way inequality $\lesssim$ holds.

Write $K_\theta = \phi R_{\nu,\alpha}$ and $\underline{c}_j(\theta) = \phi c_j(\nu,\alpha)$, for $\theta = (\nu, \phi, \alpha) \in (0, +\infty)^3$ and $j \in \mathbb{Z}$. All results suppose that $\Theta = N \times (0, +\infty) \times A$ with $N = [a_\nu, b_\nu]$, $A = [a_\alpha, b_\alpha]$, $0 < a_\nu \leq b_\nu < +\infty$, and $0 < a_\alpha \leq b_\alpha < +\infty$ unless explicitly stated otherwise. Without loss of generality, suppose that $a_\nu < \nu_0 - 1/2$ and define $N_\varepsilon = [\nu_0 - 1/2 + \varepsilon, b_\nu] \subset N$ for $\varepsilon > 0$.

Furthermore, for $\theta_0 = (\nu_0, \phi_0, \alpha_0)$, define $\mathbb{M}_n$ to be the stochastic process:

$$\mathbb{M}_n \colon\ (\nu, \alpha) \in N \times A \mapsto \inf_{\phi > 0} \mathbb{L}_n(\nu, \phi, \alpha) + 2\nu_0 \ln(n) - 1.$$

Its expression is given almost surely by the following Lemma.

Lemma 9. *(see, e.g., Santner et al., 2003, Section 3.3.2) Let $\nu, \alpha > 0$. If $Z \neq 0$, then*

$$\inf_{\phi>0} \mathbb{L}_n(\nu, \phi, \alpha) = 1 + n^{-1} \ln(\det(R_{\nu,\alpha})) + \ln\left(\frac{Z^\mathsf{T} R_{\nu,\alpha}^{-1} Z}{n}\right), \tag{15}$$

where the supremum is reached by $\widehat{\phi}_n = Z^\mathsf{T} R_{\nu,\alpha}^{-1} Z / n$.

Note that $\mathbb{M}_n$ has therefore almost surely continuous sample paths. The proofs mostly consist in studying $\mathbb{M}_n$, which is often called the profiled likelihood. For a compact $A \subset (0, +\infty)$, define now $\mathbb{U}_n \colon \nu \in N \mapsto \inf_{\alpha \in A} \mathbb{M}_n(\nu, \alpha)$. The object $\mathbb{U}_n$ is still a collection of measurable maps and is also almost surely continuous since A is compact.

Also, define

$$g_\nu = \ln\left(\gamma(2\nu + 1; \cdot)\right),$$

for $\nu > 0$ and

$$h_{\nu;\nu_0} = \frac{\gamma(2\nu_0 + 1; \cdot)}{\gamma(2\nu + 1; \cdot)},$$

for $\nu > \nu_0 - 1/2$. These functions are smooth and integrable.

Finally, suppose $n \geq 3$ and use the notation $l = \lfloor (n-1)/2 \rfloor$ throughout the following.

A.2 Circulant matrices and useful facts

Let us remind how the framework from Section 2.1 yields simplifications (see, e.g., Craven and Wahba, 1979).

First, let W be the $n \times n$ matrix with entries $W_{j,m} = \sqrt{n}^{-1} e^{2\pi i j m/n}$, for $0 \leq j, m \leq n-1$. For every $\theta = (\nu, \phi, \alpha) \in (0, +\infty)^3$, the matrix $R_{\nu,\alpha}$ is circulant and thus $R_{\nu,\alpha} = W \Delta_{\nu,\alpha} W^*$ (see, e.g., Brockwell and Davis, 1987, p. 130) with

$$\Delta_{\nu,\alpha} = \operatorname{diag}\left(\lambda_{0,n}, \ldots, \lambda_{n-1,n}\right)$$

and

$$\lambda_{m,n} = \sum_{j=0}^{n-1} e^{-2\pi i j m/n} k_{\nu,1,\alpha}(j/n) = n \sum_{j \in \mathbb{Z}} c_{m+nj}(\nu, \alpha),\ 0 \leq m \leq n-1. \tag{16}$$

Note that $\lambda_{m,n}$ depends on ν and α but the symbols are dropped to avoid cumbersome expressions. These coefficients verify

$$\lambda_{m,n} = \lambda_{n-m,n}, \text{ for } 0 \leq m \leq n-1. \tag{17}$$

Furthermore, the matrices $R_{\nu,\alpha}$ are symmetric so one also has $R_{\nu,\alpha} = P\Delta P^{\mathsf{T}}$ for a unitary matrix P, obtained, e.g., by taking the real and imaginary parts of W. For $\theta_0 = (\nu_0, \phi_0, \alpha_0)$, write

$$P^{\mathsf{T}} Z = \sqrt{\phi_0}\left(\lambda^0_{0,n} U_{0,n}, \cdots, \lambda^0_{n-1,n} U_{n-1,n}\right),$$

with $\lambda^0_{0,n}, \ldots, \lambda^0_{n-1,n}$ the eigenvalues of R_{ν_0,α_0} and $U_{0,n}, \cdots, U_{n-1,n}$ drawn independently from a standard Gaussian. We have

$$Z^{\mathsf{T}} R^{-1}_{\nu,\alpha} Z = \phi_0 \sum_{m=0}^{n-1} \frac{U^2_{m,n} \lambda^0_{m,n}}{\lambda_{m,n}}. \tag{18}$$

Finally, the following rough approximation will sometimes be used.

Lemma 10. *The quantities $\lambda_{0,n}/n$ and $\lambda_{m,n} m^{2\nu+1}/n$ are bounded away from zero and infinity uniformly in $\nu \in N$, $\alpha \in A$, n and $1 \leq m \leq \lfloor n/2 \rfloor$.*

Proof. Let $0 \leq m \leq \lfloor n/2 \rfloor$, we have using (16)

$$c_m(\nu,\alpha) \leq \lambda_{m,n}/n \leq 2c_m(\nu,\alpha) + 2\sum_{j=1}^{+\infty} c_{m+nj}(\nu,\alpha).$$

Moreover

$$\sum_{j=1}^{+\infty} c_{m+nj}(\nu,\alpha)/c_m(\nu,\alpha) \leq \sum_{j=1}^{+\infty} (b^2_\alpha + 1/4)^{\nu+1/2}/j^{2\nu+1} \leq C_1^{b_\alpha,a_\nu,b_\nu} \tag{19}$$

where $C_1^{b_\alpha,a_\nu,b_\nu} = \max(1, (b^2_\alpha + 1/4)^{b_\nu+1/2})\zeta(2a_\nu + 1)$. The result follows from elementary operations. □

Nevertheless, our results will require refined approximations.

A.3 Proofs of Section 2.2

Proof of Proposition 1. The kriging equations yield $\widehat{f}_n(x) = k^{\mathsf{T}}_{\theta,x} K^{-1}_\theta Z$, with $k_{\theta,x} = (k_\theta(m/n - x))_{0\leq m\leq n-1}$, for $x \in [0, 1]$. The simplifications described in Section A.2 give:

$$W^* Z = \sqrt{n}\left(\sum_{j\in m+n\mathbb{Z}} c_j(f)\right)_{0\leq m\leq n-1},$$

and

$$W^* k_{\theta,x} = \sqrt{n}\left(\sum_{j\in m+n\mathbb{Z}} \underline{c}_j(\theta) e^{-2\pi i x j}\right)_{0\leq m\leq n-1}.$$

The result follows from elementary manipulations. □

A.4 Proof of Theorem 2

A.4.1 Proof of the theorem

Proof of Theorem 2. Let $0 < \beta < 1/4$, Lemma 9 gives

$$\ln\left(\widehat{\phi}_n\right) = \ln(\phi_0) + \ln\left(\frac{Z^\top R^{-1}_{\widehat{\nu}_n,\widehat{\alpha}_n} Z}{n^{1+2(\widehat{\nu}_n - \nu_0)}}\right) + 2(\widehat{\nu}_n - \nu_0)\ln(n).$$

So

$$\frac{n^\beta}{\ln(n)}\left(\ln\left(\widehat{\phi}_n\right) - \ln(\phi_0)\right) = \frac{n^\beta}{\ln(n)}\ln\left(\frac{Z^\top K^{-1}_{\widehat{\nu}_n,\widehat{\alpha}_n} Z}{n^{1+2(\widehat{\nu}_n - \nu_0)}}\right) + 2n^\beta(\widehat{\nu}_n - \nu_0)$$

$$= \frac{n^\beta}{\ln(n)}\ln\left(\int_0^1 h_{\widehat{\nu}_n;\nu_0}\right) + \frac{n^\beta}{\ln(n)}\left(\ln\left(\frac{Z^\top K^{-1}_{\widehat{\nu}_n,\widehat{\alpha}_n} Z}{n^{1+2(\widehat{\nu}_n - \nu_0)}}\right) - \ln\left(\int_0^1 h_{\widehat{\nu}_n;\nu_0}\right)\right) + 2n^\beta(\widehat{\nu}_n - \nu_0).$$

The latter converges to zero in probability thanks to Lemma 15, Slutsky's Lemma, Lemma 12, and to the univariate Delta method, observing that the mapping $\nu \mapsto \ln\left(\int_0^1 h_{\nu;\nu_0}\right)$ is C^1 in a neighborhood of ν_0. □

Lemma 11. *The convergence $\widehat{\nu}_0 \to \nu_0$ holds in probability.*

Proof. First for $\nu \in N$ and $\alpha \in A$, we have

$$\mathbb{M}_n(\nu,\alpha) = \ln(\det(R_{\nu,\alpha}))/n + \ln\left(\frac{Z^\top R^{-1}_{\nu,\alpha} Z}{n^{1-2\nu_0}}\right)$$

$$= \int_0^1 g_\nu(x)dx + \mathcal{O}(\ln(n)/n) + \ln\left(\frac{Z^\top R^{-1}_{\nu,\alpha} Z}{n^{1+2(\nu-\nu_0)}}\right)$$

$$= \mathcal{O}(1) + \ln\left(\frac{Z^\top R^{-1}_{\nu,\alpha} Z}{n^{1+2(\nu-\nu_0)}}\right)$$

uniformly thanks to Lemma 14 and by observing that the mapping $\nu \in [a_\nu, b_\nu] \mapsto \int_0^1 g_\nu$ is continuous. Now let $\nu \in [a_\nu,\ \nu_0 - 1/2 + \varepsilon]$ and $\alpha \in A$, Lemma 10 yields

$$\frac{Z^\top R^{-1}_{\nu,\alpha} Z}{n^{1+2(\nu-\nu_0)}} \gtrsim \frac{1}{n}\sum_{m=1}^{l} U^2_{m,n}\left(\frac{m}{n}\right)^{2(\nu-\nu_0)} \gtrsim \frac{1}{n}\sum_{m=1}^{l} U^2_{m,n}\left(\frac{m}{n}\right)^{-1+2\varepsilon}$$

$$= \frac{1}{n^{2\varepsilon}}\sum_{m=1}^{l} U^2_{m,n} m^{-1+2\varepsilon} \to \frac{1}{2^{1+2\varepsilon}\varepsilon}$$

in probability using a similar argument as the one from Lemma 25. Considering $\mathbb{U}_n$ from Section A.1, Lemma 15 gives $\mathbb{U}_n(\nu_0) \to \int_0^1 g_{\nu_0}$ in probability, so we have

$$\inf_{\nu\in[a_\nu,\ \nu_0-1/2+\varepsilon]} \mathbb{U}_n(\nu) - \mathbb{U}_n(\nu_0)$$

$$\geq C + \inf_{\nu\in[a_\nu,\ \nu_0-1/2+\varepsilon],\alpha\in A} \ln\left(\frac{Z^\top R^{-1}_{\nu,\alpha} Z}{n^{1+2(\nu-\nu_0)}}\right) - \mathbb{U}_n(\nu_0) \to C - \ln\left(2^{1+2\varepsilon}\varepsilon\right) - \int_0^1 g_{\nu_0},$$

with the later convergence holding in probability, for a universal constant C. Letting $\varepsilon \to 0$ shows that the above limit can be made arbitrarily high.

Finally let $0 < \varepsilon < 1/2$, we have

$$\sup_{\nu \in N_\varepsilon} \left| \mathbb{U}_n(\nu) - \int_0^1 g_\nu(x)dx - \ln\left(\int_0^1 h_{\nu;\nu_0}\right)\right|$$

$$\leq \sup_{\nu \in N_\varepsilon, \alpha \in A} \left| \mathbb{M}_n(\nu, \alpha) - \int_0^1 g_\nu(x)dx - \ln\left(\int_0^1 h_{\nu;\nu_0}\right)\right| = o_P(1),$$

thanks to Lemma 14, Lemma 17 and the continuous mapping theorem applied from the space of functions that are bounded away from zero and infinity to the space of bounded functions. Moreover the function

$$\mathbb{U} : \nu \in N_\varepsilon \mapsto \int_0^1 g_\nu(x)dx + \ln\left(\int_0^1 h_{\nu;\nu_0}\right)$$

is continuous, and strictly maximized by taking $\nu = \nu_0$ thanks to Jensen inequality and the fact that $h_{\nu;\nu_0}$ is constant only if $\nu = \nu_0$. □

Lemma 12. *Let $0 < \beta < 1/4$. The bound $\widehat{\nu}_n - \nu_0 = o_P\left(n^{-\beta}\right)$ holds in probability.*

Proof. Let $2/5 < \varepsilon < 1/2$, $0 < \beta < 1/2$ and let us reuse the function $\mathbb{U}$ from the proof of Lemma 11. We have

$$\mathbb{M}_n(\nu, \alpha) - \mathbb{U}(\nu) = \mathscr{O}\left(\frac{\ln(n)}{n}\right) + \ln\left(\frac{Z^\top R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}}\right) - \ln\left(\int_0^1 h_{\nu;\nu_0}\right), \ (\nu,\ \alpha) \in N_\varepsilon \times A.$$

So this yields

$$n^\beta \sup_{\nu \in N_\varepsilon} |\mathbb{U}_n(\nu) - \mathbb{U}(\nu)| \leq n^\beta \sup_{\nu \in N_\varepsilon,\ \alpha \in A} |\mathbb{M}_n(\nu, \alpha) - \mathbb{U}(\nu)|,$$

which converges in probability to zero thanks to Lemma 15 and Slutsky's Lemma. Moreover, $\mathbb{U}$ is C^3 and some calculations show that

$$\mathbb{U}''(\nu_0) = \int_0^1 \left(\frac{\tilde{\gamma}_{2\nu_0+1}}{\gamma_{2\nu_0+1}}\right)^2 - \left(\int_0^1 \frac{\tilde{\gamma}_{2\nu_0+1}}{\gamma_{2\nu_0+1}}\right)^2 > 0,$$

with $\tilde{\gamma}$ the derivative of $(\nu, x) \mapsto \gamma_{2\nu+1}(x)$ with respect to ν. Then, Lemma 11 and the fact that the maximum of $\mathbb{U}$ on the compact N_ε is unique give the rate $n^{-\beta/2}$ thanks to a standard Taylor expansion around ν_0. □

A.4.2 Approximating $\ln(\det(R_{\nu,\alpha}))$

Lemma 13. *Let $\nu \in N$, $\alpha \in A$, and $1 \leq m \leq l$, we have*

$$c_{m+nj}(\nu, \alpha) = \frac{(1 + u_{n,m,j}(\nu, \alpha))}{|jn + m|^{2\nu+1}}, \tag{20}$$

with $-1 < v_m \leq u_{n,m,j}(\nu, \alpha) \leq 0$ and $v_m = \mathscr{O}(m^{-2})$.

Proof. Using (16), we have

$$c_{m+nj}(\nu,\alpha) = \frac{1}{(\alpha^2+(jn+m)^2)^{\nu+1/2}}$$
$$= \frac{(1+u_{n,m,j}(\nu,\alpha))}{|jn+m|^{2\nu+1}},$$

with $u_{n,m,j}(\nu,\alpha) = \left(1+(\alpha/(jn+m))^2\right)^{-\nu-1/2} - 1$. Elementary operations show that

$$0 \geq u_{n,m,j}(\theta_1) \geq \left(\left(\frac{b_\alpha}{m}\right)^2+1\right)^{-b_\nu-1/2} - 1,$$

which gives the desired result thanks to the Taylor inequality. □

Lemma 14. *Uniformly in $\nu \in N$ and $\alpha \in A$, we have*

$$\ln(\det(R_{\nu,\alpha})) = -2\nu n\ln(n) + n\int_0^1 g_\nu(x)dx + \mathcal{O}(\ln(n)).$$

Proof. Using (16) and Lemma 13, we have

$$\lambda_{m,n}/n = \sum_{j\in\mathbb{Z}} c_{m+nj}(\nu,\alpha)$$
$$= \sum_{j\in\mathbb{Z}} \frac{(1+u_{n,m,j}(\nu,\alpha))}{|jn+m|^{2\nu+1}},$$

and therefore

$$\sum_{m=1}^{l} \ln(\lambda_{m,n}/n) = -(2\nu+1)l\ln(n) + a_n + \sum_{m=1}^{l} g_\nu(m/n),$$

with

$$|a_n| \leq \left|\sum_{m=1}^{l} \ln(1+v_m)\right|$$

which is thus $\mathcal{O}(1)$ uniformly in $\nu \in N$ and $\alpha \in A$.

One has $g_\nu(x) = g_\nu(1-x)$ for $x \in (0,1)$ and it is easy to see that g_ν is non-increasing on $(0,1/2]$, so

$$\int_{1/n}^{(l+1)/n} g_\nu(x)dx \leq \frac{1}{n}\sum_{m=1}^{l} g_\nu(m/n) \leq \int_0^{l/n} g_\nu(x)dx.$$

First we have $\left|\int_{l/n}^{1/2} g_\nu\right| \leq |g_\nu(1/4)|/(2n) = \mathcal{O}(n^{-1})$ uniformly in $\nu \in N$ thanks to a compacity argument and similarly for $\left|\int_{(l+1)/n}^{1/2} g_\nu\right|$.

Moreover observe that

$$\gamma(2\nu+1;x) = \frac{1}{x^{2\nu+1}} + \frac{1}{(1-x)^{2\nu+1}} + \mathcal{O}(1) \tag{21}$$

uniformly in $\nu \in N$ and $x \in (0,1)$. Thus the equivalence $\gamma_{2\nu+1}(x) \sim_{x\to 0} 1/x^{2\nu+1}$ is uniform in $\nu \in N$ and so is $g_\nu(x) \sim_{x\to 0} -(2\nu+1)\ln(x)$. Therefore, we have

$$\int_0^{1/n} g_\nu(x)dx = \mathcal{O}(\ln(n)/n),$$

uniformly in $\nu \in N$ and

$$\sum_{m=1}^{l} g_\nu(m/n) = n\int_0^{1/2} g_\nu(x)dx + \mathcal{O}(\ln(n)).$$

Moreover, Lemma 10 shows that $\ln(\lambda_{0,n}/n) = \mathcal{O}(1)$ and $\ln\left(\lambda_{n/2,n}/n\right) = \mathcal{O}(\ln(n))$ uniformly for n even. One can then conclude using (17). □

A.4.3 Approximating $Z^\top R_{\nu,\alpha} Z$

Lemma 15. *Let $2/5 < \varepsilon < 1/2$. Then, the sequence*

$$(\nu,\alpha) \in N_\varepsilon \times A \mapsto \sqrt{n}\left(\mathbb{M}_n(\nu,\alpha) - \int_0^1 g_\nu - \ln\left(\int_0^1 h_{\nu;\nu_0}\right)\right)$$

of processes converges weakly to

$$\mathrm{GP}\left(0, (\nu_1,\nu_2) \mapsto \frac{2\int_0^1 h_{\nu_1;\nu_0} h_{\nu_2;\nu_0}}{\int_0^1 h_{\nu_1;\nu_0}\int_0^1 h_{\nu_2;\nu_0}}\right)$$

in $L^\infty(N_\varepsilon \times A)$.

Proof. Let $\mathbb{E}$ be the set of continuous real-valued functions on the compact $N_\varepsilon \times A$ and $\mathbb{D}_\psi \subset \mathbb{E}$ the subset of positive functions bounded away from zero, both endowed with the supremum norm. One has

$$(\nu,\alpha) \in N_\varepsilon \times A \mapsto \frac{Z^\top R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}\int_0^1 h_{\nu;\nu_0}} \in \mathbb{D}_\psi,$$

surely. Furthermore, the mapping $\psi : g \in \mathbb{D}_\psi \subset \mathbb{E} \mapsto \ln \circ g \in \mathbb{E}$ is Fréchet-differentiable at the unit function and $\psi'(1) : h \in \mathbb{E} \to h$. Then, Theorem 3.9.4 from Van Der Vaart and Wellner (1996) and Lemma 16 show that

$$\sup_{(\nu,\alpha)\in N_\varepsilon\times A}\left|\sqrt{n}\ln\left(\frac{Z^\top R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}\int_0^1 h_{\nu;\nu_0}}\right) - \sqrt{n}\left(\frac{Z^\top R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}\int_0^1 h_{\nu;\nu_0}} - 1\right)\right|$$

is $o_P(1)$. Conclude with Lemma 9, Lemma 14, and Slutsky's Lemma. □

Lemma 16. *Let $2/5 < \varepsilon < 1/2$. The sequence*

$$(\nu,\alpha) \in N_\varepsilon \times A \mapsto \sqrt{n}\left(\frac{Z^\top R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}\int_0^1 h_{\nu;\nu_0}} - 1\right)$$

of processes converges weakly to

$$\mathrm{GP}\left(0, (\nu_1,\nu_2) \mapsto \frac{\int_0^1 h_{\nu_1;\nu_0} h_{\nu_2;\nu_0}}{\int_0^1 h_{\nu_1;\nu_0}\int_0^1 h_{\nu_2;\nu_0}}\right),$$

in $L^\infty(N_\varepsilon \times A)$.

Proof.

$$\sqrt{n}\left(\frac{Z^\top R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}} - \int_0^1 h_{\nu;\nu_0}(x)dx\right)$$

$$= \sqrt{n}U_{0,n}^2\left(\frac{\lambda_{0,n}^0}{n^{1+2(\nu-\nu_0)}\lambda_{0,n}}\right) + \frac{1}{\sqrt{n}}\sum_{m=1}^{n-1} U_{m,n}^2\left(\frac{\lambda_{m,n}^0}{n^{2(\nu-\nu_0)}\lambda_{m,n}} - h_{\nu;\nu_0}(m/n)\right)$$

$$+\frac{1}{\sqrt{n}}\sum_{m=1}^{n-1} B_{m,n} h_{\nu;\nu_0}(m/n) + \sqrt{n}\left(\frac{1}{n}\sum_{m=1}^{n-1} h_{\nu;\nu_0}(m/n) - \int_0^1 h_{\nu;\nu_0}(x)dx\right),$$

with $B_{m,n} = U_{m,n}^2 - 1$. Then a classical Borel-Cantelli argument (see, e.g., Ying, 1991, Lemma 4) shows that $\sup_{0\le m\le n-1}|U_{m,n}^2| = o_P(n^\delta)$ for every $\delta > 0$, so

$$\sup_{\nu\in N_\varepsilon,\ \alpha\in A}\left|\sqrt{n}U_{0,n}^2\left(\frac{\lambda_{0,n}^0}{n^{1+2(\nu-\nu_0)}\lambda_{0,n}}\right)\right| = o_P(1),$$

and

$$\sup_{\nu\in N_\varepsilon,\ \alpha\in A}\left|\frac{1}{\sqrt{n}}\sum_{m=1}^{n-1} U_{m,n}^2\left(\frac{\lambda_{m,n}^0}{n^{2(\nu-\nu_0)}\lambda_{m,n}} - h_{\nu;\nu_0}(m/n)\right)\right| = o_P(1).$$

thanks to Lemma 21. Finally Lemma 19, Lemma 24, Slutsky's lemma, and the continuous mapping theorem give the claim observing that the mapping $\nu \in N_\varepsilon \mapsto \int_0^1 h_{\nu;\nu_0}$ is bounded away from zero by continuity. □

Lemma 17. *Let $0 < \varepsilon < 1/2$, we have*

$$\sup_{\nu\in N_\varepsilon,\ \alpha\in A}\left|\frac{Z^\top R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}} - \int_0^1 h_{\nu;\nu_0}\right| = o_P(1).$$

Proof. First we have

$$\sup_{\nu\in N_\varepsilon}\left|\frac{1}{n}\sum_{m=1}^{n-1} B_{m,n} h_{\nu;\nu_0}\left(\frac{m}{n}\right)\right| = o_P(1),$$

thanks to Lemma 25, Lemma 26, and (Van Der Vaart and Wellner, 1996, Theorem 1.5.4, Theorem 1.5.7, and Lemma 1.10.2). (Note that no outer probabilities are needed here since the mappings $\nu \in N_\varepsilon \mapsto h_{\nu;\nu_0}(m/n)$ are continuous.) Then, proceed as for Lemma 16. □

Lemma 18. *The function $h_{\nu;\nu_0}$ is non-decreasing (resp. non-increasing) on $(0,1/2]$ when $\nu \ge \nu_0$ (resp. $\nu \le \nu_0$).*

Proof. Suppose that $\nu \ge \nu_0$. Use (6) along with the fact that the Hurwitz Zeta function verifies (Postnikov, 1988)

$$\frac{\partial \zeta_H}{\partial x}(\alpha;\, x) = -\alpha\zeta_H(\alpha+1;\, x),\ x > 0,\ \alpha > 1 \tag{22}$$

and has the representation

$$\zeta_H(\alpha;\, x) = \frac{1}{\Gamma(\alpha)}\int_0^{+\infty}\frac{t^{\alpha-1}e^{-tx}}{1-e^{-t}}\mathrm{d}t,\ x > 0,\ \alpha > 1,$$

where Γ is the classical Gamma function. So we have

$$\gamma_{2\nu+1}(x) = \frac{1}{\Gamma(2\nu+1)} \int_0^{+\infty} \frac{t^{2\nu}(e^{-tx}+e^{-t(1-x)})}{1-e^{-t}} \mathrm{d}t,\ x \in (0,1)\,,$$

and

$$\frac{\mathrm{d}\gamma_{2\nu+1}}{\mathrm{d}x}(x) = \frac{1}{\Gamma(2\nu+1)} \int_0^{+\infty} \frac{t^{2\nu+1}(e^{-t(1-x)}-e^{-tx})}{1-e^{-t}} dt,\ x \in (0,1)\,.$$

Now let $x \in [1/2, 1)$, the derivative of $h_{\nu;\nu_0}$ at x has the sign of

$$\gamma_{2\nu+1}(x)\frac{\mathrm{d}\gamma_{2\nu_0+1}}{\mathrm{d}x}(x) - \gamma_{2\nu_0+1}(x)\frac{\mathrm{d}\gamma_{2\nu+1}}{\mathrm{d}x}(x)$$

$$= \frac{1}{\Gamma(2\nu+1)\Gamma(2\nu_0+1)} \int_0^{+\infty}\int_0^{+\infty} \frac{t^{2\nu}s^{2\nu_0}(\eta(s,t;x)-\eta(t,s;x))}{\kappa(s,t)} \mathrm{d}t\mathrm{d}s$$

with $\eta(s,t;x) = s(e^{-tx}+e^{-t(1-x)})(e^{-s(1-x)}-e^{-sx})$ and $\kappa(s,t) = (1-e^{-t})(1-e^{-s}) = \kappa(t,s)$ thanks to the Fubini-Lebesgue theorem. Then, one can split the integral to have:

$$\frac{1}{\Gamma(2\nu+1)\Gamma(2\nu_0+1)} \left(\int_0^{+\infty}\int_t^{+\infty} \frac{t^{2\nu}s^{2\nu_0}(\eta(s,t;x)-\eta(t,s;x))}{\kappa(s,t)} \mathrm{d}t\mathrm{d}s \right.$$
$$\left. + \int_0^{+\infty}\int_t^{+\infty} \frac{s^{2\nu}t^{2\nu_0}(\eta(t,s;x)-\eta(s,t;x))}{\kappa(t,s)} \mathrm{d}t\mathrm{d}s \right)$$
$$= \frac{1}{\Gamma(2\nu+1)\Gamma(2\nu_0+1)} \int_0^{+\infty}\int_t^{+\infty} \frac{(t^{2\nu}s^{2\nu_0}-s^{2\nu}t^{2\nu_0})(\eta(s,t;x)-\eta(t,s;x))}{\kappa(s,t)} \mathrm{d}t\mathrm{d}s \leq 0$$

since $t^{2\nu}s^{2\nu_0} \leq s^{2\nu}t^{2\nu_0}$ when $s \geq t$, $\kappa(s,t) \geq 0$ and $\eta(s,t;x) \geq \eta(t,s;x)$ when $s \geq t$ and $x \geq 1/2$.

So we proved that $h_{\nu;\nu_0}$ is non-increasing on $[1/2, 1)$ and the first claim is due to the symmetry with respect to $1/2$. Observe that $h_{\nu;\nu_0} = 1/h_{\nu_0;\nu}$ for the second claim. □

Lemma 19. *Let $\varepsilon > 0$, we have*

$$\frac{1}{n}\sum_{m=1}^{n-1} h_{\nu;\nu_0}(m/n) = \int_0^1 h_{\nu;\nu_0}(x)dx + \mathscr{O}\left(\frac{1}{n^{\min(1,\,2\varepsilon)}}\right),$$

uniformly in $\nu \in N_\varepsilon$.

Proof. Using Lemma 18 and the symmetry w.r.t. $1/2$, we have again the classical

$$\int_{1/n}^{(l+1)/n} h_{\nu;\nu_0}(x)dx \leq \frac{1}{n}\sum_{m=1}^{l} h_{\nu;\nu_0}(m/n) \leq \int_0^{l/n} h_{\nu;\nu_0}(x)dx, \tag{23}$$

when $\nu \geq \nu_0$ and the reversed inequality when $\nu \leq \nu_0$.

The behaviour near $1/2$ is elucided as for Lemma 14 with a compacity argument to bound the values of $h_{\nu;\nu_0}$ uniformly at $1/4$ and $1/2$, thus giving a uniform $\mathscr{O}\left(n^{-1}\right)$.

Moreover, using (21), one has $h_{\nu;\nu_0}(x) \sim_{x\to 0} x^{2(\nu-\nu_0)}$ uniformly in $\nu \in N$ and then

$$\int_0^{1/n} h_{\nu;\nu_0}(x)dx = \mathscr{O}(n^{-2\varepsilon}), \tag{24}$$

uniformly in $\nu \in N_\varepsilon$. Therefore, we have

$$\sum_{m=1}^{l} h_{\nu;\nu_0}(m/n) = n\int_0^{1/2} h_{\nu;\nu_0}(x)dx + \mathcal{O}(1) + \mathcal{O}\left(n^{1-2\varepsilon}\right),$$

uniformly and using the symmetry of $h_{\nu;\nu_0}$ gives the desired result. □

Lemma 20. *Let $1 \le m \le l$, we have*

$$\frac{\lambda^0_{m,n}}{n^{2(\nu-\nu_0)}\lambda_{m,n}} = \left(1+\mathcal{O}(m^{-2})\right) h_{\nu;\nu_0}(m/n)$$

uniformly in $\nu \in N$ and $\alpha \in A$.

Proof. Lemma 13 gives

$$\frac{\lambda^0_{m,n}}{n^{2(\nu-\nu_0)}\lambda_{m,n}} = \frac{\sum_{j\in\mathbb{Z}} \frac{(1+u_{n,m,j}(\nu_0,\alpha_0))}{|j+m/n|^{2\nu_0+1}}}{\sum_{j\in\mathbb{Z}} \frac{(1+u_{n,m,j}(\nu,\alpha))}{|j+m/n|^{2\nu+1}}},$$

with

$$\sum_{j\in\mathbb{Z}} \frac{(1+u_{n,m,j}(\nu,\alpha))}{|j+m/n|^{2\nu+1}} = \gamma_{2\nu+1}(m/n)\left(1+\mathcal{O}\left(m^{-2}\right)\right),$$

uniformly. The desired result follows then from elementary manipulations. □

Lemma 21. *Let $0<\varepsilon<1/2$, we have*

$$\sup_{(\nu,\alpha)\in N_\varepsilon\times A} \frac{1}{n}\left|\sum_{m=1}^{n-1}\left(\frac{\lambda^0_{m,n}}{n^{2(\nu-\nu_0)}\lambda_{m,n}} - h_{\nu;\nu_0}\left(\frac{m}{n}\right)\right)\right| = \mathcal{O}\left(\frac{1}{n^{\min(3/4,\ 5\varepsilon/4)}}\right).$$

Proof. Let $1\le m\le l$ and $\beta\in(0,1)$. Using Lemma 20 and the correct sign for the big-$\mathcal{O}$ term we have

$$\frac{1}{n}\left|\sum_{m=1}^{l}\left(\frac{\lambda^0_{m,n}}{n^{2(\nu-\nu_0)}\lambda_{m,n}} - h_{\nu;\nu_0}\left(\frac{m}{n}\right)\right)\right| \le \frac{1}{n}\sum_{m=1}^{l}\mathcal{O}\left(m^{-2}\right)h_{\nu;\nu_0}\left(\frac{m}{n}\right)$$

$$= \frac{1}{n}\sum_{m=1}^{\lfloor n^\beta\rfloor}\mathcal{O}(m^{-2})h_{\nu;\nu_0}(m/n) + \frac{1}{n}\sum_{m=\lfloor n^\beta\rfloor+1}^{l}\mathcal{O}(m^{-2})h_{\nu;\nu_0}(m/n),$$

with

$$\frac{1}{n}\sum_{m=\lfloor n^\beta\rfloor+1}^{l}\mathcal{O}(m^{-2})h_{\nu;\nu_0}(m/n) = \mathcal{O}\left(\frac{1}{n^{2\beta}}\right)\frac{1}{n}\sum_{m=1}^{l}h_{\nu;\nu_0}(m/n) = \mathcal{O}\left(\frac{1}{n^{2\beta}}\right)$$

uniformly in $\nu\in N_\varepsilon$ and $\alpha\in A$ thanks to (23) and by observing that the mapping $\nu\in N_\varepsilon \mapsto \int_0^{1/2} h_{\nu,\nu_0}$ is continuous. Moreover we have

$$\sum_{m=1}^{\lfloor n^\beta\rfloor}\mathcal{O}(m^{-2})h_{\nu;\nu_0}(m/n) \le \mathcal{O}(1)n\int_0^{\frac{\lfloor n^\beta\rfloor+1}{n}} h_{\nu;\nu_0}(x)dx = \mathcal{O}\left(\frac{1}{n^{2\varepsilon(1-\beta)}}\right)$$

uniformly thanks to (23) and (24). Using (17), the symmetry of $h_{\nu;\nu_0}$ and taking $\beta = 3/8$ then gives the desired result. □

Lemma 22. *The family $\mathscr{F} = \{h_{\nu;\nu_0},\ \nu \in N\}$ is a VC-subgraph class of functions (Van Der Vaart and Wellner, 1996, Section 2.6.2).*

Proof. We have $h_{\nu;\nu_0} = \gamma_{2\nu_0+1}/\gamma_{2\nu+1}$, with

$$\gamma_{2\nu+1}(x) = \sum_{j=0}^{+\infty} \frac{1}{(j+1+x)^{2\nu+1}} + \sum_{j=0}^{+\infty} \frac{1}{(j+2-x)^{2\nu+1}} + \frac{1}{x^{2\nu+1}} + \frac{1}{(1-x)^{2\nu+1}},\ 0 < x < 1.$$

The four families of functions involved in the last expression of $\gamma_{2\nu+1}$ are monotonous with respect to $\nu \in N$ so they are VC-subgraph classes of functions. Indeed, let $(x_1,y_1),(x_2,y_2) \in (0,1)\times\mathbb{R}$, there cannot be $f,g \in \mathscr{F}$ such that $f(x_1) < y_1$, $f(x_2) \geq y_2$, $g(x_1) \geq y_1$, and $g(x_2) < y_2$, since we have either $g \leq f$ or $f \leq g$.

The class $\mathscr{F}$ is therefore a VC-subgraph class of functions thanks to Lemma 2.6.18 from Van Der Vaart and Wellner (1996) by observing that the statement *(viii)* remains valid for a monotonous function defined only on a subset of $\mathbb{R}$ containing the ranges of the elements of $\mathscr{F}$. □

Let us give some definitions before continuing. For, $\varepsilon > 0$, (21) can be used to show that there exists some $C > 0$ such that

$$h_{\nu;\nu_0}(x) \leq F_\varepsilon(x) = C\min(x,1-x)^{-1+2\varepsilon},\ \text{for all } 0 < x < 1 \text{ and } \nu \in N_\varepsilon. \tag{25}$$

The function F_ε will be called the envelope of the family $\mathscr{F}_\varepsilon = \{h_{\nu;\nu_0},\ \nu \in N_\varepsilon\}$ of functions.

Lemma 23. *For all $\varepsilon > 1/4$, we have*

$$\frac{1}{n}\sum_{m=1}^{n-1}\left(h_{\nu_1;\nu_0}(m/n) - h_{\nu_2;\nu_0}(m/n)\right)^2 \to \int_0^1 \left(h_{\nu_1;\nu_0} - h_{\nu_2;\nu_0}\right)^2,$$

uniformly in $\nu_1,\nu_2 \in N_\varepsilon$.

Proof. Let $\delta > 0$, there exists $\alpha > 0$ such that:

$$\int_0^\alpha \left(h_{\nu_1;\nu_0} - h_{\nu_2;\nu_0}\right)^2 \leq 4\int_0^\alpha F_\varepsilon^2 \leq \delta/5$$

and

$$\frac{1}{n}\sum_{m=1}^{\lfloor \alpha n\rfloor}\left(h_{\nu_1;\nu_0}(m/n) - h_{\nu_2;\nu_0}(m/n)\right)^2 \leq \frac{4}{n}\sum_{m=1}^{\lfloor \alpha n\rfloor} F_\varepsilon^2(m/n) \leq \delta/5,$$

uniformly in $\nu_1,\nu_2 \in N_\varepsilon$. The same bounds also hold by symmetry for similar quantities related to $[1-\alpha,\ 1]$. Furthermore, a compacity argument shows that the derivative of $x \in (0,\ 1) \mapsto \left(h_{\nu_1;\nu_0}(x) - h_{\nu_2;\nu_0}(x)\right)^2$ is bounded on $[\alpha,\ 1-\alpha]$ uniformly in $\nu_1,\nu_2 \in N_\varepsilon$. Consequently, the standard technique for bounding the approximation error of Riemann sums give

$$\left|\frac{1}{n}\sum_{m=\lfloor \alpha n\rfloor+1}^{\lfloor (1-\alpha)n\rfloor-1}\left(h_{\nu_1;\nu_0}(m/n) - h_{\nu_2;\nu_0}(m/n)\right)^2 - \int_\alpha^{1-\alpha}\left(h_{\nu_1;\nu_0} - h_{\nu_2;\nu_0}\right)^2\right| \leq \varepsilon/5,$$

uniformly in $\nu_1,\nu_2 \in N_\varepsilon$, for a sufficiently large n.

□

For $n \geq 3$ and $0 \leq m \leq n-1$, define $B_{m,n} = U_{m,n}^2 - 1$.

Lemma 24. *Let $3/8 < \varepsilon < 1/2$, we have*

$$\frac{1}{\sqrt{n}} \sum_{m=1}^{n-1} B_{m,n} h_{\cdot;\nu_0}\left(\frac{m}{n}\right) \rightsquigarrow \mathrm{GP}\left(0, (\nu_1, \nu_2) \mapsto 2\int_0^1 h_{\nu_1;\nu_0} h_{\nu_2;\nu_0}\right),$$

in $L^\infty(N_\varepsilon)$.

Proof. Since $\varepsilon > 3/8$, then the envelope F_ε defined by (25) is in $L^4(0,1)$. Moreover, Lemma 22 shows that $\mathscr{F}_\varepsilon \subset \mathscr{F}$ satisfies the uniformly integrable entropy condition (Van Der Vaart and Wellner, 1996, Theorem 2.6.7). Then, with $d = \|\cdot - \cdot\|_{L^2(0,1)}$, we have $\|F_\varepsilon\|_{L^2(0,1)} < +\infty$ and thus $(\mathscr{F}_\varepsilon, d)$ is totally bounded thanks to the uniformly integrable entropy condition.

With $Y_{m,n} \colon g \in \mathscr{F}_\varepsilon \mapsto n^{-1/2} B_{m,n} g(m/n)$, the measurability conditions from (Van Der Vaart and Wellner, 1996, p. 205) are met since the suprema can be replaced by ones on countable sets. Indeed, using the surjection $\rho\colon \nu \in N_\varepsilon \mapsto h_{\nu;\nu_0} \in \mathscr{F}_\varepsilon$, the suprema on subsets of $\mathscr{F}_\varepsilon \times \mathscr{F}_\varepsilon$ are suprema on subsets of $(N_\varepsilon \times N_\varepsilon, \|\cdot - \cdot\|_2)$, with $\|\cdot\|_2$ standing for the euclidean norm. A subset of a separable metric space is separable. Then, the sample path continuity of $\nu \in N_\varepsilon \mapsto Y_{m,n}(\rho(\nu))$ is inherited from the continuity of $\nu \in N_\varepsilon \mapsto h_{\nu;\nu_0}(x)$, for $0 < x < 1$.

Since $F_\varepsilon \in L^4(0,1)$ is monotonous on $(0, 1/2]$, one has $n^{-1}\sum_{m=1}^{n-1} F_\varepsilon^4 \to \int_0^1 F_\varepsilon^4$. Then, for $\delta > 0$, the Lindeberg condition holds:

$$\begin{aligned}
\sum_{m=1}^{n-1} \mathsf{E}\left(\sup_{g\in\mathscr{F}_\varepsilon} Y_{m,n}^2(g) \mathbb{1}_{\sup_{g\in\mathscr{F}_\varepsilon}|Y_{m,n}(g)|>\delta}\right) &= \delta^2 \sum_{m=1}^{n-1} \mathsf{E}\left(\delta^{-2} \sup_{g\in\mathscr{F}_\varepsilon} Y_{m,n}^2(g) \mathbb{1}_{\delta^{-1}\sup_{g\in\mathscr{F}_\varepsilon}|Y_{m,n}(g)|>1}\right) \\
&\leq \delta^2 \sum_{m=1}^{n-1} \mathsf{E}\left(\delta^{-4} \sup_{g\in\mathscr{F}_\varepsilon} Y_{m,n}^4(g) \mathbb{1}_{\delta^{-1}\sup_{g\in\mathscr{F}_\varepsilon}|Y_{m,n}(g)|>1}\right) \\
&\leq \delta^2 \sum_{m=1}^{n-1} \mathsf{E}\left(\delta^{-4} \sup_{g\in\mathscr{F}_\varepsilon} Y_{m,n}^4(g)\right) \\
&\leq \frac{\mathsf{E}\left(B_{1,1}^4\right)}{\delta^2 n^2} \sum_{m=1}^{n-1} F_\varepsilon^4(m/n) = \mathscr{O}\left(\frac{1}{n}\right).
\end{aligned}$$

Furthermore, for $\delta_n \to 0$, we have

$$\begin{aligned}
\sup_{d(g_1,\, g_2)\leq\delta_n} \sum_{m=1}^{n-1} \mathsf{E}\left((Y_{m,n}(g_1) - Y_{m,n}(g_2))^2\right) &= \mathsf{E}\left(B_{1,1}^2\right) \sup_{d(g_1,\, g_2)\leq\delta_n} \frac{1}{n} \sum_{m=1}^{n-1} (g_1(m/n) - g_2(m/n))^2 \\
&= o(1) + \mathsf{E}\left(B_{1,1}^2\right) \delta_n^2 \to 0
\end{aligned}$$

on $\mathscr{F}_\varepsilon$ thanks to Lemma 23.

Then, for $g_1, g_2 \in \mathscr{F}_\varepsilon$, one has

$$\mathrm{Cov}\left(\sum_{m=1}^{n-1} Y_{m,n}(g_1), \sum_{m=1}^{n-1} Y_{m,n}(g_2)\right) \to 2\int_0^1 g_1 g_2,$$

using Lemma 23 and Lemma 18 to show that $\frac{1}{n}\sum_{m=1}^{n-1} h_{\nu;\nu_0}^2(m/n) \mapsto \int_0^1 h_{\nu;\nu_0}^2$.

Finally, with $\mu_{n,m} = n^{-1} B_{m,n}^2 \delta_{m/n}$, one has $0 < \mu_{n,m} F_\varepsilon^2 < +\infty$ almost surely and $\sum_{m=1}^{n-1} \mu_{n,m} F_\varepsilon^2 = \mathscr{O}_P(1)$ using Markov's inequality.

We can then conclude using Lemma 2.11.6 and Theorem 2.11.1 from Van Der Vaart and Wellner (1996) and the reformulation from $L^\infty(\mathscr{F}_\varepsilon)$ to $L^\infty(N_\varepsilon)$ is an application of the continuous mapping theorem. ☐

Lemma 25. *Let $\nu > \nu_0 - 1/2$, we have*

$$g_n(\nu) = \frac{1}{n}\sum_{m=1}^{n-1} B_{m,n} h_{\nu;\nu_0}\left(\frac{m}{n}\right) = o_P(1).$$

Proof. Let $\delta = \nu - \nu_0 + 1/2 > 0$. Then, (25) yields

$$\mathsf{E}\left(g_n^2(\nu)\right) = \frac{\mathrm{Var}(B_{0,n})}{n^2}\sum_{m=1}^{n-1} h^2_{\nu;\nu_0}\left(\frac{m}{n}\right) \lesssim \frac{1}{n^2}\sum_{m=1}^{\lfloor n/2\rfloor}\left(\frac{m}{n}\right)^{4\delta-2} = \frac{1}{n^{4\delta}}\sum_{m=1}^{\lfloor n/2\rfloor} m^{4\delta-2},$$

which converges to zero no matter how δ compares with $1/4$. ☐

Lemma 26. *Let $0 < \varepsilon < 1/2$ and define*

$$g_n\colon \nu \in N_\varepsilon \mapsto \frac{1}{n}\sum_{m=1}^{n-1} B_{m,n} h_{\nu;\nu_0}\left(\frac{m}{n}\right).$$

The sequence $(g_n)_{n\geq 2}$ is asymptotically uniformly equicontinuous in probability for $(x,y)\mapsto |x-y|$ (Van Der Vaart and Wellner, 1996, Chapter 1.5).

Proof. Let $n \geq 2$ and $1 \leq m \leq \lfloor n/2 \rfloor$ and define

$$\eta_{m,n}\colon \nu \in N_\varepsilon \mapsto \gamma_{2\nu+1}(m/n) = \sum_{j\in\mathbb{Z}} \frac{1}{|j+m/n|^{2\nu+1}}.$$

These functions are smooth. Furthermore, (21) shows that $\eta_{m,n}(\nu) \approx (n/m)^{2\nu+1}$ and similar derivations yield $\frac{\mathrm{d}\eta_{m,n}}{\mathrm{d}\nu}(\nu) \approx (n/m)^{2\nu+1}\ln(n/m)$, for $\nu \in N_\varepsilon$, $1 \leq m \leq \lfloor n/2 \rfloor$, and n.

Write $\kappa_{m,n}\colon \nu \in N_\varepsilon \mapsto h_{\nu;\nu_0}(m/n)$. Then, for all $\delta > 0$ and $\nu \in N_\varepsilon$, we have

$$\left|\frac{\mathrm{d}\kappa_{m,n}}{\mathrm{d}\nu}(\nu)\right| \lesssim \left(\frac{n}{m}\right)^{1-2\varepsilon+2\delta}. \tag{26}$$

Now let $\nu_1, \nu_2 \in N_\varepsilon$. If one chooses $p > 1$ and $\delta > 0$ such that $p(1-2\varepsilon+2\delta) < 1$, then we have by Hölder's inequality with $1/q + 1/p = 1$

$$|g_n(\nu_1) - g_n(\nu_2)|$$

$$\leq \left(\frac{1}{n}\sum_{m=1}^{n-1}|B_{m,n}|^q\right)^{1/q}\cdot\left(\frac{1}{n}\sum_{m=1}^{n-1}\sup_{\nu\in N_\varepsilon}\left|\frac{\mathrm{d}\kappa_{m,n}}{\mathrm{d}\nu}(\nu)\right|^p\right)^{1/p}\cdot|\nu_1-\nu_2|,$$

which is enough to prove asymptotic uniform equicontinuity using (26), the symmetry of $h_{\nu;\nu_0}$ with respect to $1/2$, and the fact that the $B_{m,n}$s admit moments of every order. ☐

A.5 Proof of Theorem 3

We start by providing a few cumbersome expressions. Remember that the quantities $\lambda_{m,n}$ depends on ν and α but have their argument dropped to avoid cumbersome notations. The function $\mathbb{L}_n$ is smooth for any realization, and

$$\mathbb{L}_n(\nu,\phi,\alpha) = \ln(\phi) + \frac{1}{n}\sum_{m=0}^{n-1}\ln(\lambda_{m,n}) + \frac{\phi_0}{n\phi}\sum_{m=0}^{n-1}\frac{\lambda^0_{m,n}U^2_{m,n}}{\lambda_{m,n}}$$

$$\frac{\partial\mathbb{L}_n}{\partial\nu}(\nu,\phi,\alpha) = \frac{1}{n}\sum_{m=0}^{n-1}\frac{\frac{\partial\lambda_{m,n}}{\partial\nu}}{\lambda_{m,n}} - \frac{\phi_0}{n\phi}\sum_{m=0}^{n-1}\frac{\lambda^0_{m,n}\frac{\partial\lambda_{m,n}}{\partial\nu}U^2_{m,n}}{\lambda^2_{m,n}}$$

$$\frac{\partial^2\mathbb{L}_n}{\partial\nu^2}(\nu,\phi,\alpha) = \frac{1}{n}\sum_{m=0}^{n-1}\frac{\frac{\partial^2\lambda_{m,n}}{\partial\nu^2}\lambda_{m,n} - \left(\frac{\partial\lambda_{m,n}}{\partial\nu}\right)^2}{\lambda^2_{m,n}}$$

$$-\frac{\phi_0}{n\phi}\sum_{m=0}^{n-1}\frac{\lambda^0_{m,n}\left(\frac{\partial^2\lambda_{m,n}}{\partial\nu^2}\lambda_{m,n} - 2\left(\frac{\partial\lambda_{m,n}}{\partial\nu}\right)^2\right)U^2_{m,n}}{\lambda^3_{m,n}}$$

$$\frac{\partial^3\mathbb{L}_n}{\partial\nu^3}(\nu,\phi,\alpha) = \frac{1}{n}\sum_{m=0}^{n-1}\lambda^{-3}_{m,n}\left(\frac{\partial^3\lambda_{m,n}}{\partial\nu^3}\lambda_{m,n} + \frac{\partial^2\lambda_{m,n}}{\partial\nu^2}\frac{\partial\lambda_{m,n}}{\partial\nu}\right.$$

$$\left.-2\frac{\partial^2\lambda_{m,n}}{\partial\nu^2}\frac{\partial\lambda_{m,n}}{\partial\nu}\right)\lambda_{m,n} - 2\frac{\partial\lambda_{m,n}}{\partial\nu}\left(\frac{\partial^2\lambda_{m,n}}{\partial\nu^2}\lambda_{m,n} - \left(\frac{\partial\lambda_{m,n}}{\partial\nu}\right)^2\right)$$

$$-\frac{\phi_0}{n\phi}\sum_{m=0}^{n-1}\lambda^{-4}_{m,n}\lambda^0_{m,n}\left(\lambda_{m,n}\left(\frac{\partial^3\lambda_{m,n}}{\partial\nu^3}\lambda_{m,n} + \frac{\partial^2\lambda_{m,n}}{\partial\nu^2}\frac{\partial\lambda_{m,n}}{\partial\nu} - 4\frac{\partial^2\lambda_{m,n}}{\partial\nu^2}\frac{\partial\lambda_{m,n}}{\partial\nu}\right)\right.$$

$$\left.-3\frac{\partial\lambda_{m,n}}{\partial\nu}\left(\frac{\partial^2\lambda_{m,n}}{\partial\nu^2}\lambda_{m,n} - 2\left(\frac{\partial\lambda_{m,n}}{\partial\nu}\right)^2\right)\right)U^2_{m,n}$$

Lemma 10 and the following will help to analyze roughly the previous expressions. Exceptionally, the arguments of the $\lambda_{m,n}$s are not dropped.

Lemma 27. *Let $\delta > 0$, $0 \le m \le \lfloor n/2\rfloor$, $\nu \in N$, $\alpha \in A$ and $k \in \{1,2,3\}$. The functions $\lambda_{m,n}$ are smooth and:*

$$\frac{1}{n}\left|\frac{\partial^k\lambda_{m,n}}{\partial\nu^k}(\nu,\alpha)\right| \lesssim \frac{1}{m^{2\nu+1-\delta}}, \quad \text{if } 1 \le m \le \lfloor n/2\rfloor$$

and

$$1 \lesssim \frac{1}{n}\left|\frac{\partial^k\lambda_{0,n}}{\partial\nu^k}(\nu,\alpha)\right| \lesssim 1.$$

Proof. We have

$$\left|\frac{\partial^k\lambda_{m,n}}{\partial\nu^k}(\nu,\alpha)\right| = \sum_{j\in\mathbb{Z}}\frac{\left|\ln^k\left(\alpha^2+(m+jn)^2\right)\right|}{\left(\alpha^2+(m+jn)^2\right)^{\nu+1/2}} \lesssim \sum_{j\in\mathbb{Z}}\frac{1}{\left(\alpha^2+(m+jn)^2\right)^{\nu+1/2-\delta/2}}.$$

Lemma 10 gives the result. □

Proof of Theorem 3. Remember that $A = \{\alpha_0\}$ and consider $\mathbb{L}_n$ as a function of (ν, ϕ) only. We start by elucidating the behavior of the score function.

Let $e_2(\nu) = \mathsf{E}\left(\psi_\nu^2(U)\right)$ for $\nu > 0$. Stein (1999, proof of Theorem 1, Section 6.7) shows that

$$\operatorname{Cov}\left(\frac{\sqrt{n}}{2}\nabla\mathbb{L}_n(\nu_0, \phi_0)\right) = \underbrace{\begin{pmatrix} 2\ln^2(n) + 4\ln(n)e(\nu_0) + 2e_2(\nu_0) & -\frac{\ln(n)}{\phi_0} - \frac{e(\nu_0)}{\phi_0} \\ -\frac{\ln(n)}{\phi_0} - \frac{e(\nu_0)}{\phi_0} & \frac{1}{2\phi_0^2} \end{pmatrix}}_{C_n} + \mathscr{O}\left(n^{-\varepsilon}\right),$$

for some $\varepsilon > 0$.

Define

$$A_n = \frac{2\phi_0}{\sqrt{V(\nu_0)}} \begin{pmatrix} \frac{1}{2\phi_0} & 0 \\ \ln(n) + e(\nu_0) & \sqrt{V(\nu_0)} \end{pmatrix}.$$

One has $A_n^\top C_n A_n = 2I_2$ and so

$$\operatorname{Cov}\left(\frac{\sqrt{n}}{2}A_n^\top\nabla\mathbb{L}_n(\nu_0, \phi_0)\right) = A_n^\top \operatorname{Cov}\left(\frac{\sqrt{n}}{2}\nabla\mathbb{L}_n(\nu_0, \phi_0)\right)A_n \to 2I_2.$$

Let us show the Lyapunov condition. To do so, write

$$\mathbb{L}_{m,n}(\nu, \phi) = \ln(\phi) + \ln(\lambda_{m,n}) + \frac{\phi_0}{\phi}\frac{\lambda_{m,n}^0 U_{m,n}^2}{\lambda_{m,n}},$$

so $\mathbb{L}_n(\nu, \phi) = n^{-1}\sum_{m=0}^{n-1}\mathbb{L}_{m,n}(\nu, \phi)$. Let $\delta > 0$, we have

$$\sum_{m=0}^{n-1}\mathsf{E}\left(\left|\frac{\partial\mathbb{L}_{m,n}}{\partial\phi}(\nu_0, \phi_0)\right|^{2+\delta}\right) = \frac{1}{\phi_0^{2+\delta}n^{2+\delta}}\sum_{m=0}^{n-1}\mathsf{E}\left(\left|1 - U_{m,n}^2\right|^{2+\delta}\right) = \frac{\mathsf{E}\left(\left|1 - U_{0,1}^2\right|^{2+\delta}\right)}{\phi_0^{2+\delta}n^{1+\delta}},$$

and

$$\sum_{m=0}^{n-1}\mathsf{E}\left(\left|\frac{\partial\mathbb{L}_{m,n}}{\partial\nu}(\nu_0, \phi_0)\right|^{2+\delta}\right) = \frac{\mathsf{E}\left(\left|1 - U_{0,1}^2\right|^{2+\delta}\right)}{n^{2+\delta}}\sum_{m=0}^{n-1}\left(\frac{\frac{\partial\lambda_{m,n}^0}{\partial\nu}}{\lambda_{m,n}^0}\right)^{2+\delta} \lesssim \frac{1}{n^{1+3\delta/4}},$$

using Lemma 10, Lemma 27, and the derivatives expressions. Therefore,

$$\sum_{m=0}^{n-1}\mathsf{E}\left(\|\sqrt{n}A_n^\top\nabla\mathbb{L}_{m,n}(\nu_0, \phi_0)\|^{2+\delta}\right) \lesssim n^{1+\delta/2}\ln(n)^{2+\delta}\sum_{m=0}^{n-1}\mathsf{E}\left(||\nabla\mathbb{L}_{m,n}(\nu_0, \phi_0)||^{2+\delta}\right) \to 0.$$

The Lyapunov condition is fulfilled, and it is straightforward to verify that the score function is centered, yielding

$$\frac{\sqrt{n}}{2\sqrt{2}}A_n^\top\nabla\mathbb{L}_n(\nu_0, \phi_0) \rightsquigarrow \mathscr{N}(0, I_2).$$

Consider now the Hessian matrix. One has

$$\mathsf{E}\left(\nabla^2\mathbb{L}_n(\nu_0, \phi_0)\right) = \frac{n}{2}\operatorname{Cov}\left(\nabla\mathbb{L}_n(\nu_0, \phi_0)\right),$$

so

$$\frac{1}{2}\mathsf{E}\left(\nabla^2\mathbb{L}_n(\nu_0,\,\phi_0)\right) - C_n = \mathscr{O}\left(n^{-\varepsilon}\right).$$

Furthermore, using Lemma 10, Lemma 27, and the previous expressions, it is straightforward to check that

$$\operatorname{Var}\left(\frac{\partial^2\mathbb{L}_n}{(\partial\nu)^p\,(\partial\nu)^{2-p}}(\nu_0,\,\phi_0)\right) = \mathscr{O}\left(n^{-2\varepsilon}\right),$$

for $\varepsilon > 0$ small enough, so $\nabla^2\mathbb{L}_n(\nu_0,\,\phi_0)/2 - C_n = \mathscr{O}_P\left(n^{-\varepsilon}\right)$ for some $\varepsilon > 0$ and therefore $A_n^{\mathsf{T}}\nabla^2\mathbb{L}_n(\nu_0,\,\phi_0)A_n \to 4I_2$, in probability.

Let us now roughly bound the third derivatives. For $\varepsilon > 0$, $m \geq 1$, $k \in \{0,1,2,3\}$, and $|\nu - \nu_0| \leq \varepsilon$, one has

$$\left|\frac{\partial^k\lambda_{m,n}}{\partial\nu^k}(\nu_0+\varepsilon,\,1)\right| \leq \left|\frac{\partial^k\lambda_{m,n}}{\partial\nu^k}(\nu,\,1)\right| \leq \left|\frac{\partial^k\lambda_{m,n}}{\partial\nu^k}(\nu_0-\varepsilon,\,1)\right|.$$

Therefore, using again the previous expressions, Lemma 10, and Lemma 27, it is straightforward to show that

$$\mathsf{E}\left(\sup_{p\in\{0,1,2,3\},|\nu-\nu_0|\leq\varepsilon,|\phi-\phi_0|\leq\varepsilon}\left|\frac{\partial^3\mathbb{L}_n}{(\partial\nu)^p\,(\partial\phi)^{3-p}}(\nu,\,\phi)\right|\right) = \mathscr{O}\left(n^{17\varepsilon}\right). \tag{27}$$

Now we have

$$0 = \Delta\mathbb{L}_n(\nu_0,\,\phi_0) + \Delta^2\mathbb{L}_n(\nu_0,\,\phi_0)\begin{pmatrix}\widehat{\nu}_n - \nu_0\\ \widehat{\phi}_n - \phi_0\end{pmatrix} + \mathscr{O}_P\left(n^{17\varepsilon}\left|\left|\begin{pmatrix}\widehat{\nu}_n - \nu_0\\ \widehat{\phi}_n - \phi_0\end{pmatrix}\right|\right|^2\right),$$

thanks to (27) and using Theorem 2 leads to

$$0 = A_n^{\mathsf{T}}\Delta\mathbb{L}_n(\nu_0,\,\phi_0) + \left(A_n^{\mathsf{T}}\Delta^2\mathbb{L}_n(\nu_0,\,\phi_0)A_n + o_P(1)\right)A_n^{-1}\begin{pmatrix}\widehat{\nu}_n - \nu_0\\ \widehat{\phi}_n - \phi_0\end{pmatrix},$$

and

$$\sqrt{2n}A_n^{-1}\begin{pmatrix}\widehat{\nu}_n - \nu_0\\ \widehat{\phi}_n - \phi_0\end{pmatrix} = -\left(A_n^{\mathsf{T}}\Delta^2\mathbb{L}_n(\nu_0,\,\phi_0)A_n + o_P(1)\right)^{-1}A_n^{\mathsf{T}}\Delta\mathbb{L}_n(\nu_0,\,\phi_0) \rightsquigarrow \mathscr{N}(0,\,I_2),$$

thanks to Slutsky's Lemma. This gives the result. □

A.6 Proofs of Theorem 4 and Theorem 5

The posterior mean and the error of the GP does not depend on ϕ, so all derivations will be written with $\phi = 1$. We also suppose that $\phi_0 = 1$ without loss of generality. Furthermore, we will write $\omega = (\nu,\alpha)$ and use the coefficients $c_j(\omega) = c_j(\nu,\alpha)$ defined in Section A.1 to avoid cumbersome expressions in this section.

The sample paths are square-integrable almost surely so the convergence of the Fourier expansion (7) holds with probability one also in $L^2\,[0,\,1]$. The proofs will rely on using Parseval's identity. Note that we assume $\nu_0 > 1/2$ to avoid dealing with conditionally convergent series. Indeed, in this case the expansion (7) converges almost surely absolutely pointwise, so the hypothesis of Proposition 1 are fulfilled.

Let $\alpha > 0$, $\nu > 0$ and $j \in \mathbb{Z}$, we have

$$2\left|c_j(\xi - \widehat{\xi}_n)\right|^2 = \left(\frac{c_j(\omega)\sum_{j_1\in\mathbb{Z}\setminus\{0\}}\sqrt{c_{j+j_1 n}(\omega_0)}U_{1,|j+j_1 n|}}{\sum_{j_1\in\mathbb{Z}} c_{j+nj_1}(\omega)} - \frac{\sqrt{c_j(\omega_0)}U_{1,|j|}\sum_{j_1\in\mathbb{Z}\setminus\{0\}} c_{j+nj_1}(\omega)}{\sum_{j_1\in\mathbb{Z}} c_{j+nj_1}(\omega)}\right)^2 + \left(\frac{c_j(\omega)\sum_{j_1\in\mathbb{Z}\setminus\{0\}}\sqrt{c_{j+nj_1}(\omega_0)}U_{2,|j+nj_1|}\operatorname{sign}(j+nj_1)}{\sum_{j_1\in\mathbb{Z}} c_{j+nj_1}(\omega)} - \frac{\sqrt{c_j(\omega_0)}U_{2,|j|}\operatorname{sign}(j)\sum_{j_1\in\mathbb{Z}\setminus\{0\}} c_{j+nj_1}(\omega)}{\sum_{j_1\in\mathbb{Z}} c_{j+nj_1}(\omega)}\right)^2. \tag{28}$$

The two terms in (28) are idenpendent and they are also identically distributed if $m \notin \{0, n/2\}$.

Proof of Theorem 4. Consider $m \notin \{0,\ n/2\}$ and the indexes of the form $m+nj$, with $j \in \mathbb{Z}$. The two terms in (28) are independent and identically distributed, so there exists a χ^2_2 distributed variable $A_{m,j,n}$ such that

$$\left|c_{m+jn}(\xi - \widehat{\xi}_n)\right|^2 = a_{m,j,n}(\nu)A_{m,j,n}/2,$$

with

$$a_{m,j,n}(\nu) = c^2_{m+jn}(\omega)\frac{\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega_0) - c_{m+jn}(\omega_0)}{\left(\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega)\right)^2} + c_{m+jn}(\omega_0)\left(1 - \frac{c_{m+jn}(\omega)}{\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega)}\right)^2, \tag{29}$$

and the dependence of $a_{m,j,n}$ on (the fixed) α be removed for readability. Using Lemma 10 gives, for $1 \le m \le l$ and $j \ne 0$:

$$a_{m,j,n}(\nu) \lesssim (jn)^{-4\nu-2}m^{4\nu+2-2\nu_0-1} + (jn)^{-2\nu_0-1}, \tag{30}$$

and

$$a_{m,0,n}(\nu) \lesssim n^{-2\nu_0-1} + m^{4\nu-2\nu_0+1}n^{-4\nu-2}, \tag{31}$$

uniformly in j and m. So, this yields

$$\sum_{j\in\mathbb{Z}} a_{m,j,n}(\nu) \lesssim n^{-2\nu_0-1} + n^{-4\nu-2}m^{4\nu-2\nu_0+1}. \tag{32}$$

The first two statements then follow from Lemma 28, Lemma 29, the identity

$$\sum_{j\in\mathbb{Z}}\left|c_{m+jn}(\xi - \widehat{\xi}_n)\right|^2 = \sum_{j\in\mathbb{Z}}\left|c_{n-m+jn}(\xi - \widehat{\xi}_n)\right|^2, \tag{33}$$

for every $0 \le m \le n-1$, and the Fubini-Tonelli thereom.

For the last statement, let $1 \le m \le l$. We have

$$\mathsf{E}\left(\left|c_{m+jn}(\xi - \widehat{\xi}_n)\right|^2\right) = c^2_{m+jn}(\omega)\frac{\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega_0)}{\left(\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega)\right)^2}$$

$$+c_{m+jn}(\omega_0)\left(1-2\frac{c_{m+jn}(\omega)}{\sum_{j_1\in\mathbb{Z}}c_{m+nj_1}(\omega)}\right),$$

for every $j\in\mathbb{Z}$ and therefore

$$\sum_{j\in\mathbb{Z}}\mathsf{E}\left(\left|c_{m+jn}(\xi-\widehat{\xi}_n)\right|^2\right)$$

$$=\left(1+\mathscr{O}(m^{-2})\right)\sum_{j\in\mathbb{Z}}|m+jn|^{-4\nu-2}\frac{\sum_{j_1\in\mathbb{Z}}|m+nj_1|^{-2\nu_0-1}}{\left(\sum_{j_1\in\mathbb{Z}}|m+nj_1|^{-2\nu-1}\right)^2}$$

$$+|m+jn|^{-2\nu_0-1}\left(1-2\frac{|m+jn|^{-2\nu-1}}{\sum_{j_1\in\mathbb{Z}}|m+nj_1|^{-2\nu-1}}\right)$$

$$=\frac{\left(1+\mathscr{O}(m^{-2})\right)}{n^{2\nu_0+1}}\vartheta_{\nu;\nu_0}(m/n),$$

using Lemma 13 and after a few algebraic manipulations. If $\nu\geq(\nu_0-1/2)/2$, then the function $\vartheta_{\nu;\nu_0}$ has a finite limit in zero. If $(\nu_0-1)/2<\nu<(\nu_0-1/2)/2$, then a taylor expansion of its derivative shows that it is decreasing in a neighborhood of zero. This shows:

$$\frac{1}{n}\sum_{m=1}^{l}\vartheta_{\nu;\nu_0}(m/n)\to\int_0^{1/2}\vartheta_{\nu;\nu_0},\tag{34}$$

and Parseval's identity and the Fubini-Tonelli thereom gives

$$n^{2\nu_0}\mathsf{E}\left(\mathrm{ISE}_n(\nu,\alpha;\xi)\right)=o(1)+\frac{2}{n}\sum_{m=1}^{l}\left(1+\mathscr{O}(m^{-2})\right)\vartheta_{\nu;\nu_0}(m/n)\to\int_0^1\vartheta_{\nu;\nu_0},$$

killing the $\mathscr{O}(m^{-2})$ term by splitting by how m compares with $\sqrt{n}$ as in the proof of Lemma 21. □

Lemma 28. *Let $n\geq 2$ be even, one has*

$$\mathsf{E}\left(\sup_{\nu\in N,\alpha\in A}\sum_{j\in\mathbb{Z}}\left|c_{n/2+jn}(\xi-\widehat{\xi}_n)\right|^2\right)\lesssim n^{-2\nu_0-1}.$$

Proof. First, observe that $\sum_{j\in\mathbb{Z}}\left|c_{n/2+jn}(\xi-\widehat{\xi}_n)\right|^2=2\sum_{j\geq 0}\left|c_{n/2+jn}(\xi-\widehat{\xi}_n)\right|^2$.

Let $\nu\in N$, $\alpha\in A$, and $j\geq 0$. One has

$$2\left|c_{n/2+jn}(\xi-\widehat{\xi}_n)\right|^2\left(\sum_{j_1\in\mathbb{Z}}c_{n/2+nj_1}(\omega)\right)^2$$

$$=\left(\left(c_{n/2+jn}(\omega)\sqrt{c_{n/2+jn}(\omega_0)}-\sqrt{c_{n/2+jn}(\omega_0)}\sum_{j_1\in\mathbb{Z}\setminus\{j\}}c_{n/2+nj_1}(\omega)\right)U_{1,|n/2+jn|}\right.$$

$$\left.+2c_{n/2+jn}(\omega)\sum_{j_1=0,j_1\neq|j|}^{+\infty}\sqrt{c_{n/2+j_1n}(\omega_0)}U_{1,|n/2+j_1n|}\right)^2$$

$$+\left(\sqrt{c_{n/2+jn}(\omega_0)}\sum_{j_1\in\mathbb{Z}} c_{n/2+nj_1}(\omega)U_{2,|n/2+jn|}\right)^2$$

So

$$2\left|c_{n/2+jn}(\xi-\widehat{\xi}_n)\right|^2 = d_{j,n}^2 D_{j,n}^2 + b_{j,n}^2 B_{j,n}^2$$

with $D_{j,n}$ and $B_{j,n}$ standard Gaussians and

$$d_{j,n}^2 \le c_{n/2+jn}(\omega_0) + 4c_{n/2+jn}^2(\omega)\frac{\sum_{j_1=0}^{+\infty} c_{n/2+j_1n}(\omega_0)}{\left(\sum_{j_1\in\mathbb{Z}} c_{n/2+nj_1}(\omega)\right)^2}$$

$$\lesssim ((j+1/2)n)^{-2\nu_0-1} + (j+1/2)^{-4\nu-2}n^{-2\nu_0-1}$$

uniformly in ν and α using Lemma 10 and the inequality $c_{n/2+jn}(\omega) \le ((j+1/2)n)^{-2\nu-1}$. Similarly, one has $b_{j,n}^2 \lesssim ((j+1/2)n)^{-2\nu_0-1}$. The statement then follows using the Fubini-Tonelli theorem. □

Lemma 29. *Let $n\ge 2$ and $p\ge 1$. Then, there exists standard Gaussians $\{A_{j,n},\ j\in\mathbb{Z}\}$ and $\varepsilon>0$ such that*

$$\sum_{j\in n\mathbb{Z}}\left|c_{jn}(\xi-\widehat{\xi}_n)\right|^{2p} \lesssim A_{0,n}^{2p}\left(n^{-2\nu_0-1}+n^{-4\nu-2}\right)^p + \sum_{j\in\mathbb{Z}\setminus\{0\}} A_{j,n}^{2p} j^{-1-\varepsilon}\left(n^{-2\nu_0-1}+n^{-4\nu-2}\right)^p,$$

uniformly in $\nu\in N$ and $\alpha\in A$.

Proof. Let $\nu\in N$ and $\alpha\in A$. For $j\in\mathbb{Z}\setminus\{0\}$, one has

$$2\left|c_{jn}(\xi-\widehat{\xi}_n)\right|^2\left(\sum_{j_1\in\mathbb{Z}} c_{nj_1}(\omega)\right)^2$$

$$=\left(\left(c_{jn}(\omega)\sqrt{c_{jn}(\omega_0)} - \sqrt{c_{jn}(\omega_0)}\sum_{j_1\in\mathbb{Z}\setminus\{j\}} c_{nj_1}(\omega)\right)U_{1,|jn|}\right.$$

$$\left.+c_{jn}(\omega)\sqrt{c_0(\omega_0)}U_{1,0} + 2c_{jn}(\omega)\sum_{j_1=1,j_1\neq j}^{+\infty}\sqrt{c_{j_1n}(\omega_0)}U_{1,|j_1n|}\right)^2$$

$$+\left(\sqrt{c_{jn}(\omega_0)}\sum_{j_1\in\mathbb{Z}} c_{nj_1}(\omega)U_{2,|jn|}\right)^2$$

Moreover,

$$2\left|c_0(\xi-\widehat{\xi}_n)\right|^2\left(\sum_{j_1\in\mathbb{Z}} c_{nj_1}(\omega)\right)^2 = \left(\sqrt{c_0(\omega_0)}\sum_{j_1\in\mathbb{Z}\setminus\{0\}} c_{nj_1}(\omega)U_{1,|0|}\right.$$

$$\left.+2c_0(\omega)\sum_{j_1=1}^{+\infty}\sqrt{c_{j_1n}(\omega_0)}U_{1,|j_1n|}\right)^2$$

Then, proceed as in Lemma 28 to conclude, observing that $c_{nj_1}(\omega) \lesssim (|j_1|n)^{-2\nu-1}$ for $j_1\neq 0$. □

Lemma 30. *Let $\varepsilon > 0$. With the notations of Theorem 5, we have:*

$$\mathsf{P}\left(\widehat{\nu}_n \leq \nu_0 - 1/2 - \varepsilon\right) \lesssim e^{-C\sqrt{n}},$$

for some $C > 0$.

Proof. First, one has

$$\begin{aligned}
\mathsf{P}\left(\widehat{\nu}_n \leq \nu_0 - 1/2 - \varepsilon\right) &\leq \mathsf{P}\left(\inf_{a_\nu \leq \nu \leq \nu_0 - 1/2 - \varepsilon} \mathbb{M}_n(\nu, \alpha) \leq \inf_{\nu_0 - 1/2 - \varepsilon \leq \nu \leq b_\nu} \mathbb{M}_n(\nu, \alpha)\right) \\
&\leq \mathsf{P}\left(\inf_{a_\nu \leq \nu \leq \nu_0 - 1/2 - \varepsilon} \mathbb{M}_n(\nu, \alpha) - \mathbb{M}_n(\nu_0, \alpha) \leq 0\right).
\end{aligned}$$

Then, let $a_\nu \leq \nu \leq \nu_0 - 1/2 - \varepsilon$, we have, using Lemma 14, Lemma 10, and Jensen inequality:

$$\begin{aligned}
\mathbb{M}_n(\nu, \alpha) &= \mathscr{O}\left(\frac{\ln(n)}{n}\right) + \ln\left(\frac{Z^\mathsf{T} R_{\nu,\alpha}^{-1} Z}{n}\right) + \int_0^1 g_\nu + 2(\nu_0 - \nu)\ln(n) \\
&= \mathscr{O}(1) + \ln\left(\frac{Z^\mathsf{T} R_{\nu,\alpha}^{-1} Z}{n}\right) + 2(\nu_0 - \nu)\ln(n) \\
&\geq \mathscr{O}(1) + \ln\left(\frac{\sum_{m=1}^{l} U_{m,n}^2 m^{2(\nu - \nu_0)}}{n^{1+2(\nu-\nu_0)}}\right) \\
&= \mathscr{O}(1) + \ln\left(\frac{1}{n}\sum_{m=1}^{l} U_{m,n}^2 \left(\frac{m}{n}\right)^{2(\nu-\nu_0)}\right) \\
&\geq \mathscr{O}(1) + \ln\left(\frac{1}{n}\sum_{m=1}^{l} U_{m,n}^2 \left(\frac{m}{n}\right)^{-1-2\varepsilon}\right) \\
&\geq \mathscr{O}(1) + 2\varepsilon\ln(n) + \ln\left(\sum_{m=1}^{l} U_{m,n}^2 m^{-1-2\varepsilon}\right) \\
&\geq \mathscr{O}(1) + 2\varepsilon\ln(n) + \ln\left(\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} U_{m,n}^2 m^{-1-2\varepsilon}\right) \\
&\geq \mathscr{O}(1) + 2\varepsilon\ln(n) + \ln\left(\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} U_{m,n}^2 \lfloor\sqrt{n}\rfloor^{-1-2\varepsilon}\right) \\
&\geq \mathscr{O}(1) + \varepsilon\ln(n) + \ln\left(\frac{1}{\lfloor\sqrt{n}\rfloor}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} U_{m,n}^2\right) \\
&\geq \mathscr{O}(1) + \varepsilon\ln(n) + \frac{1}{\lfloor\sqrt{n}\rfloor}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} \ln\left(U_{m,n}^2\right)
\end{aligned}$$

with a uniform big-$\mathscr{O}$. Let $\delta > 0$ and $t > 0$, we have

$$\begin{aligned}
\mathsf{P}\left(-\frac{1}{\lfloor\sqrt{n}\rfloor}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor}\ln(U_{m,n}^2) \geq \delta\right) &= \mathsf{P}\left(e^{-\frac{t}{\lfloor\sqrt{n}\rfloor}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor}\ln(U_{m,n}^2)} \geq e^{t\delta}\right) \\
&= \mathsf{P}\left(\left(\prod_{m=1}^{\lfloor\sqrt{n}\rfloor}U_{m,n}^{-2}\right)^{\frac{t}{\lfloor\sqrt{n}\rfloor}} \geq e^{t\delta}\right) \\
&= \mathsf{P}\left(\prod_{m=1}^{\lfloor\sqrt{n}\rfloor}|U_{m,n}|^{-2t} \geq e^{t\delta\lfloor\sqrt{n}\rfloor}\right) \\
&\leq e^{-\delta\lfloor\sqrt{n}\rfloor/4}\mathsf{E}\left(|U_{1,1}|^{-1/2}\right)^{\lfloor\sqrt{n}\rfloor},
\end{aligned}$$

with $t = 1/4$ and $C = \mathsf{E}\left(U_{1,1}^{-1/2}\right) < +\infty$. Therefore:

$$C^{\lfloor\sqrt{n}\rfloor}e^{-\delta\lfloor\sqrt{n}\rfloor/4} = e^{(\ln(C)-\delta/4)\lfloor\sqrt{n}\rfloor},$$

which converges exponentially to zero if δ is high enough.

Furthermore, Lemma 10 gives

$$\mathbb{M}_n(\nu_0,\alpha) = \mathscr{O}(1) + \ln\left(\frac{Z^\mathsf{T}R_{\nu_0,\alpha}^{-1}Z}{n}\right) = \mathscr{O}(1) + \ln\left(n^{-1}\sum_{m=0}^{n-1}U_{m,n}^2\right),$$

and we have

$$\mathsf{P}\left(\ln\left(n^{-1}\sum_{m=0}^{n-1}U_{m,n}^2\right) \geq \delta\right) = \mathsf{P}\left(n^{-1}\sum_{m=0}^{n-1}U_{m,n}^2 \geq e^{\delta}\right) \leq e^{-Cn},$$

for some $C > 0$ if $\delta > 0$ is high enough, using also a Chernoff bound argument. Now, putting all the pieces together yields:

$$\inf_{a_\nu \leq \nu \leq \nu_0 - 1/2 - \varepsilon} \mathbb{M}_n(\nu,\alpha) - \mathbb{M}_n(\nu_0,\alpha)$$

$$\geq \mathscr{O}(1) + \varepsilon\ln(l) + \frac{1}{\lfloor\sqrt{n}\rfloor}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor}\ln\left(U_{m,n}^2\right) - \ln\left(n^{-1}\sum_{m=0}^{n-1}U_{m,n}^2\right)$$

giving the result thanks to the pigeonhole principle. □

Proof of Theorem 5. Let $\varepsilon > 0$ and use the notations (29). The functions $a_{m,j,n}$ are C^1 and, for $\nu \in [\nu_0 - 1/2 - \varepsilon, b_\nu]$, the derivatives of the $a_{m,j,n}$s lead to tedious expressions, which in turns lead, with Lemma 10 and Lemma 27, to

$$\begin{aligned}
\left|a'_{m,0,n}(\nu)\right| &\lesssim m^{\delta}n^{-2\nu_0-1} + m^{4\nu-2\nu_0+1}n^{-4\nu-2+\delta} \\
&\lesssim m^{\delta}n^{-2\nu_0-1} + m^{-2\nu_0+1}n^{-2+\delta}(m/n)^{2\nu_0-2+4\varepsilon} \\
&= m^{\delta}n^{-2\nu_0-1} + m^{-1+4\varepsilon}n^{-2\nu_0+\delta-4\varepsilon}
\end{aligned}$$

and, for $j \neq 0$, to

$$\begin{aligned}\left|a'_{m,j,n}(\nu)\right| &\lesssim (|j|n)^{-4\nu-2+\delta} m^{-2\nu_0-1} + (|j|n)^{-2\nu_0-2\nu-2+\delta} m^{2\nu+1} \\ &\lesssim |j|^{-3/2} n^{-2\nu_0-4\varepsilon+\delta} m^{-2\nu_0-1} + |j|^{-3/2} n^{-2\nu_0-2+\delta} m^1 (m/n)^{\nu_0-1+2\varepsilon} \\ &= |j|^{-3/2} n^{-2\nu_0-4\varepsilon+\delta} m^{-2\nu_0-1} + |j|^{-3/2} n^{-3\nu_0-1+\delta-2\varepsilon} m^{\nu_0+2\varepsilon}\end{aligned}$$

for $\delta > 0$ small enough and uniformly in $1 \leq m \leq l$, $j \neq 0$ and $\nu \in [\nu_0 - 1/2 - \varepsilon, b_\nu]$. Then,

$$\begin{aligned}&\sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \mathsf{E}\left(A_{m,j,n}\left|a_{m,j,n}(\widehat{\nu}_n) - a_{m,j,n}(\nu_0)\right| \mathbb{1}_{\widehat{\nu}_n \geq \nu_0-1/2-\varepsilon}\right) \\ &\leq \sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \mathsf{E}\left(A_{m,j,n}\left|\widehat{\nu}_n - \nu_0\right| \sup_{\nu_0-1/2-\varepsilon\leq\nu\leq b_\nu} \left|a'_{m,j,n}(\nu)\right|\right) \\ &= \sqrt{\mathsf{E}\left(A_{1,0,1}^2\right)} \sqrt{\mathsf{E}\left((\widehat{\nu}_n - \nu_0)^2\right)} \sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \sup_{\nu_0-1/2-\varepsilon\leq\nu\leq b_\nu} \left|a'_{m,j,n}(\nu)\right| = o\left(n^{-2\nu_0}\right),\end{aligned}$$

using the above inequalities and Theorem 2. Therefore, Lemma 28, Lemma 29, the identity (33), and the Fubini-Tonelli theorem shows that

$$\mathsf{E}\left(\left|\mathrm{ISE}_n(\widehat{\nu}_n, \alpha; \xi) - \mathrm{ISE}_n(\nu_0, \alpha; \xi)\right| \mathbb{1}_{\widehat{\nu}_n \geq \nu_0-1/2-\varepsilon}\right) = o\left(n^{-2\nu_0}\right).$$

Furthermore, using again the Fubini-Tonelli theorem yields

$$\begin{aligned}&\mathsf{E}\left(\sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \left|c_{m+jn}(\xi - \widehat{\xi}_n)\right|^2 \mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\varepsilon}\right) \\ &= \sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \mathsf{E}\left(a_{m,j,n}(\widehat{\nu}_n) A_{m,j,n} \mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\varepsilon}\right) \\ &\leq \sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \sup_{a_\nu\leq\nu\leq\nu_0-1/2-\varepsilon} a_{m,j,n}(\nu) \mathsf{E}\left(A_{m,j,n} \mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\varepsilon}\right) \\ &\leq \sqrt{\mathsf{E}\left(A_{1,0,1}^2\right)} \sqrt{\mathsf{E}\left(\mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\varepsilon}\right)} \sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \sup_{a_\nu\leq\nu\leq\nu_0-1/2-\varepsilon} a_{m,j,n}(\nu) \\ &\leq \sqrt{\mathsf{E}\left(A_{1,0,1}^2\right)} \sqrt{\mathsf{E}\left(\mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\varepsilon}\right)} n^{-4a_\nu-2+4\varepsilon} \\ &= o\left(n^{-2\nu_0}\right),\end{aligned}$$

using inequalities (30) and (31), and Lemma 30. Then, the case $m = n/2$ is treated by Lemma 28 and the case $m = 0$ can be managed similarly using Lemma 29.

Finally, one can treat

$$\mathsf{E}\left(\mathrm{ISE}_n(\nu_0, \alpha; \xi) \mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\varepsilon}\right)$$

the same way and the desired result follows. □

A.7 Proofs of Section 4

Note that the finiteness of $\nu_0(f)$ is assumed so that f is necessarily nonzero. Consequently, the hypothesis of Lemma 9 is ultimately verified under the observation model (2) since f is continuous.

Proof of Proposition 6. The proof is based on the observation that

$$Z^{\mathsf{T}} R_{\nu,\alpha}^{-1} Z = \sum_{m=0}^{n-1} \frac{\left|\sum_{j\in m+n\mathbb{Z}} c_j(f)\right|^2}{\sum_{j\in m+n\mathbb{Z}} c_j(\nu,\alpha)}.$$

We give a full proof only for the third assumption. The proof for the second assumption is similar and the first is a particular case of the second.

Let $\varepsilon > 0$, $a_\nu \leq \nu \leq \nu_0(f) - \varepsilon$, $\alpha \in A$, and $p \in \mathbb{Z}$ such that $c_p(f) \neq 0$. Then, Lemma 14 gives

$$\begin{aligned} \mathbb{M}_n^f(\nu,\alpha) &= 2(\nu_0(f)-\nu)\ln(n) + \mathscr{O}(1) + \ln\left(\sum_{m=0}^{n-1} \frac{\left|\sum_{j\in m+n\mathbb{Z}} c_j(f)\right|^2}{\sum_{j\in m+n\mathbb{Z}} c_j(\nu,\alpha)}\right) \\ &\geq 2\varepsilon\ln(n) + \mathscr{O}(1) + \ln\left(\frac{\left|\sum_{j\in p+n\mathbb{Z}} c_j(f)\right|^2}{\sum_{j\in p+n\mathbb{Z}} c_j(\nu,\alpha)}\right) \\ &= 2\varepsilon\ln(n) + \mathscr{O}(1) \end{aligned}$$

uniformly by Lemma 10 and since $\left|\sum_{j\in p+n\mathbb{Z}} c_j(f)\right| \to c_p(f)$.

Moreover, for any $\alpha \in A$, we have:

$$\begin{aligned} \mathbb{M}_n^f(\nu_0(f)-\varepsilon/2,\alpha) &= \varepsilon\ln(n) + \mathscr{O}(1) + \ln\left(Z^{\mathsf{T}} R_{\nu,\alpha}^{-1} Z\right) \\ &\leq \varepsilon\ln(n) + \mathscr{O}(1), \end{aligned}$$

since $f \in H^{\beta+1/2}[0,1]$ for $\beta = \nu_0(f) - \varepsilon/2$. Indeed, this last Sobolev space is norm-equivalent to the reproducing kernel Hilbert space attached to the covariance function for $\nu = \nu_0(f) - \varepsilon/2$ and the quadratic term $Z^{\mathsf{T}} R_{\nu,\alpha}^{-1} Z$ is the squared norm of a projection of f (see, e.g., Wendland, 2004, Theorem 13.1). This completes the proof. □

Proof of Proposition 8. Let $\nu > \nu_0(f)$ and $\alpha > 0$. Then, Lemma 13 and Lemma 14 yield:

$$
\begin{aligned}
\mathbb{M}_n^f(\nu,\alpha) &= (2\nu_0(f)+1)\ln(n)+n^{-1}\ln(\det(R_{\nu,\alpha}))+\ln\left(\frac{Z^\top R_{\nu,\alpha}^{-1}Z}{n}\right)\\
&= (2\nu_0(f)+1)\ln(n)+\int_0^1 g_\nu+\mathscr{O}\left(\frac{\ln(n)}{n}\right)+\ln\left(n^{-2\nu-1}\sum_{m=0}^{n-1}\frac{\left|\sum_{j\in m+n\mathbb{Z}}c_j(f)\right|^2}{\sum_{j\in m+n\mathbb{Z}}c_j(\nu,\alpha)}\right)\\
&= (2\nu_0(f)+1)\ln(n)+\int_0^1 g_\nu+\mathscr{O}\left(\frac{\ln(n)}{n}\right)\\
&+\ln\left(\mathscr{O}\left(n^{-2\nu-1}\right)+\mathscr{O}\left(n^{-2\nu_0(f)-2}\right)+2\sum_{m=1}^{l}\frac{\left(1+\mathscr{O}(m^{-2})\right)\left|\sum_{j\in m+n\mathbb{Z}}c_j(f)\right|^2}{\gamma(2\nu+1;m/n)}\right)\\
&= \int_0^1 g_\nu+\mathscr{O}\left(\frac{\ln(n)}{n}\right)\\
&+\ln\left(\mathscr{O}\left(n^{2(\nu_0(f)-\nu)}\right)+\mathscr{O}\left(n^{-1}\right)+\frac{2}{n}\sum_{m=1}^{l}\frac{\left(1+\mathscr{O}(m^{-2})\right)\left|n^{\nu_0(f)+1}\sum_{j\in m+n\mathbb{Z}}c_j(f)\right|^2}{\gamma(2\nu+1;m/n)}\right)\\
&= \int_0^1 g_\nu+\mathscr{O}\left(\frac{\ln(n)}{n}\right)\\
&+\ln\left(\mathscr{O}\left(n^{2(\nu_0(f)-\nu)}\right)+\mathscr{O}\left(n^{-1}\right)+\mathscr{O}\left(n^{\nu_0(f)-\nu}\right)+\frac{2}{n}\sum_{m=1}^{l}\frac{\gamma^2\left(\nu_0(f)+1;m/n\right)}{\gamma(2\nu+1;m/n)}\right),
\end{aligned}
$$

by proceeding as in the proof of Lemma 21. To conclude, observe that $\gamma^2(\nu_0(f)+1;\cdot)/\gamma(2\nu+1;\cdot)$ is non-increasing in a neighborhood of zero if $\nu<\nu_0(f)+1/2$ and has a finite limit otherwise. □

References

M. L. Stein. *Interpolation of spatial data*. Springer Series in Statistics. Springer-Verlag, New York, 1999. Some theory for Kriging.

Thomas Santner, Brian Williams, and William Notz. *The Design and Analysis Computer Experiments*. 01 2003.

C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, USA, 2006.

B. Matérn. *Spatial Variation*. Springer-Verlag New York, 1986.

F. Bachoc. Asymptotic analysis of maximum likelihood estimation of covariance parameters for Gaussian processes: an introduction with proofs. In *Advances in Contemporary Statistics and Econometrics*, pages 283–303. Springer, 2021.

K. V. Mardia and R. J. Marshall. Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, 71(1):135–146, 1984.

F. Bachoc. Asymptotic analysis of the role of spatial sampling for covariance parameter estimation of Gaussian processes. *Journal of Multivariate Analysis*, 125:1–35, 2014.

Z. Ying. Asymptotic properties of a maximum likelihood estimator with data from a Gaussian process. *Journal of Multivariate analysis*, 36(2):280–296, 1991.

Z. Ying. Maximum likelihood estimation of parameters under a spatial sampling scheme. *The Annals of Statistics*, pages 1567–1590, 1993.

A. W. van der Vaart. Maximum likelihood estimation under a spatial sampling scheme. *The Annals of Statistics*, pages 2049 – 2057, 1996.

H. Zhang. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99(465):250–261, 2004.

W.-L. Loh. Fixed-domain asymptotics for a subclass of matérn-type Gaussian random fields. *The Annals of Statistics*, 33, 2006.

E. Anderes. On the consistent separation of scale and variance for Gaussian random fields. *The Annals of Statistics*, 38(2):870–893, 2010.

Michael L. Stein. Spline smoothing with an estimated order parameter. *The Annals of Statistics*, 21(3):1522–1544, 1993a.

Y. Chen, H. Owhadi, and A. Stuart. Consistency of empirical bayes and kernel flow for hierarchical parameter estimation. *Mathematics of Computation*, 90(332):2527–2578, 2021.

T. Karvonen. Asymptotic bounds for smoothness parameter estimates in gaussian process regression. *arXiv preprint arXiv:2203.05400*, 2022.

B. Szabó, A. W. van Der Vaart, and J. H. van Zanten. Frequentist coverage of adaptive nonparametric Bayesian credible sets. *The Annals of Statistics*, 43(4):1391–1428, 2015.

B. T. Knapik, B. T. Szabó, A. W. van Der Vaart, and J. H. van Zanten. Bayes procedures for adaptive inference in inverse problems for the white noise model. *Probability Theory and Related Fields*, 164(3):771–813, 2016.

T. Karvonen, G. Wynne, F. Tronarp, C. Oates, and S. Sarkka. Maximum likelihood estimation and uncertainty quantification for Gaussian process approximation of deterministic functions. *SIAM/ASA Journal on Uncertainty Quantification*, 8(3):926–958, 2020.

Michael L. Stein. A simple condition for asymptotic optimality of linear predictions of random fields. *Statistics and Probability Letters*, 17(5):399–404, August 1993b.

G. Matheron. The theory of regionalized variables and its applications. Technical Report Les cahiers du CMM de Fontainebleau, Fasc. 5, Ecole des Mines de Paris, 1971.

W. Wang and B.-Y. Jing. Gaussian process regression: Optimality, robustness, and relationship with kernel ridge regression. *Journal of Machine Learning Research*, 23(193):1–67, 2022.

P. Craven and G. Wahba. Smoothing noisy data with spline functions. Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 1979.

P. J. Brockwell and R. A. Davis. *Time series: theory and methods*. Springer science & business media, 1987.

A. W. Van Der Vaart and J. A. Wellner. *Weak convergence and empirical processes: with applications to statistics*. Springer Science & Business Media, 1996.

A. G. Postnikov. *Introduction to analytic number theory*, volume 68. American Mathematical Soc., 1988.

H. Wendland. *Scattered data approximation*, volume 17. Cambridge university press, 2004.