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

Sébastien Petit

MAXIMUM LIKELIHOOD ESTIMATION AND PREDICTION ERROR FOR A MATÉRN MODEL ON THE CIRCLE

BY SÉBASTIEN J. PETIT[1,a],

[1]*Laboratoire National de Métrologie et d'Essais, 78197 Trappes Cedex, France,* [a]*sebastien.petit@lne.fr*

This work considers Gaussian process interpolation with a periodized version of the Matérn covariance function introduced by Stein [22, Section 6.7]. Convergence rates are studied for the joint maximum likelihood estimation of the regularity and the amplitude parameters 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 fixed deterministic element of a Sobolev space of continuous functions is also considered, suggesting that bounding assumptions on some parameters can lead to different estimates.

1. Introduction. Gaussian process interpolation or kriging is a common technique for inferring an unknown function from noiseless data, which has applications in geostatistics [22], computer experiments [19], and machine learning [18]. Stein [22] stresses the importance of choosing a covariance function that fits the problem, promotes the use of the Matérn [15] 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., 3, for a review]. While several increasing-domain asymptotic frameworks have been exhaustively studied [see, e.g., 14, 2], fixed-domain frameworks are studied only with simplifications and for a restrained number of parameters to our knowledge [28, 29, 24, 30, 13, 1, 12].

Considering fixed-domain asymptotics, the regularity parameter of the Matérn covariance function seems to have been little studied, although Stein [22] presents numerous results suggesting it as the most impactful from the point of view of prediction error. To study parameter estimation, Stein [22, Section 6.7] proposes an asymptotic framework with equispaced observations on the torus and makes a conjecture about the asymptotic behavior of the maximum likelihood estimate based on the Fisher information matrix [see also 20, in the case of noisy observations]. This topic has only recently regained popularity. Indeed, Chen, Owhadi and Stuart [5] used the previous framework to show that the estimation of the regularity parameter is consistent if the other parameters remain fixed. Moreover, Karvonen [8] 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 fixed deterministic function from a Sobolev space. Other results were obtained in similar frameworks [23, 11] with observations corrupted by noise.

This article presents three main contributions. First, we focus on the one-dimensional case of the framework proposed by Stein [22, Section 6.7] to give an asymptotic normality result for the joint maximum likelihood estimation of the regularity and the amplitude parameters. 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

MSC2020 subject classifications: Primary 62E20, 62G20, 62M30.

Keywords and phrases: Fixed-domain asymptotics, Gaussian random field, Matérn-type covariance function, Regularity.

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. Section 2 introduces the asymptotic framework and our notations and Section 3 discusses relations to existing works. Then, Section 4 gives the main results. Finally, Section 5 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 introduced by Stein [22, Section 6.7] and defined by the uniformly absolutely convergent Fourier serie $k_\theta\colon x \in \mathbb{R} \mapsto \sum_{j\in\mathbb{Z}} \underline{c}_j(\theta) e^{2\pi i x j}$ with coefficients:

$$\underline{c}_j(\theta) = \frac{\phi}{(\alpha^2 + j^2)^{\nu+1/2}}, \quad \text{for } j \in \mathbb{Z} \text{ and } \theta = (\nu, \phi, \alpha) \in (0, +\infty)^3. \tag{1}$$

The function k_θ is continuous and strictly positive definite [see, e.g., 7, Theorem 1].

Assuming a centered process, the usual task in Gaussian process interpolation—also known as kriging—is to use the model $\xi \sim \mathrm{GP}\,(0,\, k_\theta)$ to infer the function f from the noiseless data

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

The parameter α is not identifiable as different values yield equivalent probability measures. However, ν and ϕ are identifiable [see, e.g., 22, Chapter 4 and Section 6.7]. Stein [22] advocates the *regularity* parameter ν as the key quantity governing the asymptotics of the prediction error. The amplitude parameter ϕ does not impact the posterior mean predictions but matters for uncertainty quantification [see, e.g., 21] and can be consistently estimated by maximum likelihood if ν is known [see, e.g., 30, who use a different parametrization].

2.2. *Best linear prediction.* The function f is usually predicted using the posterior mean function given by the kriging equations [16]. This predictor can be written simply in the framework presented above.

PROPOSITION 1. *Let $n \ge 1$ 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:*

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

where the $c_j(f)$s are the Fourier coefficients of f. The convergence of (3) holds uniformly absolutely.

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 their 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 $\Theta \subset (0, +\infty)^3$, a 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:

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

with ties broken arbitrarily and K_θ the covariance matrix of Z according to k_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, meaning that α is enforced to a fixed value (which will not necessarily be the ground truth). This type of assumption is more or less standard in the field [see, e.g., 28, 13, 5].

However, keeping ϕ unbounded is key to our results and for discussing the deterministic case in Section 5. Write $K_\theta = \phi R_{\nu,\alpha}$ for $\theta = (\nu, \phi, \alpha) \in (0, +\infty)^3$. The following proposition gives an expression for the *profile* likelihood, i.e., the infimum of $\mathbb{L}_n(\nu, \phi, \alpha)$ with respect to $\phi \in (0, +\infty)$ for fixed ν and α.

PROPOSITION 2. *[see, e.g., 19, Section 3.3.2] Let $\nu, \alpha > 0$. It holds that*

$$\inf_{\phi>0} \mathbb{L}_n(\nu, \phi, \alpha) = 1 + n^{-1}\ln(\det(R_{\nu,\alpha})) + \ln\left(\frac{Z^\mathsf{T} R_{\nu,\alpha}^{-1} Z}{n}\right). \tag{6}$$

Moreover, if Z is nonzero, then the infimum is uniquely reached by $\widehat{\phi}_n = Z^\mathsf{T} R_{\nu,\alpha}^{-1} Z/n$.

(The case $Z = 0$ is covered since both sides of (6) match.)

3. Relation to prior works.

3.1. *Linking the spectra of k_θ and K_θ.* As Craven and Wahba [6] and Stein [22, Section 6.7] point out, the framework introduced in Section 2.1 is convenient. In particular, it gives a closed-form identity

$$n^{-1}\phi\lambda_{m,n} = \sum_{j\in\mathbb{Z}} \underline{c}_{m+nj}(\theta)$$

linking the eigenvalues $\phi\lambda_{0,n}, \ldots, \phi\lambda_{n-1,n}$ of K_θ to those of k_θ given by (1). Furthermore, the matrices K_θ share the same eigenvectors (see Appendix A.2 for details). One has $n^{-1}\phi\lambda_{m,n} \to \underline{c}_m(\theta)$ for a fixed m but the ratio $n^{-1}\phi\lambda_{m,n}/\underline{c}_m(\theta)$ remains bounded away from one for m close to $n/2$.

3.2. *The consistency of $\widehat{\nu}_n$ for fixed ϕ and α [5].* Assuming observations from $\xi \sim \mathrm{GP}\left(0, k_{\theta_0}\right)$ under a similar model with $\theta_0 = (\nu_0, \phi_0, \alpha_0) \in (0, +\infty)^3$, Chen, Owhadi and Stuart [5] show the consistency of $\widehat{\nu}_n$ for equispaced observations on the d-dimensional torus for fixed parameters ϕ and α. A sketch of their reasoning for $d = 1$ is provided in this paragraph. The spectrum of K_θ is studied by showing that

$$n^{-1}\phi\lambda_{m,n} = e^{\mathcal{O}(1)}\underline{c}_m(\theta) = e^{\mathcal{O}(1)} m^{-2\nu-1}$$

uniformly in ν and $1 \le m \le n/2$.[1] This approximation yields:

$$\begin{cases} \ln\left(\det\left(K_\theta\right)\right) &= -2\nu n \ln(n) + n\ln(\phi) + n\mathcal{O}(1), \\ Z^\mathsf{T} K_\theta^{-1} Z &= \phi^{-1}\phi_0\ \mathcal{O}_\mathsf{P}\left(\ln(n)\right) \text{ if } \nu \le \nu_0 - 1/2, \\ Z^\mathsf{T} K_\theta^{-1} Z &= \phi^{-1}\phi_0\ e^{\mathcal{O}_\mathsf{P}(1)} n^{1+2(\nu-\nu_0)} \text{ if } \nu > \nu_0 - 1/2, \end{cases} \tag{7}$$

[1] The $\lambda_{m,n}$ satisfy $\lambda_{m,n} = \lambda_{n-m,n}$.

with uniform $\mathcal{O}_{\mathsf{P}}$-terms on some regularity ranges. The consistency for fixed parameters ϕ and α follows by observing that ν_0 is the turning point where the quadratic form starts dominating the log-determinant. The latter claims are also true if ν is estimated jointly with $\phi \in F$ for a set F bounded away from zero and infinity.

3.3. *Our contributions.* Consider now the case $F = (0, +\infty)$ by plugging (7) into (6), for $\nu > \nu_0 - 1/2$, to get

$$\inf_{\phi>0} \mathbb{L}_n(\nu, \phi, \alpha) = 1 + \mathcal{O}(1) + \ln(\phi_0) + \mathcal{O}_{\mathsf{P}}(1) - 2\nu_0 \ln(n) = \mathcal{O}_{\mathsf{P}}(1) - 2\nu_0 \ln(n),$$

which is not sharp enough. Therefore, a more precise analysis of how the spectrum of K_θ fluctuates around the one of k_θ is needed to study the profile likelihood. The following section provides an ingredient for this purpose. Coordination with tools for proving uniform central limit theorems makes it possible to study convergence rates for parameter estimation and prediction error in Section 4. Developments for studying the profile likelihood are used to provide insights in the case of a fixed deterministic function from a Sobolev space in Section 5, which also discusses related works in this setting.

3.4. *A symmetrized version of the Hurwitz zeta function.* Stein [22, Section 6.7] uses the function

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

for deriving the asymptotics of the Fisher information matrix of the model presented in Section 2.1. It will also play a major role in our analysis of the likelihood criterion.

The function γ is (jointly) smooth and related to the Hurwitz zeta function ζ_H by:

$$\gamma(\alpha; x) = \zeta_H(\alpha; x) + \zeta_H(\alpha; 1-x), \quad (\alpha, x) \in (1, +\infty) \times (0, 1). \tag{8}$$

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

Integrals involving $\gamma(\alpha; \cdot)$, for some $\alpha > 1$, will appear in the main body of the article. All of them are well-defined, and integrability statements are postponed to the proofs in the Appendix.

4. Main results.

4.1. *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:

$$\xi\colon x \in [0, 1] \mapsto \frac{1}{\sqrt{2}} \sum_{j\in\mathbb{Z}} \sqrt{\underline{c}_j(\theta_0)} \left(U_{1,|j|} + iU_{2,|j|}\mathrm{sign}(j)\right) e^{2\pi i x j}, \tag{9}$$

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 $j \geq 1$. Let P be the measure defined on the underlying probability space (assumed to be the completion of the product space, so the $U_{q,j}$s are coordinate projections). The convergence of the expansion (9) is meant pointwise both in $L^2(\mathsf{P})$ and almost surely. We further assume that $\nu_0 > 1/2$, so the almost sure uniform absolute-convergence of (9) is ensured. It holds that $\xi \sim \mathrm{GP}(0, k_{\theta_0})$.

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 $A, N \subset (0, +\infty)$ compact intervals and $\nu_0 \in N$. Additional hypotheses will sometimes be made on A, while keeping this level of generality on the interval N, which will no longer be mentioned. The following sections give convergence rates for parameter estimation and prediction error.

4.2. *Convergence rates of maximum likelihood estimation.* The following result states the consistency for the identifiable parameters and an upper bound on the convergence rates.

THEOREM 3. *Let $0 < \beta < 1/4$ and A an interval bounded away from zero and infinity. It holds that $\widehat{\nu}_n - \nu_0 = o_{\mathsf{P}}\left(n^{-\beta}\right)$ and $\widehat{\phi}_n - \phi_0 = o_{\mathsf{P}}\left(n^{-\beta}\right)$.*

A key step for proving Theorem 3 is to show that (a shift of) the profile likelihood converges in probability to

$$\int_0^1 \ln\left(\gamma\left(2\nu+1;\cdot\right)\right) + \ln\left(\int_0^1 \frac{\gamma\left(2\nu_0+1;\cdot\right)}{\gamma\left(2\nu+1;\cdot\right)}\right), \tag{10}$$

for $\nu > \nu_0 - 1/2$. The first term is a refinement of the $\mathcal{O}(1)$ appearing in (7) for the log-determinant. The second term is a refinement of the $\mathcal{O}_{\mathsf{P}}(1)$ appearing for the quadratic form. Jensen inequality shows that (10) is minimized by taking $\nu = \nu_0$.

Furthermore, similarly to Stein [22, Section 6.7], let us define

$$\psi_\nu \colon x \in (0,\, 1) \mapsto \frac{\sum_{j\in\mathbb{Z}} |x+j|^{-2\nu-1} \ln|x+j|}{\sum_{j\in\mathbb{Z}} |x+j|^{-2\nu-1}}, \quad \text{for } \nu > 0, \tag{11}$$

which is square integrable on $(0,\, 1)$. The following result proves the conjecture made by Stein [22, Section 6.7] when $d = 1$, $\widehat{\nu}_n$ is bounded, and α_0 is known.

THEOREM 4. *Suppose that $A = \{\alpha_0\}$. Then*

$$\sqrt{2n}\begin{pmatrix} \frac{\widehat{\phi}_n - \phi_0}{2\phi_0} - \left(\ln(n) + \mathsf{E}\left(\psi_{\nu_0}(V)\right)\right)\left(\widehat{\nu}_n - \nu_0\right) \\ \sqrt{\operatorname{Var}\left(\psi_{\nu_0}(V)\right)}\left(\widehat{\nu}_n - \nu_0\right) \end{pmatrix} \rightsquigarrow \mathcal{N}\left(0,\, I_2\right),$$

where V is a random variable distributed uniformly on $(0,\, 1)$.

4.3. *Convergence rates of the integrated squared error.* This section states our results about the expectation of (4) with estimated parameters. We begin with the case of fixed parameters.

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

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

which is smooth and integrable when $\nu > (\nu_0 - 1)/2$.

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

THEOREM 5. *Let $(\nu,\, \alpha) \in (0,\, +\infty)^2$. Then,*

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

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

and

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

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 [22, 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 6. *Let $A = \{\alpha\}$ with $0 < \alpha < +\infty$. Then,*

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

This last result shows that estimating the parameters yields asymptotically the same error as if the ground truth was known.

5. The deterministic case. Let $\beta > 0$ and define the Sobolev space

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

of (continuous) periodic functions. This section studies maximum likelihood estimation with equispaced observations (2) from a fixed deterministic periodic function $f\colon\ [0,\, 1] \to \mathbb{R}$ lying in a Sobolev space. Proceed by defining the smoothness

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

of f as Wang and Jing [26] and Karvonen [8]. We will assume that $\nu_0(f) \in (1/2,\, +\infty)$.

Suppose that $\widehat{\theta}_n = \left(\widehat{\nu}_n, \widehat{\phi}_n, \widehat{\alpha}_n \right)$ is a maximum likelihood estimate defined 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

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

will be used.

On "nice" bounded regions of $\mathbb{R}^d$, Karvonen [8] shows that $\liminf \widehat{\nu}_n \geq \nu_0(f)$ if α and ϕ are fixed. (See [10] and [9] for asymptotic studies in the deterministic case on the estimation of other parameters for a fixed regularity.) The following result shows that it holds on the circle no matter the assumption on F.

PROPOSITION 7. *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 [8] 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 $|c_j(f)| \approx |j|^{-\nu_0(f)-1}$—which is compatible with the definition of $\nu_0(f)$—for simplicity. The following result is a minor adaptation of the reasoning used by Chen, Owhadi and Stuart [5] and Karvonen [8], showing that $\widehat{\nu}_n \to \nu_0(f) + 1/2$ more generally in this setting if ϕ is bounded. The proof is omitted for brevity.

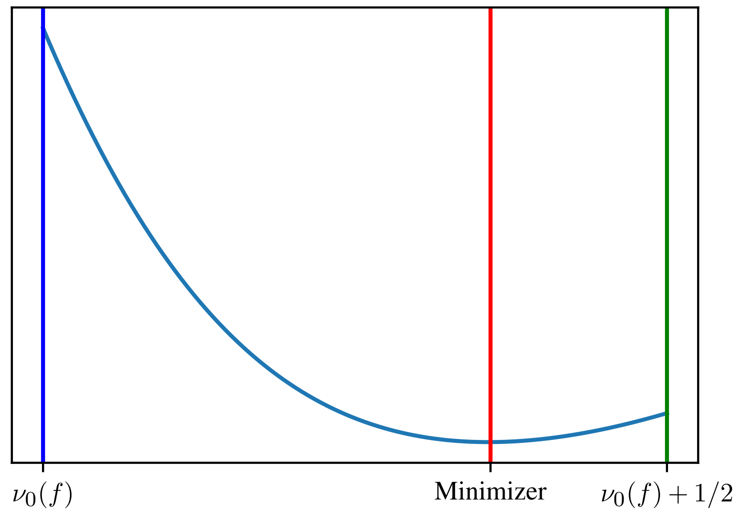

FIG 1. *The function* $\mathbb{M}_{\infty}^{f}$, *for* $\nu_0(f) = 1$. *A numerical approximation of the minimizer is about* 1.359. *Note that* $\mathbb{M}_{\infty}^{f}$ *is approximated numerically using finite sums for* γ *and discretizations for the integrals.*

PROPOSITION 8. *Suppose that* $|c_j(f)| \approx |j|^{-\nu_0(f)-1}$ *for nonzero* j *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 result suggests it does not hold when $F = (0, +\infty)$.

PROPOSITION 9. *Suppose that* $c_j(f) = |j|^{-\nu_0(f)-1}$ *for nonzero* j. *Then, for* $\nu > \nu_0(f)$ *and* $\alpha > 0$, *we have*

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

It is possible to derive 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 $\mathbb{M}_{\infty}^{f}$.

APPENDIX: PROOFS

A.1. 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 $N = [a_\nu, b_\nu]$, $A = [a_\alpha, b_\alpha]$, $0 < a_\nu < \nu_0 < b_\nu < +\infty$, 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 $N_\epsilon = [\nu_0 - 1/2 + \epsilon, +\infty) \cap N$ for $\epsilon > 0$. The notation $l = \lfloor (n-1)/2 \rfloor$ will often be used throughout the following.

A.2. Circulant matrices and useful facts. The framework introduced in Section 2.1 is convenient for analyzing kernel-based regression methods [see, e.g., 6]. This section reviews the properties needed for our purposes.

Let W be the $n \times n$ matrix with entries $W_{j,m} = n^{-1/2} e^{2\pi i j m/n}$, for $0 \leq j, m \leq n-1$. For every $\theta = (\nu, \phi, \alpha) \in (0, +\infty)^3$, the periodicity of k_θ implies that

$$K_\theta = \begin{pmatrix} k_\theta(0) & k_\theta\left(\frac{1}{n}\right) & \dots & k_\theta\left(\frac{n-1}{n}\right) \\ k_\theta\left(\frac{n-1}{n}\right) & k_\theta(0) & \dots & k_\theta\left(\frac{n-2}{n}\right) \\ \dots & \dots & \dots & \dots \\ k_\theta\left(\frac{1}{n}\right) & k_\theta\left(\frac{2}{n}\right) & \dots & k_\theta(0) \end{pmatrix}$$

is a circulant matrix and so is $R_{\nu,\alpha}$. Consequently [see, e.g., 4, p. 130], it holds that $R_{\nu,\alpha} = W \Delta_{\nu,\alpha} W^*$ with $\Delta_{\nu,\alpha} = \operatorname{diag}(\lambda_{0,n}, \dots, \lambda_{n-1,n})$ and

$$\lambda_{m,n} = \sum_{j=0}^{n-1} e^{-2\pi i j m/n} k_{\nu,1,\alpha}(j/n) = n \sum_{j \in \mathbb{Z}} c_{m+nj}(\nu, \alpha),\ 0 \leq m \leq n-1. \tag{12}$$

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

$$\lambda_{m,n} = \lambda_{n-m,n},\ \text{for } 1 \leq m \leq n-1. \tag{13}$$

The eigenvalue $\lambda_{0,n}$ is simple and there are l pairs $(\lambda_{m,n}, \lambda_{n-m,n})$, for $m \in [\![1, l]\!]$, where l is the shortcut defined in Section A.1. If n is even, then the eigenvalue $\lambda_{n/2,n}$ is also simple.

Furthermore, combining each pair of eigenvectors of W shows that $R_{\nu,\alpha} = P \Delta_{\nu,\alpha} P^\mathsf{T}$ for a unitary matrix P written using sines and cosines functions. Then, with $\theta_0 = (\nu_0, \phi_0, \alpha_0)$ the ground truth introduced in Section 4.1, write

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

with $\lambda_{0,n}^{(0)}, \dots, \lambda_{n-1,n}^{(0)}$ the eigenvalues of R_{ν_0,α_0} and $U_{0,n}, \dots, U_{n-1,n}$ drawn independently from a standard Gaussian. We have

$$Z^\mathsf{T} R_{\nu,\alpha}^{-1} Z = \phi_0 \sum_{m=0}^{n-1} \frac{U_{m,n}^2 \lambda_{m,n}^{(0)}}{\lambda_{m,n}}.$$

Our strategy to analyze this kind of expression will often consist of: 1) studying the sum for $m \in [\![1, l]\!]$; 2) using the equality (13); and 3) treating the remaining terms for $m = 0$ and eventually $m = n/2$ separately.

The following approximation discussed in Section 3.2 will sometimes be used.

LEMMA 10. *One has $n^{-1}\lambda_{0,n} \approx c_0(\nu,\alpha) \approx 1$ and $n^{-1}\lambda_{m,n} \approx c_m(\nu,\alpha) \approx m^{-2\nu-1}$ 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 (12)

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

Moreover

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

where $C_1^{b_\alpha,a_\nu,b_\nu} = \max(1,(b_\alpha^2+1/4)^{b_\nu+1/2})\zeta(2a_\nu+1)$. This shows $n^{-1}\lambda_{m,n} \approx c_m(\nu,\alpha)$. Then, $c_0(\nu,\alpha) \approx 1$ obviously and we have $(b_\alpha^2+1)^{-b_\nu-1/2} m^{-2\nu-1} \leq c_m(\nu,\alpha) \leq m^{-2\nu-1}$ for $m \geq 1$. □

Nevertheless, our results will require refined approximations, as explained in Section 3.3.

A.3. More notations and properties. For each n, it is straightforward to prove that the $\lambda_{m,n}$s are smooth functions of $(\nu,\,\alpha) \in (0,\,+\infty)^2$ by bounding the derivatives of the c_js uniformly on compacta (up to third-order derivatives suffice for our purposes). Using the formulas from Section A.2 then shows that $\mathbb{L}_n$ is also smooth for any realization.

Furthermore, define:

$$\mathbb{M}_n\colon\ (\nu,\alpha) \in N \times A \mapsto \inf_{\phi>0} \mathbb{L}_n\left(\nu,\phi,\alpha\right) + 2\nu_0 \ln(n) - \ln\left(\phi_0\right) - 1,$$

with ν_0 the ground truth introduced in Section 4.1. Its expression is given by Proposition 2 so it is a stochastic process which is smooth on the almost sure event $Z \neq 0$. The proofs mostly consist in studying $\mathbb{M}_n$.

For a compact interval $A \subset (0,\,+\infty)$, define now $\mathbb{U}_n\colon \nu \in N \mapsto \inf_{\alpha\in A} \mathbb{M}_n\left(\nu,\alpha\right)$. The object $\mathbb{U}_n$ is a stochastic process since the infima can be replaced by countable ones. Its almost sure continuity follows from the almost sure smoothness of $\mathbb{M}_n$ and the compacity of A.

Also, write $g_\nu = \ln\left(\gamma\left(2\nu+1;\cdot\right)\right)$ for $\nu > 0$ and

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

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

$$H\colon \nu \in (\nu_0 - 1/2,\,+\infty) \mapsto \int_0^1 h_{\nu;\nu_0},\quad G\colon \nu \in (0,\,+\infty) \mapsto \int_0^1 g_\nu,$$

and $\mathbb{U}\colon \nu \in (\nu_0 - 1/2,\,+\infty) \mapsto G(\nu) + \ln\left(H(\nu)\right)$. The smoothness of these functions is ensured by dominated convergence arguments (three derivatives suffice for our purposes).

A.4. Proofs of Section 2.2.

PROOF OF PROPOSITION 1. For $x \in [0,1]$, the kriging equations yield $\widehat{f}_n(x) = k_{\theta,x}^\intercal K_\theta^{-1} Z$, with $k_{\theta,x} = \left(k_\theta\left(m/n - x\right)\right)_{0\leq m\leq n-1}$. Then, using the matrix W defined in Section A.2, we have

$$W^*Z = \sqrt{n}\left(\sum_{j\in m+n\mathbb{Z}} c_j(f)\right)_{0\leq m\leq n-1} \tag{14}$$

and

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

where the sums converge absolutely. Then, the pointwise absolute convergence of (3) follows from elementary manipulations. The uniform absolute-convergence is ensured since the sum of the moduli does not depend on x. □

A.5. Proof of Theorem 3.

A.5.1. *Proof of the theorem.*

PROOF OF THEOREM 3. Let $0<\beta<1/4$, Proposition 2 gives almost surely

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

So

$$\frac{n^\beta}{\ln(n)}\left(\ln\left(\widehat{\phi}_n\right) - \ln(\phi_0)\right)$$

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

The latter converges to zero in probability thanks to (19), Slutsky's lemma [25, p. 32], Lemma 12, and the univariate delta method since the mapping $\ln\circ H$ is smooth. This implies that $\ln(\widehat{\phi}_n) - \ln(\phi_0) = o_{\mathsf{P}}\left(n^{-\beta}\right)$ for all $0<\beta<1/4$. Conclude using again the delta method. □

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

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

$$\begin{aligned}\mathbb{M}_n(\nu,\alpha) &= G(\nu) + \mathcal{O}\left(\ln(n)/n\right) + \ln\left(\frac{\phi_0^{-1} Z^{\mathsf{T}} R^{-1}_{\nu,\alpha} Z}{n^{1+2(\nu-\nu_0)}}\right)\\ &= \mathcal{O}(1) + \ln\left(\frac{\phi_0^{-1} Z^{\mathsf{T}} R^{-1}_{\nu,\alpha} Z}{n^{1+2(\nu-\nu_0)}}\right)\end{aligned}$$

uniformly in $\nu\in N$ and $\alpha\in A$ thanks to Lemma 15 and the continuity of G.

Now, let $0<\epsilon<1/4$, $\nu\in[a_\nu,\nu_0-1/2+\epsilon]$, and $\alpha\in A$ and write $l=\lfloor (n-1)/2\rfloor$. Lemma 10 yields

$$\frac{\phi_0^{-1} Z^{\mathsf{T}} R^{-1}_{\nu,\alpha} Z}{n^{1+2(\nu-\nu_0)}} \gtrsim \frac{1}{n}\sum_{m=1}^{l} U^2_{m,n}\left(\frac{m}{n}\right)^{2(\nu-\nu_0)} \ge \frac{1}{n}\sum_{m=1}^{l} U^2_{m,n}\left(\frac{m}{n}\right)^{-1+2\epsilon} \to \frac{1}{2^{1+2\epsilon}\epsilon}$$

in probability. Use the notations from Section A.3. Lemma 18 gives $\mathbb{U}_n(\nu_0) \to \int_0^1 g_{\nu_0}$ in probability, so we have

$$\inf_{\nu\in[a_\nu,\nu_0-1/2+\epsilon]} \mathbb{U}_n(\nu) - \mathbb{U}_n(\nu_0) \ge C + \ln\left(\frac{1}{n}\sum_{m=1}^{l} U^2_{m,n}\left(\frac{m}{n}\right)^{-1+2\epsilon}\right) - \mathbb{U}_n(\nu_0)$$

$$= C - \ln\left(2^{1+2\epsilon}\epsilon\right) - \int_0^1 g_{\nu_0} + o_{\mathsf{P}}(1),$$

thanks to the continuous mapping theorem, for a universal constant C. Letting $\epsilon \to 0$ shows that the above limit can be made arbitrarily high.

Moreover, let $0 < \epsilon < 1/2$. We have $\sup_{\nu \in N_\epsilon} |\mathbb{U}_n(\nu) - \mathbb{U}(\nu)| = o_{\mathsf{P}}(1)$ thanks to Lemma 18. Finally, the function $\mathbb{U}$ is continuous and strictly minimized 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 $\widehat{\nu}_n - \nu_0 = o_{\mathsf{P}}\left(n^{-\beta}\right)$ holds in probability.*

PROOF. Let $2/5 < \epsilon < 1/2$ and $0 < \beta < 1/2$ and use the notations from Section A.3. Lemma 16 implies $\sup_{\nu \in N_\epsilon} |\mathbb{U}_n(\nu) - \mathbb{U}(\nu)| = o_{\mathsf{P}}\left(n^{-\beta}\right)$. Moreover, the function $\mathbb{U}$ is C^3-smooth and we have $\mathbb{U}'(\nu_0) = 0$ and, with the notation given by (11):

$$\mathbb{U}''(\nu_0) = 4\left(\int_0^1 (\psi_{\nu_0})^2 - \left(\int_0^1 \psi_{\nu_0}\right)^2\right) > 0,$$

thanks to Jensen inequality. Finally, Lemma 11 and a second-order Taylor expansion around ν_0 give the rate $n^{-\beta/2}$. □

A.5.2. *Approximating* $\ln(\det(R_{\nu,\alpha}))$.

LEMMA 13. *Let $\nu \in N$, $\alpha \in A$, $1 \le m \le \lfloor n/2 \rfloor$, and $j \in \mathbb{Z}$. We have:*

$$c_{m+nj}(\nu,\alpha) = \frac{1 + u_{n,m,j}(\nu,\alpha)}{|jn+m|^{2\nu+1}}, \tag{15}$$

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

PROOF. Using (12), 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

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

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

LEMMA 14. *It holds that*

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

uniformly in $\nu \in N$ and $x \in (0,1)$. In particular, we have

$$\gamma(2\nu+1; x) \approx x^{-2\nu-1},$$

uniformly in $\nu \in N$ and $0 < x \le 1/2$.

PROOF. We have $0 \le \gamma(2\nu+1; x) - x^{-2\nu-1} - (1-x)^{-2\nu-1} \le 2\zeta(2a_\nu + 1)$. Deducing the second claim makes no difficulty. □

LEMMA 15. *Uniformly in $\nu \in N$ and $\alpha \in A$, we have*

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

PROOF. Let $\nu \in N$ and $\alpha \in A$. Using (12) and Lemma 13, we have

$$\lambda_{m,n}/n = \sum_{j\in\mathbb{Z}} c_{m+nj}(\nu,\alpha)$$

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

Therefore, using the notation $l = \lfloor (n-1)/2 \rfloor$, we have

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

with

$$|a_n(\nu,\,\alpha)| \le \left|\sum_{m=1}^{l} \ln(1+v_m)\right| = \mathcal{O}(1)$$

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

The function g_ν is symmetric with respect to $1/2$. Moreover, a direct consequence of Lemma 14 is that

$$g_\nu(x) = -(2\nu+1)\ln(x) + \mathcal{O}(1)\,, \tag{16}$$

uniformly in $\nu \in N$ and $0 < x \le 1/2$. For $\nu \in N$, the function g_ν is thus integrable on $(0,\,1)$. Furthermore, verifying that it is non-increasing on $(0,\,1/2]$ is straightforward using the derivative of $\gamma\,(2\nu+1;\,\cdot)$, so we have:

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

Use then (16) to get $\int_0^{1/n} g_\nu = \mathcal{O}(\ln(n)/n)$, uniformly in $\nu \in N$. The remainders $\int_{l/n}^{1/2} g_\nu$ and $\int_{1/2}^{(l+1)/n} g_\nu$ are $\mathcal{O}\left(n^{-1}\right)$ uniformly in $\nu \in N$ since the mapping $(\nu,\,x) \mapsto |g_\nu(x)|$ is bounded on $N \times [1/4,\,3/4]$ as it inherits the continuity of γ.

Therefore, we have

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

uniformly in $\nu \in N$. Moreover, Lemma 10 shows that $\ln\left(\lambda_{0,n}/n\right) = \mathcal{O}(1)$ and $\ln\left(\lambda_{n/2,n}/n\right) = \mathcal{O}(\ln(n))$ uniformly for n even. One can then conclude using (13). $\square$

A.5.3. *Approximating $Z^\mathsf{T} R_{\nu,\alpha} Z$.* Let us first give some definitions. For $\epsilon > 0$, Lemma 14 can be used to show that there exists some $C > 0$ such that

$$h_{\nu;\nu_0}(x) \le F_\epsilon(x) = C\min(x, 1-x)^{-1+2\epsilon}, \text{ for all } 0 < x < 1 \text{ and } \nu \in N_\epsilon. \tag{17}$$

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

LEMMA 16. *Let $2/5 < \epsilon < 1/2$. Then, the sequence*

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

of processes converges weakly in $L^\infty(N_\epsilon \times A)$ to

$$\mathrm{GP}\left(0, (\nu_1, \alpha_1; \nu_2, \alpha_2) \mapsto \frac{2\int_0^1 h_{\nu_1;\nu_0} h_{\nu_2;\nu_0}}{\int_0^1 h_{\nu_1;\nu_0} \int_0^1 h_{\nu_2;\nu_0}}\right) \tag{18}$$

which can be seen as a tight Borel probability measure. In particular, for all $\beta < 1/2$, we have

$$\sup_{\nu \in N_\epsilon, \alpha \in A} \left| \mathbb{M}_n(\nu, \alpha) - \int_0^1 g_\nu - \ln\left(\int_0^1 h_{\nu;\nu_0}\right)\right| = o_{\mathsf{P}}\left(n^{-\beta}\right).$$

PROOF. Use the notation $H\colon N_\epsilon \times A \mapsto \int_0^1 h_{\nu;\nu_0}$ for this proof.

Let $\mathbb{D}_\psi \subset L^\infty(N_\epsilon \times A)$ be the subset of positive functions bounded away from zero. One has $H \in \mathbb{D}_\psi$ and $(\nu, \alpha) \in N_\epsilon \times A \mapsto n^{-1-2(\nu-\nu_0)} \phi_0^{-1} Z^\top R_{\nu,\alpha}^{-1} Z$ lying also in $\mathbb{D}_\psi$ almost surely by continuity on the compact $N_\epsilon \times A$.

Furthermore, the mapping $\psi : g \in \mathbb{D}_\psi \subset L^\infty(N_\epsilon \times A) \mapsto \ln \circ g \in L^\infty(N_\epsilon \times A)$ is Fréchet-differentiable at H with $\psi'(H) : g \in L^\infty(N_\epsilon \times A) \mapsto g/H \in L^\infty(N_\epsilon \times A)$. The weak limit given by Lemma 17 is tight and hence separable, so we can use Theorem 3.9.4 from [25] to show that

$$\sqrt{n}\left(\ln\left(\frac{\phi_0^{-1} Z^\top R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}}\right) - \ln\left(\int_0^1 h_{\nu;\nu_0}\right)\right) \tag{19}$$

converges weakly to (18) in $L^\infty(N_\epsilon \times A)$. The tightness of the limit follows from the continuity of $\psi'(H)$. Conclude with Proposition 2, Lemma 15, and Slutsky's lemma [25, p. 32]. The last claim also follows from Slutsky's lemma. □

LEMMA 17. *Let $2/5 < \epsilon < 1/2$. The sequence*

$$(\nu, \alpha) \in N_\epsilon \times A \mapsto \sqrt{n}\left(\frac{\phi_0^{-1} Z^\top R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}} - \int_0^1 h_{\nu;\nu_0}\right)$$

of processes converges weakly in $L^\infty(N_\epsilon \times A)$ to

$$\mathrm{GP}\left(0, (\nu_1, \alpha_1; \nu_2, \alpha_2) \mapsto 2\int_0^1 h_{\nu_1;\nu_0} h_{\nu_2;\nu_0}\right),$$

which can be seen as a tight Borel probability measure.

PROOF. Let $\nu \in N_\epsilon$ and $\alpha \in A$. We have:

$$\sqrt{n}\left(\frac{\phi_0^{-1} Z^\top R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}} - \int_0^1 h_{\nu;\nu_0}\right)$$

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

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

with $B_{m,n} = U_{m,n}^2 - 1$. Then a classical Borel-Cantelli argument [see, e.g., 28, Lemma 4] shows that $\sup_{0\le m\le n-1} U_{m,n}^2 = o_{\mathsf{P}}(n^\delta)$ for every $\delta > 0$, so

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

thanks to Lemma 10 and

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

thanks to Lemma 22. Using the continuous mapping theorem for the linear isometry $\psi\colon L^\infty(N_\epsilon) \to L^\infty(N_\epsilon\times A)$ mapping $g\in L^\infty(N_\epsilon)$ to the function $(\nu,\alpha)\in N_\epsilon\times A \mapsto g(\nu)$ makes it possible to rephrase the convergence given by Lemma 25 in $L^\infty(N_\epsilon\times A)$. (The limit is a tight and hence separable measure.) Then, use Lemma 20 and Slutsky's lemma [25, p. 32] to conclude. □

LEMMA 18. *Let $0<\epsilon<1/2$, we have*

$$\sup_{\nu\in N_\epsilon,\alpha\in A}\left|\mathbb{M}_n(\nu,\alpha) - \int_0^1 g_\nu - \ln\left(\int_0^1 h_{\nu;\nu_0}\right)\right| = o_{\mathsf{P}}(1).$$

PROOF. First we have $\sup_{\nu\in N_\epsilon}\left|n^{-1}\sum_{m=1}^{n-1} B_{m,n}h_{\nu;\nu_0}(m/n)\right| = o_{\mathsf{P}}(1)$ using Lemma 26, Lemma 27, and [25, Theorem 1.5.4, Theorem 1.5.7, and Lemma 1.10.2]. Then, proceed as for Lemma 17 to get

$$\sup_{\nu\in N_\epsilon,\ \alpha\in A}\left|\frac{\phi_0^{-1}Z^{\mathsf{T}}R_{\nu,\alpha}^{-1}Z}{n^{1+2(\nu-\nu_0)}} - \int_0^1 h_{\nu;\nu_0}\right| = o_{\mathsf{P}}(1).$$

Conclude using Proposition 2, Lemma 15, and the continuous mapping theorem with ψ used in the proof of Lemma 16 since differentiability implies continuity at H. □

LEMMA 19. *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 (8) along with the fact that the Hurwitz Zeta function verifies

$$\frac{\partial\zeta_H}{\partial x}(\alpha;\,x) = -\alpha\zeta_H(\alpha+1;\,x),\quad \text{for } x>0, \text{ and } \alpha>1, \tag{20}$$

and has the representation

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

where Γ is the classical Gamma function [see, e.g., 17]. So, for $x\in(0,\,1)$, we have

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

and

$$\frac{\partial \gamma}{\partial x}(2\nu+1;\,x) = \frac{1}{\Gamma(2\nu+1)}\int_0^{+\infty}\frac{t^{2\nu+1}(e^{-t(1-x)}-e^{-tx})}{1-e^{-t}}\mathrm{d}t.$$

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

$$\gamma(2\nu+1;\,x)\frac{\partial \gamma}{\partial x}(2\nu_0+1;\,x) - \gamma(2\nu_0+1;\,x)\frac{\partial \gamma}{\partial x}(2\nu+1;\,x)$$

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

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:

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

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

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 20. *Let $\epsilon > 0$, we have*

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

uniformly in $\nu \in N_\epsilon$.

PROOF. For $\nu \in N_\epsilon$, the inequality (17) shows that $h_{\nu;\nu_0}$ is integrable on $(0, 1)$. Write $l = \lfloor (n-1)/2 \rfloor$. Using Lemma 19 and the symmetry with respect to $1/2$, we have:

$$\int_{1/n}^{(l+1)/n} h_{\nu;\nu_0} \le \frac{1}{n}\sum_{m=1}^{l} h_{\nu;\nu_0}(m/n) \le \int_0^{l/n} h_{\nu;\nu_0}, \tag{21}$$

when $\nu \le \nu_0$ and the reversed inequality when $\nu \ge \nu_0$. Moreover, using again (17) yields

$$\int_0^{1/n} h_{\nu;\nu_0} \le \int_0^{1/n} F_\epsilon = \mathcal{O}(n^{-2\epsilon}), \quad \text{uniformly in } \nu \in N_\epsilon. \tag{22}$$

The remainders $\int_{l/n}^{1/2} h_{\nu;\nu_0}$ and $\int_{1/2}^{(l+1)/n} h_{\nu;\nu_0}$ are $\mathcal{O}\left(n^{-1}\right)$ uniformly in $\nu \in N_\epsilon$ since the mapping $(\nu,\,x) \mapsto h_{\nu;\nu_0}(x)$ is bounded on $N_\epsilon \times [1/4,\,3/4]$ as it inherits the smoothness of γ.

Therefore, we have

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

uniformly. Use the symmetry of $h_{\nu;\nu_0}$ to conclude. The eventual term $h_{\nu;\nu_0}(1/2)$ for $m = n/2$ and n even is uniformly bounded on the compact N_ϵ. □

LEMMA 21. *Let $1 \le m \le \lfloor n/2 \rfloor$, we have*

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

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

PROOF. Lemma 13 gives

$$\frac{\lambda_{m,n}^{(0)}}{n^{2(\nu-\nu_0)}\lambda_{m,n}} = \frac{\sum_{j\in\mathbb{Z}} \frac{1+u_{n,m,j}(\nu_0,\alpha_0)}{|j+m/n|^{2\nu_0+1}}}{\sum_{j\in\mathbb{Z}} \frac{1+u_{n,m,j}(\nu,\alpha)}{|j+m/n|^{2\nu+1}}} = \frac{\gamma\left(2\nu_0+1;\, m/n\right)\left(1+\mathcal{O}\left(m^{-2}\right)\right)}{\gamma\left(2\nu+1;\, m/n\right)\left(1+\mathcal{O}\left(m^{-2}\right)\right)},$$

with uniform big-$\mathcal{O}$s. The desired result follows. □

LEMMA 22. *Let $0<\epsilon<1/2$, we have*

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

PROOF. Let $1 \le m \le \lfloor n/2 \rfloor$, $\nu \in N_\epsilon$, $\alpha \in A$, and $\beta \in (0,1)$. Using Lemma 21, we have

$$\frac{1}{n}\sum_{m=1}^{\lfloor n/2 \rfloor}\left|\frac{\lambda_{m,n}^{(0)}}{n^{2(\nu-\nu_0)}\lambda_{m,n}} - h_{\nu;\nu_0}\left(\frac{m}{n}\right)\right| = \frac{1}{n}\sum_{m=1}^{\lfloor n/2 \rfloor}\mathcal{O}\left(m^{-2}\right) h_{\nu;\nu_0}\left(\frac{m}{n}\right)$$

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

with

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

uniformly in $\nu \in N_\epsilon$ and $\alpha \in A$ thanks to (17) since $n^{-1}\sum_{m=1}^{\lfloor n/2 \rfloor} F_\epsilon(m/n) = \mathcal{O}(1)$. Moreover, using also F_ϵ, we have

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

uniformly in $\nu \in N_\epsilon$ and $\alpha \in A$. Using (13) and the symmetry of $h_{\nu;\nu_0}$ and taking $\beta = 3/8$ then gives the desired result. □

LEMMA 23. *Let $1/4<\epsilon<1/2$. The family $\mathcal{F}_\epsilon$ of functions equiped with the envelope F_ϵ defined by (17) verifies the uniform entropy condition [25, Section 2.5.1].*

PROOF. For $x \in (0,1)$ and $\nu \in N$, write $\gamma\left(2\nu+1;\, x\right) = \gamma_\uparrow\left(2\nu+1;\, x\right) + \gamma_\downarrow\left(2\nu+1;\, x\right)$, with

$$\gamma_\downarrow\left(2\nu+1;\, x\right) = \sum_{j=1}^{+\infty}(j+x)^{-2\nu-1} + \sum_{j=1}^{+\infty}(j+1-x)^{-2\nu-1},$$

and $\gamma_\uparrow(2\nu+1;x) = x^{-2\nu-1} + (1-x)^{-2\nu-1}$. Let $h_\uparrow(\nu;x) = \gamma(2\nu_0+1,x)/\gamma_\downarrow(2\nu+1,x)$ and $h_\downarrow(\nu;x) = \gamma(2\nu_0+1,x)/\gamma_\uparrow(2\nu+1,x)$.[2] The families $\mathcal{F}_\epsilon^\downarrow = \{h_\downarrow(\nu;\cdot),\, \nu\in N_\epsilon\}$ and $\mathcal{F}_\epsilon^\uparrow = \{1/h_\uparrow(\nu;\cdot),\, \nu\in N_\epsilon\}$ of functions are non-increasing with respect to the parameter ν so they are VC-subgraph classes. Indeed, let $(x_1,y_1),(x_2,y_2)\in(0,1)\times\mathbb{R}$, there cannot be two functions f and g in one of these families 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 $f\leq g$.

Equip $\mathcal{F}_\epsilon^\downarrow$ and $\mathcal{F}_\epsilon^\uparrow$ respectively with the envelopes F_ϵ (by increasing eventually the constant C in (17)) and $F_\epsilon^\uparrow\colon x\in(0,1)\mapsto C_2\min(x,1-x)^{1+2\nu_0}$, for some constant $C_2>0$. Theorem 2.6.7 from [25] shows that these families satisfy the uniform entropy condition.

Consider $\varsigma\colon x,y\in(0,+\infty)\mapsto\left(x^{-1}+y\right)^{-1}$. It holds that $\left|\frac{\partial\varsigma}{\partial x}(x,y)\right|\leq 1$ and $\left|\frac{\partial\varsigma}{\partial y}(x,y)\right| = \varsigma^2(x,y)$. Observe that $\varsigma(h_\downarrow(\nu_1;\cdot),1/h_\uparrow(\nu_2;\cdot))\lesssim F_\epsilon$, for $\nu_1,\nu_2\in N_\epsilon$. Consequently, for $\nu_1,\nu_2,\nu_3,\nu_4\in N_\epsilon$ and $x\in(0,1)$, we have:

$$\begin{aligned}&\left(\varsigma(h_\downarrow(\nu_1;x),1/h_\uparrow(\nu_3;x)) - \varsigma(h_\downarrow(\nu_2;x),1/h_\uparrow(\nu_4;x))\right)^2\\ &\lesssim (h_\downarrow(\nu_1;x)-h_\downarrow(\nu_2;x))^2 + F_\epsilon^4(x)\left(\frac{1}{h_\uparrow(\nu_3;x)} - \frac{1}{h_\uparrow(\nu_4;x)}\right)^2.\end{aligned}$$

Observe that $\varsigma(h_\downarrow(\nu;\cdot),1/h_\uparrow(\nu;\cdot)) = h_{\nu;\nu_0}$ and use Theorem 2.10.20 from [25] to conclude that the family

$$\mathcal{F}_\epsilon^{(\pi)} = \{\varsigma(h_\downarrow(\nu_1;\cdot),1/h_\uparrow(\nu_2;\cdot))-1,\ \nu_1,\nu_2\in N_\epsilon\}\quad(\text{note that } h_{\nu_0;\nu_0}=1)$$

with envelope $F_\epsilon^{(\pi)} = 2\sqrt{F_\epsilon^2 + F_\epsilon^4(F_\epsilon^\uparrow)^2}$ satisfy the uniform entropy condition. Concluding the proof is straightforward since $\mathcal{F}_\epsilon\subset\mathcal{F}_\epsilon^{(\pi)}+1$ and $F_\epsilon^{(\pi)}\lesssim F_\epsilon$. □

LEMMA 24. *For all $\epsilon>1/4$, 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 $\nu_1,\nu_2\in N_\epsilon$.

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

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

and

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

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

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

[2] The symbols $\downarrow$ and $\uparrow$ account for the monotonicity with respect to ν for fixed x.

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

$\square$

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

LEMMA 25. *Let $1/4 < \epsilon < 1/2$. Then, the sequence*

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

of processes converges weakly in $L^\infty(N_\epsilon)$ to

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

which can be seen as a tight Borel probability measure.

PROOF. Let $2 < \alpha < 1/(1-2\epsilon)$. It holds that $F_\epsilon \in L^\alpha(0,1) \subset L^2(0,1)$. Moreover, Lemma 23 shows that $\mathcal{F}_\epsilon$ satisfies the uniform entropy condition [25, Section 2.5.1].

Let us show that $(\mathcal{F}_\epsilon, \|\cdot\|_{L^2(0,1)})$ is totally bounded. Use the shortcut $Q_n = n^{-1}\delta_{1/2} + n^{-1}\sum_{m=1}^{n-1}\delta_{m/n}$. Since $\epsilon > 1/4$, then $\int F_\epsilon^2 \mathrm{d}Q_n$ is bounded uniformly in n by, say, M^2. The uniform entropy condition implies that $\mathcal{F}_\epsilon$ is totally bounded for the $L^2(Q_n)$-norm for any n. Let $\mathcal{G}_n$ be an $(M\delta)$-internal covering, for $\delta > 0$. Lemma 24 makes it possible to choose n such that

$$\sup_{g1, g2 \in \mathcal{F}_\epsilon} \left| \int (g_1 - g_2)^2 \mathrm{d}Q_n - \int_0^1 (g_1 - g_2)^2 \right| \leq \delta^2.$$

Therefore, $\mathcal{G}_n$ is a $(\delta\sqrt{M^2+1})$-covering of $(\mathcal{F}_\epsilon, \|\cdot\|_{L^2(0,1)})$.

With $Y_{m,n}\colon g \in (\mathcal{F}_\epsilon, \|\cdot\|_{L^2(0,1)}) \mapsto n^{-1/2} B_{m,n} g(m/n)$, the measurability conditions from [25, p. 205] are met since the suprema can be replaced by ones on countable sets. Indeed, using the surjection $\varrho\colon \nu \in N_\epsilon \mapsto h_{\nu;\nu_0} \in \mathcal{F}_\epsilon$, the suprema on subsets of $\mathcal{F}_\epsilon \times \mathcal{F}_\epsilon$ are suprema on subsets of $(N_\epsilon \times N_\epsilon, \|\cdot\|_2)$, with $\|\cdot\|_2$ standing for the euclidean norm. A subset of a separable metric space is separable. The sample path continuity of $\nu \in N_\epsilon \mapsto Y_{m,n}(\varrho(\nu))$ is inherited from the continuity of $\nu \in N_\epsilon \mapsto h_{\nu;\nu_0}(x)$, for $0 < x < 1$.

Since $2 < \alpha < 1/(1-2\epsilon)$, we have $n^{-1}\sum_{m=1}^{n-1} F_\epsilon^\alpha(m/n) = \mathcal{O}(1)$ and, for $\delta > 0$, the Lindeberg condition on suprema holds:

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

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

$$\begin{aligned}
&\sup_{\|g_1-g_2\|_{L^2(0,1)}<\delta_n} \sum_{m=1}^{n-1} \mathsf{E}\left((Y_{m,n}(g_1) - Y_{m,n}(g_2))^2 \right) \quad (\text{with } g_1, g_2 \in \mathcal{F}_\epsilon) \\
&= \mathsf{E}\left(B_{1,2}^2\right) \sup_{\|g_1-g_2\|_{L^2(0,1)}<\delta_n} \frac{1}{n} \sum_{m=1}^{n-1} (g_1(m/n) - g_2(m/n))^2 \\
&= o(1) + \mathcal{O}(\delta_n^2) \to 0
\end{aligned}$$

thanks to Lemma 24.

Now, let us show the pointwise convergence of the sequence of covariance functions. For a fixed $\nu \in N_\epsilon$, the convergence $n^{-1}\sum_{m=1}^{n-1} h_{\nu;\nu_0}^2(m/n) \to \int_0^1 h_{\nu;\nu_0}^2$ is ensured using Lemma 19 and the same reasoning as in the proof of Lemma 20. This fact and Lemma 24 shows that

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

for fixed $g_1, g_2 \in \mathcal{F}_\epsilon$.

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

We can then conclude using Lemma 2.11.6 and Theorem 2.11.1 from van Der Vaart and Wellner [25], which also imply the tightness of the limit [see 25, Lemma 1.3.8 and Theorem 1.5.7]. The reformulation from $L^\infty(\mathcal{F}_\epsilon)$ to $L^\infty(N_\epsilon)$ is an application of the continuous mapping theorem. □

LEMMA 26. *Let* $\nu \in (\nu_0 - 1/2, b_\nu]$. *Then,* $g_n(\nu) = n^{-1}\sum_{m=1}^{n-1} B_{m,n} h_{\nu;\nu_0}(m/n) = o_{\mathsf{P}}(1)$.

PROOF. Let $\delta = \nu - \nu_0 + 1/2 > 0$. Then, (17) yields

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

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

LEMMA 27. *Let* $0 < \epsilon < 1/2$ *and define*

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

The sequence $(g_n)_{n\geq 2}$ *is asymptotically uniformly equicontinuous in probability for* $(x, y) \mapsto |x-y|$ *[see 25, p. 37].*

PROOF. Proceeding as in the proof of Lemma 14, it is straightforward to show that

$$\left| \frac{\partial \gamma}{\partial \alpha}(2\nu+1; x) \right| \lesssim -x^{-2\nu-1}\ln(x) \lesssim x^{-2(\nu+\delta)-1} \quad (\text{with the notation } \gamma(\alpha; x)),$$

holds uniformly in $x \in (0, 1/2]$ and $\nu \in N_\epsilon$, for any $\delta > 0$. With a slight abuse of notation, the latter fact and Lemma 14 yield:

$$\left| \frac{\partial h_{\nu_0}}{\partial \nu}(\nu; m/n) \right| \lesssim \left(\frac{n}{m}\right)^{1-2\epsilon+2\delta}, \tag{23}$$

uniformly in n, $1 \le m \le \lfloor n/2 \rfloor$, and $\nu \in N_\epsilon$.

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

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

$$\le \left(\frac{1}{n}\sum_{m=1}^{n-1} |B_{m,n}|^q\right)^{1/q} \cdot \left(\frac{1}{n}\sum_{m=1}^{n-1} \sup_{\nu \in N_\epsilon} \left|\frac{\partial h_{\nu_0}}{\partial \nu}(\nu; m/n)\right|^p\right)^{1/p} \cdot |\nu_1 - \nu_2|,$$

which is enough to prove asymptotic uniform equicontinuity using (23), 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.6. Proof of Theorem 4. Remember (see Section A.2 and Section A.3) that the $\lambda_{m,n}$s depend smoothly on ν and α. Thus, the function $\mathbb{L}_n$ is smooth for any realization and can be written as:

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

Expressions for some derivatives are given in the following. These expressions are cumbersome, but rough approximations will suffice: we only need to ensure the $\partial^p \lambda_{m,n}/\partial \nu^p$s do not grow too fast compared to $\lambda_{m,n}$. The first-order derivative with respect to ν writes:

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

Then, the second-order derivative with respect to ν writes:

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

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

Finally, the third-order derivative with respect to ν writes:

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

$$-\frac{\phi_0}{n\phi}\sum_{m=0}^{n-1} \lambda_{m,n}^{-4}\lambda_{m,n}^{(0)} \left(\frac{\partial^3 \lambda_{m,n}}{\partial \nu^3}\lambda_{m,n}^2 - 4\frac{\partial^2 \lambda_{m,n}}{\partial \nu^2}\frac{\partial \lambda_{m,n}}{\partial \nu}\lambda_{m,n} + 6\left(\frac{\partial \lambda_{m,n}}{\partial \nu}\right)^3\right) U_{m,n}^2.$$

Bounding all terms independently will suffice for our purposes. The necessary approximations are given by Lemma 10 and the following. Exceptionally, the arguments of the $\lambda_{m,n}$s are not dropped.

LEMMA 28. *Let $\delta > 0$, $0 \le m \le \lfloor n/2 \rfloor$, $\nu \in N$, $\alpha \in A$ and $p \in \{1, 2, 3\}$. We have uniformly:*

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

and

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

PROOF. We have

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

which equals $n^{-1}\lambda_{m,n}(\nu-\delta/2,\,\alpha)$, so Lemma 10 gives the result.[3] □

PROOF OF THEOREM 4. Remember that $A=\{\alpha_0\}$ and consider $\mathbb{L}_n$ as a function of $(\nu,\,\phi)$ only. We start by elucidating the behavior of the score function.

The model under study is—up to reformulations that do not impact the likelihood criterion—the particular case for $d=1$ of the framework introduced by Stein [22, Section 6.7], who gives the asymptotics of the Fischer information matrix. Note that we have $p(Z\,|\,\nu,\,\phi)\propto \exp\{-n\mathbb{L}(\nu,\,\phi)/2\}$, so the score function is $-n\nabla\mathbb{L}_n/2$. Stein [22, p. 191 and p. 196] shows that

$$\text{(24)}\qquad \operatorname{Cov}\left(\sqrt{n}\nabla\mathbb{L}_n(\nu_0,\,\phi_0)/2\right) = n^{-1}\operatorname{Cov}\left(-n\nabla\mathbb{L}_n(\nu_0,\,\phi_0)/2\right) = C_n + \mathcal{O}\left(n^{-\epsilon}\right)$$

for some $\epsilon>0$ and

$$C_n = \begin{pmatrix} 2\ln^2(n)+4\ln(n)\mathsf{E}\left(\psi_{\nu_0}(V)\right)+2\mathsf{E}\left(\psi_{\nu_0}^2(V)\right) & -\ln(n)\phi_0^{-1}-\mathsf{E}\left(\psi_{\nu_0}(V)\right)\phi_0^{-1} \\ -\ln(n)\phi_0^{-1}-\mathsf{E}\left(\psi_{\nu_0}(V)\right)\phi_0^{-1} & 2^{-1}\phi_0^{-2}\end{pmatrix},$$

where V is a random variable distributed uniformly on $(0,\,1)$. Define

$$A_n = \frac{2\phi_0}{\sqrt{\operatorname{Var}\left(\psi_{\nu_0}(V)\right)}}\begin{pmatrix} 2^{-1}\phi_0^{-1} & 0 \\ \ln(n)+\mathsf{E}\left(\psi_{\nu_0}(V)\right) & \sqrt{\operatorname{Var}\left(\psi_{\nu_0}(V)\right)}\end{pmatrix}.$$

One has $A_n^{\mathsf{T}} C_n A_n = 2I_2$ and thus

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

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

$$\mathbb{L}_{m,n}(\nu,\,\phi) = n^{-1}\left(\ln(\phi)+\ln\left(\lambda_{m,n}\right)+\frac{\phi_0\lambda_{m,n}^{(0)}U_{m,n}^2}{\phi\lambda_{m,n}}\right),$$

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

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

and

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

[3] Here, δ is taken small enough, and the lower bound of N is reduced.

using Lemma 10, Lemma 28, and the expressions given in the beginning of this section. Therefore,

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

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

$$\frac{\sqrt{n}}{2\sqrt{2}}A_n^{\mathsf{T}}\nabla\mathbb{L}_n(\nu_0,\,\phi_0) \rightsquigarrow \mathcal{N}\left(0,\,I_2\right).$$

Consider now the Hessian matrix. The identity (24) and the well-known link with the covariance matrix of the score show that

$$\frac{1}{2}\mathsf{E}\left(\nabla^2\mathbb{L}_n(\nu_0,\,\phi_0)\right) = \mathrm{Cov}\left(\sqrt{n}\nabla\mathbb{L}_n(\nu_0,\,\phi_0)/2\right) = C_n + \mathcal{O}\left(n^{-\epsilon}\right),$$

for some $\epsilon > 0$. Furthermore, using Lemma 10, Lemma 28, and the expressions given in the beginning of this section, it is straightforward to check that

$$\mathrm{Var}\left(\frac{\partial^2\mathbb{L}_n}{(\partial\phi)^p\,(\partial\nu)^{2-p}}(\nu_0,\,\phi_0)\right) = \mathcal{O}\left(n^{-2\epsilon}\right),$$

for $p \in \{0,\,1,\,2\}$ and some $\epsilon > 0$. Therefore, we have $\nabla^2\mathbb{L}_n(\nu_0,\,\phi_0)/2 - C_n = \mathcal{O}_{\mathsf{P}}\left(n^{-\epsilon}\right)$ for some $\epsilon > 0$ and thus $A_n^{\mathsf{T}}\nabla^2\mathbb{L}_n(\nu_0,\,\phi_0)A_n \to 4I_2$, in probability.

We are now left to bound the third derivatives uniformly locally around $(\nu_0,\,\phi_0)$. For $\epsilon > 0$ small enough, bounding the terms individually with Lemma 10 and Lemma 28 makes it straightforward to show that

$$\text{(25)}\qquad \mathsf{E}\left(\sup_{p\in\{0,1,2,3\},\,|\nu-\nu_0|\le\epsilon,\,|\phi-\phi_0|\le\epsilon}\left|\frac{\partial^3\mathbb{L}_n}{(\partial\nu)^p\,(\partial\phi)^{3-p}}(\nu,\,\phi)\right|\right) = \mathcal{O}\left(n^{5\epsilon}\right).$$

Theorem 3 shows that $(\widehat{\nu}_n,\,\widehat{\phi}_n) \in [\nu_0-\epsilon,\,\nu_0+\epsilon]\times[\phi_0-\epsilon,\,\phi_0+\epsilon]$ with high probability. On this event, we have:

$$0 = \nabla\mathbb{L}_n\left(\nu_0,\,\phi_0\right) + \nabla^2\mathbb{L}_n\left(\nu_0,\,\phi_0\right)\begin{pmatrix}\widehat{\nu}_n-\nu_0\\ \widehat{\phi}_n-\phi_0\end{pmatrix} + \mathcal{O}_{\mathsf{P}}\left(n^{5\epsilon}\left|\left|\begin{pmatrix}\widehat{\nu}_n-\nu_0\\ \widehat{\phi}_n-\phi_0\end{pmatrix}\right|\right|^2\right),$$

thanks to (25). Multiplying by A_n^{T} and using Theorem 3 again leads to

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

where the preceding $\mathcal{O}_{\mathsf{P}}$-term has been reformulated using a few algebraic manipulations. (The largest singular value of A_n is dominated by $\|A_n\|_\infty \lesssim \ln(n)$ as norms are equivalent.) Then, multiply by $\sqrt{2n}$ and use Slutsky's lemma to get the weak convergence of

$$\sqrt{2n}A_n^{-1}\begin{pmatrix}\widehat{\nu}_n-\nu_0\\ \widehat{\phi}_n-\phi_0\end{pmatrix} = -\left(A_n^{\mathsf{T}}\nabla^2\mathbb{L}_n\left(\nu_0,\,\phi_0\right)A_n/4 + o_{\mathsf{P}}(1)\right)^{-1}\frac{\sqrt{n}}{2\sqrt{2}}A_n^{\mathsf{T}}\nabla\mathbb{L}_n\left(\nu_0,\,\phi_0\right)$$

to $\mathcal{N}\left(0,\,I_2\right)$. This gives the result. □

A.7. Proofs of Theorem 5 and Theorem 6. The posterior mean does not depend on ϕ, so all derivations will be written with $\phi = 1$. 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. Also, we assume that $\phi_0 = 1$ without loss of generality and write $\omega_0 = (\nu_0, \alpha_0)$.

We avoid dealing with conditionally convergent series since we assume that $\nu_0 > 1/2$. Indeed, in this case, the coefficients of the expansion (9) are almost surely summable (take the expectation of the sum of the moduli using the Fubini–Tonelli theorem). Then, on the corresponding almost sure event: the expansion (9) converges uniformly absolutely, so the hypothesis of Proposition 1 are fulfilled. Consequently, the expansions (9) and (3) with $f = \xi$ are almost surely Fourier expansions in $L^2[0, 1]$.

The proofs will rely on using Parseval's identity. Let $\omega = (\nu, \alpha) \in (0, +\infty)^2$ and $j \in \mathbb{Z}$, we have

$$2\left|c_j(\xi - \widehat{\xi}_n)\right|^2 = \left(\frac{c_j(\omega)\sum_{j_1\in\mathbb{Z}\setminus\{0\}}\sqrt{c_{j+j_1 n}(\omega_0)}U_{1,|j+j_1 n|}}{\sum_{j_1\in\mathbb{Z}} c_{j+nj_1}(\omega)} - \frac{\sqrt{c_j(\omega_0)}U_{1,|j|}\sum_{j_1\in\mathbb{Z}\setminus\{0\}} c_{j+nj_1}(\omega)}{\sum_{j_1\in\mathbb{Z}} c_{j+nj_1}(\omega)}\right)^2 + \left(\frac{c_j(\omega)\sum_{j_1\in\mathbb{Z}\setminus\{0\}}\sqrt{c_{j+nj_1}(\omega_0)}U_{2,|j+nj_1|}\mathrm{sign}(j+nj_1)}{\sum_{j_1\in\mathbb{Z}} c_{j+nj_1}(\omega)} - \frac{\sqrt{c_j(\omega_0)}U_{2,|j|}\mathrm{sign}(j)\sum_{j_1\in\mathbb{Z}\setminus\{0\}} c_{j+nj_1}(\omega)}{\sum_{j_1\in\mathbb{Z}} c_{j+nj_1}(\omega)}\right)^2 \tag{26}$$

after a few algebraic manipulations. The expression (26) is a sum of two independent terms. If $j \in m + n\mathbb{Z}$ with $m \notin \{0, n/2\}$, then they are identically distributed and involve independent Gaussian variables. Lemma 30 and Lemma 31 deal with the cases $m = 0$ and $m = n/2$ for n even respectively.

PROOF OF THEOREM 5. Let $m \in [\![0, n-1]\!]$ such that $m \notin \{0, \ n/2\}$ and consider indexes $m + nj$, with $j \in \mathbb{Z}$. The previous discussion shows that 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

$$a_{m,j,n}(\nu) = c^2_{m+jn}(\omega)\frac{\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega_0) - c_{m+jn}(\omega_0)}{\left(\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega)\right)^2} + c_{m+jn}(\omega_0)\left(1 - \frac{c_{m+jn}(\omega)}{\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega)}\right)^2 \tag{27}$$

and the dependence of $a_{m,j,n}$ on the (fixed) parameter α removed for readability. Use the notation $l = \lfloor (n-1)/2 \rfloor$. For $1 \le m \le l$, Lemma 29 and Lemma 10 make it straightforward to show that

$$a_{m,j,n}(\nu) \lesssim (|j|\,n)^{-4\nu-2} m^{4\nu+2-2\nu_0-1} + (|j|\,n)^{-2\nu_0-1}, \quad \text{for } j \neq 0, \tag{28}$$

and

$$a_{m,0,n}(\nu) \lesssim n^{-2\nu_0-1} + m^{4\nu-2\nu_0+1}n^{-4\nu-2}, \tag{29}$$

uniformly in j and m. So, this yields

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

The first two statements then follow from Lemma 31, Lemma 30, the identity

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

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

For the last statement, let $1 \le m \le l$. A few algebraic steps lead to:

$$\mathsf{E}\left(\left| c_{m+jn}(\xi-\widehat{\xi}_n)\right|^2\right) = c_{m+jn}^2(\omega) \frac{\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega_0)}{\left(\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega)\right)^2}$$
$$+ c_{m+jn}(\omega_0)\left(1 - 2\frac{c_{m+jn}(\omega)}{\sum_{j_1\in\mathbb{Z}} c_{m+nj_1}(\omega)}\right),$$

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

$$\sum_{j\in\mathbb{Z}} \mathsf{E}\left(\left| c_{m+jn}(\xi-\widehat{\xi}_n)\right|^2\right)$$
$$= \left(1+\mathcal{O}(m^{-2})\right)\sum_{j\in\mathbb{Z}} |m+jn|^{-4\nu-2} \frac{\sum_{j_1\in\mathbb{Z}} |m+nj_1|^{-2\nu_0-1}}{\left(\sum_{j_1\in\mathbb{Z}} |m+nj_1|^{-2\nu-1}\right)^2}$$
$$+ |m+jn|^{-2\nu_0-1}\left(1 - 2\frac{|m+jn|^{-2\nu-1}}{\sum_{j_1\in\mathbb{Z}} |m+nj_1|^{-2\nu-1}}\right)$$
$$= \frac{\left(1+\mathcal{O}(m^{-2})\right)}{n^{2\nu_0+1}} \vartheta_{\nu;\nu_0}(m/n),$$

using Lemma 13 and after a few algebraic manipulations. Using the definition of γ, it is straightforward to show that

(32) $$\vartheta_{\nu;\nu_0}(x) \sim C_1 x^{4\nu-2\nu_0+1} + C_2$$

for some nonzero constants C_1, C_2, when $x \to 0$. Since $\vartheta_{\nu;\nu_0}$ is smooth and symmetric with respect to $1/2$, it is therefore integrable if $\nu > (\nu_0-1)/2$. Furthermore, if $\nu \ge (\nu_0 - 1/2)/2$, then (32) shows that $\vartheta_{\nu;\nu_0}$ has a finite limit in zero and therefore that

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

(If $(\nu_0-1)/2 < \nu < (\nu_0-1/2)/2$, then it is possible to show that the derivative of ϑ is negative in a neighborhood of zero, so (33) holds as well thanks to the standard monotonous approach.) Then, Lemma 31, Lemma 30, the identity (31), the Fubini-Tonelli thereom, and Parseval's identity gives

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

killing the $\mathcal{O}(m^{-2})$ term by splitting by how m compares with $\sqrt{n}$ and using (32) as in the proof of Lemma 22. □

LEMMA 29. *Let $\nu, \alpha > 0$, $0 \le m \le \lfloor n/2 \rfloor$, and $j \neq 0$. We have:*

$$c_{m+nj}(\nu, \alpha) \le 2^{2\nu+1} (n\,|j|)^{-2\nu-1}.$$

PROOF. Using the fact that $m \le n/2$ leads to:

$$c_{m+nj}(\nu, \alpha) \le (n\,(|j| - 1/2))^{-2\nu-1} \le 2^{2\nu+1} (n\,|j|)^{-2\nu-1}.$$

□

For $m \in \{0, n/2\}$ and $j \in \mathbb{Z}$, the two terms in (26) are not identically distributed. Moreover, for $q \in \{1, 2\}$ and $m \in \{0, n/2\}$, there are duplicates among the variables $\left\{U_{q,|m+nj|},\, j \in \mathbb{Z}\right\}$. Nevertheless, the two terms are sums of independent Gaussian variables so there exists χ^2_1 distributed variables $D_{m,j,n}$ and $B_{m,j,n}$ and functions $d_{m,j,n}$ and $b_{m,j,n}$ such that:

$$\left|c_{m+jn}(\xi - \widehat{\xi}_n)\right|^2 = d_{m,j,n}(\omega) D_{m,j,n} + b_{m,j,n}(\omega) B_{m,j,n}, \quad \text{for } \omega = (\nu,\, \alpha) \in (0, +\infty)^2.$$

The presence of duplicates makes the expressions of $d_{m,j,n}$ and $b_{m,j,n}$ are a bit more complex than (27). However, we have $\sum_{j_1 \in \mathbb{Z}} \sqrt{c_{m+nj_1}(\omega_0)} < +\infty$ since $\nu_0 > 1/2$ and we can bound $\sqrt{d_{m,j,n}(\omega)}$ and $\sqrt{b_{m,j,n}(\omega)}$ by

$$\frac{c_{m+nj}(\omega) \sum_{j_1 \in \mathbb{Z}\setminus\{j\}} \sqrt{c_{m+nj_1}(\omega_0)} + \sqrt{c_{m+nj}(\omega_0)} \sum_{j_1 \in \mathbb{Z}\setminus\{j\}} c_{m+nj_1}(\omega)}{\sum_{j_1 \in \mathbb{Z}} c_{m+nj_1}(\omega)}$$

corresponding to full redundancies among the variables appearing in the two terms of (26). Then, using the identity $\sqrt{c_j(\nu_0, \alpha_0)} = c_j\left(\frac{\nu_0 - 1/2}{2}, \alpha_0\right)$, Lemma 10, and Lemma 29 yield the following lemmata. The statements are made uniform with respect to ranges of regularity to be used in the proof of Theorem 6.

LEMMA 30. *Let $n \ge 1$ and $N \subset (0, +\infty)$ be a compact interval. There exists χ^2_2 distributed variables $\{C_{j,n},\ j \in \mathbb{Z}\}$ and $\epsilon > 0$ such that*

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

uniformly in $\nu \in N$.

LEMMA 31. *Let $n \ge 2$ be even and $N \subset (0, +\infty)$ be a compact interval. Then:*

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

The following lemma bounds the rate at which ν falls within the range $[\nu_0 - 1/2,\, b_\nu]$ of values giving reproducing kernel Hilbert spaces almost surely not containing ξ. It will be useful for proving Theorem 6.

LEMMA 32. *Let $\epsilon > 0$. With the notations of Theorem 6, we have:*

$$\mathsf{P}\left(\widehat{\nu_n} \le \nu_0 - 1/2 - \epsilon\right) \lesssim e^{-C\sqrt{n}},$$

for some $C > 0$.

PROOF. First, one has

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

Then, let $a_\nu \le \nu \le \nu_0 - 1/2 - \epsilon$, we have:

$$\begin{aligned}\mathbb{M}_n(\nu, \alpha) &= \mathcal{O}(1) + \ln\left(\frac{Z^\mathsf{T} R_{\nu,\alpha}^{-1} Z}{n^{1+2(\nu-\nu_0)}}\right) \quad \text{(Lemma 15)} \\ &\ge \mathcal{O}(1) + \ln\left(\frac{\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} U_{m,n}^2 \lambda_{m,n}^{(0)}/\lambda_{m,n}}{n^{1+2(\nu-\nu_0)}}\right) \\ &= \mathcal{O}(1) + \ln\left(\frac{\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} U_{m,n}^2 m^{2(\nu-\nu_0)}}{n^{1+2(\nu-\nu_0)}}\right) \quad \text{(Lemma 10)} \\ &= \mathcal{O}(1) + \ln\left(\frac{1}{n}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} U_{m,n}^2 \left(\frac{m}{n}\right)^{2(\nu-\nu_0)}\right) \\ &\ge \mathcal{O}(1) + \ln\left(\frac{1}{n}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} U_{m,n}^2 \left(\frac{m}{n}\right)^{-1-2\epsilon}\right) \\ &= \mathcal{O}(1) + 2\epsilon\ln(n) + \ln\left(\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} U_{m,n}^2 m^{-1-2\epsilon}\right) \\ &\ge \mathcal{O}(1) + 2\epsilon\ln(n) + \ln\left(\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} U_{m,n}^2 \lfloor\sqrt{n}\rfloor^{-1-2\epsilon}\right) \\ &\ge \mathcal{O}(1) + \epsilon\ln(n) + \ln\left(\frac{1}{\lfloor\sqrt{n}\rfloor}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} U_{m,n}^2\right) \\ &\ge \mathcal{O}(1) + \epsilon\ln(n) + \frac{1}{\lfloor\sqrt{n}\rfloor}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} \ln\left(U_{m,n}^2\right) \quad \text{(Jensen inequality)}\end{aligned}$$

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

$$\mathsf{P}\left(-\frac{1}{\lfloor\sqrt{n}\rfloor}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor} \ln(U_{m,n}^2) \ge \delta\right) = \mathsf{P}\left(e^{-\frac{t}{\lfloor\sqrt{n}\rfloor}\sum_{m=1}^{\lfloor\sqrt{n}\rfloor}\ln(U_{m,n}^2)} \ge e^{t\delta}\right)$$

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

with $t = 1/4$ and $\mathsf{E}\left(|U_{1,1}|^{-1/2}\right) < +\infty$. This gives the desired convergence rate if δ is high enough.

Furthermore, Lemma 10 and (13) yield

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

and we have

$$
\mathsf{P}\left(\ln\left(n^{-1}\sum_{m=0}^{n-1} U_{m,n}^2\right) \geq \delta\right) = \mathsf{P}\left(n^{-1}\sum_{m=0}^{n-1} U_{m,n}^2 \geq e^{\delta}\right) \leq e^{-C_2 n},
$$

for some $C_2 > 0$ if $\delta > 0$ is high enough, using also a Chernoff bound argument. Now, putting all the pieces together yields:

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

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

giving the result thanks to the pigeonhole principle. □

PROOF OF THEOREM 6. Let $\epsilon > 0$ and $1 \leq m \leq l = \lfloor (n-1)/2 \rfloor$ and use the notation (27). The functions $a_{m,j,n}$ are smooth, but computing the derivatives of the $a_{m,j,n}$s leads cumbersome formulas. However, for $\delta > 0$ and $\nu \in [\nu_0 - 1/2 - \epsilon,\, b_\nu]$, the uniform inequality

$$
\left|\frac{\partial c_j}{\partial \nu}(\nu,\, \alpha)\right| \lesssim c_j(\nu - \delta,\, \alpha), \quad \text{for } j \in \mathbb{Z},
$$

Lemma 10, and Lemma 29 make it possible to bound them by

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

and, for $j \neq 0$, by

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

uniformly in $1 \leq m \leq l$, $j \neq 0$ and $\nu \in [\nu_0 - 1/2 - \epsilon,\, b_\nu]$. Then,

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

for δ and ϵ small enough and using the above inequalities and Theorem 3. Therefore, Lemma 31, Lemma 30, the identity (31), and the Fubini-Tonelli theorem shows that

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

Furthermore, using again the Fubini-Tonelli theorem yields

$$\begin{aligned}&\mathsf{E}\left(\sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \left|c_{m+jn}(\xi - \widehat{\xi}_n)\right|^2 \mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\epsilon}\right) \\ &= \sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \mathsf{E}\left(a_{m,j,n}(\widehat{\nu}_n) A_{m,j,n} \mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\epsilon}\right) \\ &\leq \sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \sup_{a_\nu\leq\nu\leq\nu_0-1/2-\epsilon} a_{m,j,n}(\nu) \mathsf{E}\left(A_{m,j,n} \mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\epsilon}\right) \\ &\leq \sqrt{\mathsf{E}\left(A_{1,0,1}^2\right)} \sqrt{\mathsf{E}\left(\mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\epsilon}\right)} \sum_{m=1}^{l} \sum_{j\in\mathbb{Z}} \sup_{a_\nu\leq\nu\leq\nu_0-1/2-\epsilon} a_{m,j,n}(\nu) \\ &\leq \sqrt{\mathsf{E}\left(A_{1,0,1}^2\right)} \sqrt{\mathsf{E}\left(\mathbb{1}_{\widehat{\nu}_n \leq \nu_0-1/2-\epsilon}\right)}\, n^{\alpha} \quad \text{for some } \alpha \text{ given by (28) and (29)} \\ &= o\left(n^{-2\nu_0}\right),\end{aligned}$$

using Lemma 32. Then, the sum for $j \equiv 0 \pmod{n}$ can be bounded similarly using Lemma 30 and the sum for $j \equiv n/2 \pmod{n}$ is controlled by Lemma 31 for n even.

Finally, the previous reasoning is easily applied to bound

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

and the desired result follows. □

A.8. Proofs of Section 5. Note that the finiteness of $\nu_0(f)$ is assumed so that f is necessarily nonzero. Consequently, the data Z is ultimately nonzero under the observation model (2) since f is continuous. Furthermore, we assume that $\nu_0(f) > 1/2$, so $f \in H^{\beta}[0, 1]$ for some $\beta > 1$. Consequently, the Sobolev embedding theorem implies that f has Hölder regularity strictly greater than $1/2$. Hence, f has absolutely summable Fourier coefficients.

The proofs are based on the observation that

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

using (14) and elements from Section A.2.

PROOF OF PROPOSITION 7. 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 $\epsilon > 0$, $a_\nu \le \nu \le \nu_0(f) - \epsilon$, $\alpha \in A$, and $p \in \mathbb{Z}$ such that $c_p(f) \neq 0$. Then, Proposition 2 and Lemma 15 gives

$$\begin{aligned}
\mathbb{M}_n^f(\nu,\alpha) &= 2(\nu_0(f)-\nu)\ln(n) + \mathcal{O}(1) + \ln\left(\sum_{m=0}^{n-1} \frac{\left|\sum_{j\in m+n\mathbb{Z}} c_j(f)\right|^2}{\sum_{j\in m+n\mathbb{Z}} c_j(\nu,\alpha)}\right) \\
&\ge 2\epsilon\ln(n) + \mathcal{O}(1) + \ln\left(\frac{\left|\sum_{j\in p+n\mathbb{Z}} c_j(f)\right|^2}{\sum_{j\in p+n\mathbb{Z}} c_j(\nu,\alpha)}\right) \quad (\text{consider } m = p \mod n) \\
&= 2\epsilon\ln(n) + \mathcal{O}(1)
\end{aligned}$$

uniformly since $\sum_{j\in p+n\mathbb{Z}} c_j(f) \to c_p(f)$ and by Lemma 10. (If $p < 0$, then use the symmetry of the $c_j(\nu,\alpha)$s.)

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

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

since $f \in H^{\beta+1/2}[0, 1]$ for $\beta = \nu_0(f) - \epsilon/2$. Indeed, this last Sobolev space is norm-equivalent to the reproducing kernel Hilbert space attached to the covariance function for $\nu = \nu_0(f) - \epsilon/2$ and the quadratic term $Z^{\mathsf{T}} R_{\nu,\alpha}^{-1} Z$ is the squared norm of a projection of f [see, e.g., 27, Theorem 13.1]. This completes the proof. □

PROOF OF PROPOSITION 9. Let $\nu > \nu_0(f)$ and $\alpha > 0$. Then, Proposition 2 and Lemma 15 yield:

$$\mathbb{M}_n^f(\nu,\alpha) = \int_0^1 g_\nu + \mathcal{O}\left(\frac{\ln(n)}{n}\right) + \ln\left(n^{2(\nu_0(f)-\nu)} \sum_{m=0}^{n-1} \frac{\left|\sum_{j\in m+n\mathbb{Z}} c_j(f)\right|^2}{\sum_{j\in m+n\mathbb{Z}} c_j(\nu,\alpha)}\right).$$

Focus now on the term inside the logarithm. Write $l = \lfloor (n-1)/2 \rfloor$. We have $c_j(f) = c_{-j}(f)$, and so $\sum_{j\in m+n\mathbb{Z}} c_j(f) = \sum_{j\in n-m+n\mathbb{Z}} c_j(f)$, for $m \in [\![1, l]\!]$. As usual, consider the l pairs $(m, n-m)$ and isolate the terms corresponding to $m = 0$ and eventually $m = n/2$, for n

even. For these, we have $\sum_{j\in n\mathbb{Z}} c_j(f) \to c_0(f)$ and $\sum_{j\in n/2+n\mathbb{Z}} c_j(f) \lesssim n^{-\nu_0(f)-1}$ by proceeding similarly as for Lemma 10. Then, use Lemma 10 and Lemma 13 to get:

$$
\begin{aligned}
&n^{2(\nu_0(f)-\nu)} \sum_{m=0}^{n-1} \frac{\left|\sum_{j\in m+n\mathbb{Z}} c_j(f)\right|^2}{\sum_{j\in m+n\mathbb{Z}} c_j(\nu,\alpha)} \\
&= \mathcal{O}\left(n^{2(\nu_0(f)-\nu)}\right) + \mathcal{O}\left(n^{-1}\right) + \frac{2}{n}\sum_{m=1}^{l} \frac{\left(1+\mathcal{O}(m^{-2})\right)\gamma^2\left(\nu_0(f)+1; m/n\right)}{\gamma\left(2\nu+1; m/n\right)} \\
&= \mathcal{O}\left(n^{2(\nu_0(f)-\nu)}\right) + \mathcal{O}\left(n^{-1}\right) + \mathcal{O}\left(n^{\nu_0(f)-\nu}\right) + \frac{2}{n}\sum_{m=1}^{l} \frac{\gamma^2\left(\nu_0(f)+1; m/n\right)}{\gamma\left(2\nu+1; m/n\right)}
\end{aligned}
$$

bounding the $\mathcal{O}(m^{-2})$ term by splitting by how m compares with $\sqrt{n}$ as in the proof of Lemma 22. To conclude, observe that the function $\gamma^2\left(\nu_0(f)+1;\cdot\right)/\gamma\left(2\nu+1;\cdot\right)$ is non-increasing in a neighborhood of zero if $\nu < \nu_0(f)+1/2$ and has a finite limit otherwise. □

Acknowledgments. Part of this article was written while the author was a PhD student in the Laboratoire des signaux et systèmes, Université Paris-Saclay, CNRS, CentraleSupélec, 91190, Gif-sur-Yvette, France.

REFERENCES

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

[2] BACHOC, F. (2014). Asymptotic analysis of the role of spatial sampling for covariance parameter estimation of Gaussian processes. *Journal of Multivariate Analysis* **125** 1-35.

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

[4] BROCKWELL, P. J. and DAVIS, R. A. (1987). *Time series: theory and methods*. Springer science & business media.

[5] CHEN, Y., OWHADI, H. and STUART, A. (2021). Consistency of empirical Bayes and kernel flow for hierarchical parameter estimation. *Mathematics of Computation* **90** 2527–2578.

[6] CRAVEN, P. and WAHBA, G. (1979). Smoothing Noisy Data with Spline Functions. Estimating the Correct Degree of Smoothing by the Method of Generalized Cross-Validation. *Numerische Mathematik*.

[7] GNEITING, T. (2011). Strictly and non-strictly positive definite functions on spheres. *Bernoulli* **19** 1327-1349.

[8] KARVONEN, T. (2022). Asymptotic Bounds for Smoothness Parameter Estimates in Gaussian Process Regression. *arXiv preprint arXiv:2203.05400v2*.

[9] KARVONEN, T. and OATES, C. J. (2022). Maximum Likelihood Estimation in Gaussian Process Regression is Ill-Posed. *arXiv preprint arXiv:2203.09179*.

[10] KARVONEN, T., WYNNE, G., TRONARP, F., OATES, C. and SARKKA, S. (2020). Maximum likelihood estimation and uncertainty quantification for Gaussian process approximation of deterministic functions. *SIAM/ASA Journal on Uncertainty Quantification* **8** 926–958.

[11] KNAPIK, B. T., SZABÓ, B. T., VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2016). Bayes procedures for adaptive inference in inverse problems for the white noise model. *Probability Theory and Related Fields* **164** 771–813.

[12] LI, C. (2020). Bayesian fixed-domain asymptotics: Bernstein-von Mises theorem for covariance parameters in a Gaussian process model. *arXiv preprint arXiv:2010.02126*.

[13] LOH, W. L. (2006). Fixed-domain asymptotics for a subclass of Matérn-type Gaussian random fields. *The Annals of Statistics* **33**.

[14] MARDIA, K. V. and MARSHALL, R. J. (1984). Maximum Likelihood Estimation of Models for Residual Covariance in Spatial Regression. *Biometrika* **71** 135–146.

[15] MATÉRN, B. (1986). *Spatial Variation*. Springer-Verlag New York.

[16] MATHERON, G. (1971). The theory of regionalized variables and its applications Technical Report No. Les cahiers du CMM de Fontainebleau, Fasc. 5, Ecole des Mines de Paris.

[17] POSTNIKOV, A. G. (1988). *Introduction to analytic number theory* **68**. American Mathematical Soc.

[18] RASMUSSEN, C. E. and WILLIAMS, C. K. I. (2006). *Gaussian Processes for Machine Learning*. *Adaptive Computation and Machine Learning*. MIT Press, Cambridge, MA, USA.

[19] SANTNER, T., WILLIAMS, B. and NOTZ, W. (2003). *The Design and Analysis Computer Experiments*.

[20] STEIN, M. L. (1993a). Spline Smoothing with an Estimated Order Parameter. *The Annals of Statistics* **21** 1522–1544.

[21] STEIN, M. L. (1993b). A simple condition for asymptotic optimality of linear predictions of random fields. *Statistics and Probability Letters* **17** 399-404.

[22] STEIN, M. L. (1999). *Interpolation of spatial data*. *Springer Series in Statistics*. Springer-Verlag, New York. Some theory for Kriging.

[23] SZABÓ, B., VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2015). Frequentist coverage of adaptive nonparametric Bayesian credible sets. *The Annals of Statistics* **43** 1391–1428.

[24] VAN DER VAART, A. W. (1996). Maximum likelihood estimation under a spatial sampling scheme. *The Annals of Statistics* 2049 – 2057.

[25] VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak convergence and empirical processes: with applications to statistics*. Springer Science & Business Media.

[26] WANG, W. and JING, B. Y. (2022). Gaussian process regression: Optimality, robustness, and relationship with kernel ridge regression. *Journal of Machine Learning Research* **23** 1–67.

[27] WENDLAND, H. (2004). *Scattered data approximation* **17**. Cambridge university press.

[28] YING, Z. (1991). Asymptotic properties of a maximum likelihood estimator with data from a Gaussian process. *Journal of Multivariate analysis* **36** 280–296.

[29] YING, Z. (1993). Maximum likelihood estimation of parameters under a spatial sampling scheme. *The Annals of Statistics* 1567–1590.

[30] ZHANG, H. (2004). Inconsistent Estimation and Asymptotically Equal Interpolations in Model-Based Geostatistics. *Journal of the American Statistical Association* **99** 250–261.