The Annals of Statistics
2010, Vol. 38, No. 1, 181–214
DOI: 10.1214/09-AOS705

ASYMPTOTIC EQUIVALENCE OF SPECTRAL DENSITY ESTIMATION AND GAUSSIAN WHITE NOISE

BY GEORGI K. GOLUBEV, MICHAEL NUSSBAUM[1] AND HARRISON H. ZHOU[1]

Université de Provence, Cornell University and Yale University

We consider the statistical experiment given by a sample $y(1), \dots, y(n)$ of a stationary Gaussian process with an unknown smooth spectral density f. Asymptotic equivalence, in the sense of Le Cam's deficiency Δ-distance, to two Gaussian experiments with simpler structure is established. The first one is given by independent zero mean Gaussians with variance approximately $f(\omega_i)$, where ω_i is a uniform grid of points in $(-\pi, \pi)$ (nonparametric Gaussian scale regression). This approximation is closely related to well-known asymptotic independence results for the periodogram and corresponding inference methods. The second asymptotic equivalence is to a Gaussian white noise model where the drift function is the log-spectral density. This represents the step from a Gaussian scale model to a location model, and also has a counterpart in established inference methods, that is, log-periodogram regression. The problem of simple explicit equivalence maps (Markov kernels), allowing to directly carry over inference, appears in this context but is not solved here.

1. Introduction and main results. Estimation of the spectral density $f(\omega)$, $\omega \in [-\pi, \pi]$, of a stationary process is an important and traditional problem of mathematical statistics. We observe a sample $y^{(n)} = (y(1), \dots, y(n))'$ from a real Gaussian stationary sequence $y(t)$ with $\mathrm{E}y(t) = 0$ and autocovariance function $\gamma(h) = \mathrm{E}y(t)y(t+h)$. Consider the spectral density, defined on $[-\pi, \pi]$ by

$$f(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) \exp(ih\omega), \tag{1.1}$$

where it is assumed that $\sum_{h=-\infty}^{\infty} \gamma^2(h) < \infty$. Let Γ_n be the $n \times n$ Toeplitz covariance matrix associated with $\gamma(\cdot)$, that is, the matrix with entries

$$(\Gamma_n)_{j,k} = \gamma(k-j) = \int_{-\pi}^{\pi} \exp\bigl(i(k-j)\omega\bigr) f(\omega)\, d\omega, \qquad j, k = 1, \dots, n. \tag{1.2}$$

Received January 2005; revised March 2009.
[1]Supported in part by NSF Grants DMS-03-06497 and DMS-08-05632.
AMS 2000 subject classifications. 62G07, 62G20.
Key words and phrases. Stationary Gaussian process, spectral density, Sobolev classes, Le Cam distance, asymptotic equivalence, Whittle likelihood, log-periodogram regression, nonparametric Gaussian scale model, signal in Gaussian white noise.

Write $\Gamma_n(f)$ for the covariance matrix corresponding to spectral density f and note that $y^{(n)}$ has a multivariate normal distribution $N_n(0, \Gamma_n(f))$. Let Σ be a nonparametric set of spectral densities to be described below. We are interested in the approximation of the statistical experiment

$$\mathcal{E}_n = (N_n(0, \Gamma_n(f)), f \in \Sigma), \tag{1.3}$$

in the sense of Le Cam's deficiency pseudodistance $\Delta(\cdot, \cdot)$; see the end of this section for a precise definition. The statistical interpretation of the Le Cam distance is as follows. For two experiments $\mathcal{E}$ and $\mathcal{F}$ having the same parameter space, $\Delta(\mathcal{E}, \mathcal{F}) < \varepsilon$ implies that for any decision problem with loss bounded by 1, and any statistical procedure with the experiment $\mathcal{E}$, there is a (randomized) procedure with $\mathcal{F}$, the risk of which evaluated in $\mathcal{F}$ nearly matches (within ε) the risk of the original procedure evaluated in $\mathcal{E}$. In this statement the roles of $\mathcal{E}$ and $\mathcal{F}$ can also be reversed. Two sequences, $\mathcal{E}_n$ and $\mathcal{F}_n$, are said to be *asymptotically equivalent* if $\Delta(\mathcal{E}_n, \mathcal{F}_n) \to 0$.

As a guide to what can be expected, consider first the case where f_ϑ, $\vartheta \in \Theta$, is a smooth parametric family of spectral densities. Assume that Θ is a real interval; under some regularity conditions, the model is well known to fulfill the standard LAN conditions with localization rate $n^{-1/2}$ and normalized Fisher information at ϑ,

$$\frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\frac{\partial}{\partial \vartheta} \log f_\vartheta(\omega) \right)^2 d\omega$$

(Davies [13], Dzhaparidze [15], Chapter I.3, cf. also the discussion in van der Vaart [32], Example 7.17). Consider the parametric Gaussian white noise model where the signal is the log-spectral density,

$$dZ_\omega = \log f_\vartheta(\omega)\, d\omega + 2\pi^{1/2} n^{-1/2}\, dW_\omega, \qquad \omega \in [-\pi, \pi], \tag{1.4}$$

and note that in the family $(f_\vartheta, \vartheta \in \Theta)$, this model has the same asymptotic Fisher information. This is in agreement with the LAN result for the spectral density model, but it suggests that the above white noise approximation might also be true for larger (i.e., nonparametric) spectral density classes Σ.

As a second piece of evidence for the white noise approximation in the nonparametric case, we take known results about the approximate spectral decomposition of the Toeplitz covariance matrix $\Gamma_n(f)$. It is a classical difficulty in time series analysis that the exact eigenvalues and eigenvectors of $\Gamma_n(f)$ cannot easily be found and used for inference about f; in particular, the eigenvectors depend on f. However, for an approximation which is a circulant matrix [denoted $\tilde{\Gamma}_n(f)$ below], the eigenvectors are independent of f and the eigenvalues are approximately $f(\omega_j)$, where ω_j are the points of an equispaced grid of size n in $[-\pi, \pi]$. If the approximation by $\tilde{\Gamma}_n(f)$ were justified, one could apply an orthogonal transformation to the data $y^{(n)}$ and obtain a Gaussian scale model,

$$z_j = f^{1/2}(\omega_j)\xi_j, \qquad j = 1, \dots, n, \tag{1.5}$$

where ξ_j are independent standard normal. For this model, nonparametric asymptotic equivalence theory was developed in Grama and Nussbaum [18]. Results there, for certain smoothness classes $f \in \Sigma$, with f bounded away from 0, lead to the nonparametric version of the white noise model (1.4),

$$dZ_\omega = \log f(\omega)\, d\omega + 2\pi^{1/2} n^{-1/2}\, dW_\omega, \qquad \omega \in [-\pi, \pi],\ f \in \Sigma. \tag{1.6}$$

Our proof of asymptotic equivalence will, in fact, be based on the approximation of the covariance matrix $\Gamma_n(f)$ by the circulant $\tilde{\Gamma}_n(f)$ (cf. Brockwell and Davis [3], Section 4.5). However, we shall see that this tool does not enable a straightforward approximation of the data $y^{(n)}$ in total variation or Hellinger distance. Therefore, our argument for asymptotic equivalence will be somewhat indirect, involving "bracketing" of the experiment $\mathcal{E}_n$ by upper and lower bounds (in the sense of informativity) and also a preliminary localization of the parameter space.

To formulate our main result, define a parameter space Σ of spectral densities as follows. For $M > 0$, define a set of real-valued even functions on $[-\pi, \pi]$,

$$\mathcal{F}_M = \{f : M^{-1} \le f(\omega),\ f(\omega) = f(-\omega), \omega \in [-\pi, \pi]\}. \tag{1.7}$$

Thus our spectral densities are assumed uniformly bounded away from 0. Let $L_2(-\pi, \pi)$ be the usual (real) L_2-space on $[-\pi, \pi]$; for any $f \in L_2(-\pi, \pi)$, let $\gamma_f(k)$, $k \in \mathbb{Z}$, be the Fourier coefficients according to (1.1). For any $\alpha > 0$ and $M > 0$ let

$$W^\alpha(M) = \left\{ f \in L_2(-\pi, \pi) : \gamma_f^2(0) + \sum_{k=-\infty}^{\infty} |k|^{2\alpha} \gamma_f^2(k) \le M \right\}. \tag{1.8}$$

These sets correspond to balls in the periodic fractional Sobolev scale with smoothness coefficient α. Note that for $\alpha > 1/2$, by an embedding theorem (Lemma 5.6 in [17]), functions in $W^\alpha(M)$ are also uniformly bounded. Define an a priori set for given $\alpha > 0$, $M > 0$,

$$\Sigma_{\alpha,M} = W^\alpha(M) \cap \mathcal{F}_M.$$

Consider also a Gaussian scale model (1.5), where the values $f(\omega_j)$ are replaced by local averages,

$$J_{j,n}(f) = n \int_{(j-1)/n}^{j/n} f(2\pi x - \pi)\, dx, \qquad j = 1, \dots, n.$$

THEOREM 1.1. *Let Σ be a set of spectral densities contained in $\Sigma_{\alpha,M}$ for some $M > 0$ and $\alpha > 1/2$. Then the experiments given by observations*:

$y(1), \dots, y(n)$, *a stationary centered Gaussian sequence with spectral density* f;

$z_1, \dots, z_n$, *where z_j are independent $N(0, J_{j,n}(f))$*;

with $f \in \Sigma$ are asymptotically equivalent.

Let $\|\cdot\|_{B^{\alpha}_{p,q}}$ be the Besov norm on the interval $[-\pi,\pi]$ with smoothness index α (for a summary of analytical topics, cf. Section 5 in the technical report [17]). For the second main result, we impose a smoothness condition involving this norm for the $\alpha > 1/2$ from above, and $p = q = 6$.

THEOREM 1.2. *Let Σ be a set of spectral densities as in Theorem 1.1, fulfilling additionally $\|f\|_{B^{\alpha}_{6,6}} \le M$ for all $f \in \Sigma$. Then the experiments given, respectively, by observations*:

$$z_1,\dots,z_n, \textit{ where } z_j \textit{ are independent } N(0, J_{j,n}(f));$$

$$dZ_\omega = \log f(\omega)\, d\omega + 2\pi^{1/2} n^{-1/2}\, dW_\omega, \qquad \omega \in [-\pi,\pi];$$

with $f \in \Sigma$ are asymptotically equivalent.

The proof of this result is in the thesis of Zhou [33]. The present paper is devoted to the proof of Theorem 1.1.

In nonparametric asymptotic equivalence theory, some constructive results have recently been obtained, that is, explicit equivalence maps have been exhibited which allow to carry over optimal decision functions from one sequence of experiments to the other. Brown and Low [4] and Brown, Low and Zhang [5] obtained constructive results for white noise with drift and Gaussian regression with nonrandom and random design. Brown et al. [6] found such equivalence maps (Markov kernels) for the i.i.d. model on the unit interval (density estimation) and the model of Gaussian white noise with drift (cf. also Carter [7]). The theoretical (nonconstructive) variant of this result had earlier been established in [26], in the sense of an existence proof for pertaining Markov kernels. This indirect approach relied on the well-known connection to likelihood processes of experiments (cf. Le Cam and Yang [23]). In the present paper, the proof of Theorem 1.1 is of nonconstructive type, using a variety of methods for bounding the Δ-distance between the time series experiment and the model of independent zero mean Gaussians. Similarly, the proof of Theorem 1.2 in Zhou [33] is nonconstructive, but it appears likely in that a second step, relatively simple "workable" equivalence maps can be found, at least for the case of Theorem 1.1 which is related to the classical result about asymptotic independence of discrete Fourier transforms.

To further discuss the context of the main results, we note the following points:

1. *Asymptotic independence of discrete Fourier transforms.* Let

$$d_n(\omega) = \sum_{k=1}^{n} \exp(-ik\omega) y(k), \qquad \omega \in (-\pi,\pi),$$

be the discrete Fourier transform of the time series $y(1),\dots,y(n)$. Assume n is uneven and let η_j be complex standard normal variables. It is well known that for the Fourier frequencies $\omega_j = 2\pi j/n,\ j = 1, \dots, (n-1)/2$ in $(0, \pi)$, there is an asymptotic distribution

$$(\pi n)^{-1/2} d_n(\omega_j) \approx \exp(i\omega_j) f^{1/2}(\omega_j)\eta_j \tag{1.9}$$

and the values are asymptotically uncorrelated for distinct ω_j, ω_k. For a precise formulation, see relation (2.12) below, or [3], Proposition 4.5.2. This fact is the basis for many inference methods (cf. Dahlhaus and Janas [11]; see also Lahiri [21] for an extended discussion of the asymptotic independence). A linear transformation to $n-1$ independent real normals, and adding a real normal according to $(2\pi n)^{-1/2} d_n(0) \approx N(0, f(0))$ suggests the Gaussian scale model (1.5).

2. *Log-periodogram regression*. Consider also the periodogram

$$I_n(\omega) = \frac{1}{2\pi n} |d_n(\omega)|^2.$$

Note the equality in distribution $|\eta_j|^2 \sim \chi_2^2 \sim 2e_j$, where e_j is standard exponential. As a consequence of the above result about $d_n(\omega_j)$, we have for $j = 1, \dots, (n-1)/2$,

$$I_n(\omega_j) \approx f(\omega_j) e_j \tag{1.10}$$

with asymptotic independence. Assuming this model is exact, taking a logarithm gives rise to the inference method of log-periodogram regression (for an account, cf. Fan and Gijbels [16], Section 6.4).

3. *The Whittle approximation.* This is an approximation to $-n^{-1}$ times the log-likelihood of the time series $y(1), \dots, y(n)$. In a parametric model f_ϑ, $\vartheta \in \Theta$, with multivariate normal law $N_n(0, \Gamma_n(f_\vartheta))$, computation of the MLE involves inverting the covariance matrix $\Gamma_n(f_\vartheta)$, which is difficult since both eigenvectors and eigenvalues depend on ϑ in general. Replacing $\Gamma_n^{-1}(f_\vartheta)$ by $\Gamma_n(1/4\pi^2 f_\vartheta)$ and using an approximation to $n^{-1} \log \Gamma_n(f_\vartheta)$ leads to an expression $L^W(f) + \log 2\pi$, where

$$L^W(f) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\log f_\vartheta(\omega) + \frac{I_n(\omega)}{f_\vartheta(\omega)} \right) d\omega \tag{1.11}$$

is the Whittle likelihood (cf. Dahlhaus [10] for a brief exposition and references). A closely related expression is obtained by assuming the model (1.10) exact: then $-n^{-1}$ times the log-likelihood is

$$L_n^W(f) = n^{-1} \sum_{j=1}^{(n-1)/2} \left(\log f_\vartheta(\omega_j) + \frac{I_n(\omega_j)}{f_\vartheta(\omega_j)} \right),$$

that is, a discrete approximation to (1.11). For applications of the Whittle likelihood to nonparametric inference, confer Dahlhaus and Polonik [12].

4. *Asymptotics for* $L^W(f)$. The accuracy of the Whittle approximation has been described as follows (Coursol and Dacunha-Castelle [9], Dzhaparidze [15], Theorem 1, page 52): let $L_n(f)$ be the log-likelihood in the experiment (1.3); then

$$L_n(f) = -nL^W(f) - n \log 2\pi + O_P(1), \tag{1.12}$$

uniformly over $f \in \Sigma_{1/2,M}$. This justifies use of $L^W(f)$ as a contrast function, for example, it yields asymptotic efficiency of the Whittle MLE in parametric models ([15], Chapter II), but falls short of providing asymptotic equivalence in the Le Cam sense. Indeed, if (1.12) were true with $o_P(1)$ in place of $O_P(1)$ and with $L^W(f)$ replaced by $L_n^W(f)$ [on a suitable common probability space for $L_n(f)$ and $L_n^W(f)$], then this would already imply total variation equivalence, up to an orthogonal transform, of the exact model (1.10) with $f \in \Sigma_{1/2,M}$ (via the Scheffé lemma argument of Delattre and Hoffmann [14]). In Section 2 below [cf. relation (2.18)], we note a corresponding negative result, essentially, that this total variation approximation over $f \in \Sigma_{1/2,M}$ does not take place.

Define the *Whittle measure* as the law of the sample $y^{(n)}$ when (1.9) holds exactly, such that $d_n(\omega_j)_{j=1,\dots,(n-1)/2}$ are independent centered Gaussian. Choudhouri, Ghosal and Roy [8] establish that under some smoothness conditions on f, the Whittle measure is contiguous to the original law of $y^{(n)}$, with some decision theoretic consequences. A further discussion of this result in the present context can be found at the end of Section 2.

5. *Conditions for Theorem* 1.2. For a narrower parameter space, that is, a Hölder ball with smoothness index $\alpha > 1/2$, the result of Theorem 1.2 has been proved in [18]. Note that the Sobolev balls $W^\alpha(M)$ figuring in Theorem 1.1 are natural parameter sets of spectral densities since the smoothness condition is directly stated in terms of the autocovariance function $\gamma_f(\cdot)$. The Besov balls $B^\alpha_{p,p}(M)$, given in terms of the norm $\|\cdot\|_{B^\alpha_{p,p}}$, are intermediate between L_2-Sobolev and Hölder balls. For the white noise approximation of the i.i.d. (density estimation) model, Brown et al. [6] succeeded in weakening the Hölder ball condition in [26] to a condition that Σ is compact both in the Besov spaces $B^{1/2}_{2,2}$ and $B^{1/2}_{4,4}$ on the unit interval. This is immediately implied by $\Sigma \subset B^\alpha_{4,4}(M)$ for some $\alpha > 1/2$. Our condition for Theorem 1.2 is slightly stronger, that is, $\Sigma \subset B^\alpha_{6,6}(M)$ for some $\alpha > 1/2$. In Remark 5.8 of the technical report [17], we note a sufficient condition in terms of the autocovariance function $\gamma_f(\cdot)$, that is, we give a description of the periodic version of the Besov ball.

Throughout this paper, we adopt the notation that C represents a constant independent of n and the parameter (spectral density) $f \in \Sigma$, and that the value of C may change at each occurrence, even on the same line.

Relations between experiments. All measurable sample spaces are assumed to be Polish (complete separable) metric spaces equipped with their Borel sigma algebra. For measures P, Q on the same sample space, let $\|P - Q\|_{\mathrm{TV}}$ be the total variation distance. For the general case, where P, Q are not necessarily on the same sample space, suppose K is a Markov kernel such that KP is a measure on the same sample space as Q. In that case, $\|Q - KP\|_{\mathrm{TV}}$ is defined and will be used as generic notation for a Markov kernel K.

Consider now experiments (families of measures) $\mathcal{F} = (Q_f, f \in \Sigma)$ and $\mathcal{E} = (P_f, f \in \Sigma)$, with the same parameter space Σ. All experiments here are assumed dominated by a sigma-finite measure on their respective sample space. If $\mathcal{E}$ and $\mathcal{F}$ are on the same sample space, define their total variation distance,

$$\Delta_0(\mathcal{E}, \mathcal{F}) = \sup_{f \in \Sigma} \|Q_f - P_f\|_{\mathrm{TV}}.$$

In the general case, the deficiency of $\mathcal{E}$ with respect to $\mathcal{F}$ is defined as

$$\delta(\mathcal{E}, \mathcal{F}) = \inf_K \sup_{f \in \Sigma} \|Q_f - KP_f\|_{\mathrm{TV}},$$

where inf extends over all appropriate Markov kernels. Le Cam's pseudodistance $\Delta(\cdot, \cdot)$ between $\mathcal{E}$ and $\mathcal{F}$ then is

$$\Delta(\mathcal{E}, \mathcal{F}) = \max(\delta(\mathcal{E}, \mathcal{F}), \delta(\mathcal{F}, \mathcal{E})).$$

Furthermore, we will use the following notation involving experiments $\mathcal{E}, \mathcal{F}$ or sequences of such $\mathcal{E}_n = (P_{n,f}, f \in \Sigma)$ and $\mathcal{F}_n = (Q_{n,f}, f \in \Sigma)$.

Notation.

$\mathcal{E} \preceq \mathcal{F}$ ($\mathcal{F}$ more informative than $\mathcal{E}$): $\delta(\mathcal{F}, \mathcal{E}) = 0$

$\mathcal{E} \sim \mathcal{F}$ (equivalent): $\Delta(\mathcal{E}, \mathcal{F}) = 0$

$\mathcal{E}_n \simeq \mathcal{F}_n$ (asymptotically total variation equivalent): $\Delta_0(\mathcal{F}_n, \mathcal{E}_n) \to 0$

$\mathcal{E}_n \precsim \mathcal{F}_n$ ($\mathcal{F}_n$ asymptotically more informative than $\mathcal{E}_n$): $\delta(\mathcal{F}_n, \mathcal{E}_n) \to 0$

$\mathcal{E}_n \approx \mathcal{F}_n$ (asymptotically equivalent): $\Delta(\mathcal{F}_n, \mathcal{E}_n) \to 0$

Note that "more informative" above is used in the sense of a semi-ordering, that is, its actual meaning is "at least as informative." We shall also write the relation $\simeq$ in a less formal way between data vectors such as $x^{(n)} \simeq y^{(n)}$, if it is clear from the context which experiments the data vectors represent.

2. The periodic Gaussian experiment. From now on, we shall assume that n is uneven. Our argument for asymptotic equivalence is such that it easily allows extension to the case of general sequences $n \to \infty$ (cf. Remark 4.10 for details).

Recall that the covariance matrix $\Gamma_n = \Gamma_n(f)$ has the Toeplitz form $(\Gamma_n)_{j,k} = \gamma(k - j)$, $j, k = 1, \ldots, n$, that is,

$$\Gamma_n = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-2) & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \cdots & \cdots & \gamma(n-2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma(n-2) & \cdots & \cdots & \gamma(0) & \gamma(1) \\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(1) & \gamma(0) \end{pmatrix}.$$

Following [3], Section 4.5, we shall define a circulant matrix approximation by

$$\tilde{\Gamma}_n = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(2) & \gamma(1) \\ \gamma(1) & \gamma(0) & \cdots & \cdots & \gamma(2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma(2) & \cdots & \cdots & \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(2) & \cdots & \gamma(1) & \gamma(0) \end{pmatrix},$$

where in the first row, the central element and the one following it coincide with $\gamma((n-1)/2)$. More precisely, for given uneven n, define a function on integers h with $|h| < n$,

$$\tilde{\gamma}_{(n),f}(h) = \begin{cases} \gamma_f(h), & |h| \le (n-1)/2, \\ \gamma_f(n-|h|), & (n+1)/2 \le |h| \le n-1, \end{cases}$$

and set

$$(\tilde{\Gamma}_n)_{j,k}(f) = \tilde{\gamma}_{(n),f}(k-j), \qquad j,k = 1,\dots,n. \tag{2.1}$$

We shall also write $\tilde{\Gamma}_n(f)$ for the corresponding $n \times n$ matrix, or simply $\tilde{\Gamma}_n$ and $\tilde{\gamma}_{(n)}(h)$, if the dependence on f is understood. Define

$$\omega_j = \frac{2\pi j}{n}, \qquad |j| \le (n-1)/2. \tag{2.2}$$

It is well known (see [3], Relation 4.5.5) that the spectral decomposition of $\tilde{\Gamma}_n$ can be described as follows. We have

$$\tilde{\Gamma}_n = \sum_{|j| \le (n-1)/2} \lambda_j \mathbf{u}_j \mathbf{u}_j', \tag{2.3}$$

where λ_j are real eigenvalues, and $\mathbf{u}_j$ are real orthonormal eigenvectors. The eigenvalues are

$$\lambda_j = \sum_{|k| \le (n-1)/2} \gamma(k) \exp(-i\omega_j k), \qquad |j| \le (n-1)/2.$$

Note that $\lambda_j = \lambda_{-j}$, $j \ne 0$, and that the λ_j are approximate values of $2\pi f$ in the points ω_j. Indeed, define

$$\tilde{f}_n(\omega) = \frac{1}{2\pi} \sum_{|k| \le (n-1)/2} \gamma(k) \exp(ik\omega), \qquad \omega \in [-\pi, \pi], \tag{2.4}$$

a truncated Fourier series approximation to f; then $\tilde{f}_n$ is an even function on $[-\pi, \pi]$ and

$$\lambda_j = 2\pi \tilde{f}_n(\omega_j), \qquad |j| \le (n-1)/2. \tag{2.5}$$

The eigenvectors are:

$$\mathbf{u}_0' = n^{-1/2}(1, \dots, 1); \tag{2.6}$$

$$\mathbf{u}_j' = (2/n)^{1/2}\bigl(1, \cos(\omega_j), \cos(2\omega_j), \dots, \cos\bigl((n-1)\omega_j\bigr)\bigr); \tag{2.7}$$

$$\begin{aligned} \mathbf{u}_{-j}' = (2/n)^{1/2}\bigl(0, \sin(\omega_j), \sin(2\omega_j), \dots, \sin\bigl((n-1)\omega_j\bigr)\bigr),& \\ j = 1, \dots, (n-1)/2.& \end{aligned} \tag{2.8}$$

In our setting, the circulant matrix $\tilde{\Gamma}_n$ is positive definite for n large enough. Indeed, Lemma 5.6 in [17] implies that $\tilde{f}_n \ge M^{-1}/2$ uniformly over $f \in \Sigma$, for n large enough, so that $\tilde{\Gamma}_n(f)$ is a covariance matrix. Define the experiment, in analogy to (1.3),

$$\tilde{\mathcal{E}}_n = \bigl(N_n(0, \tilde{\Gamma}_n(f)), f \in \Sigma\bigr) \tag{2.9}$$

with data $\tilde{y}^{(n)}$, say. The sequence $\tilde{y}^{(n)}$ may be called a "periodic process" since it can be represented in terms of independent standard Gaussians ξ_j, as a finite sum

$$\tilde{y}^{(n)} = \sum_{|j| \le (n-1)/2} \lambda_j^{1/2} \mathbf{u}_j \xi_j, \tag{2.10}$$

where the vector $\mathbf{u}_j$ describes a deterministic oscillation [cf. (2.6)–(2.8)]. Accordingly, $\tilde{\mathcal{E}}_n$ will be called a *periodic Gaussian experiment*.

The periodic process $\tilde{y}^{(n)}$ is known to approximate the original time series $y^{(n)}$ in the following sense. Define the $n \times n$-matrix

$$U_n = \bigl(\mathbf{u}_{-(n-1)/2}, \dots, \mathbf{u}_{(n-1)/2}\bigr) \tag{2.11}$$

and consider the transforms

$$z^{(n)} = (2\pi)^{-1/2} U_n' y^{(n)}, \qquad \tilde{z}^{(n)} = (2\pi)^{-1/2} U_n' \tilde{y}^{(n)}.$$

Denote $\operatorname{Cov}(z^{(n)})$, the covariance matrix of the random vector $z^{(n)}$. Then we have ([3], Proposition 4.5.2), for given $f \in \Sigma$,

$$\sup_{1 \le i, j \le n} \bigl|\operatorname{Cov}\bigl(z^{(n)}\bigr)_{i,j} - \operatorname{Cov}\bigl(\tilde{z}^{(n)}\bigr)_{i,j}\bigr| \to 0 \qquad \text{as } n \to \infty. \tag{2.12}$$

Since $\operatorname{Cov}(\tilde{z}^{(n)})$ is diagonal with diagonal elements $\lambda_j/2\pi$, this means that the elements of $z^{(n)}$ are approximately uncorrelated for large n.

Note that $\tilde{z}^{(n)}$ can also be written, in accordance with (2.10) and (2.5),

$$\tilde{z}^{(n)} = (\tilde{f}_n^{1/2}(\omega_j)\xi_j)_{|j| \le (n-1)/2}, \tag{2.13}$$

which is nearly identical with the Gaussian scale model (1.5). Thus the question appears whether the approximation (2.12) can be strengthened to a total variation approximation of the respective laws $\mathcal{L}(z^{(n)}|f)$ and $\mathcal{L}(\tilde{z}^{(n)}|f)$.

The answer to that is negative; let us introduce some notation. For $n \times n$ matrices $A = (a_{jk})$, define the Euclidean norm $\|A\|$ by

$$\|A\|^2 := \mathrm{tr}[A'A] = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk}^2.$$

If A is symmetric, we denote the largest and smallest eigenvalues by $\lambda_{\max}(A)$, $\lambda_{\min}(A)$. For later use, we also define the operator norm of (not necessarily symmetric) A by

$$|A| := (\lambda_{\max}(A'A))^{1/2}.$$

If A is symmetric nonnegative definite, then $|A| = \lambda_{\max}(A)$. The following lemma shows that the Hellinger distance between the laws of $y^{(n)}$ and $\tilde{y}^{(n)}$ depends crucially on the total Euclidean distance $\|\Gamma_n(f) - \tilde{\Gamma}_n(f)\|$ between the covariance matrices, so that an element-wise convergence as in (2.12) is not enough.

LEMMA 2.1. *Let A, B be $n \times n$ covariance matrices, and suppose that for some $M > 1$,*

$$0 < M^{-1} \leq \lambda_{\min}(A) \quad \textit{and} \quad \lambda_{\max}(A) \leq M.$$

Then there exist $\epsilon = \epsilon_M > 0$ and $K = K_M > 1$ not depending on A, B and n such that $\|A - B\| \leq \epsilon$ implies

$$K^{-1}\|A - B\|^2 \leq H^2(N_n(0, A), N_n(0, B)) \leq K\|A - B\|^2,$$

where $H(\cdot, \cdot)$ is the Hellinger distance.

The proof can be found in [17], Section 5. To apply this lemma, set $A = \Gamma_n(f)$, $B = \tilde{\Gamma}_n(f)$ and note that, since $f \in \Sigma$ is bounded and bounded away from 0 (both uniformly over $f \in \Sigma$), the condition on the eigenvalues of $\Gamma_n(f)$ is fulfilled, also uniformly over $f \in \Sigma$ ([3], Proposition 4.5.3). We shall see that the expression $\|\Gamma_n(f) - \tilde{\Gamma}_n(f)\|^2$ is closely related to a Sobolev-type seminorm for smoothness index 1/2. For any $f \in L_2(-\pi, \pi)$ given by (1.1), set

$$|f|_{2,\alpha}^2 := \sum_{k=-\infty}^{\infty} |k|^{2\alpha} \gamma_f^2(k), \qquad \|f\|_{2,\alpha}^2 := \gamma_f^2(0) + |f|_{2,\alpha}^2, \tag{2.14}$$

provided the right-hand side is finite; the Sobolev ball $W^\alpha(M)$ given by (1.8) is then described by $\|f\|_{2,\alpha}^2 \leq M$. Also, for any natural m, define a finite-dimensional linear subspace of $L_2(-\pi, \pi)$

$$L_m = \left\{ f \in L_2(-\pi, \pi) : \int f(\omega) \exp(ik\omega)\, d\omega = 0, |k| > m \right\}.$$

LEMMA 2.2. (i) *For any* $f \in \Sigma$ *we have*

$$\|\Gamma_n(f) - \tilde{\Gamma}_n(f)\|^2 \le 2|f|^2_{2,1/2}, \tag{2.15}$$

and for $f \in \Sigma \cap L_{(n-1)/2}$

$$|f|^2_{2,1/2} = \|\Gamma_n(f) - \tilde{\Gamma}_n(f)\|^2.$$

(ii) *For any* $f, f_0 \in \Sigma$ *we have*

$$\big\|\Gamma_n(f) - \Gamma_n(f_0) - \big(\tilde{\Gamma}_n(f) - \tilde{\Gamma}_n(f_0)\big)\big\|^2 \le 2|f - f_0|^2_{2,1/2}. \tag{2.16}$$

PROOF. (i) From the definition of $\Gamma_n(f)$ and $\tilde{\Gamma}_n(f)$ in terms of $\gamma(\cdot), \tilde{\gamma}_{(n)}(\cdot)$, we immediately obtain

$$\begin{aligned}
\|\Gamma_n(f) - \tilde{\Gamma}_n(f)\|^2 &= \sum_{|k| \le n-1} (n - |k|)\big(\gamma(k) - \tilde{\gamma}_{(n)}(k)\big)^2 \\
&= \sum_{|k|=(n+1)/2}^{n-1} (n - |k|)\big(\gamma(k) - \gamma(n - |k|)\big)^2 \\
&= 2 \sum_{k=1}^{(n-1)/2} k\big(\gamma(k) - \gamma(n-k)\big)^2 \qquad (2.17) \\
&\le 2 \sum_{k=1}^{(n-1)/2} 2k\big(\gamma^2(k) + \gamma^2(n-k)\big) \\
&\le 4 \sum_{k=1}^{n-1} k\gamma^2(k) \le 2|f|^2_{2,1/2}.
\end{aligned}$$

The first inequality is proved. The second one follows immediately from (2.17).

(ii) Note that for any n, the mapping $f \to \Gamma_n(f)$ if it is defined by (1.2) for any $f \in L_2(-\pi, \pi)$ is linear, and the same is true for $f \to \tilde{\Gamma}_n(f)$ defined by (2.1). Hence

$$\Gamma_n(f) - \Gamma_n(f_0) = \Gamma_n(f - f_0), \qquad \tilde{\Gamma}_n(f) - \tilde{\Gamma}_n(f_0) = \tilde{\Gamma}_n(f - f_0).$$

Now the argument is completely analogous to (i), if $\gamma(k) = \gamma_f(k)$ is replaced by $\gamma_{f-f_0}(k)$. $\square$

Our assumption $f \in \Sigma$, that is, $\|f\|^2_{2,\alpha} \le M$ for some $\alpha > 1/2$, provides an upper bound M for $|f|^2_{2,1/2}$ but does guarantee that this term is uniformly small. Thus we are not able to utilize Lemma 2.1 to approximate $\mathcal{E}_n$ by $\tilde{\mathcal{E}}_n$ in Hellinger distance. In fact, this Hellinger distance approximation does not take place: take a fixed m, select $f \in \Sigma \cap L_m$ such that $\|f\|^2_{2,1/2} < \epsilon$ with ϵ from Lemma 2.1 and use the lower bound in this lemma to show that

$$H^2(N_n(0, \Gamma_n(f)), N_n(0, \tilde{\Gamma}_n(f))) \geq K^{-1}\epsilon^2 \tag{2.18}$$

for all sufficiently large n. Thus the direct approximation of the time series data $y^{(n)}$ by the periodic process $\tilde{y}^{(n)}$ in total variation distance fails.

However, that does not contradict asymptotic equivalence since the latter allows for a randomization mapping (Markov kernel) applied to $\tilde{y}^{(n)}$ and $y^{(n)}$, respectively, before total variation distance of the laws is taken. We will show the existence of appropriate Markov kernels in an indirect way, via a bracketing of the original time series experiment by upper and lower bounds in the sense of informativity.

Let now $\mathcal{E}_n$ again be the time series experiment (1.3); we shall find an asymptotic bracketing, that is, two sequences $\mathring{\mathcal{E}}_{l,n}$, $\mathring{\mathcal{E}}_{u,n}$ such that

$$\mathring{\mathcal{E}}_{l,n} \precsim \mathcal{E}_n \precsim \mathring{\mathcal{E}}_{u,n},$$

and such that both $\mathring{\mathcal{E}}_{l,n}$ and $\mathring{\mathcal{E}}_{u,n}$ are asymptotically equivalent to $\tilde{\mathcal{E}}_n$ given by (2.9), and to $\mathring{\mathcal{E}}_n$, representing the independent Gaussians $z_1, \ldots, z_n$ in Theorem 1.1.

REMARK (Contiguity of the Whittle measure). Choudhouri, Ghosal and Roy [8] establish contiguity of the sequence of laws $N_n(0, \Gamma_n(f)$ to the sequence $N_n(0, \tilde{\Gamma}_n(f))$ as $n \to \infty$. In fact, this is shown for an inessential modification of the circulant $\tilde{\Gamma}_n(f)$, where the eigenvalues are $2\pi f(\omega_j)$ instead of (2.5). Contiguity would be implied by a total variation approximation of the two sequences of measures, which we have shown to fail in (2.18). On the other hand, contiguity implies a lack of entire asymptotic separation, which can easily be confirmed by a Hellinger affinity computation similar to Lemma 2.1. Note that contiguity is a property of the sequence of binary experiments $(N_n(0, \Gamma_n(f), N_n(0, \tilde{\Gamma}_n(f)))$ for fixed f (cf. Le Cam and Yang [23], page 20), whereas our result concerns approximation in Δ-distance of the full nonparametric experiments $(N_n(0, \Gamma_n(f)), f \in \Sigma)$ and $(N_n(0, \tilde{\Gamma}_n(f)), f \in \Sigma)$. Thus neither of these results implies the other. However, contiguity of the Whittle measure is an interesting fact with possible application in the problem of simple explicit equivalence maps.

3. Upper informativity bracket. The spectral representation (2.10) of the periodic sequence $\tilde{y}^{(n)} = (\tilde{y}(1), \ldots, \tilde{y}(n))'$ can be written

$$\begin{aligned} \tilde{y}(t) &= (2\pi/n)^{1/2} \tilde{f}_n^{1/2}(0)\xi_0 \\ &\quad + 2(\pi/n)^{1/2} \sum_{j=1}^{(n-1)/2} \tilde{f}_n^{1/2}(\omega_j) \cos\big((t-1)\omega_j\big)\xi_j \\ &\quad + 2(\pi/n)^{1/2} \sum_{j=-(n-1)/2}^{1} \tilde{f}_n^{1/2}(\omega_j) \sin\big((t-1)\omega_j\big)\xi_j, \qquad t = 1, \ldots, n. \end{aligned} \tag{3.1}$$

We see that here, $\tilde{y}^{(n)}$ is a one-to-one function $\tilde{y}^{(n)} = U\tilde{z}^{(n)}$ of the n-vector of independent Gaussians $\tilde{z}^{(n)}$ [cf. (2.13)], but the approximation of $\tilde{y}^{(n)}$ to $y^{(n)}$ is not in the total variation sense [cf. (2.18)]. Now take a limit in (3.1) for $n \to \infty$ and fixed t, and observe that (heuristically) this yields the spectral representation of the original stationary sequence $y(t)$,

$$
\begin{aligned}
y(t+1) = {} & \int_{[0,\pi]} \sqrt{2} f^{1/2}(\omega) \cos(t\omega)\, dB_\omega \\
& + \int_{[-\pi,0]} \sqrt{2} f^{1/2}(\omega) \sin(t\omega)\, dB_\omega, \qquad t = 0, 1, \dots,
\end{aligned} \tag{3.2}
$$

where dB_ω is standard Gaussian white noise on $[-\pi, \pi]$ (cf. [3], Problem 4.31). Here for any n, the vector $y^{(n)} = (y(1), \dots, y(n))'$ is represented as a functional of the continuous time process

$$
dZ^*_\omega = f^{1/2}(\omega)\, dB_\omega, \qquad \omega \in [-\pi, \pi].
$$

Thus a completely observed process Z^*_ω, $\omega \in [-\pi, \pi]$ would represent an upper informativity bracket for any sample size n, but this experiment is statistically trivial since the observation here identifies the parameter f.

Our approach now is to construct an intermediate series $\tilde{y}^{(m,n)}$ of size n in which the uniform size n grid of points ω_j, $|j| \le (n-1)/2$, is replaced by a finer uniform grid of $m > n$ points in the representation (3.1). Thus $\tilde{y}^{(n,m)}$ is a functional, not of n independent Gaussians, but of $m > n$ of these; call their vector $\tilde{z}^{(m)}$. The random vector $\tilde{z}^{(m)}$ now represents an upper informativity bracket which remains nontrivial (asymptotically) if $m - n \to \infty$ not too quickly. An equivalent description of that idea is as follows. Consider $m > n$ and the periodic process $\tilde{y}^{(m)}$ given by (2.10), where the original sample size n is replaced by m. Then define $\tilde{y}^{(n,m)}$ as the vector of the first n components of $\tilde{y}^{(m)}$. The law of $\tilde{y}^{(n,m)}$ is $N_n(0, \tilde{\Gamma}_{n,m}(f))$, where $\tilde{\Gamma}_{n,m}(f)$ is the upper left $n \times n$ submatrix of $\tilde{\Gamma}_m(f)$.

We now easily observe the improved approximation quality of $\tilde{y}^{(n,m)}$ for $y^{(n)}$. Assume that m is also uneven. First note that for $(m+1)/2 \ge n$, we already obtain $\tilde{\Gamma}_{n,m}(f) = \Gamma_n(f)$. This follows immediately from the definition of the circular matrix $\tilde{\Gamma}_m(f)$ via the autocovariance function $\tilde{\gamma}_{(m)}(\cdot)$. However, we would like to limit the increase of sample size, that is, require $m/n \to 1$; therefore, in what follows, we assume $m < 2n - 1$.

LEMMA 3.1. *Assume m is uneven, $n < m < 2n - 1$. Then for any $f \in \Sigma$ we have*

$$
\|\Gamma_n(f) - \tilde{\Gamma}_{n,m}(f)\|^2 \le 4(m - n + 1)^{1-2\alpha} |f|^2_{2,\alpha}
$$

and hence if $m = m_n$ is such that $m - n \to \infty$ as $n \to \infty$ then

$$
\sup_{f \in \Sigma} H^2(N_n(0, \Gamma_n(f)), N_n(0, \tilde{\Gamma}_{n,m}(f))) \to 0. \tag{3.3}
$$

PROOF. From the definition of $\Gamma_n(f)$ and $\tilde{\Gamma}_{n,m}(f)$, we immediately obtain

$$\begin{aligned}\|\Gamma_n(f)-\tilde{\Gamma}_{n,m}(f)\|^2 &= \sum_{|k|\le n-1}(n-|k|)\big(\gamma(k)-\tilde{\gamma}_{(m)}(k)\big)^2\\ &= 2\sum_{k=(m+1)/2}^{n-1}(n-k)\big(\gamma(k)-\gamma(m-k)\big)^2\\ &\le 4\sum_{k=(m+1)/2}^{n-1}(n-k)\big(\gamma^2(k)+\gamma^2(m-k)\big).\end{aligned}$$

Now note that for $m>n$, the relation $(m+1)/2\le k\le n-1$ implies $k\ge(n+1)/2$, and therefore, $n-k<k$, and note also, $n-k<m-k$. We obtain an upper bound,

$$\begin{aligned}&\le 4\sum_{k=(m+1)/2}^{n-1}k\gamma^2(k)+4\sum_{k=(m+1)/2}^{n-1}(m-k)\gamma^2(m-k)\\ &=4\sum_{k=(m+1)/2}^{n-1}k\gamma^2(k)+4\sum_{k=m-n+1}^{(m-1)/2}k\gamma^2(k)=4\sum_{k=m-n+1}^{n-1}k\gamma^2(k)\\ &\le 4(m-n+1)^{1-2\alpha}\sum_{k=m-n+1}^{n-1}k^{2\alpha}\gamma^2(k)\le 4(m-n+1)^{1-2\alpha}|f|^2_{2,\alpha},\end{aligned}$$

where $\alpha>1/2$. This proves the first relation. For the second, recall that $|f|^2_{2,\alpha}\le M$ for $f\in\Sigma$, and invoke Lemma 2.1 together with the subsequent remark on the eigenvalues of $\Gamma_n(f)$. $\square$

Define the experiment

$$\tilde{\mathcal{E}}_{n,m}=\big(N_n(0,\tilde{\Gamma}_{n,m}(f)),\ f\in\Sigma\big),$$

then (3.3) implies $\mathcal{E}_n\simeq\tilde{\mathcal{E}}_{n,m}$ if $m-n\to\infty$. Moreover, we have $\tilde{\mathcal{E}}_{n,m}\preceq\tilde{\mathcal{E}}_m$ by definition, thus

$$\mathcal{E}_n\precsim\tilde{\mathcal{E}}_m$$

in case $m-n\to\infty$. We know that $\tilde{\mathcal{E}}_m$ is equivalent [via the linear transformation $(2\pi)^{-1/2}U'$] to observing data $\tilde{z}^{(n)}$ given by (2.13). Define $\mathring{\mathcal{E}}_n$ by

$$\mathring{\mathcal{E}}_n=\big(N_n(0,\mathring{\Gamma}_n(f)),\ f\in\Sigma\big),\tag{3.4}$$

where

$$\mathring{\Gamma}_n(f)=\mathrm{Diag}(J_{j,n}(f))_{j=1,\dots,n}.$$

Note that the data $z_1,\dots,z_n$ in Theorem 1.1 are represented by $\mathring{\mathcal{E}}_n$. We shall also write $\mathring{z}^{(n)}$ for their vector, so that $\mathcal{L}(\mathring{z}^{(n)}|f)=N_n(0,\mathring{\Gamma}_n(f))$.

PROPOSITION 3.2. *We have $\mathring{\mathcal{E}}_n \approx \tilde{\mathcal{E}}_n$, with corresponding equivalence maps (Markov kernels) as follows. Let $\tilde{y}^{(n)}$ and $\mathring{z}^{(n)}$ be data in $\tilde{\mathcal{E}}_n$ and $\mathring{\mathcal{E}}_n$, respectively. Then, for the orthogonal matrix U_n given by* (2.11),

$$(2\pi)^{-1/2} U_n' \tilde{y}^{(n)} \simeq \mathring{z}^{(n)} \quad \textit{and} \quad (2\pi)^{1/2} U_n \mathring{z}^{(n)} \simeq \tilde{y}^{(n)}.$$

PROOF. Note that our first claim can also be written $\tilde{z}^{(n)} \simeq \mathring{z}^{(n)}$, where $\tilde{z}^{(n)}$ is from (2.13). To describe $\mathcal{L}(\tilde{z}^{(n)}|f)$, define $\delta_j = \tilde{f}_n(\omega_{j-(n+1)/2})$ for $j = 1, \dots, n$ and a $n \times n$ covariance matrix

$$\Delta_n(f) = \mathrm{Diag}(\delta_j)_{j=1,\dots,n}.$$

Then $\mathcal{L}(\tilde{z}^{(n)}|f) = N_n(0, \Delta_n(f))$. The conditions on f (see also Lemma 5.6 in the technical report [17]) imply that uniformly over $j = 1, \dots, n$,

$$J_{j,n}(f) \geq C^{-1}, \qquad J_{j,n}(f) \leq C$$

for some $C > 0$ not depending on f and n. Now apply Lemma 2.1 to obtain

$$\begin{aligned} H^2(N_n(0, \mathring{\Gamma}_n(f)), N_n(0, \Delta_n(f))) &\leq C \|\mathring{\Gamma}_n(f) - \Delta_n(f)\|^2 \\ &= C \sum_{j=1}^{n} \bigl(J_{j,n}(f) - \delta_j\bigr)^2. \end{aligned}$$

By Lemma 5.7 in [17] this is $o(1)$ uniformly in f. This implies the first relation $\simeq$. The second relation is an obvious consequence. $\square$

For a choice $m = n + r_n$, $r_n = 2[\log(n/2)]$, we immediately obtain the following result. Define the upper bracket Gaussian scale experiment $\mathring{\mathcal{E}}_{u,n}$ by

$$\mathring{\mathcal{E}}_{u,n} := \mathring{\mathcal{E}}_{n+r_n}. \tag{3.5}$$

COROLLARY 3.3. *Consider experiments $\mathcal{E}_n$ and $\mathring{\mathcal{E}}_{u,n}$ given, respectively, by* (1.3) *and* (3.5), (3.4) *with parameter space $\Sigma = \Sigma_{\alpha,M}$, where $M > 0$, $\alpha > 1/2$. Then as $n \to \infty$,*

$$\mathcal{E}_n \precsim \mathring{\mathcal{E}}_{u,n}.$$

4. Lower informativity bracket. The upper bound (2.15) for the Hellinger distance of $y^{(n)}$ and the periodic process $\tilde{y}^{(n)}$ which does not tend to 0, can be improved in a certain sense if f is restricted to a shrinking neighborhood, $\Sigma_n(f_0)$ say, of some $f_0 \in \Sigma$. At this stage, f_0 is assumed known, so the covariance matrices $\Gamma_n(f)$ and $\tilde{\Gamma}_n(f)$ can be used for a linear transformation of $y^{(n)}$ which brings it closer to the periodic process $\tilde{y}^{(n)}$. The linear transformation of $y^{(n)}$, which depends on f_0, can be construed as a Markov kernel mapping which yields asymptotic equivalence $\mathcal{E}_n(f_0) \approx \tilde{\mathcal{E}}_n(f_0)$, if these are the versions of $\mathcal{E}_n$ and $\tilde{\mathcal{E}}_n$ with f restricted to $f \in \Sigma_n(f_0)$.

Such a local asymptotic equivalence can be globalized in a standard way (cf. [18, 26]) if *sample splitting* were available in both global experiments $\mathcal{E}_n$ and $\tilde{\mathcal{E}}_n$. For the original stationary process, that would mean that observing a series of size n is equivalent to observing two independent series of size approximately $n/2$. We will establish an asymptotic version of sample splitting for $y^{(n)}$ which involves omitting a fraction of the sample in the center of the series, that is, omitting terms with index near $n/2$. The ensuing loss of information means that the globalization procedure only yields a lower asymptotic informativity bracket for $\mathcal{E}_n$, that is, a sequence $\tilde{\mathcal{E}}^{\#}_{3,n}$ such that $\tilde{\mathcal{E}}^{\#}_{3,n} \precsim \mathcal{E}_n$. The experiment $\tilde{\mathcal{E}}^{\#}_{3,n}$ will be made up of two independent periodic processes with the same parameter f and with a sample size $m \sim (n - \log n)/2$. Each of these is equivalent to a Gaussian scale model (2.13) with n replaced by m; further arguments show that observing these two is asymptotically equivalent to a Gaussian scale model $\mathring{\mathcal{E}}_{l,n} := \mathring{\mathcal{E}}_{2m}$ with grid size $2m \sim n - \log n$.

A crucial step now consists in showing that in the Gaussian scale models $\mathring{\mathcal{E}}_n$, the grid size n can be replaced by $n - \log n$ or $n + \log n$. This step is an analog, for the special regression model, of the well-known reasoning in the i.i.d. case that additional observations may be asymptotically negligible (cf. Mammen [25] for parametric i.i.d. models, Low and Zhou [24] for the nonparametric case). Thus it follows that the lower and upper bracketing experiments $\mathring{\mathcal{E}}_{l,n}$, $\mathring{\mathcal{E}}_{u,n}$ are both asymptotically equivalent to $\mathring{\mathcal{E}}_n$, and the relations

$$\mathring{\mathcal{E}}_{l,n} \precsim \mathcal{E}_n \precsim \mathring{\mathcal{E}}_{u,n},$$

then imply $\mathcal{E}_n \approx \mathring{\mathcal{E}}_n$, that is, Theorem 1.1.

4.1. *Local experiments*. Let $\varkappa_n$ be a sequence $\varkappa_n \searrow 0$, fixed in the sequel. A specific choice of $\varkappa_n$ will be made in Section 4.4 below [see (4.11)]. Let $\|\cdot\|_\infty$ be the sup-norm for real functions defined on $[-\pi, \pi]$, that is,

$$\|f\|_\infty = \sup_{\omega \in [-\pi,\pi]} |f(\omega)|$$

and for $f_0 \in \Sigma$ define shrinking neighborhoods

$$\Sigma_n(f_0) = \{f \in \Sigma : \|f - f_0\|_\infty + \|f - f_0\|_{2,1/2} \le \varkappa_n\}. \tag{4.1}$$

The restricted experiments are

$$\mathcal{E}_n(f_0) = \bigl(N_n(0, \Gamma_n(f)), f \in \Sigma_n(f_0)\bigr),$$
$$\tilde{\mathcal{E}}_n(f_0) = \bigl(N_n(0, \tilde{\Gamma}_n(f)), f \in \Sigma_n(f_0)\bigr).$$

For shortness write $\Gamma = \Gamma_n(f)$, $\Gamma_0 = \Gamma_n(f_0)$ and similarly $\tilde{\Gamma} = \tilde{\Gamma}_n(f)$, $\tilde{\Gamma}_0 = \tilde{\Gamma}_n(f_0)$. Define a matrix

$$K_n = K_n(f_0) = \tilde{\Gamma}_0^{1/2}\Gamma_0^{-1/2} \tag{4.2}$$

and in experiment $\mathcal{E}_n(f_0)$, consider transformed observations

$$\check{y}^{(n)} := K_n(f_0)y^{(n)}.$$

Consider also the experiment $\mathcal{E}_n^*(f_0)$ given by the laws of $\check{y}^{(n)}$, that is,

$$\mathcal{E}_n^*(f_0) = \big(N_n(0, K_n(f_0)\Gamma_n(f)K_n'(f_0)), f \in \Sigma_n(f_0)\big).$$

Clearly $\mathcal{E}_n(f_0) \sim \mathcal{E}_n^*(f_0)$; the next result proves that $\mathcal{E}_n^*(f_0) \simeq \tilde{\mathcal{E}}_n(f_0)$, and thus $\mathcal{E}_n(f_0) \approx \tilde{\mathcal{E}}_n(f_0)$.

LEMMA 4.1. *We have*

$$\sup_{f_0 \in \Sigma} \sup_{f \in \Sigma_n(f_0)} H^2(N_n(0, K_n(f_0)\Gamma_n(f)K_n'(f_0)), N_n(0, \tilde{\Gamma}_n(f)) \le C\varkappa_n.$$

PROOF. In view of Lemma 2.1, it suffices to show that

$$\sup_{f \in \Sigma}\big(\lambda_{\max}(\tilde{\Gamma}_n) + \lambda_{\min}^{-1}(\tilde{\Gamma}_n)\big) \le C \tag{4.3}$$

and that

$$\|K_n\Gamma_nK_n' - \tilde{\Gamma}_n\|^2 \le C\varkappa_n.$$

Note that

$$\lambda_{\max}(\tilde{\Gamma}) = \max_{|j| \le (n-1)/2} |\tilde{f}_n(\omega_j)|, \qquad \lambda_{\min}(\tilde{\Gamma}) = \min_{|j| \le (n-1)/2} |\tilde{f}_n(\omega_j)|$$

and that Lemma 5.6. in [17] implies

$$\sup_{f \in \Sigma} \|f - \tilde{f}_n\|_\infty \to 0.$$

Hence (4.3) follows immediately from $f \in \Sigma$, more specifically, the fact that values of f are uniformly bounded and bounded away from 0. According to Proposition 4.5.3 in [3], the assumption $f \in \Sigma$ also implies a corresponding property for Γ, that is,

$$\sup_{f \in \Sigma}\big(\lambda_{\max}(\Gamma_n) + \lambda_{\min}^{-1}(\Gamma_n)\big) \le C. \tag{4.4}$$

Note that the eigenvalues of Γ_0 and $\tilde{\Gamma}_0$ share property (4.3) since $f_0 \in \Sigma$.

Set $G = \Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2}$ and $\tilde{G} = \tilde{\Gamma}_0^{-1/2}\tilde{\Gamma}\tilde{\Gamma}_0^{-1/2}$. Since

$$\|K_n\Gamma_nK_n' - \tilde{\Gamma}_n\| \le |\tilde{\Gamma}_0|\|G - \tilde{G}\|,$$

it now suffices to show that

$$\|G - \tilde{G}\| \le C\varkappa_n. \tag{4.5}$$

To establish (4.5), denote $\Delta = \Gamma - \Gamma_0$, $\tilde{\Delta} = \tilde{\Gamma} - \tilde{\Gamma}_0$ and observe

$$\begin{aligned}\|G - \tilde{G}\| &= \|\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2} - \tilde{\Gamma}_0^{-1/2}\tilde{\Gamma}\tilde{\Gamma}_0^{-1/2}\| \\ &= \|\Gamma_0^{-1/2}\Delta\Gamma_0^{-1/2} - \tilde{\Gamma}_0^{-1/2}\tilde{\Delta}\tilde{\Gamma}_0^{-1/2}\| \\ &\le \|\Gamma_0^{-1/2}(\Delta - \tilde{\Delta})\Gamma_0^{-1/2}\| + \|\Gamma_0^{-1/2}\tilde{\Delta}\Gamma_0^{-1/2} - \tilde{\Gamma}_0^{-1/2}\tilde{\Delta}\tilde{\Gamma}_0^{-1/2}\|.\end{aligned} \tag{4.6}$$

We shall now estimate the two terms on the right-hand side separately. By elementary properties of eigenvalues, we obtain

$$\|\Gamma_0^{-1/2}(\Delta - \tilde{\Delta})\Gamma_0^{-1/2}\| \le |\Gamma_0^{-1}|\|\Delta - \tilde{\Delta}\|,$$

where $|\Gamma_0^{-1}| \le C$, and according to Lemma 2.2(ii),

$$\|\Delta - \tilde{\Delta}\|^2 \le 2|f - f_0|_{2,1/2}^2.$$

Furthermore,

$$\begin{aligned}&\|\Gamma_0^{-1/2}\tilde{\Delta}\Gamma_0^{-1/2} - \tilde{\Gamma}_0^{-1/2}\tilde{\Delta}\tilde{\Gamma}_0^{-1/2}\| \\ &\quad = \|(\Gamma_0^{-1/2} - \tilde{\Gamma}_0^{-1/2})\tilde{\Delta}\Gamma_0^{-1/2} + \tilde{\Gamma}_0^{-1/2}\tilde{\Delta}(\Gamma_0^{-1/2} - \tilde{\Gamma}_0^{-1/2})\| \\ &\quad \le 2C|\tilde{\Delta}|\|\Gamma_0^{-1/2} - \tilde{\Gamma}_0^{-1/2}\| = C|\tilde{\Delta}|\|\Gamma_0^{-1/2}(\tilde{\Gamma}_0^{1/2} - \Gamma_0^{1/2})\tilde{\Gamma}_0^{-1/2}\| \\ &\quad \le C|\tilde{\Delta}|\|\Gamma_0^{1/2} - \tilde{\Gamma}_0^{1/2}\|.\end{aligned}$$

Applying Lemma 5.1 (Section 5.2) in the technical report [17] and Lemma 2.2(i), we obtain

$$\|\Gamma_0^{1/2} - \tilde{\Gamma}_0^{1/2}\|^2 \le C\|\Gamma_0 - \tilde{\Gamma}_0\|^2 \le C|f_0|_{2,1/2}^2.$$

Here $|f_0|_{2,1/2}^2 \le |f_0|_{2,\alpha}^2 \le M$. Collecting these estimates yields

$$\|G - \tilde{G}\|^2 \le C(|f - f_0|_{2,1/2}^2 + |\tilde{\Delta}|^2).$$

To complete the proof, it suffices to note that, since $\tilde{\Gamma}$ and $\tilde{\Gamma}_0$ have the same set of eigenvectors [cf. (2.3) and (2.6)–(2.8)],

$$\begin{aligned}|\tilde{\Delta}|^2 = \lambda_{\max}(\tilde{\Gamma} - \tilde{\Gamma}_0)^2 &= (2\pi)^2 \max_{|j| \le (n-1)/2} \bigl(|\tilde{f}_n(\omega_j) - \tilde{f}_{0,n}(\omega_j)|^2\bigr) \\ &\le C\|\tilde{f}_n - \tilde{f}_{0,n}\|_\infty^2 \le C\|f - f_0\|_\infty^2 + Cn^{1-2\alpha}\log n\|f - f_0\|_{2,\alpha}^2,\end{aligned}$$

where the last inequality is a consequence of Lemma 5.6 in [17]. Hence $|\tilde{\Delta}| \le C\varkappa_n$, which establishes (4.5). $\square$

4.2. *Sample splitting*. Consider sample splitting for a stationary process: take the observed $y^{(n)} = (y(1), \dots, y(n))$ and omit r observations in the center of the series. Recall that n was assumed uneven; assume now also r to be uneven and set $m = (n-r)/2$, then the result is the series $y(1), \dots, y(m), y(n-m+1), \dots, y(n)$. The total covariance matrix for these reduced data is

$$\Gamma_{n,0}^{(m)}(f) := \begin{pmatrix} \Gamma_m(f) & A_{n,m} \\ A'_{n,m} & \Gamma_m(f) \end{pmatrix},$$

where the $m \times m$ matrix $A_{n,m} = A_{n,m}(f)$ contains only covariances $\gamma_f(r+1)$, $\gamma_f(r+2)$ and of higher order. In fact, A is the upper right $m \times m$ submatrix of $\Gamma_n(f)$, that is,

$$A_{n,m} = \begin{pmatrix} \cdots & \gamma(n-2) & \gamma(n-1) \\ \gamma(r+2) & \cdots & \gamma(n-2) \\ \gamma(r+1) & \gamma(r+2) & \cdots \end{pmatrix}.$$

In the sequel, we set $r_n = 2[\log n/2] + 1$, and thus $r_n \sim \log n$, $m = (n - r_n)/2$. In the corresponding experiment, we denote

$$\mathcal{E}_{0,n}^{\#} = \bigl(N_{2m}\bigl(0, \Gamma_{n,0}^{(m)}(f)\bigr), f \in \Sigma\bigr).$$

Consider also the experiment where two independent stationary series of length m are observed, $y_1^{(m)}$ and $y_2^{(m)}$, say. The corresponding experiment is

$$\mathcal{E}_{1,n}^{\#} := \bigl(N_{2m}\bigl(0, \Gamma_{n,1}^{(m)}(f)\bigr), f \in \Sigma\bigr), \tag{4.7}$$

where

$$\Gamma_{n,1}^{(m)}(f) := \begin{pmatrix} \Gamma_m(f) & 0_{m\times m} \\ 0_{m\times m} & \Gamma_m(f) \end{pmatrix}.$$

Clearly, we have $\mathcal{E}_{0,n}^{\#} \preceq \mathcal{E}_n$.

PROPOSITION 4.2. $\mathcal{E}_{0,n}^{\#} \simeq \mathcal{E}_{1,n}^{\#}$.

PROOF. Use Lemma 2.1 to compute the Hellinger distance. Take $A = \Gamma_{n,1}^{(m)}$; then the eigenvalues of A are those of $\Gamma_m(f)$, so that (4.4) can be invoked. The squared distance of the covariance matrices $\Gamma_{n,0}^{(m)}$ and $\Gamma_{n,1}^{(m)}$ is

$$\begin{aligned} \|\Gamma_{n,0}^{(m)} - \Gamma_{n,1}^{(m)}\|^2 = 2\|A_{n,m}\|^2 &\le 2 \sum_{k=r+1}^{n-1} (k-r)\gamma^2(k) \\ &\le 2 \sum_{k=r+1}^{n-1} k\gamma^2(k) \le (r+1)^{1-2\alpha} |f|_{2,\alpha}^2. \end{aligned}$$

Since $r_n \to \infty$, the result follows. $\square$

We have shown that two independent stationary sequences of length $m = (n - r_n)/2$ are asymptotically less informative than one sequence of length n. Having obtained a method of sample splitting for stationary sequences (with some loss of information), we can now use a localization argument to complete the proof of the lower bound.

4.3. *Preliminary estimators.* For the globalization procedure, we need existence of an estimator $\hat{f}_n$, in both of the global experiments $\mathcal{E}_n$ and $\tilde{\mathcal{E}}_n$ (or $\mathring{\mathcal{E}}_n$), such that $\hat{f}_n$ takes values in Σ and

$$\|\hat{f}_n - f\|_\infty + \|\hat{f}_n - f\|_{2,1/2} = o_p(1)$$

uniformly over $f \in \Sigma$. More specifically, a rate $o_p(\kappa_n)$ with κ_n from (4.1) is needed in the above result, but κ_n has not been selected so far, and will be determined based on the results of this section [cf. (4.11) below]. Select $\beta \in (1/2, \alpha)$ and consider the norm $\|f\|_{2,\beta}$ according to (2.14). Note that $\|f\|_{2,1/2} \le C\|f\|_{2,\beta}$ and that according to Lemma 5.6 in [17], we have $\|f\|_\infty \le C\|f\|_{2,\beta}$; therefore, it suffices to show

$$\|\hat{f}_n - f\|_{2,\beta}^2 = o_p(1). \tag{4.8}$$

For this, we shall use a standard truncated orthogonal series estimator and then modify it to take values in Σ. The empirical autocovariance function is

$$\hat{\gamma}_n(k) = \frac{1}{n-k}\sum_{j=1}^{n-k} y(j)y(k+j), \qquad k = 0, \dots, n-1.$$

We have unbiasedness: $E\hat{\gamma}_n(k) = \gamma_f(k)$. The following two lemmas concerning this spectral density estimator are elementary; proofs can be found in [17] (cf. also [28], (VI.4.5–6)).

LEMMA 4.3. *For any spectral density $f \in L_2(-\pi, \pi)$, and any $k = 0, \dots, n-1$,*

$$\operatorname{Var}\hat{\gamma}_n(k) \le \frac{5}{n-k}\sum_{j=0}^{n-1}\gamma_f^2(j).$$

For the orthogonal series estimator, define a truncation index $\tilde{n} = [n^{1/(2\alpha+1)}]$ and set

$$\hat{f}_n(\omega) = \sum_{|k|\le\tilde{n}} \hat{\gamma}_n(k)\exp(ik\omega), \qquad \omega \in [-\pi, \pi]. \tag{4.9}$$

LEMMA 4.4. *In the experiment $\mathcal{E}_n$ the estimator $\hat{f}_n$ fulfills for any $\beta \in (1/2, \alpha)$, and any $\gamma \in (0, \frac{\alpha-\beta}{2\alpha+1})$*

$$\sup_{f\in\Sigma} P(\|\hat{f}_n - f\|_{2,\beta}^2 > n^{-\gamma}) \to 0. \tag{4.10}$$

We now turn to preliminary estimation in the periodic experiment $\tilde{\mathcal{E}}_n$ with data vector $\tilde{y}^{(n)}$. Note that this data vector can be construed as coming from a stationary sequence with autocoviance function $\tilde{\gamma}_{(n)}(\cdot)$ given by (2.1) for $|k| \le n-1$ and $\tilde{\gamma}_{(n)}(k) = 0$ for $|k| > n-1$, that is, the stationary sequence having spectral density $\tilde{f}_n$. Thus if $\hat{\gamma}_n(k)$ again denotes the empirical autocoviance function in this series, then we can apply Lemma 4.3 to obtain

$$\operatorname{Var} \hat{\gamma}_n(k) \le \frac{5}{n-k} \sum_{j=0}^{n-1} \tilde{\gamma}^2_{(n),f}(j), \qquad k = 0, \dots, n-1.$$

Obviously,

$$\sum_{k=0}^{n-1} \tilde{\gamma}^2_{(n),f}(k) = \sum_{k=0}^{(n-1)/2} \gamma_f^2(k) + \sum_{k=1}^{(n-1)/2} \gamma_f^2(k) \le 2\|f\|_2^2.$$

Now use the estimator (4.9) with $\tilde{n}$ as above; since $\tilde{n} = o((n-1)/2)$, we have the unbiasedness

$$E\hat{\gamma}_n(k) = \gamma_f(k), \qquad k = 0, \dots, \tilde{n}.$$

Thus the proof of the following result is entirely analogous to Lemma 4.4; the estimator $\hat{f}_n$ is also formally the same function of the data.

LEMMA 4.5. *In the experiment $\tilde{\mathcal{E}}_n$, the estimator $\hat{f}_n$ fulfills* (4.10) *for any $\beta \in (1/2, \alpha)$ and any $\gamma \in (0, \frac{\alpha-\beta}{2\alpha+1})$.*

Finally, consider modifications such that the estimator takes values in $\Sigma_{\alpha,M}$. Consider the space $W^\beta = \{f \in L_2(-\pi,\pi) : \|f\|^2_{2,\beta} < \infty\}$; this is a periodic fractional Sobolev space which is Hilbert under the norm $\|f\|_{2,\beta}$. There the set $\Sigma_{\alpha,M}$ is compact and convex; hence there exists a ($\|\cdot\|_{2,\beta}$-continuous) projection operator Π onto $\Sigma_{\alpha,M}$ in W^β (cf. [2], Definition 1.4.1). Then

$$\|\Pi(\hat{f}_n) - f\|_{2,\beta} \le \|\hat{f}_n - f\|_{2,\beta}.$$

The modified estimators $\Pi(\hat{f}_n)$, thus, again fulfill (4.10). A summary of results in this section is the following.

PROPOSITION 4.6. *In both experiments, $\mathcal{E}_n$ and $\tilde{\mathcal{E}}_n$, there are estimators $\hat{f}_n$ taking values in Σ and fulfilling for any $\gamma \in (0, \frac{\alpha-1/2}{2\alpha+1})$,*

$$\sup_{f\in\Sigma} P(\|\hat{f}_n - f\|_\infty + \|\hat{f}_n - f\|_{2,1/2} > n^{-\gamma}) \to 0.$$

4.4. *Globalization*. In this section, we denote

$$P_{f,n} := \mathcal{L}(y^{(n)}|f) = N_n(0, \Gamma_n(f)), \qquad \tilde{P}_{f,n} := \mathcal{L}(\tilde{y}^{(n)}|f) = N_n(0, \tilde{\Gamma}_n(f)).$$

Consider again the experiment $\mathcal{E}^{\#}_{1,n}$ of (4.7), where two independent stationary series $y_1^{(m)}$ and $y_2^{(m)}$ of length $m = (n - r_n)/2$ are observed. In modified notation, we now write

$$\mathcal{E}^{\#}_{1,n} = \mathcal{E}_m \otimes \mathcal{E}_m = (P_{f,m} \otimes P_{f,m}, f \in \Sigma).$$

We shall compare this with the experiments

$$\mathcal{E}^{\#}_{2,n} := \mathcal{E}_m \otimes \tilde{\mathcal{E}}_m = (P_{f,m} \otimes \tilde{P}_{f,m}, f \in \Sigma),$$

$$\mathcal{E}^{\#}_{3,n} := \tilde{\mathcal{E}}_m \otimes \tilde{\mathcal{E}}_m = (\tilde{P}_{f,m} \otimes \tilde{P}_{f,m}, f \in \Sigma).$$

At this point, select the shrinking rate κ_n of the neighborhoods $\Sigma_n(f_0)$ [cf. (4.1)] as

$$\kappa_n = n^{-\gamma}, \qquad \gamma = \frac{\alpha - 1/2}{2(2\alpha + 1)}. \tag{4.11}$$

PROPOSITION 4.7. *We have* $\mathcal{E}^{\#}_{2,n} \precsim \mathcal{E}^{\#}_{1,n}$.

PROOF. We shall construct a sequence of Markov kernels M_n such that

$$\sup_{f \in \Sigma} H^2(P_{f,m} \otimes \tilde{P}_{f,m}, M_n(P_{f,m} \otimes P_{f,m})) \to 0.$$

Define M_n as follows: given $y_1^{(m)}$ and $y_2^{(m)}$, and A, a measurable subset of $\mathbb{R}^{2m}$, set

$$M_n(A, y_1^{(m)}, y_2^{(m)}) = \mathbf{1}_A(y_1^{(m)}, K_m(\hat{f}_m(y_1^{(m)}))y_2^{(m)}),$$

where $K_m(f)$ is the matrix defined by (4.2), that is, for $f \in \Sigma$ by

$$K_m(f) = \tilde{\Gamma}_m^{1/2}(f)\Gamma_m^{-1/2}(f)$$

and $\hat{f}_m$ is the estimator in $\mathcal{E}_m$ of Proposition 4.6 applied to data $y_1^{(m)}$. Thus the Markov kernel M_n is in fact a deterministic map, that is, given $y_1^{(m)}, y_2^{(m)}$, it defines a one-point measure on $\mathbb{R}^{2m}$ concentrated in $(y_1^{(m)}, K_m(\hat{f}_m(y_1^{(m)}))y_2^{(m)})$. Thus the law $M_n(P_{f,m} \otimes P_{f,m})$ is the joint law of $y_1^{(m)}$ and $K_m(\hat{f}_m(y_1^{(m)}))y_2^{(m)}$ under f. The latter, we split up into the marginal law of $y_1^{(m)}$, that is, $P_{f,m}$ and the conditional law of $K_m(\hat{f}_m(y_1^{(m)}))y_2^{(m)}$ given $y_1^{(m)}$; write $P^K_{f,m}|y_1^{(m)}$ for the latter. We have

$$P^K_{f,m}|y_1^{(m)} = N_n(0, K\Gamma_m(f)K') \qquad \text{for } K = K_m(\hat{f}_m(y_1^{(m)})).$$

Now clearly,

$$H^2(P_{f,m} \otimes \tilde{P}_{f,m}, M_n(P_{f,m} \otimes P_{f,m})) = E_f H^2(\tilde{P}_{f,m}, P^K_{f,m}|y_1^{(m)}), \tag{4.12}$$

where E_f is taken w.r.t. $y_1^{(m)}$ under $P_{f,m}$. Define

$$B_{f,m} := \{y \in \mathbb{R}^m : \|\hat{f}_m(y) - f\|_\infty + \|\hat{f}_m(y) - f\|_{2,1/2} \le \kappa_m\}.$$

By definition of $\Sigma_m(f_0)$ [cf. (4.1)], we have $f \in \Sigma_m(\hat{f}_m(y))$ if $y \in B_{f,m}$. Thus Lemma 4.1 implies

$$\sup_{y \in B_{f,m}, f \in \Sigma} H^2(\tilde{P}_{f,m}, P^K_{f,m}|y) = o(1).$$

Moreover, by Proposition 4.6,

$$P_{f,m}(B^c_{f,m}) = o(1), \qquad \text{uniformly over } f \in \Sigma. \tag{4.13}$$

Hence

$$\begin{aligned} E_f H^2\big(\tilde{P}_{f,m}, P^K_{f,m}|y_1^{(m)}\big) &= \int_{B_{f,m}} H^2(\tilde{P}_{f,m}, P^K_{f,m}|y)\, dP_{f,m}(y) + o(1) \\ &= o(1) P_{f,m}(B_{f,m}) + o(1) = o(1) \end{aligned} \tag{4.14}$$

uniformly over $f \in \Sigma$. In conjunction with (4.12), the last relation proves the claim. □

The next result is entirely analogous if we replace the estimator $\hat{f}_m$ based on data $y^{(m)}$ by the one based on data $\tilde{y}^{(m)}$, and formally reverse the order in the product $P_{f,m} \otimes \tilde{P}_{f,m}$.

PROPOSITION 4.8. *We have* $\mathcal{E}^{\#}_{3,n} \precsim \mathcal{E}^{\#}_{2,n}$.

PROOF. We construct a sequence of Markov kernels $\tilde{M}_n$ such that

$$\sup_{f \in \Sigma} H^2\big(\tilde{P}_{f,m} \otimes \tilde{P}_{f,m}, \tilde{M}_n(P_{f,m} \otimes \tilde{P}_{f,m})\big) \to 0.$$

Define $\tilde{M}_n$ as follows: given $y_1^{(m)}$ and $\tilde{y}_2^{(m)}$, and A, a measurable subset of $\mathbb{R}^{2m}$, set

$$\tilde{M}_n\big(A, y_1^{(m)}, \tilde{y}_2^{(m)}\big) = \mathbf{1}_A\big(K_m\big(\hat{f}_m(\tilde{y}_2^{(m)})\big) y_1^{(m)}, \tilde{y}_2^{(m)}\big),$$

where $\hat{f}_m$ is the estimator defined in the previous subsection, applied to data $\tilde{y}_2^{(m)}$. Analogously to (4.13), we have

$$\tilde{P}_{f,m}(B^c_{f,m}) = o(1), \qquad \text{uniformly over } f \in \Sigma.$$

A reasoning as in (4.14) completes the proof. □

For the experiment $\mathcal{E}^{\#}_{3,n}$ which consists of product measures $\tilde{P}_{f,m} \otimes \tilde{P}_{f,m}$, we can invoke Proposition 3.2, applying the equivalence map given there componentwise [i.e., to independent components $(\tilde{y}_1^{(m)}, \tilde{y}_2^{(m)})$ in $\mathcal{E}^{\#}_{3,n}$]. A summary of the lower informativity bound results so far can thus be given as follows. For $r_n = 2[\log(n/2)]$ define the lower bracket Gaussian scale experiment $\mathring{\mathcal{E}}_{l,n}$ by

$$\mathring{\mathcal{E}}_{l,n} := \mathring{\mathcal{E}}_{(n-r_n)/2} \otimes \mathring{\mathcal{E}}_{(n-r_n)/2}. \tag{4.15}$$

COROLLARY 4.9. *Consider experiments $\mathcal{E}_n$ and $\mathring{\mathcal{E}}_{l,n}$ given, respectively, by* (1.3) *and* (4.15), (3.4) *with parameter space $\Sigma = \Sigma_{\alpha,M}$, where $M > 0$, $\alpha > 1/2$. Then as $n \to \infty$,*

$$\mathring{\mathcal{E}}_{l,n} \precsim \mathcal{E}_n.$$

4.5. *Bracketing the Gaussian scale model.* The proof of Theorem 1.1 is complete if the lower and upper informativity bounds $\mathring{\mathcal{E}}_{l,n}$ and $\mathring{\mathcal{E}}_{u,n}$ coincide in an asymptotic sense. Since we already established the relation $\mathring{\mathcal{E}}_{l,n} \precsim \mathcal{E}_n \precsim \mathring{\mathcal{E}}_{u,n}$ (Corollaries 3.3, 4.9), it now suffices to show that $\mathring{\mathcal{E}}_{u,n} \precsim \mathring{\mathcal{E}}_{l,n}$. This essentially means that in the special nonparametric regression model $\mathring{\mathcal{E}}_n$ of Gaussian scale type, having r_n additional observations does not matter asymptotically. "Additional observations" here refers to an equidistant design of higher grid size. The problem of additional observations for i.i.d. models has been discussed by Le Cam [22] and Mammen [25] under parametric assumptions. For nonparametric i.i.d. models, one can use the approximation by Gaussian white noise or Poisson models to bound the influence of additional observations. For simplicity, consider a Gaussian white noise model on [0, 1]

$$dZ_t = f(t)\,dt + n^{-1/2}\,dW_t, \qquad t \in [0,1], f \in \Sigma,$$

with parameter space Σ. Consider this experiment $\mathcal{F}_n$, say, and also $\mathcal{F}_{n+r_n}$. Multiplying the data by $n^{1/2}$ gives an equivalent experiment

$$dZ_t^* = n^{1/2} f(t)\,dt + dW_t, \qquad t \in [0,1], f \in \Sigma,$$

and the corresponding one for $(n+r_n)^{1/2}$. Now, for given f, the squared Hellinger distance of the two respective measures is bounded by

$$\begin{aligned} &C\big((n+r_n)^{1/2} - n^{1/2}\big)^2 \|f\|_2^2 \\ &\quad = C\frac{r_n^2}{n}\big(1+o(1)\big)\|f\|_2, \end{aligned}$$

if $r_n = o(n)$. Thus if $r_n = o(n^{1/2})$ and $\sup_{f\in\Sigma} \|f\|_2 \le C$, then we have $\mathcal{F}_n \approx \mathcal{F}_{n+r_n}$.

Comparable results can be obtained for nonparametric i.i.d. and regression models if these can be approximated by $\mathcal{F}_n$. In the present case, conversely, for the nonparametric Gaussian scale regression $\mathring{\mathcal{E}}_n$, a result of type $\mathring{\mathcal{E}}_n \approx \mathring{\mathcal{E}}_{n+r_n}$ is a prerequisite for the Gaussian location (white noise) approximation. Note that for a narrower parameter space, given by a Lipschitz class, the white noise approximation of $\mathring{\mathcal{E}}_n$ has been established (cf. [18]).

REMARK 4.10. The relation

$$\mathring{\mathcal{E}}_{l,n} \precsim \mathcal{E}_n \precsim \mathring{\mathcal{E}}_{u,n} \tag{4.16}$$

has been proved under the technical assumption that n is uneven. If n is even, note first that $\mathcal{E}_{n-1} \precsim \mathcal{E}_n \precsim \mathcal{E}_{n+1}$ (omitting one observation from $\mathcal{E}_{n+1}$ and $\mathcal{E}_n$), and apply (4.16) to obtain

$$\mathring{\mathcal{E}}_{l,n-1} \precsim \mathcal{E}_n \precsim \mathring{\mathcal{E}}_{u,n+1}.$$

The relation $\mathcal{E}_{u,n} \precsim \mathcal{E}_{l,n}$, which will be proved for uneven n in the remainder of this section, is easily seen to extend to $\mathcal{E}_{u,n+2} \precsim \mathcal{E}_{l,n}$. This suffices to establish the main result Theorem 1.1 for general sample size $n \to \infty$.

4.5.1. *First part of the bracketing argument.* Denote again $m = (n - r_n)/2$, where $r_n = 2[(\log n)/2] + 1$.

LEMMA 4.11. *For $\mathring{\mathcal{E}}_{l,n} = \mathring{\mathcal{E}}_m \otimes \mathring{\mathcal{E}}_m$, we have*

$$\mathring{\mathcal{E}}_m \otimes \mathring{\mathcal{E}}_m \approx \mathring{\mathcal{E}}_{2m}.$$

PROOF. Note that the measures in $\mathring{\mathcal{E}}_m \otimes \mathring{\mathcal{E}}_m$ are product measures, which can be described, after a rearrangement of components, as

$$Q_{1,m} := \bigotimes_{j=1}^{m} \bigl(N(0, J_{j,m}(f)) \otimes N(0, J_{j,m}(f))\bigr),$$

whereas the measures in $\mathring{\mathcal{E}}_{2m}$ are

$$Q_{2,m} := \bigotimes_{j=1}^{m} \bigl(N(0, J_{2j-1,2m}(f)) \otimes N(0, J_{2j,2m}(f))\bigr).$$

Now Lemma 2.1 yields

$$H^2(Q_{1,m}, Q_{2,m}) \le C \sum_{j=1}^{m} \bigl(\bigl(J_{2j-1,2m}(f) - J_{j,m}(f)\bigr)^2 + \bigl(J_{2j,2m}(f) - J_{j,m}(f)\bigr)^2\bigr).$$

Define a partition of $(-\pi, \pi)$ into n intervals $W_{j,n}$, $j = 1, \ldots, n$, of equal length, and for any $f \in L_2(-\pi, \pi)$, let

$$\bar{f}_n = \sum_{j=1}^{n} J_{j,n}(f) \mathbf{1}_{W_{j,n}} \tag{4.17}$$

be the L_2-projection of f onto piecewise constant functions w.r.t. the partition. Note that we have

$$\|\bar{f}_{2m} - \bar{f}_m\|_2^2 = \frac{2\pi}{m} \sum_{j=1}^{m} \bigl(\bigl(J_{2j-1,2m}(f) - J_{j,m}(f)\bigr)^2 + \bigl(J_{2j,2m}(f) - J_{j,m}(f)\bigr)^2\bigr),$$

so that

$$H^2(Q_{1,m}, Q_{2,m}) \le Cm\|\bar{f}_{2m} - \bar{f}_m\|_2^2 \le Cm(\|f - \bar{f}_{2m}\|_2^2 + \|f - \bar{f}_m\|_2^2).$$

The result now follows from

$$\sup_{f\in\Sigma} m\|f - \bar{f}_m\|_2^2 \to 0, \tag{4.18}$$

which is a consequence of Lemmas 5.3 and 5.5 in [17]. □

4.5.2. *Second part of the bracketing argument.* In view of $\mathring{\mathcal{E}}_{2m} = \mathring{\mathcal{E}}_{n-r_n}$, our next aim is to show

$$\mathring{\mathcal{E}}_{n-r_n} \approx \mathring{\mathcal{E}}_n,$$

where r_n does not grow too quickly. Previously we defined $r_n = 2[(\log n)/2] + 1$, but we will assume more generally now that $r_n = o(n^{1/2})$.

Consider the gamma density with shape parameter $a > 0$,

$$g_a(x) = \frac{1}{\Gamma(a)} x^{a-1}\exp(-x), \qquad x \ge 0,$$

where $\Gamma(a)$ is the gamma function, and more generally, the density with additional scale parameter $s > 0$,

$$g_{a,s}(x) = \frac{1}{\Gamma(a)} s^{-a} x^{a-1}\exp(-xs^{-1}), \qquad x \ge 0.$$

We will call the respective law the $\Gamma(a, s)$ law. Clearly, if $X \sim \Gamma(a, 1)$, then $sX \sim \Gamma(a, s)$. It is well known that $\Gamma(n/2, 2) = \chi_n^2$ and that the following result holds. Assume $X \sim \Gamma(a, s)$ and $Y \sim \Gamma(b, s)$; then $X + Y$, $X/(X + Y)$ are independent random variables, and $X + Y \sim \Gamma(a + b, s)$ while $X/(X + Y)$ has a Beta(a, b) distribution ([1], Theorem B.2.3, page 489).

Furthermore, for fixed $a > 0$, consider the family of laws,

$$\big(\Gamma(a, s), s > 0\big). \tag{4.19}$$

Clearly this is a one parameter exponential family; the shape of this exponential family implies that in a product family,

$$\big(\Gamma^{\otimes n}(a, s), s > 0\big)$$

with n i.i.d. observations $X_1, \dots, X_n$, the sum $\sum_{i=1}^n X_i$ is a sufficient statistic. This sufficient statistic has law $\Gamma(na, s)$; hence for any subset $S \subset (0, \infty)$, we have the equivalence of experiments

$$\big(\Gamma^{\otimes n}(a, s), s \in S\big) \sim \big(\Gamma(na, s), s \in S\big). \tag{4.20}$$

The next two technical results are proved in [17], Lemmas 4.12 and 4.13.

LEMMA 4.12. *For all $a > 0$ and for $s, t > 0$*

$$H^2(\Gamma(a,s),\Gamma(a,t)) = 2\biggl(1 - \biggl(1 - \frac{(s^{1/2}-t^{1/2})^2}{s+t}\biggr)^a\biggr).$$

LEMMA 4.13. *We have, for all $s > 0$ and $a, b > 0$*

$$H^2(\Gamma(a,s),\Gamma(b,s)) = 2\biggl(1 - \frac{\Gamma((a+b)/2)}{(\Gamma(a)\Gamma(b))^{1/2}}\biggr).$$

Now in $\mathring{\mathcal{E}}_n$, we observe [cf. (3.4)]

$$z_j = J_{j,n}^{1/2}(f)\xi_j, \qquad j = 1,\dots,n,$$

for independent standard normals ξ_j, which by sufficiency is equivalent to observing $z_j^2 = J_{j,n}(f)\xi_j^2$. Thus $\mathring{\mathcal{E}}_n$ is equivalent to

$$\mathring{\mathcal{E}}_{n,1} := \Biggl(\bigotimes_{j=1}^{n}\Gamma\bigl(1/2, 2J_{j,n}(f)\bigr), f \in \Sigma\Biggr). \tag{4.21}$$

Set again $m = n - r_n$. The above experiment in turn is equivalent, by the sufficiency argument for the scaled gamma law invoked in (4.20), to

$$\mathring{\mathcal{E}}_{n,m} := \Biggl(\bigotimes_{j=1}^{n}\Gamma^{\otimes m}\bigl(1/2m, 2J_{j,n}(f)\bigr), f \in \Sigma\Biggr).$$

Analogously, we have

$$\mathring{\mathcal{E}}_m \sim \mathring{\mathcal{E}}_{m,1} \sim \mathring{\mathcal{E}}_{m,n} := \Biggl(\bigotimes_{j=1}^{m}\Gamma^{\otimes n}\bigl(1/2n, 2J_{j,m}(f)\bigr), f \in \Sigma\Biggr). \tag{4.22}$$

Introduce an intermediate experiment,

$$\mathring{\mathcal{E}}^*_{m,n} := \Biggl(\bigotimes_{j=1}^{m}\Gamma^{\otimes n}\bigl(1/2m, 2J_{j,m}(f)\bigr), f \in \Sigma\Biggr).$$

LEMMA 4.14. *We have the total variation asymptotic equivalence,*

$$\mathring{\mathcal{E}}^*_{m,n} \simeq \mathring{\mathcal{E}}_{n,m} \qquad \textit{as } n \to \infty.$$

PROOF. Write the measures in $\mathring{\mathcal{E}}_{n,m}$ as a product of mn components, that is, as $\bigotimes_{i=1}^{mn} Q_{1,i}$, where the component measures $Q_{1,i}$ are defined as follows. For every $i = 1,\dots,mn$, let $j(1,i)$ be the unique index $j \in \{1,\dots,n\}$ such that there exists $k \in \{1,\dots,m\}$ for which $i = (j-1)m + k$. Then

$$Q_{1,i} := \Gamma\bigl(1/2m, 2J_{j(1,i),n}(f)\bigr), \qquad i = 1,\dots,mn.$$

Analogously, let $j(2,i)$ be the unique index $j \in \{1, \dots, m\}$ such that there exists $k \in \{1, \dots, n\}$ for which $i = (j-1)n + k$. Then the measures in $\mathring{\mathcal{E}}^*_{m,n}$ are a product of mn components, that is, are $\bigotimes_{i=1}^{mn} Q_{2,i}$, where

$$Q_{2,i} = \Gamma\bigl(1/2m, 2J_{j(2,i),m}(f)\bigr), \qquad i = 1, \dots, mn.$$

Then the Hellinger distance between measures in $\mathring{\mathcal{E}}_{n,m}$ and $\mathring{\mathcal{E}}^*_{m,n}$ is, using Lemma 2.19 in [31] and then Lemma 4.12,

$$\begin{aligned} H^2\Biggl(\bigotimes_{i=1}^{mn} Q_{1,i}, \bigotimes_{i=1}^{mn} Q_{2,i}\Biggr) \\ &\le 2\sum_{i=1}^{mn} H^2(Q_{1,i}, Q_{2,i}) \qquad (4.23)\\ &= 4\sum_{i=1}^{mn}\Biggl(1 - \Biggl(1 - \frac{(J^{1/2}_{j(1,i),n}(f) - J^{1/2}_{j(2,i),m}(f))^2}{J_{j(1,i),n}(f) + J_{j(2,i),m}(f)}\Biggr)^{1/2m}\Biggr). \end{aligned}$$

By using the inequality

$$\frac{(s^{1/2} - t^{1/2})^2}{s+t} = \frac{(s-t)^2}{(s+t)(s^{1/2} + t^{1/2})^2} \le \frac{(s-t)^2}{s^2}$$

and observing that for $f \in \Sigma$, we have $J_{j,n}(f) \ge M^{-1}$, we obtain an upper bound for (4.23),

$$4\sum_{i=1}^{mn}\bigl(1 - \bigl(1 - M^2\bigl(J_{j(1,i),n}(f) - J_{j(2,i),m}(f)\bigr)^2\bigr)^{1/2m}\bigr). \tag{4.24}$$

The expression $J_{j(1,i),n}(f) - J_{j(2,i),m}(f)$ can be described as follows. For any $x \in (\frac{i-1}{mn}, \frac{i}{mn})$, $i = 1, \dots, mn$, we have

$$J_{j(1,i),n}(f) - J_{j(2,i),m}(f) = \bar{f}_n(x) - \bar{f}_m(x), \tag{4.25}$$

where $\bar{f}_n$ is defined by (4.17). Now as a consequence of Lemmas 5.4 and 5.5 in [17],

$$\sup_{f\in\Sigma} \|\bar{f}_n - \bar{f}_m\|_\infty \le \sup_{f\in\Sigma} \|f - \bar{f}_n\|_\infty + \sup_{f\in\Sigma} \|f - \bar{f}_m\|_\infty = o(1). \tag{4.26}$$

Note that for $m \to \infty$ and $z \to 0$ we have

$$\begin{aligned} (1 - Cz^2)^{1/2m} &= \exp\biggl(\frac{1}{2m}\log(1 - Cz^2)\biggr) \\ &= \exp\biggl(-\frac{1}{2m}\bigl(Cz^2 + O(z^4)\bigr)\biggr) \\ &= 1 - \frac{1}{2m}(Cz^2 + O(z^4)) + o\biggl(\frac{z^2}{m}\biggr). \end{aligned}$$

Thus from (4.24) we obtain in view of (4.26),

$$H^2\left(\bigotimes_{i=1}^{mn} Q_{1,i}, \bigotimes_{i=1}^{mn} Q_{2,i}\right) \le C \sum_{i=1}^{mn} \frac{1}{m}\big(J_{j(1,i),n}(f) - J_{j(2,i),m}(f)\big)^2\big(1+o(1)\big).$$

As a consequence of (4.25), we obtain

$$\|\bar{f}_n - \bar{f}_m\|_2^2 = \sum_{i=1}^{mn} \frac{1}{mn}\big(J_{j(1,i),n}(f) - J_{j(2,i),m}(f)\big)^2,$$

which implies

$$H^2\left(\bigotimes_{i=1}^{mn} Q_{1,i}, \bigotimes_{i=1}^{mn} Q_{2,i}\right) \le Cn\|\bar{f}_n - \bar{f}_m\|_2^2 \le Cn\|f - \bar{f}_m\|_2^2 + Cn\|f - \bar{f}_n\|_2^2.$$

Now as in (4.18), this upper bound is $o(1)$ uniformly over $f \in \Sigma$. $\square$

LEMMA 4.15. *We have the asymptotic equivalence*

$$\mathring{\mathcal{E}}^*_{m,n} \simeq \mathring{\mathcal{E}}_{m,n} \qquad \text{as } n \to \infty.$$

PROOF. We know [cf. (4.22), (4.21)] that $\mathring{\mathcal{E}}_{m,n} \sim \mathring{\mathcal{E}}_{m,1}$, where

$$\mathring{\mathcal{E}}_{m,1} = \left(\bigotimes_{j=1}^{m} \Gamma\big(1/2, 2J_{j,m}(f)\big), f \in \Sigma\right).$$

Analogously, using (4.20) again, we obtain

$$\mathring{\mathcal{E}}^*_{m,n} \sim \overline{\mathcal{E}}^*_{m,1} := \left(\bigotimes_{j=1}^{m} \Gamma\big(n/2m, 2J_{j,m}(f)\big), f \in \Sigma\right).$$

For given $f \in \Sigma$, the Hellinger distance between the two respective product measures is bounded by (using Lemma 2.19 in [31] and then Lemma 4.13)

$$\begin{aligned} &2\sum_{j=1}^{m} H^2\Big(\Gamma\big(1/2, 2J_{j,m}(f)\big), \Gamma\big(n/2m, 2J_{j,m}(f)\big)\Big) \\ &\quad = 4\sum_{j=1}^{m}\left(1 - \frac{\Gamma(1/4 + n/4m)}{(\Gamma(1/2)\Gamma(n/2m))^{1/2}}\right). \end{aligned}$$

Note that this bound does not depend on $f \in \Sigma$. Write $n/m = 1 + \delta$, where $\delta = r_n/m$; the above is

$$4\sum_{j=1}^{m} \frac{(\Gamma(1/2)\Gamma(1/2+\delta/2))^{1/2} - \Gamma(1/2+\delta/4)}{(\Gamma(1/2)\Gamma(1/2+\delta/2))^{1/2}}. \tag{4.27}$$

The Gamma function is infinitely differentiable on $(0, \infty)$; by a Taylor expansion we obtain

$$\Gamma(1/2+\delta/4) = \Gamma(1/2) + \Gamma'(1/2)\frac{\delta}{4} + O(\delta^2),$$
$$\Gamma^{1/2}(1/2+\delta/2) = \Gamma^{1/2}(1/2) + \frac{1}{2}\Gamma^{-1/2}(1/2)\Gamma'(1/2)\frac{\delta}{2} + O(\delta^2).$$

Consequently,

$$\big(\Gamma(1/2)\Gamma(1/2+\delta/2)\big)^{1/2} - \Gamma(1/2+\delta/4) = O(\delta^2),$$

so that (4.27) becomes

$$\sum_{j=1}^{m} \frac{O(\delta^2)}{\Gamma(1/2)(1+o(1))} = m\delta^2 O(1) \le \frac{r_n^2}{m} O(1).$$

The condition $r_n = o(n^{1/2})$ now implies that this upper bound is $o(1)$. We thus established total variation asymptotic equivalence $\mathring{\mathcal{E}}_{m,1} \simeq \overline{\mathcal{E}}^*_{m,1}$. $\square$

5. An application to hypothesis testing. In this section, we apply asymptotic equivalence theory to the problem of adaptive testing for the spectral density model. The minimax rate of testing for the spectral density model was studied in Ingster [19, 20], but the adaptive testing rate remains open. Spokoiny [30] obtained the adaptive minimax rate of testing for the Gaussian white noise model and conjectured that his method is applicable to the spectral density model. We will see that a parallel adaptive result for spectral density testing is just an immediate consequence of the asymptotic equivalence to Gaussian white noise.

For the white noise model,

$$dZ_t = f(t)\,dt + n^{-1/2}\,dW_t, \qquad t \in [0,1],$$

Spokoiny [30] considered the following testing problem:

$$H_0 : f = 0 \quad \text{vs.} \quad H_1 : f \in \mathcal{F}_{\beta,M}(\rho) := \{f : \|f\|_{B^\beta_{p,q}} \le M, \|f\|_2 \ge \rho\}$$

and obtained the adaptive minimax testing rate when p, q, β, M are unknown. Here $\|f\|_{B^\beta_{p,q}}$ is a Besov norm of f (cf. Section 5.3 in [17]); the result represents an adaptive version of the minimax rate of testing given in Ingster [20]. We will assume here $p = q = 2$, and thus limit ourselves to the implied adaptation result on the Sobolev scale where the smoothness parameter $\sigma := (\beta, M)$ is assumed unknown. Assume that σ is known to vary in a set $\mathcal{T}$, where

$$\mathcal{T} = \{(\beta, M) : \beta_l \le \beta \le \beta_u, M_l \le M \le M_u\} \tag{5.1}$$

for some prescribed $0 < \beta_l < \beta_u$, and $0 < M_l < M_u$. Let ϕ_n denote a (possibly randomized) test; we use notation $E_{n,f}$ for the expectation operator under a given

f in the white noise model. Denote the supremal error of second kind of the test over the alternatives of given smoothness σ and distance ρ by

$$\pi_\sigma(\phi_n, \rho) = \sup_{f \in \mathcal{F}_{\beta,M}(\rho)} E_{n,f}(1-\phi_n).$$

PROPOSITION 5.1. *Assume that* $\beta_l > 1/2$, *and let*

$$\rho_\sigma(n) = n^{-4\beta/(4\beta+1)}, \qquad t_n = (\log\log n)^{1/4}.$$

(i) *For any sequence* $t_n' = o_n(t_n)$ *one has*

$$\inf_{\phi_n}\Big[E_{n,0}(\phi_n) + \sup_{\sigma\in\mathcal{T}} \pi_\sigma(\phi_n, \rho_\sigma(n)t_n')\Big] \geq 1 - o_n(1).$$

(ii) *There exists a constant* $c_1(\beta_l, \beta_u, M_l, M_u) > 0$ *and a test* ϕ_n^* *such that*

$$E_{n,0}(\phi_n^*) + \sup_{\sigma\in\mathcal{T}} \pi_\sigma(\phi_n^*, c_1\rho_\sigma(n)t_n) = o_n(1).$$

Consider now stationary Gaussian observations with spectral density f (assumed to be a function on the unit interval for the present purpose). We wish to test the null hypothesis $H_0: f = 1$ against $H_1: f \neq 1$, or equivalently $H_0: \log f = 0$ against $H_1: \log f \neq 0$. Let $\mathcal{F}_M$ and $W^\beta(M)$ be function classes as defined in (1.7) and (1.8), and let for some $\beta > 1/2$ and $M > 0$, $\rho > 0$

$$\mathcal{G}_{\beta,M}(\rho) := \{f \in W^\beta(M) \cap \mathcal{F}_M : \|f-1\|_2 \geq \rho\}.$$

Note that $W^\beta(M)$ is essentially a periodic Besov–Sobolev smoothness class $\{f : \|f\|_{\tilde{B}^\beta_{2,2}} \leq M'\}$ (cf. Remark 5.8 in [17]) up to equivalence of the norms used. We will consider adaptation to a class of spectral densities $f \in \mathcal{G}_{\beta,M}(\rho)$, where the smoothness parameter $\sigma = (\beta, M)$ is unknown. Again we assume that σ varies in a set $\mathcal{T}$ defined in (5.1). In order to apply Theorem 1.2, we have to verify that the union of all smoothness classes $\mathcal{G}_{\beta,M}(\rho)$ considered is in the parameter set Σ over which the equivalence holds. Thus for some $\alpha > 1/2$ and $M' > 0$, we have to verify

$$\bigcup_{\sigma\in\mathcal{T}} \mathcal{G}_{\beta,M}(\rho) \subset \mathcal{F}_{M'} \cap W^\alpha(M') \cap \{f : \|f\|_{B^\alpha_{6,6}} \leq M'\} \tag{5.2}$$

(here ρ may be taken 0). Using the standard embedding theorems summarized in Proposition 5.2 of [17] (which hold analogously on the periodic scale $\tilde{B}^\alpha_{p,q}$), it is easy to see that the condition $\beta_l > 5/6$ guarantees (5.2) for sufficiently large M'. Thus Theorem 1.2 on asymptotic equivalence is applicable. We refer to Brown et al. [6] for a basic discussion of how risk bounds transfer from one model to another, under asymptotic equivalence.

PROPOSITION 5.2. *Assume that* $\beta_l > 5/6$. *Then Proposition* 5.1 *holds in the spectral density model, with pertaining set* $\mathcal{T}$, *and with smoothness classes* $\mathcal{F}_{\beta,M}(\rho)$ *replaced by* $\mathcal{G}_{\beta,M}(\rho)$.

PROOF. Since $f \in \mathcal{F}_M$, there are $c_2 > 0$ and $M' > 0$ such that for any $\rho > 0$,

$$\{f : \|f - 1\|_2^2 \geq c_2\rho\} \subset \{f : \|\log f\|_2^2 \geq c_1\rho\}$$

and

$$\mathcal{G}_{\beta,M}(\rho)(c_2\rho) \subset \{f : \log f \in \mathcal{F}_{\beta,M'}(\rho)(c_1\rho)\} \tag{5.3}$$

(cf. Runst [27] and Sickel [29]). Part (ii) of Proposition 5.1 implies that there exists a test ϕ_n^{**} in the the spectral density model such that

$$E_{n,0}(\phi_n^{**}) = o_\varepsilon(1),$$

$$\sup_{\sigma\in\mathcal{T}} \sup_{f\in\mathcal{G}_{\beta,M}(c_2\rho_\sigma(n))} E_{n,f}(1 - \phi_n^{**}) = o_\varepsilon(1).$$

For part (i) of Proposition 5.1, Spokoiny ([30], pages 2493–2495) constructs for a finite subset $\mathcal{T}_n$ of $\mathcal{T}$ and all $(\beta, M) \in \mathcal{T}_n$, prior measures on the set $\mathcal{A}_{\beta,M} := \mathcal{F}_{\beta,M}(c_1\rho_\sigma(n)t_n')$. An inessential modification of the prior measures near the boundaries of the interval [0, 1] shows that $\mathcal{A}_{\beta,M}$ can be replaced by its subset of functions with compact support in (0, 1), say $\mathcal{A}^0_{\beta,M}$. Then $\mathcal{A}^0_{\beta,M}$ is a subset of the periodic Besov space $\tilde{B}^\beta_{2,2}$, and similarly to (5.3), it can be shown that for every $M > 0$ there exists constant $c_3 > 0$ (depending on the bounds $\beta_l, \beta_u, M_l, M_u$ in the spectral density model) and $M'' > 0$ such that $\log(\mathcal{G}_{\beta,M}(c_3\rho_\sigma(n)t_n')) \supset \mathcal{A}^0_{\beta,M''}$. Thus part (i) of Proposition 5.1 implies for the spectral density model,

$$\inf_{\phi_n}\Big[E_{n,0}(\phi_n) + \sup_{\sigma\in\mathcal{T}} \sup_{f\in\mathcal{G}_{\beta,M}(c_3\rho_\sigma(n)t_n')} E_{n,f}(1 - \phi_n)\Big] \geq 1 - o_n(1). \qquad \square$$

REFERENCES

[1] BICKEL, P. and DOKSUM, K. (2001). *Mathematical Statistics* **1**, 2nd ed. Prentice Hall, Upper Saddle River, NJ.

[2] BALAKRISHNAN, A. V. (1976). *Applied Functional Analysis.* Springer, New York. MR0470699

[3] BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods*, 2nd ed. Springer, New York. MR1093459

[4] BROWN, L. D. and LOW, M. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384–2398. MR1425958

[5] BROWN, L. D., LOW, M. and ZHANG, C.-H. (2002). Asymptotic equivalence theory for nonparametric regression with random design. *Ann. Statist.* **30** 688–707. MR1922538

[6] BROWN, L. D., CARTER, A. V., LOW, M. G. and ZHANG, C.-H. (2004). Equivalence theory for density estimation, Poisson processes and Gaussian white noise with drift. *Ann. Statist.* **32** 2074–2097. MR2102503

[7] CARTER, A. V. (2002). Deficiency distance between multinomial and multivariate normal experiments. *Ann. Statist.* **30** 708–730. MR1922539

[8] CHOUDHOURI, N., GHOSAL, S. and ROY, A. (2004). Contiguity of the Whittle measure for a Gaussian time series. *Biometrika* **91** 211–218. MR2050470

[9] COURSOL, J. and DACUNHA-CASTELLE, D. (1982). Remarks on the approximation of the likelihood function of a stationary Gaussian process. *Theory Probab. Appl.* **27** 162–167. MR0645138

[10] DAHLHAUS, R. (1988). Small sample effects in time series analysis: A new asymptotic theory and a new estimate. *Ann. Statist.* **16** 808–841. MR0947580

[11] DAHLHAUS, R. and JANAS, D. (1996). A frequency domain bootstrap for ratio statistics in time series analysis. *Ann. Statist.* **24** 1934–1963. MR1421155

[12] DAHLHAUS, R. and POLONIK, W. (2002). Empirical spectral processes and nonparametric maximum likelihood estimation for time series. In: *Empirical Process Techniques for Dependent Data* (H. Dehling, T. Mikosch and M. Sørensen, eds.) 275–298. Birkhäuser, Boston. MR1958786

[13] DAVIES, R. B. (1973). Asymptotic inference in stationary Gaussian time-series. *Adv. in Appl. Probab.* **5** 469–497. MR0341699

[14] DELATTRE, S. and HOFFMANN, M. (2002). Asymptotic equivalence for a null recurrent diffusion. *Bernoulli* **8** 139–174. MR1895888

[15] DZHAPARIDZE, K. (1986). *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*. Springer, New York. MR0812272

[16] FAN, J. and GIJBELS, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London. MR1383587

[17] GOLUBEV, G. K., NUSSBAUM, M. and ZHOU, H. H. (2005). Asymptotic equivalence of spectral density estimation and Gaussian white noise. Technical report. Available at arXiv:0903.1314v1 [math.ST].

[18] GRAMA, I. and NUSSBAUM, M. (1998). Asymptotic equivalence for nonparametric generalized linear models. *Probab. Theory Related Fields* **111** 167–214. MR1633574

[19] INGSTER, Y. I. (1985). An asymptotic minimax test of nonparametric hypotheses about spectral density. *Theory Probab. Appl.* **29** 846–847.

[20] INGSTER, Y. I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives I–III. *Math. Methods Statist.* **2** 85–114; **3** 171–189; **4** 249–268.

[21] LAHIRI, S. N. (2003). A necessary and sufficient condition for asymptotic independence of discrete Fourier transforms under short- and long-range dependence. *Ann. Statist.* **31** 613–641. MR1983544

[22] LE CAM, L. (1974). On the information contained in additional observations. *Ann. Statist.* **2** 630–649. MR0436400

[23] LE CAM, L. and YANG, G. (2000). *Asymptotics in Statistics*, 2nd ed. Springer, New York. MR1784901

[24] LOW, M. G. and ZHOU, H. H. (2007). A complement to Le Cam's theorem. *Ann. Statist.* **35** 1146–1165. MR2341701

[25] MAMMEN, E. (1986). The statistical information contained in additional observations. *Ann. Statist.* **14** 665–678. MR0840521

[26] NUSSBAUM, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.* **24** 2399–2430. MR1425959

[27] RUNST, T. (1986). Mapping properties of nonlinear operators in spaces of Triebel–Lizorkin and Besov type. *Anal. Math.* **12** 313–346. MR0877164

[28] SHIRYAEV, A. N. (1996). *Probability*, 2nd ed. Springer, New York. MR1368405

[29] SICKEL, W. (1996). Composition operators acting on Sobolev spaces of fractional order—a survey on sufficient and necessary conditions. In *Function Spaces, Differential Operators and Nonlinear Analysis* 159–182. Prometheus, Prague. MR1480936

[30] SPOKOINY, V. G. (1996). Adaptive hypothesis testing using wavelets. *Ann. Statist.* **24** 2477–2498. MR1425962

[31] STRASSER, H. (1985). *Mathematical Theory of Statistics.* Walter de Gruyter, Berlin. MR0812467

[32] VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. Cambridge Univ. Press, Cambridge. MR1652247

[33] ZHOU, H. H. (2004). Minimax estimation with thresholding and asymptotic equivalence for Gaussian variance regression. Ph.D. thesis, Cornell Univ. Press, Ithaca, NY. Available at http://www.stat.yale.edu/~hz68/.

G. K. GOLUBEV
UNIVERSITÉ DE PROVENCE CMI
39, RUE F. JOLIOT-CURIE
13453 MARSEILLE, CEDEX 13
FRANCE
E-MAIL: golubev@gyptis.univ-mrs.fr

M. NUSSBAUM
DEPARTMENT OF MATHEMATICS
MALOTT HALL
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853
USA
E-MAIL: nussbaum@math.cornell.edu

H. H. ZHOU
DEPARTMENT OF STATISTICS
YALE UNIVERSITY
P.O. BOX 208290
NEW HAVEN, CONNECTICUT 06520
USA
E-MAIL: huibin.zhou@yale.edu