INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2010-38

Bayesian Nonparametric Estimation of the Spectral Density of a Long or Intermediate Memory Gaussian Process

N. CHOPIN[1] - B. LISEO[2]
J. ROUSSEAU[3]

[1] CREST – ENSAE, Paris.

[2] Sapienza, Università di Roma, Italy.

[3] Université Paris-Dauphine et CREST, Laboratoire de Statistique. *Corresponding author :* *rousseau@ceremade.dauphine.fr*

Bayesian nonparametric estimation of the spectral density of a long or intermediate memory Gaussian process

J. Rousseau[1,2*], N. Chopin[2] and B. Liseo[3]

1: Université Paris Dauphine, 2: CREST-ENSAE, Paris, 3: Sapienza Universit di Roma

July 23, 2010

Abstract

A stationary Gaussian process is said to be long-range dependent (resp. anti-persistent) if its spectral density $f(\lambda)$ can be written as $f(\lambda) = |\lambda|^{-2d}g(|\lambda|)$, where $0 < d < 1/2$ (resp. $-1/2 < d < 0$), and g is continuous. We propose a novel Bayesian nonparametric approach for the estimation of the spectral density of such processes. Within this approach, we prove posterior consistency for both d and g, under appropriate conditions on the prior distribution. We establish the rate of convergence for a general class of priors, and apply our results to the family of fractionally exponential priors. Our approach is based on the true likelihood function, and does not resort to Whittle's approximation, which is not valid in a long memory set-up.

Key-words: Bayesian nonparametric; consistency; FEXP priors; Gaussian long memory processes; rates of convergence

1 Introduction

Let $\mathbf{X} = \{X_t,\ t = 1, 2, \dots\}$ be a real-valued stationary zero-mean Gaussian random process, with spectral density f, and covariance function $\gamma_f(\tau) = E(X_t X_{t+\tau})$, so that

$$\gamma_f(\tau) = \int_{-\pi}^{\pi} f(\lambda)e^{i\tau\lambda}d\lambda \qquad (\tau = 0, \pm1, \pm2, \dots). \tag{1.1}$$

This process is long-range dependent (resp. anti-persistent) if there exist $C > 0$ and a value d, $0 < d < 1/2$ (resp. $-1/2 < d < 0$), such that $f(\lambda)|\lambda|^{2d} \to C$ when $\lambda \to 0$. This may be conveniently rewritten as $f(\lambda) = \lambda^{-2d}g(|\lambda|)$, where $g : [0, \pi] \to \mathbb{R}^+$ is a continuous function.

Interest in long-range dependent and anti-persistent time series has increased steadily in the last fifteen years; see Beran (1994) for a comprehensive introduction and Doukhan et al. (2003) for a review of theoretical aspects and fields of applications, including telecommunications, economics, finance, astrophysics, medicine and hydrology. Research in parametric inference for long and intermediate memory processes have been pioneered by Mandelbrot and Van Ness (1968), Mandelbrot and Wallis (1969), and continued by Fox and Taqqu (1986), Dahlhaus (1989), Giraitis and Taqqu (1999), Geweke and Porter-Hudak (1983), and Beran (1993), among others. Unfortunately, parametric inference can be highly biased under mis-specification of the true model. This limitation makes semiparametric approaches particularly appealing (Robinson, 1995).

*Corresponding author: rousseau@ceremade.dauphine.fr

For instance, under the representation $f(\lambda) = |\lambda|^{-2d} g(|\lambda|)$, one would like to estimate d as a measure of long-range dependence, without resorting to parametric assumptions on the nuisance parameter g; see Liseo et al. (2001) for a Bayesian approach to this problem, and Bardet et al. (2003) for an exhaustive review of classical approaches. However, practically all the existing procedures either exploit the regression structure of the log-spectral density in a small neighborhood of the origin (Robinson, 1995), or use an approximate likelihood function based on Whittle's approximation (Whittle, 1962), where the original vector of observations $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ gets transformed into the periodogram $I(\lambda)$ computed at the Fourier frequencies $\lambda_j = 2\pi j/n$, $j = 1, 2, \ldots, n$, and the artificial observations $I(\lambda_1), \ldots, I(\lambda_n)$ are, under short range dependence, approximately independent. Unfortunately, Whittle's approximation is not valid in in the presence of long range dependence, at least for the smallest Fourier frequencies.

We propose a Bayesian nonparametric approach to the estimation of the spectral density of the stationary Gaussian process based on the true likelihood, without resorting to Whittle's approximation. We study the asymptotic properties of our procedure, including consistency and rates of convergence. Our study is based on standard tools for an asymptotic analysis of Bayesian approaches, e.g. Ghosal et al. (2000), i.e. quantities of interest are the prior probability of a small neighborhood around the true spectral density, and some kind of entropy measure for the prior distribution. Most technical details differ however, as the observed process is long-range dependent.

The paper is organised as follows. In Section 2, we introduce the model and the notations. In Section 3, we provide a general theorem that states sufficient conditions to ensure consistency of the posterior distribution, and of several Bayes estimators. We also introduce the class of FEXP (Fractional Exponential) priors, based on the FEXP representation of Robinson (1991), and show that such prior distributions fulfil these sufficient conditions for posterior consistency. In Section 4, we study the rate of convergence of the posterior in the general case, and specialise our results for the FEXP class. Section 5 gives the proofs of the main theorems of the two previous Sections. Section 6 discusses further research. The Appendix contains several technical lemmas.

2 Model and notations

The model consists of an observed vector $\mathbf{X}_n = (X_1, \ldots, X_n)$ of n consecutive realizations from a zero-mean Gaussian stationary process, with spectral density f, which is either long-range dependent, short-range dependent, or anti-persistent. The likelihood function is

$$\varphi(\mathbf{X}_n; f) = (2\pi)^{-n/2} |T_n(f)|^{-1/2} \exp\{-\frac{1}{2} \mathbf{X}_n^t T_n(f)^{-1} \mathbf{X}_n\} \tag{2.1}$$

where $T_n(f)$ is the Toeplitz matrix associated to γ_f, see (1.1), i.e. $T_n(f) = [\gamma_f(j-k)]_{1 \le j,k \le n}$.

This model is parametrised by the pair (d, g), which defines $f = F(d, g)$ through the factorisation

$$\begin{array}{rcl} F : (-1/2, 1/2) \times \mathcal{C}^0[0, \pi] & \rightarrow & \mathcal{F} \\ (d, g) & \rightarrow & f : f(\lambda) = |\lambda|^{-2d} g(|\lambda|), \end{array}$$

where $\mathcal{C}^0[0, \pi]$ is the set of continuous functions over $[0, \pi]$, and $\mathcal{F}$ denotes the set of spectral densities, that is, the set of even functions $f : [-\pi, \pi] \rightarrow \mathbb{R}^+$ such that $\int_{-\pi}^{\pi} |f(\lambda)| \, d\lambda < +\infty$.

The model is completed with a nonparametric prior distribution π for $(d, g) \in (-1/2, 1/2) \times \mathcal{C}^0[0, \pi]$. (There should be no confusion whether π refers to either the constant or the prior distribution in the rest of the paper.) All our results will assume that the model is valid for some true' parameter (d_0, g_0), associated to some 'true' spectral density $f_0 = F(d_0, g_0)$, where $d_0 \in (-1/2, 1/2)$; conditions on g_0 are detailed in the next section.

We introduce several pseudo-distances on $\mathcal{F}$. The Kullback-Leibler divergence for finite n is defined as

$$\begin{aligned} KL_n(f_0; f) &= \frac{1}{n}\int_{\mathbb{R}^n} \varphi(\mathbf{X}_n; f_0)\left\{\log\varphi(\mathbf{X}_n; f_0) - \log\varphi(\mathbf{X}_n; f)\right\} d\mathbf{X}_n \\ &= \frac{1}{2n}\left\{\mathrm{tr}\left[T_n(f_0)T_n^{-1}(f) - \mathbf{I}_n\right] - \log\det(T_n(f_0)T_n^{-1}(f))\right\} \end{aligned}$$

where $\mathbf{I}_n$ represents the identity matrix of order n. Letting $n \to \infty$, we can define, when it exists, the quantity

$$KL_\infty(f_0; f) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left[\frac{f_0(\lambda)}{f(\lambda)} - 1 - \log\frac{f_0(\lambda)}{f(\lambda)}\right] d\lambda.$$

We also define a symmetrised version of KL_n, i.e.

$$h_n(f_0, f) = KL_n(f_0; f) + KL_n(f; f_0);$$

and its limit as $n \to \infty$:

$$h(f_0, f) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left[\frac{f_0(\lambda)}{f(\lambda)} + \frac{f(\lambda)}{f_0(\lambda)} - 2\right] d\lambda = \frac{1}{2\pi}\int_0^{\pi}\left(\frac{f_0(\lambda)}{f(\lambda)} - 1\right)^2 \frac{f(\lambda)}{f_0(\lambda)}\, d\lambda.$$

For technical reasons, we define also the pseudo-distance

$$b_n(f_0, f) = \frac{1}{n}\mathrm{tr}\left[\left(T_n(f)^{-1}T_n(f_0 - f)\right)^2\right]$$

and its limit as $n \to +\infty$,

$$b(f_0, f) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{f_0(\lambda)}{f(\lambda)} - 1\right)^2 d\lambda.$$

Finally, we consider the L^2 distance between the spectral log-densities (Moulines and Soulier, 2003),

$$\ell(f_0, f) = \int_{-\pi}^{\pi}\left\{\log f_0(\lambda) - \log f(\lambda)\right\}^2 d\lambda. \tag{2.2}$$

For the models considered in this paper, this distance always exists, whereas the L^2 distance may not.

3 Consistency

We first state and prove the strong consistency of the posterior distribution under very general conditions on both π and $f_0 = F(d_0, g_0)$, i.e. as $n \to \infty$, and for $\varepsilon > 0$ small enough,

$$P^\pi[\mathcal{A}_\varepsilon | \mathbf{X}_n] \to 1, \quad \text{a.s.},$$

where $P^\pi[.|\mathbf{X}_n]$ denotes posterior probabilities associated with the prior π, and

$$\mathcal{A}_\varepsilon = \{(d, g) \in (-1/2, 1/2) \times \mathcal{C}^0[0, \pi] : h(f_0, F(d, g)) \le \varepsilon\}.$$

From this, we shall deduce the consistency of Bayes estimators of f and d. Finally, we shall introduce the class of FEXP priors, and show that they allow for posterior consistency.

3.1 Main result

Consider the following sets:

$$\begin{aligned}
\mathcal{G}(m,M) &= \left\{g \in \mathcal{C}^0[0,\pi] : m \le g \le M\right\} \\
\mathcal{G}(m,M,L,\rho) &= \{g \in \mathcal{G}(m,M) : |g(\lambda) - g(\lambda')| \le L|\lambda - \lambda'|^\rho\} \\
\mathcal{G}(t,m,M,L,\rho) &= [-1/2+t, 1/2-t] \times \mathcal{G}(m,M,L,\rho)
\end{aligned}$$

for $\rho \in (0,1]$, $L > 0$, $m \le M$, $t \in (0,1/2)$. Restricting the parameter space to such sets makes the model identifiable (boundedness of g, provided $m > 0$), and ensures that normalized traces of products of Toeplitz matrices that appear in the distances defined in the previous section converge (Hölder inequality).

We now state our main consistency result.

Theorem 3.1. *For $\varepsilon > 0$ small enough*

$$P^\pi\left[\mathcal{A}_\varepsilon | \mathbf{X}_n\right] \to 1, \quad a.s.$$

provided the following conditions are fulfilled:

1. *There exist $t, m, M, L > 0$, $\rho \in (0,1]$, such that the set $\mathcal{G}(t,m,M,L,\rho)$ contains both the pair (d_0, g_0) that defines the true spectral density $f_0 = F(d_0, g_0)$ and the support of the prior distribution π.*
2. *For all $\varepsilon > 0$, $\pi(\mathcal{B}_\varepsilon) > 0$, where $\mathcal{B}_\varepsilon$ is defined by*
$$\mathcal{B}_\varepsilon = \{(d,g) \in \mathcal{G}(t,m,M,L,\rho) : h(f_0, F(d,g)) \le \varepsilon, 16|d_0 - d| < \rho + 1 - t\}.$$
3. *For $\varepsilon > 0$ small enough, there exists a sequence $\mathcal{F}_n$ such that $\pi(\mathcal{F}_n) \ge 1 - e^{-nr}$, $r > 0$, and a net (i.e. a finite collection)*
$$\mathcal{H}_n \subset \{(d,g) \in [-1/2+t, 1/2-t] \times \mathcal{G}(m,M,L,\rho) : h(f_0; F(d,g)) > \varepsilon/2\}$$
such that, for n large enough, for all $(d,g) \in \mathcal{F}_n \cap A_\varepsilon^c$, $f = F(d,g)$, there exists $(d_i, g_i) \in \mathcal{H}_n$, $f_i = F(d_i, g_i)$, such that $8(d_i - d) \le \rho + 1 - t$, $f \le f_i$, and:
   - *(a) if $8|d_i - d_0| \le \rho + 1 - t$,*
$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{(f_i - f)(\lambda)}{f_0(\lambda)} d\lambda \le h(f_0, f_i)/4;$$
   - *(b) if $8(d_i - d_0) > \rho + 1 - t$,*
$$b(f_i, f) \le b(f_0, f_i)|\log\varepsilon|^{-1};$$
   - *(c) otherwise, if $8(d_0 - d_i) > \rho + 1 - t$,*
$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{(f_i - f)(\lambda)}{f_i(\lambda)} d\lambda \le b(f_i, f_0)|\log\varepsilon|^{-1}.$$
4. *The cardinality $\mathcal{C}_n$ of the net $\mathcal{H}_n$ defined above is such that $\log \mathcal{C}_n \le n\varepsilon/log(\varepsilon)$.*

A proof is given in Section 5.1. Note that, in the above definition of the net $\mathcal{H}_n$, the $|\log\varepsilon|$ terms are here only to avoid writing inequalities in terms of awkward constants in the form m/M. If need be, we can replace the $|\log\varepsilon|$ by the correct constants as expressed in Appendix B. The definition of the above *entropy* is non-standard. The interest in expressing it in this general but non-standard form lies in the difficulty in dealing with spectral densities which diverge at 0. In practise, the way one constructs the net $\mathcal{H}_n$ should vary according to the form of the prior on the short memory part g.

3.2 Consistency of point estimates

As explained in §2, we focus on the quadratic loss function ℓ with respect to the logarithm of the spectral density. The corresponding Bayes estimator is

$$\hat{d} = E^{\pi}[d|\mathbf{X}_n], \quad \hat{g} : \lambda \to \exp\left\{E^{\pi}[\log g(\lambda)|\mathbf{X}_n]\right\}, \quad \hat{f} = F(\hat{d}, \hat{g}).$$

Often, the real parameter of interest is d, and g is a nuisance parameter. Consistency for $\hat{d}$ can be deduced from Theorem 3.1.

Corollary 1. *Under the assumptions of Theorem 3.1, for $\varepsilon > 0$ small enough,*

$$P^{\pi}\left[\{|d - d_0| > \varepsilon\}|\mathbf{X}_n\right] \to 0$$

and $\hat{d} \to d_0$ as $n \to \infty$.

Proof. Lemma 10, see Appendix D, implies that

$$P^{\pi}[\mathcal{A}^c_{\varepsilon}|\mathbf{X}_n] \geq P^{\pi}\left[\{|d - d_0| > \varepsilon'\}|\mathbf{X}_n\right] \to 0 \quad \text{a.s.}$$

as $n \to +\infty$, for some $\varepsilon' > 0$ and, by Jensen's inequality,

$$(\hat{d} - d_0)^2 \leq E^{\pi}[(d - d_0)^2|\mathbf{X}_n] \to 0, \quad \text{a.s.}$$

□

Consistency results for a point estimate of f can also be deduced:

Corollary 2. *Under the assumptions of Theorem 3.1, one has, as $n \to \infty$,*

$$\ell(f_0, \hat{f}) \to 0, \quad a.s.$$

Proof. For $f = F(d, g)$, $f_0 = F(d_0, g_0)$, one has $\ell(f_0, f) \leq h(f_0, f)$, since $x^2 \leq e^x + e^{-x} - 2$ for all x, and $\ell(f_0, f) \leq C$ for some well chosen constant C, provided g, $g_0 \in \mathcal{G}(m, M)$. Thus, by Jensen inequality, and for all $\varepsilon > 0$,

$$\ell(f_0, \hat{f}) \leq E^{\pi}[\ell(f_0, f)|\mathbf{X}_n] \leq \varepsilon + CP^{\pi}[A^c_{\varepsilon}|\mathbf{X}_n].$$

□

3.3 The FEXP prior

Following Hurvich et al. (2002), we consider the FEXP parameterisation of spectral densities, i.e. $f = \widetilde{F}(d, k, \theta)$, where

$$\begin{aligned} \widetilde{F} : \mathcal{T} &\to \mathcal{F} \\ (d, k, \theta) &\to f : f(\lambda) = |1 - e^{i\lambda}|^{-2d} \exp\left\{\sum_{j=0}^{k} \theta_j \cos(j\lambda)\right\}. \end{aligned} \tag{3.1}$$

and $\mathcal{T} = (-1/2 + t, 1/2 - t) \times \left\{\cup_{k=0}^{+\infty}\{k\} \times \mathbb{R}^{k+1}\right\}$, for some fixed $t \in (0, 1/2)$. This FEXP representation is equivalent to our previous representation $f = F(d, g)$, provided $g = \psi^{-d}e^{w}$, $w(\lambda) = \left\{\sum_{j=0}^{k} \theta_j \cos(j\lambda)\right\}$ and $\psi(\lambda) = |1 - e^{i\lambda}|^2/\lambda^2 = 2(1 - \cos\lambda)/\lambda^2$ for $\lambda \neq 0$, $\psi(0) = 1$.

The function ψ is bounded, infinitely differentiable and positive for $\lambda \in [0, \pi]$. Thus g and w share the same regularity properties, i.e. w is bounded and Hölder with exponent ρ implies that g is bounded and Hölder with exponent ρ, and vice versa. Under this parameterisation, the prior distribution π is expressed as a trans-dimensional prior distribution on the random vector (d, k, θ), which, for convenience, factorises as $\pi_d(d)\pi_k(k)\pi_\theta(\theta|k)$.

We assume that π puts mass one on the following Sobolev set:

$$\mathcal{S}(\beta, L) = \left\{ (d, k, \theta) \in \mathcal{T} : \sum_{j=0}^{k} \theta_j^2 (j+1)^{2\beta} \leq L \right\} \tag{3.2}$$

for some $\beta > 1/2$, $L > 0$. This ensures that the Fourier sum w, and thus the short-memory component g of the spectral density f, as explained above, belong to some set $\mathcal{G}(m, M, L', \rho)$, i.e., both w and g are bounded and Hölder, for $\rho < \beta - 1/2$. To see this, note that, for $(d, k, \theta) \in \mathcal{S}(\beta, L)$:

$$\begin{aligned} \sum_{j=0}^{k} |\theta_j| \, j^r &\leq \sum_{j=0}^{k} \theta_j^2 (j+1)^{2\beta} + \sum_{j=0}^{k} |\theta_j| \, j^r \mathbb{1}\left(|\theta_j| j^r \geq \theta_j^2 (j+1)^{2\beta} \right) \\ &\leq L + \sum_{j=0}^{+\infty} (j+1)^{2r-2\beta} < +\infty, \end{aligned} \tag{3.3}$$

provided $2r - 2\beta < -1$. By taking $r = 0$, one sees that w is bounded, and by taking $r = \rho$, for any ρ, $0 < \rho < \beta - 1/2$, one sees that w is Hölder, with coefficient ρ, since, for λ, $\lambda' \in [-\pi, \pi]$,

$$\begin{aligned} |w(\lambda) - w(\lambda')| &\leq 2 \sum_{j=0}^{k} |\theta_j| \times |\{\cos(\lambda j) - \cos(\lambda' j)\}/2|^\rho \\ &\leq 2^{1-\rho} \left(\sum_{j=0}^{k} |\theta_j| j^\rho \right) |\lambda - \lambda'|^\rho . \end{aligned}$$

Finally, we assume that π assigns positive prior probability to the intersection of $\mathcal{S}(\beta, L)$ with any rectangle set of the form

$$(a_d, b_d) \times \{k\} \times \prod_{j=1}^{k} (a_{\theta_j}, b_{\theta_j}).$$

Alternatively, one could assume that the support of π is included in a set of the form $\{(d, k, \theta) \in \mathcal{T} : \sum_{j=0}^{k} |\theta_j| j^\rho \leq L\}$. However Sobolev sets are more natural when dealing with rates of convergence, see Section 4.2, and are often considered in the non parametric literature, so we restrict our attention to these sets.

In the same spirit, we assume that the true spectral density admits a FEXP representation associated to an infinite Fourier series,

$$f_0(\lambda) = |1 - e^{i\lambda}|^{-2d_0} \exp\left\{ \sum_{j=0}^{+\infty} \theta_{0j} \cos(j\lambda) \right\},$$

i.e., $f_0 = F(d_0, g_0)$ with $g_0 = \psi^{-d_0} e^{w_0}$ and $w_0(\lambda) = \left\{ \sum_{j=0}^{+\infty} \theta_{0j} \cos(j\lambda) \right\}$. In addition, we assume that w_0 satisfies the same type of Sobolev inequality, namely

$$L_0 = \sum_{j=0}^{+\infty} \theta_{0j}^2 (j+1)^{2\beta} < L < +\infty, \tag{3.4}$$

which, as explained above, implies that $g_0 \in \mathcal{G}(m, M, L, \rho)$, for some well chosen constants m, M, L, ρ. Note that it is essential to have a strict inequality in (3.4), i.e. $L_0 < L$.

Theorem 3.2. *Let π be a prior distribution $\pi_d(d)\pi_k(k)\pi_\theta(\theta|k)$ which fulfils the above conditions, and, in addition, such that $\pi_k(k) \leq \exp(-Ck \log k)$ for some $C > 0$ and k large enough. Then the conditions of Theorem 3.1 are fulfilled, and the posterior distribution is consistent.*

Proof. Condition 1 of Theorem 3.1 is a simple consequence of (3.4) and (3.2), as explained above. For Condition 2, we noted, see (3.3), that $\sum_{j=0}^{+\infty} \theta_{0j}^2 (j+1)^{2\beta} \leq L$ implies that $\sum_{j=0}^{+\infty} |\theta_{0j}| \leq L' < +\infty$. Let k such that $\sum_{j=k+1}^{\infty} |\theta_{0j}| \leq \varepsilon/14$, $\theta = (\theta_0, ..., \theta_k)$ such that $\sum_{j=0}^{k} |\theta_{0j} - \theta_j| \leq \varepsilon/14$, d such that $|d - d_0| \leq \varepsilon/7$, and let $f = \tilde{F}(d, k, \theta)$. Using Lemma 14, see Appendix D, one has $h(f, f_0) \leq \varepsilon$. Note that it is sufficient to prove that $\pi(\mathcal{B}_\varepsilon) > 0$ for ε small enough, hence we assume that $\varepsilon/7 < (\rho + 1 - t)/16$. Thus, Condition 2 is verified as soon as the intersection of $\mathcal{S}(\beta, L)$ and the rectangle set

$$[d_0 - \varepsilon/7, d_0 + \varepsilon/7] \times \{k\} \times \prod_{j=1}^{k} [\theta_{0j} - \varepsilon/14k, \theta_{0j} - \varepsilon/14k]$$

is assigned positive prior probability. Now consider Condition 3. Let $\varepsilon > 0$ and take

$$\mathcal{F}_n = \{(d, k, \theta) \in \mathcal{S}(\beta, L) :\, k \leq k_n\},$$

where $k_n = \lfloor \alpha n / \log n \rfloor$, for some $\alpha > 0$, so that, for some r depending on α, $\pi(\mathcal{F}_n^c) \leq \pi_k(k > k_n) \leq e^{-nr}$. Let $f = \tilde{F}(d, k, \theta)$, $f_i = (2e)^{c\varepsilon} \tilde{F}(d_i, k, \theta_i)$, such that $k \leq k_n$, $d_i - c\varepsilon \leq d \leq d_i$, and $\sum_{j=0}^{k} |\theta_j - \theta_{ij}| \leq c\varepsilon$, for some $c > 0$, then

$$\frac{f(\lambda)}{f_i(\lambda)} = (2e)^{-c\varepsilon} \left[2(1 - \cos\lambda)\right]^{d_i - d} \exp\left\{\sum_{j=0}^{k} (\theta_j - \theta_{ij}) \cos(j\lambda)\right\} \leq 1,$$

and

$$\frac{f(\lambda)}{f_i(\lambda)} \quad \geq \quad (1 - \cos\lambda)^{c\varepsilon} 2^{-c\varepsilon} e^{-2c\varepsilon}.$$

If c is small enough, $f_i - f$ verifies the three inequalities considered in Condition 3. The number $\mathcal{C}_n$ of functions f_i necessary to ensure that, for any f in the support of π, at least one of them verify the above inequalities, can be bounded by, for n large enough, and some well chosen constant C,

$$\begin{aligned} \mathcal{C}_n \leq k_n (C k_n/\varepsilon)^{k_n + 2} &\leq k_n^{3k_n} \\ &\leq \exp\left\{3\alpha n \left[1 + (\log\alpha - \log\log n)/\log n\right]\right\} \\ &\leq \exp\{6\alpha n\} \end{aligned}$$

so Condition 4 is satisfied, provided one takes $\alpha = \varepsilon/6 \log\varepsilon$.

□

A convenient default choice for π is as follows: π_d is uniform over $(-1/2 + t, 1/2 - t)$, π_k is Poisson, and $\pi_{\theta|k}$ has the following structure: the sum $S = \sum_{j=0}^{k} \theta_j^2 (j+1)^{2\beta}$ has a Gamma distribution truncated to interval $[0, L]$, independently of S, the vector $(\theta_0^2, \theta_1^2 2^{2\beta}, \ldots, \theta_k^2 (k+1)^{2\beta})/S$ is Dirichlet with some coefficients $\alpha_{1,k}, \ldots, \alpha_{k,k}$, and the signs of $\theta_0, \ldots, \theta_k$ have equal probabilities. In particular one may take $\alpha_{j,k} = 1$ for all $j \leq k$, or, if one needs to generate more regular spectral densities, $\alpha_{j,k} = j^{-\kappa}$, for some fixed or random $\kappa > 0$. Another interesting choice for the prior on θ is the following truncated Gaussian process: for each k, and each $j \leq k$, $\theta_j \sim \mathcal{N}(0, \tau_0^2(1+j)^{-2\beta})$ independently apart from the constraint, for some fixed, large $L > 0$:

$$\sum_{j=1}^{k}(1+j)^{2\beta}\theta_j^2 \leq L.$$

Note that we can easily restrict ourselves to the important case $d \geq 0$, i.e. processes having long or short memory but not intermediate memory.

4 Rates of convergence

In this section we first provide a general theorem relating rates of convergence of the posterior distribution to conditions on the prior. These conditions are, in essence, similar to the conditions obtained in the i.i.d. case (e.g. Ghosal et al., 2000): i.e. a condition on the prior mass of Kullback-Leibler neighborhoods of the true spectral density, and an entropy condition on the support of the prior. We then present results specialised to the FEXP prior case.

4.1 Main result

Theorem 4.1. *Let (u_n) be a sequence of positive numbers such that $u_n \to 0$, $nu_n \to +\infty$, and $\bar{\mathcal{B}}_n$ a sequence of balls belonging to $\mathcal{G}(t, m, M, L, \rho)$, and defined as*

$$\bar{\mathcal{B}}_n = \{(d,g) : KL_n(f_0; F(d,g)) \leq u_n/4,\, b_n(f_0, F(d,g)) \leq u_n,\, d_0 \leq d \leq d_0 + \delta\},$$

for some $\delta, L > 0$, $0 < m \leq M$, $\rho \in (0,1]$. Let π be a prior which satisfies all the conditions of Theorem 3.1, and, in addition, such that:

1. *For n large enough, $\pi(\bar{\mathcal{B}}_n) \geq \exp(-nu_n/2)$.*
2. *There exists $\varepsilon > 0$ and a sequence of sets $\bar{\mathcal{F}}_n \subset \{(d,g) :\, h(F(d,g), f_0) \leq \varepsilon\}$, such that, for n large enough,*

$$\pi\left(\bar{\mathcal{F}}_n^c \cap \{(d,g) :\, h(F(d,g), f_0) \leq \varepsilon\}\right) \leq \exp(-2nu_n).$$

3. *There exists a positive sequence (ε_n), $\varepsilon_n^2 \geq u_n$, $\varepsilon_n^2 \to 0$, $n\varepsilon_n^2 \geq C \log n$, for some $C > 0$, satisfying the following conditions. Let*

$$\mathcal{V}_{n,l} = \{(d,g) \in \bar{\mathcal{F}}_n;\, \varepsilon_n^2 l \leq h_n(f_0, F(d,g)) \leq \varepsilon_n^2(l+1)\},$$

with $l_0 \leq l \leq l_n$, with fixed $l_0 \geq 2$ and $l_n = \lceil \varepsilon^2/\varepsilon_n^2 \rceil - 1$. For each $l = l_0, \cdots, l_n$, there exists a net (i.e. a finite collection) $\bar{\mathcal{H}}_{n,l} \subset \mathcal{V}_{n,l}$, with cardinality $\bar{\mathcal{C}}_{n,l}$, such that for all $f = F(d,g)$, $(d,g) \in \mathcal{V}_{n,l}$, there exists $f_{i,l} = F(d_{i,l}, g_{i,l}) \in \bar{\mathcal{H}}_{n,l}$ such that $f_{i,l} \geq f$ and

$$0 \leq g_{i,l}(x) - g(x) \leq l\varepsilon_n^2 g_{i,l}/32 \quad 0 \leq d_{i,l} - d \leq l\varepsilon_n^2(\log n)^{-1},$$

where

$$\log \bar{\mathcal{C}}_{n,l} \leq n\varepsilon_n^2 l^{\alpha}, \quad \textit{with} \quad \alpha < 1.$$

Then, there exist $C, C' > 0$ such that, for n large enough,

$$E_0^n \left[P^\pi \left(h_n(f_0, F(d,g)) \ge l_0 \varepsilon_n^2 \middle| \mathbf{X}_n \right) \right] \le C n^{-3} + 2 e^{-C' n \varepsilon_n^2} + e^{-n u_n / 16}. \tag{4.1}$$

A proof is given in Section 5.2.

The conditions given in Theorem 4.1 are similar in spirit to those considered for rates of convergence of the posterior distribution in the i.i.d. case. The first condition is a condition on the prior mass of Kullback-Leibler neighborhoods of the true spectral density, the second one is necessary to allow for sets with infinite entropy (some kind of non compactness) and the third one is an entropy condition. The inequality (4.1) obtained in Theorem 4.1 is non asymptotic, in the sense that it is valid for all n. However, the distances considered in Theorem 4.1 heavily depend on n and, although they express the impact of the differences between f and f_0 on the observations, they are not of great practical use. For these reasons, the entropy condition is awkward and cannot be directly transformed into some more common entropy conditions. To state a result involving distances between spectral densities that might be more useful, we need to consider some specific class of priors. In the next section, we obtain rates of convergence in terms of the ℓ distance for the class of FEXP priors introduced in Section 3.3. The rates obtained are the optimal rates up to a $(\log n)$ term, at least on certain classes of spectral densities. It is to be noted that the calculations used when working on these classes of priors are actually more involved than those used to prove Theorem 4.1. This is quite usual when dealing with rates of convergence of posterior distributions, however this is emphasized here by the fact that distances involved in Theorem 4.1 are strongly dependent on n. The method used in the case of the FEXP prior can be extended to other types of priors.

4.2 Rates of convergence for the FEXP prior

We apply Theorem 4.1 to the class of FEXP priors introduced in Section 3.3. Recall that under such a prior a spectral density f is parametrised as $f = \widetilde{F}(d,k,\theta)$, see (3.1). We make the same assumptions as in Section 3.3. In particular, the prior $\pi(d,k,\theta)$ factorises as $\pi_d(d)\pi_k(k)\pi_\theta(\theta|k)$, the right tail of π_k is such that

$$\exp\{-Ck\log k\} \le \pi_k(k) \le \exp\{-C'k\log k\},$$

for some C, $C' > 0$, and for k large enough, and there exists $\beta > 1/2$ such that the Sobolev set $S(\beta, L)$ contains the support of π. The last condition means that $S = \sum_{j=0}^k \theta_j^2 (j+1)^{2\beta} \in [0, L]$ with prior probability one. In addition, we assume that the support of π_d is $[-1/2+t, 1/2-t]$, and, for $d \in [-1/2+t, 1/2-t]$, $\pi_d(d) \ge c_d > 0$. Similarly, we assume that $\pi_{\theta|k}$ is such that the random variable $S = \sum_{j=0}^k \theta_j^2 (j+1)^{2\beta}$ is independent of k, and admits a probability density $\pi_S(s)$ with support $[0, L]$, and such that $\pi_S(s) \ge c_s > 0$ for $s \in [0, L]$.

Theorem 4.2. *For the FEXP prior described above, there exist $C, C' > 0$ such that, for n large enough*

$$E_0^n \left\{ P^\pi \left[\ell(f, f_0) > \frac{C \log n}{n^{2\beta/(2\beta+1)}} \middle| \mathbf{X}_n \right] \right\} \le \frac{C}{n^2} \tag{4.2}$$

where $f = \widetilde{F}(d,k,\theta)$ and

$$E_0^n \left[\ell(\hat{f}, f_0) \right] \le \frac{C'(\log n)}{n^{2\beta/(2\beta+1)}}, \tag{4.3}$$

where $\log \hat{f}(\lambda) = E^\pi \left[\log f(\lambda) | \mathbf{X}_n \right]$.

A proof is given in Appendix C.

5 Proofs of Theorems 3.1 and 4.1

5.1 Proof of Theorem 3.1

For the sake of conciseness, we introduce the following notations: for any pair (f, f_0) of spectral densities,

$$\begin{aligned} A(f_0, f) &= T_n(f)^{-1} T_n(f_0), \\ B(f_0, f) &= T_n(f_0)^{1/2} [T_n(f)^{-1} - T_n(f_0)^{-1}] T_n(f_0)^{1/2}. \end{aligned}$$

The proof borrows ideas from Ghosal et al. (2000). The main difficulty is to formulate constraints on quantities such as $h_n(f, f_0)$ or $KL_n(f, f_0)$ in terms of distances between f, f_0, independent on n, and uniformly over f. One has

$$P^\pi \left[\mathcal{A}_\varepsilon^c | \mathbf{X}_n\right] = \frac{\int 1\!\!1_{\mathcal{A}_\varepsilon^c}(f) \varphi(\mathbf{X}_n; f)/\varphi(\mathbf{X}_n; f_0) d\pi(f)}{\int \varphi(\mathbf{X}_n; f)/\varphi(\mathbf{X}_n; f_0) d\pi(f)} \triangleq \frac{N_n}{D_n}.$$

Let $\delta \in (0, \varepsilon)$ and P_0^n be a generic notation for probabilities associated to the distribution of $\mathbf{X}_n$, under the true spectral density $f_0 = F(d_0, g_0)$. One has

$$P_0^n \left\{ P^\pi \left[\mathcal{A}_\varepsilon^c | \mathbf{X}_n\right] \geq e^{-n\delta} \right\} \leq P_0^n \left[D_n \leq e^{-n\delta} \right] + P_0^n \left[N_n \geq e^{-2n\delta} \right] \tag{5.1}$$

The following Lemma bounds the first term.

Lemma 1. *There exists $C > 0$ such that*

$$P_0^n \left[D_n \leq e^{-n\delta} \right] \leq C n^{-3}. \tag{5.2}$$

Proof. Lemma 4 implies that, when n is large enough, $\tilde{\mathcal{B}}_n \supset \mathcal{B}_{\delta/8}$, where

$$\tilde{\mathcal{B}}_n = \{(d, g) \in [-1/2 + t, 1/2 - t] \times \mathcal{G}(m, M, L, \rho) : KL_n(f_0, F(d, g)) \leq \delta/4\}.$$

and Condition 2 implies that, for n large enough, $\pi(\tilde{\mathcal{B}}_n) \geq \pi(\mathcal{B}_{\delta/8}) \geq 2e^{-n\delta/2}$. Consider the indicator function

$$\Omega_n = 1\!\!1 \left[-\mathbf{X}_n^t \left\{ T_n(f)^{-1} - T_n(f_0)^{-1} \right\} \mathbf{X}_n + \log\det A(f_0, f) > -n\delta \right],$$

with implicit arguments $(f, \mathbf{X}_n)$, then, following Ghosal et al. (2000),

$$\begin{aligned} P_0^n \left[D_n \leq e^{-n\delta} \right] &\leq P_0^n \left(\int \Omega_n 1\!\!1_{\tilde{\mathcal{B}}_n}(f) \frac{\varphi(\mathbf{X}_n; f)}{\varphi(\mathbf{X}_n; f_0)} d\pi(f) \leq e^{-n\delta/2} \frac{\pi(\tilde{\mathcal{B}}_n)}{2} \right) \\ &\leq P_0^n \left(E^\pi \left\{ \Omega_n 1\!\!1_{\tilde{\mathcal{B}}_n}(f) \right\} \leq \pi(\tilde{\mathcal{B}}_n)/2 \right) \\ &\leq P_0^n \left(E^\pi \left\{ (1 - \Omega_n) 1\!\!1_{\tilde{\mathcal{B}}_n}(f) \right\} \geq \pi(\tilde{\mathcal{B}}_n)/2 \right) \\ &\leq \frac{2}{\pi(\tilde{\mathcal{B}}_n)} \int_{\tilde{\mathcal{B}}_n} E_0^n \left\{ 1 - \Omega_n \right\} d\pi(f). \end{aligned}$$

by Markov inequality. Besides,

$$\begin{aligned} E_0^n \left\{ 1 - \Omega_n \right\} &= P_0^n \left\{ \mathbf{X}_n^t \left\{ T_n(f)^{-1} - T_n(f_0)^{-1} \right\} \mathbf{X}_n - \log\det A(f_0, f) > n\delta \right\} \\ &= P_{\mathbf{Y}} \left\{ \mathbf{Y}^t B(f_0, f) \mathbf{Y} - \mathrm{tr}\left[B(f_0, f) \right] > D(f_0, f) \right\} \end{aligned}$$

where $\mathbf{Y} \sim N_n(\mathbf{0}_n, \mathbf{I}_n)$, and, for $f \in \tilde{\mathcal{B}}_n$

$$D(f_0, f) \triangleq n\delta + \log\det A(f_0, f) - \mathrm{tr}\left[B(f_0, f)\right] > n\delta/2$$

thus

$$\begin{aligned} E_0^n[1 - \Omega_n] &\leq P_{\mathbf{Y}}\left\{\mathbf{Y}^t B(f_0, f)\mathbf{Y} - \mathrm{tr}[B(f_0, f)] > n\delta/2\right\} \\ &\leq \frac{16}{n^4\delta^4} E_{\mathbf{Y}}\left[\left\{\mathbf{Y}^t B(f_0, f)\mathbf{Y} - \mathrm{tr}[B(f_0, f)]\right\}^4\right] \\ &\leq \frac{C}{n^3\delta^4}, \end{aligned}$$

which concludes the proof.

□

A bound for the second term in (5.1) is obtained as follows:

$$\begin{aligned} P_0^n\left[N_n \geq e^{-2n\delta}\right] &\leq 2e^{2n\delta}\pi(\mathcal{F}_n^c) + p \\ &\leq 2e^{-n(r-2\delta)} + p \end{aligned} \tag{5.3}$$

using Condition 3, where

$$p \triangleq P_0^n\left[\int 1\!\!1(A_\varepsilon^c \cap \mathcal{F}_n)\frac{\varphi(\mathbf{X}_n; f)}{\varphi(\mathbf{X}_n; f_0)}d\pi(f) \geq e^{-2n\delta}/2\right].$$

Assuming $2\delta < r$, we consider the following likelihood ratio tests for each $f_i \in \mathcal{H}_n$, and for some arbitrary values ρ_i,

$$\phi_i = 1\!\!1\left\{\mathbf{X}_n^t\left[T_n^{-1}(f_0) - T_n^{-1}(f_i)\right]\mathbf{X}_n \geq n\rho_i\right\}.$$

Lemmas 7, 8 and 9 given in Appendix B prove that, for each of the three cases in Condition 3 of Theorem 3.1, and well-chosen values of ρ_i, one has

$$E_0^n[\phi_i] \leq e^{-nC_1\varepsilon}, \quad E_f^n[1 - \phi_i] \leq e^{-nC_1\varepsilon}, \tag{5.4}$$

for all f_i, for f close to f_i (in the sense defined in cases a,b, and c in Condition 3), where $C_1 > 0$ is a constant that does not depend on f_i, and E_f^n stands for the expectation with respect to the likelihood $\varphi(\mathbf{X}_n; f)$. Then one concludes easily as follows. Let $\phi^{(n)} = \max_i \phi_i$, then, using Markov inequality, for n large enough,

$$\begin{aligned} p &\leq E_0^n\left[\phi^{(n)}\right] + 2e^{2n\delta}\int_{A_\varepsilon^c \cap \mathcal{F}_n} E_f\left[1 - \phi^{(n)}\right]d\pi(f) \\ &\leq \mathcal{C}_n e^{-nC_1\varepsilon} + 2e^{2n\delta - nC_1\varepsilon} \\ &\leq e^{-nC_1\varepsilon/2}, \end{aligned} \tag{5.5}$$

provided $\delta < C_1\varepsilon/4$. Combining (5.2), (5.3) and (5.5), there exists $\delta > 0$ such that

$$P_0^n\left[P^\pi[A_\varepsilon^c|\mathbf{X}_n] > e^{-n\delta}\right] \leq Cn^{-3}$$

for n large enough, which implies that $P^\pi[A_\varepsilon^c|\mathbf{X}_n] \to 0$ a.s.

5.2 Proof of Theorem 4.1

This proof uses the same notations as the previous Section, e.g. C, C' denote generic constants, f, $d\pi(f)$ are short-hands for $f = F(d,g)$, $d\pi(d,g)$, respectively, $A(f,f_0)$ and $B(f,f_0)$ have the same definition, and so on. In the proof of Theorem 3.1, we showed that $E_0^n\left[P^\pi(h(f,f_0) \geq \varepsilon|\mathbf{X}_n)\right] \leq Cn^{-3}$ for ε small enough, n large enough. Thanks to the uniform convergence Lemmas 3 and 4 in Appendix A, one sees that the same inequality holds if h is replaced by h_n. Therefore, to obtain inequality (4.1), it is sufficient to bound the expectation of the sum of the following probabilities:

$$P^\pi\left((d,g) \in \mathcal{W}_{n,l}|\mathbf{X}_n\right) = \frac{\int 1\!\!1_{\mathcal{W}_{n,l}}(d,g)\frac{\varphi(\mathbf{X}_n;f)}{\varphi(\mathbf{X}_n;f_0)}d\pi(f)}{\int \frac{\varphi(\mathbf{X}_n;f)}{\varphi(\mathbf{X}_n;f_0)}d\pi(f)} = \frac{N_{n,l}}{D_n},$$

for $l_0 \leq l \leq l_n$, where

$$\mathcal{W}_{n,l} = \left\{(d,g) : \, h(f,f_0) \leq \varepsilon, \, \varepsilon_n^2 l \leq h_n(f_0,f) \leq \varepsilon_n^2(l+1)\right\},$$

and $\mathcal{V}_{n,l} = \mathcal{W}_{n,l} \cap \bar{\mathcal{F}}_n$. Following the same lines as in Section 5.1, one has

$$\begin{aligned} E_0^n\left[\sum_{l=l_0}^{l_n} \frac{N_{n,l}}{D_n}\right] &\leq P_0^n\left(D_n \leq e^{-nu_n}/2\right) \\ &+ E_0^n\left[\sum_{l=l_0}^{l_n} \frac{N_{n,l}}{D_n} 1\!\!1\left(D_n \geq e^{-nu_n}/2\right)\right]. \end{aligned} \tag{5.6}$$

The first term is bounded as in Lemma 1, see Section 5.1:

$$\begin{aligned} P_0^n\left(D_n \leq e^{-nu_n}/2\right) &\leq P_0^n\left(D_n \leq \frac{e^{-nu_n/2}\pi(\bar{\mathcal{B}}_n)}{2}\right) \\ &\leq \frac{2\int_{\mathcal{B}_n} E_0^n\left[(1-\Omega_n(f))\right]d\pi(f)}{\pi(\bar{\mathcal{B}}_n)}, \end{aligned}$$

where Ω_n is the indicator function of

$$\left\{(\mathbf{X}_n, f); \mathbf{X}_n^t(T_n^{-1}(f) - T_n^{-1}(f_0))\mathbf{X}_n - \log\det[A(f_0,f)] \leq nu_n\right\},$$

and, for $f \in \bar{\mathcal{B}}_n$, using Chernoff-type inequalities as in Lemma 7, together with the fact that there exists $s_0 > 0$ fixed such that for all $s \leq s_0$

$$\mathbf{I}_n(1+2s) - 2sT_n(f_0)^{1/2}T_n(f)^{-1}T_n(f_0)^{1/2} \geq \mathbf{I}_n/2,$$

for $f = F(d,g)$, $d \geq d_0$, $g > 0$, we have for all $0 < s \leq s_0$

$$\begin{aligned} & E_0^n\left[1-\Omega_n\right] \\ &\leq \exp\Big\{-snu_n - s\log|T_n(f_0)T_n(f)^{-1}| \\ &\qquad -\frac{1}{2}\log\left|\mathbf{I}_n(1+2s) - 2sT_n(f_0)^{1/2}T_n(f)^{-1}T_n(f_0)^{1/2}\right|\Big\} \\ &\leq \exp\left\{-snu_n + 2snKL_n(f_0,f) + 4s^2nb_n(f_0,f)\right\} \\ &\leq \exp\left\{-\frac{snu_n}{2}(1-8s)\right\} \\ &\leq e^{-Cnu_n}, \end{aligned}$$

where the second inequality comes from a Taylor expansion in s of $\log|\mathbf{I}_n+2s(\mathbf{I}_n-T_n(f_0)^{1/2}T_n(f)^{-1}T_n(f_0)^{1/2})|$, the third from the definition of $\bar{\mathcal{B}}_n$ and the last from choosing $s = \min(s_0, 1/16)$. Note that $s_0 \geq m/(M\pi)$ and that the constant C in the above inequality can be chosen as $C = m/(32M\pi)$. The second term of (5.6) equals

$$\begin{aligned} &E_0^n\left[\sum_{l=l_0}^{l_n}\frac{N_{n,l}}{D_n}1\!\!1\left(D_n \geq e^{-nu_n}/2\right)(\bar{\phi}_l+1-\bar{\phi}_l)\right] \\ &\leq \sum_{l=l_0}^{l_n}E_0^n\left[\bar{\phi}_l\right]+2e^{nu_n}\sum_{l=l_0}^{l_n}E_0^n\left[N_{n,l}(1-\bar{\phi}_l)\right] \end{aligned} \quad (5.7)$$

where $\bar{\phi}_l = \max_{i:f_{i,l}\in\bar{\mathcal{H}}_{n,l}}\phi_{i,l}$, $\phi_{i,l}$ is a test function defined as in Section 5.1,

$$\phi_{i,l} = 1\!\!1\left\{\mathbf{X}_n'(T_n^{-1}(f_0)-T_n^{-1}(f_{i,l}))\mathbf{X}_n \geq \mathrm{tr}\left[\mathbf{I}_n - T_n(f_0)T_n^{-1}(f_{i,l})\right]+nh_n(f_0,f_{i,l})/4\right\}.$$

Using inequality (B.2) in Lemma 7, one obtains:

$$\log E_0^n\left[\phi_{i,l}\right] \quad \leq \quad -Cnh_n(f_0,f_i)\min\left(\frac{h_n(f_0,f_i)}{b_n(f_0,f_i)},1\right), \quad (5.8)$$

for some universal constant C, and n large enough. In addition, one has

$$\begin{aligned} \frac{b_n(f_0,f_i)}{h_n(f_0,f_i)} &\leq \left\|T_n(f_0)^{1/2}T_n(f_i)^{-1/2}\right\|^2 \\ &\leq C'n^{2\max(d_0-d_i),0)}, \end{aligned}$$

where the first inequality comes from Lemma 2, see Appendix A.1, and the second inequality comes from Lemma 3 in Lieberman et al. (2009). Hence for all $C > 0$, if $2|d_0-d_i| \leq C/\log n$, $b_n(f_0,f_i) \leq C'e^C h_n(f_0,f_i)$. Moreover for all $\delta > 0$, there exists $C_\delta > 0$ such that if $2|d_0 - d_i| > C_\delta(\log n)^{-1}$ then $h_n(f_0,f_i) \geq n^{-\delta}$. Indeed, equation (A.3) of Lemma 6 implies that if $h_n(f_0,f_i) \geq \varepsilon_n^2$, then

$$h_n(f_0,f_i) \geq \frac{C}{n}\mathrm{tr}\left[T_n(f_0^{-1})T_n(f_i-f_0)T_n(f_i^{-1})T_n(f_i-f_0)\right]$$

and Lemma 5, see Appendix A.3, implies that, for all $a > 0$,

$$\left|\frac{1}{n}\mathrm{tr}\left[T_n(f_0^{-1})T_n(f_i-f_0)T_n(f_i^{-1})T_n(f_i-f_0)\right]-(2\pi)^3\int_{-\pi}^{\pi}\frac{(f_i-f_0)^2}{f_if_0}d\lambda\right| \leq n^{-\rho+a}.$$

Lemma 11, see Appendix D, implies that there exists $a > 0$ such that if $2|d_0-d_i| > C_\delta(\log n)^{-1}$,

$$\int_{-\pi}^{\pi}\frac{(f_i-f_0)^2}{f_if_0}dx \geq Ce^{-a\log n/C_\delta} \geq n^{-\delta}$$

as soon as C_δ is large enough. Choosing $\delta < \rho$ we finally obtain that

$$h_n(f_0,f_i) \geq C'n^{-\delta}.$$

This and the definition of $\bar{\mathcal{H}}_{n,l}$ implies that $l \geq C'n^{-\delta}\varepsilon_n^{-2}$, and therefore $ln^{-\max(d_0-d_i),0)} \geq 2l^\alpha/C'$, for all $\alpha < 1$ as soon as $|d_0-d_i|$ is small enough. (5.8) becomes

$$\log E_0^n\left[\phi_{i,l}\right] \leq -cl\varepsilon_n^2 n^{1-\max(d_0-d_i,0)} \leq -2n\varepsilon_n^2 l^\alpha.$$

Condition 3 implies that

$$E_0^n\left[\bar{\phi}_l\right] \leq \sum_i E_0^n\left[\phi_{i,l}\right] \leq \bar{\mathcal{C}}_{n,l}\exp\{-2n\varepsilon_n^2 l^\alpha\} \leq \exp\{-n\varepsilon_n^2 l^\alpha\}$$

so that $\sum_l E_0^n\left[\bar{\phi}_l\right] \leq 2\exp\{-n\varepsilon_n^2 l_0^\alpha\}$ for n large enough.

For the second term of (5.7), since condition 3 on $f, f_{i,l}$ implies that

$$0 \leq f_{i,l} - f \leq h_n(f_0, f_{i,l})f_{i,l}\left(\frac{\pi^{2(d_i-d)}}{32} + \frac{2|\log|\lambda||}{\log n}\right),$$

when n is large enough, hence $\mathrm{tr}A(f_{i,l}-f, f_0) \leq nh_n(f_0, f_{i,l})/4$ and we obtain the first part of equation B.3:

$$\log E_f^n\left[1-\phi_{i,l}\right] \quad \leq \quad -\frac{n}{64}\min\left(\frac{h_n(f_0, f_{i,l})^2}{b_n(f, f_0)}, 4h_n(f_0, f_{i,l})\right).$$

We also have

$$b_n(f, f_0) \leq b_n(f_{i,l}, f_0) + \frac{h_n^2(f_{i,l}, f_0)}{32} + 2\sqrt{b_n(f_0, f_{i,l})}h_n(f_{i,l}, f_0),$$

hence $\log E_f^n\left[1-\phi_{i,l}\right] \leq -cnl^\alpha\varepsilon_n^2$, using the same arguments as before, and

$$\begin{aligned}
\sum_{l=l_0}^{l_n} E_0^n\left[(1-\bar{\phi}_l)N_{n,l}\right] \quad &= \quad \int\left\{\sum_{l=l_0}^{l_n} 1\!\!1_{\mathcal{W}_{n,l}}(f)E_f(1-\bar{\phi}_l)\right\}d\pi(f) \\
&\leq \quad P^\pi(f \in \mathcal{F}_n^c \cap \{h(f, f_0) \leq \varepsilon\}) \\
&\quad\quad + \sum_{l=l_0}^{l_n}\int 1\!\!1_{\mathcal{V}_{n,l}}(f)E_f^n(1-\bar{\phi}_l)\,d\pi(f) \\
&\leq \quad e^{-n\varepsilon_n^2} + \sum_{l=l_0}^{l_n} e^{-Cn\varepsilon_n^2 l^\alpha} \leq 2e^{-n\varepsilon_n^2}.
\end{aligned}$$

6 Discussion

In this paper we have considered the theoretical properties of Bayesian non parametric estimates of the spectral density for Gaussian long memory processes. Some general conditions on the prior and on the true spectral density are provided to ensure consistency and to determine concentration rates of the posterior distributions in terms of the pseudo-metric $h_n(f_0, f)$. To derive a posterior concentration rate in terms of a more common metric such as l_2, we have considered a specific family of priors based of the FEXP models and also used in the frequentist literature. Gaussian long memory processes lead to complex behaviours, which makes the derivation of concentration rates a difficult task. This paper is thus a step in the direction of better understanding the asymptotic behaviour of the posterior distribution in such models and could be applied to various types of priors on the short memory part - other than the FEXP priors.

The rates we have derived are optimal (up to a $\log n$ term) in Sobolev balls but not adaptive since the estimation procedure depends on the smoothness β. Another constraint in the paper is that the prior needs to be restricted to Sobolev balls with fixed though large radius, forbidding the use of Gaussian distributions on the coefficients appearing in the FEXP representation.

However, it is to be noted that even in the parametric framework existing results on the asymptotic behaviour of likelihood approaches, whether maximum likelihood estimators or Bayesian estimators are all assuming that the parameter space is compact, for the same reason that we have had to constraint the prior on fixed Sobolev balls in the FEXP example. The reason is that the short memory part of the spectral density needs to be uniformly bounded.

A related and fundamental problem is the practical implementation of the model described in the paper. Liseo and Rousseau (2006) adopted a Population MC algorithm which easily deals with the trans-dimensional parameter space issue. We are currently working on alternative computational approaches.

Acknowledgements

This work was supported by the 800-2007–2010 grant ANR-07-BLAN-0237-01 "SP Bayes". Part of this work was done while the third Author was visiting the Université Paris Dauphine, CEREMADE. He thanks for warm hospitality and financial support.

A Technical Lemmas on convergence rates of products of Toeplitz matrices

We first give a set of inequalities on norms of matrices that are useful throughout the proofs. We then give three technical lemmas on the uniform convergence of traces of products of Toeplitz matrices, in the spirit of Lieberman et al. (2003) and Lieberman et al. (2009), but extending those previous results to functional classes instead of parametric classes.

A.1 Some matrix inequalities

Let A and B be n dimensional matrices. We consider the following two norms:

$$|A|^2 = \mathrm{tr}\left[AA^t\right], \quad \|A\|^2 = \sup_{|x|=1} x^t AA^t x.$$

We first recall that:

$$|\mathrm{tr}[AB]| \le |A||B|, \quad |AB| \le \|A\||B|, \quad |A| \le \|A\|, \quad \|AB\| \le \|A\|\|B\|.$$

Using these inequalities we prove the following basic Lemma:

Lemma 2. *Let f_1, f_2 be two spectral densities, then*

$$2nb_n(f_1, f_2) \le n\|T_n(f_2)^{-1/2}T_n(f_1)^{1/2}\|^2 h_n(f_1, f_2)$$

Proof. One has

$$\begin{aligned}
&2nb_n(f_1, f_2)\\
&= \mathrm{tr}\left[T_n(f_1)^{1/2}T_n(f_2)^{-1}T_n(f_1)^{1/2}\left(T_n(f_1)^{-1/2}T_n(f_1 - f_2)T_n(f_2)^{-1/2}\right)^2\right]\\
&= \left|T_n(f_2)^{-1/2}T_n(f_1)^{1/2}\left(T_n(f_1)^{-1/2}T_n(f_1 - f_2)T_n(f_2)^{-1/2}\right)\right|^2\\
&\le \|T_n(f_2)^{-1/2}T_n(f_1)^{1/2}\|^2\left|T_n(f_2)^{-1/2}T_n(f_1 - f_2)T_n(f_2)^{-1/2}\right|^2\\
&= n\|T_n(f_2)^{-1/2}T_n(f_1)^{1/2}\|^2 h_n(f_1, f_2).
\end{aligned}$$

□

A.2 Uniform convergence: Lemmas 3 and 4

We state two technical lemmas, which are extensions of Lieberman et al. (2003) on uniform convergence of traces of Toeplitz matrices, and which are repeatedly used in the paper.

Lemma 3. *Let $t > 0$, $M, L > 0$ and $\rho \in (0, 1]$, let p be a positive integer, we have, as $n \to +\infty$:*

$$\sup_{\substack{f_i=F(d_1,g_i),\, f_i'=F(d_2,g_i')\\ 2p(d_1+d_2)\le 1-t\\ g_i\in\mathcal{G}(-M,M,L,\rho)\\ g_i'\in\mathcal{G}(-M,M,L,\rho)}} \left| \frac{1}{n}\mathrm{tr}\left[\prod_{i=1}^{p} T_n(f_i)T_n(f_i')\right] - \frac{\int_{-\pi}^{\pi}\prod_{i=1}^{p} f_i(\lambda)f_i'(\lambda)\,d\lambda}{(2\pi)^{1-2p}} \right| \to 0.$$

This lemma is a direct adaptation from Lieberman et al. (2003); the only non obvious part is the change from the condition of continuous differentiability in that paper to the Lipschitz condition of order ρ. This different assumption affects only equation (30) of Lieberman et al. (2003), with η_n replaced by η_n^{ρ}, which does not change the convergence results.

Lemma 4. *Let $t > 0$, $M, L, m > 0$ and $\rho_1, \rho_2 \in (0, 1]$, let p be a positive integer, we have, as $n \to +\infty$:*

$$\sup_{\substack{f_i=F(d_1,g_i)\, f_i'=F(d_2,g_i')\\ 4p(d_1-d_2)\le \rho_2+1-t\\ g_i\in\mathcal{G}(-M,M,L,\rho_1)\\ g_i'\in\mathcal{G}(m,M,L,\rho_2)}} \left| \frac{1}{n}\mathrm{tr}\left[\prod_{i=1}^{p} T_n(f_i)T_n(f_i')^{-1}\right] - \frac{1}{2\pi}\int_{-\pi}^{\pi}\prod_{i=1}^{p}\frac{f_i(\lambda)}{f_i'(\lambda)}d\lambda \right| \to 0,$$

Proof. This result is a direct consequence of Lemma 3, as in Lieberman et al. (2003). The only difference is in the proof of Lemma 5.2. of Dahlhaus (1989), i.e. in the study of terms in the form

$$|\mathbf{I}_n - T_n(f)^{1/2}T_n\left((4\pi^2 f)^{-1}\right)T_n(f)^{1/2}|,$$

with $f = F(d_2, g_i')$ for any $i \le p$. For simplicity's sake we write $f = F(d, g)$ in the following calculations. Following Dahlhaus's Dahlhaus (1989) proof, we obtain an upper bound of

$$\left|\frac{f(\lambda_1)}{f(\lambda_2)} - 1\right|$$

which is different from Dahlhaus (1989). If $g \in \mathcal{G}(m, M, L, \rho_2)$, the Lipschitz condition in ρ_2 implies that

$$\left|\frac{f(\lambda_1)}{f(\lambda_2)} - 1\right| \le K\left(|\lambda_1-\lambda_2|^{\rho_2} + \frac{|\lambda_1-\lambda_2|^{1-\delta}}{|\lambda_1|^{1-\delta}}\right).$$

Calculations as in Lemma 5.2 of Dahlhaus (1989) imply that

$$|I - T_n(f)^{1/2}T_n\left((4\pi^2 f)^{-1}\right)T_n(f)^{1/2}|^2 = O(n^{1-\rho_2}\log n^2) + O(n^{\delta}), \quad \forall\delta > 0.$$

From this we prove the Lemma following Lieberman et al. (2009) Lemma 7, the bounds being uniform over the considered class of functions. □

A.3 Order of approximation: Lemma 5

In this section we recall a result given in Kruijer and Rousseau (2010) which is a generalization of Lieberman and Phillips (2004) concerning the convergence rate of

$$\frac{1}{2}\left|\mathrm{tr}\left[\prod_{j=1}^{p} T_n(f_j)T_n(g_j)\right]/n - (2\pi)^{-1}\int_{-\pi}^{\pi}\prod_j f_j(\lambda)g_j(\lambda)d\lambda\right|.$$

Lemma 5. *Let $1/2 > a > 0$, $L > 0$, $M > 0$ and $0 < \rho \le 1$, then for all $\delta > 0$ there exists $C > 0$ such that for all $n \in \mathbb{N}^*$*

$$\sup_{\substack{p(d_1+d_2)\le a \\ g_j, g_j' \in \mathcal{G}(-M,M,L,\rho)}} \left| \frac{1}{n}\mathrm{tr}\left[\prod_{j=1}^{p} T_n(F(d_1,g_j))T_n(F(d_2,g_j'))\right] \right.$$

$$\left. -(2\pi)^{2p-1}\int_{-\pi}^{\pi}\prod_{j=1}^{p} F(d_1,g_j)F(d_2,g_j') \right| \le Cn^{-\rho+\delta+2pa_+}, \tag{A.1}$$

where $d_1, d_2 > -1/2$ and $a_+ = \max(a, 0)$.

A.4 Some other approximations: Lemma 6

Lemma 6. *Let f_j, $j \in \{1,2\}$ be such that $f_j(\lambda) = F(d_j, g_j)$, where $d_j \in (-1/2, 1/2)$, $0 < m \le g_j \le M < +\infty$ for some positive constant m, M and consider b a bounded function on $[-\pi, \pi]$. Assume that $|d_1 - d_2| < \delta$, with $\delta \in (0, 1/4)$, then, provided $d_1 > d_2$,*

$$\frac{1}{n}\mathrm{tr}\left[T_n(f_1)^{-1}T_n(f_1 b)T_n(f_2)^{-1}T_n(f_1 b)\right] \le C(\log n)\left[|b|_2^2 + \delta|b|_\infty^2\right], \tag{A.2}$$

and, without assuming $d_1 > d_2$,

$$\begin{aligned} &\frac{1}{n}\mathrm{tr}\left[T_n(f_1^{-1})T_n(f_1 - f_2)T_n(f_2^{-1})T_n(f_1 - f_2)\right] \\ &\le\ C\left[h_n(f_1,f_2) + n^{\delta-1/2}\sqrt{h_n(f_1,f_2)}\right]. \end{aligned} \tag{A.3}$$

Proof. Throughout the proof C denotes a generic constant. We first prove (A.2). To do so, we first obtain an upper bound on the following quantity:

$$\gamma(b) \;=\; \frac{1}{n}\mathrm{tr}\left[T_n(f_1^{-1})T_n(f_1 b)T_n(f_2^{-1})T_n(f_1 b)\right]. \tag{A.4}$$

First note that b can be replaced by $|b|$ so that we can assume that it is positive. Since the functions g_i are bounded from below and above, we can prove (A.2) by replacing f_i by $|\lambda|^{-2d_i}$. Thus, without loss of generality, we assume that $f_i = |\lambda|^{-2d_i}$. Let $\Delta_n(\lambda) = \sum_{j=1}^{n}\exp(-i\lambda j)$ and L_n be the 2π-periodic function defined by $L_n(\lambda) = n$ if $|\lambda| \le 1/n$ and $L_n(\lambda) = |\lambda|^{-1}$ if $1/n \le |\lambda| \le \pi$. Then $|\Delta_n(\lambda)| \le CL_n(\lambda)$,

$$\int_{-\pi}^{\pi}\Delta_n(\lambda_1 - \lambda_2)\Delta_n(\lambda_2 - \lambda_3)d\lambda_2 = 2\pi\Delta_n(\lambda_1 - \lambda_3), \tag{A.5}$$

and we can express traces of products of Toeplitz matrices in the following way. Let the symbol $d\boldsymbol{\lambda}$ denote the quantity $d\lambda_1 d\lambda_2 d\lambda_3 d\lambda_4$; the conditions on the g_j's imply

$$\begin{aligned} \gamma(b) &= \frac{1}{n}\int_{[-\pi,\pi]^4} b(\lambda_1)b(\lambda_3)\frac{f_1(\lambda_1)f_1(\lambda_3)}{f_2(\lambda_2)f_1(\lambda_4)} \times \\ &\qquad \Delta_n(\lambda_1-\lambda_2)\Delta_n(\lambda_2-\lambda_3)\Delta_n(\lambda_3-\lambda_4)\Delta_n(\lambda_4-\lambda_1)d\boldsymbol{\lambda} \\ &= \frac{(2\pi)^2}{n}\int_{[-\pi,\pi]^2} b(\lambda_1)b(\lambda_3)|\lambda_3|^{-2\delta}\Delta_n(\lambda_1-\lambda_3)\Delta_n(\lambda_3-\lambda_1)d\lambda_1 d\lambda_3 \\ &\quad + \frac{1}{n}\int_{[-\pi,\pi]^4} b(\lambda_1)b(\lambda_3)|\lambda_3|^{-2\delta}\left[\left|\frac{\lambda_3}{\lambda_2}\right|^{-2d_2}\left|\frac{\lambda_1}{\lambda_4}\right|^{-2d_1} - 1\right] d\boldsymbol{\lambda}, \end{aligned} \tag{A.6}$$

as $d_1 - d_2 \le \delta$. We decompose the following factor in the integrand:

$$\begin{aligned}\left|\frac{\lambda_3}{\lambda_2}\right|^{-2d_2}\left|\frac{\lambda_1}{\lambda_4}\right|^{-2d_1} &= \left(\left|\frac{\lambda_3}{\lambda_2}\right|^{-2d_2}-1\right)\left(\left|\frac{\lambda_1}{\lambda_4}\right|^{-2d_1}-1\right) \\ &\quad+\left(\left|\frac{\lambda_3}{\lambda_2}\right|^{-2d_2}-1\right)+\left(\left|\frac{\lambda_1}{\lambda_4}\right|^{-2d_1}-1\right)+1\end{aligned} \tag{A.7}$$

and treat each corresponding integral separately. Starting with the first term, replacing Δ_n by L_n, we obtain:

$$\begin{aligned}&\frac{1}{n}\int_{[-\pi,\pi]^4} b(\lambda_1)b(\lambda_3)|\lambda_3|^{-2\delta}\Delta_n(\lambda_1-\lambda_3)\Delta_n(\lambda_3-\lambda_1)d\lambda_1 d\lambda_3 \\ &\le \frac{1}{n}\int_{[-\pi,\pi]^2} b(\lambda_1)b(\lambda_3)|\lambda_3|^{-2\delta}L_n^2(\lambda_1-\lambda_3)d\lambda_1 d\lambda_3 \\ &\le C\int_{[-\pi,\pi]^2} b(\lambda_1)b(\lambda_3)|\lambda_3|^{-2\delta}L_n(\lambda_1-\lambda_3)d\lambda_1 d\lambda_3 \\ &\le C\Big\{\int_{\{b(\lambda_1)>b(\lambda_3)|\lambda_3|^{-2\delta}\}} b^2(\lambda_1)L_n(\lambda_1-\lambda_3)d\lambda_1 d\lambda_3 \\ &\qquad+\int_{\{b(\lambda_1)\le b(\lambda_3)|\lambda_3|^{-2\delta}\}} b^2(\lambda_3)|\lambda_3|^{-4\delta}L_n(\lambda_1-\lambda_3)d\lambda_1 d\lambda_3\Big\} \\ &\le C\Big\{\int b^2(\lambda_1)L_n(\lambda_1-\lambda_3)d\lambda_1 d\lambda_3 \\ &\qquad+\int b^2(\lambda_3)\left||\lambda_3|^{-4\delta}-1\right|L_n(\lambda_1-\lambda_3)d\lambda_1 d\lambda_3\Big\} \\ &\le C(\log n)\left\{|b|_2^2+\delta\,|b|_\infty^2\right\},\end{aligned}$$

using calculations similar to Dahlhaus (1989, Lemma 5.2).

For the integral corresponding to the second term in (A.6), we note first that for $0 < a < 1 - d_1 < 1 - d_2$,

$$\left|\frac{\lambda_1}{\lambda_4}\right|^{-2d_1} - 1 \le C\frac{|\lambda_1-\lambda_4|^{1-a}}{|\lambda_1|^{1-a}},$$

and the same inequality holds if λ_1, λ_4 and d_1 are replaced, respectively, by λ_3, λ_2 and d_2. Using the same calculations as the proof of Lemma 5.2 in Dahlhaus (1989), one has

$$\begin{aligned}&\int_{[-\pi,\pi]^4} b(\lambda_1)b(\lambda_3)|\lambda_3|^{-2\delta}\left(\left|\frac{\lambda_3}{\lambda_2}\right|^{-2d_2}-1\right)\left(\left|\frac{\lambda_1}{\lambda_4}\right|^{-2d_1}-1\right)\times \\ &\qquad L_n(\lambda_1-\lambda_2)L_n(\lambda_2-\lambda_3)L_n(\lambda_3-\lambda_4)L_n(\lambda_4-\lambda_1)d\boldsymbol{\lambda} \\ &\le C|b|_\infty^2\int_{[-\pi,\pi]^4}\frac{L_n(\lambda_1-\lambda_2)L_n(\lambda_2-\lambda_3)^a L_n(\lambda_3-\lambda_4)L_n(\lambda_4-\lambda_1)^a}{|\lambda_1|^{1-a}|\lambda_3|^{1-a+2\delta}}d\boldsymbol{\lambda} \\ &\le C|b|_\infty^2 n^{2a}(\log n)^2,\end{aligned}$$

provided $a > 2\delta$. Taking $a = 3\delta < 1/2$ and doing the same calculations for the integrals corresponding to the two intermediate terms in (A.6), one eventually obtains, when n is large enough

$$\gamma(b) \le C(\log n)\left\{|b|_2^2+\delta\,|b|_\infty^2\right\}. \tag{A.8}$$

We now prove that, for large n and $\forall a > 0$,

$$\frac{1}{n}\mathrm{tr}\left[T_n(f_1)^{-1}T_n(f_1 b)T_n(f_2)^{-1}T_n(f_1 b)\right] \leq C\left\{\gamma(b) + n^{a-1}\right\}.$$

Let

$$\begin{aligned}
\delta_n &= \mathrm{tr}\left[T_n(f_1 b)T_n^{-1}(f_2)T_n(f_1 b)T_n^{-1}(f_1)\right] \\
&= \mathrm{tr}\left[T_n(f_1 b)T_n(f_2^{-1}/4\pi^2)T_n(f_1 b)T_n(f_1^{-1}/4\pi^2)\right] \\
&\quad +\mathrm{tr}\left[T_n(f_1 b)T_n^{-1}(f_2)T_n(f_1 b)T_n^{-1/2}(f_1)R_1 T_n^{-1/2}(h_1)\right] \\
&\quad +\mathrm{tr}\left[T_n(f_1 b)T_n(f_2)^{-1/2}R_2 T_n(f_2)^{-1/2}T_n(f_1 b)T_n(f_1^{-1}/4\pi^2)\right],
\end{aligned}$$

where $R_i = \mathbf{I}_n - T_n(f_i)^{1/2}T_n(f_i^{-1}/4\pi^2)T_n(f_i)^{1/2}$, $i = 1, 2$. We bound the first term with (A.8):

$$\frac{1}{n}\mathrm{tr}\left[T_n(f_1 b)T_n(f_2^{-1})T_n(f_1 b)T_n(f_1^{-1})\right] \quad \leq \quad C(\log n)^3\left\{|b_n|_2^2 + \delta|b|_\infty\right\}.$$

Moreover

$$\begin{aligned}
&\left|\mathrm{tr}\left[T_n(f_1 b)T_n^{-1}(f_2)T_n(f_1 b)T_n^{-1/2}(f_1)R_1 T_n^{-1/2}(f_1)\right]\right| \\
&\leq |R_1||T_n^{-1/2}(g_1)T_n(f_1 b)T_n^{-1}(f_2)T_n(f_1 b)T_n^{-1/2}(f_1)| \\
&\leq \delta_n^{1/2}|R_1|\|T_n^{-1/2}(f_2)T_n(f_1 b)^{1/2}\|\|T_n(f_1 b)^{1/2}T_n^{-1/2}(f_1)\|
\end{aligned}$$

Lemmas 5.2 and 5.3 in Dahlhaus (1989) lead to, $\forall a > 0$,

$$\begin{aligned}
\left|\mathrm{tr}\left[T_n(f_1 b)T_n^{-1}(f_2)T_n(f_1 b)T_n^{-1/2}(f_1)R_1 T_n^{-1/2}(f_1)\right]\right| &\leq Cn^{a+(d_1-d_2)}|b|_\infty\delta_n^{1/2} \\
&\leq Cn^{2\delta}|b|_\infty\delta_n^{1/2}
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\left|\mathrm{tr}\left[T_n(f_1 b)T_n(f_2)^{-1/2}R_2 T_n(f_2)^{-1/2}T_n(f_1 b)T_n(f_1^{-1}/(4\pi^2))\right]\right| \\
&\leq |R_2|\delta_n^{1/2}\|T_n(f_2)^{-1/2}T_n(f_1 b)^{1/2}\|^2 \\
&\leq n^{a+2\delta}|b|_\infty\delta_n^{1/2}
\end{aligned}$$

for all $a > 0$. Finally we obtain, when n is large enough

$$\delta_n \leq Cn^{3\delta}|b|_\infty\delta_n^{1/2} + C(\log n)^3\left\{|b_n|_2^2 + \delta|b|_\infty\right\},$$

which ends the proof of (A.2).

We now prove (A.3). Since $f_j \geq mh_j := m|\lambda|^{-2d_i}$, $T_n^{-1}(f_j) \leq T_n^{-1}(h_j)$, i.e. $T_n^{-1}(h_j) - T_n^{-1}(f_j)$

is positive semidefinite, and

$$
\begin{aligned}
& h_n(f_1,f_2) && \text{(A.9)}\\
= \; & \frac{1}{2n}\mathrm{tr}\left[T_n(f_1-f_2)T_n^{-1}(f_2)T_n(f_1-f_2)T_n^{-1}(f_1)\right] \\
\geq \; & \frac{1}{2n}\mathrm{tr}\left[T_n(f_1-f_2)T_n^{-1}(h_2)T_n(f_1-f_2)T_n^{-1}(h_1)\right] \\
\geq \; & \frac{1}{2n}\mathrm{tr}\left[T_n(f_1-f_2)T_n^{-1}(h_2)T_n(f_1-f_2)T_n^{-1/2}(h_1)R_1T_n^{-1/2}(h_1)\right] \\
& +\frac{1}{2n}\mathrm{tr}\left[T_n(f_1-f_2)T_n^{-1}(h_2)T_n(f_1-f_2)T_n\left(\frac{h_1^{-1}}{4\pi^2}\right)\right] \\
= \; & \frac{1}{n(8\pi^2)}\mathrm{tr}\left[T_n(f_1-f_2)T_n^{-1/2}(h_2)R_2T_n^{-1/2}(h_2)T_n(f_1-f_2)T_n\left(h_1^{-1}\right)\right] \\
+ \; & \frac{1}{2n}\mathrm{tr}\left[T_n(f_1-f_2)T_n^{-1}(h_2)T_n(f_1-f_2)T_n^{-1/2}(h_1)R_1T_n^{-1/2}(h_1)\right] \\
+ \; & \frac{1}{n(32\pi^4)}\mathrm{tr}\left[T_n(f_1-f_2)T_n(h_2^{-1})T_n(f_1-f_2)T_n\left(h_1^{-1}\right)\right] && \text{(A.10)}
\end{aligned}
$$

where $R_j = \mathbf{I}_n - T_n^{1/2}(h_j)T_n(h_j^{-1}/(4\pi^2))T_n^{1/2}(h_j)$. We first bound the second term of the r.h.s. of (A.9). Let $\delta > 0$ and $\varepsilon < \varepsilon_0$ such that $|d-d_0| \leq \delta$ (Corollary 1 implies that there exists such a value ε_0). Then using Lemmas 5.2 and 5.3 of Dahlhaus (1989)

$$
\begin{aligned}
& \left|\mathrm{tr}\left[T_n(f_1-f_2)T_n^{-1}(h_2)T_n(f_1-f_2)T_n^{-1/2}(h_1)R_1T_n^{-1/2}(h_1)\right]\right| \\
\leq \; & 2|R_1||T_n^{-1/2}(h_1)T_n(f_1-f_2)T_n^{-1/2}(h_2)|\|T_n(|f_1-f_2|)^{1/2}T_n^{-1/2}(h_2)\| \\
\times \; & \|T_n(|f_1-f_2|)^{1/2}T_n^{-1/2}(h_1)\| \\
\leq \; & Cn^{3\delta}|T_n^{-1/2}(h_1)T_n(f_1-f_2)T_n^{-1/2}(h_2)|.
\end{aligned}
$$

Since $h_i \leq Cf_i$,

$$
\begin{aligned}
|T_n^{-1/2}(h_1)T_n(f_1-f_2)T_n^{-1/2}(f_2)|^2 &= \mathrm{tr}\left[T_n^{-1}(h_1)T_n(f_1-f_2)T_n^{-1}(f_2)T_n(f_1-f_2)\right] \\
&\leq C\mathrm{tr}\left[T_n^{-1}(f_1)T_n(f_1-f_2)T_n^{-1}(f_2)T_n(f_1-f_2)\right] \\
&= Cnh_n(f_1,f_2),
\end{aligned}
$$

and

$$
\frac{1}{n}\left|\mathrm{tr}\left[T_n(f_1-f_2)T_n^{-1}(f_2)T_n(f_1-f_2)T_n^{-1/2}(g_1)R_1T_n^{-1/2}(g_1)\right]\right| \leq Cn^{3\delta-1/2}\sqrt{h_n(f_1,f_2)}.
$$

We now bound the first term of the r.h.s. of (A.9).

$$
\begin{aligned}
= \; & \left|\frac{1}{n}\mathrm{tr}\left[T_n(f_1-f_2)T_n^{-1/2}(h_2)R_2T_n^{-1/2}(h_2)T_n(f_1-f_2)T_n(h_1^{-1})\right]\right| \\
\leq \; & \frac{1}{n}|R_2||T_n^{-1/2}(h_2)T_n(f_1-f_2)T_n(h_1)^{-1/2}|\|T_n(h_1)^{1/2}T_n(h_1^{-1})T_n(|f_1-f_2|)T_n^{-1/2}(h_2)\| \\
\leq \; & \frac{Cn^{\delta}\sqrt{nh_n(f_2,f_1)}}{n}\|T_n(h_1)^{1/2}T_n(h_1^{-1})T_n(|f_1-f_2|)T_n^{-1/2}(f_2)\| \\
\leq \; & \frac{Cn^{\delta+1/2}\sqrt{h_n(f_2,f_1)}}{n}\|T_n(h_1)^{1/2}T_n(h_1^{-1})^{1/2}\|^2 \\
& \times\|T_n(h_1)^{-1/2}T_n(|f_1-f_2|)^{1/2}\|\|T_n(|f_1-f_2|)^{1/2}T_n^{-1/2}(f_2)\| \\
\leq \; & Cn^{3\delta-1/2}\sqrt{h_n(f_1,f_2)},
\end{aligned}
$$

Where the latter inequality comes from Lemma 5.3 of Dahlhaus (1989) and from the fact that

$$\|T_n(h_1)^{1/2}T_n(h_1^{-1})^{1/2}\|^2 = 4\pi^2\|T_n(h_1)^{1/2}T_n(h_1^{-1}/(4\pi^2))T_n(h_1)^{1/2}\| \leq |R_1| + 1$$

Therefore,

$$\frac{1}{n}\mathrm{tr}\left[T_n(f_1{-}f_2)T_n(h_2^{-1})T_n(f_1{-}f_2)T_n(h_1^{-1})\right] \leq C\,\left[h_n(f_1,f_2) + n^{-1/2+3\delta}\sqrt{h_n(f_1,f_2)}\right],$$

and, using the fact that $C\,g_j > f_j$, for $j = 1, 2$ this proves (A.9).

□

B Construction of tests: Lemmas 7, 8 and 9

Lemma 7. *If* $8|d_0 - d_i| \leq \rho + 1 - t$ *(case a of Condition 1), the inequalities in* (5.4) *are verified provided* $\rho_i = \mathrm{tr}\left[\mathbf{I}_n - T_n(f_0)T_n^{-1}(f_i)\right]/n + h_n(f_0, f_i)$, $f \leq f_i$ *and*

$$\frac{1}{2\pi}\int_0^{\pi}\frac{f_i(\lambda) - f(\lambda)}{f_0(\lambda)}\,d\lambda \leq h(f_0, f_i)/4. \tag{B.1}$$

Proof. For all $s \in (0, 1/4)$, using Markov inequality,

$$\begin{aligned}
E_0^n[\phi_i] &\leq \exp\{-sn\rho_i\}\,E_0^n\left[\exp\left\{-s\mathbf{X}_n^t\left\{T_n^{-1}(f_i) - T_n^{-1}(f_0)\right\}\mathbf{X}_n\right\}\right] \\
&= \exp\left\{-sn\rho_i - \frac{1}{2}\log\det\left[\mathbf{I}_n + 2sB(f_0, f_i)\right]\right\} \\
&\leq \exp\left\{-sn\rho_i - s\mathrm{tr}\left[B(f_0,f_i)\right] + s^2\mathrm{tr}\left[((\mathbf{I}_n + 2s\tau B(f_0,f_i))^{-2}B(f_0,f_i))^2\right]\right\} \\
&\leq \exp\left\{-sn\rho_i - s\mathrm{tr}\left[B(f_0,f_i)\right] + 4s^2\mathrm{tr}\left[B(f_0,f_i)^2\right]\right\},
\end{aligned}$$

where $\tau \in (0, 1)$, using a Taylor expansion of the log-determinant around $s = 0$, and the following inequality:

$$\mathbf{I}_n + 2s\tau B(f_0, f_i) = (1 - 2s\tau)\mathbf{I}_n + 2s\tau T_n(f_0)^{1/2}T_n(f)^{-1}T_n(f_0) \geq \frac{1}{2}\mathbf{I}_n,$$

since $s\tau < 1/4$. Substituting ρ_i with its expression, the polynomial above is minimal for $s_{\min} = h_n(f_0, f_i)/8b_n(f_0, f_i)$. According to $s_{\min} \in (0, 1/4)$ or not, that is, whether $h_n(f_0, f_i) < 2b_n(f_0, f_i)$ or not, one has:

$$\begin{aligned}
\frac{1}{n}\log E_0^n[\phi_i] &\leq -\frac{h_n(f_0,f_i)^2}{16b_n(f_0,f_i)}\mathbb{1}\left\{h_n(f_0,f_i) < 2b_n(f_0,f_i)\right\} \\
&\quad -\frac{h_n(f_0,f_i) - b_n(f_0,f_i)}{4}\mathbb{1}\left\{h_n(f_0,f_i) \geq 2b_n(f_0,f_i)\right\}, \\
&\leq -\frac{h_n(f_0,f_i)}{16}\min\left\{\frac{h_n(f_0,f_i)}{b_n(f_0,f_i)}, 2\right\}.
\end{aligned} \tag{B.2}$$

Since $8|d_0 - d_i| \leq \rho + 1 - t$, the convergences $b_n(f_0, f_i) \to b(f_0, f_i)$ and $h_n(f_0, f_i) \to h(f_0, f_i)$ are unifom on the support of the prior π, see Lemma 2. One deduces that, for any $a > 0$ and n large enough,

$$\frac{1}{n}\log E_0^n[\phi_i] \leq -\frac{n}{16}\min\left\{\frac{h(f_0,f_i)^2 - a}{b(f_0,f_i) + a}, 2h(f_0,f_i) - a\right\}.$$

Since $f_i \in \mathcal{A}_\varepsilon^c$, $h(f_0, f_i) > \varepsilon$, and one may take $a = \varepsilon^2/2$ to obtain

$$\frac{1}{n} \log E_0^n[\phi_i] \leq -\frac{nh(f_0, f_i)}{32} \min\left\{\frac{h(f_0, f_i)}{b(f_0, f_i) + \varepsilon^2/2}, 2\right\}.$$

Since $|d_0 - d_i| \leq (\rho + 1 - t)/8 \leq 1/4$, Lemma 12, see Appendix D, implies that there exists $C_1 > 0$ such that

$$E_0^n[\phi_i] \leq \exp(-nC_1\varepsilon)$$

for ε small enough.

If f is in the support of π and satisfies $f \leq f_i$, and $8(d_i - d) \leq \rho + 1 - t$, using the same kind of calculations and the fact that

$$\mathbf{I}_n - 2sT_n^{1/2}(f)\left\{T_n^{-1}(f_i) - T_n^{-1}(f_0)\right\}T_n^{1/2}(f) \geq \mathbf{I}_n + 2sB(f, f_0),$$

as $T_n(f) \leq T_n(f_i)$, we obtain for $s \in (0, 1/4)$,

$$\begin{aligned}
E_f^n[1 - \phi_i] &\leq \exp\left\{ns\rho_i - s\mathrm{tr}\left[B(f, f_0)\right] + 4s^2\mathrm{tr}\left[B(f, f_0)^2\right]\right\} \\
&\leq \exp\left\{-nsh_n(f_0, f_i) + s\mathrm{tr}\left[A(f_i - f, f_0)\right] + 4s^2\mathrm{tr}\left[B(f, f_0)^2\right]\right\} \\
&\leq \exp\left\{-nsh_n(f_0, f_i)/2 + 4s^2\mathrm{tr}\left[B(f, f_0)^2\right]\right\}
\end{aligned}$$

where the last inequality comes from (B.1), which implies $\mathrm{tr}\left[A(f_i - f, f_0)\right]/n \leq h_n(f_0, f_i)/2$ for n large enough, uniformly in f, using Lemma 2. Doing the same calculations as above, for n large enough

$$\begin{aligned}
\frac{1}{n} \log E_f^n[1 - \phi_i] &\leq -\frac{1}{64} \min\left\{\frac{h_n(f_0, f_i)^2}{b_n(f, f_0)}, 4h_n(f_0, f_i)\right\} \\
&\leq -\frac{1}{64} \min\left\{\frac{h(f_0, f_i)^2/2}{b(f, f_0) + \varepsilon^2/2}, 2h(f_0, f_i)\right\}. \qquad \text{(B.3)}
\end{aligned}$$

To conclude, note that $f \leq f_i$ and (B.1) implies that

$$\begin{aligned}
b(f, f_0) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{\frac{f^2}{f_0^2} + 1 - 2\frac{f}{f_0}\right\} d\lambda \\
&\leq b(f_i, f_0) + h(f_0, f_i)/2 \\
&\leq (C + 1/2)h(f_0, f_i)
\end{aligned}$$

according to Lemma 12. One concludes that there exists $C_1 > 0$ such that $E_f^n[1 - \phi_i] \leq e^{-nC_1\varepsilon}$. $\square$

Lemma 8. *If $8(d_i - d_0) > \rho + 1 - t$ (case b of Condition 3), the inequalities (5.4) are verified provided $\rho_i = \mathrm{tr}\left[\mathbf{I}_n - T_n(f_0)T_n^{-1}(f_i)\right]/n + 2KL_n(f_0; f_i)$, for any f such that $f \leq f_i$ and*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{f_i}{f} - 1\right) d\lambda \leq \left(\frac{M}{\pi^2 m}\right)^4 \frac{b(f_0, f_i)}{64}, \quad b(f_i, f) \leq b(f_0, f_i) \qquad \text{(B.4)}$$

Note that for ε small enough, if $b(f_i, f) \leq b(f_0, f_i)|\log\varepsilon|^{-1}$, (B.4) is satisfied.

Proof. The upper bound of $E_0^n[\phi_i]$ is computed similarly to (B.2) so that

$$\frac{1}{n}\log E_0^n[\phi_i] \leq -\frac{1}{4}\min\left\{\frac{KL_n(f_0,f_i)^2}{b_n(f_0,f_i)}, KL_n(f_0,f_i)\right\}.$$

According to Lemma 11 and since $8(d_i-d_0) \geq \rho+1-t$, there exists $C>0$, such that $b(f_0,f_i) \geq C$. Using the uniform convergence results of Appendix A, this means that $b_n(f_0,f_i) \geq C/2$, for n large enough, independently of f_i. Using Lemma 13, there exists a constant $C_1 \leq 1$ such that $KL_n(f_0,f_i) \geq C_1 b_n(f_0,f_i)$. Thus, there exists $C_2>0$ such that

$$\frac{1}{n}\log E_0^n[\phi_i] \ \leq -nC_2 b(f_0,f_i),$$

and, for ε small enough, and some $C_3>0$,

$$E_0^n[\phi_i] \leq \exp\{-nC_3\varepsilon\}.$$

As in the previous Lemma, let $h \in (0,1)$:

$$\begin{aligned}
\log E_f^n\left[1-\phi_i\right] &\leq (1-h)n\rho_i/2 \\
&\quad -\frac{1}{2}\log\det\left[\mathbf{I}_n-(1-h)T_n(f)^{1/2}\left\{T_n^{-1}(f_i)-T_n^{-1}(f_0)\right\}T_n(f)^{1/2}\right] \\
&\leq (1-h)n\rho_i/2-\frac{1}{2}\log\det\left[\mathbf{I}_n+(1-h)B(f,f_0)\right] \\
&= (1-h)n\rho_i/2-\log\det[A(f,f_0)]/2 \\
&\quad -\frac{1}{2}\log\det\left[\mathbf{I}_n(1-h)+hT_n^{-1/2}(f)T_n(f_0)T_n^{-1/2}(f)\right].
\end{aligned}$$

Substituting ρ_i with its expression, i.e. $n\rho_i - \log\det A(f,f_0) = \log\det A(f_i,f)$ and using the same kind of expansions as in the previous lemma, one obtains

$$\begin{aligned}
\frac{1}{n}\log E_f^n\left[1-\phi_i\right] &\leq \frac{1}{n}\log\det[A(f_i,f)]+(h/2)\mathrm{tr}\left[T_n(f_0)\left\{T_n^{-1}(f_i)-T_n^{-1}(f)\right\}\right] \\
&\quad -hnLK_n(f_0;f_i)+h^2\mathrm{tr}\left[\left\{\mathbf{I}_n-T_n^{-1}(f)T_n(f_0)\right\}^2\right] \\
&\leq \frac{1}{n}\log\det[A(f_i,f)] \\
&\quad -hnLK_n(f_0;f_i)+h^2\mathrm{tr}\left[\left\{\mathbf{I}_n-T_n^{-1}(f)T_n(f_0)\right\}^2\right] \\
&\leq -\frac{1}{n}\log\det[A(f_i,f)]+ \\
&\quad -n\min\left(\frac{KL_n(f_0,f_i)^2}{4\mathrm{tr}B(f_0,f)^2/n}, \frac{KL_n(f_0,f_i)}{4}\right).
\end{aligned}$$

Note that we use the fact $f \leq f_i$ in the second line.

Since $\log\det A(f_i,f) = \log\det\left\{\mathbf{I}_n+T_n(f_i-f)T_n(f)^{-1}\right\}$, using a Taylor expansion of $\log\det$ around $\mathbf{I}_n$, we obtain that for n large enough

$$\frac{1}{n}\log\det A(f_i,f) \leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{f_i-f}{f}d\lambda + a$$

where a can be chosen as small as necessary. In addition, we use Lemma 13 and the uniform convergence results of Lemmas 3, 4 to obtain that:

$$\frac{(nKL_n(f_0, f_i))^2}{\mathrm{tr}[B(f_0, f)^2]} \geq \frac{nm^4(b(f_0, f_i)^2 - a)^2}{16\pi^8 M^4(b(f_0, f) + a)}$$

and, since $d \geq d_0$ and (B.4),

$$\begin{aligned} b(f_0, f) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{f_0}{f} - 1\right)^2 d\lambda \\ &\leq 2\left(b(f_0, f_i) + \frac{M^2\pi^4}{m^2}b(f_i, f)\right), \\ &\leq 2b(f_0, f_i)\left(1 + \frac{M^2\pi^4}{m^2}\right). \end{aligned}$$

hence, under the constraint (B.4), there exists $C_1 > 0$ such that, for n large enough, ε small enough,

$$E_f^n[1 - \phi_i] \leq \exp\{-nC_1 b(f_0, f_i)\} \leq e^{-n\varepsilon}.$$

□

Lemma 9. *If $8(d_0 - d_i) > \rho + 1 - t$ (case c of Condition 3), the inequalities* (5.4) *are verified provided $\rho_i = \log\det[T_n(f_i)T_n(f_0)^{-1}]/n$ if*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{f_i - f}{f_0}(\lambda)d\lambda \leq \frac{m^2}{4M^2\pi^4}b(f_i, f_0), \quad b(f, f_i) \leq b(f_i, f_0) \tag{B.5}$$

Note that for $\varepsilon > 0$ small enough if $\int(f_i - f)f_i^{-1}d\lambda \leq b(f_i, f_0)|\log\varepsilon|^{-1}$, (B.5) is satisfied.

Proof. For $0 < h < 1$, following the same calculations as in the two previous lemmas, we obtain

$$\begin{aligned} \frac{1}{n}\log E_0^n[\phi_i] &\leq -(1-h)n\rho_i/2 + \log\det[A(f_0, f_i)]/2 \\ &\quad -\frac{1}{2}\log\det\left[\mathbf{I}_n(1-h) + hT_n^{-1/2}(f_0)T_n(f_i)T_n^{-1/2}(f_0)\right] \\ &\leq -nhKL_n(f_i, f_0) + h^2\mathrm{tr}[B(f_i, f_0)^2] \leq -\varepsilon. \end{aligned}$$

Moreover, for all $f \leq f_i$, satisfying $8(d_i - d) \leq \rho + 1 - t$, using the same calculations as in the proof of Lemma 7, we bound $\log E_f^n[1 - \phi_i]$ by the maximum of

$$-\frac{\{nKL_n(f_i, f_0) - \mathrm{tr}[A(f_i - f, f_0)]/2\}^2}{4n\{b(f, f_0) + a\}}$$

and

$$-\frac{n}{4}KL_n(f_i, f_0) + \frac{1}{8}\mathrm{tr}[A(f_i - f, f_0)],$$

where a is any positive constant and n is large enough. Using Lemma 13, one has

$$nKL_n(f_i, f_0) \geq \frac{nm^2}{2\pi^4 M^2}b(f_i, f_0)$$

and the constraints (B.5) we finally obtain that there exists constant $c_1, C_1 > 0$ such that

$$\begin{aligned} E_f^n[1 - \phi_i] &\leq \exp\{-2n(KL_n(f_i, f_0) - \mathrm{tr}[A(f_i - f, f)]/2n) + 4s^2 nb_n(f, f_0)\} \\ &\leq e^{-nc_1 b(f_i, f_0)} \leq e^{-nC_1\varepsilon} \end{aligned}$$

for ε small enough.

□

C Proof of Theorem 4.2

We re-use some of the notations of Section 5.1; in particular, C, C' denote generic constants.

The proof of the theorem is divided in two parts. First, we show that

$$E_0^n \left[P^\pi \left\{ f : h_n(f, f_0) \geq \frac{\log n}{n^{2\beta/(2\beta+1)}} \middle| \mathbf{X}_n \right\} \right] \leq \frac{C}{n^2}. \tag{C.1}$$

Second, we show that, for $f \in \bar{\mathcal{F}}_n$, and n large enough,

$$h_n(f, f_0) \leq C n^{-\frac{2\beta}{2\beta+1}} \log n \Rightarrow h(f, f_0) \leq C' n^{-\frac{2\beta}{2\beta+1}} \log n. \tag{C.2}$$

Since $\ell(f, f_0) \leq h(f, f_0)$, see the proof of Corollary 2 in Section 3, the right-hand side inequality of (C.2) implies that

$$\begin{aligned} E_0^n \left\{ E^\pi \left[\ell(f, f_0) | \mathbf{X}_n \right] \right\} &\leq C \frac{\log n}{n^{2\beta/(2\beta+1)}} \\ &\quad + \bar{\ell} E_0^n \left\{ P^\pi \left(h_n(f, f_0) > \frac{\log n}{n^{2\beta/(2\beta+1)}} | \mathbf{X}_n \right) \right\} \\ &\leq C n^{-\frac{2\beta}{2\beta+1}} \log n + C' n^{-2}, \end{aligned}$$

for large n, where $\bar{\ell} < +\infty$ is an upper bound for $\ell(f, f_0)$ which is easily deduced from the fact that f, f_0 belongs to some Sobolev class of functions. This implies Theorem 4.2.

To prove (C.1), we show that Conditions 1 and 2 of Theorem 4.1 are fulfilled for

$$u_n = n^{-2\beta/(2\beta+1)} (\log n).$$

In order to establish Condition 1, we show that, for n large enough, $\bar{\mathcal{B}}_n \supset \hat{\mathcal{B}}_n$, the set containing all the $f = \tilde{F}(d, k, \theta)$ such that $k \geq \bar{k}_n$, for $\bar{k}_n = k_0 n^{1/(2\beta+1)}$, $d - u_n n^{-a} \leq d_0 \leq d$ and, for $j = 0, \ldots, k$,

$$|\theta_j - \theta_{0j}| \leq (j+1)^{-2\beta} u_n n^{-a}, \tag{C.3}$$

where $a > 0$ is some small constant. Then it is easy to see that $\pi(\bar{\mathcal{B}}_n) \geq \pi(\hat{\mathcal{B}}_n) \geq \exp\{-n u_n / 2\}$, provided k_0 is small enough, since $\pi_k(k \geq \bar{k}_n) \geq \exp\{-C \bar{k}_n \log \bar{k}_n\}$, and (C.3) for all j implies that

$$\begin{aligned} \sum_{j=0}^{k} \theta_j^2 (j+1)^{2\beta} &= \sum_{j=0}^{k} (\theta_{0j} - \theta_{0j} + \theta_j)^2 (j+1)^{2\beta} \\ &\leq L_0 + u_n^2 n^{-2a} \sum_{j=0}^{k} (1+j)^{-2\beta} + 2 u_n n^{-a} \left(\sum_{j=1}^{k} |\theta_{0j}| \right) \\ &< L \end{aligned}$$

for n large enough, since $L_0 = \sum_j \theta_{0j} (j+1)^{2\beta} < L$, and $\sum_{j=1}^{k} |\theta_{0j}|$ is bounded according to (3.3).

Let $f = \tilde{F}(d, k, \theta)$, with $(d, k, \theta) \in \hat{\mathcal{B}}_n$. To prove that $(d, k, \theta) \in \bar{\mathcal{B}}_n$, it is sufficient to prove that $h_n(f, f_0) \leq u_n / 4$, since $h_n(f, f_0) = KL_n(f_0; f) + KL_n(f; f_0)$, and $KL_n(f; f_0) \geq C b_n(f_0, f)$, using the same calculation as in Dahlhaus (1989, p. 1755) and the fact that $d \leq d_0$.

Since $f_0 \in \mathcal{S}(\beta, L)$, and for the particular choice of $\bar{k}_n$ above,

$$\sum_{j=\bar{k}_n}^{+\infty} \theta_{0j}^2 \leq L (\bar{k}_n + 1)^{-2\beta} \tag{C.4}$$

and

$$\begin{aligned}\sum_{j=\bar{k}_n}^{+\infty} |\theta_{0j}| &\leq \left(\sum_{j=\bar{k}_n}^{+\infty} \theta_{0j}^2 (j+1)^{2\beta}\right)^{1/2} \left(\sum_{j=\bar{k}_n}^{+\infty} (j+1)^{-2\beta}\right)^{1/2} \\ &\leq C\bar{k}_n^{1/2-\beta}. \end{aligned} \tag{C.5}$$

Let

$$\begin{aligned} f_{0n}(\lambda) &= |1-e^{i\lambda}|^{-2d_0} \exp\left(\sum_{j=0}^{\bar{k}_n} \theta_{0j}\cos(j\lambda)\right), \\ b_n(\lambda) &= \exp\left(-\sum_{j\geq \bar{k}_n+1} \theta_{0j}\cos(j\lambda)\right) - 1, \end{aligned}$$

and $g_n = 1 - f_{0n}/f$. Then $f - f_0 = f_0 b_n + f g_n$, where b_n and g_n are bounded as follows. From (C.5), one gets that, for n large enough, $|b_n|_\infty \leq C\bar{k}_n^{1/2-\beta}$, and

$$|b_n|_2^2 = \int_{-\pi}^{\pi} b_n(\lambda)^2\, d\lambda \leq 2 \sum_{j=\bar{k}_n+1}^{\infty} \theta_{0j}^2 \leq 2L\bar{k}_n^{-2\beta} \leq 2Lk_0^{-2\beta}\frac{u_n}{\log n}$$

according to (C.4). In addition since $1 - x \leq -\log x$, for $x > 0$,

$$\begin{aligned} g_n(\lambda) &\leq (d_0 - d)\log(1-\cos\lambda) + \sum_{j\leq \bar{k}_n} |\theta_{0j} - \theta_j| \\ &\leq Cu_n n^{-a}\left(|\log|\lambda|| + 1\right). \end{aligned}$$

Moreover, since $\mathrm{tr}\left\{(A+B)^2\right\} \leq 2\mathrm{tr}A^2 + 2\mathrm{tr}B^2$ for square matrices A and B, one has

$$\begin{aligned} h_n(f_0, f) &\leq \frac{1}{n}\mathrm{tr}\left[T_n(f_0 b_n)T_n^{-1}(f)T_n(f_0 b_n)T_n^{-1}(f_0)\right] \\ &\quad + \frac{1}{n}\mathrm{tr}\left[T_n(f g_n)T_n^{-1}(f)T_n(f g_n)T_n^{-1}(f_0)\right] \\ &\leq C\log n\left\{|b_n|_2^2 + u_n n^{-a}|b_n|_\infty^2\right\} \\ &\quad + Cu_n^2 n^{-1-2a}\mathrm{tr}\left[\left(T_n(f\left(|\log|\lambda|| + 1\right))T_n^{-1}(f)\right)^2\right] \\ &\leq cu_n \end{aligned} \tag{C.6}$$

where c may be chosen as small as necessary, since k_0 is arbitrarily large. Note that the first two terms above come from (A.2) in Lemma 6, and the third term comes from Lemma 4.

To establish Condition 2 is straightforward, since the prior has the same form as in Section 3.3, and we can use the same reasoning as in the proof of Theorem 3.2, that is, take, for some well chosen δ,

$$\bar{\mathcal{F}}_n = \left\{(d, k, \theta) \in \mathcal{S}(\beta, L) :\, |d - d_0| \leq \delta,\, k \leq \tilde{k}_n\right\}$$

where $\tilde{k}_n = k_1 n^{1/(2\beta+1)}$ so that, using Lemma 10,

$$\pi\left(\bar{\mathcal{F}}_n^c \cap \{f, h(f, f_0) < \varepsilon\}\right) \leq \pi_k(k \geq \tilde{k}_n) \leq e^{-C\tilde{k}_n \log \tilde{k}_n}$$

for n large enough. Choosing k_1 large enough leads to Condition 2.

We now verify Condition 3 of Theorem 4.2. Let $\varepsilon_n^2 \geq u_n$ and $l_0 \leq l \leq l_n$, and consider $f = \widetilde{F}(d,k,\theta)$, $(d,k,\theta) \in \mathcal{V}_{n,l}$, as defined in Theorem 4.1, and $f_{i,l} = (2e)^{l\varepsilon_n^2}\widetilde{F}(d_i,k,\theta_i)$, where dependencies on l in d_i and θ_i are dropped for convenience. If for some positive $c > 0$ to be chosen accordingly $|\theta_j - \theta_{ij}| \leq cl\varepsilon_n^2/(k+1)$, for $j = 0,\dots,k$, one obtains

$$
\begin{aligned}
\frac{g_{i,l}(\lambda)}{g(\lambda)} &= (2e)^{l\varepsilon_n^2} \exp\left\{\sum_{j=0}^{k}(\theta_j - \theta_{ij})\cos(j\lambda)\right\} \\
&\leq (2e^2)^{cl\varepsilon_n^2}
\end{aligned}
$$

and $f_{i,l}/f \geq 1$ so that the constraints of Condition 3 of Theorem 4.2 are verified by choosing c small enough. The cardinal of the smallest possible net under these constraints needed to cover $\mathcal{V}_{n,l}$ is bounded by

$$
\bar{\mathcal{C}}_{n,l} \leq k_n \left(\frac{1}{cl\varepsilon_n^2}\right)\left(\frac{L'k_n}{cl\varepsilon_n^2}\right)^{k_n+1}
$$

since for all l $|\theta_l| \leq L$. This implies that

$$
\log \bar{\mathcal{C}}_{n,l} \leq Cnu_n
$$

and Condition 3 is verified with $\varepsilon_n^2 = \varepsilon_0^2 u_n$. This achieves the proof of (C.1), which provides a rate of convergence in terms of the distance $h_n(\cdot,\cdot)$.

Finally, we prove (C.2) to obtain a rate of convergence in terms of the distance $h(\cdot,\cdot)$. Consider f such that

$$
h_n(f_0,f) = \frac{1}{2n}\mathrm{tr}\left[T_n^{-1}(f_0)T_n(f-f_0)T_n^{-1}(f)T_n(f-f_0)\right] \leq \varepsilon_n^2.
$$

Equation (A.3) of Lemma 6 implies that

$$
\begin{aligned}
\frac{1}{2n}\mathrm{tr}\left[T_n(f_0^{-1})T_n(f-f_0)T_n(f^{-1})T_n(f-f_0)\right] &\leq C\varepsilon_n[\varepsilon_n + n^{-1/2+\delta}] \\
&\leq C\varepsilon_n^2. \qquad \text{(C.7)}
\end{aligned}
$$

We now prove that

$$
\begin{aligned}
&\mathrm{tr}\left[T_n(f_0^{-1})T_n(f-f_0)T_n(f^{-1})T_n(f-f_0)\right] \\
&\quad -\mathrm{tr}\left[T_n(f_0^{-1}(f-f_0))T_n(f^{-1}(f-f_0))\right] \leq \frac{C(\log n)^2}{n^{1-2a}}.
\end{aligned}
$$

for some small $a > 0$. By symmetry we consider only the case $d \geq d_0$. Let $h_0 = (1-\cos\lambda)^{d_0}$, $h = (1-\cos\lambda)^d$, then $fh \leq C$, $f_0h_0 \leq C$ and $|f-f_0|h \leq C$ for some $C \geq 0$, and it is sufficient to study the difference below. Note that the calculations below follow the same lines and the same notations as the treatment of $\gamma(b)$ in Lemma 6, see Appendix A.

$$
\begin{aligned}
&\frac{1}{n}\mathrm{tr}\left[T_n(h_0(f-f_0))T_n(h(f-f_0))\right] \\
&-\frac{1}{n}\mathrm{tr}\left[T_n(h_0)T_n(f-f_0)T_n(h)T_n(f-f_0)\right] \\
&= \quad -\frac{1}{n}\int_{[-\pi,\pi]^3}(f-f_0)(\lambda_2)h_0(\lambda_2)(f-f_0)(\lambda_4)h(\lambda_4)\left(\frac{h_0(\lambda_1)}{h_0(\lambda_2)}-1\right) \\
&\qquad\times\Delta_n(\lambda_1-\lambda_2)\Delta_n(\lambda_2-\lambda_4)\Delta_n(\lambda_4-\lambda_1)d\boldsymbol{\lambda} \\
&\qquad-\frac{1}{n}\int_{[-\pi,\pi]^3}(f-f_0)(\lambda_2)h_0(\lambda_1)(f-f_0)(\lambda_4)h(\lambda_4)\left(\frac{h(\lambda_3)}{h(\lambda_4)}-1\right) \\
&\qquad\times\Delta_n(\lambda_1-\lambda_2)\Delta_n(\lambda_2-\lambda_3)\Delta_n(\lambda_3-\lambda_4)\Delta_n(\lambda_4-\lambda_1)d\boldsymbol{\lambda} \\
&\leq \quad \frac{C(\log n)}{n}\int_{[-\pi,\pi]^2}|\lambda_2|^{-2(d-d_0)}|\lambda_1|^{-1+a}L_n(\lambda_1-\lambda_2)^{1+a}d\boldsymbol{\lambda} \\
&\qquad+\frac{C}{n}\int_{[-\pi,\pi]^4}\frac{|\lambda_1|^{2d}}{|\lambda_2|^{2d}|\lambda_3|^{1-a}} \\
&\qquad\times L_n(\lambda_1-\lambda_2)L_n(\lambda_2-\lambda_3)L_n(\lambda_3-\lambda_4)^a L_n(\lambda_4-\lambda_1)d\boldsymbol{\lambda} \\
&\leq \quad \frac{C(\log n)^2}{n^{1-a}}\int_{[-\pi,\pi]^2}|\lambda_2|^{-2(d-d_0)}|\lambda_1|^{-1+a}L_n(\lambda_2-\lambda_1)d\boldsymbol{\lambda} \\
&\qquad+\frac{C(\log n)}{n^{1-a}}\int_{[-\pi,\pi]^3}\frac{|\lambda_1|^{2d}}{|\lambda_2|^{2d}|\lambda_3|^{1-a}}L_n(\lambda_1-\lambda_2)L_n(\lambda_2-\lambda_3)d\boldsymbol{\lambda} \\
&\leq \quad \frac{C(\log n)^2}{n^{1-2a}},
\end{aligned}
$$

provided $d-d_0\leq a/4$, using standard calculations and inequality (A.6). Combined with (C.7), this result implies that

$$
\frac{1}{n}\mathrm{tr}\left[T_n\left(h_0(f-f_0)\right)T_n\left(h(f-f_0)\right)\right]\leq C\epsilon_n^2.
$$

Finally, to obtain (C.2), we bound

$$
\begin{aligned}
&\left|\mathrm{tr}\left[T_n(h_0(f-f_0))T_n(h(f-f_0))\right]-\mathrm{tr}\left[T_n(h_0h(f-f_0)^2)\right]\right| \\
&= \quad C\Big|\int_{[-\pi,\pi]^2}\left\{h_0(f-f_0)\right\}(\lambda_1) \\
&\qquad\times\left[\left\{h(f-f_0)\right\}(\lambda_2)-\left\{h(f-f_0)\right\}(\lambda_1)\right]\Delta_n(\lambda_1-\lambda_2)\Delta_n(\lambda_2-\lambda_1)d\boldsymbol{\lambda}\Big| \\
&\leq \quad C\left|\int_{[-\pi,\pi]^2}\left\{h(f-f_0)\right\}(\lambda_1)(f-f_0)(\lambda_2)[h(\lambda_2)-h(\lambda_1)]\Delta_n(\lambda_1-\lambda_2)\Delta_n(\lambda_2-\lambda_1)d\boldsymbol{\lambda}\right| \\
&\qquad+C\left|\int_{[-\pi,\pi]^2}\left\{hh_0(f-f_0)\right\}(\lambda_1)\left[f_0(\lambda_2)-f_0(\lambda_1)\right]\Delta_n(\lambda_1-\lambda_2)\Delta_n(\lambda_2-\lambda_1)d\boldsymbol{\lambda}\right| \\
&\qquad+C\left|\int_{[-\pi,\pi]^2}\left\{hh_0(f-f_0)\right\}(\lambda_1)\left[f(\lambda_2)-f(\lambda_1)\right]\Delta_n(\lambda_1-\lambda_2)\Delta_n(\lambda_2-\lambda_1)d\boldsymbol{\lambda}\right|.
\end{aligned}
$$

The first term is of order $O(n^{2a}\log n)$, from the same calculations as above. We consider the last term, but the calculations for the second term follow exactly the same lines. Recall that

$f = he^w$, where $w(\lambda) = \sum_{j=0}^k \theta_j \cos(j\lambda)$ is not necessarily continuously differentiable, e.g. when $\beta < 1$. Thus

$$f(\lambda_2) - f(\lambda_1) = \left[h(\lambda_2)^{-1} - h(\lambda_1)^{-1}\right] e^{w(\lambda_2)} + h(\lambda_1)^{-1}\left[e^{w(\lambda_2)} - e^{w(\lambda_1)}\right].$$

The first term is dealt with using (A.6), leading to a bound of order $(\log n)^2 n^{2a}$. For the second term, and $k \le k_n$,

$$\begin{aligned}
&\left|\int_{[-\pi,\pi]^2} h_0(f-f_0)(\lambda_1)[g(\lambda_2)-g(\lambda_1)]\Delta_n(\lambda_1-\lambda_2)\Delta_n(\lambda_2-\lambda_1)d\boldsymbol{\lambda}\right| \\
&\le\ C\int_{[-\pi,\pi]^2} h_0|f-f_0|(\lambda_1)\left|\sum_{j=0}^k \theta_j\left(\cos(j\lambda_2)-\cos(j\lambda_1)\right)\right| L_n(\lambda_1-\lambda_2)L_n(\lambda_2-\lambda_1)d\boldsymbol{\lambda} \\
&\le\ C(\log n)\left(\sum_{j=0}^k |\theta_j| j\right)\int_{-\pi}^{\pi}\left\{h_0|f-f_0|\right\}(\lambda_1)d\lambda_1 \\
&\le\ C(\log n)\left(\sum_{j=0}^k |\theta_j| j\right)\left(\int_{-\pi}^{\pi}\left\{hh_0(f-f_0)^2\right\}(\lambda)d\boldsymbol{\lambda}\right)^{1/2},
\end{aligned}$$

where the latter inequality holds because $\int_{-\pi}^{\pi}\{h_0/h\}(\lambda)d\boldsymbol{\lambda}$ is bounded when $|d-d_0|$ is small enough. The same computations can be made on f_0 so that for all $a > 4|d-d_0|$ we finally obtain that

$$\begin{aligned}
&\left|\mathrm{tr}\left[T_n(h_0(f-f_0))T_n(h(f-f_0))\right] - \mathrm{tr}\left[T_n(h_0h(f-f_0)^2)\right]\right| \\
&\le\ C(\log n)n^{2a} + (\log n)\sum_{j=0}^k j(|\theta_j|+|\theta_{0j}|)\left(\int_{[-\pi,\pi]} g_0g(f-f_0)^2(\lambda)d\lambda\right)^{1/2}.
\end{aligned}$$

Splitting the indices of the sum above into into $\left\{j : j|\theta_j| \le j^{2\beta+r}\theta_j^2\right\}$ and its complementary, for some r, we get that

$$\begin{aligned}
\sum_{j=0}^k j|\theta_j| &\le \sum_{j=0}^k j^{2\beta+r}\theta_j^2 + \sum_{j=0}^k j^{1-2\beta-r} \\
&\le C\left(k^r + k^{2-2\beta-r}\right) \le Ck_n,
\end{aligned}$$

provided we take $r = 3/2-\beta$. Using the same computation for f_0, one obtains eventually that, provided $\beta \ge 1/2$,

$$\int_{[-\pi,\pi]} h_0h(f_0-f)^2 d\lambda \le C\varepsilon_n^2,$$

which achieves the proof.

D Technical lemmas

The three following lemmas provide inequalities involving

$$b(f,f_0) = \frac{1}{2\pi}\int_0^\pi (f/f_0-1)^2 d\lambda, \quad h(f,f_0) = \frac{1}{2\pi}\int_0^\pi (f/f_0-1)^2\frac{f_0}{f}d\lambda,$$

for $f = F(d,g)$, $f_0 = F(d_0,g_0)$, $d, d_0 \in (0,1/2)$, $g, g_0 \in \mathcal{G}(m,M)$, for $0 < m < M$.

Lemma 10. *For any $\varepsilon > 0$,*

$$|d - d_0| \geq \varepsilon \Rightarrow h(f, f_0) \geq \frac{1}{\pi}\left(\frac{4M}{m}\right)^{-1/2\varepsilon}.$$

Proof. Without loss of generality, take $d \geq d_0$, then, since $(x-1)^2/x \geq x/2$ for $x \geq 4$,

$$\begin{aligned} h(f, f_0) &\geq \frac{m}{4\pi M}\int_0^\pi 1\!\!1\left\{\lambda^{-2(d-d_0)} \geq 4M/m\right\}\lambda^{-2(d-d_0)}\,d\lambda \\ &\geq \frac{1}{\pi}\left(\frac{4M}{m}\right)^{-1/2\varepsilon}. \end{aligned}$$

$\square$

Lemma 11. *There exists $C > 0$ such that, for any $\varepsilon > 0$,*

$$|d - d_0| \geq \varepsilon \Rightarrow b(f, f_0) \geq C^{-1/2\varepsilon}.$$

Proof. If $d \geq d_0$, then, since $(x-1)^2 \geq x^2/2$ for $x \geq 4$,

$$\begin{aligned} b(f, f_0) &\geq \frac{m^2}{4\pi M^2}\int_0^\pi 1\!\!1\left\{\lambda^{-2(d-d_0)} \geq 4M/m\right\}\lambda^{-4(d-d_0)}\,d\lambda \\ &\geq \frac{4}{\pi}\left(\frac{4M}{m}\right)^{-1/2\varepsilon}. \end{aligned}$$

Otherwise, if $d < d_0$, one has $(x-1)^2 \geq 1/4$ for $0 \leq x \leq 1/2$, so

$$\begin{aligned} b(f, f_0) &\geq \frac{1}{8\pi}\int_0^\pi 1\!\!1\left\{\lambda^{2(d_0-d)} \leq m/2M\right\}\,d\lambda \\ &\geq \frac{1}{8\pi}\left(\frac{2M}{m}\right)^{-1/2\varepsilon}. \end{aligned}$$

$\square$

Lemma 12. *For any $\tau \in (0, 1/4)$, there exists $C > 0$ such that*

$$d - d_0 < \frac{1}{4} - \tau \Rightarrow b(f, f_0) \leq Ch(f, f_0).$$

Proof. If $d \leq d_0$, the bound is trivial, since $f/f_0 \leq M/m\pi^{2(d_0-d)}$. Assume $d > d_0$, and let $A \geq 1/2$ some arbitrary large constant. Since $(x-1)^2 \leq x^2$ for $x \geq 1/2$, one has

$$\begin{aligned} b(f, f_0) &\leq Ah(f, f_0) + \frac{M^2}{2\pi m^2}\int_0^\pi 1\!\!1\left\{f(\lambda)/f_0(\lambda) \geq A\right\}\lambda^{-4(d-d_0)}d\lambda \\ &\leq Ah(f, f_0) + \frac{M^2}{2\pi m^2}\int_0^\pi 1\!\!1\left\{\lambda^{-2(d-d_0)} \geq Am/M\right\}\lambda^{-4(d-d_0)}d\lambda \\ &\leq Ah(f, f_0) + \frac{C'(Am/M)^{2-1/2(d-d_0)}}{1-4t}, \end{aligned} \tag{D.1}$$

provided $A \geq M/m$ and $C' = M^2/2\pi m^2$. In turn, since $(x-1)^2 \geq x^2/2$ for $x \geq 4$, and assuming $A \geq 4M^2/m^2$, then $\lambda^{-2(d-d_0)} \geq Am/M$ implies that $f/f_0 \geq Am^2/M^2 \geq 4$, and $(f/f_0 - 1)^2 f_0/f \geq f/2f_0 \geq Am^2/2M^2$. Therefore

$$\begin{aligned} h(f, f_0) &\geq \frac{1}{2\pi}\int_0^\pi \mathbb{1}\left\{\lambda^{-2(d-d_0)} \geq Am/M\right\}(f/f_0 - 1)^2 \frac{f_0}{f}\, d\lambda && \text{(D.2)} \\ &\geq (Am/M)^{2-1/2(d-d_0)}/4\pi A. && \text{(D.3)} \end{aligned}$$

One concludes the proof by combining (D.1) with (D.3) and taking $A = 4M^2/m^2$.

□

The lemma below makes the same assumptions with respect to f and f_0, but it involves finite n distances.

Lemma 13. *One has:*

$$d > d_0 \Rightarrow KL_n(f_0; f) \geq \frac{m^2}{M^2\pi^2} b_n(f_0, f).$$

Proof. Dahlhaus (1989, p. 1755) proves that $KL_n(f_0; f) \geq C^{-2} b_n(f_0, f)$ where C is the largest eigenvalue of $T_n(f_0)T_n^{-1}(f)$. In our case, $f_0/f \leq M\pi^{2(d-d_0)}/m$, hence $C^{-2} = m^2/M^2\pi^{2(d-d_0)}$.

□

The last lemma in this section applies to the FEXP formulation of Section 3.3.

Lemma 14. *Let*

$$f_0(\lambda) = (2 - 2\cos\lambda)^{-d_0} \exp\{w_0(\lambda)\}, \quad f(\lambda) = (2 - 2\cos\lambda)^{-d} \exp\{w(\lambda)\},$$

then, for $\varepsilon \in (0, 1/4)$,

$$|d - d_0| \leq \varepsilon, |w - w_0| \leq \varepsilon \Rightarrow h(f, f_0) \leq 7\varepsilon.$$

Proof. Without loss of generality, take $d - d_0 \geq 0$. Then $f_0/f - 1 \leq 2^\varepsilon e^\varepsilon - 1 \leq (1 + \log 2)\varepsilon$, since $e^x \leq 1 + 2x$ for $x \in [0, 1]$. Moreover, since $2(1 - \cos\lambda) \geq \lambda^2/3$ for $\lambda \in (0, \pi)$, one has

$$\int_0^\pi \frac{f(\lambda)}{f_0(\lambda)}\, d\lambda = e^\varepsilon 3^{(d-d_0)} \int_0^\pi \lambda^{-2(d-d_0)}\, d\lambda \leq \frac{\pi e^\varepsilon 3^\varepsilon}{1 - 2\varepsilon},$$

and, to conclude, as again $e^x \leq 1 + 2x$ for $x \in [0, 1]$, and $e^{\varepsilon(1+\log 3)}(1 - 2\varepsilon)^{-1} - 1 \leq 10\varepsilon$, for $\varepsilon \leq 1/4$,

$$h(f, f_0) = \frac{1}{2\pi}\int_0^\pi \left(\frac{f(\lambda)}{f_0(\lambda)} + \frac{f_0(\lambda)}{f(\lambda)} - 2\right) d\lambda \leq (6 + \log 2)\varepsilon.$$

□

References

Bardet, J., Philippe, A., Oppenheim, G., Taqqu, M., Stoev, S., and Lang, G. (2003). Semi-parametric estimation of the long-range dependence parameter: A survey. In *Theory and applications of long-range dependence*, pages 557–577. Birkhäuser, Boston, M A.

Beran, J. (1993). Fitting long-memory models by generalized linear regression. *Biometrika*, 80(4):817.

Beran, J. (1994). *Statistics for long-memory processes.* Chapman & Hall/CRC.

Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Stat.*, 17(4):1749–1766.

Doukhan, P., Oppenheim, G., and Taqqu, M. (2003). *Theory and applications of long-range dependence*. Birkhäuser.

Fox, R. and Taqqu, M. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Stat.*, 14(2):517–532.

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *J. Time Ser. Anal.*, 4(4):221–238.

Ghosal, S., Ghosh, J., and van der Vaart, A. (2000). Convergence rates of posterior distributions. *Ann. Stat.*, 28(2):500–531.

Giraitis, L. and Taqqu, M. (1999). Whittle estimator for finite-variance non-Gaussian time series with long memory. *Ann. Stat.*, 27(1):178–203.

Hurvich, C., Moulines, E., and Soulier, P. (2002). The FEXP estimator for potentially non-stationary linear time series. *Stochastic processes and their applications*, 97(2):307–340.

Kruijer, W. and Rousseau, J. (2010). Asymptotic behaviour of the posterior distribution of the long-memory parameter in gaussian stationary models. Technical report.

Lieberman, O. and Phillips, P. (2004). Error bounds and asymptotic expansions for Toeplitz product functionals of unbounded spectra. *J. Time Ser. Anal.*, 25:733–753.

Lieberman, O., Rosemarin, R., and Rousseau, J. (2009). Asymptotic theory for maximum likelihood estimation in stationary fractional gaussian processes, under short, long and intermediate memory. *preprint*.

Lieberman, O., Rousseau, J., and Zucker, D. (2003). Valid asymptotic expansions for the maximum likelihood estimator of the parameter of a stationary, Gaussian, strongly dependent process. *Ann. Stat.*, 31(2):586–612.

Liseo, B., Marinucci, D., and Petrella, L. (2001). Bayesian semiparametric inference on long-range dependence. *Biometrika*, 88(4):1089.

Liseo, B. and Rousseau, J. (2006). Sequential importance sampling algorithm for Bayesian nonparametric long range inference. *Atti della XLIII Riunione Scientifica della Societa Italiana di Statistica*, 2(4):43–46.

Mandelbrot, B. and Van Ness, J. (1968). Fractional brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437.

Mandelbrot, B. and Wallis, J. (1969). Some long run properties of geophysical records. *Water Resources research*, 5:321–340.

Moulines, E. and Soulier, P. (2003). Semiparametric spectral estimation for fractional processes. *Theory and applications of long-range dependence*, pages 251–301.

Robinson, P. (1991). Nonparametric function estimation for long memory time series. In *Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proceedings of the 5th International Symposium in Economic Theory and Econometrics*, pages 437–457.

Robinson, P. (1995). Gaussian semiparametric estimation of long range dependence. *Ann. Stat.*, pages 1630–1661.

Whittle, P. (1962). Gaussian estimation in stationary time series. *Bull. Inst. Internat. Statist*, 39:105–129.