The Annals of Statistics
1999, Vol. 27, No. 4, 1415–1439

BROADBAND LOG-PERIODOGRAM REGRESSION OF TIME SERIES WITH LONG-RANGE DEPENDENCE

By Eric Moulines and Philippe Soulier

Ecole Nationale Supérieure des Télécommunications and Université d'Evry–Val d'Essonne

This paper discusses the properties of an estimator of the memory parameter of a stationary long-memory time-series originally proposed by Robinson. As opposed to "narrow-band" estimators of the memory parameter (such as the Geweke and Porter-Hudak or the Gaussian semiparametric estimators) which use only the periodogram ordinates belonging to an interval which degenerates to zero as the sample size n increases, this estimator builds a model of the spectral density of the process over all the frequency range, hence the name, "broadband." This is achieved by estimating the "short-memory" component of the spectral density, $f^*(x) = |1 - e^{ix}|^{2d} f(x)$, where $d \in (-1/2, 1/2)$ is the memory parameter and $f(x)$ is the spectral density, by means of a truncated Fourier series estimator of $\log f^*$. Assuming Gaussianity and additional conditions on the regularity of f^* which seem mild, we obtain expressions for the asymptotic bias and variance of the long-memory parameter estimator as a function of the truncation order. Under additional assumptions, we show that this estimator is consistent and asymptotically normal. If the true spectral density is sufficiently smooth outside the origin, this broadband estimator outperforms existing semiparametric estimators, attaining an asymptotic mean-square error $O(\log(n)/n)$.

1. Introduction. Let $\{X_t\}$, $t = 0, \pm 1, \ldots,$ be a covariance stationary process with spectral density

$$f(x) = |1 - e^{ix}|^{-2d} f^*(x),$$

where f^* is a 2π-periodic positive continuous function.

The parameter d controls the behavior (and possibly, the singularity) of the spectral density in a neighborhood of the zero frequency, whereas f^* controls the short-memory behavior. When $0 < d < 1/2$, the process $\{X_t\}$ is said to be "long-range dependent." When $-1/2 < d < 0$, the spectral density at zero frequency is zero, but the process is still invertible; such a situation occurs, for example, when modeling the first differences of a process which is nonstationary but less so than a unit root process. The case $d = 0$ and f^* sufficiently smooth corresponds to the usual "weak dependence" situation. The importance of such models in virtually all fields of statistical applications has been demonstrated in numerous situations [see Robinson (1994) or Beran (1994)].

Received July 1997; revised April 1999.

AMS 1991 *subject classifications*. Primary 62G05; secondary 60F05.

Key words and phrases. Long-range dependence, log-periodogram regression, central limit theorems for dependent variables.

In the *parametric* approach, a finite-dimensional parametric model is assumed to hold for f^*. A key example is the ARFIMA(p, d, q) model [Granger and Joyeux (1980)] in which f^* is assumed to be rational. Another example is the FEXP(p) model [Robinson (1994), Beran (1993)], where $\log f^*$ is a finite order trigonometric polynomial. The parameters of f, including d, may then be estimated using, for example, Gaussian maximum likelihood [see Dahlhaus (1989)] or the Whittle likelihood [see Fox and Taqqu (1986), Giraitis and Surgailis (1990)]. These estimators have been shown to be asymptotically normal (under appropriate conditions) with the usual rate of convergence $\sqrt{n}$, where n is the sample size, provided that the *parametric model is correctly specified*. However, it has now been well documented that the estimator may be inconsistent if the model is *misspecified*. This drawback calls for *semiparametric* estimation of d. The short-memory component f^* is allowed to belong to a wider class of functions, which is specified either by the behavior at the zero-frequency or by a given regularity over the whole frequency range. The former provides motivation for *local methods* while the latter calls for *global methods*.

The local methods aim at constructing estimators of d that are consistent without any restrictions on f^* away from zero frequency (apart from integrability on $(-\pi, \pi]$ a consequence of covariance stationarity). Examples of such local estimators include the so-called GPH estimator introduced by Geweke and Porter-Hudak (1983) [see also Robinson (1995a) and Hurvich, Deo and Brodsky (1998)], or the Gaussian semiparametric estimate suggested by Künsch (1986) [see Robinson (1995b)]. Local methods are based on a subset of periodogram ordinates $[m_n, M_n]$ depending on the sample size n. The upper bound of this interval M_n tends to infinity more slowly than n, so that the proportion of the frequency band $(-\pi, +\pi]$ involved in the estimation degenerates (relatively slowly) to zero as n increases. For these reasons, local estimators are also referred in the literature as "narrow band." The choice of the trimming numbers m_n and M_n is crucial since it determines the bias, the variance and, in general, asymptotic distribution of the estimator [see Robinson (1995a, b) and Hurvich, Deo and Brodsky (1998)]. From a theoretical standpoint, the major drawback of narrow band methods is that, as shown by Giraitis, Robinson and Samarov (1997) and Hurvich, Deo and Brodsky (1998), the typical rate of convergence of the narrow band estimators of the LRD parameter d is limited to $n^{2/5}$, whatever the regularity of f^* is. The rate $n^{2/5}$ is achieved when $f^*(x) = f^*(0) + O(x^2)$ in a neighborhood of the origin. This rate can only be improved by the unnatural assumption that f^* has several vanishing derivatives at zero.

The global or broadband method presented in this paper aims to construct an estimator of f over all the frequency range. This is done by computing a nonparametric estimator of the short-memory component f^*. Perhaps surprisingly, this kind of approach [suggested and informally discussed by Robinson (1994) in his survey on LRD] has not yet been thoroughly investigated. There are many ways to construct a nonparametric estimator of f^*. In this contribution, we use a truncated Fourier expansion-type estimator of the log-spectral density. The method works as follows: under appropriate regularity conditions, the logarithm of the short-memory component, $l^* = \log f^*$ may be expanded on the cosine basis $l^* = \sum_{j=0}^{\infty} \theta_j h_j$, where $h_0(x) = 1/\sqrt{2\pi}$ and $h_j(x) = \cos(jx)/\sqrt{\pi}$. This expansion has, in general, an infinite number of nonzero coefficients. A nonparametric estimator of $f^*(x)$ is obtained by estimating only a finite number of Fourier coefficients. Heuristically, projection type estimators exploit the fact that the Fourier coefficients of a function go to zero at a rate linked to the regularity of the function. Hence, provided that l^* is smooth enough, a finite number of coefficients gives a reasonable approximation of the function. Truncated Fourier expansion has its root in nonparametric smoothing and curve estimation [see Hart (1997) and the references therein].

There are several ways to estimate these finitely many coefficients (e.g., approximate Gaussian likelihood). We study here the simplest to implement, namely the linear regression of the log-periodogram ordinates on log frequencies. The advantage of this method over local methods is that we take the regularity assumption in its full force; the rate of convergence of the estimator is related to the regularity of the spectral density outside the origin, that is, the regularity of f^*. In particular, if the coefficients of the Fourier expansion of $\log f^*(x)$ go to zero at an exponential rate, that is, $\theta_k = O(\rho^k)$, with $0 < \rho < 1$, then the rate $\sqrt{n/\log(n)}$ is achievable. This property is of interest since analyticity of $\log f^*$ holds for many models, including invertible ARIMA models, and thus the estimator presented here is nearly as efficient (in terms of rate of convergence) as an estimator based on a parametric model, but avoids the pitfalls (inconsistency) resulting from the use of a misspecified parametric model.

Our estimator is defined and our main results are presented in Section 2. The proofs are based on a theorem which describes the dependence structure of the log-periodogram ordinates over all the Fourier frequencies: a decomposition of the log-periodogram ordinates is introduced, which, together with a moment inequality, allows obtaining a central limit theorem and other related results. This result is of independent interest and is presented in Section 3.

2. Semiparametric estimation of the fractional differencing coefficient. Let m be a fixed integer, and for all n, let $n_m = 2m[n/2m]$ and $K_n = [n/2m]$. Define the discrete Fourier transform and the periodogram of $(X_1, \dots, X_n)$ as

$$\omega_n(x) = (2\pi n_m)^{-1/2} \sum_{t=1}^{n_m} X_t e^{itx} \qquad I_n(x) = |\omega_n(x)|^2.$$

For notational simplicity, we drop the last few data. However, this is asymptotically irrelevant since $n_m/n \to 1$. We evaluate the discrete Fourier transform and the periodogram at the Fourier frequencies $x_s = 2\pi s/n_m$, $1 \le s \le$

$n_m/2$. Since $\sum_{t=1}^{n_m} \exp(itx_i) = 0$ for $1 \le i \le n_m$, the mean of $\{X_t\}$ need not be estimated. Following the procedure proposed by Robinson (1995a), the frequency axis is divided into nonoverlapping segments of size m, and the periodogram is averaged over each segment. For $k = 1, \dots, K_n$, denote $J_n = \{m(k-1)+1, \dots, mk\}$ and

$$Y_{n,k} = \log\left(\exp(-\psi(m)) \sum_{i \in J_k} I_n(x_i)\right),$$

where ψ denotes the digamma function, that is, $\psi(m) = \Gamma'(m)/\Gamma(m)$, where Γ denotes the Gamma function [see Johnson and Kotz (1970)]. It is well known that if χ^2_{2m} denotes a random variable distributed as a central chi-square with $2m$ degrees of freedom, then $E(\log(\chi^2_{2m})) = \log(2) + \psi(m)$ and $\operatorname{var}(\log(\chi^2_{2m})) = \psi'(m)$. For instance, $\psi(1) = -\gamma$ (Euler's constant) and $\psi'(1) = \pi^2/6$.

Let $\hat{d}_{p,n}$ be the least square estimator of d when only $(p+1)$ Fourier coefficients are estimated,

$$\hat{d}_{p,n} = \arg \min_{\bar{d}, \bar{\theta}_0, \dots, \bar{\theta}_p} \sum_{k=1}^{K_n} \left(Y_{n,k} - \bar{d} g(y_k) - \sum_{j=0}^{p} \bar{\theta}_j h_j(y_k)\right)^2, \tag{2.1}$$

where $y_k = (2k-1)\pi/2K_n$ and $g(x) := -2\log|1 - e^{ix}|$. Note that there is no restriction on $\bar{d}$, and $\hat{d}_{p,n}$ might be outside $]-1/2, 1/2[$.

We now precisely state our assumptions.

Assumption 1. The process $\{X_t\}$ is Gaussian.

Assumption 2. The spectral density f of $\{X_t\}$ satisfies

$$f(x) = |1 - e^{ix}|^{-2d} f^*(x),$$

where $-1/2 < d < 1/2$, and f^* is positive and differentiable on $[-\pi, \pi] \setminus \{0\}$ and there exists a finite constant C^* such that

$$\forall\, x \in [-\pi, \pi] \setminus \{0\}, \qquad |f^{*\prime}(x)| \le C^*/|x|.$$

Assumption 3. There exists a real $\beta > 0$ and a finite real L such that

$$|\theta_0| + \sum_{j=1}^{\infty} j^{\beta} |\theta_j| \le L. \tag{2.2}$$

Mean square error and consistency of the estimator of d. Denote $\alpha_j = 2\sqrt{\pi}/j$, $j \ge 1$ and $\gamma_p = \sum_{j=p+1}^{\infty} \alpha_j^2$. Then $g(x) = \sum_{j=1}^{\infty} \alpha_j h_j(x)$.

PROPOSITION 1. *Under Assumptions* 1, 2 *and* 3,

$$E\left(\hat{d}_{p,n} - d\right)^2 = \frac{4\pi m \psi'(m)}{n\gamma_p} + \frac{\left(\sum_{j=p+1}^{\infty} \alpha_j \theta_j\right)^2}{\gamma_p^2} + \varepsilon_{p,n}, \tag{2.3}$$

where there exists a constant C such that for all $1 \le p \le K_n$,

$$|\varepsilon_{p,n}| \le Cpn^{-1}\left(p^{-\beta}\log^3(n) + p\log^6(n)n^{-1}\right).$$

As usual, the mean square error (MSE) is composed of a bias term and a variance term. Under Assumption 3, the bias term is bounded by

$$\gamma_p^{-1}\left|\sum_{j=p+1}^{\infty} \alpha_j \theta_j\right| \le \left(L/2\sqrt{\pi}\right)p^{-\beta}. \tag{2.4}$$

Thus, the bias term goes to zero as the truncation number p goes to infinity. It should be stressed that the bias term is not necessarily a monotone nonincreasing function of the truncation order p (contrary to the "traditional" theory of truncated Fourier series estimation).

The variance term is proportional to γ_p^{-1} which is monotonically increasing and asymptotically equivalent to $p/4\pi$. Thanks to (2.4), $\hat{d}_{p,n}$ is thus a consistent estimator of d as soon as $p := p_n$ is chosen in such a way that $p_n \to \infty$, $p_n/n \to 0$ and $(p_n^{-\beta}\log^3(n) + p_n\log^6(n)n^{-1}) \to 0$. Under Assumption 2, the best bound we can get for the MSE is of order $n^{-2\beta/(2\beta+1)}$, which is obtained by choosing p_n proportional to $n^{1/(2\beta+1)}$. In practice, this means that the rate of convergence of the estimator increases with the rate of decay to zero of the Fourier coefficients of the "short-memory" component (or equivalently, with the regularity of this function. If the coefficients θ_j decrease exponentially fast to zero, more precisely if for some $r > 0$,

$$\sum_{j=0}^{\infty} e^{rj}|\theta_j| \le L, \tag{2.5}$$

and if we take $p_n = [r^{-1}\log(n)]$, then

$$\lim_{n\to\infty} (n/\log(n))E\left(\hat{d}_{p_n,n} - d\right)^2 = m\psi'(m)/r.$$

Note that, when either $p_n \simeq n^{1/(2\beta+1)}$ or $p_n \simeq [r^{-1}\log(n)]$, the remainder term $\varepsilon_{p,n}$ is negligible compared with both the first and the second term on the right-hand side of (2.3) and thus does not affect the mean squared error.

On the other hand, the fastest rate of convergence which can be achieved for the GPH estimator is $O(n^{-4/5})$ [see Hurvich, Deo and Brodsky (1998)]. The reason for this relatively slow rate is that, although the short-memory component is extremely smooth, it does not in general satisfy the condition $f''^*(0) = 0$. Indeed, if the process is, say, ARFIMA(p, d, q), then the Fourier coefficients θ_j decays exponentially, but $f''^*(0)$ is most often nonzero so that the "broadband" method outperforms the "local" methods even with an optimal choice of the tuning constants.

Though the discussion above may serve as guideline for the choice of the truncation number, in any practical situation, p_n must be estimated from the data. Two approaches are usually considered. The classical one [already suggested in Robinson (1994)] is based on Mallow's C_L criterion [see Hart (1997)]. Properties of this criterion are developed in a companion paper [Moulines and Soulier (1998)]. The other approach is based on the so-called adaptive estimation theory, which has been investigated by Giraitis, Robinson and Samarov (1998) for the GPH estimator. This approach for the estimator developed in this contribution is currently under study.

Central limit theory. The mean square consistency results above can be extended by the following central limit theorem.

THEOREM 1. *Let $\{p_n\}$ be an increasing sequence of integers such that*

$$\lim_{n \to \infty} p_n^3 \log^2(n) n^{-2} = 0, \qquad \lim_{n \to \infty} n \log^2(n) p_n^{-1-2\beta} = 0. \tag{2.6}$$

If Assumption 2 *holds and if Assumption* 3 *holds with* $\beta > 1/4$, *then* $\sqrt{n/p_n}(\hat{d}_{p_n,n} - d)$ *is asymptotically normal,*

$$\sqrt{\frac{n}{p_n}}\left(\hat{d}_{p_n,n} - d\right) \to_d \mathcal{N}(0, m\psi'(m)),$$

$$\lim_{n \to \infty} n p_n^{-1} E\left(\hat{d}_n - d\right)^2 = m\psi'(m).$$

The proof is given in Appendix C.

REMARKS. (a) The condition $\lim_{n\to\infty} n \log^2(n) p_n^{-1-2\beta} = 0$ implies that the squared bias term is asymptotically negligible relative to the variance term. The condition $\lim_{n\to\infty} p_n^3 \log^2(n) n^{-2} = 0$ is a technicality involved in the proof of asymptotic normality.

(b) If $p_n = [n^\delta]$, the assumptions of Theorem 1 hold for $1/(1+2\beta) < \delta < 2/3$. The restriction $\beta > 1/4$ ensures that the assumptions on p_n are not exclusive. We do not know if this assumption is really necessary.

(c) Since $m\psi'(m)$ decreases to one as m tends to infinity, m should be chosen as large as possible. On the other hand, for finite sample size, too large m will increase the bias. In practice, a choice of $m = 4$ reduces the asymptotic variance from $\pi^2/6 \approx 1.6449$ to 1.1354.

3. Asymptotics of log-periodogram ordinates. As mentioned in the introduction, the local methods involve a trimming number for very low frequencies. The problem with the very low frequency periodogram ordinates is that their asymptotic behavior departs strongly from that under weak dependence (i.e., when f is regular over the whole interval $[-\pi, \pi]$), where they can be approximated by independent exponentially distributed variables with mean $f(0)/2$. It was first shown by Künsch (1986) and then exhaustively investigated by Hurvich and Beltrao (1993) and Robinson (1995a) that the periodogram ordinates computed at Fourier frequencies are asymptotically biased and correlated. Since one of our aims is to eliminate these trimming numbers, we need first to develop a theory for the periodogram ordinates which allows all the Fourier frequencies to be taken into account. We only consider the log-periodogram in this paper, but the next theorem is valid for any function H of the periodogram such that $E(H^2(Y)) < \infty$ and $E(H(Y)) = 0$ where the distribution of Y is central chi-square with $2m$ degrees of freedom. It is quite natural, when dealing with periodogram ordinates to use the method of moments to derive central limit theorems, since the dependence structure of the periodogram ordinates cannot be easily described (except in the case of Gaussian white noise). Thus what we need is a bound for moments of products of log-periodogram ordinates and we use this bound (Theorem 2) to derive a criterion for a central limit theorem (Theorem 3).

THEOREM 2. *Under Assumption* 2,

$$\forall\, 1 \le k \le K_n, \qquad Y_{n,k} = \log(f(y_k)) + \eta_{n,k} + r_{n,k}, \tag{3.1}$$

for each $1 \le k \le K_n$, $\eta_{n,k}$ *is distributed as* $\log(\chi^2_{2m})) - E(\log(\chi^2_{2m}))$, *where* χ^2_{2m} *is distributed as a central chi-square with* $2m$ *degrees of freedom.*

There exist constants $c_d < \infty$ *and* $C_d < \infty$, *such that, for all* $n \ge 2m$ *and for all* $1 \le k < j \le K_n$,

$$|r_{n,k}| \le c_d \log(1+k)/k,\ w.p.\ 1, \tag{3.2}$$

$$|\mathrm{cov}(\eta_{n,k}, \eta_{n,j})| \le C_d \log^2(j) k^{-2|d|} j^{2|d|-2}. \tag{3.3}$$

Let u *be a positive integer, and let* $(r_1, \dots, r_u)$ *be a* u*-tuple of positive integers among which exactly* s *are equal to* 1, *and let* $r = r_1 + \cdots + r_u$. *Then there exist a constant* $c_r < \infty$ *and an integer* K_r *depending only on* r *and not on* n *such that, for all* u*-tuple* $(k_1, \dots, k_u)$ *of pairwise distinct integers, if* $\min_{1 \le i \le n} k_i > K_r$, *then*

$$\left| E\left(\prod_{i=1}^{u} \eta_{n,k_i}^{r_i} \right) \right| \le c_r (\log(K_r)/K_r)^s. \tag{3.4}$$

REMARKS. (a) Theorem 1 means that the noise term in (3.1) can be approximated by $\eta_{n,k}$ with $E(\eta_{n,k}) = 0$ and $\mathrm{var}(\eta_{n,k}) = \psi'(m))$.

(b) Note that this theorem is not in contradiction with the result of Hurvich and Beltrao (1993), who proved that the log-periodogram ordinates for fixed k are asymptotically neither independent nor equidistributed. For any fixed k and j, the remainder term $r_{n,k}$ in Theorem 2 cannot be neglected and $\mathrm{cov}(\eta_{n,k}, \eta_{n,j})$ does not tend to zero as n tends to infinity. But we will show that for any bounded sequence of reals $h_{n,k}$, the following bounds hold (with $h_n^* = \max_{1 \le k \le K_n} |h_{n,k}|$),

$$\sum_{1 \le k \le K_n} |h_{n,k}|\,|r_{n,k}| = O\big(h_h^* \log^2(n)\big), \tag{3.5}$$

$$\sum_{1 \le k \le j \le K_n} |h_{n,k}|\,|h_{n,j}|\,|\mathrm{cov}(\eta_{n,k}, \eta_{n,j})| = O\big((h_n^*)^2 \log^3(n)\big). \tag{3.6}$$

(c). Recent related results consider tapered periodogram ordinates [Velasco (1999)]. Tapering reduces the covariance that appears in (3.3). For our purpose, this reduction would make the proof only marginally simpler. Using a cosine bell taper (operating on three adjacent frequencies) would remove the $\log^3(n)$ term in (3.6), which is not really significant in our computations.

The proof is given in Appendix A.

We now state a central limit theorem for weighted sums of variables $\{\eta_{n,k}\}$ which is used in the following sections to prove central limit theorems for log-periodogram ordinates. It is proved in Appendix B.

THEOREM 3. *Let $(v_n)_{n \ge 0}$ and $(w_n)_{n \ge 0}$ be two nondecreasing sequences of integers such that $0 < v_n \le w_n < K_n$. Let $(\beta_{n,k})_{v_n \le k \le w_n}$ be a triangular array of nonidentically vanishing real numbers. Define*

$$S_n = \sum_{k=v_n}^{w_n} \beta_{n,k} \eta_{n,k}, \qquad s_n^2 = \sum_{k=v_n}^{w_n} \beta_{n,k}^2,$$

$$a_n = \sum_{k=v_n}^{w_n} |\beta_{n,k}|, \qquad b_n = \max_{v_n \le k \le w_n} |\beta_{n,k}|.$$

Assume that:

(i) $\lim_{n\to\infty} b_n/s_n = 0$,
(ii) $\lim_{n\to\infty} a_n \log(n)/s_n v_n = 0$. *Then*,

$$s_n^{-1} S_n \to_d \mathcal{N}(0, \psi'(m)).$$

REMARKS. (a) For a triangular array of i.i.d. square integrable variables $(\eta'_{n,k})_{v_n \le k \le w_n}$, condition (i) implies the Lindeberg condition that ensures asymptotic normality of $s_n^{-1} \sum_{k=v}^{w_n} \beta_{n,k} \eta'_{n,k}$. Condition (ii) is a technicality needed to compensate for the dependence of the variables $\{\eta_{n,k}\}_{v_n \le k \le w_n}$.

(b) Conditions (i) and (ii) imply that v_n must tend to infinity.

(c) The proof of Theorem 3 is based on the method of moments. This is quite natural since the variables $\eta_{n,k}$ have moments of all order, and more profoundly, because the deeper dependence structure of the triangular array $\eta_{n,k}$ cannot be easily described.

(d) These two theorems are of independent interest and may be applied to any problem involving spectral estimation for long-range dependent Gaussian process such as nonlinear log-periodogram regression for parameter estimation or estimation of nonlinear functionals of the spectral density related to estimation or hypothesis testing.

4. Concluding remarks and open questions. In Giraitis, Robinson and Samarov (1997), it is shown that the optimal rate for memory parameter estimators in semiparametric long-memory models with degree of "local smoothness" β is $n^{-\beta/(2\beta+1)}$ and that the GPH estimator with maximum frequency $M_n = n^{2\beta/(2\beta+1)}$ is rate optimal. More precisely, these authors consider the following class of spectral densities:

$$\mathscr{F}(\gamma, C_0, K_0, \delta) = \{f: f(x) = C|x|^{-2d}[1 + \Delta(x)], 0 < C \le C_0, \\ -1/2 < d < 1/2 - \delta, |\Delta(x)| \le K_0|x|^{\gamma}, x \in [-\pi, \pi]\},$$

where C_0, K_0 and $\delta \in (0,1)$ are independent of γ, and they show that there exists a positive constant c such that

$$\liminf_n \inf_{d_n} \sup_{f \in \mathscr{F}(\gamma, C_0, K_0)} P_f\big(n^{\gamma/(2\gamma+1)}|\hat{d}_n - d| \ge c\big) > 0,$$

where inf is taken over all estimators of d and P_f stands for the distribution of a covariance stationary process with spectral density f.

As discussed in Giraitis, Robinson and Samarov (1997), the parameter γ is related to the local-to-zero smoothness σ of the short-memory component f^* of the spectral density, which could be defined as follows. For $0 < \sigma \le 1$, $f^*(x)$ has smoothness σ if $f^*(x)$ satisfies a Lipschitz condition of degree σ around 0. For $\sigma > 1$, $f^*(x)$ has smoothness σ if $f^*(x)$ is $s = [\sigma]$ times differentiable at zero, and the sth derivative satisfies a Lipshitz condition of degree $\sigma - s$ around $x = 0$. Then $\beta = \sigma$ for $\sigma \le 2$ [noting that $f^*(x)$ is an even function], whereas $\beta \le \sigma$ for $\sigma > 2$, with $\beta = \sigma$ if the first s derivatives of $f^*(x)$ at $x = 0$ are all zero. In general therefore, for $\sigma > 2$, we have $\beta = 2$ only. This is the case, for example, with fractionally integrated autoregressive moving average processes. Thus, even for a class of spectral densities with a very smooth short-range component, the rate of convergence of the local regression estimator is bounded by $2/5$.

We are currently generalizing the minimax result of Giraitis, Robinson and Samarov (1997) to the following classes of spectral densities. For $r > 1$, $\beta > 1$ and $0 < L < \infty$, define

$$\mathscr{S}(\beta, L, \delta) = \left\{f: f(x) = |1 - e^{ix}|^{-2d} \exp\left\{\sum_{j=0}^{\infty} \theta_j h_j(x)\right\}, \right. \\ \left. |d| \le 1/2 - \delta, |\theta_0| + \sum_{j=1}^{\infty} j^{\beta}|\theta_j| \le L\right\},$$

$$\mathscr{A}(r, L, \delta) = \left\{f: f(x) = |1 - e^{ix}|^{-2d} \exp\left\{\sum_{j=0}^{\infty} \theta_j h_j(x)\right\}, \right. \\ \left. |d| \le 1/2 - \delta, \sum_{j \ge 0} \theta_j^2 r^j \le L^2\right\}.$$

For these classes, the following lower bounds hold. There exists a constant $c' > 0$ such that

$$\liminf_n \inf_{\hat{d}_n} \sup_{f \in \mathscr{S}(\beta, \mathscr{L}, \delta)} P_f\left(n^{\beta/(2\beta+1)}|\hat{d}_n - d| \geq c'\right) > 0,$$

$$\liminf_n \inf_{\hat{d}_n} \sup_{f \in \mathscr{A}(r, L, \delta)} P_f\left(\sqrt{n/\log(n)}\,|\hat{d}_n - d| \geq c'\right) > 0.$$

This proves that the broadband estimator presented in this paper is rate optimal in the Sobolev and analytic classes. It also shows that the "narrow band" estimators are not rate-optimal in the Sobolev classes of regularity $\beta > 2$, and are of course never rate optimal in the analytic class. An adaptive version of this estimator (also under investigation) would be highly desirable and would even compete with parametric estimators, because of the problem of misspecification, as already mentioned.

APPENDIX A

Proof of Theorem 2. Recall that $w_n(x) = (2\pi n_m)^{-1/2}\sum_{t=1}^{n_m} X_t e^{itx}$ and $I_n(x) = |\omega_n(x)|^2$. The choice of frequencies x_s makes the correction for the unknown mean μ unnecessary because $\sum_{t=1}^{n_m} \exp(itx_s) = 0$ for $0 < s < n_m$. It is thus assumed in the sequel that $E(X_t) = 0$. Define

$$\xi_{n,s} = \left[E\left(\sum_{t=0}^{n_m-1} X_t \cos(tx_s)\right)^2\right]^{-1/2} \sum_{t=0}^{n_m-1} X_t \cos(tx_s),$$

$$\zeta_{n,s} = \left[E\left(\sum_{t=0}^{n_m-1} X_t \sin(tx_s)\right)^2\right]^{-1/2} \sum_{t=0}^{n_m-1} X_t \sin(tx_s),$$

$$a_{n,s} = \tfrac{1}{2}E(I_n(x_s)), \qquad b_{n,s} + ic_{n,s} = \tfrac{1}{2}E\left(\omega_n(x_s)^2\right), \qquad \gamma_{n,s} = b_{n,s}/a_{n,s}.$$

With these notations, we get

$$I_n(x_s)/a_{n,s} = \left(\xi_{n,s}^2 + \zeta_{n,s}^2\right)\left(1 + \gamma_{n,s}\left(\xi_{n,s}^2 - \zeta_{n,s}^2\right)/\left(\xi_{n,s}^2 + \zeta_{n,s}^2\right)\right),$$

$$\begin{aligned} Y_{n,k} &= \log\left(\sum_{i \in J_k} I_n(x_i)\right) - \psi(m) \\ &= \log\left(\sum_{i \in J_k} a_{n,i}\left(\xi_{n,i}^2 + \zeta_{n,i}^2\right)\right) - \psi(m) \\ &\quad + \log\left(1 + \frac{\sum_{i \in J_k} a_{n,i}\gamma_{n,i}\left(\xi_{n,i}^2 - \zeta_{n,i}^2\right)}{\sum_{i \in J_k} a_{n,i}\left(\xi_{n,i}^2 + \zeta_{n,i}^2\right)}\right) \end{aligned}$$

$$= \log\Bigl(\sum_{i \in J_k} (\xi_{n,i}^2 + \zeta_{n,i}^2)\Bigr) - \log(2) - \psi(m) + \log(f(y_k))$$
$$+ \log\left(1 + \frac{\sum_{i \in J_k}((2a_{n,i} - f(y_k))/f(y_k))(\xi_{n,i}^2 + \zeta_{n,i}^2)}{\sum_{i \in J_k}(\xi_{n,i}^2 + \zeta_{n,i}^2)}\right)$$
$$+ \log\left(1 + \frac{\sum_{i \in J_k} a_{n,i}\gamma_{n,i}(\xi_{n,i}^2 - \zeta_{n,i}^2)}{\sum_{i \in J_k} a_{n,i}(\xi_{n,i}^2 + \zeta_{n,i}^2)}\right).$$

Consider now the $2m$-dimensional Gaussian vector $\Xi_{n,k} = [\xi_{n,m(k-1)+1}, \zeta_{n,m(k-1)+1}, \dots, \xi_{n,mk}, \zeta_{n,mk}]$ and let $\Gamma_{n,k}$ be its covariance matrix. $\Gamma_{n,k}$ is invertible for all n, so we can define a standard $2m$-dimensional Gaussian vector $W_{n,k} = \Gamma_{n,k}^{-1/2}\Xi_{n,k}$. Define also $\Delta_{n,k} = \Gamma_{n,k} - I_{2m}$. We can now write

$$\log\Bigl(\sum_{i \in J_k} (\xi_{n,i}^2 + \zeta_{n,i}^2)\Bigr) = \log(W_{n,k}^T W_{n,k}) + \log\left(1 + \frac{W_{n,k}^T \Delta_{n,k} W_{n,k}}{W_{n,k}^T W_{n,k}}\right).$$

We finally have $Y_{n,k} = \eta_{n,k} + \log(f(y_k)) + r_{n,k}$ with

$$\eta_{n,k} = \log(W_{n,k}^T W_{n,k}) - \log(2) - \psi(m), \tag{A.1}$$

$$r_{n,k} = \log\left(1 + \frac{\sum_{i \in J_k} a_{n,i}\gamma_{n,i}(\xi_{n,i}^2 - \zeta_{n,i}^2)}{\sum_{i \in J_k} a_{n,i}(\xi_{n,i}^2 + \zeta_{n,i}^2)}\right) \tag{A.2}$$

$$+\log\left(1 + \frac{\sum_{i \in J_k}((2a_{n,i} - f(y_k))/f(y_k))(\xi_{n,i}^2 + \zeta_{n,i}^2)}{\sum_{i \in J_k}(\xi_{n,i}^2 + \zeta_{n,i}^2)}\right) \tag{A.3}$$

$$+\log\left(1 + \frac{W_{n,k}^T \Delta_{n,k} W_{n,k}}{W_{n,k}^T W_{n,k}}\right). \tag{A.4}$$

(A.1) shows that $\eta_{n,k}$ has the claimed distribution. From this point, the proof of Theorem 2 has two components. We will need the following:

1. Technical lemmas that prove that $r_{n,k}$ satisfies (3.2) and that describe the covariance structure of the jointly Gaussian vectors $(W_{n,k})_{1 \le k \le K_n}$;
2. A general inequality for moments of functions of jointly Gaussian vectors that will be used to derive (3.3) and (3.4) from the covariance structure of $(W_{n,k})_{1 \le k \le K_n}$.

Note that the remainder terms (A.2), (A.3) and (A.4) all are of the form $\log(1 + \psi_{n,k})$. In order to derive uniform bounds for these terms, we must prove the required bound for $\psi_{n,k}$ and that $\psi_{n,k}$ is uniformly bounded (wrt both n and k) away from -1. Thus (3.2) will follow straightforwardly from the following lemmas.

LEMMA 1. *There exists a finite constant c_1 such that for all $n \ge 2m$ and all $1 \le k \le K_n$,*

$$\forall\ i \in J_k,\ |\gamma_{n,i}| \le c_1 \log(1 + k)/k. \tag{A.5}$$

Moreover, $|\gamma_{n,i}|$ is uniformly bounded away from 1.

LEMMA 2. *There exists a finite constant c_3 such that for all $n \geq 2m$ and all $1 \leq k \leq K_n$,*

$$\forall\ i \in J_k, \left|\frac{2a_{n,i} - f(y_k)}{f(y_k)}\right| \leq c_3 \log(1+k)/k.$$

Moreover, $2a_{n,i}/f(y_k)$ is positive and uniformly bounded away from 0 for all $i \in J_k$.

LEMMA 3. *The smallest eigenvalue of $\Delta_{n,k}$ is uniformly bounded away from -1 and there exists a finite constant c_2 such that for all $n \geq 2m$ and all $1 \leq k \leq K_n$, the spectral radius $\delta_{n,k}$ satisfies the following bound:*

$$\delta_{n,k} \leq c_2 \log(1+k)/k.$$

The covariance structure of the jointly Gaussian vectors $(W_{n,k})_{1 \leq k \leq K_n}$ is exhibited in the following lemma.

LEMMA 4. *For $n \geq 2m$ and $1 \leq k < j \leq K_n$, let $\rho_{n,k,j}$ denote the supremum of the absolute value of the entries of $E(W_{n,k}W_{n,j}^T)$. Then there exists a finite constant c_4 such that for all $n \geq 2m$ and $1 \leq k < j \leq K_n$,*

$$\rho_{n,k,j} \leq c_4 \log(j) k^{-|d|} j^{|d|-1}.$$

Lemmas 1–4 are proved in Appendix D. We now state a theorem that gives the bounds for moments of products of functions of jointly Gaussian vectors that we need to prove (3.3) and (3.4). Let $\varphi(x_1, \dots, x_{2m}) = \log(\sum_{i=1}^m (x_{2i-1}^2 + x_{2i}^2)) - \log(2) - \psi(m)$, then $\eta_{n,k} = \varphi(W_{n,k})$. We define the Hermite rank of a function Ψ as in Arcones (1994). Let X be a ν-dimensional standard Gaussian vector. If $E(\Psi^2(X)) < \infty$, the Hermite rank of Ψ is the smallest integer τ for which there exists a polynomial $P(X)$ of degree τ such that $E(P(X)\Psi(X)) \neq 0$. Let u, q be positive integers. Let $X_1, \dots, X_u$ be jointly Gaussian and marginally q-dimensional standard Gaussian vectors. Denote $X_i = (X_{i,1}, \dots, X_{i,q})^T$ and $\rho^* = \max_{1 \leq i \neq i' \leq u, 1 \leq j, j' \leq q} |E(X_{i,j}X_{i',j'})|$. Consider real-valued functions ϕ_i with Hermite rank τ_i, $1 \leq i \leq u$.

THEOREM 4. *Let s and τ be positive integers. Assume that at least s among the functions ϕ_i have Hermite rank $\tau_i \geq \tau$. If $\rho^* \leq (1-\varepsilon)/(qu-1)$ for some $\varepsilon > 0$, then there exists a constant $c(\varepsilon, q, s, \tau) < \infty$ such that*

$$\left|E\left(\prod_{i=1}^u \phi_i(X_i)\right)\right| \leq c(\varepsilon, q, s, \tau) \prod_{i=1}^u \left(E\left(\phi_i^2(X_i)\right)\right)^{1/2} \rho^{*s\tau/2}. \tag{A.6}$$

This theorem can be proved by reformulating in a multidimensional setting Lemmas 3.1, 3.2 and 3.4 in Taqqu (1977). An alternative (and more straightforward) proof can be found in Soulier (1998). For $u = 2$, it is Lemma 1 of Arcones (1994). In our context, we can apply Theorem 4 with $q = 2m$ and $\tau = 2$. Indeed, φ is an even function and $E^0(\varphi(X_1, \dots, X_{2m})) = 0$, so the Hermite rank of φ is 2.

APPENDIX B

Proof of Theorem 3. The proof is based on the method of moments. Let $r \geq 2$ be an integer. The rth moment of $s_n^{-1}\mathscr{S}_n$ may be expanded as

$$s_n^{-r} E(S_n^r) = \sum_{u=1}^{r} \sum{}' \frac{r!}{r_1! \cdots r_u!} \frac{1}{u!} A_n(r_1, \dots, r_u) \tag{B.1}$$

with

$$A_n(r_1, \dots, r_u) = s_n^{-r} \sum{}'' \prod_{i=1}^{u} \beta_{n,k_i}^{r_i} \left(\prod_{i=1}^{u} \eta_{n,k_i}^{r_i} \right),$$

where Σ' extends over all u-tuples of positive integers $(r_1, \dots, r_u)$, such that $r_1 + \cdots + r_u = r$, and Σ'' extends over the u-tuples $(k_1, \dots, k_u)$ of pairwise distinct integers in the range $1 \leq k_i \leq K_n$. To prove that (B.1) converges to the rth moment of the normal distribution with mean zero and variance σ^2, it is enough to show that

$$\lim_{n \to \infty} A_n(r_1, \dots, r_u) = \begin{cases} \sigma^r, & \text{if } r_1 = \cdots = r_u = 2, \\ 0, & \text{otherwise.} \end{cases}$$

Let $(r_1, \dots, r_u)$ be a u-tuple of strictly positive integers satisfying $r_1 + \cdots + r_u = r$. Denote by s the number of elements in $(r_1, \dots, r_u)$ equal to 1. If $s = 0$ and $u < r/2$, Hölder inequality and assumption (i) yield

$$|A_n(r_1, \dots, r_u)| \leq \mu_r (b_n / s_n)^{r-2u} = o(1),$$

with $\mu_r = E(|\eta_{n,k}|^r)$. Consider now $s > 0$. In Theorem 2 (3.4) implies that

$$|A_n(r_1, \dots, r_u)| \leq \kappa_r s_n^{-r} \left(\prod_{\{i;\, r_i \geq 2\}} \sum_{k_i=1}^{n} |\beta_{n,k_i}|^{r_i} \right) a_n^s (\log(n)/v_n)^s.$$

By definition of b_n and s_n, we have for any i such that $r_i \geq 2$,

$$\sum_{k_i=1}^{n} |\beta_{n,k_i}|^{r_i} \leq b_n^{r_i - 2} s_n^2.$$

Since s is the number of indices i such that $r_i = 1$, we always have $2u - s \leq r$, and the number of indices $i \geq 2$ is $u - s$ and $\sum_{\{i;\, r_i \geq 2\}} r_i = r - s$, so assumptions (i) and (ii) imply

$$\prod_{\{i;\, r_i \geq 2\}} \sum_{k_i=1}^{n} |\beta_{n,k_i}|^{r_i} \leq (s_n^2)^{u-s} b_n^{r-s-2(u-s)},$$

$$|A_n(r_1, \dots, r_u)| \leq \kappa_r (b_n/s_n)^{r-2u+s} (a_n \log(n)/v_n s_n)^s = o(1).$$

Since it is impossible that both $s = 0$ and $u > r/2$, there only remains to consider the case $s = 0$ and $u = r/2$, which implies that r is even and $r_1 = \cdots = r_u = 2$. Denote $Z_{n,k} = \eta_{n,k}^2 - \psi'(m)$. $Z_{n,k}$ is a function of

$(U_{n,k}, V_{n,k})$ of Hermite rank at least 1 since $E(Z_{n,k}) = 0$. We can write

$$E\left(\prod_{i=1}^{r/2} \eta_{n,k_i}^2\right) = E\left(\prod_{i=1}^{r/2}\left(Z_{n,k_i} + \sigma^2\right)\right)$$
$$= \sigma^r + \sum_{I\subset\{1,\dots,r/2\},\,|I|\geq 2} \sigma^{r-2|I|} E\left(\prod_{i\in I} Z_{n,k_i}\right),$$

where $|I|$ denotes the number of elements of I. The summation starts at $|I| = 2$ since $E(Z_{n,k}) = 0$ for all $1 \leq k \leq K_n$. Applying Theorem 4, we get

$$\left|E\left(\prod_{i=1}^{r/2}\left(Z_{n,k_i} - \sigma^r\right)\right)\right| \leq \kappa_r \sum_{u=2}^{r/2} (\log(n)/v_n)^{u/2} = o(1),$$

$$A_n(2,\dots,2) = s_n^{-r}\sigma^r \sum{}'' \prod_{i=1}^{r/2} \beta_{n,k_i}^2 + o(1).$$

Finally, since $\sum'' \prod_{i=1}^{u} \beta_{n,k_i}^2$ extends over all $u = r/2$-tuples of pairwise distinct integers in the range $1 \leq k_i \leq K_n$, it differs from $(\sum_{k=1}^{K_n} \beta_{n,k}^2)^{r/2}$ by $\sum''' \beta_{n,k_i}^2 \cdots \beta_{n,k_u}^2$, the sum extending over the $(k_1,\dots,k_u)$ with at least one repeated index. Assumption (i) yields

$$s_n^{-r}\left|\sum{}'' \prod_{i=1}^{r/2} \beta_{n,k_i}^2 - \left(\sum_{k=1}^{K_n} \beta_{n,k}^2\right)^{r/2}\right| = O\left((b_n/s_n)^2\right) = o(1).$$

So $A_n(2,\dots,2) = \sigma^r + o(1)$. This concludes the proof of Theorem 3. □

APPENDIX C

Proof of Proposition 1 and Theorem 1. The functions h_j form an orthonormal basis of $L^2([-\pi,\pi])$ and they also have the following orthogonality property:

$$\forall\, 0 \leq j, j' \leq K_n, \qquad \frac{2\pi}{K_n}\sum_{k=1}^{K_n} h_j(y_k)h_{j'}(y_k) = \delta_{j,j'}. \tag{C.1}$$

Let $\tilde{\alpha}_j = 2\pi K_n^{-1}\sum_{k=1}^{K_n} g(y_k)h_j(y_k)$, $\tilde{g}_{p,n} = g - \sum_{j=0}^{p} \tilde{\alpha}_j h_j$ and $\tilde{\gamma}_{p,n} = 2\pi K_n^{-1}\sum_{k=1}^{K_n} \tilde{g}_{p,n}^2(y_k)$. Our estimator of d can then be written

$$\hat{d}_{p,n} = \frac{2\pi}{K_n\tilde{\gamma}_{n,p}} \sum_{k=1}^{K_n} \tilde{g}_{p,n}(y_k) Y_{n,k}.$$

Using the decomposition (3.1) and the orthogonality property (C.1) we get

$$\begin{aligned}\sqrt{K_n\tilde{\gamma}_{p,n}/2\pi}\left(\hat{d}_{p,n} - d\right) &= \sum_{k=1}^{K_n} \beta_{n,k}\eta_{n,k} + \sum_{k=1}^{K_n} \beta_{n,k} r_{n,k} \\ &\quad + \sum_{k=1}^{K_n} \beta_{n,k} l_p^*(y_k).\end{aligned} \tag{C.2}$$

with $\beta_{n,k} = \sqrt{2\pi K_n \tilde{\gamma}_{n,p}}\,\tilde{g}_{p,n}(y_k)$ and $l_p^* = \sum_{j=p+1}^{\infty} \theta_j h_j$. By construction, $\sum_{k=1}^{K_n} \beta_{n,k}^2 = \pi$.

We will need the following bounds that can be found in Lemma 1 in Moulines and Soulier (1998). Define $w_n = \min(K_n^2/(p_n^2 \log(n)), K_n)$ and let v_n be any sequence of integers such that $v_n < w_n$,

$$\max_{1 \le k \le K_n} |\beta_{n,k}| \le C(p/K_n)^{1/2} \log(n), \tag{C.3}$$

$$\forall\ w_n \le k \le K_n,\ |\beta_{n,k}| \le C(p_n/n)^{3/2} \log(n), \tag{C.4}$$

$$\sum_{k=v_n}^{w_n} |\beta_{n,k}| = O\big((n/p)^{1/2} \log(n)\big). \tag{C.5}$$

We must first evaluate the two bias terms. Equation (2.2) implies the following bounds:

$$\sum_{j=p+1}^{\infty} |\theta_j| = o(p^{-\beta}), \qquad \sum_{j=p+1}^{\infty} \theta_j^2 = o(p^{-2\beta}). \tag{C.6}$$

Applying (3.2), (C.6) and (C.3), we get that there exists a deterministic constant $C < \infty$ such that

$$\left| \sum_{k=1}^{2K_n} \beta_{n,k} r_{n,k} \right| \le C(p/K_n)^{1/2} \log^3(n), \text{ w.p.1},$$

$$\sum_{k=1}^{2K_n} \left| \beta_{n,k} l_p^*(y_k) \right| \le C K_n^{1/2} \log(n) p^{-1/2-\beta}.$$

Altogether, since $E(\eta_{n,k}) = 0$ for all k and n, we get

$$\begin{aligned} \sqrt{K_n \tilde{\gamma}_{p,n}} \left| E\big(\hat{d}_{p,n}\big) - d \right| &\le C\big((p/K_n)^{1/2} \log^3(n) + K_n^{1/2} \log(n) p^{-1/2-\beta}\big) \\ &= o(1), \end{aligned}$$

under the assumptions of Theorem 1.

We must now prove a central limit theorem for $\sum_{k=1}^{K_n} \beta_{n,k} \eta_{n,k}$. In view of Theorem 3, we define two sequences $v_n < w_n$ such that $E(\sum_{k=1}^{v_n} \beta_{n,k} \eta_{n,k})^2 = o(1)$, $E(\sum_{k=w_n}^{K_n} \beta_{n,k} \eta_{n,k})^2 = o(1)$ and such that Assumptions (i) and (ii) of Theorem 3 hold.

Let v_n be an increasing sequence of integers such that $(n \log^4(n)/p_n)^{1/2} = o(v_n)$ and $v_n = o(n/p_n \log^2(n))$. These assumptions are nonexclusive. Applying (C.3) and Theorem 2 and (C.4), we get

$$E\left(\sum_{k=1}^{v_n} \beta_{n,k} \eta_{n,k} \right)^2 = O\big(v_n \log^2(n) p/n\big) = o(1),$$

$$E\left(\sum_{k=w_n}^{K_n} \beta_{n,k} \eta_{n,k} \right)^2 = O\big(p_n^3 \log^2(n)/n^2\big) = o(1).$$

With the notation of Theorem 3, using (C.5), we get

$$a_n \log(n)/v_n = O\Big((n/p)^{1/2} \log^2(n)/v_n\Big) = o(1),$$

by assumption on v_n. Finally, Lemma 1 in Moulines and Soulier (1998) also yields that $\lim_{n\to\infty} p_n \tilde{\gamma}_{p_n,n} = 4\pi$, and Theorem 1 is proved. □

The proof of Proposition 1 uses the decomposition of the error (C.2), and the bounds (3.2), (3.3) and (C.3)–(C.6). Technical details can be found in Moulines and Soulier (1998), Appendix A1 and A2.

APPENDIX D

Technical lemmas. The proofs of these lemmas are based on the techniques presented in Robinson (1995a) These proofs are carried out in the frequency domain and rely upon truncation arguments to find uniform bounds of convolution integrals involving the spectral density and the Dirichlet kernel. We preface these proofs by recalling some useful results for $H_n(x)$, the (nonsymmetric) Dirichlet kernel,

$$H_n(x) = \sum_{t=1}^{n} e^{itx} = \exp\left(i(n+1)\frac{x}{2}\right)\frac{\sin(nx/2)}{\sin(x/2)}$$

It is easily seen that, for $-\pi \le x,\ y \le \pi$, we have

$$\begin{aligned} E(\omega_n(x)\omega_n(y)) &= \frac{1}{2\pi n}\int_{-\pi}^{\pi} f(t)H_n(x+t)H_n(y-t)\,dt, \\ &= \frac{1}{2\pi n}\int_0^{\pi}\left(f\left(\frac{y-x}{2}-t\right)+f\left(\frac{y-x}{2}+t\right)\right) \\ &\qquad \times H_n\left(\frac{x+y}{2}-t\right)H_n\left(\frac{x+y}{2}+t\right)dt. \end{aligned} \tag{D.1}$$

Let $0 < \beta < 2\pi$; the following results are repeatedly used in the sequel:

$$|H_n(x)| \le c_H(\beta)/x, \qquad 0 \le x \le \beta, \tag{D.2}$$

$$L_n(x) = \int_{-x}^{x} |H_n(t)|\,dt \le d_H(1+\log(1+nx)), \qquad 0 \le x \le \pi, \tag{D.3}$$

$$\int_u^{\pi} |H_n(t+z)H_n(t-z)|\,dt \le c_\Delta(z+u)^{-1}L_n(z+u), \qquad 0 \le z+u \le \delta < 2\pi, \tag{D.4}$$

$$\int_{-\pi}^{\pi} H_n(x+t)H_n(y-t)\,dt = 2\pi H_n(x+y). \tag{D.5}$$

LEMMA 5. *Let $z \ne w$ be two real numbers in $[0,\pi]$ such that $0 < z+w < \pi$. Let $g(w;t)$ be a function defined on $[0,\pi]\times[0,\pi]$; assume that*:

(i) *For all $0 \le w \le \pi$, the function $t \to g(w;t)$ is differentiable on $(0,\pi)\setminus\{w\}$ and admits left and right derivative at 0 and π;*

(ii) *There exist constants* $c_g < \infty$ *and* $c_{g'} < \infty$, *such that, for all* $0 \le t, w \le \pi$ *we have*

$$|g(w;t)| \le c_g\big(|t-w|^{-2d} + |t+w|^{-2d}\big), \tag{D.6}$$

$$|g_t'(w;t)| \le c_{g'}|t-w|^{-1-2d}, \qquad t \in (0,\pi)\setminus\{w\}, \tag{D.7}$$

where $g_t'(w;t)$ *is the differential of* $g(w;t)$ *w.r.t* t. *Then, there exists a constant* $c < \infty$ *such that*

$$\begin{aligned}&\int_0^{\pi} |(g(w;t) - g(w;z))\,\Delta_n(z;t)|\,dt\\ &\qquad \le c\frac{|z-w|^{-2d} + (z+w)^{-2d}}{(z+w)} L_n(z+w),\end{aligned} \tag{D.8}$$

where $\Delta_n(z;t) = H_n(t+z)H_n(t-z)$.

PROOF. We prove Lemma 5 only in the case $d > 0$. The case $d < 0$ is dealt with using the same techniques.

CASE 1 ($0 \le z < w < \pi$). Note that since $z + w < \pi$, $z < w$ implies that $z < \pi/2$. The proof consists in splitting the integral in several parts and showing that each part is uniformly bounded by

$$\varepsilon_n(w;z) = (z+w)^{-1}\big(|z-w|^{-2d} + (z+w)^{-2d}\big)L_n(w+z).$$

We decompose the integral on $(0,(w+z)/2), ((w+z)/2, \min((3w-z)/2,\pi), (\min((3w-z)/2,\pi),\pi)$. For the first part [the integral on $(0,(w+z)/2]$, we need to distinguish between close and distant z and w. Equations (D.2), (D3) and (D7) imply that, for all $0 < w < \pi$,

$$\begin{aligned}\sup_{0\le t\le (w+z)/2} |g(w;t) - g(w;z)|\,|H_n(t-z)| &\le c_H(\pi) \sup_{0\le t\le (w+z)/2} |g_t'(w;t)|\\ &\le c(w-z)^{-1-2d},\end{aligned}$$

$$\int_0^{(w+z)/2} |H_n(t+z)|\,dt \le \int_z^{(w+z)} |H_n(u)|\,du \le L_n(w+z).$$

These relations together imply

$$\begin{aligned}&\int_0^{(w+z)/2} |g(w;t) - g(w;z)|\,\Delta_n(z;t)\,dt\\ &\qquad \le c(w-z)^{-1-2d}\int_0^{(w+z)/2} |H_n(t+z)|\,dt \le c(w-z)^{-1-2d} L_n(w+z).\end{aligned}$$

When $2z \le w$ (distant z, w), we have $1/3(w+z) \le w - z$ and the factor $(w-z)^{-1-2d}$ in the previous expression is bounded by $3^{1+2d}(w+z)^{-1-2d}$. Thus, there exists $c < \infty$, such that for all $0 < w < \pi$ and all z such that

$2z \le w$, we have

$$\int_0^{(w+z)/2} |g(w;t) - g(w;z)| \, |\Delta_n(z;t)| \, dt \le c\varepsilon_n(w;z).$$

Equations (D.2), (D.3) and (D.6) imply

$$\sup_{0 \le t \le (w+z)/2} |H_n(t+z)| \le c_H(\pi) z^{-1} \text{ and } \int_0^{(w+z)/2} |H_n(t-z)| \, dt \le L_n(w),$$

$$\sup_{0 \le t \le (w+z)/2} (|g(w;t)) + |g(w;z)|) \le 5c_g\big((w-z)^{-2d} + (w+z)^{-2d}\big).$$

It follows that

$$\int_0^{(w+z)/2} |g(w;t) - g(w;z)| \, |\Delta_n(z;t)| \, dt$$
$$\le cz^{-1}\big((w-z)^{-2d} + (w+z)^{-2d}\big) L_n(w).$$

For $z \le w \le 2z$ (close z, w), we have $(w+z) \le 3z$ and $z^{-1} \le 3(w+z)^{-1}$. The previous relation implies that, for all $0 < w < \pi$ and $0 \le z \le w \le 2z$, there exists a constant $c < \infty$ such that

$$\int_0^{(w+z)/2} |g(w;t) - g(w;z)| \, |\Delta_n(z;t)| \, dt \le c\varepsilon_n(w;z).$$

The remaining terms are easier to work out because there is no need to distinguish between close and distant z, w. Note that, by (D.4) and (D.6),

$$|g(w;z)| \int_{(z+w)/2}^{\pi} |\Delta_n(z;t)| \, dt \le 2c_\Delta c_g \varepsilon_n(w;z).$$

On the interval $I(w;z) = ((w+z)/2, \min((3w-z)/2, \pi))$, we have $|g(w;t)| \le 2c_g|t-w|^{-2d}$ and we bound the integral as follows:

$$\int_{(w+z)/2}^{\min((3w-z)/2,\pi)} |g(w;t)| \, \Delta_n(t;z) \, dt$$
$$\le 2c_g \sup_{t \in I(w;z)} |\Delta_n(z;t)| \int_{(w+z)/2}^{(3w-z)/2} |t-w|^{-2d} \, dt.$$

The desired bound then follows from the relations

(D.9) $$\sup_{t \in I(w;z)} |\Delta_n(z;t)| \le 4c_H^2(3\pi/2)(w+z)^{-1}(w-z)^{-1},$$

(D.10) $$\int_{(w+z)/2}^{(3w-z)/2} |t-w|^{-2d} \, dt = 2^{2d}(1-2d)^{-1}(w-z)^{1-2d}.$$

When $(3w-z)/2 \ge \pi$ the last term is zero. When $(3w-z)/2 \le \pi$, (D.6) implies that $|g(w;t)| \le g(w;(3w-z)/2) \le 2c_g(w-z)^{-2d}$. By (D.4) we have

$$\int_{(3w-z)/2}^{\pi} |g(w;t)| \Delta_n(z;t) \, dt \le 2c_g(w-z)^{-2d} \int_{(3w-z)/2}^{\pi} \Delta_n(z;t) \, dt$$
$$\le c(w-z)^{-2d}(w+z)^{-1} L_n(w+z).$$

CASE 2 ($0 \le w < z < \delta$, with $\pi/2 < \delta < \pi$). The derivations are much as before, and are thus presented in an abbreviated form. By (D.2), we have $|H_n(2z)| \le c_H(2\delta)/(2z)$. Since $w + z < 2z$, this implies: $H_n(2z) \le c_H(2\delta)(w+z)^{-1}$ and

$$|g(w;z)H_n(2z)| \le c_H(2\delta)c_g(w+z)^{-1}\big((z-w)^{-2d} + (z+w)^{-2d}\big).$$

We split the integral in three parts, according to the following partition of the interval $(0,\pi)$: $(0,(w+z)/2)$, $((w+z)/2, \min((3z+w)/2,\pi)$, $(\min((3z+w)/2,\pi),\pi)$. On the interval $(0,(w+z)/2))$, we have

$$|H_n(t+z)| \le 2c_H(\pi+\delta)(w+z)^{-1}, \tag{D.11}$$

$$|H_n(t-z)| \le c_H(\pi)(z-t)^{-1}, \tag{D.12}$$

$$\int_0^{(w+z)/2}|H_n(t-z)|\,dt \le \int_{-z}^{z}|H_n(u)|\,du \le L_n(z). \tag{D.13}$$

Note that we have $|g(w;t)| \le 2c_g|w-t|^{-2d}$, which together with (D.11) implies

$$\begin{aligned}
&\int_0^{(w+z)/2}|g(w;t)||\Delta_n(t;z)\,dt\\
&\quad\le 4c_H(\pi+\delta)c_H(\pi)c_g(w+z)^{-1}\int_0^{(w+z)/2}|w-t|^{-2d}(z-t)^{-1}\,dt\\
&\quad\le 4c_H(\pi+\delta)c_H(\pi)c_g(w+z)^{-1}(z-w)^{-2d}\int_{1/2}^{\infty}|v-1|^{-2d}v^{-1}\,dv.
\end{aligned}$$

On the interval $I(w;z) = ((z+w)/2, \min((3z+w)/2,\pi))$, we have

$$\sup_{t\in I(w;z)}|g(w;t)| \le 4c_g\big((z-w)^{-2d} + (z+w)^{-2d}\big).$$

which, with (D.4) implies

$$\begin{aligned}
&\int_{(z+w)/2}^{\min((3z+w)/2,\pi)}|g(w;t)|\,|\Delta_n(t;z)|\,dt\\
&\quad\le \sup_{t\in I(w;z)}|g(w;t)|\int_{(z+w)/2}^{\pi}|\Delta_n(t;z)|\,dt \le c\varepsilon_n(w;z).
\end{aligned}$$

For $d > 0$, we have [it is assumed that $(3z+w)/2 \le \pi$; the conclusion is otherwise trivial]

$$\sup_{(3z+w)/2\le t\le\pi}|g(w;t)| \le |g(w;(3z+w)/2))| \le 4c_g(z-w)^{-2d}$$

and we conclude the proof by applying (D.4),

$$\begin{aligned}
&\int_{(3z+w)/2}^{\pi}|g(w;t)|\,\Delta_n(z;t)\,dt\\
&\quad\le c(z-w)^{-2d}\int_w^{\pi}\Delta_n(z;t)\,dt \le c(z-w)^{-2d}(z+w)^{-1}L_n(z+w).
\end{aligned}$$

CASE 3 ($0 \le w < z$ and $z > \delta > \pi/2$). We must consider this case separately because if z is allowed to be arbitrarily close to π, $|H_n(t+z)|$ cannot be uniformly bounded on $[0, \pi]$. Since $z + w < \pi$, we have $w < \pi - \delta$ and $z - w > 2\delta - \pi$. Let β be such that $\pi - \delta < \beta < \delta$. By (D.6) and (D.7), the functions $g(w;t)$ and $g'_t(w;t)$ are uniformly bounded on $[\beta, \pi]$, that is, there exists $c_\beta < \infty$, such that for all $w < \pi - \delta$ and all $\beta \le t \le \pi$, we have

$$|g(w;t)| \le c_\beta \quad \text{and} \quad |g'_t(w;t)| \le c_\beta.$$

Similarly, $H_n(z - t)$ and $H_n(t + z)$ are uniformly bounded on $[0, \beta]$,

$$|H_n(z-t)| \le c'_\beta, \qquad |H_n(z+t)| \le c'_\beta, \qquad 0 \le qt \le \beta,\ \delta \le z \le \pi.$$

Note that the choice of δ and β is arbitrary. For instance, we can take $\delta = 3\pi/4$ and $\beta = \pi/2$. We split the integral at β. On the interval $(0, \beta)$, we have

$$\int_0^\beta |g(w;t) - g(w;z)|\,|\Delta_n(z;t)|\,dt \le c\int_0^\beta |g(w;t)|\,dt + \beta|g(w;z)|.$$

The integral on the right-hand side is uniformly bounded by (D.6). On the interval (β, π), we have

$$\int_\beta^\pi |g(w;t) - g(w;z)|\,|\Delta_n(z;t)|\,dt$$
$$\le c_H(\pi) \sup_{\beta \le t \le \pi} |g'_t(w;t)| \int_\beta^\pi |H_n(t+z)|\,dt.$$

Since $\pi > z \ge \delta$ and $\pi > z - w > 2\delta - \pi$, these last bounds imply (D.8).

We now apply the preceding result to $g(w;t) = f(w-t) + f(w+t)$. It is easily checked that, under Assumption 2, $g(w;t)$ satisfies the assumptions (D.6) and (D.7) of the preceding lemma.

LEMMA 6. *There exist finite constants c_1 and c_2 such that, for all $n \ge 1$, and all $0 < s < n_m/2$,*

$$|2a_{n,s} - f(x_s)| \le c_1 \frac{\log(s)}{s} x_s^{-2d}, \qquad a_{n,s} \ge c_2 x_s^{-2d}.$$

PROOF. Note that, $E(I_n(x)) = E(\omega_n(x)\omega_n(-x))$. By (D.1), we have

$$E(I_n(x)) = \frac{1}{2\pi n}\int_0^\pi g(x;t)\,\Delta_n(0;t)\,dt.$$

Equations (D.8), (D.3) and (D.5) imply that, for $x > 0$ and $1 < k < n/2$,

$$|E(I_n(x)) - f(x)| \le cn^{-1}x^{-1-2d}L_n(x),$$
$$|E(I_n(x_k)) - f(x_k)| \le ck^{-1}\log(k)\,x_k^{-2d}.$$

Under Assumption 2, $c_1 x^{-2d} \le f(x) \le c_2 x^{-2d}$. Thus

$$\begin{aligned}
a_{n,k} &= (2\pi n)^{-1} \int_0^{\pi} \Delta_n(0;t) f(x_k - t)\, dt \\
&\ge (2\pi n)^{-1} \int_0^{2x_k/3} \Delta_n(0;t) f(x_k - t)\, dt \\
&\ge c x_k^{-2d} \left(1 - \int_{2x_k/3}^{\pi} \Delta_n(0;t)\, dt\right) \ge c x_k^{-2d}(1 - 3/4(2k+1)) \ge c/\big(4x_k^{2d}\big).
\end{aligned}$$

□

LEMMA 7. *There exists a constant $c < \infty$, such that, for all $n \ge 1$ and all $0 < s < n_m/2$,*

$$|b_{n,s}| \le c \frac{\log(s)}{s} x_s^{-2d}, \qquad |c_{n,s}| \le c \frac{\log(s)}{s} x_s^{-2d}.$$

PROOF. For $0 < x \le \pi$, (D.1) and (D.8) imply that

$$E\big(\omega_n(x)^2\big) = \frac{1}{\pi n} \int_0^{\pi} g(0;t)\, \Delta_n(x;t)\, dt,$$

$$\Big| E\big(\omega_n(x)^2\big) - n^{-1} f(x) H_n(2x) \Big| \le c n^{-1} x^{-1-2d} L_n(x).$$

The proof is concluded by noting that $|f(x_k) H_n(2x_k)| = 0$ for $0 < k < n/2$. □

LEMMA 8. *There exists some $0 < \delta < 1$, such that for all $0 < s \le n_m/2$,*

$$2a_{n,s}/f(x_s) > \delta, \qquad \gamma_{n,s} \in [-1+\delta, 1-\delta].$$

PROOF. Lemmas 6 and 7 imply that there exists an integer K_0 such that for all $k \ge K_0$ and all $n \ge 2k$, $|\gamma_{n,k}| \le 1/2$. Moreover, Hurvich and Beltrao (1993) have shown that, for any fixed $k > 0$,

$$\begin{aligned}
\lim_{n\to\infty} n^{2d} a_{n,k} &= \frac{f^*(0)}{\pi} \int_{-\infty}^{\infty} u^{-2d} \frac{\sin^2(k\pi - u)}{(k\pi - u)^2}\, du, \\
\lim_{n\to\infty} n^{2d} b_{n,k} &= -\frac{f^*(0)}{\pi} \int_{-\infty}^{\infty} u^{-2d} \frac{\sin(k\pi - u)\sin(k\pi + u)}{(k\pi - u)(k\pi + u)}\, du, \\
\lim_{n\to\infty} n^{2d} c_{n,k} &= 0.
\end{aligned}$$

This implies that, for each $k > 0$, and all $n \ge 0$, the sequence $\{\gamma_{n,k}\}_{n \ge k}$ has a finite limit γ_k^*; this limit lies in the interval $[-1+2\delta, +1-2\delta]$ for some $0 < \delta < 1$. So for $k < K_0$, there exists an integer N_0 such that for all $n \ge N_0$, $\gamma_{n,k} \in [-1+\delta, 1-\delta]$. Note that for all k and $n \ge 2k$, $|b_{n,k}| < a_{n,k}$ and $|c_{n,k}| < a_{n,k}$, so there exists $0 < \delta' < 1$ such that for all $k < K_0$ and $n < N_0$, $\gamma_{n,k} \in [-1+\delta', 1-\delta']$. This concludes the proof of Lemma 8. □

LEMMA 9. *Assume Assumption* 2. *Then there exists a constant* $\kappa_d < \infty$ *such that for all* $1 \le s < t \le n_m/2$,

$$|E(\xi_{n,s}\zeta_{n,t})| + |E(\zeta_{n,s}\zeta_{n,t})| + |E(\zeta_{n,s}\xi_{n,t})| + |E(\xi_{n,s}\xi_{n,t})| \\ \le \kappa_d \log(t) s^{-|d|} t^{|d|-1}.$$

PROOF. Let $0 < s < t < n/2$. The terms $E(\xi_{n,s}\zeta_{n,t})$ may be directly evaluated as functions of $E(\omega_n(\pm x_s)\omega_n(\pm x_t))$,

$$\begin{aligned} \operatorname{cov}(\xi_{n,s}, \xi_{n,t}) = {} & \big[E(\omega_n(x_s)\omega_n(x_t) + \omega_n(-x_s)\omega_n(x_t) \\ & + \omega_n(x_s)\omega_n(-x_t) + \omega_n(-x_s)\omega_n(-x_t))\big] \\ & \times \Big[4\sqrt{a_{n,s}a_{n,t}(1-\gamma_{n,s})(1-\gamma_{n,t})}\,\Big]^{-1} \\ \operatorname{cov}(\xi_{n,s}, \zeta_{n,t}) = {} & \big[E(\omega_n(x_s)\omega_n(x_t) - \omega_n(x_s)\omega_n(-x_t) \\ & + \omega_n(-x_s)\omega_n(x_t) - \omega_n(-x_s)\omega_n(-x_t))\big] \\ & \times \Big[4i\sqrt{a_{n,s}a_{n,t}(1-\gamma_{n,s})(1-\gamma_{n,t})}\,\Big]^{-1}. \end{aligned}$$

So it is clear that we need only give bounds for $E(\omega_n(x_s)\omega_n(\pm x_t))$, $0 < s < t < n/2$. By (D.1), we have

$$E(\omega_n(x_s)\omega_n(x_t)) = \frac{1}{2\pi n}\int_0^\pi g((x_t - x_s)/2; u)\Delta_n((x_t + x_s)/2; u)\, du,$$

$$E(\omega_n(x_s)\omega_n(-x_t)) = \frac{1}{2\pi n}\int_0^\pi g((x_t + x_s)/2; u)\Delta((x_t - x_s)/2; u)\, du.$$

Note also that $H_n(x_t + x_s) = H_n(x_t - x_s) = 0$. Applying Lemma 5,

$$|E(\omega_n(x_s)\omega_n(x_t))| \le cn^{-1}x_t^{-1}(x_s^{-2d} + x_t^{-2d})L_n(x_t),$$

$$|E(\omega_n(x_s)\omega_n(-x_t))| \le cn^{-1}x_t^{-1}(x_s^{-2d} + x_t^{-2d})L_n(x_t).$$

By Lemmas 6 and 8, we have

$$\sqrt{a_{n,s}a_{n,t}(1-\gamma_{n,s})(1-\gamma_{n,t})} \ge cx_s^{-d}x_t^{-d}.$$

Dividing these bounds, we obtain

$$|\operatorname{cov}(\xi_{n,s}, \xi_{n,t})| \le cs^{-|d|}t^{|d|-1}\log(t).$$

The other terms are treated similarly, and this concludes the proof of Lemma 9. □

We can now prove Lemmas 1–4. Lemmas 6 and 7 yield the following bounds:

(D.14) $$\max_{n \ge 2s}|\gamma_{n,s}| \le c\log(s)/s,$$

(D.15) $$\max_{n \ge 2s}|2a_{n,s}/f(x_s) - 1| \le c\log(s)/s.$$

Lemma 1 is now a consequence of (D.14) and Lemma 8: for all $1 \le k \le K_n$ and all $i \in J_k$,

$$|\gamma_{n,i}| = |b_{n,k}|/a_{n,k} \le c \log(1+i)/i \le c \log(1+mk)/(1+m(k-1))$$
$$\le c' \log(1+k)/k.$$

Lemma 2 is also a straightforward consequence of Lemmas 6 and 8. Indeed, we can write

$$\frac{2a_{n,i} - f(y_k)}{f(y_k)} = \frac{2a_{n,i} - f(x_i)}{f(x_i)} \frac{f(x_i)}{f(y_k)} + \frac{f(x_i) - f(y_k)}{f(y_k)},$$

and Lemma 6 and assumption (A1) imply that for all $1 \le k \le K_n$ and all $i \in J_k$,

$$\left| \frac{2a_{n,i} - f(y_k)}{f(y_k)} \right| \le c \frac{\log(1+k)}{k}.$$

Finally, since $2a_{n,i}/f(y_k) = (2a_{n,i}/f(x_i))(f(x_i)/f(y_k))$, Assumption 2 and Lemma 8 imply that $2a_{n,i}/f(y_k)$ is uniformly bounded away from 0. This concludes the proof of Lemma 2. □

Define $\rho_{n,s} = \mathrm{cov}(\xi_{n,s}, \zeta_{n,s})$. It is easily seen that

$$\rho_{n,s} = c_{n,s}/a_{n,s}\sqrt{1-\gamma_{n,s}^2},$$

so the following inequality is a straightforward consequence of Lemmas 6, 7 and 8. There exists a finite constant c such that for all $1 \le k \le K_n$ and all $\in J_k$,

$$|\rho_{n,i}| \le c \log(1+k)/k. \tag{D.16}$$

Moreover, $\rho_{n,i}$ is uniformly bounded away from 1.

We can now prove Lemma 3. Equation (D.16) and Lemma 9 imply that the off-diagonal terms of $\Delta_{n,k}$ are all of order $\log(1+k)/k$, uniformly wrt n; thus the spectral radius $\delta_{n,k}$ of $\Delta_{n,k}$ is of order $\log(1+k)/k$, uniformly wrt n. We must now prove that the smallest eigenvalue of $\Delta_{n,k}$ is uniformly bounded away from -1. First note that there exists an integer K_0 such that for all $k \ge K_0$ and all $n \ge 2mk$, $\delta_{n,k} \le 1/2$. Consider now $k \le K_0$. Since $\Delta_{n,k} = \Gamma_{n,k} - I_{2m}$, it is equivalent to prove that the smallest eigenvalue of $\Gamma_{n,k}$ is bounded away from 0. Define $(\phi, \psi)_d = \int_{-\infty}^{\infty} |u|^{-2d} \phi(u)\psi(u)\, du$. Let

$$s_k^+(u) = \frac{\sin(k\pi + u)}{k\pi + u} + \frac{\sin(k\pi - u)}{k\pi - u},$$
$$s_k^-(u) = \frac{\sin(k\pi + u)}{k\pi + u} - \frac{\sin(k\pi - u)}{k\pi - u}$$

and let $\alpha_k(u) = s_k^+(u)/(s_k^+, s_k^+)_d^{1/2}$ and $\beta_k(u) = s_k^-(u)/(s_k^-, s_k^-)_d^{1/2}$. Note that $(\alpha_k, \alpha_k)_d = (\beta_k, \beta_k)_d = 1$ and $(\alpha_k, \beta_k)_d = 0$. As in the proof of Lemma 8, it is easily seen that $\lim_{n\to\infty} \Gamma_{n,k} = \Gamma_k^*$, where Γ_k^* is the Gram–Schmidt matrix related to the functions α_t and β_t, $m(k-1)+1 \le t \le mk$ and the scalar product $(\cdot,\cdot)_d$. Since the functions α_t and β_t, $m(k-1)+1 \le t \le mk$ are linearly independent, Γ_k^* is invertible, which implies that its smallest eigenvalue is positive. Thus there exists a real $\varepsilon > 0$ and an integer n_0 such that for all $n \ge n_0$ and all $k \le K_0$, the smallest eigenvalue of $\Gamma_{n,k}$ is greater than ε. For each $n < n_0$ and $k \le k_0$, $\Gamma_{n,k}$ is invertible so the smallest eigenvalue of each $\Gamma_{n,k}$, $n < n_0, k \le k_0$ is greater than some $\varepsilon' > 0$. This finally concludes the proof of Lemma 3. □

To conclude the proof of Lemma 4, note that for $1 \le k < j \le K_n$, $E(W_{n,k}W_{n,j}^T) = \Gamma_{n,k}^{-1/2}(E(\Xi_{n,k}\Xi_{n,j}^T)\Gamma_{n,j}^{-1/2}$. Lemma 9 implies that the entries of $E(\Xi_{n,k}\Xi_{n,j}^T)$ are of order $\log(j)k^{-|d|}j^{|d|-1}$ and since the smallest eigenvalue of $\Gamma_{n,k}$ is uniformly bounded away from zero, the spectral radius of $\Gamma_{n,k}^{-1/2}$ is uniformly bounded, so that the entries of $E(W_{n,k}W_{n,j}^T)$ are also of order $\log(j)k^{-|d|}j^{|d|-1}$. This concludes the proof of Lemma 4. □

Acknowledgments. The authors thank the Editor and an Associate Editor for their helpful comments that improved the presentation. The authors also thank P. M. Robinson for helpful discussions and for providing reference Robinson (1995a), which was a major source of inspiration. Finally, the authors are indebted to C. Hurvich for providing them with a technical report on a work in progress on a broad band estimator related to our construction and for pointing out that the assumption on f^* used in an earlier version of the paper could be weakened to Assumption 2.

REFERENCES

ARCONES, M. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.* **22** 2243–2274.

BERAN, J. (1993). Fitting long memory models by generalized regression. *Biometrika* **80** 817–822.

BERAN, J. (1994). Statistics for long memory processes. Chapman and Hall, New York.

DAHLHAUS, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17** 1749–1766.

FOX, R. and TAQQU, M. S. (1986). Large samples properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14** 517–532.

GEWEKE, J. and PORTER-HUDAK, S. (1983). The estimation and application of long memory time series models. *J. Time Anal.* **4** 221–238.

GIRAITIS, L., ROBINSON, P. and SAMAROV, A. (1997). Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long range dependence. *J. Time Series Anal.* **18** 49–61.

GIRAITIS, L. and SURGAILIS, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittles's estimate. *Probab. Theory Related Fields* **86** 87–104.

GRANGER, C. W. J. and JOYEUX, R. (1980). An introduction to long memory time series and fractional differencing. *J. Time Series Anal.* **1** 15–30.

HART, J. D. (1997). Nonparametric smoothing and lack-of-fit tests. *Springer Series in Statistics*. New York, NY: Springer.

HURVICH, C. and BELTRAO, K. (1993). Asymptotics of the low-frequency ordinates of the periodogram of a long-memory time series. *J. Time Series Anal.* **14** 455–472.

HURVICH, C., DEO, R. and BRODSKY, J. (1998). The mean squared error of Geweke and Porter-Hudak's estimator of a long-memory time-series. *J. Time Ser. Anal.* **19** 19–46.

JOHNSON, N. L. and KOTZ, S. (1970). *Continuous Univariate Distributions* **1** Wiley, New York.

KÜNSCH, H. R. (1986). Discrimination between monotonic trends and long-range dependence. *J. Appl. Probab.* **23** 1025–1030.

MOULINES, E. and SOULIER, P. (1998). Data-driven order selection for long range dependent time series. *J. Time Ser. Anal.* To appear.

ROBINSON, P. M. (1994). Time series with strong dependence. In *Advances in Econometrics. Proceedings of the Sixth World Congress* **1** 47–95. Cambridge Univ. Press.

ROBINSON, P. M. (1995a). Log-periodogram regression of time series with long range dependence. *Annal. Statist.* **23** 1048–1072.

ROBINSON, P. M. (1995b). Gaussian semi-parametric estimation of long range dependence. *Annal. Statist.* **23** 1630–1661.

SOULIER, P. (1998). Some new bounds and a central limit theorem for nonstationary sequences of Gaussian vectors. Université d'Evry-Val d'Essonne. Submitted to *Statistics and Probability Letters*.

TAQQU, M. S. (1977). Law of the iterated logarithm for sums of nonlinear functions of Gaussian variables that exhibit long range dependence. *Z. Wahrsch. Verw. Gebiete.* **40** 203–238.

VELASCO, C. (1999). Gaussian semiparametric estimation of non-stationary time series. *J. Time Ser. Anal.* **20** 87–127.

ECOLE NATIONALE SUPÉRIEURE
DES TELECOMMUNICATIONS
CNRS-URA 820, 75634 PARIS CEDEX
FRANCE

DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITY D'EVRY
VAL D'ESSONNE
91025 EVRY CEDEX
FRANCE
E-MAIL: soulier@lami.univ-evry.fr