The Annals of Statistics
1995, Vol. 23, No. 5, 1630–1661

GAUSSIAN SEMIPARAMETRIC ESTIMATION OF LONG RANGE DEPENDENCE[1]

BY P. M. ROBINSON

London School of Economics

Assuming the model $f(\lambda) \sim G\lambda^{1-2H}$, as $\lambda \to 0+$, for the spectral density of a covariance stationary process, we consider an estimate of $H \in (0,1)$ which maximizes an approximate form of frequency domain Gaussian likelihood, where discrete averaging is carried out over a neighbourhood of zero frequency which degenerates slowly to zero as sample size tends to infinity. The estimate has several advantages. It is shown to be consistent under mild conditions. Under conditions which are not greatly stronger, it is shown to be asymptotically normal and more efficient than previous estimates. Gaussianity is nowhere assumed in the asymptotic theory, the limiting normal distribution is of very simple form, involving a variance which is not dependent on unknown parameters, and the theory covers simultaneously the cases $f(\lambda) \to \infty$, $f(\lambda) \to 0$ and $f(\lambda) \to C \in (0, \infty)$, as $\lambda \to 0$. Monte Carlo evidence on finite-sample performance is reported, along with an application to a historical series of minimum levels of the River Nile.

1. Introduction. Several estimates are now available for the slope of the logged spectral density of a long range dependent covariance stationary scalar process x_t, $t = 0, \pm 1, \dots,$ which is observed at times $t = 1, \dots, n$. Denote by γ_j the lag-j autocovariance of x_t and by $f(\lambda)$ the spectral density of x_t, such that $\gamma_j = E(x_0 - Ex_0)(x_j - Ex_0) = \int_{-\pi}^{\pi} \cos(j\lambda) f(\lambda)\, d\lambda$. It is assumed that

$$f(\lambda) \sim G\lambda^{1-2H} \quad \text{as } \lambda \to 0+, \tag{1.1}$$

for $G \in (0, \infty)$ and $H \in (0,1)$. The parameter H is sometimes called the self-similarity parameter. In case $H = \frac{1}{2}$, $f(\lambda)$ tends to a finite positive constant at zero frequency, whereas if $H \in (\frac{1}{2}, 1)$ it tends to infinity and if $H \in (0, \frac{1}{2})$ it tends to zero. A recently published survey of relevant literature up to about 1990 is Robinson (1994a).

Finite-parameter models for $f(\lambda)$ over the full frequency band $(-\pi, \pi]$ which are consistent with the property (1.1), have been considered (such as fractional autoregressive moving average models), as have methods of estimating an unknown H and additional parameters. In particular, the asymptotic distributional properties of Gaussian parameter estimates have been derived by Fox and Taqqu (1986), Dahlhaus (1989) and Giraitis and Surgailis

Received October 1993; revised December 1994.
[1] Research funded by the Economic and Social Research Council (ESRC) reference number R000233609.
AMS 1991 *subject classifications*. Primary 62M15; secondary 62G05, 60G18.
Key words and phrases. Long range dependence, Gaussian estimation.

(1990), in case $H \in (\frac{1}{2}, 1)$ and under regularity conditions. These properties are highly desirable ones: $n^{1/2}$-consistency and asymptotic normality and, when x_t is actually Gaussian, asymptotic efficiency. However, these properties also depend on correct specification of $f(\lambda)$ over $(-\pi, \pi]$, and in the event of any misspecification, estimates will, in general, be inconsistent. In particular, misspecification of $f(\lambda)$ at high frequencies can lead to an inconsistent estimate of the parameter H, which characterizes low frequency behaviour. The spectral density of a fractional autoregressive moving average model has the rather strange mathematical property of being infinitely differentiable at all frequencies in $[-\pi, \pi]$ except at zero frequency, where (for $H > \frac{1}{2}$) it is discontinuous and unbounded.

To overcome such criticisms, "semiparametric" estimates of H in (1.1) have been proposed which can be justified as consistent in the absence of full parametric or other global assumptions on $f(\lambda)$. One of these, due to Robinson (1994b), is consistent when $H \in (\frac{1}{2}, 1)$ under mild conditions which do not include Gaussianity or even loose restrictions on $f(\lambda)$ away from zero frequency (apart from integrability on $(-\pi, \pi]$, a consequence of covariance stationarity). The limiting distributional properties of this estimate are rather complicated, however [see Lobato and Robinson (1994)]. One alternative semiparametric estimate of H is due to Geweke and Porter-Hudak (1983). Robinson (1995) recently established desirable asymptotic properties of modified and more efficient versions of this estimate, but under the restrictive condition of Gaussianity, which may not be necessary but seems difficult to replace by weaker but reasonably comprehensible conditions because the estimate involves nonlinear transformation of the periodogram.

The present paper discusses a semiparametric "Gaussian" estimate of H which, unlike the estimates of Geweke and Porter-Hudak (1983) and Robinson (1995), is not defined in closed form, but which dominates these estimates in several respects. It is asymptotically more efficient. Much weaker assumptions than Gaussianity are imposed. [Giraitis and Surgailis (1990) likewise avoided Gaussianity in their study of a parametric Gaussian estimate.] Trimming out low frequency components [introduced in Robinson's (1995) estimate following a suggestion of Künsch (1986)] is avoided, as is the additional user-chosen number needed for Robinson's (1994b) estimate. The current estimate was suggested by Künsch (1987), but he did not establish or conjecture any statistical properties. The estimate is described in the following section. In Section 3 it is shown to be weakly consistent under only slight additional conditions to (1.1), with only finite second moments of x_t assumed. In Section 4 the estimate is shown to be asymptotically normally distributed under somewhat stronger conditions, including a fourth moment condition on x_t. Three technical lemmas are proved in Section 5. The proof of consistency and asymptotic normality involves some relatively unusual features. The objective function optimized by the (implicitly defined) estimate behaves in a nonuniform way over the parameter space. The mean-value theorem argument used in the central limit theorem requires not just a preliminary consistency proof for the estimate, but a rate of consistency. Throughout, we aim for minimal conditions on the spectrum away from zero frequency and on the bandwidth number involved in the semiparametric estimation, and our proofs cover simultaneously cases $0 < H < \frac{1}{2}$, $H = \frac{1}{2}$ and $1 > H > \frac{1}{2}$, whether or not H is known a priori to be consistent with an infinite or a zero spectrum at zero frequency. Section 6 contains Monte Carlo evidence of finite-sample performance, and comparison with an estimate similar to those of Geweke and Porter-Hudak (1983), Künsch (1986) and Robinson (1995), as well as an application to the series of Nile minima.

It has been stated in the literature that fractional autoregressive moving average processes which satisfy (1.1) also satisfy the time domain property

$$\gamma_j \sim g j^{2H-2} \quad \text{as } j \to \infty, \tag{1.2}$$

where $g < 0$ for $0 < H < \frac{1}{2}$ and $g > 0$ for $\frac{1}{2} < H < 1$. Condition (1.2) (extended to include a slowly varying function) has been stressed by Taqqu (1975), for example. With $g = 2G\Gamma(2-2H)\cos \pi H$, it is known that for $0 < H < \frac{1}{2}$, (1.2) implies (1.1) [Yong (1974), page 90], whereas for $\frac{1}{2} < H < 1$, (1.1) and (1.2) are equivalent if the γ_j are quasimonotonically convergent to zero, that is, $\gamma_j \to 0$ as $j \to \infty$ and for some $C < \infty$, $\gamma_{j+1} \le \gamma_j(1 + C/j)$ for all large enough j [Yong (1974), page 75]. In general, (1.1) does not imply (1.2).

2. Semiparametric Gaussian estimate. Define the discrete Fourier transform and periodogram of x_t:

$$w(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^{n} x_t e^{it\lambda}, \qquad I(\lambda) = |w(\lambda)|^2, \tag{2.1}$$

where correction for an unknown mean of x_t is unnecessary because the statistics (2.1) will be computed only at frequencies $\lambda_j = 2\pi j/n$ for $j = 1, \dots, m$, where m is an integer less than $\frac{1}{2}n$. Because our estimate is not defined in closed form, it is convenient to denote by G_0 and H_0 the true parameter values, and by G and H any admissible values.

Consider the objective function [see Künsch (1987)]

$$Q(G, H) = \frac{1}{m} \sum_{j=1}^{m} \left\{ \log G\lambda_j^{1-2H} + \frac{\lambda_j^{2H-1}}{G} I_j \right\}, \tag{2.2}$$

writing $I_j = I(\lambda_j)$. Define the closed interval of admissible estimates of H_0, $\Theta = [\Delta_1, \Delta_2]$, where Δ_1 and Δ_2 are numbers picked such that $0 < \Delta_1 < \Delta_2 < 1$. We can choose Δ_1 and Δ_2 arbitrarily close to 0 and 1, respectively, or we can choose them to reflect weak prior knowledge on H_0, for example, $\Delta_1 = \frac{1}{2}$ if we are confident that $f(\lambda) \not\to 0$ as $\lambda \to 0$. Clearly the estimate

$$(\hat{G}, \hat{H}) = \arg \min_{\substack{0 < G < \infty \\ H \in \Theta}} Q(G, H)$$

exists. We can also write

$$\hat{H} = \arg \min_{H \in \Theta} R(H),$$

where

$$R(H) = \log \hat{G}(H) - (2H-1)\frac{1}{m}\sum_{1}^{m}\log \lambda_j, \qquad \hat{G}(H) = \frac{1}{m}\sum_{1}^{m}\lambda_j^{2H-1}I_j.$$

Were we to take $m = \frac{1}{2}n$ for n even, or $m = \frac{1}{2}(n-1)$ for n odd, $\hat{H}$ would be a Gaussian estimate of H_0 in the parametric model $f(\lambda) = G_0|\lambda|^{1-2H_0}$, $\lambda \in (-\pi, \pi]$, and the asymptotic theory of $\hat{H}$ would be effectively covered by that of Fox and Taqqu (1986) and Giraitis and Surgailis (1990) in case $H_0 \in (\frac{1}{2}, 1)$ and $\Delta_1 > \frac{1}{2}$, although these authors considered an integral form in place of the summation form (2.2). In our asymptotic theory, m 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. The derivation of such asymptotic theory is quite different from that of Fox and Taqqu (1986), Dahlhaus (1989) and Giraitis and Surgailis (1990) for parametric Gaussian estimates. Our $\hat{H}$ is only $m^{1/2}$-consistent and is thus much less efficient than these estimates when they happen to be based on a correct parametric model. Like Giraitis and Surgailis (1990), we avoid the Gaussianity assumption of Fox and Taqqu (1986) and Dahlhaus (1989), and unlike any of these authors, we allow H_0 to be $\frac{1}{2}$ or less than $\frac{1}{2}$, as well as greater than $\frac{1}{2}$, also permitting the set of admissible values of $\hat{H}$ to include ones in $(0, \frac{1}{2}]$ as well as $(\frac{1}{2}, 1)$. An integral form of (2.2) could be considered, but we prefer the discrete form (2.2), partly for its computational convenience and partly because, in case of an unknown mean of x_t, it avoids dependence on the sample mean; for parametric Gaussian estimates, Cheung and Diebold (1994) have found that the slow convergence of the sample mean when $H_0 > \frac{1}{2}$ can produce inferior finite sample behaviour in estimates of H_0 besed on the integral form of objective function.

3. Consistency of estimates.

The following assumptions are introduced.

ASSUMPTION A1. As $\lambda \to 0+$,

$$f(\lambda) \sim G_0 \lambda^{1-2H_0},$$

where $G_0 \in (0, \infty)$ and $H_0 \in [\Delta_1, \Delta_2]$.

ASSUMPTION A2. In a neighbourhood $(0, \delta)$ of the origin, $f(\lambda)$ is differentiable and

$$\frac{d}{d\lambda}\log f(\lambda) = O(\lambda^{-1}) \quad \text{as } \lambda \to 0+.$$

ASSUMPTION A3. We have

$$x_t - Ex_0 = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}, \qquad \sum_{j=0}^{\infty} \alpha_j^2 < \infty, \tag{3.1}$$

where

$$E(\varepsilon_t | F_{t-1}) = 0, \qquad E(\varepsilon_t^2 | F_{t-1}) = 1, \quad \text{a.s., } t = 0, \pm 1, \dots,$$

in which F_t is the σ-field of events generated by ε_s, $s \le t$, and there exists a random variable ε such that $E\varepsilon^2 < \infty$ and for all $\eta > 0$ and some $K > 0$, $P(|\varepsilon_t| > \eta) \le KP(|\varepsilon| > \eta)$.

ASSUMPTION A4. As $n \to \infty$,

$$\frac{1}{m} + \frac{m}{n} \to 0.$$

Assumption A1 is just the basic model (1.1), with H_0 contained in the interval of admissible estimates $[\Delta_1, \Delta_2]$, while Assumption A2 is a regularity condition, analogous to ones imposed in the parametric case by Fox and Taqqu (1986) and Giraitis and Surgailis (1990). Assumption A3 takes the innovations in the Wold representation (3.1) to be a square-integrable martingale difference sequence that need not be strictly stationary, but satisfies a mild homogeneity restriction. Assumption A4 is minimal, because m must tend to infinity for consistency, while it must do so more slowly than n because A1 specifies $f(\lambda)$ only as $\lambda \to 0+$. The estimate of Robinson (1994b) was shown to be consistent under conditions which are in some respects weaker, but the latter estimate may be less attractive for practical use than $\hat{H}$ because it is sensitive not only to m, but to an additional user-chosen constant.

THEOREM 1. *Let Assumptions* A1–A4 *hold. Then*

$$\hat{H} \to_p H_0 \quad \textit{as } n \to \infty.$$

PROOF. For $\frac{1}{2} > \delta > 0$ let $N_\delta = \{H: |H - H_0| < \delta\}$ and $\overline{N}_\delta = (-\infty, \infty) - N_\delta$. Then for $S(H) = R(H) - R(H_0)$,

$$\begin{aligned} P\big(|\hat{H} - H_0| \ge \delta\big) &= P\big(\hat{H} \in \overline{N}_\delta \cap \Theta\big) \\ &= P\Big(\inf_{\overline{N}_\delta \cap \Theta} R(H) \le \inf_{N_\delta \cap \Theta} R(H)\Big) \le P\Big(\inf_{\overline{N}_\delta \cap \Theta} S(H) \le 0\Big), \end{aligned}$$

because $H_0 \in N_\delta \cap \Theta$. Now define $\Theta_1 = \{H: \Delta \le H \le \Delta_2)$, where $\Delta = \Delta_1$ when $H_0 < \frac{1}{2} + \Delta_1$ and $H_0 \ge \Delta > H_0 - \frac{1}{2}$ otherwise. When $H_0 \ge \frac{1}{2} + \Delta_1$, define $\Theta_2 = \{H: \Delta_1 \le H < \Delta\}$, and otherwise take Θ_2 to be empty. It follows that

$$P\big(|\hat{H} - H_0| \ge \delta\big) \le P\Big(\inf_{\overline{N}_\delta \cap \Theta_1} S(H) \le 0\Big) + P\Big(\inf_{\Theta_2} S(H) \le 0\Big). \tag{3.2}$$

It is necessary to treat Θ_1 and Θ_2 separately because of the nonuniform behaviour of $R(H)$ around $H = H_0 - \frac{1}{2}$. The set Θ_2 is empty when, for example, $\Delta_1 \ge \frac{1}{2}$ (so knowledge that $H_0 \ge \frac{1}{2}$ is used in fixing Θ) or when, on the other hand, $H_0 < \frac{1}{2}$. The first probability on the right of (3.2) is bounded by

$$P\Big(\sup_{\Theta_1}|T(H)| \ge \inf_{\overline{N}_\delta \cap \Theta_1} U(H)\Big) \tag{3.3}$$

with the definitions

$$T(H) = \log\left\{\frac{\hat{G}(H)}{G_0}\right\} - \log\left\{\frac{\hat{G}(H_0)}{G(H)}\right\} - \log\left\{\frac{1}{m}\sum_1^m j^{2(H-H_0)} \Big/ \frac{m^{2(H-H_0)}}{2(H-H_0)+1}\right\}$$
$$+ 2(H-H_0)\left\{\frac{1}{m}\sum_1^m \log j - (\log m - 1)\right\},$$
$$U(H) = 2(H-H_0) - \log\{2(H-H_0)+1\},$$
$$G(H) = G_0 \frac{1}{m}\sum_1^m \lambda_j^{2(H-H_0)},$$

so that $S(H) = U(H) - T(H)$. Here, $U(H)$ is the deterministic part of $S(H)$ obtained by replacing I_j by $G\lambda_j^{1-2H}$ and Riemann sums by integrals, $T(H)$ being the remainder. Because the function $x - \log(1+x)$ achieves a unique relative and absolute minimum on $(-1, \infty)$ at $x = 0$, and because $\log(1+x) \le x - \frac{1}{6}x^2$ and $-\log(1-x) \ge x + \frac{1}{2}x^2$ for $0 < x < 1$, it follows that

$$\inf_{\overline{N}_\delta \cap \Theta_1} U(H) \ge \min(2\delta - \log(1+2\delta), -2\delta - \log(1-2\delta)) > \tfrac{1}{2}\delta^2. \tag{3.4}$$

On the other hand, from the inequality $|\log(1+x)| \le 2|x|$ for $|x| \le \frac{1}{2}$ we deduce that, for any nonnegative random variable Y,

$$P(|\log Y| \le \varepsilon) \le 2P(|Y-1| \le 2\varepsilon) \quad \text{when } \varepsilon \le 1, \tag{3.5}$$

and thus that $\sup_{\Theta_1}|T(H)| \to_p 0$ if

$$\sup_{\Theta_1}\left|\frac{\hat{G}(H) - G(H)}{G(H)}\right| \tag{3.6}$$

is $o_p(1)$, while

$$\sup_{\Theta_1}\left|\frac{2(H-H_0)+1}{m}\sum_1^m \left(\frac{j}{m}\right)^{2(H-H_0)} - 1\right| \tag{3.7}$$

and

$$\left|\frac{1}{m}\sum_1^m \log j - (\log m - 1)\right| \tag{3.8}$$

are both $o(1)$. Now (3.7) is $O(m^{-2(\Delta - H_0)-1}) \to 0$ as $m \to \infty$ from Lemma 1 (see Section 5), while (3.8) is $O(\log m/m)$ from Lemma 2 (see Section 5). Write

$$\frac{\hat{G}(H) - G(H)}{G(H)} = \frac{A(H)}{B(H)}, \tag{3.9}$$

where

$$A(H) = \frac{2(H-H_0)+1}{m}\sum_1^m \left(\frac{j}{m}\right)^{2(H-H_0)}\left(\frac{I_j}{g_j} - 1\right),$$
$$B(H) = \frac{2(H-H_0)+1}{m}\sum_1^m \left(\frac{j}{m}\right)^{2(H-H_0)}, \tag{3.10}$$

for $g_j = G_0\lambda_j^{1-2H_0}$. Now

$$\inf_{\Theta_1} B(H) \geq 1 - \sup_{\Theta_1}\left|\frac{2(H-H_0)+1}{m}\sum_1^m\left(\frac{j}{m}\right)^{2(H-H_0)} - 1\right| \geq \frac{1}{2}, \tag{3.11}$$

for all sufficiently large m, by Lemma 1. By summation by parts

$$\begin{aligned}|A(H)| \leq \frac{3}{m}\Bigg|\sum_{r=1}^{m-1}\Bigg\{\left(\frac{r}{m}\right)^{2(H-H_0)} &- \left(\frac{r+1}{m}\right)^{2(H-H_0)}\Bigg\}\sum_{j=1}^{r}\left(\frac{I_j}{g_j}-1\right)\Bigg| \\ &+ \frac{3}{m}\left|\sum_1^m\left(\frac{I_j}{g_j}-1\right)\right|.\end{aligned} \tag{3.12}$$

Because $|(1+1/r)^{2(H-H_0)} - 1| \leq 2/r$ on Θ_1 when $r > 0$, the first term on the right of (3.12) has supremum on Θ_1 bounded by

$$\begin{aligned}6\sup_{\Theta_1}\sum_1^{m-1}\left(\frac{r}{m}\right)^{2(H-H_0)+1}\frac{1}{r^2}\left|\sum_1^r\left(\frac{I_j}{g_j}-1\right)\right| \\ \leq 6\sum_1^{m-1}\left(\frac{r}{m}\right)^{2(\Delta-H_0)+1}\frac{1}{r^2}\left|\sum_1^r\left(\frac{I_j}{g_j}-1\right)\right|,\end{aligned} \tag{3.13}$$

the inequality being due to $0 < 2(\Delta - H_0) + 1 \leq 2(H - H_0) + 1$ on Θ_1. Now

$$\frac{I_j}{g_j} - 1 = \left(1 - \frac{g_j}{f_j}\right)\frac{I_j}{g_j} + \frac{1}{f_j}\left(I_j - |\alpha_j|^2 I_{\varepsilon j}\right) + (2\pi I_{\varepsilon j} - 1), \tag{3.14}$$

where $I_{\varepsilon j} = I_\varepsilon(\lambda_j) = |w_\varepsilon(\lambda_j)|^2$, $w_\varepsilon(\lambda) = (2\pi n)^{-1/2}\sum_{t=1}^n \varepsilon_t e^{it\lambda}$, $f_j = f(\lambda_j)$ and $\alpha_j = \alpha(\lambda_j) = \sum_{l=0}^\infty \alpha_l e^{il\lambda_j}$. For any $\eta > 0$, Assumptions A1 and A4 imply that n can be chosen such that

$$\left|1 - \frac{g_j}{f_j}\right| \leq \eta, \qquad j = 1, \dots, m. \tag{3.15}$$

Let C be a generic finite positive constant. Assumptions A1, A2 and A4 imply that for n sufficiently large,

$$E\left|\frac{I_j}{g_j}\right| \leq C, \qquad j = 1, \dots, m, \tag{3.16}$$

in view of the proof of Theorem 2 of Robinson (1995). Thus

$$\begin{aligned}E\left\{\sum_1^{m-1}\left(\frac{r}{m}\right)^{2(\Delta-H_0)+1}\frac{1}{r^2}\left|\sum_1^r\left(1-\frac{g_j}{f_j}\right)\frac{I_j}{g_j}\right|\right\} &\leq \frac{C\eta}{m}\sum_1^m\left(\frac{r}{m}\right)^{2(\Delta-H_0)} \\ &\leq \frac{C\eta}{2(\Delta - H_0)+1},\end{aligned}$$

because $\sum_1^m r^a \le m^{a+1}/(a+1)$ for $a > -1$. Next,

$$\begin{aligned} E\left|I_j - |\alpha_j|^2 I_{\varepsilon j}\right| &\le E\left\{|w_j - \alpha_j w_{\varepsilon j}|\,|w_j + \alpha_j w_{\varepsilon j}|\right\} \\ &\le \left(EI_j - \alpha_j E w_{\varepsilon j}\overline{w}_j - \overline{\alpha}_j E\overline{w}_{\varepsilon j} w_j + |\alpha_j|^2 EI_{\varepsilon j}\right)^{1/2} \\ &\quad \times \left(EI_j + \alpha_j E w_{\varepsilon j}\overline{w}_j + \overline{\alpha}_j E\overline{w}_{\varepsilon j} w_j + |\alpha_j|^2 EI_{\varepsilon j}\right)^{1/2}, \end{aligned} \tag{3.17}$$

where the first inequality is due to $\left|\,|a|^2 - |b|^2\right| = |\mathrm{Re}\{(a-b)(\overline{a}+\overline{b})\}| \le |(a-b)(\overline{a}+\overline{b})| \le |a-b|\,|a+b|$ and the second to the Schwarz inequality. It follows from the proof of Theorem 2 of Robinson (1995) that, as $n \to \infty$,

$$EI_j = f_j\left(1 + O\left(\frac{\log(j+1)}{j}\right)\right),$$

$$Ew_j\overline{w}_{\varepsilon j} = \frac{\alpha_j}{2\pi} + O\left(\frac{\log(j+1)}{j}\lambda_j^{1/2-H_0}\right),$$

$$EI_{\varepsilon j} = \frac{1}{2\pi} + O\left(\frac{\log(j+1)}{j}\right)$$

uniformly in $j = 1,\dots,m$. Thus (3.17) is $O(f_j(\log(j+1)/j)^{1/2})$ as $n \to \infty$ and

$$\begin{aligned} &E\left\{\sum_1^{m-1}\left(\frac{r}{m}\right)^{2(\Delta-H_0)+1}\frac{1}{r^2}\left|\sum_1^r \frac{1}{f_j}\left(I_j - |\alpha_j|^2 I_{\varepsilon j}\right)\right|\right\} \\ &\le C\sum_1^m\left(\frac{r}{m}\right)^{2(\Delta-H_0)+1}\frac{1}{r^2}\sum_1^r\left(\frac{\log j}{j}\right)^{1/2} \\ &\le \frac{C}{m^{2(\Delta-H_0)+1}}\sum_1^m r^{2(\Delta-H_0)-1/2}(\log r)^{1/2} \\ &= O\left(m^{2(H_0-\Delta)-1} + (\log m)^{3/2} m^{-1/2}\right) \\ &= o(1) \quad \text{as } n \to \infty, \end{aligned}$$

where the penultimate equality results from separate consideration of the cases $2(\Delta - H_0) - \frac{1}{2} < -1$ and $2(\Delta - H_0) - \frac{1}{2} \ge -1$. To deal with the final contribution to (3.14), write

$$2\pi I_{\varepsilon j} - 1 = \frac{1}{n}\sum_1^n(\varepsilon_t^2 - 1) + \frac{1}{n}\sum\sum_{s\ne t}\cos\{(s-t)\lambda_j\}\varepsilon_s\varepsilon_t,$$

so that

$$\begin{aligned} &\sum_1^{m-1}\left(\frac{r}{m}\right)^{2(\Delta-H_0)+1}\frac{1}{r^2}\left|\sum_1^r(2\pi I_{\varepsilon j} - 1)\right| \\ &\le \left|\frac{1}{n}\sum_1^n(\varepsilon_t^2 - 1)\right|\sum_1^m\left(\frac{r}{m}\right)^{2(\Delta-H_0)+1}\frac{1}{r} \\ &\quad + \sum_1^m\left(\frac{r}{m}\right)^{2(\Delta-H_0)+1}\frac{1}{r^2 n}\left|\sum\sum_{s\ne t}\varepsilon_s\varepsilon_t\sum_1^r\cos\{(s-t)\lambda_j\}\right|. \end{aligned} \tag{3.18}$$

Under Assumption A3,

$$\frac{1}{n}\sum_{1}^{n}(\varepsilon_t^2 - 1) \to_p 0 \tag{3.19}$$

from Theorem 1 of Heyde and Seneta (1972), so the first term on the right of (3.18) is $o_p(1)$. Assumption A3 also implies that

$$\begin{aligned} &E\Biggl(\sum\sum_{s\neq t}\varepsilon_s\varepsilon_t\sum_{1}^{r}\cos\{(s-t)\lambda_j\}\Biggr)^2 \\ &\quad = 2\sum\sum_{s\neq t}\Biggl(\sum_{1}^{r}\cos\{(s-t)\lambda_j\}\Biggr)^2 \\ &\quad = 2\sum_{1}^{r}\sum_{1}^{r}\Biggl[\sum_{1}^{n}\sum_{1}^{n}\cos\{(s-t)\lambda_j\}\cos\{(s-t)\lambda_k\} - n\Biggr] \\ &\quad = rn^2 - 2r^2 n \end{aligned} \tag{3.20}$$

for $1 \le r < \frac{1}{2}n$, so that the second term on the right of (3.18) is

$$O\Biggl(\sum_{1}^{m}\biggl(\frac{r}{m}\biggr)^{2(\Delta - H_0)+1} r^{-3/2}\Biggr) = O\bigl(m^{2(H_0-\Delta)-1} + (\log m)m^{-1/2}\bigr) = o(1)$$

as $n \to \infty$. Because η is arbitrary, we have proved that (3.13) is $o_p(1)$. Using the same techniques, as $n \to \infty$,

$$\left|\frac{1}{m}\sum_{1}^{m}\biggl(\frac{I_j}{g_j} - 1\biggr)\right| = O_p\Biggl(\eta + \frac{1}{m}\sum_{1}^{m}\biggl(\frac{\log j}{j}\biggr)^{1/2}\Biggr) + o_p(1) = o_p(1),$$

so the second term on the right of (3.12) is $o_p(1)$. Thus as $n \to \infty$, $\sup_{\Theta_1}|A(H)| \to_p 0$ and, with (3.9) and (3.11), $\sup_{\Theta_1}|\hat{G}(H)/G(H) - 1| \to_p 0$. In view of (3.4) it follows that (3.3) $\to 0$ as $n \to \infty$.

In case $H_0 < \frac{1}{2} + \Delta_1$, the proof of the theorem is completed. In case $H_0 \ge \frac{1}{2} + \Delta_1$, the second probability on the right of (3.2) can be nonzero. Put $p = p_m = \exp(m^{-1}\sum_1^m \log j)$ and $S(H) = \log\{\hat{D}(H)/\hat{D}(H_0)\}$, where

$$\hat{D}(H) = \frac{1}{m}\sum_{1}^{m}\biggl(\frac{j}{p}\biggr)^{2(H-H_0)} j^{2H_0-1}I_j.$$

Because $1 \le p \le m$ and $\inf_{\Theta_2}(j/p)^{2(H-H_0)} \ge (j/p)^{2(\Delta-H_0)}$ for $1 \le j \le p$, while $\inf_{\Theta_2}(j/p)^{2(H-H_0)} \ge (j/p)^{2(\Delta_1-H_0)}$ for $p < j \le m$, it follows that

$$\inf_{\Theta_2}\hat{D}(H) \ge \frac{1}{m}\sum_{1}^{m}a_j j^{2H_0-1}I_j,$$

where

$$a_j = \begin{cases} \biggl(\dfrac{j}{p}\biggr)^{2(\Delta-H_0)}, & 1 \le j \le p, \\ \biggl(\dfrac{j}{p}\biggr)^{2(\Delta_1-H_0)}, & p < j \le m. \end{cases}$$

Thus,

$$P\Big(\inf_{\Theta_2} S(H) \le 0\Big) \le P\bigg(\frac{1}{m}\sum_{1}^{m}(a_j - 1) j^{2H_0-1} I_j \le 0\bigg). \tag{3.21}$$

As $m \to \infty$, $p \sim \exp(\log m - 1) = m/e$ and

$$\begin{aligned}\sum_{1\le j\le p} a_j &\sim p^{2(H_0-\Delta)}\int_0^p x^{2(\Delta-H_0)}\,dx \\ &= \frac{p}{2(\Delta - H_0) + 1} \sim \frac{m/e}{2(\Delta - H_0) + 1}.\end{aligned} \tag{3.22}$$

It follows that

$$\frac{1}{m}\sum_{1}^{m}(a_j - 1) \ge \frac{1}{m}\sum_{1\le j\le p} a_j - 1 \sim \frac{1}{e(2(\Delta - H_0) + 1)} - 1 \quad \text{as } m \to \infty.$$

Choose $\Delta < H_0 - \frac{1}{2} + \frac{1}{4}e$, which we may do with no loss of generality. Then for all sufficiently large m, $m^{-1}\Sigma_1^m(a_j - 1) \ge 1$ and thus (3.21) is bounded by

$$P\bigg(\bigg|\frac{1}{m}\sum_{1}^{m}(a_j - 1)\bigg(\frac{I_j}{g_j} - 1\bigg)\bigg| \ge 1\bigg).$$

Now apply (3.14) again and first note from (3.15) and (3.16) that

$$\bigg|\frac{1}{m}\sum_{1}^{m}(a_j - 1)\bigg(1 - \frac{g_j}{f_j}\bigg)\frac{I_j}{g_j}\bigg| = O_p\bigg(\frac{\eta}{m}\sum_{1}^{m}(a_j + 1)\bigg) = O_p(\eta),$$

using also (3.22) and

$$\sum_{p<j\le m} a_j \sim p^{2(H_0-\Delta_1)}\int_p^m x^{2(\Delta_1-H_0)}\,dx = O(m). \tag{3.23}$$

Next we have from (3.17) and Theorem 2 of Robinson (1995) that

$$\begin{aligned}\bigg|\frac{1}{m}\sum_{1}^{m}\frac{(a_j - 1)}{f_j}\Big(I_j - |\alpha_j|^2 I_{\varepsilon j}\Big)\bigg| &= O_p\bigg(\frac{1}{m}\sum_{1}^{m}(a_j + 1)\bigg(\frac{\log(j+1)}{j}\bigg)^{1/2}\bigg) \\ &= O_p\bigg(\frac{\log m}{m}\bigg(\sum_{1}^{m}a_j^2 + m\bigg)^{1/2}\bigg)\end{aligned} \tag{3.24}$$

by the Cauchy inequality. Because

$$\begin{aligned}\sum_{1}^{m}a_j^2 &\le p^{4(H_0-\Delta)}\sum_{1}^{p} j^{4(\Delta-H_0)} + p^{4(H_0-\Delta_1)}\sum_{p}^{m} j^{4(\Delta_1-H_0)} \\ &= O\big(m^{4(H_0-\Delta)} + m\log m\big),\end{aligned}$$

it follows that (3.24) is $O_p((\log m)^{3/2}(m^{2(H_0-\Delta)-1} + m^{-1/2})) = o_p(1)$. Finally,

$$\begin{aligned}\frac{1}{m}\sum_{1}^{m}(a_j - 1)(2\pi I_{\varepsilon j} - 1) &= \frac{1}{n}\sum_{1}^{n}(\varepsilon_t^2 - 1)\frac{1}{m}\sum_{1}^{m}(a_j - 1) \\ &\quad + \frac{1}{n}\sum\sum_{s\ne t}\varepsilon_s\varepsilon_t\frac{1}{m}\sum_{1}^{m}(a_j - 1)\cos\{(s-t)\lambda_j\}.\end{aligned}$$

From (3.19), (3.20), (3.22) and (3.23) the first term on the right is $o_p(1)$, while the second has variance

$$\frac{2}{n^2}\sum\sum_{s\neq t}\left(\frac{1}{m}\sum_1^m (a_j-1)\cos\{(s-t)\lambda_j\}\right)^2$$
$$=\frac{1}{m^2}\sum_1^m (a_j-1)^2-\frac{2}{m^2 n}\left(\sum_1^m (a_j-1)\right)^2\to 0$$

as $n\to\infty$. The proof is complete. □

4. Asymptotic normality of estimates. We now show that under conditions somewhat stronger than those of the previous section,

$$m^{1/2}(\hat{H}-H_0)\to_d N(0,\tfrac{1}{4}). \tag{4.1}$$

The variance in the limiting distribution is thus constant over H_0 and indeed completely free of unknown parameters, so (4.1) is simple to use in approximate rules of inference. (4.1) indicates that $\hat{H}$ is asymptotically more efficient (for the same m sequence) than the estimate of Geweke and Porter-Hudak (1983) as modified by Robinson (1992). Robinson (1992) developed a class of estimates whose limiting variance after $m^{1/2}$ norming has upper bound $\pi^2/24$ and lower bound $\frac{1}{4}$, though the lower bound is not precisely attainable by members of this class. Moreover, the limiting distribution theory of Robinson (1992) employed Gaussianity of x_t, and it seems unlikely that this assumption could be replaced by assumptions which are as mild or comprehensible as those under which we shall establish (4.1). The analogy with parametric problems suggests that $\frac{1}{4}$ can be identified with an efficiency bound for semiparametric estimation of H_0 in the Gaussian case with a given m sequence, but the verification of this conjecture remains an open question. The question of optimal choice of m as a function of n is also important and left for future research; some formulae for optimally choosing m in another statistic of interest in the study of long range dependence were derived by Robinson (1994c).

We have found it necessary to strengthen each of the assumptions A1–A4 in order to obtain (4.1). The new assumptions are described as follows.

ASSUMPTION A1′. For some $\beta\in(0,2]$,

$$f(\lambda)\sim G_0\lambda^{1-2H_0}(1+O(\lambda^\beta))\quad\text{as }\lambda\to 0+,$$

where $G_0\in(0,\infty)$ and $H_0\in[\Delta_1,\Delta_2]$.

ASSUMPTION A2′. In a neighbourhood $(0,\delta)$ of the origin, $\alpha(\lambda)$ is differentiable and

$$\frac{d}{d\lambda}\alpha(\lambda)=O\left(\frac{|\alpha(\lambda)|}{\lambda}\right)\quad\text{as }\lambda\to 0+.$$

ASSUMPTION A3′. Assumption A3 holds and also

$$E\left(\varepsilon_t^3 | F_{t-1}\right) = \mu_3, \quad \text{a.s.}, \qquad E\left(\varepsilon_t^4\right) = \mu_4, \qquad t = 0, \pm 1, \dots,$$

for finite constants μ_3 and μ_4.

ASSUMPTION A4′. As $n \to \infty$,

$$\frac{1}{m} + \frac{m^{1+2\beta}(\log m)^2}{n^{2\beta}} \to 0.$$

Assumption A1′ is equivalent to one employed by Robinson (1995), imposing a rate of convergence on (1.1) analogous to the smoothness conditions used in the asymptotic theory of power spectral density estimates. Assumption A2′ implies A2 because $f(\lambda) = |\alpha(\lambda)|^2/2\pi$. Assumption A3′ implies x_t is fourth-order stationary, and holds in case of independent and identically distributed ε_t with finite fourth moment. Assumption A4′, like A4, allows m to increase arbitrarily slowly with m, but it also imposes an upper bound on the rate of increase of m with n, the weakest version arising when $\beta = 2$.

THEOREM 2. *Let Assumptions* A1′–A4′ *hold. Then*

$$m^{1/2}\left(\hat{H} - H_0\right) \to_d N\left(0, \tfrac{1}{4}\right) \quad \text{as } n \to \infty.$$

PROOF. Theorem 1 holds under the current conditions and implies that with probability approaching 1, as $n \to \infty$, $\hat{H}$ satisfies

$$0 = \frac{dR(\hat{H})}{dH} = \frac{dR(H_0)}{dH} + \frac{d^2R(\tilde{H})}{dH^2}\left(\hat{H} - H_0\right), \tag{4.2}$$

where $|\tilde{H} - H_0| \le |\hat{H} - H_0|$. Now

$$\frac{dR(H)}{dH} = 2\frac{\hat{G}_1(H)}{\hat{G}(H)} - \frac{2}{m}\sum_1^m \log \lambda_j,$$

$$\frac{d^2R(H)}{dH^2} = \frac{4\left\{\hat{G}_2(H)\hat{G}(H) - \hat{G}_1^2(H)\right\}}{\hat{G}^2(H)},$$

where

$$\hat{G}_k(H) = \frac{1}{m}\sum_1^m (\log \lambda_j)^k \lambda_j^{2H-1} I_j.$$

Defining also

$$\hat{F}_k(H) = \frac{1}{m}\sum_1^m (\log j)^k \lambda_j^{2H-1} I_j, \qquad \hat{E}_k(H) = \frac{1}{m}\sum_1^m (\log j)^k j^{2H-1} I_j,$$

elementary calculation gives

$$\frac{d^2R(H)}{dH^2} = \frac{4\{\hat{F}_2(H)\hat{F}_0(H) - \hat{F}_1^2(H)\}}{\hat{F}_0^2(H)} = \frac{4\{\hat{E}_2(H)\hat{E}_0(H) - \hat{E}_1^2(H)\}}{\hat{E}_0^2(H)}. \tag{4.3}$$

Fix $\varepsilon > 0$ and choose n so that $2\varepsilon < (\log m)^2$. On the set $M = \{H: (\log m)^3 \times |H - H_0| \le \varepsilon\}$,

$$\begin{aligned}|\hat{E}_k(H) - \hat{E}_k(H_0)| &\le \frac{1}{m}\sum_1^m |j^{2(H-H_0)} - 1|(\log j)^k j^{2H_0-1} I_j \\ &\le 2e|H - H_0|\hat{E}_{k+1}(H_0) \\ &\le 2e\varepsilon(\log m)^{k-2}\hat{E}_0(H_0),\end{aligned}$$

where the second inequality uses $\frac{1}{2}|j^{2(H-H_0)} - 1|/|H - H_0| \le (\log j)m^{2|H-H_0|} \le (\log j)m^{1/\log m} = e \log j$ on M. Thus for $\eta > 0$,

$$\begin{aligned}&P\left(|\hat{E}_k(\tilde{H}) - \hat{E}_k(H_0)| > \eta\left(\frac{2\pi}{n}\right)^{1-2H_0}\right) \\ &\quad \le P\left(\hat{G}(H_0) > \frac{\eta}{2e\varepsilon}(\log m)^{2-k}\right) + P\left((\log m)^3|\tilde{H} - H_0| > \varepsilon\right).\end{aligned} \tag{4.4}$$

For any $\eta > 0$ and $k = 0, 1, 2$, the first probability on the right-hand side tends to 0 for ε sufficiently small because $\hat{G}(H_0) \to_p G_0 \in (0, \infty)$ is implied by the proof that $\sup_{\Theta_1}|A(H)| \to_p 0$ in the proof of Theorem 1. The second probability is bounded by

$$P\left(\inf_{\Theta_1 \cap \overline{N}_\delta \cap \overline{M}} S(H) \le 0\right) + P\left(\inf_{\Theta_1 \cap \overline{N}_\delta} S(H) \le 0\right) + P\left(\inf_{\Theta_2} S(H) \le 0\right),$$

where $\overline{M} = (-\infty, \infty) - M$. We have already shown in the proof of Theorem 1 that the last two probabilities tend to 0. The first probability is bounded by

$$P\left(\sup_{\Theta_1 \cap N_\delta} |T(H)| \ge \inf_{\Theta_1 \cap N_\delta \cap \overline{M}} U(H)\right). \tag{4.5}$$

By applying (3.4), and then (3.5) and the orders of magnitude for (3.7) and (3.8) noted in the proof of Theorem 1, it follows from (3.9) that (4.5) tends to 0 if

$$\sup_{\Theta_1 \cap N_\delta}\left|\frac{\hat{G}(H) - G(H)}{G(H)}\right| = o_p\left((\log m)^{-6}\right). \tag{4.6}$$

Making further use of the notation of (3.10) we have $\inf_{\Theta_1 \cap N_\delta} B(H) \geq \inf_{\Theta_1} B(H) \geq \frac{1}{2}$ for all large enough m from (3.11), and [cf. (3.12) and (3.13)]

$$\sup_{\Theta_1 \cap N_\delta} |A(H)| \leq 6 \sum_1^m \left(\frac{r}{m}\right)^{1-2\delta} \frac{1}{r^2} \left| \sum_1^r \left(\frac{I_j}{g_j} - 1\right) \right| + \frac{3}{m} \left| \sum_1^m \left(\frac{I_j}{g_j} - 1\right) \right|. \tag{4.7}$$

We now state the following properties, to be established later:

$$\sum_1^r \left(\frac{I_j}{g_j} - 2\pi I_{\varepsilon j}\right) = O_p\left(r^{1/3}(\log r)^{2/3} + r^{\beta+1} n^{-\beta} + r^{1/2} n^{-1/4}\right) \quad \text{as } n \to \infty, \tag{4.8}$$

and

$$\sum_1^r (2\pi I_{\varepsilon j} - 1) = O_p(r^{1/2}) \quad \text{as } n \to \infty, \tag{4.9}$$

for $1 \leq r \leq m$. Using A4′, we then deduce that (4.7) is $O_p(m^{-1/2})$, so it follows from (3.9) and (4.5) that $P(\inf_{\Theta_1 \cap \overline{N}_\delta \cap \overline{M}} S(H) \leq 0) \to 0$. The detailed nature of the bounds in (4.8) and (4.9) will be of further use below. We have established that (4.4) tends to 0 as $n \to \infty$. Thus

$$\begin{aligned} \frac{d^2 R(\tilde{H})}{dH^2} &= \frac{4\left[\left\{\hat{E}_2(H_0) + o_p(n^{2H_0-1})\right\}\left\{\hat{E}_0(H_0) + o_p(n^{2H_0-1})\right\} - \left\{\hat{E}_1(H_0) + o_p(n^{2H_0-1})\right\}^2\right]}{\left\{\hat{E}_0(H_0) + o_p(n^{2H_0-1})\right\}^2} \\ &= \frac{4\{\hat{F}_2(H_0)\hat{F}_0(H_0) - \hat{F}_1^2(H_0)\}}{\hat{F}_0^2(H_0)} + o_p(1) \quad \text{as } n \to \infty. \end{aligned}$$

For $k \geq 0$,

$$\begin{aligned} &\left| \hat{F}_k(H_0) - G_0 \frac{1}{m} \sum_1^m (\log j)^k \right| \\ &\leq \frac{G_0}{m} \sum_1^{m-1} \left| (\log r)^k - (\log(r+1))^k \right| \left| \sum_1^r \left\{ \frac{I_j}{g_j} - 1 \right\} \right| \\ &\quad + \frac{G_0}{m} (\log m)^k \left| \sum_1^m \left\{ \frac{I_j}{g_j} - 1 \right\} \right| \\ &= O_p\left(\frac{(\log m)^2}{m^{1/2}} \right) \quad \text{as } n \to \infty, \end{aligned}$$

using summation by parts, (4.8), (4.9) and $|(\log r)^k - (\log(r+1))^k| \le (\log(r+1))^{k-1}/r$ for $k = 1, 2$. Thus from (4.3), as $n \to \infty$,

$$\frac{d^2 R(\tilde{H})}{dH^2} = 4\left\{\frac{1}{m}\sum_1^m (\log j)^2 - \left(\frac{1}{m}\sum_1^m \log j\right)^2\right\}\left(1 + o_p(1)\right) + o_p(1) \to_p 4. \tag{4.10}$$

Now because $\hat{G}(H_0) \to_p G_o$,

$$\begin{aligned} m^{1/2}\frac{dR(H_0)}{dH} &= 2m^{-1/2}\sum_1^m \nu_j \frac{\lambda_j^{2H_0-1} I_j}{G_0 + o_p(1)} \\ &= 2m^{-1/2}\left(1 - \frac{o_p(1)}{G_0 + o_p(1)}\right)\sum_1^m \nu_j \frac{I_j}{g_j} \\ &= 2m^{-1/2}\sum_1^m \nu_j\left(\frac{I_j}{g_j} - 1\right)\left(1 + o_p(1)\right), \end{aligned} \tag{4.11}$$

where $\nu_j = \log j - m^{-1}\Sigma_1^m \log j$ satisfies $\Sigma_1^m \nu_j = 0$. From summation by parts, (4.8) and $\Sigma_1^m \log j = O(m \log m)$, (4.11) is

$$\left\{2m^{-1/2}\sum_1^m \nu_j(2\pi I_{\varepsilon j} - 1) + O_p\left(\left((\log m)^{2/3} m^{-1/6} + m^{\beta+1/2} n^{-\beta} + n^{-1/4}\right)\log m\right)\right\}\left(1 + o_p(1)\right),$$

which from Assumption A4′ and $\Sigma_1^m \nu_j = 0$ is $(2\Sigma_1^n z_t + o_p(1))(1 + o_p(1))$, where $z_1 = 0$ and, for $t \ge 2$,

$$z_t = \varepsilon_t \sum_{s=1}^{t-1} \varepsilon_s c_{t-s},$$

$$c_s = 2n^{-1} m^{-1/2} \sum_1^m \nu_j \cos(s\lambda_j),$$

supressing reference to n in z_t and c_s. Now the z_t form a zero-mean martingale difference array, and from a standard martingale CLT we can deduce that $\Sigma_1^n z_t$ tends in distribution to a $N(0,1)$ random variable if

$$\sum_1^n E\left(z_t^2 | F_{t-1}\right) - 1 \to_p 0, \tag{4.12}$$

$$\sum_1^n E\left(z_t^2 I(|z_t| > \delta)\right) \to 0 \quad \text{for all } \delta > 0. \tag{4.13}$$

The left-hand side of (4.12) is

$$\left\{\sum_{t=2}^n \sum_{s=1}^{t-1} \varepsilon_s^2 c_{t-s}^2 - 1\right\} + \sum_{t=2}^n \sum\sum_{r \ne s} \varepsilon_r \varepsilon_s c_{t-r} c_{t-s}. \tag{4.14}$$

The term in braces is

$$\left\{\sum_{1}^{n-1}\left(\varepsilon_t^2-1\right)\sum_{s=1}^{n-t}c_s^2\right\}+\left\{\sum_{1}^{n-1}\sum_{1}^{n-t}c_s^2-1\right\}. \tag{4.15}$$

Now

$$\sum_{1}^{n-1}\sum_{1}^{n-t}c_s^2=\frac{4}{mn^2}\sum\sum_{j,k=1}^{m}\nu_j\nu_k\sum_{1}^{n-1}\sum_{1}^{n-t}\cos(s\lambda_j)\cos(s\lambda_k) \tag{4.16}$$

$$\begin{aligned}&=\frac{4}{mn^2}\sum_{1}^{m}\nu_j^2\sum_{1}^{n-1}\sum_{1}^{n-t}\cos^2(s\lambda_j)+\frac{2}{mn^2}\sum\sum_{j\neq k}\nu_j\nu_k \\ &\quad\times\sum_{1}^{n-1}\sum_{1}^{n-t}\left[\cos\{s(\lambda_j+\lambda_k)\}+\cos\{s(\lambda_j-\lambda_k)\}\right].\end{aligned} \tag{4.17}$$

From the trigonometric identities [see, e.g., Zygmund (1977), page 49],

$$\sum_{t=1}^{r}\cos\theta t=\frac{\sin\left(r+\frac{1}{2}\right)\theta}{2\sin\frac{1}{2}\theta}-\frac{1}{2},$$

$$\sum_{t=1}^{r}\sin\theta t=\frac{\cos\frac{1}{2}\theta-\cos\left(r+\frac{1}{2}\right)\theta}{2\sin\frac{1}{2}\theta},\qquad \theta\neq 0,\ \mathrm{mod}(2\pi),$$

and trigonometric addition formulae, we deduce

$$\sum_{r=1}^{q-1}\sum_{t=1}^{q-r}\cos\theta t=\frac{(\cos\theta-\cos q\theta)}{4\sin^2\frac{1}{2}\theta}-\frac{q-1}{2}. \tag{4.18}$$

Thus, for $j=1,\dots,m<\frac{1}{2}n$,

$$\begin{aligned}\sum_{1}^{n-1}\sum_{1}^{n-t}\cos^2(s\lambda_j)&=\frac{1}{2}\sum_{1}^{n-1}\sum_{1}^{n-t}\left\{1+\cos(2s\lambda_j)\right\}\\&=\frac{n(n-1)}{4}-\frac{n-1}{4}=\frac{(n-1)^2}{4}.\end{aligned}$$

Because

$$\frac{1}{m}\sum_{1}^{m}\nu_j^2=1+O\left(\frac{(\log m)^2}{m}\right),$$

it follows that (4.16) tends to 1 as $n\to\infty$. Using (4.18) again, for $j,k=1,\dots,m<\frac{1}{2}n$, $j\neq k$,

$$\sum_{1}^{n-1}\sum_{1}^{n-t}\left[\cos\{s(\lambda_j+\lambda_k)\}+\cos\{s(\lambda_j-\lambda_k)\}\right]=-n,$$

so that (4.17) is

$$\frac{-2}{mn}\sum\sum_{j\neq k}\nu_j\nu_k=O\left(\frac{\sum_1^m\nu_j^2}{n}\right)=O\left(\frac{m}{n}\right)\to 0$$

as $n \to \infty$. Thus the second component of (4.15) tends to 0 as $n \to \infty$. The first component of (4.15) has zero mean and variance

$$O\left(\sum_{1}^{n-1}\left(\sum_{s=1}^{n-t} c_s^2\right)^2\right). \tag{4.19}$$

Now

$$|c_s| \le 2m^{-1/2}n^{-1}\sum_{1}^{m}|\nu_j| = O\left(\frac{m^{1/2}\log m}{n}\right), \tag{4.20}$$

whereas $c_s = c_{n-s}$ and, by summation by parts,

$$\begin{aligned}|c_s| &= \left|2m^{-1/2}n^{-1}\sum_{1}^{m-1}(\nu_j - \nu_{j+1})\sum_{1}^{j}\cos(s\lambda_l) + 2m^{-1/2}n^{-1}\nu_m\sum_{1}^{m}\cos(s\lambda_j)\right| \\ &\le Cm^{-1/2}s^{-1}\left(\sum_{1}^{m-1}\left|\log\left(1+\frac{1}{j}\right)\right| + \log m\right) \\ &= O(m^{-1/2}s^{-1}\log m),\end{aligned} \tag{4.21}$$

for $1 \le s \le n/2$, because $|\sum_1^j \cos(s\lambda_l)| = O(n/s)$ for such s [Zygmund (1977), page 2], and because $|\log(1 + 1/j)| \le 1/j$ for $j \ge 1$. The bound in (4.21) is at least as good as that in (4.20) for $n/m < s \le n/2$. From (4.20) and (4.21),

$$\begin{aligned}\sum_{s=1}^{n} c_s^2 &= O\left(\frac{n}{m}\left(\frac{m^{1/2}\log m}{n}\right)^2 + \frac{(\log m)^2}{m}\sum_{s>n/m} s^{-2}\right) \\ &= O\left(\frac{(\log m)^2}{n}\right)\end{aligned} \tag{4.22}$$

and so (4.19) $= O((\log m)^4/n)$. We have shown that (4.15) is $o_p(1)$. The second component of (4.14) has mean zero and variance

$$\begin{aligned}&2\sum\sum_{t,u=2}^{n}\ \sum\sum_{r\ne s}^{\min(t-1,u-1)}(c_{t-r}c_{t-s}c_{u-r}c_{u-s}) \\ &\quad = 2\sum_{t=2}^{n}\sum\sum_{r\ne s} c_{t-r}^2 c_{t-s}^2 + 4\sum_{t=3}^{n}\sum_{u=2}^{t-1}\sum\sum_{r\ne s}^{u-1} c_{t-r}c_{t-s}c_{u-r}c_{u-s}.\end{aligned}$$

From (4.22) the first term on the right is $O((\log m)^4/n)$. The second term has absolute value bounded by

$$\begin{aligned}&4\sum_{3}^{n}\sum_{2}^{t-1}\left(\sum_{1}^{u-1} c_{t-r}^2 \sum_{1}^{u-1} c_{u-r}^2\right) \\ &\quad \le 4\left(\sum_{1}^{n} c_t^2\right)\left(\sum_{3}^{n}\sum_{2}^{t-1}\sum_{t-u+1}^{t-1} c_r^2\right),\end{aligned} \tag{4.23}$$

and the final bracketed factor is

$$\sum_{j=1}^{n-2} j(n-j-1)c_j^2 \le 2n \sum_{1}^{[n/2]} jc_j^2$$

$$\le 2n \sum_{2}^{[n/m^{2/3}]} jc_j^2 + 2n \sum_{[n/m^{2/3}]+1}^{[n/2]} jc_j^2$$

$$= O\left(\frac{n^3}{m^{4/3}}\left(\frac{m^{1/2}\log m}{n}\right)^2 + n^2\left(\frac{\log m}{m^{1/2}}\right)^2 \sum_{[n/m^{2/3}]}^{\infty} s^{-2}\right)$$

$$= O\left(\frac{n(\log m)^2}{m^{1/3}}\right)$$

as $n \to \infty$, so (4.23) is $O((\log m)^4/m^{1/3})$ in view of (4.22). We have thus verified (4.12). We shall prove (4.13) by checking the sufficient condition

$$\sum_{1}^{n} E(z_t^4) \to 0 \quad \text{as } n \to \infty.$$

The left-hand side of this equals

$$\mu_4 \sum_{2}^{n} E\left(\sum_{1}^{t-1} \varepsilon_s c_{t-s}\right)^4$$

$$\le C \sum_{2}^{n} E\left(\sum\sum\sum_{1}^{t-1}\sum \varepsilon_s \varepsilon_r \varepsilon_q \varepsilon_p c_{t-s} c_{t-r} c_{t-q} c_{t-p}\right)$$

$$\le C \sum_{1}^{n}\left(\sum_{1}^{n} c_{t-s}^4\right) + C \sum_{1}^{n} \sum_{1}^{t-1}\sum c_{t-s}^2 c_{t-r}^2$$

$$\le Cn\left(\sum_{1}^{n} c_t^2\right)^2$$

$$= O\left(\frac{(\log m)^4}{n}\right)$$

from (4.22). We have shown that $2\sum_1^n z_t$, and thus (4.11), is asymptotically $N(0, 4)$, so that in view of (4.2) and (4.10) the proof of the theorem will be completed on verifying (4.8) and (4.9). To prove (4.9), we have

$$(4.24) \qquad \sum_{1}^{r} (2\pi I_{\varepsilon j} - 1) = \frac{r}{n} \sum_{1}^{n} (\varepsilon_t^2 - 1) + \sum_{2}^{n} \varepsilon_t \sum_{s=1}^{t-1} \varepsilon_s d_{t-s},$$

where $d_s = (2/n)\sum_1^r \cos(s\lambda_j)$. We have $|d_s| \le 2r/n$, on the one hand, and $d_s = d_{n-s}$, $|d_s| \le 2/\pi s$ for $1 \le s \le n/2$. Both terms on the right of (4.24) have zero mean, and they have variances, respectively, $O(r^2/n)$ and

$$O\left(n \sum_{1}^{n} d_s^2\right) = O\left(n \frac{n}{r}\left(\frac{r}{n}\right)^2 + n \sum_{[n/r]}^{\infty} s^{-2}\right) = O(r).$$

To prove (4.8), first choose an integer $l < r$. From (3.16) and $E(2\pi I_{\varepsilon j}) = 1$,

$$E\left|\sum_{1}^{l}\left(\frac{I_j}{g_j} - 2\pi I_{\varepsilon j}\right)\right| = O(l) \quad \text{as } n \to \infty. \tag{4.25}$$

From (3.16) and A1′,

$$E\left|\sum_{l+1}^{r}\left(\frac{I_j}{g_j} - \frac{I_j}{f_j}\right)\right| \le C\sum_{l+1}^{r}\left|1 - \frac{g_j}{f_j}\right|$$
$$= O\left(\frac{r^{\beta+1}}{n^{\beta}}\right) \quad \text{as } n \to \infty.$$

Write $u_j = w_j/|\alpha_j|$, $v_j = w_{\varepsilon j}$. Then we are left with consideration of

$$E\left\{\sum_{l+1}^{r}\left(\frac{I_j}{f_j} - 2\pi I_{\varepsilon j}\right)\right\}^2 = (2\pi)^2(a+b),$$

where

$$a = \sum_{l+1}^{r}\left(E|u_j|^4 - 2E|u_jv_j|^2 + E|v_j|^4\right),$$
$$b = 2\sum_{\substack{j<k\\ l+1}}^{r}\sum\left(E|u_ju_k|^2 - E|u_jv_k|^2 - E|u_kv_j|^2 + E|v_jv_k|^2\right).$$

We then have $a = a_1 + a_2$ and $b = b_1 + b_2$, where

$$a_1 = \sum_{l+1}^{r}\Big\{2\left(E|u_j|^2\right)^2 + |E(u_j^2)|^2 - 2|E(u_jv_j)|^2 - 2|E(u_j\bar{v}_j)|^2$$
$$-2E|u_j|^2E|v_j|^2 + 2\left(E|v_j|^2\right)^2 + |E(v_j^2)|^2\Big\},$$

$$a_2 = \sum_{l+1}^{r}\left\{\operatorname{cum}(u_j, u_j, \bar{u}_j, \bar{u}_j) - 2\operatorname{cum}(u_j, v_j, \bar{u}_j, \bar{v}_j) + \operatorname{cum}(v_j, v_j, \bar{v}_j, \bar{v}_j)\right\},$$

$$b_1 = 2\sum_{\substack{j<k\\ l+1}}^{r}\sum\Big\{E|u_j|^2E|u_k|^2 + |E(u_ju_k)|^2 + |E(u_j\bar{u}_k)|^2 - E|u_j|^2E|v_k|^2$$
$$-|E(u_jv_k)|^2 - |E(u_j\bar{v}_k)|^2 - E|u_k|^2E|v_j|^2 - |E(u_kv_j)|^2$$
$$-|E(u_k\bar{v}_j)|^2 + E|v_j|^2E|v_k|^2 + |E(v_jv_k)|^2 + |E(v_j\bar{v}_k)|^2\Big\},$$

$$b_2 = 2\sum_{\substack{j<k\\ l+1}}^{r}\sum\Big\{\operatorname{cum}(u_j, u_k, \bar{u}_j, \bar{u}_k) - \operatorname{cum}(u_j, v_k, \bar{u}_j, \bar{v}_k)$$
$$-\operatorname{cum}(v_j, u_k, \bar{v}_j, \bar{u}_k) + \operatorname{cum}(v_j, v_k, \bar{v}_j, \bar{v}_k)\Big\},$$

where $\operatorname{cum}(\cdot,\cdot,\cdot,\cdot)$ is the joint cumulant of the argument random variables. Because $E|v_j|^2 \equiv 1$,

$$a_1 = \sum_{l+1}^{r} \Big\{ 2\big(E|u_j|^2 - 1\big)^2 + 2\big(E|u_j|^2 - 1\big) + |E(u_j^2)|^2 - 2|E(u_j v_j)|^2$$

$$- 2|E(u_j \bar{v}_j) - 1|^2 - 2\big(E u_j \bar{v}_j - 1\big) - 2\big(E\bar{u}_j v_j - 1\big) + |E(v_j^2)|^2 \Big\}$$

$$= O\left(\sum_{l+1}^{r} \left(\frac{\log j}{j} \right) \right) = O\big((\log r)^2\big),$$

$$b_1 = 2 \sum\sum_{\substack{j<k \\ j=l+1}}^{r} \Big\{ \big(E|u_j|^2 - 1\big)\big(E|u_k|^2 - 1\big) + |E(u_j u_k)|^2 + |E(u_j \bar{u}_k)|^2$$

$$(4.26) \qquad - |E(u_j v_k)|^2 - |E(u_j \bar{v}_k)|^2 - |E(u_k v_j)|^2 - |E(u_k \bar{v}_j)|^2$$

$$+ |E(v_j v_k)|^2 + |E(v_j \bar{v}_k)|^2 \Big\}$$

$$= O\left(\sum\sum_{\substack{j<k \\ j=l+1}}^{r} \left\{ \left(\frac{\log j}{j} \right)\left(\frac{\log k}{k} \right) + \left(\frac{\log k}{j} \right)^2 \right\} \right)$$

$$= O\left((\log r)^2 \sum_{k=l+2}^{r} \sum_{j=l+1}^{k-1} j^{-2} \right) = O\left(\frac{r(\log r)^2}{l} \right)$$

as $n \to \infty$, using Theorem 2 of Robinson (1995). Choosing $l \sim r^{1/3}(\log r)^{2/3}$ in (4.25) and (4.26) gives the $O_p(r^{1/3}(\log r)^{2/3})$ component in (4.8). Now consider a_2 and b_2. Applying formulae of Brillinger [(1975), (2.6.3), page 26, and (2.10.3), page 39], we deduce after straightforward calculation that the summand of $(2\pi)^5\, b_2$ is

$$(4.27) \quad \frac{\kappa}{n^2} \iiint_{-\pi}^{\pi} \left\{ \frac{\alpha(\lambda + \mu + \zeta)\alpha(-\mu)}{|\alpha_j|^2} - 1 \right\} \left\{ \frac{\alpha(-\lambda)\alpha(-\zeta)}{|\alpha_k|^2} - 1 \right\} \times E_{jk}(\lambda,\mu,\zeta)\, d\lambda\, d\mu\, d\zeta,$$

where $\kappa = \mu_4 - 3$ and

$$E_{jk}(\lambda,\mu,\zeta) = D(\lambda_j - \lambda - \mu - \zeta) D(\lambda_k + \lambda) D(\mu - \lambda_j) D(\zeta - \lambda_k),$$

$$D(\lambda) = \sum_{t=1}^{n} e^{it\lambda}.$$

Now use the identity

$$(c_1 c_2 - 1)(c_3 c_4 - 1) = \prod_{1}^{4} (c_j - 1) + \sum_{i=1}^{4} \prod_{\substack{j=1 \\ j \neq i}}^{4} (c_j - 1)$$

$$+ \sum\sum_{i,j=1}^{2} (c_i - 1)(c_{j+2} - 1)$$

to observe that (4.27) has components of three types. The first is

$$\frac{\kappa}{n^2}\iiint_{-\pi}^{\pi}\left\{\frac{\alpha(\lambda+\mu+\zeta)}{\alpha_j}-1\right\}\left\{\frac{\alpha(-\mu)}{\bar{\alpha}_j}-1\right\}\left\{\frac{\alpha(-\lambda)}{\alpha_k}-1\right\}\left\{\frac{\alpha(-\zeta)}{\bar{\alpha}_k}-1\right\}$$
$$\times E_{jk}(\lambda,\mu,\zeta)\,d\lambda\,d\mu\,d\zeta.$$

By the Schwarz inequality and periodicity, this is bounded in absolute value by

$$(2\pi)^3\kappa P_j P_k, \tag{4.28}$$

where

$$P_j=\int_{-\pi}^{\pi}\left|\frac{\alpha(\lambda)}{\alpha_j}-1\right|^2 K(\lambda-\lambda_j)\,d\lambda$$

and

$$K(\lambda)=\frac{|D(\lambda)|^2}{2\pi n}$$

is Fejèr's kernel. The second sort of component of (4.27) is typified by

$$\frac{\kappa}{(2\pi n)^2}\iiint_{-\pi}^{\pi}\left\{\frac{\alpha(\lambda+\mu+\zeta)}{\alpha_j}-1\right\}\left\{\frac{\alpha(-\mu)}{\bar{\alpha}_j}-1\right\}\left\{\frac{\alpha(-\zeta)}{\bar{\alpha}_k}-1\right\}$$
$$\times E_{jk}(\lambda,\mu,\zeta)\,d\lambda\,d\mu\,d\zeta.$$

The modulus of this is bounded by

$$(2\pi)^3\kappa P_j P_k^{1/2} \tag{4.29}$$

because

$$\int_{-\pi}^{\pi}K(\lambda)\,d\lambda=1. \tag{4.30}$$

An example of the third type of component of (4.27) is

$$\frac{\kappa}{(2\pi n)^2}\iiint_{-\pi}^{\pi}\left\{\frac{\alpha(\lambda+\mu+\zeta)}{\alpha_j}-1\right\}\left\{\frac{\alpha(-\lambda)}{\alpha_k}-1\right\}$$
$$\times E_{jk}(\lambda,\mu,\zeta)\,d\lambda\,d\mu\,d\zeta$$
$$=\frac{\kappa}{(2\pi n)^2}\iiint_{-\pi}^{\pi}\left\{\frac{\alpha(\theta)}{\alpha_j}-1\right\}\left\{\frac{\alpha(-\lambda)}{\alpha_k}-1\right\}$$
$$\times E_{jk}(\lambda,\theta-\lambda-\zeta,\zeta)\,d\lambda\,d\mu\,d\zeta$$

$$=\frac{\kappa}{n^2}\iint_{-\pi}^{\pi}\left\{\frac{\alpha(\theta)}{\alpha_j}-1\right\}\left\{\frac{\alpha(-\lambda)}{\alpha_k}-1\right\} \tag{4.31}$$
$$\times D(\lambda_j-\theta)D(\lambda_k+\lambda)D(\theta-\lambda-\lambda_j-\lambda_k)\,d\theta\,d\lambda,$$

because

$$\int_{-\pi}^{\pi} D(u+\lambda)D(v-\lambda)\,d\lambda = 2\pi D(u+v).$$

The absolute value of (4.31) is bounded by

$$\frac{(2\pi)^2\kappa}{n^{1/2}} P_j^{1/2} P_k^{1/2}. \tag{4.32}$$

The summand of a_2 is (4.27) with $j = k$ and has components bounded by (4.28), (4.29) or (4.32), with $j = k$. Applying Lemma 3, we deduce that

$$a_2 = O\left(\sum_1^r (j^{-2} + j^{-3/2} + n^{-1/2}j^{-1})\right) = O(1),$$

$$\begin{aligned} b_2 &= O\left(\sum_{\substack{j<k\\ l+1}}^{r}\sum (j^{-1}k^{-1} + j^{-1}k^{-1/2} + j^{-1/2}k^{-1} + n^{-1/2}j^{-1/2}k^{-1/2})\right) \\ &= O\big((\log r)^2 + r^{1/2}\log r + rn^{-1/2}\big), \end{aligned}$$

to complete the proof of the theorem. □

5. Technical lemmas.

LEMMA 1. *For $\varepsilon \in (0, 1]$ and $C \in (\varepsilon, \infty)$, as $m \to \infty$,*

$$\sup_{C \geq \gamma \geq \varepsilon} \left| \frac{\gamma}{m} \sum_1^m \left(\frac{j}{m}\right)^{\gamma-1} - 1 \right| = O\left(\frac{1}{m^{\varepsilon}}\right).$$

PROOF. Because $\int_0^a x^{\gamma-1}\,dx = a^{\gamma}/\gamma$ for $\gamma > 0$,

$$\begin{aligned} \left| \frac{\gamma}{m} \sum_1^m \left(\frac{j}{m}\right)^{\gamma-1} - 1 \right| &\leq \gamma \int_0^{1/m} \left\{ \left(\frac{1}{m}\right)^{\gamma-1} + x^{\gamma-1} \right\} dx \\ &\quad + \gamma \sum_2^m \left| \int_{(j-1)/m}^{j/m} \left\{ \left(\frac{j}{m}\right)^{\gamma-1} - x^{\gamma-1} \right\} dx \right| \\ &\leq \frac{\gamma}{m^{\gamma}} + \frac{1}{m^{\gamma}} + \frac{\gamma|\gamma-1|}{m^2} \sum_1^m \left(\frac{j}{m}\right)^{\gamma-2} \end{aligned}$$

by the mean-value theorem. The last term is $O(\gamma^2 m^{-1})$ for $\gamma > 1$, zero for $\gamma = 1$ and $O(m^{-\gamma})$ for $0 < \gamma < 1$, whence the conclusion follows straightforwardly. □

LEMMA 2. *For all $m \geq 2$,*

$$\left| \frac{1}{m} \sum_1^m \log j - \log m + 1 \right| \leq \frac{2 + \log(m-1)}{m}. \tag{5.1}$$

PROOF. Because $\int_0^m \log x\,dx = m(\log m - 1)$, the left-hand side of (5.1) is

$$\left| \frac{1}{m} \int_0^1 \log x\,dx - \frac{1}{m} \sum_2^m \int_{j-1}^j \log\left(\frac{j}{x}\right) dx \right|$$
$$\leq \frac{1}{m} + \frac{1}{m} \sum_1^{m-1} \frac{1}{j} \leq \frac{2 + \log(m-1)}{m}. \qquad \square$$

LEMMA 3. *Under* A1 *and* A2′, *uniformly in integer* j *such that* $j/n \to 0$ *as* $n \to \infty$,

$$\int_{-\pi}^{\pi} \left| \frac{\alpha(\lambda)}{\alpha(\lambda_j)} - 1 \right|^2 K(\lambda - \lambda_j)\, d\lambda = O\left(\frac{1}{j}\right) \quad \text{as } n \to \infty.$$

PROOF. We split the integral up as follows:

$$\int_{-\pi}^{-\delta} + \int_{-\delta}^{-\lambda_j/2} + \int_{-\lambda_j/2}^{\lambda_j/2} + \int_{\lambda_j/2}^{2\lambda_j} + \int_{2\lambda_j}^{\delta} + \int_{\delta}^{\pi},$$

for $\delta \in (2\lambda_j, \pi)$. For n sufficiently large, A1 and A2′ imply that we can choose δ such that, for some $C < \infty$,

$$|\alpha(\lambda)| \leq C|\lambda|^{1/2 - H_0}, \qquad |\alpha'(\lambda)| \leq C|\lambda|^{-1/2 - H_0}, \qquad 0 < |\lambda| < \delta.$$

Now

$$\left|\int_{-\pi}^{-\delta}\right| \leq \frac{2}{f_j}\left\{\int_{-\pi}^{-\delta} f(\lambda) K(\lambda - \lambda_j)\, d\lambda + f_j \int_{-\pi}^{-\delta} K(\lambda - \lambda_j)\, d\lambda\right\}$$
$$\leq \frac{1}{\pi f_j}\left\{\frac{1}{n\delta^2}\int_{-\pi}^{\pi} f(\lambda)\, d\lambda + f_j \frac{2\pi}{n\delta^2}\right\}$$
$$= O\left(\frac{1}{n}\left(\frac{1}{f_j} + 1\right)\right) = O\left(\frac{\lambda_j^{2H}}{j} + \frac{1}{n}\right) = O\left(\frac{1}{j}\right),$$

where the second inequality uses the property [Zygmund (1977), pages 49–51]

$$|D(\lambda)| < 2/|\lambda|, \qquad 0 < \lambda < \pi; \tag{5.2}$$

also $|\int_\delta^\pi|$ has the same bound by the same proof. Using (5.2) again,

$$\left|\int_{-\delta}^{-\lambda_j/2}\right| \leq \frac{1}{f_j}\left\{\int_{-\delta}^{-\lambda_j/2} f(\lambda) K(\lambda - \lambda_j)\, d\lambda + f_j \int_{-\delta}^{-\lambda_j/2} K(\lambda - \lambda_j)\, d\lambda\right\}$$
$$\leq \frac{1}{2\pi n f_j}\left\{\int_{\lambda_j/2}^{\pi} \lambda^{-1-2H}\, d\lambda + f_j \int_{\lambda_j/2}^{\pi} \lambda^{-2}\, d\lambda\right\}$$
$$= O\left(\frac{\lambda_j^{-2H}}{n f_j}\right) = O\left(\frac{1}{j}\right);$$

$|\int_{2\lambda_j}^{\delta}|$ has the same bound by the same proof. Next,

$$\left|\int_{-\lambda_j/2}^{\lambda_j/2}\right| \le f_j^{-1} \max_{|\lambda| \le \lambda_j/2} K(\lambda - \lambda_j) \left\{ \int_{-\lambda_j/2}^{\lambda_j/2} f(\lambda)\, d\lambda + \lambda_j f_j \right\}$$

$$= O\left(f_j^{-1} \frac{1}{n \lambda_j^2} \lambda_j f_j \right) = O\left(\frac{1}{j} \right).$$

Finally, by the mean-value theorem,

$$\left|\int_{\lambda_j/2}^{2\lambda_j}\right| \le \frac{1}{|\alpha_j|^2} \max_{\lambda_j/2 \le \lambda \le 2\lambda_j} \left| \frac{d\alpha(\lambda)}{d\lambda} \right|^2 \int_{\lambda_j/2}^{2\lambda_j} |\lambda - \lambda_j|^2 K(\lambda - \lambda_j)\, d\lambda$$

$$= O\left(\frac{1}{n \lambda_j} \right) = O\left(\frac{1}{j} \right). \qquad \square$$

6. Numerical work. The finite-sample behaviour of $\hat{H}$ was investigated in a small Monte Carlo study. $\hat{H}$ was also compared with a simple closed-form estimate:

$$\tilde{H} = \frac{1}{2} \left\{ 1 - \frac{\sum_{j=1}^{m} \log I_j (\log j - (1/m) \sum_{l=1}^{m} \log l)}{\sum_{j=1}^{m} \log j (\log j - (1/m) \sum_{l=1}^{m} \log l)} \right\}. \tag{6.1}$$

This estimate is a slightly simplified version of the one proposed by Geweke and Porter-Hudak (1983). Robinson (1995) derived asymptotic theory for a modified form of $\tilde{H}$, in which the contribution from a slowly increasing (with m) number l of the lowest frequencies $\lambda_1, \dots, \lambda_l$ are deleted from (6.1). [Künsch (1986) earlier suggested such a trimming.] It is not known whether the trimming is necessary in order to achieve the desirable asymptotic property

$$m^{1/2}(\tilde{H} - H) \to_d N\left(0, \frac{\pi^2}{24}\right) \tag{6.2}$$

obtained by Robinson (1995); the bulk of the many empirical applications of the Geweke–Porter-Hudak approach have used all of the m lowest frequencies. (We omit the zero subscript on H throughout this section and Tables 1–8.) In view of this, and for simplicity and ease of comparison with respect to degrees of freedom, we employ the untrimmed estimate (6.1) here.

Using an algorithm of Davies and Harte (1987) and random number generator GO5DDF from the NAG library, Gaussian time series were generated with mean zero, variance unity and lag-j autocovariance

$$\gamma_j = \tfrac{1}{2}\left(|j+1|^{2H} - 2|j|^{2H} + |j-1|^{2H}\right). \tag{6.3}$$

We call this Model A. The corresponding spectral density satisfies (1.1); indeed it satisfies Assumption A1′ with $\beta = 2$. Five values of H were employed: $H = 0.1$, 0.3, 0.5, 0.7 and 0.9; $H = 0.5$ corresponds to white noise in this model. The sample sizes chosen were $n = 64$, 128 and 256, and for each of these, three values of m were tried: $n/16$, $n/8$ and $n/4$. For each (H, n, m) combination, 5000 replications were generated. From each of these, $\tilde{H}$ and $\hat{H}$ were computed, in the latter case using a simple golden section search applied to the first derivative of the objective function. No difficulties were encountered in computing $\hat{H}$, and for selected replications $R(H)$ was plotted and always found either to have a single relative minimum or (in some cases when $H = 0.1$ or 0.9) to be minimized at 0.001 or 0.999, these being our chosen values of Δ_1 and Δ_2. Tables 1–6 give the Monte Carlo biases, standard deviations, mean squared errors (MSE), relative efficiencies (ratios of the $\hat{H}$ and $\tilde{H}$ MSE's) and 95% and 99% coverage frequencies based on the limit distributions in (4.1) and (6.2). In Tables 1–3 and 5–6, "log" refers to the log-periodogram estimate $\tilde{H}$ and "eff" refer to the more efficient estimate $\hat{H}$.

For the most part, $\hat{H}$ seems more biased than $\tilde{H}$, though this effect tends to be reversed when $m = n/8$ and $n/4$ for $n = 256$, and when $m = n/4$ for the other values of n. $\tilde{H}$ is apt to be negatively biased for small H and positively biased for large H, but a negative bias in $\hat{H}$ is more pervasive. Unsurprisingly, bias tends to increase with m. The standard deviations in Table 2 all diminish as both m and n increase, and the $\hat{H}$ standard deviations are always decisively the smaller. Table 3 presents a similar picture, and in only one case does MSE show an increase with m. The entries in Table 4 are to be compared with 0.608, which is the asymptotic relative efficiency [from (4.1) and (6.2)]. Though the finite-sample superiority of $\hat{H}$ tends not to be as great for the central values of H, for $H = 0.9$ and especially $H = 0.1$, it is much better. Much the same can be said of the relative accuracy of the 95% and 99% interval estimates, but whereas $\hat{H}$ is best in 24 of the 45 cases in Table 5, in Table 6 it is best on 29 occasions, with one tie.

Model A, (6.3), is favourably disposed toward both $\hat{H}$ and $\tilde{H}$ for any value of m, because the approximation

$$f(\lambda) \doteq G\lambda^{1-2H} \tag{6.4}$$

is good for all $\lambda \in (0, \pi]$. However, both $\hat{H}$ and $\tilde{H}$ are motivated by the far wider range of circumstances (1.1). In many of these there is the possibility that (6.4) is not good over $(0, 2\pi m/n]$, and this is a potential source of bias.

TABLE 1
Bias, Model A

H	n = 64			n = 128			n = 256		
	m = n / 16	m = n / 8	m = n / 4	m = n / 16	m = n / 8	m = n / 4	m = n / 16	m = n / 8	m = n / 4
0.1 log	−0.008	−0.064	−0.123	−0.015	−0.066	−0.125	−0.016	−0.070	−0.124
eff	0.132	0.018	−0.049	0.048	−0.022	−0.071	0.006	−0.044	−0.081
0.3 log	−0.009	−0.008	−0.038	−0.004	−0.013	−0.032	−0.001	−0.014	−0.033
eff	0.041	−0.011	−0.050	−0.006	−0.028	−0.044	−0.016	−0.026	−0.037
0.5 log	−0.001	−0.005	−0.003	0.007	0.001	0.002	−0.001	−0.004	0.001
eff	−0.031	−0.026	−0.020	−0.022	−0.020	−0.011	−0.022	−0.013	−0.005
0.7 log	0.012	0.010	0.018	0.012	0.003	0.015	0.002	0.012	0.014
eff	−0.099	−0.044	−0.004	−0.041	−0.017	−0.004	−0.020	−0.001	0.009
0.9 log	0.008	0.029	0.035	0.025	0.021	0.031	0.016	0.015	0.027
eff	−0.181	−0.084	−0.019	−0.088	−0.030	0.005	−0.034	−0.007	0.016

TABLE 2
Standard deviation, Model A

H	n = 64			n = 128			n = 256		
	m = n / 16	m = n / 8	m = n / 4	m = n / 16	m = n / 8	m = n / 4	m = n / 16	m = n / 8	m = n / 4
0.1 log	0.610	0.349	0.218	0.350	0.215	0.139	0.214	0.143	0.093
eff	0.307	0.173	0.084	0.190	0.104	0.051	0.121	0.071	0.034
0.3 log	0.613	0.344	0.213	0.347	0.209	0.136	0.215	0.136	0.091
eff	0.345	0.239	0.159	0.244	0.161	0.109	0.166	0.109	0.073
0.5 log	0.606	0.344	0.210	0.348	0.209	0.135	0.209	0.134	0.089
eff	0.366	0.270	0.175	0.268	0.175	0.111	0.173	0.110	0.071
0.7 log	0.606	0.347	0.211	0.347	0.216	0.136	0.209	0.133	0.090
eff	0.362	0.266	0.173	0.264	0.175	0.111	0.172	0.109	0.073
0.9 log	0.597	0.345	0.210	0.340	0.212	0.135	0.211	0.136	0.089
eff	0.333	0.222	0.134	0.221	0.138	0.091	0.142	0.094	0.065

TABLE 3
MSE, Model A

H	$n=64$			$n=128$			$n=256$		
	$m=n/16$	$m=n/8$	$m=n/4$	$m=n/16$	$m=n/8$	$m=n/4$	$m=n/16$	$m=n/8$	$m=n/4$
0.1 log	0.372	0.126	0.063	0.123	0.050	0.035	0.046	0.025	0.024
eff	0.112	0.030	0.010	0.039	0.011	0.008	0.015	0.007	0.008
0.3 log	0.376	0.118	0.047	0.120	0.044	0.019	0.046	0.019	0.009
eff	0.121	0.057	0.028	0.060	0.027	0.014	0.028	0.013	0.007
0.5 log	0.367	0.118	0.044	0.121	0.044	0.018	0.044	0.018	0.008
eff	0.135	0.073	0.031	0.073	0.031	0.012	0.031	0.012	0.005
0.7 log	0.367	0.121	0.045	0.121	0.047	0.019	0.044	0.018	0.008
eff	0.141	0.073	0.030	0.072	0.031	0.012	0.030	0.012	0.005
0.9 log	0.356	0.120	0.045	0.116	0.045	0.019	0.045	0.019	0.009
eff	0.144	0.056	0.018	0.057	0.020	0.008	0.021	0.009	0.004

TABLE 4
Relative efficiency, Model A

H	$n=64$			$n=128$			$n=256$		
	$m=n/16$	$m=n/8$	$m=n/4$	$m=n/16$	$m=n/8$	$m=n/4$	$m=n/16$	$m=n/8$	$m=n/4$
0.1	0.300	0.240	0.151	0.314	0.226	0.220	0.317	0.276	0.322
0.3	0.321	0.487	0.594	0.494	0.611	0.706	0.601	0.677	0.725
0.5	0.368	0.620	0.705	0.600	0.714	0.679	0.701	0.694	0.645
0.7	0.385	0.604	0.669	0.593	0.661	0.662	0.687	0.669	0.647
0.9	0.403	0.470	0.406	0.488	0.443	0.433	0.474	0.475	0.517

TABLE 5
95% coverage probabilities, Model A

H	$n = 64$			$n = 128$			$n = 256$		
	$m = n/16$	$m = n/8$	$m = n/4$	$m = n/16$	$m = n/8$	$m = n/4$	$m = n/16$	$m = n/8$	$m = n/4$
0.1 log	0.725	0.816	0.800	0.806	0.851	0.760	0.869	0.846	0.648
eff	0.841	0.938	0.991	0.907	0.975	0.997	0.949	0.988	1.000
0.3 log	0.722	0.825	0.865	0.806	0.874	0.888	0.862	0.895	0.900
eff	0.847	0.915	0.823	0.912	0.841	0.859	0.831	0.873	0.864
0.5 log	0.733	0.817	0.870	0.816	0.870	0.900	0.879	0.907	0.924
eff	0.637	0.766	0.839	0.762	0.840	0.885	0.843	0.883	0.914
0.7 log	0.728	0.813	0.868	0.813	0.868	0.900	0.871	0.902	0.922
eff	0.797	0.852	0.840	0.860	0.837	0.883	0.838	0.889	0.910
0.9 log	0.738	0.814	0.860	0.814	0.869	0.890	0.869	0.895	0.909
eff	0.798	0.865	0.924	0.861	0.915	0.954	0.909	0.940	0.974

TABLE 6
99% coverage probabilities, Model A

H	$n = 64$			$n = 128$			$n = 256$		
	$m = n/16$	$m = n/8$	$m = n/4$	$m = n/16$	$m = n/8$	$m = n/4$	$m = n/16$	$m = n/8$	$m = n/4$
0.1 log	0.832	0.903	0.903	0.908	0.932	0.885	0.946	0.932	0.817
eff	0.900	0.972	0.997	0.956	0.990	1.000	0.984	0.997	1.000
0.3 lg	0.837	0.917	0.942	0.908	0.953	0.957	0.946	0.966	0.966
eff	0.905	0.960	0.988	0.956	0.985	0.934	0.981	0.951	0.948
0.5 log	0.847	0.910	0.948	0.913	0.953	0.966	0.950	0.969	0.981
eff	1.000	0.868	0.926	0.873	0.928	0.954	0.931	0.958	0.977
0.7 log	0.849	0.911	0.947	0.905	0.948	0.967	0.946	0.972	0.976
eff	0.857	0.909	0.955	0.923	0.948	0.956	0.952	0.964	0.973
0.9 log	0.843	0.913	0.947	0.917	0.950	0.967	0.952	0.969	0.977
eff	0.862	0.924	0.965	0.921	0.962	0.983	0.956	0.978	0.991

TABLE 7
Bias, Model B

	$n = 64$			$n = 128$			$n = 256$		
H	$m = n/16$	$m = n/8$	$m = n/4$	$m = n/16$	$m = n/8$	$m = n/4$	$m = n/16$	$m = n/8$	$m = n/4$
0.1 log	0.193	−0.175	0.048	0.230	−0.050	0.055	0.253	0.013	0.058
eff	0.218	−0.042	−0.034	0.212	−0.025	−0.042	0.231	0.003	−0.036
0.3 log	0.031	−0.260	−0.044	0.045	−0.160	−0.044	0.054	−0.113	−0.047
eff	0.040	−0.196	−0.166	0.024	−0.170	−0.169	0.033	0.136	−0.161
0.5 log	−0.028	−0.279	−0.083	−0.026	−0.191	−0.084	−0.021	−0.151	−0.088
eff	−0.074	−0.289	−0.228	−0.062	-0.229	−0.213	−0.045	−0.179	−0.197
0.7 log	−0.039	−0.285	−0.104	−0.137	−0.198	−0.102	−0.030	−0.158	−0.104
eff	−0.154	−0.337	−0.259	−0.091	−0.244	−0.232	−0.057	−0.186	−0.214
0.9 log	−0.061	−0.381	−0.169	−0.049	−0.270	−0.161	−0.038	−0.218	−0.160
eff	−0.244	−0.450	−0.375	−0.137	−0.324	−0.334	−0.080	−0.252	−0.306

TABLE 8
Bias, Model C

	$n = 64$			$n = 128$			$n = 256$		
H	$m = n/16$	$m = n/8$	$m = n/4$	$m = n/16$	$m = n/8$	$m = n/4$	$m = n/16$	$m = n/8$	$m = n/4$
0.1 log	0.193	0.050	−0.254	0.208	0.092	−0.205	0.222	0.115	−0.177
eff	0.218	0.070	−0.092	0.192	0.081	−0.095	0.201	0.104	−0.096
0.3 log	0.045	−0.027	−0.275	0.038	−0.007	−0.237	0.041	0.002	−0.219
eff	0.048	−0.042	−0.256	0.019	−0.030	−0.258	0.020	−0.012	−0.254
0.5 log	0.002	−0.045	−0.278	−0.008	−0.031	−0.242	−0.008	−0.025	−0.225
eff	−0.055	−0.089	−0.347	−0.046	−0.061	−0.318	−0.032	−0.039	−0.290
0.7 log	−0.004	−0.048	−0.299	−0.012	−0.033	−0.259	−0.011	−0.027	−0.241
eff	−0.131	−0.110	−0.398	−0.069	−0.065	−0.344	−0.038	−0.041	−0.310
0.9 log	−0.010	−0.080	−0.426	−0.012	−0.054	−0.371	−0.009	−0.044	−0.344
eff	−0.214	−0.167	−0.539	−0.111	−0.097	−0.464	−0.056	−0.062	−0.415

Consider, in particular, the process

$$x_t = y_t + z_t, \tag{6.5}$$

where $\{y_t\}$ is Gaussian with autocovariances given by (6.3), and $\{z_t\}$ is a second-order autoregressive (AR) process

$$z_t + a_1 z_{t-1} + a_2 z_{t-2} = \eta_t, \tag{6.6}$$

such that η_t is an iid $N(0,1)$ sequence. When the quadratic $x^2 + a_1 x + a_2$ has complex zeroes, the spectral density of z_t has a peak at a frequency in $(0, \pi]$, in particular, at $\tilde{\lambda} = \arccos\{-a_1(1 + a_2)/4a_2\}$. For example,

$$a_1 = -1.34, \qquad a_2 = 0.9 \quad \Rightarrow \quad \tilde{\lambda} = \pi/4, \tag{6.7}$$

$$a_1 = -0.725, \qquad a_2 = 0.9 \quad \Rightarrow \quad \tilde{\lambda} = 3\pi/8 \tag{6.8}$$

and the amplitude of the peak is very sharp (the zeroes of $x^2 + a_1 x + a_2$ have moduli 0.948 in both cases). Because it is just the sum of the spectra of y_t and z_t, $f(\lambda)$ will have a peak at around $\tilde{\lambda}$, while at $\lambda = 0$ it will be infinite when $H > 0.5$ and small but positive when $H < 0.5$. The values of m and n chosen in the Monte Carlo study reported above entail the frequency bands $(0, \pi/8]$, $(0, \pi/4]$ and $(0, \pi/2]$, respectively, over which $f(\lambda)$ corresponding to (6.5), (6.6) and (6.7) or (6.8) would be influenced in various ways by the peak at $\tilde{\lambda}$.

The Monte Carlo experiment was repeated for (6.5)–(6.8), using the same values of H, m and n and the same number of replications as before. Only biases are reported. Tables 7 and 8, respectively contain results for the parameter values in (6.7) and (6.8), referring to these as, respectively, Models B and C. For Model B, (6.7), the AR peak happens to occur around the $m = n/8$ cutoff point, and the serious negative biases when $H > 0.5$ are not unexpected. For $m = n/4$, the whole peak is included and the biases are less, owing to some cancellation effect, but evidently meaningful estimates of H are not achieved. For $m = n/16$, the AR peak has some negative impact on bias when $H > 0.5$, as comparison with Table 1 indicates, but it is not nearly as significant as in the other cases. When $H < 0.5$ in (6.3) for Models B and C, (1.1) is actually misspecified (or rather, it has $H = 0.5$), and this appears to be dominant in producing the positive biases in Table 7 when $H = 0.1$ and $m = n/16$, while the AR peak is responsible for the large negative biases when $H = 0.3$ and $m = n/8$. In Table 8, the more distant AR peak at $3\pi/8$ produces a larger bias for $m = n/4$, but a smaller one for $m = n/8$. In both Tables 7 and 8, the overwhelming tendency when $H \geq 0.5$ is for $\tilde{H}$ to be less biased than $\hat{H}$ and for bias to decrease in n, though the reverse phenomenon predominates when $H = 0.1$. To place these results in perspective, it should be observed that if m had been chosen according to some optimal bandwidth scheme [see, e.g., Robinson (1994c)], it would increase more slowly with n than the proportionate increase employed in the Monte Carlo, while a preliminary plot of the logged periodogram can help to avoid pitfalls.

In the long range dependence literature, various parametric and semiparametric estimates of H have been reported for the time series of 663 annual minimum water levels of the River Nile measured at the Roda Gorge near Cairo during the years 622 through 1284. [The data are in Toussoun (1925); for subsequent years there are missing observations.] For $m = 41$, we obtained $\tilde{H} = 1.033$ and $\hat{H} = 0.941$, so that $\tilde{H}$ is outside the stationary region. For $m = 82$, $\tilde{H} = 0.920$ and $\hat{H} = 0.905$. For $m = 164$, the estimates are again smaller but the order is reversed: $\tilde{H} = 0.855$ and $\hat{H} = 0.866$.

Acknowledgments. I thank two referees for helpful comments, Joris Pinkse, Ignacio Lobato and Paolo Zaffaroni for carrying out the numerical work and Ignacio Lobato for correcting a number of errors.

REFERENCES

BRILLINGER, D. R. (1975). *Time Series, Data Analysis and Theory.* Holden-Day, San Francisco.

CHEUNG, Y-W. and DIEBOLD, F. X. (1994). On maximum-likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean. *J. Econometrics* **62** 317–350.

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

DAVIES, R. B. and HARTE, D. S. (1987). Tests for Hurst effect. *Biometrika* **74** 95–101.

FOX, R. and TAQQU, M. S. (1986). Large-sample 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 Ser. Anal.* **4** 221–238.

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 Whittle's estimate. *Probab. Theory Related Fields* **86** 87–104.

HEYDE, C. C. and SENETA, E. (1972). Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Probab.* **9** 235–256.

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

KÜNSCH, H. R. (1987). Statistical aspects of self-similar processes. In *Proceedings of the First World Congress of the Bernoulli Society* (Yu. Prohorov and V. V. Sazanov, eds.) **1** 67–74. VNU Science Press, Utrecht.

LOBATO, I. and ROBINSON, P. M. (1994). The averaged periodogram estimate of long memory. *J. Econometrics.* To appear.

ROBINSON, P. M. (1994a). Time series with strong dependence. In *Advances in Econometrics. Sixth World Congress* (C. A. Sims, ed.) **1** 47–95. Cambridge Univ. Press.

ROBINSON, P. M. (1994b). Semiparametric analysis of long-memory time series. *Ann Statist.* **22** 515–539.

ROBINSON, P. M. (1994c). Rates of convergence and optimal bandwidth in spectral analysis of processes with long range dependence. *Probab. Theory Related Fields* **99** 443–473.

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

TAQQU, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrsch. Verw. Gebiete* **31** 287–302.
TOUSSOUN, O. (1925). Memoire sur l'histoire du Nil. *Memoires de l'Institut d'Egypte (Cairo)* **9**.
YONG, C. H. (1974). *Asymptotic Behaviour of Trigonometric Series*. Chinese Univ. Hong Kong.
ZYGMUND, A. (1977). *Trigonometric Series*. Cambridge Univ. Press.

LONDON SCHOOL OF ECONOMICS
HOUGHTON STREET
LONDON WC2A 2AE
ENGLAND