The Annals of Statistics
1995, Vol. 23, No. 5, 1684–1711

PERIODOGRAM-BASED ESTIMATORS OF FRACTAL PROPERTIES

BY GRACE CHAN, PETER HALL[1] AND D. S. POSKITT

Australian National University

We suggest an estimator, based on the periodogram, of the fractal index and fractal dimension of a continuous, stationary Gaussian process. We argue that the cosine part of the periodogram is more appropriate than the full periodogram for this application. The term "semiperiodogram" is used to describe the cosine component, and our estimator is based on simple linear regression of the logarithm of the semiperiodogram on the algorithm of frequency. Theoretical properties of the estimator, including its bias, variance and asymptotic distribution, are derived. Consistency is possible using only a small trace of the process, recorded over a fixed interval. We do not need to model the covariance function parametrically, and assume only mild conditions on the behaviour of the covariance in the neighbourhood of the origin. The issue of aliasing is discussed in both theoretical and numerical terms, and the numerical properties of the estimator are assessed in a simulation study.

1. Introduction.

1.1. *General motivation.* During the last decade there has been considerable interest in modelling real phenomena using self-similar processes. One particularly popular application is to the problem of describing surfaces. For example, the surface of a fractured metal bar or of a highly magnified polished metal sheet may be modelled in this way. Literature inspired by problems such as this includes Berry and Hannay (1978), Coster and Chermant (1983), Mandelbrot, Passoja and Paullay (1984), Thomas and Thomas (1988), Ogata and Katsura (1991), Taylor and Taylor (1991) and Constantine and Hall (1994). The reader is referred to Carter, Cawley and Mauldin (1988), Dubuc, Quiniou, Roques-Carmes, Tricot and Zucker (1989) and Dubuc, Zucker, Tricot, Quiniou and Wehbi (1989) for discussion of a variety of ways of estimating fractal properties in an engineering context. Often the data obtained from surfaces are in the form of line transect samples, and so the self-similar process under study is a curve in the plane rather than a surface in three-dimensional space. In this context, some recent developments in time series analysis are of interest, namely, work that has focused attention on processes exhibiting long range dependence. Such processes are characterised by a power spectrum proportional to an inverse

Received November 1992; revised August 1993.
[1]Also affiliated with CSIRO Division of Mathematics and Statistics, Sydney, Australia.
AMS 1991 *subject classifications.* Primary 62M10, 62G05; secondary 62E10.
Key words and phrases. Aliasing, bias, covariance, fractal dimension, fractal index, frequency, Gaussian process, periodogram, regression, self-similar process, semiperiodogram.

power law near the origin and have covariances that decay slowly; examples are fractional Brownian motion, due to Mandelbrot and Van Ness (1968), and fractional differences, introduced by Granger and Joyeux (1980) and Hosking (1981), both of which have been used to model strongly dependent behaviour in economics, geophysics and hydrology. Various approaches to the estimation of parameters of interest in strongly dependent processes have been proposed, including the rescaled range and frequency domain methods. The former are discussed in Mandelbrot (1975) and Mandelbrot and Taqqu (1979), while Fox and Taqqu (1986), Geweke and Porter-Hudak (1983), Janacek (1982) and Kashyap and Eom (1988) consider the latter. Cox (1984) and Hampel (1987) provide general reviews of many of these ideas. In this paper we are concerned with the analysis of line transect samples, but because we have in mind applications that involve examining the smoothness of surfaces, our focus of attention centres not on the long range properties of the process, but on its local behaviour. As will be seen below, although there are analogies with the investigation of long range dependence, this change of emphasis introduces several subtle and important differences.

1.2. *Relationship between fractal and covariance properties.* If the self-similar process under study is Gaussian, then the fractal properties of its sample paths may be described very simply in terms of the behaviour of its covariance function at the origin. For example, if the stationary Gaussian process X has covariance given by $\gamma(s - t) = \text{cov}\{X(s), X(t)\}$ and if

$$\gamma(t) = \gamma(0) - c|t|^{\alpha} + o(|t|^{\alpha}), \tag{1.1}$$

as $t \to 0$, where $0 < \alpha \leq 2$ and $c > 0$, then α is called the fractal index of X. The fractal or Hausdorff dimension of X, equivalent also (on the present occasion) to capacity, is given by

$$D = 2 - \tfrac{1}{2}\alpha. \tag{1.2}$$

See, for example, Adler [(1981), Chapter 8]. When $\alpha = 2$, the process X is differentiable with probability 1 and so $D = 2 - \frac{1}{2} \cdot 2 = 1$, as expected. For smaller values of α, the fractal dimension is fractional and lies strictly between 1 and 2. The relationship (1.2) between fractal dimension and fractal index persists for a variety of non-Gaussian processes, but not for all such processes. This issue will be discussed further in Section 1.4.

1.3. *Summary.* In the present paper we suggest a technique for estimating fractal index, and hence [using formula (1.2)] fractal dimension of the sample paths of a stationary Gaussian process, based on the periodogram. We assume that the data are in the form of a trace of the process in the continuum. The trace need not be a long one. Indeed, our asymptotic theory is based on a trace over a fixed interval, which we take without loss of generality to be $(-1, 1)$. Issues of aliasing, that is, of discretizing a continuous trace on a grid, are discussed both theoretically and numerically. We show

that the cosine part of the periodogram has certain advantages over the full periodogram in the context of of estimating α.

We should stress that all our theory is conducted under the nonparametric model (1.1), rather than a fully parametric model such as $\gamma(t) = \exp(-c|t|^\alpha)$. It is easy to show that an estimator based on the latter assumption is inconsistent if the model is misspecified.

Section 2 introduces our estimator and derives its basic properties, including variance, bias and asymptotic distribution. Section 3 discusses numerical issues and summarises the result of a simulation study. Proofs of results in Section 2 are deferred to Section 4.

1.4. *Engineering aspects of fractal estimation.* Surface data of a fractal nature may be gathered in a very wide variety of ways. All involve a certain amount of smoothing or degradation of the "true" surface data. The nature and extent of this smoothing depends very much on the method of recording, of which we shall mention only a few. A stylus, rather like a gramophone needle, may be drawn across the surface, its oscillations magnified electronically and the resulting height measurements recorded at regular intervals. The shape, and in particular the width, of the stylus influence the extent of smoothing, although sometimes there are other significant factors, such as complex electronic filters in the recording device. The type of smoothing offered by a stylus is distinctly nonlinear, which may be seen by noting that the stylus cannot record any information at all from crevices which are narrower than the stylus width. Relatively linear smoothing is provided by devices which measure surface height via the interference patterns of reflected light from a narrow beam. Often the beam is from a laser, but sometimes it is white light. The diameter of the light beam, and again aspects of the recording equipment, determine the extent of smoothing. Third, two-dimensional data may be recorded on a camera pixel grid. Fractal data obtained by electron microscopy are often of this form. Here, properties of the camera determine the nature of smoothing, which is often approximately uniform over pixels, but occasionally is very close to point measurements at pixel centres.

The effects of smoothing are related to those of discretizing, or gridding. They should be acknowledged in some way by the manner of statistical analysis. In the context of periodogram-based estimation of fractal dimension, smoothing provides yet another factor which limits the size of the frequency at which one may estimate the periodogram (or semiperiodogram). The extent of such limitations is admittedly easier to see when using methods such as box-counting or the variogram, where calculations are based directly on differencing the process at grid points. Those points should be paired at distances which are at least several stylus widths or light beam widths or pixel widths apart. A concise mathematical account of some of the effects of smoothing, in the context of recording fractal surfaces, has been given by Hall (1995).

Virtually all methods for estimating fractal dimension are based on so-called scaling laws, which involve seeking approximately linear relationships between the logarithm of a certain estimable function and the logarithm of its argument. There are many possible choices for the function, of which some will be discussed shortly. Exceptions to this rule included the stochastic modelling of concave or convex departures from linear relationships, for example, modelling the effect of data degradation by smoothing. Other matters, related to discontinuities and oscillations, may be addressed at least in part by modelling arguments. There also exist scaling laws based on fitting a linear relationship to the *iterated* logarithm of a function and the logarithm of its argument. See, for example, Ling (1987). However, while these laws have been discussed in the context of fractal properties, they do not seem to be efficacious in actually estimating those properties.

The scaling law based on the (semi)periodogram is the subject of this paper, and will be discussed at length in the next section. To place it into context, we mention here two other laws, based on the length of an approximating polygonal path and on the variogram. The first of these is arguably the "classical" scaling law, very commonly associated with physical properties. It declares that if (1.2) holds, then the length $l(s)$ of a polygonal approximation to a fractal curve, constructed on a grid of edge width s, should vary in such a way that $\log l(s)$ increases like $(D - 1)|\log s|$ as s decreases:

$$\log l(s) = (D - 1)|\log s| + \text{constant} + o(1),$$

as $s \to 0$. For discussions of the application of this law to real data, the reader is referred to Ling (1987, 1989, 1990), Majumdar and Bhushan (1991) and Brown, Charles, Johnsen and Chester (1993).

The variogram scaling law is founded on the observation that if $v(s)$ equals the mean square of the difference between two values of a fractal process at points distant s apart, then in the presence of (1.2), $|\log v(s)|$ should decrease like $2(2 - D)|\log s|$ as s decreases:

$$|\log v(s)| = 2(2 - D)|\log s| + \text{constant} + o(1),$$

as $s \to 0$. Examples of the practical application of this law include Serra (1968), Delfiner and Delhomme (1975), Journel and Huijbregts (1978) and Burrough (1981).

The power spectrum method, discussed, for example, by Dubuc, Quiniou, Roques-Carmes, Tricot and Zucker (1989), is closely related to estimation via the periodogram. However, it is tailored specifically to discretized data and so requires an adjustable level of smoothing to estimate the spectrum at all frequencies. We have deliberately avoided such an approach here, since incorporation of a smoothing parameter would considerably complicate our theoretical analysis. Thus, we consider the twin issues of discretizing and statistical smoothing (as distinct from the unwanted smoothing of the data during recording) from the viewpoint of aliasing.

An alternative approach to inference, quite different from those discussed above, is based on modelling the covariance function in greater detail. This can take a wide range of forms. At one extreme, the second-order term in (1.1), that is, the $o(|t|^\alpha)$ term, may be modelled. This amounts to modelling one aspect of the departure of the scaling law from a straight line. Alternatively, one may model the entire function γ, for example, by $\gamma(t) = \exp(-c|t|^\alpha)$. Philosophically we find ourselves uncomfortable with such an approach. It relates the long-range dependence of the process to its fractal properties in a way which is rather difficult to substantiate without extensive information. Of course, more complex models are possible, which allow more flexible relationships between long-range dependence and fractal behaviour. However, with more complex modelling comes greater difficulty of conducting inference. Nevertheless, completely parametric models have been used very successfully in some settings; see, for example, Ogata and Katsura (1991).

The assumption that the underlying process is Gaussian is of course not completely necessary to the estimation of fractal dimension. However, some restriction of this sort seems to be required. Virtually all scaling laws derive from the relationship between fractal and covariance properties discussed in Section 1.2. This is true even for scaling laws which do not explicitly involve covariance; for example, the length-based law discussed earlier is of this form. Therefore, any stochastic model must be capable of supporting a connection between the fractal dimension of sample paths and second-order moments. This is a tall order, and is usually only available for processes which are closely related to Gaussian. Even among such processes there are definite limitations to the validity of key elements of the scaling laws, such as formula (1.2). For example, if $X = |Y|^\nu$, where Y is a stationary Gaussian process, then (1.2) holds if and only if $\nu \geq \frac{1}{2}$. See Hall and Roy (1994).

2. Methodology and basic properties.

2.1. *Introduction and summary.* Our estimator of fractal index is based on a version of the periodogram for a continuous trace of a stochastic process X. The definition of the continuum periodogram I is similar to that of the periodogram computed from discrete observations of X, and is given in Section 2.2. It is composed of both cosine and sine parts. We define the semiperiodogram J to be the cosine part of the periodogram. Section 2.2 provides asymptotic formulae for expected values of both the cosine and sine components of I, assuming a regularity condition similar to (1.1). Those results imply that, when $\alpha > 1$, the value of α cannot be estimated directly from the periodogram, but that the semiperiodogram can be employed successfully whenever $0 < \alpha < 2$. Therefore we focus attention on the semiperiodogram, and define an estimator $\hat{\alpha}$ based on regression of $\log J$ on the logarithm of frequency.

Bias and variance of $\hat{\alpha}$ are discussed in Sections 2.3 and 2.4, respectively. The size of bias depends critically on behaviour of the "$o(|t|^\alpha)$" term in (1.1), and so that expansion must be elaborated upon if we are to describe bias.

Section 2.5 discusses the overall performance of $\hat{\alpha}$, explaining concisely how the accuracy of $\hat{\alpha}$ is affected simultaneously by both deterministic errors and stochastic fluctuations. We present a central limit theorem for the estimator and address the issue of recording the data on a fine grid rather than in the continuum. This matter of "aliasing," as it is often called, can influence the accuracy of our theoretical description of the properties of $\hat{\alpha}$. We state an explicit condition, on grid width, that is sufficient to ensure that aliasing does not affect the central limit theorem for $\hat{\alpha}$.

Proofs of the results in Sections 2.2–2.5 are given in Section 4.

2.2. *Methodology.* Let X denote a stationary Gaussian process observed on the interval $(-1, 1)$ and define

$$A(\omega) = \int_{-1}^{1} X(t)\cos(\omega t)\, dt, \qquad B(\omega) = \int_{-1}^{1} X(t)\sin(\omega t)\, dt,$$

$$I(\omega) = A(\omega)^2 + B(\omega)^2, \qquad J(\omega) = A(\omega)^2.$$

We call I the periodogram and J the semiperiodogram. Note that $I(\omega) = I(-\omega)$ and $J(\omega) = J(-\omega)$. If ω is an integer multiple of 2π, then, regardless of the mean of X, $EA(\omega) = EB(\omega) = 0$. Henceforth we suppose that ω is a positive integer multiple of 2π.

The theorem below states that if $0 < \alpha < 2$, then $EJ(\omega) \sim$ constant $\times$ $\omega^{-(\alpha+1)}$ as $\omega \to \infty$, and that the same is true of $EI(\omega)$ provided $0 < \alpha \le 1$. These properties form the basis of our approach to estimating α. To state the theorem we must introduce regularity conditions, which we do next.

Given $\xi > 0$, let $\Lambda(\xi)$ denote the class of functions h on $(0, 2)$ such that

$$\sup_{|u| \le \delta} \int_0^2 |h(t+u) - h(t)|\, dt = O(\delta^{\xi}),$$

as $\delta \to 0$. [We extend h from $(0, 2)$ to $(-\infty, \infty)$ by periodicity.] Note that if $0 \le \eta < 1$, the function $h(t) \equiv t^{-\eta}$ is in $\Lambda(1 - \eta)$. Given a real-valued function g, let $g', g'', \ldots$ or equivalently $g^{(1)}, g^{(2)}, \ldots$ denote its successive derivatives. Define g_1 by

$$\gamma(t) = \gamma(0) - ct^{\alpha} + g_1(t), \qquad t > 0, \tag{2.1}$$

for constants $c > 0$ and $0 < \alpha < 2$ [as in (1.1)]. We assume that, for some $\varepsilon > 0$,

$$g_1' \in \Lambda(\alpha + \varepsilon), \quad \text{if } \alpha < 1, \quad \text{or} \quad g_1'' \in \Lambda(\alpha + \varepsilon - 1) \quad \text{if } 1 \le \alpha < 2. \tag{2.2}$$

For $\delta > 0$, put

$$K_{\left\{\begin{smallmatrix}\sin\\ \cos\end{smallmatrix}\right\}}(-\delta) = \int_0^{\infty} t^{-\delta} \left\{\begin{smallmatrix}\sin\\ \cos\end{smallmatrix}\right\}(t)\, dt,$$

respectively. Here, the sine and cosine functions should be taken respectively.

THEOREM 2.1. *Assume condition* (2.2). *Then, as* $\omega \to \infty$ *through integer multiples of* 2π,

$$E\{A(\omega)^2\} = -\omega^{-1}\int_0^2 (2-t)\gamma'(t)\sin(\omega t)\,dt$$

$$\sim \begin{cases} 2c\alpha\omega^{-(\alpha+1)}K_{\sin}(\alpha-1), & \text{if } 0<\alpha<1, \\ 2c\omega^{-2}, & \text{if } \alpha=1, \\ 2c\alpha(\alpha-1)\omega^{-(\alpha+1)}K_{\cos}(\alpha-2), & \text{if } 1<\alpha<2; \end{cases} \tag{2.3}$$

$$E\{B(\omega)^2\} = -\omega^{-1}\int_0^2 (2-t)\gamma'(t)\sin(\omega t)\,dt$$

$$+ 2\omega^{-2}\int_0^2 \gamma'(t)\cos(\omega t)\,dt + 2\omega^{-2}\{\gamma(0)-\gamma(2)\}$$

$$\sim \begin{cases} 2c\alpha\omega^{-(\alpha+1)}K_{\sin}(\alpha-1), & \text{if } 0<\alpha<1, \\ 2\{c+\gamma(0)-\gamma(2)\}\omega^{-2}, & \text{if } \alpha=1, \\ 2\{\gamma(0)-\gamma(2)\}\omega^{-2}, & \text{if } 1<\alpha<2. \end{cases} \tag{2.4}$$

In view of the theorem, $\log EI(\omega) \sim -(\alpha+1)\log\omega$, for $0<\alpha\le 1$, but $\log EI(\omega) \sim -2\log\omega$, for $1<\alpha<2$. Thus, in the case $\alpha>1$ we cannot expect to be able to estimate α in terms of the slope of the regression of $\log I(\omega)$ on $\log\omega$. On the other hand, $\log EJ(\omega) \sim -(\alpha+1)\log\omega$ for $0<\alpha<2$, so that regression of $\log J(\omega)$ on $\log\omega$ is practicable for estimating α in this wider range. Therefore we focus attention on the semiperiodogram.

Let $\sigma(\omega)^2 = EA(\omega)^2$, $Z(\omega) = A(\omega)/\sigma(\omega)$, $Y(\omega) = \log A(\omega)^2$ and

$$e(\omega) = \log Z(\omega)^2 - E\log Z(\omega)^2, \qquad C_1 = E\log Z(\omega)^2.$$

Since $Z(\omega)$ has the standard normal distribution for each ω, then C_1 does not depend on ω. In this notation,

$$Y(\omega) = \log\sigma(\omega)^2 + C_1 + e(\omega).$$

Formula (2.3) may be written as $\sigma(\omega)^2 \sim C_2\omega^{-(\alpha+1)}$, where $C_2>0$. This observation prompts the approximate regression model,

$$Y(\omega) \simeq C_3 - (\alpha+1)\log\omega + e(\omega),$$

where $C_3 = C_1 + \log C_2$. That in turn suggests the following approach to estimating α. Let $\omega_1 < \cdots < \omega_k$ denote positive integer multiples of 2π. Put $x_j = \log\omega_j$, $\bar{x} = k^{-1}\Sigma x_j$, $Y_j = Y(\omega_j)$ and

$$\hat{\alpha} = -\left\{\sum_{j=1}^k (x_j-\bar{x})^2\right\}^{-1}\sum_{j=1}^k (x_j-\bar{x})Y_j - 1.$$

The systematic and random errors of this estimator are described by bias and variance, respectively. We treat them separately. To simplify exposition

we take $w_j = 2jm\pi$, $1 \le j \le k$, where the integer $m = m(k) \ge 1$ may depend on k and could diverge to $+\infty$ as $k \to \infty$.

2.3. *Bias.* The size of bias is determined by the "$o(|t|^\alpha)$" term in (1.1), corresponding to g_1 in (2.1). In order to be explicit about bias it is necessary to enter into more detail about properties of g_1. With this in mind we shall refine (2.2) to the new condition (2.6) below.

Define g_2 by

$$\gamma(t) = \gamma(0) - ct^\alpha - dt^{\alpha+\beta} + d't^2 + g_2(t), \qquad t > 0, \tag{2.5}$$

for constants $c > 0, d, d'$, $0 < \alpha < 2$, $0 < \beta < 1$, $\beta \ne 2 - \alpha$. We assume that, for some $\varepsilon > 0$,

$$\begin{aligned} &g_2' \in \Lambda(\alpha + \beta + \varepsilon) \quad \text{if } \alpha + \beta < 1, \\ &g_2'' \in \Lambda(\alpha + \beta + \varepsilon - 1) \quad \text{if } 1 \le \alpha + \beta < 2, \\ &g_2''' \in \Lambda(\alpha + \beta + \varepsilon - 2) \quad \text{if } 2 < \alpha + \beta < 3. \end{aligned} \tag{2.6}$$

Put

$$M_\xi = \begin{cases} K_{\sin}(\xi - 1), & \text{if } 0 < \xi < 1, \\ 1, & \text{if } \xi = 1, \\ (\xi - 1)K_{\cos}(\xi - 2), & \text{if } 1 < \xi < 2, \\ -(\xi - 1)(\xi - 2)K_{\sin}(\xi - 3), & \text{if } 2 < \xi < 3 \end{cases} \tag{2.7}$$

(not defined for $\xi = 2$), $\tau_1 = dc^{-1}(1 + \alpha^{-1}\beta)M_\alpha^{-1}M_{\alpha+\beta}$, $\tau = -\tau_1\beta(1 - \beta)^{-2}(2\pi)^{-\beta}$.

THEOREM 2.2. *Assume condition* (2.6). *Then, as* $k \to \infty$,

$$E(\hat{\alpha}) = \alpha + \tau(km)^{-\beta} + o\{(km)^{-\beta}\}.$$

2.4. *Variance.* If the variables $Y_j = Y(\omega_j)$ were stochastically independent, then variance would be given by

$$\operatorname{var}(\hat{\alpha}) = \left\{ \sum_{j=1}^{k} (x_j - \bar{x})^2 \right\}^{-1} \operatorname{var}(Y_1).$$

Of course, the Y_j's are not independent, but we claim that, nevertheless, this variance formula is asymptotically correct, as evidenced by the following theorem.

THEOREM 2.3. *Assume condition* (2.2) *and that* $\gamma''(t) = O(t^{\alpha-3})$ *as* $t \downarrow 0$. *Then*

$$\operatorname{var}(\hat{\alpha}) \sim \left\{ \sum_{j=1}^{k} (x_j - \bar{x})^2 \right\}^{-1} \operatorname{var}(Y_1) \sim k^{-1} \operatorname{var}(Y_1). \tag{2.8}$$

We should comment on the fact that, in Theorem 2.3, the asymptotic variance of $\hat{\alpha}$ does not depend on m, although that quantity does appear in our asymptotic formula for bias; see Theorem 2.2. The fact that m does not enter the first-order asymptotics for $\text{var}(\hat{\alpha})$ is a consequence of our definition of the design points, $x_j = \log(2\pi jm)$. Note that m vanishes from the centred design points $x_j - \bar{x}$. For other choices of design, in particular for $\omega_j/2\pi$ equal to the integer part of $\exp\{(jm)^{\varepsilon}\}$, for some $\varepsilon > 0$, and $x_j = \log \omega_j$, the first asymptotic relation in (2.8) continues to hold, but the second is no longer true.

2.5. *Discussion.* The conditions of both Theorems 2.2 and 2.3 hold if we assume that, for some $\varepsilon > 0$, the function g_2 appearing in (2.5) satisfies

$$g_2^{(4)}(t) = O(t^{\alpha+\beta+\varepsilon-4}), \tag{2.9}$$

as $t \downarrow 0$. In the event that the conditions of both theorems hold, we may deduce from those results that, as $k \to \infty$,

$$\hat{\alpha} - \alpha = \tau(km)^{-\beta} + k^{-1/2}W_k + o_p\{(km)^{-\beta} + k^{-1/2}\},$$

where the random variable W_k has zero mean and variance $s^2 = \text{var}\{\log N(0,1)^2\}$. Thus, $\hat{\alpha} \to \alpha$ in probability as $k \to \infty$. If m is chosen so large that $k^{1-2\beta}m^{-2\beta} \to 0$ as $k \to \infty$ (e.g., if $m \geq 1$ in the case $\beta > \frac{1}{2}$), then $k^{1/2}(\hat{\alpha} - \alpha)$ has asymptotically zero mean and variance s^2. Quite generally, $k^{1/2}(\hat{\alpha} - E\hat{\alpha})$ has an asymptotic normal distribution, as our next result shows.

THEOREM 2.4. *Assume condition* (2.2) *and that* $\gamma^{(4)}(t) = O(t^{\alpha-4})$ *as* $t \downarrow 0$. *Then* $k^{1/2}(\hat{\alpha} - E\hat{\alpha})$ *is asymptotically normally distributed with zero mean and variance* s^2.

Condition (2.9) is sufficient for the assumptions of Theorem 2.4.

In practice, the integrals defining $A(\omega_1), \ldots, A(\omega_k)$, and hence $\hat{\alpha}$, would usually be approximated by a series computed on points of a grid. So we should state a result which points out that if the grid points are sufficiently closely spaced, then our asymptotic results continue to hold. In the discussion above we noted that $\hat{\alpha} - \alpha$ is at least as large as $k^{-1/2}$, and so it suffices to describe a series-based approximation to $\hat{\alpha}$ which is accurate to terms of smaller order than $k^{-1/2}$, as $k \to \infty$.

To this end, define

$$\hat{A}_n(\omega) = n^{-1} \sum_{i=-n}^{n-1} X(i/n)\cos(i\omega/n),$$

our approximation to $A(\omega)$ based on a $2n$-point grid. Let $\hat{\alpha}_n$ denote the corresponding approximation to $\hat{\alpha}$, computed with $\hat{A}_n(\omega_j)$ replacing $A(\omega_j)$ for $1 \leq j \leq k$. Our next theorem shows that if $n = n(k)$ increases sufficiently

rapidly, then $\hat{\alpha} - \hat{\alpha}_n = o(k^{-1/2})$, so that (for example) $\hat{\alpha}_n - E\hat{\alpha}$ satisfies the same central limit theorem as $\hat{\alpha} - E\hat{\alpha}$.

THEOREM 2.5. *Assume condition* (2.2) *and that* γ *is decreasing on an interval* $(0, \varepsilon)$, *for* $\varepsilon > 0$ *sufficiently small. Suppose* $n = n(k) \to \infty$ *so fast that*

$$n^{-\alpha}(mk)^{\alpha+1}k^5(\log k)^2 \to 0, \tag{2.10}$$

as $k \to \infty$. *Then* $k^{1/2}(\hat{\alpha} - \hat{\alpha}_n) \to 0$ *with probability* 1.

We suspect that condition (2.10) is substantially more severe than is necessary in practice, and our simulation study indicates this too. However, in strictly rigorous theoretical terms we have not been able to improve on it. The effect of discretizing the time variable, commonly called aliasing, will be taken up in more detail in the next section. It should be stressed that the problem is aggravated by the requirement that we examine the spectral density for large values of its argument, ω; for small values the effects of aliasing are much less. However, there seems no way to avoid this difficulty.

The cosine and sine parts of the periodogram are not individually shift invariant—that is, the random processes

$$A(\omega, \lambda) = \int_{-1}^{1} X(t)\cos\{\omega(t + \lambda)\}\, dt$$

and

$$B(\omega, \lambda) = \int_{-1}^{1} X(t)\sin\{\omega(t + \lambda)\}\, dt$$

are nondegenerate functions of λ as well as ω. (The full periodogram, $\{A(\omega, \lambda)^2 + B(\omega, \lambda)^2\}^{1/2}$, does not depend on λ.) This property may be used to enhance the performance of the estimator $\hat{\alpha}$, as follows. Let $\hat{\alpha}(\lambda)$ denote the same function of $A(\omega_1, \lambda), \ldots, A(\omega_k, \lambda)$ as $\hat{\alpha}$ was of $A(\omega_1), \ldots, A(\omega_k)$, let $\nu \geq 1$ be an integer, let $\lambda_1, \ldots, \lambda_\nu$ be arbitrary real numbers and put

$$\hat{\alpha}_0 = \nu^{-1} \sum_{j=1}^{\nu} \hat{\alpha}(\lambda_j).$$

The distribution of $\hat{\alpha}(\lambda)$ does not depend on λ, and so

$$E(\hat{\alpha}_0) = E(\hat{\alpha}), \qquad \operatorname{var}(\hat{\alpha}_0) \leq \operatorname{var}(\hat{\alpha}).$$

It follows that in mean-squared error terms, $\hat{\alpha}_0$ performs at least as well as $\hat{\alpha}$ and possibly a little better.

To appreciate the extent of any improvement offered by $\hat{\alpha}_0$, we must evaluate its variance. To this end, note that $A(\omega, \lambda_1)/\sigma(\omega)$ and $A(\omega, \lambda_2)/\sigma(\omega)$ have a bivariate normal distribution with zero means, unit variances and correlation coefficient equal to

$$\cos\{\omega(\lambda_1 - \lambda_2)\}$$

uniformly in λ_1 and λ_2, as $\omega \to \infty$. Let $g(\rho)$ denote the covariance of $(\log N_1^2, \log N_2^2)$ if (N_1, N_2) are bivariate normal $N(0,0;1,1,\rho)$. Then $g(\rho)$ depends only on $|\rho|$ and is strictly positive unless $\rho = 0$, in which case $g(\rho) = 0$. It may be shown that the asymptotic variance of $\hat{\alpha}_0$ equals

$$\left\{\sum (x_j - \bar{x})^2\right\}^{-2} \sum (x_j - \bar{x})^2 \nu^{-2} \sum_{l_1} \sum_{l_2} g\left[\cos\left\{\omega_j(\lambda_{l_1} - \lambda_{l_2})\right\}\right],$$

which for appropriate choice of ν $(\to \infty)$ and $\lambda_1, \ldots, \lambda_\nu$ is asymptotic to $k^{-1}\eta s^2$, where $0 < \eta < 1$.

The variance may be further reduced by using appropriate weights in the linear regression, but it may not be reduced to $o(k^{-1})$. The specific practical problem which motivated this work, that of assessing the roughness of rollers used to produce metal sheet [see Constantine and Hall (1994)], is one where additional data values may be derived with relatively little expense. Therefore we have not felt particularly motivated to discuss issues of statistical efficiency, in terms of choice of ν and the λ's. The issue of aliasing has been of much greater concern. However, the matter of statistical inconsistency, in connection with practical estimation of fractal dimension via the periodogram when $\alpha > 1$, has considerable practical interest and actually caused practical difficulties until it was well understood.

3. Numerical results.

3.1. *Sampling design and integral approximation.* Recall from Section 2.5 that our initial approximation of semiperiodogram $A(\omega)$ is defined by

$$\hat{A}_n(\omega) = n^{-1} \sum_{i=-n}^{n-1} X(i/n)\cos(i\omega/n). \tag{3.1}$$

This, of course, amounts to using a uniform sample design of $2n$ points on the interval $[-1, 1)$. It is known, however, that for the problem of approximating a weighted integral of a stochastic process over a finite interval, uniform sampling is not necessarily optimal; see Benhenni and Cambanis (1992) and the references contained therein. More generally, therefore, we may consider estimating $A(\omega)$ using $2n + 1$ observations of the process X at the points $t_{i,n}$, $i = -n, \ldots, n$, over the interval $[-1, 1]$ and an integral approximation of the form

$$A_n(\omega) = n^{-1} \sum_{i=-n}^{n} w_i X(t_{i,n})\cos(\omega t_{i,n}),$$

where the sample design $\{t_{i,n}\}$ and weights $\{w_i\}$ are chosen so as to minimise the asymptotic mean-squared error $E(A(\omega) - A_n(\omega))^2$; see Benhenni and Cambanis (1992). This choice depends on the process covariance structure, and in the simulations reported in Section 3.2 the underlying covariance function is assumed, for convenience, to be of the form

$$\gamma(s - t) = \exp(-c|s - t|^\alpha), \tag{3.2}$$

where $c > 0$ and $0 < \alpha < 2$. We shall therefore examine this case in a little more detail.

From the two differential equations given in (3.3) below it is straightforward to verify that the regularity conditions required by Benhenni and Cambanis for the determination of the optimal sample design are only applicable if we know $\alpha = 1$. In that case we find that $\{t_{i,n}\}$ is obtained by solving

$$\int_{-1}^{t_{i,n}} \kappa_\omega |\cos(\omega t)|^{2/3}\, dt = \frac{n+i}{2n}, \qquad i = -n, \ldots, n,$$

where the constant κ_ω is such that

$$\int_{-1}^{1} \kappa_\omega |\cos(\omega t)|^{2/3}\, dt = 1.$$

We have evaluated $\{t_{i,n}\}$ for various combinations of ω and n and in Table 1 present typical examples of the minimum and maximum values of $|t_{i,n} - i/n|$, $i = -n, \ldots, n$, for a range of ω values when $n = 500$.

Table 1 shows that only for small values of $\omega = 2\pi l$ will $|t_{i,n} - i/n|$ equal or exceed the length of the uniform sampling interval; for moderate and large values, the difference between the optimal and uniform sampling points will be less than n^{-1} by an order of magnitude at worst. Although these figures relate to the particular case $\alpha = 1$, it is clear that the close proximity of $\{t_{i,n}\}$ and $\{i/n\}$ for increasing ω is brought about by the increasingly rapid oscillations in $|\cos(\omega t)|^{2/3}$ over $[-1, 1]$ and it appears reasonable to conjecture that the behaviour of the cosine function would similarly dominate such calculations for other values of α, as well as more general covariance functions obeying (1.1), could they be made. This suggests that in general very little may be lost by using a uniform sample design. Indeed, in practice the true value of α is unknown and we envisage estimating it using a range of frequency values. Thus the estimation process will be more convenient and versatile if the same set of sampling points can be employed in all circumstances, as appears to be the case.

Once the sample design is known, the determination of the weighting sequence $\{w_i\}$ is governed by the continuity properties of X. Differentiating

TABLE 1
Absolute differences of optimal and uniform design points for $\omega = 2\pi l$, $n = 500$

ω	Minimum	Maximum
2π	2.20×10^{-8}	1.99×10^{-2}
20π	3.18×10^{-7}	2.13×10^{-3}
50π	3.63×10^{-7}	9.36×10^{-4}
100π	1.52×10^{-6}	5.48×10^{-4}
150π	9.07×10^{-7}	4.15×10^{-4}
200π	1.82×10^{-7}	3.37×10^{-4}

$\gamma(t)$ in (3.2) we obtain

$$\begin{aligned}\frac{\partial\gamma(t)}{\partial t} &= (-\operatorname{sgn} t)c\alpha|t|^{\alpha-1}\exp(-c|t|^{\alpha}),\\ \frac{\partial^2\gamma(t)}{\partial t^2} &= c\alpha\{c\alpha t^{2\alpha-2} - (\alpha-1)t^{\alpha-2}\}\exp(-c|t|^{\alpha}).\end{aligned} \tag{3.3}$$

Since all α's of interest are less than 2, these formulae confirm [Benhenni and Cambanis (1992), page 166] that the process has no quadratic mean derivatives. Given that the optimal sampling design is closely approximated by the uniform sequence $\{i/n\}$, particularly when $\omega = 2\pi l$ is large, following the arguments presented immediately above, this implies that the optimal weighting is likely to differ little from the trapezoidal rule. Hence, in the simulation experiments that follow we will consider the performance of the quasioptimal integral approximation

$$\tilde{A}_n(\omega) = n^{-1}\left[\{X(-1) + X(1)\}/2 + \sum_{i=-n+1}^{n-1} X(i/n)\cos(i\omega/n)\right].$$

With regard to the latter, it is perhaps of interest to observe that $\tilde{A}_n(\omega)$ may be viewed as a conventional approximation based on tapered data, and it is known that in the frequency domain analysis of time series, tapering can produce high resolution estimates with enhanced performance; see Dahlhaus (1988). As a basis for comparison we will therefore also investigate the alternative approximation

$$\check{A}_n(\omega) = n^{-1}\sum_{i=-n}^{n} X(i/n)\cos(i\omega/n)h(i/n),$$

where $h(\cdot)$ is the split cosine taper with cosine bell applied to 20% of the data. The properties of

$$\bar{A}_n(\omega) = n^{-1}\sum_{i=-n}^{n} X(i/n)\cos(i\omega/n)$$

and, following the suggestion of a referee,

$$\grave{A}_n(\omega) = n^{-1}\sum_{i=-n}^{n-1} X(i/n + 1/2n)\cos\{(i/n + 1/2n)\omega\}$$

will also be examined.

3.2. *Simulation study.* The purpose of this study is to illustrate the methodology outlined in Section 2 and to compare the performance of the integral approximations $\hat{A}_n(\omega)$, $\tilde{A}_n(\omega)$, $\bar{A}_n(\omega)$, $\check{A}_n(\omega)$ and $\grave{A}(\omega)$. In view of the theoretical results presented above, the experiments are based on the simulation of Gaussian processes whose covariance function is as given in

(3.2). The following values for the covariance parameters c and α were considered:

$$\alpha = 0.01, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75 \text{ and } 1.99,$$

$$c = \begin{cases} 0.1(2^{\alpha}), & \alpha \leq 1, \\ 2^{\alpha}, & \text{otherwise.} \end{cases}$$

The total number of sampling time points were taken to be 3751 ($n = 1875$) and 5001 ($n = 2500$). The first of these corresponds to an actual value used to analyse data from a metal surface and the second is used to demonstrate the improvement in estimation brought about by increasing n. The number of frequency values k was taken to be $[n/2]$ and $[(2n + 1)^{\rho}]$, $\rho = \frac{1}{3}, \frac{1}{2}, \frac{2}{3}$, for each n, $[x]$ denoting the integer part of x. The first value gives the Nyquist frequency and provides an upper bound on k. This value is used as a point of reference even though, according to condition (2.10) in Theorem 2.5, it is inappropriate. This condition implies that $n > k^{1+6/\alpha}$, for $0 < \alpha < 2$, and if $k = 25$, it would mean sampling the realization of the process at at least 2×10^{10} points. Clearly this is not practicable. The last three values for k were chosen in accord with the spirit of Theorem 2.5 while maintaining feasibility.

For each combination of values 100 replications of $2n + 1$ observations of the process X on the uniform grid $\{i/n\}$ were generated, each with zero mean, unit variance and $\text{cov}(X(i/n), X(j/n)) = \gamma\{(i - j)/n\}$, $i, j = -n, \ldots, n$. The realisations were generated using the simulation method described in Wood and Chan (1993). For each realisation, values of the estimates $\bar{D}_n$, $\hat{D}_n$, $\tilde{D}_n$, $\check{D}_n$ and $\grave{D}_n$ were then obtained, the different estimates denoting that $A(\omega_j)$, $j = 1, \ldots, k$, was replaced by $\bar{A}_n(\omega_j)$, $\hat{A}_n(\omega_j)$, $\tilde{A}_n(\omega_j)$, $\check{A}_n(\omega_j)$ and $\grave{A}_n(\omega_j)$, respectively, when computing the approximation to $\hat{D}$. The frequency spacing parameter m was fixed at unity in all cases. The empirical performance of the five estimators was summarised by calculating the observed bias (B), standard deviation (SD) and mean-squared error (MSE) in the usual manner, namely, given $\tilde{\alpha}_{n,1}, \ldots, \tilde{\alpha}_{n,100}$, say, the estimated values of the true α obtained from each replication, we put

$$\tilde{D}_{n,l} = 2 - \tilde{\alpha}_{n,l}/2, \qquad l = 1, \ldots, 100,$$

$$\tilde{D}_n = 100^{-1} \sum_{l=1}^{100} \tilde{D}_{n,l},$$

$$\text{B} = \tilde{D}_n - D,$$

$$\text{SD}^2 = 100^{-1} \sum_{l=1}^{100} \left(\tilde{D}_{n,l} - \tilde{D}_n\right)^2$$

and

$$\text{MSE} = \text{B}^2 + \text{SD}^2.$$

Some representative results are presented in Tables 2, 3 and 4. The rows labelled U and L give counts of the number of times the estimator of D

TABLE 2
Comparison among different approximations when $n = 1875$, $k = 61$

		$\bar{D}_n$	$\hat{D}_n$	$\tilde{D}_n$	$\check{D}_n$	$\grave{D}_n$
$D = 1.75$	B	0.006	0.022	0.003	−0.076	−0.011
	SD	0.136	0.143	0.153	0.183	0.130
	MSE	0.019	0.021	0.023	0.039	0.017
	U	5	7	5	4	5
	L	0	0	0	0	0
$D = 1.5$	B	−0.003	−0.012	−0.014	−0.095	−0.002
	SE	0.144	0.146	0.138	0.189	0.156
	MSE	0.021	0.021	0.019	0.045	0.024
	U	0	0	0	0	0
	L	0	0	0	2	0
$D = 1.25$	B	−0.022	−0.039	−0.018	−0.061	−0.024
	SD	0.128	0.147	0.139	0.149	0.165
	MSE	0.017	0.023	0.020	0.026	0.028
	U	0	0	0	0	0
	L	5	9	6	11	5

produced a value greater than 2 and less than 1, respectively. If the estimate of D fell outside the interval $[1, 2]$, it was reset to the nearest value within the interval. The estimates of bias, standard deviation and mean-squared error include such reset values.

Table 2 presents the observed outcomes obtained using the different integral approximations when $D = 1.75, 1.5, 1.25$ ($\alpha = 0.5, 1, 1.5$), $n = 1875$

TABLE 3
Comparison between different values of n and k for $\tilde{D}_n$

		$n = 1875$ $k = 15$	**$n = 2500$ $k = 17$**
$D = 1.75$	B	−0.044	−0.002
	SD	0.278	0.238
	MSE	0.065	0.057
	U	37	24
	L	0	0
$D = 1.5$	B	−0.026	0.012
	SD	0.300	0.292
	MSE	0.090	0.085
	U	8	4
	L	11	5
$D = 1.25$	B	0.060	−0.001
	SD	0.285	0.254
	MSE	0.085	0.064
	U	4	2
	L	22	30

TABLE 4
Comparison among different values of k when n = 1875 for $\tilde{D}_n$

		$k = 937$	$k = 241$	$k = 61$	$k = 15$
$D = 1.625$	B	0.121	0.031	0.001	−0.082
	SD	0.046	0.078	0.151	0.308
	MSE	0.017	0.007	0.023	0.102
	U	0	0	1	10
	L	0	0	0	7
$D = 1.375$	B	0.088	0.006	0.002	−0.022
	SD	0.036	0.079	0.193	0.312
	MSE	0.009	0.006	0.037	0.098
	U	0	0	0	6
	L	0	0	1	21

and $k = [3751^{1/2}] = 61$. It indicates that although there is no estimator that dominates the others, the top two rankings are invariably shared between $\bar{D}_n$, $\tilde{D}_n$ and $\grave{D}_n$, whatever the performance criterion. A more detailed analysis of the relationships between the performance measures and the interaction between n and k is provided in Tables 3 and 4. For the moment we simply note that although a choice between these three estimators is problematic, on the basis of these and other results not reported here, we would express a preference for $\tilde{D}_n$.

The results given in Table 3 relate to the behaviour of the estimator $\tilde{D}_n$ when $D = 1.75, 1.5, 1.25$ and $(n, k) = (n, [2n + 1]^{1/3}) = (1875, 15), (2500, 17)$. They show that increasing n can improve the estimation of D, even though almost the same value of k is being employed. This presumably reflects the increasing accuracy of the integral approximation as n is increased and the reduced confounding brought about by using a finer grid.

The figures given in Table 4 show how the performance of the estimator $\tilde{D}_n$ changes as k is varied. They relate to the situation where $D = 1.625, 1.375$ ($\alpha = 0.75, 1.25$), $n = 1875$ and $k = [1875/2] = 937$ and $k = [3751^{\rho}]$, $(\rho = \frac{2}{3}, \frac{1}{2}, \frac{1}{3}) = 241, 61, 15$. From the table it is clear that the standard deviation is reduced by about one-half as k is increased by roughly a factor of 4. It is also apparent that a reduction in bias can be obtained by increasing k. Such results are to be anticipated, of course, from Theorems 2.2 and 2.3, but the improvement achieved is neither harmonic nor monotonic, particularly for the bias. This suggests that although $k = [(2n + 1)^{2/3}]$, like $k = [n/2]$, may be too large in relation to n, coarser discretizations than would be allowed under condition (2.10) of Theorem 2.5 may be efficacious. As intimated previously, sampling at a rate corresponding to $n > k^{1+6/\alpha}$, $1 \leq \alpha \leq 2$, as implied by (2.10) is not claimed to be optimal in any way and choosing $n > k^q$, $2 \leq q < 4$, seems to provide a fairly effective practical rule.

Finally we also investigated three processes with fractal dimension $D = 1.6, 1.5, 1.4$ ($\alpha = 0.8, 1, 1.2$) in order to compare our methodology with the one-dimensional variation method introduced by Dubuc, Quiniou, Roques-

Carmes, Tricot and Zucker (1989). Using $\hat{D}_n$ in conjunction with $k = [16{,}383^{1/2}] = 127$, $n = 8191$, we found that the largest observed bias was 0.002, which compares favourably with the smallest bias of 0.005 reported by Dubuc, Quiniou, Roques-Carmes, Tricot and Zucker (1989) for a similar exercise.

It goes without saying that extrapolating from limited and specialised Monte Carlo experiments is fraught with dangers. Nevertheless, the results we have obtained lend support to the following tentative conclusions: the reduction in both bias and variance, and hence mean-squared error, produced by increasing the number of frequency values and associated sampling time points predicted by our theoretical results can often be quite considerable. The estimator $\tilde{D}_n$ based on the quasioptimal integral approximation $\tilde{A}_n(\omega)$ performs well relative to the alternatives. Additionally, the semiperiodogram regression-based technique seems to provide a reasonable procedure for estimating fractal dimension.

4. Proofs.

4.1. *Proofs of Theorems* 2.1–2.3. Our proofs rely on the following lemmas. Lemma 4.1 is taken from Zygmund [(1959), page 46], Lemma 4.2 may be derived by elementary calculus after making the change of variable $u = \omega t$ and Lemma 4.3 follows from Euler–Maclaurin–type integral approximations to series; see Abramowitz and Stegun [(1965), page 886].

LEMMA 4.1. *If* $h \in \Lambda(\xi)$, *for some* $0 < \xi \leq 1$, *then if* $g(x)$ *denotes either* $\sin x$ *or* $\cos x$,

$$\int_0^2 h(t)g(\omega t)\,dt = O(\omega^{-\xi}),$$

as $\omega \to \infty$ *through integer multiples of* π.

LEMMA 4.2. *Let* $0 < \xi < 3$, *define*

$$L_\xi(\omega) = \int_0^2 (2 - t)t^{\xi - 1}\sin(\omega t)\,dt,$$

and (for $\xi \neq 2$*) let* M_ξ *be given by* (2.7). *In the case* $\xi = 2$ *and for* ω *an integer multiple of* π, *we have* $L_\xi(\omega) = 0$. *When* $0 < \xi < 3$ *and* $\xi \neq 2$, *we have as* $\omega \to \infty$ *through integer multiples of* π,

$$L_\xi(\omega) = 2M_\xi\omega^{-\xi} + O(\omega^{-(\xi+1)}).$$

LEMMA 4.3. *Assume that* $0 < \beta < 1$. *Then as* $k \to \infty$,

$$k^{-1}\sum_{j=1}^{k} j^{-\beta}\log j - \left(k^{-1}\sum_{j=1}^{k} j^{-\beta}\right)\left(k^{-1}\sum_{j=1}^{k}\log j\right)$$
$$= -\beta(1 - \beta)^{-2}k^{-\beta} + O(k^{-1}\log k),$$

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

To obtain the initial equation in (2.3), observe that

$$\begin{aligned}
E\{A(\omega)^2\} &= \int_{-1}^{1}\int_{-1}^{1} \gamma(t_1 - t_2)\cos(\omega t_1)\cos(\omega t_2)\, dt_1\, dt_2 \\
&= -\omega^{-1}\int_{-1}^{1} \cos(\omega t_2)\, dt_2 \int_{-1}^{1} \gamma'(t_1 - t_2)\sin(\omega t_1)\, dt_1 \\
&= -(2\omega)^{-1}\int_{-1}^{1}\int_{-1}^{1} \gamma'(t_1 - t_2)\big[\sin\{\omega(t_1 - t_2)\} \\
&\qquad\qquad + \sin\{\omega(t_1 + t_2)\}\big]\, dt_1\, dt_2 \\
&= -(2\omega)^{-1}\int_{-2}^{2} \gamma'(t)\, dt \\
&\quad \times \int_{(-1)\vee(-1-t)}^{1\wedge(1-t)} \big[\sin(\omega t) + \sin\{\omega(t + 2t_2)\}\big]\, dt_2 \\
&= -\omega^{-1}\int_{0}^{2} \gamma'(t)\, dt \int_{-1}^{1-t} \big[\sin(\omega t) + \sin\{\omega(t + 2t_2)\}\big]\, dt_2 \\
&= -\omega^{-1}\int_{0}^{2} \gamma'(t)(2 - t)\sin(\omega t)\, dt.
\end{aligned}$$

A similar sequence of identities gives the first equation (2.4). The asymptotic relations then follow from Lemmas 4.1 and 4.2. This completes the proof of Theorem 2.1.

To derive Theorem 2.2, recall that $\omega_j = 2jm\pi$, $x_j = \log \omega_j$. Put $\overline{\log k} = k^{-1}\sum \log j$ and observe that by Lemma 4.3,

$$\sum_{j=1}^{k} (x_j - \bar{x})^2 = \sum_{j=1}^{k} (\log j - \overline{\log k})^2 = k + O\{(\log k)^2\}, \tag{4.1}$$

$$\begin{aligned}
\sum_{j=1}^{k} (x_j - \bar{x})\omega_j^{-\beta} &= (2\pi m)^{-\beta} \sum_{j=1}^{k} (\log j - \overline{\log k}) j^{-\beta} \\
&= -\beta(1 - \beta)^{-2}(2\pi m)^{-\beta} k^{1-\beta} + O(m^{-\beta}\log k).
\end{aligned} \tag{4.2}$$

More simply, if $0 < \beta + \varepsilon < 1$, then

$$\sum_{j=1}^{k} |x_j - \bar{x}|\omega_j^{-\beta-\varepsilon} = O(m^{-\beta-\varepsilon} k^{1-\beta-\varepsilon}\log k). \tag{4.3}$$

By (2.5),

$$\begin{aligned}
\sigma(\omega)^2 &= -\omega^{-1}\int_{0}^{2} (2 - t)\gamma'(t)\sin(\omega t)\, dt \\
&= -\omega^{-1}\int_{0}^{2} (2 - t)\{-c\alpha t^{\alpha-1} - d(\alpha + \beta)t^{\alpha+\beta-1} + 2d't\}\sin(\omega t)\, dt \\
&\quad - \omega^{-1}\int_{0}^{2} (2 - t) g_2'(t)\sin(\omega t)\, dt.
\end{aligned}$$

In view of Lemma 4.2, the first term on the last right-hand side equals

$$2\omega^{-1}\{c\alpha M_\alpha \omega^{-\alpha} + d(\alpha+\beta)M_{\alpha+\beta}\omega^{-(\alpha+\beta)} + O(\omega^{-(\alpha+1)})\}.$$

By Lemma 4.1, the second term equals $O(\omega^{-(\alpha+\beta+\varepsilon+1)})$, provided $\varepsilon > 0$ is sufficiently small. (Integrations by parts are necessary if $\alpha + \beta \geq 1$.) Therefore,

$$\sigma(\omega)^2 = 2c\alpha M_\alpha \omega^{-(\alpha+1)}\{1 + \tau_1\omega^{-\beta} + O(\omega^{-(\beta+\varepsilon)})\},$$

$$\log\sigma(\omega)^2 = \log(2c\alpha M_\alpha) - (\alpha+1)\log\omega + \tau_1\omega^{-\beta} + O(\omega^{-(\beta+\varepsilon)} + \omega^{-2\beta}),$$

$$\begin{aligned} E(\hat{\alpha}) - \alpha &= -\left\{\sum_{j=1}^k (x_j - \bar{x})^2\right\}^{-1} \sum_{j=1}^k (x_j - \bar{x})\log\sigma(\omega_j)^2 - (\alpha+1) \\ &= -\left\{\sum_{j=1}^k (x_j - \bar{x})^2\right\}^{-1} \sum_{j=1}^k (x_j - \bar{x})\{\tau_1\omega_j^{-\beta} + O(\omega_j^{-(\beta+\varepsilon)} + \omega_j^{-2\beta})\} \\ &= \tau_1\beta(1-\beta)^{-2}(2\pi km)^{-\beta} + o\{(km)^{-\beta}\}, \end{aligned}$$

the last line following from (4.1)–(4.3). This completes the proof of Theorem 2.2.

As a prelude to establishing Theorem 2.3, we derive the following lemma. Let ω_1, ω_2 denote positive integer multiples of 2π.

LEMMA 4.4. *Under the conditions of Theorem* 2.3,

$$\begin{aligned} &\operatorname{cov}\{A(\omega_1), A(\omega_2)\} \\ &\quad = 2(\omega_1^2 - \omega_2^2)^{-1}\int_0^2 \gamma'(t)\{\cos(\omega_2 t) - \cos(\omega_1 t)\}\,dt. \end{aligned} \tag{4.4}$$

Furthermore, there exists a constant $C > 0$ *such that, for* $\omega_1 < \omega_2$,

$$|\operatorname{cov}\{A(\omega_1), A(\omega_2)\}| \leq C(\omega_2 - \omega_1)^{-1}\omega_1^{-\alpha}\omega_2^{-1}, \tag{4.5}$$

$$\begin{aligned} |\operatorname{cov}\{A(\omega_1), A(\omega_2)\}| \leq C\omega_1^{-2}\{&\omega_2^{-\alpha}I(\alpha<1) \\ &+ \omega_2^{-1}(\log\omega_2)I(\alpha=1) \\ &+ \omega_2^{-1}I(\alpha>1)\}. \end{aligned} \tag{4.6}$$

PROOF. Formula (4.4) may be obtained by change of variable in a double integral. To derive (4.5), note that when $\alpha \leq 1$, using Lemmas 4.1 and 4.2 and condition (2.2),

$$\int_0^2 \gamma'(t)\cos(\omega t)\,dt = O(\omega^{-\alpha}),$$

and when $\alpha > 1$, by the same argument,

$$\int_0^2 \gamma'(t)\cos(\omega t)\,dt = -\omega^{-1}\int_0^2 \gamma''(t)\sin(\omega t)\,dt = O(\omega^{-\alpha}).$$

Result (4.5) is now immediate from (4.4).

To establish (4.6), observe that for $0 < \alpha < 2$,

$$K(\omega_1, \omega_2) \equiv \int_0^2 \gamma'(t)\{\cos(\omega_2 t) - \cos(\omega_1 t)\}\, dt$$

$$= (\omega_1\omega_2)^{-1}(\omega_2 - \omega_1)\int_0^2 \gamma''(t)\sin(\omega_2 t)\, dt$$

$$-(\omega_1^2\omega_2)^{-1}(\omega_2 - \omega_1)\int_0^2 \gamma'''(t)\{1 - \cos(\omega_2 t)\}\, dt$$

$$+(\omega_1\omega_2)^{-2}(\omega_2 - \omega_1)K_1(\omega_1, \omega_2),$$

where, with $\theta = \omega_1/\omega_2$,

$$K_1(\omega_1, \omega_2)$$

$$= \int_0^{2\omega_2} t\gamma'''\left(\frac{t}{\omega_2}\right)\left[\frac{1 - \cos\{(\theta - 1)t\}}{(\theta - 1)t}\cos t + \frac{\sin\{(\theta - 1)t\}}{(\theta - 1)t}\sin t\right] dt.$$

Now,

$$\gamma''(t)\sin(\omega_2)\, dt$$

$$= \omega_2^{-1}\int_0^{2\omega_2} \gamma''(t/\omega_2)\sin t\, dt$$

$$= O\left\{\omega_2^{-1}\int_0^{2\omega_2}(t/\omega_2)^{\alpha-2}(1 \wedge t)\, dt\right\}$$

$$= O\left[\omega_2^{1-\alpha}\{I(\alpha < 1) + (\log \omega_2)I(\alpha = 1) + \omega_2^{\alpha-1}I(\alpha > 1)\}\right],$$

$$\int_0^2 \gamma'''(t)\{1 - \cos(\omega_2 t)\}\, dt$$

$$= \omega_2^{-1}\int_0^{2\omega_2} \gamma'''(t/\omega_2)(1 - \cos t)\, dt$$

$$= O\left\{\omega_2^{-1}\int_0^{2\omega_2}(t/\omega_2)^{\alpha-3}(1 \wedge t^2)\, dt\right\}$$

$$= O(\omega_2^{2-\alpha}),$$

$$K_1(\omega_1, \omega_2)$$

$$= O\left[\int_0^{2\omega_2} t(t/\omega_2)^{\alpha-3}\{1 \wedge (|\theta - 1|t) + 1 \wedge t\}\, dt\right]$$

$$= O\left[\omega_2^{3-\alpha}\{I(\alpha < 1) + (\log \omega_2)I(\alpha = 1) + \omega_2^{\alpha-1}I(\alpha > 1)\}\right].$$

Therefore,

$$K_1(\omega_1, \omega_2)$$

$$= O\Big((\omega_2 - \omega_1)\Big[(\omega_1\omega_2)^{-1}\omega_2^{1-\alpha}\{I(\alpha < 1) + (\log \omega_2)I(\alpha = 1) + \omega_2^{\alpha-1}I(\alpha > 1)\} + (\omega_1^2\omega_2)^{-1}\omega_2^{2-\alpha} + (\omega_1\omega_2)^{-2} \times\{\omega_2^{3-\alpha}I(\alpha < 1) + \omega_2^2(\log \omega_2)I(\alpha = 1) + \omega_2^2 I(\alpha > 1)\}\Big]\Big)$$

$$= O\big[(\omega_2 - \omega_1)\{\omega_1^{-2}\omega_2^{1-\alpha} I(\alpha < 1) + \omega_1^{-2}(\log \omega_2) I(\alpha = 1) + \omega_1^{-2} I(\alpha > 1)\}\big].$$

Hence, result (4.6) follows from (4.4). This completes the proof of Lemma 4.4. □

By Theorem 2.1, $\operatorname{var}\{A(\omega)\} \sim \text{constant} \cdot \omega^{-(\alpha+1)}$, as $\omega \to \infty$, and so inequalities (4.5) and (4.6) yield, for $\omega_1 < \omega_2$ and a constant $C_1 > 0$,

$$|\operatorname{corr}\{A(\omega_1), A(\omega_2)\}| \le C_1(\omega_2 - \omega_1)^{-1} \omega_1^{(1-\alpha)/2} \omega_2^{(\alpha-1)/2}, \tag{4.7}$$

$$\begin{aligned} |\operatorname{corr}\{A(\omega_1), A(\omega_2)\}| \le C_1 \omega_1^{(\alpha-3)/2} \omega_2^{(\alpha+1)/2} \{&\omega_2^{-\alpha} I(\alpha < 1) \\ &+ \omega_2^{-1}(\log \omega_2) I(\alpha = 1) + \omega_2^{-1} I(\alpha > 1)\}. \end{aligned} \tag{4.8}$$

It may be shown that if Z_1, Z_2 have a bivariate normal distribution with zero means, unit variance and correlation coefficient ρ, then for an absolute constant $C_2 > 0$,

$$|\operatorname{cov}(\log|Z_1|, \log|Z_2|)| \le C_2 \rho^2.$$

Therefore,

$$\begin{aligned} \tfrac{1}{4}|\operatorname{cov}\{Y(\omega_1), Y(\omega_2)\}| &= |\operatorname{cov}\{\log|Z(\omega_1)|, \log|Z(\omega_2)|\}| \\ &\le C_2[\operatorname{corr}\{Z(\omega_1), Z(\omega_2)\}]^2 \\ &= C_2[\operatorname{corr}\{A(\omega_1), A(\omega_2)\}]^2. \end{aligned}$$

Hence,

$$S_1 = \sum\sum_{1 \le j_1 < j_2 \le k} |\operatorname{cov}(Y_{j_1}, Y_{j_2})| \le 4C_2 S_2, \tag{4.9}$$

where

$$S_2 = \sum\sum_{1 \le j_1 < j_2 \le k} [\operatorname{corr}\{A(\omega_{j_1}), A(\omega_{j_2})\}]^2.$$

When $\alpha < 1$, results (4.7) (in the case $\omega_1 \le \frac{1}{2}\omega_2$) and (4.8) (for $\omega_1 > \frac{1}{2}\omega_2$) imply that

$$|\operatorname{corr}\{A(\omega_1), A(\omega_2)\}| \le C_3 \omega_1^{(1-\alpha)/2} \omega_2^{(\alpha-3)/2}. \tag{4.10}$$

Therefore,

$$\begin{aligned} S_2 &\le C_3^2 \sum\sum_{1 \le j_1 < j_2 \le k} \omega_{j_1}^{1-\alpha} \omega_{j_2}^{\alpha-3} \\ &\le C_3^2 \sum_{j_2=2}^{k} j_2^{\alpha-3} \sum_{j_1=1}^{j_2-1} j_1^{1-\alpha} \qquad (4.11a) \\ &= O\left(\sum_{j_2=1}^{k} j_2^{-1} \right) = O(\log k). \end{aligned}$$

Similarly, when $\alpha = 1$, result (4.8) implies that

$$S_2 = O\{(\log k)^2\}. \tag{4.11b}$$

When $\alpha > 1$, we employ both (4.7) and (4.8), as follows:

$$S_2 \le C_1^2 \sum\sum_{\substack{1 \le j_1 < j_2 \le k \\ j_2 \le j_1 + j_1^{2-\alpha}}} \omega_{j_1}^{\alpha-3} \omega_{j_2}^{\alpha-1} + C_1^2 \sum\sum_{\substack{1 \le j_1 < j_2 \le k \\ j_2 > j_1 + j_1^{2-\alpha}}} (\omega_{j_2} - \omega_{j_1})^{-2} \omega_{j_1} \omega_{j_2}^{\alpha-1},$$

$$S_2 \le C_1^2 \sum\sum_{\substack{1 \le j_1 < j_2 \le k \\ j_2 \le j_1 + j_1^{2-\alpha}}} j_1^{\alpha-3} j_2^{\alpha-1} + C_1^2 \sum\sum_{\substack{1 \le j_1 < j_2 \le k \\ j_2 > j_1 + j_1^{2-\alpha}}} (j_2 - j_1)^{-2} j_1^{1-\alpha} j_2^{\alpha-1}$$

$$= O\left\{ \int_1^k x^{\alpha-3}\, dx \int_x^{x+x^{2-\alpha}} y^{\alpha-1}\, dy \right.$$

$$\left. + \int_1^k x^{1-\alpha}\, dx \int_{x+x^{2-\alpha}}^k (y-x)^{-2} y^{\alpha-1}\, dy \right\} \tag{4.11c}$$

$$= O\left\{ \int_1^k x^{\alpha-3} x^{2-\alpha} x^{\alpha-1}\, dx \right.$$

$$\left. + \int_1^k x^{1-\alpha} k^{\alpha-1}\, dx \int_{x+x^{2-\alpha}}^k (y-x)^{-2}\, dy \right\}$$

$$= O\left(k^{\alpha-1} + k^{\alpha-1} \int_1^k x^{1-\alpha} x^{\alpha-2}\, dx \right)$$

$$= O(k^{\alpha-1} \log k).$$

Combining (4.1), (4.9) and (4.11) we deduce that

$$\left| \sum\sum_{j_1 \ne j_2} (x_{j_1} - \bar{x})(x_{j_2} - \bar{x}) \operatorname{cov}(y_{j_1}, y_{j_2}) \right| \le 2(\log k)^2 S_1$$

$$\le 8C_2 (\log k)^2 S_2$$

$$= o(k) = o\left\{ \sum (x_j - \bar{x})^2 \right\},$$

as had to be proved. □

4.2. *Proof of Theorem* 2.4. Our proof is by the method of moments. Note that, since $Z(\omega)$ is a linear function of a Gaussian process, it has a normal distribution. Let E, E_0 denote expectation under the true model for the joint distribution of $Z(\omega_j)$, $1 \le j \le k$, and the model where the $Z(\omega_j)$'s are stochastically independent with standard normal marginal distributions. Put $\mu = E\{\log|N(0,1)|\}$. It suffices to show that, for each integer $\nu \ge 1$,

$$E\left[\sum_{j=1}^k (x_j - \bar{x}) \{\log|Z(\omega_j)| - \mu\} \right]^\nu - E_0\left[\sum_{j=1}^k (x_j - \bar{x}) \{\log|Z(\omega_j)| - \mu\} \right]^\nu$$

$$= o(k^{\nu/2}),$$

as $k \to \infty$. The left-hand side is dominated by a constant multiple of a sum of

a bounded number of terms of the form

$$(\log k)^{\nu} \sum_{\substack{\omega_{j_1},\ldots,\omega_{j_m} \\ \text{all distinct}}} \cdots \sum \left| d(l_1,\ldots,l_m;\omega_{j_1},\ldots,\omega_{j_m}) \right|,$$

where

$$d(l_1,\ldots,l_m;\omega_{j_1},\ldots,\omega_{j_m}) = (E - E_0)\left[\prod_{i=1}^{m} \left\{ \log|Z(\omega_{j_i})| - \mu \right\}^{l_i} \right]$$

and $1 \le m \le \nu$ and $l_1,\ldots,l_m \ge 1$ denote integers satisfying $l_1 + \cdots + l_m = \nu$.

Put $Z = (Z(\omega_{j_1}),\ldots,Z(\omega_{j_m}))^T$ and let Σ denote the covariance matrix of Z. Define $\Delta = \Sigma - I$ and consider Taylor expansions of the quantities $(I + \Delta)^{-1}$ and $(I + \Delta)^{-1/2}$ in the elements of Δ. Arguing thus we may express the density of Z,

$$\phi_\Delta(z) = (2\pi)^{-m/2} |I + \Delta|^{-1/2} \exp\left\{ -\tfrac{1}{2} z^{\mathrm{T}} (I + \Delta)^{-1} z \right\}, \tag{4.12}$$

as a Taylor expansion of the form

$$\phi_\Delta(z) = \phi_0(z)\{1 + p_1(z) + p_2(z) + \cdots\}, \tag{4.13}$$

where $\phi_0(z) = (2\pi)^{-m/2} \exp(\frac{1}{2} z^T z)$ and $p_j(z)$ denotes a polynomial in the $\frac{1}{2}m(m+1)$ quantities $z_{i_1} z_{i_2}$, $1 \le i_1 < i_2 \le m$, in which each coefficient is a sum of constant multiples of products of precisely j elements of Δ, and $\int p_j \phi_0 = 0$.

In this notation,

$$\begin{aligned} &d(l_1,\ldots,l_m;\omega_{j_1},\ldots,\omega_{j_m}) \\ &\quad = \int \left\{ \prod_{i=1}^{m} (\log|z_i| - \mu)^{l_i} \right\} \{\phi_\Delta(z) - \phi_0(z)\}\, dz \qquad (4.14) \\ &\quad = \sum_{j \ge 1} \int \left\{ \prod_{i=1}^{m} (\log|z_i| - \mu)^{l_i} \right\} p_j(z)\, \phi_0(z)\, dz. \qquad (4.15) \end{aligned}$$

Now, $p_j(z)$ may be written as a linear form in terms

$$\mathrm{He}_{s_1}(z_1) \cdots \mathrm{He}_{s_m}(z_m),$$

where $\mathrm{He}_i(u)$ denotes the ith Hermite polynomial (orthogonal with respect to the weight function $e^{-u^2/2}$), the number of such terms is bounded by a constant depending only on j and m and the coefficient of each term is a product of j elements of Δ. Thus, the jth term in (4.15) may be expressed as a linear form in terms

$$\begin{aligned} &\int \left\{ \prod_{i=1}^{m} (\log|z_i| - \mu)^{l_i} \mathrm{He}_{s_i}(z_i) \right\} \phi_0(z)\, dz \\ &\quad = \prod_{i=1}^{m} \left\{ (2\pi)^{-1/2} \int (\log|u| - \mu)^{l_i} \mathrm{He}_{s_i}(u) \exp(-u^2/2)\, du \right\}. \end{aligned} \tag{4.16}$$

The quantity in (4.16) vanishes if (i) for some i, s_i is odd or (ii) for some i, $s_i = 0$ and $l_i = 1$. Therefore, we may assume that, for each i, either (a) $l_i = 1$ and $s_i = 2, 4, \ldots$, or (b) $l_i \geq 2$, and $s_i = 0, 2, 4, \ldots$, and also that (c) not each $s_i = 0$. [The case where each $s_i = 0$ corresponds to the subtracted term in (4.14).]

Without loss of generality, the first r s_i's all exceed zero and the others all equal zero. Let $\rho_{i_1 i_2} = \operatorname{corr}\{Z(\omega_{j_{i_1}}), Z(\omega_{j_{i_2}})\}$. Inspecting the way in which (4.13) arises from (4.12) we see that the coefficient of (4.15) is dominated by a linear form in terms like

$$\prod_{i=1}^{r} \rho_{ih(i)}^2,$$

where $i \neq h(i) \in \{1, \ldots, m\}$. Note that, since $l_i \geq 2$ for $r + 1 \leq j \leq m$, we have

$$r + 2(m - r) \leq \nu, \quad \text{that is, } m - r \leq \tfrac{1}{2}(\nu - r) \leq \tfrac{1}{2}(\nu - 1).$$

Therefore, it suffices to show that for each $1 \leq r \leq m$ satisfying $m - r \leq \frac{1}{2}(\nu - 1)$ and functions h: $\{1, \ldots, r\} \to \{1, \ldots, m\}$ such that $h(i) \neq i$, we have

$$\sup_h \sum_{\substack{\omega_{j_1}, \ldots, \omega_{j_m} \\ \text{all distinct}}} \cdots \sum \prod_{i=1}^{r} \rho_{ih(i)}^2 = o\{k^{\nu/2}(\log k)^{-\nu}\}. \tag{4.17}$$

If we prove that

$$\sum\sum_{j_1 \neq j_2} \left[\operatorname{corr}\{Z(\omega_{j_1}), Z(\omega_{j_2})\}\right]^2 = O\{(\log k)^2\}, \tag{4.18}$$

then it follows that the left-hand side of (4.17) equals

$$O\{(\log k)^{2r} k^{m-r}\} = O\{k^{(\nu-1)/2}(\log k)^{2r}\},$$

which establishes (4.17) and completes the proof of Theorem 2.4.

In the cases $\alpha < 1$ and $\alpha = 1$, result (4.18) follows from (4.11a) and (4.11b), respectively. However, (4.11c) does not yield (4.18) in the case $\alpha > 1$. To establish (4.18) in this instance we shall first prove (4.10) for $\alpha > 1$.

When $\omega_1 < \frac{1}{2}\omega_2$, (4.10) follows from (4.7). In the case $\omega_1 \geq \frac{1}{2}\omega_2$ we shall refine the proof of (4.6) in Lemma 4.4. By that proof, for $\frac{1}{2}\omega_2 \leq \omega_1 < \omega_2$,

$$\begin{aligned} |\operatorname{cov}\{A(\omega_1), A(\omega_2)\}| &\leq 4(\omega_1\omega_2)^{-2}\left|\int_0^2 \gamma'''(t)\{1 - \cos(\omega_2 t)\}\, dt\right| \\ &\quad + 2(\omega_1^2\omega_2^3)^{-1}|K_1(\omega_1, \omega_2)| \\ &\leq C_4(\omega_1^2\omega_2^\alpha)^{-1} + 2(\omega_1^2\omega_2^3)^{-1}|K_1(\omega_1, \omega_2)|, \end{aligned}$$

where $C_4, C_5, \ldots$ denote positive constants. To bound K_1, put

$$H(u) = \int_0^u \left[\frac{1 - \cos\{(\theta - 1)t\}}{(\theta - 1)t} \cos t + \frac{\sin\{(\theta - 1)t\}}{(\theta - 1)t} \sin t \right] dt$$

and note that $H(u)$ is bounded uniformly in $u > 0$ and $\frac{1}{2} \le \theta < 1$. Now,

$$\begin{aligned} K_1(\omega_1, \omega_2) &= \left(\int_0^1 + \int_1^{2\omega_2} \right) t\gamma'''(t/\omega_2)\, dH(t) \\ &= \int_1^{2\omega_2} t\gamma'''(t/\omega_2)\, dH(t) + O(\omega_2^{3-\alpha}) \\ &= -\int_1^{2\omega_2} \{\gamma'''(t/\omega_2) + \omega_2^{-1} t\gamma^{(4)}(t/\omega_2)\} H(t)\, dt + O(\omega_2^{3-\alpha}) \\ &= O\left(\int_1^{2\omega_2} (t/\omega_2)^{\alpha-3}\, dt + \omega_2^{3-\alpha} \right) = O(\omega_2^{3-\alpha}). \end{aligned}$$

Therefore, if $\frac{1}{2}\omega_2 \le \omega_1 < \omega_2$, then

$$|\operatorname{cov}\{A(\omega_1), A(\omega_2)\}| \le C_5 \omega_1^{-2} \omega_2^{-\alpha} \le C_6 \omega_1^{-\alpha} \omega_2^{-2},$$

which (in view of Theorem 2.1) is equivalent to (4.10). From that result and the argument at (4.11a), follows (4.18). This completes the proof of Theorem 2.4. □

4.3. *Proof of Theorem* 2.5. Theorem 2 of Garsia (1972) implies the existence of a random variable V such that, with probability 1 for all $-1 < s < t < 1$ satisfying $t - s \le \frac{1}{2}$,

$$|X(s) - X(t)| \le V\{(t - s)^{\alpha} |\log(t - s)|\}^{1/2}.$$

Therefore, if $n \ge 2$,

$$\sup_{i/n \le t \le (i+1)/n} |X(t) - X(i/n)| \le V(n^{-\alpha} \log n)^{1/2},$$

whence for $0 < \omega < 2\pi k$,

$$\begin{aligned} |A(\omega) - \hat{A}_n(\omega)| &= \left| \int_{-1}^{1} X(t) \cos(\omega t)\, dt - n^{-1} \sum_{i=-n}^{n-1} X(i/n) \cos(\omega i/n) \right| \\ &\le n^{-1} \sum_{i=-n}^{n-1} \left[\sup_{i/n \le t \le (i+1)/n} |X(t) - X(i/n)| \right. \\ &\qquad \left. + \left\{ \sup_{-1 < t < 1} |X(t)| \right\} \sup_{i/n \le t \le (i+1)/n} |\cos(\omega t) - \cos(\omega i/n)| \right] \\ &\le W\{(n^{-\alpha} \log n)^{1/2} + kn^{-1}\}, \end{aligned}$$

where the random variable W does not depend on ω or n.

If $\delta > 0$, then the complement of the event,

$$\mathscr{E} = \left\{ \min_{1 \le j \le k} |Z(\omega_j)| > \delta \right\},$$

has probability

$$\mathrm{P}(\tilde{\mathscr{E}}) \le \sum_{j=1}^{k} \mathrm{P}\{|Z(\omega_j)| \le \delta\} \le k\delta.$$

Hence, if we choose $\delta = \delta(k) = o(k^{-1})$, then $\mathrm{P}(\mathscr{E}) \to 1$. Furthermore, for a constant $C_1 > 0$, $\sigma(\omega_j) \ge C_1 \omega_j^{-(\alpha+1)/2}$, for all j, and so on the set $\mathscr{E}$,

$$\begin{aligned} &\left| |\hat{A}_n(\omega_j)| - |A(\omega_j)| \right| |A(\omega_j)|^{-1} \\ &\quad \le W\{(n^{-\alpha} \log n)^{1/2} + kn^{-1}\}\{C_1 \omega_j^{-(\alpha+1)/2}\delta\}^{-1} \\ &\quad \le C_2 W\{n^{-\alpha}(mk)^{\alpha+1} \log n + n^{-2}(mk)^{\alpha+1}k^2\}^{1/2} \delta^{-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} |\hat{\alpha} - \hat{\alpha}_n| &\le 2\left\{ \sum_{j=1}^{k} (x_j - \bar{x})^2 \right\}^{-1} \\ &\quad \times \sum_{j=1}^{k} |x_j - \bar{x}| \left| \log\left[1 + \left\{ |\hat{A}_n(\omega_j)| - |A(\omega_j)| \right\} |A(\omega_j)|^{-1} \right] \right| \\ &= O\left[\left\{ n^{-\alpha}(mk)^{\alpha+1} \log n + n^{-2}(mk)^{\alpha+1}k^2 \right\}^{1/2} \delta^{-1} \log k \right] \end{aligned} \tag{4.19}$$

with probability 1. Condition (2.10) is adequate to ensure that $\delta = \delta(k)$ may be chosen so that $\delta = o(k^{-1})$ and the right-hand side of (4.19) equals $o(k^{-1/2})$. This shows that $k^{1/2}|\hat{\alpha}_n - \hat{\alpha}| \to 0$. □

Acknowledgment. We are grateful to two referees for their particularly helpful and detailed comments.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. A. (1965). *Handbook of Mathematical Functions*. Dover, New York.

ADLER, R. J. (1981). *The Geometry of Random Fields*. Wiley, New York.

BENHENNI, K. and CAMBANIS, S. (1992). Sampling designs for estimating integrals of stochastic processes. *Ann. Statist.* **20** 161–194.

BERRY, M. V. and HANNAY, J. H. (1978). Topography of random surfaces. *Nature* **273** 573.

BROWN, C. A., CHARLES, P. D., JOHNSEN, W. A. and CHESTER, S. (1993). Fractal analysis of topographic data by the patchwork method. *Wear* **161** 61–67.

BURROUGH, P. A. (1981). Fractal dimensions of landscapes and other environmental data. *Nature* **19** 240–242.

CARTER, P. H., CAWLEY, R. and MAULDIN, R. D. (1988). Mathematics of dimension measurements of graphs of functions. In *Proceedings of the Symposium on Fractal Aspects of Materials, Disordered Systems* (D. A. Weitz, L. M. Sander and B. B. Mandelbrot, eds.) 183–186. Materials Research Society, Pittsburgh, PA.

CONSTANTINE, A. G. and HALL, P. (1994). Characterising surface smoothness via estimation of effective fractal dimension. *J. Roy. Statist. Soc. Ser. B* **56** 97–113.

COSTER, M. and CHERMANT, J. C. (1983). Recent developments in quantitative fractography. *International Metals Review* **28** 234–238.

COX, D. R. (1984). Long range dependence: a review. In *Statistics: An Appraisal* (H. A. David and H. T. David, eds.). Iowa State Univ. Press, Ames, IA.

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

DELFINER, P. and DELHOMME, J. P. (1975). Optimal interpolation by kriging. In *Display and Analysis of Spatial Data* (J. C. Davis and M. J. McCullagh, eds.) 96–114. Wiley, London.

DUBUC, B., QUINIOU, J. F., ROQUES-CARMES, C., TRICOT, C. and ZUCKER, S. W. (1989). Evaluating the fractal dimension of profiles. *Phys. Rev. A* **39** 1500–1512.

DUBUC, B., ZUCKER, S. W., TRICOT, C., QUINIOU, J. F. and WEHBI, D. (1989). Evaluating the fractal dimension of surfaces. *Proc. Roy. Soc. London Ser. A* **25** 113–127.

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.

GARSIA, A. M. (1972). Continuity properties of Gaussian processes with multidimensional time parameter. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2**. Univ. California Press, Berkeley.

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

GRANGER, C. W. J. and JOYEUX, R. (1980). An introduction to long memory time series models and fractional differencing. *J. Time Ser. Anal.* **1** 15–29.

HALL, P. (1995). On the effect of measuring a self-similar process. *SIAM J. Appl. Math.* **55** 800–808.

HALL, P. and ROY, R. (1994). On the relationship between fractal dimension and fractal index for stationary stochastic processes. *Ann. Appl. Probab.* **4** 241–253.

HAMPEL, F. R. (1987). Data analysis and self-similar processes (with discussion). In *Proceedings of the 46th Session of the ISI* **52** 235–264.

HOSKING, J. R. M. (1981). Fractional differencing. *Biometrika* **68** 165–176.

JANACEK, G. J. (1982). Determining the degree of differencing for time series via the log spectrum. *J. Time Ser. Anal.* **3** 177–183.

JOURNEL, A. G. and HUIJBREGTS, CH. J. (1978). *Mining Geostatistics*. Academic Press, London.

KASHYAP, R. L. and EOM, K. B. (1988). Estimation in long memory time series models. *J. Time Ser. Anal.* **9** 35–42.

LING, F. F. (1987). Scaling law for contoured length of engineering surfaces. *J. Appl. Phys.* **62** 2570–2572.

LING, F. F. (1989). The possible role of fractal geometry in tribology. *Tribology Trans.* **32** 497–505.

LING, F. F. (1990). Fractals, engineering surfaces and tribology. *Wear* **136** 141–156.

MAJUMDAR, A. and BHUSHAN, B. (1991). Fractal model of elastic–plastic contact between rough surfaces. *J. Tribology* **113** 1–11.

MANDELBROT, B. B. (1975). Limit theorems for the self-normalized range. *Z. Wahrsch. Verw. Gebiete* **31** 271–285.

MANDELBROT, B. B., PASSOJA, D. E. and PAULLAY, A. J. (1984). Fractal character of surfaces of metals. *Nature* **308** 721–722.

MANDELBROT, B. B. and TAQQU, M. S. (1979). Robust R/S analysis of long run serial correlation. *Bull. Inst. Internat. Statist.* **48** 69–104.

MANDELBROT, B. B. and VAN NESS, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10** 422–437.

OGATA, Y. and KATSURA, K. (1991). Maximum likelihood estimates of the fractal dimension for spatial patterns. *Biometrika* **78** 463–474.

SERRA, J. (1968). Les structures gigogles-morphologie, mathématique et interpretation metallogenique. *Mineralium Deposita* **3** 135–154.

TAYLOR, C. C. and TAYLOR, S. J. (1991). Estimating the dimension of a fractal. *J. Roy. Statist. Soc. Ser. B* **53** 353–364.

THOMAS, T. R. and THOMAS, A. P. (1988). Fractals and engineering surface roughness. *Surface Topography* **1** 143–152.

WOOD, A. T. A. and CHAN, G. (1993). Simulation of stationary Gaussian processes in $[0, 1]^d$. Research Report CMA-SR06-93, Australian National Univ.

ZYGMUND, A. (1959). *Trigonometric Series* **1**. Cambridge Univ. Press.

GRACE CHAN
D. S. POSKITT
DEPARTMENT OF STATISTICS
AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA, ACT 0200
AUSTRALIA

PETER HALL
CENTRE FOR MATHEMATICS AND ITS APPLICATIONS
AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA, ACT 0200
AUSTRALIA