The Annals of Statistics
1987, Vol. 15, No. 1, 134–151

LIMITING BEHAVIOR OF FUNCTIONALS OF HIGHER-ORDER SAMPLE CUMULANT SPECTRA[1]

BY DANIEL MACRAE KEENAN

Brown University

This paper is concerned with establishing a broad class of estimators (and the limiting behavior, thereof) for parametrizations of higher than second-order structure. This includes parametrizations which reflect, for example, such properties as nonlinearity and/or non-Gaussianity and/or time irreversibility. Asymptotic distributions, almost-sure convergence, and probability-one bounds for such estimators are established. Several applications of such estimators are discussed.

1. Introduction. Time series analysis is a body of techniques that have been developed to study the dependence structure of processes evolving over time. The processes that have typically been studied are those that are stationary (in one of its various forms) and linear; the methods which have been developed have primarily been based upon second-order structure. One future direction, however, appears to be toward broadening the framework to include more general processes (e.g., nonlinear).

This paper is concerned with developing a broad class of estimators (and the limiting behavior, thereof) for parametrizations of higher than second-order structure. This would include parametrizations which reflect, for example, such properties as nonlinearity and/or non-Gaussianity and/or time irreversibility. A very important example of such a parameter concerns that component of the asymptotic variance of the quasi-maximum likelihood estimator [see Hosoya and Taniguchi (1982)] which is an integral of a certain function w.r.t. the fourth-order cumulant spectra [see Brillinger (1975), Theorem 5.10.1]. Taniguchi (1982) has constructed a consistent estimator of this integral. Below we suggest a modification [expression (2.28)] to the estimator suggested by Taniguchi, obtaining one which is simpler both in its form and its variance. An application of Theorem 3.3 below gives a bound on the rate of its convergence. For linear processes, estimators (of parameters) defined via a minimization procedure (under regularity conditions) will not contain this integral term because the fourth-order cumulant spectra factorizes as a product of second-order spectra; otherwise there will typically be such a term. In Keenan (1985a) estimators defined via integral minimization are considered; the result of Taniguchi (1982) applies to the corresponding asymptotic variance component of these estimators. Taniguchi's result allows for the construction of confidence intervals in these cases. Weiss (1975) has shown that for linear, stationary (more specifically, ARMA) processes,

Received July 1984; revised July 1986.

[1]Research supported by an NSF Mathematical Sciences Postdoctoral Fellowship at Brown University.

AMS 1980 *subject classifications*. Primary 60G10; secondary 62M15.

Key words and phrases. Time series, cumulant spectra, functionals, limiting behavior.

Gaussianity and time reversibility are equivalent; Cox (1981) has suggested third- (and higher-) order cumulant analogues of the variogram as potential measures of time irreversibility. Such statistics also fit the framework of the present paper.

Higher-order cumulant spectra [first suggested by Kolmogorov, but see also Tukey (1953)] have been studied by various authors, e.g., Brillinger and Rosenblatt (1967a, b). Brillinger and Rosenblatt (1967a, b) develop consistent estimators of higher-order cumulant spectra via spectral windows (analogous to what is done in the second-order case). The present work is not concerned with the estimation of higher-order spectra, which is local in nature, but rather with the global concern of estimating functionals of the higher-order spectra. If we wish to consider higher-order parameters other than the cumulant spectra itself, then a framework for their estimation is needed. The parameters (estimators) which we will consider are those representable as an integral of a kernel function w.r.t the kth (≥ 2)-order (sample) cumulant spectra. In this paper we will establish the asymptotic distributions, almost-sure convergence and probability one bounds for such estimators. Two approaches are considered in this paper; one approach taken in the construction of said estimators is motivated by Taniguchi (1982). We will construct consistent estimators via recursion; the bias depends on lower-dimensional parameters and, consequently, we will build up from consistent estimators for second- and third-order parameters. The second approach excludes the contributions due to those points which cause the biases; the covariances of these estimators have a simpler form. This approach of excluding such points was also used by Brillinger and Rosenblatt (1967a, b) in a related context.

2. Background and construction of estimators. Let $\{X_i| -\infty < i < \infty\}$ be a zero-mean, real-valued, strictly stationary process satisfying [see Brillinger (1975), Section 2.6]:

ASSUMPTION 1.

$$\sum_{v_1, v_2, \dots, v_{k-1} = -\infty}^{\infty} |v_j| \, |c(v_1, v_2, \dots, v_{k-1})| < \infty, \tag{2.1}$$

for $j = 1, 2, \dots, k-1$, $k = 2, 3, \dots,$ where $c(v_1, v_2, \dots, v_{k-1})$ is the kth-order cumulant of $\{X(0), X(v_1), \dots, X(v_{k-1})\}$.

In the case of a Gaussian process the cumulants of order greater than or equal to three are all zero. The kth (≥ 2)-order cumulant spectrum is defined as

$$f^{(k)}(\lambda_1, \lambda_2, \dots, \lambda_{k-1}) = 1/(2\pi)^{k-1} \sum_{v_1, v_2, \dots, v_{k-1} = -\infty}^{\infty} \exp\left\{ i \left(\sum_{j=1}^{k-1} \lambda_j v_j \right) \right\} c(v_1, \dots, v_{k-1}). \tag{2.2}$$

We could consider R^p-valued processes, the generalization being straightforward. The kth-order cumulant spectral distribution is defined to be the complex-valued measure constructed as the integral of $f^{(k)}$:

$$F^{(k)}(A) = \int_A f^{(k)}(\lambda_1, \lambda_2, \dots, \lambda_{k-1})\, d\lambda, \tag{2.3}$$

where A is a Borel set in the $(k-1)$ toral group, $(S^1)^{k-1}$, S^1 being the unit circle. [Stationarity has reduced the integration on $[0, 2\pi]^k$ to that on the $(k-1)$-manifold, $\{\lambda | \Sigma_{i=1}^k \lambda_i \equiv 0 \pmod{2\pi}\}$, i.e., $[0, 2\pi]^{k-1}$.] The same symbol $F^{(k)}(\cdot)$, will be used to represent both the integrated kth-order cumulant spectral distribution function and the complex measure. Let h be a $\mathbb{C}^r$-valued function of bounded variation [see Hobson (1927), pages 343–346] defined on $(S^1)^{k-1}$ and $\theta(F^{(k)})$ be a parameter of interest defined by

$$\theta(F^{(k)}) =_{\text{def}} \int_{(S^1)^{k-1}} h(\lambda_1, \lambda_2, \dots, \lambda_{k-1})\, dF^{(k)}(\lambda_1, \lambda_2, \dots, \lambda_{k-1}). \tag{2.4}$$

That is, the present generalization of bounded variation to more than one dimension, which we will refer to as Hobson bounded variation, is the following. We assume that the real and imaginary parts of each of the q components of h are of bounded variation in j variables, $1 \le j \le k-1$, for all values of the remaining $k-1-j$ variables, where bounded variation in the j variables means that supremum over all rectangular decompositions of $[0, 2\pi]^j$ is finite for the absolute value of the additive rectangle function (i.e., jth difference) constructed from h. Let $\theta_j(F^{(k)})$ and $h_j(\cdot)$ correspond to the jth, $1 \le j \le r$, (complex-valued) component of $\theta(F^{(k)})$ and $h(\cdot)$, respectively.

To illustrate such parameters, consider the following examples of expression (2.4). Hosoya and Taniguchi (1982), Keenan (1985a), among others, consider second-order vector-valued parameters (θ) and their estimators ($\hat{\theta}$) defined via minimization (in t) of integral expressions (for the true and sample spectral distribution functions, respectively, substituted for G):

$$\int_0^{2\pi} \rho(\xi, t, g(\xi))\, dG(\xi), \tag{2.5}$$

where ρ is a function satisfying certain regularity conditions and g is a (generalized) derivative of G. A component of the asymptotic covariance matrix of $\hat{\theta}$ is

$$\int_{[0,2\pi]^2} \left[\frac{\partial \rho}{\partial \theta}(\alpha, \theta, f^{(4)}(\alpha, \beta, 2\pi - \beta)) \frac{\partial \rho}{\partial \theta'}(\beta, \theta, f^{(4)}(\alpha, \beta, 2\pi - \beta)) \right] \times f^{(4)}(\alpha, \beta, 2\pi - \beta)\, d\alpha\, d\beta, \tag{2.6}$$

which is of the form of expression (2.4) with

$$h(\lambda_1, \lambda_2) = \phi(\lambda_1, \lambda_2, \lambda_3) \eta(\lambda_2 + \lambda_3),$$

where

$$\phi(\lambda_1, \lambda_2, \lambda_3) = \frac{\partial \rho}{\partial \theta}\big(\lambda_1, \theta, f^{(4)}(\lambda_1, \lambda_2, 2\pi - \lambda_2)\big) \times \frac{\partial \rho}{\partial \theta'}\big(\lambda_2, \theta, f^{(4)}(\lambda_1, \lambda_2, 2\pi - \lambda_2)\big) \tag{2.7}$$

[$\eta(\cdot)$ is defined by (2.10) below]. Note that this example differs in form from (2.4) in that the integration is over $(S^1)^{k-2}$ rather than $(S^1)^{k-1}$, $k = 4$. Taniguchi (1982) handles this by using a spectral window on this extra dimension. A generalization of this can be taken for parameters defined via integration only over some submanifold; because of Lemma 2.1 below, such integrals are common in higher-order spectral theory. Another example concerns second-order parameters and estimators defined via integration of a kernel function. A component of the asymptotic variance of these estimators is again of the form of (2.4) with a different $h(\cdot)$ function than in (2.7) [see (3.9), Keenan (1983)]. Weiss (1975) and Cox (1981) have proposed diagnostics for time irreversibility of a strictly stationary process. The parameters of interest are

$$\{E(X_i - X_{i+l})^m,\ l = 1, 2, \dots, L;\ m = 3, 4\} \qquad (L \text{ fixed}) \tag{2.8}$$

appropriately standardized. For $h(\lambda_1, \lambda_2, \dots, \lambda_{k-1})$ defined as

$$h(\lambda_1, \dots, \lambda_{k-1}) = \exp\{i(\lambda \cdot \mathbf{j})\}, \tag{2.9}$$

where $\mathbf{j} = (j_1, j_2, \dots, j_{k-1}) \in \mathbb{Z}^{k-1}$, $\theta(F^{(k)})$ is the kth-order joint cumulant evaluated at $\mathbf{j}$ and thus the results of this paper apply to finite linear combinations of joint cumulants such as (2.8). That is, Corollary 2.5 and Theorem 3.3 (below) give the asymptotic normality and probability one bounds (and almost-sure convergence) for the sample versions of (2.8) used in testing time reversibility. More generally, the trigonometric polynomials in $k - 1$ variables are possible $h(\cdot)$ functions for (2.4).

CONVENTION. The following notational form [also used by Brillinger and Rosenblatt (1967a, b)] will be employed throughout Section 2 as a matter of uniformity in representation. We write the functions [e.g., $f^{(k)}$ and h] as if they were functions of k variables, $\xi_1, \dots, \xi_k$ (rather than $k - 1$), with the restriction that $\sum_{i=1}^{k} \xi_i \equiv 0 \pmod{2\pi}$ is always intact in addition to possibly other linear constraints of the same nature, e.g., $\sum_{i=1}^{p} \xi_{j_i} \equiv 0 \pmod{2\pi}$, $1 \le p \le k$. Throughout, when we refer to submanifolds it is always to submanifolds defined by these constraints. The integration (of these functions) is to be interpreted as over the l-dimensional torus where l is the number of free variables; the number of free variables in the integration is identified via the multiplication by the Dirac comb (or a product of),

$$\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda + 2\pi j), \tag{2.10}$$

where $\delta(\cdot)$ is the Dirac δ-function. For example, the integral in (2.4) defining $\theta(F^{(k)})$ will be written in its equivalent form

$$\int_{(S^1)^k} h(\xi_1,\dots,\xi_{k-1},\xi_k)\eta\left(\sum_{i=1}^{k}\xi_i\right) dF^{(k)}(\xi_1,\dots,\xi_{k-1},\xi_k).$$

Define $F_n^{(k)}(\cdot)$, based upon $\{X_1, X_2,\dots,X_n\}$, as

$$F_n^{(k)}(\lambda_1,\lambda_2,\dots,\lambda_{k-1}) = \left(\frac{2\pi}{n}\right)^{k-1} \sum_{\substack{(2\pi r_j)/(n)\le\lambda_j,\\ j=1,2,\dots,k-1}} I_n^{(k)}\left(\frac{2\pi r_1}{n},\dots,\frac{2\pi r_{k-1}}{n},\frac{2\pi r_k}{n}\right), \tag{2.11}$$

where $I_n^{(k)}(\cdot)$ is the kth-order periodogram [Brillinger and Rosenblatt (1967a, b)]

$$I_n^{(k)}(\lambda_1,\lambda_2,\dots,\lambda_k) = (2\pi)^{-k+1}n^{-1}\eta\left\{\sum_{j=1}^{k}\lambda_j\right\}\prod_{j=1}^{k}d_n(\lambda_j), \tag{2.12}$$

$\eta\{\cdot\}$ being the Kronecker comb and

$$d_n(\lambda) = \sum_{t=1}^{n} X_t\exp\{-i\lambda t\} \tag{2.13}$$

the finite Fourier transform, and

$$\theta\left(F_n^{(k)}\right) = \int h(\lambda_1,\dots,\lambda_{k-1})\, dF_n^{(k)}(\lambda_1,\dots,\lambda_{k-1})$$
$$= \left(\frac{2\pi}{n}\right)^{k-1} \sum_{r_1,r_2,\dots,r_{k-1}=1}^{n} h\left(\frac{2\pi r_1}{n},\dots,\frac{2\pi r_k}{n}\right) I_n^{(k)}\left(\frac{2\pi r_1}{n},\dots,\frac{2\pi r_k}{n}\right).$$

For $k = 2$, $F_n^{(2)}$ is a good estimator of $F^{(2)}$ [and, consequently, $\theta(F_n^{(2)})$ for $\theta(F^{(2)})$] because there is sufficient, natural smoothing built into its construction and because there are no further submanifolds of $\sum_{j=1}^{k}\lambda_j \equiv 0 \pmod{2\pi}$ besides $(0,0)$ and (π,π); for $k = 3$ the same is also true (assuming mean zero for the process). For $k > 3$, the submanifolds create inconveniences. The following two well-known lemmas are due, respectively, to Leonov and Shiryaev (1959) [see Brillinger (1975), pages 20–21] and Brillinger and Rosenblatt (1967a).

LEMMA 2.1. *For Y_{ij}, $j = 1,2,\dots,J_i$, $i = 1,2,\dots,I$, a two-way array of random variables and*

$$Z_i = \prod_{j=1}^{J_i} Y_{ij},$$

the joint cumulant of $\{Z_1, Z_2,\dots,Z_I\}$ is given by

$$\sum_{\nu}\left[\prod_{i=1}^{p}\mathrm{cum}(Y_{ij},\ j\in\nu_i)\right],$$

the summation being over all indecomposable partitions, ν, of the two-way array: $\{(i, j),\ j = 1,\dots, J_i,\ i = 1,\dots, I\}$.

PROOF. Leonov and Shiryaev (1959). □

LEMMA 2.2. *Under Assumption* 1, *we have*

$$\operatorname{cum}\{d_n(\lambda_1),\dots, d_n(\lambda_k)\} = (2\pi)^{k-1}\Delta_n\left(\sum_{j=1}^{k}\lambda_j\right)\eta\left(\sum_{j=1}^{k}\lambda_j\right) f_{xx\cdots x}(\lambda_1,\dots,\lambda_k) + O(1), \tag{2.14}$$

where

$$\lambda_j = \frac{2\pi r_j}{n}, \qquad r_j \in \mathbb{Z},$$

and

$$\Delta_n(\lambda) = \sum_{t=0}^{n-1}\exp\{-i\lambda t\} = \begin{cases} 0, & \lambda \neq 0 \pmod{2\pi}, \\ n, & \lambda = 0 \pmod{2\pi}, \end{cases} \tag{2.15}$$

and the error term $O(1)$ *is uniform for all* $\lambda_1,\dots,\lambda_{k-1},\lambda_k$.

PROOF. Lemma 1, Brillinger and Rosenblatt (1967a). □

Applying the preceding lemmas, in the case $I = 1$, one can calculate $E[\theta(F_n^{(k)})]$. Since h and $\{f^{(\alpha_j)},\ j = 1,\dots, p\}$ are of bounded variation we have

$$E\left[\theta\left(F_n^{(k)}\right)\right] = \theta(F^{(k)}) + \sum_{\nu}\int_{(S^1)^k} h(\xi_1,\xi_2,\dots,\xi_k)\left[\prod_{j=1}^{p} f^{(\alpha_j)}(\xi_{j_i};\ j_i \in \nu_j)\right] \times\left[\prod_{j=1}^{p}\eta\left(\sum_{i=1}^{\alpha_j}\xi_{j_i}\right)\right]\prod_{l=1}^{k} d\xi_l + O(n^{-1}), \tag{2.16}$$

where the outer sum is over all $\nu = (\nu_1,\dots,\nu_p)$, $p > 1$, and $\alpha_j > 1$ is equal to the number of elements in ν_j.

DEFINITION. Let $\Omega^{(k)}$ be the collection of $(\lambda_1,\dots,\lambda_k)$ in $(S^1)^k$ such that $\sum_{j=1}^{k}\lambda_j \equiv 0 \pmod{2\pi}$ but is not contained in any further submanifold of the form $\sum_{j=1}^{s}\lambda_{i_j} \equiv 0 \pmod{2\pi}$ where $(i_j,\ j = 1,\dots, s)$ is a proper subset of $\{1, 2,\dots, k\}$. Let $\hat{F}_n^{(k)}(\lambda_1,\dots,\lambda_k)$ be defined the same as $F_n^{(k)}(\lambda_1,\dots,\lambda_k)$, given by (2.11), except that the sum in (2.11) is now only over $(2\pi r_1/n,\dots,2\pi r_k/n)$ which are contained in $\Omega^{(k)}$.

By the construction of $\theta(\hat{F}_n^{(k)})$, the summation in (2.16) will be void (since the sum is over further submanifolds) and, consequently,

$$E\left[\theta\left(\hat{F}_n^{(k)}\right)\right] = \theta(F^{(k)}) + O(n^{-1}).$$

Alternatively, by constructing consistent estimators of all possible integrals w.r.t. products of cumulant spectra (of various orders), e.g.,

$$(2.17)\quad \int_{(S^1)^k} h(\xi_1,\dots,\xi_k) \prod_{j=1}^{p} f^{(\alpha_j)}(\xi_{j_i}, i = 1,\dots,\alpha_j) \prod_{j=1}^{p} \eta\left(\sum_{i=1}^{\alpha_j} \xi_{j_i}\right) \prod_{l=1}^{k} d\xi_l,$$

where $\alpha_j = |\nu_j|$, $\sum_{i=1}^{\alpha_j} \xi_{j_i} \equiv 0 \pmod{2\pi}$, $\sum_{j=1}^{p} \alpha_j = k$, we can (following the next lemma) construct consistent estimators of $\theta(F^{(k)})$ (and, in particular, of $F^{(k)}(\xi_1, \xi_2, \dots, \xi_k)$). We will first consider the latter approach.

Definition. If $\mu = (\mu_1, \mu_2, \dots, \mu_l)$ and $\nu = (\nu_1, \nu_2, \dots, \nu_p)$, $l \geq p$, are partitions, then μ is said to be finer than ν if for every $i \leq j \leq p$, there exist members of μ whose union is ν_j. The greatest integer function is denoted by $[\cdot]$.

Lemma 2.3. *If $\{X_n\}$ is a zero-mean, strictly stationary process satisfying Assumption* 1, *$h(\cdot)$ is a $\mathbb{C}^r$-function of Hobson bounded variation on $[0, 2\pi]^{k-1}$, $p, k \in \mathbb{N}^+$, $p \leq [k/2]$, $\nu = (\nu_1, \dots, \nu_p)$ is a partition of $\{1, 2, \dots, k\}$, $\sum_{i=1}^{\alpha_j}(2\pi r_{j_i}/n) \equiv 0 \pmod{2\pi}$, where $|\nu_j| = \alpha_j$, then*

$$E\left[\frac{1}{n^{k-p}} \sum_{r_1=1}^{n} \sum_{r_2=1}^{n} \cdots \sum_{r_{k-1}=1}^{n} h\left(\frac{2\pi r_1}{n}, \dots, \frac{2\pi r_k}{n}\right) \prod_{j=1}^{p} I_n^{(\alpha_j)}\left(\frac{2\pi r_{j_i}}{n}, j_i \in \nu_j\right)\right]$$

$$(2.18)\quad \begin{aligned} &= \sum_{l=p}^{[k/2]} \sum_{\mu=(\mu,\dots,\mu_l)} \left[\int_{(S^1)^k} h(\beta_1, \beta_2, \dots, \beta_k)\right. \\ &\qquad \times \left\{\prod_{j=1}^{l} \eta\left(\sum_{i=1}^{\delta_j} \beta_{j_i}\right)\right\} \\ &\qquad \left.\times \left\{\prod_{j=1}^{l} f^{(\delta_j)}(\beta_{j_i}, j_i \in \mu_j)\right\} d\beta_1\, d\beta_2 \cdots d\beta_k\right] + O(n^{-1}), \end{aligned}$$

where the summation over μ is over all partitions μ finer than ν and $|\mu_j| = \delta_j > 1$, $j = 1, \dots, l$.

Proof. The left-hand side of (2.18), because there are only $k - p$ freely varying variables, can be written as

$$(2.19)\quad \begin{aligned} &\frac{1}{n^{k-p}}\left\{\sum_{r_1=1}^{n} \sum_{r_2=1}^{n} \cdots \sum_{r_{k-1}=1}^{n} h\left(\frac{2\pi r_1}{n}, \dots, \frac{2\pi r_k}{n}\right)\right. \\ &\qquad \left.\times E\left[n^{-p} \prod_{j=1}^{p} \eta\left\{\sum_{i=1}^{\alpha_j} \frac{2\pi r_{j_i}}{n}\right\} \prod_{j=1}^{\alpha_j} d_n\left(\frac{2\pi r_{j_i}}{n}\right)\right]\right\} \end{aligned}$$

and the term in square brackets is equal to

$$n^{-p}\left\{nf^{(k)}\left(\frac{2\pi r_1}{n},\ldots,\frac{2\pi r_k}{n}\right)+O(1)\right\}$$

$$(2.20)\qquad +n^{-p}\left\{\sum_{l=2}^{p-1} n^l \sum_{\mu=(\mu_1,\ldots,\mu_l)} \prod_{s=1}^{l}\left[f^{(\mu_s)}\left(\frac{2\pi r_{j_i}}{n},\, j_i\in\mu_j\right)+O(1)\right]\right\}$$

$$+n^{-p}\left\{\sum_{l=p}^{[k/2]} n^l \sum_{\mu=(\mu_1,\ldots,\mu_l)} \prod_{s=1}^{l}\left[f^{(\mu_s)}\left(\frac{2\pi r_{j_i}}{n},\, j_i\in\mu_j\right)+O(1)\right]\right\},$$

where $\sum_{i=1}^{\delta_j} 2\pi r_{j_i}/n \equiv 0 \pmod{2\pi}$, $\delta_j = |\mu_j|$ and the summations are over all partitions, μ. The first two terms in (2.20) are $O(n^{-1})$; the remaining term is a sum of terms $O(n^{l-p})$, $p \le l \le [k/2]$. The partitions $\mu = (\mu_1,\ldots,\mu_l)$ place l linear constraints on $\{2\pi r_1/n,\ldots,2\pi r_k/n\}$, and so fewer than $k-p$ of the discrete frequencies can vary freely in (2.11). The largest number of freely varying discrete frequencies is $k-l$ and this can occur if and only if μ is a finer partition than ν. □

DEFINITION OF $\theta(F_n^{(\le k)})$. We can construct estimators of (2.17) with $\alpha_j = 2$ or 3, $j = 1,\ldots,p$, whose bias is $O(n^{-1})$, since there are no finer partitions which need be included in (2.18). We can, therefore, construct estimators of (2.17) with $\alpha_j < 4$ except for one exactly equal to 4, since estimators of terms over finer partitions were constructed in the first step and by induction we can construct, recursively, estimators of $\theta(F^{(k)})$ whose bias is $O(n^{-1})$. This estimator, given by the above recursive construction, will be denoted by $\theta(F_n^{(\le k)})$ because it uses $F_n^{(j)}(\cdot)$, $2 \le j \le k$, in the construction. For $h_\lambda(\cdot)$ defined by

$$(2.21)\qquad h_\lambda(\xi_1,\ldots,\xi_{k-1}) = \begin{cases} 1, & \xi_j \le \lambda_j,\ j = 1,2,\ldots,k-1,\\ 0, & \text{otherwise,}\end{cases}$$

this estimator of $F^{(k)}$ at λ will be denoted by $F_n^{(\le k)}$ at λ and

$$(2.22)\qquad E\big(F_n^{(\le k)}(\lambda_1,\ldots,\lambda_{k-1})\big) = F^{(k)}(\lambda_1,\ldots,\lambda_{k-1}) + O(n^{-1}),$$

where the error term is uniform in $\lambda = (\lambda_1,\ldots,\lambda_{k-1})$.

LEMMA 2.4. *If $\{X_n\}$ is a zero-mean, strictly stationary process satisfying Assumption 1, $\{h_s\}_{s=1}^m$, $m \ge 2$, are complex-valued functions of Hobson bounded variation on $[0,2\pi]^{k-1}$, $\{p_s\}_{s=1}^m$ are integers, $1 \le p_s < k$, such that for each s, $\nu_s = (\nu_1^{(s)},\nu_2^{(s)},\ldots,\nu_{p_s}^{(s)})$ is a partition of $\{1,2,\ldots,k\}$ for which*

$$\sum_{i=1}^{\alpha_j^{(s)}}\left(\frac{2\pi}{n} r_{j_i}^{(s)};\ r_{j_i}^{(s)} \in \nu_j^{(s)}\right) \equiv 0 \pmod{2\pi}, \quad \textit{where } |\nu_j^{(s)}| = \alpha_j^{(s)},$$

and $\sum_{s=1}^{m} p_s = t \geq m$, *then*

$$\operatorname{cum}\Bigg[n^{1/2}\left(\frac{2\pi}{n}\right)^{k-1} \sum_{r_1^{(s)}}^{n} \sum_{r_2^{(s)}}^{n} \cdots \sum_{r_{k-1}^{(s)}=1}^{n} h_s\left(\frac{2\pi r_1^{(s)}}{n}, \ldots, \frac{2\pi r_k^{(s)}}{n}\right) \tag{2.23}$$
$$\times \prod_{j=1}^{p_s} I_n^{(\alpha_j^{(s)})}\left(\frac{2\pi r_{j_i}^{(s)}}{n}; r_{j_i}^{(s)} \in \nu_j^{(s)}\right); 1 \leq s \leq m\Bigg]$$
$$= O(n^{-t+(m/2)+\delta}),$$

where δ *is equal to either zero or one.*

PROOF. Using the fact that the joint cumulant of several linear combinations of random variables is a linear combination of the joint cumulants [Leonov and Shiryaev (1959)] and Lemma 2.1, the summations in (2.23) are

$$\sum_{\mathbf{P}}\left[\sum_{s=1}^{m} \sum_{r_1^{(s)}=1}^{n} \cdots \sum_{r_{k-1}^{(s)}=1}^{n}\right],$$

where the outer sum is over all indecomposable partitions of an $m \times k$ array, and, consequently, the inner sums are constrained to also satisfy, in addition to its t linear constraints, the linear constraints of the given partition $\mathbf{P}$. The argument from here on is the same as that in the proof of Theorem 4, pages 185–186, Brillinger and Rosenblatt (1967a), where their proof is for $p_s = 1$, $1 \leq s \leq m$; this difference has no effect on the basic argument. □

COROLLARY 2.5. *If* $\{X_n\}$ *is a zero-mean, strictly stationary process satisfying Assumption* 1 *and* $\{h_s\}_{s=1}^{m}$, $m \geq 1$, *are Hobson bounded variation, complex-valued functions defining parameters* $\theta_s(F^{(k)})$, $1 \leq s \leq m$, *by expression* (2.4), *then the real and imaginary parts of the random vector,*

$$\left(n^{1/2}\left[\theta_1\left(F_n^{(\leq k)}\right) - \theta_1\left(F^{(k)}\right)\right], \ldots, n^{1/2}\left[\theta_m\left(F_n^{(\leq k)}\right) - \theta_m\left(F^{(k)}\right)\right]\right), \tag{2.24}$$

are asymptotically joint multivariate normal with mean zero and a covariance matrix Λ, *determined by Lemma* 2.4.

PROOF. By the construction (recursive) of each of the functionals $\theta_s(F_n^{(\leq k)})$ and $\bar{\theta}_s(F^{(\leq k)})$, $s = 1, 2, \ldots, m$ (the overbar denotes complex-conjugate), each is a sum of a finite number of terms, each term being of the form of (2.23), with varying values of $\alpha_j^{(s)}$ and p_s, $j = 1, \ldots, p_s$, $s = 1, \ldots, m$. By Lemma 2.4, the joint cumulants of all orders converge, the limits being zero for all orders greater than two. The result now follows from Lemma P4.5, page 403, Brillinger (1975). □

For example, for $k = 3$, the asymptotic covariance of two elements of (2.24), $1 \le a, b \le m$, is

$$\lim_{n\to\infty} \operatorname{Cov}\left(n^{1/2}\theta_a\left(F_n^{(3)}\right), n^{1/2}\theta_b\left(F_n^{(3)}\right)\right)$$

$$(2.25)\qquad = \sum_{\nu} \int_{[0,2\pi]^{\alpha_1-1}} \cdots \int_{[0,2\pi]^{\alpha_p-1}} h_a\left(\lambda_1^{(1)}, \lambda_2^{(1)}, \lambda_3^{(1)}\right) h_b\left(\lambda_1^{(2)}, \lambda_2^{(2)}, \lambda_3^{(2)}\right) \times \prod_{j=1}^{p} f^{(\alpha_j)}\left(\lambda_{j_i};\ j_i \in \nu_j\right) \eta\left(\sum_{i=1}^{\alpha_j} \lambda_{j_i}\right) \prod_{j=1}^{p} \prod_{i=1}^{\alpha_j} d\lambda_{j_i},$$

where $\nu = (\nu_1, \dots, \nu_p)$ varies over all indecomposable partitions of the array

$$(2.26)\qquad \begin{matrix} \lambda_1^{(1)}, \lambda_2^{(1)}, \lambda_3^{(1)} \\ \lambda_1^{(2)}, \lambda_2^{(2)}, \lambda_3^{(2)} \end{matrix}$$

for which the row sums are zero (mod 2π), $\alpha_j = |\nu_j|$, and $1 \le p \le 3$.

An analogous result holds for $F_n^{(\le k)}(\cdot)$ replaced by $\hat{F}_n^{(k)}$; because the submanifolds which introduce biases are left out in the calculation of $\hat{F}_n^{(k)}(\cdot)$, the covariances are much simpler than those of (2.24).

Corollary 2.6. *Under the assumption of Corollary 2.5 with $\theta_j(F_n^{(\le k)})$ replaced by $\theta_j(\hat{F}_n^{(k)})$, $1 \le j \le m$, in (2.24), the real and imaginary parts of the random vector are asymptotically jointly multivariate normal with mean zero and covariances of the complex components given by*

$$n \operatorname{Cov}\left(\theta_a\left(\hat{F}_n^{(k)}\right), \theta_b\left(\hat{F}_n^{(k)}\right)\right)$$

$$(2.27)\qquad = \sum_{\nu} \int_{(S^1)^{2k}} h_a\left(\lambda_1^{(1)}, \dots, \lambda_k^{(1)}\right) \bar{h}_b\left(\lambda_1^{(2)}, \dots, \lambda_k^{(2)}\right) \prod_{i=1}^{s} \eta\left(\sum \lambda_{i_j};\ i_j \in \nu_i\right) \times \prod_{i=1}^{s} f^{(\alpha_i)}\left(\lambda_{i_j};\ i_j \in \nu_i\right) d\lambda^{(1)}\, d\lambda^{(2)},$$

where the sum is over all indecomposable partitions, $\nu = (\nu_1, \dots, \nu_s)$, $|\nu| > 0$, of the $2 \times k$ table

$$\begin{matrix} \lambda_1^{(1)} & \lambda_2^{(1)} & \cdots & \lambda_k^{(1)} \\ -\lambda_1^{(2)} & -\lambda_2^{(2)} & \cdots & -\lambda_k^{(2)} \end{matrix}$$

where the rows are contained in $\Omega^{(k)}$ and such that each component of ν is the union of elements from both rows.

Proof. Same as Corollary 2.5 except that the exclusion of all further submanifolds of $\{\sum_{i=1}^{k} \lambda_i \equiv 0 \pmod{2\pi}\}$ is taken into account. □

As an application of Corollary 2.6 we can consider the estimation of that component of the asymptotic variance [(2.6) above] of the quasi-maximum likelihood estimator which is discussed in Taniguchi (1982). Taniguchi replaces the strict constraint $\eta(\lambda_2+\lambda_3)$ in (2.7) by a spectral window which asymptotically imposes the constraint. Taniguchi's estimator, given by Theorem 2, Taniguchi (1982), involves three sums, the latter two needed to remove the biases due to the inclusion of further submanifolds in the first sum. Our estimator is the first of the three sums where we restrict the summation to points in $\Omega^{(4)}$:

$$\left(\frac{2\pi}{n}\right)^3 \sum_{\Omega^{(4)}} \phi(\lambda_1,\lambda_2,\lambda_3) H_n(\lambda_2+\lambda_3) I_n^{(4)}(\lambda_1,\lambda_2,\lambda_3) \tag{2.28}$$

and where ϕ is given by (2.7), $H_n(x) = B_n^{-1}H(B_n^{-1}x)$, $H(\cdot)$ being a nonnegative function of bounded variation on $\mathbb{R}$, zero outside $[0,2\pi]$, an even function about π, with $B_n \to 0$, $B_n^2 n \to \infty$. The asymptotic variance of the estimator is given by

$$\left\{\int_{-\pi}^{\pi} H(w)\,dw\right\}^2 \sum_{\nu} \int_{(S^1)^4} h(\lambda_1^{(1)},\lambda_2^{(2)})\bar{h}(\lambda_1^{(2)},\lambda_2^{(2)})\eta(\lambda_2^{(1)}+\lambda_3^{(1)})\eta(\lambda_2^{(2)}+\lambda_3^{(2)})$$
$$\times \prod_{i=1}^{s} \eta\Big(\sum \lambda_{i_j};\ i_j \in \nu_i\Big) f^{(\alpha_i)}\big(\lambda_{i_j};\ i_j\in\nu_i\big)\, d\lambda_1^{(1)}\, d\lambda_2^{(1)}\, d\lambda_1^{(2)}\, d\lambda_2^{(2)},$$

where the sum is over the same partitions as in (2.27) with $k = 4$.

Remark 1. Grenander and Rosenblatt (1957) [generalizing a result of Bartlett (1954) for Gaussian white noise] showed that for a linear, Gaussian process (under dependency conditions), the process $\{N^{1/2}[F_n^{(2)}(\lambda) - F^{(2)}(\lambda)], 0 \le \lambda \le \pi\}$ converges weakly (in $D[0,\pi]$) to a Wiener process under transformed time (and were, consequently, able to find the distribution of the sup functional). Their Kolmogorov–Smirnov analogue in this setting cannot delineate between a misspecification of the parameters and non-Gaussianity. If one wishes to assess the appropriateness of a Gaussian assumption, one direction is to expand the perspective to that of $\{n^{1/2}F_n^{(3)}(\lambda),\ \lambda \in [0,2\pi]^2\}$.

The following two corollaries are applications of Corollary 2.5. The first concerns the (partial) construction of a test of Gaussianity of a stationary process; under a Gaussian assumption the weak limit is related to a two-dimensional analogue of a Wiener process under transformed time. The second concerns a test of time reversibility of a process. One advantage which these tests would have (over alternatives) is that natural smoothing is built in and, consequently, the consideration (and need) of spectral windows is removed.

Definition. For $\lambda = (\lambda_1,\lambda_2)$, $\mu = (\mu_1,\mu_2) \in [0,2\pi]^2$, let $r(\lambda) = (\lambda_2,\lambda_1)$, the reversal of order, $\lambda \wedge \mu = (\min\{\lambda_1,\mu_1\}, \min\{\lambda_2,\mu_2\})$, and

$$\Delta = \left\{\lambda = (\lambda_1,\lambda_2) \in [0,2\pi]^2 \middle| \lambda_2 \le \lambda_1,\ \lambda_2 \le 2\pi - 2\lambda_1\right\}.$$

Corollary 2.7. *If $\{X_n\}$ is a zero mean, strictly stationary process satisfying Assumption* 1, *then*

$$\left\{n^{1/2}\left[F_n^{(3)}(\lambda) - F^{(3)}(\lambda)\right],\ \lambda \in [0,2\pi]^2\right\} \tag{2.29}$$

converges weakly [*in* $D([0,2\pi]^2)$] *to a zero-mean process* $\{Y(\lambda),\ \lambda \in [0,2\pi]^2\}$ *whose real and imaginary parts form a two-dimensional Gaussian process with* $\operatorname{Cov}(Y(\lambda), Y(\mu))$, *given by* (2.25), *where* $h_\lambda(\cdot)$ *and* $h_\mu(\cdot)$ *are given by* (2.21).

If $\{X_n\}$ *is a zero-mean, stationary, Gaussian process satisfying Assumption* 1 [*which reduces to* $\sum|\nu|\,|c(\nu)| < \infty$], *and, consequently, linear with spectral density denoted by* $f_{xx}(\cdot)$, *then*

$$\{n^{1/2}F_n^{(3)}(\lambda),\ \lambda \in \Delta\}$$

converges weakly to a zero-mean, real-valued, Gaussian process $\{Y(\lambda),\ \lambda \in \Delta\}$ *with*

$$\operatorname{Cov}(Y(\lambda), Y(\mu)) = H(\lambda \wedge \mu) + H(\lambda \wedge r(\mu)),$$

where

$$H(\lambda) = \int_0^{\lambda_1}\int_0^{\lambda_2} f_{xx}(\alpha) f_{xx}(\beta) f_{xx}(\alpha+\beta)\, d\alpha\, d\beta. \tag{2.30}$$

Remark 2. Under the Gaussian assumption the $\{Y(\lambda),\ \lambda \in \Delta\}$ process can be, equivalently, viewed as

$$Y(\lambda) = 2^{-1/2}\left[W^{0H}(\lambda) + W^{0H}(r(\lambda))\right], \tag{2.31}$$

where W^{0H} (a two-dimensional analogue of a Wiener process under transformed time) is a zero-mean, Gaussian process on $[0,2\pi]^2$ with covariance function

$$\operatorname{Cov}\left(W^{0H}(\lambda), W^{0H}(\mu)\right) = H(\lambda \wedge \mu). \tag{2.32}$$

Although the corollary is given for $k = 3$, the proof for expression (2.27) goes through for arbitrary k.

Proof. By (2.22) $E[F_n^{(3)}(\lambda)] = O(n^{-1})$, uniformly for $\lambda \in [0,2\pi]^2$. Let $Y_n^{(3)} = \{n^{1/2}[F_n^{(3)}(\lambda) - E(F_n^{(3)}(\lambda))],\ \lambda \in [0,2\pi]^2\}$. By Corollary 2.5 the finite distributions converge to zero-mean Gaussian distributions with covariances given by (2.25) where $\lambda, \mu \in [0,2\pi]^2$ and $h_\lambda(\cdot)$ and $h_\mu(\cdot)$ are given by (2.21). By Theorem 3, Bickel and Wichura (1971), a Chentsov-type inequality (as in the one-dimensional case) is sufficient for tightness of the family $\{Y_n^{(3)}(\lambda),\ \lambda \in [0,2\pi]^2\}$. If $B = (\lambda_1', \lambda_1] \times (\lambda_2', \lambda_2]$ and $C = (\mu_1', \mu_1] \times (\mu_2', \mu_2]$ are (disjoint) neighboring rectangles [see Bickel and Wichura (1971)], where $\mu_1 \le \lambda_1$, $\mu_2 \le \lambda_2$, then it suffices to show that there exist a $L > 0$ and a finite measure, ν, on $[0,2\pi]^2$ with continuous marginals such that for all B, C

$$E\left\{\left|V^{(n)}(B)\right|^2\left|V^{(n)}(C)\right|^2\right\} \le L[\nu(B \cup C)]^2, \tag{2.33}$$

where for

$$D = (\xi_1', \xi_2] \times (\xi_1', \xi_2] \subset [0,2\pi]^2,$$

$$\begin{aligned} V^{(n)}(D) &= \left[Y_n^{(3)}(\xi_1, \xi_2) - Y_n^{(3)}(\xi_1, \xi_2') - Y_n^{(3)}(\xi_1', \xi_2) + Y_n^{(3)}(\xi_1', \xi_2')\right] \\ &= n^{1/2}\left(\frac{2\pi}{n}\right)^2 \sum_{r_1}\sum_{r_2}\left[I_n^{(3)}\left(\frac{2\pi r_1}{n}, \frac{2\pi r_2}{n}\right) - E\left(I_n^{(3)}\left(\frac{2\pi r_1}{n}, \frac{2\pi r_2}{n}\right)\right)\right], \end{aligned} \tag{2.34}$$

the summation being over $\{r_1, r_2 | \xi_j' \le 2\pi r_j/n \le \xi_j,\ j = 1,2\}$.

We will take ν to be Lebesgue measure on $[0, 2\pi]^2$. The remainder of the proof is a two-dimensional version of the one-dimensional proof for Theorem 7.6.3, page 439, Brillinger (1975), where Brillinger's reference to Theorem 7.6.1, Brillinger (1975), is replaced by Corollary 2.5 and noting that the $O(1)$ terms (due to the repeated application of Lemma 2.2), after integration, contribute at most $[O(1)(\lambda_1 - \mu_1')(\lambda_2 - \mu_2')]$ to (2.33). Consider, now, the process restricted to the compact subset, Δ, of $[0, 2\pi]^2$. Let $\{\nu_n\}_{n=1}^{\infty}$ be the restriction to $D(\Delta)$ of the distributions on $(D([0, 2\pi]^2), D)$ associated with the $\{Y_n\}$ process; it follows that the $\{\nu_n\}_{n=1}^{\infty}$ are tight on $D(\Delta)$. There are 15 indecomposable partitions of the array (2.26) for which there are three elements of the partition, each containing two members. Because of the constraints defining the set Δ, only two of the integrals in (2.25) will be nonzero in this case; the two are $H(\lambda \wedge \mu)$ and $H(\lambda \wedge r(\mu))$, for $\lambda, \mu \in \Delta$. □

REMARK 3. Since the limiting process is continuous the convergence is in fact w.r.t. the finer topology of uniform convergence, for which the sup functional is continuous. The remaining obstacle to a Kolmogorov–Smirnov type test of a Gaussianity assumption is a closed-form expression (or reasonable approximation) to the sup of the process, $\{W^{0H} + W^{0H}r(\lambda),\ \lambda \in [0, 2\pi]^2\}$, defined in Remark 2. Also, if one employs the integrated third-order periodogram instead of $F_n(\cdot)$, then $D^2[0, 2\pi]$ in Corollary 2.7, could be reduced beyond $C^2[0, 2\pi]$ to appropriate Lipschitz spaces [see Brillinger (1969)], with topologies finer than uniform convergence, with an expanded space of applicable continuous (a.e.) functionals (of the process) whose distributions one can consider.

A strictly stationary process, $\{X_s\}_{-\infty}^{\infty}$, is said to be time reversible if for all k and $(t_1, \ldots, t_k)$, $(X_{t_1}, \ldots, X_{t_k})$ and $(X_{-t_1}, \ldots, X_{-t_k})$ have the same distribution. Consequently, for a real-valued, time reversible process, all of the cumulant spectra ($k \geq 2$) are real-valued (if they exist). Consequently, the following application of Corollary 2.7 gives the framework for a potential test of time reversibility. Let $\mathrm{Im}(\cdot)$ be the function from $\mathbb{C}$ to $\mathbb{R}$ which maps the imaginary part.

COROLLARY 2.8. *If $\{X_n\}$ is a zero-mean, strictly stationary time reversible process satisfying Assumption* 1, *then*

$$\left\{n^{1/2}\mathrm{Im}\left(F_n^{(3)}(\lambda)\right),\ \lambda \in [0, 2\pi]^2\right\}$$

converges weakly to a zero-mean, real-valued Gaussian process whose covariances are given by the real part of (2.25) *with h_a and h_b given by $h_\lambda(\cdot)$ and $h_\mu(\cdot)$ defined by* (2.21).

PROOF. Define $\mathrm{Im}(\cdot)$ from $D^{\mathbb{C}}[0, 2\pi]^2$ to $D^{\mathbb{R}}[0, 2\pi]^2$ as

$$\mathrm{Im}\left(Z(t),\ t \in [0, 2\pi]^2\right) = \left(\mathrm{Im}(Z(t)),\ t \in [0, 2\pi]^2\right).$$

Since $\mathrm{Im}(\cdot)$ is a continuous mapping between the two metric spaces and since the imaginary components of the finite distribution of the complex-valued process,

$Y(\cdot)$, given by Corollary 2.7, are real-valued multivariate normal whose covariances are the real part of the complex valued covariances, the result follows. □

3. Probability 1 bounds and almost sure convergence. In this section an additional assumption will be made concerning the cumulants of the X_n process. The assumption is Assumption 7.7.2 of Brillinger (1975), page 264. We will assume for $j \in \mathbb{N}$

ASSUMPTION 2. $C_j = \sum_{\nu_1, \nu_2, \dots, \nu_{j-1} = -\infty}^{\infty} |c(\nu_1, \dots, \nu_{j-1})|$, is finite where $c(\nu_1, \dots, \nu_{j-1})$ is the jth-order cumulant of $\{X(0), X(\nu_1), \dots, X(\nu_{j-1})\}$. We will also assume that for a fixed $k \in \mathbb{N}^+$ $(k > 2)$

$$\sum_{L=1}^{\infty} \sum_{\nu} \left(C_{n_1} C_{n_2} \cdots C_{n_p}\right) \frac{Z^L}{L!} < \infty, \tag{3.1}$$

for Z in a neighborhood of zero, where the inner summation is over all indecomposable partitions $\boldsymbol{\nu} = (\nu_1, \nu_2, \dots, \nu_p)$ of the table

$$\begin{matrix} 1 & 2 & \cdots & k \\ k+1 & k+2 & \cdots & 2k \\ \vdots & \vdots & & \vdots \\ k(L-1)+1 & k(L-1)+2 & \cdots & kL \end{matrix} \tag{3.2}$$

with ν_j having $n_j > 1$ elements, $j = 1, \dots, p$. The next lemma uses an approach of Keenan (1983) which itself was based upon an approach of Brillinger (1975) for obtaining probability one bounds. We will show that

$$\sup_{(\lambda_1, \lambda_2, \dots, \lambda_{k-1})} \left| F_n^{(\leq k)}(\lambda_1, \lambda_2, \dots, \lambda_{k-1}) - F^{(k)}(\lambda_1, \lambda_2, \dots, \lambda_{k-1}) \right| \tag{3.3}$$

is $O(n^{-1/2}(\log n)^{1/2})$, w.p.1. We will then use this result to show that this same rate passes through the integration of well-behaved functions w.r.t. $d[F_n^{(\leq k)}(\cdot) - F^{(k)}(\cdot)]$ and, consequently, is inherited by our estimators.

THEOREM 3.1. *For a fixed $k \in \mathbb{N}^+$ $(k > 2)$, if Assumptions 1 and 2 are satisfied, then $\|F_n^{(\leq k)} - F^{(k)}\|_\infty = O(n^{-1/2}(\log n)^{1/2})$ almost surely. The same result holds for $F_n^{(\leq k)}$ replaced by $\hat{F}_n^{(k)}$.*

PROOF. The proof is analogous to that of Keenan (1983), Lemma 4.1, except that $\lambda \in [0, 2\pi]$ is replaced by $(\lambda_1, \lambda_2, \dots, \lambda_{k-1}) \in [0, 2\pi]^{k-1}$, the n jumps in $[F_n^{(2)}(\cdot) - EF_n^{(2)}(\cdot)]$ are replaced by n^{k-1} jumps in $[F_n^{(\leq k)}(\cdot) - E(F_n^{(\leq k)}(\cdot))]$ in the present proof; consequently, the n in the bound given by expression (4.4) in Keenan (1983) is replaced by n^{k-1}. The functions $c(n)$, $\alpha(n)$ and $a(n)$ are changed by having $(2 + \delta)$ replaced everywhere by $(k + \delta)$. By (2.22), above, $|E(F_n^{(\leq k)}(\lambda_1, \dots, \lambda_{k-1}) - F^{(k)}(\lambda_1, \dots, \lambda_{k-1}))| = O(n^{-1})$ uniformly in $(\lambda_1, \dots, \lambda_{k-1})$ and the result follows. □

The following corollary is an application of Theorem 3.1. An estimate of the kth-order joint cumulant at $\mathbf{j} = (j_1, \dots, j_{k-1})$, analogous to the circular autocovariances, is

$$\hat{c}_n(j) = \int_{[0,2\pi]^{k-1}} e^{i(\mathbf{j}\cdot\boldsymbol{\lambda})}\, dF_n^{(\le k)}(\boldsymbol{\lambda}). \tag{3.4}$$

COROLLARY 3.2. *Under Assumptions* 1 *and* 2, *if* $J_n = o(n^{1/2}(\log n)^{-1/2})$ *is an increasing sequence of positive integers, then for* $\mathbf{j} = (j_1, j_2, \dots, j_{k-1})$

$$\sup_{\{\mathbf{j} \mid |j_s| \le J_n,\, 1 \le s \le k-1\}} |\hat{c}_n(j_1, \dots, j_{k-1}) - c(j_1, \dots, j_{k-1})| = o(1) \quad w.p.1. \tag{3.5}$$

PROOF. The function $[F_n^{(\le k)}(\cdot) - E(F_n^{(\le k)}(\cdot))]$ has jumps at the n^{k-1} discrete Fourier frequencies and $\|EF_n^{(\le k)} - F^{(k)}\|_\infty = o(n^{-1/2})$. For the moment consider a given realization. Define $M_n^{(k)}(\cdot)$ as the piecewise linear version of $[F_n^{(\le k)}(\cdot) - E(F_n^{(\le k)}(\cdot))]$; the sup over any rectangle with endpoints at discrete Fourier frequencies has not changed. The real and imaginary parts of $M_n^{(k)}(\cdot)$ are the differences of two functions of normalized bounded variation (N.B.V.) and the same symbol, $M_n^{(k)}$, will also be used to represent the corresponding Borel complex measure on $[0, 2\pi]^{k-1}$. Since $M_n^{(k)}$ is absolutely continuous w.r.t. Lebesgue measure on $[0, 2\pi]^{k-1}$, denote by

$$\left[\frac{dM_n^{(k)}}{d\xi}\right]$$

the corresponding Radon–Nikodym derivative and define $m_n(\cdot)$ as

$$m_n(\lambda_1, \lambda_2, \dots, \lambda_{k-1}) = \int_0^{\lambda_1} \cdots \int_0^{\lambda_{k-1}} \left[\frac{dM_n^{(k)}}{d\xi}(\xi)\right] d\xi. \tag{3.6}$$

Since $m_n(\cdot)$ is differentiable Lebesgue-almost everywhere (w.r.t. r distinct components, $1 \le r \le k-1$), by induction we have

$$\begin{aligned} &\frac{\partial^{k-1}}{\partial\lambda_1 \cdots \partial\lambda_{k-1}}\left(e^{i(\mathbf{j}\cdot\boldsymbol{\lambda})} m_n(\boldsymbol{\lambda})\right) \\ &\quad = \sum_{r=0}^{k-1} \sum_{\binom{k-1}{r}} \frac{\partial(r) e^{i}(\mathbf{j}\cdot\boldsymbol{\lambda})}{\partial\lambda_{l_1} \cdots \partial\lambda_{l_r}} \frac{\partial^{(k-1-r)} m_n(\boldsymbol{\lambda})}{\partial\lambda_{l_{r+1}} \cdots \partial\lambda_{l_{k-1}}}, \end{aligned} \tag{3.7}$$

where the inner sum is over all $\binom{k-1}{r}$ choice of r components. Since the term being differentiated on the left-hand side of (3.7) has real and imaginary parts which are the differences of functions of N.B.V., its integral w.r.t. Lebesgue measure on $[0, 2\pi]^{k-1}$ is

$$M_n^{(k)}([0, 2\pi]^{k-1}),$$

which by Theorem 3.1 is $o(1)$, w.p.1. Since $\partial^{(k-1-r)} m_n(\boldsymbol{\lambda})/\partial\lambda_{l_{r+1}} \cdots \partial\lambda_{l_{k-1}}$, $1 \le r \le k-1$, is of N.B.V. and its absolute value is equal (w.p.1) to the Radon–Nikodym derivative (w.r.t. Lebesgue measure) of the absolute value

Borel measure associated with $m_n(\cdot)$, we have

$$\left| \int_{[0,2\pi]^{k-1}} \left(\frac{\partial^{(r)} e^i(\mathbf{j} \cdot \lambda)}{\partial \lambda_{l_1} \cdots \partial \lambda_{l_r}} \right) \left(\frac{\partial^{(k-1-r)} m_n(\lambda)}{\partial \lambda_{l_{r+1}} \cdots \partial \lambda_{l_{k-1}}} \right) d\lambda \right|$$
$$\leq |j_{l_1} \cdot j_{l_2} \cdots j_{l_r}| \|F_n^{(\leq k)} - EF_n^{(\leq k)}\|_\infty,$$

and, consequently,

$$\sup_{\{\mathbf{j} \mid |j_s| \leq \mathcal{J}_n,\, 1 \leq s \leq k-1\}} |\hat{c}_n(j) - c(j)| = \left| \int_{[0,2\pi]^{k-1}} e^{i(\mathbf{j} \cdot \lambda)} \, d\left[F_n^{(\leq k)} - F^{(\leq k)}\right](\lambda) \right|$$

and by Theorem 3.1

$$\leq o(1) + \mathcal{J}_n O\left(\|F_n^{(\leq k)} - F^{(\leq k)}\|_\infty\right) = o(1) \quad \text{w.p.1.} \qquad \square$$

The following theorem establishes probability one bounds and, consequently, almost sure convergence for our estimators, $\theta(F_n^{(\leq k)})$.

THEOREM 3.3. *For $k > 2$ $(\in \mathbb{N}^+)$, $\theta_j(F_n^{(\leq k)})$, $\theta_j(F^{(k)})$, $1 \leq j \leq q$, defined by* (2.4) *and* (2.21) *with $h_j(\cdot)$ being of Hobson bounded variation, if Assumptions* 1 *and* 2 *are satisfied, then*

$$\limsup_{n \to \infty} \frac{n^{1/2} \left| \theta_j\left(F_n^{(\leq k)}\right) - \theta_j\left(F^{(k)}\right) \right|}{\left(2\Lambda_{jj} \log n\right)^{1/2}} \leq 1, \tag{3.8}$$

almost surely where Λ_{jj} is the (j, j)th element (*assumed to be positive*) *of the matrix Λ defined by Corollary* 2.4. *The same result holds for $f_n^{(\leq k)}$ replaced by $\hat{F}_n^{(k)}$ and Λ given by* (2.27).

PROOF. The proof uses Theorem 3.1 in the same manner that Theorem 4.4, Keenan (1983) used Lemma 4.1, Keenan (1983), which was for the case $k = 2$. The proof is parallel to that of Theorem 4.4, Keenan (1983), except that

$$n\left[F_n^{(\leq k)}(\lambda_1, \lambda_2, \ldots, \lambda_{k-1}) - F^{(k)}(\lambda_1, \ldots, \lambda_{k-1})\right]$$
$$= n\left(\frac{2\pi}{n}\right)^{k-1} \sum_{(2\pi s_j/n) \leq \lambda_j,\, j=1,\ldots,k-1} \left[I_n^{(k)}\left(\frac{2\pi s_1}{n}, \ldots, \frac{2\pi s_{k-1}}{n}\right) - f^{(k)}\left(\frac{2\pi s_1}{n}, \ldots, \frac{2\pi s_{k-1}}{n}\right)\right] + O(1)$$

is replaced by

$$H_n = n\left\{\left(\frac{2\pi}{n}\right)^{k-1} \sum_{s_1,\ldots,s_{k-1}=1}^{n} h\left(\frac{2\pi s_1}{n}, \ldots, \frac{2\pi s_{k-1}}{n}\right) \times \left[I_n\left(\frac{2\pi s_1}{n}, \ldots, \frac{2\pi s_{k-1}}{n}\right) - f^{(k)}\left(\frac{2\pi s_1}{n}, \ldots, \frac{2\pi s_{k-1}}{n}\right)\right]\right\} + O(1),$$

where h and $f^{(k)}$ are of bounded variation and the difficulty with the supremum over $[0,2\pi]^{k-1}$ is avoided since we are smoothing over all of $[0,2\pi]^{k-1}$. The $\alpha(n)$, $c(n)$ and $a(n)$ are the same as in Theorem 3.1, except that $(k+\delta)$ is replaced by $(1+\delta)$. □

4. Additional applications and summary. This paper is concerned with establishing a general framework for estimators of parametrizations reflecting more than second-order structure. Asymptotic normality and probability one bounds (and, consequently, almost-sure convergence) are established for a broad class of such estimators.

Hosoya and Taniguchi (1982) consider quasi-maximum likelihood estimation (i.e., minimization of a specific integral expression). The asymptotic variance of such estimators contains an integral of a function w.r.t. the fourth-order cumulant spectra of the process. Taniguchi (1982) develops consistent estimators of such integrals, the integration is over a two-dimensional submanifold of $(S^1)^4$. Expression (2.28) defines a modification of Taniguchi's estimator, obtaining a simpler form. Taniguchi's result allows for the construction of confidence intervals in these cases. For linear processes, estimators (of parameters) defined via a minimization procedure (under regularity conditions) will not contain this integral term because the fourth-order cumulant spectra factorizes as a product of second-order spectra; otherwise there will typically be such a term. In Keenan (1985a) estimators are defined which minimize a general integral expression (including quasi-maximum likelihood). Taniguchi's result (1982) and a similar modification as above produce consistent estimators of their asymptotic variances and consequently allow for the construction of confidence intervals for the parameter under consideration. Theorem 3.3 gives a probability one bound on the convergence of the estimated integrals.

Weiss (1975) and Cox (1981) have proposed diagnostics for time irreversibility of a strictly stationary process [see (2.8) and (2.9)]. Corollary 2.5 and Theorem 3.3 give the asymptotic normality and probability one bounds for the sample versions of (2.8) used in testing time reversibility. An alternative to these time domain diagnostics would be a third-order sample integrated cumulant spectra diagnostic suggested by Corollary 2.8.

Corollary 2.7 is a first step toward a Kolmogorov–Smirnov type test for Gaussianity of a strictly stationary process. The asymptotic distribution of the sup norm for (2.29) is (typically) not known. In the analogous situation, i.i.d. observations with a two-dimensional marginal distribution, Kiefer and Wolfowitz (1958) have obtained bounds on the tail probabilities. A similar approach may be possible here.

Lastly, Keenan (1985b) has proposed a test of time series nonlinearity whose statistics (under the alternative hypotheses) fit the framework of this paper. Because Volterra expansions are to linear processes what higher-order polynomials are to linear functions, a diagnostic similar to Tukey's one degree of freedom for nonadditivity test is developed.

REFERENCES

BARTLETT, M. S. (1954). Problemes de l'analyse spectrale des series temporelles stationaires. *Publ. Inst. Statist. Univ. Paris* **III-3** 119–134.

BICKEL, P. J. and WICHURA, M. J. (1971). Convergence criteria for multi-parameter stochastic processes and some applications. *Ann. Math. Statist.* **42** 1656–1670.

BRILLINGER, D. R. (1969). Asymptotic properties of spectral estimates of second-order. *Biometrika* **56** 375–390.

BRILLINGER, D. R. (1975). *Time Series: Data Analysis and Theory.* Holt, Rinehart and Winston, New York.

BRILLINGER, D. R. and ROSENBLATT, M. (1967a). Asymptotic theory of k-th order spectra. In *Spectral Analysis of Time Series* (B. Harris, ed.) 153–188. Wiley, New York.

BRILLINGER, D. R. and ROSENBLATT, M. (1967b). Computation and interpretation of k-th order spectra. In *Spectral Analysis of Time Series* (B. Harris, ed.) 189–232. Wiley, New York.

COX, D. R. (1981). Statistical analysis of time series: some recent developments. *Scand. J. Statist.* **8** 93–115.

GRENANDER, U. and ROSENBLATT, M. (1957). *Statistical Analysis of Stationary Time Series.* Wiley, New York.

HOBSON, E. W. (1927). *The Theory of Functions of a Real Variable and the Theory of Fourier's Series* **1**, 3rd ed. Cambridge Univ. Press.

HOSOYA, Y. and TANIGUCHI, M. (1982). A central limit theorem for stationary processes and the parameter estimation of linear processes. *Ann. Statist.* **10** 132–153.

KEENAN, D. M. (1983). Limiting behavior of functionals of the sample spectral distribution. *Ann. Statist.* **11** 1206–1217.

KEENAN, D. M. (1985a). Asymptotic properties of minimization estimators of time series parameters. *Ann. Statist.* **13** 369–382.

KEENAN, D. M. (1985b). A Tukey nonadditivity-type test for time series nonlinearity. *Biometrika* **72** 39–44.

KIEFER, J. and WOLFOWITZ, J. (1958). On the deviations of the empiric distribution function of vector chance variables. *Trans. Amer. Math. Soc.* **87** 173–186.

LEONOV, V. P. and SHIRYAEV, A. N. (1959). On a method of calculation of semi-invariants. *Theory Probab. Appl.* 4 319–329.

TANIGUCHI, M. (1982). On estimation of the integrals of the fourth order cumulant spectral density. *Biometrika* **69** 117–122.

TUKEY, J. W. (1953). The spectral representation and transformation properties of the higher moments of stationary time series. Paper 4 in *The Collected Works of John W. Tukey* **1**. Wadsworth, Belmont, Calif. (1984).

WEISS, G. H. (1975). Time-reversibility of linear stochastic processes. *J. Appl. Probab.* **12** 831–836.

DIVISION OF APPLIED MATHEMATICS
BROWN UNIVERSITY
PROVIDENCE, RHODE ISLAND 02912