ELSEVIER

Signal Processing 47 (1995) 1–17

SIGNAL PROCESSING

Cumulant-based autocorrelation estimates of non-Gaussian linear processes[1]

Georgios B. Giannakis[a,*], Anastasios Delopoulos[b]

[a] *Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22903-2442, USA*
[b] *Department of Electrical Engineering, National Technical University of Athens, Zografou, GR-15773, Greece*

Received 28 January 1994; revised 18 January 1995

Abstract

Autocorrelation of linear random processes can be expressed in terms of their cumulants. Theoretical insensitivity of the latter to additive Gaussian noise of unknown covariance, is exploited in this paper to develop (within a scale) autocorrelation estimators of linear non-Gaussian time series using cumulants of order higher than two. Windowed projections of third-order cumulants are shown to yield strongly consistent estimators of the autocorrelation sequence. Both batch and recursive algorithms are derived. Asymptotic variance expressions of the proposed estimators are also presented. Simulations are provided to illustrate the performance of the proposed algorithms and compare them with conventional approaches.

Zusammenfassung

Die Autokorrelation linearer Zufallsprozesse kann durch deren Kumulanten ausgedrückt werden. In dieser Arbeit wird die theoretische Unempfindlichkeit der Kumulanten gegenüber additivem Gaußschem Rauschen unbekannter Kovarianz ausgenützt, um (bis auf einen Skalierungsfaktor) Autokorrelationsschätzer für lineare nicht-Gaußsche Zeitreihen zu entwickeln, die Kumulanten mit Ordnung größer als 2 verwenden. Es wird gezeigt, daß gefensterte Projektionen von Kumulanten dritter Ordnung streng konsistente Schätzer der Autokorrelationsfolge ergeben. Es werden sowohl Blockalgorithmen als auch rekursive Algorithmen abgeleitet. Weiters werden asymptotische Ausdrücke für die Varianz der vorgeschlagenen Schätzer präsentiert. Schließlich wird mittels Simulationen die Leistungsfähigkeit der vorgeschlagenen Algorithmen gezeigt und ein Vergleich dieser Algorithmen mit konventionellen Methoden durchgeführt.

Résumé

L'autocorrélation des processus linéaires aléatoires peut être exprimée en fonction de leurs cumulants. L'insensibilité théorique de ces derniers au bruit additif gaussien de variance inconnue est exploitée dans cet article pour développer (à une échelle près) des estimateurs de l'autocorrélation de séries temporelles linéaires non gaussiennes utilisant des cumulants d'ordre supérieur à deux. Nous montrons que des projections fenêtrées de cumulants d'ordre trois produisent des estimateurs fortement consistants de la séquence d'autocorrélation. Nous dérivons à la fois des algorithmes récursifs

* Corresponding author. Tel: (804) 924-3659. E-mail: georgios@virginia.edu.

[1] Results from this paper were presented at the SPIE Conf., San Diego, CA, 10–12 July 1990.

SSDI 0165-1684(95)00095-X

et par blocs. Nous présentons également des expressions de la variance asymptotique des estimateurs proposés. Des simulations illustrent les performances des algorithmes proposés et permettent de les comparer à celles des approches conventionnelles.

Keywords: Non-Gaussian time series: Autocorrelation estimation; Cumulants; Consistency; Asymptotic covariance

1. Introduction

In signal processing applications such as sonar and radar, non-Gaussian time series of interest are observed in additive Gaussian noise (AGN) of *unknown* covariance. Often, the AGN is assumed independent of the information signal, stationary, and colored (i.e., temporally correlated). In such cases, spectral analysis fails to characterize the noise-free signal of interest because the available parametric and non-parametric approaches provide consistent spectral estimators of the signal plus noise process, instead of the signal only component. As an alternative to spectral analysis, there has been a recent interest towards cumulants and polyspectra, primarily because in this context the latter separate Gaussian from non-Gaussian additive components, and characterize completely non-Gaussian linear processes. Despite the fact that 3rd- and 4th-order cumulant estimators usually exhibit higher variance than autocorrelation estimators, given sufficiently long data records they have been successfully applied to phase estimation, deconvolution, system identification of mixed phase models and array processing (see [3–8, 16–18, 24] and references therein).

Although it is well-known that spectral information is contained in polyspectral slices of linear processes, to the best of the authors knowledge, the time-domain relationships between 2nd- and higher-order statistics have not been thoroughly studied, especially in the presence of AGN. These relationships originally presented in [13] are used in this paper to construct non-parametric, autocorrelation estimators, which in their naive form are inconsistent. One of the principal goals of this work is to establish consistency and develop asymptotic variance expressions which allow performance analysis and comparisons with the conventional autocorrelation estimators. The development focuses on estimating autocorrelations via 3rd-order cumulants (see [21] for frequency-domain counterpart). We are motivated by the following reasons:

- Computation of 2nd- from higher-order statistics, offers the potential of noise suppression in detection, [16], and spectral estimation of linear, non-Gaussian time series observed in AGN [9, 10]. Parametric modeling assumptions are not necessary.
- The MA and ARMA identification algorithms of [11–13, 27], employ both autocorrelation and 3rd- or 4th-order cumulants, and as such are sensitive to AGN (see also [18]). Estimating the second-order statistics via higher-order statistics guarantees consistency of the ARMA parameter estimators, even in the presence of colored AGN of unknown spectral characteristics.
- Parametric cumulant-based algorithm have been reported for time-delay estimation [19], harmonic retrieval, echo cancellation and bearing estimation [24]. Two of the reasons for using cumulants in array processing are: (i) to retrieve more sources than sensors [23], and (ii) to combat AGN. For the latter, one can use cumulants to obtain 'clean' autocorrelation estimates, and then follow standard autocorrelation-based approaches.

The layout of the paper is as follows. Section 2 deals with relations between 2nd- and higher-order statistics. Strongly consistent autocorrelation estimators along with their asymptotic variance are obtained in Section 3 using windowed projections of 3rd-order cumulants. A recursive algorithm for estimating autocorrelation lags through 3rd-order cumulants is derived in Section 4. Both batch and recursive autocorrelation estimators are found within a scale ambiguity. The latter poses no problem for parameter estimation based on linear equations of the Yule–Walker type, or, optimization of non-linear matching criteria, because in both cases the solutions are autocorrelation scale invariant. In other words, this paper's methods yield normalized 'AGN-insensitive' autocorrelation estimates of non-Gaussian linear processes.

The simulations described in Section 5, demonstrate the potential of higher-order statistics for spectral analysis of signals corrupted by low signal-to-noise-ratio (SNR) AGN of unknown covariance.

2. Cumulants of linear processes and the proposed estimator

Let $h(i)$ denote an absolutely summable sequence with kth-order correlation given by

$$h_k(m_1, \dots, m_{k-1}) \triangleq \sum_{n=-\infty}^{\infty} h(n)h(n+m_1) \cdots h(n+m_{k-1}). \quad (1)$$

To relate $h_k(m_1, \dots, m_{k-1})$ with $h_l(m_1, \dots, m_{l-1})$ when $k > l$, we modulate and project the former over the indices $m_l, \dots, m_{k-1}$ to obtain

$$\sum_{m_l} \cdots \sum_{m_{k-1}} \exp\left\{-\mathrm{j} \sum_{j=1}^{k-1} \omega_j m_j\right\} h_k(m_1, \dots, m_{k-1})$$
$$= \sum_{n=-\infty}^{\infty} h(i) \cdots h(i+m_{l-1}) \times \left[\sum_{m_l} h(n+m_l) \exp\{-\mathrm{j}\,\omega_l m_l\}\right] \cdots \left[\sum_{m_{k-1}} h(n+m_{k-1}) \exp\{-\mathrm{j}\,\omega_{k-1} m_{k-1}\}\right]$$
$$= \sum_{n=-\infty}^{\infty} h(n) \cdots h(n+m_{l-1}) [H(\omega_l) \exp\{\mathrm{j} n \omega_l\}] \cdots [H(\omega_{k-1}) \exp\{\mathrm{j} n \omega_{k-1}\}]$$
$$= \prod_{s=1}^{k-1} H(\omega_s) \exp\left\{\mathrm{j} n \sum_{s=1}^{k-1} \omega_s\right\} \times \sum_{n=-\infty}^{\infty} h(n) \cdots h(n+m_{l-1})$$
$$= \left[\prod_{s=1}^{k-1} H(\omega_s)\right] h_l(m_1, \dots, m_{l-1}), \quad (2)$$

on any manifold $\sum_{r=l}^{k-1} \omega_r = 0$, where $\mathrm{j} \triangleq \sqrt{-1}$ and $H(\omega) \triangleq \sum_{n=-\infty}^{\infty} h(n) \mathrm{e}^{-\mathrm{j}\omega n}$.

Eq. (2) shows that deterministic correlations of a given order can be obtained, within a scale factor, from weighted projections of their high-order correlations. When $H(0) \neq 0$, of practical interest are the following relations between 2nd-, 3rd- and 4th-order correlations which are easily obtained from (2):

$$h_2(m) = [H(0)]^{-1} \sum_{n=-\infty}^{\infty} h_3(m,n), \quad (3a)$$

$$h_3(m,n) = [H(0)]^{-1} \sum_{l=-\infty}^{\infty} h_4(m,n,l), \quad (3b)$$

$$h_2(m) = [H(0)]^{-2} \sum_{n,l=-\infty}^{\infty} h_4(m,n,l). \quad (3c)$$

To avoid the case $H(0) = 0$, one may use weighted projections; for example, 2nd- with 4th-order correlations are related by

$$h_2(m) = |H(\omega_0)|^{-2} \sum_{n,l=-\infty}^{\infty} \mathrm{e}^{\mathrm{j}\omega_0(n-l)} h_4(m,n,l), \quad (4)$$

for real $h(n)$. Clearly in (4), $H(\omega_0)$ cannot be zero for all ω_0.

Let now $x(n)$ be a linear time series generated by a zero-mean, non-Gaussian, and independent, identically distributed (i.i.d.) input $w(n)$,

$$x(n) = \sum_{i=-\infty}^{\infty} h(i) w(n-i). \quad (5a)$$

If $w(n)$ has finite moments up to order k, and $h(i)$ is absolutely summable, then the kth-order output cumulant, $c_{kx}(m_1, \dots, m_{k-1})$, is absolutely summable, and can be expressed as a scalar multiple of the kth-order correlation $h_k(m_1, \dots, m_{k-1})$. Specifically, it holds that [2],

$$\sum_{m_1, \dots, m_{k-1}=-\infty}^{\infty} |c_{kx}(m_1, \dots, m_{k-1})| = C_{kx} < \infty \quad (5b)$$

and

$$c_{kx}(m_1, \dots, m_{k-1}) = \gamma_{kw} h_k(m_1, \dots, m_{k-1}), \quad (5c)$$

where $h_k(m_1, \dots, m_{k-1})$ is given by (1), and γ_{kw} stands for the kth-order cumulant of $w(n)$.

Combining (2) with (5c) it follows that

$$c_{lx}(m_1, \ldots, m_{l-1}) = \left[\frac{\gamma_{lw}}{\gamma_{kw}} \prod_{i=l}^{k-1} H^{-1}(\omega_i)\right]$$

$$\times \sum_{m_l, \ldots, m_{k-1}} c_{ks}(m_1, \ldots, m_{k-1}), \quad k > l, \tag{6}$$

provided that $H(\omega_i) \neq 0$ for $i = l, \ldots, k-1$ and $\sum_{i=l}^{k-1} \omega_i = 0$.

If $w(n)$ is non-Gaussian and i.i.d. with finite moments, then for at least one k, $c_{kw}(m_1, \ldots, m_{k-1}) = \gamma_{kw}\delta(m_1, \ldots, m_{k-1}) \neq 0$, where $\delta(m_1, \ldots, m_{k-1})$ denotes the $(k-1)$-dimensional Kronecker-delta function. On the contrary, for every stationary Gaussian process $v(n)$ it holds that [2]

$$c_{kv}(m_1, \ldots, m_{k-1}) \triangleq 0 \quad \text{for } k \geqslant 3, \tag{7a}$$

which explains why higher- than 2nd-order cumulants are insensitive to AGN. Indeed, if the non-Gaussian linear process $x(n)$ is observed in zero-mean AGN $v(n)$, which is independent of $w(n)$, then the noisy time series

$$y(n) = x(n) + v(n), \tag{7b}$$

has kth-order cumulant given by

$$\begin{aligned} c_{ky}&(m_1, \ldots, m_{k-1}) \\ &= c_{kx}(m_1, \ldots, m_{k-1}) + c_{kv}(m_1, \ldots, m_{k-1}) \\ &= \gamma_{kw} h_k(m_1, \ldots, m_{k-1}). \end{aligned} \tag{7c}$$

Eqs. (7a)–(7c) form the basis of the subsequent autocorrelation estimation algorithms which employ higher than 2nd-order cumulants. To verify the latter, set $k = 3$, $l = 2$ in (6) and use (7a)–(7c) to obtain

$$c_{2x}(m) = \frac{\gamma_{2w}}{\gamma_{3w} H(0)} \sum_{n=-\infty}^{\infty} c_{3x}(m,n) = \alpha \sum_{n=-\infty}^{\infty} c_{3y}(m,n), \tag{8a}$$

where $\alpha \triangleq \gamma_{2w}/[\gamma_{3w}H(0)]$, and $c_{2x}(m,n)$, $c_{3x}(m,n)$, $c_{3y}(m,n)$, are defined by

$$c_{2x}(m) \triangleq E\{x(i)x(i+m)\}. \tag{8b}$$

$$c_{3x}(m,n) \triangleq E\{x(i)x(i+m)x(i+n)\}. \tag{8c}$$

$$c_{3y}(m,n) \triangleq E\{y(i)y(i+m)y(i+n)\}. \tag{8d}$$

In the sequel, for notational simplicity, we will assume that $H(0) \neq 0$.

The last equality in (8a) shows that under the aforementioned assumptions on $x(n)$ and $v(n)$, the noise-free autocorrelation $c_{2x}(m)$ can be obtained (within a scale factor) from the 3rd-order cumulant $c_{3y}(m,n)$ of the noisy time series $y(n)$. Furthermore, (8a) suggests the following two-step autocorrelation estimation *algorithm* in practice.

Step 1. Given N samples of the zero mean noisy time series $\{y(i)\}_{i=1}^{N}$ obtain consistent 3rd-order cumulant estimates; e.g., using the biased estimator

$$c_{3y}^{(N)}(m,n) = \frac{1}{N} \sum_{i=\max(1,1-m,1-n)}^{\min(N,N-m,N-n)} y(i)y(i+m)y(i+n). \tag{9a}$$

Step 2. Use (9a) into (8a) to estimate $c_{2x}(m)$, within a scale factor, as

$$c_{2x}^{(N)}(m) = \alpha \sum_{n=-N+1}^{N-1} c_{3y}^{(N)}(m,n). \tag{9b}$$

The potential of (8a) for Gaussian noise insensitive autocorrelation estimation becomes transparent if we take into account that

$$c_{2x}(m) = c_{2y}(m) - c_{2v}(m). \tag{10}$$

Clearly, it is impossible to estimate $c_{2x}(m)$ from $c_{2y}(m)$ by this formula, unless the covariance, $c_{2v}(m)$, of the AGN is available; thus, autocorrelation estimation based on 3rd-order cumulants is well motivated in the presence of AGN of unknown covariance. In fact, with sufficiently long data records, (9b) offers the potential for noise insensitive estimators even when $v(n)$ is non-Gaussian, but i.i.d. and non-skewed (e.g., symmetrically distributed), provided that $x(n)$ is skewed.

In (8a) it was tacitly assumed that $0 < |\alpha| < \infty$, which is guaranteed provided that $H(0) \neq 0$ and $\gamma_{3w} \neq 0$. If $\gamma_{3w} = 0$, $c_{2x}(m)$ must be expressed via higher-than 3rd-order cumulants; e.g., if $\gamma_{4w} \neq 0$ substituting $l = 2$, $k = 4$ in (6) and taking into account (7c), we obtain

$$c_{2x}(m) = \frac{\gamma_{2w}}{[H(0)]^2 \gamma_{4w}} \sum_{n,l=-\infty}^{\infty} c_{4y}(m,n,l), \tag{11a}$$

where

$$c_{4y}(m,n,l) \triangleq E\{y(i)y(i+m)y(i+n)y(i+l)\} - c_{2y}(m)c_{2y}(n-l) - c_{2y}(n)c_{2y}(l-m) - c_{2y}(l)c_{2y}(m-n). \tag{11b}$$

If $H(0) = 0$ in (11a), use of (6) with $\omega \neq 0$ is necessary.

It is not necessary to assume that $h(i)$ comes from an AR, MA, or ARMA model. As discussed in the sequel a windowed version of (9b) guarantees consistent non-parametric estimators of $c_{2x}(m)$ for any (parametric or not) non-Gaussian linear time series.

3. Asymptotic performance analysis

Taking expected values on both sides of (9b) it follows that

$$E\{c_{2x}^{(N)}(m)\} = \alpha \sum_{n=-N+1}^{N-1} \psi_N(m,n)\, c_{3x}(m,n) \underset{N\to\infty}{\to} c_{2x}(m), \tag{12a}$$

where $N\psi_N(m,n)$ denotes the number of terms included in the sum of (9a), i.e.,

$$N\psi_N(m,n) \triangleq \min(N, N-m, N-n) - \max(1, 1-m, 1-n) \equiv N - 1 - \max(|m|, |n|, |m-n|).$$

To establish the convergence in (12a), we assume that $x(n)$ obeys standard mixing conditions [2], and consider

$$E\{c_{2x}^{(N)}(m)\} - c_{2x}(m) = \alpha \sum_{|n|<N-1} [\psi_N(m,n) - 1]\, c_{3x}(m,n) - \sum_{|n|\geqslant N} c_{3x}(m,n).$$

Due to (5b) the second term tends to zero as $N \to \infty$. On the other hand, quantity

$$\psi_N(m,n) - 1 = -\frac{1 + \max(|m|, |n|, |m-n|)}{N},$$

is $O(|n|/N)$; hence, the first term is equal to $(-1/N) \times \sum_{|n|<N-1} |n|\, c_{3x}(m,n) + O(1/N)$ which tends to zero if, for example, the mixing condition $\sum_{|n|} |n c_{3x}(m,n)| < \infty$ holds.

Although (12a) shows that $c_{2x}^{(N)}(m)$ is asymptotically unbiased, in its 'raw form' $c_{2x}^{(N)}(m)$ is a poor (if not useless) estimator of $c_{2x}(m)$ because no matter how large N becomes $c_{2x}^{(N)}(m)$ always involves the tail of $c_{3y}^{(N)}(m,n)$ which is an unreliable estimate of $c_{3y}(m,n)$. Despite the fact that $c_{2x}^{(N)}(m)$ is formed by $2N+1$ 3rd-order sample cumulants each having variance $O(1/N)$ [22], the cumulative effect of the $2N+1$ terms exhibits variance which is $O(1)$. From this viewpoint the time-domain estimator in (9b) resembles the statistical behavior of the frequency-domain periodogram estimator which for processes satisfying (5b) is known to be inconsistent (e.g., [20, Chapter 6]).

To reduce the variance of $c_{2x}^{(N)}(m)$, the tail samples of $c_{3y}^{(N)}(m,n)$ must be omitted. The latter will not affect the bias of $c_{2x}^{(N)}(m)$, since for every N, $c_{3y}(m,n) \underset{n\to\infty}{\to} 0$. Prompted by the modified periodogram analysis, we subsequently study the asymptotic behavior of the following windowed version of (9b):

$$c_{2x}^{(N)}(m) = \alpha \sum_{n=-M}^{M} \lambda^*(m,n)\, c_{3y}^{(N)}(m,n). \tag{12b}$$

Although the class of windows which guarantee consistency of the estimator in (12b) is large, we will restrict ourselves to the 'scale parameter' windows, which are defined by

$$\lambda^*(m,n) \triangleq \lambda\left(\frac{m}{M}, \frac{n}{M}\right), \tag{12c}$$

where $\lambda(s,t)$ is a continuous function of (s,t) which is non-zero only over the support $[-1,1] \times [-1,1]$. Further, we require $\lambda^*(m,n)$ to have the following 3rd-order cumulant symmetries:

$$\text{(C1)}\quad \lambda^*(m,n) = \lambda^*(n,m) = \lambda^*(m-n,-n) = \lambda^*(-n,m-n) = \lambda^*(-m,n-m) = \lambda^*(n-m,-m).$$

Condition (C1) assures that (12b) yields autocorrelation estimates which are symmetric. Indeed, recalling that $c_{3y}^{(N)}(m,n)$ is defined on a hexagonal support, and using (12c) and (12b) for $m \geqslant 0$, one can verify that

$$c_{2x}^{(N)}(m) = \alpha \sum_{n=-M}^{M} \lambda^*(m,n)\, c_{3y}^{(N)}(m,n)$$
$$= \alpha \sum_{n=m-M}^{M} \lambda^*(m,n)\, c_{3y}^{(N)}(m,n)$$
$$= \alpha \sum_{n=m-M}^{M} \lambda^*(-m,n-m)\, c_{3y}^{(N)}(-m,n-m)$$
$$= \alpha \sum_{l=-M}^{M-m} \lambda^*(-m,l)\, c_{3y}^{(N)}(-m,l) = c_{2x}^{(N)}(-m).$$

On the other hand, the sample autocorrelation sequence estimated as in (12b) is not guaranteed to be positive (or non-negative) definite. For example, substituting (9a) into (12b) we find: $c_{2x}^{(N)}(m) = \sum_l s_y(l) y(l) y(l+m)$, where $s_y(l) \triangleq (\alpha/N)\sum_{n=-M}^{M} \lambda^*(n) y(l+n)$, and for simplicity we have assumed that $\lambda^*(m,n) = \lambda^*(n) = \lambda(n/M)$, for all m. It thus follows that for a fixed $L < N$, the $L \times L$ sample autocorrelation matrix can be written as a product of three matrices, $\boldsymbol{C}_{2x}^{(N)} = \boldsymbol{YSY}^{\mathrm{T}}$, where T denotes transpose, $\boldsymbol{Y}$ is the $L \times (N-L+1)$ data matrix with entries $y_{ij} = y(L+j-i)$, $i = 1,\dots,L$, $j = 1,\dots,N-L+1$, and $\boldsymbol{S}$ is a diagonal matrix with entries $\{s_y(l)\}_{l=L}^{N}$. We infer that for $\boldsymbol{C}_{2x}^{(N)}$ to be positive definite, we should have $s_y(l) > 0$ for all l, provided that $\boldsymbol{Y}$ is full rank. Future research could explore the effect of window selection in the positive definiteness of the (12b) estimates.

However, as it will be proved in the sequel (see Proposition 1 in this section), the precise shape of the window is not important for consistency, as long as (C1) and the following conditions are satisfied:

(C2) $\lambda(0,0) = 1 \geqslant \lambda(s,t) \geqslant 0$, $\forall\, s,t,M,N$,
(C3) $\lim_{s,t\to 0}(1-\lambda(s,t))/s^2 \triangleq L$, $0 \leqslant L < \infty$,
(C4) $\sum_{m,n} \lambda^2(m/M, n/M) < \infty\ \forall M$,
(C5) $M \to \infty$ as $N \to \infty$, and $M^2/N \underset{N\to\infty}{\to} 0$.

Using (C1)–(C5) the following consistency result holds for the proposed estimator.

Proposition 1. *Let $x(n)$ be a linear process given by Eq. (5a) where $h(n)$ is absolutely summable and $w(n)$ i.i.d. input with non-symmetric pdf and finite moments up to order six. Assume also that $x(n)$ obeys the standard mixing condition,*

$$\sum_{|t|} |t|\, |c_{2x}(t)| < \infty.$$

Let also, $y(n)$ be the noisy version of $x(n)$ as in Eq. (7b) where $v(n)$ obeys (7a) for $k = 3$. If (C1)–(C5) are satisfied, the estimator

$$c_{2x}^{(N)}(m) = \alpha \sum_{n=-M}^{M} \lambda^*(m,n)\, c_{3y}^{(N)}(m,n), \tag{13}$$

(i) *is asymptotically unbiased, i.e.,*

$$E\{c_{2x}^{(N)}(m)\} - \alpha \sum_{n=-\infty}^{\infty} c_{3x}(m,n) \underset{N\to\infty}{\to} 0, \tag{14}$$

(ii) *is mean-square consistent with asymptotic mean-square error variance*

$$\lim_{N\to\infty} \frac{N}{M^2} E\{[c_{2x}^{(N)}(m) - c_{2x}(m)]^2\} = \xi [c_{2y}(m)]^2, \tag{15a}$$

where

$$\xi \triangleq \alpha^2 \left[\sum_{t=-\infty}^{\infty} c_{2y}(t)\right]\left[\int_{-1}^{1} \lambda(0,u)\,\mathrm{d}u\right]^2. \tag{15b}$$

Proof. (i) Following the methodology used for statistical analysis of conventional correlation estimators [1, pp. 522–532], let us consider the difference

$$E\{c_{2x}^{(N)}(m)\} - \alpha \sum_{n=-\infty}^{\infty} c_{3x}(m,n)$$
$$= \alpha \sum_{n=-N+1}^{N-1} \lambda^*(m,n)\, E\{c_{3y}^{(N)}(m,n)\} - \alpha \sum_{n=-\infty}^{\infty} c_{3x}(m,n)$$
$$= \alpha \sum_{n=-N+1}^{N-1} \lambda^*(m,n)\psi_N(m,n) c_{3y}(m,n) - \alpha \sum_{n=-\infty}^{\infty} c_{3x}(m,n)$$
$$= \alpha \sum_{n=-N+1}^{N-1} \left[\bar{\lambda}\left(\frac{m}{M},\frac{n}{M}\right) - 1\right] c_{3x}(m,n)$$
$$- \alpha \sum_{|n|\geqslant N} c_{3x}(m,n), \tag{16}$$

where we have used that $c_{3y}(m,n) = c_{3x}(m,n)$, and the definition

$$\bar{\lambda}\left(\frac{m}{M},\frac{n}{M}\right) \triangleq \lambda\left(\frac{m}{M},\frac{n}{M}\right)\psi_N(m,n). \tag{17}$$

The first term on the rhs of (16) can be written as

$$\alpha \sum_{n=-M}^{M}\left[\bar{\lambda}\left(\frac{m}{M},\frac{n}{M}\right)-1\right]c_{3x}(m,n)$$
$$+\alpha \sum_{M<|n|<N}\left[\bar{\lambda}\left(\frac{m}{M},\frac{n}{M}\right)-1\right]c_{3x}(m,n). \quad (18)$$

The first sum of (18) can be further broken into the following two sums:

$$\sum_{n=-\bar{M}}^{\bar{M}}\left[\bar{\lambda}\left(\frac{m}{M},\frac{n}{M}\right)-1\right]c_{3x}(m,n)$$
$$+\sum_{\bar{M}\leqslant|n|<M}\left[\bar{\lambda}\left(\frac{m}{M},\frac{n}{M}\right)-1\right]c_{3x}(m,n), \quad (19)$$

for any integer $\bar{M}<M$. Because of (C4), $\lambda(m/M,n/M)$ and hence $\bar{\lambda}(m/M,n/M)$ are absolutely bounded; thus,

$$\left|\bar{\lambda}\left(\frac{m}{M},\frac{n}{M}\right)-1\right|\leqslant\bar{\mu}. \quad (20)$$

Further, from (C3) and (17) we infer that

$$\lim_{N\to\infty}\frac{1-\bar{\lambda}\left(\frac{m}{M},\frac{n}{M}\right)}{(n/M)^2}=L, \quad (21)$$

where we have used that $\psi_N(m,n)$ in (17) converges to 1 as $N\to\infty$. Because of (21), for any $\varepsilon>0$, we can choose N and consequently $M=M(N)$, so that, δ, so that

$$\left|\frac{\bar{\lambda}\left(\frac{m}{M},\frac{n}{M}\right)-1}{(n/M)^2}+L\right|<\varepsilon. \quad (22a)$$

Also, for any $\delta>0$ we can define $\bar{M}=\delta M$ such that

$$\left|\frac{n}{M}\right|<\delta \quad \forall\,|n|<\bar{M}. \quad (22b)$$

Now, the first term in (19) is within $\varepsilon' \triangleq \varepsilon\sum_{|n|<\bar{M}}|n/M|^2\,|c_{3x}(m,n)|$ of

$$|L|\sum_{|n|<\bar{M}}\left|\frac{n}{M}\right|^2|c_{3x}(m,n)|$$
$$\leqslant |L|\,\delta^2\sum_{|n|<\bar{M}}\leqslant |L|\,\delta^2\,C_{3x}$$

(where C_{3x} is defined in (5b)) and hence arbitrarily small.

The second term in (19) is in absolute value no greater than $\bar{\mu}\sum_{|n|<\bar{M}}|c_{3x}(m,n)|$, which converges to zero as $N\to\infty$ [cf. (20) and (5b)]. For exactly the same reasons, the second terms in (18) and (16) converge to zero as $N\to\infty$, and the asymptotic unbiasedness of the estimator in (13) has been established.

(ii) Following the steps used for conventional autocorrelation estimators [15, pp. 313–315], in our case we seek the limit of

$$\frac{N}{M^2}\operatorname{cov}\left\{\frac{c_{2x}^{(N)}(k)}{\alpha},\frac{c_{2x}^{(N)}(m)}{\alpha}\right\}$$
$$=\frac{1}{M^2}\sum_{l=N+1}^{N-1}\sum_{n=-N+1}^{N-1}\lambda^*(k,l)\lambda^*(m,n)$$
$$\times[N\operatorname{cov}\{c_{3x}^{(N)}(k,l),c_{3x}^{(N)}(m,n)\}]. \quad (23)$$

Recall that due to (C5) the ratio $N/M^2\to\infty$ meaning that existence of a finite limit of the expressions in (23) guarantees convergence of $\operatorname{cov}\{c_{2x}^{(N)}(k)/\alpha, c_{2x}^{(N)}(m)/\alpha\}$ to zero with rate equal to M^2/N.

Taking into account the symmetries present in 3rd-order cumulants, [2], we may assume without loss of generality that $m\geqslant n\geqslant 0$, $k\geqslant l\geqslant 0$, and re-write the bracketed term in (23) as

$$N\operatorname{cov}\left\{\frac{1}{N}\sum_{t_1=0}^{N-k}y(t_1)y(t_1+k)y(t_1+l),\right.$$
$$\left.\frac{1}{N}\sum_{t_2=0}^{N-m}y(t_2)y(t_2+m)y(t_2+n)\right\}$$
$$=\frac{1}{N}\sum_{t_1=0}^{N-k}\sum_{t_2=0}^{N-m}E\{y(0)y(k)y(l)y(t_2-t_1)$$
$$\times y(t_2-t_1+m)y(t_2-t_1+n)\}$$
$$-\frac{(N-k+1)(N-m+1)}{N}c_{3y}(k,l)c_{3y}(m,n). \quad (24)$$

Because the double sum in (24) is a function $f(t_2-t_1)$, it can be reduced to a single sum according to

$$\frac{1}{N}\sum_{t_1=0}^{N-k}\sum_{t_2=0}^{N-m} f(t_2-t_1)=\sum_{t=-(N-m-1)}^{N-k-1}\phi_N(k,m:t)f(t), \tag{25a}$$

where

$$\phi_N(k,m;t) \triangleq \begin{cases} \dfrac{N-k-t}{N}, & \max(m-k,0)+1 \leqslant t \leqslant N-k-1,\\ \dfrac{N-\max(m,k)}{N}, & \min(m-k,0)\leqslant t\leqslant \max(m-k,0),\\ \dfrac{N-m-|t|}{N}, & -(N-m-1)\leqslant t\leqslant \min(m-k,0)-1,\\ 0, & \text{elsewhere.}\end{cases} \tag{25b}$$

Because

$$\sum_{t=-(N-1)}^{N-1}\phi_N(k,m;t)=\frac{(N-k)(N-m)}{N}, \tag{25c}$$

substitution of (24) into (23) leads to

$$\frac{N}{M^2}\operatorname{cov}\left\{\frac{c_{2x}^{(N)}(k)}{\alpha},\frac{c_{2x}^{(N)}(m)}{\alpha}\right\} = \frac{1}{M^2}\sum_{l=N+1}^{N-1}\sum_{n=-N+1}^{N-1}\lambda^*(k,l)\lambda^*(m,n) \times\left\{\sum_{t=-(N-1)}^{N-1}\phi_N(k,m;t) \times[E\{y(0)y(k)y(l)y(t)y(t+m)y(t+n)\} - c_{3y}(k,l)c_{3y}(m,n)]\right\}. \tag{26}$$

The 6th-order correlation on the rhs of (26) can be expressed as the sum of 41 terms which include 2nd-, 3rd-, 4th- and 6th-order cumulants. More specifically we quote from [22] and [26] that if $m_k(v_1,\dots,v_k)$ and $c_k(v_1,\dots,v_k)$ are the kth-order moment and cumulant of a process, then

$$\begin{aligned} m_6(v_1,\dots,v_6) &= c_6(v_1,\dots,v_6)\\ &+\{c_3(v_1,v_2,v_3)c_3(v_4,v_5,v_6)\}_{10}\\ &+\{c_2(v_1,v_2)c_4(v_3,v_4,v_5,v_6)\}_{15}\\ &+\{c_2(v_1,v_2)c_2(v_3,v_4)c_2(v_5,v_6)\}_{15},\end{aligned}$$

where the notation $\{\cdot\}_j$ denotes the sum of all j *different* terms obtained by interchanging the arguments of the terms in brackets (the order of the arguments of c's being immaterial).

Exploiting the aforementioned splitting, it turns out that all except one term tend to zero when we take the limit of (26) as $N\to\infty$. Let us consider first the term corresponding to the 6th-order cumulant

$$\frac{1}{M^2}\sum_{l=-(N-1)}^{N-1}\sum_{n=-(N-1)}^{N-1}\lambda^*(k,l)\lambda^*(m,n) \times\sum_{t=-(N-1)}^{N-1}\phi_N(k,m;t)c_{6y}(k,l,t,t+m,t+n).$$

Using the substitution $u=t+n$, it is easy to show that this term is in absolute value no greater than

$$\frac{\sup_{r,s}[\lambda^2(r,s)]}{M^2}\sum_{l,t,u=-\infty}^{\infty}|c_{6x}(k,l,t,t+m,u)|,$$

which because of (5b) tends to zero as $N\to\infty$.

Next, we examine the 10 terms formed by products of 3rd-order cumulants. One of them is $c_{3y}(k,l)c_{3y}(m,n)$, and cancels the last term on the rhs of (26). A typical one from the remaining nine contributes the term

$$\frac{1}{M^2}\sum_{l=-(N-1)}^{N-1}\sum_{n=-(N-1)}^{N-1}\lambda^*(k,l)\lambda^*(m,n) \times\sum_{t=-(N-1)}^{N-1}\phi_N(k,m;t)c_{3y}(k,t)c_{3y}(t+m-l,t+n-l),$$

and is dominated by

$$\frac{\sup_{r,s}[\lambda^2(r,s)]}{M^2}\sum_{u,t,n=-\infty}^{\infty}|c_{3x}(k,t)|\,|c_{3x}(u,n)|,$$

which tends to zero as $N\to\infty$. Similarly, the other eight terms can be shown to vanish as $N\to\infty$.

The next category of 15 terms includes products between 2nd- and 4th-order cumulants. One of these contributes the term

$$\frac{1}{M^2}\sum_{l=-(N-1)}^{N-1}\sum_{n=-(N-1)}^{N-1}\lambda^*(k,l)\lambda^*(m,n)$$
$$\times\sum_{t=-(N-1)}^{N-1}\phi_N(k,m;t)\,c_{2y}(k)$$
$$\times c_{4y}(t-l,t+m-l,t+n-l),$$

which is dominated by

$$\frac{\sup_{r,s}[\lambda^2(r,s)](2M+1)}{M^2}c_{2y}(k)\sum_{u,n=-\infty}^{\infty}|c_{4x}(u,m,n)|,$$

and tends to zero as $N\to\infty$, [here we have used stationarity to replace $c_{4y}(t-l,t+m-l,t-l+n)$ by $c_{4y}(l-t,m,n)$, and the Gaussianity of the noise to substitute c_{4y} by c_{4x}]. Likewise, the remaining 14 terms go to zero as $N\to\infty$.

The last category has 15 terms formed by the triple products of 2nd-order cumulants, and 14 of those can be shown to vanish as $N\to\infty$ using, as before, dominated convergence arguments. The only term which survives has the form

$$\frac{1}{M^2}\sum_{l=-(N-1)}^{N-1}\sum_{n=-(N-1)}^{N-1}\lambda^*(k,l)\lambda^*(m,n)$$
$$\times\sum_{t=-(N-1)}^{N-1}\phi_N(k,m;t)c_{2y}(k)c_{2y}(t+n-l)$$
$$\times c_{2y}(k)c_{2y}(m)\left\{\frac{1}{M^2}\sum_{l=-(M-1)}^{M-1}\sum_{n=-(M-1)}^{M-1}\lambda\left(\frac{k}{M},\frac{l}{M}\right)\right.$$
$$\left.\times\lambda\left(\frac{m}{M},\frac{n}{M}\right)\sum_{t=-\infty}^{\infty}c_{2y}(t)\phi_N(k,m;t-n+l)\right\},\tag{27}$$

where we used the fact that $\lambda(k/M,l/M)$ and $\lambda(m/M,n/M)$ are zero for $|k|,|l|,|m|,|n|>M$. For large N and since $M^2/N\to 0$ the kernel $\phi_N(k,m;t-n+l)$ tends to $1-|t|/N$. Hence as $N\to\infty$ the leftmost sum in (27) converges to

$$\sum_{t=-\infty}^{\infty}c_{2y}(t).$$

Also,

$$\frac{1}{M^2}\sum_{l=-(M-1)}^{M-1}\sum_{n=-(M-1)}^{M-1}\lambda\left(\frac{k}{M},\frac{l}{M}\right)\lambda\left(\frac{m}{M},\frac{n}{M}\right)$$
$$\to\zeta\triangleq\left[\int_{-1}^{1}\lambda(0,u)\,\mathrm{d}u\right]^2.$$

Consequently, from (27), it follows that

$$\lim_{N\to\infty}\frac{N}{M^2}\operatorname{cov}\left\{\frac{c_{2x}^{(N)}(k)}{\alpha},\frac{c_{2x}^{(N)}(m)}{\alpha}\right\}$$
$$=\zeta\left[\sum_{t=-\infty}^{\infty}c_{2y}(t)\right]c_{2y}(k)c_{2y}(m),\tag{28}$$

which establishes consistency.

To find the variance of $c_{2x}^{(N)}(m)$ consider (28) with $k=m$ to obtain

$$\lim_{N\to\infty}\frac{N}{M^2}\operatorname{var}\{c_{2x}^{(N)}(m)\}=\xi[c_{2y}(m)]^2,$$

where $\xi\triangleq\alpha^2\zeta[\sum_{t=-\infty}^{\infty}c_{2y}(t)]$. □

The asymptotic variance of (15a) is very useful for analyzing the performance of parametric and non-parametric methods which use $c_{2x}^{(N)}(m)$. For N sufficiently large, (15a) is estimated by

$$\operatorname{var}\{c_{2x}^{(N)}(m)-c_{2x}(m)]^2\}\sim\frac{M^2}{N}\xi[c_{2y}^{(N)}(m)]^2.\tag{29a}$$

It is interesting to compare (29a) with the Bartlett's result which states that the conventional estimator

$$r_{2y}^{(N)}(m)=\frac{1}{N}\sum_{i=1}^{N-m}y(i)y(i+m),\quad m>0,\tag{29b}$$

has variance (e.g., [20, p. 326])

$$\operatorname{var}\{[r_{2y}^{(N)}(m)-c_{2y}(m)]^2\}$$
$$\sim\frac{1}{N}\sum_{t=-\infty}^{\infty}\{[r_{2y}^{(N)}(t)]^2+r_{2y}^{(N)}(t-m)r_{2y}^{(N)}(t+m)$$
$$+c_{4y}^{(N)}(t,m,0)\}.\tag{29c}$$

Although $c_{2x}^{(N)}(m)$ is insensitive to AGN *in the mean*, its variance is $O(M^2/N)$, as is the variance of the biperiodogram estimates [2]. On the contrary, $r_{2y}^{(N)}(m)$ yields a biased estimate of $c_{2x}(m)$ in the presence of AGN, but its variance is always $O(1/N)$. Thus, due to the M^2 factor, it is expected that $c_{2x}^{(N)}(m)$ will outperform $r_{2y}^{(N)}(m)$ when N is large and the SNR is low. However, the trade-off between bias and variance may be in favor of $r_{2y}^{(N)}(m)$

when short data records are encountered in high SNR applications.

In addition to the mean-square convergence established by Proposition 1, the proposed estimator is also strongly consistent under alternate conditions. The following proposition specifies these conditions and establishes w.p.1 convergence of $c_{2x}^{(N)}(m)$ to $\alpha c_{2x}(m)$.

Proposition 2. *If* (C1)–(C4) *are satisfied,* (C5) *is replaced by* (C5′) $M = O(N^{1/(4+\delta)})$, *and if in addition* (C6) $v(n)$ *is a linear, zero-mean process,* $v(n) = \sum_k g(k)\zeta(n-k)$, *where* $\sum_k |c(k)| < \infty$ *and* $\zeta(n)$ *i.i.d Gaussian or symmetrically distributed with finite (up to order 6) moments, then the estimator* $c_{2x}^{(N)}(m)$ *is a strongly consistent estimator of* $c_{2x}(m)$; *i.e.,*

$$|c_{2x}^{(N)}(m) - c_{2x}(m)| \to 0 \quad \text{w.p.1 as } N \to \infty. \tag{30}$$

Proof. Before proceeding to prove Proposition 2 we quote here two lemmas (see [9] for proofs).

Lemma 1. *Let* $\phi = \sum_{k=-\infty}^{\infty} a(k)z(k)$, *where* (i) $a(k)$ *is a deterministic sequence satisfying* $\sum_{k=-\infty}^{\infty} |a(k)| \leqslant C_a$, *and* (ii) $z(k)$ *is a sequence of random variables such that* $E\{z^2(k)\} \leqslant C_z$. *Then,* $E\{\phi^2\} \leqslant C_z[\sum_{k=-\infty}^{\infty} |a(k)|]^2 \leqslant C_z C_a^2$.

Lemma 2. *Let* $\{X_t\}$ *be a sequence of zero-mean random variables and define* $Y_r^N \triangleq \sum_{t=r+1}^{N} X_t$, $0 \leqslant r \leqslant N$. *If* $E\{(Y_r^N)^2\} \leqslant CM^2(N)(N-r)$, *where* $M(N) \leqslant O(N^{1/(4+\delta)})$, $\delta > 0$, *then the sequence* $N^{-1}|Y_0^N| \to 0$ w.p.1 *as* $N \to \infty$.

We return now to the proof of (30). According to Proposition 1, $c_{2x}^{(N)}(m)$ is an asymptotically unbiased estimator of $c_{2x}(m)$, i.e., $E\{c_{2x}^{(N)}(m)\} \to c_{2x}(m)$ as $N \to \infty$. Hence, in order to establish (30) we need to show that $|c_{2x}^{(N)}(m) - E\{c_{2x}^{(N)}(m)| \to 0$ w.p.1 as $N \to \infty$. We shall prove this assertion in two stages. In the first stage we shall assume that the measurements are noise-free; in the next stage we shall show that the contribution of the AGN decays to 0 w.p.1 as $N \to \infty$. In the following we shall fix m and hence for notational simplicity and w.l.o.g we shall write $\lambda^*(n)$ instead of $\lambda^*(m,n)$.

Stage 1. Let

$$\begin{aligned} D_x(r,N) = \sum_{t=r+1}^{N} \Bigg[&\sum_{n=-N}^{N} \lambda^*(n)x(t)x(t+m)x(t+n) \\ &- E\Bigg\{ \sum_{n=-N}^{N} \lambda^*(n)x(t)x(t+m)x(t+n) \Bigg\} \Bigg] \\ = \sum_{n=-N}^{N} &\lambda^*(n) \sum_{i,j,k} h(i)h(j)h(k) \\ \times \sum_{t=r+1}^{N} &[e(t-i)e(t+m-j)e(t+n-k) \\ &- E\{e(t-i)e(t+m-j)e(t+n-k)\}]. \end{aligned} \tag{31a}$$

If

$$\begin{aligned} S(r,N) \triangleq \sum_{t=r+1}^{N} &[e(t+\eta_0)e(t+\eta_1)e(t+\eta_2) \\ &- E\{e(t+\eta_0)e(t+\eta_1)e(t+\eta_2)\}], \end{aligned} \tag{31b}$$

then (31a) yields

$$|D_x(r,N)| \leqslant \sum_{n=-N}^{N} |\lambda^*(n)| \sum_i |h(i)| \sum_j |h(j)| \times \sum_k |h(k)|\, |S(r,N)|, \tag{31c}$$

with $\eta_0 = -i$, $\eta_1 = m - j$ and $\eta_2 = n - k$. Because $e(t)$ is white it can be easily checked that

$$E\{|S(r,N)|^2\} \leqslant (N-r)m_{\max}, \tag{31d}$$

where $m_{\max} \triangleq \{m_2^3, m_2, m_4, m_6 - m_3^2\}$, m_i being the ith moment of the driving noise $e(t)$. Since we also have $\sum_{n=-N}^{N} |\lambda^*(n)| \leqslant C_l M$ and $\sum_i |h(i)| \leqslant C_h$, using Lemma 1 we obtain

$$E\{|D_x(r,N)|^2\} \leqslant C_l^2 C_h^6 M^2(N-r), \tag{31e}$$

which in view of Lemma 2 and condition (C5′) implies

$$\frac{1}{N}|D_x(0,N)| \to 0 \quad \text{w.p.1 as } N \to \infty, \tag{31f}$$

or equivalently,

$$|c_{2x}^{(N)}(m) - c_{2x}(m)| \to 0 \quad \text{w.p.1 as } N \to \infty. \tag{32}$$

This completes the proof for the noiseless case.

Stage 2. We first observe that

$$E\left\{\sum_{n=-N}^{N}|\lambda^*(n)x(t)x(t+m)x(t+n)\right\}$$

$$=E\left\{\sum_{n=-N}^{N}|\lambda^*(n)y(t)y(t+m)y(t+n)\right\}.$$

Then using obvious notation,

$$D_y(r,N)$$
$$=D_x(r,N)+\sum_{t=r+1}^{N}\sum_{n=-N}^{N}\lambda^*(n)\{x(t)x(t+m)x(t+n)$$
$$+x(t)v(t+m)x(t+n)x(t)v(t+m)$$
$$\times v(t+n)v(t)x(t+m)x(t+n)$$
$$+v(t)x(t+m)v(t+n)v(t)v(t+m)$$
$$\times x(t+n)v(t)v(t+m)v(t+n)\}. \quad (33)$$

In the sequel we shall show that the contribution of the due to the noises terms decays to zero as $N\to\infty$. We shall focus on the first term. Let

$$A_m(r,N)\triangleq\sum_{t=r+1}^{N}\sum_{n=-N}^{N}\lambda^*(n)x(t)x(t+m)v(t+n)$$
$$=\sum_{n=-N}^{N}\lambda^*(n)\sum_{l_1}h(l_1)\sum_{l_2}h(l_2)\sum_{l_3}g(l_3)Z_m(r,N), \quad (34a)$$

where we have defined

$$Z_m(r,N)\triangleq\sum_{t=r+1}^{N}e(t+\eta_0)e(t+\eta_1)\zeta(t+\eta_2), \quad (34b)$$

with $\eta_0=-l_1$, $\eta_1=m-l_2$ and $\eta_2=n-l_3$. Because $e(t)$, $\zeta(t)$ are i.i.d and mutually uncorrelated

$$E\{[Z_m(r,N)]^2\}\leqslant C_z(N-r), \quad (34c)$$

where $C_z\triangleq\gamma_{2\zeta}\max\{E\{e(t+\eta_0)^2e(t+\eta_1)^2\}\}<\infty$. Hence, in view of Lemma 1,

$$E\{[A_m(r,N)\}^2\}\leqslant C_h^4C_c^2C_\lambda^2C_2(N-r), \quad (34d)$$

which implies, due to Lemma 2, that

$$\frac{1}{N}|A_m(0,N)|^2\to 0 \quad \text{w.p.1 as } N\to\infty. \quad (34e)$$

Similar arguments hold for the contribution of the next five terms. The last term on the RHS of Eq. (33) is

$$A_m(r,N)\triangleq\sum_{t=r+1}^{N}\sum_{n=-N}^{N}\lambda^*(n)v(t)v(t+m)v(t+n)$$
$$=\sum_{n=-N}^{N}\lambda^*(n)\sum_{l_1}g(l_1)\sum_{l_2}g(l_2)\sum_{l_3}g(l_3)Z_m(r,N), \quad (35a)$$

where we have now defined

$$Z_m(r,N)\triangleq\sum_{t=r+1}^{N}\zeta(t+\eta_0)\zeta(t+\eta_1)\zeta(t+\eta_2), \quad (35b)$$

with $\eta_0=-l_1$, $\eta_1=m-l_2$ and $\eta_2=n-l_3$. Then

$$E\{[Z_m(r,N)]^2\}$$
$$=\sum_{t=r+1}^{N}\sum_{s=r+1}^{N}E\{\zeta(t+\eta_0)\zeta(s+\eta_0)\zeta(t+\eta_1)$$
$$\times\zeta(s+\eta_1)\zeta(t+\eta_2)\zeta(s+\eta_2)\}. \quad (35c)$$

Since $\zeta(t)$ is zero-mean, white, Gaussian it can be easily checked that $E\{\zeta(t+\eta_0)\zeta(s+\eta_0)\zeta(t+\eta_1)\times\zeta(s+\eta_1)\zeta(t+\eta_2)\zeta(s+\eta_2)\}=0$ for $t\neq s$. Hence,

$$E\{[Z_m(r,N)\}^2\}$$
$$=\sum_{t=r+1}^{N}E\{\zeta(t+\eta_0)^2\zeta(t+\eta_1)^2\zeta(t+\eta_2)^2\}$$
$$\leqslant C_\zeta(N-r), \quad (35d)$$

where $C_\zeta\triangleq\max_{\eta_0,\eta_1,\eta_2}E\{\zeta(t+\eta_0)^2\zeta(t+\eta_1)^2\times\zeta(t+\eta_2)^2\}<\infty$ according to assumption (C6). It is now straightforward to show that

$$\frac{1}{N}|A_m(0,N)|^2\to 0 \quad \text{w.p.1 as } N\to\infty. \quad (35e)$$

This completes the proof of the Stage 2 establishing Proposition 2.

4. Recursive estimation

The estimator in (13) is a two-step batch procedure. Sample 3rd-order cumulants are computed in the first step, and windowed projections are formed in the second step. Because on-line processing is attractive in many problems, a recursive algorithm for computing $c_{2x}^{(N)}(m)$ directly from the data is developed in this subsection. Recursions for sample 3rd-order cumulants have been reported in [24] and [11] for cumulant-based adaptive modeling. Here, we develop recursions for sample autocorrelations obtained through windowed cumulant projections. For notational simplicity the 'rectangular window' is adopted. Further, the window support is fixed at $M = q$, i.e.,

$$\lambda^*(m,n) = 1, \quad |m|, |n|, |n+m| \leqslant q, \qquad = 0, \quad \text{elsewhere}. \tag{36a}$$

If $\bar{c}_{3y}^{(N)}(m,n) \triangleq N c_{3y}^{(N)}(m,n)$ denotes the sum in (9a), it follows that

$$\bar{c}_{3y}^{(N)}(m,n) = \bar{c}_{3y}^{(N-1)}(m,n) + y(N-n)y(N-n+m)y(N), \quad n > m, \tag{36b}$$

$$= \bar{c}_{3y}^{(N-1)}(m,n) + y(N-m)y(N)y(N-m+n), \quad n \leqslant m. \tag{36c}$$

Based on (36b) and (36c), the following recursion can be established between $\bar{c}_{2x}^{(N)}(m)$ and $\bar{c}_{2x}^{(N-1)}(m)$ for $m \geqslant 0$:

$$\begin{aligned}\bar{c}_{2x}^{(N)}(m) &= \alpha \sum_{n=-q+m}^{q} \bar{c}_{3y}^{(N)}(m,n) \\ &= \bar{c}_{2x}^{(N-1)}(m) + \alpha \sum_{n=-q+m}^{m} y(N-m)y(N)y(N-m+n) + \alpha \sum_{n=m+1}^{q} y(N-n)y(N-n+m)y(N) \\ &= \bar{c}_{2x}^{(N-1)}(m) + \alpha y(N)\Big[y(N-m) \times \sum_{k=0}^{q} y(N-k) + \alpha^{(N)}(m)\Big],\end{aligned} \tag{36d}$$

where

$$\begin{aligned}\alpha^{(N)}(m) &\triangleq \sum_{n=m+1}^{q} y(N-n)y(N-n+m) \\ &= \sum_{l=m}^{q-1} y(N-1-l)y(N-1-l+m) \\ &= \alpha^{(N-1)}(m) - y(N-1-q)y(N-1-q+m) + y(N-1-m)y(N-1).\end{aligned} \tag{36e}$$

Recursion (36d) and (36e) are useful for real-time computation of the autocorrelation lags $\{c_{2x}^{(N)}(m)\}_{m=0}^{q}$, when the underlying non-Gaussian time series is (or can be approximated by) an MA process of fixed order q. Although (36d) and (36e) do not explicitly compute 3rd-order cumulant estimates, for a given N, the final autocorrelation values obtained from the batch and the recursive solutions coincide. Thus, consistency of the recursive solution is implied by that of the batch estimator.

To take full advantage of parallel computing, all the $q+1$ autocorrelation lags can be computed simultaneously after vectorizing the recursion (36d) and (36e). Let $\boldsymbol{c}_N$ stand for the $1 \times (q+1)$ normalized vector

$$\boldsymbol{c}_N \triangleq [\bar{c}_{2x}^{(N)}(0), \bar{c}_{2x}^{(N)}(1), \ldots, \bar{c}_{2x}^{(N)}(q)]/\alpha, \tag{37a}$$

and, $\boldsymbol{y}_{n_1}^{n_2}$ denote the $1 \times (n_2 - n_1 + 1)$ data vector,

$$\boldsymbol{y}_{n_1}^{n_2} \triangleq [y(n_1), y(n_1+1), \ldots, y(n_2)]. \tag{37b}$$

Then, the vector counterparts of (36d) and (36e) are given by

$$\boldsymbol{c}_N = \boldsymbol{c}_{N-1} + y(N)[s(N)\boldsymbol{y}_N^{N-q} + [\alpha_N|0]\} \tag{37c}$$

and

$$\alpha_N = \alpha_{N-1} + y(N-1)\boldsymbol{y}_{N-1}^{N-q} - y(N-1-q)\boldsymbol{y}_{N-1-q}^{N-2}, \tag{37d}$$

where

$$s(N) = s(N-1) + y(N) - y(N-1-q) \tag{37e}$$

and

$$s(N) \triangleq \sum_{k=0}^{q} y(N-k). \tag{37f}$$

To initialize (37c)–(37e) it suffices to use

$$\boldsymbol{c}_{q+1} \triangleq [0 \ldots 0]_{1\times(q+1)}, \quad \alpha_{q+1} \triangleq [0 \ldots 0]_{1\times q} \quad \text{and } s(q+1) \triangleq 0. \tag{37g}$$

5. Simulations

In this section the autocorrelation and spectral estimators developed in Sections 2–4 are tested with simulated data and compared against the

Table 1
(a) Autocorrelation estimation and (b) parameter estimation of a non-minimum phase MA(3) process. $N = 4096$ (16×256), 100 Monte Carlo runs, colored AGN/UC, SNR = 0 db

(a) True and estimated (mean ± st. dev.) autocorrelation values

Lag	0	1	2	3
True	2.553	0.950	− 0.309	− 0.771
Conventional estimates	5.098 ± 0.117	0.333 ± 0.520	− 0.809 ± 0.077	− 0.223 ± 0.097
Proposed estimates	2.477 ± 0.320	0.892 ± 0.244	− 0.337 ± 0.195	− 0.774 ± 0.166

(b) True and estimated (mean ± st. dev.) MA coefficients

True	1.000	0.900	0.385	− 0.771
Conventional estimates	1.000 ± 0.000	1.834 ± 0.751	− 0.841 ± 0.481	− 1.431 ± 0.439
Proposed estimates	1.000 ± 0.000	0.892 ± 0.239	0.210 ± 0.301	− 0.746 ± 0.164

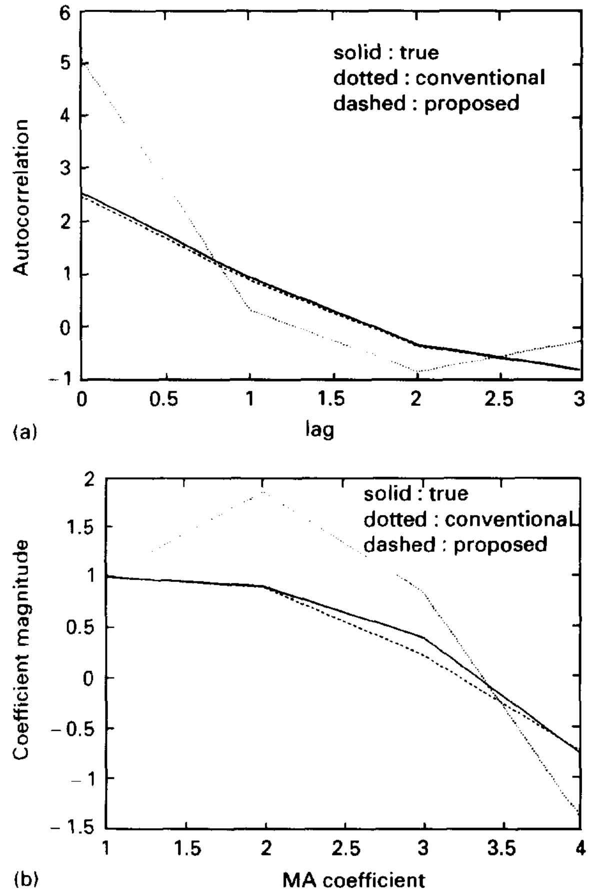

Fig. 1(a). Test case #1: MA(3), $N = 4096$, rectangular window with $M = 9$, colored AGN – SNR = 0 db. True and estimated (Monte Carlo mean, 100 runs): (a) autocorrelation estimates; (b) MA parameter estimates.

conventional estimators. Their usefulness for AR and MA modeling is also demonstrated. To verify the theoretically proved insensitivity to AGN, 100 Monte Carlo runs were performed in order to attain valid approximations of the estimators' statistics (mean ± standard deviation).

Except for test case # 2, the window was chosen, for simplicity, to be rectangular [cf. Eq. (36a)]. Therefore, the variance estimates reported correspond to worst-case analysis. Of course, the bias-variance trade-off, common to all spectral estimation approaches, is present here as well, and experience is the only guideline for choosing the 'best' window.

All three tests show that the proposed algorithms yield approximately unbiased estimators of the noise-free signal autocorrelation and spectrum, although the variance of the proposed estimators is in most cases higher than of the conventional estimators. In the presence of low SNR (0 db) colored AGN, the conventional estimators exhibit high bias but relatively low variance.

The driving noise $w(i)$ in the three experiments is zero-mean, independent and exponentially distributed, with $\gamma_{2w} = 1$ and $\gamma_{3w} = 2$. The exponential distribution for the input was adopted as an example of non-symmetric pdf corresponding to processes with non-zero γ_{3w} (cf. conditions of Propositions 1 and 2 regarding the input pdf). The AGN $v(i)$ is colored and independent of $x(i)$. When necessary, the estimates obtained are appropriately scaled for comparison purposes [note that the scaling constants γ_{2w}, γ_{3w}, and $H(0)$, are usually unknown, and hence normalized values are used].

Table 2
(a) Autocorrelation estimation and (b) identification of an AR(2) process. $N = 2048$ (8×256), 100 Monte Carlo runs, colored AGN/UC, SNR = 0 db

(a) True and estimated (mean $\pm$ st. dev.) autocorrelation values

Lag	0	1	2	3	4	5	6	7	8	9
True	1.727	− 0.538	− 0.763	0.699	0.107	− 0.462	0.161	0.183	− 0.168	− 0.021
Conventional estimates	3.420	− 1.262	− 0.916	0.452	0.768	− 0.715	0.139	0.190	− 0.163	− 0.035
	± 0.132	± 0.068	± 0.107	± 0.109	± 0.083	± 0.093	± 0.080	± 0.088	± 0.109	± 0.095
Proposed estimates	1.505	− 0.498	− 0.605	0.564	0.080	− 0.338	0.110	0.123	− 0.097	− 0.022
	± 0.351	± 0.185	± 0.287	± 0.243	± 0.224	± 0.209	± 0.204	± 0.195	± 0.195	± 0.229

(b) True and estimated (mean $\pm$ st. dev.) AR coefficients

True	1.000	0.500	0.600
Conventional estimates	1.000	0.541	0.467
	± 0.000	± 0.018	± 0.020
Proposed estimates	1.000	0.512	0.589
	± 0.000	± 0.145	± 0.129

Test case #1 (MA process). The mixed phase MA(3) time series

$$x(i) = w(i) + 0.9w(i-1) + 0.385w(i-2) - 0.771w(i-3), \quad \text{zeros: } 0.6,\ -0.75 \pm \mathrm{j}0.85,$$

was corrupted by AGN at SNR = 0 db. The AGN was generated after passing white Gaussian noise through the MA(5) model with coefficients $[1, -2.33, 0.75, 0.5, -1.3, -1.4]$. The true $\{c_{2x}(m)\}_{m=0}^{3}$ lags are shown in Table 1(a); 4096 samples of $y(i)$ were divided into 16 records each containing 256 samples. Both the conventional and the proposed autocorrelation estimator of Eq. (13) were computed for each record, and averaged over the 16 records. A rectangular window with $M = 9$ was used in the implementation of $c_{2x}^{(N)}(m)$. From the means and standard deviations shown in Table 1(a), we observe that even at 0 db, the proposed estimator is approximately unbiased, while the conventional is clearly biased but with lower standard deviation. The superiority of the proposed estimator is apparent in Fig. 1(a), where the Monte Carlo means and the true values of the first four autocorrelation lags are plotted.

In each Monte Carlo run the $c_{2x}(m)$ and $c_{3x}(m,n)$ values were also used to estimate the MA parameters using the algorithm reported in [13]. The mean and standard deviation of the parameter estimates are displayed in Table 1(b) and Fig. 1(b). Notice that the parameter estimates obtained with the conventional autocorrelational approach are biased, and at least in this particular example, they also exhibit higher variance than that of the proposed estimates. Observe also that in this parameter estimation application knowledge of the scale α is not required.

Test case #2 (AR process). The following AR(2) process was simulated:

$$x(i) + 0.5x(i-1) + 0.6x(i-2) = w(i),$$

and $x(i)$ was corrupted by 0 db AGN, $v(i)$, generated as in test case #1; 2048 samples were divided into 8 records. Table 2(a) displays the conventional and proposed autocorrelation estimates, obtained

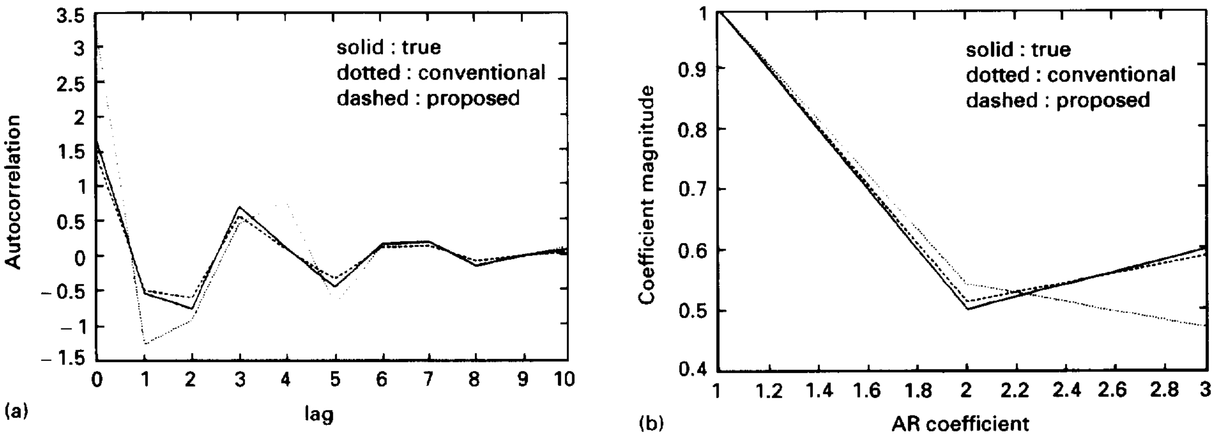

Fig. 2(a). Test case #2: AR(2), $N = 2048$, raised cosine window with $M = 23$, colored AGN – SNR = 0 db. True and estimated (Monte Carlo mean, 100 runs): (a) autocorrelation estimates; (b) AR(2) parameter estimates.

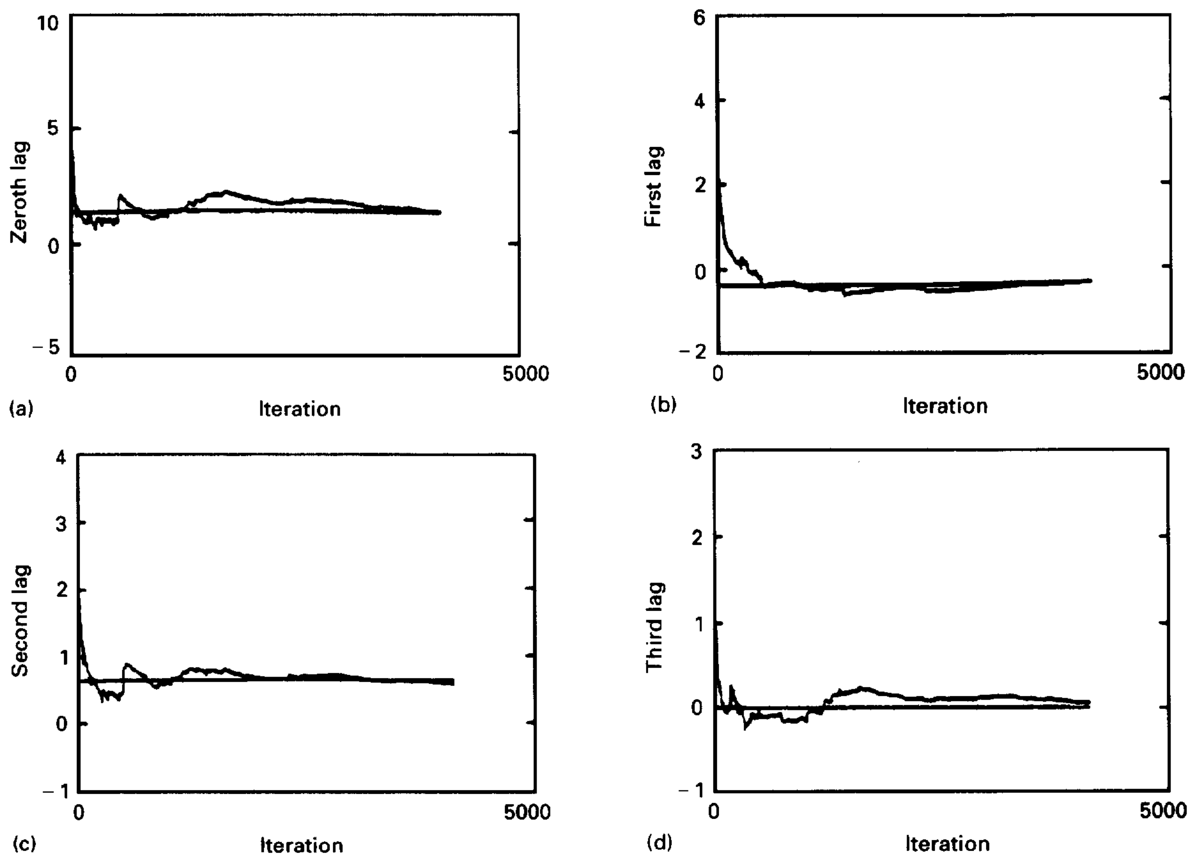

Fig. 3(a)–(d). Test case #3: recursive estimation of autocorrelation. MA(2), colored AGN – SNR = 0 db. True and estimated (Monte Carlo mean, 100 runs) autocorrelation coefficients.

by weighting 3rd-order cumulant projections with the following raised cosine window:

$$\lambda^*(m,n) = \begin{cases} \frac{1}{2}\left[1 + \cos\left(\frac{n\pi}{23}\right)\right], & (m,n)[-23,23], \\ 0, & \text{else.} \end{cases}$$

In theory (refer to Propositions 1 and 2) the precise shape of the window is immaterial. In practice though, where the user chooses windows of finite size rather than windows with size increasing as $N \to \infty$, a rule of thumb is to choose windows of length equal or greater to the essentially non-zero support of $c_{3x}(m,n)$.

The conventional estimators exhibit a bias equal to the noise autocorrelation. Thus, as depicted in Fig. 2(a), the (mean) conventional estimates fail to approximate the first six lags of $c_{2x}(m)$, although they do well with the rest autocorrelation lags. On the other hand, the standard deviation of the proposed estimators is 2–3 times higher, but the bias of all lags is smaller than that of the conventional estimator.

The first three lags of the $c_{2x}^{(N)}(m)$ estimates were subsequently used to identify the AR parameters using the Levinson–Durbin algorithm. Again, the AR parameter estimates shown in Table 2(b) and Fig. 2(b), have smaller bias, but higher variance than the conventional estimator.

Test case #3 (Recursive estimation). The MA(2) time series

$$x(i) = w(i) - 0.233w(i-1) + 0.666w(i-2),$$

was corrupted by the same AGN used in the Test cases #1 and #2. The noisy signal was used as input to the vector-recursive approach described by Eqs. 37(a)–(g). The trajectories of the $c_{2x}^{(N)}(m)$ values are depicted in Figs. 3(a)–(d). Clearly, the recursive algorithm converges to the true values $\{c_{2x}(m)\}_{m=0}^{3}$, shown with straight lines in Figs. 3(a)–(d).

6. Conclusions

Windowed projections of 3rd-order sample cumulants were shown to yield consistent autocorrelation estimators. The window requirements were presented, and the asymptotic variance of the resulting estimator was computed. The recursive algorithm developed for estimating autocorrelations from 3rd-order cumulants might be useful for recasting existing autocorrelation based adaptive algorithms through higher-order statistics. Theory and simulations verified that when dealing with sufficiently long records of stationary, non-Gaussian, linear processes observed in Gaussian noise, it may be useful to perform parametric or non-parametric second-order analysis, after estimating autocorrelation through cumulants.

References

[1] T.W. Anderson, The *Statistical Analysis of Time Series*, Wiley, New York, 1972.

[2] D.R. Brillinger and M. Rosenblatt, "Computation and interpretation of kth-order spectra", in: B. Harris, ed., *Spectral Analysis of Time Series*, Wiley, New York, 1967, pp. 189–232.

[3] J.-F. Cardoso, "Eigenstructure of the 4th-order cumulant tensor with application to the blind source separation problem", *Proc. Internat. Conf. Acoust. Speech Signal Process.*, Albuquerque, NM, April 1990, pp. 2655–2658; see also *Proc. Internat. Conf. Proc. Acoust. Speech Signal Process.*, Glasgow, Scotland, April 1989, pp. 2109–2212.

[4] C.-Y. Chi and J.-Y. Kung, "A new identification algorithm for allpass systems by higher-order statistics", *Signal Processing*, Vol. 41, No. 2, January 1995, pp. 239–256.

[5] P. Comon, "MA estimation using fourth order cumulants", *Signal Processing*, Vol. 26, No. 3, March 1992, pp. 381–388.

[6] P. Comon, "Independent component analysis: A new concept?", *Signal Processing*, Vol. 36, No. 3, April 1994, pp. 287–314.

[7] P. Comon, C. Jutten and J. Herault, "Blind separation of sources, Part II: Problems statement", *Signal Processing*, Vol. 24, No. 1, July 1991, pp. 11–20.

[8] A.G. Constantinides and P.T. Stathaki, "Robust autoregressive modeling through higher-order spectral estimation techniques with application to mammography", *Proc. 27th Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, 1–3 November 1993.

[9] A. Delopoulos and G.B. Giannakis, "Strongly consistent identification algorithms and noise insensitive MSE criteria", *IEEE Trans. Signal Process.*, Vol. 40, No. 8, 1992, pp. 1955–1970.

[10] A. Delopoulos and G.B. Giannakis, "Consistent identification of stochastic linear systems with noisy input–output data", *Automatica*, Vol. 30, No. 8, 1994, pp. 1271–1294.

[11] B. Friedlander and B. Porat, "Adaptive IIR filtering based on high-order statistics", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. 37, 1989, pp. 485–495.

[12] B. Friedlander and B. Porat, "Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from higher-order moments", *IEEE Trans. Automat. Control*, Vol. AC-35, 1990, pp. 27–35.

[13] G.B. Giannakis and A. Delopoulos, "Non-parametric estimation of autocorrelation and spectra using cumulants and polyspectra", *Proc. Soc. of Photo-Opt. Instr. Engr., Advanced Signal Processing Alg., Arch. and Implem.*, Vol. 1348, San Diego, CA, July 1990, pp. 503–517.

[14] G.B. Giannakis and J.M. Mendel, "Identification of nonminimum phase systems using higher-order statistics, *IEEE Trans. Acoust. Speech Signal Process.*, Vol. 37, 1989, pp. 360–377.

[15] E.J. Hannan, *Multiple Time Series*, Wiley, New York, 1970.

[16] D. Kletter and H. Messer, "Suboptimal detection of non-Gaussian signals by third-order analysis", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. 38, June 1990, pp. 901–909.

[17] K.-S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes", *Ann. Statist.*, 1982, pp. 1195–1208.

[18] J.M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications", *Proc. IEEE*, Vol. 79, No. 3, 1991, pp. 278–305.

[19] C.L. Nikias and R. Pan, "Time delay estimation in unknown Gaussian spatially correlated noise", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. 36, No. 11, 1988, pp. 291–325.

[20] M.B. Priestley, *Spectral Analysis and Time Series*, Academic Press, London, 1981.

[21] M. Rangoussi and G.B. Giannakis, "FIR modeling using log-bispectra: Weighted least-squares algorithms and performance analysis", *IEEE Trans. Circuits Systems*, Vol. 38, March 1991, pp. 281–296.

[22] M. Rosenblatt and J.N. Van Ness, "Estimation of the bispectrum", *Ann. Math. Statist.*, Vol. 36, 1965, pp. 1120–1136.

[23] S. Shamsunder and G.B. Giannakis, "Modeling of non-Gaussian array data using cumulants: DOA estimation of more sources with less sensors", *Signal Processing*, Vol. 30, No. 3, February 1993, pp. 279–297.

[24] A. Swami, System identification using cumulant, PhD Thesis, USC, Los Angeles, California, 1988.

[25] A. Swami and J. Mendel, "Adaptive system identification using cumulants", *Proc. Internat. Conf. Acoust. Speech Signal Process.*, New York, NY, 1988, pp. 2248–2251.

[26] J.N. Van Ness, "Asymptotic normality of bispectral estimates", *Ann. Math. Statist.*, Vol. 37, 1966, pp. 1257–1272.

[27] F.-C. Zheng, S. McLaughlin and B. Mulgrew, "Blind equalization of multilevel PAM data for non minimum phase channels via second- and fourth-order cumulants", *Signal Processing*, Vol. 31, No. 3, April 1993, pp. 313–327.