Cyclostationary Signal Processing

William A. Gardner[1]
Chad M. Spooner
Department of Electrical and Computer Engineering
University of California
Davis, CA 95616

[1]W. A. Gardner is also with Statistical Signal Processing, Inc.

CONTROL AND DYNAMIC SYSTEMS, VOL. 65

I. INTRODUCTION

Many conventional statistical signal processing methods treat random signals as if they were statistically stationary, that is, as if the parameters of the signal model do not vary with time. But for most manmade signals encountered in communication, telemetry, radar, and sonar systems, some statistical parameters do vary periodically with time. In some cases even multiple incommensurate (not harmonically related) periodicities are involved. Examples of such periodicities include sinusoidal carriers in amplitude, phase, and frequency modulation systems, periodic keying of the amplitude, phase, or frequency in digital modulation systems, periodic scanning in television, facsimile, and some radar systems, and periodic motion in rotating machinery. Although in some cases these periodicities can be ignored by signal processors, such as receivers which must detect the presence of signals of interest, estimate their parameters, and/or extract their messages, in many cases there can be much to gain in terms of improvements in performance of these signal processors by recognizing and exploiting underlying periodicity. This typically requires that the random signal be modeled as *cyclostationary*, in which case the statistical parameters vary in time with single or multiple periodicities. Cyclostationarity also arises in signals of natural origin, due to the presence of rhythmic, seasonal, or other cyclic behavior. Examples include time-series data encountered in meteorology, climatology, atmospheric science, oceanology, astronomy, hydrology, biomedicine, and economics. This article introduces the principles of the theory of cyclostationary signals and the exploitation of the cyclostationarity property.

Let us be more specific about the nature of a cyclostationary signal. A signal is cyclostationary of order n (in the wide sense) if and only if we can find some nth-order homogeneous polynomial transformation of the signal that will generate finite-strength additive sine-wave components, which result in spectral lines. For example, for $n = 2$, a quadratic transformation (like the squared signal or the product of the signal with a delayed version of itself, or the weighted sum of such products) will generate spectral lines. For $n = 3$ or $n = 4$, cubic or quartic transformations (i.e., sums of weighted

products of 3 or 4 delayed versions of the signal) will generate spectral lines. In contrast, for stationary signals, only a spectral line at frequency zero can be generated.

Another way to describe a cyclostationary signal, which is completely equivalent to the first, but does not appear to be so upon first encounter, is as follows. A signal is cyclostationary of order n (in the wide sense) if and only if the time fluctuations in n distinct spectral bands with center frequencies that sum to certain discrete nonzero values are statistically dependent in the sense that their joint nth-order moment (the infinite-time average of their product in which each factor is shifted in frequency to have a center frequency of zero) is nonzero. In contrast, for stationary signals, only those bands whose center frequencies sum to zero can exhibit statistical dependence.

In fact, for a cyclostationary signal, each distinct sum of center frequencies for which the nth-order spectral moment is nonzero is identical to the frequency of a sine wave that can be generated by putting the signal through an appropriate nth-order nonlinear transformation.

For the simplest nontrivial case, which is $n = 2$, this means that a signal $x(t)$ is cyclostationary with *cycle frequency* α if and only if at least some of its delay-product waveforms, $y(t) = x(t-\tau)x(t)$ or $z(t) = x(t-\tau)x^*(t)$ (where $*$ denotes conjugation) for some delays τ, exhibit a spectral line at frequency α, and if and only if the time fluctuations in at least some pairs of spectral bands of $x(t)$, whose two center frequencies sum (for the case of $y(t)$) or difference (for the case of $z(t)$) to α, are correlated.

If not all cycle frequencies α for which a signal is cyclostationary are multiples of a single fundamental frequency (equal to the reciprocal of a fundamental period), then the signal is said to be *polycyclostationary* (although the term cyclostationary can also be used in this more general case when the distinction is not important), which means that there is more than one periodicity associated with the statistical parameters of the signal.

Let us now consider the utility of the property of cyclostationarity. Cyclostationarity can be used to enhance the accuracy and reliability of information gleaned from data sets such as measurements of corrupted signals. This enhancement is relative to the accuracy and reliability of information

that can be gleaned from stationary data sets or from cyclostationary data sets that are treated as if they were stationary. Such information includes the following:

1. A decision as to the presence or absence of a random signal, or about the number of random signals present in a data set that also contains background noise and other modulated signals,
2. A classification of multiple received signals present in a noisy data set according to modulation types,
3. An estimate of a signal parameter, such as carrier phase, pulse timing, or direction of arrival, based on a noise-and-interference-corrupted data set,
4. An estimate of an analog or digital message being communicated by a signal over a channel corrupted by noise, interference, and distortion,
5. A prediction of a future value of a random signal,
6. An estimate of the input-output relation of a linear or nonlinear system based on measurements of the system's response to random excitation,
7. An estimate of the degree of causality between two data sets, and
8. An estimate of the parameters of a model for a data set.

In the first part of Section II of this article, the possibility of generating spectral lines by simply squaring the signal is illustrated for two types of signals: the random-amplitude modulated sine wave and the random-amplitude modulated periodic pulse train. Then in the second part of Section II, it is explained that the property that enables spectral-line generation with some type of quadratic time-invariant transformation is called *cyclostationarity of order 2* (in the wide sense), and is characterized by the *cyclic autocorrelation function*, which is a generalization of the conventional autocorrelation function. Following this, it is shown that a signal exhibits cyclostationarity if and only if the signal is correlated with certain frequency-shifted versions of itself.

In the third and last part of Section II, the correlation of frequency-shifted versions of a signal is analyzed in the frequency domain and this leads to the definition of a *spectral-correlation density function.* It is then explained that this function is the Fourier transform of the cyclic autocorrelation function. This Fourier-transform relation includes as a special case the well-known *Wiener relation* between the power spectral density function and the autocorrelation function. A normalization of the spectral-correlation density function that converts it into a spectral correlation coefficient, whose magnitude is between zero and unity, is then introduced as a convenient measure of the degree of spectral redundancy in a signal. For example, if the spectral correlation coefficient is zero, then the associated spectral components are not redundant, and if the spectral correlation coefficient is unity, then the associated spectral components are completely redundant.

Continuing in the final part of Section II, the effects on the spectral-correlation density function of several signal-processing operations are described. These operations include filtering and waveform multiplication, which in turn include the special cases of time delay and multipath propagation, bandlimiting, frequency conversion, and time sampling. These results are used to derive the spectral-correlation density function for the amplitude-modulated sine wave, the amplitude modulated pulse train, and the binary phase-shift-keyed sine wave. The spectral-correlation density functions for some other types of phase-shift-keyed signals are also described graphically.

To conclude Section II, measurement of the (estimation of the ideal) spectral-correlation density function is discussed and a particular algorithm for this purpose is described and illustrated with a simulated phase-shift-keyed signal. To complement similar treatments of this material [1, 2], attention is focused in Section II on discrete-time signals rather than continuous-time signals[2].

Section III describes some ways of exploiting the inherent spectral redundancy associated with the spectral correlation in cyclostationary signals

[2] For convenience, the notation herein is modified from that in [1, 2]: here, $\tilde{R}_x^\alpha$ and $\tilde{S}_x^\alpha$ are used for continuous time and R_x^α and S_x^α are used for discrete time.

to perform various signal processing tasks. These tasks include detecting the presence of signals buried in noise and/or severely masked by interference; recognizing such corrupted signals according to modulation type; estimating parameters such as time difference of arrival at two reception platforms and direction of arrival at a reception array on a single platform; blind-adaptive spatial filtering of signals impinging on a reception array; reduction of signal corruption due to cochannel interference and/or channel fading for single-receiver systems; linear periodically time-variant prediction and causality; and identification of linear and nonlinear systems from measurements of the input and output. The task descriptions include brief explanations of how and why the signal processors that exploit spectral redundancy can outperform their more conventional counterparts that ignore spectral redundancy or, equivalently, ignore cyclostationarity. References to more detailed treatments are given throughout.

Finally, in Section IV, the principles of second-order cyclostationarity that are surveyed in Section II are generalized to higher orders. It is shown that, whereas temporal and spectral second-order moments play the central role in the theory of second-order wide-sense cyclostationarity, temporal and spectral cumulants arise naturally in the formulation of the theory of higher-order wide-sense cyclostationarity. Some applications of the theory of higher-order cyclostationarity are described.

II. PRINCIPLES OF CYCLOSTATIONARITY

A. SPECTRAL LINE GENERATION

A discrete-time signal $x(t)$, for $t = 0, \pm 1, \pm 2, \pm 3, \ldots$, contains a *finite-strength additive sine-wave component* (an ac component) with frequency α, say

$$a\cos(2\pi\alpha t + \theta) \quad \text{with } \alpha \neq 0, \tag{1}$$

if the Fourier coefficient

$$M_x^\alpha = \left\langle x(t)e^{-i2\pi\alpha t}\right\rangle \tag{2}$$

is not zero, in which case Eq. (1) gives

$$M_x^\alpha = \frac{ae^{i\theta}}{2}.$$

In Eq. (2), the operation $\langle\cdot\rangle$ is the time-averaging operation

$$\langle g(t)\rangle \triangleq \lim_{Z\to\infty} \frac{1}{2Z+1} \sum_{t=-Z}^{Z} g(t).$$

In this case, the power spectral density (PSD) of $x(t)$ includes a spectral line at $f = \alpha$ and its image $f = -\alpha$. (The PSD is defined later in this section.) That is, the PSD in the principal domain $(-1/2, 1/2]$ contains the additive term[3]

$$|M_x^\alpha|^2 [\delta(f-\alpha) + \delta(f+\alpha)] \tag{3}$$

where $\delta(f)$ is the Dirac delta (or impulse) function. For convenience in the sequel, it is said that such a signal exhibits *first-order periodicity* with frequency α.

Let $x(t)$ be decomposed into the sum of its finite-strength sine-wave component, with frequency α, and its residual, say $n(t)$,

$$x(t) = a\cos(2\pi\alpha t + \theta) + n(t), \tag{4}$$

where $n(t)$ is defined to be that which is left after subtraction of Eq. (1) from $x(t)$. It is assumed that $n(t)$ is random. Here, the term *random* is used to denote nothing more than the vague notion of erratic or unpredictable behavior. If the sine wave is weak relative to the random residual, it might not be evident from visual inspection of $x(t)$ that it contains a periodic component. Hence, it is said to contain *hidden periodicity.* However, because of the associated spectral lines, hidden periodicity can be detected, and in some applications exploited, through techniques of spectral analysis.

This article is concerned with signals that contain more subtle types of hidden periodicity that, unlike first-order periodicity, do not give rise to spectral lines in the PSD, but that can be converted into first-order periodicity by a nonlinear time-invariant transformation of the signal. In

[3]The strength of the spectral line is $|M_x^\alpha|^2$ as indicated in Eq. (3) if and only if the limit Eq. (2) exists in the temporal mean-square sense with respect to the time parameter u obtained by replacing t with $t+u$ in Eq. (2) [1, Chapter 15, Exercise 6].

particular, we shall focus in Sections I-III on the type of hidden periodicity that can be converted by a quadratic transformation to yield spectral lines in the PSD of the transformed signal. In Section IV, we shall consider hidden periodicity that requires higher-order nonlinear transformations for conversion into first-order periodicity.

The theory presented here is couched entirely in terms of time averages, because time averages arise naturally in our inquiry into hidden periodicity. Nevertheless, it is mentioned that this theory can be reformulated in terms of stochastic processes. In fact, the theory presented here can be translated to the alternative theory (cf. [3]) simply by following the rule[4]:

> For all sinusoidally weighted time averages
>
> $$\left\langle z(t)e^{-i2\pi\alpha t}\right\rangle$$
>
> of time-series $z(t)$, replace $z(t)$ by the expected value $E\{Z(t)\}$ of the corresponding stochastic process $Z(t)$ to obtain
>
> $$\left\langle E\{Z(t)\}e^{-i2\pi\alpha t}\right\rangle.$$

Common examples of $z(t)$ appearing in this presentation include lag products $z(t) = x(t)x^*(t-\tau)$ and cross products $z(t) = u(t)v^*(t)$.

The discussion begins with two motivating examples. In the convention used here, the PSD for $x(t)$ is denoted by $S_x(f)$ and is periodic with unity period. $\tilde{S}_x(f)$ denotes the PSD restricted to the principal domain $(-1/2, 1/2]$, therefore, the PSD is given by

$$S_x(f) = \sum_{n=-\infty}^{\infty} \tilde{S}_x(f+n).$$

On occasion, continuous-time signals also are discussed herein. In such cases it is assumed that the signal is time-scaled and bandlimited so that the PSD is restricted to the band $(-1/2, 1/2]$. Consequently, the PSD of the discrete-time sampled version, restricted to the principal domain, will be identical to the PSD of the continuous-time signal and the same notation, $\tilde{S}_x(f)$, is used for both.

[4]When $\alpha = 0$, the operation $\langle\cdot\rangle$ can be omitted from $\langle E\{Z(t)\}\rangle$ *only* if $Z(t)$ is purely stationary, i.e., stationary and cycloergodic [3].

Example 1: AM. Let $a(t)$ be a random lowpass signal (say lowpass filtered thermal noise) with the PSD $S_a(f)$ shown in Figure 1a, which contains no spectral lines. If $a(t)$ is used to modulate the amplitude of a sine wave, we obtain the amplitude modulated (AM) signal

$$x(t) = a(t)\cos(2\pi f_0 t), \tag{5}$$

whose PSD $S_x(f)$ is given by [1, Chapter 3, Sec. D]

$$S_x(f) = \frac{1}{4}S_a(f + f_0) + \frac{1}{4}S_a(f - f_0) \tag{6}$$

as shown in Figure 1b.

Although the PSD is centered about $f = f_0$ and $f = -f_0$, there is no spectral line at f_0 or $-f_0$. The reason for this is that, as shown in Figure 1a, there is no spectral line in $S_a(f)$ at $f = 0$. This means that the dc component

$$M_a^0 \triangleq \langle a(t) \rangle \tag{7}$$

is zero, since the strength of any spectral line at $f = 0$ is $|M_a^0|^2$. Let us now square $x(t)$ to obtain

$$\begin{aligned} y(t) &= x^2(t) \\ &= a^2(t)\cos^2(2\pi f_0 t) \\ &= \frac{1}{2}[b(t) + b(t)\cos(4\pi f_0 t)] \end{aligned} \tag{8}$$

where

$$b(t) = a^2(t). \tag{9}$$

Since $b(t)$ is nonnegative, its dc value must be positive: $M_b^0 > 0$. Consequently, the PSD of $b(t)$ contains a spectral line at $f = 0$, as shown in Figure 1c. The PSD for $y(t)$ is given by

$$S_y(f) = \frac{1}{4}[S_b(f) + \frac{1}{4}S_b(f + 2f_0) + \frac{1}{4}S_b(f - 2f_0)] \tag{10}$$

and, as shown in Figure 1d, it contains spectral lines at $f = \pm 2f_0$ as well as at $f = 0$. Thus, by putting $x(t)$ through a quadratic transformation (a squarer in this case) we have converted the hidden periodicity resulting from the sine-wave factor $\cos(2\pi f_0 t)$ in Eq. (5) into first-order periodicity

with associated spectral lines. This is particularly easy to see if $a(t)$ is a binary sequence that switches randomly between 1 and -1 because then $b(t) = 1$ and $y(t)$ in Eq. (8) is therefore a periodic signal:

$$y(t) = \frac{1}{2} + \frac{1}{2}\cos(4\pi f_0 t),$$

which clearly contains finite-strength additive sine waves. □

Example 2: PAM. As another example, we consider the pulse-amplitude modulated (PAM) signal

$$x(t) = \sum_{n=-\infty}^{\infty} a(nT_0)p(t - nT_0) \tag{11}$$

where the pulse $p(t)$ is confined within the interval $(-T_0/2, T_0/2)$ so that the pulse translates do not overlap, as shown in Figure 2. For simplicity, we consider a continuous-time signal in this example (to avoid aliasing). The PSD of $x(t)$ is given by [1, Chapter 3, Sec. D]

$$\tilde{S}_x(f) = \frac{1}{T_0}\left|\tilde{P}(f)\right|^2 \sum_{m=-\infty}^{\infty} \tilde{S}_a(f - m/T_0), \tag{12}$$

where $\tilde{S}_a(f)$ is shown in Figure 1a, which contains no spectral lines, and where $\tilde{P}(f)$ is the Fourier transform of $p(t)$. Since there are no spectral lines in $\tilde{S}_a(f)$ (or $\tilde{P}(f)$ since $p(t)$ has finite duration), there are none in $\tilde{S}_x(f)$, as shown in Figure 3a, regardless of the periodic repetition of pulses in $x(t)$. But, let us look at the square of $x(t)$:

$$y(t) = x^2(t) = \sum_{n=-\infty}^{\infty} b(nT_0)q(t - nT_0), \tag{13}$$

where

$$b(nT_0) = a^2(nT_0) \tag{14}$$

and

$$q(t) = p^2(t). \tag{15}$$

The PSD for $y(t)$ is given by

$$\tilde{S}_y(f) = \frac{1}{T_0}\left|\tilde{Q}(f)\right|^2 \sum_{m=-\infty}^{\infty} \tilde{S}_b(f - m/T_0), \tag{16}$$

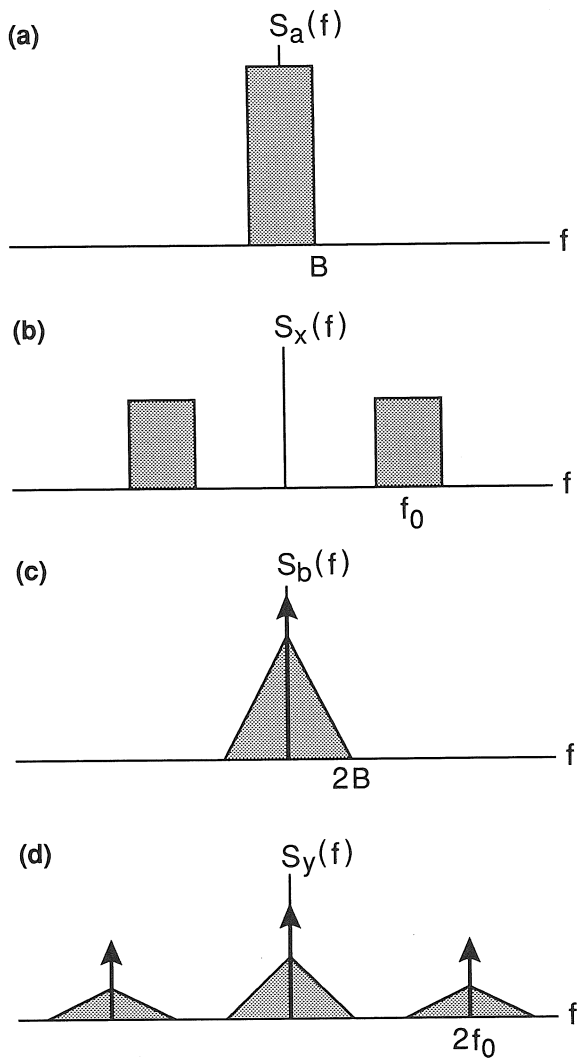

Figure 1: (a) Power spectral density (PSD) of a lowpass signal. (b) PSD of an amplitude-modulated (AM) signal. (c) PSD of a squared lowpass signal. (d) PSD of a squared AM signal.

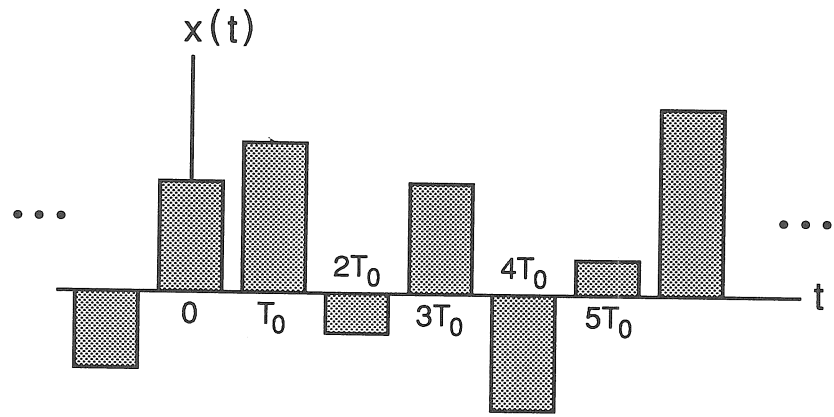

Figure 2: A pulse-amplitude-modulated (PAM) signal with pulse width less than pulse repetition time.

where $\tilde{Q}(f)$ is the Fourier transform of $q(t)$. Because of the spectral line at $f = 0$ in $\tilde{S}_b(f)$, which is shown in Figure 1c, we have spectral lines in $\tilde{S}_y(f)$ at the harmonics m/T_0 (for some integer values of m) of the pulse rate $1/T_0$, as shown in Figure 3b. Thus, again, we have converted the hidden periodicity in $x(t)$ into first-order periodicity with associated spectral lines by using a quadratic transformation. This is particularly easy to see if $a(nT_0)$ is a random binary sequence with values ± 1, because then $b(nT_0) \equiv 1$ and $y(t)$ in Eq. (13) is therefore a periodic signal

$$y(t) = \sum_{n=-\infty}^{\infty} q(t - n/T_0),$$

which clearly contains finite-strength additive sine waves. □

B. THE CYCLIC AUTOCORRELATION FUNCTION

Although the squaring transformation works in these examples, a different quadratic transformation involving delays can be required in some cases. For example, if $a(nT_0)$ in Example 2 is again binary, but $p(t)$ is flat with height 1 and width T_0, as shown in Figure 4, then $y(t) = x^2(t) = 1$, which is a constant for all t. Thus, we have a spectral line at $f = 0$ but none at the harmonics of the pulse rate. Nevertheless, if we use the quadratic

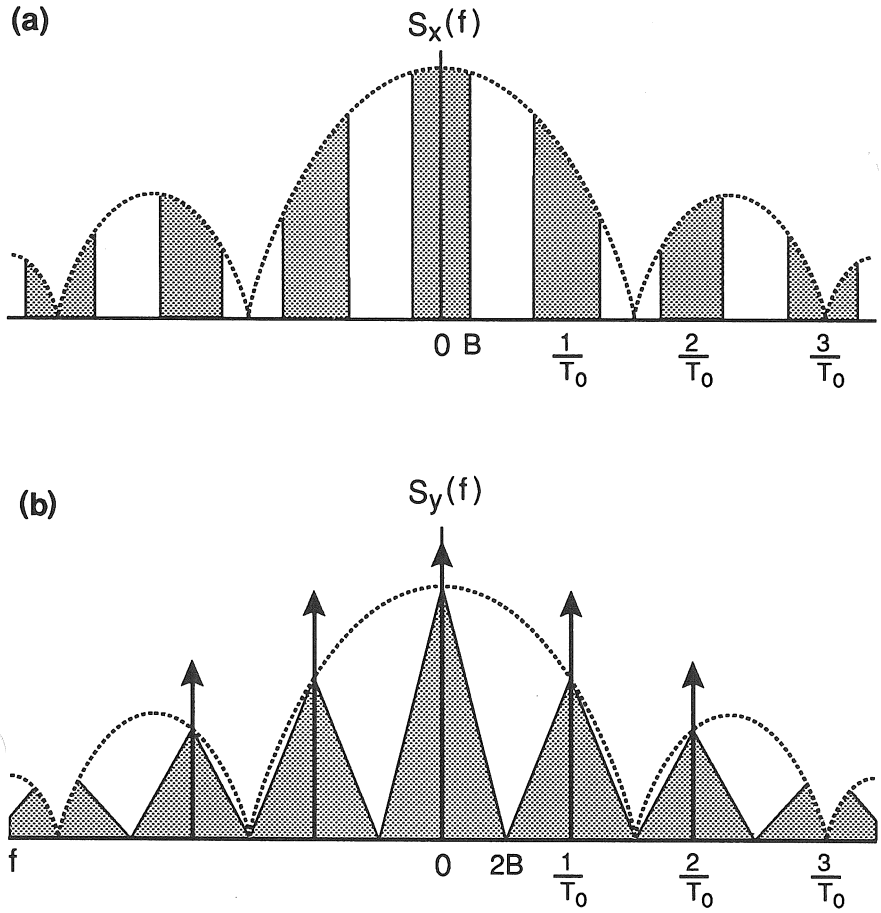

Figure 3: (a) Power spectral density (PSD) of a pulse-amplitude-modulated (PAM) signal with 67% duty-cycle pulses. (b) PSD of the squared PAM signal.

transformation

$$y(t) = x(t)x(t-\tau) \tag{17}$$

for any of a number of nonzero delays τ, we will indeed obtain spectral lines at $f = m/T_0$. That is,

$$\begin{aligned} M_y^\alpha &= \langle y(t)e^{-i2\pi\alpha t}\rangle \\ &= \langle x(t)x(t-\tau)e^{-i2\pi\alpha t}\rangle \neq 0 \end{aligned} \tag{18}$$

for $\alpha = m/T_0$ for some integers $m \neq 0$.

The most general time-invariant quadratic transformation for real signals is simply a linear combination of delay products

$$y(t) = \sum_{\tau_1,\tau_2} h(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)$$

for some weighting function $h(\tau_1,\tau_2)$ that is analogous to the impulse-response function for a linear transformation. This motivates us to define the property of second-order periodicity as follows: The real-valued signal $x(t)$ contains second-order periodicity if and only if the PSD of the delay-product signal $x(t+\tau_1)x(t+\tau_2)$ for some delays τ_1 and τ_2 contains at least one spectral line at some nonzero frequency α. But, this will be so if and only if the PSD of Eq. (17) for some delays ($\tau = \tau_2 + \tau_1$) contains at least one spectral line at some nonzero frequency α, that is, if and only if Eq. (18) is satisfied for some $\alpha \neq 0$.

In developing the continuous-time theory of second-order periodicity it has been found to be more convenient to work with the symmetric delay product

$$y_\tau(t) = x(t+\tau/2)x^*(t-\tau/2). \tag{19}$$

The complex conjugate $*$ is introduced here for generality to accommodate complex-valued signals, but it is mentioned that for some complex-valued signals, the quadratic transformation without the conjugate can also be useful [1, Chapter 10, Sec. C]. From Eq. (19), the fundamental parameter Eq. (18) of second-order periodicity for continuous time becomes

$$\tilde{R}_x^\alpha(\tau) \triangleq \langle x(t+\tau/2)x^*(t-\tau/2)e^{-i2\pi\alpha t}\rangle, \tag{20}$$

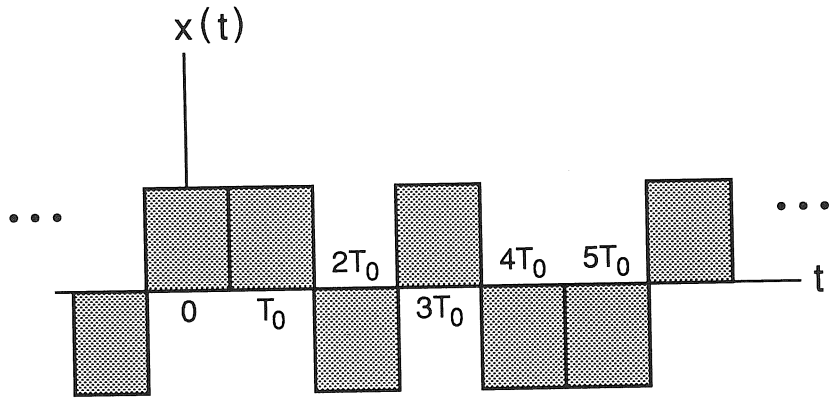

Figure 4: A binary pulse-amplitude-modulated (PAM) signal with full duty-cycle pulses.

which is the Fourier coefficient $M_{y_\tau}^{\alpha}$ of the additive sine-wave component with frequency α contained in the delay-product signal $y_\tau(t)$. However, for discrete-time signals, delays that are not equal to multiples of the sampling increment are not allowed (i.e. $\pm\tau/2$ for τ odd). Nevertheless, since

$$\left\langle x(t)x^*(t-\tau)e^{-i2\pi\alpha t}\right\rangle = \tilde{R}_x^{\alpha}(t)e^{-i\pi\alpha\tau}$$

for continuous time, then we can define the fundamental parameter of second-order periodicity for discrete-time as follows

$$R_x^{\alpha}(\tau) \triangleq \left\langle x(t)x^*(t-\tau)e^{-i2\pi\alpha t}\right\rangle e^{i\pi\alpha\tau} \tag{21}$$

in order to maintain the strongest analogy between the continuous- and discrete-time theories. Observe that since t and τ take on only integer values, then $R_x^{\alpha}(\tau)$ is periodic in α with period two, and also $R_x^{\alpha+1}(\tau) = R_x^{\alpha}(\tau)e^{i\pi\tau}$.

The notation $R_x^{\alpha}(\tau)$ is introduced for this Fourier coefficient because, for $\alpha = 0$, Eq. (21) reduces to the conventional autocorrelation function

$$R_x^0(\tau) = \langle x(t)x^*(t-\tau)\rangle,$$

for which the notation $R_x(\tau)$ is commonly used. Furthermore, since $R_x^\alpha(\tau)$ is a generalization of the autocorrelation function, in which a cyclic (sinusoidal) weighting factor $e^{-i2\pi\alpha t}$ is included before the time averaging is carried out, $R_x^\alpha(\tau)$ is called the *cyclic autocorrelation function.* Also, the *conjugate cyclic autocorrelation* for complex-valued signals obtained from Eq. (21) by deleting the conjugate,

$$R_{xx^*}^\alpha(\tau) = \left\langle x(t)x(t-\tau)e^{-i2\pi\alpha t}\right\rangle e^{i\pi\alpha\tau}, \tag{22}$$

is a further modification of the conventional autocorrelation[5].

We have given two distinct interpretations of $R_x^\alpha(t) = M_{y_\tau}^\alpha$. In fact, we have yet a third distinct interpretation, which can be obtained by factoring $e^{-i2\pi\alpha t}$ in order to reexpress Eq. (21) as

$$R_x^\alpha(\tau) = \left\langle \left[x(t)e^{-i\pi\alpha t}\right]\left[x(t-\tau)e^{i\pi\alpha(t-\tau)}\right]^*\right\rangle. \tag{23}$$

That is, $R_x^\alpha(\tau)$ is actually a conventional crosscorrelation function

$$R_{uv}(\tau) \triangleq \langle u(t)v^*(t-\tau)\rangle = R_x^\alpha(\tau), \tag{24}$$

where

$$u(t) = x(t)e^{-i\pi\alpha t} \tag{25}$$

and

$$v(t) = x(t)e^{i\pi\alpha t} \tag{26}$$

are frequency translates of $x(t)$. Recall that multiplying a signal by $e^{\pm i\pi\alpha t}$ shifts the spectral content of the signal by $\pm\alpha/2$. For example, the PSDs of $u(t)$ and $v(t)$ are

$$S_u(f) = S_x(f+\alpha/2) \tag{27}$$

and

$$S_v(f) = S_x(f-\alpha/2). \tag{28}$$

[5]Although some readers will recognize the similarity between the cyclic autocorrelation function and the radar ambiguity function, the relationship between these two functions is only superficial. The concepts and theory underlying the cyclic autocorrelation function, as summarized in this article, have little in common with the concepts and theory of radar ambiguity (cf. [1, Chapter 10, Sec. C]). For example, the radar ambiguity function has no meaning relevant to ambiguity (in Doppler) when applied to a real signal, or when applied to a complex signal without the conjugate.

It follows from Eqs. (24)–(26) that $x(t)$ exhibits second-order periodicity (Eq. (21) is not identically zero as a function of τ for some $\alpha \neq 0$) if and only if *frequency translates* of $x(t)$ are correlated with each other in the sense that Eq. (24) is not identically zero as a function of τ for some $\alpha \neq 0$ in Eqs. (25)–(26). This third interpretation of $R_x^\alpha(\tau)$ suggests an appropriate way to normalize $R_x^\alpha(\tau)$ as explained next.

As long as the mean values of the frequency translates $u(t)$ and $v(t)$ are zero (which means that $x(t)$ does not contain finite-strength additive sine-wave components[6] at frequencies $\pm\alpha/2$ and, therefore, that $S_x(f)$ has no spectral lines at $f = \pm\alpha/2$), the crosscorrelation $R_{uv}(\tau) = R_x^\alpha(\tau)$ is actually a temporal crosscovariance $K_{uv}(\tau)$. That is,

$$\begin{aligned} K_{uv}(\tau) &\triangleq \langle [u(t) - \langle u(t)\rangle][v(t-\tau) - \langle v(t-\tau)\rangle]^*\rangle \\ &= \langle u(t)v^*(t-\tau)\rangle = R_{uv}(\tau). \end{aligned} \tag{29}$$

An appropriate normalization for the temporal crosscovariance is the geometric mean of the two corresponding temporal variances. This yields a temporal correlation coefficient, the magnitude of which is upper bounded by unity. It follows from Eqs. (25)–(26) that the two variances are given by

$$K_u(0) = R_u(0) = \langle |u(t)|^2\rangle = R_x(0) \tag{30}$$

and

$$K_v(0) = R_v(0) = \langle |v(t)|^2\rangle = R_x(0). \tag{31}$$

Therefore, the temporal correlation coefficient for frequency translates is given by

$$\gamma_x^\alpha(\tau) \triangleq \frac{K_{uv}(\tau)}{[K_u(0)K_v(0)]^{1/2}} = \frac{R_x^\alpha(\tau)}{R_x(0)}. \tag{32}$$

Hence, the appropriate normalization factor for the cyclic autocovariance $R_x^\alpha(\tau)$ is simply $1/R_x(0)$ (and it is the same for the conjugate cyclic autocovariance).

This is a good point at which to introduce some more terminology. A signal $x(t)$ for which the autocorrelation $R_x(\tau)$ exists (e.g., remains finite as the averaging time goes to infinity) and is not identically zero (as it

[6] It does contain infinitesimal sine-wave components.

is for transient signals) is commonly said to be stationary (in the wide sense). But we need to refine the terminology to distinguish between those stationary signals that exhibit second-order periodicity ($R_x^\alpha(\tau) \not\equiv 0$ for some $\alpha \neq 0$) and those stationary signals that do not ($R_x^\alpha(\tau) \equiv 0$ for all $\alpha \neq 0$). Consequently, we shall call the latter for which $R_x^\alpha(\tau) \equiv 0$ *stationary of second order* (in the wide sense) and the former for which $R_x^\alpha(\tau) \not\equiv 0$ for some values of α that are integer multiples of a single fundamental frequency $1/T$ (corresponding to the period T) *cyclostationary of second order* (in the wide sense). If there is more than one fundamental frequency, then we shall call the signal *polycyclostationary.* We shall also call any nonzero value of the frequency parameter α in the principal domain $(-1/2, 1/2]$ for which $R_x^\alpha(\tau) \not\equiv 0$ a *cycle frequency.* The discrete set of cycle frequencies is called the *cycle spectrum.* For example, if a signal is cyclostationary, the cycle spectrum contains only harmonics (integer multiples) of the fundamental cycle frequency, which is the reciprocal of the fundamental period. But if the signal is polycyclostationary, then the cycle spectrum can contain harmonics of each of the incommensurate fundamental cycle frequencies.

It should be clarified at this point that *wide-sense* theory deals with moments (and their combinations that form cumulants), whereas *strict-sense* theory deals with probability distributions. A signal is cyclostationary of order n in the strict sense if and only if its nth-order fraction-of-time probability distribution changes periodically with time, whereas a signal is cyclostationary of order n in the wide sense if and only if its nth-order joint moment function changes periodically with time (cf. Section IV).

We conclude this section by determining the cyclic autocorrelation function for the AM signal considered in Example 1.

Example 1 continued: AM. Let $a(t)$ be a real random stationary signal with zero mean:

$$\langle a(t) \rangle = 0, \tag{33}$$

$$\langle a(t)a^*(t-\tau) \rangle \not\equiv 0, \tag{34}$$

$$\left\langle a(t)a^*(t-\tau)e^{-i2\pi\alpha t} \right\rangle \equiv 0 \text{ for all } \alpha \neq 0. \tag{35}$$

Equation (35) guarantees that

$$\left\langle a(t)e^{-i2\pi\alpha t} \right\rangle \equiv 0 \text{ for all } \alpha \neq 0. \tag{36}$$

We consider the amplitude-modulated sine wave

$$\begin{aligned} x(t) &= a(t)\cos(2\pi f_0 t + \theta) \\ &= \frac{a(t)}{2}\left[e^{i(2\pi f_0 t+\theta)} + e^{-i(2\pi f_0 t+\theta)}\right]. \end{aligned} \tag{37}$$

Because of Eq. (36), $a(t)$ contains no finite-strength additive sine-wave components and, therefore (together with Eq. (33)), $x(t)$ contains no finite-strength additive sine-wave components. This means that its power spectral density contains no spectral lines. However, the quadratic transformation

$$\begin{aligned} y_\tau(t) &= x(t)x^*(t-\tau) \\ &= \frac{1}{4}a(t)a^*(t-\tau)\left[e^{i2\pi f_0\tau} + e^{-i2\pi f_0\tau} + e^{i(4\pi f_0 t+2\theta)}e^{-i2\pi f_0\tau}\right. \\ &\quad \left. + e^{-i(4\pi f_0 t+2\theta)}e^{i2\pi f_0\tau}\right] \end{aligned} \tag{38}$$

does contain finite-strength additive sine-wave components with frequencies $\alpha = \pm 2f_0$, since Eq. (34) renders one or the other of the last two terms in the quantity

$$\begin{aligned} \left\langle y_\tau(t)e^{-i2\pi\alpha t}\right\rangle &= \frac{1}{4}e^{i2\pi f_0\tau}\left\langle a(t)a^*(t-\tau)e^{-i2\pi\alpha t}\right\rangle \\ &\quad + \frac{1}{4}e^{-i2\pi f_0\tau}\left\langle a(t)a^*(t-\tau)e^{-i2\pi\alpha t}\right\rangle \\ &\quad + \frac{1}{4}e^{i2\theta}e^{-i2\pi f_0\tau}\left\langle a(t)a^*(t-\tau)e^{-i2\pi(\alpha-2f_0)t}\right\rangle \\ &\quad + \frac{1}{4}e^{-i2\theta}e^{i2\pi f_0\tau}\left\langle a(t)a^*(t-\tau)e^{-i2\pi(\alpha+2f_0)t}\right\rangle \end{aligned} \tag{39}$$

nonzero for $\alpha = \pm 2f_0$. That these are the only two nonzero cycle frequencies α follows from the fact that Eq. (35) renders Eq. (39) equal to zero for all α except $\alpha = 0$ and $\alpha = \pm 2f_0$. Thus, the cycle spectrum consists of only the two cycle frequencies $\alpha = \pm 2f_0$ and the degenerate cycle frequency $\alpha = 0$.

Hence, the versions $u(t)$ and $v(t)$ of $x(t)$ obtained by frequency shifting $x(t)$ up and down by $\alpha/2 = f_0$ are correlated. This is not surprising since Eq. (37) reveals that $x(t)$ is obtained from $a(t)$ by frequency shifting up and down by f_0 and then adding. In conclusion, we have the cyclic

autocorrelation function (in the principal domain of α)

$$R_x^\alpha(\tau) = \begin{cases} \frac{1}{4}e^{\pm i2\theta}R_a(\tau) & \alpha = \pm 2f_0 \\ \frac{1}{2}R_a(\tau)\cos(2\pi f_0\tau) & \alpha = 0 \\ 0 & \text{otherwise,} \end{cases} \tag{40}$$

from which it follows that the temporal correlation coefficient is given by

$$\gamma_x^\alpha(\tau) = \begin{cases} \frac{1}{4}e^{\pm i2\theta}\gamma_a^0(\tau) & \alpha = \pm 2f_0 \\ \frac{1}{2}\gamma_a^0(\tau)\cos(2\pi f_0\tau) & \alpha = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{41}$$

Thus, the strength of correlation between $x(t)e^{-i\pi\alpha t}$ and $x(t-\tau)e^{i\pi\alpha(t-\tau)}$, which is given by

$$|\gamma_x^\alpha(\tau)| = \frac{1}{2}\left|\gamma_a^0(\tau)\right|, \quad \alpha = \pm 2f_0, \tag{42}$$

can be substantial (as large as 1/2) for this amplitude-modulated signal.

As an especially simple example of $a(t)$, we consider as before a random binary sequence that switches back and forth between 1 and -1. If we set $\tau = 0$ in Eq. (38), we obtain

$$\begin{aligned} y_0(t) &= |x(t)|^2 = |a(t)|^2\cos^2(2\pi f_0 t + \theta) \\ &= \frac{1}{2} + \frac{1}{2}\cos(4\pi f_0 t + 2\theta), \end{aligned}$$

which clearly contains finite-strength additive sine-wave components with frequencies $\alpha = \pm 2f_0$. In fact, in this very special case, there is no random component in $y_0(t)$. On the other hand, for $\tau \neq 0, y_\tau(t)$ can contain both a sine-wave component and a random component.

To illustrate the conjugate cyclic autocorrelation Eq. (22), let us consider the analytic signal for AM,

$$z(t) = \frac{1}{2}a(t)e^{i(2\pi f_0 t+\theta)}.$$

For this signal, we have

$$\begin{aligned} R_{zz^*}^\alpha(\tau) &\triangleq \left\langle z(t)z(t-\tau)e^{-i2\pi\alpha t}\right\rangle e^{i\pi\alpha\tau} \\ &= \frac{1}{4}\left\langle a(t)a(t-\tau)e^{i2\pi(2f_0-\alpha)t}\right\rangle e^{-i[2\pi(f_0-\alpha/2)\tau-2\theta]} \\ &= \begin{cases} \frac{1}{4}R_a(\tau)e^{i2\theta}, & \text{for } \alpha = 2f_0 \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Other examples of cyclostationary and polycyclostationary signals can be similarly viewed as mixtures of stationarity and periodicity. Examples are cited in Section I. Typical cycle spectra include harmonics of pulse rates, keying rates, spreading-code chipping rates, frequency-hopping rates, code-repetition rates, doubled-carrier frequencies, and sums and differences of these [1, Chapter 12]. □

C. THE SPECTRAL-CORRELATION DENSITY FUNCTION

In the same way that it is beneficial for some purposes to analyze in the frequency domain the average power $\langle |x(t)|^2 \rangle = R_x(0)$ in a stationary random signal, it can be very helpful to localize in frequency the correlation $\langle u(t)v^*(t) \rangle = \langle |x(t)|^2 e^{-i2\pi\alpha t} \rangle = R_x^\alpha(0)$ of frequency-shifted signals $u(t)$ and $v(t)$ for a cyclostationary or polycyclostationary random signal $x(t)$. In the former case of localizing the power, we simply pass the signal of interest $x(t)$ through a narrowband bandpass filter and then measure the average power at the output of the filter. By doing this with many filters whose center frequencies are separated by the bandwidth of the filters, we can partition any spectral band of interest into a set of contiguous narrow disjoint bands. In the limit as the bandwidths approach zero, the corresponding set of measurements of average power, normalized by the bandwidth, constitute the power spectral density (PSD) function. That is, at any particular frequency f (in the principal domain $(-1/2, 1/2]$), the PSD for $x(t)$ is given by

$$S_x(f) \triangleq \lim_{B \to 0} \frac{1}{B} \left\langle \left| h_B^f(t) \otimes x(t) \right|^2 \right\rangle, \tag{43}$$

where $\otimes$ denotes convolution and $h_B^f(t)$ is the discrete-impulse response of a one-sided bandpass filter with center frequency f, bandwidth B, and unity gain at the band center (see Figure 5).

In the latter case of localizing the correlation, we simply pass both of the two frequency translates $u(t)$ and $v(t)$ of $x(t)$ through the same set of bandpass filters that are used for the PSD measurement and then measure the temporal correlation of the filtered signals (see Figure 6) to obtain

$$S_x^\alpha(f) \triangleq \lim_{B \to 0} \frac{1}{B} \left\langle \left[h_B^f(t) \otimes u(t) \right] \left[h_B^f(t) \otimes v(t) \right]^* \right\rangle, \tag{44}$$

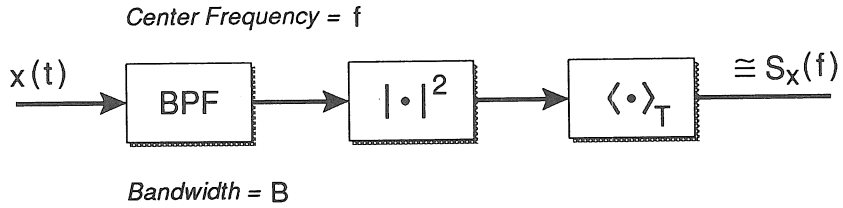

Figure 5: One channel of a spectrum analyzer for measuring the power spectral density (PSD). (The symbol $\cong$ indicates that the output only approximates the ideal function $S_x(f)$ for finite T and B.)

which is called the *spectral-correlation density* (SCD) *function.* This yields the spectral density of correlation in $u(t)$ and $v(t)$ at frequency f, which is identical to the spectral density of correlation in $x(t)$ at frequencies $f+\alpha/2$ and $f-\alpha/2$ (see Figure 7). That is, $S_x^\alpha(f)$ is the bandwidth-normalized (i.e., divided by B) correlation of the amplitude and phase fluctuations of the narrowband spectral components in $x(t)$ centered at frequencies $f+\alpha/2$ and $f-\alpha/2$, in the limit as the bandwidth B of these narrowband components approaches zero. For complex-valued signals, the conjugate SCD obtained from Eq. (44) by deleting the complex conjugate is also of interest for som signals [1, Chapter 10, Sec. C].

Strictly speaking, the SCD is not a valid density function in the usual sense, since it is not nonnegative and, in fact, not even real-valued. However, its integral over all frequencies does equal the correlation of $u(t)$ and $v(t)$ and, when $u(t)$ and $v(t)$ are decomposed into narrowband spectral components, the correlation of the components centered at f is indeed the SCD evaluated at f. Because of the lack of the nonnegativity property of the SCD, the correlation of $u(t)$ and $v(t)$ can equal zero without the SCD being identically zero because the integral of the SCD over all f can be zero even though the SCD is not identically zero. Nevertheless, because of the properties that the SCD does share with densities like the PSD, the term *density* is retained.

It is well known (see, for example, [1, Chapter 3, Sec. C] for a proof

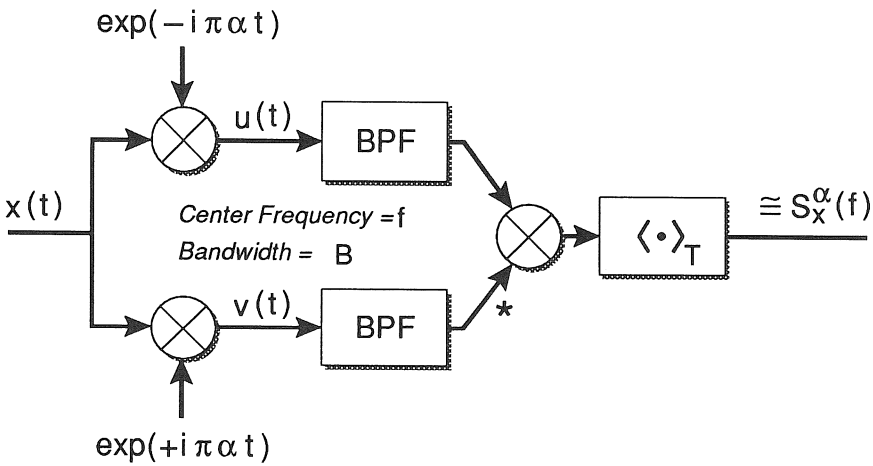

Figure 6: One channel-pair of a spectral correlation analyzer (or a cyclic spectrum analyzer) for measuring the spectral correlation density (or cyclic spectral density).

for continuous time) that the PSD obtained from Eq. (43) is equal to the Fourier transform of the autocorrelation function,

$$S_x(f) = \sum_{\tau=-\infty}^{\infty} R_x(\tau)\, e^{-i2\pi f\tau}. \tag{45}$$

Similarly, it can be shown (cf. [1, Chapter 11, Sec. C] for continuous time) that the SCD (or conjugate SCD) obtained from Eq. (44) is the Fourier transform of the cyclic autocorrelation function (or conjugate cyclic autocorrelation),

$$S_x^\alpha(f) = \sum_{\tau=-\infty}^{\infty} R_x^\alpha(\tau)\, e^{-i2\pi f\tau}, \tag{46}$$

and, therefore, $R_x^\alpha(\tau)$ is given by the inverse transform

$$R_x^\alpha(\tau) = \int_{-1/2}^{1/2} S_x^\alpha(f)\, e^{i2\pi f\tau}\, df. \tag{47}$$

Since $R_x^\alpha(\tau)$ is periodic in α with period two, so too is $S_x^\alpha(f)$. Also, since τ takes on only integer values, then $S_x^\alpha(f)$ is periodic in f with period

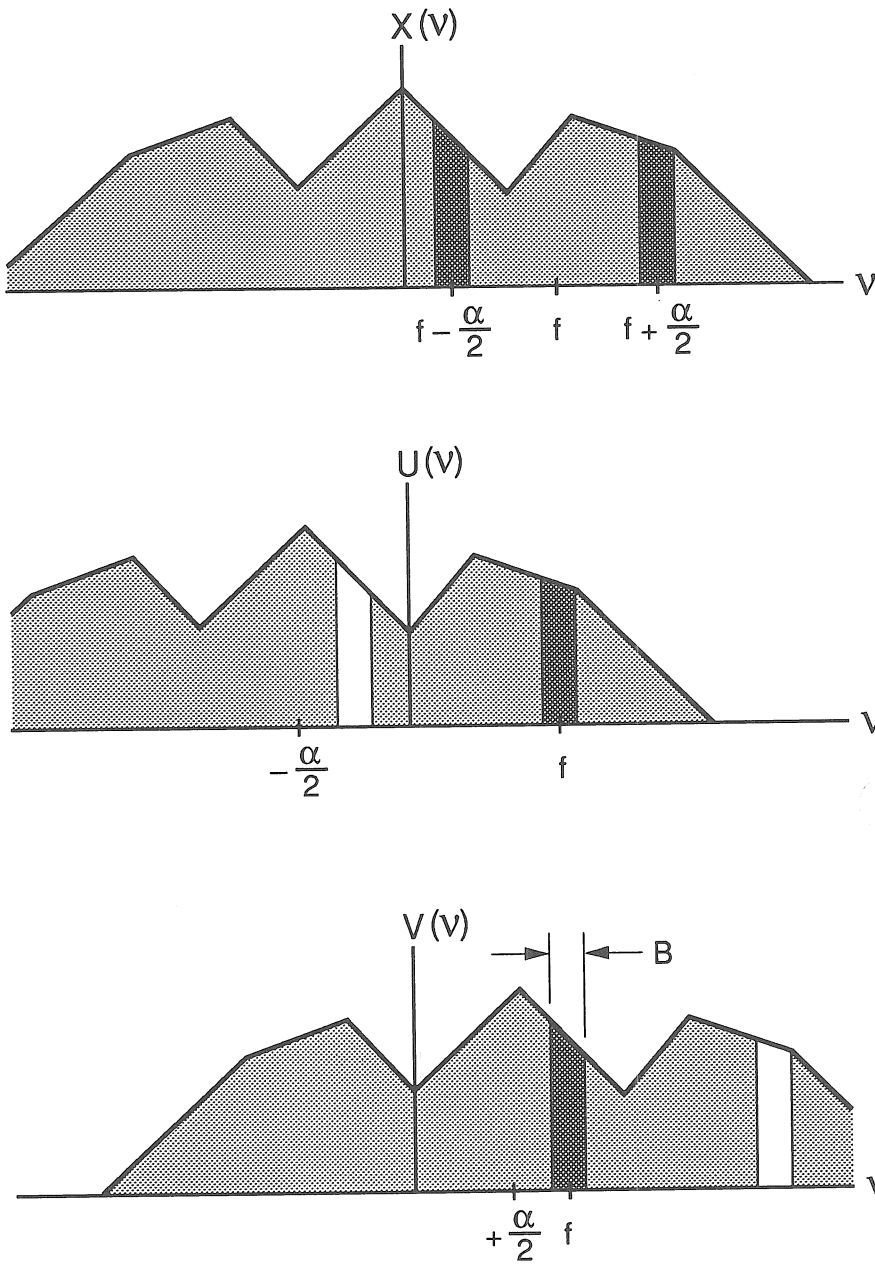

Figure 7: Illustration of spectral bands used in the measurement of the spectral-correlation density $S_x^\alpha(f)$. (ν is a dummy frequency variable; the shaded bands are the bands selected by the BPFs.)

one. Furthermore, since increasing $f \pm \alpha/2$ by ± 1 has no effect on the spectral components at these frequencies, then it follows that $S_x^\alpha(f)$ also exhibits the periodicity $S_x^{\alpha+1}(f + 1/2) = S_x^\alpha(f)$. Consequently, the principal domain for $S_x^\alpha(f)$ can be taken to be either the square with vertices $(f, \alpha) = (\pm 1/2, \pm 1/2)$ or the diamond with vertices $(f, \alpha) = (0, \pm 1)$ and $(\pm 1/2, 0)$. Relation Eq. (45) is known as the *Wiener relation* (see, for example, [1, Chapter 3, Sec. C]), and Eq. (46) is therefore called the *cyclic Wiener relation* [1, Chapter 11, Sec. C]. The cyclic Wiener relation includes the Wiener relation as the special case of $\alpha = 0$. (In the probabilistic framework of stochastic processes, which is based on expected values [ensemble averages] instead of time averages, the probabilistic counterpart of Eq. (45) is known as the *Wiener-Khinchin relation* and, therefore, the probabilistic counterpart of Eq. (46) is called the *cyclic Wiener-Khinchin relation* [4, Chapter 12, Sec. 12.2].) Because of the relation Eq. (46), the SCD is also called the *cyclic spectral density* function [1, Chapter 10, Sec. B].

It follows from Eq. (46) and the interpretation Eq. (24) of $R_x^\alpha(\tau)$ as $R_{uv}(\tau)$ that the SCD is the Fourier transform of the crosscorrelation function $R_{uv}(\tau)$ and is therefore identical to the cross spectral density function for the frequency translates $u(t)$ and $v(t)$,

$$S_x^\alpha(f) \equiv S_{uv}(f), \tag{48}$$

where $S_{uv}(f)$ is defined by the right hand side of Eq. (44) for arbitrary $u(t)$ and $v(t)$. This is to be expected since the cross spectral density $S_{uv}(f)$ is known (cf. [1, Chapter 7, Sec. A]) to be the spectral-correlation density for spectral components in $u(t)$ and $v(t)$ at frequency f, and $u(t)$ and $v(t)$ are frequency-shifted versions of $x(t)$. The identity Eq. (48) suggests an appropriate normalization for $S_x^\alpha(f)$: as long as the PSDs of $u(t)$ and $v(t)$ contain no spectral lines at frequency f, which means that the PSD of $x(t)$ contains no spectral lines at either of the frequencies $f \pm \alpha/2$, then the correlation of the spectral components Eq. (48) is actually a covariance since the means of the spectral components are zero [1, Chapter 11, Sec. C]. When normalized by the geometric mean of the corresponding variances, which are given by

$$S_u(f) = S_x(f + \alpha/2) \tag{49}$$

and

$$S_v(f) = S_x(f - \alpha/2), \tag{50}$$

the covariance becomes a correlation coefficient:

$$\frac{S_{uv}(f)}{[S_u(f)S_v(f)]^{1/2}} = \frac{S_x^\alpha(f)}{[S_x(f+\alpha/2)S_x(f-\alpha/2)]^{1/2}} \triangleq \rho_x^\alpha(f). \tag{51}$$

Since $|\rho_x^\alpha(f)|$ is bounded to the interval $[0, 1]$, it is a convenient measure of the degree of local spectral redundancy that results from spectral correlation. For example, for $|\rho_x^\alpha(f)| = 1$, we have complete spectral redundancy at $f + \alpha/2$ and $f - \alpha/2$. For conjugate spectral redundancy of complex-valued signals, Eq. (51) is modified by replacing the numerator with the conjugate SCD. Let us now return to the AM example considered previously.

Example 1 continued: AM. By Fourier transforming Eq. (40) and invoking the cyclic Wiener relation Eq. (46), we obtain the following SCD function on the principal domain for the amplitude-modulated signal Eq. (5):

$$S_x^\alpha(f) = \begin{cases} \frac{1}{4}e^{\pm i2\theta}S_a(f), & \text{for } \alpha = \pm 2f_0 \\ \frac{1}{4}S_a(f+f_0) + \frac{1}{4}S_a(f-f_0), & \text{for } \alpha = 0 \\ 0, & \text{otherwise,} \end{cases} \tag{52}$$

where it has been assumed that $S_a(f \pm f_0) = 0$ for $|f| > 1/2$ to avc aliasing effects in the principal domain. The magnitude of this SCD is graphed in Figure 8 as the height of a surface above the bifrequency plane with coordinates f and α. For purposes of illustration, $a(t)$ is assumed to have an arbitrary low-pass PSD for this graph. Observe that although the argument f of the SCD is continuous, as it always will be for a random signal, the argument α is discrete, as it always will be since it represents the harmonic frequencies of periodicities underlying the random time-series (the sine-wave carrier in this example).

It follows from Eq. (52) that the spectral correlation coefficient is given by

$$\rho_x^\alpha(f) = \frac{S_a(f)e^{\pm i2\theta}}{\{[S_a(f+2f_0) + S_a(f)][S_a(f) + S_a(f-2f_0)]\}^{1/2}} \tag{53}$$

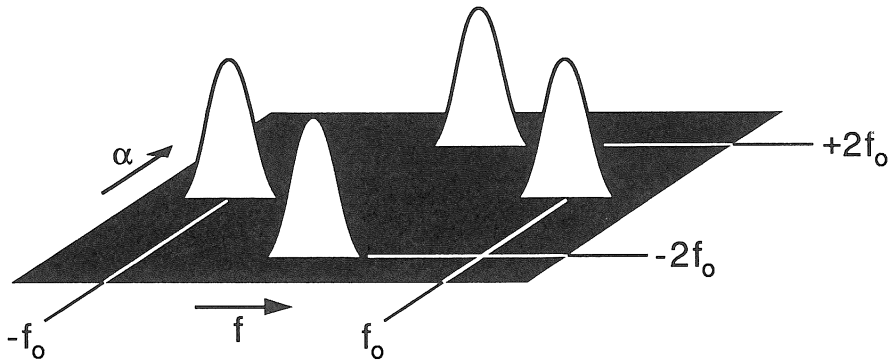

Figure 8: Magnitude of the spectral-correlation density function for an AM signal graphed as a height above the bifrequency plane with coordinates f and α.

for $\alpha = \pm 2f_0$. Thus, the strength of correlation between spectral component in $x(t)$ at frequencies $f + \alpha/2$ and $f - \alpha/2$ is unity:

$$|\rho_x^\alpha(f)| \text{ for } |f| < f_0 \text{ and } \alpha = \pm 2f_0 \tag{54}$$

provided that $a(t)$ is bandlimited to $|f| < f_0$,

$$S_a(f) = 0 \text{ for } |f| \geq f_0.$$

This is not surprising since the two spectral components in $x(t)$ at frequencies $f \pm \alpha/2 = f \pm f_0$ are obtained from the single spectral component in $a(t)$ at frequency f simply by shifting and scaling. Thus, they are perfectly correlated. That is, the upper (lower) sideband for $f > 0$ carries exactly the same information as the lower (upper) sideband for $f < 0$. Techniques for exploiting this spectral redundancy are described in Section III.

To illustrate the conjugate SCD, we consider the analytic signal $z(t)$ for AM:

$$\begin{aligned} S_{zz^*}^\alpha(f) &= \sum_{\tau=-\infty}^{\infty} R_{zz^*}^\alpha(\tau) e^{-i2\pi f\tau} \\ &= \begin{cases} \frac{1}{4} S_a(f) e^{i2\theta}, & \text{for } \alpha = 2f_0 \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Before considering other examples of the SCD, let us first gain an understanding of the effects of some basic signal processing operations on the SCD. This greatly facilitates the determination of the SCD for commonly encountered manmade signals.

D. FILTERING

When a signal $x(t)$ undergoes a linear time-invariant (LTI) transformation, (i.e., a convolution or a filtering operation),

$$\begin{aligned} z(t) &= h(t) \otimes x(t) \\ &\triangleq \sum_{u=-\infty}^{\infty} h(u)x(t-u), \end{aligned} \tag{55}$$

the spectral components in $x(t)$ are simply scaled by the complex-valued transfer function $H(f)$, which is the Fourier transform

$$H(f) = \sum_{t=-\infty}^{\infty} h(t)e^{-i2\pi ft} \tag{56}$$

of the impulse-response $h(t)$ of the transformation. As a result, the PSD gets scaled by the squared magnitude of $H(f)$ (see, for example, [1, Chapter 3, Sec. C] or [4, Chapter 10, Sec. 10.1] for continuous time)

$$S_z(f) = |H(f)|^2 S_x(f). \tag{57}$$

Equation (57) can be derived from the definition Eq. (43) of the PSD. Similarly, because the spectral components of $x(t)$ at frequencies $f \pm \alpha/2$ are scaled by $H(f \pm \alpha/2)$, the SCD gets scaled by the product $H(f + \alpha/2)H*(f - \alpha/2)$:

$$S_z^\alpha(f) = H(f + \alpha/2)H^*(f - \alpha/2)S_x^\alpha(f). \tag{58}$$

This result, called the *input-output SCD relation for filtering*, which can be derived from the definition Eq. (44) of the SCD, includes Eq. (57) as the special case of $\alpha = 0$. Observe that it follows from Eq. (58) and the definition Eq. (51) that

$$|\rho_z^\alpha(f)| \equiv |\rho_x^\alpha(f)| \tag{59}$$

That is, the magnitude of the spectral correlation coefficient is unaffected by filtering (provided that $H(f \pm \alpha/2) \neq 0$).

Example 3: Time Delay. As our first example of Eq. (58), we consider a filter that simply delays the input by some integer t_0; then $h(t) = \delta(t - t_0)$, where δ is the Kronecker delta, and $H(f) = e^{-i2\pi f t_0}$. Therefore, for $z(t) = x(t - t_0)$, we obtain from the input-output SCD relation Eq. (58)

$$S_z^\alpha(f) = S_x^\alpha(f) e^{-i2\pi\alpha t_0}, \tag{60}$$

which indicates that, unlike the PSD, the SCD of a cyclostationary signal is sensitive to the timing or phase of the signal. □

Example 4: Multipath Propagation. As a second example of Eq. (58), consider a communication signal $x(t)$ that undergoes multipath distortion during transmission to yield a received signal that is modeled by

$$z(t) = \sum_n a_n x(t - t_n),$$

where a_n and the integer t_n are the attenuation factor and delay of the nth propagation path. The transfer function corresponding to the multipath channel is given by

$$H(f) = \sum_n a_n e^{-i2\pi f t_n} \tag{61}$$

and therefore Eq. (58) yields

$$S_z^\alpha(f) = S_x^\alpha(f) \sum_{n,m} a_n a_m^* e^{-i2\pi[f(t_n - t_m) + \alpha(t_n + t_m)/2]}. \tag{62}$$

Example 5: Bandpass Signals. As a third example of the utility of the relation Eq. (58), let us determine the support region in the (f, α) plane for a bandpass signal with lowest frequency b and highest frequency B. To enforce such a spectrum, we can simply put any signal $x(t)$ through an ideal bandpass filter with transfer function (on the principal domain $(-1/2, 1/2]$)

$$H(f) = \begin{cases} 1, & \text{for } b < |f| < B \\ 0, & \text{otherwise.} \end{cases}$$

It then follows directly from the input-output SCD relation Eq. (58) that the SCD for the output of this filter can be nonzero only for $f > b$ and

$f < B$:

$$S_z^\alpha(f) = \begin{cases} 0, & \text{for } ||f| - |\alpha|/2| \le b \text{ or } |f| + |\alpha|/2 \ge B \\ S_x^\alpha(f), & \text{otherwise.} \end{cases} \tag{63}$$

This shows that the support region in the (f, α) plane for a bandpass signal is the four diamonds located at the vertices of a larger diamond, depicted in Figure 9a. By letting $b \to 0$, we obtain the support region for a lowpass signal, and by letting $B \to 1/2$, we obtain the support region for a highpass signal. This is shown in Figures 9b and 9c. □

E. SIGNAL MULTIPLICATION AND TIME SAMPLING

When finite segments of two discrete-time signals are multiplied together, we know from the convolution theorem that their Fourier transforms get circularly convolved. From this, we expect some sort of convolution relation to hold for the SCDs of signals passing through a product modulator. In fact, it can be shown (cf. [1, Chapter 11, Sec. C] or [4, Chapter 12, exc. 41] for continuous time) that if $x(t)$ is obtained by multiplying together two statistically independent[7] time-series $r(t)$ and $s(t)$,

$$x(t) = r(t)s(t), \tag{64}$$

then the cyclic autocorrelation of $x(t)$ is given by the discrete circular convolution in cycle frequency of the cyclic autocorrelations of $r(t)$ and $s(t)$:

$$R_x^\alpha(\tau) = \sum_{\beta \in (-1/2, 1/2]} R_r^\beta(\tau) R_s^{\alpha-\beta}(\tau), \tag{65}$$

where, for each α, β ranges over all values in the principal domain $(-1/2, 1/2]$ for which $R_r^\beta(\tau) \not\equiv 0$. By Fourier transforming Eq. (65), we obtain the input-output SCD relation for signal multiplication:

$$S_x^\alpha(f) = \int_{-1/2}^{1/2} \sum_{\beta \in (-1/2, 1/2]} S_r^\beta(\tau) S_s^{\alpha-\beta}(f - \nu)\, d\nu \tag{66}$$

[7] Time-series are statistically independent if their joint fraction-of-time probability densities factor into products of individual fraction-of-time probability densities, as explained in [1, Chapter 15, Sec. A].

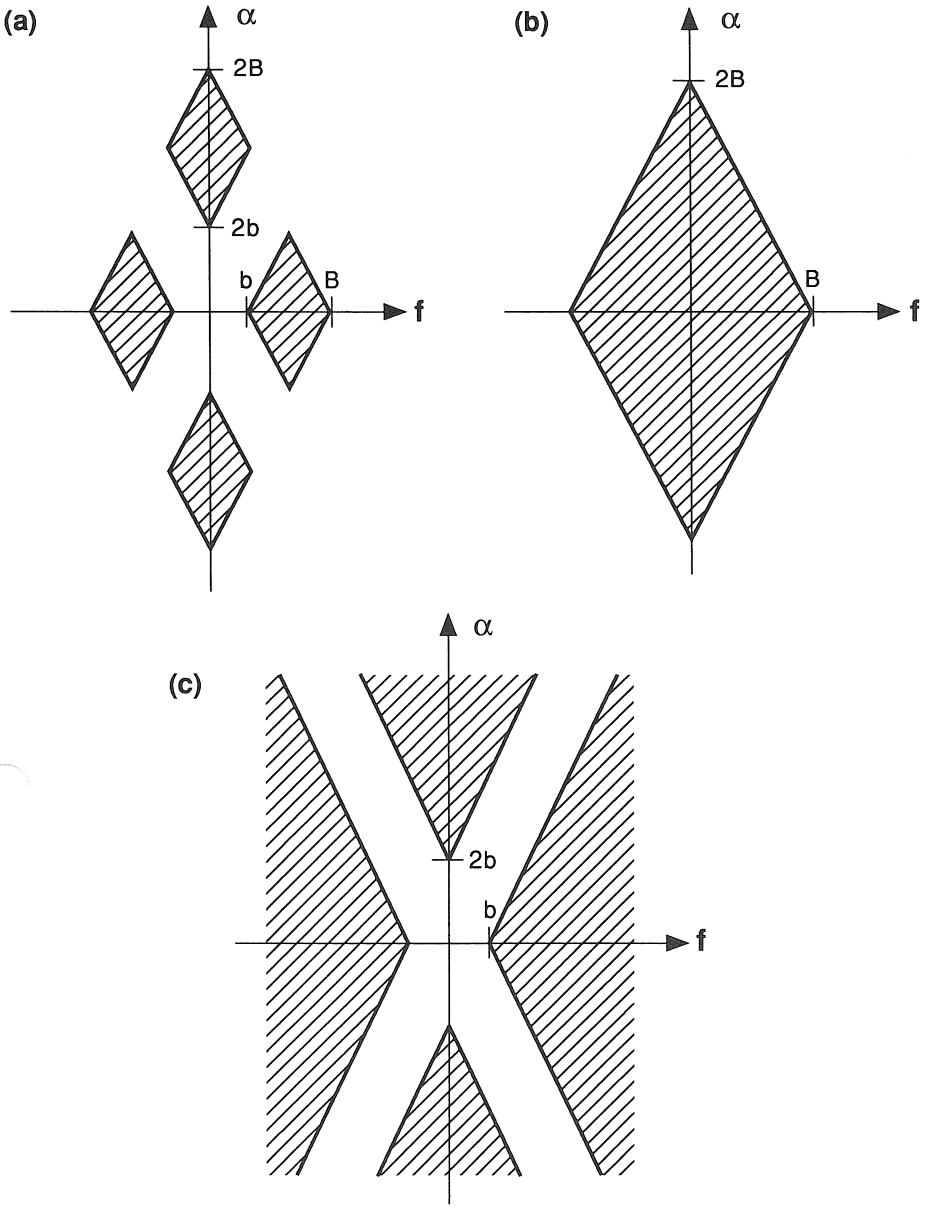

Figure 9: (a) Support region in the bifrequency plane for the spectral-correlation density function of a bandpass signal. (b) Support region for a lowpass signal. (c) Support region for a highpass signal (shown over a small fraction of the diamond-shaped principal domain).

which is a double circular convolution that is continuous in the variable f and discrete in the variable α.

Example 6: Frequency Conversion. As an example of Eq. (66), if $s(t)$ is simply a sinusoid,

$$s(t) = \cos(2\pi f_0 t + \theta),$$

the product modulator becomes a frequency converter when followed by a filter to select either the up-converted version or the down-converted version of $s(t)$. By applying first the input-output SCD relation Eq. (66) for the product modulator (which applies since a sinusoid is statistically independent of all time-series [1, Chapter 15, Sec. A]), and then Eq. (58) for the filter, we can determine the up-converted or down-converted SCD. To illustrate, we first determine the SCD for the sinusoid $s(t)$. By substituting the sinusoid $s(t)$ into the definition of the cyclic autocorrelation, we obtain

$$R_s^\alpha(\tau) = \begin{cases} \frac{1}{2}\cos(2\pi f_0 \tau), & \text{for } \alpha = 0 \\ \frac{1}{4}e^{\pm i2\theta}, & \text{for } \alpha = \pm 2f_0 \\ 0, & \text{otherwise} \end{cases} \tag{67}$$

on the principal domain of α. Fourier transforming then yields the SCD

$$S_s^\alpha(f) = \begin{cases} \frac{1}{4}\delta(f - f_0) + \frac{1}{4}\delta(f + f_0), & \text{for } \alpha = 0 \\ \frac{1}{4}e^{\pm i2\theta}\delta(f), & \text{for } \alpha = \pm 2f_0 \\ 0, & \text{otherwise} \end{cases} \tag{68}$$

on the principal domain of f and α, which is illustrated in Figure 10a. Using Eq. (66), we circularly convolve this SCD with that of a stationary signal $r(t)$, for which

$$S_r^\alpha(f) = \begin{cases} S_r(f), & \text{for } \alpha = 0 \\ 0, & \text{for } \alpha \neq 0 \end{cases} \tag{69}$$

on the principal domain (see Figure 10b). The result is that the SCD of the stationary signal simply gets replicated and scaled at the four locations of the impulses in the SCD of the sinusoid, as illustrated in Figure 10c (provided that $S_r(f \pm f_0) = 0$ for $|f| > 1/2$ to avoid aliasing effects in the principal domain).

Example 7: Time Sampling. Another important signal processing operation is periodic time sampling. It is known that for a stationary signal $x(t)$, the PSD $S_x(f)$ of the sequence of samples $\{x(nT_s) : n = 0, \pm 1, \pm 2, \ldots\}$ is related to the PSD $\tilde{S}_x(f)$ of the continuous-time waveform by the aliasing formula (cf. [1, Chapter 3, Sec. E] or [4, Chapter 11, Sec. 11.1])

$$S_x(f) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} \tilde{S}_x(f - n/T_s). \tag{70}$$

It is shown in [1, Chapter 11, Sec. C], [4, Chapter 12, Sec. 12.4] that this aliasing formula generalizes for the SCD to

$$S_x^\alpha(f) = \frac{1}{T_s} \sum_{m,n=-\infty}^{\infty} \tilde{S}_x^{\alpha+m/T_s}(f - m/2T_s - n/T_s). \tag{71}$$

Observe that, when $x(t)$ is not stationary (i.e., when $S_x^\alpha(f) \not\equiv 0$ for $\alpha = m/T$ for some nonzero integers m), the conventional PSD aliasing formula Eq. (70) must be corrected according to Eq. (71) evaluated at $\alpha = 0$:

$$S_x(f) = \frac{1}{T_s} \sum_{m,n=-\infty}^{\infty} \tilde{S}_x^{m/T_s}(f - m/2T_s - n/T_s). \tag{72}$$

This reflects the fact that, when aliased overlapping spectral components add together, their PSD values add only if they are uncorrelated. When they are correlated, as in a cyclostationary signal, the PSD value of the sum of overlapping aliased components depends on the particular magnitudes and phases of their correlations. The SCD aliasing formula Eq. (71) is illustrated graphically in Figure 11, where the support regions for the SCD $S_x^\alpha(f)$ for the sequence of samples $\{x(nT_s)\}$ is depicted in terms of the single diamond support region for a lowpass waveform $x(t)$, which is shown in Figure 9b.

When we subsample a discrete-time signal $x(t)$ with sampling rate $1/T_s$ for some integer T_s to obtain the signal $z(t)$, we obtain the discrete-time analog of Eq. (71)

$$S_z^\alpha(f) = \frac{1}{T_s} \sum_{q \in P_\alpha} S_x^{\alpha+q/T_s}\left(\frac{f + q/2}{T_s}\right), \tag{73}$$

where P_α is the set of all integers $q = \beta T_s - \alpha$ for which $\alpha \in (-1/2, 1/2]$. Similarly, when we resample a discrete-time signal $x(t)$, by (effectively)

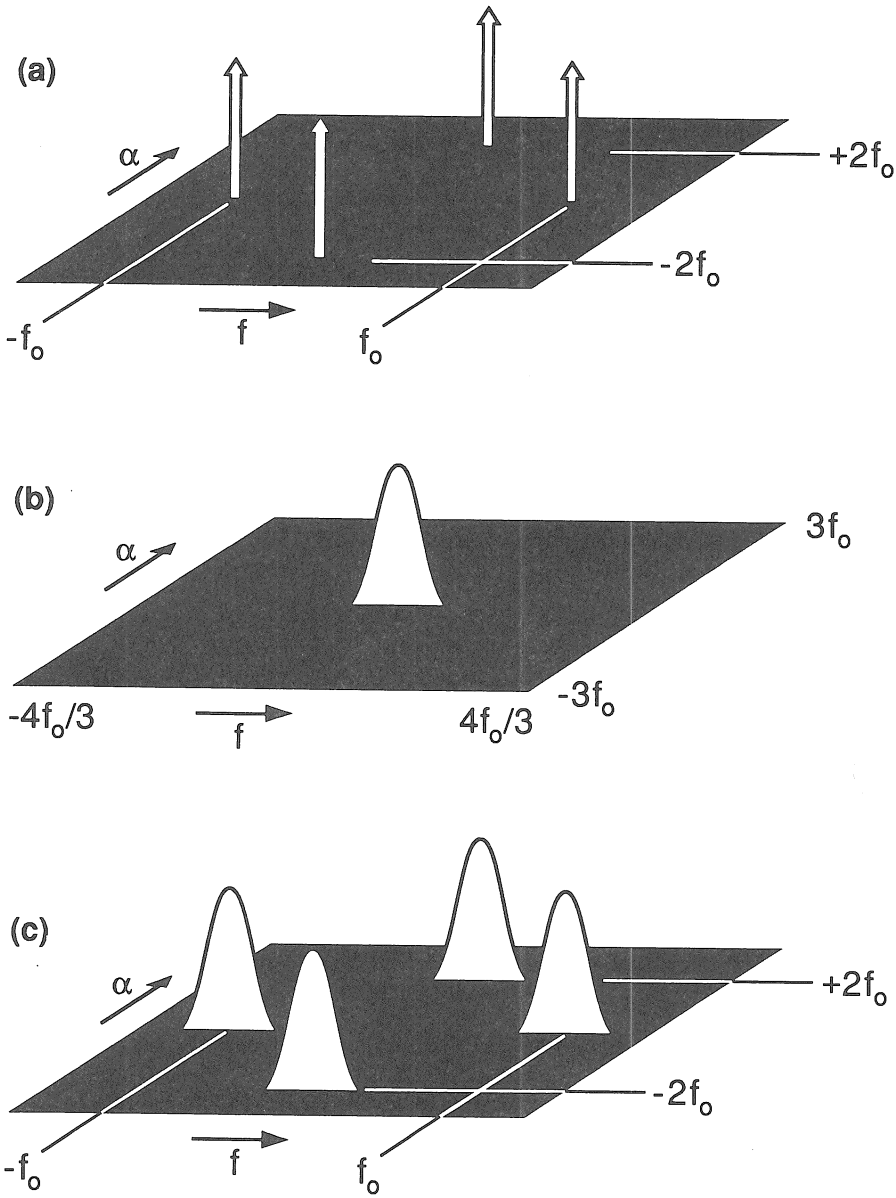

Figure 10: (a) Magnitude of the spectral correlation density (SCD) for a sine wave of frequency f_0. (b) SCD for a lowpass stationary signal. (c) SCD magnitude for the product of signals corresponding to (a) and (b), obtained by convolving the SCDs in (a) and (b).

nterpolating back to a continuous-time waveform and then time sampling at the new rate $1/T_s$ to obtain $z(t)$, the SCD is given by

$$S_z^\alpha(f) = \sum_{n,m=-\infty}^{\infty} \bar{S}_x^{\alpha+m/T_s}(f - m/2T_s - n/T_s), \tag{74}$$

where $\bar{S}_x^\alpha(f)$ is the SCD $S_x^\alpha(f)$ restricted to its principal domain. □

F. PERIODICALLY TIME-VARIANT FILTERING

Many signal processing devices such as pulse and carrier modulators, multiplexors, samplers, and scanners, can be modeled as periodically time-variant filters, especially if multiple incommensurate periodicities are included in the model. By expanding the periodically time-variant impulse response in a Fourier series as explained shortly, any such system can be represented by a parallel bank of sinusoidal product modulators followed by time-invariant filters. Consequently, the effect of any such system on the SCD of its input can be determined by using the SCD relations for filters and product modulators. In particular, it can be shown (cf. [1, Chapter 11, Sec. D] for continuous time) that the SCD of the output $z(t)$ of a multiply-periodic system with input $x(t)$ is given by

$$S_z^\alpha(f) = \sum_{\beta,\gamma \in A} G_\beta(f + \alpha/2) G_\gamma^*(f - \alpha/2) S_x^{\alpha-\beta+\gamma}\left(f - \frac{\beta - \gamma}{2}\right), \tag{75}$$

where A is the set of sinusoid frequencies associated with the product modulators in the system representation, $S_x(f + \beta) = 0$ for $|f| > 1/2$ for all $\beta \in A$ to avoid aliasing effects in the principal domain, and $G_\beta(f)$ are the transfer functions of the filters. More specifically, for the input-output equation

$$z(t) = \sum_{u=-\infty}^{\infty} h(t,u)x(u), \tag{76}$$

the multiply-periodic impulse response $h(t,u)$ can be expanded in the Fourier series

$$h(t+\tau,t) = \sum_{\beta \in A} g_\beta(\tau) e^{i2\pi\beta t}, \tag{77}$$

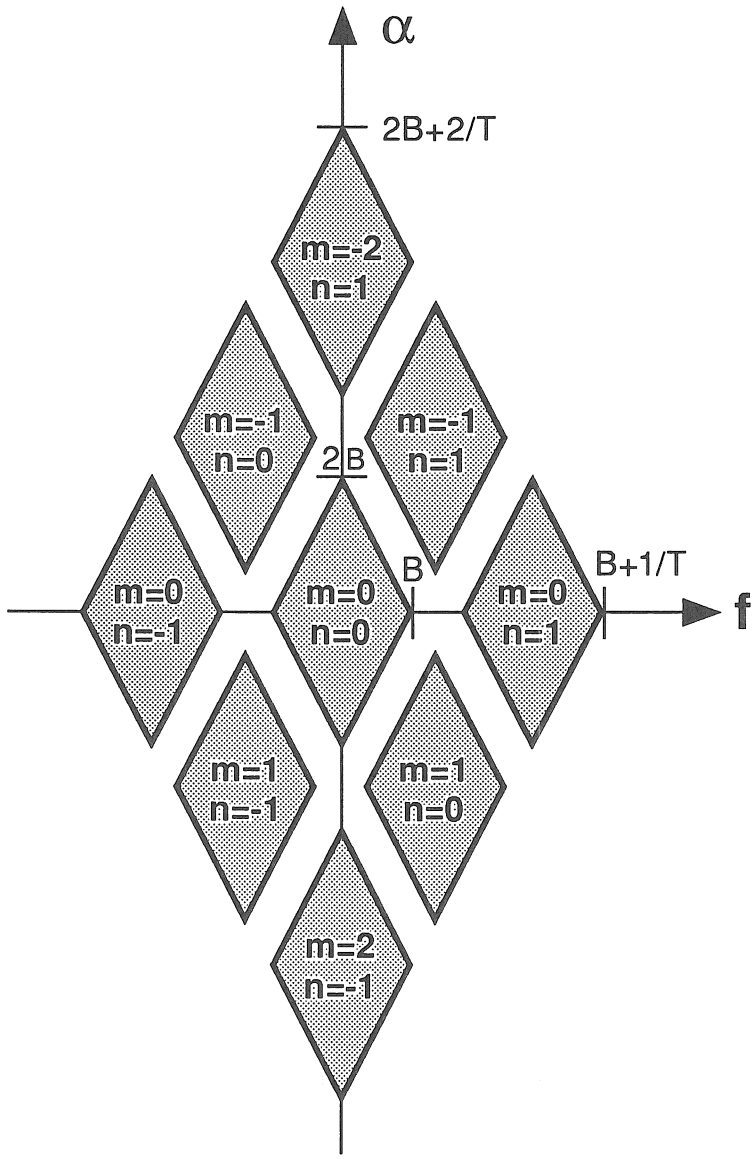

Figure 11: Illustration of support regions in the bifrequency plane for the spectral-correlation densities that are aliased by periodic time-sampling.

where the Fourier coefficients (for each τ) are given by

$$g_\beta(\tau) = \left\langle h(t+\tau, t) e^{-i2\pi\beta t} \right\rangle . \tag{78}$$

It follows from Eqs. (76) and (77) that the filter output can be expressed as

$$z(t) = \sum_{\beta \in A} \left[x(t) e^{i2\pi\beta t} \right] \otimes g_\beta(t), \tag{79}$$

where $g_\beta(t)$ are the impulse responses of the filters with corresponding transfer functions $G_\beta(f)$. Thus, periodically time-variant filters perform time-invariant filtering on frequency-shifted versions of the input, such as $x(t)e^{i2\pi\beta t}$. This results in summing scaled, frequency-shifted, and cycle-frequency-shifted versions of the SCD for the input $x(t)$ to obtain the SCD for the output $z(t)$, as indicated in Eq. (75).

Let us now consider some additional examples of modulation types, making use of the results obtained in the preceding paragraphs to determine SCDs. However, in the interest of realism and for the sake of analytical simplicity, continuous-time signal models are used.

Example 2 continued: PAM. Let $\{a_n\}$ be a stationary random sequence, and let us interpret these random variables as the time-samples of a continuous-time random waveform, $a_n = a(nT_0)$, with PSD $\tilde{S}_a(f)$. We onsider the continuous-time PAM signal

$$x(t) = \sum_{n=-\infty}^{\infty} a_n p(t - nT_0 + \epsilon), \tag{80}$$

where $p(t)$ is a deterministic finite-energy pulse and ϵ is a fixed pulse-timing phase parameter. To determine the SCD of $x(t)$, we can recognize that $x(t)$ is the output of a periodically time-variant linear system with input $a(t)$, and impulse response

$$h(t,u) = \sum_{n=-\infty}^{\infty} p(t - nT_0 + \epsilon)\delta(u - nT_0), \tag{81}$$

where δ is the Dirac δ. We can then use the continuous-time counterpart of the input-output SCD relation Eq. (75), which is identical in form except that continuous-time Fourier transforms are used (cf. [1, Chapt.

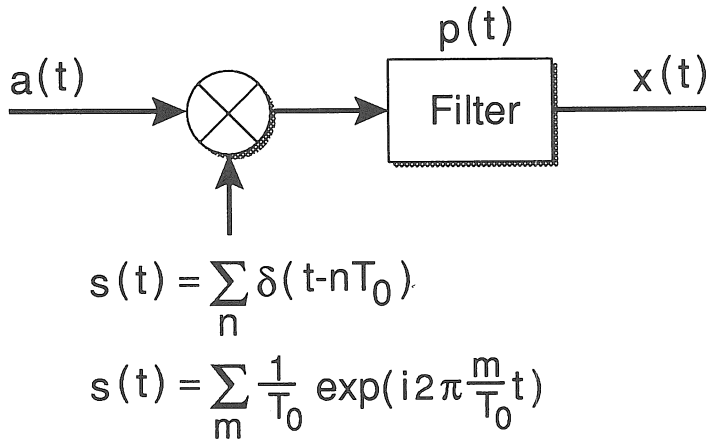

Figure 12: Interpretation of PAM signal generator as the cascade of an impulse sampler and a pulse-shaping filter.

11, Sec. D]). Or we can recognize that this particular periodically time-variant system is composed of a product modulator that implements an impulse sampler, followed by a linear time-invariant pulse-shaping filter with impulse-response function $h(t) = p(t)$, as shown in Figure 12. We can then use the continuous-time counterpart of the input-output SCD relation Eq. (66), which is identical in form except the convolutions are linear (cf. [1, Chapt. 11, Sec. C]), as it applies to impulse sampling, together wit the relation Eq. (58) for filtering. The result is

$$\tilde{S}_x^\alpha(f) = \frac{1}{T_0^2}\tilde{P}(f+\alpha/2)\tilde{P}^*(f-\alpha/2) \sum_{n,m=-\infty}^{\infty} \tilde{S}_a^{\alpha+m/T_0}(f - m/2T_0 - n/T_0)\, e^{i2\pi\alpha\epsilon}. \quad (82)$$

Using the SCD aliasing formula Eq. (71) for $a(t)$ we can re-express Eq. (82) as

$$\tilde{S}_x^\alpha(f) = \frac{1}{T_0}\tilde{P}(f+\alpha/2)\tilde{P}^*(f-\alpha/2)S_a^\alpha(f)e^{i2\pi\alpha\epsilon}, \quad (83)$$

where $S_a^\alpha(f)$ is the SCD for the pulse-amplitude sequence $\{a_n\}$. Having assumed that $\{a_n\}$ is stationary, and using the periodicity property

$$S_a^\alpha(f) = \begin{cases} S_a(f+\alpha/2), & \text{for } \alpha = k/T_0 \\ 0, & \text{otherwise,} \end{cases} \quad (84)$$

for $k = 0, \pm 1, \pm 2, \ldots$, we can express Eq. (83) as

$$\tilde{S}_x^\alpha(f) = \begin{cases} \frac{1}{T_0}\tilde{P}(f+\alpha/2)\tilde{P}^*(f-\alpha/2)S_a(f+\alpha/2)e^{i2\pi\alpha\epsilon}, & \text{for } \alpha = k/T_0 \\ 0, & \text{otherwise.} \end{cases} \tag{85}$$

A graph of the magnitude of this SCD for the full-duty-cycle rectangular pulse

$$p(t) = \begin{cases} 1, & \text{for } |t| \le T_0/2 \\ 0, & \text{otherwise,} \end{cases} \tag{86}$$

and a white-noise amplitude sequence with PSD

$$S_a(f) = 1 \tag{87}$$

is shown in Figure 13.

It follows from Eq. (87) that for all $\alpha = k/T_0$ for which $S_a(f \pm \alpha/2) \neq 0$ and $\tilde{P}(f+\alpha/2)\tilde{P}^*(f-\alpha/2) \neq 0$, the spectral correlation coefficient $\rho_x^\alpha(f)$ is unity in magnitude:

$$|\rho_x^\alpha(f)| = 1. \tag{88}$$

Thus, all spectral components outside the band $|f| < 1/2T_0$ are completely redundant with respect to those inside this band. Techniques for exploiting this spectral redundancy are described in Section III.

The conjugate SCD for the PAM signal Eq. (11) is given by Eq. (85) with $\tilde{P}^*(f-\alpha/2)$ replaced by $\tilde{P}(\alpha/2-f)$ and $S_a^\alpha(f)$ replaced by $S_{aa^*}^\alpha(f)$. For a real PAM signal, the conjugate SCD is identical to the SCD; however, for complex PAM the conjugate SCD is, in general, different and is, in fact, zero for the complex PAM that models the complex envelopes of most digital QAM signals, including QPSK. This follows from the fact that $\langle a_n a_{n+m}\rangle = 0$ for all m in such signals; consequently, $S_{aa^*}^\alpha(f) = 0$ for all α.

By inverse Fourier transforming the SCD Eq. (82), we obtain the cyclic autocorrelation function

$$\tilde{R}_x^\alpha(\tau) = \begin{cases} \frac{1}{T_0}\sum_{n=-\infty}^{\infty} \tilde{R}_a(nT_0)r_p^\alpha(\tau - nT_0)e^{i2\pi\alpha\epsilon}, & \text{for } \alpha = k/T_0 \\ 0, & \text{otherwise,} \end{cases} \tag{89}$$

where

$$r_p^\alpha(\tau) \triangleq \int_{-\infty}^{\infty} p(t+\tau/2)p^*(t-\tau/2)e^{-i2\pi\alpha t}\,dt. \tag{90}$$

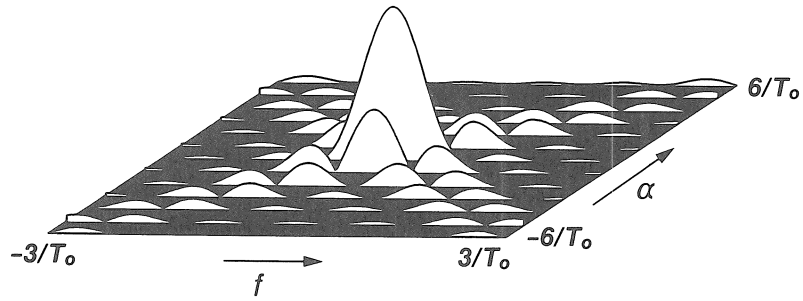

Figure 13: Magnitude of the spectral-correlation density for a PAM signal with full-duty-cycle rectangular pulses.

For a white-noise amplitude sequence as in Eq. (87), Eq. (89) reduces to

$$\tilde{R}_x^\alpha(\tau) = \frac{1}{T_0} r_p^\alpha(\tau) e^{i2\pi\alpha\epsilon} \text{ for } \alpha = k/T_0, \tag{91}$$

and, for a rectangular pulse as in Eq. (86), this yields the temporal correlation coefficient

$$\tilde{\gamma}_x^\alpha(\tau) = \frac{\sin\left[\pi\alpha(T_0 - |\tau|)\right]}{\pi\alpha T_0} e^{i2\pi\alpha\epsilon} \quad \text{for } |\tau| \leq T_0, \tag{92}$$

which peaks for $\alpha = 1/T_0$ at $t = T_0/2$, where it takes on the value

$$|\tilde{\gamma}_x^\alpha(T_0/2)| = 1/\pi \quad \text{for } \alpha = 1/T_0. \tag{93}$$

That is, the strongest possible spectral line that can be generated in a delay-product signal (cf. remark made following Eq. (21)) for this particular PAM signal occurs when the delay equals half the pulse period. In contrast to this, when the more bandwidth-efficient pulse whose transform is a raised cosine is used, the optimal delay for sine-wave generation is zero.

An especially simple example of a sequence of pulse amplitudes $\{a_n\}$ is a binary sequence with values ± 1. If we consider $\tau = 0$ in the delay-product signal, then we obtain

$$y_0(t) = |x(t)|^2 = \sum_{n,m=-\infty}^{\infty} a_n a_m p(t - nT_0 + \epsilon) p(t - mT_0 + \epsilon).$$

If the pulses do not overlap (i.e., if $p(t) = 0$ for $|t| > T_0/2$), this reduces to

$$\begin{aligned} y_0(t) &= \sum_{n=-\infty}^{\infty} a_n^2 p^2(t - nT_0 + \epsilon) \\ &= \sum_{n=-\infty}^{\infty} p^2(t - nT_0 + \epsilon), \end{aligned}$$

which is periodic with period T_0 and therefore contains finite-strength additive sine-wave components with frequencies k/T_0 (except when $p(t)$ is flat as in Eq. (86)). In this very special case where $\{a_n\}$ is binary and the pulses do not overlap, there is no random component in $y_0(t)$; but, for $\tau \neq 0$, $y_\tau(t)$ contains both sine-wave components and random components (even when $p(t)$ is flat). □

Example 8: ASK and PSK. By combining the amplitude-modulated sine wave and the digital amplitude-modulated pulse train, we obtain the amplitude-shift-keyed (ASK) signal

$$x(t) = a(t)\cos(2\pi f_0 t + \theta), \tag{94}$$

where

$$a(t) = \sum_{n=-\infty}^{\infty} a_n p(t - nT_0 + \epsilon), \tag{95}$$

and $\{a_n\}$ are digital amplitudes. By using the continuous-time counterpart of the SCD relation Eq. (66) for signal multiplication and the result Eq. (85) for the SCD of $a(t)$, we can obtain the SCD for the signal Eq. (94) by simply convolving the SCD functions shown in Figures 10a and 13. The result is shown in Figure 14a, where the cycle frequencies shown are $\alpha = \pm 2f_0 + m/T_0$ and $\alpha = m/T_0$ for integers m, and where $f_0 = 3.3/T_0$. When $f_0 T_0$ is irrational, the ASK signal is polycyclostationary with fundamental periods T_0 and $1/2f_0$.

For a binary sequence with each $a_n = \pm 1$, this amplitude-shift keyed signal, with the pulse Eq. (86), is identical to the binary phase-shift keyed (BPSK) signal

$$x(t) = \cos\left[2\pi f_0 t + \theta + \sum_{n=-\infty}^{\infty} \phi_n p(t - nT_0)\right], \tag{96}$$

where $\phi_n \triangleq (a_n - 1)/2$, since shifting the phase of a sine wave by 0 or π is the same as multiplying its amplitude by 1 or -1. Other commonly used types of phase-shift-keyed signals include quaternary phase-shift keying (QPSK) and staggered QPSK (SQPSK). The details of these signal types are available in the literature (see, for example [1, Chapter 12, Sec. E] or [4, Chapter 12, Sec. 12.5]). Only their SCD-magnitude surfaces are shown here in Figure 14b, c, where again $f_0 = 3.3/T_0$.

It is emphasized that the three signals BPSK, QPSK, and SQPSK differ only in their carrier phase shifts and pulse timing and, as a result, they have identical PSDs, as shown in Figure 14 (consider $\alpha = 0$). However, as also shown in Figure 14, these differences in phase and timing result in substantially different SCDs (consider $\alpha \neq 0$). That is, the quadrature component present in QPSK but absent in BPSK results in cancellation of the SCD at cycle frequencies associated with the carrier frequency (viz., $\alpha = \pm 2f_0 + m/T_0$ for all integers m) in QPSK. Similarly, the pulse staggering by $T_0/2$ (between the in-phase and quadrature components) present in SQPSK but absent in QPSK results in the SCDs being cancelled at $\alpha = \pm 2f_0 + m/T_0$ only for even integers m, and at $\alpha = m/T_0$ only for odd integers m in SQPSK. This again illustrates the fact that the SCD contains phase and timing information not available in the PSD. In fact, as formulas Eq. (52) and Eq. (85) reveal, the carrier phase θ in Eq. (5) and the pulse timing ϵ in Eq. (11) are contained explicitly in the SCDs for these carrier- and pulse-modulated signals. □

G. MEASUREMENT OF SPECTRAL CORRELATION

The SCD function Eq. (44) is derived by idealizing the practical spectral correlation measurement depicted in Figure 6 by letting the averaging time T in the correlation measurement approach infinity and then letting the spectral resolving bandwidth B approach zero. Consequently, the practical measurement with finite parameters T and B can be interpreted as an estimate of the ideal SCD. This estimate will be statistically reliable only if $TB >> 1$, and it will approach the ideal SCD only for sufficiently large T and sufficiently small B. Numerous alternative methods for making

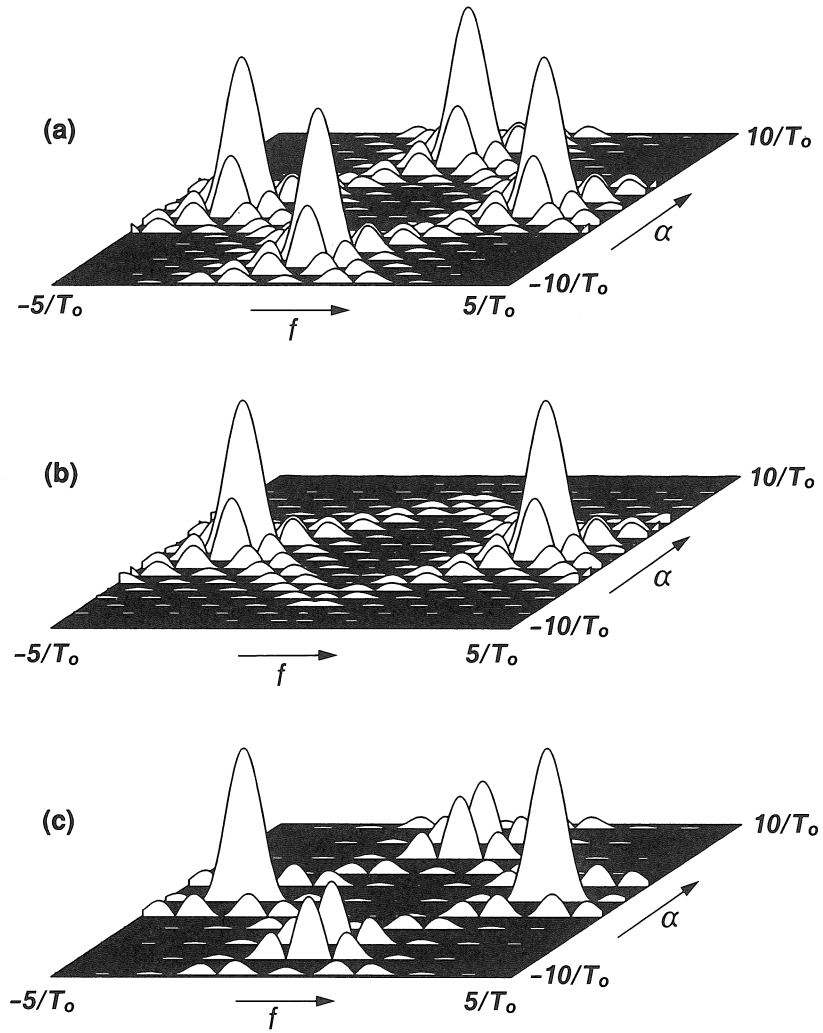

Figure 14: Magnitude of spectral-correlation densities. (a) BPSK, (b) QPSK, and (c) SQPSK. (Each signal has a rectangular keying envelope.)

this practical measurement are described in [1, Chapter 13], and computationally efficient digital algorithms and architectures for some of these are developed in [5–7]. The statistical behavior (bias and variance) of such estimates is analyzed in detail in [1, Chapter 15, Sec. B], [6,8–11]. For the purpose of making the applications described in Section III more concrete, it suffices here to simply point out that because the SCD $S_x^\alpha(f)$ is equivalent to a particular case of the conventional cross spectral density $S_{uv}(f)$ (cf. Eq. (48)), one can envision any of the conventional methods of cross spectral analysis as being used in the applications. For example, the Wiener-Daniell method [1, Chapter 7, Sec. D], based on frequency smoothing the crossperiodogram of $u(t)$ and $v(t)$ (the conjugate product of their FFTs) can be used. This is equivalent to multiplying the FFT bin for the data $x(t)$ corresponding to the frequency $f + \alpha/2$ times the conjugate of the FFT bin corresponding to the frequency $f - \alpha/2$ to obtain what is called the cyclic periodogram for $x(t)$, and then averaging in the parameter f over a band of width B for each fixed α of interest. For an FFT length of T, this yields a cycle resolution (resolution in α) of $1/T$ and a spectral resolution (resolution in f) of B.

Example 9: QPSK. As an example, the result of using the Wiener-Daniell method is illustrated in Figure 15 for a QPSK signal with carrier frequency $f_0 = 1/4T_s$ and keying rate $1/T_0 = 1/8T_s$, where $1/T_s$ is the sampling rate. An FFT of length 128 ($T = 128T_s$) was used in Figure 15a, and only four frequency bins were averaged together ($B = 4/T = 1/32T_s$) to produce each output point, whereas, in Figure 15b, the FFT length used was 32,768 ($T = 32,768T_s$) and 1,024 bins were averaged together ($B = 1,024/T = 1/32T_s$). It is easily seen by comparing with the ideal SCD in Figure 14b that without adequate spectral smoothing the variability of the SCD estimate can be very large. □

III. APPLICATIONS OF CYCLOSTATIONARITY

A. SPECTRAL REDUNDANCY

The existence of correlation between widely separated spectral components

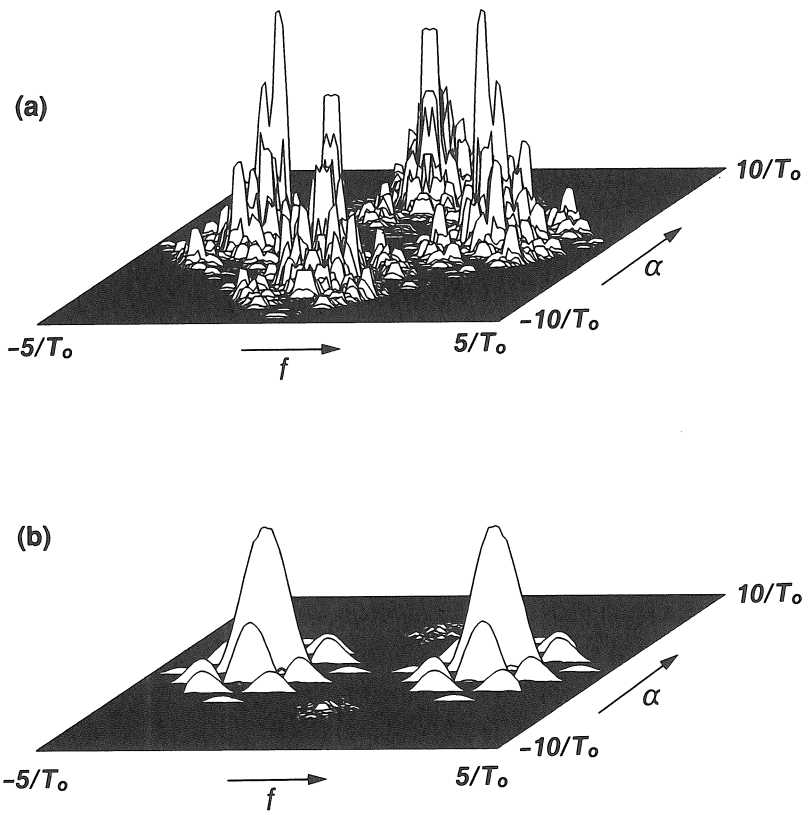

Figure 15: Magnitude of a spectral-correlation density (SCD) estimate obtained from a finite-length data record for the QPSK signal whose ideal SCD is shown in Figure 14b. (a) Record length is 128 time samples, and four adjacent frequency (f) bins are averaged together. (b) Record length is 32,768 and 1,024 adjacent frequency (f) bins are averaged together. (The sampling rate in both (a) and (b) is $10/T_0$, where $1/T_0$ is the keying rate of the QPSK signal.)

(separation equal to a cycle frequency α) can be interpreted as spectral redundancy. The meaning of the term *redundancy* that is intended here is essentially the same as that used in the field of information theory and coding. Specifically, multiple randomly fluctuating quantities (random variables) exhibit some redundancy if they are statistically dependent, for example, correlated. In coding, undesired redundancy is removed from data to increase the efficiency with which it represents information, and redundancy is introduced in a controlled manner to increase the reliability of storage and transmission of information in the presence of noise by enabling error detection and correction.

Here, redundancy that is inadvertently introduced into signals by the modulation process is to be exploited to enhance the accuracy and reliability of information gleaned from the measurements of corrupted signals, but the term *information* is interpreted in a broad sense. For instance, it includes the eight examples outlined in Section I. In all of these examples, the performance of the signal processors that make the decisions and/or produce the estimates can be substantially improved by suitably exploiting spectral redundancy. The degree of improvement relative to the performance of more commonly used signal processors that ignore spectral redundancy depends on both the severity of the signal corruption (noise, interference, distortion) and the degree of redundancy in the signal x(t), as measured by the magnitude of the spectral correlation coefficient (or its conjugate coun terpart) defined in Section II. The degree of improvement also depends on the amount of data available for processing (the collection time). The utility of exploiting spectral redundancy can also be enhanced by intentionally designing the signal to exhibit a sufficient amount of spectral redundancy.

The primary feature of spectral redundancy that enables it to be readily exploited is that it is typically a unique signature for each unique signal type. That is, most manmade signals exhibit spectral redundancy, but most noise (all noise that is not cyclostationary) does not and, more importantly, when multiple signals of interest and signals of no interest (interference) overlap in both time and frequency, the spectral redundancy functions (SCDs) of these signals are nonoverlapping in the cycle-frequency domain: their cycle frequencies α are distinct. This results from signals

aving distinct carrier frequencies and/or pulse rates or keying rates, even when occupying the same spectral band.

The uniqueness of the spectral redundancy signature makes signal-selective measurements possible. Specifically, for the received composite signal

$$x(t) = \sum_{\ell=1}^{L} s_\ell(t) + n(t), \tag{97}$$

where the set $\{s_\ell(t)\}_{\ell=1}^{L}$ includes both signals of interest and interference—all of which are statistically independent of each other—and where $n(t)$ is background noise, we have the SCD (for measurement time $T \to \infty$)

$$S_x^\alpha(f) = \sum_{\ell=1}^{L} S_{s_\ell}^\alpha(f) + S_n^\alpha(f). \tag{98}$$

But if the only signal with the particular cycle frequency α_k is $s_k(t)$, then (for $T \to \infty$) we have

$$S_x^{\alpha_k}(f) = S_{s_k}^{\alpha_k}(f), \tag{99}$$

regardless of the temporal or spectral overlap among $\{s_\ell(t)\}_{\ell=1}^{L}$ and also $n(t)$. This perfect signal-selectivity of ideal SCDs implies that practical measurements of SCDs or their parameters can be made signal-selective for measurement times T that are long enough.

Example 1: BPSK Signal in Multiple AM Interference and Noise. To illustrate the concept of signal selectivity, let us consider the situation in which a broadband BPSK signal of interest is received in the presence of white noise and five interfering AM signals with narrower bandwidths that together cover the entire band of the BPSK signal. The noise and each of the five interfering signals have equal average power. Therefore, the total signal-to-interference-and-noise ratio (SINR) is approximately -8 dB. The BPSK signal has carrier frequency $f_0 = 0.25/T_s$ and keying rate $\alpha_0 = 0.0625/T_s$. It has full-duty-cycle half-cosine envelope, which results in an approximate bandwidth of $B_0 = 0.1875/T_s$. The five AM signals have carrier frequencies $f_1 = 0.156/T_s, f_2 = 0.203/T_s, f_3 = 0.266/T_s, f_4 = 0.313/T_s, f_5 = 0.375/T_s$, and bandwidths $B_1 = 0.04/T_s$, $B_2 = 0.05/T_s$, $B_3 = 0.045/T_s$, $B4 = 0.04/T_s$, $B_5 = 0.08/T_s$. With the use of the same measurement parameters (FFT length $= 32,768$) as in Example 9 in Section II, the SCD for these six signals in noise was measured. The resultant

SCD magnitude is shown in Figure 16a. Also shown in Figures 16b and 16c are the SCD magnitudes for the BPSK signal alone and for the five AM interferences plus noise alone. Although all six signals exhibit strong spectral redundancy, the cycle frequencies α at which this redundancy exists are distinct because the carrier frequencies are all distinct. Thus, an accurate estimate of the SCD for the BPSK signal is easily extracted from the SCD for the corrupted measurements. Similarly, accurate estimates of the SCDs for each of the five AM signals can be extracted. Consequently, any information contained in these SCDs can be reliably extracted. In connection with this example, let us briefly consider some of the signal processing tasks outlined at the beginning of Section I. □

B. DETECTION AND CLASSIFICATION

We can see from Figure 16 that knowing the particular pattern of the SCDs for BPSK and AM signals (see Figures 8 and 14) enables us to detect the presence of six signals and to classify them according to modulation type. This would be impossible if only PSD (SCD at $\alpha = 0$) measurements were used. One approach to exploiting the spectral redundancy of a signal to detect its presence is to generate a spectral line at one of its cycle frequencies and then detect the presence of the spectral line (cf. Section II). It has been shown that the maximum-SNR spectral-line generator for a signal $s(t)$ in additive Gaussian noise and interference with PSD $S_n(f)$ produces the detection statistic (cf. [1, Chapter 14, Sec. E] for continuous time)

$$z = \left| \int_{-1/2}^{1/2} \hat{S}_x^\alpha(f) \frac{S_s^\alpha(f)^*}{S_n(f+\alpha/2)S_n(f-\alpha/2)}\, df \right| \tag{100}$$

for comparison to a threshold. In Eq. (100), $\hat{S}_x^\alpha(f)$ is a crude estimate of $S_x^\alpha(f)$ obtained by deleting the time-averaging operation $\langle\cdot\rangle$ and the limiting operation from Eq. (44) and choosing B equal to the reciprocal of the record length of $x(t)$ (i.e., $\hat{S}_x^\alpha(f)$ is the cyclic periodogram). It can be shown that Eq. (100) is equivalent to whitening the noise and interference using a filter with transfer function $[S_n(f)]^{-1/2}$, and then correlating the measured SCD for the noise-and-interference-whitened data with the ideal SCD of the signal, as transformed by the whitener, to be detected [1, Chapter 14,

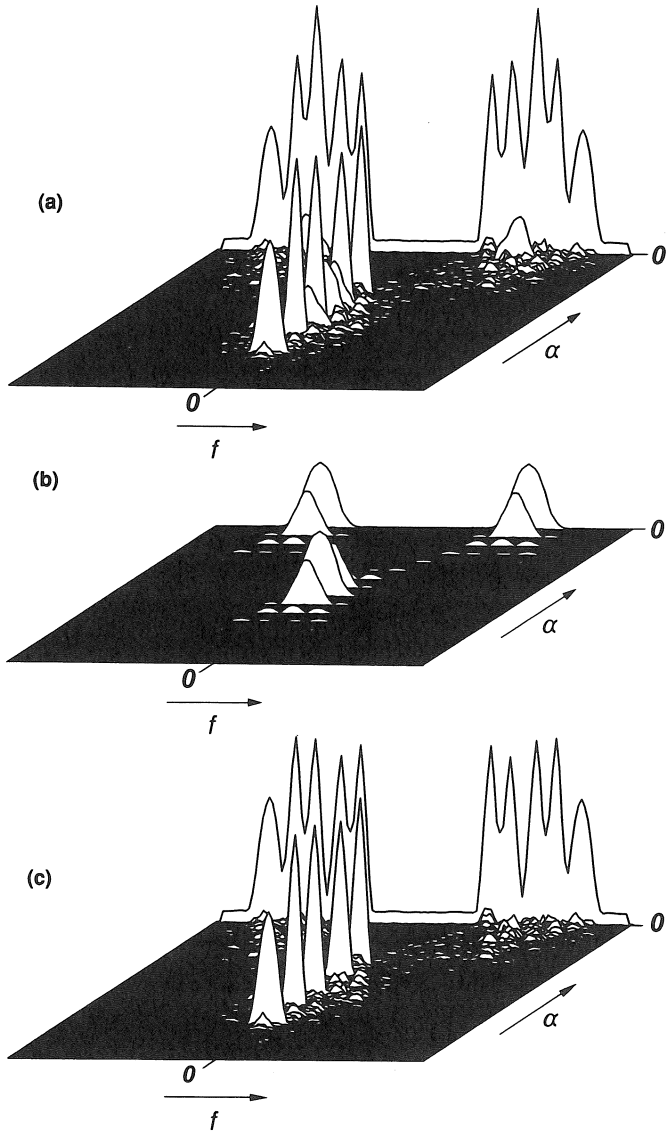

Figure 16: Estimated spectral-correlation densities (SCDs). (a) SCD magnitude for a BPSK signal corrupted by white noise and five AM interferences. (b) SCD magnitude for the BPSK signal alone. (c) SCD magnitude for the white noise and five AM interferences. (The power levels, center frequencies, and bandwidths for the signals and noise are specified in the text; the record length used is 32,768 time-samples and 1,024 adjacent frequency (f) bins are averaged together.)

Sec. E]. Equivalently, for noise consisting of a white component plus strong narrowband components, Eq. (100) corresponds to attenuating the narrowband components well below the white-noise component—that is, excising the narrowband components—using a filter with transfer function $1/S_n(f)$, and then correlating the measured SCD for the narrowband-excised data with the ideal SCD of the signal (untransformed by the excision filter).

A detailed study of both optimum (e.g., maximum-SNR and maximum-likelihood) and more practical suboptimum detection on the basis of SCD measurement is reported in [12], and receiver operating characteristics for these detectors obtained by simulation are presented in [13–14].

C. PARAMETER ESTIMATION

Once the six signals have been detected and classified, their carrier frequencies and phases and the keying rate and phase of the BPSK signal can—with sufficiently long signal duration—be accurately estimated from the magnitude and phase of the SCD (cf., f_0 and θ in Eq. (52), and T_0 and ϵ in Eq. (85)) [14]. It is clear from the theory discussed in Section II that SCD measurement is intimately related to the measurement of the amplitudes and phases of sine waves generated by quadratic transformations of the data. Thus, the fact that an SCD feature occurs at $\alpha = 2f_0$ for each carrier frequency f_0 is a direct result of the fact that a sine wave (spectral line) with frequency $\alpha = 2f_0$ and phase 2θ can be generated by putting the data through a quadratic transformation. Similarly, for the SCD feature at $\alpha = 1/T_0$, where $1/T_0$ is the keying rate, a spectral line with frequency $\alpha = 1/T_0$ and phase ϵ can be quadratically generated. Consequently, SCD measurement is useful either directly or indirectly for estimation of synchronization parameters (frequencies and phases) required for the operation of synchronized receivers. The link between synchronization problems and spectral redundancy is pursued in [15–16].

D. TIME-DIFFERENCE-OF-ARRIVAL ESTIMATION

The cross SCD $S_{wx}^{\alpha}(f)$ for two signals $x(t)$ and $w(t)$ is defined in a way

that is analogous to the definition Eq. (44) and Eqs. (25)–(26) of the SCD $S_x^\alpha(f)$. That is, $x(t)$ in Eq. (25) is simply replaced with $w(t)$. If we were to compute the cross SCD for two sets of corrupted measurements obtained from two receivers, then the cross SCD magnitude would look very similar to that in Figure 16 (except that the low flat feature at $\alpha = 0$, which represents the PSD of the receiver noise, would be absent), but the phase of the cross SCD would contain a term linear in f for each value of α where the auto SCD of one of the six signals is nonzero. The slope of this linear phase is proportional to the time difference of arrival (TDOA) of the wavefront at the two receivers for the particular signal with that feature. That is, for $x(t)$ from one receiver given by Eq. (97) and $w(t)$ from the other receiver given by

$$w(t) = \sum_{\ell=1}^{L} a_\ell s_\ell(t - t_\ell) + m(t) \tag{101}$$

where $\{t_\ell\}$ are the TDOAs, we have

$$S_{wx}^\alpha(f) = a_\ell S_{s_\ell}^\alpha(f) e^{-i2\pi(f+\alpha/2)t_\ell} \tag{102}$$

provided that $s_\ell(t)$ is the only signal with cycle frequency α. Consequently, accurate estimates of the TDOAs of each of these signals can be obtained from the cross SCD measurement, regardless of temporal and spectral overlap or of the closeness of the individual TDOAs. In other words, the signal selectivity in the α domain eliminates the problem of resolving TDOAs of overlapping as well as nonoverlapping signals.

For example, it follows from Eqs. (99) and (102) that

$$\frac{S_{wx}^\alpha(f)}{S_x^\alpha(f)} = a_\ell e^{-i2\pi(f+\alpha/2)t_\ell} \tag{103}$$

over the support band of $S_{s_\ell}^\alpha(f)$. This suggests doing a weighted least squares fit, with respect to a_ℓ and t_ℓ, of a measurement of the left side of Eq. (103) to the right side:

$$\min_{\hat{a}_\ell, \hat{t}_\ell} \left\{ \int_{-1/2}^{1/2} \left| W_\alpha(f) \left[\frac{\hat{S}_{wx}^\alpha(f)}{\hat{S}_x^\alpha(f)} - \hat{a}_\ell e^{-i2\pi(f+\alpha/2)\hat{t}_\ell} \right] \right|^2 df \right\} \tag{104}$$

where $W_\alpha(f)$ is some weighting function. After minimization with respect to $\hat{a}_\ell$, this reduces to

$$\max_{\hat{t}_\ell} \left\{ \int_{-1/2}^{1/2} |W_\alpha(f)|^2 \frac{\hat{S}_{wx}^\alpha(f)}{\hat{S}_x^\alpha(f)} e^{-i2\pi(f+\alpha/2)\hat{t}_\ell} \, df \right\}. \tag{105}$$

The two algorithms corresponding to the two choices $W_\alpha(f) = S_x^\alpha(f)$ (which yields the SPECCOA method) and $W_\alpha(f) = 1$ over some band (which yields the SPECCORR method), along with several other related algorithms are studied in detail in [14, 17, 18]. It is also shown in [19] that this approach is easily generalized to the problem of multipath channel identification where multiple t_ℓ and a_ℓ for a single signal are to be estimated using the least squares criterion Eq. (104) with a sum over ℓ included (provided that the multiple t_ℓ are resolvable, that is, spaced farther apart than the width of the inverse discrete Fourier transform of $|W_\alpha(f)|^2 S_{wx}^\alpha(f)/S_x^\alpha(f)$).

E. SPATIAL FILTERING

Continuing in the same vein, we consider receiving these same six signals in noise with an antenna array. Then we can use the signal selectivity in α to blindly adapt (without any training information other than knowledge of the cycle frequencies α of the signals of interest) a linear combiner of the complex-valued outputs from the elements in the array to perform spatial filtering. Specifically, by directing the linear combiner to enhance or restore spectral redundancy (or conjugate spectral redundancy) in its output at a particular cycle frequency α, the combiner will adapt to null out all other signals (if there are enough elements in the array to make this nulling possible). This behavior of the combiner can be seen from the fact that the spectral correlation coefficient for $x(t)$ in Eq. (97) is (from Eq. (99))

$$\rho_x^\alpha(f) = \frac{S_{s_\ell}^\alpha(f)}{[S_x(f+\alpha/2)S_x(f-\alpha/2)]^{1/2}}, \tag{106}$$

where

$$S_x(f) = S_n(f) + \sum_{k=1}^{L} S_{s_k}(f), \tag{107}$$

and, similarly, the temporal correlation coefficient for the frequency-shifted

ersions of $x(t)$ is

$$\gamma_x^\alpha(\tau) = \frac{R_{s_\ell}^\alpha(\tau)}{R_x(0)}, \tag{108}$$

where

$$R_x(0) = R_n(0) + \sum_{k=1}^{L} R_{s_k}(0). \tag{109}$$

Thus, nulling signals other than $s_\ell(t)$ in the output $x(t)$ of the linear combiner and attenuating the noise $n(t)$ in $x(t)$ reduces the denominators in Eqs. (106) and (108) but not the numerators. Hence, $|\rho_x^\alpha(f)|$ and $|\gamma_x^\alpha(\tau)|$ can be increased by attenuating the noise and nulling any of the signals other than $s_\ell(t)$. Moreover, the linear combiner needs no knowledge of the reception characteristics of the array (no calibration) to accomplish this attenuation and nulling. To be more specific, let us consider the narrowband model

$$\boldsymbol{x}(t) = \boldsymbol{a}(\theta)s(t) + \boldsymbol{i}(t)$$

for the vector of complex envelopes of the outputs of the array of antennas, where $\boldsymbol{a}(\theta)$ is the direction vector for the array corresponding to the plane wave signal $s(t)$ arriving at angle θ, and $\boldsymbol{i}(t)$ consists of noise plus interfering signals arriving at other angles. A training signal $\tilde{s}(t)$ is formed by linearly combining the array output, using a prescribed complex-valued eight vector $\boldsymbol{c}$ (which adjusts the magnitudes and phases of the complex envelopes), and then time- and frequency-shifting the result to obtain

$$\tilde{s}(t) = \boldsymbol{c}^H \boldsymbol{x}(t+\tau)e^{-i2\pi\alpha t} \tag{110}$$

where $(\cdot)^H$ denotes conjugate transpose. The output $\hat{s}(t)$ of the spatial filter is obtained by linearly combining the array outputs using the weight vector $\boldsymbol{w}$,

$$\hat{s}(t) = \boldsymbol{w}^\dagger \boldsymbol{x}(t).$$

where $(\cdot)^\dagger$ denotes transpose. The spatial filter weights $\boldsymbol{w}$ that minimize the time-averaged squared error between the filter output $\hat{s}(t)$ and the training signal $\tilde{s}(t)$ can then be sought:

$$\min_{\boldsymbol{w}} \left\langle |\hat{s}(t) - \tilde{s}(t)|^2 \right\rangle. \tag{111}$$

By setting to zero the partial derivatives of this error, with respect to eacl of the weights in the vector w, we obtain the optimum weight vector

$$w = R_x^{-1} R_{\tilde{s}x}, \tag{112}$$

where R_x is the autocorrelation matrix for $x(t)$ and $R_{\tilde{s}x}$ is the crosscorrelation vector for $\tilde{s}(t)$ and $x(t)$. Substituting Eq. (110) into Eq. (112) yields the result

$$w = bR_x^{-1} R_{sx}, \tag{113}$$

where the scalar b is given by

$$b \triangleq \gamma_s^{\alpha}(\tau) c^H a(\theta) e^{i\pi\alpha\tau}. \tag{114}$$

The weight vector w in Eq. (113) is the scaled solution to the optimization problem Eq. (111) with a perfect training signal $\tilde{s}(t) = s(t)$. Thus, as long as the correlation coefficient $\gamma_s^{\alpha}(\tau)$ is sufficiently large and the prescribed weight vector c is sufficiently colinear with $a(\theta)$, so that the scalar b in Eq. (113) is not very small, the data-derived training signal $\tilde{s}(t)$ can be used in place of $s(t)$ to obtain the optimum spatial filter. In practice, the ideal time-average $\langle\cdot\rangle$ is replaced with a finite-length time-average and the correlations in Eq. (112) therefore become finite-time estimates. The weight vector w can be computed in block form using Eq. (112) directly, or the recursiv least squares algorithm can be used. Alternatively, the solution Eq. (112, can be approximated using the LMS algorithm. Although the prescribed weight vector c can be as simple as $[1\ 0\ 0\ 0\ \ldots\ 0]^{\dagger}$, convergence to the optimum filter Eq. (113) can be accelerated by adaptively adjusting c as well as w. A thorough study of this and other spectral-coherence-restoral algorithms that perform blind adaptive spatial filtering is reported in [20] and [21] and a tutorial discussion is given in [22].

F. DIRECTION FINDING

We can take this approach one step further if we do indeed have calibration data for the reception characteristics of an antenna array because we can then also exploit signal selectivity in α to perform high-resolution direction

nding (DF) without some of the drawbacks (described below) of conventional methods for high-resolution DF, such as subspace fitting methods [23], that do not exploit spectral redundancy. In particular, let us consider the narrowband model

$$\boldsymbol{x}(t) = \sum_{\ell=1}^{L} \boldsymbol{a}(\theta_\ell) s_\ell(t) + \boldsymbol{n}(t) \tag{115}$$

for the analytic signal (or complex envelope) $\boldsymbol{x}(t)$ of the received data vector of dimension r, where $\boldsymbol{a}(\theta_\ell)$ is the direction vector associated with the ℓth received signal $s_\ell(t)$, and the function $\boldsymbol{a}(\theta)$ is specified by the calibration data for the array. Then, by working with the magnitude and phase information contained in the $r \times r$ cyclic correlation matrix

$$\boldsymbol{R}_{\boldsymbol{x}}^{\alpha}(\tau) = \boldsymbol{R}_{s_k}^{\alpha}(\tau) = \boldsymbol{a}(\theta_k) \boldsymbol{R}_{s_k}^{\alpha}(\tau) \boldsymbol{a}^H(\theta_k) \tag{116}$$

for some fixed τ, instead of working with the information contained in the conventional correlation matrix

$$\boldsymbol{R}_{\boldsymbol{x}}(0) = \sum_{\ell=1}^{L} \boldsymbol{R}_{s_\ell}(0) + \boldsymbol{R}_n(0) = \sum_{\ell=1}^{L} \boldsymbol{a}(\theta_\ell) R_{s_\ell}(0) \boldsymbol{a}^H(\theta_\ell) + \boldsymbol{R}_n(0), \tag{117}$$

we can avoid the need for advance knowledge of the correlation properties f the noise $\boldsymbol{R}_n(0)$ and interference $\boldsymbol{R}_{s_l}(0)$ for $l \neq k$, and we can avoid the constraint imposed by conventional methods that the number of elements in the array exceed the total number L of signals impinging on the array. Also, by resolving signals in α, we need not resolve them in direction of arrival. Consequently, superior effective spatial resolution is another advantage available through the exploitation of spectral redundancy. As an example of a cyclostationarity-exploiting DF method, we can use the fact that the $r \times r$ matrix in Eq. (116) has a rank of unity and the $(r-1)$-dimensional null space of this matrix is orthogonal to $\mathbf{a}(\theta_k)$. Therefore, we can choose as our estimate of θ_k that value $\hat{\theta}_k$ that renders $\mathbf{a}(\hat{\theta}_k)$ most nearly orthogonal to the null space of an estimate of the matrix $\boldsymbol{R}_{\boldsymbol{x}}^{\alpha}(\tau)$ obtained from finite-time averaging. Similar remarks apply to the conjugate cyclic autocorrelation matrix. A thorough study of this approach to signal-selective DF is reported in [24] and [25], where various algorithms are

introduced and their performances are evaluated, and a tutorial discussion is given in [22].

In the preceding paragraphs of this Section III, the signal processing tasks (with the exception of spatial filtering) involve making decisions or estimating parameters, but do not involve estimating (or extracting) an entire signal or an information-bearing message carried by the signal. Nevertheless, for the signal-extraction problem, the utility of spectral redundancy is just as apparent, as explained in the following paragraphs.

G. SIGNAL EXTRACTION

Spectrally redundant signals that are corrupted by interfering signals can be more effectively extracted in some applications by exploiting spectral correlation through the use of periodic or multiply-periodic linear time-variant filters, instead of the more common time-invariant filters. These time-variant filters enable spectral redundancy to be exploited for signal extraction, because such filters perform frequency-shifting operations (cf. Eq. (79)) as well as the frequency-dependent magnitude-weighting and phase-shifting operations performed by time-invariant filters. The utility of this is easily seen for the simple example in which interference in some portions of the spectral band of the signal is so strong that it overpowers the signal in those partial bands. In this case, a time-invariant filter can only reject both the signal and the interference in those highly corrupted bands, whereas a time-variant filter can replace the rejected spectral components of the signal of interest with spectral components from other uncorrupted (or less corrupted) bands that are highly correlated with the rejected components from the signal.

AM is an obvious example of this because of the complete redundancy that exists between its upper sideband (above the carrier frequency) and its lower sideband (below the carrier frequency). Although this redundancy is exploited in the conventional double-sideband demodulator to obtain a 3dB gain in SNR performance, it is seldom exploited properly when partial-band interference is present. The proper exploitation in this case is illustrated in Figure 17. Figure 17a shows the spectral content (Fourier transform

magnitude of a finite segment of data) for an AM signal with partial-band interference in the upper sideband. Figure 17b shows the spectral content after the interference has been rejected by time-invariant filtering. The signal distortion caused by rejection of the signal components along with the interference can be completely removed by simply shifting replicas of perfectly correlated components from the lower sideband into the upper sideband, and then properly adjusting their magnitudes and phases, as suggested in Figure 17c.

A less easily explained example involves two spectrally overlapping linearly modulated signals such as AM, PAM, ASK, PSK, or digital QAM (quadrature AM). It can be shown that, regardless of the degree of spectral and temporal overlap, each of the two interfering signals can be perfectly extracted by using frequency shifting and complex weighting, provided only that they have either different carrier frequencies or phases (AM, ASK, BPSK) or different keying rates or phases (PAM, ASK, PSK, digital QAM) and at least 100% excess bandwidth (bandwidth in excess of the minimum Nyquist bandwidth for zero intersymbol interference). A blind-adaptive frequency-shift filter for separating such signals is described in [3].

In addition, when the excess bandwidth is $(L-1)100\%$, L spectrally overlapping signals can be separated if they have the same keying rate but fferent keying phases or carrier frequencies. Also, when broadband noise is present, extraction of each of the signals can in many cases be accomplished without substantial noise amplification. To illustrate the potential for signal separation in this case, consider L digital QAM signals with $(L-1)100\%$ excess bandwidth, all sharing the same carrier frequency and keying rate, but with distinct keying phases. Then for any particular frequency f in the Nyquist band the received spectral component at that frequency is a weighted sum of the L spectral components of the L individual signals at that same frequency f and the same is true at the $L-1$ additional frequencies separated by the keying rate, except that the sets of L weights in each of these L weighted sums are distinct, although the L sets of L spectral components are all identical in the sense that their spectral correlation coefficients are unity in magnitude. Thus, for each frequency within the Nyquist band, we have L equations with L unknowns (the L keying phases).

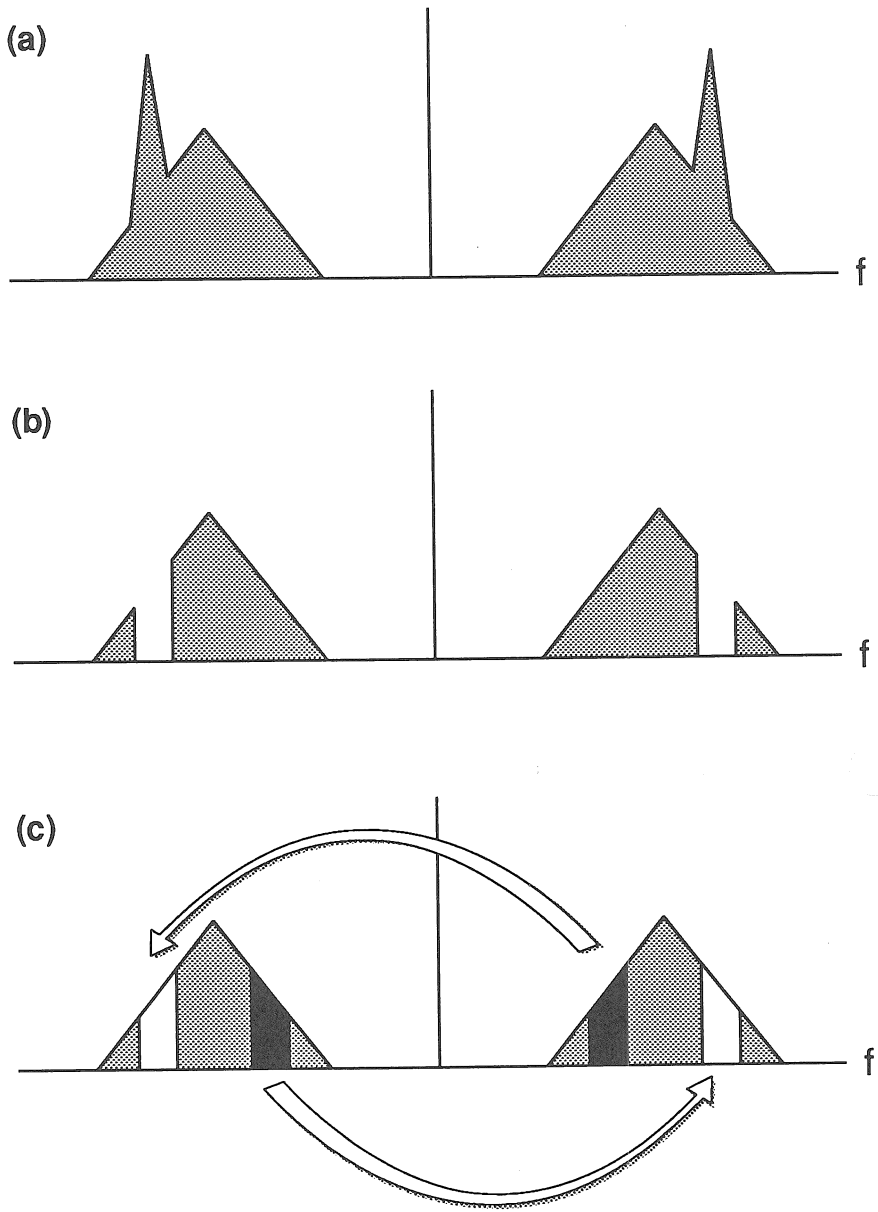

Figure 17: Illustration of power spectral densities (PSDs) for cochannel-interference removal with minimal signal distortion. (a) PSD for AM signal plus interference. (b) PSD after interference removal by time-invariant filtering. (c) PSD after distortion removal by frequency-shifting.

In practice, the $L \times L$ array of weights also will be unknown and will have to be adaptively learned. This particular problem of separating multiple digital QAM signals sharing the same carrier frequency (or baseband PAM signals) and sharing the same keying rate is explored in [26].

A final example involves the reduction of the signal distortion due to frequency-selective fading caused by multipath propagation. Straightforward amplification in faded portions of the spectrum using a time-invariant filter suffers from the resultant amplification of noise. In contrast to this, a periodically time-variant filter can replace the faded spectral components with stronger, highly correlated components from other bands. If these correlated spectral components are weaker than the original components before fading there will be some noise enhancement when they are amplified. But the amount of noise enhancement can be much less than that which would result from the time-invariant filter, which can only amplify the very weak faded components.

Detailed studies of the principles of operation and the mean-squared-error performance of both optimum and adaptive frequency-shift filters are reported in [1, Chapter 14, Secs. A, B], [4, Chapter 12, Sec. 12.8], [27–29].

H. PREDICTION AND CAUSALITY

If a signal is correlated with time-shifted versions of itself (that is, if it is not a white-noise signal), then its past can be used to predict its future. The higher the degree of temporal coherence, the better the prediction can be. A signal that exhibits cyclostationarity is also correlated with frequency-shifted versions of itself. Consequently, its future can be better predicted if frequency-shifted versions of its past are also used, so that its spectral coherence as well as its temporal coherence is exploited. For example, if $x(t)$ has cycle frequencies $\{\alpha_0, \ldots, \alpha_{N-1}\}$ then we can estimate the future value $x(t+\tau)$ for some $\tau > 0$ using a linear combination of the present and past values of the N signals

$$x_q(t) = x(t)e^{i2\pi\alpha_q t} \text{ for } q = 0, \cdots, N-1. \tag{118}$$

That is, the predicted value is given by

$$\hat{x}(t+\tau) = \sum_{u=0}^{M-1} \sum_{q=0}^{N-1} h_q(u) x_q(t-u), \tag{119}$$

where M is the memory-length of the predictor. The set of MN prediction coefficients that minimize the time-averaged (over t) squared magnitude of the prediction error $\hat{x}(t+\tau) - x(t+\tau)$ can be shown to be fully specified by the cyclic correlation functions for the N cycle frequencies. Specifically, the set of MN coefficients $\{h_q(u)\}$ is the solution to the set of MN simultaneous linear equations

$$\sum_{u=0}^{M-1} \sum_{q=0}^{N-1} h_q(u) R_x^{\alpha_p - \alpha_q}(t-u) e^{i\pi(\alpha_p - \alpha_q)(t-u)} = R_x^{\alpha_p}(t+\tau) e^{i\pi\alpha_p(t+\tau)}$$

for $t = 0, \cdots, M-1$ and $p = 0, \cdots, N-1$. Also, the percent accuracy of prediction is determined solely by the temporal coherence functions Eq. (32) for the frequency translates. It can be shown that for each cycle frequency α_q exploited, there is a corresponding increase in the percent accuracy of the prediction.

In the same way that time-invariant autoregressive model-fitting of stationary time-series data is mathematically equivalent to time-invariant linear prediction [1, Chapter 9, Sec. B], it can be shown that frequency-shif (or polyperiodic time-variant) autoregressive model-fitting is mathematically equivalent to frequency-shift linear prediction. Studies of this problem are reported in [30–41]. Also, the univariate prediction problem for cyclostationary (not polycyclostationary) time-series is equivalent to the multivariate prediction problem for stationary time-series [31]. This follows from the representation of univariate cyclostationary time-series in terms of multivariate stationary time-series [4, Chapter 12, Sec. 12.6], [42, 43].

A measure of the degree to which one time-series causes another time-series is the degree to which the present and past of the former can linearly predict the future of the latter. If the two time-series are jointly cyclostationary, then cyclic as well as constant causality is possible. In fact, by considering only time-invariant predictors, it is possible to conclude for

some pairs of time-series that no causality exists when, in fact, one time-series is perfectly cyclically caused by the other. An example of this is $x(t) = z(t)$ and $y(t) = z(t-\tau)\cos(t)$, where $t > 0$ and $z(t)$ is an independent identically distributed sequence. The best linear time-invariant predictor of $y(t)$ using the past of $x(t)$ is $\hat{y}(t) = 0$, whereas the best linear periodically time-variant predictor is $\hat{y}(t) = x(t-\tau)\cos(t) = y(t)$, which yields perfect prediction. Moreover, if $z(t)$ takes on values of only ± 1 it is particularly easy to show for this example that

$$\langle x^n(t-s)y^m(t)\rangle = \langle x^n(t-s)\rangle \langle y^m(t)\rangle$$

for all positive integers m, n and s. Consequently, even the best nonlinear time-invariant predictor is $\hat{y}(t) = 0$.

I. NONLINEAR SYSTEM IDENTIFICATION

A popular approach to the identification of nonlinear dynamical systems from input-output measurements is to model the system in terms of the Volterra series, which is a generalization of the power-series (or polynomial) representation of a memoryless system to systems with memory, and then to identify one-by-one the Volterra kernels, each one of which characterizes one term in the series representation. The first kernel is the impulse response of the linear part of the system. The second kernel is a two-dimensional generalization of the impulse response for the quadratic part of the system, and so on. Common approaches to identifying the kernels are based on crosscorrelation measurements between the unknown-system output and specially designed nonlinear functions of the system input.

Although the fundamental theory of this crosscorrelation approach to nonlinear system identification is built on the foundation of stationary random processes or time-series [44], it has recently been shown [45] that substantial advantages can be gained by using cyclostationary inputs to the unknown system and cyclic crosscorrelations. In particular, desirable orthogonality (zero-correlation) properties between the system output and nonlinear functions of the input that are not possible for stationary inputs are possible for cyclostationary inputs, and this leads to particularly convenient designs for the inputs and the nonlinear functions. Moreover, this

approach of exploiting cyclostationarity to identify time-invariant systems has recently been generalized to identify polyperiodic nonlinear systems [46]. In [45, 46], the basic theory of this new approach is presented for both a time-domain method, which directly identifies the Volterra kernels or their polyperiodic counterparts, and a frequency-domain method, which directly identifies the multidimensional Fourier transforms of the kernels—the Volterra transfer functions—and several examples of cyclostationary inputs and corresponding nonlinear functions are given. This work exploits higher-than-second-order cyclostationarity, the principles of which are given in the next section.

IV. HIGHER-ORDER CYCLOSTATIONARITY

A. INTRODUCTION

In Example 8 of Section II.F it is shown that, for QPSK signals, spectral lines with frequencies equal to the harmonics of the keying rate can be generated by using quadratic nonlinearities, but that no quadratic nonlinearity can generate spectral lines with frequencies associated with the carrier frequency—a quartic (4th-order) nonlinearity is needed. Similarly, for some PSK signals with $M \geq 4$ states, a nonlinearity of order M is needed to generate a spectral line with frequency associated with the carrier frequency (M times the carrier frequency). In Example 2 of Section II.F, the SCD (83) for PAM signals is shown to depend on the product of shifted pulse transforms:

$$\tilde{P}(f + \alpha/2)\tilde{P}^*(f - \alpha/2), \qquad \alpha = k/T_0.$$

Thus, if the pulse transform has support that is limited to the interval $[-1/2T_0, 1/2T_0)$, all such products are identically zero except that for $\alpha = 0$, which implies that the signal is second-order stationary in the wide sense: no spectral lines can be generated by using quadratic nonlinearities. An example of such a signal is the duobinary coded (or partial response) PAM signal [47, 48]. However, a spectral line with frequency equal to the keying rate can be generated from this kind of signal by using a fourth-order

nonlinearity such as

$$L_x(t,\tau)_4 \triangleq x(t+\tau_1)x(t+\tau_2)x(t+\tau_3)x(t+\tau_4),$$

provided that the support of the pulse transform is not limited to the interval $[-1/4T_0, 1/4T_0)$ (that is, the bandwidth exceeds $1/4T_0$). Thus, in the case of a QPSK signal with carrier frequency f_0, there are some delay sets $\tau \triangleq [\tau_1 \cdots \tau_4]^\dagger$ such that[8]

$$\langle L_x(t,\tau)_4 e^{-i2\pi 4 f_0 t}\rangle \neq 0,$$

whereas

$$\langle x(t+\tau_1)x(t+\tau_2)e^{-i2\pi k f_0 t}\rangle = 0$$

for all τ_1, τ_2, and $k \neq 0$. Similarly, in the case of a PAM signal with keying rate $1/T_0$ and pulse transform that has positive-frequency bandwidth less than $1/2T_0$ but greater than $1/4T_0$, there are some delay sets τ such that

$$\left\langle L_x(t,\tau)_4 e^{-i2\pi t/T_0}\right\rangle \neq 0,$$

whereas

$$\left\langle x(t+\tau_1)x(t+\tau_2)e^{-i2\pi tk/T_0}\right\rangle = 0$$

for all τ_1, τ_2, and $k \neq 0$.

The preceding observations motivate the study of both fourth-order nonlinear transformations and of signals that produce spectral lines only after undergoing nonlinear transformations of order no smaller than four because we have seen in previous sections that regenerated spectral lines can be utilized by signal processing algorithms to obtain superior performance when compared to algorithms that ignore these spectral lines. But these observations also motivate the more general study of spectral lines in the outputs of higher-order nonlinear transformations such as the nth-order delay product

$$L_x(t,\tau)_n \triangleq \prod_{j=1}^{n} x(t+\tau_j)$$

[8] The notation $\langle\cdot\rangle$ is used throughout this section to denote the continuous-time counterpart of the discrete-time-averaging operation defined in Section II.A.

and, for complex-valued signals,

$$L_x(t,\tau)_n \triangleq \prod_{j=1}^{n} x^{(*)_j}(t+\tau_j),$$

where $(*)_j$ denotes an optional conjugation for the jth factor $x(t+\tau_j)$.

There are also other important reasons to study the sine-wave components (spectral lines) in the outputs of higher-order nonlinear transformations. For instance, to determine the performance of the algorithms described in Section III, which exploit second-order cyclostationarity, higher-order moments of cyclostationary signals are required (for example, the variance of an SCD estimator depends on the spectral lines in the fourth-order delay products). As another example, unintentional or unavoidable nonlinearities (e.g., amplifier nonlinearities) in signal processing systems involving cyclostationary signals can produce undesired spectral lines in their outputs [49]. A complete analysis of the performance of such systems requires an understanding of the effects of nth-order nonlinear transformations on the input signals. As a final example, nonlinear system identification can be accomplished by using specially designed cyclostationary inputs, as described and illustrated in [45], and this entails measurement of nonlinearly generated sine waves.

The preceding examples show that a case of primary interest occurs when higher-than-second order nonlinearities are *required* to generate spectral lines with desired frequencies, or to generate any spectral lines whatsoever. However, it is also the case that if a signal is cyclostationary of order two, then it is also cyclostationary of order $2n$ for all natural numbers n. This is simply a result of the fact that if there are sine waves in second-order delay products, then there are sine waves in some delay products with an even number of factors, because sine waves multiplied by sine waves results in sine waves. This observation suggests the following question: Is there some part of the higher-order sine-wave strength that is *not* due to the multiplication of lower-order sine-wave strengths? Or, rephrased: Can the higher-order sine-wave strength be larger or smaller than the product of lower-order strengths? If so, how can we characterize the difference? We shall see that in the process of seeking an answer to this question, we are led naturally to the central parameters of the theory of higher-order cyclo-

stationarity, namely temporal and spectral cumulants. Cumulants are also a focal point in the study of the higher-order statistics of stationary signals [50]–[54], which is currently a subject of considerable research effort, but there cumulants are not derived, they are used as a tool in the design of algorithms for the solution of signal processing problems.

A brief development of the principles of higher-order cyclostationarity (originally introduced in [55]–[61]) is given in the remaining parts of this section. To minimize distracting technical issues associated with discrete-time, such as aliasing, the discussion is limited to continuous-time signals.

B. SPECTRAL LINE GENERATION

Let us consider, as motivation, all nonlinear signal processing operations that can be expanded in a Volterra series (generalized in the sense that the kernels are not necessarily causal):

$$y(t) = \sum_n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\boldsymbol{\tau}) L_x(t, \boldsymbol{\tau})_n \, d\boldsymbol{\tau},$$

where $\boldsymbol{\tau} \triangleq [\tau_1 \cdots \tau_n]^\dagger$, $L_x(t, \boldsymbol{\tau})_n$ is the nth-order delay product of the input $x(t)$

$$L_x(t, \boldsymbol{\tau})_n \triangleq \prod_{j=1}^{n} x(t + \tau_j), \tag{120}$$

and $h_n(\boldsymbol{\tau})$ are the Volterra kernels. This series can be generalized for complex-valued $x(t)$ to include multiple terms for each order to accommodate conjugated factors in Eq. (120). Such generalized versions of Eq. (120) are assumed henceforth.

We are interested in the finite-strength additive sine-wave components present in $y(t)$ but absent in $x(t)$, that is, those sine waves that are generated by the action of the nonlinear operation on $x(t)$, which is assumed to exhibit cyclostationarity. For example, the strength of the sine wave with frequency α in $y(t)$ is given by

$$\left\langle y(t) e^{-i2\pi\alpha t} \right\rangle = \sum_n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\boldsymbol{\tau}) \left\langle L_x(t, \boldsymbol{\tau})_n e^{-i2\pi\alpha t} \right\rangle d\boldsymbol{\tau},$$

where restrictions on $x(t)$ that validate the interchange of operations are assumed to apply. Thus, we need only study the statistical quantities

$$\left\langle L_x(t,\tau)_n e^{-i2\pi\alpha t}\right\rangle,$$

for arbitrary positive integers n, which are the strengths of the sine-wave components contained in the nth-order delay products of $x(t)$.

1. The Cyclic Temporal Moment Function.

The nth-order delay product is an elementary nth-order homogeneous polynomial nonlinear transformation of $x(t)$. This transformed signal can be decomposed into a polyperiodic (periodic or multiply-periodic—also called almost periodic [62]) part $p(t,\tau)_n$ and an aperiodic residual part $m(t,\tau)_n$,

$$L_x(t,\tau)_n = p(t,\tau)_n + m(t,\tau)_n,$$

where

$$\left\langle m(t,\tau)_n\, e^{-i2\pi\alpha t}\right\rangle = 0, \quad \text{for all } \alpha. \tag{121}$$

The polyperiodic portion of $L_x(t,\tau)_n$ has associated with it the Fourier series

$$p(t,\tau)_n = \sum_\alpha R_x^\alpha(\tau)_n\, e^{i2\pi\alpha t}, \tag{122}$$

where

$$R_x^\alpha(\tau)_n \triangleq \left\langle p(t,\tau)_n\, e^{-i2\pi\alpha t}\right\rangle. \tag{123}$$

It is assumed herein that the partial sums in the Fourier series Eq. (122) converge uniformly in t for each τ to $p(t,\tau)_n$. Then $p(\cdot,\tau)_n$ is an almost periodic function in the mathematical sense [62], the limit Eq. (123) exists for each τ, and the set of values of the real variable α for which $R_x^\alpha(\tau)_n \not\equiv 0$ for each τ is countable. That is, there is at most a countable set of incommensurate periods in the polyperiodic function $p(t,\tau)_n$ for each τ. It is further assumed that the union over all τ of the sets of values of α for which $R_x^\alpha(\tau)_n \not\equiv 0$ results in a countable set.

The delay-product signal Eq. (120) can be expressed as

$$L_x(t,\tau)_n = \sum_\alpha R_x^\alpha(\tau)_n\, e^{i2\pi\alpha t} + m(t,\tau)_n, \tag{124}$$

where the sum is over the countable set of real numbers α for which $R_x^\alpha(\tau)_n \not\equiv 0$. From Eqs. (121) and (124), we have

$$R_x^\alpha(\tau)_n = \left\langle L_x(t,\tau)_n \, e^{-i2\pi\alpha t} \right\rangle . \tag{125}$$

Each value of α in the representation Eq. (124) is called a *cycle frequency of order n*, and $R_x^\alpha(\tau)_n$ is called the *cyclic temporal moment function* (CTMF) *of order n*. For $n = 2$, $\tau_1 = \tau/2$, and $\tau_2 = -\tau/2$, the CTMF is identical to the continuous-time counterpart of the cyclic autocorrelation studied in previous sections (cf. Eq. (20)).

The sum of all sine waves in $L_x(t,\tau)_n$ can be interpreted as the result of applying the temporal expectation operator $\hat{E}^{\{\alpha\}}\{\cdot\}$ to the delay product,

$$\hat{E}^{\{\alpha\}}\{L_x(t,\tau)_n\} = \sum_\alpha R_x^\alpha(\tau)_n \, e^{i2\pi\alpha t}.$$

This temporal expectation operation is completely analogous to the probabilistic expectation operation $E\{\cdot\}$ [3, 63]. The result of applying the operator $\hat{E}^{\{\alpha\}}\{\cdot\}$ to the delay product $L_x(t,\tau)_n$ is called the *temporal moment function* (TMF), and is denoted by $R_x(t,\tau)_n$,

$$R_x(t,\tau)_n \triangleq \hat{E}^{\{\alpha\}}\{L_x(t,\tau)_n\} . \tag{126}$$

An individual sine-wave component of the TMF, such as $R_x^\alpha(\tau)_n e^{i2\pi\alpha t}$, is called a *moment sine wave of order n* (to distinguish it from a *cumulant sine wave of order n* which is defined in Section IV.C).

If the signal $x(t)$ exhibits second-order cyclostationarity, then at least some of the fourth-order moment sine waves are composed (in part or completely) of products of second-order sine waves whose frequencies sum to a fourth-order cycle frequency α. Generally, nth-order moment sine waves can be composed (in part or completely) of products of lower-order sine waves whose orders sum to n. In the following section, we remove these product sine waves in order to reveal that part of the nth-order moment sine wave that does not arise from sine waves contained in factors of the delay product and is therefore uncontaminated by lower-order product sine waves. If this sine-wave component is nonzero, it is called a *pure nth-order sine wave.*

2. Pure nth-Order Sine Waves.

For low orders n, it is easy to mathematically characterize a pure nth-order sine wave in a way that matches our intuition. For notational simplicity, we choose the case of no conjugated factors in Eq. (120). For $n = 1$, the moment sine waves are, by definition, pure first-order sine waves. For $n = 2$, all products of first-order moment sine waves can be subtracted from the second-order moment sine waves to obtain the pure second-order sine waves, which are denoted by $\sigma_x(t, \tau_1, \tau_2)_2$,

$$\begin{aligned}\sigma_x(t,\tau_1,\tau_2)_2 &\triangleq \hat{E}^{\{\alpha\}}\{x(t+\tau_1)x(t+\tau_2)\} \\ &\qquad -\hat{E}^{\{\alpha\}}\{x(t+\tau_1)\}\,\hat{E}^{\{\alpha\}}\{x(t+\tau_2)\} \\ &= R_x(t,\tau)_2 - R_x(t,\tau_1)_1\,R_x(t,\tau_2)_1.\end{aligned}$$

There are several interesting points to be made concerning pure second-order sine waves: (i) since $R_x(t,\tau_i)_1, i = 1, 2$, and $R_x(t,\tau)_2$ are first- and second-order moments respectively, then $\sigma_x(t,\tau_1,\tau_2)_2$ is a temporal covariance function; (ii) if $R_x(t,\tau)_1 \equiv 0$, then there are no lower-order sine waves, and the second-order moment sine waves are equal to the pure second-order sine waves; (iii) if the variables $x(t+\tau_1)$ and $x(t+\tau_2)$ are statistically independent (in the temporal sense [63]), then

$$\hat{E}^{\{\alpha\}}\{x(t+\tau_1)x(t+\tau_2)\} = \hat{E}^{\{\alpha\}}\{x(t+\tau_1)\}\,\hat{E}^{\{\alpha\}}\{x(t+\tau_2)\}$$

and therefore $\sigma_x(t,\tau_1,\tau_2)_2 = 0$, that is, there is no pure second-order sin wave for this particular pair of delays τ_1 and τ_2. An example of (iii) is the case of (poly)periodic $x(t)$, say $x(t) = \sin(2\pi f_1 t + \theta)$.

The pure third-order sine waves are obtained next. From the third-order moment sine waves, we want to subtract each possible product of lower-order sine waves, but only once each. Thus, we subtract products of *pure* second-order and *pure* first-order sine waves from the third-order moment sine waves, rather than subtracting products of first- and second-order moment sine waves:

$$\begin{aligned}\sigma_x(t,\tau)_3 &= \hat{E}^{\{\alpha\}}\left\{\prod_{j=1}^{3} x(t+\tau_j)\right\} - \sigma_x(t,\tau_1,\tau_2)_2\sigma_x(t,\tau_3)_1 \\ &\quad -\sigma_x(t,\tau_1,\tau_3)_2\sigma_x(t,\tau_2)_1 - \sigma_x(t,\tau_2,\tau_3)_2\sigma_x(t,\tau_1)_1 \\ &\quad -\sigma_x(t,\tau_1)_1\sigma_x(t,\tau_2)_1\sigma_x(t,\tau_3)_1.\end{aligned}$$

Observe that there are no other possible products of pure lower-order sine waves. The terms in the sum of products that are subtracted can be enumerated by considering the distinct partitions of the index set $\{1,2,3\}$. A partition of a set G is a collection of p subsets of G, $\{\nu_i\}_{i=1}^{p}$, with the following properties:

$$G = \bigcup_{j=1}^{p} \nu_j \quad \text{and} \quad \nu_j \bigcap \nu_k = \emptyset \quad \text{for} \quad j \neq k.$$

The set P_3 of distinct partitions of $\{1,2,3\}$ is

$$\begin{array}{lllll} p=1: & \{1,2,3\} & & & \\ p=2: & \{1,2\},\{3\} & \{1,3\},\{2\} & \{2,3\},\{1\} & \\ p=3: & \{1\},\{2\},\{3\}\,. & & & \end{array}$$

Thus, we can express the pure third-order sine wave $\sigma_x(t,\tau)_3$ as a sum over the elements of P_3:

$$\sigma_x(t,\tau)_3 = R_x(t,\tau)_3 - \sum_{\substack{P_3 \\ p\neq 1}} \left[\prod_{j=1}^{p} \sigma_x(t,\tau_{\nu_j})_{n_j} \right],$$

where τ_{ν_j} is the vector corresponding to the subset of $\{\tau_j\}_{j=1}^{3}$ with elements having indices in ν_j, and n_j is the number of elements in ν_j.

Notice that, as in the case of $n = 2$, if the first-order moment sine waves are zero, then the third-order moment sine waves are equal to the pure third-order sine waves. In this case, there are no *products* of lower-order sine waves to subtract from the moment sine waves.

The formula for the pure nth-order sine waves is

$$\sigma_x(t,\tau)_n = R_x(t,\tau)_n - \sum_{\substack{P_n \\ p\neq 1}} \left[\prod_{j=1}^{p} \sigma_x(t,\tau_{\nu_j})_{n_j} \right], \tag{127}$$

where P_n is the set of distinct partitions of the index set $\{1,2,\cdots,n\}$. The pure-sine-waves formula Eq. (127) gives all the pure nth-order sine waves associated with the delay set τ. A single pure nth-order sine wave with frequency β can be selected by using the single-sine-wave extraction operation

$$\sigma_x^{\beta}(\tau)_n\, e^{i2\pi\beta t} = \left\langle \sigma_x(u,\tau)_n\, e^{-i2\pi\beta(u-t)} \right\rangle, \tag{128}$$

and can be expressed in terms of pure lower-order sine waves by using the Fourier series for each $\sigma_x(t,\tau_{\nu_j})_{n_j}$ in Eq. (127):

$$\sigma_x(t,w)_k = \sum_{\beta_k} \sigma_x^{\beta_k}(w)_k \, e^{i2\pi\beta_k t}, \qquad w = [w_1 \cdots w_k]^\dagger, \tag{129}$$

where the sum is over all cycle frequencies β_k of order k. Thus,

$$\sigma_x^\beta(\tau)_n = R_x^\beta(\tau)_n - \sum_{\substack{P_n \\ p\neq 1}} \left[\sum_{\mathbf{1}^\dagger\boldsymbol{\beta}=\beta} \prod_{j=1}^{p} \sigma_x^{\beta_j}(\tau_{\nu_j})_{n_j} \right], \tag{130}$$

where $\boldsymbol{\beta}$ is the p-dimensional vector of cycle frequencies $[\beta_1 \ \cdots \ \beta_p]^\dagger$ and $\mathbf{1}$ is a p-dimensional vector of ones. Hence, the pure-sine-wave strength $\sigma_x^\beta(\tau)_n$ is given by the CTMF $R_x^\beta(\tau)_n$ with all products of pure lower-order sine-wave strengths, for sine waves whose frequencies sum to β and whose orders sum to n, subtracted out.

C. THE TEMPORAL PARAMETERS OF HOCS

In this section, we show that the pure nth-order sine-waves function is, in fact, an nth-order cumulant function.

1. The Cyclic Temporal Cumulant Function.

The nth-order temporal moment function $R_x(t,\tau)_n$ for the set of n variables $\{x^{(*)_j}(t+\tau_j)\}_{j=1}^n$ is defined by (126), and is also related to the nth-order joint temporal characteristic function

$$\Phi_x(t,\omega)_n \triangleq \hat{E}^{\{\alpha\}}\left\{e^{i\omega^\dagger \boldsymbol{x}(t)}\right\} \tag{131}$$

$$\boldsymbol{x}(t) \triangleq [x^{(*)_1}(t+\tau_1)\cdots x^{(*)_n}(t+\tau_n)]^\dagger,$$

as follows

$$R_x(t,\tau)_n = (i)^{-n}\frac{\partial^n}{\partial\omega_1\cdots\partial\omega_n}\Phi_x(t,\omega)_n\Big|_{\omega=\mathbf{0}}. \tag{132}$$

The nth-order temporal cumulant function (TCF) for $\{x^{(*)_j}(t+\tau_j)\}_{j=1}^n$ is defined by analogy to (132), except that the characteristic function is replaced by its logarithm (cf. [50])

$$C_x(t,\tau)_n = (i)^{-n}\frac{\partial^n}{\partial\omega_1\cdots\partial\omega_n}\ln\Phi_x(t,\omega)_n\Big|_{\omega=\mathbf{0}}. \tag{133}$$

It can be shown (cf. [50, 65] for the analogous case involving the probabilistic expectation $E\{\cdot\}$) using only (131)–(133) that the TCF can be expressed in terms of the TMF as follows

$$C_x(t,\tau)_n = \sum_{P_n} \left[(-1)^{p-1}(p-1)! \prod_{j=1}^{p} R_x(t,\tau_{\nu_j})_{n_j} \right], \tag{134}$$

where

$$R_x(t,\tau_{\nu_j})_{n_j} = \hat{E}^{\{\alpha\}} \left\{ \prod_{k\in\nu_j} x^{(*)_k}(t+\tau_k) \right\}. \tag{135}$$

Furthermore, this relation can be inverted [65] to obtain

$$R_x(t,\tau)_n = \sum_{P_n} \left[\prod_{j=1}^{p} C_x(t,\tau_{\nu_j})_{n_j} \right] \tag{136}$$

where

$$C_x(t,\tau_{\nu_j})_{n_j} = \text{Cumulant}\left\{ x^{(*)_k}(t+\tau_k) \right\}_{k\in\nu_j}. \tag{137}$$

By separating out the $p=1$ term from the sum over P_n in (136), we obtain

$$C_x(t,\tau)_n = R_x(t,\tau)_n - \sum_{\substack{P_n \\ p\neq 1}} \left[\prod_{j=1}^{p} C_x(t,\tau_{\nu_j})_{n_j} \right]. \tag{138}$$

Comparing (138) with (127), it can be seen that the pure-nth-order-sine-waves function is identical to the nth-order TCF,

$$C_x(t,\tau)_n \equiv \sigma_x(t,\tau)_n. \tag{139}$$

Since $R_x(t,\tau_{\nu_j})_{n_j}$ is (poly)periodic in t, so too is $C_x(t,\tau)_n$. A Fourier coefficient of this (poly)periodic function of t is given by

$$C_x^{\beta}(\tau)_n \triangleq \left\langle C_x(t,\tau)_n\, e^{-i2\pi\beta t} \right\rangle = \sigma_x^{\beta}(\tau)_n, \tag{140}$$

and is called a *cyclic temporal cumulant function* (CTCF). The multivariate relationships between moments and cumulants can be used to express the CTMF and the CTCF in terms of lower-order CTMFs or CTCFs. For example, from Eqs. (134) and (140) we have

$$C_x^{\beta}(\tau)_n = R_x^{\beta}(\tau)_n + \sum_{\substack{P_n \\ p\neq 1}} \left[(-1)^{p-1}(p-1)! \sum_{\boldsymbol{\alpha}^{\dagger}\mathbf{1}=\beta} \prod_{j=1}^{p} R_x^{\alpha_j}(\tau_{\nu_j})_{n_j} \right] \tag{141}$$

and, using Eqs. (123) and (136), we have

$$R_x^{\alpha}(\tau)_n = C_x^{\alpha}(\tau)_n + \sum_{\substack{P_n \\ p \neq 1}} \left[\sum_{\beta^{\dagger} \mathbf{1} = \alpha} \prod_{j=1}^{p} C_x^{\beta_j}(\tau_{\nu_j})_{n_j} \right]. \tag{142}$$

In (141) and (142), α and β are p-dimensional vectors of moment and cumulant cycle frequencies, respectively, and $\mathbf{1}$ is a p-dimensional vector of ones. Each sine-wave component of the TCF, $C_x^{\beta}(\tau)_n e^{i2\pi\beta t}$, is called a *cumulant sine wave* to distinguish it from a moment sine wave. It can be seen from Eqs. (128) and (140) that the CTCF is identical to the (complex-valued) strength of the pure nth-order sine wave with frequency β that is contained in the nth-order delay product $L_x(t,\tau)_n$ and, therefore, β is called a *pure nth-order cycle frequency* (for the particular choice of optional conjugations used in the delay product Eq. (120)). The CTMF in Eq. (125), on the other hand, gives the strength of the entire sine wave with frequency α that is contained in $L_x(t,\tau)_n$, which can be called the *impure nth-order sine wave* with *impure nth-order cycle frequency* α (for the particular set of chosen conjugations).

It should be pointed out that this (i.e., [56]) is the first instance (to the best of our knowledge) in which cumulants have arisen as the solution to a practically motivated problem—namely the problem of pure nth-order sine wave generation—rather than as a mathematical observation concerning the characteristic function.

2. Properties of the Temporal Parameters of HOCS.

a. Signal Selectivity.

Let the time-series $x(t)$ consists of the sum of M statistically independent signals $\{y_m(t)\}_{m=1}^{M}$,

$$x(t) = \sum_{m=1}^{M} y_m(t). \tag{143}$$

Then the addition rule for cumulants can be used to show that the nth-order

TCF for $x(t)$ is the sum of TCFs for $\{y_m(t)\}$,

$$C_x(t,\tau)_n = \sum_{m=1}^{M} C_{y_m}(t,\tau)_n. \tag{144}$$

Thus, the pure nth-order sine waves in the delay products of each $y_m(t)$ add to form the pure nth-order sine wave in the delay product of $x(t)$:

$$C_x^{\beta}(\tau)_n = \sum_{m=1}^{M} C_{y_m}^{\beta}(\tau)_n. \tag{145}$$

The TMF does not, in general, exhibit this useful property. That is, the nth-order TMF for $x(t)$ is not the sum of nth-order TMFs for each $y_m(t)$:

$$R_x(t,\tau)_n \neq \sum_{m=1}^{M} R_{y_m}(t,\tau)_n.$$

An exception is the case of zero-mean signals and $n \leq 3$, for which moments and cumulants are equal.

To illustrate how Eq. (145) can be applied in practice, consider the situation in which $\{y_m(t)\}_{m=1}^{M}$ represent M interfering signals that overlap in time and frequency, but each $y_m(t)$ possesses some distinct pure nth-order cycle frequency, say β_m. Then it follows from Eq. (145) that

$$C_x^{\beta_m}(\tau)_n = C_{y_m}^{\beta_m}(\tau)_n, \quad m = 1, 2, \cdots, M.$$

This indicates that the presence or absence of each of the signals $y_m(t)$ can be detected by measuring (estimating) the CTCFs of $x(t)$ for the pure cycle frequencies $\{\beta_m\}$.

As another application, let $M = 2$, $y_1(t)$ be non-Gaussian, and $y_2(t)$ be Gaussian. Then $C_{y_2}(t,\tau)_n \equiv 0$ for $n \geq 3$ and, from Eq. (144), we have

$$C_x(t,\tau)_n = C_{y_1}(t,\tau)_n, \quad n \geq 3,$$

which indicates the detectability of $y_1(t)$ with no knowledge about $y_2(t)$ except that it is Gaussian.

b. Integrability.

It can be shown that both the CTMF and the CTCF are sinusoidal jointly in the n variables $\tau_1, \cdots, \tau_n$:

$$C_x^\beta(\tau + 1\Delta)_n = C_x^\beta(\tau)_n e^{i2\pi\beta\Delta}, \tag{146}$$

$$R_x^\alpha(\tau + 1\Delta)_n = R_x^\alpha(\tau)_n e^{i2\pi\alpha\Delta}. \tag{147}$$

Hence, these functions are not absolutely integrable with respect to τ. This periodicity suggests that we might reduce the dimension of the functions and retain all the information present in the original functions. Reducing the dimension by one yields the definitions[9]

$$\bar{C}_x^\beta(u)_n \triangleq C_x^\beta([u\ 0]^\dagger)_n \text{ and } \bar{R}_x^\alpha(u)_n \triangleq R_x^\alpha([u\ 0]^\dagger)_n,$$

(where $u = [u_1 \cdots u_{n-1}]$), which are no longer sinusoidal. The value of $C_x^\beta(\tau)_n$ ($R_x^\alpha(\tau)_n$) for any τ can be obtained from the value of $\bar{C}_x^\beta(u)_n$ ($\bar{R}_x^\alpha(u)_n$) by using Eq. (146) ((147)).

It is shown in [59] that the function $\bar{R}_x^\alpha(u)_n$ (the RD-CTMF) is not in general integrable, whereas the function $\bar{C}_x^\beta(u)_n$ (the RD-CTCF) is in general integrable for signals possessing an asymptotic independence property (mixing condition). That is, if the rate of decay of the RD-CTCF $|\bar{C}_x^\beta(u)_n|$ is sufficiently high, then $\bar{C}_x^\beta(u)_n$ is absolutely integrable and, therefore, Fourier transformable, whereas the RD-CTMF is not, in general, Fourie transformable except in a generalized sense that accommodates Dirac delta functions, and products of Dirac delta functions.

D. THE SPECTRAL PARAMETERS OF HOCS

The Fourier transform of $\bar{R}_x^0(u)_2$ (with the second factor conjugated) is the power spectral density (PSD) of $x(t)$ and the Fourier transform of the corresponding second-order cumulant $\bar{C}_x^0(u)_2$ is the PSD of $x(t)$ with its sine-wave components removed, that is, it is the PSD with the spectral lines removed. The Fourier transform of a symmetrized version of $\bar{R}_x^\alpha(u)_2$ (with the second factor conjugated) for $\alpha \neq 0$ is the spectral-correlation

[9] The reason for this particular choice of dimension reduction is made clear in Section IV.D.

density function (the discrete-time counterpart of which is studied in previous sections), and the Fourier transform of a symmetrized version of the corresponding cumulant $\bar{C}_x^\alpha(u)_2$ is the spectral-correlation density function for the signal $x(t)$ with its first-order sine waves removed. Therefore, we could define (by analogy) the spectral parameters of HOCS to be the multidimensional Fourier transforms of $\bar{R}_x^\alpha(u)_n$ and $\bar{C}_x^\alpha(u)_n$, whenever such transforms exist. These transforms are indeed the central spectral parameters of the theory of HOCS, but it is more natural to derive them from a consideration of spectral moments and spectral cumulants; that is, from limiting versions (as bandwidth approaches zero) of moments and cumulants of narrowband spectral components of $x(t)$, and then to show that they can be characterized as Fourier transforms of temporal moments and cumulants.

It is assumed that $x(t)$ is absolutely integrable on finite intervals. We consider the complex envelope of the spectral component of a segment of $x(u)$ that is centered at t and has width T:

$$X_T(t,f) \triangleq \int_{t-T/2}^{t+T/2} x(v)e^{-i2\pi fv}\,dv. \tag{148}$$

The time-average of the set of n variables $\{X_T^{(*)_j}(t,f_j)\}_{j=1}^n$ is defined by

$$S_{x_T}(\boldsymbol{f})_n \triangleq \left\langle \prod_{j=1}^n X_T^{(*)_j}(t,f_j) \right\rangle, \qquad \boldsymbol{f} = [f_1 \cdots f_n]^\dagger, \tag{149}$$

and is assumed for the time being to exist. If we now let the integration time T in Eq. (148) tend to infinity in Eq. (149), we obtain the *spectral moment function* (SMF)

$$S_x(\boldsymbol{f})_n \triangleq \lim_{T\to\infty} S_{x_T}(\boldsymbol{f})_n. \tag{150}$$

However, this limit exists only in a generalized sense that accommodates products of Dirac delta functions. We shall see that Dirac deltas can be avoided by working with the cumulant counterpart of this moment.

It can be shown [58, 59] that the SMF can be characterized in terms of the transform of the RD-CTMF,

$$\bar{S}_x^\alpha(\boldsymbol{f}')_n \triangleq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \bar{R}_x^\alpha(\boldsymbol{u})_n e^{-i2\pi \boldsymbol{u}^\dagger \boldsymbol{f}'}\,d\boldsymbol{u}, \quad \boldsymbol{f}' \triangleq [f_1 \cdots f_{n-1}], \tag{151}$$

in the following way:

$$S_x(f)_n = \begin{cases} \bar{S}_x^{\alpha}(f')_n \delta(f^{\dagger}\mathbf{1} - \alpha), & f^{\dagger}\mathbf{1} = \alpha, \\ 0, & f^{\dagger}\mathbf{1} \neq \alpha, \end{cases} \tag{152}$$

for all nth-order cycle frequencies α of $x(t)$ that correspond to the chosen conjugations, which can be reexpressed as

$$S_x(f)_n = \sum_{\alpha} \bar{S}_x^{\alpha}(f')_n \delta(f^{\dagger}\mathbf{1} - \alpha). \tag{153}$$

Equation (153) reveals that the SMF is a sum of components with impulsive factors. Moreover, we can show that $\bar{S}_x^{\alpha}(f')_n$ can also be a sum of components with impulsive factors and even products of impulses (this is done at the end of this section). Thus, neither the SMF nor the reduced-dimension SMF (RD-SMF) Eq. (151) are well-behaved functions.

The *spectral cumulant function* (SCF) is better behaved than the SMF. The simple cumulant of the variables $\{X_T^{(*)_j}(t, f_j)\}_{j=1}^n$ is given by[10]

$$P_{x_T}(f)_n \triangleq \sum_{P_n} \left[(-1)^{p-1}(p-1)! \prod_{j=1}^{p} S_{X_T}(f_{\nu_j})_{n_j} \right], \tag{154}$$

where f_{ν_j} is the vector of frequencies with subscripts in ν_j. This function is well-defined for finite T since each $S_{X_T}(\cdot)$ is finite. The spectral cumulant function is formally defined to be the limit

$$P_x(f)_n \triangleq \lim_{T \to \infty} P_{x_T}(f)_n. \tag{155}$$

By analogy with the preceding argument for the SMF, it can be shown [59, 58] that the SCF is equivalent to

$$P_x(f)_n = \sum_{\beta} \bar{P}_x^{\beta}(f')_n \delta(f^{\dagger}\mathbf{1} - \beta), \tag{156}$$

where

$$\bar{P}_x^{\beta}(f')_n \triangleq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \bar{C}_x^{\beta}(u)_n e^{-i2\pi u^{\dagger} f'} \, du \tag{157}$$

[10] The cumulant of the variables $\{X_T^{(*)_j}(t, f_j)\}_{j=1}^n$ is actually given by (154) with each $S_{X_T}(f_{\nu_j})_{n_j}$ replaced by a corresponding version given by (149) with $\langle \cdot \rangle$ replaced by $\hat{E}^{\{\alpha\}}\{\cdot\}$. Nevertheless, these more complex expressions reduce to those given by (150) and (155) when $T \to \infty$.

's defined to be the *cyclic polyspectrum* (CP) or *cyclic polyspectral density function.* The transform Eq. (157) does exist in general for signals with asymptotically independent variables such that the reduced-dimension CTCF decays sufficiently rapidly in all directions (cf. Section IV.C.2.b).

It is now easy to show that the RD-SMF $\bar{S}_x^\alpha(\boldsymbol{f}')_n$ can itself contain impulses. From Eq. (142) we have

$$R_x^\alpha(\boldsymbol{\tau})_n = C_x^\alpha(\boldsymbol{\tau})_n + \sum_{\substack{P_n \\ p\neq 1}} \left[\sum_{\boldsymbol{\beta}^\dagger \mathbf{1}=\alpha} \prod_{j=1}^{p} C_x^{\beta_j}(\boldsymbol{\tau}_{\nu_j})_{n_j} \right].$$

By setting $\boldsymbol{\tau} = [\boldsymbol{u}\ 0]$ and Fourier transforming in $\boldsymbol{u}$, we obtain by definition:

$$\bar{S}_x^\alpha(\boldsymbol{f}')_n = \bar{P}_x^\alpha(\boldsymbol{f}')_n + \sum_{\substack{P_n \\ p\neq 1}} \left[\sum_{\boldsymbol{\beta}^\dagger \mathbf{1}=\alpha} \bar{P}_x^{\beta_p}(\boldsymbol{f}'_{\nu_p})_{n_p} \prod_{j=1}^{p-1} \bar{P}_x^{\beta_j}(\boldsymbol{f}'_{\nu_j})_{n_j} \delta(\mathbf{1}^\dagger \boldsymbol{f}_{\nu_j} - \beta_j) \right] \tag{158}$$

where it is assumed that the partitions in Eq. (142) are ordered so that ν_p always contains n. In Eq. (158), each coefficient $\bar{P}_x^{\beta_j}(\cdot)$ is well-behaved, which implies that there are no hidden impulse functions. For the case in which there is no lower-order cyclostationarity (cyclostationarity of order less than n) associated with $\{x^{(*)_j}(t+\tau_j)\}_{j=1}^n$, the sum over P_n in Eq. (158) is zero for $\alpha \neq 0$. If $x(t)$ exhibits cyclostationarity of order less than n so that there is at least one $\boldsymbol{\beta}$ such that $\boldsymbol{\beta}^\dagger \mathbf{1} = \alpha$, then the sum cannot be identically zero as a function of $\boldsymbol{f}'$. Since we know that $\bar{P}_x^\alpha(\boldsymbol{f}')_n$ is well-behaved (contains no impulses), then the RD-SMF must contain impulses or products of impulses and, therefore, so must the SMF.

In this section we have explained that the CP is in general the only well-behaved spectral function in the theory of HOCS. The SMF $S_x(\boldsymbol{f})_n$ and its reduced-dimension version $\bar{S}_x^\alpha(\boldsymbol{f}')_n$ in general contain products of impulses and are, therefore, not well-behaved functions. However, in the special case where the lowest order of cyclostationarity of $x(t)$ is n, the impure nth-order sine waves (with strengths given by the CTMFs) are identical to the pure nth-order sine waves (CTCFs) and, as a result, the nth-order SCF is identical to the nth-order SMF $P_x(\boldsymbol{f})_n = S_x(\boldsymbol{f})_n$, which

results in equality between the CP and the RD-SMF $\bar{P}_x^{\alpha}(f')_n = \bar{S}_x^{\alpha}(f')_n$ In the case in which there is lower-order cyclostationarity, there can be many values of f' such that $\bar{P}_x^{\beta}(f')_n \neq 0$ and either one or more of the weighting functions $\bar{P}_x^{\beta_j}(\cdot)_{n_j}$ in Eq. (158) is zero, or one or more of the Dirac delta functions in Eq. (158) is zero which results in equality between the CP and the RD-SMF for these f'. This fact can be exploited to yield estimators of the cumulant-based CP $\bar{P}_x^{\beta}(f')_n$ that actually estimate the simpler moment-based RD-SMF $\bar{S}_x^{\beta}(f')_n$ [58, 59, 67, 68].

E. APPLICATIONS OF *HOCS*

The problems considered in this section can be formulated in terms of the following two-sensor received-signal model

$$\begin{aligned} x(t) &= s(t) + m_x(t) + \sum_{k=1}^{M} i_k(t) \\ y(t) &= A_0 s(t + d_0) + m_y(t) + \sum_{k=1}^{M} A_k i_k(t + d_k), \end{aligned} \tag{159}$$

where $s(t)$ is the signal of interest (SOI), $m_x(t)$ and $m_y(t)$ are independent white Gaussian noises, the $\{i_k(t)\}$ are signals not of interest (SNOIs), $\{A_k\}$ and $\{d_k\}$ are the relative attenuation factors and delays, respectively, between the signal components in $x(t)$ and $y(t)$. As an example of the kind of problem considered here, suppose that $M = 2$ and the power levels of m_x and m_y are time-varying. Further, assume that the SOI is weak and fourth-order cyclostationary—but not second-order cyclostationary—with period T_0, and the two interferers are second-order cyclostationary with periods T_1 and T_2 such that $T_1 + T_2 = T_0$. The first problem is to detect the presence of the SOI given a finite segment of $x(t)$. The second problem is to determine the parameter d_0 given finite segments of both $x(t)$ and $y(t)$. These problems are difficult to solve using the stationary models of the various signals (which leads to radiometry [12] for the first problem and generalized cross correlation methods [17, 70] for the second) because for detection the nonstationary noise and interference complicates the threshold setting, and for time-delay estimation the interference corrupts the relevant phase

information in the cross spectrum. The theory of higher-order statistics is not helpful here because its parameters are not signal selective: all of the signals contribute to higher-order cumulants for a stationary non-Gaussian signal model [58]. The theory of second-order cyclostationarity (SOCS) is also not helpful because the SOI has no second-order cyclic features (cf. Section IV.A). However, these two tasks can be accomplished by using the higher-order cyclostationarity of the signal of interest.

Before discussing the signal selectivity of cyclic cumulants, we present the nth-order CP for the continuous-time complex-valued pulse-amplitude-modulated signal (cf. Eq. (11)). Specifically, the CP for the set of n variables $\{x^{(*)_j}(t+\tau_j)\}_{j=1}^{n}$ is given by

$$\bar{P}_x^{\beta}(\boldsymbol{f}')_n = \frac{C_{a,n}}{T_0} P((-)_n[\beta - \mathbf{1}^{\dagger}\boldsymbol{f}'])^{(*)_n} \prod_{j=1}^{n-1} P((-)_j f_j)^{(*)_j} e^{i2\pi\beta t_0}, \quad (160)$$

for $\beta = k/T_0$, where $(-)_j$ is the optional minus sign corresponding to the optional conjugation of the jth factor, and $C_{a,n}$ is the nth-order cumulant of the symbol variables. Thus, the CP is a scaled product of pulse transforms.

1. Signal Selectivity.

An important advantage of exploiting cyclostationarity in signal processing tasks is that the cyclic parameters are signal selective in that the parameters associated with the SOI can be estimated from data that also contains noise and SNOIs. As the amount of data grows, the effects of the noise and interference on the parameter estimate decrease. The nature of the signal selectivity properties of higher-order cyclic moments and cumulants is examined next (cf. Section IV.C).

Because the signals and noises in Eq. (159) are assumed to be statistically independent, the TCF for $x(t)$ is given by the simple additive formula

$$C_x(t,\tau)_n = C_s(t,\tau)_n + C_{m_x}(t,\tau)_n + \sum_{k=1}^{M} C_{i_k}(t,\tau)_n,$$

which implies that if β is a pure nth-order cycle frequency for $s(t)$, and is not for any of the other signals, then $C_x^{\beta}(\tau)_n = C_s^{\beta}(\tau)_n$. Similarly, if β_k is a unique pure nth-order cycle frequency for $i_k(t)$, then $C_x^{\beta_k}(\tau)_n = C_{i_k}^{\beta_k}(\tau)_n$.

For moments, the signal selectivity property depends on the cycle frequencies for all of the signals present for all orders $m \leq n$. This can be seen by expressing the CTMF in terms of CTCFs (cf. Eq. (142)):

$$R_x^\alpha(\tau)_n = \sum_{P_n} \left[\sum_{\beta^\dagger \mathbf{1}=\alpha} \prod_{j=1}^{p} C_x^{\beta_j}(\tau_{\nu_j})_{n_j} \right]. \tag{161}$$

If there is a vector β in Eq. (161) such that at least one of its elements is a cycle frequency that is associated with a SNOI, then the CTMF for $x(t)$ will not be equal to the CTMF for $s(t)$. Nevertheless, it can happen that the contributions to the CTMF from SNOIs do not affect the phase of the CTMF but only its magnitude. This can happen, for example, if $n = 4$ and the interference and noise contribute only second-order cumulants for $\beta_j = 0$ to (161). For moments, then, there are two kinds of signal selectivity, depending on what information in the moment is considered useful (magnitude or phase). Because of this potentially troublesome complication, the signal-selectivity properties of cumulants are deemed more useful than those for moments. Thus, the following sections describe algorithms in terms of cumulants, but in most cases, an alternative algorithm can be created by replacing cumulants with moments. The usefulness of these moment-based alternatives depends on the cycle frequencies associated with the SNOIs, which must be evaluated on a case-by-case basis.

2. Weak-Signal Detection.

In this section, the problem of detecting the presence of the signal $s(t)$ in a received data set $x(t)$ as in Eq. (159) is considered. There are several versions of the detection problem that are of interest. The first is called the *general search problem*, in which a data set is analyzed to determine if there are *any* cyclostationary signals present. No information about the received data is assumed to be known in the general search problem. In the second problem, called the *known-cycle-frequency problem*, a specific pure cycle frequency/order pair (β, n_0) is of interest, and it is desired to determine if there is a signal present in the data corresponding to this pair. In the third problem, called the *known-modulation problem*, the modulation format of the signal of interest is known, and hence the cyclic cumulants of

the signal are known (in principle); it is desired to determine the presence or absence of this particular signal.

a. The General Search Problem.

In this problem there is a maximum order N of nonlinearity that is to be used for processing. The goal of the processing is to produce a list of pure nth-order cycle frequencies $\{\beta_n\}$ for each order $n \leq N$. This list $\{\beta_n\}$ for each n characterizes the detectable cyclostationarity of order n (and only n) that is associated with $x(t)$. Thus, these lists are not contaminated by entries that are due to lower-order sine wave interactions, which result in false detections. To accomplish this task, the TCF is estimated for $x(t)$ for each order n. From the estimate of the TCF of order n, the cycle frequencies $\{\beta_n\}$, which are needed for the estimate of the TCF for order $n+1$, can be found. More explicitly, the general search problem can be tackled by using the following algorithm:

0. Let $n = 1$
1. $\hat{C}'_x(t,\tau)_n = L_x(t,\tau)_n - \sum_{\substack{P \\ p \neq 1}} \left[\prod_{j=1}^{p} \hat{C}_x(t,\tau_{\nu_j})_{n_j} \right]$
2. $Y(f) = \mathrm{FFT}_t \{\hat{C}'_x(t,\tau)_n\}$
3. Threshold detect the bins of Y to find $\{\beta_n\}$
4. $\hat{C}_x^{\beta_n}(\tau)_n = \left\langle \hat{C}'_x(t,\tau)_n \, e^{-i2\pi\beta_n t} \right\rangle$
5. $\hat{C}_x(t,\tau)_n = \sum_{\beta_n} \hat{C}_x^{\beta_n}(\tau)_n e^{i2\pi\beta_n t}$
6. $n \rightarrow n+1$; if $n \leq N$ then go to 1.

In step 4, the interval over which the average $\langle \cdot \rangle$ is performed is determined by the amount of data available. If any of the detected cycle frequencies are of particular interest, a cyclic polyspectral analysis can be performed from which the modulation type can possibly be determined [58, 59].

b. The Known-Cycle-Frequency Problem.

In this problem, one or more of the signal's modulation frequencies, such as a keying rate or carrier frequency, is assumed to be known, but the exact

functional form of the CTCF is unknown. The received-signal environment is still assumed to be unknown and, therefore, the general search algorithm is of interest. However, it can be improved for the known-cycle-frequency problem by combining it with a least-squares estimation technique. Let (β, n_0) be the cycle frequency/order pair of interest. Use the general search algorithm up to order $n_0 - 1$. Form $\hat{C}'_x(t,\tau)_{n_0}$, and use a least-squares estimator to detect the presence of the signal of interest using a threshold test on the statistic

$$Y = \left\langle \hat{w}^\dagger \hat{C}'_x(t,\tau)_{n_0} e^{-i2\pi\beta t} \right\rangle = \hat{w}^\dagger \hat{C}^\beta_x(\tau)_{n_0},$$

where

$$\hat{C}'_x(t,\tau)_{n_0} = \left[\hat{C}'_x(t,\tau_1)_{n_0} \cdots \hat{C}'_x(t,\tau_K)_{n_0}\right]^\dagger,$$
$$\hat{C}^\beta_x(\tau)_{n_0} = \left[\hat{C}^\beta_x(\tau_1)_{n_0} \cdots \hat{C}^\beta_x(\tau_K)_{n_0}\right]^\dagger,$$

and where $\hat{w}$ is the unit-norm version of the least-squares weight vector

$$\hat{w} = \arg \min_w \left\langle \left| w^\dagger \hat{C}'_x(t,\tau)_{n_0} - e^{i2\pi\beta t} \right|^2 \right\rangle. \tag{162}$$

The solution to Eq. (162) is $\hat{w} = R^{-1}\hat{C}^\beta_x(\tau)_{n_0}$, where

$$R = \left\langle \hat{C}'_x(t,\tau)_{n_0} \hat{C}'_x(t,\tau)^H_{n_0} \right\rangle,$$

in which H denotes conjugate transpose. Thus, the detection statistic is

$$Y = \hat{C}^\beta_x(\tau)^H_{n_0} R^{-1} \hat{C}^\beta_x(\tau)_{n_0},$$

which is obtained by forming the particular linear combination of data sets

$$\hat{C}'_x(t,\tau_1)_{n_0}, \cdots, \hat{C}'_x(t,\tau_K)_{n_0}$$

that optimally combines the regenerated sine waves with frequency β that are present in each set, and then correlates this composite regenerated sine wave with the stored sine wave $e^{i2\pi\beta t}$.

An alternative is to use the maximum over τ of $|\hat{C}'_x(t,\tau)_{n_0}|^2$ in a threshold test.

c. The Known-Modulation Problem.

In this problem, it is desired to determine if a signal with known modulation type is present. In particular, the CTCF of $s(t)$ for $n = n_0$ and pure cycle frequency β is known. The general search algorithm can be used to remove all lower-order sine waves up to order $n_0 - 1$. Then, from $\hat{C}'_x(t,\tau)_{n_0}$ the CTCF estimate $\bar{C}^{\beta}_{x_T}(u)_{n_0}$ for cycle frequency β can be determined by computing the Fourier coefficient as in Eq. (140). The proposed detection statistic is

$$Y = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \bar{C}'^{\beta}_{x_T}(u)_{n_0} \bar{C}^{\beta}_{s}(u)^{*}_{n_0}\, du.$$

The primary justification for this particular statistic is that when no signal is present with pure nth-order cycle frequency β, then $\bar{C}^{\beta}_{x_T}(u)_{n_0} \to 0$, which implies that $Y \to 0$; when the signal of interest is present, then

$$Y \to \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left|\bar{C}^{\beta}_{s}(u)_{n}\right|^2 du. \tag{163}$$

Thus, Y is an asymptotically noise-free statistic on both the signal-present and signal-absent hypotheses. Furthermore, the integral Eq. (163) is finite (cf. Section IV.C). Hence, this statistic is the natural generalization of the single-cycle detector that exploits SOCS and possesses several optimality properties [11].

3. TIME-DELAY ESTIMATION.

Conventional approaches to the problem of estimating the time-delay (or time difference of arrival (TDOA)) between signal components in data from two sensors are collectively referred to as *generalized cross correlation* (GCC) methods [70]. In the GCC methods, filtered versions of the sensor outputs $x(t)$ and $y(t)$ in Eq. (159) are cross correlated, and the estimate of d_0 is taken to be the location of the peak in the cross-correlation estimate. These methods suffer when spectrally overlapping interferers are present ($M \geq 1$ in Eq. (159)), because each interferer contributes a peak of its own to the cross correlation function. This causes two problems. The first is a resolution problem which, to be solved, requires that the differences in the TDOAs for each of the signals be greater than the widths of the cross correlation functions so that the peaks can be resolved. The second problem is that it is difficult to correctly associate each peak with its corresponding

signal. Both of these problems arise because the GCC methods are not signal selective; they produce TDOA peaks for all the signals in the received data unless they are spectrally disjoint and can, therefore, be separated by filtering.

Signal-selective methods that exploit the SOCS of the desired signal, which is assumed to be unique to that signal, are studied in [17, 18]. These methods have been shown to outperform the GCC methods, and have been shown to produce unbiased estimates with variance that is smaller than the Cramer-Rao lower bound on the variance of TDOA estimators that are based on the assumption that the signal and its environment are stationary. However, these methods are not applicable when there is no SOCS to exploit. In this case, the theory of HOCS can be used to develop signal-selective TDOA estimators. Following the approach in [17, 18] for SOCS, the methodology considered here for HOCS is based on least-squares estimation. The following two examples illustrate the methodology.

The cross cumulant between n-1 time-translates of $x(t)$ and one time-translate of $y(t)$ is defined by

$$C_{xy}(t,\tau)_n \triangleq \text{ Cumulant } \{y(t+\tau_n), \{x^{(*)_j}(t+\tau_j)\}_{j=1}^{n-1}\}.$$

The Fourier coefficient of this cross cumulant for the pure nth-order cycle frequency β for the signal model Eq. (159) (assuming that the noise and interference do not exhibit nth-order cyclostationarity with pure nth-order cycle frequency β) is given by

$$\begin{aligned} C_{xy}^{\beta}(\tau)_n &\triangleq \left\langle C_{xy}(t,\tau)_n e^{-i2\pi\beta t}\right\rangle \\ &= A_0 C_s^{\beta}(\tau + \delta_n d_0)_n, \end{aligned}$$

where δ_n is the unit vector along the nth coordinate, and the conjugates in C_s^{β} match those in C_{xy}^{β}. It is easy to show that the following relations involving RD-CTCFs hold:

$$\bar{C}_{xy}^{\beta}(u)_n = A_0 \bar{C}_s^{\beta}(u - \mathbf{1}d_0)_n e^{i2\pi\beta d_0}, \qquad \bar{C}_x^{\beta}(u)_n = \bar{C}_s^{\beta}(u)_n.$$

This suggests a least-squares fit of a measurement of $\bar{C}_{xy}^{\beta}$ to a measurement of $\bar{C}_x^{\beta}$ over a region G of u-space of interest:

$$\min_{A,d} \int_G \left|\bar{C}_{xy_T}^{\beta}(u)_n - A\bar{C}_{x_T}^{\beta}(u - \mathbf{1}d)_n e^{i2\pi\beta d}\right|^2 du,$$

which leads to the following estimator of the delay d_0:

$$\hat{d}_0 = \arg \max_d \Re \left\{ \int_G \bar{C}^{\beta}_{x_T}(u)_n \bar{C}^{\beta}_{xy_T}(u + \mathbf{1}d)^*_n e^{i2\pi\beta d}\, du \right\}, \tag{164}$$

where $\Re\{\cdot\}$ denotes the real part of its argument. This estimator is a higher-order generalization of the SPECtral COherence Alignment (SPECCOA) algorithm for TDOA estimation [17, 18], which exploits SOCS and has been shown to possess several optimality properties.

As an alternative, cross-sensor measurements can be avoided entirely by noting that

$$\bar{C}^{\beta}_{y}(u)_n = A^n_0 \bar{C}^{\beta}_{s}(u)_n e^{i2\pi\beta d_0}, \quad \bar{C}^{\beta}_{x}(u)_n = \bar{C}^{\beta}_{s}(u)_n,$$

which suggests the following least-squares approach:

$$\hat{d}_0 = \arg \min_{A,d} \int_G \left| \bar{C}^{\beta}_{y_T}(u)_n - A^n \bar{C}^{\beta}_{x_T}(u)_n e^{i2\pi\beta d} \right|^2 du.$$

The estimator for d_0 is given explicitly by

$$\hat{d}_0 = \frac{-1}{2\pi\beta} \text{ angle} \left\{ \int_G \bar{C}^{\beta}_{x_T}(u)_n \bar{C}^{\beta}_{y_T}(u)^*_n\, du \right\}, \tag{165}$$

which is a higher-order generalization of the SOCS-exploiting Cyclic Phase Difference algorithm for TDOA estimation without cross-sensor measurements [17, 18]. It's disadvantage relative to Eq. (164) is that d_0 can be estimated modulo $1/\beta$ only.

V. SUMMARY

In this chapter, we have given an introduction to the fundamental concepts and basic definitions in the theory of cyclostationary signals, and we have surveyed some of the many applications of this theory to signal processing problems. We have presented theory and method for both second-order cyclostationary signals, which is considered to be established but still maturing, and higher-order cyclostationary signals, which is currently being formulated. For readers interested in pursuing any of the topics discussed, we have given references to more in-depth threatments.

We have explained that the properties of cyclostationary signals that enable the design of signal processing algorithms with unique performance capabilities are the regenerative periodicity property, which enables the generation of spectral lines from the signal by subjecting it to nonlinear transformations, and the spectral redundancy property, which enables the extraction of information about the signal in some subbands of its spectrum using measurements from other subbands. We have further explained that these two properties are actually different ways of interpreting a single property, and that this property of cyclostationary signals results in a kind of signal separability that can be used to counteract the masking effects of noise and interference through the use of specially designed signal processing algorithms.

VI. ACKNOWLEDGEMENTS

Much of the work surveyed in this chapter was supported by the following programs: the Office of Naval Research under contract N00014-92-J-1218 (PIs: W. A. Gardner and S. V. Schell); E-Systems, Inc., Greenville Div.; the Army Research Office under contract DAAL03-89-C-0035 sponsored by the U. S. Army Communications Electronics Command Center for Signals Warfare (PI: W. A. Gardner); the National Science Foundation under grant MIP-88-12902 (PI: W. A. Gardner); and ESL, Inc. and the California Sta[illegible] MICRO Program.

References

[1] W. A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory,* Prentice-Hall, Englewood Cliffs, New Jersey (1987).

[2] W. A. Gardner, "Exploitation of Spectral Redundancy in Cyclostationary Signals," *IEEE Signal Processing Magazine* 8, pp. 14–36 (1991).

[3] W. A. Gardner, "An Introduction to Cyclostationary Signals," Chapter 1 in *Cyclostationarity in Communications and Signal Processing,* (W. A. Gardner, ed.) IEEE Press, New York (1993).

[4] W. A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems,* 2nd ed., McGraw-Hill, New York (1990).

[5] R. S. Roberts, W. A. Brown, and H. H. Loomis, Jr., "Computationally Efficient Algorithms for Cyclic Spectral Analysis," *IEEE Signal Processing Magazine* **8**, pp. 38–49 (1991).

[6] W. A. Brown and H. H. Loomis, Jr., "Digital Implementation of Spectral Correlation Analyzers," *IEEE Transactions on Signal Processing* **41**, pp. 703–720 (1993).

[7] R. S. Roberts, W. A. Brown, and H. H. Loomis, Jr., "A Review of Digital Spectral Correlation Analysis: Theory and Implementation," Article 6 in *Cyclostationarity in Communications and Signal Processing,* (W. A. Gardner, ed.) IEEE Press, New York (1993).

[8] H. L. Hurd, "Nonparametric Time Series Analysis for Periodically Correlated Processes," *IEEE Transactions on Information Theory* **35**, pp. 350–359 (1989).

[9] H. L. Hurd and J. Leskow, "Estimation of the Fourier Coefficient Functions and Their Spectral Densities for Φ-mixing Almost Periodically Correlated Processes," *Statistics and Probability Letters* **14**, pp. 299–306 (1992).

[10] D. Dehay and H. L. Hurd, "Representation and Estimation for Periodically Correlated and Almost Periodically Correlated Random Processes," Chapter 6 in *Cyclostationarity in Communications and Signal Processing,* (W. A. Gardner, ed.), IEEE Press, New York (1993).

[11] W. A. Gardner, "Measurement of Spectral Correlation," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **34**, pp. 1111–1123 (1986).

[12] W. A. Gardner, "Signal Interception: A Unifying Theoretical Framework for Feature Detection," *IEEE Transactions on Communications* **40**, pp. 897–906 (1988).

[13] W. A. Gardner and C. M. Spooner, "Signal Interception: Performance Advantages of Cyclic Feature Detectors," *IEEE Transactions on Communications* **40**, pp. 149–159 (1992).

[14] W. A. Gardner and C. M. Spooner, "Detection and Source Location of Weak Cyclostationary Signals: Simplifications of the Maximum-Likelihood Receiver," *IEEE Transactions on Communications* **41**, (1993).

[15] W. A. Gardner, "The Role of Spectral Correlation in Design and Performance Analysis of Synchronizers," *IEEE Transactions on Information Theory* **34**, pp. 1089–1095 (1986).

[16] N. Blachman, "Beneficial Effects of Spectral Correlation on Synchronization," Article 2 in *Cyclostationarity in Communications and Signal Processing,* (W. A. Gardner, ed.), IEEE Press, New York (1993).

[17] W. A. Gardner and C. K. Chen, "Signal-Selective Time-Difference-of-Arrival Estimation for Passive Location of Man-made Signal Sources in Highly Corruptive Environments. Part I: Theory and Method," *IEEE Transactions on Signal Processing* **40**, pp. 1168–1184 (1992).

[18] C. K. Chen and W. A. Gardner, "Signal-Selective Time-Difference-of-Arrival Estimation for Passive Location of Man-made Signal Sources in Highly Corruptive Environments. Part II: Algorithms and Performance," *IEEE Transactions on Signal Processing* **40**, pp. 1185–1197 (1992).

[19] L. Izzo, A. Napolitano, and L. Paura, "Cyclostationarity-Exploiting Methods for Multipath-Channel Identification," Article 3 in *Cyclostationarity in Communications and Signal Processing,* (W. A. Gardner, ed.), IEEE Press, New York (1993).

[20] B. G. Agee, S. V. Schell, and W. A. Gardner, "Spectral Self-Coherence Restoral: A New Approach to Blind Adaptive Signal Extraction," *Proceedings of the IEEE* **78**, pp. 756–767 (1990).

[21] S. V. Schell and W. A. Gardner, "Blind Adaptive Spatio-Temporal Filtering for Wideband Cyclostationary Signals," *IEEE Transactions on Signal Processing* **41**, 1961–1964 (1993).

[22] S. V. Schell, "An Overview of Sensor Array Processing for Cyclostationary Signals," Chapter 3 in *Cyclostationarity in Communications and Signal Processing,* (W. A. Gardner, ed.), IEEE Press, New York (1993).

[23] S. V. Schell and W. A. Gardner, "High-Resolution Direction Finding," in *Handbook of Statistics* **10** (N. K. Bose and C. R. Rao, eds.), North Holland, Amsterdam (1993).

[24] S. V. Schell and W. A. Gardner, "Estimating the Directions of Arrival of Cyclostationary Signals—Part I: Theory and Methods," Technical Report, Department of Electrical and Computer Engineering, University of California, Davis, CA, November 1991.

[25] S. V. Schell, "Exploitation of Spectral Correlation for Signal-Selective Direction Finding," Ph.D. Dissertation, Department of Electrical and Computer Engineering, University of California, Davis, CA, (1990).

[26] S. Roy, J. Yang, and P. S. Kumar, "Joint Transmitter/Receiver Optimization for Multi-User Communications," Article 1 in *Cyclostationarity in Communications and Signal Processing,* (W. A. Gardner, ed.), IEEE Press, New York (1993).

[27] W. A. Gardner and W. A. Brown, "Frequency-Shift Filtering Theory for Adaptive Co-Channel Interference Removal," *Proceedings of the Twenty-Third Asilomar Conference on Signals, Systems, and Computers,* Pacific Grove CA (1989).

[28] J. H. Reed and T. C. Hsia, "The Performance of Time-Dependent Adaptive Filters for Interference Rejection," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **38**, pp. 1373–1385 (1990).

[29] W. A. Gardner, "Cyclic Wiener Filtering: Theory and Method," *IEEE Transactions on Communications* **41**, pp. 151–163 (1993).

[30] W. M. Brelsford, "Probability Predictions and Time Series with Periodic Structure," Ph.D. Dissertation, Johns Hopkins University, Baltimore Maryland (1967).

[31] M. Pagano, "On Periodic and Multiple Regressions," *Ann. Stat.* **6**, pp. 1310–1317 (1978).

[32] A. G. Miamee and H. Salehi, "On the Prediction of Periodically Correlated Stochastic Processes," in *Multivariate Analysis—V,* (P. R. Krishnaiah, ed.), North Holland, Amsterdam (1980).

[33] G. C. Tiao and M. R. Grupe, "Hidden Periodic Autoregressive-Moving Average Models in Time Series Data," *Biometrika* **67**, pp. 365–373 (1980).

[34] H. Sakai, "Circular Lattice Filtering Using Pagano's Method," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **30**, pp. 279–287 (1982).

[35] H. Sakai, "Covariance Matrices Characterization by a Set of Scalar Partial Autocorrelation Coefficients," *Ann. Stat.* **11**, pp. 337–340 (1983).

[36] H. Sakai, "Spectral Analysis and Lattice Filter for Periodic Autoregressive Processes," *Electronics and Communications in Japan* **73**, pp. 9–15 (1990).

[37] H. Sakai, "On the Spectral Density Matrix of a Periodic ARMA Process," *Journal of Time Series Analysis* **12**, pp. 73–82 (1991).

[38] A. V. Vecchia, "Maximum Likelihood Estimation for Periodic Autoregressive Moving Average Models," *Technometrics* **27**, pp. 375–384 (1985).

[39] J. T. B. Obeysekera and J. D. Salas, "Modeling of Aggregated Hydrologic Time Series," *Journal of Hydrology* **86**, pp. 197–219 (1986).

[40] W. K. Li and Y. V. Hui, "An Algorithm for the Exact Likelihood of Periodic Autoregressive Moving Average Models," *Comm. Stat., Simulation Comput.* **17**, pp. 1483–1494 (1988).

[41] P. L. Anderson and A. V. Vecchia, "Asymptotic Results for Periodic Autoregressive Moving Average Models," *Journal of Time Series Analysis* **14**, pp. 1–18 (1993).

[42] E. G. Gladyshev, "Periodically Correlated Random Sequences," *Soviet Math. Dokl.* **2**, pp. 385–388 (1961).

[43] W. A. Gardner and L. E. Franks, "Characterization of Cyclostationary Random Signal Processes," *IEEE Transactions on Information Theory* **21**, pp. 4–14 (1975).

[44] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, 2nd ed., Krieger Publ. Co., Malabar FL (1989).

[45] W. A. Gardner and T. L. Archer, "Exploitation of Cyclostationarity for Identifying the Volterra Kernels of Nonlinear Systems," *IEE Transactions on Information Theory* **39**, 535–542 (1993).

[46] W. A. Gardner and L. Paura, "Identification of Polyperiodic Nonlinear Systems," Technical Report, December 1992, Department of Electrical and Computer Engineering, University of California, Davis.

[47] A. Lender, "The Duobinary Technique for High-Speed Data Transmission," *IEEE Transactions on Communications and Electronics* **82**, pp. 214–218 (1963).

[48] J. G. Proakis, *Digital Communications,* McGraw-Hill, New York (1983).

[49] J. P. A. Albuqurque, O. Shimbo, and L. N. Ngugen, "Modulation Transfer Noise Effects from a Continuous Digital Carrier to FDM/FM Carriers in Memoryless Nonlinear Devices," *IEEE Transactions on Communications* **32**, pp. 337–353 (1984).

[50] A. N. Shiryaev, "Some Problems in the Spectral Theory of Higher-Order Moments – I," *Theory of Probability and Its Applications* **5**, pp. 265–284 (1960).

[51] D. R. Brillinger, "An Introduction to Polyspectra," *Annals of Mathematical Statistics* **36**, pp. 1351–1374 (1965).

[52] D. R. Brillinger and M. Rosenblatt, "Asymptotic Theory of Estimates of k-Th Order Spectra," in *Spectral Analysis of Time Series,* (B. Harris, ed.), Wiley, New York (1967).

[53] D. R. Brillinger and M. Rosenblatt, "Computation and Interpretation of k-Th Order Spectra," in *Spectral Analysis of Time Series,* (B. Harris, ed.), Wiley, New York (1967).

[54] M. Rosenblatt, *Stationary Sequences and Random Fields,* Birkhauser, Boston (1985).

[55] W. A. Gardner, "Spectral Characterization of N-th Order Cyclostationarity," *Proceedings of the IEEE/ASSP Workshop on Spectrum Estimation,* Rochester New York (1990).

[56] W. A. Gardner and C. M. Spooner, "Higher-Order Cyclostationarity, Cyclic Cumulants, and Cyclic Polyspectra," *Proceedings of the International Symposium on Information Theory and its Applications (ISITA),* Honolulu Hawaii (1990).

[57] C. M. Spooner and W. A. Gardner, "An Overview of the Theory of Higher-Order Cyclostationarity," in *Proceedings of the Workshop on Nonstationary Stochastic Processes and Their Applications,* (A. G. Miamee, ed.), World Scientific, Singapore (1992).

[58] C. M. Spooner, "Theory and Application of Higher-Order Cyclostationarity," Ph.D. Dissertation, Department of Electrical and Computer Engineering, University of California, Davis CA (1992).

[59] C. M. Spooner, "Higher-Order Statistics for Nonlinear Processing of Cyclostationary Signals," Chapter 2 in *Cyclostationarity in Communications and Signal Processing,* (W. A. Gardner, ed.), IEEE Press, New York, (1993) (in press).

[60] W. A. Gardner and C. M. Spooner. "The Cumulant Theory of Cyclostationary Time-Series, Part I: Foundation," *IEEE Transactions on Signal Processing* (in press).

[61] C. M. Spooner and W. A. Gardner. "The Cumulant Theory of Cyclostationary Time-Series, Part II: Development and Applications," *IEEE Transactions on Signal Processing* (in press).

[62] C. Corduneanu, *Almost Periodic Functions, 2nd Edition,* Chelsea, New York (1989).

[63] W. A. Gardner and W. A. Brown, "Fraction-of-Time Probability for Time-Series that Exhibit Cyclostationarity," *Signal Processing* **23**, pp. 273–292 (1991).

[64] T. N. Thiele, *The Theory of Observations,* (1903), reprinted in *Annals of Mathematical Statistics* **2**, pp. 165–308 (1931).

[65] V. P. Leonov and A. N. Shiryaev, "On a Method of Calculation of Semi-Invariants," *Theory of Probability and Its Applications* **4**, pp. 319–328 (1959).

[66] M. J. Hinich, "Testing for Gaussianity and Linearity of a Stationary Time Series," *Journal of Time Series Analysis* **3**, pp. 169–176 (1982).

[67] C. M. Spooner and W. A. Gardner, "Estimation of Cyclic Polyspectra," *Proceedings of the Twenty-Fifth Annual Asilomar Conference on Signals, Systems and Computers,* Pacific Grove CA (1991).

[68] C. M. Spooner and W. A. Gardner, "Performance Evaluation of Cyclic Polyspectrum Estimators," *Proceedings of the Twenty-Sixth Annual Asilomar Conference on Signals, Systems and Computers,* Pacific Grove CA (1992).

[69] C. M. Spooner and W. A. Gardner, "Exploitation of Higher-Order Cyclostationarity for Weak-Signal Detection and Time-Delay Estimation," *Proceedings of the Sixth Workshop on Statistical Signal & Array Processing,* Victoria, British Columbia, Canada (1992).

[70] C. H. Knapp and G. C. Carter, "The Generalized Correlation Method for Estimation of Time Delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **24**, pp. 320–327 (1976).