

See discussions, stats, and author profiles for this publication at:
<https://www.researchgate.net/publication/223859277>

Cyclic spectral analysis in practice. Mech Syst Signal Process

ARTICLE *in* MECHANICAL SYSTEMS AND SIGNAL PROCESSING · FEBRUARY 2007

Impact Factor: 2.26 · DOI: 10.1016/j.ymssp.2006.08.007

CITATIONS

79

READS

151

1 AUTHOR:



Jerome Antoni

University of Lyon

194 PUBLICATIONS 2,982 CITATIONS

SEE PROFILE



Available online at www.sciencedirect.com



Mechanical Systems and Signal Processing 21 (2007) 597–630

Mechanical Systems
and
Signal Processing

www.elsevier.com/locate/jnlabr/ymssp

Cyclic spectral analysis in practice

Jérôme Antoni*

Laboratory Roberval of Mechanics, University of Technology of Compiègne, 60205 Compiègne, Cedex, France

Received 29 November 2005; received in revised form 22 August 2006; accepted 23 August 2006

Abstract

This paper addresses the spectral analysis of cyclostationary (CS) signals from a generic point of view, with the aim to provide the practical conditions of success in a wide range of applications, such as in mechanical vibrations and acoustics. Specifically, it points out the similarities, differences and potential pitfalls associated with cyclic spectral analysis as opposed to classical spectral analysis. It is shown that non-parametric cyclic spectral estimators can all be derived from a general quadratic form, which yields as particular cases “cyclic” versions of the smoothed, averaged, and multitaper periodograms. The performance of these estimators is investigated in detail on the basis of their frequency resolution, cyclic leakage, systematic and stochastic estimation errors. The results are then extended to more advanced spectral quantities such as the cyclic coherence function and the Wigner–Ville spectrum of CS signals. In particular an optimal estimator of the Wigner–Ville spectrum is found, with remarkable properties. Several examples of cyclic spectral analyses, with an emphasis on mechanical systems, are finally presented in order to illustrate the value of such a general treatment for practical applications.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Cyclic spectral analysis; Cyclostationary signals; Cyclic spectrum; Spectral correlation; Cyclic coherence function; Wigner–Ville spectrum; Cyclic leakage

1. Introduction

1.1. Objectives

During the last two decades, the theory of cyclostationarity has emerged as a new approach for characterising a certain type of nonstationary signals. It has recently lead to a number of technological breakthroughs essentially in the field of communication and similar advances are likely to arrive shortly in other engineering fields. One major property of cyclostationary (CS) signals is that they can be easily separated from other interfering signals even in the case of overlapping spectral supports. This allows new possibilities to detect, identify, characterise, and handle signals hidden in high levels of stationary noise. Surprisingly, such potentials have remained virtually unknown from engineers and scientists in vibrations and acoustics—notwithstanding some isolated precursory works on the subject—despite the fact that cyclostationarity

*Tel.: +33 3 44 23 45 25; fax: +33 3 44 23 44 77.

E-mail address: antonij@utc.fr.

perfectly describes many vibro-acoustical signals. We believe that one reason that may explain this status quo is the difficulty encountered in estimating cyclic statistics, and especially in the frequency domain. It is the object of this paper to partially fill in this gap, by proposing a somewhat general treatment of cyclic spectral analysis. As far as permitted by place limitation, we have endeavoured to cover the many aspects of the topic and detail, whenever necessary, its important mathematical content.

1.2. Cyclostationary signals

Cyclostationarity extends the class of stationary signals to those signals whose statistical properties change periodically with time. It should be clear right from the beginning that such signals are generally *not periodic but random* in their waveform, yet they are inherently generated by some periodic mechanism. To avoid confusion, the “period” of a CS signal is referred to as its *cycle*. Typical examples include meteorological signals (the cycles corresponding to daily, monthly, or annual periodicities), econometrics signals (seasonal trends), astronomical signals, biomedical signals (ECG, EMG), mechanical signals issuing from rotating machines (periodic modulations of random phenomena by the rotating parts), etc. It actually happens that the set of CS signals pertaining to a given domain of application is often much larger than expected, probably because a cyclical behaviour is the solution adopted by mother nature to reach a state of (statistical) equilibrium.

As opposed to stationary signals, CS signals contain extra information due to their hidden periodicities. It is therefore important to adopt the CS assumption whenever possible, rather than sticking to the conventional stationary assumption—incidentally, the assumption of stationarity is most often an assumption of convenience and facility rather than a realistic one. In the time domain, the extra information afforded by cyclostationarity is carried by the periodical variations of statistical descriptors such as the instantaneous auto-correlation function and its related quantities (instantaneous power, the envelope function, etc.) [1]. Very interestingly, the same information relates in the frequency domain to correlations between spectral components spaced apart by specific frequencies, referred to as *cyclic frequencies* [2]. Taking advantage of these properties obviously leads to more powerful processing than is possible with the stationary approach [3]. Moreover, this extra information most often counter-balances the complication it may involve as compared to the “simple” stationary approach.

1.3. Cyclic spectral analysis

As opposed to general nonstationary signals, CS signals also enjoy a well-understood theory which actually extends all the signal processing tools historically developed for stationary signals. This is because cyclostationarity is a well-defined property, contrary to nonstationarity in the wide sense. In particular, the very powerful spectral analysis of stationary signals finds an interesting generalisation, which is the topic of this paper.

The spectral analysis of random signals is usually recognised to be a difficult ill-posed problem. A fortiori, the same difficulty holds with CS signals. Actually many aspects of the spectral analysis of CS signals were already investigated in the specialised literature, some of which nearly two decades ago [4–7]; however these works remained isolated and mainly addressed to an audience of statisticians. It is the author’s feeling that there is still an existing need to expose the basics of the experimental spectral analysis of CS signals—which we shall coin “cyclic spectral analysis”—to a wider audience. Therefore, the aim of the present paper is to provide an unifying view of the issue that points out the important connections between theory and practice. This should make available to engineers and researchers in transdisciplinary fields—in particular in mechanics—a set of useful guidelines so as to directly and efficiently perform the spectral analysis of CS signals without bothering too much about mathematical technicalities. In particular, we investigate the generalisation of conventional spectral estimators to cyclic analysis, their frequency resolution, their cyclic frequency resolution, their bias, their variance, the optimal setting of their parameters, and we point out the differences and potential pitfalls as compared to conventional spectral analysis. In this sense our contribution is more than simply a review paper as it brings together a number of fully original results in order to provide a more comprehensive view of previously published material. In particular most of the content of Section 3 is new—or

at least generalises earlier works—as well as its interesting connection with the Wigner–Ville spectrum of Section 4. Also the statistical analyses of Section 4 are fully original and are published here for the first time. Another goal of this paper is to demonstrate some of the possibilities of cyclic spectral analysis by means of illustrative examples on various actual signals, with an emphasis on those issued from mechanical systems.

1.4. Organisation of the paper

The paper is organised as follows. In Section 2, the basic definitions of CS signals are reviewed and the theory of their spectral analysis is introduced. In Section 3, a unifying framework is build for cyclic spectral estimation, where conventional spectral estimators are deduced as particular cases of a general quadratic estimator. This general approach then enables a comprehensive treatment of the estimation performance in all cases. In Section 4, the spectral estimators introduced in Section 3 are extended to more advanced spectral quantities such as the *cyclic coherence* and the *Wigner–Ville spectrum* which may prove useful for specific purposes. Finally, several illustrative examples of application are provided in Section 5.

2. CS signals and their descriptors

In the following we will consider sampled signals with Δ denoting their sampling period and n their temporal index: $X(n\Delta) \rightarrow X[n]$.

2.1. Basic definitions

Stricto sensu, a (quasi-) cyclostationarity signal $X[n]$ is a signal whose joint probability density function is a (quasi-)¹ periodic function of time. This entails that any statistical descriptor (mean, RMS value, instantaneous mean frequency, etc.) of signal $X[n]$ is also a (quasi-) periodic function of time. It is important to recall that this definition implies *ensemble* statistics as opposed to *time* statistics as often understood for stationary and ergodic signals.² In particular a non-zero (quasi-) periodic mean ($E[X[n]] \neq 0$) then necessarily implies that signal $X[n]$ contains some periodic components—while at the same time the time-average \bar{X} of the signal may indeed be zero. For simplicity, we shall only consider signals with zero mean and therefore restrict our analysis to purely random signals. This assumption is without loss of generality since efficient techniques exist for centring non-zero mean CS signals—see Ref. [8] for a discussion on the subject.

Example 1. A typical example of a CS signal is the periodically amplitude modulated white noise:

$$X[n] = p[n] \cdot W[n] \quad (1)$$

where $p[n] = p[n + N]$ is a N -periodic function and $W[n]$ is a strict (i.e. stationary) white noise of variance σ_W^2 . A trajectory of signal $X[n]$ is shown in Fig. 1(a), which clearly exhibits cyclic amplitude modulations. This first example makes it clear that a CS signal does not in general have a periodic waveform, but has a random behaviour synchronised on some periodic mechanism, the period N of which is purposely named the *cycle* of the signal in order to avoid confusion with the period of a periodic signal.

Example 2. Another less obvious example of a CS signal is the response of a periodically time-varying system to a stationary excitation. For instance, consider the output of a second-order system excited by a strict white noise:

$$X[n] = a_1[n] \cdot X[n - 1] + a_2[n] \cdot X[n - 2] + W[n] \quad (2)$$

where $a_i[n] = a_i[n + N]$, $i = 1, 2$, are N -periodic autoregressive coefficients and $W[n]$ is as in Example 1. This model may serve as a good approximation to many real-life signals, e.g. in meteorology, econometrics,

¹Quasi-periodicity generalises periodicity by including functions made of a sum of periodicities with incommensurable periods.

²Gardner [4] has developed a “non-probabilistic” approach of CS signals based on fraction-of-time probability densities, i.e. time statistics. However his approach is difficult and will not be addressed here.

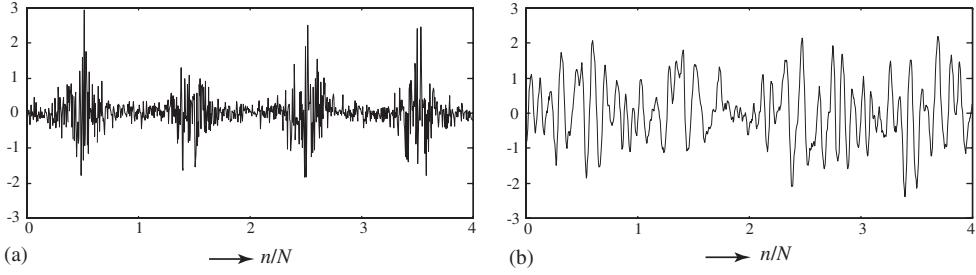


Fig. 1. Examples of cyclostationary signals: (a) periodic amplitude-modulation of a random noise; (b) periodic frequency-modulation of a random noise.

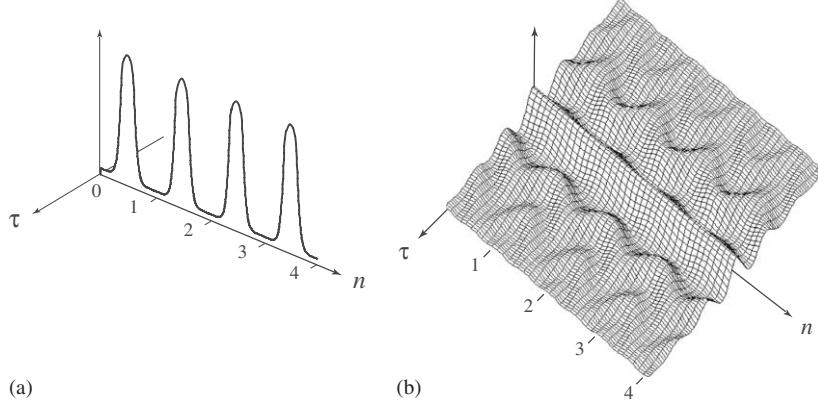


Fig. 2. Instantaneous auto-correlation functions $\mathcal{R}_{2X}[n, \tau]$ of the signals of: (a) Example 1; and (b) Example 2. Note the periodic structure in the time variable n .

astronomy. Fig. 1(b) shows a trajectory of signal $X[n]$ where careful inspection is necessary to distinguish the slight periodic frequency modulations.

2.2. Second-order temporal descriptors

2.2.1. The instantaneous auto-correlation function

In practice it is often sufficient—or convenient—to restrict the description of a random signal to its second-order statistics, i.e. those statistics which relate to energy intensity and interactions. Indeed, second-order statistics are all embodied in the *instantaneous auto-correlation function*

$$\mathcal{R}_{2X}[n, \tau] = \mathbb{E}\{X[n + \beta\tau]X[n - \bar{\beta}\tau]^*\}, \quad \beta + \bar{\beta} = 1, \quad (3)$$

where the parameter β allows for a general formulation of various equivalent definitions found in the literature (typical value of β are $\beta = \frac{1}{2}$ for the symmetric instantaneous auto-correlation function, and $\beta = 1$ or $\beta = 0$ for the asymmetric instantaneous auto-correlation function). Random signals having a periodic instantaneous auto-correlation function

$$\mathcal{R}_{2X}[n, \tau] = \mathcal{R}_{2X}[n + N, \tau] \quad (4)$$

are referred to as *second-order CS* signals. For example, in the case of Example 1,

$$\mathcal{R}_{2X}[n, \tau] = |p[n]|^2 \cdot \sigma_W^2 \cdot \delta[\tau] = \mathcal{R}_{2X}[n + N, \tau] \quad (5)$$

because $p[n] = p[n + N]$ ($\delta[\tau] = 1$ if $\tau = 0$ and is zero elsewhere). This is illustrated in Fig. 2(a). The cyclic auto-correlation function of Example 2 is illustrated in Fig. 2(b). Signals with periodic second-order statistics are also sometimes referred to as *periodically correlated* [5,9].

2.2.2. The cyclic auto-correlation function

By definition the instantaneous auto-correlation function of a (quasi-) CS signal is (quasi-) periodic; therefore it accepts a Fourier series

$$\mathcal{R}_{2X}[n, \tau] = \sum_{\alpha_i \in \mathcal{A}} \mathcal{R}_{2X}[\tau; \alpha_i] e^{j2\pi\alpha_i n \Delta} \quad (6)$$

over the spectrum $\mathcal{A} = \{\alpha_i\}$ of *cyclic frequencies* α_i associated with the non-zero Fourier coefficients

$$\mathcal{R}_{2X}[\tau; \alpha_i] = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{n=-L/2}^{L/2} \mathcal{R}_{2X}[n, \tau] e^{-j2\pi\alpha_i n \Delta}. \quad (7)$$

Because they are functions of the time-lag τ only and are indexed by the cyclic frequency α_i , the Fourier coefficients $\mathcal{R}_{2X}[\tau; \alpha_i]$ are called the *cyclic auto-correlation functions* or *cyclo-correlation functions* of signal X . For Example 1, the cyclic auto-correlation function is given by:

$$\mathcal{R}_{2X}[\tau; \alpha_i] = c_i \cdot \sigma_W^2 \cdot \delta[\tau], \quad (8)$$

where $\alpha_i = i/N$, $i \in \mathbb{Z}$, and c_i is the i th Fourier coefficient of $|p[n]|^2$.

Note that for $\alpha_i = 0$, the cyclic auto-correlation function boils down to the conventional (time-averaged) auto-correlation function $\mathcal{R}_{2X}[\tau]$. In general, the cyclic auto-correlation function enjoys the following symmetry properties

$$\mathcal{R}_{2X}[\tau; \alpha_i]^* = \mathcal{R}_{2X}[-\tau; -\alpha_i] e^{j2\pi(\bar{\beta}-\beta)\alpha_i \tau \Delta} \quad (9)$$

$$= \mathcal{R}_{2X}[\tau; -\alpha_i] \quad \text{if } X \in \mathbb{R}. \quad (10)$$

2.3. Second-order frequency descriptors

2.3.1. The spectral correlation

Second-order temporal descriptors of CS signals³ such as the cyclic auto-correlation function have proven very useful in signal processing applications. For an overview of applications in communication engineering see for instance Refs. [1–3]. However, for practitioners concerned with the analysis of physical measurements, it is often more convenient and natural to apprehend the structure of a CS signal in the frequency domain. Since the instantaneous auto-correlation function $\mathcal{R}_{2X}[n, \tau]$ is a function of two time variable n and τ , its frequency domain counterpart is a 2D Fourier transform with two frequency variables α and f

$$\mathcal{SC}_{2X}(\alpha, f) = \Delta^2 \sum_{n=-\infty}^{\infty} \sum_{\tau=-\infty}^{\infty} \mathcal{R}_{2X}[n, \tau] e^{-j2\pi\alpha n \Delta} e^{-j2\pi f \tau \Delta}. \quad (11)$$

In the above equation $\mathcal{SC}_{2X}(\alpha, f)$ is known as the *spectral correlation* and displays the power distribution of the signal with respect to both the *spectral frequency* f —linked to the waveforms in the signal—and the *cyclic frequency* α —linked to the cyclic evolution of the waveforms. Hence, contrary to the classical spectral analysis of stationary signals, the spectral correlation displays an additional dimension related to the nonstationarity of the signal features. This is reminiscent to time-frequency analysis and, indeed, the quantity

$$\begin{aligned} \mathcal{WV}_{2X}[n, f] &= \int_{-1/2\Delta}^{1/2\Delta} \mathcal{SC}_{2X}(\alpha, f) e^{j2\pi\alpha n \Delta} d\alpha \\ &= \Delta \sum_{\tau=-\infty}^{\infty} \mathcal{R}_{2X}[n, \tau] e^{-j2\pi f \tau \Delta}, \quad \beta = \frac{1}{2} \end{aligned} \quad (12)$$

defines a valid time-frequency representation (surface-wise energy distribution) known as the *Wigner–Ville spectrum* [10,11]. More will be said on the Wigner–Ville spectrum in Section 4.

In order to give an explanation to the terminology “spectral correlation” for $\mathcal{SC}_{2X}(\alpha, f)$, let us consider the Cramér’s spectral decomposition which decomposes a random signal $X[n]$ over a basis of complex

³For notational simplicity we will drop from on the “suffix-” quasi when not explicitly required.

exponential [12]

$$X[n] = \int_{-1/2\Delta}^{1/2\Delta} e^{j2\pi fn\Delta} dX(f), \quad (13)$$

where $dX(f)$ is a spectral increment which accounts for the weight of exponential $e^{j2\pi fn\Delta}$ in signal $X[n]$. Upon inserting Cramér's decomposition (13) into definition (3), it readily comes that $\mathcal{SC}_{2X}(\alpha, f)$ has the alternative expression

$$\mathcal{SC}_{2X}(\alpha, f) df d\alpha = \mathbb{E}\{dX(f + \bar{\beta}\alpha) dX(f - \beta\alpha)^*\}. \quad (14)$$

This expression shows that $\mathcal{SC}_{2X}(\alpha, f)$ is a measure of the *correlation* between the spectral increments $dX(f + \bar{\beta}\alpha)$ and $dX(f - \beta\alpha)$, that is between the frequency components of signal X at frequencies $f + \bar{\beta}\alpha$ and $f - \beta\alpha$ spaced apart by α .

2.3.2. Properties

The spectral correlation enjoys the following symmetries:

$$\mathcal{SC}_{2X}(\alpha, f)^* = \mathcal{SC}_{2X}(-\alpha, f + (\bar{\beta} - \beta)\alpha) \quad (15)$$

$$= \mathcal{SC}_{2X}(-\alpha, -f) \quad \text{if } X \in \mathbb{R} \quad (16)$$

$$\mathcal{SC}_{2X}(-\alpha, f) = \mathcal{SC}_{2X}(\alpha, f + (\bar{\beta} - \beta)\alpha)^* \quad (17)$$

$$= \mathcal{SC}_{2X}(\alpha, -f)^* \quad \text{if } X \in \mathbb{R}. \quad (18)$$

Because $dX(f + \bar{\beta}\alpha)$ and $dX(f - \beta\alpha)$ are periodic functions of frequency f with period $1/\Delta$ (due to sampling in the time domain), the frequency support of the spectral correlation is restricted to the principal quadrant

$$\begin{cases} -1/(2\Delta) \leq f + \bar{\beta}\alpha \leq 1/(2\Delta), \\ -1/(2\Delta) \leq f - \beta\alpha \leq 1/(2\Delta). \end{cases} \quad (19)$$

The other quadrants are obtained by symmetry, as shown in Fig. 3(c).

2.3.3. The cyclic power spectrum

Thus far the so-defined spectral correlation (11) can potentially characterise any nonstationary signal, be it CS or not. In the case of CS signals, the spectral correlation has a very specific structure. Upon inserting the Fourier series (6) into definition (11), it actually comes that

$$\mathcal{SC}_{2X}(\alpha, f) = \sum_{\alpha_i \in \mathcal{A}} \mathcal{S}_{2X}(f; \alpha_i) \delta(\alpha - \alpha_i), \quad (20)$$

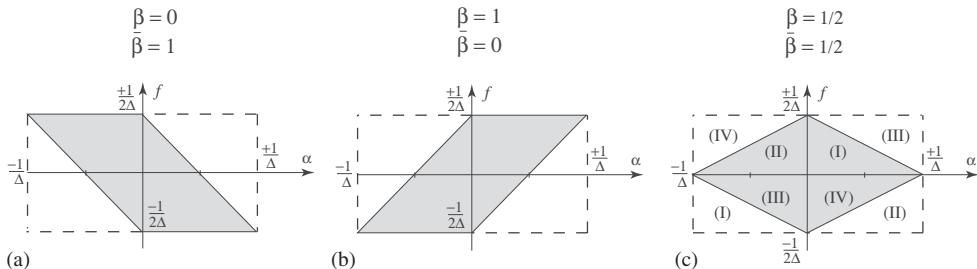


Fig. 3. Frequency support (α, f) of the spectral correlation for: (a) $\beta = 0$; (b) $\beta = 1$; and (c) $\beta = \frac{1}{2}$.

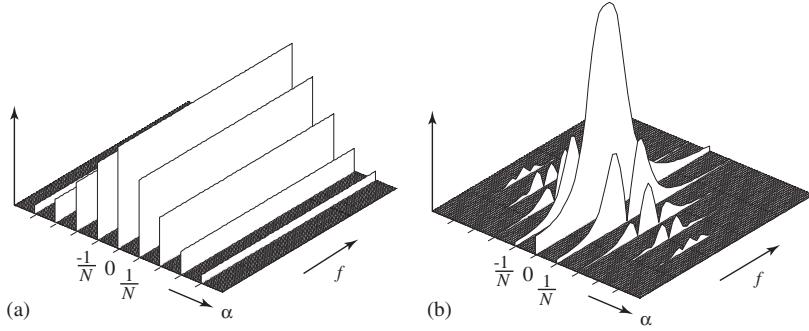


Fig. 4. Spectral correlations $\mathcal{S}_{2X}(\alpha, f)$ of the signals of Examples 1 and 2.

where

$$\mathcal{S}_{2X}(f; \alpha_i) = \Delta \sum_{\tau=-\infty}^{\infty} \mathcal{R}_{2X}[\tau; \alpha_i] e^{-j2\pi f \tau \Delta}. \quad (21)$$

Eq. (20) says that the power of CS signals is distributed along spectral lines parallel to the f -axis and positioned on the cyclic frequencies $\alpha = \alpha_i \in A$. In short, the spectral correlation is continuous in the f -frequency and discrete in the α -frequency. The typical continuous/discrete structure of the spectral correlation of CS signals reveals the central role played by the spectral quantity $\mathcal{S}_{2X}(f; \alpha)$ known as the *cyclic power spectrum* or *cyclo-spectrum*, i.e. the Discrete-Time Fourier Transform (DTFT) of the cyclic auto-correlation function⁴: whereas the spectral correlation $\mathcal{S}_{2X}(\alpha, f)$ is a surface-wise power density (units = power/Hz²), the cyclic power-spectrum $\mathcal{S}_{2X}(f; \alpha)$ is a line-wise power density (units = power/Hz) which is a function of the f -frequency only and is indexed by the cyclic frequency α_i . As a matter of fact, the difference between the spectral correlation and the cyclic power-spectrum is probably made even clearer through the alternative definitions

$$\mathcal{S}_{2X}(\alpha, f) = \lim_{L \rightarrow \infty} \mathbb{E}\{X_L(f + \bar{\beta}\alpha) X_L(f - \beta\alpha)^*\} \quad (22)$$

and

$$\mathcal{S}_{2X}(f; \alpha_i) = \lim_{L \rightarrow \infty} \frac{1}{L\Delta} \mathbb{E}\{X_L(f + \bar{\beta}\alpha_i) X_L(f - \beta\alpha_i)^*\} \quad (23)$$

where $X_L(f) = \Delta \sum_{n=0}^{L-1} X[n] e^{-j2\pi fn\Delta}$ is the DTFT of signal $X[n]$ over L points. Either definition (20) or (23) makes it clear that the physical interpretation of the cyclic power spectrum is similar to that of the conventional power spectrum of stationary signals. Indeed, the cyclic power-spectrum equals the classical power spectrum (i.e. the Fourier transform of $\mathcal{R}_{2X}[n; 0]$) at $\alpha_i = 0$

$$\mathcal{S}_{2X}(f; 0) \equiv \mathcal{S}_{2X}(f). \quad (24)$$

As an illustration to the relationship between the spectral correlation and the cyclic power spectrum, let us consider again the signal of Example 1. Its spectral correlation reads

$$\mathcal{S}_{2X}(\alpha, f) = \Delta \sum_{\alpha_i \in \mathcal{A}} c_i \cdot \sigma_W^2 \cdot \delta(\alpha - \alpha_i) \quad (25)$$

and therefore its cyclic power-spectrum $\mathcal{S}_{2X}(f; \alpha_i) = \Delta \cdot c_i \cdot \sigma_W^2$ is a constant. This is illustrated in Fig. 4(a). The spectral correlation of the signal of Example 2 is shown in Fig. 4(b).

Finally, the cyclic power-spectrum shares the same symmetry properties as the spectral correlation after substituting $\mathcal{S}_{2X}(f; \alpha)$ for $\mathcal{S}_{2X}(\alpha, f)$ by in Eqs. (15)–(19).

⁴This quantity is sometimes referred to as the “spectral correlation”. In order to avoid confusion with the quantity defined by Eq. (11), we discourage this usage.

Table 1

Connections between 2nd order statistical descriptors of CS signals^a

<i>Instantaneous correlation</i>		\Rightarrow	<i>Cyclic correlation</i>
$\leftarrow \mathcal{F}_{\tau \rightarrow f}$	<i>function</i>	$\mathcal{F} \mathcal{S}_{n \rightarrow \alpha_i}$	<i>function</i>
	$\mathcal{R}_{YX}[n, \tau] \quad (\text{U}^2)$		$\mathcal{R}_{YX}[\tau; \alpha_i] \quad (\text{U}^2)$
<i>Wigner–Ville</i>			
<i>spectrum</i>	$\Downarrow 2\mathcal{F}_{\frac{n}{\tau} \rightarrow \frac{\alpha}{f}}$		$\Downarrow \mathcal{F}_{\tau \rightarrow f}$
$\mathcal{W}\mathcal{V}_{YX}[n, f] \quad (\text{U}^2/\text{Hz})$			
$\searrow \mathcal{F}_{n \rightarrow \alpha}$	<i>Spectral correlation</i>	\Leftarrow	<i>Cyclic spectrum</i>
	$\mathcal{SC}_{YX}(\alpha, f) \quad (\text{U}^2/\text{Hz}^2)$		$\mathcal{S}_{YX}(f; \alpha_i) \quad (\text{U}^2/\text{Hz})$

^a \mathcal{F} stands for “Fourier transform”, $\mathcal{F}\mathcal{S}$ for “Fourier series”, and U for the units of the signals.

2.4. Jointly CS signals

So far definitions of temporal and frequency descriptors have been introduced with respect to a single signal $X[n]$. Of particular interest is when these definitions are extended to couples of CS signals. Two signals $Y[n]$ and $X[n]$ are said to be *jointly CS* if their cross-statistics are quasi-periodic functions of time. In particular, this implies a quasi-periodic instantaneous cross-correlation function, from which *cyclic cross-correlation functions* and *cyclic cross-spectra* can be defined by simply replacing subscript $2X$ by subscript YX in all previous definitions. From now on, in an effort to be general, we will consider cross-statistics only—from which auto-statistics will immediately follow by setting $Y = X$. We will also use the shortened labels “cyclic spectrum”, “cyclic correlation”, and so on, wherein the missing suffix “cross-” or “auto-” will be understood from the context.

2.5. Discussion

This brief introduction to CS signals and their descriptors did not pretend to be exhaustive—interested readers will find more complete introductions to the theory of CS signals in the excellent references [1–3] for instance. One objective was to highlight the tight connections that exist between the various descriptors encountered in the context of CS signals. These are summarised in Table 1. Of particular interest among these descriptors is the cyclic spectrum which provides an elegant generalisation to the classical spectrum of stationary signals. Into addition, the cyclic spectrum forms the core from which more advanced descriptors such as the spectral correlation $\mathcal{SC}_{YX}(\alpha, f)$, the cyclic coherence function $\gamma_{YX}(f; \alpha)$ or the Wigner–Ville spectrum $\mathcal{W}\mathcal{V}_{YX}[n, f]$ can be constructed (see Section 4). The cyclic spectrum contains the same information as the instantaneous correlation function and the cyclic correlation function, but likely in a more explicit way as is often the case for spectral descriptors as opposed to temporal descriptors of random signals. For all these reasons, the next section will focus on estimating the cyclic spectrum only.

3. Cyclic spectral estimation: A unifying approach

Although the estimation of the cyclic correlation function is not so much a problem—it has been investigated in reference [13] for instance—this is not the case for the cyclic spectrum. As well-known in spectral analysis, the finite-sample version of neither formula (21) nor (23) can provide a consistent estimator (i.e. whose variance tends to zero as the amount of available data becomes large); thus other formulations must be devised [14]. Although some early papers have been published on the subject [4,5], none of them show specifically how the cyclic spectrum can be estimated from extended versions of the conventional and simple non-parametric spectral estimators commonly used in commercial data analysers and signal processing softwares. It is the object of this section to fill in this gap by providing an unifying framework which encompasses conventional spectral methods ($\alpha = 0$) as particular cases.

3.1. A general quadratic form

The issue at hand is to find a valid non-parametric estimator $\hat{\mathcal{S}}_{YX}(f; \alpha; L)$ of the cyclic spectrum $\mathcal{S}_{YX}(f; \alpha)$ given two finite-length signals $\{Y[n]\}_{n=0}^{L-1}$ and $\{X[n]\}_{n=0}^{L-1}$. We claim that most conventional estimators can be deduced from the general quadratic form

$$\hat{\mathcal{S}}_{YX}(f; \alpha; L) = \Delta \sum_{p=0}^{L-1} \sum_{q=0}^{L-1} Q_L[p, q] Y[p] X[q]^* e^{-j2\pi(f+\bar{\beta}\alpha)p\Delta} e^{j2\pi(f-\beta\alpha)q\Delta}, \quad (26)$$

where Q_L is a suitably chosen positive semi-definite kernel. Inserting Cramér's spectral decomposition (13) into definition (26), the proposed quadratic estimator also has expression

$$\hat{\mathcal{S}}_{YX}(f; \alpha; L) = \frac{1}{\Delta} \int \int_{-1/2\Delta}^{1/2\Delta} \mathcal{Q}_L(\lambda, \eta) dY(f + \bar{\beta}\alpha - \lambda) dX(f - \beta\alpha - \lambda + \eta)^*, \quad (27)$$

where

$$\mathcal{Q}_L(\lambda, \eta) = \Delta^2 \sum_{p=0}^{L-1} \sum_{q=0}^{L-1} Q_L[p, q] e^{-j2\pi\lambda(p-q)\Delta} e^{-j2\pi\eta q\Delta} \quad (28)$$

is the double DTFT of kernel Q_L . The latter expression reveals that the quadratic estimator $\hat{\mathcal{S}}_{YX}(f; \alpha; L)$ may be interpreted on the basis of definition (14) where the theoretical expected value operator $\mathbb{E}\{\cdot\}$ is replaced by convolving (smoothing) the product $dY(f + \bar{\beta}\alpha) dX(f - \beta\alpha)^*$ with the 2D kernel $\mathcal{Q}_L(\lambda, \eta)$. There are an infinite number of degrees of freedom in designing the smoothing kernel $\mathcal{Q}_L(\lambda, \eta)$ but it must fulfil certain calibration constraints to be mentioned later. Different kernels obviously result in different spectral estimators, the most typical of which are now reviewed.

3.2. Particular cases

3.2.1. The cyclic periodogram

The cyclic periodogram estimator is simply

$$\hat{\mathcal{S}}_{YX}^{(P)}(f; \alpha; L) = \frac{1}{L\Delta} Y_L(f + \bar{\beta}\alpha) X_L(f - \beta\alpha)^* \quad (29)$$

with $Y_L(f) = \Delta \sum_{n=0}^{L-1} y[n] e^{-j2\pi f n}$ the DTFT of sequence $\{Y[n]\}_{n=0}^{L-1}$ and similarly for $X_L(f)$. The cyclic periodogram is obtained from (26) wherein

$$Q_L[p, q] = \frac{1}{L} \Leftrightarrow \mathcal{Q}_L(\lambda, \eta) = \Delta^2 \cdot L \cdot D_L^\Delta(\lambda) D_L^\Delta(\eta - \lambda) e^{-j\pi\eta\Delta(L-1)} \quad (30)$$

with $D_L^\Delta(\lambda) = L^{-1} \sin(\pi\lambda\Delta L) / \sin(\pi\lambda\Delta)$ the Dirichlet kernel. Kernel $\mathcal{Q}_L(\lambda, \eta)$ is illustrated in Fig. 5(a). The cycle periodogram is the most simple estimator possible. It turns out to be a direct transcription of definition (23) to finite-length signals. However, it is well-known that the periodogram is not a consistent estimator (its variance does not converge to zero as integration time L becomes large) and therefore cannot be used as such.

3.2.2. The smoothed cyclic periodogram

The smoothed (cyclic) periodogram is classically the approach adopted to make the (cyclic) periodogram consistent. When $\alpha = 0$ it is also known as the “lag-window” or Blackman and Tukey’s method. Let $\{g[\tau]\}_{\tau=-L+1}^{L-1}$ be a positive-definite and symmetric lag-window $g[\tau] = g[-\tau]$ and such that $g[0] \geq g[\tau]$. Then the smoothed cyclic periodogram is

$$\hat{\mathcal{S}}_{YX}^{(B)}(f; \alpha; L) = \Delta \sum_{\tau=-L+1}^{L-1} g[\tau] \cdot \hat{\mathcal{R}}_{YX}[\tau; \alpha; L] e^{-j2\pi f \tau \Delta}, \quad (31)$$

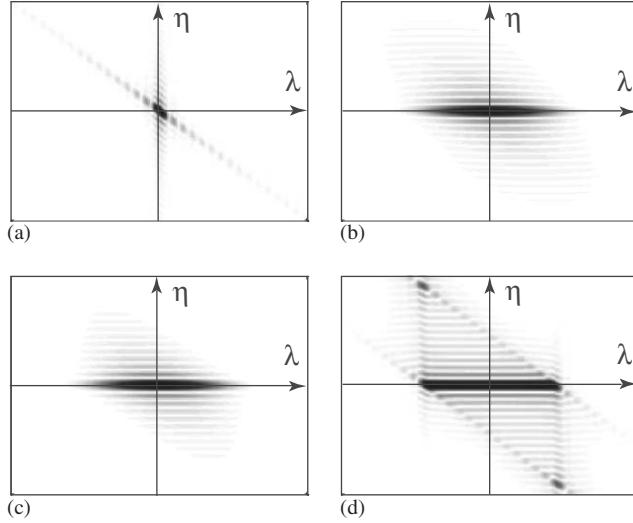


Fig. 5. Magnitude of the smoothing kernel $\mathcal{Q}_L(\lambda, \eta)$ in case of: (a) the cyclic periodogram; (b) the smoothed cyclic periodogram; (c) the averaged cyclic periodogram; and (d) the multitaper cyclic periodogram.

where

$$\widehat{\mathcal{R}}_{YX}[\tau; \alpha; L] = \frac{1}{L} \sum_{n=n_1(\tau)}^{n_2(\tau)} Y[n + \tau] X[n]^* e^{-j2\pi n \Delta(\alpha + \bar{\beta}\tau)} \quad (32)$$

with $n_1(\tau) = \max(0, -\tau)$, $n_2(\tau) = \min(L - 1, L - 1 - \tau)$ is the empirical cyclic correlation function. The smoothed cyclic periodogram is obtained from (26) wherein

$$\mathcal{Q}_L[p, q] = \frac{g[p - q]}{L} \Leftrightarrow \mathcal{Q}_L(\lambda, \eta) = \Delta^2 \cdot L \cdot \int_{-1/2\Delta}^{1/2\Delta} G(\lambda - \gamma) D_L^\Delta(\gamma) D_L^\Delta(\eta - \gamma) e^{-j\pi\eta\Delta(L-1)} d\gamma \quad (33)$$

with $G(f) = \Delta \sum_{\tau=-L+1}^{L-1} g[\tau] e^{-j2\pi f \tau \Delta}$ the “spectral window” (a useful approximation to $\mathcal{Q}_L(\lambda, \eta)$ is simply $\Delta G(\lambda) D_L^\Delta(\eta) e^{-j\pi\eta\Delta(L-1)}$ when the temporal dispersion of $g[\tau]$ is very small as compared to L). Kernel $\mathcal{Q}_L(\lambda, \eta)$ is illustrated in Fig. 5(b). The smoothed cyclic periodogram is a direct transcription of formula (21) where consistency is achieved by forcing the empirical cyclic correlation function $\widehat{\mathcal{R}}_{YX}[\tau; \alpha; L]$ to decay rapidly to zero by tapering it with the lag-window $g[n]$. As well-known, this corresponds in the frequency domain to a convolution of the (cyclic) periodogram (29) with the spectral window $G(f)$, viz $\hat{\mathcal{S}}_{YX}^{(B)}(f; \alpha; L) = \hat{\mathcal{S}}_{YX}^{(P)}(f; \alpha; L) * G(f)$ —note that similarly $\mathcal{Q}_L^{(B)}(\lambda, \eta) = \mathcal{Q}_L^{(P)}(\lambda, \eta) * G(f)$ —hence the name “smoothed” (cyclic) periodogram.

3.2.3. Daniell's estimator

As a particular case to the smoothed (cyclic) periodogram is Daniell's estimator where smoothing is directly applied on the (cyclic) periodogram $\hat{\mathcal{S}}_{YX}^{(P)}(f; \alpha; L)$ by convolution with a rectangular spectral window $G(f) = B^{-1} \mathbf{1}_{[-B/2; +B/2]}$ of bandwidth B Hz. This corresponds to setting

$$g[\tau] = \frac{\sin(B\pi\tau\Delta)}{(B\pi\tau\Delta)} = \text{sinc}(B\tau\Delta) \quad (34)$$

into formula (31).

3.2.4. The averaged cyclic periodogram

The averaged (cyclic) periodogram has probably become the most popular technique in recent years due its high computational efficiency. When $\alpha = 0$ it is also known as the “weighted overlapped segment averaging”

or Welch's method. Let $\{w[n]\}_{n=0}^{N_w-1}$ be a positive and smooth N_w -long data-window and let $w_k[n] = w[n - kR]$ be its shifted version by R samples so that $w_k[n]y[n]$ selects a segment of $y[n]$ at times $kR, \dots, kR + N_w + 1$. The increment R is set between 1 and N_w so as to allow possible overlap between adjacent segments. Then the averaged cyclic periodogram is

$$\hat{\mathcal{S}}_{YX}^{(W)}(f; \alpha; L) = \frac{1}{K\Delta} \sum_{k=0}^{K-1} Y_{N_w}^{(k)}(f + \bar{\beta}\alpha) X_{N_w}^{(k)}(f - \beta\alpha)^*, \quad (35)$$

where $Y_{N_w}^{(k)}(f) = \Delta \sum_{n=kR}^{kR+N_w-1} w_k[n]y[n]e^{-j2\pi fn\Delta}$ is the short-time DTFT of the k th weighted sequence $\{w_k[n]y[n]\}_{n=kR}^{kR+N_w-1}$ and similarly for $X_{N_w}^{(k)}(f)$, and $K = \lfloor (L - N_w)/R \rfloor + 1$ (where $\lfloor x \rfloor$ stands for the greatest integer smaller than or equal to x). The averaged cyclic periodogram is obtained from (26) wherein

$$Q_L[p, q] = \frac{1}{K} \sum_{k=0}^{K-1} w_k[p]w_k[q] \Leftrightarrow \mathcal{Q}_L(\lambda, \eta) = W(\lambda)W(\lambda - \eta)^* D_K^{R\Delta}(\eta)e^{-j\pi\eta R\Delta(K-1)} \quad (36)$$

with $W(f) = \Delta \sum_{n=0}^{N_w-1} w[n]e^{-j2\pi fn\Delta} = \mathcal{F}\{w[n]\}$ the DTFT of $w[n]$ and $D_K^{R\Delta}(\eta) = K^{-1} \sin(\pi\eta R\Delta K)/\sin(\pi\eta R\Delta)$. Kernel $\mathcal{Q}_L(\lambda, \eta)$ is illustrated in Fig. 5(c). Formula (35) can be very efficiently implemented by means of the FFT algorithm by imposing N_w to be a power of 2. By comparing Eqs. (29) and (35), it is clear that the averaged (cyclic) periodogram achieves consistency by averaging K short-time (cyclic) periodograms.

3.2.5. Bartlett's estimator

Bartlett's estimator is a special case of the averaged (cyclic) periodogram where the data-windows are not allowed to overlap, i.e. where $R = N_w$. It implies a smaller number K of short-time periodograms to average and therefore has less variance reduction.

3.2.6. The multitaper cyclic periodogram

The multitaper (cyclic) periodogram, also known as Thomson's method, has been proposed in an attempt to find a spectral estimator with optimal properties (minimum local bias for a given spectral resolution). It is in principle very similar to the averaged cyclic periodogram except that the short-term data-windows are replaced by L -long functions $\{v_k[n]\}_{n=0}^{L-1}$, $k = 0, \dots, K - 1$, which form an orthogonal basis (the data-tapers). Specifically,

$$\hat{\mathcal{S}}_{YX}^{(T)}(f; \alpha; L) = \frac{1}{\Delta} \sum_{k=0}^{K-1} \beta_k Y_L^{(k)}(f + \bar{\beta}\alpha) X_L^{(k)}(f - \beta\alpha)^* \quad (37)$$

with $Y_L^{(k)}(f) = \Delta \sum_{n=0}^{L-1} v_k[n]y[n]e^{-j2\pi fn\Delta} = \mathcal{F}\{v_k[n]y[n]\}$ the DTFT of the weighted sequence $\{v_k[n]y[n]\}_{n=0}^{L-1}$ and similarly for $X_L^{(k)}(f)$, and β_k the weight assigned to the k th taper. Thomson originally proposed to use discrete prolate spheroidal functions for the $\{v_k[n]\}$; however this turns out to be relevant only for nearly white signals. In the more general case, it has been shown that a much better choice is to use the sine tapers

$$v_k[n] = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi k(n+1)}{L}\right), \quad n = 0, \dots, L-1 \quad (38)$$

which yields nearly minimum bias estimates and are also more efficient from the computational point of view [15]. Just as with the averaged (cyclic) periodogram, the consistency of the multitaper (cyclic) periodogram is achieved by increasing the number of averages K . The multitaper cyclic periodogram is obtained from (26) wherein

$$Q_L[p, q] = \sum_{k=0}^{K-1} \beta_k v_k[p]v_k[q] \Leftrightarrow \mathcal{Q}_L(\lambda, \eta) = \sum_{k=0}^{K-1} \beta_k V_k(\lambda) V_k(\lambda - \eta)^* \quad (39)$$

with $V_k(f) = \Delta \sum_{n=0}^{L-1} v_k[n]e^{-j2\pi fn\Delta} = \mathcal{F}\{v_k[n]\}$ the DTFT of $v_k[n]$. Kernel $\mathcal{Q}_L(\lambda, \eta)$ is illustrated in Fig. 5d.

3.2.7. Discussion

This section has suggested that classical non-parametric estimators of the cyclic spectrum can all be devised from a general quadratic form parameterised with some smoothing kernel Q_L . Consideration of specific structures of Q_L is verified to produce valid extensions to the CS case of the conventional estimators used in empirical spectral analysis of stationary signals. This proves that cyclic spectral analysis can benefit from the very efficient implementations of some of these estimators, in particular the averaged cyclic periodogram which as far as we know has rarely been advocated in the CS literature. As a matter of fact, early works on cyclic spectral analysis seem to have mainly focused on the smoothed cyclic periodogram whose statistical analysis is probably simpler [5,7], but yet is computationally less attractive when dealing with long signals. Conversely, for short signals the multitaper cyclic periodogram offers an interesting alternative in applications where minimum bias estimates are required. Obviously, many other spectral estimators can be designed on the same lines according to their fields of application by suitably designing kernel Q_L . In this view, it now remains to state which constraints must fulfil kernel Q_L such as to yield a valid spectral estimate in terms of: (i) low estimation bias; (ii) accurate frequency resolution, and (iii) low estimation variance.

3.3. Bias analysis

For the proposed quadratic estimator to have a small bias, it must hold that $\mathbb{E}\{\hat{\mathcal{S}}_{YX}(f; \alpha; L)\} \simeq \mathcal{S}_{YX}(f; \alpha)$ as closely as possible. Combining Eqs. (27), (14), and (20), it comes that

$$\mathbb{E}\{\hat{\mathcal{S}}_{YX}(f; \alpha; L)\} = \frac{1}{\Delta} \sum_{\alpha_i \in \mathcal{A}} \int_{-1/2\Delta}^{1/2\Delta} \mathcal{Q}_L(\lambda, \alpha - \alpha_i) \mathcal{S}_{YX}(f - \lambda + \bar{\beta}(\alpha - \alpha_i); \alpha_i) d\lambda \quad (40)$$

wherein the smoothing kernel Q_L must be designed such that the right-hand side equals $\mathcal{S}_{YX}(f; \alpha)$ as closely as possible. This implies two conditions to be met, which we now investigate in detail.

3.3.1. Condition I (power calibration)

The first condition is to impose that the leading term in the right-hand side of (40) indexed by $\alpha_i = \alpha$ equals $\mathcal{S}_{YX}(f; \alpha)$. Although this is not possible in general—because $\mathcal{Q}_L(\lambda, 0)$ cannot be a Dirac—a strong requirement is to force the equality at least when $\mathcal{S}_{YX}(f; \alpha)$ is a constant function in the f -frequency (i.e. in the case of CS white noise). This yields the condition:

$$\frac{1}{\Delta} \int_{-1/2\Delta}^{1/2\Delta} \mathcal{Q}_L(\lambda, 0) d\lambda = 1 \Leftrightarrow \sum_{p=0}^{L-1} Q_L[p, p] = 1 \quad (41)$$

which entails the well-known calibrations used in the spectral analysis of stationary signals:

- smoothed cyclic periodogram: $g[0] = 1$;
- averaged cyclic periodogram: $\|w\|^2 = \sum_n w[n]^2 = 1$;
- multitaper cyclic periodogram: $\sum_k \beta_k = 1$ (where the property $\|v_k\|^2 = 1$ has been used).

3.3.2. Condition II (cyclic leakage minimisation)

The second condition is to impose that the terms indexed by $\alpha_i \neq \alpha$ in Eq. (40) are zero. When estimating the cyclic spectrum at cyclic frequency α , such terms may be interpreted as interferences resulting from the action of the smoothing kernel $\mathcal{Q}_L(\lambda, \eta)$ on the other cyclic spectra positioned at cyclic frequencies $\alpha_i \neq \alpha$. The phenomena is akin to *energy leakage in the cyclic frequency* and is illustrated in Fig. 6; we shall refer to it as *cyclic leakage* as first discovered in Ref. [4]. It is in general impossible to make these interferences exactly zero, however it is still possible to seek kernel Q_L such as to minimise them. To see this, let us consider again the situation where $\mathcal{S}_{YX}(f; \alpha)$ is a constant function in the f -frequency, so that condition II becomes that of making function

$$B_Q(\eta) = \frac{1}{\Delta} \int_{-1/2\Delta}^{1/2\Delta} \mathcal{Q}_L(\lambda, \eta) d\lambda \quad (42)$$

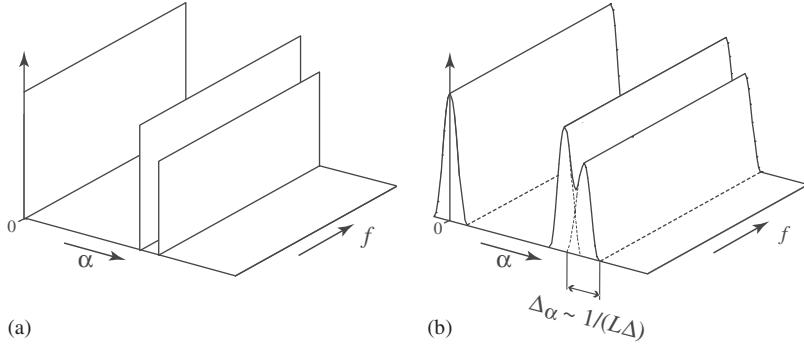


Fig. 6. Illustration of the effect of cyclic leakage.

as close as possible to zero for *any* $\eta \neq 0$. Together with condition I, this boils down to seeking kernel Q_L such as to minimise the bandwidth of $|B_Q(\eta)|$. Let us investigate how this condition is treated in each of the particular cases listed above.

3.3.2.1. Smoothed cyclic periodogram.

In the case of the smoothed cyclic periodogram

$$|B_Q(\eta)| = |D_L^\Delta(\eta)|. \quad (43)$$

In the vicinity of $\eta = 0$, the Dirichlet kernel $D_L^\Delta(\eta)$ has a central lobe of bandwidth $1/(L\Delta)$; thus the following rule applies:

Proposition 1. *For the smoothed cyclic periodogram, Condition II requires that the acquisition time be long enough so that $L\Delta$ is significantly larger than the inverse of the minimum α -spacing—say Δ_α^{\min} —to be resolved, that is $L\Delta \gg 1/\Delta_\alpha^{\min}$.*

3.3.2.2. Averaged cyclic periodogram.

In the case of the averaged cyclic periodogram:

$$|B_Q(\eta)| = \frac{1}{\Delta} |D_K^{R\Delta}(\eta)| \cdot |W_2(\eta)|, \quad (44)$$

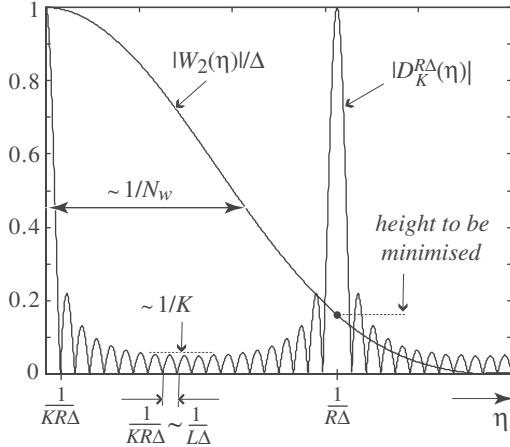
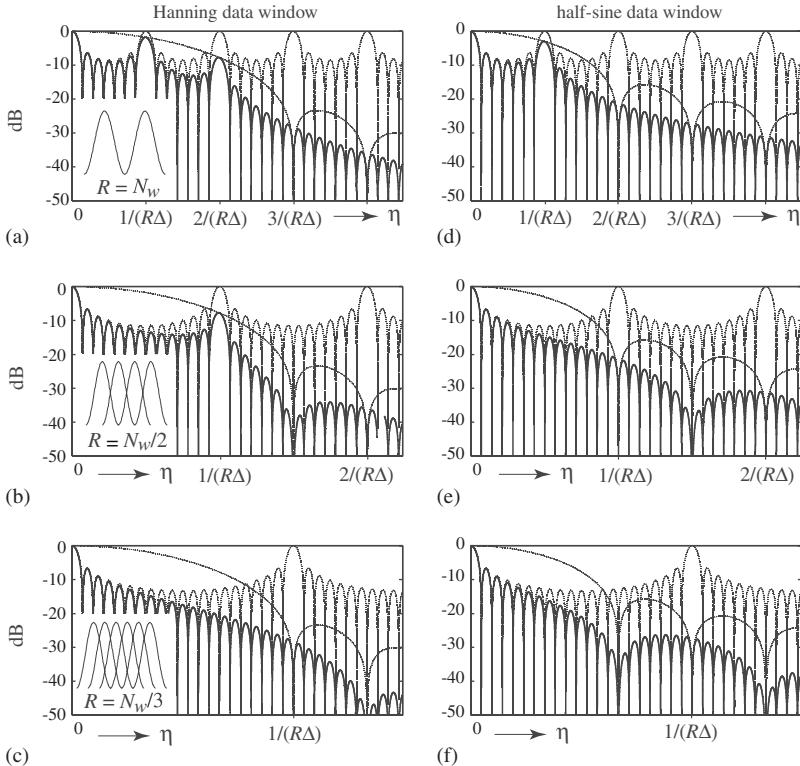
where $W_2(f) = \Delta \sum_{n=0}^{N_w-1} w[n]^2 e^{-j2\pi fn\Delta} = \mathcal{F}\{w[n]^2\}$. Although the Dirichlet kernel $D_K^{R\Delta}(\eta)$ has a narrow central lobe of bandwidth $1/(KR\Delta) \sim 1/(L\Delta)$, it also has several other lobes with period $1/(R\Delta)$ —see Fig. 7. It is therefore important that function $W_2(\eta)$ decreases fast enough in order to keep the effect of those secondary lobes as small as possible. Since the bandwidth of $W_2(\eta)$ is on the order of $1/(\Delta N_w)$, the condition is that $R \ll N_w$, i.e. important overlap must be placed on adjacent data-windows. In practice, $R = N_w/3$ (67% overlap) with a Hanning or a Hamming data-window, and $R = N_w/2$ (50% overlap) with a half-sine data-window will produce an excellent cyclic leakage minimisation because the secondary lobes of $D_K^{R\Delta}(\eta)$ then happen to fall on the nulls of $W_2(\eta)$. These results are illustrated in Fig. 8 and summarised in Table 2.

Magnitudes of the secondary lobes relatively to the main lobe of $|B_Q(\eta)|$ for some typical data-windows used in the averaged cyclic periodogram.

	Hamming	Hanning	half-sine
height of 2nd lobe	-2.0 dB	-1.7 dB	-3.0 dB
height of 3rd lobe	-8.8 dB	-7.8 dB	$-\infty$ dB
height of 4th lobe	$-\infty$ dB	$-\infty$ dB	$-\infty$ dB

Therefore, the following rule applies

Proposition 2. *For the averaged cyclic periodogram, Condition II requires that $L\Delta \ll 1/\Delta_\alpha^{\min}$ and $R \leq N_w/3$ with a Hanning or Hamming data-window, or $R \leq N_w/2$ with a half-sine data-window. Note that setting exactly $R = N_w/3$ and $R = N_w/2$ achieves excellent cyclic leakage rejection with the former respective data-windows.*

Fig. 7. Structure of function $|B_Q(\eta)|$ in the averaged cyclic periodogram.Fig. 8. Illustration of the effect of segments overlapping on cyclic leakage in the averaged cyclic periodogram: continuous lines = function $|B_Q(\eta)|$; dotted lines = functions $|D_K^{RA}(\eta)|/\Delta$ and $|W_2(\eta)|$. Hanning window: (a) no overlap; (b) 50% overlap; (c) 67% overlap. Half-sine window: (d) no overlap; (e) 50% overlap; (f) 67% overlap.

3.3.2.3. Multitaper cyclic periodogram. In the case of the multitaper cyclic periodogram:

$$|B_Q(\eta)| = \left| e^{-j\pi\eta\Delta(L-1)} D_L^\Delta(\eta) - \sum_{k=1}^K \frac{\beta_k}{2} e^{-j\pi(\eta\Delta \pm f_k)(L-1)} e^{\pm j2\pi f_k} D_L^\Delta\left(\eta \pm \frac{f_k}{\Delta}\right) \right| \quad (45)$$

with $f_k = k/L$. It is easy to check that $|B_Q(\eta)|$ has a series of secondary lobes positioned at $\eta = k/(\Delta L)$, $k \in \mathbb{Z}$, with one prominent lobe at $\eta = K/(\Delta L)$ with height on the order of β_k . By setting $\beta_k = 1/K$ —which we will

Table 2
Comparison of asymptotic results (large L)

	Smoothed cyclic periodogram	Averaged cyclic periodogram	Multitaper cyclic periodogram
Characteristic parameters	$g[\tau]$	$w[n], N_w, R$	$\{v_k[n]\}, \beta_k, K$
Power calibration (Condition I)	$g[0] = 1$	$\mathcal{R}_w[0] = 1$	$\sum_{k=1}^K \beta_k = 1$
Cyclic leakage minimisation (Condition II)		Hanning/Hamming: $R \leq N_w/3$ half-sine window: $R \leq N_w/2$	$K \gg 1$
Optimal data-windows ^a :	Papoulis ^b	half-sine ^c	sine tapers
Optimal variance reduction:		$R \leq N_w/3$	$\beta_k = 1/K$
Frequency resolution (Δ_f):	$\sim 1/(\Delta \ g\ ^2)$	$\sim 1/(\Delta \ \mathcal{R}_w\ ^2)$	$\sim K/(L\Delta)$
Cyclic frequency resolution (Δ_α):	$\sim 1/(L\Delta)$	$\sim 1/(L\Delta)$	$\sim 1/(L\Delta)$
Variance reduction factor (\mathcal{E}_{Q1}):	$\sim \ g\ ^2/L$	$\sim \ \mathcal{R}_w\ ^2/L$	$\sim 1/K$

^aSolution to requirement (1) subject to a given variance reduction factor.

^bAuto-correlation of the half-sine window.

^cEqual to the first sine taper.

^dMeasured by $|\int \mathcal{Q}_L(\lambda, 0) d\lambda|^2 / \int |\mathcal{Q}_L(\lambda, 0)|^2 d\lambda$ for simplicity.

later show to be the optimal choice—function $|B_Q(\eta)|$ converges rapidly to $D_L^\Delta(\eta)$ when K becomes large. Hence the following rule applies:

Proposition 3. *For the multitaper cyclic periodogram, Condition II requires that $L\Delta \geq 1/\Delta_\alpha^{\min}$ and $K \gg 1$, i.e. long signals with a large number of tapers have to be used.*

3.3.3. General bias evaluation

The following proposition quantifies the residual bias in the general case.

Proposition 4. *Provided that Conditions I and II are met, the proposed quadratic spectral estimator has bias⁵:*

$$b\{\hat{\mathcal{S}}_{YX}(f; \alpha; L)\} \stackrel{\text{def}}{=} \mathbb{E}\{\hat{\mathcal{S}}_{YX}(f; \alpha; L)\} - \mathcal{S}_{YX}(f; \alpha) \simeq \frac{I_Q}{2} \frac{\partial^2}{\partial f^2} \mathcal{S}_{YX}(f; \alpha) \quad (46)$$

with $I_Q = \Delta^{-1} \int_{-1/2\Delta}^{1/2\Delta} \lambda^2 \mathcal{Q}_L(\lambda, 0) d\lambda$ the inertia of kernel $\mathcal{Q}_L(\lambda, 0)$.

The proof follows from a trivial Taylor expansion of Eq. (40). Proposition 4 generalises the bias formula known in the spectral analysis of stationary signals [14,16]. However, contrary to the stationary case, care should be taken to meet condition II which protects against the effect of cyclic leakage, otherwise Eq. (46) becomes much more intricate. To our knowledge this point has rarely been addressed before. Once again, in order to reduce cyclic leakage, it is very important that the acquisition time be long enough.

3.4. Frequency resolution

Assessment of the frequency resolution of the quadratic estimator follows immediately from the previous section. Upon considering formula (40) it is clear that the smoothing kernel Q_L imposes a limit to the frequency resolutions in both the f and the α frequencies. Interestingly, whereas solving the resolution in the f -frequency is reminiscent to the classical spectral analysis of stationary signals, that in the α -frequency is reminiscent to the spectral analysis of periodic signals. This is made more precise hereafter.

⁵The notation $o(1/L^2)$ means that the approximation error decreases faster than $1/L^2$. In the present case this is true for any “smooth” data-window/lag-window/taper.

3.4.1. Resolution in the f -frequency (Δ_f)

Formula (40) shows that Δ_f is linked to the spectral bandwidth of $\mathcal{Q}_L(\lambda, 0)$. Particular cases lead to the following classical results [16]:

- smoothed cyclic periodogram: $\mathcal{Q}_L(\lambda, 0) = G(\lambda)$ where $G(\lambda) = \mathcal{F}\{g[\tau]\}$;
- averaged cyclic periodogram: $\mathcal{Q}_L(\lambda, 0) = |W(\lambda)|^2$ where $W(\lambda) = \mathcal{F}\{w[n]\}$;
- multitaper cyclic periodogram: $\mathcal{Q}_L(\lambda, 0) = \sum_k \beta_k |V_k(\lambda)|^2$ where $V_k(\lambda) = \mathcal{F}\{v_k[n]\}$.

3.4.2. Resolution in the α -frequency (Δ_α)

Formula (40) shows that Δ_α is linked to the spectral bandwidth of $|B_Q(\eta)|$. Hence the results of Section 3.3 lead to the following proposition:

Proposition 5. *Provided that conditions I and II are met, the resolution in the α -frequency essentially depends on the acquisition time $L\Delta$ irrespectively of the estimator type, i.e. $\Delta_\alpha \sim 1/(L\Delta)$.*

Proposition 5 establishes an unexpected result which refutes the natural intuition that Δ_α should be lower-bounded by the data-window bandwidth, i.e. cruder for the averaged cyclic periodogram than for the smoothed and multitaper cyclic periodograms. These conclusions are supported by Fig. 5 where cyclic-frequency resolution is directly related to the vertical width of kernel $\mathcal{Q}_L(\lambda, \eta)$.

3.5. Variance analysis

It finally remains to prove that the proposed quadratic estimator is consistent, i.e. has estimation variance which tends to zero as the number of available data becomes large. It is the object of this subsection to derive the general expression of the variance as a function of kernel Q_L , and then to deduce its expression in all particular cases.

We start with the following general result (see proof in Appendix A):

Proposition 6. *Let $\hat{\mathcal{S}}_{YX}(f; \alpha_1; L)$ and $\hat{\mathcal{S}}_{VU}(f; \alpha_2; L)$ be two estimators of the cyclic spectra $\mathcal{S}_{YX}(f; \alpha_1)$ and $\mathcal{S}_{VU}(f; \alpha_2)$ at the spectral frequency f and cyclic frequencies α_1 and α_2 , respectively. Then for large L :*

$$\text{Cov}\{\hat{\mathcal{S}}_{YX}(f; \alpha_1; L); \hat{\mathcal{S}}_{VU}(f; \alpha_2; L)\} \simeq \mathcal{E}_{Q1} \cdot \mathcal{S}_{YV}(f + \bar{\beta}\alpha_1 - \bar{\beta}\gamma_1; \gamma_1) \mathcal{S}_{XU}(f - \beta\alpha_2 + \beta\gamma_2; \gamma_2)^* \\ + \mathcal{E}_{Q2} \cdot \mathcal{S}_{YU^*}(f + \bar{\beta}\alpha_1 - \bar{\beta}\gamma_3; \gamma_3) \mathcal{S}_{VX^*}(f + \bar{\beta}\alpha_2 - \bar{\beta}\gamma_4; \gamma_4)^*,$$

where

$$\begin{cases} \gamma_1 = \bar{\beta}\alpha_1 - \bar{\beta}\alpha_2 \\ \gamma_2 = \beta\alpha_2 - \beta\alpha_1 \end{cases} \quad \text{and} \quad \begin{cases} \gamma_3 = \bar{\beta}\alpha_1 + \beta\alpha_2, \\ \gamma_4 = \beta\alpha_1 + \bar{\beta}\alpha_2 \end{cases} \quad (47)$$

and

$$\mathcal{E}_{Q1} = \frac{1}{\Delta^2} \int \int_{-1/2\Delta}^{1/2\Delta} |\mathcal{Q}_L(\lambda, \eta)|^2 d\lambda d\eta = \|Q_L\|^2 \quad (48)$$

and

$$\mathcal{E}_{Q2} = \frac{1}{\Delta^2} \int \int_{-1/2\Delta}^{1/2\Delta} \mathcal{Q}_L(\lambda, \eta) \mathcal{Q}_L(\lambda - \eta, -\eta)^* d\lambda d\eta \quad (49)$$

The above formula reveals the existence of two terms carried by \mathcal{E}_{Q1} and \mathcal{E}_{Q2} , respectively. However, it happens that in many instances the second term is non-zero only on a countable set of frequencies. For instance in the case of *real* signals, we have $dY(f)^* = dY(-f)$ and similarly for X , V , and U so that

$$\begin{aligned} & \mathcal{S}_{YU^*}(f + \bar{\beta}\alpha_1 - \bar{\beta}\gamma_3; \gamma_3) \mathcal{S}_{VX^*}(f + \bar{\beta}\alpha_2 - \bar{\beta}\gamma_4; \gamma_4)^* \\ &= \mathcal{S}_{YU}(f(\beta - \bar{\beta}) + \beta\bar{\beta}(\alpha_1 + \alpha_2); 2f + \bar{\beta}\alpha_1 - \beta\alpha_2) \mathcal{S}_{VX}(f(\beta - \bar{\beta}) + \beta\bar{\beta}(\alpha_1 + \alpha_2); 2f + \bar{\beta}\alpha_2 - \beta\alpha_1)^* \end{aligned} \quad (50)$$

which is nonzero only at some *discrete* values of frequency f such that $2f + 2\bar{\beta}\alpha_1 - \beta\gamma_3 \in \mathcal{A} \bmod(\Delta^{-1})$ and $2f - 2\bar{\beta}\alpha_2 - \gamma_4 \in \mathcal{A} \bmod(\Delta^{-1})$. For the sake of simplicity and because knowledge of the (co)-variance almost everywhere (except on a countable set of frequency points) turn out to be good enough in most applications, we will deliberately neglect the second term carried by \mathcal{E}_{Q2} from now on (upon mentioning the word “almost everywhere”). Hence the following result:

Proposition 7. *For large L , the variance of $\hat{\mathcal{S}}_{YX}(f; \alpha; L)$ is*

$$\text{Var}\{\hat{\mathcal{S}}_{YX}(f; \alpha; L)\} \simeq \mathcal{E}_{Q1} \cdot \mathcal{S}_{2Y}(f + \bar{\beta}\alpha) \mathcal{S}_{2X}(f - \beta\alpha) \quad (51)$$

almost everywhere.

This very simple formula can be checked to generalise the result obtained for stationary signals ($\alpha = 0$) [14,16,17] and it also accepts the results of references [5,7] as particular cases. It states that the variance of $\hat{\mathcal{S}}_{YX}(f; \alpha; L)$ is proportional to the spectra $\mathcal{S}_{2Y}(f + \bar{\beta}\alpha)$ and $\mathcal{S}_{2X}(f - \beta\alpha)$ at frequencies $f + \bar{\beta}\alpha$ and $f - \beta\alpha$, and to the energy $\mathcal{E}_{Q1} = \|Q_L\|^2$ of kernel Q_L .

We now establish the particular expressions taken by \mathcal{E}_{Q1} in the case of the conventional spectral estimators listed so far. Although these calculations boils down to well-known results, we think that regrouping them in a unified approach is interesting in its own right. Since these results are somehow scattered in the specialised literature the reader may find convenient to have hereafter their comparative exposition.

3.5.1. Variance of the cyclic periodogram

For the cyclic periodogram, formula (48) implies that

$$\mathcal{E}_{Q1} = 1 \quad (52)$$

which is independent of the signal length L . This corroborates the well-known fact that the periodogram is not a consistent spectral estimator, i.e. its variance does not diminish when increasing the integration time [14,16,17].

3.5.2. Variance of the smoothed cyclic periodogram

For the smoothed cyclic periodogram, formula (48) implies after some manipulations that:

$$\mathcal{E}_{Q1} = \sum_{\tau=-L+1}^{L-1} g[\tau]^2 \cdot \frac{L - |\tau|}{L^2}. \quad (53)$$

Assuming that the temporal dispersion of $g[\tau]$ is very small as compared to L , result (53) is very well approximated by $\|g\|^2/L$ which is the usual (approximate) formula mentioned in the literature [14,16]. As well-known, a low variance thus requires a long integration time L and a short lag-window $g[\tau]$.

3.5.3. Variance of the averaged cyclic periodogram

For the averaged cyclic periodogram, formula (48) implies after some manipulations that:

$$\mathcal{E}_{Q1} = \sum_{k=-K+1}^{K-1} \mathcal{R}_w[kR]^2 \cdot \frac{K - |k|}{K^2}, \quad (54)$$

where $\mathcal{R}_w[k] = \sum_n w[n - k]w[n]$ is the auto-correlation function of the data-window $w[n]$. This result is valid for any value of R and is in accordance with that of Ref. [18]. A small variance thus requires a large number of averages K , that is a short data-window $w[n]$ as compared to the acquisition time L . Into addition, assessment of formula (54) with classical data-windows (e.g. Hanning, Hamming, half-sine) shows that \mathcal{E}_{Q1} is a decreasing function of R/N_w where the minimum is virtually reached as soon as $R \leq N_w/3$. For large L we have found that the minimum of \mathcal{E}_{Q1} tends rapidly to $\|\mathcal{R}_w\|^2/L$. This is illustrated in Fig. 9.

3.5.4. Variance of Bartlett's estimator

The variance reduction factor of Bartlett's estimator is a particular case of Eq. (54), viz

$$\mathcal{E}_{Q1} = \frac{\mathcal{R}_w[0]^2}{K} = \frac{1}{K}, \quad (55)$$

where we have used $\mathcal{R}_w[0] = \|w\|^2 = 1$ as required by condition I.

3.5.5. Variance of the multitaper cyclic periodogram

For the multitaper cyclic periodogram, formula (48) implies after some manipulations that

$$\mathcal{E}_{Q1} = \sum_{k=0}^{K-1} \beta_k^2 \quad (56)$$

The simplicity of this result stems from the orthonormality of tapers $\{v_k[n]\}_{k=0}^{K-1}$. Hence, a low variance requires averaging many multi-tapered periodograms. This formula can now be used to find the optimal weights β_k such as to minimise \mathcal{E}_{Q1} subject to constraint $\sum_k \beta_k = 1$ imposed by condition I. This is a trivial optimisation problem whose solution is

$$\beta_k^{\text{opt}} = \frac{1}{K}. \quad (57)$$

3.6. Synthesis

The previous subsections have provided the conditions to be met by kernel Q_L in order for the proposed quadratic estimator to provide valid cyclic spectra with minimum bias, accurate frequency resolution, and minimum variance. Overall, three major requirements have emerged (summarised in Fig. 10)

- (1) According to Proposition 4, minimise the inertia I_Q of $\mathcal{Q}_L(\lambda, 0)$ subject to $B_Q(0) = 1$ (Condition I) in order to reduce bias and at the same time improve the f -frequency resolution, that is make $\mathcal{Q}_L(\lambda, 0)$ as close as possible to a Dirac.
- (2) According to the analysis of Section 3.3, minimise the bandwidth of $|B_Q(\eta)|$ (Condition II) subject to $B_Q(0) = 1$ (Condition I) in order to reduce cyclic leakage and at the same time improve the α -frequency resolution, that is make $\mathcal{Q}_L(\lambda, \eta)$ as close as possible to 0 for $\eta \neq 0$.
- (3) According to Proposition 6, minimise the overall energy of $\mathcal{Q}_L(\lambda, \eta)$ subject to $B_Q(0) = 1$ in order to reduce variance, that is make $\mathcal{Q}_L(\lambda, \eta)$ as close as possible to a flat function in the variable λ .

It is well-known from the spectral analysis of stationary signals that requirements (1) and (3) are conflicting and that a compromise has to be found between bias and variance minimisation.

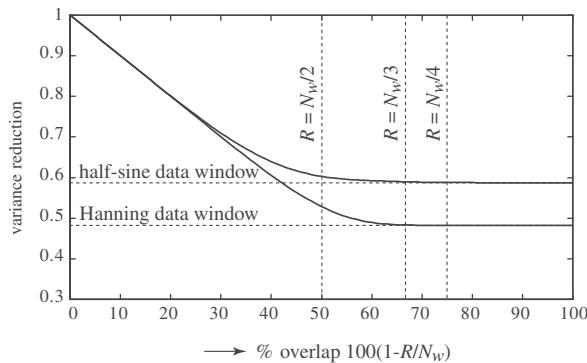
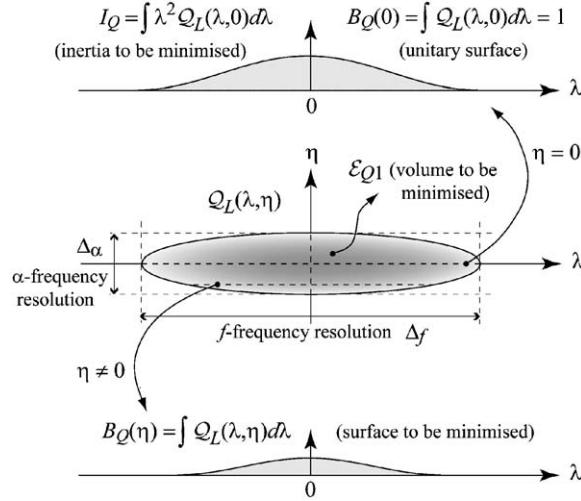


Fig. 9. Variance reduction factor (normalised to 1 at $R = N_w$) of the averaged cyclic periodogram as a function of the percentage of overlap.

Fig. 10. Summary of the requirements imposed on kernel Q_L .

As guidelines we provide a synopsis of results in Table 2. It is clear from this table that provided that the optimal parameters are selected, the three cyclic spectral estimators are all *asymptotically equivalent* since they all share the remarkable asymptotic property

$$\Delta_\alpha \simeq \Delta_f \cdot \mathcal{E}_Q. \quad (58)$$

This statement closes some lengthy debates sometimes initiated in the past as which spectral estimator is the best. It appears here that the best choice depends more on the context of their application and the ease of their implementation than on their actual performance.

4. Other cyclic spectral descriptors

The previous section has discussed in great details the estimation of the cyclic spectrum of CS signals. A number of general results have been established that can now be applied to the estimation of other related spectral descriptors such as the *magnitude* and *phase* of the cyclic spectrum, the *cyclic coherence function* and the *Wigner–Ville spectrum*—together with original results on their estimation. Examples of application will be provided in the next section.

4.1. Magnitude and phase of the cyclic spectrum

The magnitude $|\mathcal{S}_{YX}(f; \alpha)|$ and the phase $\phi_{YX}(f; \alpha)$ of the cyclic spectrum happen to be the quantities of interest in many practical applications. This is because the cyclic spectrum is generally a complex quantity ($\alpha \neq 0$), so that its magnitude and phase have easier physical interpretations when considered separately; they are also more appropriate for graphical displays. The statistical performance of the estimated magnitude and phase can be directly assessed from Propositions 4 and 7 using standard perturbation calculus.

4.1.1. Magnitude of the cyclic spectrum

Proposition 8. For large L , the bias of $|\hat{\mathcal{S}}_{YX}(f; \alpha; L)|$ is

$$\frac{b\{|\hat{\mathcal{S}}_{YX}(f; \alpha; L)|\}}{|\mathcal{S}_{YX}(f; \alpha)|} \simeq \operatorname{Re} \left\{ \frac{b\{\hat{\mathcal{S}}_{YX}(f; \alpha; L)\}}{\mathcal{S}_{YX}(f; \alpha)} \right\} + \frac{\mathcal{E}_{Q1}}{4} \cdot \left[\frac{\mathcal{S}_{2Y}(f + \bar{\beta}\alpha)\mathcal{S}_{2X}(f - \beta\alpha)}{|\mathcal{S}_{YX}(f; \alpha)|^2} - 1 \right] \quad (59)$$

almost everywhere.

Note that even though $\hat{\mathcal{S}}_{YX}(f; \alpha; L)$ was found unbiased in the special case of white signals, its *magnitude* is in general a biased quantity.

Proposition 9. For large L , the variance of $|\hat{\mathcal{S}}_{YX}(f; \alpha; L)|$ is

$$\text{Var}\{|\hat{\mathcal{S}}_{YX}(f; \alpha; L)|\} \simeq \frac{\mathcal{E}_{Q1}}{2} \cdot [|\mathcal{S}_{YX}(f; \alpha)|^2 + \mathcal{S}_{2Y}(f + \bar{\beta}\alpha)\mathcal{S}_{2X}(f - \beta\alpha)] \quad (60)$$

almost everywhere.

4.1.2. Phase of the cyclic spectrum

Proposition 10. For large L , the bias of $\hat{\phi}_{YX}(f; \alpha; L)$ is

$$b\{\hat{\phi}_{YX}(f; \alpha; L)\} \simeq 2\text{Im}\left\{\frac{b\{\hat{\mathcal{S}}_{YX}(f; \alpha; L)\}}{\mathcal{S}_{YX}(f; \alpha)}\right\} \quad (61)$$

almost everywhere.

Proposition 11. For large L , the variance of $\hat{\phi}_{YX}(f; \alpha; L)$ is

$$\text{Var}\{\hat{\phi}_{YX}(f; \alpha; L)\} \simeq \frac{\mathcal{E}_{Q1}}{2} \cdot \left[\frac{\mathcal{S}_{2Y}(f + \bar{\beta}\alpha)\mathcal{S}_{2X}(f - \beta\alpha)}{|\mathcal{S}_{YX}(f; \alpha)|^2} - 1 \right] \quad (62)$$

almost everywhere.

Proposition 12. For large L , the covariance between $|\hat{\mathcal{S}}_{YX}(f; \alpha; L)|$ and $\hat{\phi}_{YX}(f; \alpha; L)$ is

$$\text{Cov}\{|\hat{\mathcal{S}}_{YX}(f; \alpha; L)|; \hat{\phi}_{YX}(f; \alpha; L)\} \simeq 0 \quad (63)$$

almost everywhere.

Again these results generalise those obtained for stationary signals ($\alpha = 0$) (e.g see [14,17]). The proofs essentially rely on exploiting Proposition 6.

4.2. The cyclic coherence function

4.2.1. Definition

As opposed to stationary signals, CS signals have been shown to exhibit correlations in their spectral components spaced apart by cyclic frequencies $\alpha_i \in \mathcal{A}$. The *cyclic coherence* function between signals Y and X is a convenient tool to measure the strength of these correlations. Let $dY(f + \bar{\beta}\alpha)$ and $dX(f - \beta\alpha)$ be the spectral increments of signals Y and X at frequencies $f + \bar{\beta}\alpha$ and $f - \beta\alpha$, respectively. Then the cyclic coherence function $\gamma_{YX}(f)$ is defined as

$$\gamma_{YX}(f; \alpha) = \frac{\mathbb{E}\{dY(f + \bar{\beta}\alpha)dX(f - \beta\alpha)\}}{[\mathbb{E}\{|dY(f + \bar{\beta}\alpha)|^2\}\mathbb{E}\{|dX(f - \beta\alpha)|^2\}]^{1/2}}. \quad (64)$$

It is easy to check that $|\gamma_{YX}(f; \alpha)|^2$ is normalised between 0 and 1 like a regular coefficient of correlation; therefore it provides a measure of correlation that has the remarkable property of being independent of the signal power spectrum. At first glance the cyclic coherence function is a straightforward generalisation of the ordinary coherence function $\gamma_{YX}(f; 0)$ of stationary processes [14]. However, it also has an interpretation when $\alpha \neq 0$: a cyclic coherence close to 1 for some cyclic frequency α indicates that the two signals are “strongly” jointly CS at that cyclic frequency; on the other hand, a cyclic coherence of zero indicates that signals Y and X shares no cyclostationarity at that cyclic frequency (even though they may be strongly coherent at $\alpha = 0$). Another difference is that the cyclic coherence function also applies on a single signal ($Y = X$) in order to measure the strength of its cyclostationarity at cyclic frequency α . For instance, in the case of the CS signal of Example 1

$$\gamma_{2X}(f; \alpha) = \frac{c_i}{|c_0|} \cdot \delta[\alpha - \alpha_i] \quad (65)$$

which clearly points out the cyclostationarity of the signal at any $\alpha = \alpha_i \in \mathcal{A}$, the strength of which is measured by the ratio $|c_i|/|c_0|$.

4.2.2. Estimation issues

Given two finite-length signals $\{Y[n]\}_{n=0}^{L-1}$ and $\{X[n]\}_{n=0}^{L-1}$, an estimator of the cyclic coherence function is

$$\hat{\gamma}_{YX}(f; \alpha; L) = \frac{\hat{\mathcal{S}}_{YX}(f; \alpha; L)}{[\hat{\mathcal{S}}_{2Y}(f + \bar{\beta}\alpha; L)\hat{\mathcal{S}}_{2X}(f - \beta\alpha; L)]^{1/2}}, \quad (66)$$

where $\hat{\mathcal{S}}_{YX}(f; \alpha; L)$ is any of the estimators introduced in Section 3. We summarise hereafter the most relevant results concerning the bias and variance of the squared modulus $|\hat{\gamma}_{YX}(f; \alpha; L)|^2$ which turns out to be the main quantity of interest in practice; the phase $\hat{\phi}_{YX}(f; \alpha; L)$ of $\hat{\gamma}_{YX}(f; \alpha; L)$ is identical to the phase of the cyclic spectrum $\hat{\mathcal{S}}_{YX}(f; \alpha; L)$ addressed in Section 4.1.2. For reasons to become clear later, the statistical analysis of $|\hat{\gamma}_{YX}(f; \alpha; L)|^2$ requires that two cases be distinguished:

- (1) the case H_1 where the analysed signals are jointly CS such that $\gamma_{YX}(f; \alpha) \neq 0$,
- (2) the case H_0 where the analysed signals are jointly stationary such that $\gamma_{YX}(f; \alpha) \equiv 0$.

Proposition 13. Under H_1 and for large L , the bias of $|\hat{\gamma}_{YX}(f; \alpha; L)|^2$ is

$$\frac{b\{|\hat{\gamma}_{YX}(f; \alpha; L)|^2\}}{|\gamma_{YX}(f; \alpha)|^2} \simeq 2Re\left\{\frac{b\{\hat{\mathcal{S}}_{YX}(f; \alpha; L)\}}{\mathcal{S}_{YX}(f; \alpha)}\right\} - \frac{b\{\hat{\mathcal{S}}_{YX}(f; L)\}}{\mathcal{S}_{YX}(f)} - \frac{b\{\hat{\mathcal{S}}_{YX}(f - \alpha; L)\}}{\mathcal{S}_{YX}(f - \alpha)} + \mathcal{E}_{Q1} \frac{[1 - |\gamma_{YX}(f; \alpha)|^2]^2}{|\gamma_{YX}(f; \alpha)|^2}$$

almost everywhere, with $b\{\hat{\mathcal{S}}_{YX}(f; \alpha; L)\}$ given in Proposition (4). Under H_0 , the bias simplifies to

$$b\{|\hat{\gamma}_{YX}(f; \alpha; L)|^2\} \simeq \mathcal{E}_{Q1} \quad (67)$$

almost everywhere.

Proposition 14. Under H_1 and for large L , the variance of $|\hat{\gamma}_{YX}(f; \alpha; L)|^2$ is

$$Var\{|\hat{\gamma}_{YX}(f; \alpha; L)|^2\} \simeq 2\mathcal{E}_{Q1}|\gamma_{YX}(f; \alpha)|^2 \cdot [1 - |\gamma_{YX}(f; \alpha)|^2]^2 \quad (68)$$

almost everywhere. Under H_0 and for large L , the variance simplifies to

$$Var\{|\hat{\gamma}_{YX}(f; \alpha; L)|^2\} \simeq \mathcal{E}_{Q1}^2 \quad (69)$$

almost everywhere.

Proposition 15. Under H_1 and for large L , the covariance between $|\hat{\gamma}_{YX}(f; \alpha; L)|^2$ and $\hat{\phi}_{YX}(f; \alpha; L)$ is

$$Cov\{|\hat{\gamma}_{YX}(f; \alpha; L)|^2; \hat{\phi}_{YX}(f; \alpha; L)\} \simeq 0 \quad (70)$$

almost everywhere.

Remark 1. Propositions 13–15 exclude the useless case where $Y = X$ and $\alpha = 0$ which leads to the trivial (deterministic) result $\hat{\gamma}_{2X}(f; 0) \equiv 1$.

4.2.3. Using the cyclic coherence function for testing for cyclostationarity

We are now in a position to use the previous results for designing a simple and yet powerful hypothesis test for testing for the presence of cyclostationarity at a given cyclic frequency α . Given two finite-length signals $\{Y[n]\}_{n=0}^{L-1}$ and $\{X[n]\}_{n=0}^{L-1}$, we want to test the null hypothesis H_0 : “ Y and X are jointly stationary” against the

alternative hypothesis H_1 : “ Y and X are jointly CS with cyclic frequency α ”, at the level of significance of $100\lambda\%$. This can be achieved by means of the following rule:

Proposition 16. *Reject H_0 if there exists some (non-singleton) frequency interval I where*

$$|\hat{\gamma}_{YX}(f; \alpha; L)|^2 \geq \frac{\mathcal{E}_{Q1}}{2} \chi_{1-\lambda, 2}^2, \quad \forall f \in I \quad (71)$$

with $\chi_{1-\lambda, 2}^2$ the $100(1 - \lambda)\%$ percentile of the χ_2^2 distribution.

The proof of Proposition 16 is based on observing that under the null hypothesis the quantity $2|\hat{\gamma}_{YX}(f; \alpha; L)|^2/\mathcal{E}_{Q1}$ asymptotically follows a chi-square distribution with 2 degrees of freedom with mean and variance given by Eqs. (67) and (69), respectively. To be relevant the proposed hypothesis test requires that sufficient data is available (large L). It should also be remembered that the test is true “almost everywhere”, meaning that point-wise intervals I should not be taken into consideration (for the reasons invoked in Section 6). Very interestingly the so-obtained threshold is independent of frequency.

4.3. The Wigner–Ville spectrum

4.3.1. Definition

The Wigner–Ville spectrum $\mathcal{WV}_{YX}[n, f]$ was introduced in Section 2 and was mentioned to provide a valid joint time-frequency energy distribution for CS signals. As a matter of fact, the Wigner–Ville *spectrum* is related to the expected value of the Wigner–Ville *distribution* that was originally introduced for the time-frequency analysis of finite energy *deterministic* signals [10,11], viz

$$\mathcal{WV}_{YX}[n, f] = \mathbb{E} \left\{ \mathcal{F}_{\tau \rightarrow f} \{ Y[n + \tau/2] X[n - \tau/2]^* \} \right\}. \quad (72)$$

The similarity of the above definition with a time-dependent version of the Wiener–Khintchin theorem supports the interpretation of $\mathcal{WV}_{YX}[n, f]$ as an *instantaneous* spectrum at time n . The use of the symmetric correlation function ($\beta = \frac{1}{2}$) is essential here, because it guarantees that $\mathcal{WV}_{2X}[n, f](Y = X)$ is a real function as required for an instantaneous *power* spectrum.

The Wigner–Ville spectrum turns out to be a very useful tool for analysing CS signals and its interpretation is certainly more intuitive than that of the other spectral descriptors addressed so far. However, it is important to bear in mind that it contains exactly the same information as the cyclic spectrum (and the spectral correlation) but displayed in a different way. This is made explicit through the following formula obtained from combining Eqs. (12) and (20)

$$\mathcal{WV}_{YX}[n, f] = \sum_{\alpha_i \in \mathcal{A}} \mathcal{S}_{YX}(f; \alpha_i) e^{j2\pi\alpha_i n \Delta}, \quad \beta = \frac{1}{2}. \quad (73)$$

For instance, in the case of Example 1,

$$\mathcal{WV}_{2X}[n, f] = \Delta \cdot \sigma_W^2 \cdot |p[n]|^2 \quad (74)$$

which is illustrated in Fig. 11(a); the periodic amplitude modulation of the signal (uniformly spread over the frequency axis) is very obvious from this representation. Similarly Fig. 11(b) displays the Wigner–Ville spectrum associated with the signal of Example 2, where the periodic frequency modulation is now perfectly evidenced, although it was hardly visible in the signal itself.

4.3.2. Estimation issues

The definition and estimation of relevant time-frequency distributions has been an active field of research during the last two decades; a central preoccupation has been to find solutions that can overcome the Fourier uncertainty principle while still providing an energy density with reduced negative interferences⁶ [12]. Among a large variety of candidates, the so-called Cohen’s class of time-frequency distributions has now become widely

⁶Briefly stated, interferences are due to the fact that the Wigner–Ville spectrum of the sum of two signals X_1 and X_2 is not equal to the sum of the Wigner–Ville spectra of each signal; it involves additional cross-terms due to interferences between X_1 and X_2 , which can be

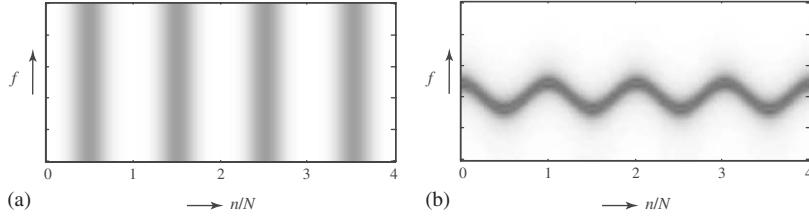


Fig. 11. Wigner–Ville spectra $\mathcal{WV}_{2X}[n,f]$ of the signals of Examples 1 and 2.

acknowledged due to its several good properties. It is defined in the discrete (ideally infinite) case as

$$\mathcal{C}_{YX}[n,f] = \Delta \sum_{\tau=-\infty}^{\infty} \sum_{u=-\infty}^{\infty} \Phi[u,\tau] Y[n-u+\tau/2] X[n-u-\tau/2]^* e^{-j2\pi f \tau \Delta}, \quad (75)$$

where Φ is a “parameterisation” kernel to be optimally designed [12]. The proper selection of kernel Φ in Eq. (75) gives as particular cases most of the classical time-frequency distributions such as the smoothed pseudo-Wigner–Ville distribution, the spectrogram, the Choi–Williams distribution, etc. We now prove that the Wigner–Ville spectrum is a member of the Cohen’s class and that in the CS case, its relationship with the cyclic spectrum can be taken into advantage to define a very efficient time-frequency estimator. Let us start by introducing into Eq. (73) a positive and symmetric lowpass spectral window $H(\alpha)$ such that $0 \leq H(\alpha) \leq H(0) = 1$ for possible smoothing in the time direction, viz:

$$\widehat{\mathcal{WV}}_{YX}[n,f] = \sum_{\alpha_i \in \mathcal{A}} H(\alpha_i) \cdot \widehat{\mathcal{S}}_{YX}(f; \alpha_i) e^{j2\pi \alpha_i n \Delta}, \quad \beta = \frac{1}{2}. \quad (76)$$

Replacing $\widehat{\mathcal{S}}_{YX}(f; \alpha_i)$ by its general expression (26) and identifying with Eq. (75), it then comes that

$$\Phi[u, p-q] = Q_L[p, q] \cdot \sum_{\alpha_i \in \mathcal{A}} H(\alpha_i) e^{-j2\pi \alpha_i u \Delta}, \quad 0 \leq p, q \leq L-1 \quad (77)$$

is the *optimal* structure of the parameterisation kernel to be used in the joint time-frequency analysis of CS signals. Note that in this case Eq. (77) forces kernel Φ to be separable in time and therefore to be *quasi-periodic* with the same cyclic frequencies $\alpha_i \in \mathcal{A}$ as the analysed signals. Finally, using the results of the previous section, the following proposition applies:

Proposition 17. *The Wigner–Ville spectrum estimated by Eq. (76) is a valid time-frequency representation pertaining to the Cohen’s class. It is optimal in the sense that the (negative) interference terms are reduced proportionally to the variance reduction factor \mathcal{E}_{Q1} given in (48)—i.e. they can be made as small as desired by increasing the acquisition time.*

In other words, thanks to its optimal quasi-periodic structure, kernel Φ as given by Eq. (77) provides a consistent estimator of the Wigner–Ville spectrum with the property of minimising the effect of interferences while still offering an accurate time-frequency resolution. This is a remarkable—and unique—result in clear contrast with other kernels classically used in time-frequency analysis where the reduction of interferences is always at the cost of losing time-frequency resolution [12].

Following the lines of the previous subsections, it is now possible to find the variance of the proposed Wigner–Ville spectrum estimator under the two hypotheses H_0 : “the signals are jointly stationary” and H_1 : “the signals are jointly CS”. The variance under H_0 will turn out useful in many practical applications to set a threshold on the Wigner–Ville spectrum when displayed as an image.

(footnote continued)

shown to oscillate between positive and negative values. The presence of those interferences tainted with non-physical meaning (negative energy!) is the main problem of the Wigner–Ville distribution (deterministic signals).

Proposition 18. Under H_1 and for large L , the variance of $\widehat{\mathcal{WV}}_{YX}[n,f]$ is

$$\begin{aligned} \text{Var}\left\{\widehat{\mathcal{WV}}_{YX}[n,f]\right\} &\simeq \mathcal{E}_{Q1} \cdot \sum_{\alpha_i \in \mathcal{A}} \sum_{\alpha_j \in \mathcal{A}} H(\alpha_i) H(\alpha_j)^* \mathcal{S}_{YX}\left(f + \frac{\alpha_i + \alpha_j}{4}, \frac{\alpha_i - \alpha_j}{2}\right) \\ &\quad \times \mathcal{S}_{YX}\left(f - \frac{\alpha_i + \alpha_j}{4}, \frac{\alpha_j - \alpha_i}{2}\right)^* e^{j2\pi(\alpha_i - \alpha_j)n}, \quad \beta = \frac{1}{2} \end{aligned}$$

almost everywhere. Under H_0 and for large L , the variance simplifies to

$$\text{Var}\left\{\widehat{\mathcal{WV}}_{YX}[n,f]\right\} \simeq \mathcal{E}_{Q1} \cdot \sum_{\alpha_i \in \mathcal{A}} |H(\alpha_i)|^2 \mathcal{S}_{YX}\left(f + \frac{\alpha_i}{2}\right) \mathcal{S}_{YX}\left(f - \frac{\alpha_i}{2}\right)^*, \quad \beta = \frac{1}{2} \quad (78)$$

almost everywhere. A useful approximation to the latter result is when the bandwidth of $H(\alpha)$ is short as compared to that of $\mathcal{S}_{YX}(f)$; then:

$$\text{Var}\left\{\widehat{\mathcal{WV}}_{YX}[n,f]\right\} \simeq \mathcal{E}_{Q1} \cdot |\mathcal{S}_{YX}(f)|^2 \sum_{\alpha_i \in \mathcal{A}} |H(\alpha_i)|^2. \quad (79)$$

The proof of these results follows directly from applying Proposition 6.

4.3.3. Time-frequency support and time-frequency resolution

One recognised difficulty when estimating a time-frequency distribution from the direct Fourier transform (75) is the impossibility to handle half samples $\tau/2$, $\tau = \pm 1, \pm 3, \pm 5, \dots$. Two solutions are commonly resorted to in order to disregard odd values of τ without introducing frequency aliasing: either oversampling the signal by a factor two or replacing the signal by its analytic version [12]. Exactly the same difficulty occurs when estimating the Wigner–Ville spectrum from the cyclic spectra in Eq. (76) even though the proposed approach does not explicitly involve a Fourier transform over the time-lag variable τ . To see this, let us consider Fig. 3(c) again which displays the spectral support of the (symmetric) cyclic spectrum. According to Eq. (76) the Wigner–Ville is estimated from a Fourier series over the cyclic frequencies α_i . That is, for a given frequency f , this process evidences that cyclic spectra pertaining to triangles (IV) and (III) will be summed together with cyclic spectra from triangles (II) and (I). This means that *cyclic frequency aliasing* will occur. In order to avoid aliasing, triangles (I) and (II) must be placed such that their supports on the f -axis do not intersect with the support of triangles (IV) and (III). As can be seen in Fig. 12(a)–(b), oversampling the signal or taking its analytic version both meet this requirement.

However, as a side-effect to the necessity of oversampling the signal or taking its analytic version, the time resolution Δ_t becomes frequency dependent and, more dramatically, becomes extremely crude at some frequencies. This is illustrated in Fig. 12(c)–(d). The oversampled signal makes Δ_t cruder when f approaches the Nyquist frequency $\pm 1/(2\Delta)$, which is exactly the opposite behaviour as expected from a relevant time-frequency distribution.⁷ This phenomenon is less problematic with the analytic signal. Nevertheless, these side-effects are hardly perceptible when a short-bandwidth spectral window $H(\alpha)$ is used in Eq. (76), since then the finest time resolution Δ_t of $\widehat{\mathcal{WV}}_{YX}[n,f]$ is dictated by the temporal dispersion of $h[n] = F^{-1}\{H(\alpha)\}$. Conversely, the frequency resolution Δ_f of $\widehat{\mathcal{WV}}_{YX}[n,f]$ is a constant equal to the spectral bandwidth $|B_Q(\eta)|$.

5. Examples of application

5.1. Experiment 1

This first experiment has for objective to prove the consistency of the results derived so far. A CS signal Y was synthesised ($\Delta = 1s$) by passing the periodically modulated white noise X of Example 1 into a linear system H with a pole at $f = 0.15$ Hz and a zero at $f = 0.23$ Hz. The modulation period was $N = 50$ samples,

⁷High frequency components are short in time and ideally require a fine time resolution whereas low frequency components are long in time and can cope with a crude time resolution.

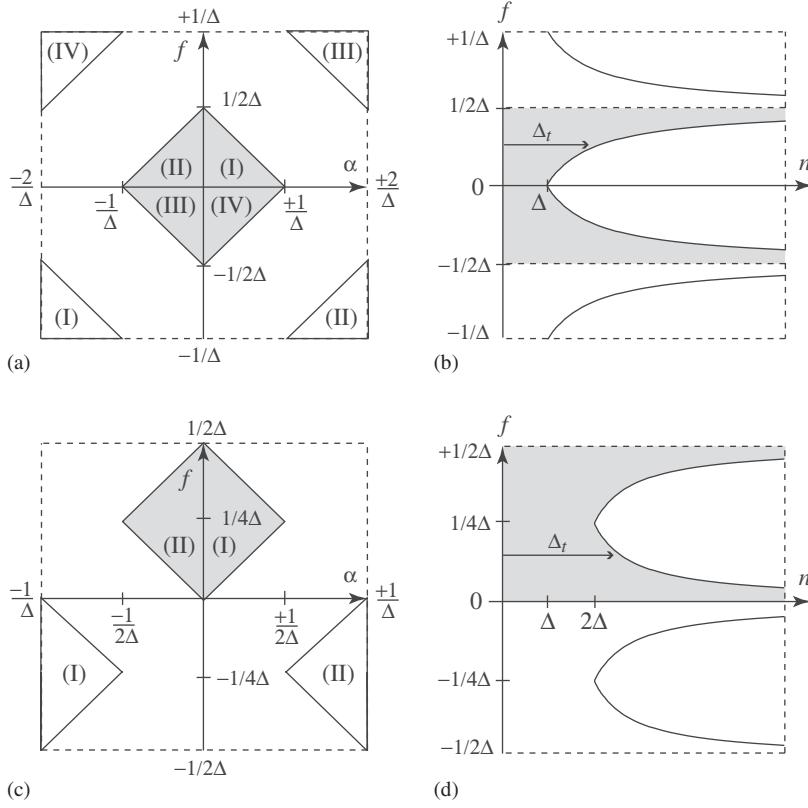


Fig. 12. (a) Principal domain of the spectral correlation after oversampling the signal by a factor 2 and (b) corresponding time resolution Δ_t as a function of frequency in the Wigner–Ville spectrum. (c) Principal domain of the spectral correlation after taking the analytic signal and (d) corresponding time resolution Δ_t as a function of frequency in the Wigner–Ville spectrum.

and the number of available samples $L = 50\,000$. Fig. 13 displays 10 cycles of the synthesised signal, where the hidden periodicity of the signal is hardly noticeable due to the very light damping of the system. The cyclic power spectrum $\mathcal{S}_{2Y}(f; \alpha)$ with $\beta = \frac{1}{2}$ was then estimated at $\alpha = \frac{1}{\Delta N}$ Hz using the three different estimators, namely: (i) the averaged cyclic periodogram, (ii) the smoothed cyclic periodogram, and (iii) the multitaper estimator. For the averaged cyclic periodogram we used the following settings: $N_w = 1000$, $R = N_w/3$, a half-sine data-window, and a FFT size of 2048 samples. This provided a frequency resolution Δ_f of 0.002 Hz, and a common variance reduction factor \mathcal{E}_{Q1} of $\frac{1}{83}$. In order to make comparisons with the other estimators possible, we selected the other parameters according to Table 1 such as to yield identical frequency resolution and variance reduction, that is we used $g[\tau] = \mathcal{R}_w[\tau]$ in the smoothed cyclic periodogram and $K = \lceil L/\|g\|^2 \rceil$ in the multitaper estimator. The magnitudes and the phases of the cyclic power spectrum estimates are displayed in Fig. 14 together with the theoretical cyclic power spectrum given by:

$$\mathcal{S}_{2Y}(f; \alpha) = H(f + \alpha/2)H(f - \alpha/2)^* \cdot \mathcal{S}_{2X}(f; \alpha), \quad \beta = \frac{1}{2}, \quad (80)$$

where $\mathcal{S}_{2X}(f; \alpha) = c_1 \cdot \sigma_X^2$ and $H(f)$ denotes the transfer function of the system. It can be checked that the three estimates are nearly identical, as argued in Section 3.5. Using Propositions 9 and 11, the 6- σ intervals on the estimated log-magnitude and phase were found to be 3.67 dB and 0.52 rad, respectively, which agree very well with the experimental results. The cyclic coherence function $\gamma_{YX}(f; \alpha)$, $\alpha = 1/\Delta N$ was also estimated on the same lines and is displayed in Fig. 15 together with its theoretical value given by:

$$\gamma_{YX}(f; \alpha) = e^{j\phi_H(f+\alpha/2)} \cdot \gamma_{2X}(f; \alpha), \quad \beta = \frac{1}{2}, \quad (81)$$

where $\gamma_{2X}(f; \alpha) = c_1/|c_0|$ and $\phi_H(f)$ stands for the phase of the transfer function $H(f)$. Using Propositions 9, the 6- σ interval on the estimated squared magnitude of the cyclic coherence function was found to be 0.34—

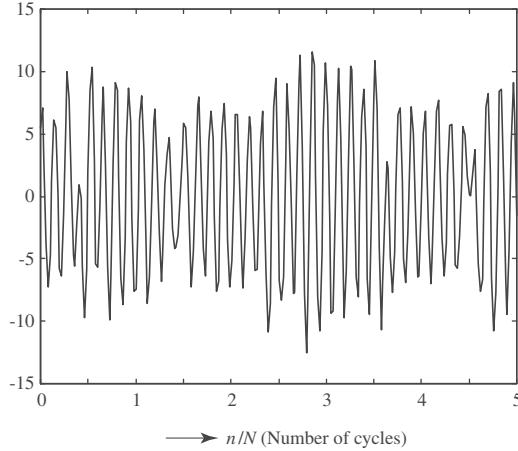
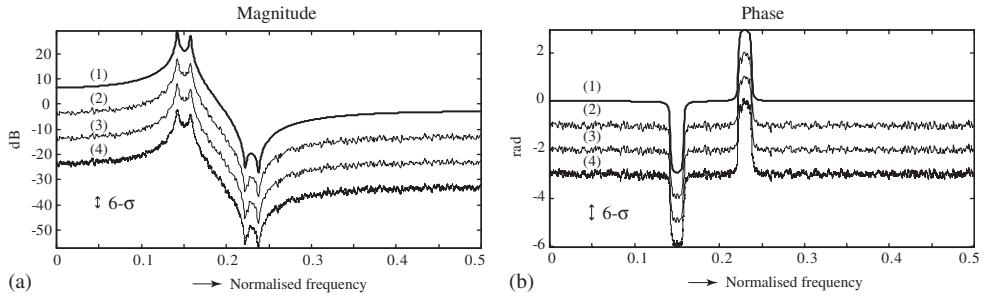
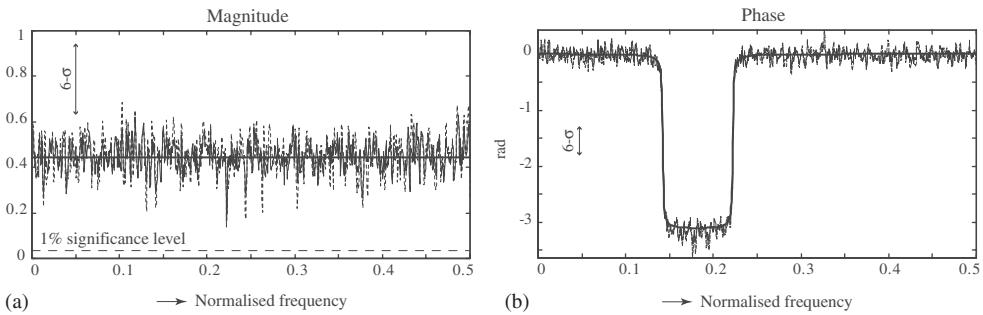


Fig. 13. Synthesised cyclostationary signal.

Fig. 14. (a) Magnitude and (b) phase of the cyclic power-spectrum of the synthesised signal: (1) theoretical values; (2) averaged cyclic periodogram estimate; (3) smoothed averaged cyclic periodogram estimate; (4) multitaper estimate. The 6σ confidence intervals were computed according to Eqs. (60) and (62).Fig. 15. (a) Squared magnitude and (b) phase of the cyclic coherence function of the synthesised signal. The theoretical values (continuous lines) are superposed with those of the three proposed estimates. The 6σ confidence intervals were computed from Eqs. (68) and (62). The 1% level of significance on the squared magnitude is 0.036 according to Eq. (71).

that of the phase is as for the cyclic spectrum—which agrees with the experimental results. Again the three estimates agree very well. Nevertheless if attempting to establish a ranking is deemed meaningful, then the multitaper estimator is found to produce the least bias (as expected by construction), and then the averaged cyclic periodogram.

Although this first experiment aimed only at validating the proposed estimators, it also points towards some original applications. For instance Eq. (80) shows that the phase of the transfer function is perfectly preserved

in the cyclic power spectrum of signal Y ; this is in clear contrast with the conventional power spectrum ($\alpha = 0$), and suggests that inference on the phase is possible from output cyclic statistics only. An additional advantage to this approach is that it is immune to additive stationary noise since cyclic statistics of stationary components are zero for any $\alpha \neq 0$. Applications concern blind deconvolution and equalisation, blind identification, and blind separation of sources [3]. Alternatively, Eq. (80) shows similar potentials when the input to the system is available. An original application to system identification is detailed in Ref. [8].

5.2. Experiment 2

This second experiment has for objective to illustrate the necessity of properly setting the percentage of overlap in the averaged cyclic periodogram. A stationary signal was synthesised ($\Delta = 1s$) and its cyclic coherence function $\gamma_{2X}(f; \alpha)$ ($\beta = \frac{1}{2}$)—theoretically nil at $\alpha \neq 0$ —was computed for several values of alpha in the range $[0.001; 0.08]$ Hz in order to check for the presence of cyclostationarity. We used the averaged cyclic periodogram with the following settings: $N_w = 60$, a Hanning data-window, and FFT sizes of 128 samples; this provided a 1% level of significance of 0.002 to the cyclic coherence function. Different values of the increment parameter R were tested ($100(1 - R/N_w)$ is the overlap between adjacent windows), i.e. $R = N_w$, $R = N_w/2$, $R = N_w/3$, and $R = N_w/4$. Results are displayed in Fig. 16 for the four tested cases. As expected from the analysis of Section 3.3, the case $R = N_w$ —Fig. 16(a)—produces significant cyclic leakage: instead of being statistically zero as expected from the theory, $\hat{\gamma}_{2X}(f; \alpha)$ carries a significant amount of “energy” at $\alpha = 1/\Delta N_w$ leaking from $\alpha = 0$. Similarly, a small amount of leakage—but still statistically significant (!)—is present at $\alpha = 2/\Delta N_w$. The same phenomena occurs at $\alpha = 2/\Delta N_w$ when $R = N_w/2$ —Fig. 16(b). However, as soon as $R \geq N_w/3$, it is checked that cyclic leakage virtually disappears as established in Section 3.3—Fig. 16(c)–(d). Although not shown here very similar observations were obtained with a half-sine data-window except that cyclic leakage then disappeared for $R \geq N_w/2$ in accordance with the discussion of Section 3.3. This experiment illustrates the importance of correctly setting the amount of overlap in the averaged cyclic periodogram, a requirement that is truly specific to cyclic spectral analysis.

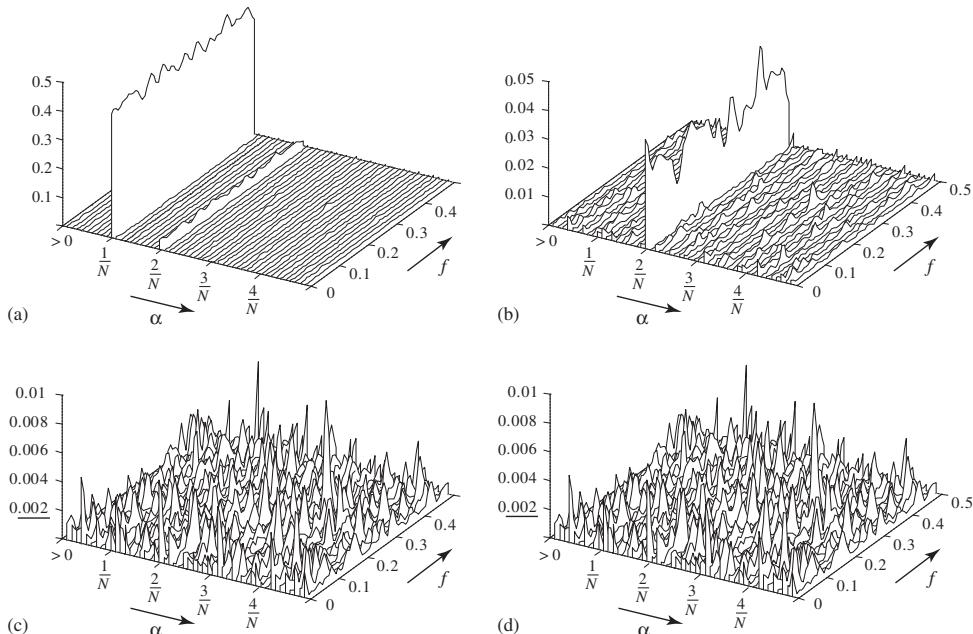


Fig. 16. Illustration of the effect of cyclic leakage on the estimated cyclic coherence function. The cyclic coherence function is that of a stationary signal and is theoretically zero everywhere ($\alpha \neq 0$). A Hanning data-window is used with: (a) $R = N_w$; (b) $R = N_w/2$; (c) $R = N_w/3$, and (d) $R = N_w/4$. The 1% level of significance on the squared magnitude is 0.002 according to Eq. (71).

5.3. Application 1

This first application illustrates the use of cyclic spectral analysis on an industrial system. The system of interest is a hydraulic pump rotating at $\Omega \simeq 3000$ rpm to be monitored by means of vibration analysis. The vibration signal is captured by an accelerometer mounted on the pump casing and is sampled at a rate of 50 kHz. The first objective is to assess the origin of the pump vibrations, and secondly to characterise the “mechanical signature” of these vibrations. Figs. 17(a) and (b) display 0.2 s of the vibration signal and the corresponding power spectrum, respectively. Inspection of these plots suggests that the signal essentially has a random structure: no clear harmonic structure is revealed in the spectrum, but two wide resonance peaks are found around 4.5 and 10.5 kHz, the origin of which is to be investigated. To this purpose, the cyclic coherence function $\gamma_{2X}(f; \alpha)$ ($\beta = \frac{1}{2}$) was computed for different values of α ranging from 1 to 500 Hz with cyclic resolution $\Delta_\alpha = 1$ Hz. The result is displayed in Fig. 18. There is clearly a region with high cyclic coherence at $\alpha = 391$ Hz, i.e. at eight times the pump rotation speed. Note that this harmonic is not present in the power spectrum (even when using a very fine resolution), so it does not relate to periodic vibrations but to some kind of periodic modulation of a random carrier. This modulation is likely due to the passing frequency of the eight blades of the pump propeller, i.e. $8 \times \Omega = 391$ Hz, with $\Omega = 48.9$ Hz. Because the cyclic coherence function at $\alpha = 391$ Hz is above the 1% level of significance, it confirms that the vibration signal effectively exhibits a certain amount of cyclostationarity at this cyclic frequency. The final step is to characterise more finely the structure of this CS source by means of the Wigner–Ville spectrum $\mathcal{W}\mathcal{V}_{2X}(f; \alpha)$. The Wigner–Ville spectrum was computed by using α_i ’s equal to all multiples of 391 Hz so that the blade time-frequency signature could be extracted from the remaining signal. The result is displayed in Fig. 19(a), revealing that the two resonance peaks initially detected in the power spectrum are linked to the blade activity; more specifically they are two random carriers which are periodically amplitude modulated by the blade passing frequency. Incidentally, Fig. 19(b) displays the smoothed pseudo-Wigner–Ville distribution computed on one rotation of the pump with similar time-frequency resolution in order to stress the fact that the proposed estimator of the Wigner–Ville spectrum drastically helps to suppress interference terms and consequently improve the quality of the time-frequency lecture.

5.4. Application 2

The second application illustrates another potential use of cyclic spectral analysis for industrial purpose. The system of interest is a one-stage gearbox composed of two 20-tooth gears rotating at $\Omega \sim 1000$ rpm, submitted to an accelerated fatigue-test at the CETIM (Technical Centre for Mechanical Industries, France). The objective was to investigate the feasibility of acoustic emission (elastic stress waves) for early detection of fatigue damage. Accordingly, the gearbox was run during 15 days until complete breakdown, and acoustic

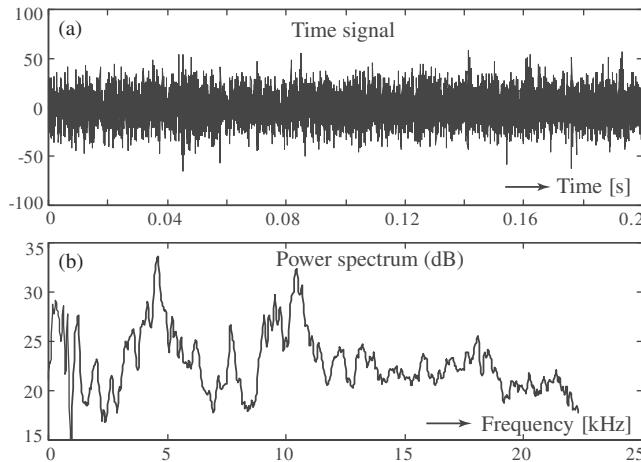


Fig. 17. (a) Time signal and (b) corresponding power spectrum computed from the averaged periodogram with a half-sine window of 1024 samples and 67% overlap. The frequency resolution is $\Delta_f \sim 0.1$ kHz. The 6- σ confidence interval is $\simeq 2.6$ dB according to Eq. (51).

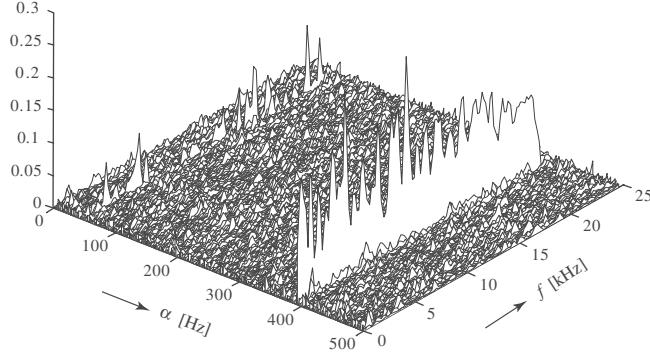


Fig. 18. Cyclic coherence function, computed from the averaged cyclic periodogram with a half-sine window of 256 samples and 67% overlap. The frequency resolution is $\Delta_f \sim 0.3$ kHz. The cyclic frequency resolution is $\Delta_\alpha \sim 1$ Hz. The 1% level of significance is 0.017 according to Eq. (71).

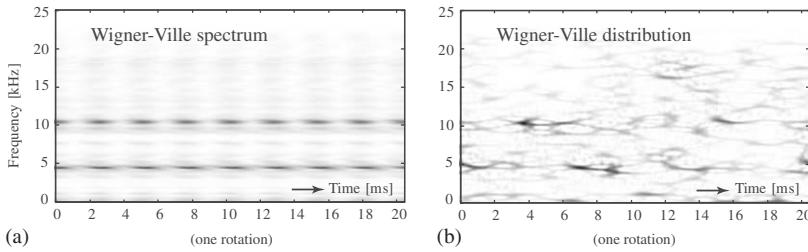


Fig. 19. (a) Wigner–Ville spectrum from the analytic signal, synchronised on the blade passing frequency and displayed over one rotation. The frequency resolution is $\Delta_f \sim 0.3$ kHz. The finest time resolution is ~ 1 ms. Only values above 3 standard deviations of the stationary hypothesis (computed from Eq. (79)) are displayed. (b) Smoothed pseudo-Wigner–Ville distribution displayed over one rotation with similar time-frequency resolution ~ 1 ms.

emission signals were regularly captured during that period. Fig. 20(a) and (b) display 0.4 s of an acoustic emission signal captured at 125 kHz on the first day of the test, together with its corresponding power spectrum. Inspection of these figures reveals a general random behaviour, as to be expected from acoustic emission. This precludes the use of conventional spectral analysis where incipient damage is classically detected from an increase of the gears harmonics (i.e. periodic components) [19]. Nevertheless, a similar strategy can be devised based on cyclic spectral analysis. Fig. 21 displays the cyclic coherence function $\gamma_{2X}(f; \alpha)$ ($\beta = \frac{1}{2}$) computed on the signal of Fig. 20(a) for cyclic frequencies ranging from 1 to 500 Hz, with cyclic resolution $\Delta_\alpha = 0.5$ Hz. The cyclic coherence function reveals the existence of a very rich structure of hidden periodicities linked with the gears rotation speed, i.e. $\alpha_i = i \cdot \Omega$, with $\Omega = 16.9$ Hz. Noteworthy also is the strong cyclic harmonic at $20 \times \Omega = 338$ Hz which relates to the gearmesh frequency of the 20-tooth wheels. This harmonic structure is exactly similar to that of gears vibration signals [19], yet it occurs in the α -frequency instead of the f -frequency. Therefore a natural idea is to use the cyclic coherence function at $\alpha = \Omega$ as an indicator of damage intensity of the gears. This is checked in Fig. 22 where the power spectra and the cyclic coherence functions are compared on the first day and the last day of the test (just before breakdown). Whereas the power spectrum shows some changes in the energy distribution that can hardly be related to anything without further information, the cyclic coherence function evidences a net increase whose origin can be attributed unambiguously to an abnormally strong cyclic phenomena. The characterisation of the detected damage can be brought up one step further by means of the Wigner–Ville spectrum $\mathcal{W}\mathcal{V}_{2X}(f; \alpha)$ as shown in Fig. 23, where the temporal information allows the exact location of the fatigue damage on the 11th tooth.

5.5. Application 3

The third application illustrates the use of cyclic spectral analysis in a very different context than the previous examples. The signals of interest are monthly sea surface temperatures averaged over several Nino

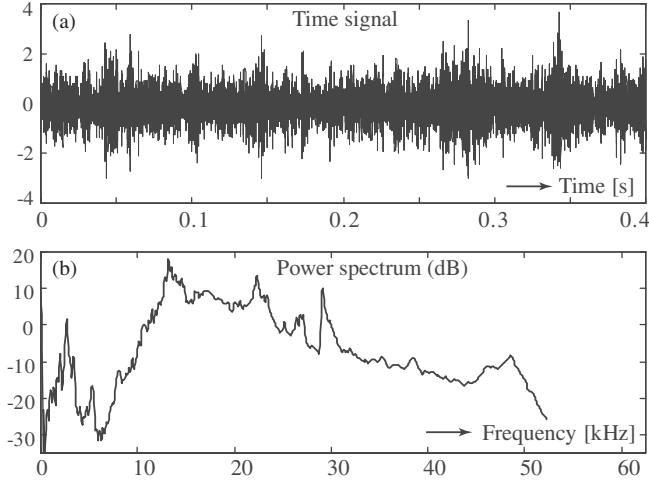


Fig. 20. (a) Time signal and (b) power spectrum computed from the averaged periodogram with a half-sine window of 1024 samples and 67% overlap. The frequency resolution is $\Delta_f \sim 0.25$ kHz. The 6σ confidence interval is ≈ 1.1 dB according to Eq. (51).

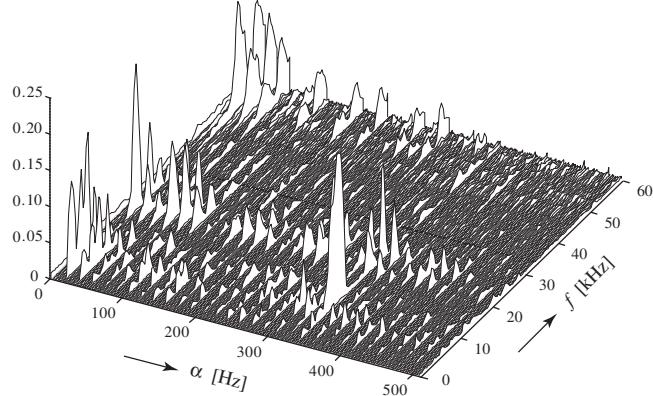


Fig. 21. Cyclic coherence function, computed from the averaged cyclic periodogram with a half-sine window of 128 samples and 67% overlap. The frequency resolution is $\Delta_f \sim 2$ kHz. The cyclic frequency resolution is $\Delta_\alpha \sim 0.5$ Hz. The 1% level of significance is 0.001 according to Eq. (71).

regions, from 1871 to 1996. The objective is to check for the existence of linear relationships between these time series for use in climate prediction models. Fig. 24 displays two temperature time series from two different regions, where the seasonal effects have been extracted from the anomalies, i.e. the periodic parts from the residual random parts. Analysis of the correlation between the two seasonal parts is an easy matter and is not of concern here; only the correlation between the anomalies is investigated. From a stationary point of view there is obviously a strong coherence in the low frequency patterns of the anomalies, but cyclic correlations ($\alpha \neq 0$) should be tested as well. For this purpose the cyclic coherence function $\gamma_{YX}(f; \alpha)$ ($\beta = \frac{1}{2}$) between the two time series was estimated for cyclic frequencies ranging from 0 to 4 cycles/year. Results are displayed in Fig. 25; they clearly confirm the existence of a strong linear relationship between the two time series at $\alpha = 0$ (ordinary coherence function), but also at $\alpha = 1$ cycle/year. This analysis proves that the anomalies still contain some hidden seasonal effects (yearly correlations), however not in an additive form but in a multiplicative form. Further investigation with the cross-Wigner–Ville spectrum $\mathcal{W}\mathcal{V}_{YX}(f; \alpha)$ between the two time series helps to better characterise these effects, as shown in Fig. 26, where a significant periodic amplitude modulation plus a slight frequency modulation are clearly visible. This extra information could be advantageously taken into account when designing a prediction model, for instance by inclusion of periodically time-varying coefficients.

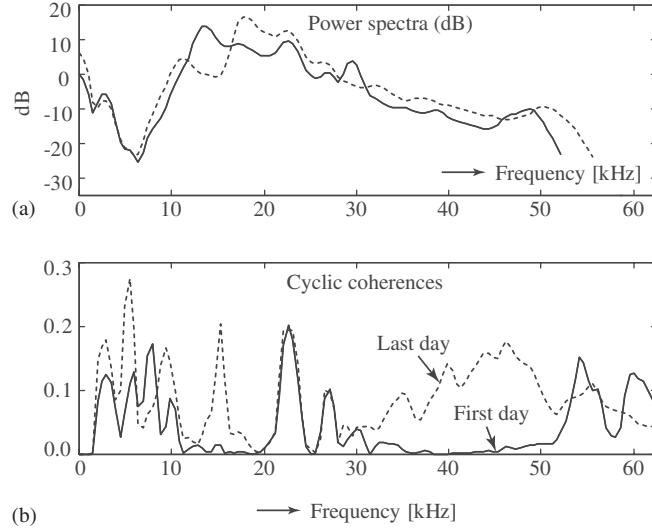


Fig. 22. (a) Power spectra and (b) cyclic coherence functions at $\alpha = \Omega = 16.67$ Hz on the first day (continuous line) and last day (dotted line) of the test. The 6σ confidence interval in (a) is $\simeq 0.4$ dB (Eq. (51)), and the 1% level of significance in (b) is 0.001 according to Eq. (71).

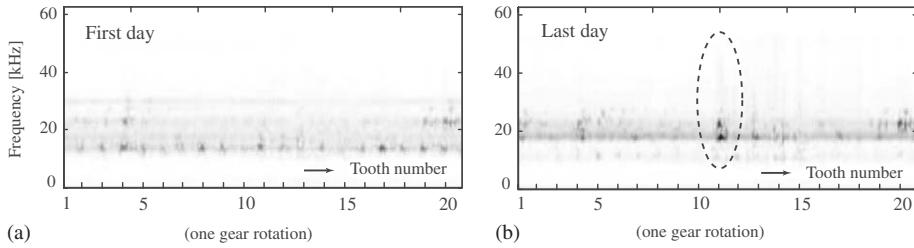


Fig. 23. Wigner–Ville spectra from the analytic signals, synchronised on the gears rotation frequency and displayed over one rotation. The frequency resolution is $\Delta_f \sim 2$ kHz. The finest time resolution is ~ 0.3 ms, i.e. a tenth of a tooth. Only values above three standard deviations of the stationary hypothesis (computed from Eq. (79)) are displayed on a gray scale. (a) First day and (b) last day of the test.

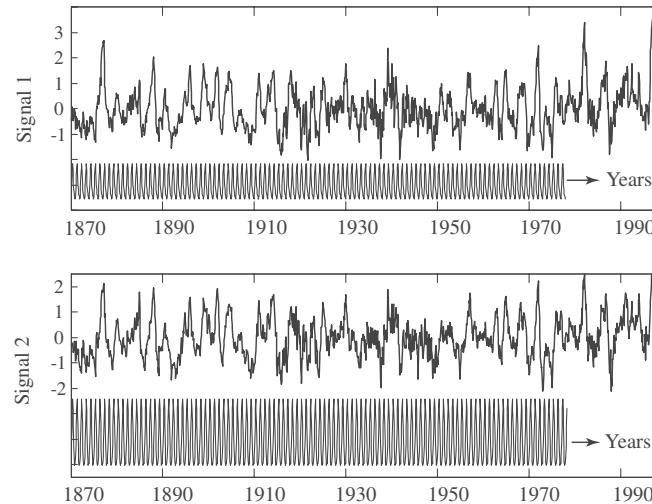


Fig. 24. Temperature time series at two different Nino regions from 1870 to 1996. Anomalies (upper traces in each panel) have been extracted from the seasonal (yearly) effects (bottom traces in each panel).

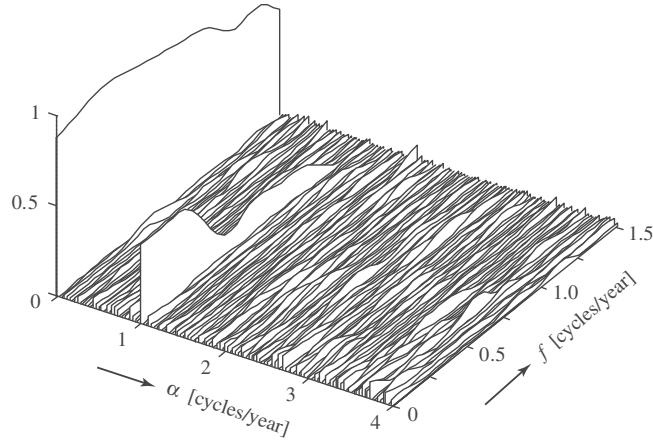


Fig. 25. Cyclic coherence function, computed from the averaged cyclic periodogram with a half-sine window of 64 samples and 67% overlap. The frequency resolution is $\Delta_f \sim 0.3$ cycle/year. The cyclic frequency resolution is $\Delta_\alpha \sim 0.03$ cycle/year. The 1% level of significance is 0.1 according to Eq. (71).

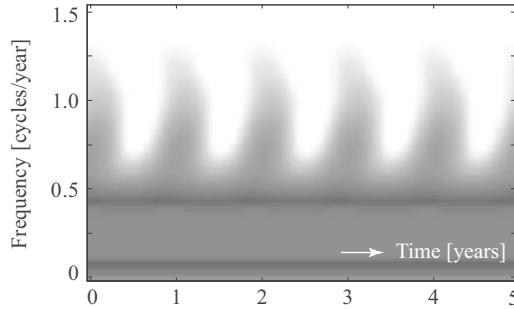


Fig. 26. Wigner–Ville spectrum from the analytic signal, synchronised on the yearly effects. The frequency resolution is $\Delta_f \sim 0.3$ cycle/year. The finest time resolution is ~ 0.5 year. Only values above 3 standard deviations of the stationary hypothesis (computed from Eq. (79)) are displayed on a gray scale.

6. Conclusion

The purpose of this paper was to provide a comprehensive exposition of the spectral analysis of cyclostationary signals (CS). Although some facets of the subject have already been discussed in the specialised literature, our originality lies in an effort to propose an unifying approach, from which all previously published results can be retrieved as particular cases. First, the empirical estimation of the cyclic spectrum has been investigated in detail, and it has been shown that conventional estimators can all be derived from a general quadratic form parameterised with a smoothing kernel Q_L . The conditions for this kernel to yield minimum bias and consistent estimators have been addressed. In particular, a new type of bias, coined cyclic leakage, has been identified and remedies against it have been proposed. We have then arrived to the conclusion that conventional cyclic spectral estimators end up with similar asymptotic results, so that the important question is not so much which one gives the best results, but which one is the most convenient for use in a specific application. In short, we advocate the use of the averaged cyclic periodogram due to its ease of implementation and high computational efficiency, whilst surprisingly it has rarely been suggested in previous works. Into addition by setting at least 50% overlap with a half-sine data window, or 67% overlap with a Hanning or a Hamming data-window, the averaged cyclic periodogram offers excellent cyclic leakage rejection, while at the same time this precaution guarantees a nearly minimum estimation variance. Moreover, by providing the relevant conditions in designing kernel Q_L , the material presented herein leaves open a wide scope of possibilities for designing new and original cyclic spectral estimators. A second originality of this

paper is to show how the above results can be extended to more advanced spectral quantities such as the cyclic coherence function and the Wigner–Ville spectrum. In particular we have shown how to design an optimal estimator of the Wigner–Ville spectrum of CS signals that is able to reduce negative interference terms while still preserving the finest possible time-frequency resolution. Finally, the results presented in the paper have been intensively illustrated on simulated and actual examples of application.

As a final conclusion, cyclic spectral analysis has been shown to be no more difficult than conventional spectral analysis, provided that a number of elementary guidelines are followed. The only recognised drawback so far is that when the set of cyclic frequencies is unknown a large number of values of α should be scanned, a task which can rapidly become computationally intensive.

Appendix A. Proof of Proposition 7

The multilinearity of the covariance function implies that

$$\begin{aligned} \text{Cov}\{\hat{\mathcal{S}}_{YX}(f; \alpha_1; L); \hat{\mathcal{S}}_{VU}(f; \alpha_2; L)\} &= \frac{1}{\Delta^2} \int \int \int \int \mathcal{Q}_L(\lambda, \eta) \mathcal{Q}_L(\xi, \zeta)^* \\ &\quad \times \text{Cov}\{dY(f + \bar{\beta}\alpha_1 - \lambda) dX(f - \beta\alpha_1 - \lambda + \eta)^*; \\ &\quad \times dV(f + \bar{\beta}\alpha_2 - \xi) dU(f - \beta\alpha_2 - \xi + \zeta)^*\}, \end{aligned} \quad (\text{A.1})$$

where

$$\begin{aligned} &\text{Cov}\{dY(f + \bar{\beta}\alpha_1 - \lambda) dX(f - \beta\alpha_1 - \lambda + \eta)^*; dV(f + \bar{\beta}\alpha_2 - \xi) dU(f - \beta\alpha_2 - \xi + \zeta)^*\} \\ &= \text{Cum}\{dY(f + \bar{\beta}\alpha_1 - \lambda); dX(f - \beta\alpha_1 - \lambda + \eta)^*; dV(f + \bar{\beta}\alpha_2 - \xi)^*; dU(f - \beta\alpha_2 - \xi + \zeta)\} \\ &\quad + \text{Cov}\{dY(f + \bar{\beta}\alpha_1 - \lambda); dV(f + \bar{\beta}\alpha_2 - \xi)\} \text{Cov}\{dX(f - \beta\alpha_1 - \lambda + \eta); dU(f - \beta\alpha_2 - \xi + \zeta)\}^* \\ &\quad + \text{Cov}\{dY(f + \bar{\beta}\alpha_1 - \lambda); dU(f - \beta\alpha_2 - \xi + \zeta)^*\} \text{Cov}\{dV(f + \bar{\beta}\alpha_2 - \xi); dX(f - \beta\alpha_1 - \lambda + \eta)^*\}. \end{aligned} \quad (\text{A.2})$$

It can be shown that the fourth order cumulant is one order of magnitude (in $1/L$) lower than the other covariance terms; therefore it can be neglected for large L . Using definitions (14) and (20), the first covariance term becomes

$$\text{Cov}\{dY(f + \bar{\beta}\alpha_1 - \lambda); dV(f + \bar{\beta}\alpha_2 - \xi)\} = \sum_{\alpha_i} \mathcal{S}_{YV}(f + \bar{\beta}(\alpha_1 - \gamma_1) - \lambda; \alpha_i) \delta(\gamma_1 - \alpha_i) d\lambda d\xi \quad (\text{A.3})$$

where $\gamma_1 = \bar{\beta}(\alpha_1 - \alpha_2) - \lambda + \xi$. Repeating the same procedure for the other terms and using the property of the Dirac impulses, it comes:

$$\begin{aligned} \text{Cov}\{\hat{\mathcal{S}}_{YX}(f; \alpha_1; L); \hat{\mathcal{S}}_{VU}(f; \alpha_2; L)\} &\simeq \frac{1}{\Delta^2} \sum_{\alpha_i} \sum_{\alpha_j} \int \int \mathcal{Q}_L^1(\lambda, \eta) \mathcal{S}_{YV}(f + \bar{\beta}(\alpha_1 - \alpha_i) - \lambda; \alpha_i) \\ &\quad \times \mathcal{S}_{XU}(f - \beta(\alpha_2 - \alpha_j) - \xi + \zeta; \alpha_j)^* d\lambda d\eta \\ &\quad + \frac{1}{\Delta^2} \sum_{\alpha_i} \sum_{\alpha_j} \int \int \mathcal{Q}_L^2(\lambda, \eta) \mathcal{S}_{YU^*}(f + \bar{\beta}(\alpha_1 - \alpha_i) \\ &\quad - \lambda; \alpha_i) \mathcal{S}_{VX^*}(f + \bar{\beta}(\alpha_2 - \alpha_j) - \xi; \alpha_j)^* d\lambda d\eta, \end{aligned}$$

where $\mathcal{Q}_L^1(\lambda, \eta) = \mathcal{Q}_L(\lambda, \eta) \mathcal{Q}_L(\alpha_i - \bar{\beta}(\alpha_1 - \alpha_2) + \lambda, \alpha_i - \alpha_j - \alpha_1 + \alpha_2 + \eta)^*$ and $\mathcal{Q}_L^2(\lambda, \eta) = \mathcal{Q}_L(\lambda, \eta) \mathcal{Q}_L(-\alpha_j + \beta\alpha_1 + \bar{\beta}\alpha_2 + \lambda - \eta, -\alpha_j - \alpha_i + \alpha_1 + \alpha_2 - \eta)^*$. A simplification is possible when L is sufficiently large so that $\mathcal{Q}_L(\lambda, \eta)$ behaves like a product of Dirac impulses as compared to the other quantities in the

integration sign:

$$\begin{aligned} & \text{Cov}\{\hat{\mathcal{S}}_{YX}(f; \alpha_1; L); \hat{\mathcal{S}}_{VU}(f; \alpha_2; L)\} \\ & \simeq \sum_{\alpha_i} \sum_{\alpha_j} \mathcal{S}_{YV}(f + \bar{\beta}(\alpha_1 - \alpha_i); \alpha_i) \mathcal{S}_{XU}(f - \beta(\alpha_2 - \alpha_j); \alpha_j)^* \frac{1}{\Delta^2} \int \int \mathcal{Q}_L^1(\lambda, \eta) d\lambda d\eta \\ & \quad + \sum_{\alpha_i} \sum_{\alpha_j} \mathcal{S}_{YU^*}(f + \bar{\beta}(\alpha_1 - \alpha_i); \alpha_i) \mathcal{S}_{VX^*}(f + \bar{\beta}(\alpha_2 - \alpha_j); \alpha_j)^* \frac{1}{\Delta^2} \int \int \mathcal{Q}_L^2(\lambda, \eta) d\lambda d\eta \end{aligned}$$

A final simplification occurs when the bandwidths of $\mathcal{Q}_L(\lambda, \eta)$ in both λ and η are significantly smaller than $\min_i |\alpha_i|$, $\alpha_i \in \mathcal{A}$. Hence it must be (i) that $\alpha_i = \bar{\beta}(\alpha_1 - \alpha_2)$ and $\alpha_j = \beta(\alpha_2 - \alpha_1)$ for $\mathcal{Q}_L^1(\lambda, \eta)$ to be different from 0, and (ii) $\alpha_i = \bar{\beta}\alpha_1 + \beta\alpha_2$ and $\alpha_j = \beta\alpha_1 + \bar{\beta}\alpha_2$ for $\mathcal{Q}_L^2(\lambda, \eta)$ to be different from 0. Reporting these values in the latter equation then yields Proposition 6.

References

- [1] W. Gardner, Exploitation of spectral redundancy in cyclostationary signals, IEEE Signal Processing Magazine, April 1991, 14–36.
- [2] W. Gardner, Introduction to Random Processes, second ed., McGraw-Hill, New York, 1990 (Chapter 12).
- [3] G. Giannakis, Cyclostationary signal analysis, in: The Digital Signal Processing Handbook, IEEE Press, 1998 (Chapter 5).
- [4] W. Gardner, Measurement of spectral correlation, IEEE Transactions on Acoustics, Speech, and Signal Processing 34 (5) (1986) 1111–1123.
- [5] H. Hurd, Nonparametric time series analysis for periodically correlated processes, IEEE Transactions on Information Theory 35 (2) (1989) 350–359.
- [6] R. Roberts, W. Brown, H. Loomis, Computationally efficient algorithms for cyclic spectral analysis, IEEE Signal Processing Magazine, April 1991, 38–49.
- [7] A. Dandawate, G. Giannakis, Nonparametric polyspectral estimators for k th-order (almost) cyclostationary processes, IEEE Transactions on Information Theory 40 (1) (1994) 67–84.
- [8] J. Antoni, F. Bonnardot, A. Raad, M. El Badaoui, Cyclostationary modelling of rotating machine vibration signals, Mechanical Systems and Signal Processing 18 (6) (November 2004) 1285–1314.
- [9] J. Antoni, P. Wagstaff, J.-C. Henrio, H_x —A consistent estimator for frequency response functions with input and output noise, IEEE Instrumentation and Measurement 53(2) 457–465.
- [10] W. Martin, Time-frequency analysis of random signals, International Conference on Acoustics, Speech, and Signal Processing, Paris, 1982, 1325–1328.
- [11] W. Martin, P. Flandrin, Wigner–Ville spectral analysis of nonstationary processes, IEEE Transactions on Acoustics, Speech, and Signal Processing 33 (6) (1985) 1461–1470.
- [12] P. Flandrin, J. Stockler, Time-Frequency/Time-scale Analysis, Academic Press, New York, 1999.
- [13] M. Genosar, H. Lev-Ari, T. Kailath, Consistent estimation of the cyclic auto-correlation, IEEE Transactions on Acoustics, Speech, and Signal Processing 42 (3) (1994) 595–603.
- [14] G. Jenkins, D. Watts, Spectral Analysis and its Applications, Emerson-Adams Press, 1996.
- [15] K. Riedel, A. Sidorenko, Minimum bias multiple taper spectral estimation, IEEE Transactions on Signal Processing 43 (1) (1995) 188–195.
- [16] D. Percival, A. Walden, Spectral Analysis for Physical Applications, Cambridge University Press, Cambridge, 1993.
- [17] J. Bendat, A. Piersol, Random Data—Analysis and Measurement Procedures, second ed., Wiley Interscience, New York, 1986.
- [18] P. Welch, The use of the fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Transactions on Audio and Electroacoustics, 15 (1967) 70–73.
- [19] R.B. Randall, A new method for modeling gear faults, Journal of Mechanical Design 104 (1986) 259–267.