The infogram: Entropic evidence of the signature of repetitive transients

Jérôme Antoni

HAL Id: hal-01714496
https://hal.science/hal-01714496

The infogram: Entropic evidence of the signature of repetitive transients

Jerome Antoni

Laboratoire Vibrations Acoustique (LVA), University of Lyon, INSA-Lyon, F-69621 Villeurbanne, France

A B S T R A C T

A classical symptom of rotating machines faults in vibration signals is the presence of repetitive transients, whose distinctive signature is both impulsive and cyclostationary. Typical approaches for their detection proceed in the time or frequency domains, with tools such as the spectral kurtosis, the kurtogram, or the envelope spectrum. The object of this paper is to extend and somehow connect these concepts in order to capture the signature of repetitive transients in both domains. Motivated by ideas borrowed from the field of thermodynamics where transients are seen as departures from a state of equilibrium, it is proposed to measure the negentropy of the squared envelope (SE) and of the squared envelope spectrum (SES) of the signal. This defines the SE infogram, the SES infogram, and their average which is theoretically maximum for a Dirac comb according to Hirschman's uncertainty principle. It is demonstrated that the joint consideration of the infograms significantly extends the domain of applicability of the kurtogram, in particular to situations corrupted with impulsive noise or when the relaxation time of the transients is low as compared to their rate of repetition. This is illustrated on both synthetic and actual vibration signals. This paper is part of a special issue in honor of Professor Simon Braun and pays tribute to his early contribution to the field of mechanical signature analysis.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

One of the fundamental premises of vibration-based condition monitoring is that the abnormal states of a machine lead to symptomatic vibrations that can be easily distinguished from a healthy reference. This is the concept of mechanical signature analysis that was early investigated in Simon Braun's pioneer work of among others. When concerned with rotating machines, incipient faults often produce series of transients, which leaves good hope for their early detection. Typical examples are rolling element bearing faults and gear faults—as explained in Simon Braun's book and early papers [1,2]. Indeed, whenever a fault produces impacts in the machinery (i.e. when some mechanical parts strike each other), the corresponding mechanical signature is likely to take the form of repetitive transients.

Because of their distinctive characteristics, the detection of repetitive transients has been approached from several angles. One of the earliest strategies was to quantify the intensity of the impacts by means of scalar indicators such as the crest factor, the peak-to-peak,

the kurtosis [3], or "vector" indicators such as the shock spectrum [4]. In essence, these techniques characterize the impulsiveness of a series of transients. With the advent of digital spectral analyzers, it was then recognized that repetitive transients could also be detected from their characteristic frequencies. Based on empirical grounds, the envelope spectrum was early proposed as a powerful diagnostic tool [5]. A related approach introduced by Simon Braun in Ref. [6] proposed to synchronously average the energy of repetitive transients; incidentally, this led the foundations of cyclostationary analysis – without naming it so at that time – which aims at characterizing the cyclic nature of signals. These techniques have been recently formalized by making explicit use of the cyclostationary framework developed in the field of communication engineering [7]. Another family of approaches is based on time-frequency and time-scale analyses; incidentally, the latter often happen to boil down to particular cases of envelope spectrum analysis where various bandpass filters are designed by means of wavelet or other kinds of related transforms.

The specialized literature on these matters is nowadays considerable and an exhaustive survey has become out of reach. Two recent contributions will be considered only which are in the direct continuation of the aforementioned historical works. The first one is the spectral kurtosis (SK) and its related kurtogram and second one is the spectral correlation.

The SK was introduced in Refs. [8–10] to detect and localize in frequency the presence of impulsiveness in a signal. It is essentially a distribution of the kurtosis as a function of frequency, notwithstanding the theoretical technicalities that this definition implies. Because the SK strongly depends on the frequency resolution, it has been suggested to display it as a function of both frequency and frequency resolution, which leads to the concept of kurtogram; in short, a high value of the kurtogram indicates high impulsiveness in the corresponding frequency band. Several research works have recently investigated various implementations (e.g. based on wavelets, wavelet packets, dual-tree wavelets, etc. [11–19]) of the kurtogram in order to improve its efficiency. However, one of the most serious limitations of the kurtogram is its inability to recognize whether a series of transients is repetitive or not. Actually, it is known that the value of the kurtosis decreases when the repetition rate of the transients increases. One extreme is when the transients become so frequent that they overlap over each other, in which case the kurtosis vanishes [20]. The other extreme is when the transients become so spaced apart that there is only one left in the recorded signal, in which case the kurtosis is maximal. This is also the exact reason why the kurtogram is unfortunately found so sensitive to impulsive noise and can sometimes lead to misleading interpretations. This issue was addressed in Ref. [21] which proposed an interesting alternative – the protrugram – based on the kurtosis of the envelope spectrum rather than the signal. However, despite its merits, the protrugram has the dual limitations of the SK and it raises the question of which mathematical expression to use in its definition among many possible choices.

Besides the SK, the spectral correlation (or its normalized version, the spectral coherence) is probably one of the most efficient tools nowadays available to detect the presence of repetitive transients in a signal. Such an ability is rigorously demonstrated within the framework of cyclostationary signals [22–25]. In brief, the spectral correlation may be seen as a waterfall of envelope spectra for all carrier frequencies of the signal. Its application to the diagnosis of rolling element bearings has been shown to return excellent results in a variety of configurations [24–31]. Unfortunately, the computation of the spectral correlation might still be prohibitive in certain applications, which places it as a second choice behind the kurtogram. A very illuminating link between the spectral correlation and the SK was recently elucidated in Ref. [32]: given a frequency band, the kurtosis is nothing but the sum of all the peaks of the squared envelope spectrum normalized by the zero-frequency peak. High kurtosis values therefore mean that the envelope spectrum is non-flat and possesses high non-zero frequency peaks. In words, not only can the spectral correlation detect repetitiveness of a phenomenon, but it can also characterize its impulsiveness through the number of its harmonics.

The aim of the present research work is to proceed with the concept of the kurtogram and to modify its definition in order to fix its shortcomings. The idea is to make it sensitive to the cyclic property of repetitive transients – actually to cyclostationarity in general – while still preserving its simplicity. The proposed solution is motivated by both the protrugram [21] and the link between the kurtosis and the envelope spectrum elucidated in Ref. [32]. Inspired by concepts of thermodynamics, a transient is seen as a departure from the state of equilibrium of the system of interest. In the time domain, this corresponds to local fluctuations of the signal energy that comes with a decrease of the entropy of its squared envelope (SE). If the energy fluctuations are cyclic, this also comes with local fluctuations in the envelope spectrum (in the form of peaks) and therefore with a decrease of the entropy of the Fourier transform of the SE. Interestingly, the sum of the entropies in the time and frequency domains achieves a lower bound, as first shown by Hirschman in 1957 [35], a result known as the "entropic uncertainty principle" [36]. In the case of discrete-time signals – which is of concern here – it has recently been shown that the entropic bound is achieved by a Dirac comb which, incidentally, stands as a perfect idealization of a series of transients. These theoretical results suggest the use of the entropy – or its opposite, the "negentropy" – as a substitute of the kurtosis. The negentropy provides (i) two rigorous measures of impulsiveness, one in the time domain (on the SE) and another one in the frequency domain (on the envelope spectrum) which (ii) are additive and (iii) whose average is indicative of a distance to a Dirac comb. The related quantities, when displayed as functions of frequency and frequency resolutions, are coined the "SE infogram", the "SES infogram", and the "average infogram". The first one reflects the impulsiveness of a signal, the second one its cyclic content, and the last one its resemblance to a Dirac comb.

The paper is organized as follows. Section 2 revisits the interpretation of the SK from a thermodynamical point of view and addresses its limitation for the detection of repetitive transients. Section 3 introduces the spectral negentropy and the related infograms together with the entropic uncertainty principle. Section 4 provides some recommendations as how to compute and interpret the infograms. Finally, Section 5 illustrates the use of the infograms on synthetic and actual vibration signals.

2. Characterization of repetitive transients: The interplay between non-stationarity and non-Gaussianity

The type of faults whose detection is of interest in this paper produces a repetition of transients. This is a very distinctive signature that combines the characteristic of impulsiveness on the one hand with that of cyclostationarity – i.e. periodicity of the energy – on the other hand. It therefore makes sense to exploit these two features conjointly in designing a detection strategy. Interestingly, impulsiveness is one manifestation of nonstationarity that is usually detected by means of nonlinear tools – i.e. tools dedicated to the analysis of non-Gaussian signals – such as the SK. This section attempts to draw connections between the properties of impulsiveness, nonlinearity, and cyclostationarity through the idea of energy fluctuations of a signal.

2.1. Informative events

Just as for music or speech signals, information in vibration signals is usually conveyed by nonstationary events. In the context of condition monitoring, it means that any change in the structure of the signal will possibly disclose information about the underlying system. Of concern here are abnormal transients generated by incipient faults, which will locally change the frequency spectrum and the amplitude of the signal at the times of their occurrence. That is, the energy flow of the signal will suddenly experience unexpectedly large fluctuations in some frequency bands. At the same time, this is accompanied by a gain of information. This vision has interesting thermodynamical roots. Let us first start by a short reminder of the concept of entropy which is central to both thermodynamics and the theory of information. For a discrete stochastic process that is allowed to explore a set of states indexed by i, with probabilities p_i, the entropy is generally defined as $H = -\sum_i p_i \ln p_i$. This definition is common to statistical thermodynamics (Boltzmann's entropy) and signal processing (Shannon's entropy). It might subsequently differ in how the probabilities p_i are defined in each case. In thermodynamics, p_i is the probability of occupation of state i in the phase space spanned by the degrees of freedom of the system, for instance the positions and the velocities of all particles in the system. In signal processing, p_i is usually the probability that the signal takes a given amplitude A_i.

In thermodynamics, the less-informative macro-state of a system is that which allows its degrees of freedom to explore the maximum number of configurations, or "completions". This is the state of equilibrium – or thermalization – where the strength of the energy fluctuations is uniform over the whole system, characterized by a constant temperature and maximum entropy—see Fig. 1(a). In signal processing terminology, the system is stationary. Information is produced when, for some reason, the system is momentarily driven outside its state of equilibrium and in particular when this is accompanied by the production of structured patterns—Fig. 1(b). In that situation, the system experiences deviations away from its mean energy level until the perturbation eventually dies out. At the same time, entropy varies away from its maximum level by an amount that is indicative of the degree of structure that has been produced. Because any departure from equilibrium must have a cause, the decrease of entropy is also indicative of the action of an external force acting on the system and thus of an increase of knowledge about its status. In signal processing terminology, the system has undergone a transient and therefore has become nonstationary. This brief analogy with thermodynamics unveils the profound links between nonstationarity, energy fluctuations, entropy, and information that underline this paper. It will also justify the specific definition of entropy to be introduced in Section 3.

2.2. The signature of transients

2.2.1. Out-of-equilibrium fluctuations

In light of the above discussion, a transient corresponds to a departure from equilibrium of a system away from its state of equipartition of energy and maximum entropy. In order to assess the production of information associated with it, it thus

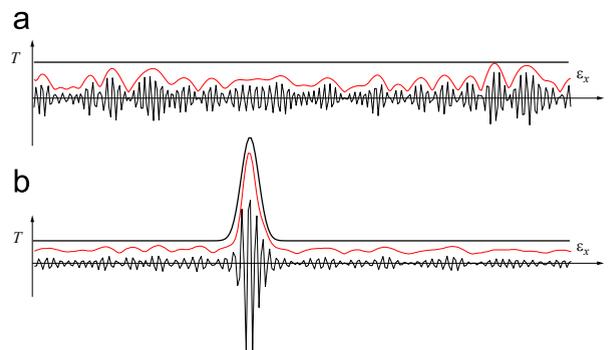**Fig. 1.** (a) Thermalization corresponds to the state of equilibrium of a system, characterized by the equipartition of the energy ε_x, a uniform temperature T, and maximum entropy: the system is stationary. (b) A transient drives the system out of its equilibrium state, generates energy fluctuations, lowers the entropy and produces information associated with structured patterns: the system is nonstationary.
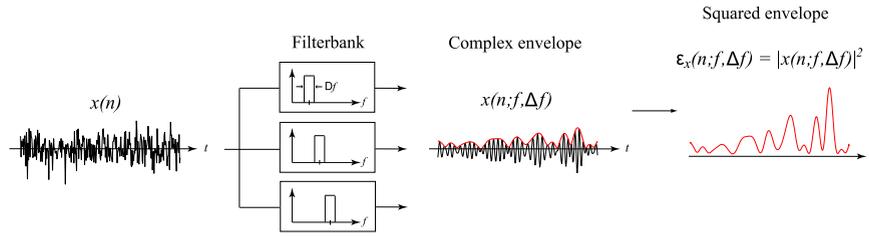
Fig. 2. Illustration of the concept of squared envelope as a mean of measuring the energy fluctuations in a signal (reproduced from Ref. [26]).

makes sense to measure the energy fluctuation in the signal. Formally, let consider a discrete-time signal, $x(n)$, $n = 0, ..., L-1$, of length L and let $x(n; f, \Delta f)$ denotes its complex envelope in a frequency band $[f - \Delta f/2; \ f + \Delta f/2]$—see Fig. 2. The instantaneous flow of energy in the band is returned by the SE

$$\varepsilon_x(n; f, \Delta f) = |x(n; f, \Delta f)|^2 \tag{1}$$

and its average value is

$$\bar{\varepsilon}_x(f, \Delta f) = \langle \varepsilon_x(n; f, \Delta f) \rangle = \frac{1}{L} \sum_{n=0}^{L-1} \varepsilon_x(n; f, \Delta f) \tag{2}$$

which in the thermodynamic analogy, equals $k_B T_S$, the system temperature T_S times the Boltzmann constant k_B. Therefore, the strength of the energy fluctuations may be measured by the variance of the energy flow, that is

$$\text{Var}\{\varepsilon_x(n; f, \Delta f)\} = \langle \varepsilon_x(n; f, \Delta f)^2 \rangle - \bar{\varepsilon_x}(f, \Delta f)^2, \tag{3}$$

or its normalized version, the (squared) coefficient of variation

$$V_x(f; \Delta f) = \frac{\text{Var}\{\varepsilon_x(n; f, \Delta f)\}}{\bar{\varepsilon}_x(f, \Delta f)^2} = \frac{\langle \varepsilon_x(n; f, \Delta f)^2 \rangle}{\bar{\varepsilon}_x(f, \Delta f)^2} - 1. \tag{4}$$

When the system has thermalized (i.e. it has reached its thermodynamical equilibrium), it is known from the principle of maximum entropy [34] that the energy fluctuations are distributed according to a Chi-squared probability law—or equivalently the complex envelope is distributed according to a circular complex Gaussian law. Therefore, $V_x(f; \Delta f) = 1$. Alternatively, in the non-equilibrium regime resulting from the presence of a transient, the energy fluctuations are characterized by $V_x(f; \Delta f) > 1$. This is the signature to be tracked in the signal as a function of frequency f and frequency resolution Δf.

2.2.2. Connection of nonstationarity and nonlinearity: The spectral kurtosis

Interestingly, for a zero-mean signal the normalized variance of the energy flow is closely linked to the SK; specifically,

$$K_x(f; \Delta f) = V_x(f; \Delta f) - 1 = \frac{\langle |x(n; f, \Delta f)|^4 \rangle}{\langle |x(n; f, \Delta f)|^{22} \rangle} - 2 \tag{5}$$

where it is reminded that the kurtosis of a complex variable subtracts 2 instead of 3 for a real variable. It is seen that the kurtosis is advantageously nil when the signal is stationary. This equality draws a correspondence between nonstationarity on the one hand – which is measured by the variance of the energy flow, $V_x(f; \Delta f)$ – and nonlinearity on the other hand – which is measured by the SK $K_x(f; \Delta f)$. The origin of this observation is the following. In the proposed setting, all statistics have been defined according to the time averaging operator instead of the ensemble average (i.e. the expected value), which is justified under the ergodic assumption (ergodicity guarantees that all time averages converge to deterministic quantities, yet not necessarily the same ones as in the ensemble average). As a consequence, nonstationary events such as transients squeeze the time histogram of the signal into a super-Gaussian density (leptokurtic) with a positive excess kurtosis. Loosely speaking, the type of nonstationarity addressed in this work is thus equivalent to non-Gaussianity (i.e. nonlinearity) "along the time axis". The concept was formalized by means of "conditionally non-stationary" processes in Ref. [8]. The conclusion of this subsection is that the information released by transient energy fluctuations may be equivalently addressed in the nonstationary or the nonlinear settings depending on the points of view.

2.2.3. The kurtogram

In practice, the SK is found strongly dependent on the frequency resolution Δf. Ideally, $1/\Delta f$ should be longer than the relaxation time (or duration) τ of transients and shorter than the mean spacing $\bar{T}$ between them, viz: $\tau < 1/\Delta f < \bar{T}$. For this reason, it is often convenient to scrutinize the SK – or any related quantity such as those to come later—in two dimensions as a function of both f and Δf. This leads to the concept of "kurtogram" [10] and, in Section 4, of "infogram", which will play a central role in this paper.

2.3. The signature of repetitive events

2.3.1. The squared envelope spectrum

So far the former discussion did not specifically account for a possible structure into the occurrence of transients. Indeed, the SK is ideally suited to the detection of a single or a few transients, but not necessarily to the situation when their rate of repetition becomes high. By way of an example, let us consider a series of transients with relaxation time $1/\Delta f$ and mean rate of occurrence equal to p events per second. Then it can be shown (see Appendix A.1) that the SK is

$$K_x(f; \Delta f) \sim \frac{\Delta f}{p}, \tag{6}$$

which falls to zero as p increases (see Ref. [20] for a related discussion on the scalar kurtosis). Into addition, no account is taken of the *periodic* timing of repetitive transients (the above formula applies equally to a Poisson random point process or to a Dirac comb), although this is a distinctive feature of faults in rotating machines. Note that repetitive faults are not periodic in general, but rather cyclostationary, which means that their energy flow instead of their waveform is periodic. An extreme example is given by white noise endorsing a periodic amplitude modulation, which is surely not a periodic signal, yet still exhibiting a marked rhythm. As demonstrated in several publications [29,30], the discrete Fourier transform (DFT) of the energy flow – the "squared envelope spectrum" (SES) – thus evidences a spectrum of harmonics equispaced by the fault frequency α_0, i.e.

$$E_x(\alpha; f, \Delta f) = \sum_{n=0}^{L-1} \varepsilon_x(n; f, \Delta f) e^{-j2\pi \frac{\alpha n}{F_s}} = \sum_{i \in \mathbb{Z}} E_{x,i}(f, \Delta f) \delta(\alpha - i\alpha_0) \tag{7}$$

where δ is the discrete Dirac symbol (i.e. $\delta(\alpha)$ is nil everywhere except at $\alpha = 0$ where it is one), $E_{x,i}(f, \Delta f)$ stands for the Fourier coefficient of $\varepsilon_x(n; f, \Delta f)$ at frequency $i\alpha_0$, and F_s is the sampling frequency. This opens the door to a family of cyclostationary techniques based on the exploration of the energy distribution $E_x(\alpha; f, \Delta f)$ spanned by the "spectral" (or "carrier") frequency f and the "cyclic" (or "modulation") frequency α [26,31]. A different route is taken in this paper after unveiling the connection between the SES and the SK.

2.3.2. Connection between the SES and the SK

The SES aims at detecting periodicity in the energy fluctuations of a signal and therefore computes the Fourier coefficients of $\varepsilon(n; f, \Delta f)$ instead of its variance as done for the SK—compare Eqs. (5) and (7). Interestingly, the SES amounts to a decomposition of the SK in the cyclic frequency domain as recently shown by Borghesani et al. [32]. Using Parseval identity on the DFT, it readily comes that

$$\frac{\left\langle |\varepsilon_x(n; f, \Delta f)|^2 \right\rangle}{\left\langle \varepsilon_x(n; f, \Delta f)^2 \right\rangle} = \frac{\left\langle |E_x(\alpha; f, \Delta f)|^2 \right\rangle}{|E_x(0; f, \Delta f)|^2} \tag{8}$$

where the summation $\left\langle |E_x(\alpha; f, \Delta f)|^2 \right\rangle = L^{-1} \sum_{k=0}^{L-1} \left\langle |E_x(\alpha_k; f, \Delta f)|^2 \right\rangle$ is taken over the DFT frequencies $\alpha_k = kF_s/L$. Formula (8) draws a remarkable connection between the SK and the SES for cyclostationary signals: upon inserting Eq. (7) into Eq. (8), it is seen that

$$K_x(f; \Delta f) = \sum_{i \in \mathbb{Z}} \frac{|E_{x,i}(f, \Delta f)|^2}{|E_{x,0}(f, \Delta f)|^2} - 2 = 2 \sum_{i > 0} \frac{|E_{x,i}(f, \Delta f)|^2}{|E_{x,0}(f, \Delta f)|^2} - 1, \tag{9}$$

that is the SK is the sum of the harmonics in the SES normalized by the squared mean energy. Since it grows with the number of harmonics in the SES, the SK actually returns a measure of impulsiveness in the time domain. For the sake of coherence with the discussion of the previous subsection, let us take the example of a series of repetitive transients with relaxation time $1/\Delta f$ and rate of repetition p (the signal is now cyclostationary). Since the corresponding SES is a Dirac comb with spacing p and bandwidth Δf, it comes that there are $\Delta f/p$ peaks of magnitude $|E_{x,i}(f, \Delta f)|^2 \sim |E_{x,0}(f, \Delta f)|^2$ in the frequency band $[-\Delta f/2 < \alpha \le \Delta f/2]$; thus, according to Eq. (9), $K_x(f; \Delta f) \sim \Delta f/p$, which is identical to result (6) and means that the SK actually counts the number of prominent harmonics in the SES. The latter interpretation of the SK is one step forward to the notion of a measure of information to be discussed in Section 3.2.

2.3.3. The protrugram

As demonstrated in the previous subsection, the detection of repetitive transients may be well achieved by testing for the presence of harmonics in the SES. Of obvious interest is to proceed in this direction without having to know explicitly the fault frequencies entering in Eq. (7). In this respect, advantage can be taken of the property that harmonics in the SES may be seen as transients associated with high fluctuations of the energy flow along the cyclic frequency (α) axis. As a consequence, the harmonic structure of the SES may be detected – and the information coming with it measured – following similar lines as for transients in the time domain. The idea was first formulated in Ref. [21] where it was proposed to compute the kurtosis of the SES, this leading to the concept of "protrugram". A consensual definition of the protrugram is still an open question, since the kurtosis may be applied either on the complex Fourier transform, or on its magnitude, its squared magnitude, etc. The protrugram also suffers from the

same drawback as the kurtogram: it is maximum for a single harmonic in the SES, that is a pure tone (or purely sinusoidal modulation) in the time domain. Therefore it is still not ideal for the detection of repetitive transient, but it suggests an interesting direction towards this objective.

3. Spectral negentropy: Measuring information in energy fluctuations

3.1. Motivations

The previous sections have pointed out that the SK, despite being ideally suited to the detection of a single impulse, is not optimal for the detection of repetitive transients. Besides, the concept of protrugram has highlighted that the signature of repetitive transients in time is also a series of repetitive transients in frequency (i.e. harmonics in the SES). Formally, the invariant is a Dirac comb which has the same structure in both domains.

One could think of devising a combination of time-domain and frequency-domain SKs (of the kurtogram and protrugram), such that a maximum value is reached when one (or both) of them is (are) high. However, as far as the author knows, a mathematical basis for such a construction is missing and there is no justification that its maximization would generally correspond to the requested signal. Instead, this paper proposes to replace the kurtosis by another measure of impulsiveness, the negentropy, which happens to provide an uncertainty principle on the balance of impulsiveness in the time and frequency domains. It will also happen to be maximized by the ideal Dirac comb.

3.2. Entropy and information

Entropy is a versatile concept that measures the disorder in a system. It was first introduced by Clausius in the 1850s to appraise the quantity of energy "burnt" as heat and remained quite an abstract concept of thermodynamics until Boltzmann showed in the 1870s that entropy actually counts the number of configurations the degrees of freedom of a system can take given some constant macroscopic variables such as temperature. In particular, it is a fundamental principle of statistical thermodynamics that a system reaches maximum entropy when it is in equilibrium. The concept was later rediscovered in the 1940s by Shannon in the context of communication theory to measure the amount of information transmitted by a channel. The equivalence between thermodynamical entropy and information was eventually recognized by Brillouin [33]. Referring back to our introductory discussion in Section 2.2, it makes sense that entropy could be just as good a measure as kurtosis to detect out-of-equilibrium perturbations in a system and to measure the information associated with them.

The proposed definition of entropy proceeds from interpreting the (square of the) instantaneous energy flow in a signal as a probability distribution. Note in passing that this is coherent with the quantum mechanics interpretation of the squared magnitude of a wave function as the probability density of finding a particle in a given place at a given time. Hence, the "spectral entropy" in frequency band $[f - \Delta f/2; f + \Delta f/2]$ reads

$$H_\varepsilon(f; \Delta f) = -\left\langle \frac{\varepsilon_x(n; f, \Delta f)^2}{\left\langle \varepsilon_x(n; f, \Delta f)^2 \right\rangle} \ln\left(\frac{\varepsilon_x(n; f, \Delta f)^2}{\left\langle \varepsilon_x(n; f, \Delta f)^2 \right\rangle} \right) \right\rangle \tag{10}$$

where care has been taken to normalize the energy flow by its averaged value as requested for $\varepsilon_x(n; f, \Delta f)^2 / \left\langle \varepsilon_x(n; f, \Delta f)^2 \right\rangle$ to be interpreted as a probability distribution (the reason for taking the square of the energy in the definition of entropy is to allow the introduction of the Hirschman's uncertainty principle in the next section). With this definition, it can be shown using calculus of variation that the maximum value of entropy $H_\varepsilon(f; \Delta f)$ is reached when the energy flow is constant, which is fully coherent with the thermodynamic interpretation of Section 2.2— see Fig. 1. On the contrary, the lowest entropy (an unbounded quantity according to the proposed definition) happens when the energy flow condense to a single impulse. The spectral entropy thus shows an opposite behavior to the SK; according to Eq. (10), it can actually be seen as a version of the SK weighted by $-\ln\left(\varepsilon_x(n; f, \Delta f)^2 / \left\langle \varepsilon_x(n; f, \Delta f)^2 \right\rangle \right)$. To be completely analogous to the SK, the negative of the spectral entropy should thus be considered—see Fig. 3. This defines the "spectral negentropy":

$$\Delta I_\varepsilon(f; \Delta f) = -H_\varepsilon(f; \Delta f). \tag{11}$$

The concept of negentropy—a contraction for "negative entropy" – was initially elaborated by Schrödinger in the 1940s. It represents the inclination of a system to increase its degree of organization. As demonstrated by Brillouin [33], negentropy is equivalent to a gain of information – hence the proposed notation ΔI – which may be construed as the logarithm of the number of configurations that the experimenter has learned in attempting to describe the state of the system. This brings us back to the discussion of Section 2.3.2 where information was approached with the occurrence of a transient. Spectral negentropy can be similarly defined in the frequency domain on the SES, i.e. as

$$\Delta I_E(f; \Delta f) = -H_E(f; \Delta f) = \left\langle \frac{|E_x(\alpha; f, \Delta f)|^2}{\left\langle |E_x(\alpha; f, \Delta f)|^2 \right\rangle} \ln\left(\frac{|E_x(\alpha; f, \Delta f)|^2}{\left\langle |E_x(\alpha; f, \Delta f)|^2 \right\rangle} \right) \right\rangle, \tag{12}$$
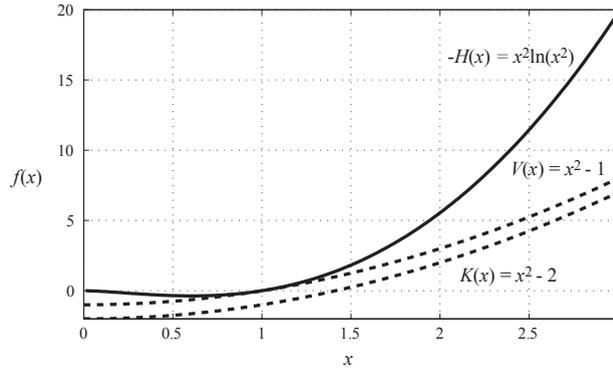
Fig. 3. Comparison of the functional forms of the energy variance **V(x)**, the kurtosis **K(x)**, and the negentropy **–H(x)**.

Table 1
Expressions of the spectral negentropies and their average for some typical signals sampled at rate F_s (see Appendix A.3). ($\delta_{\Delta f}$ stands for an impulse of temporal width $1/\Delta f$, $\Pi_{\Delta f}$ for the unitary rectangle function with width Δf, and $\gamma \approx 0.5772$ is Euler constant)

Squared envelope (SE)	Squared envelope spectrum (SES)	$\Delta I_\varepsilon(f;\Delta f)$	$\Delta I_E(f;\Delta f)$	Infogram $\Delta I_{1/2}(f;\Delta f)$
$\varepsilon_x(n;\Delta f)$ is WCCGN[a]	$E_x(\alpha)$ is WCCGN	$1-\gamma \approx 0.422$	$1-\gamma \approx 0.4228$	$1-\gamma \approx 0.4228$
$\varepsilon_x(n;\Delta f)$ is the SE of WCCGN	No simple expression, but asymptotically WCCGN	$3-2\gamma-\ln 2 \approx 1.1525$	≈ 0.4228 (asymptotically)	≈ 1.5752 (asymptotically)
SE is sinusoidal (i.e. the signal is cyclostationary): $\varepsilon_x(n;\Delta f) = A\,\cos\left(2\pi\frac{\alpha_0}{F_s}n+\phi\right)$	$E_x(\alpha) = \frac{1}{2}LAe^{j\phi}\delta(\alpha-\alpha_0),\ \alpha \geq 0$	$\ln\left(\frac{\Delta f}{F_s}\right)$	$\ln\left(L\frac{F_s}{\Delta f}\right)$	$\frac{1}{2}\ln L$
SE of a single impulse: $\varepsilon_x(n;\Delta f) = A\delta_{\Delta f}(n-n_0)$	$E_x(\alpha) = Ae^{-j2\pi n_0\frac{\alpha}{F_s}}\Pi_{\Delta f}(\alpha)$	$\ln\left(L\frac{\Delta f}{F_s}\right)$	$\ln\left(\frac{F_s}{\Delta f}\right)$	$\frac{1}{2}\ln L$
SE of a Dirac comb: $\varepsilon_x(n;\Delta f) = A\sum_{m=0}^{K-1}\delta_{\Delta f}(n-mN)$ with $L=KN$	$E_x(\alpha) = KA\sum_{p=0}^{N-1}\delta(\alpha-pKF_s)\Pi_{\Delta f}(\alpha)$	$\ln\left(N\frac{\Delta f}{F_s}\right)$	$\ln\left(K\frac{F_s}{\Delta f}\right)$ with $K=L/N$	$\frac{1}{2}\ln L$

[a] WCCGN: white circular complex Gaussian noise. A circular complex Gaussian random variable, $z=a+jb$, is such that a and b are independently and identically distributed according to a real Gaussian.

where the average over the cyclic frequency range α is as defined in Eq. (8). The minimum value of $\Delta I_E(f;\Delta f)$ corresponds to a flat SES (i.e. an impulse in the time domain) and the maximum to a single harmonic in the SES (i.e. a pure sinusoidal SE). This is the reverse behavior of $\Delta I_\varepsilon(f;\Delta f)$.

Expressions of $\Delta I_\varepsilon(f;\Delta f)$ and $\Delta I_E(f;\Delta f)$ are given in Table 1 for some typical signals. It is seen that they are measures of the degree of structure (or organization) of a signal in the time and in the frequency domains, respectively. As a consequence, the spectral negentropy $\Delta I_\varepsilon(f;\Delta f)$ and $\Delta I_E(f;\Delta f)$ should be considered jointly in order to have a full discrimination between different types of signals.

3.3. Hirschman's entropic uncertainty principle

From the above discussion, it is now clear that a signal made of repetitive transients will produce high values of the spectral negentropy in both the time and the frequency domains. Because entropies are extensive quantities, it makes sense to consider the weighted average

$$\Delta I_\rho(f;\Delta f) = \rho\Delta I_\varepsilon(f;\Delta f) + (1-\rho)\Delta I_E(f;\Delta f) \tag{13}$$

for some $0 \leq \rho \leq 1$. Interestingly, this quantity achieves an upper bound for $\rho = 1/2$. Using a result due to Hirschman [35–39], it can be shown that the average spectral negentropy has a bounded information capacity that grows with the signal length, i.e.

$$\Delta I_{1/2}(f;\Delta f) \leq \frac{1}{2}\ln L \tag{14}$$

where the upper-bound is achieved by a Dirac comb[1] (see Appendix A.2). This result is of fundamental importance, since a Dirac comb is the idealization of the repetitive transients that are of interest herein. Although it is very unlikely to occur in practice (because it would require its period to be an exact multiple of the sampling period and the acquisition time to be an

---

[1] It is discussed in Refs. [38,39] that for continuous-time (finite-energy) signals, the upper-bound is achieved by the Gaussian pulse whereas, quite surprisingly, for discrete signals it is achieved by a Dirac comb.

exact multiple of the period), it gives an indication that repetitive transients are close to the maximum information possibly returned by $\Delta I_{1/2}(f;\Delta f)$, that is $0.5\ln L$ [nats][2].

Expressions of the spectral entropies are reported in Table 1 for some typical signals. It is noteworthy that, in accordance with the theorem of Appendix A.2, the single impulse and the sinusoidal SE (as returned by a cyclostationary signal) also produce maximum information. The single impulse is actually a particular instance of a Dirac comb whose period equals the signal length; the sinusoidal SE is the dual signal in the frequency domain. These two cases can be easily discriminated by comparing their spectral negentropies, $\Delta I_\varepsilon(f;\Delta f)$ and $\Delta I_E(f;\Delta f)$, as illustrated in Fig. 4. In particular, $\Delta I_\varepsilon(f;\Delta f)$ is found independent of the signal duration for a pure sine while it grows with the number of samples L for a single impulse, whereas the opposite behavior holds true for $\Delta I_E(f;\Delta f)$.

Next, the behaviors of the spectral entropies are illustrated by means of numerical simulations. The following signals are synthesized for various lengths L:

1) a white Gaussian random noise,
2) a non-linear noise defined as the 5th power of white Gaussian random noise,
3) a second-order cyclostationary signal defined as the multiplication of white Gaussian random noise with a pure sine with period $N=30$,
4) a pure Dirac impulse,
5) a Dirac impulse in white Gaussian noise with crest factor 50,
6) a Dirac comb with period $N=30$,
7) a Dirac comb with period $N=30$ in white Gaussian noise with SNR$=20$ dB, and
8) a Dirac comb with 3% random fluctuations around the mean period $N=30$ (see Ref. [40]).

The corresponding spectral negentropies ΔI_ε and ΔI_E are displayed in full band $\Delta f = F_s$ in Fig. 4(a) and (b), respectively, together with their average $\Delta I_{1/2}$, in Fig. 4(c). It is seen that the simulations correctly follows the theoretical results of Tab. 1. White Gaussian noise correctly takes values $\Delta I_\varepsilon \approx 1.15$ and $\Delta I_E \approx 0.42$, whereas non-linear noise exhibits a dependence of ΔI_ε on the signal length L that reveals its impulsive nature. Cyclostationary noise has a constant ΔI_ε but evidences an increase of ΔI_E with the signal length because of the presence of harmonics in its SES. The Dirac impulse has ΔI_ε linearly increasing with $\ln L$ and a constant ΔI_E. Given a constant crest-factor, the linear increase of ΔI_ε drops down in the presence of noise after a certain signal duration that correspond to $SNR \approx -25dB$. The Dirac comb correctly shows a linear increase of ΔI_E like $\ln K = \ln L - \ln N \approx \ln L - 3.4$, which is slightly affected by moderate noise (SNR$=20$ dB). However, even a small random fluctuation (3%) of the comb period significantly lowers the theoretical value of ΔI_E; this is because it has a low-pass effect in the SES that limits the number of harmonics. Finally, the average $\Delta I_{1/2}$ clearly singles out the Dirac comb and the Dirac impulse as its maximizers, as expected from Hirschman's
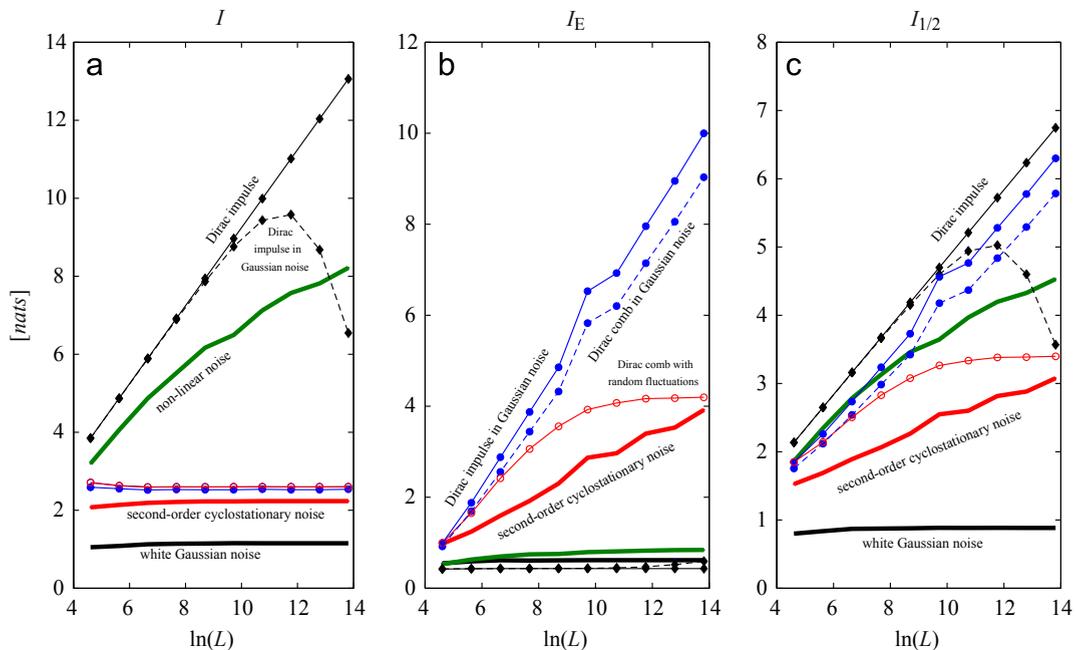

Fig. 4. Spectral negentropies: (a) ΔI_ε, (b) ΔI_E and (c) their average $\Delta I_{1/2} = (\Delta I_\varepsilon + \Delta I_E)/2$ in full band $\Delta f = F_s$. Black thick line: white Gaussian noise; green thick line: non-linear noise; Red thick line: second-order cyclostationary noise ($N=30$); black line with diamonds: Dirac impulse; black dotted line with diamonds: Dirac impulse in Gaussian noise (crest-factor$=50$); blue line with circles: Dirac comb ($N=30$); blue dotted line with circles: Dirac comb in Gaussian noise ($N=30$, SNR$=20$ dB); red line with circles: Dirac comb with 3% random fluctuations around the mean period $N=30$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

uncertainty principle. Non-linear noise and cyclostationary noise also correspond to high values in $\Delta I_{1/2}$, which evidences the presence of informative transients and structured non-stationarities, respectively.

4. Recommendations for computing and interpreting the infogram

4.1. Analysis strategy: The infogram

Just as the kurtogram displays the values of the spectral kurtosis for different frequency resolutions Δf's, the proposed strategy is to display the spectral negentropies $\Delta I_\epsilon(f; \Delta f)$, $\Delta I_E(f; \Delta f)$, and $\Delta I_{1/2}(f; \Delta f)$ in the $(f, \Delta f)$ plane as an image. This defines the "SE infogram", the "SES infogram", and the "average infogram", respectively. This way, impulsive events can be detected and localized in frequency by high values of the SE infogram $\Delta I_\epsilon(f; \Delta f)$ in some frequency bands $[f - \Delta f/2; f - \Delta f/2;]$, repetitive (cyclostationary) events by high values of the envelope spectrum $\Delta I_E(f; \Delta f)$, and events that are closed to maximize Hirschman's uncertainty principle by high values of the average infogram $\Delta I_{1/2}(f; \Delta f)$.

4.2. Algorithmic implementation

Technically speaking, the algorithmic implementation of the infogram follows similar lines as for the SK. The first task is to design a filterbank that decomposes the signal through a series of subbands. Various architectures are possible as demonstrated in the literature, such as multirate filters, wavelets, wavelet packets, dual tree wavelets, etc [12–18]. In this paper, an implementation based on the short-time Fourier transform (STFT) is used due its simplicity and high flexibility. Namely, let $x(n)$ be sampled at frequency F_s and $w(n)$ an analysis window (symmetric and positive) of length $N_w = \lceil F_s/\Delta f \rceil$ (where the brackets extract the closest integer to the number they are embracing). Then, the squared magnitude of the short-time Fourier coefficient,

$$\epsilon_x(n; f_k, \Delta f) = \left| \sum_{m=nR}^{N_w+nR-1} x(m)w(m-nR)e^{-j2\pi\frac{f_k}{F_s}m} \right|^2, \tag{15}$$

returns an estimate of the energy flow at time $t = nR/F_s$ and frequency $f_k = kF_s/N_w$, with $\Delta f \sim F_s/N_w$. The advantage of the STFT is to easily allow a cascade of different resolutions Δf's to be explored in the kurtogram or infogram. In this work, the frequency axis is successively segmented according the following law

$$(\Delta f)_i = \frac{F_s}{2^i} \tag{16}$$

where index i stands for the level of the decomposition and is allowed to take non-integer values. For instance, typical values for the STFT lengths are $N_w = 1, 2, 4, 6, 8, 12, 16, 24, 32, 48, \ldots$ (where even values are retained only to speed up the computation of the STFT) which corresponds to a cascade of levels $i = 0, 1, 2, 2.585, 3, 3.585, 4, 4.585, 5, 5.585, \ldots$ The proposed decomposition interleaves divisions of the frequency bands by two and three [10]. In addition, since the bandpass filters of the STFT filterbank are not ideal, there is a significant fraction of overlap between contiguous frequency bands. For all these reasons there is little risk to miss an informative band, even though the paving is rather coarse as compared to a "continuous" decomposition.

At this stage, attention should be paid to the following points:

1) the subband signals should be in the analytic form (i.e. with nil components at negative frequencies), and
2) the energy flow should be critically sampled such that $\epsilon_x(n; f_k, \Delta f)$ remains a white process when the underlying signal is white.

The STFT guarantees point (1) by construction, except at the zero and folding frequencies where the signal is real. One way to fix it is to consider the analytic signal right from the beginning before passing it through the filterbank. Point (2) is fulfilled by setting the value of increment R according to the window length in the STFT: with a Hann window w, this is achieved by taking $R = N_w/4$ except at the zero and folding frequencies where $R = N_w/2$ when the signal is analytic (this is because the spectrum has half-bandwidth there).

The energy flow estimated by Eq. (15) is then used to compute the spectral negentropies according to Eqs. (10)–(12), where the SES is defined as the DFT of the SE as in Eq. (7). Note that it is wise to remove the zero-frequency component, $E_x(0; f, \Delta f) = L\overline{\epsilon_x}$, from the computation of the spectral negentropy (12) because it generally corresponds to a transient in the SES.

---

[2] The nat is the unit of information based on the natural logarithm. It corresponds to $1/\log 2 \approx 1.44$ bits.

4.3. Normalization

Just as the SK is theoretically zero for stationary Gaussian signals, it may be convenient to normalize the infogram such that it is zero under the same circumstance. This amounts to systematically subtracting the theoretical values $\Delta I_\varepsilon \approx 1.1525$ and $\Delta I_E \approx 0.4228$ given in Table 1 for a stationary Gaussian signal. In practice, the computation of the spectral entropies where found extremely sensitive to departure of the filterbank frequency gains from ideal filters and to the discontinuity imposed by the analytical signal at the zero and folding frequencies of the spectrum. In such cases, it may be better recommended to learn the expected values of the infogram under the stationary Gaussian assumption by Monte–Carlo simulations.

Finally, confidence intervals can possibly be computed to test against the null hypothesis of the infogram based on the following approximate formulae of the standard deviation:

$$\begin{cases} \mathrm{std}(\Delta I_\varepsilon) \approx \frac{6.0}{\sqrt{L}} \\ \mathrm{std}(\Delta I_E) \approx \frac{1.5}{\sqrt{L}}. \end{cases} \qquad (17)$$

4.4. Discrete random separation

Because the spectral negentropy of the SES is sensible to cyclostationary signals, it matters to decide whether first-order cyclostationarity (periodic components) or second-order cyclostationarity (random components) is of concern. As discussed in Refs. [23] and [26], the former type usually characterizes gears whilst the latter better characterizes bearings. Thus, whether gear or bearing faults are of concern, the SES infogram should be either computed on the deterministic part of the signal or on its random residual part. Techniques as how to separate these two parts are summarized in Ref. [41] (see also [42] for a modern view of the synchronous average).

4.5. Pre-whitening

It was empirically found that pre-whitening the signals before computing the spectral negentropy can make a real difference. This practice is well-known in detection theory and its benefit was clearly demonstrated in the cyclostationary context where the pre-whitened version of the spectral correlation – the "spectral coherence" – usually better enhances the signal of interest [24]. Recently, Ref. [43] showed that pre-whitening can also help to unmask the weak bearing signals from the strong interfering vibrations of gears. For all these reasons, pre-whitening the signals prior to the computation of the infogram is surely a good recommendation. Pre-whitening can be achieved using several efficient implementations which are not discussed here, such as linear prediction, cepstrum editing, FFT-filters, etc.

A flowchart of the algorithm is given in Fig. 5.

5. Experimental results

5.1. Numerical experiments

This section aims at illustrating the use of the infogram and in particular at demonstrating some of its advantages as compared to the kurtogram. There are at least two situations which the kurtogram has trouble with. The first one is in the presence of impulsive noise and the second one when the rate of repetition of the transients is high with respect to their bandwidth.

5.1.1. Impulsive noise

The following example illustrates the failure of the kurtogram in the presence of impulsive noise. A series of transients is synthesized according to the model of Ref. [40] as the impulse responses of a single-degree-of-freedom system with normalized resonance frequency at 0.2 ($F_s = 1$) and damping ratio of 8.3% to a series of Diracs with mean spacing $N = 120$ with 5% random jitter and 10% random amplitude modulation. This is known to simulate rather realistically the vibration produced by a faulty rolling element bearing. The impulsive noise is similarly modeled as the response of a single-degree-of-freedom system, with normalized resonance frequency at 0.35 and damping ratio of 1.6% to a Dirac impulse located at time instant $n = 1000$. The resulting synthetic signal of length $L = 10^5$ is displayed in Fig. 6(a) and (b) after addition of white Gaussian noise with SNR$= -12$ dB. Fig. 6(c) displays the corresponding power spectral density with highlights of the resonances that carry the series of transients and the impulsive noise.

The kurtogram $K_x(f; \Delta f)$ and the infograms $\Delta I_\varepsilon(f; \Delta f)$, $\Delta I_E(f; \Delta f)$, and $\Delta I_{1/2}(f; \Delta f)$ are displayed in Fig. 7(a–d). It is seen that the impulsive noise at frequency 0.35 dominates both the kurtogram and the SE infogram $\Delta I_\varepsilon(f; \Delta f)$ about levels 7 to 8 and masks the signature of the series of transients. However, the SES infogram $\Delta I_E(f; \Delta f)$ shows a radically different behavior: it is insensitive to the impulsive noise and clearly reveals the series of transients at frequency 0.35 from levels 4 to 6. The comparison of $\Delta I_\varepsilon(f; \Delta f)$ and $\Delta I_E(f; \Delta f)$ is thus indicative of the existence of two informative events with different structures.
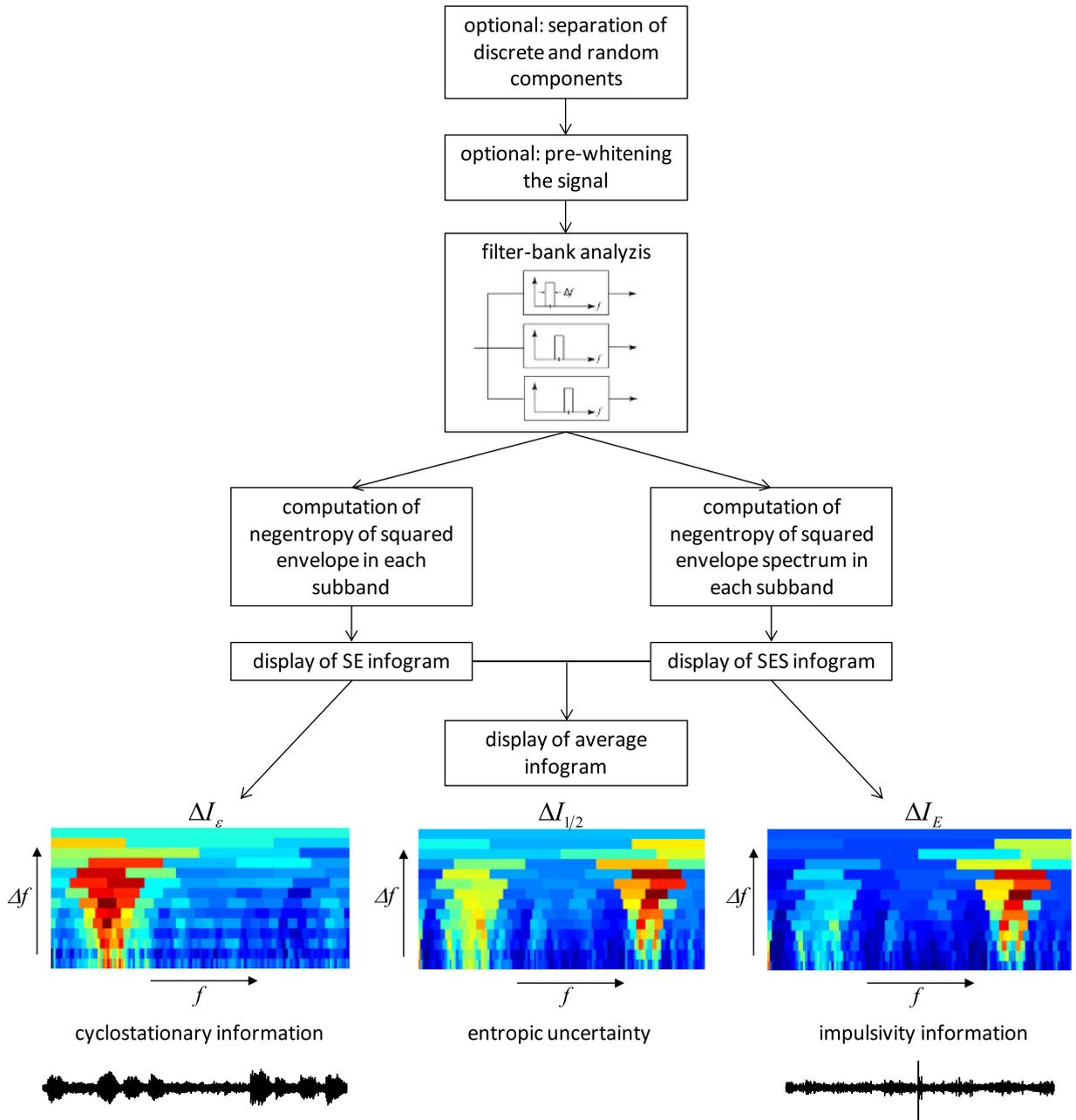
Fig. 5. Flowchart of the algorithm for computing the infograms.

Eventually, the average infogram $\Delta I_{1/2}(f; \Delta f)$ evidences the two events, yet indicating that impulsive noise is closer to the upper bound of the Hirschman's uncertainty principle in this example.

5.1.2. High repetition rate of the transients

Another drawback of the kurtogram is to take smaller and smaller values as the repetition rate of the transient increases. This is reflected by formula (6): $K_x(f; \Delta f) \sim \Delta f/p$. The same behavior holds true for the SE infogram $\Delta I_\varepsilon(f; \Delta f)$ which, according to Tab.1, equals $\ln(\Delta f/p)$ for a Dirac comb with fundamental frequency $p = F_s/N$. However, the opposite trend is observed in the SES infogram $\Delta I_E(f; \Delta f)$ which equals $\ln(Lp/\Delta f)$ in this case so that, in theory, the average infogram $\Delta I_{1/2}(f; \Delta f) = (1/2)\ln L$ is found independent of the comb frequency p. This attracting property is illustrated here on synthetic signals.
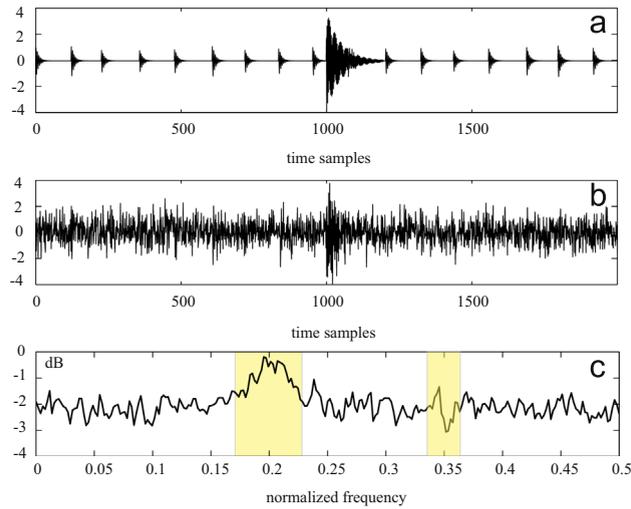
Fig. 6. (a) Synthetic signal simulating the series of transients produced by a faulty rolling element bearing and further corrupted by impulsive noise at time instant $\boldsymbol{n}=1000$. (b) Same signal in additive white Gaussian noise (**SNR** $= -12$ **dB**). (c) Power spectral density of noisy signal (frequency resolution$=0.002$) with highlights of the resonances at 0.2 and 0.35 that carry the series of transients and the impulsive noise, respectively.
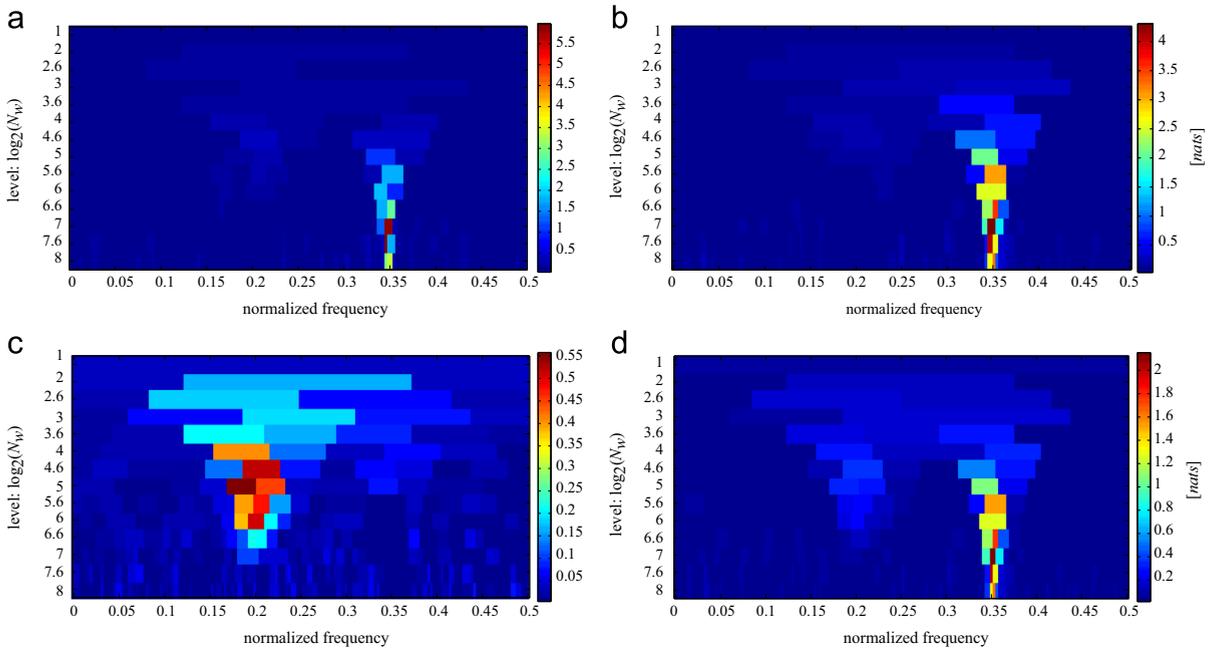

Fig. 7. (a) Kurtogram, (b) SE infogram $\Delta I_{\varepsilon}(\boldsymbol{f};\Delta\boldsymbol{f})$, (c) SES infogram $\Delta I_{E}(\boldsymbol{f};\Delta\boldsymbol{f})$, (d) average infogram $\Delta I_{1/2}(\boldsymbol{f};\Delta\boldsymbol{f})$ of the signal in Fig. 6.

Series of transients are simulated according to the same model as in the previous example with resonance frequency 0.2, damping ratio 8%, and increasing mean periods $N = 50$, 200, 1000. The signals are further buried in white Gaussian noise with constant SNR$= -6$ dB, as displayed in Fig. 8.

The corresponding kurtogram and infograms are displayed in Figs. 9–11. It is seen that with $N = 50$, neither the kurtogram nor the SE infogram can detect the presence of the series of transients, whereas it appears clearly in the SES infogram and in the average infogram—see Fig. 8. As seen in Fig. 8(a), this is a situation with high overlap between successive transients, such that $p \gg \Delta f$. The same conclusion basically holds for $N = 200$, although a close inspection of the kurtogram and SE infogram shows that some information is present—see Fig. 10. As seen in Fig. 8(c), this is the situation where the effective length of the transients happens to coincide with the period, that is $p \sim \Delta f$. Finally, when the transients become well separated, $p \ll \Delta f$—see Fig. 8(e)—the kurtogram and the SE infogram become fully efficient and the average infogram is virtually identical to the latter—see Fig. 11.
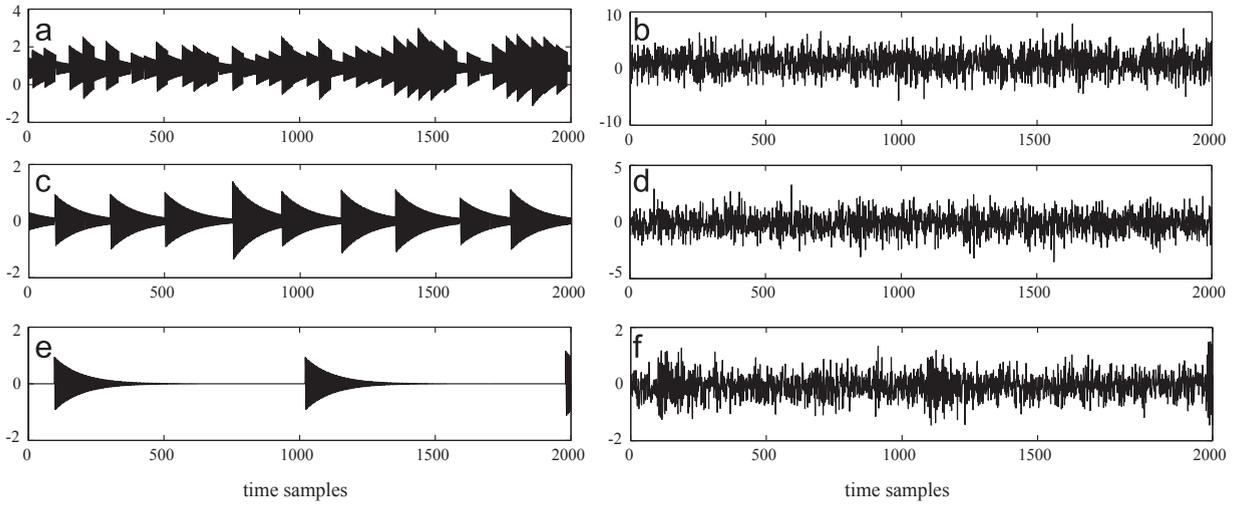
Fig. 8. Synthetic signals simulating the series of transients produced by a faulty rolling element bearing with mean period (a) $N = 50$, (c) $N = 200$, and (e) $N = 1000$. (b, d and f) Corresponding signals after addition of white Gaussian noise with constant **SNR** $= -6$**dB**.
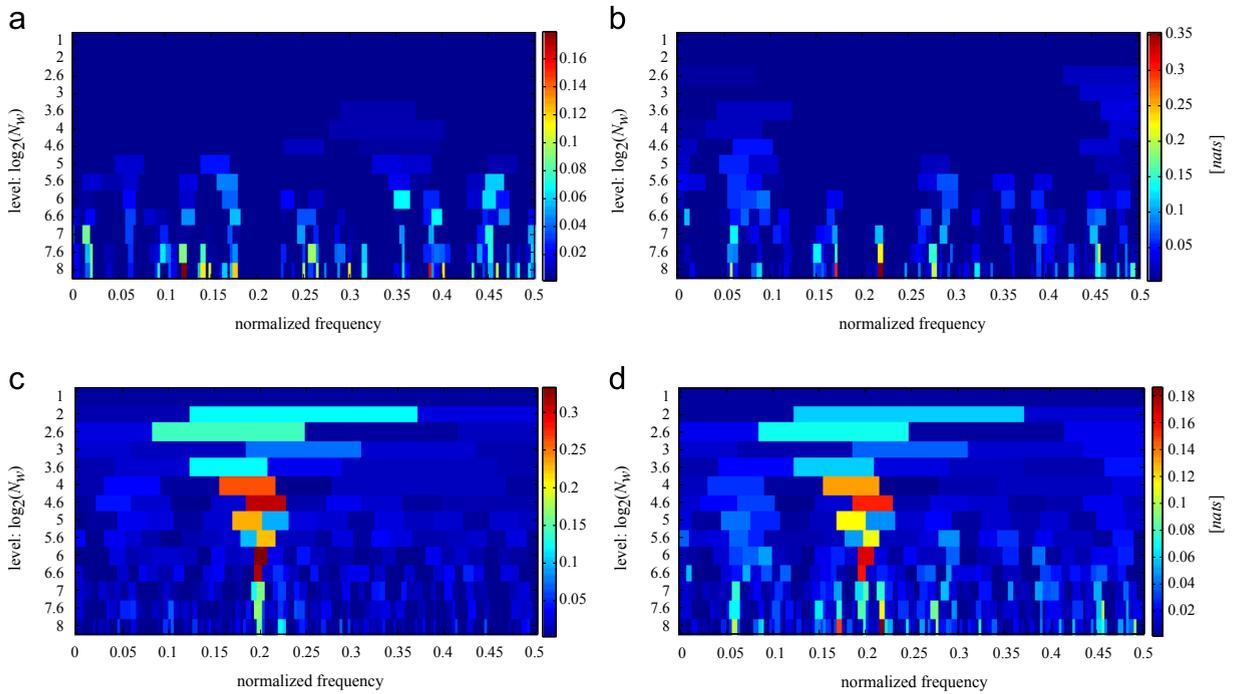

Fig. 9. (a) Kurtogram, (b) SE infogram $\Delta I_\varepsilon(f; \Delta f)$, (c) SES infogram $\Delta I_E(f; \Delta f)$, (d) average infogram $\Delta I_{1/2}(f; \Delta f)$. Case with $N = 50$.

5.2. Vibration-based diagnosis of rolling element bearing: Experiment 1

The kurtogram and infograms are now illustrated on actual vibration signals. In order to allow fair comparison with previously published results, the same data as used in Section 4.2 of Ref. [10] are analyzed here, except that a STFT implementation is used instead of a multirate filterbank, which leads to slightly different results and definition of the analysis levels in Eq. (16)). The signals have been recorded on a test-rig at the University of New South Wales, Sydney. The system is a one-stage gearbox with primary and secondary shafts supported by ball bearings (12 elements, ball diameter = 7.12 mm, ball pitch diameter = 38.5 mm, load angle = 0°). The rotational speed of the secondary shaft is 10 Hz. The sampling frequency is 48 kHz. Two bearing faults are investigated here: an inner race fault with theoretical BPFI = 71.10 Hz and an outer race fault with theoretical BPFO = 48.90 Hz. All signals have been preprocessed as advocated in Sections 4.4 and 4.5: first the synchronous average has been subtracted in order to remove all periodic components due to
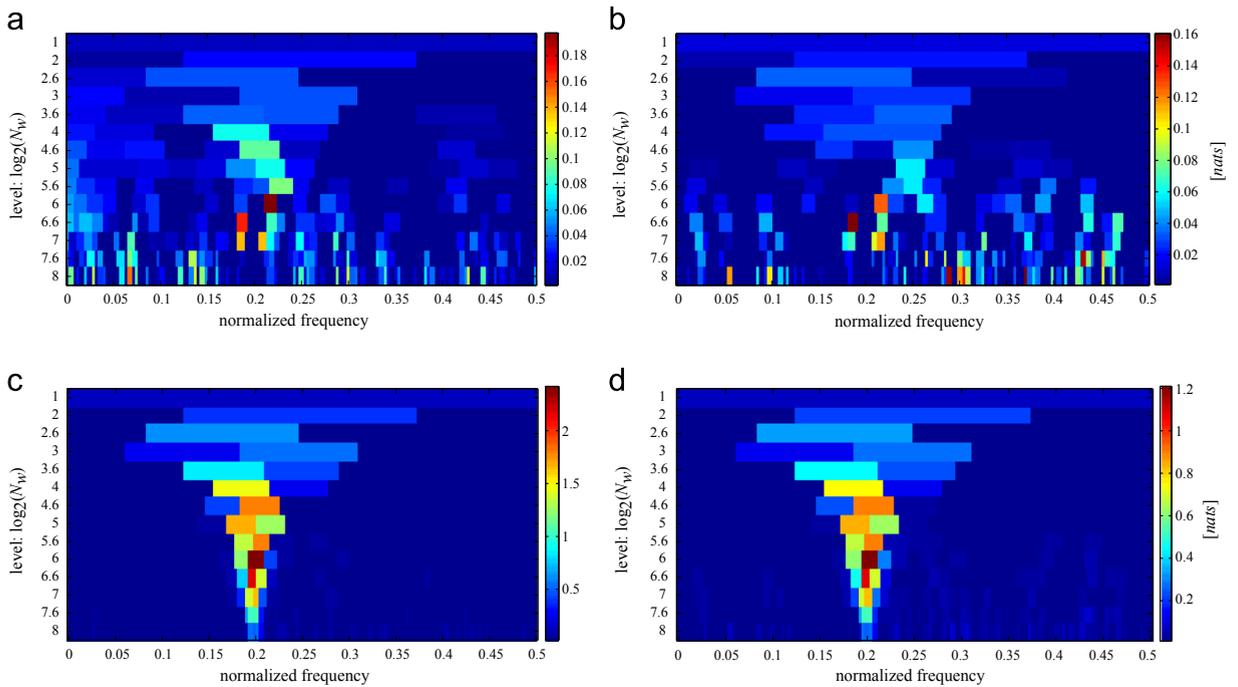
Fig. 10. (a) Kurtogram, (b) SE infogram $\Delta I_e(f; \Delta f)$, (c) SES infogram $\Delta I_E(f; \Delta f)$, (d) average infogram $\Delta I_{1/2}(f; \Delta f)$. Case with $N = 200$.
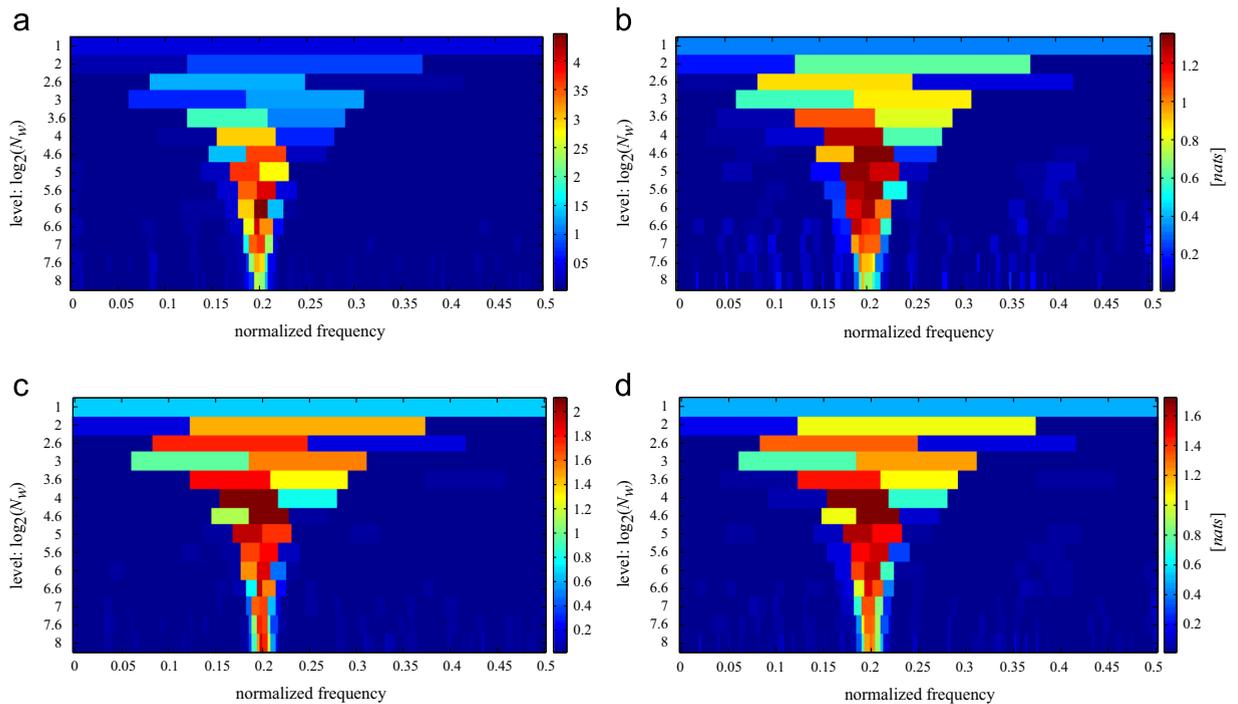

Fig. 11. (a)Kurtogram, (b) SE infogram $\Delta I_e(f; \Delta f)$, (c) SES infogram $\Delta I_E(f; \Delta f)$, (d) average infogram $\Delta I_{1/2}(f; \Delta f)$. Case with $N = 1000$.

gears and therefore to enhance the bearing vibrations; second, the residual signals have been pre-whitened (note these two steps lead to further differences as compared to the results of Ref. [10]).

The kurtogram and infograms in healthy state are displayed in Fig. 12(a–d). On the one hand, the kurtogram and especially the SE infogram evidence high values below level 3 around 20 kHz. Because there is no equivalence in the SES infogram, the corresponding event is expected to be impulsive noise. This is confirmed by filtering the signal in band
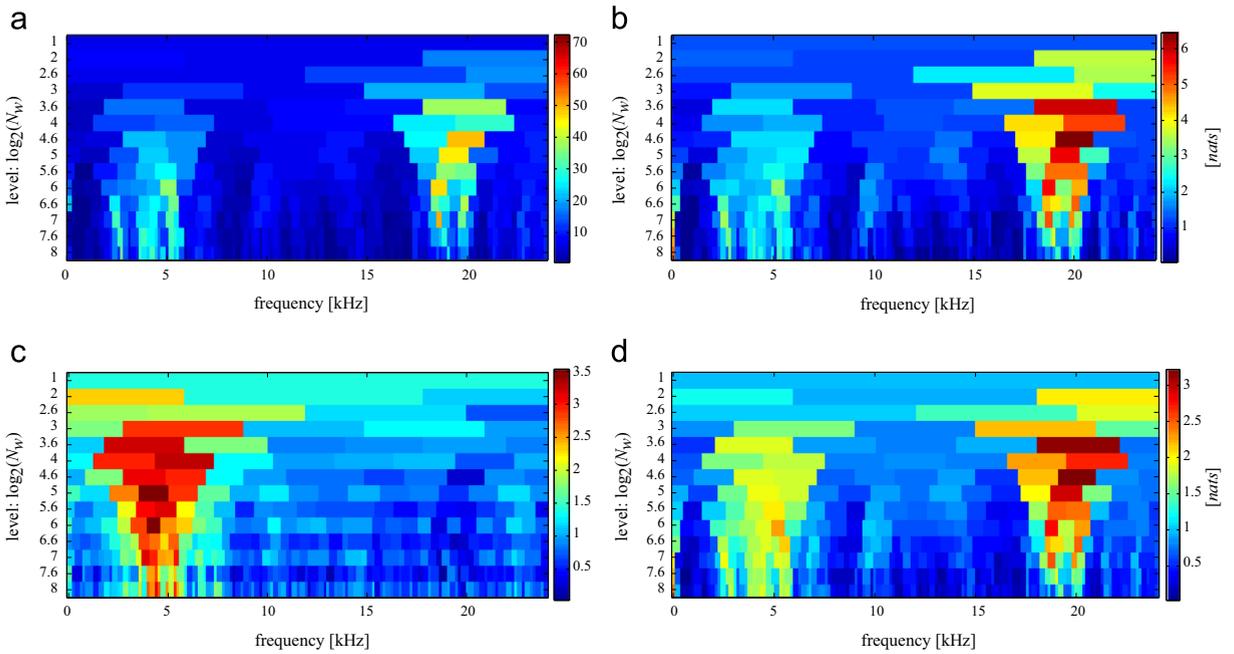
Fig. 12. Analysis of a vibration signal from a healthy bearing. (a) Kurtogram, (b) SE infogram $\Delta I_\epsilon(f; \Delta f)$, (c) SES infogram $\Delta I_E(f; \Delta f)$, (d) average infogram $\Delta I_{1/2}(f; \Delta f)$.
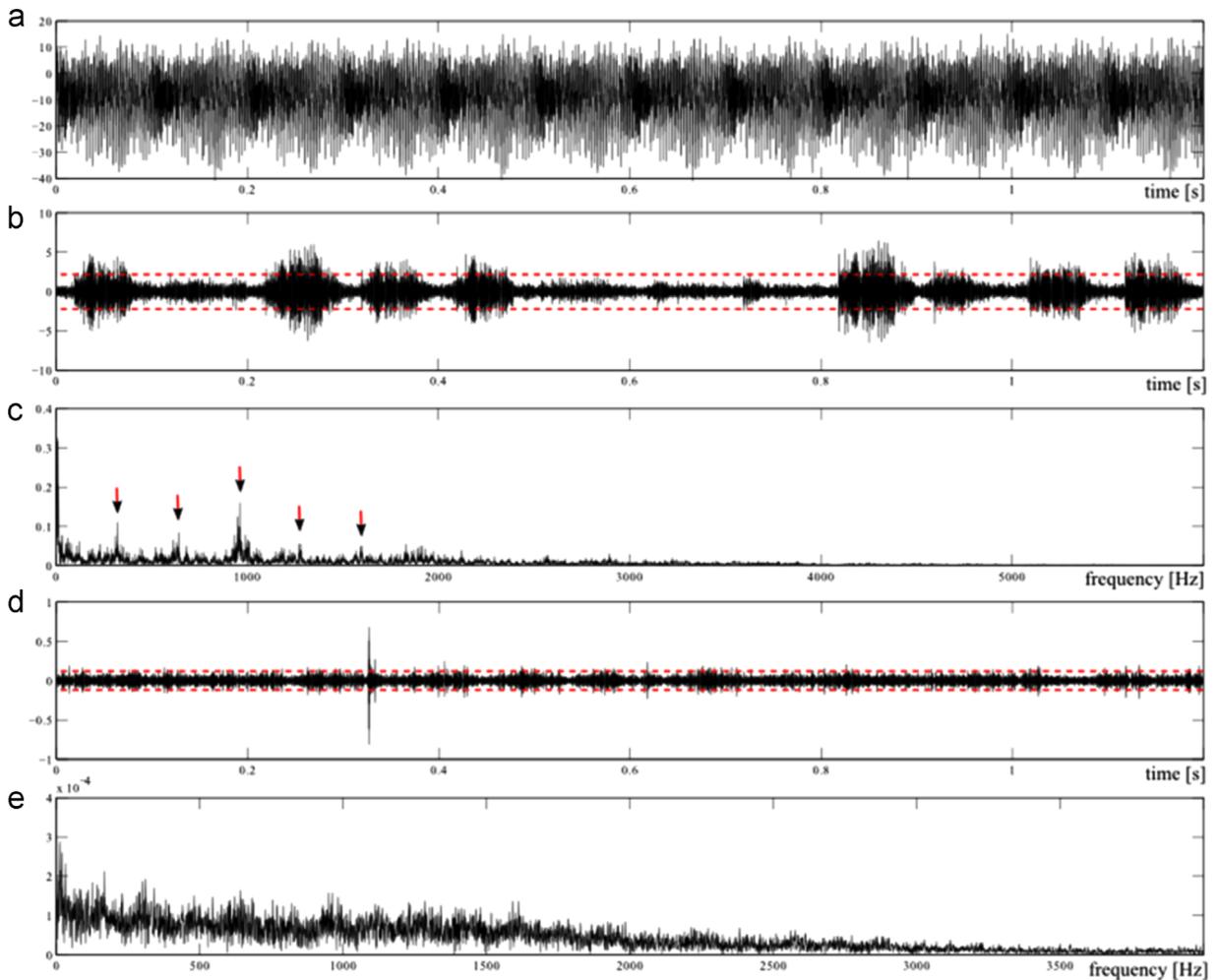

Fig. 13. (a) Healthy vibration signal and its filtered versions in bands (b) [4500;7500] Hz and (d) [19,750;20,250] Hz together with their respective squared envelope spectra (c) and (e).
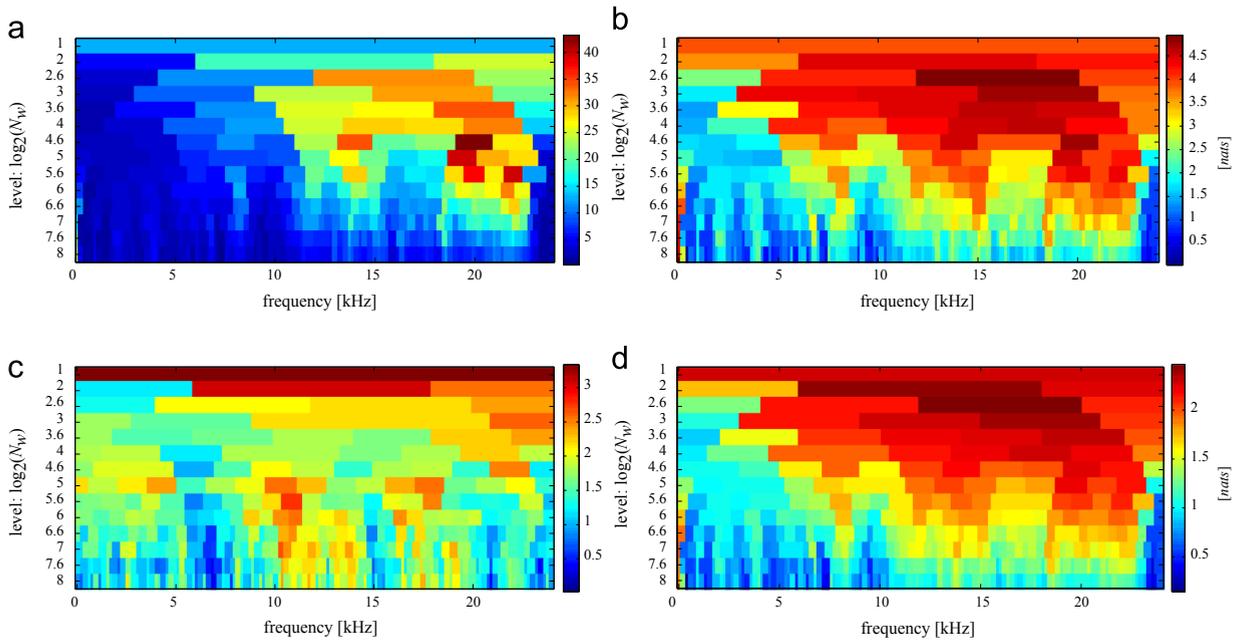
Fig. 14. Analysis of a vibration signal with an inner race fault. (a) Kurtogram, (b) SE infogram $\Delta I_\varepsilon(f; \Delta f)$, (c) SES infogram $\Delta I_E(f; \Delta f)$, (d) average infogram $\Delta I_{1/2}(f; \Delta f)$.

[19,750;20,250] Hz which maximizes the SE infogram, as shown in Fig. 13(d–e). On the other hand, the SES infogram shows a marked distribution around 5 kHz which is also noticeable in the kurtogram and SE infogram; this is therefore symptomatic of a repetitive and slightly impulsive event. The corresponding filtered signal is shown in Fig. 13(b–c) which clearly evidences second-order cyclostationarity: the envelope spectrum shows harmonics of the gearmesh frequency at 320 Hz (especially the third one). Although no fault was known in the gearbox, such a strong modulation is indicative of abnormal stress on the gears.

None of the above events are actually symptomatic of a rolling element bearing fault, yet they are informative in their own right. At the same time, this example illustrates the robustness of the SES infogram against impulsive noise and its capability to detect cyclostationarity.

A similar analysis is now conducted on the signal captured with an inner race fault. Fig. 14 evidences high values of the kurtogram and infograms in several regions. On the one hand, the kurtogram and SE infogram clearly indicate presence of impulsivity; on the other hand, the SES infogram confirms the corresponding events are cyclostationary in several frequency bands. This suggests there are several regions which are likely to contain repetitive transients. The filtered signals in bands [19,000;21,000] Hz, [23,625;24,000], and [9500;11,000] are displayed in Fig. 15 together with their SES. They all reveal the signature of an inner race fault with estimated BPFI=71.12 Hz, although the information is extracted from somewhat different frequency bands; actually, even the full-band signal clearly shows up the fault as indicated by the high value of the SES infogram at level 1. This shows that the kurtogram and infogram may be complementary rather than redundant. Besides, the SES infogram singles out some cyclostationary events below 5 kHz which do not find equivalence in the kurtogram and SE infogram; this is suspected to corresponds to gear vibrations and is readily confirmed by inspection of the bandpass signals (not displayed here).

The analysis of the vibration signal with an outer race fault is displayed in Fig. 16. On the one hand, the kurtogram and the SE infogram have very similar distributions – although their local maxima are different – with high values above 8 kHz. On the other hand, the SES infogram reveals a completely different distribution mostly concentrated below 8 kHz; the most prominent intersection with the SE infogram is in the frequency band [15,500;17,500] Hz, which might be of special interest. Fig. 17 returns the filtered in this band and evidences the outer race fault at BPFO=49.10 Hz. Another region of interest is around 5 kHz, where the bandpass signal (not displayed) evidences again strong modulation by the gears. Interestingly, the average infogram returns a complete picture of all nonstationary events in the signal which are seen to occupy the full frequency band in this case.

In this example, the kurtogram and SE infogram have efficiently detected the bearing faults, whereas the SES infogram have evidenced strong gear modulations. This situation is in accordance with the analysis of Section 5.1.2 where the rate of repetition of the transient is much lower than their spectral content, i.e. when condition $p \ll \Delta f$ holds. This example illustrates the capacity of the spectral and cyclic negentropies to identify nonstationary events of different nature, which might be of considerable importance for differential diagnosis.
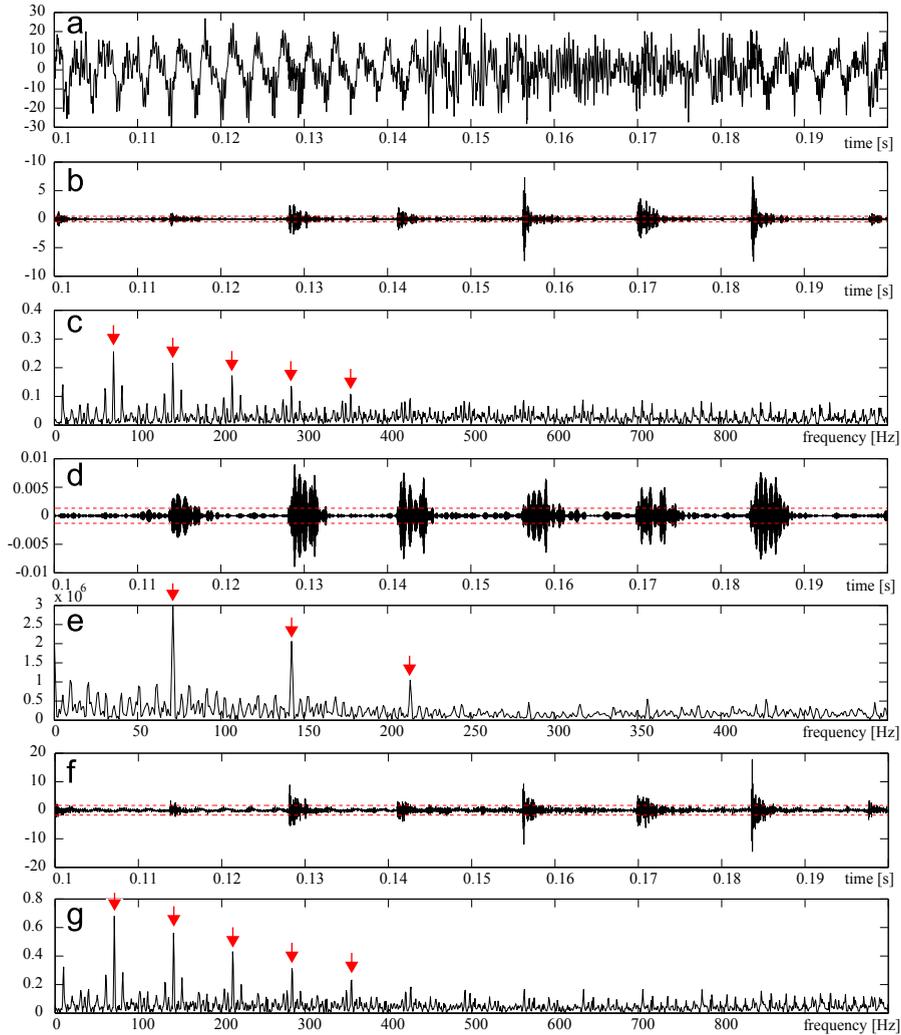
Fig. 15. (a)Faulty vibration signal (inner race fault) and its filtered versions in bands (b) [19,000;21,000] Hz, (d) [23,625;24,000] Hz, and (f) [9500;11,000] Hz together with their respective squared envelope spectra in (c), (e) and (g). Inner race fault is diagnosed at 71.12 Hz.

5.3. Vibration-based diagnosis of rolling element bearing: Experiment 2

The last experiment now demonstrates an instance where the SES infogram can detect a fault where the kurtogram and SE infogram fail. An inner race fault was introduced in the bearing of a high speed machinery by producing 8 small indentations distributed along the inner race of width 0.2 mm each. The estimated BPFI is around 1400 Hz. The rotational speed of the shaft is 200 Hz and the sampling frequency is 50 kHz. The signal was pre-whitened before processing as advocated in Section 4.5. As demonstrated in Section 5.1.2, the very high repetition rate of the fault is troublesome for the kurtogram or the SE infogram in this instance. This observation is confirmed in Fig. 18(a–b). On the contrary, the SES infogram evidences a marked maximum in the frequency band [18,000;23,000] Hz where, after bandpass filtering the signal, the inner race fault appears as a highly ringing signal with modulation at the expected BPFI—see Fig. 19.

6. Conclusion

The aim of this paper was to improve the state-of-the-art in detecting repetitive transients, a typical signature of faults in rotating machines. A distinctive feature of repetitive transients is to be impulsive in both the time and the frequency domains: the envelope spectrum shows a series of harmonics which is not accounted for in traditional approaches based on the kurtosis and its extensions such as the kurtogram. Motivated by ideas borrowed from the field of thermodynamics where transients are seen as departures from a state of equilibrium, it has been proposed to detect repetitive transients by the associated increase of negentropy in the SE of the signal and in its envelope spectrum. The spectral negentropy of the SE and of the SES return two new quantities: the "SE infogram" and "SES infogram", respectively. Whereas the first one has an
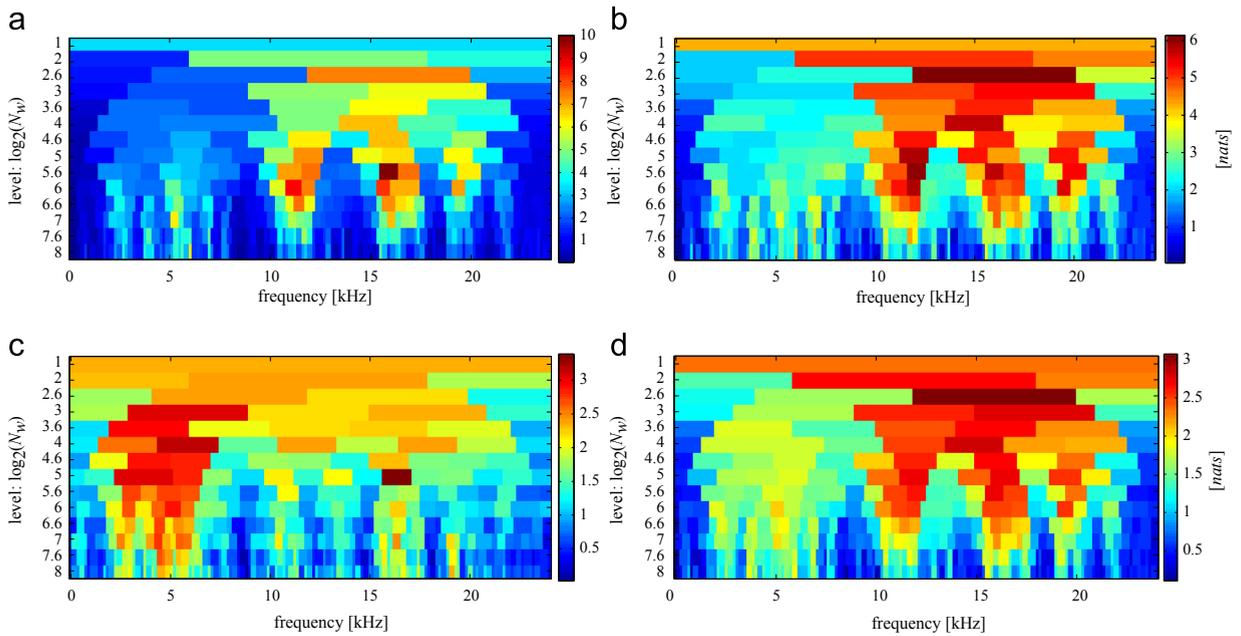
Fig. 16. Analysis of a vibration signal with an outer race fault: (a) Kurtogram, (b) SE infogram $\Delta I_e(f; \Delta f)$, (c) SES infogram $\Delta I_E(f; \Delta f)$, (d) average infogram $\Delta I_{1/2}(f; \Delta f)$.
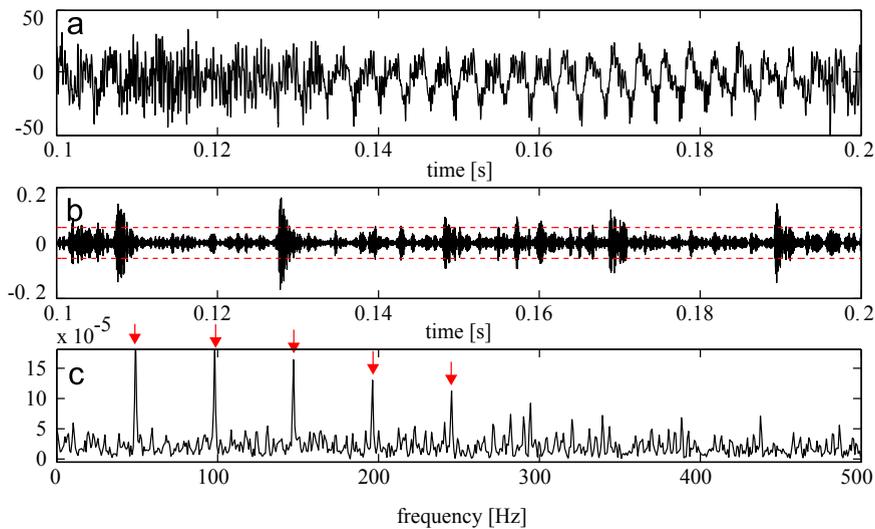

Fig. 17. (a) Faulty vibration signal (outer race fault) and (b) its filtered version in band [15,500;17,500] Hz together with (c) its squared envelope spectrum. Outer race fault is diagnosed at 49.10 Hz.

interpretation in all points similar to that of the kurtogram, the second one brings new information on the repetitiveness – or cyclostationarity – of the transients that have been missing so far. One advantage of the SE infogram and SES infogram is to share the same unit – "nat", a unit of information – and thus they are additive. Among all possible linear combinations, the classical average – coined the average infogram – enjoys a special status: as a consequence of Hirschman's entropic uncertainty principle, it is theoretically maximized by a Dirac comb—an idealization of the signature of repetitive transients. In practice, the joint analysis of the SE and SES infograms makes possible the discrimination between transient events of different natures. In particular, as compared to the kurtogram, repetitive transients are detected (1) even in the presence of strong impulsive noise and (2) independently of their rate of repetition (whereas the kurtogram becomes less and less effective as the repetition rate of the transients increases). In practice, the usefulness of the average infogram was yet found less obvious and it could probably be ignored in some applications. More generally, the proposed infograms may also become difficult to interpret – at least in an automated way – when high values appear in several different frequency bands. The analysis effort thus grows proportionally with the number of frequency bands where the signal has to be demodulated and the envelope spectra computed. This drawback is shared by the kurtogram and, in such a situation, it might be
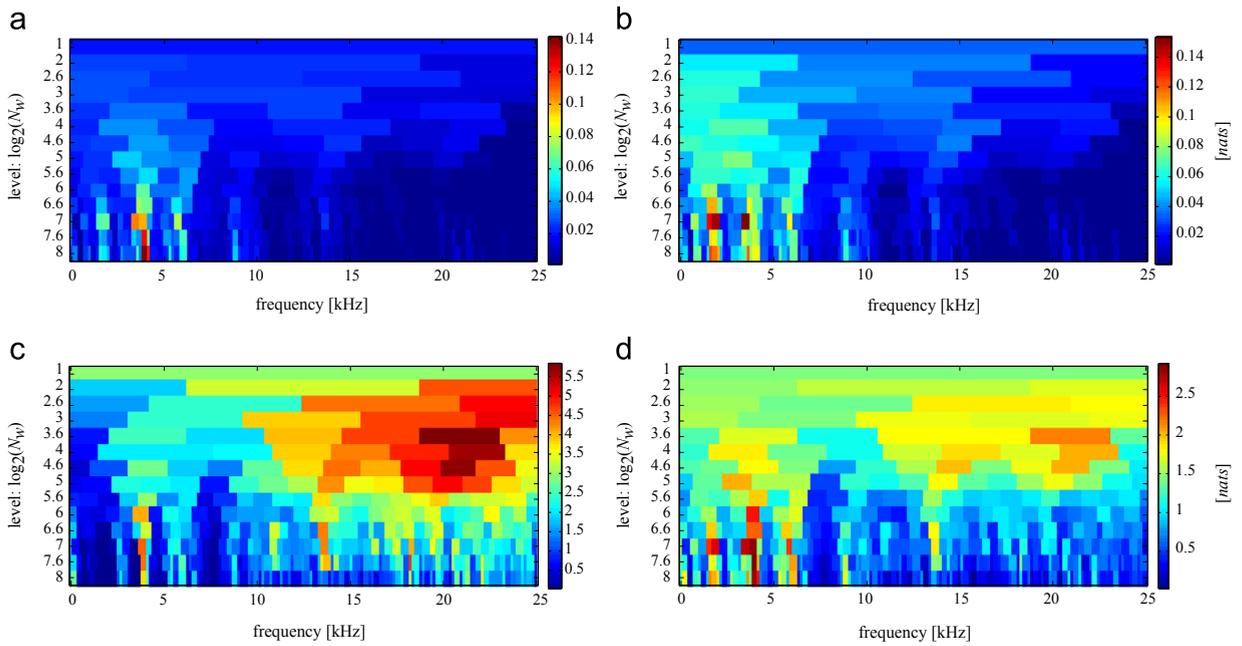
Fig. 18. Analysis of a vibration signal with a high-frequency inner race fault. (a) Kurtogram, (b) SE infogram $\Delta I_{\epsilon}(f;\Delta f)$, (c) SES infogram $\Delta I_E(f;\Delta f)$, (d) average infogram $\Delta I_{1/2}(f;\Delta f)$.
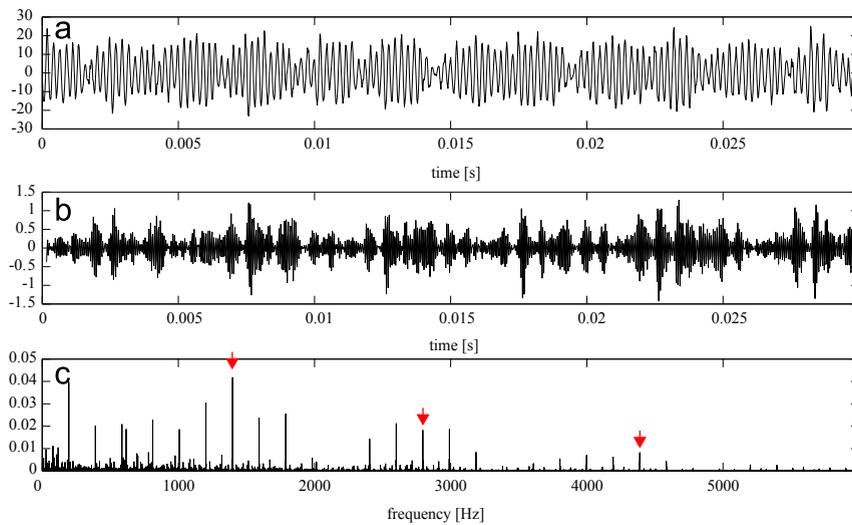

Fig. 19. (a) Faulty vibration signal and (b) its filtered version in band [18,000;23,000] Hz together with (c) its squared envelope spectrum. Inner race fault is diagnosed at 1400 Hz (note the strong modulation by shaft speed).

questionable whether an analysis by means of the spectral correlation or coherence (which is nothing else than a cascade of envelope spectra in all possible frequency bands) should not be preferred. A less obvious drawback of the infogram is to depend on the record length and on the sampling frequency (see Table 1); although this property can be used to recognize the signature of repetitive transients from other types of signals, it may also prevent the direct comparison of infograms obtained by different experimenters (i.e. on different record lengths or with different sampling frequencies).

This research work was initially motivated by a number of preliminary ideas such as the protrugram [21] and the relationship between the kurtosis and the SES [32]. One of the objectives was to attempt to combine them in a unified framework. In particular, unexpected connections have been foreseen between the concepts of nonstationarity (cyclostationarity) and nonlinearity (kurtosis). The capability of the SES infogram to detect and filter out cyclostationary components is surely worth further investigation, especially because it is much faster to compute than tools such as the spectral correlation. Another perspective is to investigate the possibility of improving the definition of the spectral negentropy of the SES, for instance by using spectral smoothing or windowing in the SES domain. Comparison with other criteria apt to

appraise harmonic structures in the SES would also be of interest [44]: the author has found fairly good results with as simple a criterion as the crest factor of the SES, which is unfortunately not additive in the sense of the Hirschman uncertainty principle.

Acknowledgements

Appendices

A.1. Proof of Eq. (6)

A frequency resolution Δf imposes a temporal width of $1/\Delta f$ to the impulses. Given a sampling frequency F_s and a total acquisition time $T = L/F_s$, the signal is thus equivalent to a point process with $L\Delta f/F_s$ samples, out of which a fraction $p/\Delta f$ is non-zero on the average. Let A be the impulses height; therefore,

$$K_x(f; \Delta f) \approx \frac{\frac{p}{\Delta f}A^4}{\left(\frac{p}{\Delta f}A^2\right)^2} - 2 \sim \frac{\Delta f}{p} \tag{18}$$

A.2. Proof of Eq. (14)

The proof is based on the following theorem due to DeBrunner et al. [38]:

Theorem: The only sequences $x(n) \in \mathbb{C}$, $n = 0, \ldots, L-1$, with $\sum_{n=0}^{L-1} |x(n)|^2 = 1$ and $X(k) = L^{-1/2}x(n)e^{-j2\pi nk/L}$ for which

$$H_{\frac{1}{2}} = -\frac{1}{2}\left(\sum_{n=0}^{L-1} |x(n)|^2 \ln\left(|x(n)|^2\right) + \sum_{k=0}^{L-1} |X(k)|^2 \ln\left(|X(k)|^2\right)\right) \tag{19}$$

is minimal are obtained from the Kronecker delta sequence by applying any composition of periodization, translation, modulation, the DFT, and multiplication by a complex number of unit magnitude.

Accommodating Eq. (19) with the definitions (10) and (12) used in this paper, it immediately results that the quantity in Eq. (13) is upper bounded by a Dirac comb when $\rho = 1/2$. The bound is found by calculating $\Delta I_\epsilon(f; \Delta f)$ and $\Delta I_E(f; \Delta f)$ in this special instance. Setting

$$x(n) = \sum_{m=0}^{K-1} A\delta(n - mN) \tag{20}$$

with $L = KN$, one has

$$\Delta I_\epsilon(f; \Delta f) = \ln N \tag{21}$$

and

$$\Delta I_E(f; \Delta f) = \ln K \tag{22}$$

Therefore $\Delta I_\epsilon(f; \Delta f) + \Delta I_\epsilon(f; \Delta f) = \ln N + \ln K = \ln L$.

A.3. Results in Table 1

After passing through a bandpass filter of bandwidth Δf, the Dirac comb is equivalent to another comb with period N/r, sampling rate F_s/r, and length L/r, with $r = F_s/\Delta f$. The results then follow from the proof in Appendix A.2. The results for the sinusoidal SE and the single impulse follows from those of the Dirac comb by setting $N = 1$ and $K = 1$, respectively. The spectral negentropy $\Delta I_\epsilon(f; \Delta f)$ for a white circular complex Gaussian envelope is obtained from considering the expected value

$$\mathbb{E}\left\{\frac{|\epsilon_x(n; \Delta f)|^2}{\mathbb{E}|\epsilon_x(n; \Delta f)|^2} \ln \frac{|\epsilon_x(n; \Delta f)|^2}{\mathbb{E}|\epsilon_x(n; \Delta f)|^2}\right\} = \mathbb{E}\left\{|z|^2 \ln |z|^2\right\} \tag{23}$$

where z is circular complex Gaussian with unit standard deviation. Setting $u = |z|^2$ whose half follows a Chi-squared distribution with 2 degrees of freedom, Eq. (23) is thus returned by the integral

$$\int_0^{+\infty} u \ln u e^{-u} du = 1 - \gamma \tag{24}$$

Since the Fourier transform of a Gaussian process is also Gaussian, the same result holds for $\Delta I_E(f; \Delta f)$. The spectral negentropy of the SE of a white circular complex Gaussian noise is similarly calculated as

$$\mathbb{E}\left\{ \frac{|\varepsilon_x(n; \Delta f)|^4}{\mathbb{E}|\varepsilon_x(n; \Delta f)|^4} \ln \frac{|\varepsilon_x(n; \Delta f)|^4}{\mathbb{E}|\varepsilon_x(n; \Delta f)|^4} \right\} = \frac{\mathbb{E}\{|z|^4 \ln |z|^4\}}{\mathbb{E}\{|z|^4\}} - \ln \mathbb{E}\{|z|^4\} \tag{25}$$

wherein $\mathbb{E}\{|z|^4\} = \int_0^{+\infty} u^2 e^{-u} du = 2$ and $\mathbb{E}\{|z|^4 \ln |z|^4\} = \int_0^{+\infty} u^2 \ln u^2 e^{-u} du = 6 - 4\gamma$. Thus $\Delta I_\varepsilon(f; \Delta f) = 3 - 2\gamma - \ln 2$.

References

[1] S.G. Braun, Mechanical Signature Analysis, Theory and Applications, Academic Press, London, 1986.
[2] S.G. Braun, The signature analysis of sonic bearing vibrations, EEE Trans. Sonics Ultrason. 27 (6) (1980) 317–327.
[3] D. Dyer, R.M. Stewart, Detection of rolling element bearing damage by statistical vibration analysis, Mech. Des. 100 (1978) 229–235.
[4] D.E. Butler, The Shock-pulse method for the detection of damaged rolling bearings, Non-Destr. Test. 6 (2) (1973) 92–95.
[5] J.L. Frarey, The history and application of the envelope detector, in: Proceedings of Technology Showcase: Integrated Monitoring Diagnostics and Failure Prevention, Mobile, Alabama, April 22–26, 1996.
[6] S.G. Braun, B. Seth, Analysis of repetitive mechanism signatures, J. Sound Vib. 70 (4) (1980) 513–522.
[7] W.A. Gardner, A. Napolitano, L. Paura, Cyclostationarity: half a century of research, Signal Process. 86 (4) (2006) 639–697.
[8] J. Antoni, The spectral kurtosis: a useful tool for characterising non-stationary signals, Mech. Syst. Sig. Process. 20 (2) (2006) 282–307.
[9] J. Antoni, R.B. Randall, The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines, Mech. Syst. Sig. Process. 20 (2) (2006) 308–331.
[10] J. Antoni, Fast computation of the kurtogram for the detection of transient faults, Mech. Syst. Sig. Process. 21 (1) (2007) 108–124.
[11] Y. Zhang, R.B. Randall, Rolling element bearing fault diagnosis based on the combination of genetic algorithms and fast kurtogram, Mech. Syst. Sig. Process. 23 (5) (2009) 1509–1517.
[12] Y. Wang, Z. He, Y. Zi, Enhancement of signal denoising and multiple fault signatures detecting in rotating machinery using dual-tree complex wavelet transform, Mech. Syst. Sig. Process. 24 (1) (2010) 119–137.
[13] Y. Wang, M. Liang, An adaptive SK technique and its application for fault detection of rolling element bearings, Mech. Syst. Sig. Process. 25 (5) (2011) 1750–1764.
[14] Y. Lei, J. Lin, Z. He, Y. Zi, Application of an improved kurtogram method for fault diagnosis of rolling element bearings, Mech. Syst. Sig. Process. 25 (5) (2011) 1738–1749.
[15] D. Wang, P.W. Tse, K.L. Tsui, An enhanced Kurtogram method for fault diagnosis of rolling element bearings, Mech. Syst. Sig. Process. 35 (1–2) (2013) 176–199.
[16] B.i. Chen, Z. Zhang, Y. Zi, Z. He, C. Sun, Detecting of transient vibration signatures using an improved fast spatial–spectral ensemble kurtosis kurtogram and its applications to mechanical signature analysis of short duration data from rotating machinery, Mech. Syst. Sig. Process. 40 (1) (2013) 1–37.
[17] P.W. Tse, Dong Wang, Two automatic vibration-based fault diagnostic methods using the novel sparsity measurement—Parts 1: The design of a new sparsogram for fast bearing fault diagnosis, Mech. Syst. Sig. Process. 40 (2) (2013) 499–519.
[18] P.W. Tse, D. Wang, Two automatic vibration-based fault diagnostic methods using the novel sparsity measurement—Parts 2: The automatic selection of an optimal wavelet filter and its enhancement by the new sparsogram for bearing fault detection, Mech. Syst. Sig. Process. 40 (2) (2013) 520–544.
[19] H. Sun, Z. He, Y. Zi, J. Yuan, X. Wang, J. Chen, S. He, Multiwavelet transform and its applications in mechanical fault diagnosis—a review, Mech. Syst. Sig. Process. 43 (1–2) (2014) 1–24.
[20] C. Pachaud, R. Salvetat, C. Fray, Crest factor and kurtosis contributions to identify defects inducing periodical impulsive forces, Mech. Syst. Sig. Process. 11 (6) (1997) 903–916.
[21] T. Barszcz, A. Jabłoński, A novel method for the optimal band selection for vibration signal demodulation and comparison with the kurtogram, Mech. Syst. Sig. Process. 25 (1) (2011) 431–451.
[22] C. Capdessus, M. Sidahmed, J.L. Lacoume, Cyclostationary processes: application in gear faults early diagnosis, Mech. Syst. Sig. Process. 14 (3) (2000) 371–385.
[23] J. Antoni, F. Bonnardot, A. Raad, M. El Badaoui, Cyclostationary modelling of rotating machine vibration signals, Mech. Syst. Sig. Process. 18 (6) (2004) 1285–1314.
[24] J. Antoni, Cyclic spectral analysis of rolling-element bearing signals: facts and fictions, J. Sound Vib. 304 (3-5) (2007) 497–529.
[25] J. Antoni, Cyclic spectral analysis in practice, Mech. Syst. Sig. Process. 21 (2) (2007) 597–630.
[26] J. Antoni, Cyclostationarity by examples, Mech. Syst. Sig. Process. 23 (4) (2009) 987–1036.
[27] J. Urbanek, J. Antoni, T. Barszcz, Detection of signal component modulations using modulation intensity distribution, Mech. Syst. Sig. Process. 28 (2012) 399–413.
[28] V. Girondin, K.M. Pekpe, H. Morel, J.-P. Cassar, Bearings fault detection in helicopters using frequency readjustment and cyclostationary analysis, Mech. Syst. Sig. Process. 38 (2) (2013) 499–514.
[29] R.B. Randall, J. Antoni, S. Chobsaard, The relationship between spectral correlation and envelope analysis for cyclostationary machine signals, application to ball bearing diagnostics, Mech. Syst. Syst. Process. 15 (5) (2001) 945–962.
[30] P. Borghesani, P. Pennacchi, R. Ricci, S. Chatterton, Testing second order cyclostationarity in the squared envelope spectrum of non-white vibration signals, Mech. Syst. Sig. Process. 40 (1) (2013) 38–55.
[31] J. Urbanek, T. Barszcz, J. Antoni, Integrated modulation intensity distribution as a practical tool for condition monitoring, Appl. Acoust. 77 (2014) 184–194.
[32] P. Borghesani, P. Pennacchi, S. Chatterton, The relationship between kurtosis- and envelope-based indexes for the diagnostic of rolling element bearings, Mech. Syst. Sig. Process. 43 (1–2) (2014) 25–43.
[33] L. Brillouin, Science and Information Theory, second ed. Academic Press, New York, NY, 1962.
[34] Bernard H. Lavenda, Statistical Physics. A Probabilistic Approach, Wiley-Interscience, Chichester, 1991.
[35] I.I. Hirschman, A note on entropy, Am. J. Math. (1957) 152–156.
[36] S. Wehner, A. Winter, Entropic uncertainty relations—a survey, New J. Phys.—Spec. Issue Quantum Inf. Many-Body Theor. 12 (025009) (2010).
[37] D.C. Brody, B.K. Meister, Discrete uncertainty relations, J. Phys. A: Math. Gen. 32 (26) (1999) 4921–4930.
[38] V. DeBrunner, M. Ozaydin, T. Przebinda, J. Havlicek, The optimal solutions to the continuous and discrete-time versions of the Hirschman uncertainty principle, in: IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 1(81), 2000.
[39] V. DeBrunner, J.P. Havlicek, T. Przebinda, M. Ozaydin, Entropy-based uncertainty measures for $L^2(\mathbb{R}^n)$, $\ell^2(\mathbb{Z})$, and $\ell^2(\mathbb{Z}/N\mathbb{Z})$ with a Hirschman optimal transform for $\ell^2(\mathbb{Z}/N\mathbb{Z})$, IEEE Trans. Signal Process. 53 (8) (2005) 2690–2699.
[40] J. Antoni, R.B. Randall, A stochastic model for simulation and diagnostics of rolling element bearings with localized faults, ASME J. Vibr. Acoust. 125 (3) (2003) 282–289.
[41] R.B. Randall, J. Antoni, Rolling element bearing diagnostics—a tutorial, Mech. Syst. Sig. Process. 25 (2) (2011) 485–520.
[42] S.G. Braun, The synchronous (time domain) average revisited, Mech. Syst. Sig. Process. 25 (4) (2011) 1087–1102.

[43] P. Borghesani, P. Pennacchi, R.B. Randall, N. Sawalhi, R. Ricci, Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions, Mech. Syst. Sig. Process. 36 (2) (2013) 370–384.
[44] J. Obuchowski, A. Wyłomańska, R. Zimroz, Selection of informative frequency band in local damage detection in rotating machinery, Mech. Syst. Sig. Process. 48 (1–2) (2014) 138–152.