Journal of Sound and Vibration 330 (2011) 1225–1243

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

Time–frequency analysis of time-varying modulated signals based on improved energy separation by iterative generalized demodulation

Zhipeng Feng [a,*], Fulei Chu [b], Ming J. Zuo [c]

[a] Institute of Vehicular Engineering, University of Science and Technology Beijing, Beijing 100083, China
[b] Department of Precision Instruments, Tsinghua University, Beijing 100084, China
[c] Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, Canada T6G 2G8

ARTICLE INFO

Article history:
Received 15 April 2010
Received in revised form
18 September 2010
Accepted 26 September 2010
Handling Editor: K. Shin
Available online 20 October 2010

ABSTRACT

Energy separation algorithm is good at tracking instantaneous changes in frequency and amplitude of modulated signals, but it is subject to the constraints of mono-component and narrow band. In most cases, time-varying modulated vibration signals of machinery consist of multiple components, and have so complicated instantaneous frequency trajectories on time–frequency plane that they overlap in frequency domain. For such signals, conventional filters fail to obtain mono-components of narrow band, and their rectangular decomposition of time–frequency plane may split instantaneous frequency trajectories thus resulting in information loss. Regarding the advantage of generalized demodulation method in decomposing multi-component signals into mono-components, an iterative generalized demodulation method is used as a preprocessing tool to separate signals into mono-components, so as to satisfy the requirements by energy separation algorithm. By this improvement, energy separation algorithm can be generalized to a broad range of signals, as long as the instantaneous frequency trajectories of signal components do not intersect on time–frequency plane. Due to the good adaptability of energy separation algorithm to instantaneous changes in signals and the mono-component decomposition nature of generalized demodulation, the derived time–frequency energy distribution has fine resolution and is free from cross term interferences. The good performance of the proposed time–frequency analysis is illustrated by analyses of a simulated signal and the on-site recorded nonstationary vibration signal of a hydroturbine rotor during a shut-down transient process, showing that it has potential to analyze time-varying modulated signals of multi-components.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Time-varying modulation is a common phenomenon in the vibration responses of machinery, such as the vibration signals of gears and rolling element bearings, and the vibration signals of rotors during speed and/or load varying processes, which are typically amplitude modulation (AM) and/or frequency modulation (FM) signals. This is mainly due to changes in the internal elements of machinery and/or in the external excitation from environment.

* Corresponding author. Tel.: +86 10 62332865; fax: +86 10 82381628.
E-mail addresses: zhipeng.feng@yahoo.com.cn (Z. Feng), chufl@mail.tsinghua.edu.cn (F. Chu), ming.zuo@ualberta.ca (M.J. Zuo).

0022-460X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2010.09.030

Slow time variation of vibration system parameters, such as the mass, stiffness and damping, can result in the AM or FM of simple harmonic vibration response. For example, for a vibration system under harmonic excitation, the time variation in mass and/or stiffness will cause not only the FM of free vibration response due to the time-varying natural frequency, but also the AM of forced vibration response due to the time-varying vibration amplitude. Similarly, the time variation of damping will cause the AM of free vibration response due to the time-decaying vibration amplitude, and both the AM and FM of forced vibration response due to the time-varying vibration amplitude and phase.

As well, changes in the running condition, such as the speed and load, cause modulation in the vibration responses of machinery. For example, the vibration signals of rotors during the speed varying process are typical AM–FM processes, because: (1) usually the rotating frequency and its multiples dominate rotor vibration signals, and they follow the rotating frequency during the speed varying process, so the vibration responses have an FM characteristic; (2) the vibration of a rotor in normal status is mainly excited by the centrifugal force which is proportional to the square of rotating frequency, so the vibration amplitude is closely related to the rotating frequency, and the speed variation results in a time-varying amplitude envelope, i.e. an AM effect on the vibration responses.

In addition, the signal propagation from vibration sources to transducers may also cause modulation effect on signals. In general, there are multiple vibration excitations in machinery. The transfer path for a vibration signal propagating from each excitation source to a vibration transducer consists of multiple components, and each component has its own dynamic characteristics and operational condition. During the propagation, these components will have an effect on the vibration signal, and cause it to be a modulated signal.

In summary, the modulation phenomenon of vibration signals is closely related to the internal components and their dynamic characteristics, as well as the external environment of machinery. In this sense, time-varying modulated vibration responses contain the health or running condition information of machinery. In essence, vibration monitoring and diagnosis relies on the detection of changes in vibration system parameters and running condition to reveal the health status of machinery. Therefore, how to effectively extract AM and FM information from time-varying modulated signals is an important issue in vibration monitoring and diagnosis of machinery.

At present, the most commonly used method to demodulate time-varying modulated signals is based on the integration of Hilbert transform and analytic signal, such as the well known envelope demodulation method [1,2]. However, Hilbert transform is a linear integral approach, and is usually implemented via fast Fourier transform. The stationarity assumption on signals of Fourier transform and the integral over a long duration cause it to lose the adaptability to instantaneous signal changes.

Recently, it is found that energy separation algorithm based on energy operator is effective in analyzing modulated signals [3–13]. Energy operator is a nonlinear differential operator. It can estimate the energy required to generate a signal by means of nonlinear combination of the instantaneous signal values and its derivatives. Its derivative, so called energy separation algorithm can separate AM–FM signals into AM and FM components, and obtain their amplitude envelope and instantaneous frequency. It does not need to construct any basis functions, and is a completely data driven algorithm adaptive to the local structure of a signal. Compared with Hilbert transform based demodulation, it is a local algorithm in time domain, and it can compute the instantaneous frequency and amplitude envelope of a time-varying AM–FM signal. It has attractive features such as high time–frequency resolution, adaptability to instantaneous nature, and low computational complexity.

Kaiser [3–5] and Maragos and coworkers [6–9] presented energy operator based demodulation, and applied it to speech analysis. Cheng et al. [10] used energy separation algorithm to separate the amplitude envelope and instantaneous frequency from intrinsic mode functions obtained with empirical mode decomposition, and thereby detected the localized damage of rolling element bearings. Bassiuny and Li [11] separated the AM and FM characteristics from the current signal of feed motor in a machine using empirical mode decomposition, computed the amplitude envelope and instantaneous frequency based on energy separation algorithm, and constructed a time–frequency distribution to monitor the condition of machine tools. Liang and Soltani Bozchalooi [12] used energy separation algorithm to extract both the amplitude and frequency modulations of vibration signals for bearing fault detection. In these researches, the good performance of energy separation algorithm in processing modulated signals of narrow band is shown.

Energy separation algorithm is subject to the constraints of mono-component and narrow band. It usually assumes AM and FM (amplitude envelope and instantaneous frequency) signals do not vary too fast or too greatly with time compared to the carrier frequency. To fulfill this requirement, a signal is usually preprocessed by bandpass filtering, but for many time-varying modulated signals, this does not address the limitation caused by the requirement of mono-component and narrow band on signals yet. Regarding large frequency deviation and wideband signals, Santhanam [13] proposed a generalized energy separation method based on multi-rate conversion and frequency shift. Its idea is useful to analyze modulated signals of wideband.

However, the issue on how to process multi-component signals with spectral overlap so as to fulfill the requirement by the energy separation algorithm, have not been addressed yet. In applications, many signals are composed of multiple components, and their frequency contents are nonlinearly time-varying and widely ranging on the time–frequency plane. In most cases, the instantaneous frequency trajectories of different components on the time–frequency plane are so curved that they have spectral overlap when they are projected onto the frequency axis. For such signals, conventional filtering will split the instantaneous frequency trajectories on the time–frequency plane due to its nature of narrow band decomposition parallel to the time axis, thus resulting in energy leakage and information loss, and cannot obtain mono-components. So how to generalize the energy separation algorithm to such signals is an interesting issue of important value.

Recently, Olhede and Walden [14] proposed a generalized demodulation approach to multi-component signals. By means of this method, some modulated signals with curvilinear instantaneous frequency trajectories on the time–frequency plane can be demodulated and decomposed into mono-components. Cheng et al. [15] improved the generalized demodulation so that the waveform of each separated mono-component can be obtained. However, in order to separate a multi-component signal into mono-components, both the original generalized demodulation and its improved version require that the demodulated components do not have spectral overlaps in the frequency domain, since it is accomplished by a single step of generalized demodulation following by a rectangular multi-band decomposition of time–frequency plane. This requirement is somewhat too strict to be fulfilled for many complicated time-varying signals. For example, rotor vibration signals during speed varying processes often have spectral overlap, because the rotating frequency and its multiples are usually dominant and they follow the time variation of the rotating frequency, thus ranging in a wideband and overlapping in the frequency domain. For such signals, the original generalized demodulation method or its improved version often cannot extract the constituting mono-components. Hence, the problem with the original and improved generalized demodulation method is not negligible. Fortunately, this limitation can be overcome by our proposed method, i.e. iterative application of the generalized demodulation. The components of interest will be separated one by one with suitable demodulation vector adapted to it. This improvement makes generalized demodulation a potential solution to the problems with the energy separation algorithm caused by multi-components and their spectral overlap. With such method as a preprocessing tool, the energy separation algorithm is expected to be generalized to wideband modulated signals of multi-components with spectral overlap.

Time-varying modulated signals are essentially nonstationary, because their amplitude and/or frequency change along with time. So time–frequency analysis is expected to be useful to process such signals. However, the inherent shortcomings or drawbacks of conventional time–frequency analysis methods limit their performance. As linear transforms, both short time Fourier transform and wavelet transform are subject to the constraint by Heisenberg uncertainty principle, i.e. the time localization and frequency resolution cannot be obtained at their highest simultaneously, either of them can only be enhanced at the expense of the other one, so that their time–frequency resolution is limited [16–18]. In addition, as basis expansion based methods, the basis in either Fourier or wavelet transform is fixed, therefore they lack adaptability in simultaneously matching the complicated components of signals. As a typical representative of Cohen class distributions, Wigner–Ville distribution has the best time–frequency resolution, but it suffers from the inevitable cross term interferences (including both inner and outer interferences) for multi-component signals which hinders the interpretation of signal features in the time–frequency domain and causes it not suitable to analyze multi-component signals. Various reduced interference distributions may suppress the negative effect of cross terms, but at the expense of a worse time–frequency resolution or suppression of auto terms [16–18]. Therefore, how to construct a time–frequency representation for arbitrary time-varying modulated signals in a fine resolution and without cross term interferences, so as to resolve the time–frequency feature effectively, is another important issue.

Observing the issue with the time–frequency analysis of time-varying modulated signals of multi-components, the mono-component and narrow band requirements on energy separation algorithm, and the issue with the original generalized demodulation in extracting mono-components from signals with spectral overlapping, the solutions to address them are proposed, respectively: (1) an iterative generalized demodulation method for extracting mono-components from time varying modulated signals; (2) improved energy separation algorithm via iterative generalized demodulation method, which is suitable to analyze many complicated modulated signals; and (3) a new time–frequency energy distribution based on the improved energy separation algorithm, which is suitable to analyze time varying modulated signals of multi-components. Three main contributions are made in this paper correspondingly: (1) The generalization of generalized demodulation to arbitrary time-varying modulated signals as long as their instantaneous frequency trajectories do not intersect on the time–frequency plane. By iterative generalized demodulation, the components of interest will be separated one by one, even though the other components still overlap in the frequency domain after a single generalized demodulation. (2) The generalization of energy separation algorithm to arbitrary time-varying modulated signals even though they do not satisfy the two requirements by the original energy separation algorithm, i.e. (a) mono-component and (b) narrow band. (3) A new time–frequency energy distribution for analysis of complicated nonstationary signals of multi-components, which has a fine time–frequency resolution and is free from cross term interferences. By integrating the above solutions and exploiting their respective merits, a time–frequency analysis method based on improved energy separation via iterative generalized demodulation is proposed for analyzing almost arbitrary time-varying modulated signals.

This paper is organized as follows. Firstly, the basics on the energy operator and the energy separation algorithm, together with the construction of time–frequency energy distribution, are introduced in Section 2, and the idea of iterative generalized demodulation is introduced in Section 3. Then, in Section 4, the proposed method based on generalized demodulation and the energy separation algorithm is applied to analyze the rotor vibration signal of a hydroturbine during a transient process. Finally, conclusions are drawn in Section 5.

2. Time–frequency analysis based on energy separation algorithm

Any modulated signal can be modeled as a superposition of AM–FM processes. The instantaneous frequency and amplitude envelope characterize such modulated signals, and they can be estimated using the energy separation

algorithm. Time-varying modulated signals are essentially nonstationary, so it is a better way to analyze them in the joint time–frequency domain based on the estimated instantaneous frequency and energy.

2.1. Energy operator

For any signal $x(t)$, the energy operator Ψ is defined as [3–6]

$$\Psi[x(t)] = [\dot{x}(t)]^2 - x(t)\ddot{x}(t), \tag{1}$$

where $\dot{x}(t)$ and $\ddot{x}(t)$ are the first and the second derivative of $x(t)$ with respect to time t, respectively. Actually, the output of energy operator tracks the energy required to generate the signal $x(t)$.

Consider an undriven linear undamped vibration system which is composed of a mass m suspended by a spring of stiffness k. Applying Newton's law of motion to it yields a second-order differential motion equation

$$m\ddot{x}(t) + kx(t) = 0, \tag{2}$$

where $x(t)$ is the displacement of mass m measured from its equilibrium position, and $\ddot{x}(t)$ the second derivative of $x(t)$ with respect to time t, i.e. acceleration. This equation governs the motion of the mass, and its solution is a harmonic oscillation

$$x(t) = A\cos(\omega t + \varphi). \tag{3}$$

Accordingly its first and second derivatives, i.e. velocity and acceleration, can be deduced as

$$\dot{x}(t) = -A\omega\sin(\omega t + \varphi), \tag{4}$$

$$\ddot{x}(t) = -A\omega^2\cos(\omega t + \varphi), \tag{5}$$

where A is the oscillation amplitude, $\omega = \sqrt{k/m}$ is the natural angular frequency, and ϕ is an arbitrary initial phase.

The total energy of this vibration system is the sum of the potential energy in the spring and the kinetic energy of the mass, i.e.

$$E = \frac{1}{2}k[x(t)]^2 + \frac{1}{2}m[\dot{x}(t)]^2. \tag{6}$$

Substituting for $x(t)$ and its derivative $\dot{x}(t)$, then

$$E = \frac{1}{2}mA^2\omega^2. \tag{7}$$

Thus the energy of a simple harmonic oscillation is proportional to the square of both the amplitude and the frequency of the oscillation.

Applying the energy operator Ψ to the harmonic oscillation $x(t)$, and substituting for its first and second derivatives, $\dot{x}(t)$ and $\ddot{x}(t)$, yields

$$\Psi[x(t)] = \Psi[A\cos(\omega t + \varphi)] = A^2\omega^2. \tag{8}$$

Comparing Eq. (8) with (7), it can be seen that the output of energy operator tracks the source energy (in the sense of per half unit mass) generating the harmonic oscillation.

The energy operator in Eq. (1) is defined for continuous time signals. Its counterpart for discrete time signals is defined as

$$\Psi[x(n)] = [x(n)]^2 - x(n-1)x(n+1). \tag{9}$$

This operator only needs three samples to calculate the signal source energy at any time, so it has a good adaptability to the instantaneous changes in signals and an excellent ability to resolve transient event.

2.2. Energy separation algorithm

Applying the energy operator Ψ to the derivative of $x(t)$, i.e. $\dot{x}(t)$, produces

$$\Psi[\dot{x}(t)] = \Psi[-A\omega\sin(\omega t + \varphi)] = A^2\omega^4. \tag{10}$$

The absolute amplitude and the frequency can be obtained by solving the combination of Eqs. (8) and (10) [7–9]

$$|A| = \frac{\Psi[x(t)]}{\sqrt{\Psi[\dot{x}(t)]}}, \tag{11}$$

$$\omega = \sqrt{\frac{\Psi[\dot{x}(t)]}{\Psi[x(t)]}}. \tag{12}$$

Slow time variation of oscillation elements will result in the AM or FM of the simple oscillator harmonic response. Energy operator can also be generalized to signals with arbitrary time-varying amplitude and frequency.

Any modulated signals, including both AM components and FM components, can be modeled as a superposition of AM–FM processes. Any AM or FM is a special case of time-varying modulated signals.

Assume an AM–FM signal in the form

$$x(t) = a(t)\cos[\underbrace{\omega_c t + \omega_m \int_0^t q(\tau)d\tau + \varphi}_{\phi(t)}], \tag{13}$$

where $a(t)$ is the instantaneous amplitude envelope, $\phi(t)$ is the instantaneous phase, ω_c is the carrier frequency, ω_m is the maximum frequency deviation from ω_c, and φ is an arbitrary phase offset. The instantaneous angular frequency equals the derivative of the instantaneous phase

$$\omega(t) = \dot{\phi}(t) = \omega_c + \omega_m q(t). \tag{14}$$

If the instantaneous amplitude envelope $a(t)$ and the instantaneous frequency $\omega(t)$ do not vary too fast (in terms of changing rate) or too greatly (in terms of changing range) compared to the carrier frequency ω_c, then the output of the energy operator Ψ applied to the AM–FM signal will track the squared product of the instantaneous amplitude envelope $a(t)$ and the instantaneous frequency $\omega(t)$, i.e.

$$\Psi\left\{a(t)\cos\left[\int_0^t \omega(\tau)d\tau + \varphi\right]\right\} = [a(t)\omega(t)]^2. \tag{15}$$

Using the energy separation algorithm, the absolute of instantaneous amplitude envelope $a(t)$ and the instantaneous frequency $\omega(t)$ can be estimated as

$$|a(t)| = \frac{\Psi[x(t)]}{\sqrt{\Psi[\dot{x}(t)]}}, \tag{16}$$

$$\omega(t) = \sqrt{\frac{\Psi[\dot{x}(t)]}{\Psi[x(t)]}}. \tag{17}$$

The above energy operator and the energy separation algorithm are associated with continuous-time signals. They can also be generalized to discrete-time signals. Different algorithms [7] have been derived by using difference to approximate derivatives, such as DESA-1a which uses a single sample asymmetric difference to approximate derivatives, DESA-1 which uses a single sample symmetric (by averaging the sum output of the energy operator applied to the backward and forward difference) difference, and DESA-2 which uses symmetric difference (between samples with time indices differing by 2) to approximate derivatives.

2.3. Time–frequency energy distribution

Time-varying modulated signals are essentially nonstationary, so it is useful to investigate the time variation of each frequency component in the joint time–frequency domain.

Given the instantaneous frequency $f_i(t)=\omega_i(t)/(2\pi)$ and energy $E_i(t)$ of the ith signal component at each instant t, a time–frequency energy distribution can be constructed as

$$E(t,f) = \sum_i E_i(t)\delta[f - f_i(t)]. \tag{18}$$

Note that the energy here refers to the energy required to produce the signal, rather than the conventional energy of the signal itself which is simply related to the square of signal amplitude. For a vibration signal, if it is viewed as a generalized 'displacement' or 'velocity', then its energy in the conventional sense indicates the potential or kinetic energy of the vibration system producing it. But the energy calculated by the energy operator is completely different, it implies the total mechanical energy which is the sum of potential and kinetic energy required to generate the signal. The time variation of energy indicates changes in the internal parameters of a vibration system (such as the mass and/or stiffness associated with frequency, and the damping associated with energy dissipation and thereby the amplitude envelope) and in the energy excitation or consumption by external environment. The instantaneous energy is a function of both instantaneous frequency and amplitude, which is proportional to the product of the squared instantaneous frequency and the squared amplitude envelope, so the time–frequency energy distribution reveals not only the time evolution of each frequency component, but also the time variation of the vibration amplitude in terms of amplitude envelope associated with each frequency component.

3. Generalized demodulation

The mono-component requirement on signals by the energy separation algorithm should be satisfied to enable the estimated instantaneous frequency to have physical meaning. In most cases, the signal is composed of multiple components

with complicated time–frequency structure, so it is necessary to decompose the signal into mono-components. Generalized demodulation is very useful to separate mono-component from time-varying modulated signals of multi-components.

3.1. Principle

Arbitrary modulated signals of multi-components can be demodulated by means of generalized demodulation [14]. The superiority of generalized demodulation lies in that it can transform the instantaneous frequency trajectory of an interested component, whatever form the instantaneous frequency trajectory may be, into a line parallel to the time axis, thus avoiding overlap with other components in the frequency domain. This enables it to be possible to separate any interested component from other components by conventional filters, as long as its instantaneous frequency trajectory does not intersect with others. If projected onto the time–frequency plane, the demodulated components behave in such a manner that (1) each component contributes exclusively to a different 'tile' in a specific tiling cell of the time–frequency plane as closely as possible and (2) at any time instant, the contribution to each tile definitely comes from no more than one component. In essence, the signal is decomposed into mono-components. Because of its advantage in separating mono-component from multi-component signals, this method is expected to be an effective tool to preprocess signals so as to satisfy the narrow band and mono-component requirements by the energy separation algorithm.

The generalized demodulation method is motivated from the modulation or frequency shift property of Fourier transform. If a signal $x(t)$ is frequency modulated by $\exp[-j2\pi\upsilon(t)]$, where $\upsilon(t)$ is a real-valued function and specifies the time-varying phase, then its Fourier transform can be expressed as

$$\int_{-\infty}^{\infty} x(t)\exp[-j2\pi\upsilon(t)]\exp[-j2\pi ft]dt = X[f+\dot{\upsilon}(t)]. \tag{19}$$

If $\exp\{-j2\pi[\upsilon(t)+ft]\}$ is viewed as a basis, then the generalized Fourier transform [14] is defined as

$$\int_{-\infty}^{\infty} x(t)\exp[-j2\pi\upsilon(t)]\exp[-j2\pi ft]dt = \int_{-\infty}^{\infty} x(t)\exp\{-j2\pi[\upsilon(t)+ft]\}dt = X_G(f). \tag{20}$$

Accordingly its counterpart, the inverse generalized Fourier transform [14], is derived as

$$x(t) = \int_{-\infty}^{\infty} X_G(f)\exp\{j2\pi[\upsilon(t)+ft]\}df = \exp[j2\pi\upsilon(t)]\int_{-\infty}^{\infty} X_G(f)\exp[j2\pi ft]df. \tag{21}$$

Therefore, if $X_G(f)=\delta(f-f_0)$, then $x(t)=\exp\{j2\pi[\upsilon(t)+f_0t]\}$, i.e. a signal with instantaneous frequency $\theta(t)=f_0+\dot{\upsilon}(t)$ is mapped to the frequency point $f=f_0$. If the instantaneous frequency trajectory of a signal on the time–frequency plane is specified by a function $\theta(t)=f_0+\dot{\upsilon}(t)$, and the signal is expected to be mapped into a passband on the time–frequency plane, then a demodulation phase function is needed to approximate the time-varying phase $\upsilon(t)$. This demodulation approach is very flexible, because the demodulation phase function, i.e. the approximation function fitting the time-varying phase $\upsilon(t)$, is constrained neither by linearity nor quadric, and it can vary arbitrarily with respect to time.

The original method applies the generalized demodulation only once, and it is only effective in such cases that all the instantaneous frequency trajectories after a single step of generalized demodulation do not overlap in the frequency domain. For example, if all the interested instantaneous frequency trajectories on the time–frequency plane have almost the same slope at every instant, then after a single step of generalized demodulation with an appropriate demodulation phase function, they will become parallel to the time axis and separable by conventional bandpass filtering. However, this situation is not true in many cases. If some of the instantaneous frequency trajectories after a single step of generalized demodulation still overlap in the frequency domain, then the original generalized demodulation will fail to decompose such complicated signals into mono-components. For such signals, an improved generalized demodulation is proposed in this section to tackle the problem with the original method. It repeats the generalized demodulation, each time with a different demodulation phase function which is properly designed for the interested component, thus making the interested component separable from others in the frequency domain by conventional filters. By such iterative generalized demodulation, all the interested components will be extracted one by one, and the signal will be decomposed into mono-components.

Arbitrary time-varying modulated signals of multi-components can be modeled as

$$x(t) = \sum_{i=1}^{N} x_i(t) = \sum_{i=1}^{N} a_i(t)\cos[2\pi f_c t + \phi_i(t)], \tag{22}$$

where $a_i(t)$ is the real-valued instantaneous amplitude envelope, $\phi_i(t)$ is the instantaneous phase, and f_c is the carrier frequency. The basic steps to decompose it into mono-components based on iterative generalized demodulation are summarized as follows.

(1) Create the analytic signal $y(t)=x(t)+jH[x(t)]$, to avoid interference on the time–frequency plane caused by meaningless negative frequency, where $H(\cdot)$ denotes the Hilbert transform. For the signal as expressed by Eq. (22), its Hilbert

transform can be approximated by the corresponding quadrature part

$$H[x(t)] \approx \sum_{i=1}^{N} a_i(t)\sin[2\pi f_c t + \phi_i(t)]. \tag{23}$$

So its analytic version can be expressed as

$$y(t) = x(t) + jH[x(t)] = \sum_{i=1}^{N} a_i(t)\exp\{j[2\pi f_c t + \phi_i(t)]\}, \tag{24}$$

as long as a simple condition is satisfied, i.e. Fourier spectrum of the modulations

$$M(f) = \int_{-\infty}^{\infty} \sum_{i=1}^{N} a_i(t)\exp[j\phi_i(t)]\exp(-j2\pi f t)dt \tag{25}$$

be zero for $f < -f_c$. This requirement can usually be approximately satisfied because: for many modulated signals, Fourier spectrum of the modulations will decrease monotonically for $f < -f_c$ [19].

(2) Forward demodulate the analytic signal $y(t)$ using a demodulation vector $\exp[-j2\pi\tilde{v}_i(t)]$ which is constructed from a properly designed demodulation phase function $\tilde{v}_i(t)$ for the interested component $x_i(t)$, resulting in a demodulated signal

$$z(t) = \exp[-j2\pi\tilde{v}_i(t)]y(t) = \sum_{i=1}^{N} a_i(t)\exp(j\{2\pi[f_c t - \tilde{v}_i(t)] + \phi_i(t)\}), \tag{26}$$

which is still analytic. This is to frequency demodulate the signal with the complex exponential of the demodulation phase function $\tilde{v}_i(t)$. It is equivalent to transforming the instantaneous frequency trajectory of an interested signal component on the time–frequency plane into a desired shape or position, such that the frequency component of interest concentrates exclusively in a specific rectangular cell on the time–frequency plane, so as to satisfy the mono-component requirement. The change in the instantaneous frequency trajectory depends on the derivative of the demodulation phase function used $\tilde{v}_i(t)$, i.e. the instantaneous frequency of the demodulation vector $\dot{\tilde{v}}_i(t)$, because the instantaneous frequency of the demodulated component differs from its original one $f_c + (1/2\pi)\dot{\phi}_i(t)$ by the difference $\dot{\tilde{v}}_i(t)$ and becomes $f_c + (1/2\pi)\dot{\phi}_i(t) - \dot{\tilde{v}}_i(t)$. For example, the instantaneous frequency trajectory is simply shifted along the frequency axis if $\dot{\tilde{v}}_i(t)$ is constant (i.e. the demodulation vector $\exp[-j2\pi\tilde{v}_i(t)]$ is a harmonic); it is rotated by an angle if $\dot{\tilde{v}}_i(t)$ is linear (i.e. $\exp[-j2\pi\tilde{v}_i(t)]$ is a chirp); and it is curved if $\dot{\tilde{v}}_i(t)$ is nonlinear, e.g. quadratic or logarithmic (i.e. $\exp[-j2\pi\tilde{v}_i(t)]$ is nonlinear chirp). The shapes and positions of instantaneous frequency trajectories on the time–frequency plane are changed, but the relative distances between each other along the frequency axis are unchanged. Because all the components are demodulated using an identical demodulation phase function $\tilde{v}_i(t)$, at each time instant t their instantaneous frequency trajectories differ from their respective original ones by the same one $\dot{\tilde{v}}_i(t)$, so that the spacing between each other remains the same. If the instantaneous frequency trajectories of interested components do not intersect with others on the time–frequency plane, then after a single step of generalized demodulation, they still do not intersect with others. In particular, the demodulated component which is expected to be separated will become almost parallel to the time axis and perpendicular to the frequency axis after such generalized demodulation using a suitable demodulation phase function $\tilde{v}_i(t)$, thus concentrating exclusively in a narrow rectangular cell on the time–frequency plane and readily being separable by conventional filters. The selection of demodulation phase function $\tilde{v}_i(t)$ is very flexible, it may be arbitrary even without a closed form expression as long as it can be fitted numerically and differentiable so that the instantaneous frequency can be approximated by its derivative with respect to time. However, it should be carefully selected to ensure the instantaneous frequency trajectory of the interested demodulated signal component lies exclusively in a rectangular cell on the time–frequency plane and does not overlap with the other components if they are projected onto the frequency axis, such that the signal component can be isolated from the other components by conventional filters. In applications, the instantaneous frequency of a interested signal component is usually unknown, while it can be estimated by a numerical approximation (e.g. polynomial or exponential fitting) to its time–frequency distribution (e.g. short time Fourier transform spectrogram), then the demodulation phase function can be derived by integral of the estimated instantaneous frequency with respect to time. In most cases, the instantaneous frequency trajectory of interested components is expected to be 'shifted' to a specific frequency f_0, and to be parallel to the time axis as much as possible. This can be accomplished by using a $\dot{\tilde{v}}_i(t)$ which differs from the estimated instantaneous frequency $\tilde{f}(t) \approx f_c + (1/2\pi)\dot{\phi}_i(t)$ by a constant f_0 at every instant. The constant difference f_0 will determine the frequency offset of the demodulated component from zero frequency on the time–frequency plane.

(3) Filter the frequency demodulated signal $z(t)$ so as to obtain a demodulated component

$$z_i(t) = a_i(t)\exp\big(j\{2\pi[f_c t - \tilde{v}_i(t)] + \phi_i(t)]\}\big). \tag{27}$$

Any filtering manipulation with real coefficients on analytic signals results in analytic ones, since it only removes the undesired components and does not change the frequency of interest. So the demodulated and filtered component $z_i(t)$ is still analytic approximately. After such filtering manipulation, only the frequency component of interest is kept and

separated from the signal, and the other temporarily undesired components are removed. Note that the filter center should be around the specific frequency f_0, and the filter bandwidth should be carefully selected according to the minimum frequency spacing between the instantaneous frequency trajectories of the desired component and the others.

(4) Reverse demodulate the component $z_i(t)$, using a reverse demodulation vector $\exp[j2\pi\tilde{v}_i(t)]$, resulting in

$$y_i(t) = \exp[j2\pi\tilde{v}_i(t)]z_i(t) = a_i(t)\exp\{j[2\pi f_c t + \phi_i(t)]\}, \tag{28}$$

which is still analytic. This is to reverse frequency demodulate the filtered signal with the conjugate complex exponential of the same demodulation phase function $\tilde{v}_i(t)$ as used in step (2). It is equivalent to recover the instantaneous frequency trajectory into its original shape and position on the time–frequency plane, but the recovered component has exclusively one frequency at any instant, such that a mono-component of interest in the signal is obtained. After such preprocessing, the separated component $y_i(t)$ satisfies the mono-component requirement by the energy separation algorithm. If necessary, e.g. to satisfy the high carrier frequency requirement by the energy separation algorithm, a linear term with regards to time ct may be added to the reverse demodulation phase function $\tilde{v}_i(t) = \tilde{v}_i(t) + ct$, such that the recovered component will be shifted along the frequency axis to a desired position, where the constant c controls the frequency offset from the original position at any instant.

(5) If it is necessary, e.g. the instantaneous frequency trajectory of the other components, except the one separated, still have overlap between each other in the frequency domain, then repeat steps (2)–(4) to separate the other components of interest. Otherwise, repeat steps (3) and (4) to separate them.

(6) Take the real part of each separated component $y_i(t)$

$$x_i(t) = \mathrm{Re}[y_i(t)] = a_i(t)\cos[2\pi f_c t + \phi_i(t)]. \tag{29}$$

Now all the mono-components of interest $x_i(t)$ constituting the modulated signal of multi-components $x(t)$ are separated, and their amplitude envelope and instantaneous frequency can be readily estimated using the energy separation algorithm.

This method is effective for a broad class of signals, as long as the instantaneous frequency trajectories of different components do not intersect on the time–frequency plane. In most cases, signals satisfy such a simple requirement, and the components of interest can be separated by filtering after suitable generalized demodulation.

3.2. Example

The performance of energy separation based on iterative generalized demodulation algorithm will be illustrated by analysis of a simulation AM–FM signal. The synthetic signal consists of three AM–FM components which are quadratic chirps multiplied by exponentially decaying functions, and of a white Gaussian noise which is added to simulate the background noise interference

$$x(t) = \sum_{i=1}^{3} \exp(-c_i t)\cos\left\{2\pi\left[\frac{f_{0i}-f_{1i}}{3t_{1i}^2}(t-0.511)^3 + f_{1i}(t-0.511)\right]\right\} + n(t), \tag{30}$$

where $t = 0, 0.001, \ldots, 0.511$ s, i.e. the sampling frequency is 1000 Hz, $n(t)$ is a white Gaussian process with zero mean and variance of 0.0969 (whereas the variance of the summed quadratic chirps is 0.6911, and in this case, the signal-to-noise ratio is 8.5 dB), and the value for the other parameters are listed in Table 1. It simulates the AM–FM characteristics of rotor vibration signals during speed varying process. The feasibility of the proposed time–frequency analysis based on improved energy separation by generalized demodulation in analyzing nonstationary rotor vibration signals will be tested.

Fig. 1 shows the ideal time–frequency energy distribution of the pure signal (sum of the three quadratic chirps), which is obtained by separately applying the energy separation algorithm to each AM–FM component and followed by a summation of the time–frequency energy distribution of the three AM–FM components. (Note: In all the figures showing the time–frequency distribution of a signal, the signal waveform is on the top, its power spectrum on the left, a colorbar denoting the time–frequency distribution magnitude on the right, and the time–frequency distribution in the middle.) It can be seen that the instantaneous frequency trajectory of each component is curvilinear and changes in a wide frequency range on the time–frequency plane. Worst of all, the three AM–FM components overlap in the frequency domain, if their instantaneous frequency trajectories are projected onto the frequency axis. Any horizontal lines between the signal

Table 1
AM–FM parameter.

Component i	c_i	f_{0i} (Hz)	f_{1i} (Hz)	t_{1i} (s)
1	2	150	40	0.511
2	3	200	60	0.511
3	4	250	80	0.511

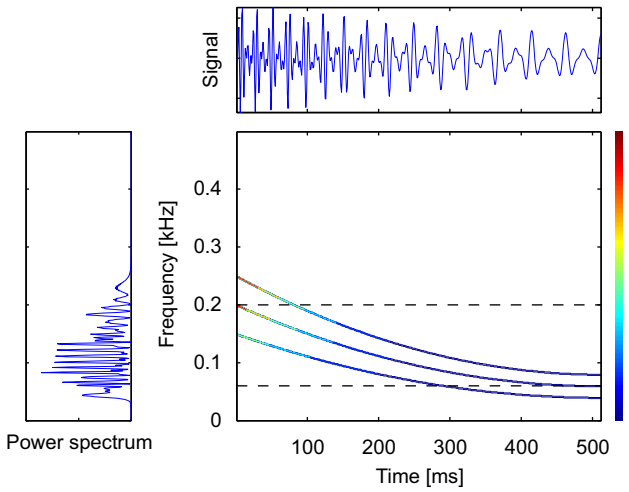

Fig. 1. Ideal time–frequency energy distribution of a pure signal.

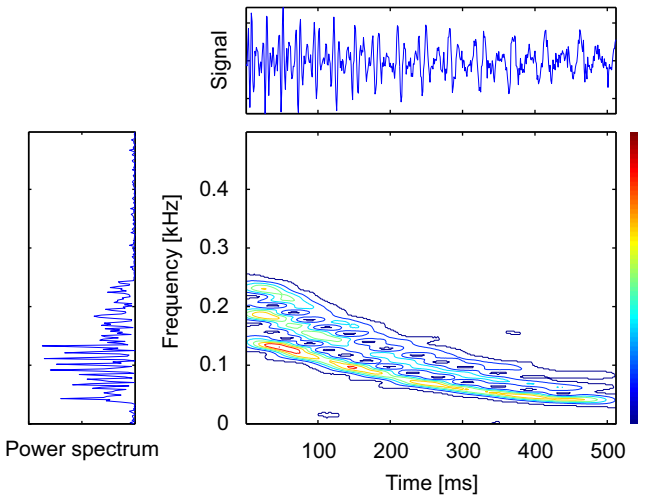

Fig. 2. Short time Fourier transform spectrogram of a noisy signal with SNR=8.5 dB.

distribution band (40, 250) Hz split the instantaneous frequency trajectory on the time–frequency plane, so any filtering manipulation on the signal with a frequency associated with the horizontal lines as the cut-off frequency will cause information loss. In addition, the mono-component condition which is required by the energy operator cannot be satisfied by filtering or narrow band decomposition applied directly to the synthetic signal, no matter how narrow the bandwidth is. A narrow band filtering or decomposition divides the time–frequency plane along the frequency axis, and results in a rectangular division parallel to the time axis and perpendicular to the frequency axis on the time–frequency plane. Any rectangular division that covers the full instantaneous frequency trajectory of one component also covers part of the adjacent instantaneous frequency trajectory. So the obtained filtered signal is not a mono-component. For example, the subband [60,200] Hz covers the full instantaneous frequency trajectory of the middle component, but it also contains part of both the top and bottom component instantaneous frequency trajectories. If a filter with too narrow a bandwidth is used to fulfill the mono-component requirement, it cannot cover the full instantaneous frequency trajectory of a component. For example, the subband [0,60] Hz and [200,250] Hz contain only the bottom and the top component, respectively, but unfortunately they cover only the end and the beginning of their respective components, so the time-varying feature over the whole duration cannot be revealed. This issue makes it an intractable problem to separate such signals into mono-components with traditional filtering or decomposition method.

Fig. 2 shows the short time Fourier transform spectrogram of the simulated noisy signal, with a Hamming window of length 101 which is well adapted to resolve the signal time–frequency structure. By comparing the signal waveforms in Figs. 1 and 2, the noise effect can be clearly seen. Even though short time Fourier transform spectrogram has a lower

time–frequency resolution in comparison to Fig. 1, it reveals the time–frequency features of the signal: the signal has three components, and their instantaneous frequencies decay exponentially with time evolution.

According to the findings based on the short time Fourier transform spectrogram, an exponential model

$$\dot{v}(t) = a\exp(bt), \tag{31}$$

where a and b are coefficients to be estimated, is used to fit the instantaneous frequency trajectory of each component on the time–frequency plane. Selection of the fitting models depends on the shape of instantaneous frequency trajectories. Other curve fitting models like polynomials can also be used, as long as they are integrable so that the demodulation phase function can be obtained by integrating the fitting function with respect to time. Here, the coefficients a and b are regressed based on visually selected peaks at three time instants on the time–frequency plane rather than using the known simulation parameters in Table 1. Table 2 lists the time–frequency location of the peaks at the three instants for each component and the regressed coefficients of corresponding exponential model. Fig. 3 shows the theoretical instantaneous frequencies, their respective estimations based on the peaks of the spectrogram in Fig. 2 and the corresponding exponentially fittings, indicating that the exponential model tracks approximately the theoretical instantaneous frequency. Now the demodulation phase function for each component can be obtained by integrating the exponential function in Eq. (31) fitting the instantaneous frequency trajectory with respect to time

$$\tilde{v}(t) = \frac{a}{b}\exp(bt). \tag{32}$$

Fig. 4 shows the time–frequency distribution of the signal demodulated with the demodulation phase function for demodulating the bottom component. The demodulation phase function used is $\exp\{j2\pi[52.2352\exp(-2.9699t)+150t]\}$. Note that a linear term $150t$ is added to the exponential. It can be seen that the instantaneous frequency trajectory of the bottom component is de-curved so that it is almost parallel to the time axis and perpendicular to the frequency axis, and it is shifted along the frequency axis to the 150 Hz position. Although the instantaneous frequency trajectories of the other two components are also changed, the spacing between each other remains the same as original in Fig. 1. Most of all, the bottom one neither overlaps nor intersects with them on the time–frequency plane, so it can be isolated from the other components by a horizontal line, e.g. the bottom black dashed line. Now it can be separated from the signal with a simple lowpass or bandpass filter, and as a result, it is definitely a mono-component containing only one frequency content at any instant.

Yet, the top and middle components cannot be separated temporarily, because they still overlap in the frequency domain after the foregoing single step of generalized demodulation. For example, the subband between the two dashed lines covers the full instantaneous frequency trajectory of the middle component, but it also contains more than half the

Table 2
Location of peaks and regressed coefficients of exponential fitting model.

		Time instant			Exponential coefficients	
		0.064	0.256	0.448	a	b
Frequency	Bottom	128.9063	70.3125	42.9688	155.1333	−2.9699
	Middle	169.9219	97.6563	64.4531	199.8198	−2.6420
	Top	214.8438	126.9531	85.9375	250.1848	−2.4954

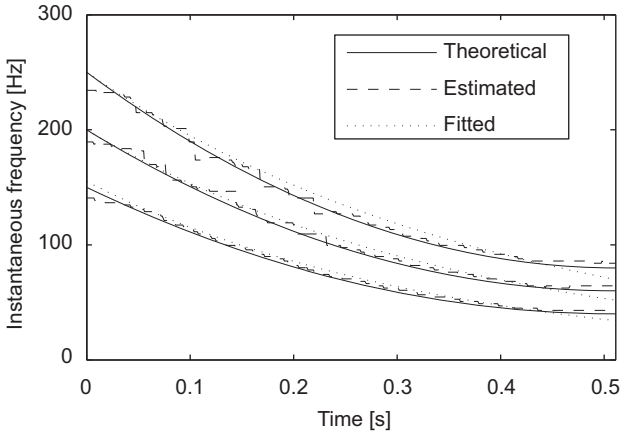

Fig. 3. Instantaneous frequencies and exponential fittings of a noisy signal with SNR=8.5 dB.

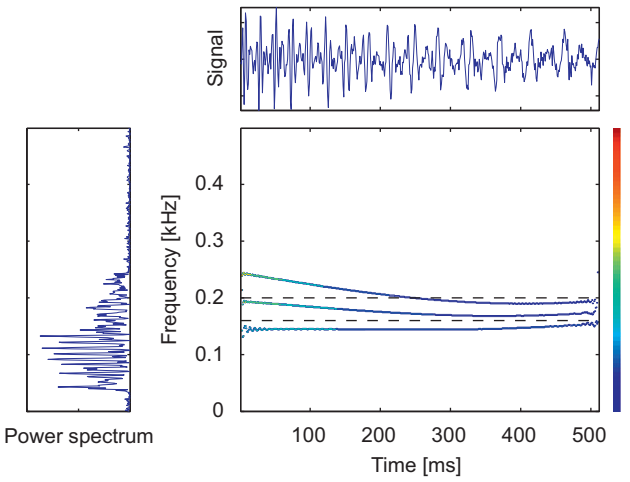

Fig. 4. Time–frequency energy distribution with bottom component de-curved.

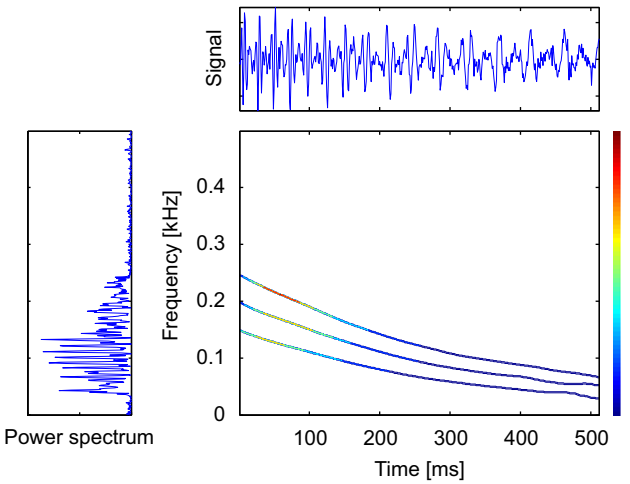

Fig. 5. Time–frequency energy distribution of a noisy signal with SNR=8.5 dB.

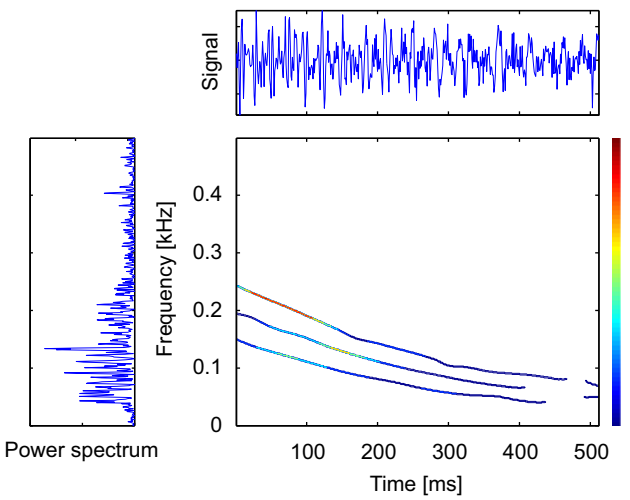

Fig. 6. Time–frequency energy distribution of a noisy signal with SNR=2 dB.

end part of the top component instantaneous frequency trajectory. Such a filter with the interval between the two dashed lines as passband cannot separate the middle component into mono-component. Therefore, for such multi-component signals with complicated time–frequency structures, it is necessary to iteratively apply the generalized demodulation with different properly designed demodulation phase functions to separate them into mono-components.

After separating each generalized demodulated component, apply reverse demodulation to them and take their real parts separately, then calculate the instantaneous frequency and energy with the energy separation algorithm. The reverse demodulation phase function used for the bottom component is $\exp\{-j2\pi[52.2352\exp(-2.9699t)]\}$. By comparing it with the forward demodulation phase function $\exp\{j2\pi[52.2352\exp(-2.9699t)+150t]\}$, it can be seen that they differ in the linear term with regards to time by a term $150t$ in the exponential. This is equivalent to shifting the separated mono-component of interest along the frequency axis by an offset of 150 Hz, so as to increase its carrier frequency and thereby to fulfill the high carrier frequency requirement by the energy separation algorithm. Remember to reverse frequency shift the estimated instantaneous frequency after energy separation.

Fig. 5 shows the time–frequency energy distribution constructed according to the instantaneous frequency and energy. It reveals the overall time–frequency feature of the signal with respect to the time variation of the frequency content and its energy. By comparing Fig. 5 with Fig. 1, it can be seen that it accords well with the ideal one, and reflects the theoretical instantaneous frequency trajectory. Though the estimated instantaneous frequency trajectories differ somewhat from the theoretical ones, for example, in the interval 0.4–0.5 s for the middle and bottom ones, this can be overcome by improving the fitting model or using more points of data (in this example only three points are used) for curve fitting.

In order to further illustrate the robustness of the proposed method to noise, the noise level is increased with a variance of 0.4359 (in this case, the signal-to-noise ratio is reduced to 2.0 dB). Fig. 6 shows the time–frequency energy distribution of the noisy signal under this scenario. The signal components and their time evolution are extracted successfully too. This again validates the performance of the method in presence of noise.

By comparing the new time–frequency energy distribution as shown in Figs. 5 or 6 with the conventional short time Fourier transform as shown in Fig. 2, some attractive features of the time–frequency energy distribution can be noted. First of all, it has a fine time–frequency resolution. This advantage is mainly attributed to the good adaptability of both the energy operator and the energy separation algorithm to instantaneous changes in signals. The energy operator based energy separation algorithm uses only a few adjacent signal samples to calculate the instantaneous energy of signals at any instant, and thereby separates the instantaneous frequency and amplitude envelope, so it emphasizes the local time–frequency structure of signals.

In addition, the time–frequency energy distribution is free of both inner and outer cross term interferences, because the instantaneous frequency is calculated from mono-components, and it does not involve integral of bilinear/quadratic terms.

4. Analysis of a nonstationary hydroturbine vibration signal

Hydroturbines in hydroelectric power stations often switch between different running states, especially during transient processes, such as start-up, shut-down, and load rejection processes. The nonstationary response during these processes contains special information about the condition of hydroturbine generator sets which cannot be obtained during stable processes. Therefore, it is important to analyze the nonstationary signals for health monitoring and fault diagnosis purposes. The radial vibration displacement of hydroturbine rotors is a key indicator. In this section, the proposed time–frequency analysis method based on energy separation and generalized demodulation is applied to analyzing the nonstationary rotor vibration signal of a hydroturbine during a shut-down process.

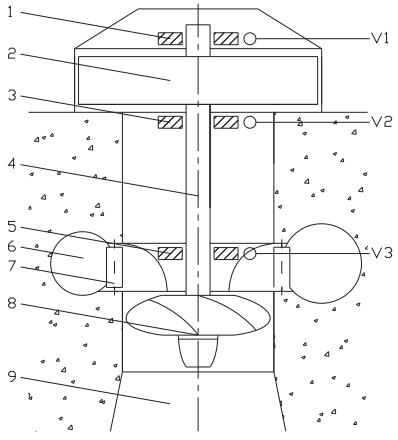

Fig. 7. Hydroturbine sketch. Hydroturbine schematic: 1—upper guide bearing, 2—generator rotor, 3—thrust bearing, 4—main shaft, 5—water turbine guide bearing, 6—spiral case, 7—guide vane, 8—turbine runner, 9—draft tube. Measurement point configuration: V1—at upper guide bearing, V2—at thrust bearing, V3—at water turbine guide bearing.

4.1. Hydroturbine parameters and measurement configuration

Fig. 7 shows the sketch of a vertical axial-flow hydroturbine in a hydroelectric power station and the on-site vibration measurement point configuration.

The main parameters of the hydroturbine are as follows.

Model: ZD(F23)-LH-700
Rated power: 75 MW
Rated speed: 88 rev/min (1.467 Hz).

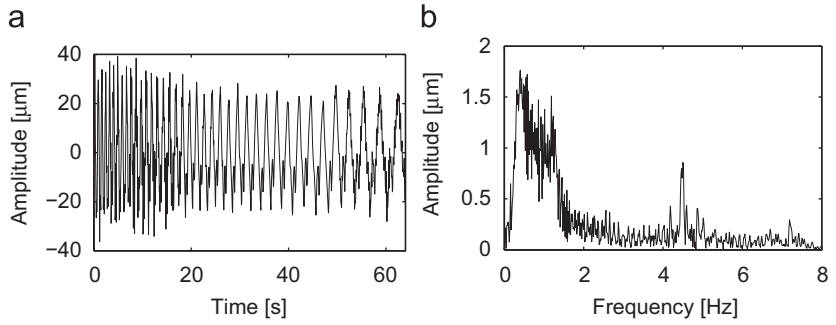

Fig. 8. Rotor vibration displacement signal: (a) waveform and (b) Fourier spectrum.

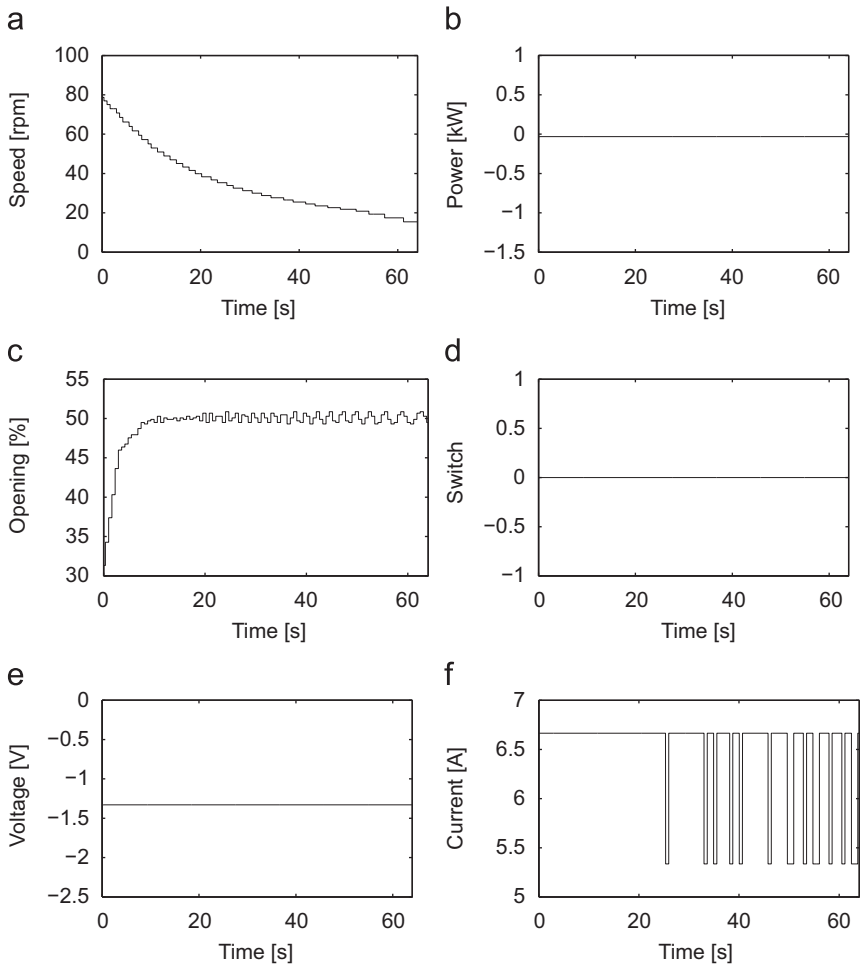

Fig. 9. Running parameters. (a) speed, (b) power, (c) guide vane opening, (d) switch, (e) excitation voltage and (f) excitation current.

The rotor vibration signals are measured at the positions of upper guide bearing, thrust bearing, and water turbine bearing, and are sampled at a frequency of 16 Hz.

4.2. Time–frequency analysis

A shut-down transient process of the hydroturbine lasts for about 64 s without power output. Fig. 8 shows the radial rotor vibration displacement signal measured at the position of upper guide bearing, and Fig. 9(a)–(f) shows the corresponding running parameters, such as the speed, power, guide vane opening, switch, excitation voltage and current, respectively.

During the transient processes of hydroturbines, the operation parameters vary in a wide range continuously, thus resulting in the inevitable nonstationarity of rotor vibration.

In most cases, the rotating frequency and its multiples are dominant in rotor vibration signals. The rotating frequency, its multiples or fractional multiples, together with their respective amplitude, are key indicators for monitoring and diagnosis of rotating machines. During the speed varying process, such as start-up and shut-down processes, these dominant frequency components follow the time variation of rotating frequency, and they never intersect each other on the time–frequency plane except at the instant when the speed reaches zero. So they satisfy the condition to be separated using the iterative generalized demodulation method.

Fig. 10 shows the short time Fourier transform spectrogram of the signal with Hamming window of length 257 well adapted to resolve the signal time–frequency structure. It can be noted that the signal consists mainly of the rotating frequency and its multiples up to four times, and a transient component of almost constant frequency around 4.5 Hz. The rotating frequency and its multiples have overlap in the frequency domain. Based on this preliminary knowledge about the signal components, an idea to separate these components is outlined as follows: construct a demodulation vector for each of the four multiples of rotating frequency and the transient component of constant frequency, respectively, then separate them one by one by applying the generalized demodulation iteratively.

To separate these components, a suitable demodulation phase function is carefully designed for each component. For the rotating frequency and its multiples, the demodulation phase function can be approximated according to the recorded rotating frequency which is usually monitored in real applications, or according to the estimated rotating frequency based on usual time–frequency distribution like short time Fourier transform if the rotating speed is unavailable. Here, to illustrate the flexibility of the proposed method, the demodulation phase functions are constructed based on the estimated

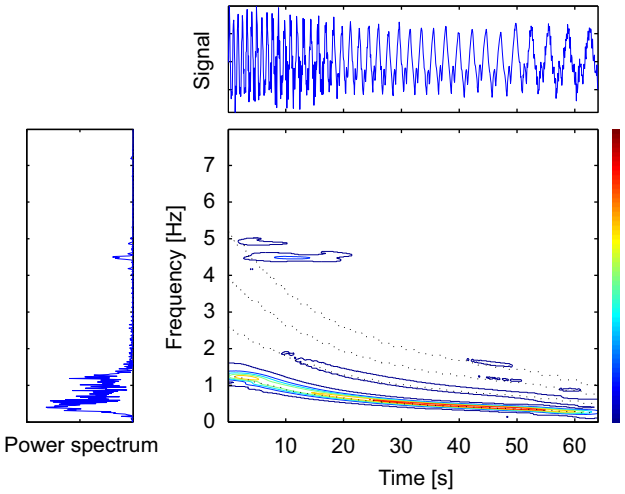

Fig. 10. Short time Fourier transform spectrogram.

Table 3
Regressed coefficients of polynomial fitting model.

Harmonic order	c_0	c_1	c_2	c_3
1	1.3428	−0.0461	0.8265×10^{-3}	-5.7648×10^{-6}
2	2.6856	−0.0922	1.6530×10^{-3}	-11.5295×10^{-6}
3	4.0284	−0.1383	2.4794×10^{-3}	-17.2943×10^{-6}
4	5.3712	−0.1844	3.3059×10^{-3}	-23.0590×10^{-6}

rotating frequency according to the short time Fourier transform spectrogram Fig. 10, rather than using the recorded rotating speed data during the shut-down process. A third-order polynomial model $\dot{v}(t)=c_3t^3+c_2t^2+c_1t+c_0$ is used to approximate the instantaneous frequency trajectory of the rotating frequency component on the time–frequency plane. The polynomial coefficients are regressed exploiting the time–frequency locations of all the peaks along the bottom trajectory (i.e. the rotating frequency trajectory) in Fig. 10 to ensure a high precision for the polynomial curve fitting, and their regressions are listed in Table 3. Fig. 11 shows the recorded, estimated and fitted rotating frequency. It can be seen that the polynomial fitting tracks the time evolution of the real rotating frequency very well. For the higher (up to forth) order harmonics of rotating frequency components, the polynomial coefficients are regressed using the corresponding order harmonic of time–frequency location data. The frequency location at very time instant is obtained by multiplying the estimated instantaneous rotating frequency with the corresponding order, rather than is estimated from Fig. 10 since the higher order harmonics are not very significant over the whole shut-down process. Then the polynomial fitting function is integrated with respect to time to get a corresponding demodulation phase function $\tilde{v}(t)=(1/4)c_3t^4+(1/3)c_2t^3+(1/2)c_1t^2+c_0t$, and a demodulation vector is constructed from the conjugate complex exponential of the demodulation phase function as $\exp\{-j2\pi[(1/4)c_3t^4+(1/3)c_2t^3+(1/2)c_1t^2+c_0t]\}$. For the

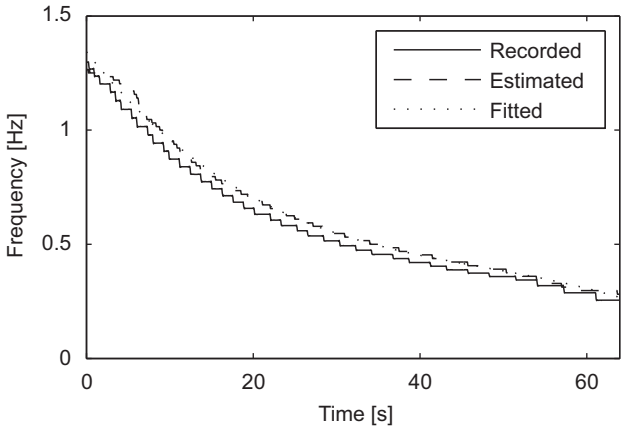

Fig. 11. Polynomial fitting of rotating frequency.

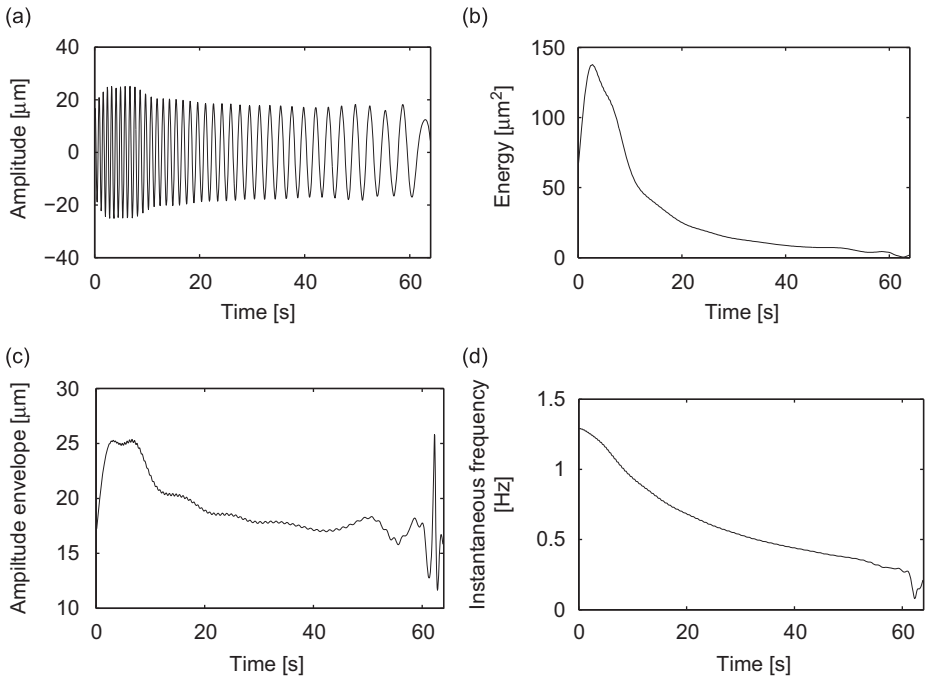

Fig. 12. Rotating frequency component: (a) waveform, (b) energy, (c) amplitude envelope and (d) instantaneous frequency.

transient component of constant frequency, a linear phase function with regards to time is used, i.e. $\exp(-j2\pi ct)$ where c=4.5.

The generalized demodulation algorithm is repeated until all the components of interest are separated into mono-components. Figs. 12–16 show the separated component waveform, as well as their estimated instantaneous energy, amplitude envelope and instantaneous frequency, respectively.

Fig. 17 shows the time–frequency energy distribution constructed with the above five separated components. It is obvious that all the prominent components are clearly identified. The time variation of each component is tracked in a fine

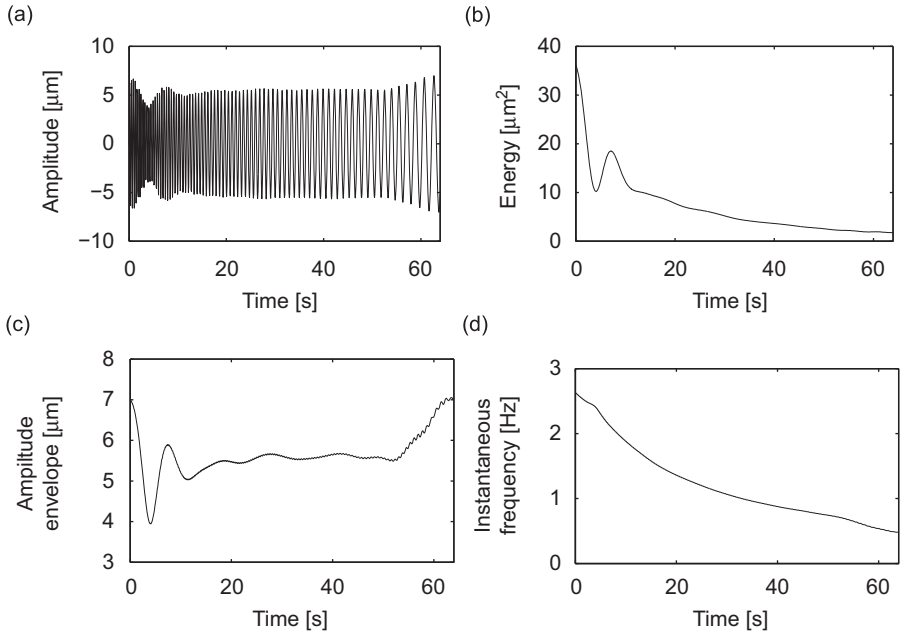

Fig. 13. Twice rotating frequency component: (a) waveform, (b) energy, (c) amplitude envelope and (d) instantaneous frequency.

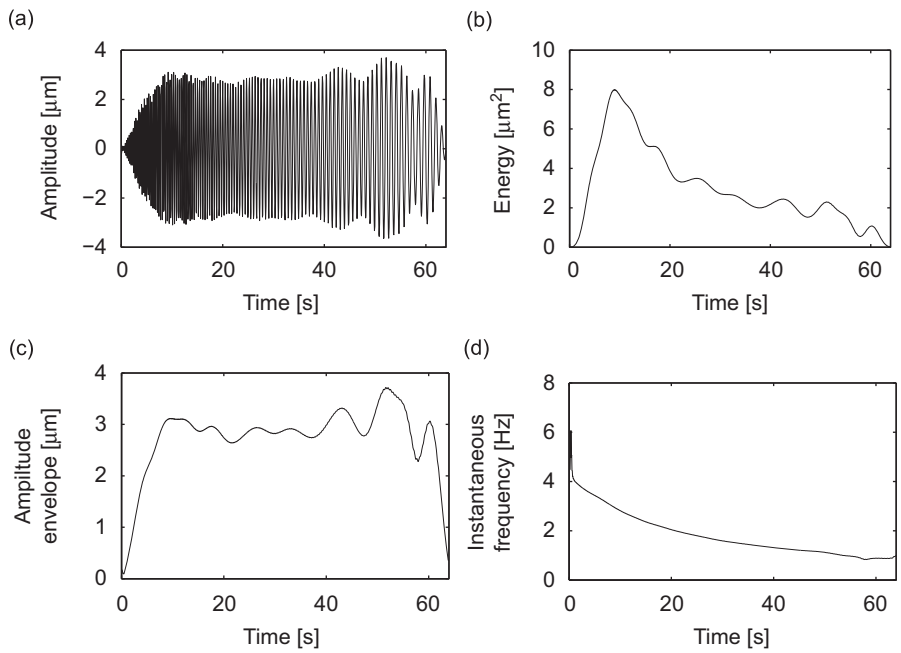

Fig. 14. Three times rotating frequency component: (a) waveform, (b) energy, (c) amplitude envelope and (d) instantaneous frequency.

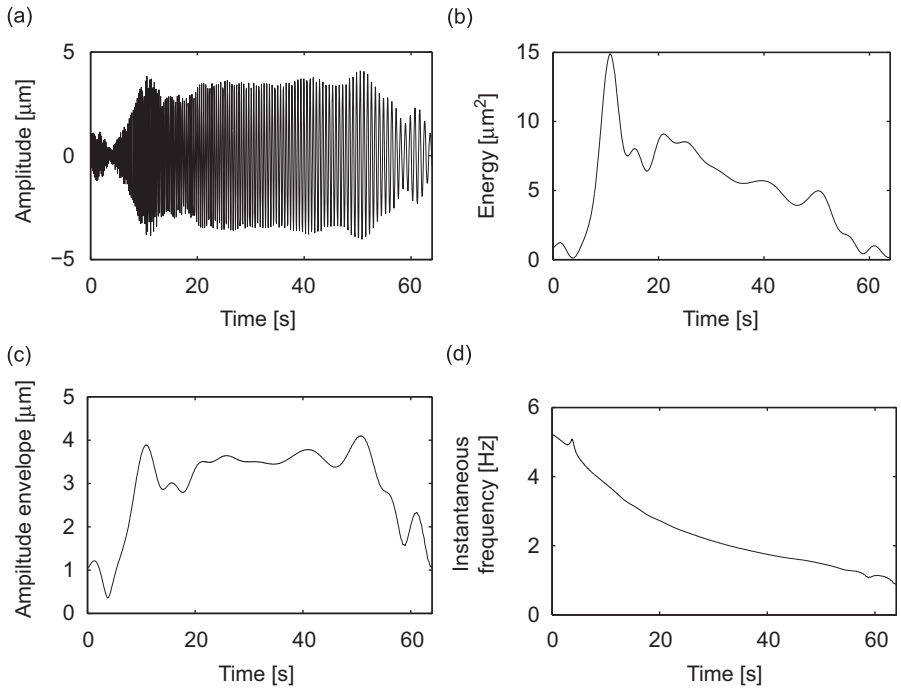

Fig. 15. Four times rotating frequency component: (a) waveform, (b) energy, (c) amplitude envelope and (d) instantaneous frequency.

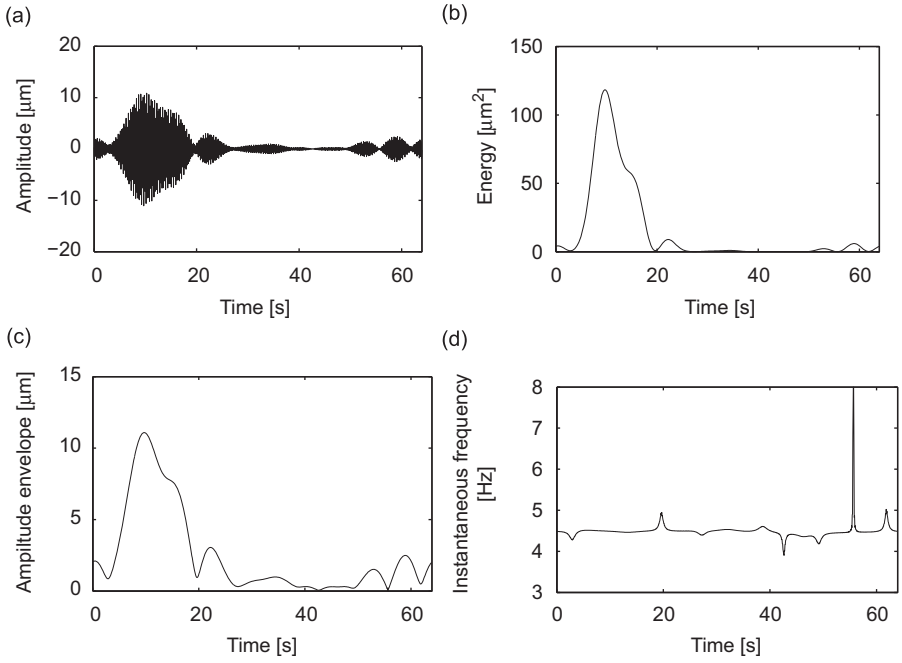

Fig. 16. Constant frequency transient component: (a) wave form, (b) energy, (c) ampiltude envelope and (d) instantaneous frequency.

resolution. The rotating frequency and its multiples up to four times exist throughout the shut-down process. The transient component around 4.5 Hz appears during the first half and near the end of the process. The vibration source energy is mainly associated with the rotating frequency during the beginning of the shut-down process, i.e. [0,12] s, and with the transient component of almost constant frequency around 4.5 Hz during the interval [7,17] s.

The time–frequency feature shown in Fig. 17 is different from that in Fig. 10. It is mainly due to reason that the energy defined in the two methods are completely different. In the energy operator, the energy refers to the total mechanical

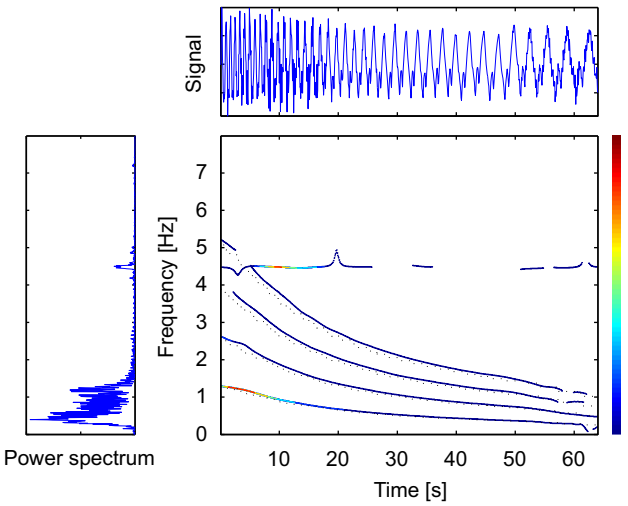

Fig. 17. Time–frequency energy distribution.

energy needed to generate the vibration signal, and is proportional to the product of the squared amplitude and the squared frequency, whereas in many traditional time–frequency analysis methods, the energy is simply defined as directly proportional to the square of signal amplitude. The time evolution of the rotating frequency multiples over the whole process is fully revealed in Fig. 17, because although the amplitudes of the 2–4 times rotating frequency are smaller compared to the rotating frequency, their frequencies are bigger so that their energy is not negligible. But in Fig. 10, part of the time evolution information is lost, because the amplitudes of the 2–4 times rotating frequency are small so that their energy in terms of squared amplitude is not comparable to that of the rotating frequency.

During shut-down processes, the guide vane opening controls the water flow amount and plays an important role on the rotation of hydroturbines. With the guide vane closing gradually, less and less water flows though and acts on the hydroturbine, i.e. less and less energy input into the hydroturbine, so the total source energy generating the vibration is also becoming smaller and smaller. In addition, the damping caused by the friction and water, etc, also absorbs the energy and thereby attenuates the vibration. Therefore, the energy decreases along time from a global point of view, so does the rotating speed, despite the almost constant amplitude (e.g. in the interval of 20–50 s) of the rotating frequency and its multiple components. This fact is fully revealed in the subplots (b) of Figs. 12–16.

In summary, the time–frequency features of the rotor vibration signal is revealed in the time–frequency energy distribution, by virtue of the advantage of generalized demodulation in separating mono-components from signals, and the advantage of the energy separation algorithm in finely tracking the instantaneous changes in signals. It shows the potential of the proposed time–frequency method in analyzing nonstationary rotor vibration signals.

5. Conclusions

The energy separation algorithm is improved with the iterative generalized demodulation method as a preprocessing tool, so it can be generalized to a broad range of signals as long as the instantaneous frequency trajectories of desired signal components do not intersect the others on the time–frequency plane.

The time–frequency energy distribution based on the proposed method has fine time–frequency resolution, due to the good adaptability of the energy operator to instantaneous changes in signals. It is also free of cross term interferences, because the instantaneous frequency is calculated from mono-components, and the energy separation algorithm does not involve integral of quadratic terms.

The good performance of the proposed method, including the iterative generalized demodulation based energy separation algorithm and the time–frequency energy distribution, in analyzing time-varying modulated multi-component signals, are illustrated by both simulation signals and an on-site measured hydroturbine rotor vibration signal during a shut-down transient process. This method has potential to analyze many time-varying modulated multi-component signals, especially nonstationary rotor vibration signals.

Acknowledgment

This work is supported by National Natural Science Foundation of China (51075028, 50705007), Scientific Research Foundation for Returned Overseas Chinese Scholars, Ministry of Education, Beijing Natural Science Foundation (3102022), China, and Natural Sciences and Engineering Research Council of Canada.

References

[1] P.D. McFadden, J.D. Smith, Vibration monitoring of rolling element bearings by the high-frequency resonance technique-a review, *Tribology International* 17 (1984) 1–18.
[2] P.D. McFadden, Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration, *Journal of Vibration Acoustics Stress and Reliability in Design-Transactions of the ASME* 108 (1986) 165–170.
[3] J.F. Kaiser, On a simple algorithm to calculate the 'energy' of a signal, *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 1, 1990, pp. 381–384.
[4] J.F. Kaiser, On Teager's energy algorithm and its generalization to continuous signals, *Proceedings of 4th IEEE Digital Signal Processing Workshop* (1990).
[5] J.F. Kaiser, Some useful properties of Teager's energy operators, *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 3, 1993, pp. 149–152.
[6] P. Maragos, J.F. Kaiser, T.F. Quatieri, On amplitude and frequency demodulation using energy operators, *IEEE Transactions on Signal Processing* 41 (4) (1993) 1532–1550.
[7] P. Maragos, J.F. Kaiser, T.F. Quatieri, Energy separation in signal modulations with application to speech analysis, *IEEE Transactions on Signal Processing* 41 (10) (1993) 3024–3051.
[8] A.C. Bovik, P. Maragos, T.F. Quatieri, AM-FM energy detection and separation in noise using multiband energy operators, *IEEE Transactions on Signal Processing* 41 (12) (1993) 3245–3265.
[9] A. Potamianos, P. Maragos, A comparison of the energy operator and Hilbert transform approaches for signal and speech demodulation, *Signal Processing* 37 (1) (1994) 95–120.
[10] J. Cheng, D. Yu, Y. Yang, The application of energy operator demodulation approach based on EMD in machinery fault diagnosis, *Mechanical Systems and Signal Processing* 21 (2) (2007) 668–677.
[11] A.M. Bassiuny, X. Li, Flute breakage detection during end milling using Hilbert-Huang transform and smoothed nonlinear energy operator, *International Journal of Machine Tools and Manufacture* 47 (6) (2007) 1011–1020.
[12] M. Liang, I. Soltani Bozchalooi, An energy operator approach to joint application of amplitude and frequency-demodulations for bearing fault detection, *Mechanical Systems and Signal Processing* doi:10.1016/j.ymssp.2009.12.007.
[13] B. Santhanam, Generalized energy demodulation for large frequency deviations and wideband signals, *IEEE Signal Processing Letters* 11 (3) (2004) 341–344.
[14] S. Olhede, A.T. Walden, A generalized demodulation approach to time frequency projections for multicomponent signals, *Proceedings of the Royal Society (A)* 461 (2005) 2159–2179.
[15] J. Cheng, Y. Yang, D. Yu, Application of the improved generalized demodulation time–frequency analysis method to multi-component signal decomposition, *Signal Processing* 89 (2009) 1205–1215.
[16] L. Cohen, Time frequency distributions: a review, *Proceedings of the IEEE* 77 (1989) 941–981.
[17] F. Hlawatsch, G.F. Boudreaux-Bartels, Linear and quadratic time frequency signal representations, *IEEE Signal Processing Magazine* 9 (2) (1992) 21–67.
[18] J. Hammond, P. White, The analysis of non-stationary signals using time frequency methods, *Journal of Sound and Vibration* 190 (1996) 419–447.
[19] A.H. Nuttalll, E. Bedrosian, On the quadrature approximation on the Hilbert transform of modulated signals, *Proceedings of the IEEE* 54 (1966) 1458–1459.