Signal Processing 138 (2017) 146–158

Contents lists available at ScienceDirect

Signal Processing

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

A time varying filter approach for empirical mode decomposition

Heng Li[a,b,1,*], Zhi Li[a,c], Wei Mo[a]

[a] School of Mechano-Electronic Engineering, Xidian University, Xi'an 710071, China
[b] Guangxi Transportation Research Institute Company Limited, Nanning 530007, China
[c] Guilin University of Aerospace Technology, Guilin 541004, China

ARTICLE INFO

Article history:
Received 2 November 2016
Revised 4 February 2017
Accepted 18 March 2017
Available online 20 March 2017

Keywords:
Empirical mode decomposition
Time varying filter
Adaptive signal analysis
Time-frequency analysis
Mode mixing

ABSTRACT

A modified version of empirical mode decomposition (EMD) is presented to solve the mode mixing problem. The sifting process is completed using a time varying filter technique. In this paper, the local cut-off frequency is adaptively designed by fully facilitating the instantaneous amplitude and frequency information. Then nonuniform B-spline approximation is adopted as a time varying filter. In order to solve the intermittence problem, a cut-off frequency realignment algorithm is also introduced. Aimed at improving the performance under low sampling rates, a bandwidth criterion for intrinsic mode function (IMF) is proposed. The proposed method is fully adaptive and suitable for the analysis of linear and non-stationary signals. Compared with EMD, the proposed method is able to improve the frequency separation performance, as well as the stability under low sampling rates. Besides, the proposed method is robust against noise interference.

1. Introduction

For many type of signals, such as vibration signals, water waves and biological signals, it is often desirable to decompose the given signal into simple components (or mono-component signals) for further analysis. One of the most common examples is in the field of time-frequency analysis, where the Hilbert spectrum exhibits valuable performance for mono-component signals, but fails to provide meaningful information for multi-component signals [1]. Empirical mode decomposition (EMD) [2] is a method for decomposing complex, multi-component signals into several elementary oscillations, called intrinsic mode functions (IMFs). An IMF has to satisfy the following two conditions [2]: (1) the number of extrema and zero-crossings must either be equal or differ at most by one and (2) the local average of the upper and lower envelopes is zero. Although EMD has been proved remarkably effective in many applications [3–6], some problems still remain. One of these problems is that EMD fails to distinguish components whose frequencies lie within an octave [7] (we refer to this as the separation problem). Another problem is that EMD is vulnerable to intermittence such as noise [8] (we refer to this as the intermittence problem). As a consequence, the resulting IMFs may contain widely spread scales, known as the mode mixing problem. Since both the separation problem and the intermittence problem can cause mode mixing, a number of methods are proposed to address the two problems.

For the separation problem, it is reported that EMD can separate two components only when their frequency ratio is below a particular cut-off (about 0.65). In order to achieve a better separation performance, specific models, such as linear models [9,10] and AM-FM models [11–13], are studied. However, these specialized models are difficult to use for real-life signals. Since EMD involves the interpolation of local extrema, some methods are proposed by finding alternative points for interpolation [14,15], and achieving a separation performance better than 0.8. However, these methods are often complicated and difficult to interpret. Later, the focus of research was on more comprehensive ways, where analytic methods [16,17] are proposed. The analytic methods defined analytical expressions, and some achieved a separation performance better than 0.9.

For the intermittence problem, noise-assisted techniques such as ensemble EMD (EEMD) [18] and noise-assisted MEMD (N-A MEMD) [19,20] are popular methods, which are based on the filter bank structure of EMD [21]. With EEMD and N-A MEMD, the decomposed results were uniformly aligned according to the cut-off frequency bands. However, the noise-assisted methods have some disadvantages: (1) Parameters such as amplitude of noise and ensemble number are quite difficult to choose, making it no longer adaptive. (2) It still fails to separate modes whose frequencies lie within an octave. (3) If we regard EMD as a filtering

* Corresponding author. Guangxi Transportation Research Institute Company Limited, No.6, Gaoxin 2 road, Nanning, Guangxi, China.
E-mail address: wudaomana@gmail.com (H. Li).

[1] This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

http://dx.doi.org/10.1016/j.sigpro.2017.03.019
0165-1684/

process, its cut-off frequency is time-varying. On the contrary, the cut-off frequency of the noise-assisted methods is constant, which is not suitable for non-stationary signals. Besides the noise-assisted techniques, filtering based methods such as variational mode decomposition (VMD) [22] and empirical wavelet transform (EWT) [23] have been recently proposed, which are based on Wiener filtering and wavelet filter bank respectively. Although VMD and EWT have been proved to be robust to noise, these methods require an appropriate choice of parameters.

In this paper, we alleviate the mode mixing problem by the use of time varying filter (TVF). The cut-off frequency of a TVF is time varying, which is suitable for non-stationary signals. Our idea is quite simple: find out the local cut-off frequency and then perform time varying filtering. During the sifting process, the given signal is filtered and divided into two parts, i.e., the local higher frequency (LHF) component and the local lower frequency (LLF) one. By successively applying the TVF, the obtained LHF component is a local narrow-band signal, which has similar properties to the IMF.

Compared with existing methods, our proposed method has the following characteristics. Firstly, we fully address the separation problem [7] as well as the intermittence problem [8]. By contrast, many of the existing methods deal with the two problems separately, such as [9–13] and [18–20,23]. Secondly, our method adopts TVF in the sifting process. Thus, it is able to solve the mode mixing problem while simultaneously preserving the time varying feature, which is a very important aspect of EMD. By contrast, the filter cut-off frequencies of the filtering based methods, such as EEMD, N-A MEMD, VMD and EWT, are constant with respect to time, which is not suitable for non-stationary signals. Thirdly, the stopping criterion is improved, by considering signals under low sampling rates. We require the resulting IMF to be of a local narrow-band, and do not require the upper and lower envelopes to be symmetric. This is different from most of the existing methods. As a result, our method is able to achieve a robust performance under low sampling rates. Finally, the parameters in our method have clear physical meanings and are easy to be selected. By contrast, some existing methods such as EEMD, VMD and EWT, require many parameters, and the parameters are not easy to be selected adaptively. The proposed method is referred to as time varying filtering based EMD (TVF-EMD). As a supplement, the MATLAB source code for TVF-EMD is available from http://samd.site90.com/tvfemd.php.

The remainder of the paper is organized as follows. Section 2 briefly introduces the principle of EMD and the limitations of IMF. In Section 3 we show that B-spline approximation is a special form of TVF. In Section 4, we focus on the central topic of our work, i.e., how to estimate the filter cut-off frequency, which is capable of dealing with both the separation problem and the intermittence problem. In Section 5, based on the concepts outlined in Sections 3 and 4, we present the sifting process of our proposed method. In this section, we also discuss the stopping criterion and some properties of the proposed method. Section 6 provides experimental validation of our method along with performance comparisons. Finally in Section 7, we will offer a conclusion to our paper.

2. EMD, IMF and local narrow-band signal

2.1. Empirical mode decomposition

EMD decompose a given signal $x(t)$ into a finite set of mono-component like components (IMFs), plus a nonzero mean residual $r(t)$.

$$x(t) = \sum_{i=1}^{N} imf_i(t) + r(t) \tag{1}$$

where $imf_i(t)$ is the i-th IMF. To obtain each IMF, an iterative procedure called the sifting process is used. The sifting process of EMD is mainly carried out by two steps: (1) Estimate the "local mean". (2) Recursively subtract the local mean from the input signal until the resulting signal becomes an IMF. Given an input signal$x(t)$, let$h(t) = x(t)$, the sifting process of EMD can be summarized as Algorithm 1.

Due to the lack of an exact definition of the mono-component, IMF is used as an alternative. However, IMF does not guarantee that it has only one oscillatory mode. As we will discuss later, it cannot even guarantee that IMF is narrow-band. In EMD, the local mean is defined as the average of the upper and the lower envelopes, and the upper and lower envelopes are obtained by cubic spline interpolation. This process is difficult to express as an analytic formula. Besides, the obtained IMF may not be meaningful due to its limitations.

2.2. Limitations of IMF

In EMD, IMF is used to replace the mono-component signal based on the observations of the narrow band signal [2]. Recalling the second condition of IMF, the local average of the upper and lower envelope of an IMF should be zero. Although the second condition of IMF is to make the instantaneous frequency of an IMF meaningful, it has its limitations.

The first limitation is that the second condition of IMF is too rigid for stopping criterion. The local mean may not be zero everywhere, even the input signal is narrow-band. In order to make the stopping criterion easier to be realized, Huang [2] defined a stopping criterion called standard deviation (SD), by limiting the variation between two consecutive sifting results. Rilling [24] proposed a 3-threshold criterion and Damerval [25] proposed a criterion based on the number of iterations. None of the aforementioned criteria explore the physical information and therefore have no reasonable interpretation.

The second limitation is that the second condition of IMF may not be valid under low sampling rates. Consider the continuous-time signals, the definition of IMF seems intuitive and reasonable. In practical applications, signals are always sampled and presented in digital time series. According to the second condition of IMF, EMD requires the upper and lower envelopes of an IMF to be symmetric [2]. However, under low sampling rates, the upper and lower envelopes of a mono-component signal cannot be guaranteed to be symmetric. As a consequence, errors inevitably occur. For more detailed information about the sampling influence, one can refer to Rilling's work [26]. In order to deal with the low sampling problem, interpolation (or up-sampling) [27,28] is employed. If EMD is implemented without up-sampling, the sampling frequency is required to be 5 times higher than the Nyquist frequency [29].

2.3. Local narrow-band signal

Since IMF has the aforementioned shortcomings, it seems that the definition of IMF still requires further development. A substitution for the mono-component signal is the local narrow-band signal, which is able to provide a meaningful Hilbert spectrum [30]. The definition of local narrow-band is intrinsically related to the instantaneous bandwidth. If the local instantaneous bandwidth is small enough, then we can describe the signal as being local narrow-band. In Section 5.1, we will discuss in detail how to determine a local narrow-band signal. To obtain a local narrow-band signal, TVF is used in this paper. By successively applying the filter and filtering out the LLF component, the obtained LHF signal is local narrow-band.

3. Construction of a time varying filter

In EMD, the estimation of local mean can be viewed as a special form of low pass filtering [21]. Thus it is natural to resort to low pass filtering methods. Traditional linear filters are not suitable candidates due to the group delay and pass band ripple. Moreover, the local cut-off frequency of the linear filter is constant, making it difficult to deal with nonlinear and non-stationary signals.

In this work, B-spline approximation is adopted as a TVF. We believe that B-spline approximation filter is quite interesting and deserves further investigation. Besides, it is easy to be constructed from the filter cut-off frequency. It should be pointed out that B-spline filter is far from the only choice. In our opinion, many TVF techniques [31–33] are potential candidates. One can adopt these techniques and compare their performances, however, this is beyond the scope of our work. Since the use of B-splines for EMD improvement has been reported, it is necessary to present some related works here. Chen et al. [34] and Yang et al. [35] used B-spline interpolation to provide a better analytic expression for EMD. It can be seen that most present works use B-splines as an interpolation tool. In our work, B-spline approximation is used as a filter, whose cut-off frequency is time varying.

First it is worthwhile to review some mathematical results related to B-spline approximation of equally spaced knots (uniform B-spline approximation). B-spline approximation is a problem of constructing polynomial splines that approximate the input signal. We use B-spline functions, which are piecewise polynomials. The polynomial segments are then joined together to form the desired signal. The joining points of the polynomial segments are called knots. To show how B-spline functions form a signal, let $\beta^n(t)$ be the B-spline function of order n and m be the step size of the knot sequence, any signal in this B-spline space is defined by [36,37]

$$g_m^n(t) = \sum_{k=-\infty}^{\infty} c(k)\beta^n(t/m - k) \tag{2}$$

where $c(k)$are the B-spline coefficients. It can be seen that the B-spline function are enlarged by a factor of m. Note that the signal (or approximation result) is determined by n, m and $c(k)$. It means that given the B-spline order and knots, B-spline approximation is to determine the B-spline coefficients $c(k)$ that minimizes the approximation error. Let $b_m^n(t) := \beta^n(t/m)$ and the asterisk denotes the convolution operator, Eqs. (3)–(6) have been derived by Unser et al. [36]. For a given signal $x(t)$, $c(k)$ is determined by minimizing the approximation error ε_m^2:

$$\varepsilon_m^2 = \sum_{t=-\infty}^{+\infty} \left(x(t) - [c]_{\uparrow m} * b_m^n(t)\right)^2 \tag{3}$$

where $[\cdot]_{\uparrow m}$ is the up-sampling operation (adding zeros between each sample) by m.

After introducing the concept of B-spline approximation, we further reveal its low-pass filtering property. The solution of $c(k)$ is

$$c(k) = [p_m^n * x]_{\downarrow m}(k) \tag{4}$$

where $[\cdot]_{\downarrow m}$is the down-sampling operation (decimation) by m and p_m^n is the pre-filter denoted by

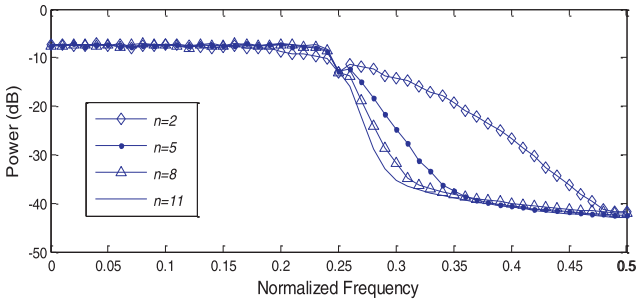
$$p_m^n = \left[\left([b_m^n * b_m^n]_{\downarrow m}\right)^{-1}\right]_{\uparrow m} * b_m^n \tag{5}$$

We can rewrite Eq. (2) as

$$g_m^n(t) = [p_m^n * x]_{\downarrow m} * b_m^n(t). \tag{6}$$

By observing (6), the B-spline approximation of a signal is carried out in three steps [38]. The signal x is first band-limited through a pre-filterp_m^n. Then the band-limited signal is decimated by a factor of m. Finally, the approximation is reconstructed using a post-filterb_m^n. It appears that the use of B-spline approximation is a special form of low pass filtering. Fig. 1(a) shows the frequency response for $m = \{2\}$ and $n = \{2, 5, 8\}$ and it is apparent that the filter approaches an ideal low pass filter for increasing values of spline order n. Fig. 1(b) shows the frequency response for $m = \{2, 4, 8, 16, 32\}$,$n = \{14\}$. It is also worth noting that the cut-off frequency is exactly $1/2\,m$. This means that the cut-off frequency is determined by the knot spacing m.

a



b

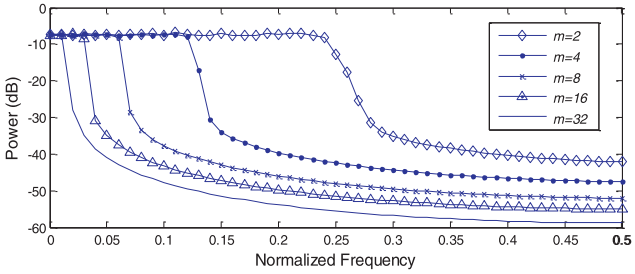


Fig. 1. Frequency responses of B-spline approximation for various values of m and n. (a) $m = 2$, $n = 2$, 5,8,11. (b) $m = 2$, 4, 8, 16, 32, $n = 14$.

For non-uniform splines, an analytic expression is still missing and merits further study. Yang [39] has recently revealed that non-uniform B-splines behave similarly to a TVF. Thus we believe that the non-uniform B-spline approximation is closely related to the uniform B-spline approximation. An evidence indicates that after solving Eq. (3), the local cut-off frequency will asymptotically converge according to the knot distribution. We present an example to further validate its TVF property. It is illustrated by applying B-spline approximation on a noisy non-stationary signal. Let us consider a noisy chirp signal $x(t)$, which is defined as

$$x(t) = \cos(12\pi t + 3.2\pi t^2) + n(t) \tag{7}$$

where $t \in [0, 4]$ and $n(t)$ is the white Gaussian noise with a standard deviation $\sigma = 0.2$. Fig. 2(b) shows the time domain waveform of $x(t)$. In order to retrieve the chirp signal from the surrounding noise, it is feasible to construct a TVF according to the information of the clean signal. From this point of view, the timings corresponding to the extrema of the clean signal are taken as knots. The approximation result is shown in Fig. 2(a). For the given knots, B-spline approximation successfully extracts the chirp signal. This means that the local cut-off frequency of B-spline filter is determined by the knot distribution. In this example, the information from the clean signal is used to build the TVF. However, the clean signal cannot be known as a priori. In practice, the information used to build a TVF is estimated from the input signal. In Section 4, we will focus on how to achieve such a goal.

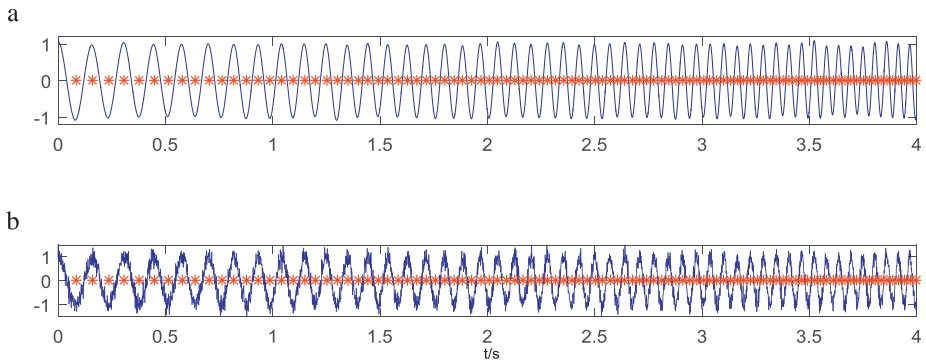


Fig. 2. Approximation result, knots are shown in asterisk. (a) Approximation result. (b) Noisy signal.

4. Estimation of the local cut-off frequency

4.1. Addressing the separation problem

We focus on how to estimate the local cut-off frequency which is capable of separating close components within the frequency domain. The estimated local cut-off frequency is used to build a TVF. By iteratively filtering out the LLF component, the obtained signal is local narrow-band. Thus, the key concern is how to estimate a meaningful local cut-off frequency. Consider four multi-component signals: (1) a two-tone signal comprised of 50 Hz and 15 Hz, (2) a three-component signal comprised of 50 Hz, 42.5 Hz and 15 Hz

$$x(t) = \cos(2\pi \times 50t) + \cos(2\pi \times 42.5t) + \cos(2\pi \times 15t) \tag{8}$$

(3) a realization of white Gaussian noise, (4) a noisy chirp signal.

In these cases, what should the local cut-off frequencies be? Ideally, via filtering, the unbiased LHF and LLF components should be close to local narrow-band, indeed the closer the better. In the case of the two-tone signal, it is obvious that the bisecting frequency should be taken as the local cut-off frequency f_c, i.e., $f_c = (50+15)/2 = 32.5$Hz. For the three-component signal, note that Eq. (8) can be represented as a two-component signal consisting of 46.25 Hz and 15 Hz

$$x(t) = 2\cos(2\pi \times 3.75t)\cos(2\pi \times 46.25t) + \cos(2\pi \times 15t) \tag{9}$$

Both of these two components are narrow-band signals whose amplitude varies much slower than the phase, which is good. Thus, the local cut-off frequency can be $f_c = (46.25+15)/2 = 30.625$Hz. In the case of white Gaussian noise, it is difficult to filter the noise into two local narrow-band signals. For an unbiased estimation of LHF and LLF components, we hope the average local cut-off frequency to be 0.25 (normalized frequency). For the noisy chirp signal, since there are distinguished components, i.e., noise and chirp signal, we hope to extract the chirp from noise. Thus, as in Section 3, the local cut-off frequency should be able to "track" the local characteristic time scale of the chirp signal.

From now on, we will discuss how to estimate the local cut-off frequency in an analytical way. In this paper, instantaneous amplitude and frequency are used. Since the instantaneous amplitude and frequency are well-defined following the HHT approach, we employ the analytic signal (complex valued signal). Given a real valued signal $x(t)$, the analytic signal is obtained using Hilbert transform. The Hilbert transform of $x(t)$ is denoted by $\hat{x}(t)$, then the analytic signal corresponding to $x(t)$ is

$$z(t) = x(t) + j\hat{x}(t) = A(t)e^{j\varphi(t)} \tag{10}$$

where $A(t) = \sqrt{x^2 + \hat{x}(t)^2}$is the instantaneous amplitude and $\varphi(t) = \arctan[\hat{x}(t)/x(t)]$is the instantaneous phase. Consider a multi-component signal comprised of N narrow-band components

$$x(t) = \sum_{m=1}^{N} a_m(t)\cos[\varphi_m(t)] \tag{11}$$

where $a_m(t)$ and $\varphi_m(t)$ are respectively the amplitude and phase of the m-th component. Note that $a_m(t)\cos[\varphi_m(t)]$ can be used to model many type of signals. For example, $\cos\omega t$ is a stationary signal, $\cos\omega t^2$is a non-stationary signal and $\cos(\omega t + sin\omega t)$ is a model for nonlinear signal [2]. Obviously,$a_m(t)$is non-negative and $\varphi_m(t)$ is non-decreasing, more importantly, both $a_m(t)$ and $\varphi'_m(t)$ vary very slowly [2]. By using Eq. (10), an analytic signal for $x(t)$ is obtained. For a band-limited signal $x(t)$, Xuan et al. [30] have derived its instantaneous amplitude $A(t)$ and frequency $\varphi'(t)$ as follows

$$A^2(t) = \sum_{m=1}^{N}\sum_{n=1}^{N} a_m(t)a_n(t)\cos[\varphi_m(t) - \varphi_n(t)] \tag{12}$$

$$\varphi'(t) = \frac{1}{A^2(t)}\sum_{m=1}^{N}\left(\varphi'_m(t)\sum_{n=1}^{N} a_m(t)a_n(t)\cos[\varphi_m(t) - \varphi_n(t)]\right) + \sum_{m=1}^{N}\sum_{n=1}^{N}\frac{a'_m(t)a_n(t)}{A^2(t)}\sin[\varphi_m(t) - \varphi_n(t)] \tag{13}$$

If $A(t)e^{j\varphi(t)}$ is narrow-band, then $\varphi_m(t)$ are closely spaced. From Eqs. (12) and (13), it can be verified that both $A(t)$ and $\varphi'(t)$ are slow varying components. Because $x(t)$ is a multi-component signal, Eqs. (12) and (13) may not be meaningful in interpreting its instantaneous amplitude and frequency [1]. However, it is mathematically valuable. For example, in [40], the instantaneous amplitude and frequency of a multi-component signal are used for demodulation. In [41], the instantaneous information is used for estimating the frequency of the major component. In this paper, we use this information for cut-off frequency estimation.

Mathematically, a multi-component signal can also be expressed as a combination of two signals, i.e.

$$z(t) = A(t)e^{j\varphi(t)} = a_1(t)e^{j\varphi_1(t)} + a_2(t)e^{j\varphi_2(t)} \tag{14}$$

Our goal is to find out the bisecting frequency between $\varphi_1(t)$ and $\varphi_2(t)$, so that we can build a TVF. Obviously, there are infinite possible combinations. For example, a possible combination of Eq. (8) is illustrated in Eq. (9). Ideally, the two components should be local narrow-band. For a N-component signals with $N > 2$, it seems difficult to achieve such a goal. Thus we hope the two components to be close to local narrow-band, the closer the better. According to the observation of local narrow-band signal, $a_1(t)$ and $a_2(t)$ should vary as slow as possible [2]. Similarly, $\varphi'_1(t)$ and $\varphi'_2(t)$ should vary as slow as possible. From Eq. (12) and Eq. (13), by setting$N = 2$, we have

$$A^2(t) = a_1^2(t) + a_2^2(t) + 2a_1(t)a_2(t)\cos[\varphi_1(t) - \varphi_2(t)] \tag{15}$$

$$\varphi'(t) = \frac{1}{A^2(t)}\big(\varphi'_1(t)(a_1^2(t) + a_1(t)a_2(t)\cos[\varphi_1(t) - \varphi_2(t)]) + \varphi'_2(t)(a_2^2(t) + a_1(t)a_2(t)\cos[\varphi_1(t) - \varphi_2(t)])\big) + \frac{1}{A^2(t)}\big(a'_1(t)a_2(t)\sin[\varphi_1(t) - \varphi_2(t)] - a'_2(t)a_1(t)\sin[\varphi_1(t) - \varphi_2(t)]\big) \quad (16)$$

Suppose that both $a_1(t)$ and $a_2(t)$ vary much slower than $\cos[\varphi_1(t) - \varphi_2(t)]$, according to Eq. (15), the local extrema of $A(t)$ is approximately determined by the local extrema of $\cos[\varphi_1(t) - \varphi_2(t)]$. This means that a minimum of $A(t)$is obtained when $\cos[\varphi_1(t) - \varphi_2(t)] = -1$. Similarly, a maximum of $A(t)$is obtained when $\cos[\varphi_1(t) - \varphi_2(t)] = 1$. Suppose a local minimum of $A(t)$is obtained at $t_{\min}$, then

$$\cos[\varphi_1(t_{\min}) - \varphi_2(t_{\min})] = -1 \quad (17)$$

Substituting Eq. (17) into Eqs. (15) and (16) yields

$$|a_1(t_{\min}) - a_2(t_{\min})| = A(t_{\min}) \quad (18)$$

$$\varphi'_1(t_{\min})\big(a_1{}^2(t_{\min}) - a_1(t_{\min})a_2(t_{\min})\big) + \varphi'_2(t_{\min}) \times \big(a_2{}^2(t_{\min}) - a_1(t_{\min})a_2(t_{\min})\big) = \varphi'(t_{\min})A^2(t_{\min}) \quad (19)$$

Because $A(t_{\min})$ is a local minimum of $A(t)$, we have $A'(t_{\min}) = 0$, then

$$a'_1(t_{\min}) - a'_2(t_{\min}) = 0 \quad (20)$$

Similarly, suppose a local maximum of $A(t)$is obtained at $t_{\max}$, then

$$\cos[\varphi_1(t_{\max}) - \varphi_2(t_{\max})] = 1 \quad (21)$$

$$\varphi'_1(t_{\max})\big(a_1{}^2(t_{\max}) + a_1(t_{\max})a_2(t_{\max})\big) + \varphi'_2(t_{\max}) \times \big(a_2{}^2(t_{\max}) + a_1(t_{\max})a_2(t_{\max})\big) = \varphi'(t_{\max})A^2(t_{\max}) \quad (22)$$

$$a_1(t_{\max}) + a_2(t_{\max}) = A(t_{\max}) \quad (23)$$

$$a'_1(t_{\max}) + a'_2(t_{\max}) = 0 \quad (24)$$

$a_1(t_{\min})$, $a_2(t_{\min})$, $\varphi_1(t_{\min})$ and $\varphi_2(t_{\min})$ can be computed by solving Eq. (17)~Eq. (20) while $a_1(t_{\max})$, $a_2(t_{\max})$, $\varphi_1(t_{\max})$ and $\varphi_2(t_{\max})$ can be computed by solving Eq. (21)~Eq. (24). After that, $a_1(t)$is obtained by interpolating the set of point $\{a_1(t_{\min}), a_1(t_{\max})\}$. Then $a_2(t)$, $\varphi_1(t)$ and $\varphi_2(t)$ can be obtained in a similar way. However, solving these equations is not an easy task. As an alternative, we propose an approximate method. Our goal is to find out the bisecting frequency of $\varphi_1(t)$ and $\varphi_2(t)$. A rough estimation is sufficient for the design of TVF, because slight errors in the cut-off frequency will not result in huge variations.

Let

$$\beta_1(t) = |a_1(t) - a_2(t)|$$
$$\beta_2(t) = a_1(t) + a_2(t) \quad (25)$$

$\beta_1(t)$ and $\beta_2(t)$ are yet to be determined. From Eq. (15), we have

$$\beta_1(t_{\min}) = A(t_{\min}) = |a_1(t_{\min}) - a_2(t_{\min})|$$
$$\beta_2(t_{\max}) = A(t_{\max}) = a_1(t_{\max}) + a_2(t_{\max}) \quad (26)$$

Because $a_1(t)$ and $a_2(t)$ are slow varying components, $\beta_1(t)$ can be estimated by finding the curve going through all the minima of $A(t)$. This is achieved via an interpolation from the set of point $A(t_{\min})$, i.e., $A(\{t_{\min}\})$. In the same way, $\beta_2(t)$ can be estimated by interpolating $A(\{t_{\max}\})$. Suppose $a_1(t) \geq a_2(t)$, $a_1(t)$ and $a_2(t)$ are able to be obtained by solving Eq. (25):

$$a_1(t) = [\beta_1(t) + \beta_2(t)]/2$$
$$a_2(t) = [\beta_2(t) - \beta_1(t)]/2 \quad (27)$$

Let

$$\eta_1(t) = \varphi'_1(t)[a_1^2(t) - a_1(t)a_2(t)] + \varphi'_2(t)[a_2^2(t) - a_1(t)a_2(t)]$$
$$\eta_2(t) = \varphi'_1(t)[a_1^2(t) + a_1(t)a_2(t)] + \varphi'_2(t)[a_2^2(t) + a_1(t)a_2(t)] \quad (28)$$

$\eta_1(t)$ and $\eta_2(t)$ are yet to be determined. From Eq. (16) we have

$$\eta_1(t_{\min}) = \varphi'(t_{\min})A^2(t_{\min}) = \varphi'_1(t_{\min})[a_1^2(t_{\min}) - a_1(t_{\min})a_2(t_{\min})] + \varphi'_2(t_{\min})[a_2^2(t_{\min}) - a_1(t_{\min})a_2(t_{\min})]$$
$$\eta_2(t_{\max}) = \varphi'(t_{\max})A^2(t_{\max}) = \varphi'_1(t_{\max})[a_1^2(t_{\max}) + a_1(t_{\max})a_2(t_{\max})] + \varphi'_2(t_{\max})[a_2^2(t_{\max}) + a_1(t_{\max})a_2(t_{\max})] \quad (29)$$

Because $a_1(t)$, $a_2(t)$, $\varphi'_1(t)$ and $\varphi'_2(t)$ are slow varying components, $\eta_1(t)$ can be estimated by interpolating a set of point $\varphi'(\{t_{\min}\})A^2(\{t_{\min}\})$ while $\eta_2(t)$ can be estimated by interpolating $\varphi'(\{t_{\max}\})A^2(\{t_{\max}\})$. Then $\varphi'_1(t)$ and $\varphi'_2(t)$ are obtained by solving Eq. (28)

$$\varphi'_1(t) = \frac{\eta_1(t)}{2a_1^2(t) - 2a_1(t)a_2(t)} + \frac{\eta_2(t)}{2a_1^2(t) + 2a_1(t)a_2(t)}$$
$$\varphi'_2(t) = \frac{\eta_1(t)}{2a_2^2(t) - 2a_1(t)a_2(t)} + \frac{\eta_2(t)}{2a_2^2(t) + 2a_1(t)a_2(t)} \quad (30)$$

The bisecting frequency is given by

$$\varphi'_{bis}(t) = \frac{\varphi'_1(t) + \varphi'_2(t)}{2} = \frac{\eta_2(t) - \eta_1(t)}{4a_1(t)a_2(t)} \quad (31)$$

It is easy to be validated that although we assume $a_1(t) \geq a_2(t)$, $\varphi'_{bis}(t)$ is unchangeable in the case of $a_1(t) < a_2(t)$. Given $\varphi'_{bis}(t)$, a signal $h(t)$ can be retrieved using Eq. (32). In this paper, B-spline approximation is used as a TVF. According to Section 3, extrema timings of $h(t)$ are taken as the knots. By doing this, the filter cut-off frequency is in accordance with $\varphi'_{bis}(t)$.

$$h(t) = \cos[\int \varphi'_{bis}(t)dt] \quad (32)$$

4.2. Addressing the intermittence problem

The cut-off frequency estimated in Section 4.1 may be affected by intermittence such as noise. Intermittence may occur in both the LHF component and LLF component. Fig. 3 shows some examples in a two-component model. For the case depicted in Fig. 3(a) and (b), the two components can be successfully separated. However, for Fig. 3(c) and (d), mode mixing occurs because the LLF component is segmented by the local cut-off frequency.

When intermittency occurs, the cut-off frequency changes abruptly, i.e., changes from a floor value to a peak value or vice versa. In an attempt to solve the mode mixing problem, the cut-off frequency should be realigned. A simple approach is to adjust the floor to remain aligned with the peak, so that the LLF component is less likely to be segmented by the cut-off frequency. The first step is to find out where intermittency occurs, i.e., the region where cut-off frequency changes dramatically. This can be achieved by setting a threshold on its changing rate within a certain time span. In this paper, the time span is chosen as the time interval between two adjacent maxima. In the time span where intermittency occurs, the floor and peak are then to be determined. If $\varphi'_{bis}(t)$ is increasing within the time span, it is on its rising edge, then $\varphi'_{bis}(t)$ to the left of the intermittence is considered to be a floor.

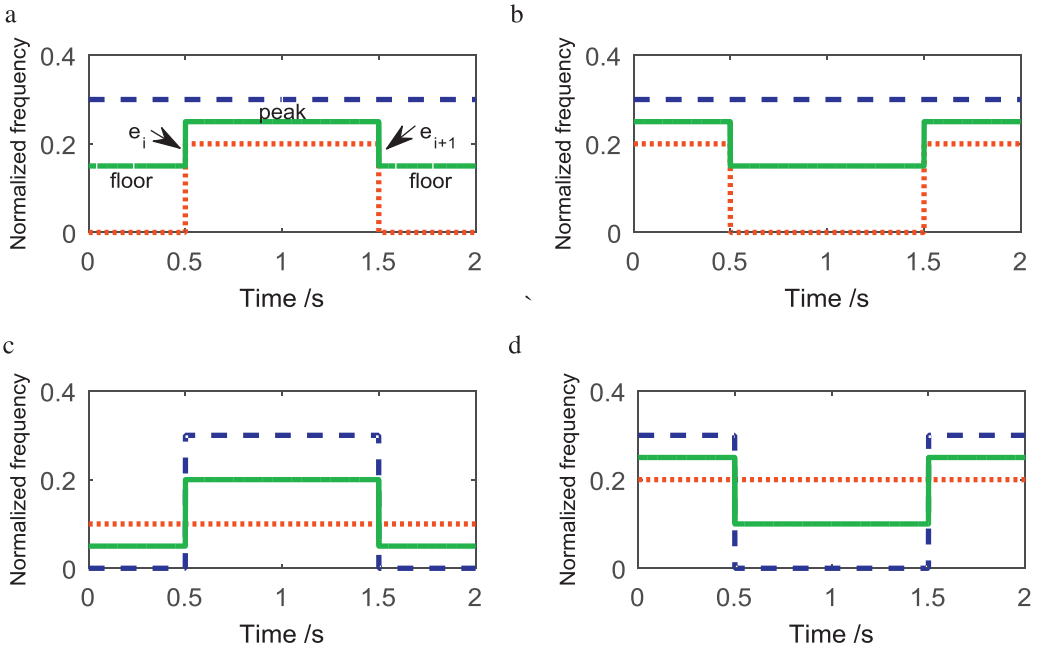


Fig. 3. Examples of mode mixing in a two-component model. The dotted lines correspond to the LLF component, the dashed lines correspond to the LHF component and the solid lines correspond to the local cut-off frequency. (a) (b) Intermittency occurs at the LLF component (c) (d) Intermittency occurs at the LHF component.

Algorithm 1. Sifting process of EMD.

Step 1. Extract local maxima and minima of $h(t)$.
Step 2. Compute upper envelope $u(t)$ and lower envelope $l(t)$ by cubic interpolating the local maxima and minima respectively.
Step 3. Obtain the local mean $m(t) = (u(t) + l(t))/2$.
Step 4. Subtract the local mean from $h(t)$ and update $h(t) = h(t) - m(t)$.
Step 5. Repeat steps (1) to (4) until $h(t)$ becomes an IMF.

Conversely, if $\varphi'_{bis}(t)$ is decreasing within the time span, it is on its falling edge, then $\varphi'_{bis}(t)$ to the right of the intermittence is considered to be a floor. Actually, the identification of a floor is achieved by examining the neighboring intermittences. Suppose e_{j-1}, e_j and e_{j+1} are three neighboring intermittences. If $\varphi'_{bis}(e_j)$ is on its rising edge, then $\varphi'_{bis}(e_{j-1} : e_j)$ is considered to be a floor. For the colon operator, $\varphi'_{bis}(e_{j-1} : e_j)$ is the same as $\varphi'_{bis}(\{t|t \in [e_{j-1}, e_j)\})$. Similarly, if $\varphi'_{bis}(e_j)$ is on its falling edge, then $\varphi'_{bis}(e_j : e_{j+1})$ is considered to be a floor. Finally, the cut-off frequency is realigned, by interpolating between the peaks. Given an input signal$x(t)$ and the cut-off frequency $\varphi'_{bis}(t)$, the algorithm for local cut-off frequency realignment is shown in Algorithm 2.

In order to solve the intermittence problem, Algorithm 2. must be performed before applying TVF. In step 2, the frequency changing rate between two adjacent maxima is limited by ρ. For a signal whose frequency content varies rapidly, an intermittence may be declared incorrectly. However, only for very strong FM signals will the limitation become a problem. Therefore, we believe this algorithm is applicable to modestly modulated signals.

Algorithm 2. Local cut-off frequency realignment.

Step 1. Locate the maxima timing of $x(t)$, denoted as u_i, $i = 1, 2, 3...$
Step 2. Find out all intermittences, denoted as e_j $j = 1, 2, 3...$, that satisfy
$$\frac{\max(\varphi'_{bis}(u_i : u_{i+1})) - \min(\varphi'_{bis}(u_i : u_{i+1}))}{\min(\varphi'_{bis}(u_i : u_{i+1}))} > \rho \quad (\rho = 0.25 \text{ is used in this paper})$$
then the timing of u_i is taken as an intermittence, i.e., $e_j = u_i$.
If $\varphi'_{bis}(u_{i+1}) - \varphi'_{bis}(u_i) > 0$, e_j is on a rising edge of $\varphi'_{bis}(t)$.
If $\varphi'_{bis}(u_{i+1}) - \varphi'_{bis}(u_i) < 0$, e_j is on a falling edge of $\varphi'_{bis}(t)$.
Step 3. For each e_j, if it is on a rising edge according to step 2, then $\varphi'_{bis}(e_{j-1} : e_j)$ is regarded as a floor. If it is on a falling edge, then $\varphi'_{bis}(e_j : e_{j+1})$ is regarded as a floor. The remaining parts of the $\varphi'_{bis}(t)$ are regarded as peaks.
Step 4. Interpolate between the peaks to obtain the resulting local cut-off frequency.

5. Time varying filtering based empirical mode decomposition

5.1. Stopping criterion

A criterion based on instantaneous bandwidth has been reported [30]. It takes into account Cohen bandwidth information and has a clear physical meaning. However, the bandwidth threshold is quite confusing because the bandwidth may contain oscillatory terms. The instantaneous bandwidth was firstly developed by Cohen [42,43]. For an analytic signal $A(t)e^{j\varphi(t)}$, the instantaneous frequency is defined as the derivative of its phase, i.e. $\varphi'(t)$ [44]. The Cohen instantaneous bandwidth is

$$B_{Cohen}(t) = \left| \frac{d}{dt} \log A(t) \right| = \left| \frac{A'(t)}{A(t)} \right| \tag{33}$$

This definition indicates that a signal is narrow-band if its instantaneous amplitude varies very slowly. However, this expression has its defects. Consider a simple two-component signal

$$x(t) = \cos[\varphi_1(t)] + \cos[\varphi_2(t)] \tag{34}$$

Obtaining its analytic signal and its instantaneous amplitude (see Eqs. (10) and (15)), then the Cohen instantaneous bandwidth is

$$B_1(t) = \left| \frac{A'(t)}{A(t)} \right| = \left| \frac{\left(\varphi'_1(t) - \varphi'_2(t)\right) \sin[\varphi_1(t) - \varphi_2(t)]}{2 + 2\cos[\varphi_1(t) - \varphi_2(t)]} \right| \tag{35}$$

For a two-tone signal whose frequency content is changeless, the frequency separation between its components is constant. Thus, its instantaneous bandwidth should be constant. However, the oscillatory terms in Eq. (35) will cause difficulty in interpreting its instantaneous bandwidth.

In order to eliminate the oscillatory terms, Loughlin et al. [45] derived an alternative expression for instantaneous bandwidth. According to Loughlin, instantaneous bandwidth is defined

Algorithm 3. Sifting process based on time varying filtering.

Step 1. Calculate the instantaneous amplitude $A(t)$ and instantaneous frequency $\varphi'(t)$ of $x(t)$ using Hilbert transform (see Eq. (10)).
Step 2. Locate the local minima and maxima of $A(t)$, denoted as $\{t_{\min}\}$ and $\{t_{\max}\}$ respectively.
Step 3. Interpolate the set of points $A(\{t_{\min}\})$ to obtain $\beta_1(t)$. In the same manner, interpolate $A(\{t_{\max}\})$ to obtain $\beta_2(t)$. Compute $a_1(t)$ and $a_2(t)$ via Eq. (27).
Step 4. Interpolate $\varphi'(\{t_{\min}\})A^2(\{t_{\min}\})$ and $\varphi'(\{t_{\max}\})A^2(\{t_{\max}\})$ to obtain $\eta_1(t)$ and $\eta_2(t)$. Compute $\varphi'_1(t)$ and $\varphi'_2(t)$ via Eq. (30).
Step 5. Calculate the local cut-off frequency $\varphi'_{bis}(t) = [\varphi'_1(t) + \varphi'_2(t)]/2$ according to Eq. (31).
Step 6. Realign $\varphi'_{bis}(t)$ according to algorithm 2 to deal with the intermittence problem.
Step 7. Compute $h(t)$ according to Eq. (32). Then apply B-spline approximation filter on $x(t)$, by taking the extrema timings of $h(t)$ as knots, i.e., $\{t_{\min}\}$ and $\{t_{\max}\}$ of $h(t)$. The approximate result is denoted as $m(t)$.
Step 8. Calculate the stopping criterion $\theta(t)$ via Eq. (39). If $\theta(t) \le \xi$, $x(t)$ is taken to be an IMF, else let $x(t) = x(t) - m(t)$ and repeat from step 1 to step 7.
Unless noted, we set the bandwidth threshold as $\xi = 0.1$ throughout this paper.

as the standard deviation in the weighted average instantaneous frequency (WAIF) at a given time. WAIF is considered to be a very useful expression. For a two-component signal given in Eq. (14), WAIF [46] is defined as a weighted average of the instantaneous frequencies of the individual components:

$$\varphi_{avg}(t) = \frac{a_1^2(t)\varphi'_1(t) + a_2^2(t)\varphi'_2(t)}{a_1^2(t) + a_2^2(t)} \tag{36}$$

Then the Loughlin instantaneous bandwidth for the two component signal is given by [45]:

$$B_{Loughlin}(t) = \sqrt{\frac{a'^2_1(t) + a'^2_2(t)}{a_1^2(t) + a_2^2(t)} + \frac{a_1^2(t)a_2^2(t)(\varphi'_1(t) - \varphi'_2(t))^2}{\left(a_1^2(t) + a_2^2(t)\right)^2}} \tag{37}$$

It can be seen from Eq. (37) that as $|\varphi'_1(t) - \varphi'_2(t)|$ increases, the separation between the two components increases, so does the Loughlin instantaneous bandwidth. As $|a'_1(t)|$ or $|a'_2(t)|$ increase, the amplitude modulation become more pronounced, and the instantaneous bandwidth again increases. By the use of Eq. (37), the Loughlin instantaneous bandwidth for the signal in Eq. (34) is

$$BW_2(t) = \left|\frac{\varphi'_1(t) - \varphi'_2(t)}{2}\right| \tag{38}$$

The oscillatory terms are eliminated. It offers a good interpretation of how this two components separate in frequency.

With the Loughlin's definition of instantaneous bandwidth, if $B_{Loughlin}(t)$ is small enough, then we call the signal local narrow-band. Note that $B_{Loughlin}(t)$ is an absolute value (in rad/s). In order to measure how much it deviates from the WAIF, we define a relative criteria rather than an absolute one

$$\theta(t) = \frac{B_{Loughlin}(t)}{\varphi_{avg}(t)} \tag{39}$$

When $\theta(t)$ gets smaller, the bandwidth of the signal becomes smaller. Given a threshold ξ, we call a signal local narrow-band if $\theta(t) \le \xi$.

5.2. Sifting process of TVF-EMD

We will now introduce the process for local narrow-band signal extraction, i.e., the sifting process of TVF-EMD, based on the concepts outlined in the previous sections. The sifting process of TVF-EMD is mainly carried out in three steps: (1) Estimate the bisecting frequency (local cut-off frequency). (2) Filter the input signal using TVF to obtain the local mean. (3) Check if the residual signal satisfies the stopping criterion. Given a real valued signal$x(t)$, the sifting process of our method is shown in Algorithm 3:

There are two parameters in the proposed method, i.e., the bandwidth threshold ξ and the B-spline order n. Bandwidth threshold ξ is detailed in Section 5.1 and is used to determine whether the input signal should be filtered or not. In Section 6, we will show that ξ determines the separation performance. On the other hand, n is a parameter that determines the roll-off of TVF (see Fig. 1(a)). Note that the choice of TVF only affects step 7. Thus the parameter n is entirely unrelated to the cut-off frequency estimation. As mentioned in Section 3, one can choose a TVF freely if the filter is able to be constructed using $\varphi'_{bis}(t)$. Once a different TVF used, n should be replaced accordingly. In this paper, $n = 28$ is adopted for gaining better filtering performance without losing much computing performance.

In step 3 and step 4, we assume that $A(t)$ have a succession of local minima and maxima. If $A(t)$ does not contain enough minima and maxima to continue interpolation, $A(t)$ must be slowly varying. In this case, the input signal is considered to be local narrow-band and should not be further decomposed.

5.3. Discussions

5.3.1. Convergence

An example is provided to study the convergence property of the proposed method. In the filtering procedure, the input signal is split into two components, i.e., the LHF and LLF components. For a multi-component signal, suppose that $f_{\min}(t)$ is the local minimum frequency and $f_{\max}(t)$ the local maximum frequency. The cut-

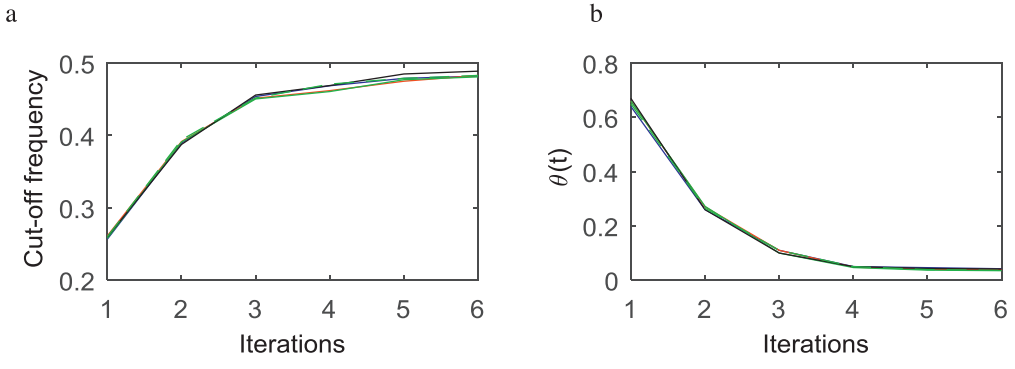


Fig. 4. Convergence tested on noise. (a) Cut-off frequency obtained for 5 realizations of white Gaussian noise via TVF-EMD (b) Stopping criterion $\theta(t)$ obtained for 5 realizations of white Gaussian noise via TVF-EMD.

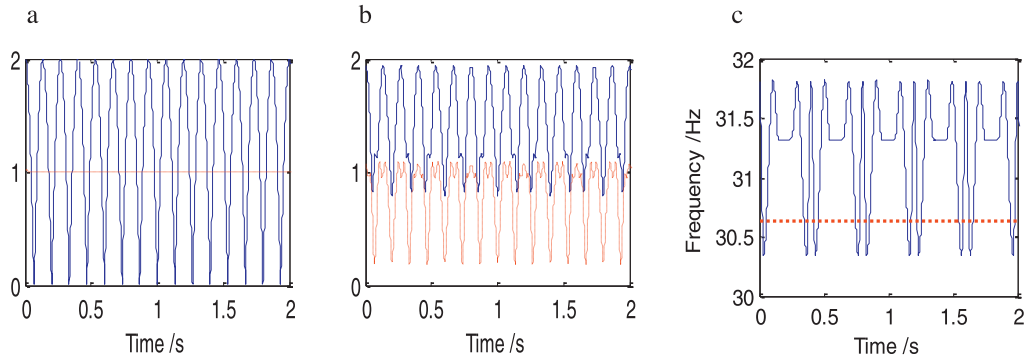


Fig. 5. Three-component signal decomposition with sampling frequency as 1000sps. (a) Instantaneous amplitudes of the 46.25 Hz (blue line) and 15 Hz (red line) components. (b) Instantaneous amplitudes of the two obtained components, $a_1(t)$ (blue line) and $a_2(t)$ (red line). (c) Theoretical (red line) and estimated (blue line) cut-off frequencies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

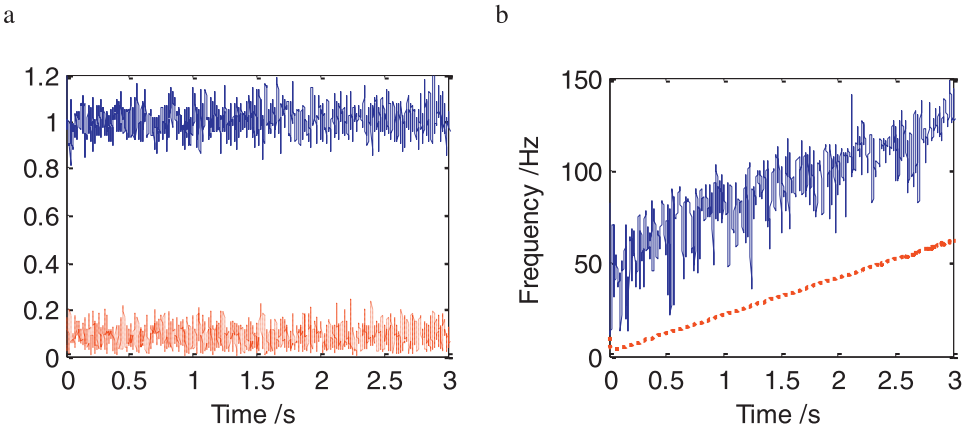


Fig. 6. Noisy chirp signal decomposition. (a) Instantaneous amplitudes of the two estimated components, $a_1(t)$ (above) and $a_2(t)$ (below). (b) Estimated (solid line) cut-off frequency and the instantaneous frequency of the noise free chirp signal (dotted line).

off frequency of the B-spline filter is expected to locate between $f_{\min}(t)$ and $f_{\max}(t)$. By successively applying the filter, the obtained higher frequency component is a local narrow-band signal. In order to validate the convergence property, TVF-EMD is applied on 5 realizations of white Gaussian noise. In Fig. 4 we show the trajectories of the averaged local cut-off frequency and $\theta(t)$ (see Eq. (39)). The results appear to have a good consistency and achieve a quick convergence.

5.3.2. Examples for local cut-off frequency estimation

In the beginning of Section 4.1, we have discussed four signals and the expected cut-off frequencies. Here we will test whether TVF-EMD meets our expectations. The cut-off frequency is without realignment.

(1) For the two-tone signal case, it can be easily validated that the obtained local cut-off frequency is exactly the bisecting frequency.
(2) For the three component signal in Eq. (8), Fig. 5 shows the decomposition result. Comparatively, it is obvious that the estimated instantaneous amplitudes shown in Fig. 5(b) matches closely with that shown in Fig. 5(a), although there are slight errors. Notice that the estimated instantaneous amplitudes in Fig. 5(b) are segmented because we force $a_1(t) \geq a_2(t)$. Fig. 5(c) shows the theoretical bisecting frequency and the estimated one. As we can see, due to the estimation error, the bisecting frequency obtained by TVF-EMD appears to oscillate. However, the variation in the estimated local bisecting frequency is so small that it is unlikely to cause significant variations in the filtering result.
(3) For white Gaussian noise, from Fig. 4(a), we can see that the average cut-off frequency is about 0.25.
(4) For a noisy chirp signal $x(t) = \cos(4\pi t + 20\pi t^2) + n(t)$, where $n(t)$ is the additive noise with a standard deviation of 0.1. As depicted in Fig. 6(b), for $t \in [0, 3]$, its instantaneous frequency varies linearly between 2 Hz and 62 Hz. The estimated instantaneous amplitudes $a_1(t)$ and $a_2(t)$ are shown in Fig. 6(a). One can see that $a_1(t)$ oscillates around its mean of 1 and $a_2(t)$ oscillates around its mean of 0.1. It indicates that TVF-EMD successfully extract the chirp's amplitude from noise. In Fig. 6(b), the local cut-off frequency varies in the same way as the chirp signal's instantaneous frequency. It is obvious that our method is able to adjust its cut-off frequency adaptively.

6. Simulation and comparison

In this section, we compare TVF-EMD with EMD, EEMD [18], VMD [22] and EWT [23]. EEMD is proved to be robust to noise, it consists of a large number of trials in which realizations of white noise are added. VMD is based on Wiener filtering and it is able to address both the separation problem and the noise problem. EWT is based on adaptive wavelet filter bank and it is robust to noise.

The key parameters of the above methods are as follows. For TVF-EMD, unless noted, we set the bandwidth threshold $\xi = 0.1$ and B-spline order $n = 28$. The stopping criterion for EMD is the 3-threshold criterion proposed in [24] with threshold1 = 0.05, threshold2 = 0.5 and tolerance = 0.05. For EEMD, the noise

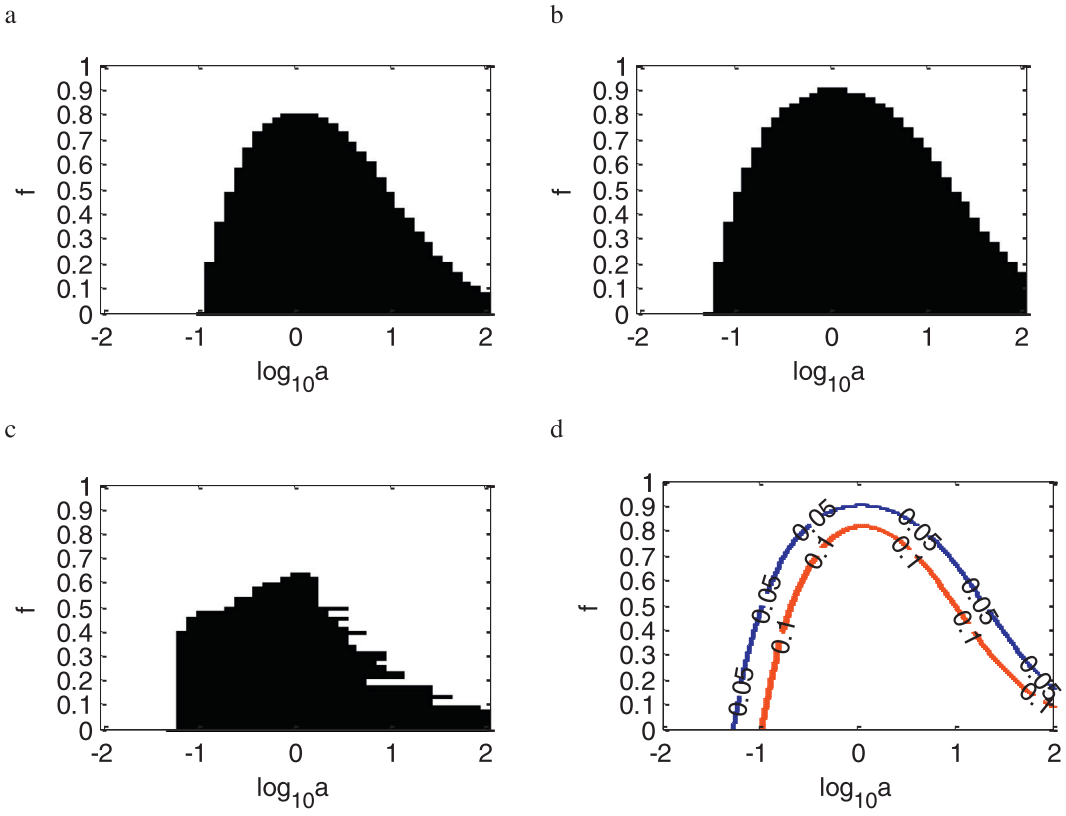


Fig. 7. Separation performance with respect to (a, f). (a) TVF-EMD with $\xi = 0.1$. (b) TVF-EMD with $\xi = 0.05$. (c) EMD. (d) Theoretical performance expectation for $\xi = 0.1$ and $\xi = 0.05$.

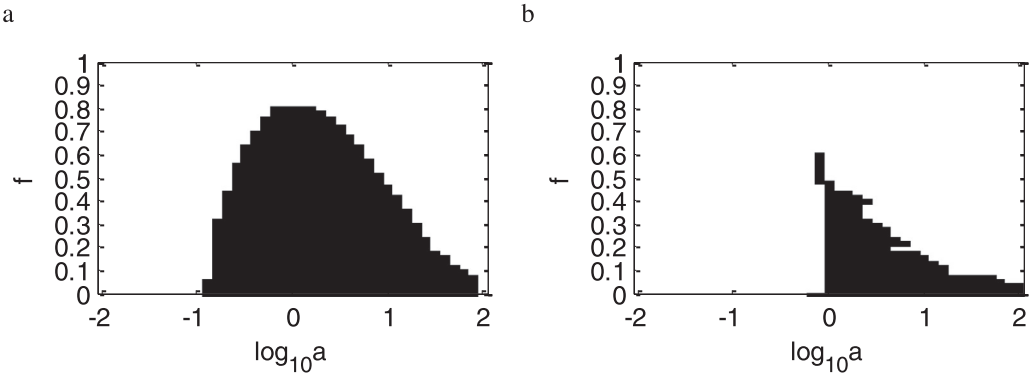


Fig. 8. Separation performance with respect to (a, f) under low sampling rate. (a) TVF-EMD with $\xi = 0.1$. (b) EMD.

standard deviation is selected to be 0.3 and the number of trials is 500. For VMD and EWT, a very important parameter is the number of mode N. Since N is still difficult to be chosen adaptively, we use different N for different signals.

6.1. Two-tone signal example

We evaluate the separation performance of the proposed method, by considering a two-tone signal.

$$x(t) = \cos(2\pi t) + a\cos(2\pi f t) \tag{40}$$

where a is the amplitude ratio and f is the frequency ratio. For $f \in [0, 1)$, the first term is the LHF component. The frequency ratio of the two components is set from 0 to 1 and the amplitude ratio is set from 0.01 to 100. A quantity used in [7] that measures the separation performance is given by

$$c = \frac{\|imf_1(t) - s_1(t)\|}{\|s_2(t)\|} \tag{41}$$

A value of c approaching zero indicates a perfect separation while a value close to 1 indicates that the two components are badly separated. A threshold of 0.5 is used to determine whether the decomposition is successful. Fig. 7 shows a comparison between EMD and our proposed method. The black areas ($c \leq 0.5$) indicate that the two components are correctly separated, while the white areas ($c > 0.5$) mean that they are not correctly separated. It can be seen that both EMD and TVF-EMD resolve the given signal only when the frequency ratio is below a particular cut-off. The cut-off of EMD is about 0.65, while TVF-EMD extends it to 0.8 (for $\xi = 0.1$) and 0.9 (for $\xi = 0.05$). The numerical results of TVF-EMD are consistent with the theoretical expectation (see

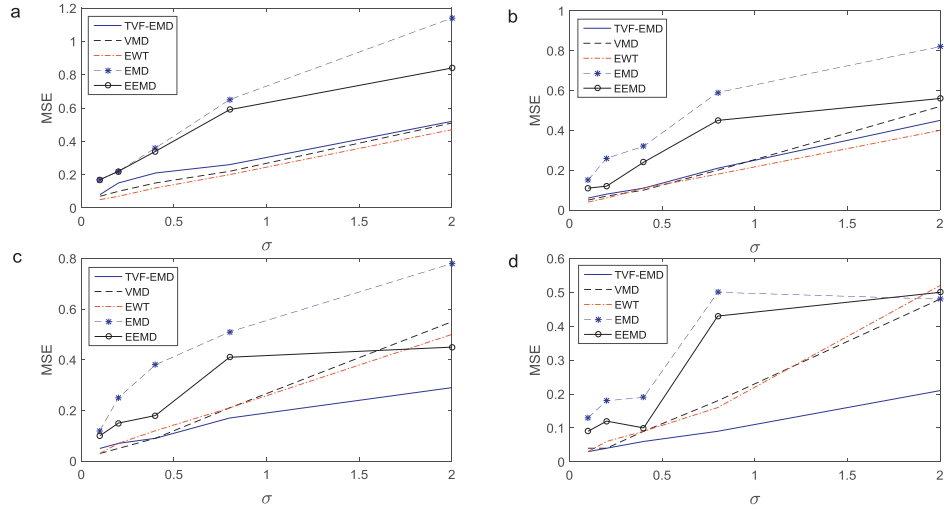


Fig. 9. Performance of noise resistance on a single tone signal. (a) Frequency = 350 Hz. (b) Frequency = 100 Hz. (c) Frequency = 50 Hz. (d) Frequency = 10 Hz.

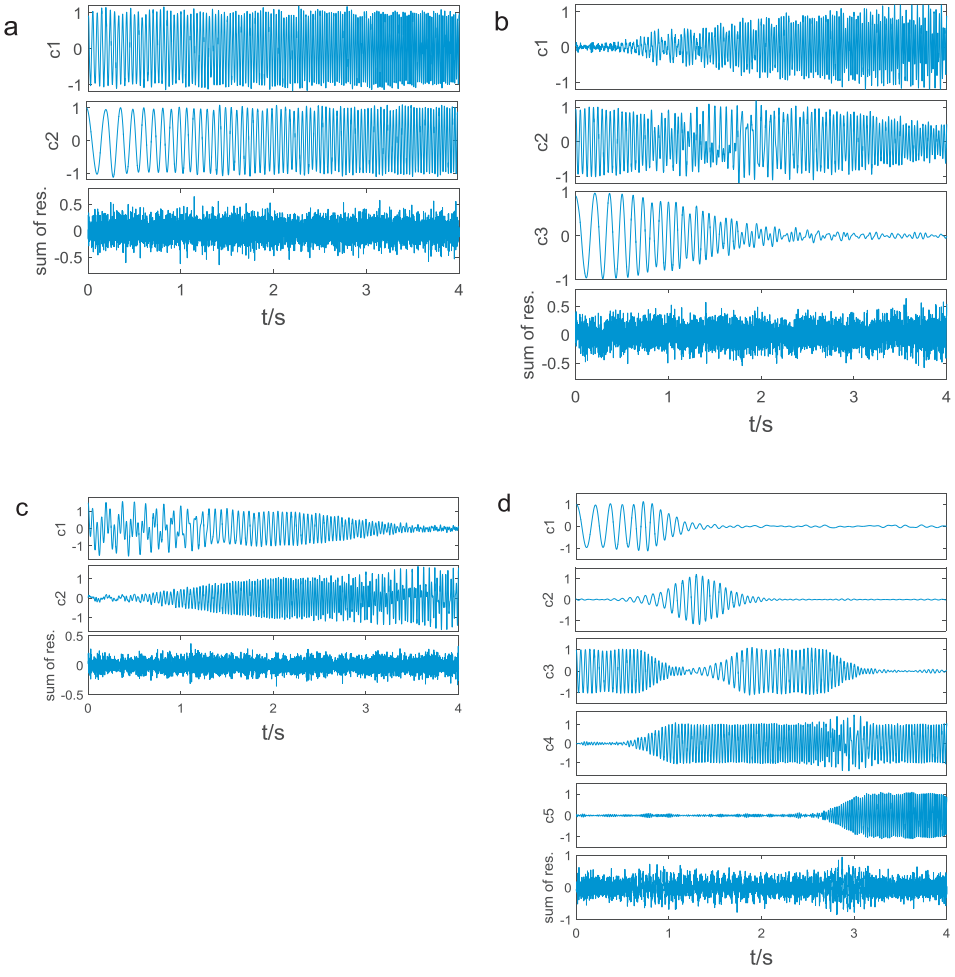


Fig. 10. The analysis of noisy chirp signal. (a) TVF-EMD. (b) EEMD. (c) VMD. (d) EWT.

Fig. 7(d)). In this example, a two-tone signal model is considered. What will happen if multi-component signals are analyzed? Recalling that TVF-EMD is an iteratively filtering method, the local bandwidth criterion will decrease and eventually reach the stopping criterion. Thus, for multi-component signals, we believe TVF-EMD can also meet the theoretical expectation.

EEMD, VMD and EWT are also used to evaluate their performance. The cut-off frequency ratio is about 0.65 for EEMD and EWT, and 0.95 for VMD. It seems that VMD yields the best performance. However, for VMD and EWT, it is sensible to ensure that the number of modes N is properly chosen. In this example, N is chosen to be 2 because the input is a two-tone signal. Once an im-

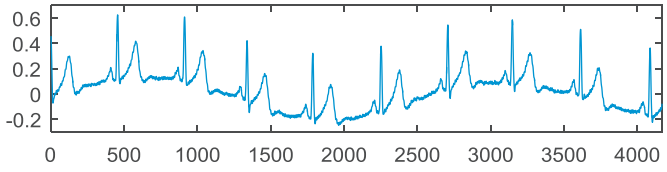


Fig. 11. Real ECG signal.

proper number is chosen, VMD and EWT may fail to provide the correct result.

6.2. Performance under low sampling rate

In Section 6.1, oversampling is adopted so that the given signal is considered to be continuous-time. In this example, we address the case where the sampling frequency is close to the Nyquist frequency. Clearly, for the two-tone model given in Eq. (40), the Nyquist frequency is $f_N = 2$. Here we process the simulation under a low sampling rate $f_s = 2.5$.

The separation performance of TVF-EMD and EMD are shown in Fig. 8(a) and (b) respectively. It can be seen that $f_s = 2.5$is quite a low sampling rate under which EMD cannot work efficiently. A sharp reduction of the black area is clearly observed, indicating that sampling rate has a large effect on the separation performance. By comparing Figs. 7(a) and 8(a), no noticeable change is found, which means that TVF-EMD can work effectively. As mentioned before, under low sampling rates the extrema cannot be located exactly by searching the time series. On the other hand, EMD requires the upper and lower envelopes of IMF to be symmetric, which is limited under low sampling rates. By contrast, TVF-EMD does not require an exact information of the local extrema. In our method, even a rough estimation of the cut-off frequency is acceptable for filtering. Besides, TVF-EMD require the modes to be local narrow-band, and does not require the upper and lower envelopes to be symmetric. For these reasons, TVF-EMD can guarantee a reliable performance under low sampling rates.

Since EEMD also requires the upper and lower envelopes to be symmetric, the separation performance under low sampling rate is exactly the same as EMD. For VMD and EWT, the filtering based methods are able to keep a high performance under low sampling rates. However, as mentioned, N must be properly chosen (2 in this example).

6.3. Performance of noise resistance

We evaluate the noise tolerance by decomposing a single tone signal plus an additive Gaussian noise (noise standard deviations:$\sigma = 0.1$, 0.2, 0.4, 0.8 and 2). The amplitude of the single tone is fixed to be 1 and the frequency is selected to be 10, 20, 50 and 350 Hz (the sampling frequency is 1000 Hz.). The decomposition is carried out under different frequencies and noise standard deviations. After that, mean square error (MSE) is calculated for the noise free signal and its corresponding mode.

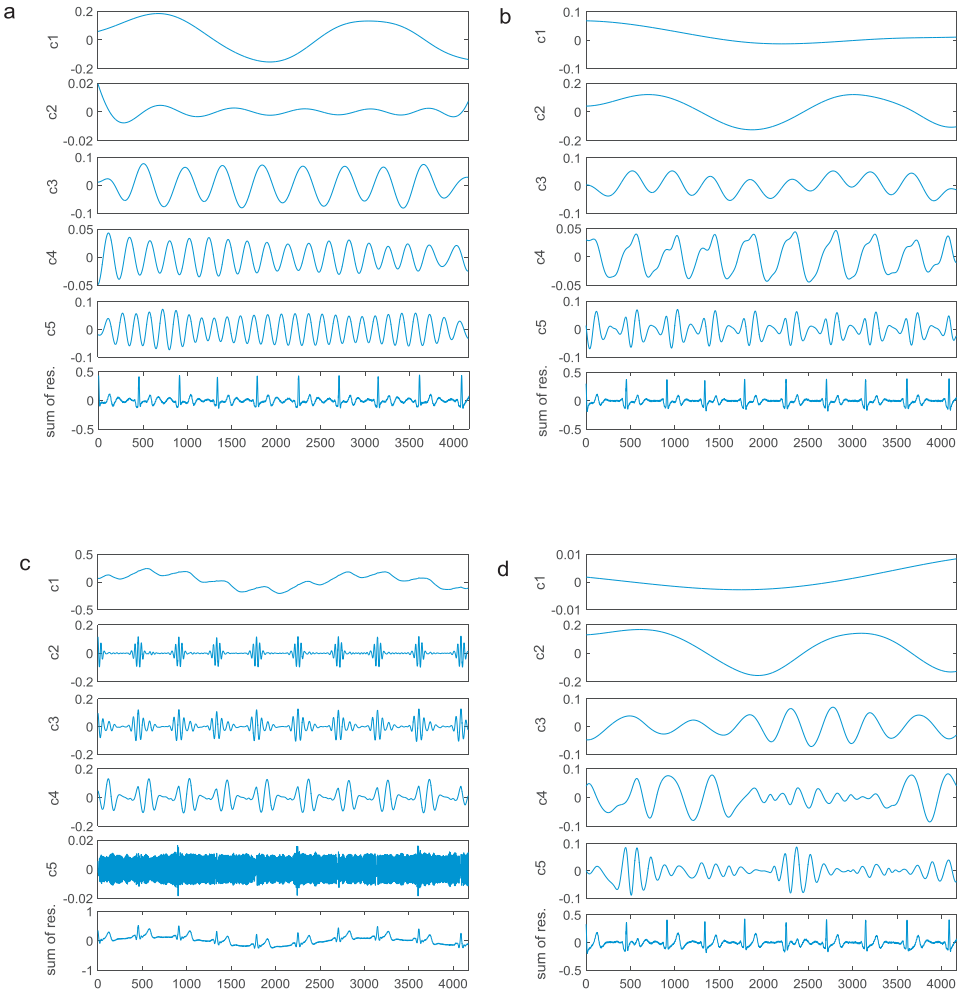


Fig. 12. The analysis of ECG signal. (a) TVF-EMD. (b) EEMD. (c) VMD. (d) EMD.

The resulting MSEs are shown in Fig. 9. We can observe that the performance depends on the noise level. For lower noise levels, TVF-EMD has a comparable performance with VMD and EWT. It also shows that all the methods achieve a better performance when the frequency gets lower. It indicates that an oversampling will help to improve the decomposition performance. In the presence of stronger noise especially in the case of $\sigma = 2$, TVF-EMD fails to extract the single tone signal when the frequency is high (see the case of 350 and 100 Hz). However, in the case of 50 Hz and 10 Hz, TVF-EMD works effectively. By contrast, VMD and EWT seem more sensitive to the noise level, and fail to pick up the tone when the signal is properly oversampled.

6.4. Non-stationary example

One of the potential applications of EMD is the analysis of the non–stationary signal, whose frequency content is time varying. We evaluate the proposed method in a noisy chirp signal ($\sigma = 0.2$).

$$x(t) = \cos(20\pi t + 4\pi t^2) + \cos(4\pi t + 4\pi t^2) + n(t) t \in [0, 4] \tag{42}$$

Since the results are a number of modes, in Fig. 10, we only show modes corresponding to the chirps and the sum of the remaining rest modes. For TVF-EMD, it is clear that the two dominant components, i.e., c1 and c2 are intact modes. Based on the TVF, a reasonably effective separation is achieved. For VMD ($N = 5$), distortions in the time domain are clearly visible (see$t \in [0, 1]$ and$t \in [3, 4]$). In addition, both c1 and c2 resulting from VMD do not meet the criteria of IMF. For EEMD and EWT ($N = 5$), the two chirps are split into several modes. If we take EEMD or EWT as filtering, the cut-off frequency is constant with respect to time, which is not suitable for non-stationary signals. It seems that the feature of local characteristic time scale decomposition is lost, which is a very important aspect of EMD. In contrast, the use of TVF not only improves the stability against noise, but also finely preserves the interesting properties of EMD.

6.5. Real ECG signal

The real signal illustrated in Fig. 11 is an electrocardiogram (ECG) signal which has been used in [23]. Since the resulting modes using EWT with $N = 5$ are presented in [23], we will not repeat the relative work here. The modes are obtained via TVF-EMD, EEMD, VMD ($N = 5$) and EMD.

As shown in Fig. 12, modes corresponding to the five lowest frequency components are given. The EMD modes c1 to c5 are difficult to interpret because of mode mixing. EEMD and VMD are able to extract the oscillating patterns that are visible in the ECG signal. However, there are modes containing “unwanted” extrema and requiring further decomposition (see c4 in EEMD and c1 in VMD). On the contrary, TVF-EMD successfully extract the oscillating modes without mode mixing. In fact, the result obtained using TVF-EMD is very similar to the EWT output, as shown in [23]. Thus, experiments on the real ECG signal seem give the advantage to TVF-EMD and EWT.

7. Conclusion

In this paper, the performance of EMD is improved based on the time varying filter approach. Our research is a step towards a clearer analytic method. The proposed method decomposes the input signal into a series of local narrow-band components through filtering. With the TVF, it is able to separate components that are closely spaced in the frequency domain. It is also found that the local narrow-band signal is a good alternative to IMF, especially under low sampling rates. Besides, the mode mixing problem can be alleviated by realigning the cut-off frequency. Simulations and comparisons are carried out to demonstrate the superiority of the proposed method. It shows that the use of TVF can significantly improve the performance of EMD, with regards to tone separation, stability under low sampling rates and noise robustness. In future studies, strict mathematical proofs for the TVF property of B-spline approximation should be carried out.

References

[1] B. Boashash, Estimating and interpreting the instantaneous frequency of a signal 0.1. fundamentals, Proc. IEEE 80 (4) (1992) 520–538.
[2] N.E. Huang, Z. Shen, S.R. Long, et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc. Lond. Series A 454 (1998) 903–995.
[3] C. Guo, M.A. Al-Shudeifat, J. Yan, et al., Application of empirical mode decomposition to a Jeffcott rotor with a breathing crack, J. Sound Vib. 332 (16) (2013) 3881–3892.
[4] J. Zheng, J. Cheng, Y. Yang, Generalized empirical mode decomposition and its applications to rolling element bearing fault diagnosis, Mech. Syst. Sig. Process. 40 (1) (2013) 136–153.
[5] R. Ghanati, M. Fallahsafari, M.K. Hafizi, Joint application of a statistical optimization process and empirical mode decomposition to magnetic resonance sounding noise cancelation, J. Appl. Geophys. 111 (2014) 110–120.
[6] X. Xie, Illumination preprocessing for face images based on empirical mode decomposition, Signal Process. 103 (2014) 250–257.
[7] G. Rilling, P. Flandrin, One or two frequencies? The empirical mode decomposition answers, IEEE Trans. Signal Process. 56 (1) (2008) 85–95.
[8] N.E. Huang, Z. Shen, S.R. Long, A new view of nonlinear water waves: the Hilbert spectrum, Ann. Rev. Fluid Mech. 31 (1) (2003) 417–457.
[9] Z. Xu, B. Huang, K. Li, An alternative envelope approach for empirical mode decomposition, Digit. Signal Process. 20 (1) (2010) 77–84.
[10] R. Deering, J.F. Kaiser, The use of a masking signal to improve empirical mode decomposition, ICASSP 2005, IEEE 4 (2005) 485–488.
[11] X. Guanlei, W. Xiaotong, X. Xiaogang, Time-varying frequency-shifting signal-assisted empirical mode decomposition method for AM–FM signals, Mech. Syst. Sig. Process. 23 (8) (2009) 2458–2469.
[12] Z. Li, X. Hao, P. Li, et al., A local normalized method to eliminate riding waves appearing in the empirical AM/FM decomposition, Digit. Signal Process. 25 (1) (2014) 104–113.
[13] X. Hu, S. Peng, W.L. Hwang, EMD Revisited, A new understanding of the envelope and resolving the mode-mixing problem in AM-FM Signals, IEEE Trans. Signal Process. 60 (3) (2012) 1075–1086.
[14] Y. Kopsinis, S. Mclaughlin, Investigation and performance enhancement of the empirical mode decomposition method based on a heuristic search optimization approach, IEEE Trans. Signal Process. 56 (1) (2008) 1–13.
[15] H. Hong, X. Wang, Z. Tao, Local integral mean-based sifting for empirical mode decomposition, IEEE Signal Process Lett. 16 (10) (2009) 841–844.
[16] T. Oberlin, S. Meignen, V. Perrier, in: An alternative formulation for the empirical mode decomposition, 60, 2012, pp. 2236–2246.
[17] M.A. Colominas, G. Schlotthauer, M.E. Torres, An unconstrained optimization approach to empirical mode decomposition, Digit. Signal Process. 40 (2015) 164–175.
[18] Z. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Analy. 1 (01) (2009) 1–41.
[19] N.U. Rehman, D.P. Mandic, Filter bank property of multivariate empirical mode decomposition, IEEE Trans. Signal Process. 59 (5) (2011) 2421–2426.
[20] N. Rehman, D.P. Mandic, Multivariate empirical mode decomposition, Proc. R. Soc. Lond. Series A 466 (2117) (2010) 1291–1302.
[21] P. Flandrin, G. Rilling, P. Goncalves, Empirical mode decomposition as a filter bank, IEEE Signal Process Lett. 11 (2) (2004) 112–114.
[22] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, IEEE Trans. Signal Process. 62 (3) (2014) 531–544.
[23] J. Gilles, Empirical wavelet transform, IEEE Trans. Signal Process. 61 (16) (2013) 3999–4010.
[24] G. Rilling, P. Flandrin, P. Goncalves, On empirical mode decomposition and its algorithms, in: IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, Grado, 2003, pp. 8–11.
[25] C. Damerval, S. Meignen, V. Perrier, A fast algorithm for bidimensional EMD, IEEE Signal Process Lett. 12 (10) (2005) 701–704.
[26] G. Rilling, P. Flandrin, Sampling effects on the empirical mode decomposition, Adv. Adapt. Data Analy. 1 (01) (2009) 43–59.
[27] Z. Xu, B. Huang, F. Zhang, Improvement of empirical mode decomposition under low sampling rate, Signal Process. 89 (11) (2009) 2296–2303.
[28] J. Mo, W. Hu, S. Le, Improving Empirical Mode Decomposition based on up--sampling, in: ICIS 2014, IEEE, 2014, pp. 425–428.
[29] N. Stevenson, M. Mesbah, B. Boashash, A sampling limit for the empirical mode decomposition, in: ISSPA 2005, IEEE, 2005, pp. 647–650.
[30] B. Xuan, Q. Xie, S. Peng, EMD sifting based on bandwidth, IEEE Signal Process Lett. 14 (8) (2007) 537–540.

[31] G. Wang, The most general time-varying filter bank and time-varying lapped transforms, IEEE Trans. Signal Process. 54 (10) (2006) 3775–3789.
[32] S.M. Phoong, P.P. Vaidyanathan, Time-varying filters and filter banks: some basic principles, IEEE Trans. Signal Process. 44 (12) (1997) 2971–2987.
[33] F. Hlawatsch, G. Matz, H. Kirchauer, et al., Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation, IEEE Trans. Signal Process. 48 (5) (2000) 1417–1432.
[34] Q. Chen, N. Huang, S. Riemenschneider, et al., A B-spline approach for empirical mode decompositions, Adv. Comput. Math. 24 (1-4) (2006) 171–195.
[35] Y. Yang, C. Miao, J. Deng, An analytical expression for empirical mode decomposition based on B-spline interpolation, Circuits Syst. Signal Process. 32 (6) (2013) 2899–2914.
[36] M. Unser, A. Aldroubi, M. Eden, B-spline signal processing. I. Theory, IEEE Trans. Signal Process. 41 (2) (1993) 821–833.
[37] M. Unser, Splines: a perfect fit for signal and image processing, IEEE Signal Process Mag. 16 (6) (1999) 22–38.
[38] M. Unser, A. Aldroubi, M. Eden, Polynomial spline signal approximations: filter design and asymptotic equivalence with Shannon's sampling theorem, IEEE Trans. Inf. Theory 38 (1) (1992) 95–103.
[39] Y. Yang, Empirical mode decomposition as a time-varying multirate signal processing system, Mech. Syst. Sig. Process. 76-77 (2016) 759–770.
[40] N. Senroy, S. Suryanarayanan, P.F. Ribeiro, An improved hilbert–huang method for analysis of time-varying waveforms in power quality, IEEE Trans. Power Syst. 22 (4) (2007) 1843–1850.
[41] M. Feldman, Time-varying vibration decomposition and analysis based on the Hilbert transform, J. Sound Vib., 295 (s 3–5) (2006) 518-530.
[42] L. Cohen, C. Lee, Instantaneous frequency, its standard deviation and multicomponent signals, Proc. SPIE Adv. Signal Process. II (1988) 186–208.
[43] L. Cohen, C. Lee, Instantaneous bandwidth for signals and spectrogram, in: ICASSP 1990, IEEE, 1990, pp. 2451–2454.
[44] L. Cohen, Time-Frequency Analysis, Prentice Hall, 1995.
[45] P.J. Loughlin, K.L. Davidson, Modified Cohen-Lee time-frequency distributions and instantaneous bandwidth of multicomponent signals, IEEE Trans. Signal Process. 49 (6) (2001) 1153–1165.
[46] G. Jones, B. Boashash, Instantaneous frequency, instantaneous bandwidth and the analysis of multicomponent signals, in: ICASSP 1990, IEEE, 1990, pp. 2467–2470.