IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 9, 2018, accepted December 29, 2018, date of publication January 11, 2019, date of current version February 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2892622

A Minimum Arclength Method for Removing Spikes in Empirical Mode Decomposition

HUI-WEN YANG[1,2], SHYH-KANG JENG[1], (Senior Member, IEEE), HSU-WEN V. YOUNG[3], CHEN LIN[2], YUNG-HUNG WANG[4], KUN HU[5,6], AND MEN-TZUNG LO[2]

[1]Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan
[2]Department of Biomedical Sciences and Engineering, National Central University, Taoyuan 32001, Taiwan
[3]Department of Applied Mathematics, Chung Yuan Christian University, Taoyuan 32023, Taiwan
[4]Research Center for Adaptive Data Analysis, National Central University, Taoyuan 32001, Taiwan
[5]Division of Sleep Medicine, Harvard Medical School, Boston, MA 02115, USA
[6]Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Boston, MA 02115, USA

Corresponding authors: Shyh-Kang Jeng (skjeng@ntu.edu.tw) and Men-Tzung Lo (mzlo@ncu.edu.tw)

This work was supported by the Ministry of Science and Technology in Taiwan under Grant 106-2221-E-008-032-MY2, Grant 103-2221-E-008-006-MY3, and Grant 106-2221-E-008-023.

ABSTRACT Empirical mode decomposition (EMD) is an extensively utilized tool in a time–frequency analysis. However, disturbances, such as impulse noise, can result in a mode-splitting effect, in which one physically meaningful component is split into two or more intrinsic mode functions (IMFs). In this paper, we propose a novel method, minimum arclength EMD (MA-EMD), to robustly decompose time series data with impulse-like noises. The idea is to apply a minimum arclength criterion to adjust the knot positions of impulses during the sifting process in EMD. In this way, the impulse-like artifact is extracted with the first IMF, and the mode splitting effect of the latter decomposition is alleviated. Furthermore, when the first IMF contains the desired information, we separate the spikes and the first IMF by adding a pair of masking signals. For using this masking-aided MA-EMD (MAMA-EMD) method, we also mathematically derived the appropriate ranges of the frequency and the amplitude of the masking signal. The MAMA-EMD is utilized to deal with the simulated Duffing wave and four real-world data, including electrical current, vibration signals, the cyclic alternating pattern in sleep EEG (electroencephalography), and circadian of core body temperature. The results show that the MA-EMD and MAMA-EMD have a sound improvement when encountering impulse noises.

INDEX TERMS Adaptive filters, empirical mode decomposition, impulse noise, masking-EMD, mode splitting; spike.

I. INTRODUCTION

The non-stationary signal can be encountered frequently in various fields in nature, such as speech processing, marine biologic sound analysis, and physiological rhythm [1]–[3]. One powerful and popular tool is empirical mode decomposition (EMD) which decomposes a time series into several intrinsic mode functions (IMFs). Due to the non-parametric nature of EMD, these IMFs often preserves non-linear and non-stationary properties which cannot be derived from conventional linear approaches such as Fourier or wavelet transform. Recent studies in its mathematical structure unmasked its outstanding ability in dealing with nonlinear and non-stationary signal [4], [5] and confirmed its computational efficiency [6].

Albeit powerful, one common problem in signal processing encumbering the decomposition in EMD is the spikes. Spikes are extremely high/low values in very short periods in time-domain but contain wide spectrums of frequency. Therefore, when decomposed by EMD, the energy of a single spike would be scattered in several IMFs with different frequencies. Moreover, the effect of spikes would propagate to nearby signal. Even though locality characteristic of EMD permits its effect to decay exponentially [4], the relatively strong magnitude of spike still results in perturbation of the IMF. Fig. 1 shows an example of Duffing wave with spikes decomposed by EMD. Compared to the IMFs from the same Duffing wave without spikes, the ∼0.1 Hz signal is split into IMF 1 and 2. This is called the mode-splitting effect.

Physically, a single spike may result from a single cause, such as the collective neuronal activity in EEG (Electroencephalography), or the electrical current surge caused by the switch. Transforming the single spike in frequency domain

2169-3536 © 2019 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

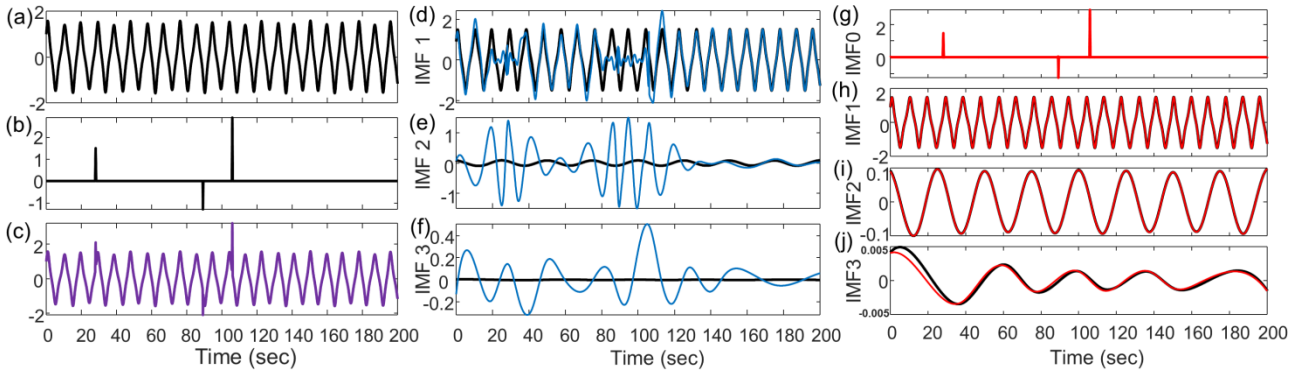

FIGURE 1. Surrogated nonstationary Duffing signal contaminated by triangular spikes, and its decomposition by EMD and our proposed MAMA-EMD, respectively. (a) Duffing wave. (b) The spike signal as a perturbation. (c) The spike-contaminated signal as the input for EMD and MAMA-EMD. (d-f) The blue lines are results of EMD on (c), and the black lines are from EMD on pure Duffing wave in (a) as the ground truth. (g-h) The red lines are MAMA-EMD on (c). The black lines in (h-j) are the same as the black lines in (d-f) but in different scales. Note that with MAMA-EMD method, the triangles are extracted, and the mode-splitting effect in (d-e), in which the 0.1Hz component in the first IMF of black line is split into IMF 1 and 2, is alleviated.

by Fourier or wavelet analysis results in many to infinite harmonic terms, loses its parsimonious nature, and contaminates frequency components of the signals. For single point spikes, median filter or its adaptive versions have been widely used [7]–[9]. For spikes in the shape of triangles where the median filter cannot be used, wavelet methods are popular. It would require choosing the appropriate wavelet functions similar to the spike shape, and decomposing the signal into different scales. The spikes can then be detected or removed in certain scales [10]–[12], and the new signal is reconstructed from the modified coefficients. Even though these methods can themselves be treated as a pre-process for EMD [13], this frequency domain approach is inefficient and ineffective, and may loses its nonlinear and nonstationary property.

In EMD, the decomposition is through a sifting process, in which the relative low-frequency fluctuations are subtracted iteratively. This low-frequency baseline is determined through averaging the upper and the lower envelops formed through cubic spline interpolation of extrema points. These extrema points determines the filtering property of baseline: the extrema intervals determine frequency response of the spline [4], [14], [15], and the amplitude of signal determines the magnitude of frequency response to its neighboring points [4]. Undoubtedly, changing the extrema points changes the baseline function and therefore the sifting results in EMD. Different studies have been proposed to adjust the extrema position by an offset Δt_i [16]–[18] or increasing extrema rate by adding assistant signals (e.g. masking EMD [19] or EEMD [20]) for different purposes. Nevertheless, to the authors' knowledge, no method had been proposed to remove effects of spike-like noises directly during the decomposition process.

In this paper, we aim to develop a method that can effectively extract spike morphology by adjusting the height of extrema on the spike during sifting of EMD. We propose a novel method, the minimum arclength EMD (MA-EMD) method, to replace the spike point by minimizing the arclength of the upper/lower envelop. Furthermore, the spike extraction in MA-EMD can also be aided with masking EMD, becoming the masking aided MA-EMD (MAMA-EMD), where a pair of assistant signals is added in the decomposition. The proposed MAMA-EMD for spike extraction can be treated as a noise filter to improve time-frequency analysis or the later decomposition in EMD; the extracted spikes themselves can also be a feature for physical/physiological explanation.

The remaining part of this paper is organized as follows. Section II is a brief review of EMD, masking EMD and spike detection algorithms. Then, the proposed minimum arclength EMD (MA-EMD) and masking aided MA-EMD (MAMA-EMD) methods are presented in Section III. The validation of our method and its capability to maintain the nonlinearity in EMD are verified by two numerical experiments in Section IV. In Section V, we apply the MAMA-EMD on realistic experimental data to demonstrate its effectiveness in extracting spike signal. The discussion and conclusion are in the last section.

II. METHOD REVIEWS

A. THE EMD ALGORITHM

EMD is an adaptive data-driven algorithm aiming to decompose a signal into several intrinsic mode functions (IMFs) successively through a repeated sifting process. The sifting process can be regarded as a high-pass filter that iteratively removes low-frequency trends, which is determined by the mean curve estimated by averaging the upper and lower envelops. Given a signal $x(t)$ ($t > 0$), we define k as the IMF index and p the sifting step. Then, the notation $x^{k,p}$ represents the k_{th} proto-IMF at p_{th} sifting step. The EMD algorithm is given in **Algorithm 1**.

B. THE MASKING EMD

To solve the problem of mode mixing, Deering and Kaiser [19] proposed to insert a single tone sinusoid $w(t) = a_m sin(2\pi f_m t)$ during the decomposition. This sinusoid,

Algorithm 1 EMD Algorithm

(1) Define $x^{0,0}x(t)$. Starting with $k = 1$ and $p = 0$. Set $x^{1,0}(t) = x(t)$.
(2) Identify all the maximum $\{(t_a, x_a)\}$ and minimum $\{(t_b, x_b)\}$ of $x^{k,p}(t)$.
(3) Connect maximum (respectively minimum) points with natural cubic spline to derive upper (lower) envelope $U(t)$ (and $L(t)$, respectively).
(4) Obtain the local mean of the upper and lower envelopes $m(t) = (U(t) + L(t))/2$.
(5) Subtract local mean from the temporal signal $x^{k,p+1}(t) = x^{k,p}(t) - m(t)$.
(6) Repeat (2)-(5) n_{sp} times, i.e. $p = 0, \cdots, n_{sp} - 1$. Derive $x^{k,n_{sp}}(t)$.
(7) Assign the k_{th} IMF as $c_k(t) = x^{k,n_{sp}}(t)$.
(8) Calculate residual $x^{k+1,0} = x^{k,0}(t) - c_k(t)$.
(9) Increment k and repeat steps (2)-(8) to generate series of IMFs and a residue until that the residue contains no more than one extrema

whose frequency is relatively higher, is a "masking signal" and serves as an assisted disturbance to avoid extraction of low frequency components during sifting. The algorithm is summarized in **Algorithm 2**.

Algorithm 2 Masking EMD Algorithm

(1) For step k in EMD, generate masking signal $w(t) = a_m \sin(2\pi f_m t)$.
(2) Perform Steps (2)-(5) in **Algorithm 1** on $x^+(t) = x^{k,p}(t) + w(t)$. In other words, substitute $x(t)$ by $x^+(t)$) to obtain IMF c_k^+. Similarly, perform steps (2)-(5) in **Algorithm 1** on $x^-(t) = x^{k,p}(t) - w(t)$ and obtain c_k^-.
(3) The resultant IMF is defined as $c_k = (c_k^+ + c_k^-)/2$.

C. THE SPIKE DETECTION

Mathematically, a signal with occasional artifacts can be modeled as

$$x(t) = \hat{x}(t) + v(t) \tag{1}$$

where $\hat{x}(t)$ is the signal of interest, $v(t)$ models the noise term and $x(t)$ is the observed signal. We suggest that the noise term contains two components

$$v(t) = \omega(t) + z(t) * i(t) \tag{2}$$

where $\omega(t)$ represents the white Gaussian process and $i(t)$ is the random process generating impulsive artifact which is convolved with a spike-like function $z(t)$. Here, $z(t)$ can be of different shapes, such as a single-point spike or a triangular spike.

Many types of impulse rejection filters have been designed for different types of signals, such as wavelet for speech signals [12] and Raman spectra [21], and median filters for images [7], [8]. We adopt one of the simplest designs, the median filter, as a tool for spike detection. Similar to other impulse detection algorithms, our spike detector is based on the prior assumption that the signal should be smooth. Therefore, the extrema that differ too much from nearby extrema is regarded as spike points. The maximum and minimum are dealt with separately. For each maximum $\{(t_{a_i}, x_{a_i})\}$, we first find the set containing D nearby maximum values (D is an even value) in a window centered about x_{a_i}.

$$W_i^D = \{x_{a_d} \mid i - D/2 \le d \le i + D/2\} \tag{3}$$

The median and standard deviation of this set are

$$m_i^D = \text{Med}\left\{x_{a_d} \mid x_{a_d} \in W_i^D\right\} \tag{4}$$

and

$$S_i^D = \text{std}\left\{x_{a_d} \mid x_{a_d} \in W_i^D\right\} \tag{5}$$

respectively. Then, the extrema values larger than the median by more than T standard deviation are classified as maximum impulses, i.e. the set of maximum impulses, and is defined as

$$G^M = \left\{(t_{a_i}, x_{a_i}) \mid x_{a_i} > m_i^D + S_i^D \cdot T\right\} \tag{6}$$

Similarly, the set of minimum impulses is obtained as

$$G^N = \left\{(t_{b_j}, x_{b_j}) \mid x_{b_j} < m_j^D - S_j^D \cdot T\right\}, \tag{7}$$

where $m_i^D(m_j^D)$ and $S_i^D(S_j^D)$ are the median and the standard deviation of the minimum values within window $W_i^D(W_j^D)$ centered at $x_{b_i}(x_{b_j})$. Finally, we have the subsets G^M and G^N for the subsequent analysis.

III. MINIMUM ARCLENTH METHOD

A. THE MINIMUM ARCLENGTH

Our proposed method aims to find an optimal replacement of the extrema value while maintaining the position of it as a knot. In this way, the magnitude of impulse response of the spike point to the rest of the point is decreased. On the aspect of time domain, the estimated upper/lower envelop would become smoother, and so does the mean envelop, $m(t)$. Thus, after sifting, the morphology of the spikes is left in the first IMF with the high frequency term. Here, we propose to minimize the arclength of the resultant upper or lower envelop. For brevity, we take the upper envelop as an example to illustrate our method; the lower envelop is processed similarly. An illustration of the minimum-arclength method is shown in Fig. 2. Given a spike point $(t_{a_s}, x_{a_s}) \in G^M$, we find a minimizer $(t_{a_s}, \hat{x}_{a_s})$ such that

$$\hat{x}_{a_s} = \arg\min_y \left\{F\left(U\left(t|y_a = x_a, a \ne a_s; y_{a_s} = y\right)\right)\right\} \tag{8}$$

where

$$(U(t \mid \cdot)) = \int_0^{t_N} \sqrt{1 + \left(\frac{dU(t \mid \cdot)}{dt}\right)^2}\, dt \tag{9}$$

is the arclength of $U(t|\cdot)$. Then, the modified spline is created by the new series of maxima with impulse point replaced

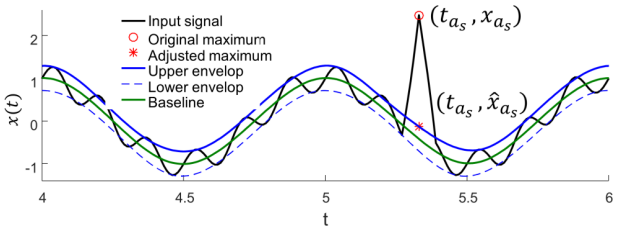

FIGURE 2. An illustration of our proposed method. The input signal **x(t)** (black) is first processed to detect spike points $(\mathbf{t_{a_s}}, \mathbf{x_{a_s}})$ (o). Then, the height of this point is adjusted to become a new knot $(\mathbf{t_{a_s}}, \hat{\mathbf{x}}_{\mathbf{a_s}})$ (*) that minimizes the arc length of the upper envelop. The new upper envelope **U (t)** (blue) is calculated by the new knot and the other original extrema. This creates a baseline **m (t)** (green) that is smooth and depicts the relatively low frequency of the signal.

by $(t_{a_s}, \hat{x}_{a_s})$. Here, the arclength is calculated over all the maxima points. However, this can be time consuming when the signal is long. In our experience, minimizing the arclength over 10 maxima points near the spike points is sufficient.

For all spike points in G^M, each point is processed consecutively. For the case of multipoint spike, the spike detection algorithm will mark two or more consecutive points on one single spike. (Here, we regard the whole structure as one single spike.) Minimizing the arclength point by point may not achieve the optimal result. The solution to this problem can be studied in future works.

The modified algorithm of EMD is shown as **Algorithm 3**. Comparing it to **Algorithm 1**, we added our protocol of detecting spike points and replacing them only in the first IMF (i.e. when $k = 1$). The rest of the IMFs are processed in the same way as EMD.

B. THE MASKING-AIDED MINIMUM ARCLENGTH EMD

The above method extracts spikes in the first IMF (IMF 1). However, if IMF 1 contains information of interest, the spikes should be removed to obtain a clear signal. The idea is to insert a high frequency masking signal, the single sine tone $w(t)$, to the original signal to prevent lower frequency components from being included in this IMF. Then, perform sifting algorithm on $x^+(t) = x(t) + w(t)$, the resultant mode c^+ contains only spikes and the single sine wave. We repeat this algorithm on $x^-(t) = x(t) - w(t)$ and derive c^-. When averaging c^+ and c^-, the added masking signals were compensated. Thus, we have the new "first IMF" (IMF 0) that contains only spikes and some very-high-frequency components (most of time noises), leaving the IMF 1 free of spikes.

The rest of this section discusses how to find the appropriate amplitude a_M and frequency f_M for the masking signal. A numerical simulation in Section 4.1 is performed to validate the effect of different a_M and f_M on separating a spike signal from a single tone sinusoid.

The inserted sinusoid should create no extrema on the spike, and allow EMD to separate itself from the original signal. According to [14], given two sinusoid $S_L(t) = a_L \cos(2\pi f_L t)$ and $S_M(t) = a_M \cos(2\pi f_M t)$, the necessary

Algorithm 3 Spike Extraction by EMD With Minimum Arclength Method (MA-EMD) Algorithm

(1) Define $x^{0,0}x(t)$. Starting with $k = 1$ and $p = 0$. Set $x^{1,0}(t) = x(t)$
(2) Identify all the maximum $\{(t_a, x_a)\}$ and minimum $\{(t_b, x_b)\}$ of $x^{k,0}(t)$
(3) If $k = 1$, find the subset of maximum impulses $G^M = \{(t_{a_i}, x_{a_i})\}$, $i = 1 \ldots n_{spx}$, and also find the subset of minimum impulses $G^N = \{(t_{a_j}, x_{a_j})\}$, $j = 1 \ldots n_{spn}$, where n_{spx} and n_{spn} are the total number of points regarded as maximum impulses and minimum impulses, respectively.
(4) Start with $i = 1$, find minimizer $(t_{a_i}, \hat{x}_{a_i})$ to minimize the arclength of upper envelope $U(t|x_a)$.
(5) Replace (t_{a_i}, x_{a_i}) with $(t_{a_i}, \hat{x}_{a_i})$, and form the new set of maximum points $(t_a, \hat{x}_a)$.
(6) Repeat (4) and (5) for $i = 1 \ldots n_{spx}$.
(7) For each $j = 1 \ldots n_{spn}$, find minimizers $(t_{b_j}, \hat{x}_{b_j})$ to minimize the arclength of lower envelope $L(t|x_b)$, consecutively. Then, replace (t_{b_j}, x_{b_j}) with $(t_{b_j}, \hat{x}_{b_j})$. Form the new set of minimum points $\{(t_b, \hat{x}_b)\}$
(8) Derive the new upper and lower envelope $U(t|\hat{x}_a)$ and $L(t|\hat{x}_b)$, respectively, according to the new set of maxima $\{(t_a, \hat{x}_a)\}$ and minima $\{(t_b, \hat{x}_b)\}$.

Perform step (4)-(5) in **Algorithm 1**, which is to derive the local mean $m(t)$ and subtract it from the present signal to form the temporal signal

$$x^{1,p+1}(t) = x^{1,p}(t) - m(t)$$

(9) Repeat (3)-(8) for $p = 0 \cdots n_{sp} - 1$, and derive the first IMF, $c_1(t)$.
(10) For $k > 1$, the steps are the same as steps (2)-(9) in **Algorithm 1** of EMD.

conditions to separate these two signals are

$$a_r f_r < 1 \quad \text{and} \, f_r < \frac{2}{3} \tag{10}$$

where $a_r = a_L/a_M$ and $f_r = f_L/f_M$.

Similarly, there should be no extrema on the morphology of spike. For simplicity, we assume that the spike is a triangular-shaped signal ascending within the time range $[t_a, t_p]$ at a slope s_t. In other words, the signal is $T(t) = s(t - t_a)$, when $t \in [t_a, t_p]$. Finding the extrema point is equivalent to solving the equation

$$\frac{d}{dt}[T(t) + S_M] = 0 \tag{11}$$

In other words,

$$\frac{dS_M}{dt} + \frac{dT}{dt} = 2\pi a_M f_M \cos(2\pi f_M t) + s = 0 \tag{12}$$

No extrema points means that the above equation has no solutions. Namely,

$$\cos(2\pi f_L t) = \frac{-s}{2\pi a_M f_M} \tag{13}$$

is not solvable. This leads to

$$\frac{s}{2\pi a_M f_M} > 1 \tag{14}$$

Thus,

$$2\pi a_M f_M < \mathrm{s}. \tag{15}$$

The derivation above is based on an ideal situation where the signal to be separated is a pure sinusoid. In practice, we aim to use this method to separate spikes from the first IMF derived from EMD. Thus, the a_L and f_L can be the peak-power frequency and power of the first IMF. Therefore, to remove the effect of spike, we first detect the spikes and then find a proper masking frequency by analyzing the slope of the spike and the frequency of the first IMF from EMD to meet both equations (10) and (15).

The algorithm of our proposed masking-aided minimum arclength EMD (MAMA-EMD) is in **Algorithm 4.**

Algorithm 4 Spike Extraction by Masking Aided Minimum Arclength EMD

(1) Perform steps (1) – (3) in **Algorithm 3** to detect spike points when $k = 1$.
(2) Perform EMD to derive the first IMF. Analyze its peak-power frequency f_L and power a_L.
(3) Analyze the slope of each spike, choose the smallest, call it s_t
(4) Find the proper frequency a_M and amplitude f_M, such that they meet both (10) and (15).
(5) Generate masking signal $w\,(\mathrm{t}) = a_M \sin(2\pi f_M t)$
(6) Perform steps (4)-(5) in **Algorithm 3** on $x^+(t) = x(t) + w(t)$ to obtain IMF c_k^+, and similarly on $x^-(t) = x(t) - w(t)$ and obtain c_k^-
(7) The resultant IMF is defined as $c_k = (c_k^+ + c_k^-)/2$.
(8) For $k > 1$, the steps are the same as **Algorithm 3**.

IV. SIMULATION VERIFICATION

A. SINGLE SINUSOID

To validate our derivation of appropriate masking signals, we test the effect of different amplitude and frequency of masking signal in MAMA-EMD in extracting a spike from a single tone sinusoid. Without loss of generality, we set the frequency of the single-tone sinusoidal signal to be 1, since the filtering property of EMD is only related with the ratio of two frequencies, f_r, and amplitudes, a_r, of the pure sinusoidal signal and the added masking signal. The simulated signal and its components are $S(t) = S_L(t) + S_p(t)$, where

$$S_L(t) = \cos(2\pi t) \tag{16}$$

and

$$S_p(t) = \begin{cases} 200t - 380 & if\ 2 < t < 2.05 \\ -200t + 420 & if\ 2.05 < t < 2.1 \\ 0 & otherwise \end{cases} \tag{17}$$

Then, a masking signal $w\,(\mathrm{t}) = a_M \sin(2\pi f_M t)$ is added to assist separating $S_L(t)$ and $S_p(t)$. Next, we vary the masking amplitude a_M from 0.01 to 100, and frequency f_M from 0.95 to 20. The sampling frequency is 100Hz. Fig. 3 shows the time-domain waveforms of $S(t)$ and its two components. The result of separation is evaluated by the mean squared error (MSE) between the extracted IMF2 (sinusoidal) and $S_L(t)$.

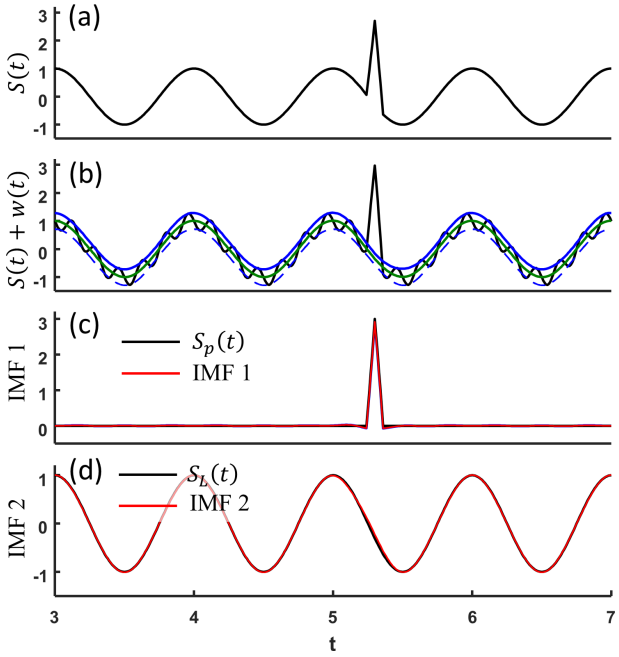

FIGURE 3. The simulated signal and its decomposed IMFs. (a) The signal S(t) is a combination of the triangular spike $S_p(t)$ (c, black line) and pure sinusoid $S_L(t)$ (d, black line). (b) A demonstration of summation of the signal and added masking sig (black solid line). The upper envelope (blue line) is connected by the adjusted extrema. The baseline (green line) is the average of the upper and lower (dashed blue line) envelop. (c) The triangular spike $S_p(t)$ (black line) and the first IMF from MAMA-EMD (red line). (d) The pure sinusoid $S_L(t)$ (black line) and the second IMF from MAMA-EMD (red line).

Fig. 4 demonstrates the MSE of IMF2 versus f_M and a_M. As expected, the proper frequency and amplitude of masking signals is bounded by equations (10) and (15). At the left hand side of the curve of $2\pi a_M f_M = s$, the spike is not separable from the $w(t)$. Meanwhile, at the right hand side of the curve, $a_r f_r = 1$ and $f_r < 2/3$, the masking signal is not separable from $S_L(t)$.

B. DUFFING WAVE

We use nonstationary Duffing wave with artificially added spikes to demonstrate that the proposed method extracts, maintains the non-linear and non-stationary properties of EMD and suppresses the mode-splitting effect. Duffing wave can be understood by the motion of a pendulum with non-linear stiffness. The Duffing equation has the form

$$\ddot{x}(t) = x - x^3 + 0.1\cos(\frac{2\pi}{25}) \tag{18}$$

with initial conditions $x(0) = \dot{x}(0) = 0$. Here, x is the displacement, and $\ddot{x}$ is the acceleration. We add three spike

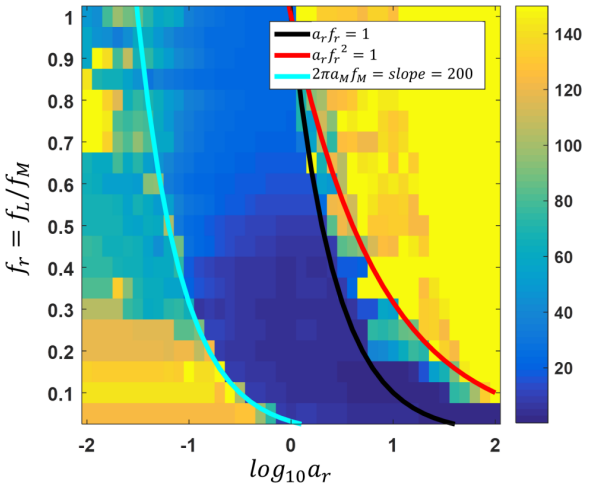

FIGURE 4. The MSE of IMF2 extracted by MAMA-EMD for different f_M and a_M. The ground truth is a pure sinusoidal signal.

signals with slope 5, -8 and 10, and height 1.5, -1.3 and 2.9, respectively, and get spike-contaminated signals. The original Duffing wave, the spike signal as an artifact, and the spike-contaminated signal is shown in Fig. 1(a-c).

The IMFs from EMD by decomposing the Duffing wave serve as the ground truth of the decomposition. Three IMFs were derived from the decomposition (Fig. 1). The first IMF corresponds to the intrinsic frequency around 0.1 Hz which shows strong intra-wave frequency modulation structure; IMF 2 corresponds to a uniform intermediate frequency component representing the forcing function. The sub-harmonic term is the evidence for the non-linearity of the system. Moreover, its amplitude is very small, which means any error will destroy the waveform of the sub-harmonic motion; IMF 3 represents a very low-intensity sub-harmonics.

The decomposition results by EMD on the perturbed signal is shown in Fig. 1(d-f). Under the influence of spikes, the IMFs are disturbed, resulting in a mode-splitting effect, where the original 0.1 Hz signal is distributed in both IMF 1 and 2. Hilbert spectrum also shows the frequency shift from 0.1 Hz to three times larger in IMF 1 (Fig. 5a). The 0.1Hz during this period (80-100 sec) is split into IMF 2.

Then, we decompose the perturbed signal by the proposed method. Since the input signal itself is smooth and without noise, we use MAMA-EMD where a high frequency sine wave is added in the first step to increase extrema points, so that the baseline, i.e. the signals except spike, can be depicted during sifting. The masking signal here has a frequency $f_M = 200$ and an amplitude $a_M = 3.5$. The result of our decomposition is depicted in Fig. 1 (g-j). In comparison to EMD, our MAMA-EMD method extracted the spike signal in the first mode, which is labeled as IMF 0 to avoid confusing with the original IMF 1. Afterwards, the later IMFs can be successfully recovered from the decomposition. Compared to the ground truth, our method only differs from the ground truth around both edges. The MAMA-EMD derived IMFs clearly depict the intra-wave frequency modulation in Hilbert spectrum (Fig. 5b.).

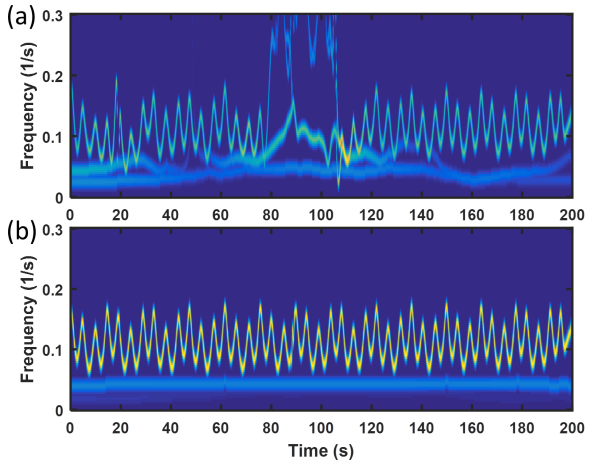

FIGURE 5. Hilbert spectrum showing the frequency overlapping caused by spikes. (a) IMF 2-4 by EMD. (b) IMF 1-3 by MAMA-EMD.

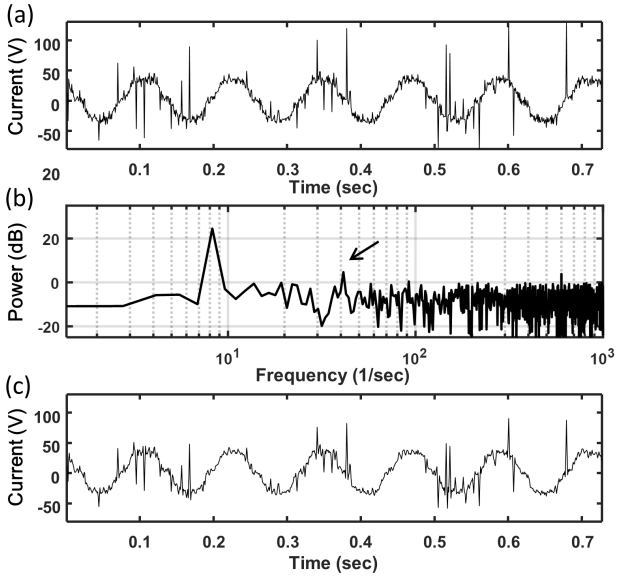

FIGURE 6. Current data. (a) The time domain signal. (b) Frequency spectrum. The arrow indicates the harmonic at 40Hz. The spikes cannot be removed by a 15 Hz low-pass filter (c).

V. APPLICATIONS

A. ELECTRICAL CURRENT

Electrical current surge is a common problem in automatic control system. Conventional method of processing this signal is to use a linear low-pass filter, which does not remove the spikes effectively. Here, we demonstrate the performance of our proposed MAMA-EMD in solving this problem comparing to a Fourier-based low pass filter and traditional EMD.

The data were phase currents measured from three-phase AC servomotor (YBL-9D, Ye Li Electric & Machinery Co., LTD) at 300rpm. The current values were transferred to voltage values by the current sensors. After filtered by RC low pass filter, the analogue voltage was converted to digital data by the microcontroller (STM32F103). A 3.3V 12-bit 1KS/s analog-to-digital converter was used.

The time domain signal and the Fourier spectrum are shown in Fig. 6(a,b). The spikes on the signals were cause

by MOSFET switch on the three-phase inverter of the motor drive, and the frequency of switching was 10 KHz. The instantaneously switching would cause the current surges, resulting in the spikes on the signals. Furthermore, the microcontroller and the peripheral circuits would also generate high frequency noise. The phase current cycle represents the rotation of the engine. The harmonic (∼40 Hz) in the Fourier spectrum showed that the signal is not pure sinusoidal.

First, the signal is filtered by a low-pass FIR filter with pass band equals to 15Hz. Fig. 6(c) shows the results of filtering. Since a spike has a very wide band in Fourier spectrum, the designed filter can only decrease the height of spikes.

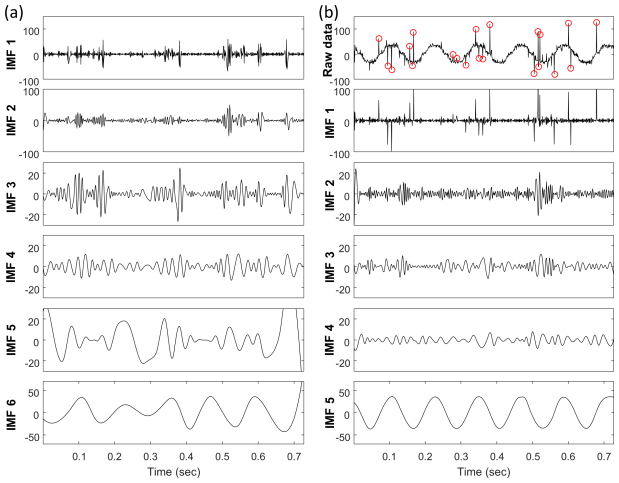

FIGURE 7. Intrinsic mode functions (IMFs) derived from EMD (a) and MAMA-EMD (b). In (b), the first panel showed the original signal (black line) and the points detected as spikes (red circles). Panel 2-6 demonstrate the decomposition by our propose method. The first IMF contains the spikes and some high-frequency noise. The last panel shows IMF 5 that depicts the low frequency wave.

The decomposition result derived by EMD is shown in Fig. 7(a), where only the first 6 IMFs are given. The spikes even in the non-spike region disturb EMD, resulting in the mode splitting problem during the 0.2-0.4 second.

Then, the data is processed by our proposed MAMA-EMD. The spikes were first detected by the median filter described in Section 2.2 with window size 20 and threshold 1.5. The extrema detected as spikes are shown in Fig. 7(b, panel 1). Then, the first mode is derived from adding a masking signal with frequency $f_M = 200$ and amplitude $a_M = 3.5$. It can be observed that, our MAMA-EMD extracted the spikes and some high-frequency noise in the first IMF Fig. 7b, panel 2). The rest of the signal contains no spikes can be further processed by traditional linear filters or EMD. Here, we demonstrate the IMF 1-5 of decomposition results by EMD. In summary, it shows that our algorithm is capable of removing the spikes, and improves the decomposition of the rest modes.

B. ROTOR TEST RIG

Vibration signals from rolling element bearing is adopted to verify our proposed method on realistic data. The faulty mechanical components often result in impulses-like vibration signals. These spikes, although including useful information, may cause mode mixing effects when decomposed by EMD. In fact, moderate de-noising before decomposition or filtering has been proposed to improve the fault detection algorithm [10], [13]. In the example demonstrated in this section, we show that our method can extract spikes in the first IMF and improve the accuracy of later analysis.

The experimental data are provided by Center on Intelligent Maintenance Systems (IMS), University of Cincinnati [22]. In this run-to-failure test, four Rexnord ZA-2115 double roll bearings were installed on one shaft. Each bearing was equipped with two PCB 353B33 High Sensitivity Quarts ICP®Accelerometers (x and y axis). Vibration data was collected for 1 second every 20 minutes for 164 hours with a sampling rate of 20 kHz, and the length of each data is 20480 points. The rotation speed is kept constant at 2000 rpm (rotation frequency $f_r = 33.3$ Hz), and a radial load of 6000 lb. was applied onto the shaft and bearings by a spring mechanism. At the end of the test-to-failure experiment, an inner race defect occurred on bearing 3. The inner race fault frequency f_i is 296.9 Hz.

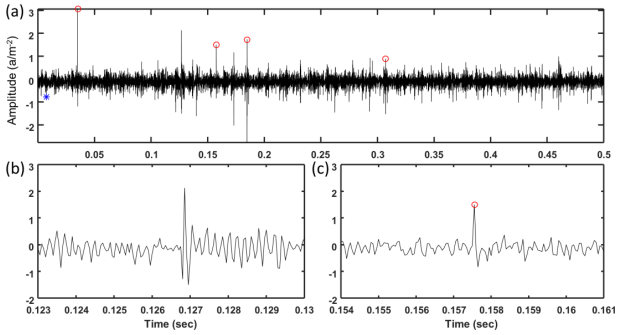

FIGURE 8. A vibration signal from bearing 3 with inner race defect. Maxima spikes are denoted as red circles; minima spikes are blue stars. (a) The raw data. (b) and (c) are partial zoom-in of the signal.

Fig. 8(a) shows the time-domain waveform of the vibration signal. Note that our spike detection is defined on differences to extreme values of nearby extrema, not the absolute value of the spike point, and thus some of those seemingly large values are not detected as spike if its nearby extrema is also large (Fig. 8b and c). This provides an advantage to maintain the resonance excited by the impact of default.

The decomposition results derived by EMD and MAMA-EMD are shown in Fig. 9, where only the first 5 IMFs are presented. We performed envelop spectrum on IMF2 and 3. The envelop spectrum is the Fourier transform of the envelope of the signal. From the envelope spectrum derived from IMFs of both methods (Fig. 10), the inner race fault frequency ($f_i = 296.9$) and its modulation with rotation frequency (296.9 ± 33) can be found from IMF 2. However, our proposed method has a lower noise level and clearer peaks in both IMF 3, and clear peaks on the rotation-related frequencies $2f_r$ and $4f_r$.

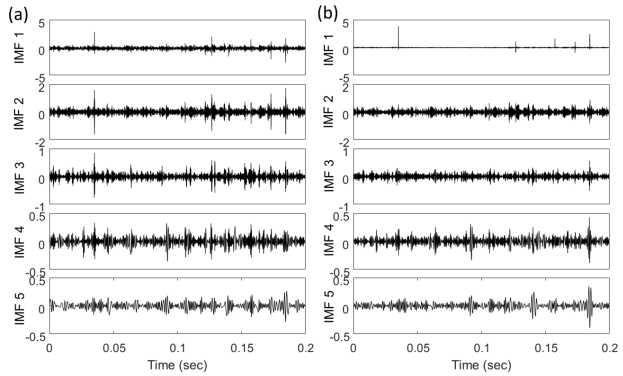

FIGURE 9. Intrinsic mode functions (IMFs) derived from EMD (a) and MAMA-EMD (b).

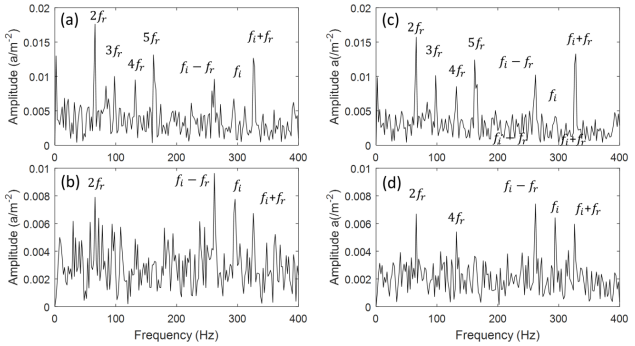

FIGURE 10. Envelope spectra of components derived from EMD (a-b) and MAMA-EMD (c-d). (a) and (c) are components from IMF 2; (b) and (d) are components from IMF 3.

C. CYCLIC ALTERNATION PATTERN (CAP) SUBTYPE CLASSIFICATION IN SLEEP EEG

Neuronal signal often presents different shapes of spikes. In this section, we showed that the extracted spikes can be used as a feature preserving its physiological significance. The cyclic alternating pattern (CAP) is a periodic EEG (Electroencephalography) activity, which is characterized by sequences of transient electrocortical events that are distinct from background EEG activity. The CAP may signify sleep instability, sleep disturbance, or both [23]. CAP is composed of transitions between Phase A, identified by high-voltage slow waves, and the low-voltage irregular activity of at least 2 seconds (Phase B) (Fig. 11). Phase A activity can be classified into three subtypes based on the reciprocal proportion of high-voltage slow waves (EEG synchrony) and low-amplitude fast rhythms (EEG desynchrony) throughout the entire phase A duration. According to the standard, the proportion of EEG desynchrony occupies $<20\%$, 20-50% and $>50\%$ of the entire phase A duration in subtype A1, A2 and A3, respectively ([24]. Subtype A1 marks the brain's attempt to preserve sleep; subtypes A2 and A3 often coincide with a frank EEG arousal. Specifically, 85% of subtypes A3 and 62% of subtypes A2 meet the AASM (American Academy of Sleep Medicine) criteria for arousals.

Here, we regard the high-voltage slow waves as spikes. By separating them from the background EEG, we can define the relative proportion of time between EEG synchrony and

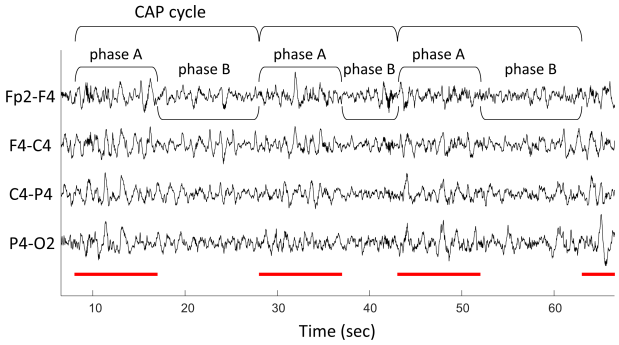

FIGURE 11. An example of CAP cycles in 4 EEG channels (Fp2-F4, F4-C4, C4-P4 and P4-O2). A CAP cycle is defined as a sequence of 2 alternating EEG patterns called phase A (indicated by red line) and phase B. Phase A is composed of high-amplitude EEG bursts which stand out from the background rhythm (phase B) in all the EEG channels.

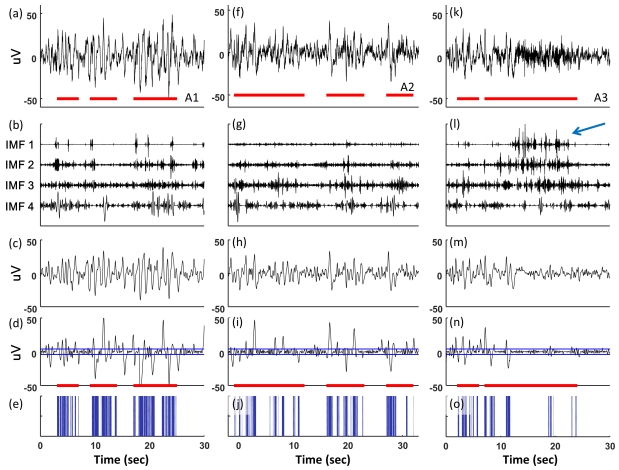

FIGURE 12. An example of CAP cycles. A CAP cycle is defined as a sequence of 2 alternating EEG patterns called phase A (indicated by red line) and phase B. Phase A is composed of high-amplitude EEG bursts which stand out from the background rhythm (phase B). Decomposition of the three different phase A subtypes, including subtype A1 (a-e), A2 (f-j) and A3 (k-o). The red horizontal lines indicate occurrences of A phase. The original signals (a,f,k) are first decomposed by masking EMD to remove the first 4 IMFs (b,g,i). The residual signals (c,h,m), derived by subtracting IMF1-4 from the original signal, are then processed by MAMA-EMD to extract the high-amplitude spikes (d,i,n). We then set up a threshold (±4uV, blue lines in d,i,n) to identify whether the extracted spikes are above/under this threshold. The proportion of time that the spikes are above/under this threshold (indicated by blue areas in e,j,o) is distinguishable among different phase A subtypes. The blue arrow in IMF 1 of (l) indicates EEG arousal, which is often observed in subtype A3.

desynchrony and distinguish different phase A subtypes. The EEG is first processed by masking EMD to remove the first 4 IMFs which contain high frequency information (Fig. 12). The residual signal is then processed by MAMA-EMD to extract spikes. This extracted signal, called IMF 4_{sp}, which contains spikes is used to calculate proportion of EEG synchrony, defined as the proportion of time IMF 4_{sp} is above or below a threshold (±4uV). We test our algorithm on the CAP Sleep Database [24], [25] (https://physionet.org/pn6/capslpdb/#ref02). This database includes polysomnography recordings, and the sleep microstructure is labeled by a team of trained neurologists. We use EEG recordings of the bipolar

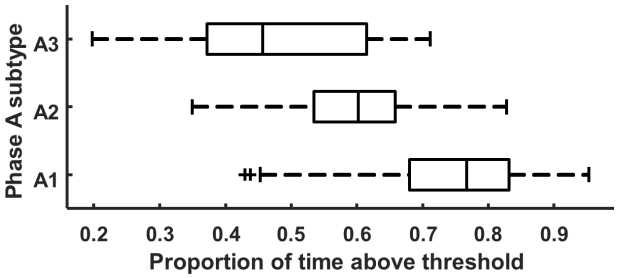

FIGURE 13. Proportion of time that the extracted spikes in a phase A exceed the threshold.

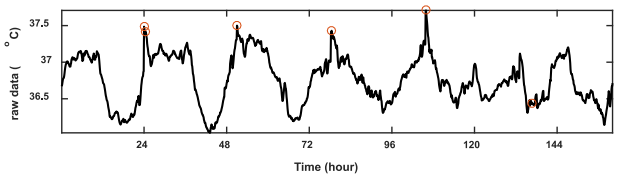

FIGURE 14. A 7-day recording of core body temperature. Red circles indicate extrema detected as spikes.

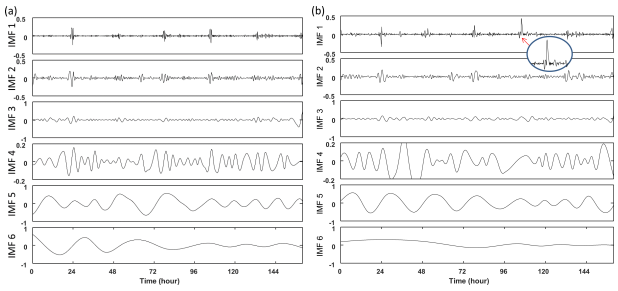

FIGURE 15. Recoding of core body temperature decomposed by EMD (a) and MAMA-EMD (b). The red arrow indicates the spike morphology that have been extracted by our method.

derivation C3-A2 from one of the normal subjects (n9) as an example. A total of 317 phase-A segments were analyzed.

Fig. 13 summarizes the result of our thresholding method on the three subtypes. The proportions of time with EEG synchrony are significantly different among the three subtypes ($P < 0.05$).

Fig. 12 demonstrates the results of our decomposition and thresholding method on different phase A subtypes. As expected, the proportion of time that the extracted spike is above or under the threshold is the largest in subtype A1, and the least in subtype A3. In addition, the EEG arousal, which is characterized by an abrupt frequency shift to 16Hz or higher, can be identified in the first IMF (Fig. 12(l)).

D. CORE BODY TEMPERATURE

In this section, we apply our MAMA-EMD on the long-term core body temperature to demonstrate the capability to extract the irregular shapes of spikes, and improve the later decomposition of EMD. The data is from rectal temperature recorded every 6 minutes for 7 days. The circadian rhythm can be observed (Fig. 14). The triangular spikes in each 24-hr cycle is from showering, where the body temperature increases 0.2-0.3°C (20-30% of daily circadian changes in normal core body temperature) between two consecutive samples, and decreases in around 15 minutes. We aim to extract the daily rhythm in this data. The spikes, although naturally produced by subject's behavior, interfere decomposition of IMFs (Fig. 15a). The circadian rhythm with cycle length 24 hours is split into IMFs 5 and 6, which shows the mode-splitting effect. When applying our MAMA-EMD with masking frequency $= 0.8$ and amplitude $= 0.05$, the morphology of spikes is extracted in the first IMF (Fig. 15 b), and alleviate the mode-splitting problem.

VI. CONCLUSION

A novel method has been proposed to sparsely decompose the spike noise in time-domain by a modified EMD. Aided with a masking signal with proper frequency and amplitude, we isolate the spikes in the first IMF, and improve the performance of decomposition the later IMFs. We also provided a mathematical induction and numerical experiment to find the proper amplitude and frequency of masking signal. Four real-world examples, the electrical current of an engine, vibration signals from rolling element bearing, cyclic alternating pattern in sleep EEG, and circadian of core body temperature are dealt with to demonstrate the wide application of our method. Served as a preprocedure for denoising in EMD or a tool for spike extraction, this method can also be applied to a wide variety of signals, where non-stationary information is to be extracted from spike contaminated data.

DATA ACCESSIBILITY

The source code of our MA-EMD and MAMA-EMD can be downloaded from the website of "Laboratory of Integrated Biomedical Signal Applications" in National Central University (http://in.ncu.edu.tw/~mzlo/drLo.html).

ACKNOWLEDGMENT

The authors would like to thank Dr. Hong-Tsu Young and his student, Ms. Ling-Jia Wu, for their generous advices on rolling-bearing signal processing. They would like to thank Dr. Kuo-Kai Shyu and his student, Guan-Xie Du, for their electrical current data. They would also like to thank Dr. Frank A. J. L. Scheer for providing core body temperature data.

REFERENCES

[1] L. Deng, "A generalized hidden Markov model with state-conditioned trend functions of time for the speech signal," *Signal Process.*, vol. 27, no. 1, pp. 65–78, 1992.
[2] O. Adam, "Advantages of the Hilbert Huang transform for marine mammals signals analysis," *J. Acoust. Soc. Amer.*, vol. 120, no. 5, pp. 2965–2973, 2006.
[3] B. Pittman-Polletta, W.-H. Hsieh, S. Kaur, M.-T. Lo, and K. Hu, "Detecting phase-amplitude coupling with high frequency resolution using adaptive decompositions," *J. Neurosci. Methods*, vol. 226, pp. 15–32, Apr. 2014.
[4] Y.-H. Wang, H.-W. V. Young, and M.-T. Lo, "The inner structure of empirical mode decomposition," *Phys. A, Statist. Mech. Appl.*, vol. 462, no. 300, pp. 1003–1017, 2016.
[5] E. H. S. Diop, R. Alexandre, and A. O. Boudraa, "Analysis of intrinsic mode functions: A PDE approach," *IEEE Signal Process. Lett.*, vol. 17, no. 4, pp. 398–401, Apr. 2010.

[6] Y.-H. Wang, C.-H. Yeh, H.-W. V. Young, K. Hu, and M.-T. Lo, "On the computational complexity of the empirical mode decomposition algorithm," *Phys. A, Statist. Mech. Appl.*, vol. 400, pp. 159–167, Apr. 2014.
[7] H. Hwang and R. A. Haddad, "Adaptive median filters: New algorithms and results," *IEEE Trans. Image Process.*, vol. 4, no. 4, pp. 499–502, Apr. 1995.
[8] S. R. Kim and A. Efron, "Adaptive robust impulse noise filtering," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 1855–1866, Aug. 1995.
[9] G.-X. Liu and L.-F. Shi, "Adaptive algorithm of magnetic heading detection," *Meas. Sci. Technol.*, vol. 28, no. 11, p. 115101, 2017.
[10] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *J. Sound Vib.*, vol. 289, nos. 4–5, pp. 1066–1090, 2006.
[11] Z. Nenadic and J. W. Burdick, "Spike detection using the continuous wavelet transform," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 1, pp. 74–87, Jan. 2005.
[12] R. C. Nongpiur, "Impulse noise removal in speech using wavelets," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, no. 3, Mar./Apr. 2008, pp. 1593–1596.
[13] J.-H. Ahn, D.-H. Kwak, and H.-B. Koh, "Fault detection of a roller-bearing system through the EMD of a wavelet denoised signal," *Sensors*, vol. 14, no. 8, pp. 15022–15038, 2014.
[14] G. Rilling and P. Flandrin, "One or two frequencies? The empirical mode decomposition answers," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 85–95, Jan. 2008.
[15] M. Unser, "Splines: A perfect fit for signal processing," in *Proc. Eur. Signal Process. Conf.*, 2000, pp. 27–28.
[16] X. Hu, S. Peng, and 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.*, vol. 60, no. 3, pp. 1075–1086, Mar. 2012.
[17] Y. Kopsinis and S. McLaughlin, "Investigation and performance enhancement of the empirical mode decomposition method based on a heuristic search optimization approach," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 1–13, Jan. 2008.
[18] J. Zheng, H. Pan, T. Liu, and Q. Liu, "Extreme-point weighted mode decomposition," *Signal Process.*, vol. 142, pp. 366–374, Jan. 2018.
[19] R. Deering and J. F. Kaiser, "The use of a masking signal to improve empirical mode decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2005, pp. iv/485–iv/488.
[20] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, 2008.
[21] F. Ehrentreich and L. Sümmchen, "Spike removal and denoising of Raman spectra by wavelet transform methods," *Anal. Chem.*, vol. 73, no. 17, pp. 4364–4373, 2001.
[22] J. Lee, H. Qiu, G. Yu, and J. Lin, "IMS bearing data," Bearing Data Set, IMS, Univ. Cincinnati, NASA Ames Prognostics Data Repository, Rexnord Tech. Services, Tech. Rep., 2007.
[23] L. Parrino, R. Ferri, O. Bruni, and M. G. Terzano, "Cyclic alternating pattern (CAP): The marker of sleep instability," *Sleep Med. Rev.*, vol. 16, no. 1, pp. 27–45, Feb. 2012.
[24] M. G. Terzano *et al.*, "Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep," *Sleep Med.*, vol. 2, no. 6, pp. 537–553, Nov. 2001.
[25] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

HUI-WEN YANG received the M.S. degree in ecology and evolutionary biology from National Taiwan University, Taipei, Taiwan, in 2011, where she is currently pursuing the Ph.D. degree with the Graduate Institute of Communication Engineering. She is also a Research Assistant with the Department of Biomedical Sciences, National Central University.

SHYH-KANG JENG (M'86–SM'98) received the B.S.E.E. and Ph.D. degrees from National Taiwan University, Taipei, Taiwan, in 1979 and 1983, respectively, where he is currently a Professor. In 1981, he joined the Department of Electrical Engineering, National Taiwan University, as a Faculty Member. From 1985 to 1993, he was a Visiting Research Associate Professor and a Visiting Research Professor with the University of Illinois at Urbana–Champaign. In 1999, he visited the Center for Computer Research in Music and Acoustics, Stanford University, Stanford, CA, USA, for six months. His research interests include time-domain electromagnetic field computation techniques, antenna design, multimedia signal processing, computational neuroscience, and computational cognitive neuroscience. He received the 1998 Outstanding Research Award from the National Science Council and the 2004 Outstanding Teaching Award.

HSU-WEN V. YOUNG received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, and the Ph.D. degree in mathematics from the University of Michigan, Ann Arbor, MI, USA. He completed a Postdoctoral training in mathematics at National Taiwan University and in adaptive and nonlinear data analysis at National Central University, Taiwan. He is currently an Assistant Professor of applied mathematics with Chung Yuan Christian University, Taoyuan, Taiwan. His current research interests include data analysis, mathematical problems in data science and signal processing, and mathematical modeling.

CHEN LIN received the M.A.Sc. degree in physiology from National Yang Ming University, Taiwan, and the Ph.D. degree in systems biology and bioinformatics from National Central University, Taiwan, where he is currently an Assistant Professor with the Department of Biomedical Sciences and Engineering. He completed one-year training in sleep medicine at Stanford Sleep Center, CA, USA. Due to his multidisciplinary training backgrounds, his research interests include applying different nonlinear signal analysis techniques to explore the alterations of underlying mechanisms in different physiological and pathological statuses and to develop indexes for clinically relevant outcomes in various cardiovascular diseases and sleep disorder.

YUNG-HUNG WANG received the Ph.D. degree in aeronautics and astronautics from Stanford University, Stanford, CA, USA, in 1998. He is currently an Associate Research Fellow with the Research Center for Adaptive Data Analysis, National Central University, Taiwan. His research interests include signal processing, biomedical engineering, electronic design automation, and fluid mechanics.

KUN HU received the Ph.D. degree in statistical physics from Boston University, Boston, MA, USA, in 2005. He completed the multidisciplinary Postdoctoral training in human physiology, neuroscience, and system biology at the Harvard Medical School, Boston.

He is currently an Associate Professor of medicine with the Harvard Medical School, and a Physiologist and the Director of the Medical Biodynamics Program, Brigham and Women's Hospital, Boston. His research interests include behavioral, cardiovascular, and circadian/sleep regulations in health and disease.

Dr. Hu is a member of the American Academy of Sleep Medicine, the Sleep Research Society, the Society for Neuroscience, and the International Cerebral Autoregulation Research Network Committee.

MEN-TZUNG LO received the Ph.D. degree in communication engineering from National Taiwan University, with a focus on biomedical signal and image processing and biomedical imaging and drug delivery systems.

He completed a Postdoctoral training at the Taipei Veteran General Hospital and the Syncope and Falls in the Elderly Laboratory, Beth Israel Deaconess Medical Center, Harvard Medical School, for applying novel nonlinear signal analysis to multiple biomedical signals from different disease groups to explore their underlying properties and quantifying the properties altered from a normal as parameters for severity or prognosis of diseases. He is currently a Professor with the Department of Biomedical Sciences and the Institute of Translational and Interdisciplinary Medicine, Engineering National Central University, Taiwan. His current research interests include time-varying interactions between multiple biological signals of human subjects and the changes of nonlinear properties in different physiological and pathological statuses.