Advances in Adaptive Data Analysis
Vol. 2, No. 2 (2010) 135–156

DOI: 10.1142/S1793536910000422

COMPLEMENTARY ENSEMBLE EMPIRICAL MODE DECOMPOSITION: A NOVEL NOISE ENHANCED DATA ANALYSIS METHOD

JIA-RONG YEH and JIANN-SHING SHIEH*
Department of Mechanical Engineering,
Yuan Ze University, 135 Yuan-Tung Road,
Chung-Li, Taoyuan, 320, Taiwan
jsshieh@saturn.yzu.edu.tw

NORDEN E. HUANG
Research Center for Adaptive Data Analysis,
National Central University, Chungli, Taiwan

The phenomenon of mode-mixing caused by intermittence signals is an annoying problem in Empirical Mode Decomposition (EMD) method. The noise assisted method of Ensemble EMD (EEMD) has not only effectively resolved this problem but also generated a new one, which tolerates the residue noise in the signal reconstruction. Of course, the relative magnitude of the residue noise could be reduced with large enough ensemble, it would be too time consuming to implement. An improved algorithm of noise enhanced data analysis method is suggested in this paper. In this approach, the residue of added white noises can be extracted from the mixtures of data and white noises via pairs of complementary ensemble IMFs with positive and negative added white noises. Though this new approach yields IMF with the similar RMS noise as EEMD, it effectively eliminated residue noise in the IMFs. Numerical experiments were conducted to demonstrate the new approach and also illustrate the problems of mode splitting and translation.

Keywords: Ensemble empirical mode decomposition (EEMD); intermittence; noise enhanced method; complementary ensemble empirical mode decomposition (CEEMD).

1. Introduction

The Empirical Mode Decomposition (EMD) [Huang *et al.* (1998)] is an adaptive time–frequency data analysis method designed for nonlinear and nonstationary signal analysis. When applied to a uniformly distributed white noise signal, EMD was proved to be a dyadic filter bank [Flandrin *et al.* (2004); Wu and Huang (2004)]. As it is, EMD has been used for a wide variety of applications covering engineering, sciences and even financial data. There is, however, a serious drawback in the EMD as originally proposed: the mode mixing problem, where widely disparate

*Corresponding author.

scales could appear in one Instrinsic Mode Function (IMF) component. Or, from a different view, the consequence is to have a coherent signal fragmented with the separated parts appear in more than one IMF component.

To deal with the mode mixing problem, a subjective intermittence test algorithms had been proposed by Huang *et al.* [1999]. The shortcomings for this approach had been discussed in great detail by Wu and Huang [2009], when they proposed the Ensemble EMD (EEMD), which essentially resolved the mode mixing problem associated with EMD through the help of added noises. Adding noise to the signal to be analyzed is not new. Gledhill [2003] first employed noise to investigate the effect of adding noise as a tool for checking stability of EMD. He made a critical assumption, that the true decomposition is a result from the original data without any added noise. In contrast to this view, Wu and Huang [2009] proposed that the true decomposition should be the limit when the number of ensemble approaches infinity. Undoubtedly, EEMD is a breakthrough in the development of EMD algorithm and it works well to enhance the stability of EMD algorithm remarkably. Practicality, however, limited the number that one could employ in the ensemble; therefore, the resulting IMFs derived from EEMD would inevitably be contaminated by the added noise especially when the number of ensemble was relatively low. This is especially true in the reconstruction of the signal from the IMF components. In this paper, we proposed a novel noise enhanced algorithm to improve the efficiency of the original noise assisted algorithm of EEMD by using each noise in pairs with plus and minus signs. Contrary to the requirement of EEMD, which calls for independent and identically distributed (IID) noise, the present paired noises are perfectly anti-correlated. The advantage for this new approach, however, is to have an exact cancellation of the residual noise in the reconstruction of the signal. We designate this new approach as the Complementary EEMD (CEEMD). As the perfectly correlated noise would not directly contribute to the reduction of RMS statistics, one would still need the same number of ensemble to achieve the desired tolerable level of RMS noise in the decomposed results. The rest of this paper is organized as follows. Section 2 gives the CEEMD algorithm, which is followed by numerical experiments to check the efficiency of the new approach plus an application on the blood pressure data in Sec. 3; the conclusion and discussion will be given in the final section.

2. CEEMD Algorithm

2.1. *EMD algorithm*

Empirical Mode Decomposition (EMD) is an adaptive method to remove oscillation successively through repeated subtraction of the envelope means. For a signal $x(t)$, the EMD algorithm consisted of the following steps:

(1) Connecting the sequential local maxima (respective minima) to derive the upper (respective lower) envelop using cubic spline.

(2) Derive the median of envelope, $m(t)$, by averaging the upper and lower envelopes.
(3) Extract the temporary local oscillation $h(t) = x(t) - m(t)$.
(4) Repeat Steps (1)–(3) on the temporary local oscillation $h(t)$ until $m(t)$ is close to zero. Then, $h(t)$ is an IMF noted as $c(t)$.
(5) Compute the residue $r(t) = x(t) - c(t)$.
(6) Repeat steps from (1) to (5) using $r(t)$ for $x(t)$ to generate the next IMF and residue.

Therefore, the original signal $x(t)$ can be reconstructed by the following formula:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{1}$$

where $c_i(t)$ is the ith IMF (i.e. local oscillation) and $r_n(t)$ is the nth residue (i.e. local trend).

As the algorithm use all the local extrema to construct the envelopes, the mode mixing would be inevitable when the signal contains intermittent processes. As discussed by Wu and Huang [2009], the intermittence would cause the resulting true physical processes to be obscured by the fragmentation of a given signal.

2.2. *The noise assisted method of EEMD*

EEMD is a marked milestone in the development of EMD algorithm. Proposed by Wu and Huang [2009], they utilized the fact that the white noise could provide a uniformly distributed scale in the time–frequency space. The intrinsic oscillations in the signal with different scales would automatically associate with the similar scales of reference gridings provided by white noise. Therefore, the intrinsic local oscillations can be filtered adaptively to proper scales via the natural filter bank of EMD associated with the added uniformly distributed white noise. The problem of mode-mixing problem is solved through this elegant use of noise.

The residue of added white noises should decrease following the well-established statistical rule,

$$\varepsilon_n = \frac{\varepsilon}{\sqrt{N}} \tag{2}$$

where N is the number of trials used to derive the ensemble IMFs, ε is the RMS amplitude of added noises, and ε_n is the final standard deviation of error, which is defined as the difference between the original data and the reconstructed data via the ensemble IMFs and the final residue by EEMD.

Here, a numerical simulation was used to illustrate the mode-mixing problem caused by intermittent processes. In this experiment, the simulated signal contains three sinusoid waves having different initial phases, amplitudes and frequencies and an intermittent signal. The main components of the simulated signal are given as:

$$x(t) = \sin\left(2\pi 10t + \frac{\pi}{2}\right) + 0.2\sin\left(2\pi 4t + \frac{\pi}{3}\right) + 0.1\sin\left(2\pi t - \frac{\pi}{4}\right) \tag{3}$$

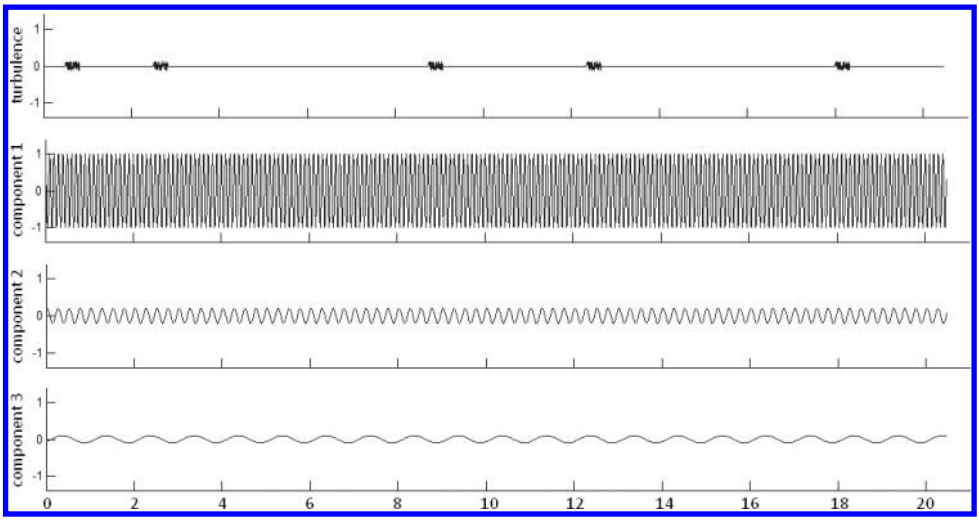


Fig. 1. The components of the simulated signal, which contains three sinusoidal waves and the simulated intermittent turbulences.

In addition, the intermittent sections of stochastic signal with zero-mean and appear in different time points. The sinusoid waves and intermittent turbulence of the simulated signals are shown in Fig. 1. The simulated signal was decomposed by EMD with the results shown in Fig. 2. Clearly, many parts of IMF 1 were replaced by the intermittent signal. The replaced parts of IMF 1 are shifted to IMF 2 resulting in the phenomenon of mode-mixing in the second and all the following IMFs. This result illustrated the mode-mixing problem caused by intermittent turbulences as discussed in detail by Wu and Huang [2009].

The same simulated signal decomposed by EEMD, with 1000 ensemble members and the RMS noise similar to that of first intermittent signal, resulted in the IMFs as shown in Fig. 3. There, IMF 1 shows the mixture of the intermittent signals contaminated to certain degree by the added noise; IMFs 2–4 recovered the sinusoidal waves as the original constituting the components of the simulated signal. This result proved the efficiency of EEMD for solving the mode-mixing problem. The residue of added white noise can be determined by the difference between the original and the reconstructed signals. The reconstructed signal is the sum of all IMFs and the final residue in EEMD. Now, we will introduce the new CEEMD.

2.3. *A novel noise enhanced method of CEEMD*

In the new CEEMD, white noise is added in pairs to the original data (i.e. one positive and one negative) to generate two sets of ensemble IMFs. Therefore, we can derive two mixtures composed of the original data and added noise by

$$\begin{bmatrix} M_1 \\ M_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} S \\ N \end{bmatrix} \tag{4}$$

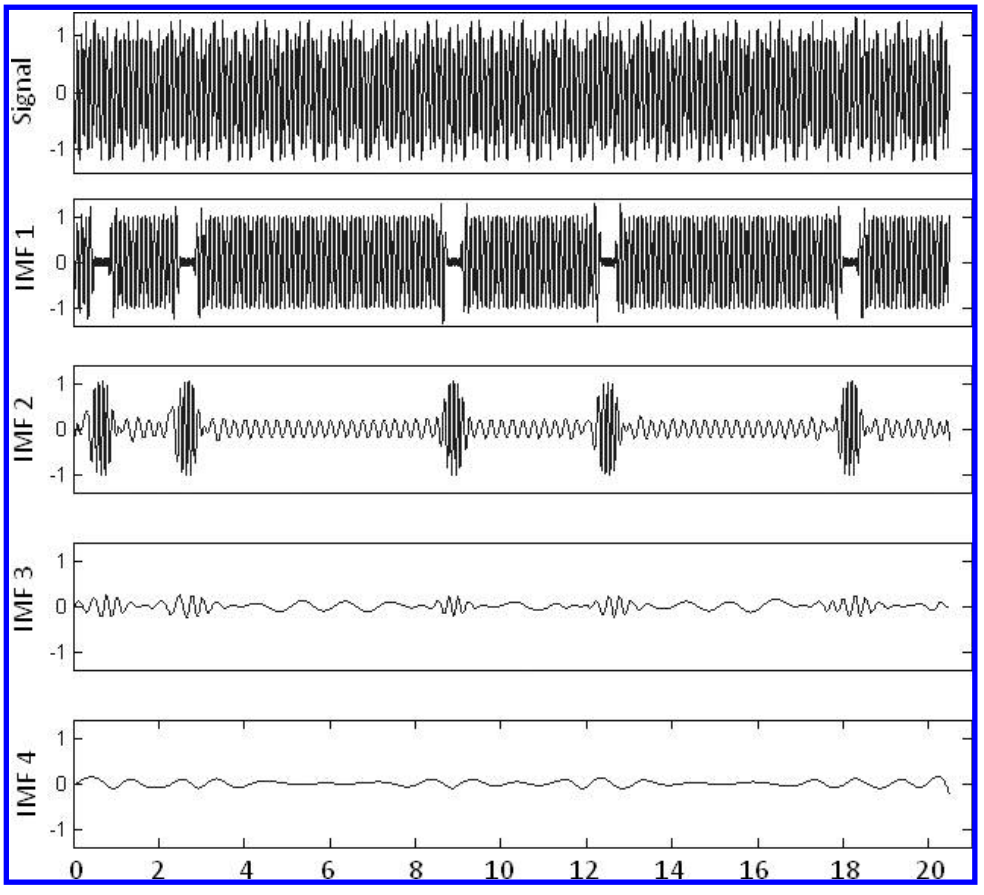


Fig. 2. The simulated signal and the decomposed IMFs by EMD. Here, the mode-shifting happens on the time periods, while the simulated turbulences appear.

where S is the original data; N is the added white noise; M_1 is the sum of the original data with positive noise, and M_2 is the sum of the original data with the negative noise.

Then, the ensemble IMFs obtained from those positive mixtures contribute to a set of IMFs with positive residues of added white noises. Similarly, the ensemble IMFs obtained from those negative mixtures contribute to another set of ensemble IMFs with negative residue of added white noises. Thus, the final IMF is the ensemble of both the IMFs with positive and negative noises.

Figure 4 shows the IMFs decomposed from the simulated signal using 20 pairs of added white noises. Through CEEMD we also obtained four IMFs, which are all similar to those produced by EEMD. Again, IMF 1 shows the mixture of intermittent signal and some residue of the white noises added; IMFs 2–4 are the components of sinusoid waves of the simulated signal. Visual comparison of the results from EEMD

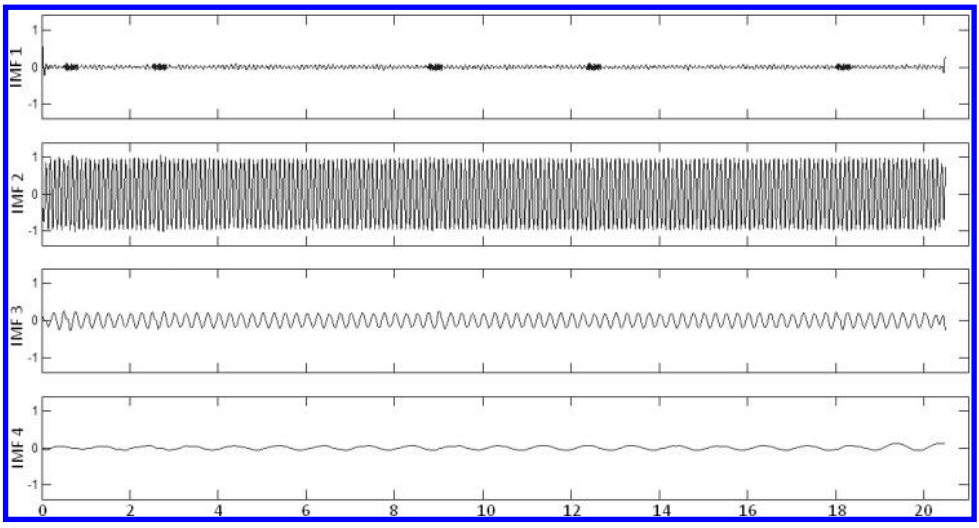


Fig. 3. The decomposed IMFs of the simulated signal by EEMD.

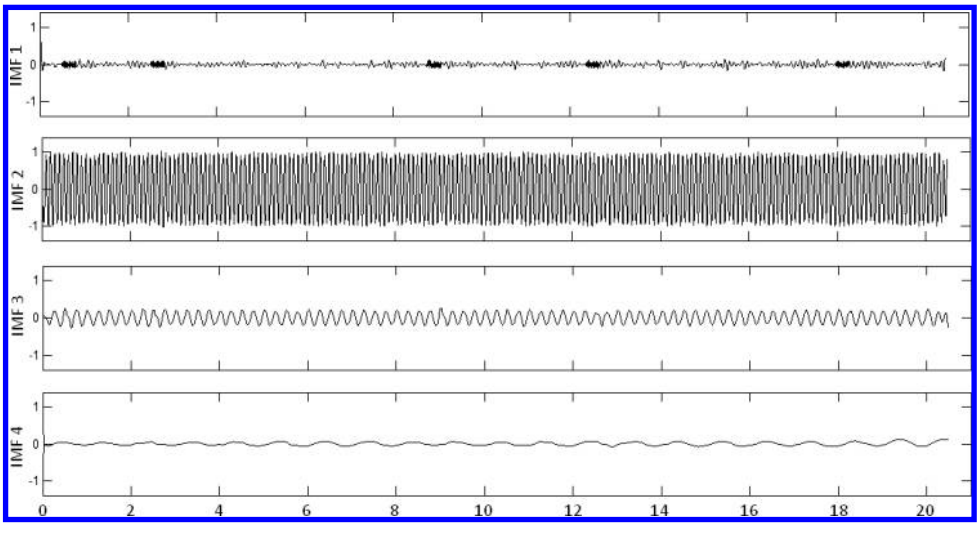


Fig. 4. The decomposed IMFs of the simulated signal by CEEMD.

and CEEMD shows no significant difference. However, there is a significant difference between the reconstructed signals via the IMFs and the original signal. The final residue derived from EEMD and CEEMD defined as the difference between the original and the reconstructed signals is very different and is shown in Fig. 5. While the residue from EEMD has an average amplitude around 0.03, the corresponding residue from CEEMD has an average amplitude close to 0 (of the order

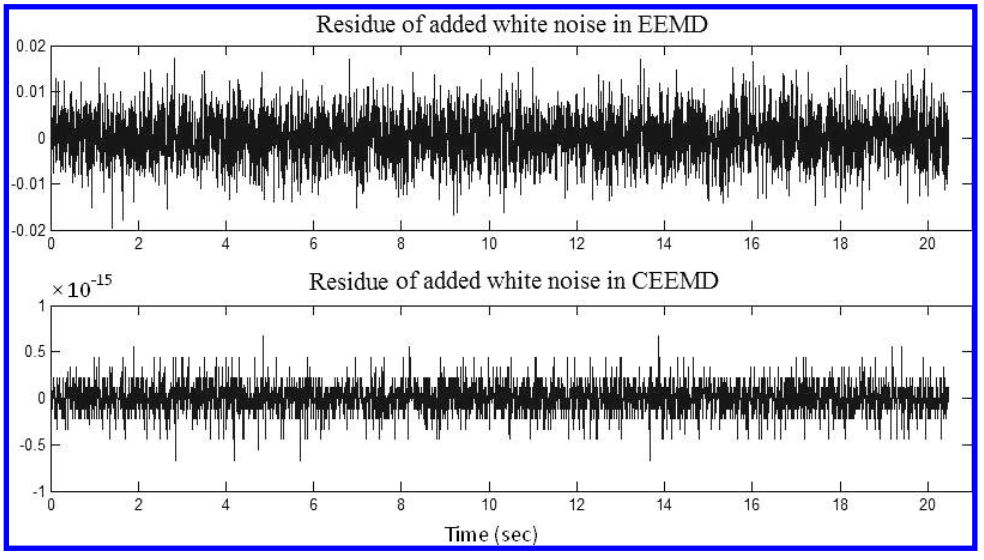


Fig. 5. Residues of added white noises derived by EEMD and CEEMD.

of 10^{-15}). Such an error could be very well attributed to the numerical error generated in the calculation. Thus, CEEMD can improve the results of decomposition by eliminating the residue of added white noise.

2.4. *Advantages of CEEMD*

Comparing results from CEEMD with those of EEMD, we can conclude that there might be a computational time saving if the reconstruction of the final result is a concern, because the paired noises could effectively reduce the final white noise residue. As the proof of this advantage of CEEMD, we have conducted a numerical experiment. In this numerical experiment, different numbers (from $10^{0.2}$ to 10^{4}) of added white noises were used to evaluate the residue of added white noises expressed in term of percentage. The results of our experiment, shown in Fig. 6, indicate that for EEMD the residue depends on the members in the ensemble as expected. In contrast with EEMD, CEEMD can eliminate the residue of added white noises totally no matter how many noises were used.

Other than the elimination of the final residue noise, the performances of the EEMD and CEEMD are comparable in terms of RMS errors for each IMF. Table 1 summarizes the results of mean squared errors of every dominant components and the residue of added white noises in this test. Of course, a large value of mean squared error indicates a significant difference between the decomposed and the original components, and hence a poor performance of decomposition. According to the results shown in Table 1, EEMD and CEEMD have similar performances in the decomposition of single component when the same number of white noise was used.

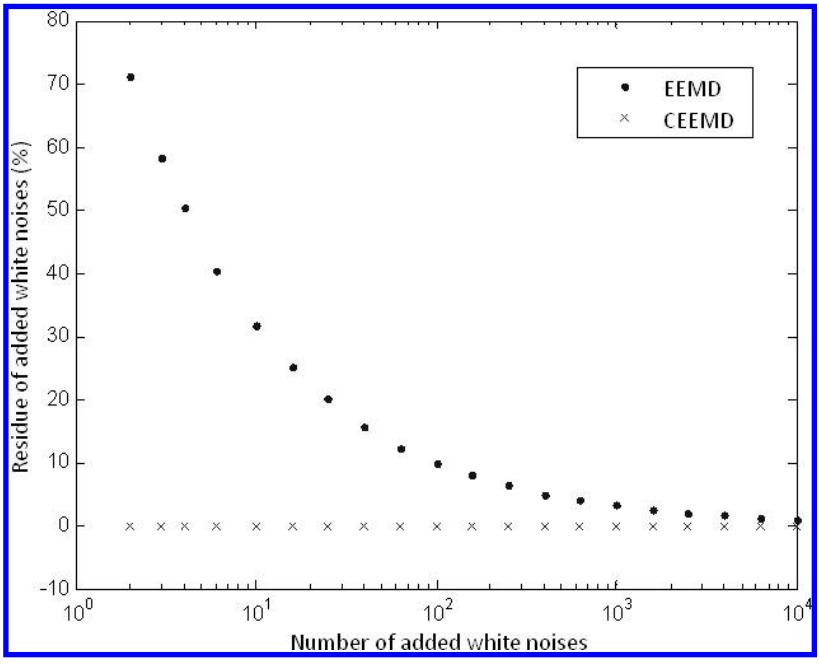


Fig. 6. The correlation between numbers and residues of added white noises in EEMD and CEEMD.

Table 1. The residue of added white noises and the percentage mean square errors (PMSE) between the original components/their corresponding IMFs by EEMD and CEEMD using different numbers of added white noises.

Number of added white noises	PMSE of component 1 (%)		PMSE of component 2 (%)		Residue of added white noises (%)	
	EEMD	CEEMD	EEMD	CEEMD	EEMD	CEEMD
40	2.08	1.89	19.78	19.56	13.43	0
80	1.83	1.67	19.56	19.67	9.68	0
160	1.69	1.60	19.13	19.62	6.80	0
320	1.62	1.58	18.59	18.59	4.79	0
640	1.57	1.56	18.88	15.50	3.45	0
1280	1.54	1.57	18.48	18.59	2.37	0
2560	1.55	1.56	18.52	18.48	1.69	0

Note: The residue of added white noise is shown as the ratio between energy densities of the residue and the added white noise. The mean square error is shown as percentage correlated to mean square energy of the component. The number of added white noises used in CEEMD is a half of that used in EEMD to take comparisons using the same scale of calculations.

3. A Study of the Effects of Intrinsic Noise

By now, it is well established that, in EEMD and the newly proposed CEEMD, added noise could help us to extract useful information from a data. Left unexplored is the effect of intrinsic noise, the noise contained in the data. As the intrinsic noise

could not be removed by EEMD or CEEMD, the question of how the pre-existing noise interacts with the added noise begs detailed investigation. This is a complicate problem requiring extensive studies. Here, we will only make a limited trial in two numerical experiments to highlight the problem, but not the solution. The limited goal here is to investigate the correlations among intrinsic noise, added noise and dominant components of signal. Here, the dominant components are designed to be the main information containing part of the signal. The test signals consist of the known dominant component and added intrinsic noise. To control the energy level of the intrinsic noise, it is the simulated with a pre-determined S/N ratio to the total energy of signal contributed by dominant components. To measure the performance of CEEMD, we use the coherence between the original dominant components and their corresponding IMFs.

3.1. *Test data for numerical experiments*

In this study, two different types of signals were used: one simulated mixed regular sinusoidal signal and another from a biomedical experiment.

3.1.1. *Determining the dominant components from a biomedical experiment*

Monte Carlo verification is an efficient method used to determine the characteristics of IMFs according to the distribution of energy densities and their corresponding average periods. In the study of the characteristics of white noise, Wu and Huang (2004) explored the relationship between the energy density and the average period by the Monte Carlo test. They proposed a method to test the information content of a data set relative to the unknown noise level. This significant test method provides the criterion for determining which IMFs from a noisy data set contain statistically significant information, and which IMFs are purely noise. Here, we will follow the same approach. In this Monte Carlo verification, two parameters, energy density and its corresponding averaged period, were defined to characteristicize the targeted IMF. As suggested by Wu and Huang [2009], the energy density and average period are calculated by the following equations:

$$E_n = \frac{1}{N}\sum_{j=1}^{N}[C_n(j)]^2 \tag{5}$$

$$\bar{T}_n = \int S_{\ln T,n} d\ln T \left(\int S_{\ln T,n}\frac{d\ln T}{T}\right)^{-1} \tag{6}$$

where E_n is the energy density of the nth IMF; $S_{\ln T,n}$ is the Fourier spectrum of the nth IMF as a function of $\ln T$; T is period, and $\bar{T}_n$ is the averaged period of the nth IMF.

In a separate study by Flandrin *et al.* [2004] of fractal Gaussian noise (fGn) using EMD, the distribution plot of energy densities versus their corresponding averaged

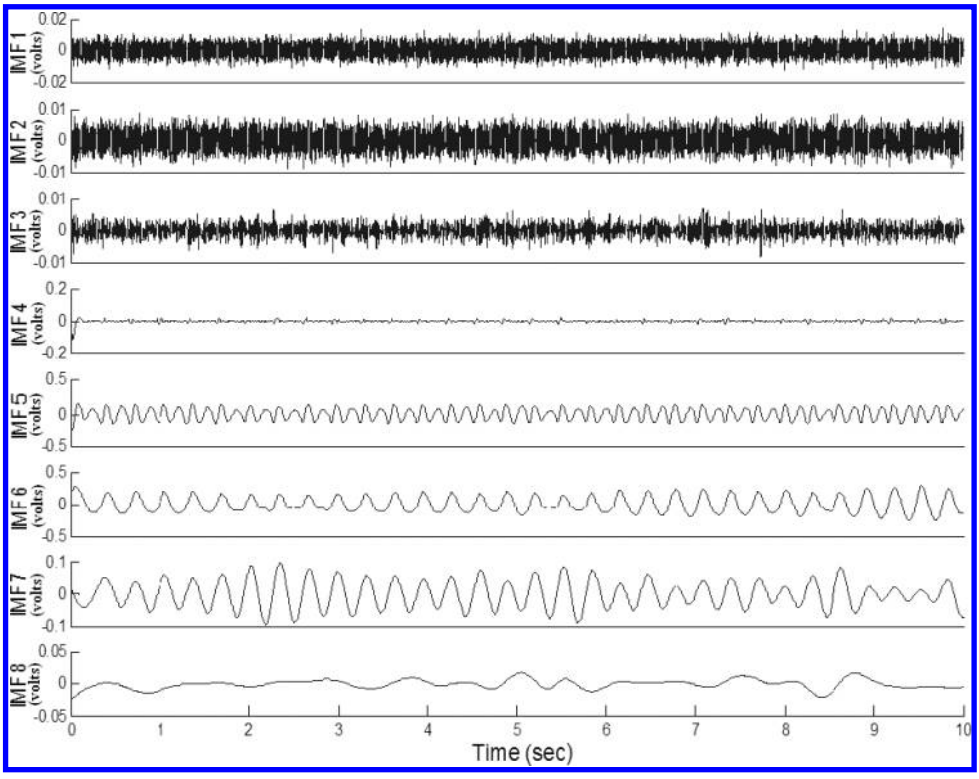


Fig. 7. The first 8 IMFs decomposed from a pig's blood pressure by CEEMD using 20 added white noises.

periods resulted in straight lines with the slope reflecting the fractal property of fGn.

However, the distribution of energy densities versus their corresponding average periods for the IMFs decomposed from our experimental signal indicates that the signal contains many statistically significant components according to the significant test proposed by Wu and Huang [2009]. In this study, a pig's blood pressure signal in a biomedical experiment is studied. The original data were decomposed through CEEMD as shown in Fig. 7. Then, we calculated the energy densities and average periods of IMFs and plotted the distribution of energy density versus their corresponding average periods as shown in Fig. 8. The energy density/averaged period plot for the first three IMFs presents a distribution similar to a straight line similar to the result from white noise as shown in the same figure. Therefore, these three IMFs were identified as the noisy components according to Wu and Huang [2009]. IMFs 4–8, however, are way above the significant limit for white noise; therefore, they are identified as the dominant components of signal. Of all the significant components, IMFs 4–7 are periodic mode functions that contribute

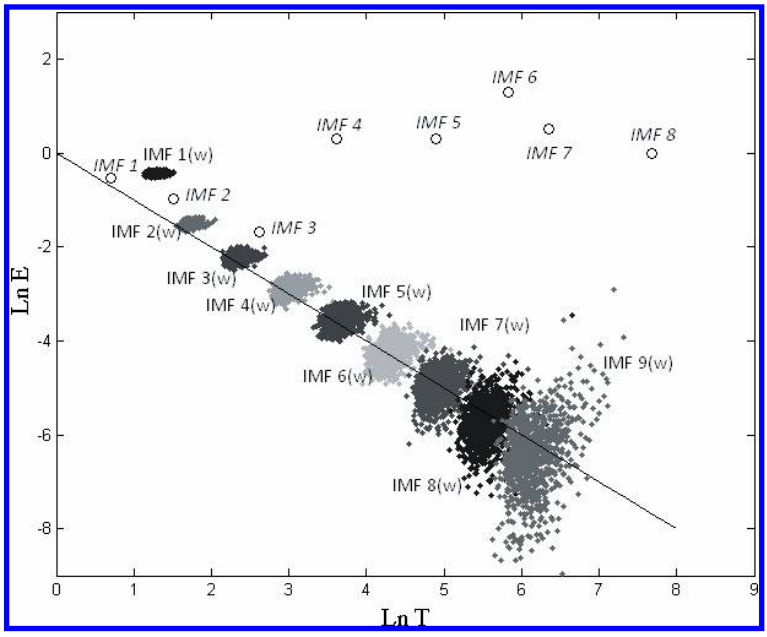


Fig. 8. Illustration of Monte Carlo verification. In this presentation, the distribution of energy densities versus their corresponding averaged periods for the first 8 IMFs decomposed from a pig's blood pressure noted as italic IMF 1–IMF 8. The straight line and the points noted as IMF 1(w)–IMF 9(w) present the distributions of the first 9 IMFs for 2000 simulated white noises.

to the main waveform of a blood pressure signal. IMF 8 presents a related long-term component compared with the period of cardiac cycle. Therefore, the main waveform of the pig's blood pressure can be reconstructed via IMFs 4–7, which is used as our dominant signal in the numerical test with controlled noise as follows.

To separate signal and noise strictly, the sum of IMFs 4–7 were smoothed by median filters to eliminate any fluctuations riding on themselves from the EEMD operation. Hence, the smoothed dominant components of a pig's blood pressure were used as the test signal of the numerical experiment. Furthermore, we simulated the natural noisy components using a stochastic time series with pre-determined energy scale. This simulated noisy component is treated as intrinsic noise, which is different from the added noise used in EEMD or CEEMD. To determine the energy scale of intrinsic noise, signal/noise (S/N) ratio between energy densities of main waveform (i.e. mixture of the dominant components) and intrinsic noise was calculated. S/N ratio between main waveform and intrinsic noise of the pig's blood pressure is 194.36 (i.e. the total energy density of dominant components is 194.36 times of that of noisy components). Thus, we determined energy scale of intrinsic noise for the simulated signals is in a range of S/N ratio from 2 to 5000 (i.e. from 6 to 74 dB) including the S/N ratio of the experimental signal (i.e. a pig's blood pressure).

3.1.2. *Test data of simulated nonstationary sine waves*

The dominant components used in this numerical experiment were simulated using sinusoidal waves with varying amplitudes and instantaneous frequencies. These sine waves can be expressed as the following formulas (7):

$$\begin{aligned} x_i &= a\sin(\phi+\theta) \\ a &= a_0 + b_1\sin(\phi_1 + w_1 t) \\ \theta &= \theta_0 + b_2\sin(\phi_2 + w_2 t) \end{aligned} \tag{7}$$

where x_i is the ith component of the simulated signal; a is the amplitude of the component and θ is the instantaneous frequency of component; ϕ is the initial phase angle of component; $a_0, \theta_0, b_1, b_2, \phi_1, \phi_2, w_1$, and w_2 are constants.

Basically, the simulated component is a sine wave with periodic amplitude and instantaneous frequency fluctuations. Amplitude a has a fixed value of a_0 and fluctuating range of b_1. Frequency of component also has a fixed value of θ_0 and fluctuating range of b_2. ϕ_1 and ϕ_2 are initial phase angles of fluctuations for amplitude and frequency of component. w_1 and w_2 are angle speeds of fluctuations for amplitude and frequency of component. Values of all parameters of Eq. (7) for two simulated components are shown in Table 2. Time–frequency–amplitude distributions of two simulated components were derived from Hilbert transform. Time series and time–frequency–amplitude distributions are shown in Fig. 9.

3.2. *The simulated signals with various intrinsic noises*

Since IMFs decomposed by EMD had been proven to satisfy the condition of orthogonal, the energy density of the main waveform of material can be derived by superposition. Therefore, the energy density of main waveform is defined as the total energy density of these original dominant components. Then, intrinsic noises were simulated using the stochastic time series generated by the random number generator of MATLAB with pre-given energy scales, which induce the ratios between the energy densities of intrinsic noise and main waveform of signal. Signal/noise ratio is presented in logarithmic scale. Figure 10 shows the simulated signals of a pig's blood pressure with different intrinsic noises.

3.3. *Pearson's correlation coefficient and weighted Pearson's correlation coefficient*

Pearson's correlation coefficient (PCC) is often used to estimate the consistence between two signals. In this study, we used it to check consistence between the

Table 2. Values of all parameters used in two simulated components for the second numerical experiment.

Parameters	ϕ	a_0	b_1	ϕ_1	w_1	θ_0	b_2	ϕ_2	w_2
Component 1	0.6	0.1	0.005	$\pi/2$	0.5	5	0.05	$3\pi/2$	1.67
Component 2	1.0	0.1	0.1	$\pi/12$	0.2	2	0.1	$\pi/2$	0.125

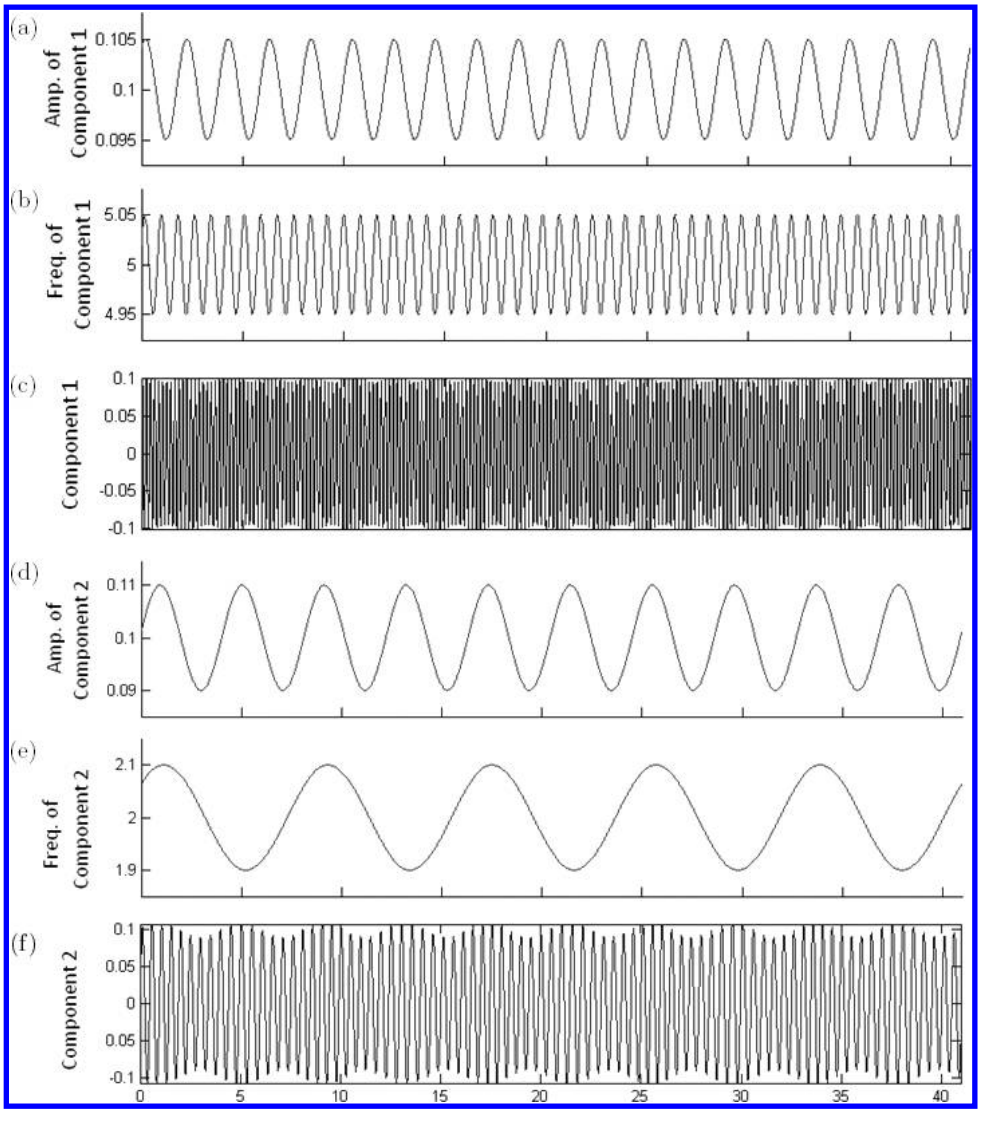


Fig. 9. Time series and time–frequency–amplitude distributions of two simulated components. (a, d) Time–amplitude distributions of two components. (b, e) Time–frequency distributions of two components. (c, f) Simulated time series of two components.

original component and its corresponding IMF derived by CEEMD. The corresponding IMF of a dominant component of a simulated signal is determined as the IMF performs the highest value of PCC to the dominant component. Moreover, the simulated signals also have several dominant components but not single component. Value of PCC just reflects the consistence between one component and its corresponding IMF. An overall consistence between all dominant components and their corresponding IMFs can be expressed by a weighted PCC. Weighting factors are

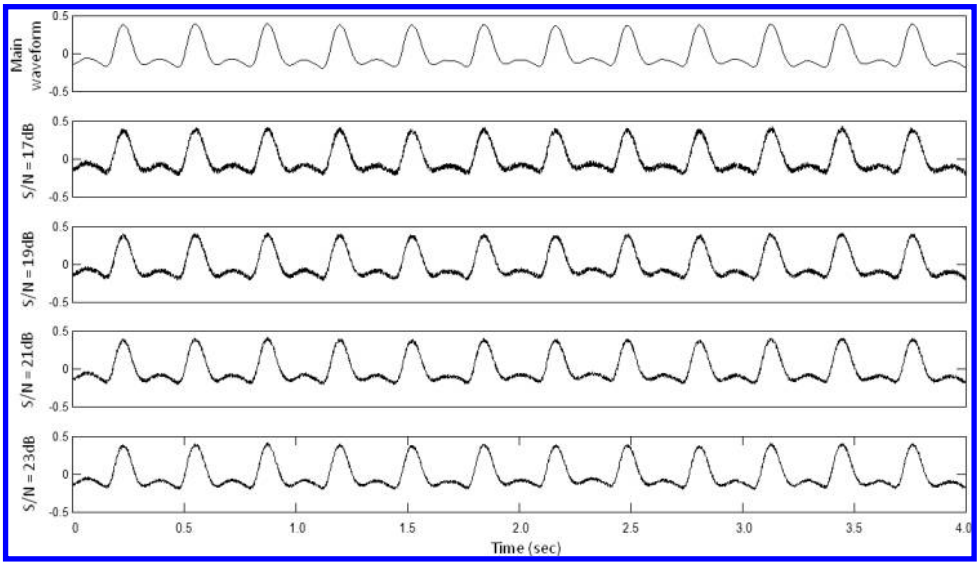


Fig. 10. Main waveform and the simulated signals of a pig's blood pressure with different intrinsic noises.

energy densities of dominant components. Thus, weighted PCC can be calculated by the following equation:

$$P_{\mathrm{o}} = \frac{\sum_{i=1}^{n} P_i \times E_i}{\sum_{i=1}^{n} E_i} \tag{8}$$

where P_{o} is the overall PCC; P_i is the value of PCC between the ith dominant component and its corresponding IMF; E_i is the energy density of the ith dominant component.

4. Results Form the Numerical Experiments

In our numerical experiments, intrinsic noises were simulated using random time series with different energy scales. S/N ratio was used to identify the proportion of intrinsic noise to main waveform of the simulated signal. Simulated signals containing different intrinsic noises with S/N ratio from 2 to 5000 (i.e. 6–74 dB) were decomposed by CEEMD using added white noises with S/N ratio from 2 to 5000. Figures 11 and 12 show the results of weighted PCC in two numerical experiments. Both figures show the trenches with correlatively low values of weighted PCC in blue areas. As shown in Fig. 11, when the intrinsic noise is high (S/N ratio < 200) and added noise is smaller than intrinsic noise, weighted PCC would be low. It shows that a larger energy scale of added noise should be used in CEEMD for signals with large intrinsic noises. On contrast to CEEMD for signals with large intrinsic noise, no optimal energy scale of added noise could be found.

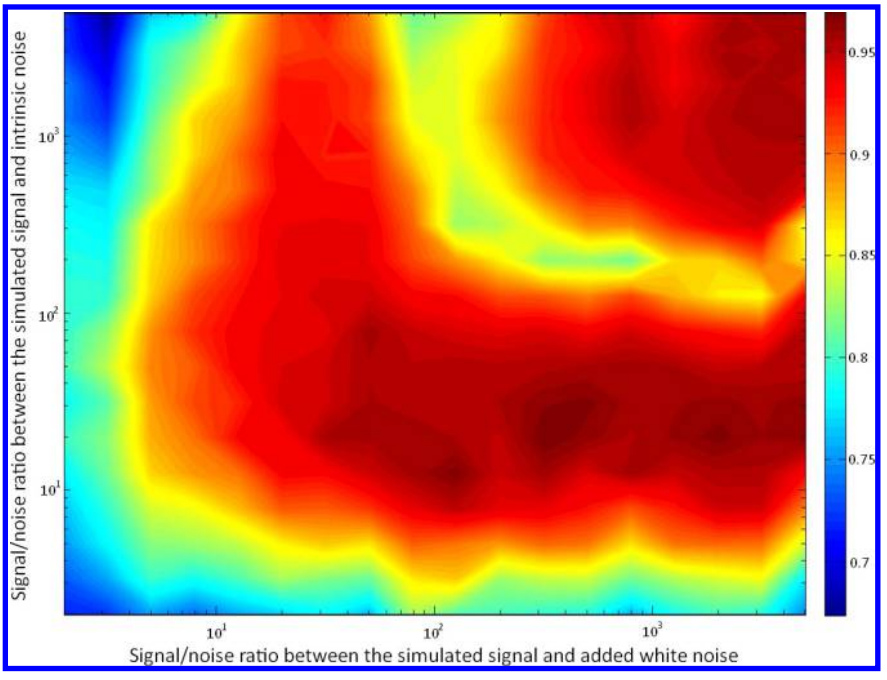


Fig. 11. Result of the first numerical experiment using the dominant components decomposed from a pig's blood pressure. Values of weighted PCC between the dominant components and their corresponding IMFs are shown in colors. x-axis shows S/N ratio between the dominant components of the simulated signal and the simulated intrinsic noise. y-axis shows S/N ratio between the dominant components of the simulated signal and the added noise used in CEEMD.

Figure 13 shows PCC values between the dominant components and their corresponding IMFs. Figures 13(a) & 13(d) shows the results for a simulated signal with low intrinsic noise energy scale (S/N ratio = 1259); Figs. 13(b) & 13(e) shows the results for a simulated signal with median intrinsic noise energy scale (S/N ratio = 50 dB); and Figs. 13(c) & 13(f) shows the results for a simulated signal with high intrinsic noise energy scale (S/N ratio = 79.43). There are two subplots for each simulated signal. Figure 13(b) shows the PCC values between the first component of the simulated signal with median intrinsic noise and its corresponding IMF and Fig. 13(e) shows those between the second component and its corresponding IMF.

According to results shown in these two subplots (i.e. Figs. 13(b) and 13(e)), IMF 2 has high PCC value with the first simulated component and IMF 3 has high PCC with the second simulated component with median intrinsic noise in CEEMD and the added white noises of S/N ratio > 79.43. It should be noted that this correspondence is not fixed for all the intrinsic noise or the added noise energy levels. If the simulated signal with median intrinsic noise is decomposed by CEEMD using added noise of S/N ratio = 17, IMF 2 corresponds to the first simulated component (PCC > 0.9) and both IMFs 3 & 4 correspond to the second component

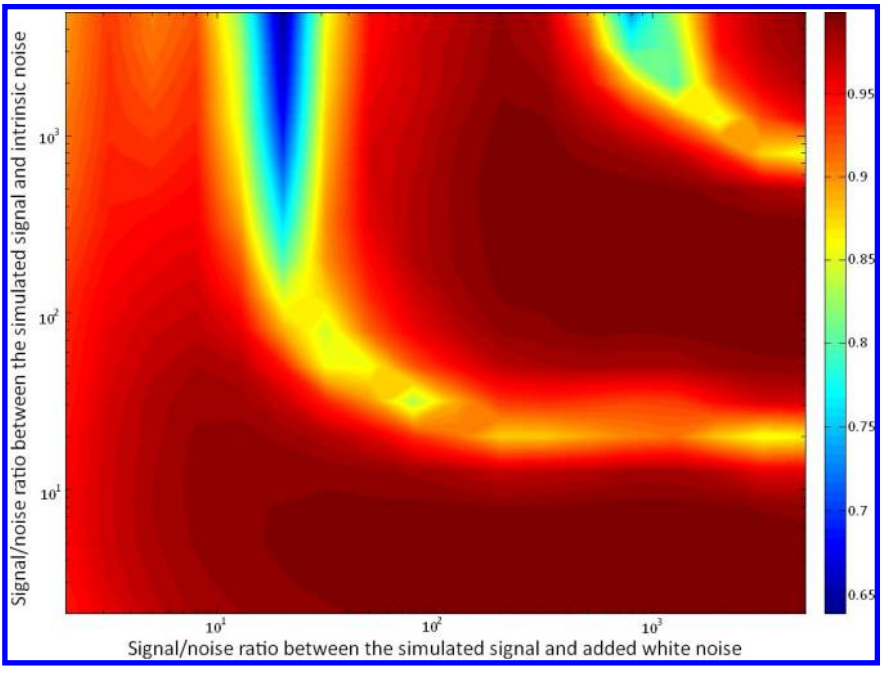


Fig. 12. Result of the second numerical experiment using two sinusoidal waves with modulating periods and amplitudes. Values of weighted PCC between the dominant components and their corresponding IMFs are shown in colors. x-axis shows S/N ratio between the dominant components of the simulated signal and the simulated intrinsic noise. y-axis shows S/N ratio between the dominant components of the simulated signal and the added noise used in CEEMD.

(PCC > 0.9). If the simulated signal with median intrinsic noise is decomposed by CEEMD using added noise of S/N ratio $= 11$, IMFs 2 & 3 correspond to the first component (PCC > 0.9) and IMF 4 corresponds to the second component (PCC close to 1). Moreover, when the simulated signal is decomposed by CEEMD using added noise with S/N ratio < 7.9, IMF 3 corresponds to the first component and IMF 4 corresponds to the second component. Thus, we found that a mode-translation phenomenon associates with the increase of energy scale of added noise in CEEMD. This is easy to understand, for as the added noise energy increases, it would over power the signal and produce some artificial noise components without corresponding significant signal near the scale.

In addition to the mode-translation phenomenon, the leakage of the dyadic filter bank as discussed in Wu and Huang [2009] could also cause a single dominant component in the simulated signal to reside in 2 IMFs, or mode splitting. This leakage caused mode splitting becomes more severe when the energy level of the added noise is high, for the higher added noise in EEMD or CEEMD would cause more perturbations and hence there are more chances for mode splitting. As the added noise energy level changes, the leakage pattern would also change accordingly. Consequently, the mode splitting of a single signal into two components could

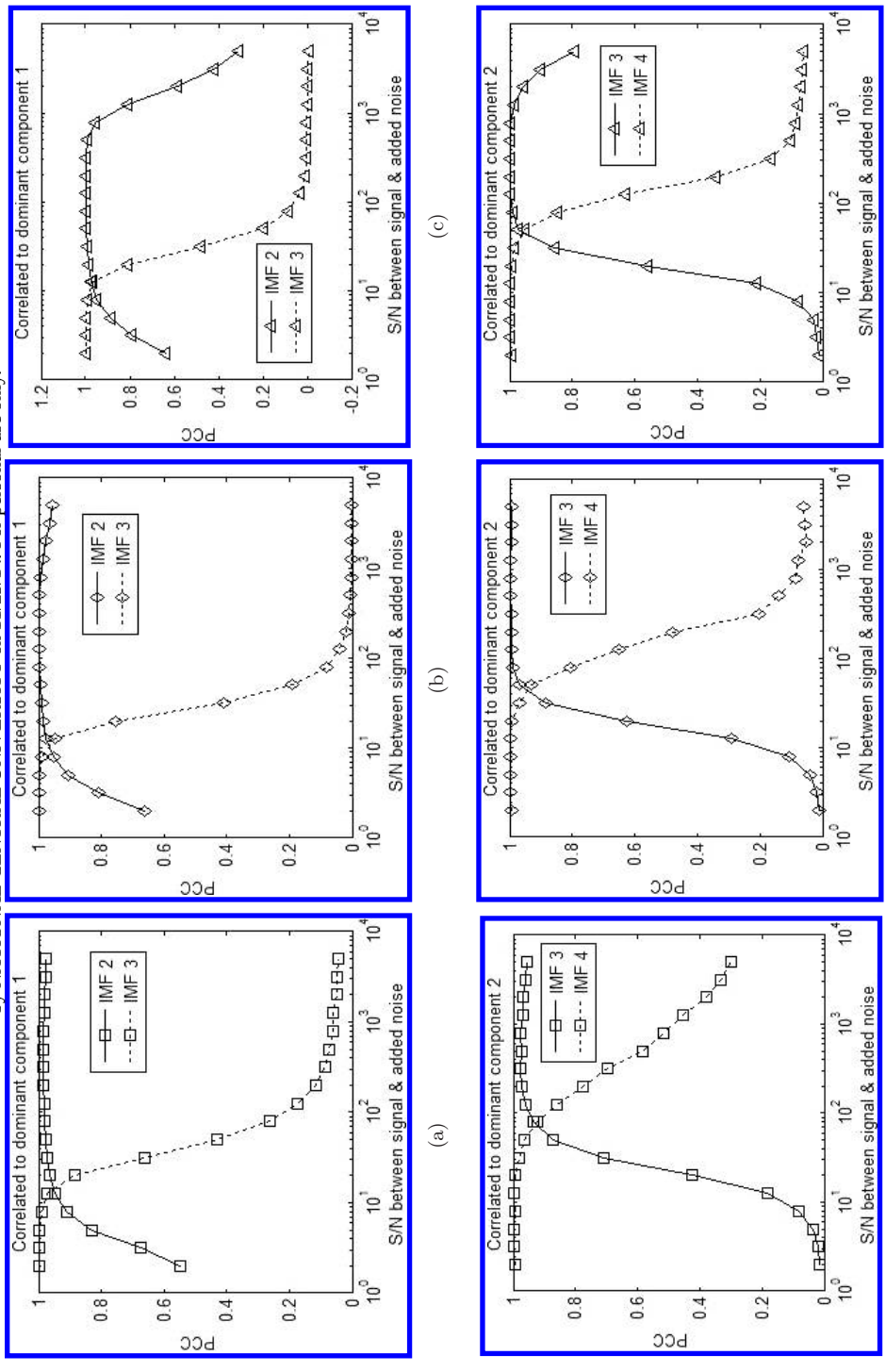


Fig. 13. PCC values between the dominant components and their corresponding IMFs in the second numerical experiment. (a, d) Results for a simulated signal with low energy scale of intrinsic noise (S/N ratio = around 1259, 62 dB). (b, e) Results for a simulated signal with median energy scale of intrinsic noise (S/N ratio = around 316.23, 50 dB). (c, f) Results for a simulated signal with high energy scale of intrinsic noise (S/N ratio = around 79.43, 38 dB).

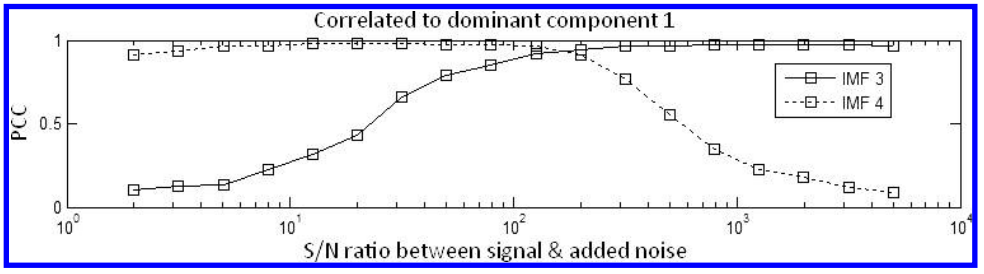
happen at unpredictable locations. Fortunately, an orthogonal check could spot this mode splitting easily. The method to remedy this mode splitting is to add the two adjacent nonorthogonal IMF components to form a single one as suggested by Wu and Huang [2009].

In the calculation of weighted PCC, we consider that the dominant components have a relationship of one-to-one mapping to their corresponding IMFs. In fact, when an inappropriate added noise is used in CEEMD, low values of weighted PCC appear in the area that phenomena of mode-translation and leakage caused mode splitting to occur. These are the main causes for trenches of low weighted PCC values in Figs. 11 and 12.

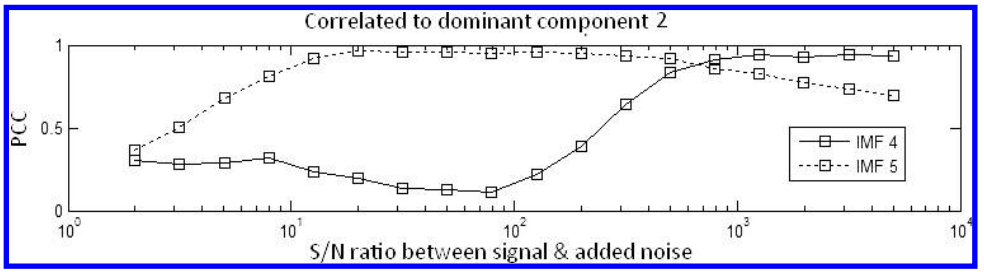
Figure 14 shows the PCC values between the dominant components and their corresponding IMFs in the biomedical numerical experiment. Signal is simulated using dominant components from a pig's blood pressure and median energy scale with S/N ratio of 316.23 (i.e. 50 dB). Similarly to results of the second numerical experiment, phenomenon of mode-translation also happens during the increasing of energy scale of added noise. Moreover, phenomenon of mode-translation happens within a wider range of energy scale of added noise, because the natural signal had a wider frequency band than the simulated sinusoidal signals.

To clarify the phenomena of mode-translation and leakage caused mode splitting, the simulated signal with intrinsic noise of S/N ratio of 316.23 (i.e. 50 dB) had been decomposed by CEEMD using three different energy scales of added noises. These three energy scales of added noises present three situations of before, during, and after mode-translation and leakage caused mode splitting as shown in Fig. 15. In this illustration, Fig. 15(a) shows the simulated signals associating with two dominant components and intrinsic noise, and Figs. 15(b)–15(d) show IMFs 1–4 decomposed by CEEMD using added noise with different energy scales of S/N ratio around 126, 50 and 12.59 (i.e. 42, 34 and 22 dB). When added noise with energy scale of S/N = 126 (i.e. 42 dB) is used in CEEMD, it is clear that IMF 2 is the corresponding IMF for the simulated component 1 and IMF 3 for component 2. When the energy scale of added noise raises the S/N ratio to 50 (i.e. 34 dB), IMF 2 also is the corresponding IMF for component 1 as shown in Fig. 15(b), but IMFs 3 and 4 both have the similar time scale (albeit with different energy densities) to component 2 as shown in Fig. 15(c). As the added noise energy level raises, IMF 4 gradually becomes the only corresponding IMF for component 2 as shown in Fig. 15(d). But under this condition, IMFs 2 and 3 both have the same time scale as the first simulated component. All this happened without a firm rule to determine when this mode splitting would occur. Therefore, it is not possible to select the correct added energy level to implement either EEMD or CEEMD. The remedy is to check orthogonality index diligently to spot the splitting and re-combine them a posteriori, if any two adjacent IMF components are grossly nonorthogonal.

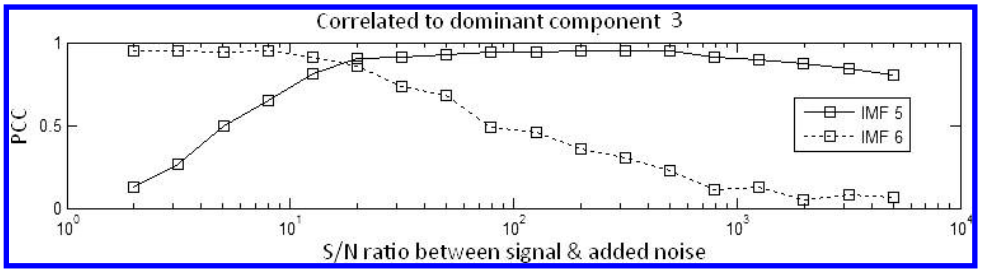
Alternatively, as shown in the case studied here with the simulated signal consisted of intrinsic noise energy scale of S/N = 316.23 (i.e. 50 dB), we can implement an extensive test of the parameter space for the added noise and look for the



(a)



(b)



(c)

Fig. 14. PCC values between the dominant components and their corresponding IMFs in the first numerical experiment. Signal is simulated using median energy scale of intrinsic noise (S/N ratio = around 316.23, 50 dB). (a) The PCC values between the first component and its corresponding IMF (IMF 3 or 4). (b) The PCC values between the second component and its corresponding IMF (IMF 4 or 5). (c) The PCC values between the third component and its corresponding IMF (IMF 5 or 6).

optimal. As shown in Fig. 15(b), energy scale S/N = 126 (i.e. 42 dB) of added noise results in better performance (value of weighted PCC is higher) than the other two approaches using energy scale with S/N = 50 and 12.59 (i.e. 34 and 22 dB). It shows that an appropriate energy scale of added noise could improve performance of CEEMD in dominant decomposition. This approach is too time-consuming to be used routinely; consequently, this is not the approach we would recommend. There, however, is a general rule: high energy level of added noise usually causes deterioration of the high-frequency component and its corresponding IMF.

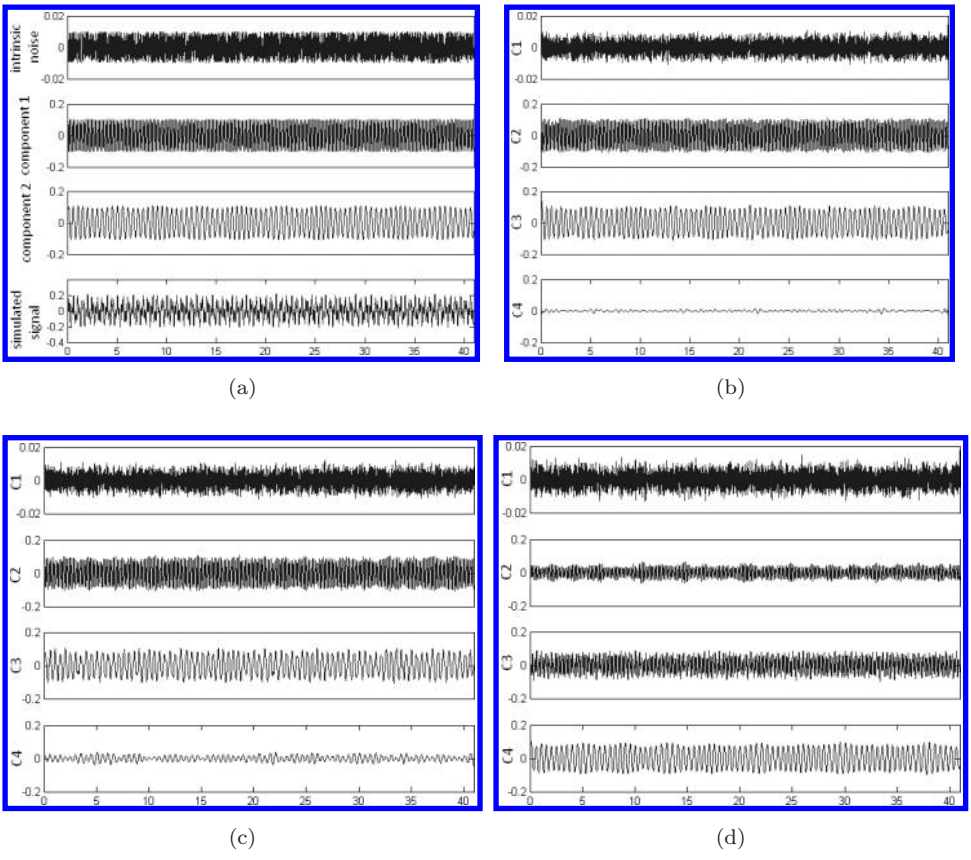


Fig. 15. Illustration of mode-transfer phenomenon. (a) The components and signal of the simulated signal used in this illustration. (b) IMFs 1–4 decomposed by CEEMD using added noise with energy scale of S/N = around 126 (i.e. 42 dB). (c) IMFs 1–4 decomposed by CEEMD using added noise with energy scale of S/N = around 50 (i.e. 34 dB). (d) IMFs 1–4 decomposed by CEEMD using added noise with energy scale of S/N = around 12.59 (i.e. 22 dB).

The above examples clearly indicate that the interactions between added white noises and real-world signals are complicated. As real signal may contain many dominant components with different energy densities and intrinsic frequency bands, it would be hard, if not possible, to determine an appropriate energy scale of added noise for general cases of CEEMD. Our experience suggests that an energy level of added noise similar to the intrinsic noise of signal is the best trial for CEEMD.

5. Discussions and Conclusion

A complementary process using both positive and negative white noise is proposed to enhance processing efficiency of EEMD. According to our numerical simulations, the residue of added white noises is completely removed in CEEMD. Therefore, if the goal is to keep the noise in the reconstructed signal, CEEMD should be the

preferred method. Our computation, however, reveals that the RMS noise level is comparable for EEMD and CEEMD. As the CEEMD could reduce the residue noise, and it would not cause more computation time, we would recommend CEEMD as the standard form of EEMD.

We have also conducted a simple numerical experiment to illustrate the effect of intrinsic noise in EEMD or CEEMD decomposition. Our experiments revealed the phenomena of mode translation (new IMF would be generated with increasing added white noise level) and leakage induced mode splitting (a significant signal resides in more than one IMF). Both of these phenomena could occur at unpredictable combinations of intrinsic noise and added noise in data analysis. Furthermore, it is also impossible to predict when and at which IMFs the phenomena would happen. Although an exhaustive search could determine the optimal added white noise level for a specific problem, this approach is not practical because it is too time consuming. The recommended method is to conduct orthogonality test diligently and recombine the adjacent nonorthogonal components into one. Based on our experience, we have reached the following general rules in EEMD and CEEMD:

(1) Higher added noise in EEMD could cause the high frequency component to be masked, if the number of ensemble is not large enough.
(2) Higher added noise level could also cause a more severe mode splitting from the leakage.
(3) Based on these observations, we recommend that the energy level of the added noise should be keep at the level of the intrinsic noise, if the level is known. Otherwise, try not to use the added noise level with RMS value more than 20% of the signal.

Acknowledgments

The authors wish to thank the National Science Council (NSC) of Taiwan (Grant Number NSC96-2221-E-155-015-MY3-2) for supporting this research. NEH is supported in part by grants NSC 98-2627-B-008-004 (Biology) and NSC 98-2611-M-008-004 (Geophysical) from the National Science Council and a grant from the Federal Highway.

References

Flandrin, P., Rilling, G. and Goncalces, P. (2004). Empirical mode decomposition as a filter bank. *IEEE Signal Process. Lett.*, **11**: 112–114.

Flandrin, P., Concalves, P. and Rilling, G. (2005). EMD equivalent filter banks, from interpretation to application. *Hilbert–Huang Transform: Introduction and Applications*, ed. N. E. Huang and S. S. P. Shen, World Scientific, Singapore.

Gledhill, R. J. (2003). *Methods for Investigating Conformational Change in Biomolecular Simulation.* A dissertation for the degree of Doctor of Philosophy at Department of Chemistry, The University of Southampton.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. and Liu, H. H. (1998). The empirical mode decomposition and the Hilbert

spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A*, **454**: 903–995.

Huang, N. E., Shen Z. and Long, R. S. (1999). A new view of nonlinear water waves — the Hilbert spectrum. *Ann. Rev. Fluid Mech.*, **31**: 417–457.

Huang, N. E., Wu, Z., Long, S. R., Arnold, K. C., Chen, X. and Blank, K. (2009). On instantaneous frequency. *Adv. Adapt. Data Anal.*, **1**(2): 177–229.

Wu, Z. and Huang, N. E. (2004). A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. R. Soc. Lond. A*, **460**: 1597–1611.

Wu, Z. and Huang, N. E. (2009). Ensemble Empirical Mode Decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.*, **1**: 1–41.

This article has been cited by:

1. Marcelo A. Colominas, Gastón Schlotthauer, María E. Torres. 2014. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control* **14**, 19-29. [CrossRef]
2. Ling Tang, Wei Dai, Lean Yu, Shouyang Wang. 2014. A Novel CEEMD-Based EELM Ensemble Learning Paradigm for Crude Oil Price Forecasting. *International Journal of Information Technology & Decision Making* **18**, 1-29. [CrossRef]
3. WENYING ZHANG, XINGMING GUO, ZHIHUI YUAN, XINGHUA ZHU. 2014. HEART SOUND CLASSIFICATION AND RECOGNITION BASED ON EEMD AND CORRELATION DIMENSION. *Journal of Mechanics in Medicine and Biology* **14**:04. . [Abstract] [PDF] [PDF Plus]
4. Jinde Zheng, Junsheng Cheng, Yu Yang. 2014. Partly ensemble empirical mode decomposition: An improved noise-assisted method for eliminating mode mixing. *Signal Processing* **96**, 362-374. [CrossRef]
5. Xuejun Chen, Jing Zhao, Wenchao Hu, Yufeng Yang. 2014. Short-Term Wind Speed Forecasting Using Decomposition-Based Neural Networks Combining Abnormal Detection Method. *Abstract and Applied Analysis* **2014**, 1-21. [CrossRef]
6. Meijiao Li, Huaqing Wang, Gang Tang, Hongfang Yuan, Yang Yang. 2014. An Improved Method Based on CEEMD for Fault Diagnosis of Rolling Bearing. *Advances in Mechanical Engineering* **2014**, 1-10. [CrossRef]
7. Cheng-Wei Huang, Pei-Der Sue, Maysam Abbod, Bernard Jiang, Jiann-Shing Shieh. 2013. Measuring Center of Pressure Signals to Quantify Human Balance Using Multivariate Multiscale Entropy by Designing a Force Platform. *Sensors* **13**, 10151-10166. [CrossRef]
8. Artūras Janušauskas, Vaidotas Marozas, Arūnas Lukoševičius. 2013. Ensemble empirical mode decomposition based feature enhancement of cardio signals. *Medical Engineering & Physics* **35**, 1059-1069. [CrossRef]
9. NAVEED UR REHMAN, CHEOLSOO PARK, NORDEN E. HUANG, DANILO P. MANDIC. 2013. EMD VIA MEMD: MULTIVARIATE NOISE-AIDED COMPUTATION OF STANDARD EMD. *Advances in Adaptive Data Analysis* **05**:02. . [Citation] [PDF] [PDF Plus]
10. CHIH-YU KUO, SHAO-KUAN WEI, PI-WEN TSAI. 2013. ENSEMBLE EMPIRICAL MODE DECOMPOSITION WITH SUPERVISED CLUSTER ANALYSIS. *Advances in Adaptive Data Analysis* **05**:01. . [Citation] [PDF] [PDF Plus]
11. CHIA-CHI CHANG, HUNG-YI HSU, TZU-CHIEN HSIAO. 2013. QUANTITATIVE NON-STATIONARY ASSESSMENT OF CEREBRAL HEMODYNAMICS BY EMPIRICAL MODE DECOMPOSITION OF CEREBRAL DOPPLER FLOW VELOCITY. *Advances in Adaptive Data Analysis* **05**:01. . [Citation] [PDF] [PDF Plus]
12. Wei Guo, Peter W. Tse. 2013. A novel signal compression method based on optimal ensemble empirical mode decomposition for bearing vibration signals. *Journal of Sound and Vibration* **332**, 423-441. [CrossRef]

13. Shou-Zen Fan, Qin Wei, Pei-Fung Shi, Yu-Jen Chen, Quan Liu, Jiann-Shing Shieh. 2012. A comparison of patients' heart rate variability and blood flow variability during surgery based on the Hilbert–Huang Transform. *Biomedical Signal Processing and Control* **7**, 465-473. [CrossRef]
14. GANG WANG, XIAN-YAO CHEN, FANG-LI QIAO, ZHAOHUA WU, NORDEN E. HUANG. 2010. ON INTRINSIC MODE FUNCTION. *Advances in Adaptive Data Analysis* **02**:03, 277-293. [Citation] [References] [PDF] [PDF Plus]