

Optimizing Approach for Sifting Process to Solve a Common Type of Empirical Mode Decomposition Mode Mixing

Saad Al-Baddai, Karema Al-Subari, Elmar Lang, Bernd Ludwig

Abstract—Empirical mode decomposition (EMD), a new data-driven of time-series decomposition, has the advantage of supposing that a time series is non-linear or non-stationary, as is implicitly achieved in Fourier decomposition. However, the EMD suffers of mode mixing problem in some cases. The aim of this paper is to present a solution for a common type of signals causing of EMD mode mixing problem, in case a signal suffers of an intermittency. By an artificial example, the solution shows superior performance in terms of cope EMD mode mixing problem comparing with the conventional EMD and *Ensemble Empirical Mode decomposition* (EEMD). Furthermore, the over-sifting problem is also completely avoided; and computation load is reduced roughly six times compared with EEMD, an ensemble number of 50.

Keywords—Empirical mode decomposition, mode mixing, sifting process, over-sifting.

I. INTRODUCTION

ROUGHLY a decade ago, an empirical nonlinear analysis tool for complex, non-stationary temporal signal variations has been introduced by N. E. Huang et al. [1]. Afterwards, such techniques are commonly referred to as *Empirical Mode Decomposition* (EMD), and if combined with Hilbert spectral analysis they are called *Hilbert - Huang Transform* (HHT). They adaptively and locally decompose any non-stationary signal in a sum of Intrinsic Mode Functions (IMF) which represent zero-mean, amplitude and (spatial-) frequency modulated components. EMD represents a fully data-driven, unsupervised signal decomposition which does not need any *a priori* defined basis system. Since EMD is fully data-driven, not mathematical-based and only defined as the extracted components of an iterative algorithm, it is an open question to know what sort of separation can (or cannot) be performed for two-signals or more composite signals when using the method. Other than competing *Exploratory Matrix Factorization* (EMF) techniques like Independent Component Analysis (ICA) [2], [3], EMD also satisfies the perfect reconstruction property, i.e. superimposing all extracted IMFs together with the residual slowly varying trend reconstructs the original signal without information loss or distortion. Thus EMD lacks the scaling and permutation indeterminacy familiar from blind source separation techniques [4]. Because EMD operates on sequences of local extremes, and the decomposition is carried out by direct extraction of the local energy associated with the intrinsic time scales of the signal itself, the method is thus similar to traditional Fourier or

Wavelet decompositions. It differs from the wavelet-based multi-scale analysis, however, which characterizes the scale of a signal event using pre-specified basis functions. Owing to this feature, EMD, and even more so its noise-assisted variant called *Ensemble Empirical Mode decomposition* (EEMD), is highly promising in dealing with other problems of a multi-scale nature. But the interpretation of IMFs is not straightforward, and it is still a challenging task to identify and/or combine extracted IMFs in a proper way so as to yield physically meaningful components. However, one can find more details about EMD in [5]–[7].

The goal of this paper is therefore to contribute a better method and to improve experimentally extracted modes depend on the original method. This work shows the probable solution for a common type of EMD mode mixing which produces by using conventional EMD.

This paper is organized as follows: A background about standard EMD algorithm is shortly introduced. Then the mode mixing problem in Section II is explained. Afterwards, Section III introduces the improvements and solutions of EMD mode mixing problem. Finally, a short conclusion is drawn.

II. EMPIRICAL MODE DECOMPOSITION

EMD is a fully data-driven method for the different scales analysis of complex, nonlinear and non-stationary real-world signals. It decomposes the original signal into a finite set of amplitude-modulated (AM) components, which are called *Intrinsic Mode Functions* (IMFs). IMFs represent zero-mean amplitude and frequency modulated components. The EMD represents a fully data-driven, unsupervised signal decomposition and does not need any *a priori* defined basis system. EMD also assures perfect reconstruction, i.e. superimposing all extracted IMFs together with the residual trend reconstructs the original signal without information loss or distortion. The empirical nature of EMD offers the advantage over other signal decomposition techniques like *Exploratory Matrix Factorization* (EMF) [8] of not being constrained by conditions which often only apply approximately. Especially with cognitive signal processing, one often has only a rough idea about the underlying modes or component images, and frequently their number is unknown.

Eventually, the original signal $x(t)$ can be expressed as

$$\begin{aligned}
 x(t) &= \sum_j c^{(j)}(t) + r(t) \\
 c^{(j)}(t) &= \text{Re} \{ a_j(t) \exp(i\phi_j(t)) \} \\
 &= \text{Re} \left\{ a_j(t) \exp \left(i \int_{-\infty}^t \omega_j(t') dt' \right) \right\} \quad (1)
 \end{aligned}$$

where the $c^{(j)}(t)$ represents the IMFs and $r(t)$ the remaining non-oscillating trend. Furthermore, $a_j(t)$ denotes a time-dependent amplitude, $\phi_j(t) = \int \omega_j(t) dt$ represents a time-dependent phase and $\omega_j[\text{rad/s}] = \frac{d\phi_j(t)}{dt}$ denotes the related instantaneous frequency. Plotting both amplitude $a_j(t)$ and phase $\phi_j(t)$ as a function of time for each extracted IMF represents a *Hilbert - Huang spectrogram* [9].

During sifting, mode mixing as well as boundary artifacts can be avoided by a variant called *Ensemble Empirical Mode Decomposition* (EEMD) which has been introduced by [10]. It represents a noise-assisted data analysis method. First white noise of finite amplitude is added to the data, and then the EMD algorithm is applied. This procedure is repeated many times, and the IMFs are calculated as the mean of an ensemble, consisting of the signal and added white noise. With a growing ensemble number, the IMF converges to the true IMF [10]. Adding white noise to the data can be considered a physical experiment which is repeated many times. The added noise is treated as random noise, which appears in the measurement. In this case, the n -th noisy observation will be

$$x_n(t) = x(t) + \epsilon_n(t) = \sum_j c_n^{(j)}(t) + r_n(t), \quad (2)$$

where $x(t)$ is the true signal, $\epsilon_n(t)$ is the random noise and $c_n^{(j)} = c^{(j)} + \epsilon_n(t)$ represents the IMF obtained for the n -th noise observation. For the sake of simplicity, following we denote the residuum as $r_n(t) \equiv c_n^{(j)}(t)$, hence formally include it into the summation over the IMFs.

III. SIMULATION RESULTS

A. Mode Mixing

As shown by Huang et al. [1], [10] in many times the mode mixing is a consequence of signal intermittency. The intermittency leaves its fingerprint in the extracted modes which leads to a serious distortion the time-frequency distribution. In addition, it also makes the individual IMF entirely lacking of physical meaning. To reduce this kind of mode mixing problem, Huang et al. [1] suggested an approach to intermittence test, which can cope some of the difficulties. However, the approach has some problems because of scaling separation. Afterward, a new approach was introduced by Wu et al. [11], which is called *Ensemble Empirical Mode Decomposition* (EEMD). It represents a noise-assisted data analysis method. First white noise of 0.1 amplitude of the standard deviation of the original signal is added to the data, and then the EMD algorithm is applied. This procedure is repeated many times, and the IMFs are calculated as the mean of an ensemble, consisting of the signal and added white noise. With a growing ensemble number, the IMF

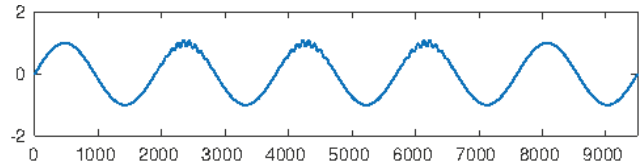


Fig. 1 An artificial signal obtained as the superposition of the waveforms and an intermittent signal according to (3)

converges to the true IMF [10], [11]. Adding white noise to the data can be considered a physical experiment which is repeated many times. The added noise is treated as random noise, which appears in the measurement. This technique is based on the studies of the statistical properties of white noise [12], [13], which showed that the EMD is effectively an adaptive dyadic filter bank when applied to white noise. Although this approach has been succeed to alleviate this kind of problem, but an additional efforts must be made to choice the meaningful modes. Furthermore, this leads to over-sifting problem. To see the effect of EEMD, an intermittent signal is generated as (see Fig. 1):

$$\begin{aligned}
 S &= 0.1 * \sin(2 * \pi * t + 5 * \sin(\pi * t/100)) * \\
 &\quad (\exp(-(t - 25).^2/10) + \exp(-(t - 45).^2/10) \\
 &\quad \quad + \exp(-(t - 65).^2/10)) \\
 &\quad \quad + \sin(\pi * t/10) \quad (3)
 \end{aligned}$$

Separation combining of these signals using conventional EMD has failed as presented in Fig. 2 because of relatively small amplitude of intermittent signal compared to the pure signal and disappearing it periodically. From Fig. 2 (a) one can see clearly the effect of mode mixing in IMF1 and IMF2. IMF1 is the mixture of both the low frequency fundamental and the high frequency intermittent waves, this make it difficult to interpret and determine the underlying physical meaning. Beside, one can see in Fig. 2 (b) the needing to a bit effort to recognize on the correct extracted modes by EEMD, IMF7 and IMF5, which has a physical meaning and interpretable.

B. Mode Mixing Solution

The solution for this kind of mode mixing can be done by optimizing the sifting process as the following:

- 1) Normalize input signal $S = [s_1, s_2, \dots, s_n] \rightarrow nS = (s - \min(s))/(\max(s) - \min(s))$
- 2) Add a helping white noise $\epsilon(x)$ with a small amplitude $nS = \frac{1}{A}\epsilon(x) + nS$, where A should be in the range, rule of thumb, between 1000 and 9000
- 3) Initialize $r_1(n) = nS$ and set $j = 1$;
- 4) Extract j th IMF by the following sifting process (SP) procedure:
 - a. Set the intermediate signal $h_j(n) = r_j(n)$ and $i = 0$ and $k = N - j$; where N represents the number of modes.
 - b. Extract the local extremes (maxima and minima) of $h_j(n)$;
 - c. Generate upper and lower envelopes $u_j^+(n)$ and $u_j^-(n)$, based on the generated extremes of $h_j(n)$;
 - d. Take the average of envelopes $m_j(n) = \frac{1}{2}(u_j^+(n) + u_j^-(n))$;

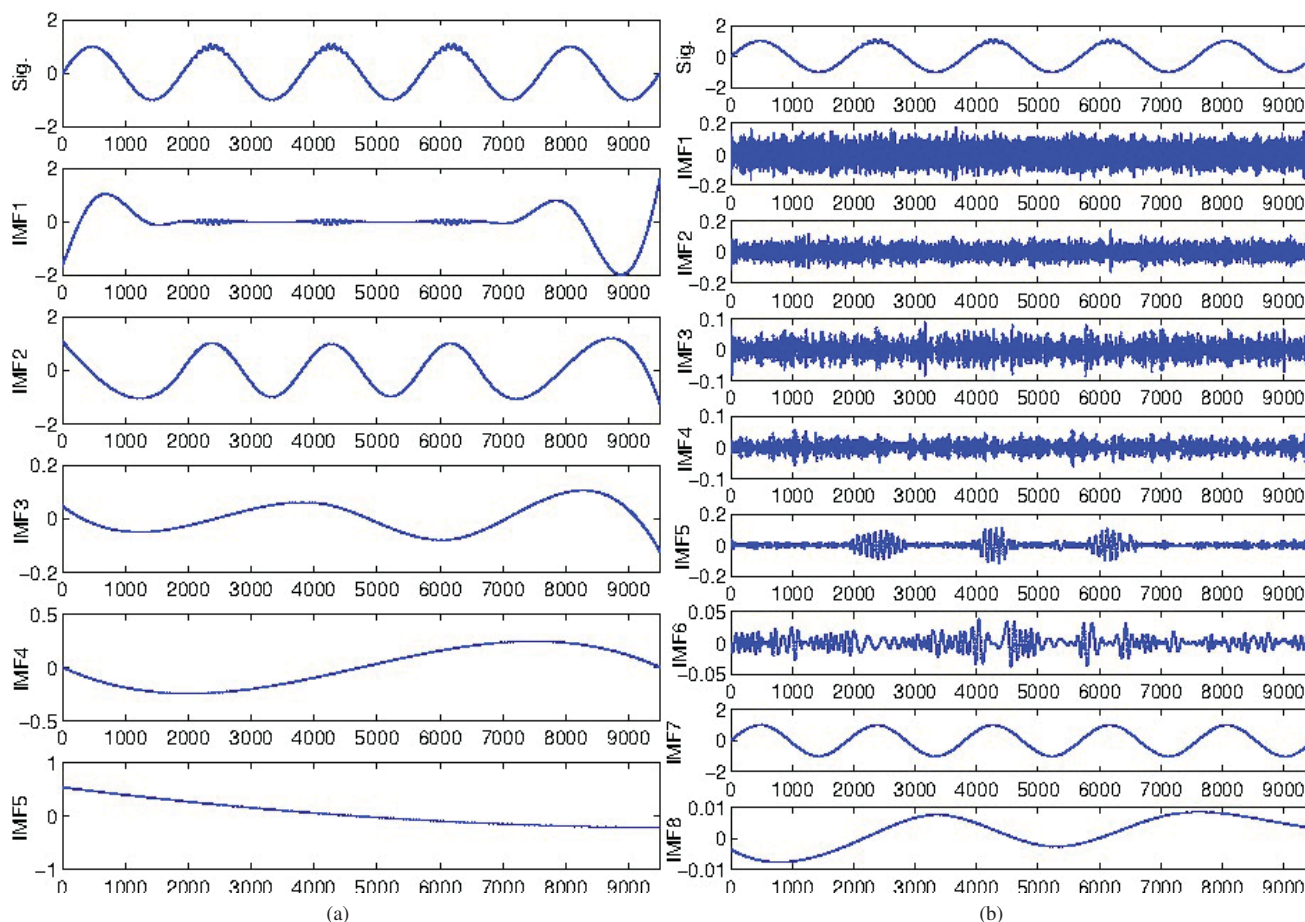


Fig. 2 (a) the intermittent signal and extracted modes by conventional EMD decomposition with 10 sift iteration. (b) the intermittent signal and extracted modes by EEMD, an ensemble member of 50 is used, 10 sift iteration is used, and the added white noise in each ensemble member has a standard deviation of 0.1

- e. Update $h_j(n) = h_j(n) - m_j(n)$ and $i = i + 1$;
- f. Repeat, steps (b) to (e) until stopping criteria (SD) is met, $SD_i \geq A \times N$ and $j = 1$; obtain the first noisy mode (IMF1), $c_1(n) := h_1(n)$;
- 5) Update $r_{j+1}(n) = r_j - c_j(n)$;
- 6) Update stopping criteria (SD), $SD_i \geq \text{round}(\log(k^{2*k^2})) + 1$ and $j \neq 1$;
- 7) Repeat steps from 3 to 6; obtain the rest of j IMFs; repeat till a predned IMF index is met, i.e $j = J$

As shown on Fig. 2 the standard method completely failed to extract correct components. Besides, EEMD extracted the interesting modes but suffered of over-sifting problem, still. Hence according to algorithm in Fig. 3, after decomposition process with original signal we have the following: when a white noise with a tiny amplitude is added without updating the number of sifting iteration (step 6 of the proposed improvements), EMD successes to estimate modes rather correctly, but it fails to show up the added noise. Also the over-sifting is appeared because the intermittency is repeated in IMF1 and IMF2. While applying the whole proposed method, see Fig. 3 (b), clearly succeed to obtain modes perfectly matches with the original one; the added noise in IMF1, the intermittent signal in IMF2 and the pure signal in

IMF3. So the diapason between the amplitude of the added noise and the sifting iterations is required as we suggested to solve this kind of problem. In addition to the performance of this proposed improvements in coping mode mixing problem compared to EEMD-50, the computation load is reduced roughly six times.

IV. CONCLUSION

This paper addressed a common type of EMD mode mixing problem and specialize existing knowledge of EMD algorithm performance. Hence, this solution can cope such mode mixing problem without over-sifting problem and save much computation load as well. However, the aim of presented method is to extend scope field for research and application.

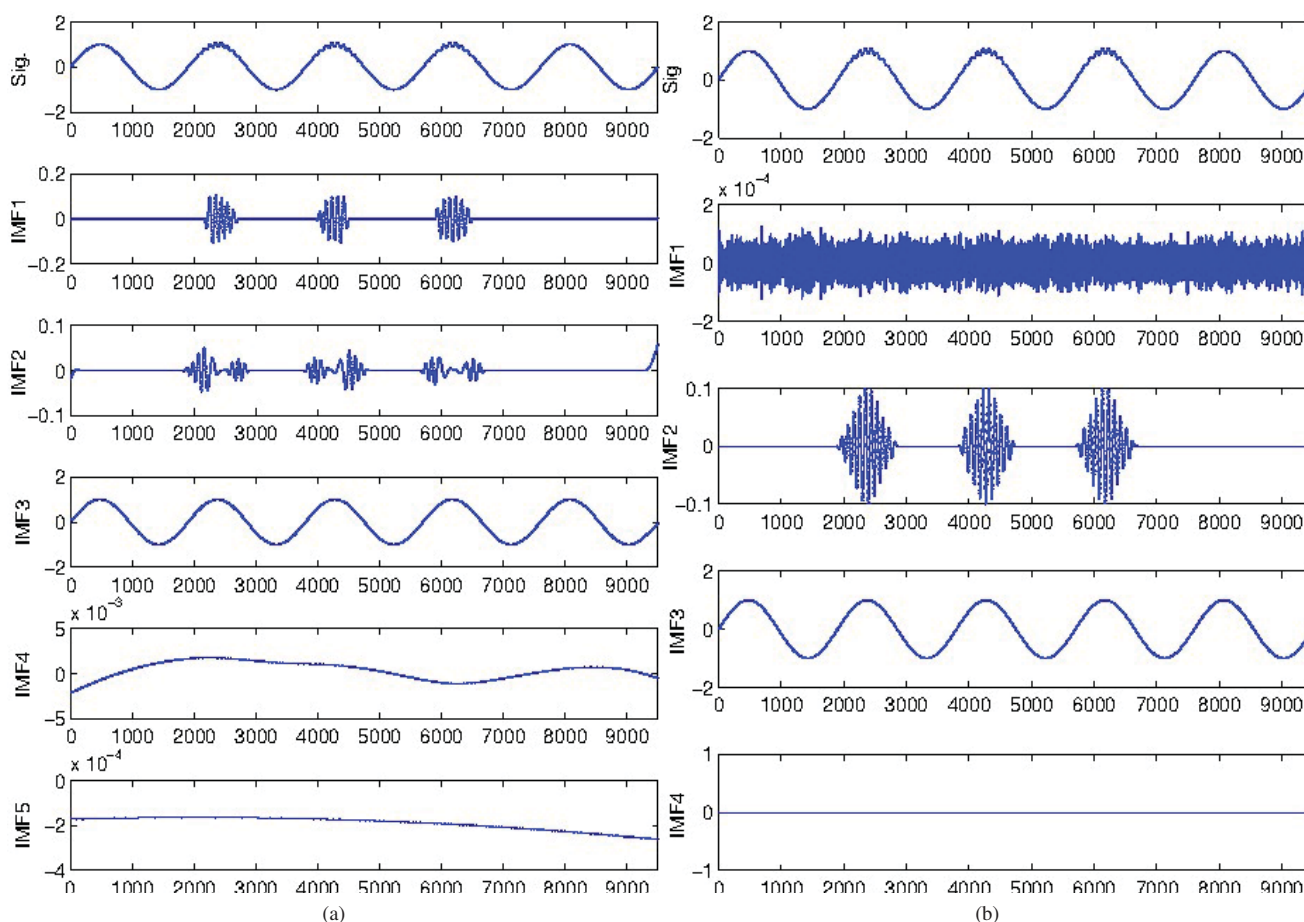


Fig. 3 (a) the signal and extracted modes by EMD algorithm decomposition with a proposed tiny assisted noise, the added white noise has a standard deviation of 0.0001. (b) the signal and extracted modes by proposed EMD algorithm decomposition

REFERENCES

- [1] N. E. Huang, Z. Shen, S. R. Long, M. L. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis," *Proc. Roy. Soc. London A*, vol. 454, pp. 903–995, 1998.
- [2] V. D. Calhoun, T. Adali, L. K. Hansen, J. Larsen, and J. J. Pekar, "ICA of functional MRI data: An overview," in *Proceedings of the International Workshop on Independent Component Analysis and Blind Signal Separation*, 2003, pp. 281–288.
- [3] A. Cichocki, S. Amari, K. Siwek, T. Tanaka, and A. H. P. et al., "ICALAB Toolboxes," 2007. (Online). Available: <http://www.bsp.brain.riken.jp/ICALAB>.
- [4] P. Common and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and its Applications*. Academic Press, 2010.
- [5] S. Al-Baddai, K. Al-Subari, A. Tomé, J. J. Solé-Casals, and E. Lang, "A green's function-based bi-dimensional empirical mode decomposition," *Information Sciences*, vol. 348, pp. 305–321, 2016.
- [6] K. Al-Subari, S. Al-Baddai, A. Tomé, G. Volberg, R. Hammwöhner, and E. Lang, "Ensemble empirical mode decomposition analysis of EEG data collected during a contour integration task," *PLoS ONE*, vol. 10, no. 4, p. e0119489, 04 2015.
- [7] K. Al-Subari, S. Al-Baddai, A. Tomé, M. Goldhacker, R. Faltermeier, and E. Lang, "Emdlab: a toolbox for analysis of single-trial eeg dynamics using empirical mode decomposition," *Journal of Neuroscience Methods*, vol. 253C, pp. 193–205, 07 2015.
- [8] E. W. Lang, R. Schachtner, D. Lutter, D. Herold, A. Kodewitz, F. Blöchl, F. J. Theis, I. R. Keck, J. M. G. Sáez, P. G. Vilda, and A. M. Tomé, *Exploratory Matrix Factorization Techniques for Large Scale Biomedical Data Sets*. Bentham Science Publishers, 2010.
- [9] N. Attoh-Okine, K. Barner, D. Benti, and R. Zhang, "The Empirical Mode Decomposition and the Hilbert-Huang Transform," *EURASIP J. Advances in Signal Processing*, 2008.
- [10] Z. Wu and N. E. Huang, "Ensemble Empirical Mode Decomposition: a noise-assisted data analysis method," *Adv. Adaptive Data Analysis*, vol. 1(1), pp. 1–41, 2009.
- [11] Z. Wu, N. E. Huang, and X. Chen, "The Multidimensional Ensemble Empirical Mode Decomposition Method," *Adv. Adaptive Data Analysis*, vol. 1, pp. 339–372, 2009.
- [12] P. Flandrin, G. Rilling, and P. Goncalves, "Empirical mode decomposition as a filter bank," *Signal Processing Letters, IEEE*, vol. 11(2), pp. 112–114, 2004.
- [13] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," in *Proceedings of the Royal Society*, vol. 460, 2004, pp. 1597–1611.