Empirical mode decomposition (EMD) is an adaptive and data-driven approach for analyzing multicomponent nonlinear and non-stationary signals. The mode-mixing problem is one of the most important topics for the improvement of the EMD algorithm. In this paper, we study the reasons of the mode mixing phenomenon. And then, we propose a new method to resolve this problem relying on the assumption that each IMF should be locally orthogonal to the others. We experiment on several signals, including simulated and real life signals, to demonstrate the efficacy of the proposed method to resolve the mode mixing problem.

Keywords: Empirical mode decomposition (EMD); intrinsic mode function (IMF); mode

*The corresponding author of this paper.
mixing phenomenon, locally orthogonal.

AMS Subject Classification: 22E46, 53C35, 57S20

1. Introduction

Single channel signal separation has attracted a great deal of attention in recent years since it has affected many applications. Many approaches, such as Ref. 1, 2, 3, 4, have been proposed to separate a single channel signal into a mixture of several additive coherent subcomponents. Among those single channel signal separation approaches, the Empirical Mode Decomposition (EMD), proposed by Huang et al.\textsuperscript{1}, has attracted much attention owing to its appealing features. It models the signal to be analyzed as a series of Intrinsic Mode Functions (IMFs) plus a remainder trend signal. Since the physical meaning underlying the complex signal can often be preserved in the decomposed IMFs, the IMFs have always been applied to a broad range of such fields as mechanic vibration and image analysis.\textsuperscript{5,6,7}

To obtain an IMF, EMD uses a sifting procedure based on the estimated upper and lower envelope, which are interpolated from the local extremum points of the input signal. Those extremum points are so susceptible to the noise and intermittency, etc, that lead to the extracted IMFs often enclosed some defects, which prevent the further applications of EMD. One of these problems is the mode mixing phenomenon, as firstly noted by Huang et al\textsuperscript{5}, which means, at a particular time scale, the component either coming into existence or disappearing from a signal entirely. Furthermore, Wu and Huang\textsuperscript{8} define the mode mixing problem as a single IMF either consisting of signals of widely disparate scales, or a signal of a similar scale residing in different IMF components. In order to overcome the mode mixing problem, many solutions have been proposed and we will present a brief review about them in Section 2.

In this work, we firstly analyze the reasons for causing the mode mixing phenomenon. In order to resolve the mode mixing problem, for instance, a signal of a similar scale residing in different IMFs, we assume that each extracted IMF should be locally orthogonal to the others and the residue. This assumption means that, at any time $t$, different IMFs should have different instantaneous frequencies. Based on this assumption, we propose a back projection strategy to implement the locally orthogonal condition and resolve the mode mixing problem in both intermittent signals and AM-FM signals. To evaluate the performance of our proposed algorithm, we compare the separation results derived by our algorithm and the Ensemble EMD (EEMD) algorithm on some simulated and real life signals in the experimental part.

The remainder of this paper is organized as follows. In Section 2, we review the EMD algorithm and some related works on the mode mixing problem. In Section 3, we propose our algorithm for resolving the mode mixing problem using the back projection strategy. And then, some numerical results of experiments on simulated and real-life signals have been presented in Section 4. Finally, Section 5 contains some concluding remarks.
2. EMD and Mode Mixing

Here, we will present EMD method in brief. The more precise, implementation details and performance of the algorithm can be referred to Ref. 1, 9, 10, 11, 12.

2.1. IMF Definition and Sifting Process

The essence of EMD algorithm is to identify the intrinsic oscillatory mode which is defined as a function that satisfies two conditions: (1) in the whole data set, the number of extrema and the number of zero-crossings must either be equal or differ, at most, by one and (2) at any point, the local average of upper and lower envelopes is zero. Based on these two conditions, Huang et al. designed a sifting process to decompose the given signal into multiple IMFs. The first IMF $h_1(t)$, of a real valued signal $x(t)$, is acquired by the following iteration process.

1. Find all local minima and maxima of $x(t)$.
2. Interpolate (using a cubic spline) between minima and maxima to obtain the lower and upper envelope, as $e_{\min}(t)$ and $e_{\max}(t)$, respectively.
3. Compute the mean envelope $m(t) = (e_{\min}(t) + e_{\max}(t))/2$.
4. Let $i = i + 1$, subtract from the signal to obtain $p_i(t) = x(t) - m(t)$.
5. Repeat steps 1) to 4) on $p_i(t)$ until it is an IMF, then record the IMF $h_1(t) = p_i(t)$.

Since, in each iteration, this procedure removes the subcomponent with the lower frequency, the first extracted IMF $h_1(t)$ is the subcomponent with the highest frequency of $x(t)$. The sifting process can then be performed on the residual $x_1(t) = x(t) - h_1(t)$, iteratively. So, for any one-dimensional discrete signal $x(t)$, EMD can finally present it as:

$$x(t) = \sum_{k=1}^{K} h_k(t) + r(t).$$

(2.1)

2.2. The Mode Mixing Phenomenon

Mode mixing is a major obstacle on the use of EMD to successfully decompose signals. The primary reason for causing mode mixing is the intermittency of the detected extrema, which means the extrema detected in the sifting process belong to the different subcomponents. For example, considering a two-component signal $s(t) = s_h(t) + s_l(t)$, with $s_h(t)$ and $s_l(t)$ denote the subcomponents with high frequency and low frequency, respectively. To decompose the signal $s(t)$ by the EMD, we need to interpolate the extrema of $s(t)$ to acquire the upper and lower envelope. If the high frequency subcomponent $s_h(t)$ is a intermittent signal, or, in some part of $s_h(t)$, its energy is extremely smaller than the energy of $s_l(t)$, we can only detect the extrema belonging to the $s_l(t)$ from $s(t)$. However, in other parts of $s(t)$, the sifting procedure can detect the extrema of $s_h(t)$. Under this situation, the
extracted IMF is a signal comprised of different scales; some parts are from signal \( s_h(t) \) and the other parts are from signal \( s_l(t) \).

Up to now, several solutions have been proposed to solve the mode mixing problem. One kind of these solutions, for example in Ref. 5, 13, changes the choice of extrema, which is based on some intermittency test, for the envelope in the sifting process. For detecting the intermittency in the extracted IMF, some frequency tracking methods such as Hilbert transform or T-K energy operator\(^{14}\) have been used. However, for complicated multicomponent amplitude and frequency modulated signals, those approaches using some single intermittency test criteria to detect intermittency are rather difficult. And hence, the other kind of approaches using masking signals has appeared. Deering and Kaiser\(^{15}\) adopted a pure sine wave, whose frequency is calculated using an energy weighted mean method, as the masking signal to avoid the mode mixing for two pure sine wave signals. The approach, obviously, is not suitable for the frequency modulated (FM) signals. Recently, the Ensemble EMD (EEMD) algorithm\(^{8}\) has been proposed, which uses a large number of random noise signals as the masking signals to overcome the mode mixing phenomenon. This approach, however, depends much on the input parameters and lacks of theoretical analysis.

All the approaches reviewed above aim to solve the mode mixing problem caused by the intermittent signal. However, the mode mixing may still occur in the extracted IMFs when the input signal does not contain any intermittent components. Furthermore, none of the before-mentioned algorithms uses the information of the residual signal and its relation to the extracted IMF.

3. Back Projection method for Mode Mixing

As described in the previous section, although the mode mixing is mainly caused by the intermittency in signals, the intermittency is not the only reason. Figure 1 illustrates such a mode mixing example which is not caused by the intermittency. The input signal and its two subcomponents are shown in Figure 1(a), (b) and (c), respectively. The first extracted IMF by the EMD is shown in Figure 1(d). Obviously, the IMF has mode mixing problem since it contains segments of different scales from different signals. And also, the frequency tracked from this extracted IMF will jump at those time segments \( S_i, i = 1, 2, 3, 4 \) begins and ends. The reason of this kind of mode mixing is that the EMD will detect extremum points belonged to the different subcomponents because of the large energy difference between those subcomponents. In this example, as shown in Figure 1(a), the extrema of the signal \( s_1(t) \) is totally immersed in the signal \( s_2(t) \) in time segments from \( S_1 \) to \( S_4 \), which directly leads to the result shown in Figure 1(d).

From Figure 1, we can find that the subcomponent \( s_2(t) \) (in Figure 1(c)) resides in both the extracted IMF (in Figure 1(d)) and the residual signal (in Figure 1(e)). And hence, we propose a locally orthogonal constraint to overcome it, which assumes that each extracted IMF should be locally orthogonal to the others and the residue.
In order to analyze and exhibit the proposed algorithm conveniently, we give the definition of \textit{locally orthogonal} as follows.

\textbf{Definition 3.1.} For two signals, $s_1(t)$ and $s_2(t)$, \textit{locally orthogonal} can be defined as, for each point $\tau_i$, in its neighborhood $[\tau_{i,l}, \tau_{i,r}]$, the inner product of $s_1(t)$ and $s_2(t)$ is zero, which is \(\langle s_1(t), s_2(t) \rangle = 0\) where $t \in [\tau_{i,l}, \tau_{i,r}]$.

The \textit{locally orthogonal} condition means that, in every time intervals, the frequencies of $s_1$ and $s_2$ should be different. However, in real signals, neither the instantaneous frequency nor the width of neighborhood can be estimated accurately and chosen suitably. Therefore, we use Gabor transform to approximate this locally orthogonal definition. With the help of Gabor transform, signals $s_1(t)$ and $s_2(t)$ being \textit{locally orthogonal} equals to \(|\mathcal{G}(s_1(t))| \cdot |\mathcal{G}(s_2(t))| = 0\), where $\bullet$ denotes the
entrywise product and $G$ denotes the Gabor transform.

We use a multicomponent AM-FM signal to model the input signal as

$$s(t) = \sum_{i=1}^{N} s_i(t) = \sum_{i=1}^{N} a_i(t) \cos(\phi_i(t)), \quad (3.1)$$

with subcomponents’ instantaneous frequencies (IF) $f_i(t) = d\phi_i(t)/dt$ and assuming that $f_i(t) > f_j(t)$ for any $i < j$. We further assume that, after $n$th sifting iterations, the IMF $h(t)$ and residual $r(t)$ have been separated from the input signal $s(t)$. Since EMD extracts subcomponent with the highest frequency first, the frequency of $h(t)$ should be higher than the residual $r(t)$. However, because of the mode mixing problem, $h(t)$ may contain more than one monocomponent signals. Therefore, we can model $h(t)$ and $r(t)$ as

$$h(t) = \tilde{h}(t) + \tilde{v}(t) = s_1(t) + \sum_{i=2}^{k} \alpha_i(t)s_i(t), \quad (3.2)$$

$$r(t) = \tilde{u}(t) + \tilde{r}(t) = \sum_{i=2}^{k} (1 - \alpha_i(t))s_i(t) + \sum_{i=k+1}^{N} s_i(t), \quad (3.3)$$

where $\tilde{h}(t)$ denotes the desired IMF and the frequency of $\tilde{v}(t)$ is higher than the frequency of $\tilde{u}(t)$. The function $\alpha_i(t)$ in Eq. (3.2) is a smooth varying function which represents that, in some time segments, some subcomponents with lower frequencies may retain in the extracted IMF. This may be caused by the time intermittency or large energy difference, etc. Under this situation, apparently, the extracted IMF $h(t)$ and the residual $r(t)$, for example as shown in Figure 1(d) and (e), respectively, are not locally orthogonal since $\tilde{v}(t)$ and $\tilde{u}(t)$ have some common subcomponents. This can be seen more clear in Figure 1(f), which shows the Gabor transform of $h(t)$ (in color red) and $r(t)$ (in color green), and then superpose them together. From which, in the time segments from $S_1$ to $S_4$, the color trends to be yellow, which means that $h(t)$ and $r(t)$ have some common components in those regions.

In order to acquire the correct monocomponent $\tilde{h}(t)$, we project the extracted IMF $h(t)$ back to the residue $r(t)$ and compute the locally orthogonal component to the $r(t)$ as:

$$y(t) = G^{-1}(G(h(t)) \bullet \mathcal{P}(\mathcal{G}(r(t)); T)), \quad (3.4)$$

where $G^{-1}$ denotes the inverse Gabor transform, $\mathcal{P}$ is a thresholding function with $\mathcal{P}(x; T) = 0$ when $x > T$ and otherwise $\mathcal{P}(x; T) = 1$. After the back projection, we can ensure that the derived $y(t)$ is locally orthogonal to the residual signal. However, because of the scale and resolution limits of Gabor transform, this method can not ensure that $y(t) = s_1(t)$. And hence, using $y(t)$ to replace $\tilde{h}(t)$ directly is inappropriate. On the other hand, albeit we can not ensure that $y(t)$ equals to $\tilde{h}(t)$,
we can assume that \( y(t) = \alpha \tilde{h}(t) + \beta \tilde{v}(t) \), where \( \alpha \approx 1 \) and \( \beta \approx 0 \). Then, if we add \( \gamma y(t) \) to the IMF \( h(t) \), we will have

\[
    h_+(t) = h(t) + \gamma y(t) = (1 + \gamma \alpha) \tilde{h}(t) + (1 + \gamma \beta) \tilde{v}(t).
\] (3.5)

Since \( \alpha \approx 1 \) and \( \beta \approx 0 \), we can have \( h_+(t) \approx (1 + \gamma) \tilde{h}(t) \) and \( \tilde{v}(t) \). This means that we can use \( \gamma \) as the energy factor, which can be estimated by the energy of \( \tilde{h}(t) \) and \( \tilde{v}(t) \). With the help of \( \gamma \), we can use EMD to extract \( (1 + \gamma \alpha) \tilde{h}(t) \) from \( h_+(t) \). Furthermore, we can derive the EMD on the signal \( h_-(t) = h(t) - \gamma y(t) \) to extract \( (1 - \gamma \alpha) \tilde{h}(t) \) to cancel out the energy factor \( \gamma \) and derive the final desired IMF \( \tilde{h}(t) \). We summarize the previous discussions into the Algorithm 1 for solving the mode mixing problem.

**Algorithm 1** Back Projection Algorithm

1: Input signal \( s(t) \), perform EMD on \( s(t) \) to obtain the IMF \( h(t) \) and residual \( r(t) \).
2: Construct the locally orthogonal signal of \( r(t) \) as \( y(t) \) using (3.4).
3: while \(|y(t)| \neq 0\) do
   4: Perform sifting operation on \( h_+(t) = h(t) + \gamma y(t) \) and \( h_-(t) = h(t) - \gamma y(t) \) to obtain the IMF \( \tilde{h}_+(t) \) and \( \tilde{h}_-(t) \), respectively.
   5: Update the IMF \( h(t) = (\tilde{h}_+(t) + \tilde{h}_-(t))/2 \) and \( r(t) = s(t) - h(t) \).
   6: Project \( h(t) \) to \( r(t) \) using (3.4) to get \( y(t) \).
4: end while
8: return The desired IMF \( \tilde{h}(t) = h(t) \).

**Remarks** In most cases, only one time of the projection operation (as described in step 6) is enough for resolving the mode mixing phenomenon. For example, Figure 1 (g) and (h) show the extracted IMF \( \tilde{h}(t) \) after only one time back projection operation and the difference with the real subcomponent \( s_1(t) \), respectively. However, for some intermittency signals or complicated multicomponent signals, we may need to repeat this back projection operation several times to ensure the local orthogonal condition.

### 4. Implementation and Experiment Results

In this section, we consider some implementation issues of the proposed algorithm and present the results of experiments on various signals.

In Eq. (3.4), the threshold \( T \) is the parameter for controlling the local orthogonal constraint between the extracted IMF and the residual signal. Therefore, for some complicated signals, a localized parameter \( T(t, \omega) \) will be better than the global uniformed parameter. The value of localized parameters \( T(t, \omega) \) can be chosen according to the ridges in the Gabor transform. To detect ridges, we can use some similar approaches proposed in Ref. 16, 17 for extracting multiple ridges in
a continuous wavelet transform. But here, we use a more simple method. Firstly, we compute the Gabor transform of the residual signal as \( G_r(t, \omega) \). And then, for each time location \( t_0 \), we compute the local maxima points as \((t_0, \omega_i)\) according to the amplitude of Gabor coefficients. Finally, at time location \( t_0 \), the threshold value \( T(t_0, \omega) \) is chosen as \( 0.05|G(t_0, \omega_k)| \), where \((t_0, \omega_k)\) is the closest local maxima points to \((t_0, \omega)\).

In the cases when the input signal contains intermittent signals, we need to do some modifications in the step 4 of the proposed algorithm. If the desired IMF \( h(t) \) is an intermittent signal, the obtained signal \( y(t) \) should also be an intermittent signal after the projection. For an intermittent signal, however, the energy factor \( \gamma \) is meaningless for the time interval where the signal disappears. And hence, under this situation, we can only perform the EMD on the time interval where the signal \( y(t) \) exists. This kind of example can be found in Figure 2.

In the following, we provide some examples to demonstrate the results achieved by our algorithm. The first example shows that the algorithm can separate a two-component AM-FM signal with one of its subcomponents is an intermittent signal. The simulated input signal and the extracted IMFs derived by our algorithm and the EMD algorithm are shown in Figure 2. The EMD algorithm fails to separate this kind of signal because that the detected extrema belong to the different subcomponents. The extracted IMF and the residue derived by the EMD are shown in Figure 2(e) and (f). In our approach, after one time projection operation, we can separate the two subcomponents well in the time interval from \(-1\) to \(1\), which can be seen in Figure 2(g). However, because the subcomponent \( s_2(t) \) is an intermittent signal, we may need to perform the projection operation several times to get rid of the component belonging to \( s_1(t) \). Figure 2(i) shows the extracted IMF derived by our approach after 7 times of back projection.

And then, we compare the separation results derived by implementing the ensemble EMD (EEMD) algorithm and our proposed algorithm on the airline passenger data obtained from Ref. 18. The data has 144 points corresponding to the 144 months in a 12-year period. The input data and its Gabor spectrogram can be seen in the top row in Figure 3. Since the standard EMD algorithm will encounter difficulties in dealing with this data with noise, we use the ensemble EMD algorithm instead. In the EEMD algorithm, we set the standard deviation of noise to be 2.5 and the ensemble number to be 1000. The left column and right column of Figure 3, from second row to bottom, show the separation results acquired by the EEMD algorithm and our proposed algorithm, respectively. Although the EEMD has a strong power to deal with signals contaminated by noise, our proposed algorithm performs slightly better than it in this example. Both algorithms have similar results in the third extracted subcomponent, which can represent the peak periods for air travel each year. However, for the second extracted subcomponent, the result by using EEMD seems to have some mode mixing problem, while the result acquired by our proposed algorithm is more like a monocomponent AM-FM signal than the EEMD’s result. In order to analyze the frequency properties more clearly,
Back Projection Strategy for Solving Mode Mixing Problem

Fig. 2. An example of mode mixing phenomenon caused by the intermittency in the subcomponent signals. (a) and (b) the input simulated signal $s(t) = s_1(t) + s_2(t)$ and its Gabor spectrogram, respectively. (c) and (d) the two subcomponents $s_1(t)$ and $s_2(t)$, respectively. (e) and (f) the first extracted IMF and the residue derived by the EMD algorithm. (g), (h) and (i) the acquired IMFs of our algorithm after 1 time, 4 times, and 7 times of projection, respectively. (j) the difference between the IMF in subfigure (i) and the real subcomponent $s_2(t)$.

we perform the Gabor transform of the first extracted subcomponent of the EEMD algorithm and our algorithm, and plot their Gabor spectrogram in Figures 4(a) and (b), respectively. Although the first IMF derived by the EEMD and our algorithm are not narrow band signals, if we compare the spectrograms of Figure 4(a) and (b), the pattern of our IMF, as shown in Figure 4(b), is more regular than that of the EEMD’s IMF shown Figure 4(a). The spectrograms of the Gabor transform for the second IMFs derived by the EEMD and our approach are shown in Figure 4(c) and (d), respectively. As shown in Figure 4(c), in the rectangle enclosed by the dotted lines, the IMF of the EEMD has about three distinctive frequencies in the interval from 60 to 120; while our IMF, shown in Figure 4(d) only contains one dominant
Fig. 3. Analyzing the airline passenger data. The first row: the input signal and its Gabor spectrum. The frequency axis is normalized from 0 to 0.5. The left column (from second row to the bottom) are four extracted subcomponents by the EEMD algorithm and the residual signal, respectively. Similarly, shown in the right column (from second row to the bottom) are the subcomponents extracted by our proposed algorithm and the residual signal, respectively. The parameters used by the EEMD algorithm are: $\sigma = 2.5$ and the ensemble number is 1000.

In the third example, we demonstrate the efficacy of separating the phonics structure of a speech signal by our algorithm. The input signal, a digitalized sound of the word, /zai/, at 16KHz digitization, and its Gabor spectrogram are shown in the top row in Figure 5. We compare the separation results of our algorithm with the ensemble EMD (EEMD) algorithm of this example, which can be found in Figure 5 from second row to the bottom. The left column and right column of Figure 5, from second row to the bottom, show the separation results acquired by the EEMD algorithm and our algorithm, respectively. From which, we can find that all the subcomponents derived by our algorithm seem more like monocomponent AM-FM signals than the subcomponents extracted by the EEMD. In order to investigate the frequency properties of the subcomponents derived by different approaches, we display the Gabor spectrogram of the first five subcomponents derived by EEMD algorithm, corresponding to the left column in Figure 6; and the subcomponents derived by our algorithm, corresponding to the right column in Figure 6. The mode
Fig. 4. The spectrograms of the Gabor transform for IMFs shown in Figure 3. (a) and (b) are the spectrograms of the first IMFs derived by the EEMD algorithm and our algorithm, respectively. (c) and (d) are the spectrograms of the second IMFs derived by the EEMD algorithm and our algorithm, respectively. The frequency axis is normalized between 0 and 0.5. Comparing the regions enclosed by the dotted lines, the IMF derived by EEMD is affected by the mode-mixing phenomenon because (c) has three distinctive frequencies in the corresponding interval from 60 to 120.

mixing problem is very clear in the Gabor spectrogram of the IMFs derived by the EEMD algorithm since some signals with a similar scale are reside in many different IMFs; and also, some extracted IMFs consist of widely disparate scales. The reasons for causing mode mixing problem in this example is primarily the large energy difference between different subcomponents, which can be found in the right column in Figure 5 in the time interval from 0.02s to 0.06s. From the right column in Figure 6, we can find that the second to fifth subcomponents derived by our algorithm have their own distinctive frequencies.

5. Conclusion

In this paper, we firstly discuss the reasons for causing the mode mixing phenomenon in the EMD algorithm. As the definition of mode mixing says, when the mode mixing occurs, a signal of a similar scale often resides in different IMF components, which means that the extracted IMFs are not orthogonal. Therefore, based on the assumption that each IMF should have its individual frequency, we use a locally orthogonal projection manner to avoid the mode mixing by keeping each IMF locally orthogonal to the others. Then, several simulated and real life signals have been used to evaluate the efficacy of our proposed algorithm. However, in our method, we rely on the time-frequency techniques to perform the locally orthogonal projection which suffers a high computational complexity. And thus, to find some fast locally orthogonal projection methods requires further study.
Fig. 5. Analyzing the digitalized speech data. The first row: the input speech signal /zài/ and the Gabor spectrogram of the input signal. The frequency axis is normalized from 0 to 0.5. The left column (from second row to the bottom) are five extracted subcomponents by the EEMD algorithm and the residual signal, respectively. Similarly, shown in the right column (from second row to the bottom) are the subcomponents extracted by our proposed algorithm and the residual signal. The parameters used by the EEMD algorithm are: $\sigma = 1.5$ and the ensemble number is 2000.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No.60972126 and 61062011), the Joint Funds of the National Natural Science Foundation of China (U0935002/L05), the Beijing Natural Science Foundation (4102060) and the State Key Program of National Natural Science of China (Grant No.61032007).

References

Fig. 6. The spectrograms of the Gabor transform for IMFs shown in Figure 5. Left column: spectrograms of the Gabor transform for the first five extracted subcomponents by the EEMD algorithm. Right column: spectrograms of the Gabor transform for the first five extracted subcomponents by our proposed algorithm. The frequencies of all the subfigures are normalized between 0 and 0.5. Comparing the second to fifth subcomponents extracted by both algorithms, the IMFs derived by the EEMD is obviously affected by the mode-mixing phenomenon.

7. Z. Yang and L. Yang, A New Bidimensional Empirical Mode Decomposition by us-
14 X. Hu, S. Peng, J. Tan, Weiping Hu


