THE USE OF A MASKING SIGNAL TO IMPROVE EMPIRICAL MODE DECOMPOSITION

Ryan Deering *

Duke University
Department of Mathematics
Box 90320
Durham, NC 27708
trd1@math.duke.edu

James F. Kaiser

Duke University
Department of ECE
Box 90291
Durham, NC 27708
j.kaiser@ieee.org

ABSTRACT

Empirical Mode Decomposition (EMD) provides a new method for analyzing signals from a nonlinear viewpoint. EMD is defined by an algorithm requiring experimental investigation instead of rigorous mathematical analysis. We show that EMD yields its own interpretation of combinations of pure tones. We present the problem of mode mixing and give a solution involving a masking signal. The masking signal method also allows EMD to be used to separate components that are similar in frequency that would be inseparable with standard EMD techniques.

1. INTRODUCTION

Empirical Mode Decomposition(EMD), introduced by Huang [1], is a method for decomposing nonlinear, multicomponent signals. The components resulting from EMD, called Intrinsic Mode Functions (IMFs), each admit an unambiguous definition of instantaneous frequency and amplitude through the Hilbert transform. As discussed in [2, 3], EMD is defined by the algorithm and has **no** analytical formulation. Hence, our understanding of EMD comes from experimental rather than analytical results [2]. In this paper we discuss the problem of intermittency and a new method for dealing with mode mixing. We also discuss the issue of amplitude modulations from pure tones and the implications in evaluating the effectiveness of EMD.

2. EMPIRICAL MODE DECOMPOSITION

By definition an Intrinsic Mode Function (IMF) satisfies two conditions

1. the number of extrema and the number of zero crossings may differ by no more than one, and
2. the local average is zero

where the local average is defined by the average of the maximum and minimum envelopes discussed in the following section. These properties of IMFs allow for instantaneous frequency and amplitude to be defined unambiguously.

2.1. The Sifting Process

In order to obtain the separate components called IMFs, we perform a process call sifting. The goal of sifting is to subtract away the large-scale features of the signal repeatedly until only the fine-scale features remain. A signal $x(t)$ is thus divided into the fine-scale detail, $d(t)$, and the residual, $m(t)$, so $x(t) = m(t) + d(t)$. This detail becomes the first IMF and the sifting process is repeated on the residual, $m(t) = x(t) - d(t)$.

The sifting process requires that a local average of the function be defined. If we knew the components a priori we would naturally define the local average to be the lowest frequency component. Since the goal of EMD is to discover these components, we must approximate the local average of the signal. Huang's solution to finding a local average creates maximum and minimum envelopes around the signal using natural cubic splines through the respective local extrema. The local average is approximated as the mean of the two envelopes.

The first IMF, $y_1(t)$, of a signal, $x(t)$, is found by iterating through the following loop.

1. Find the local extrema of $x(t)$.
2. Find the maximum envelope $e_+(t)$ of $x(t)$ by passing a natural cubic spline through the local maxima. Similarly find the minimum envelope $e_-(t)$ with the local minima.
3. Compute an approximation to the local average, $m(t) = (e_+(t) + e_-(t))/2$.
4. Find the proto-mode function $z_i(t) = x(t) - m(t)$
5. Check whether $z_i(t)$ is an IMF. If $z_i(t)$ is not an IMF, repeat the loop on $z_i(t)$. If $z_i(t)$ is an IMF then set $y_1(t) = z_i(t)$.

The name, sifting, indicates the process of removing the lowest frequency information until only the highest frequency remains. The sifting procedure performed on $x(t)$ can then be performed on the residual $x_1(t) = x(t) - y_1(t)$ to obtain $x_2(t)$ and $y_2(t)$ and repeated until

$$x(n) = \sum_{i=1}^{k} x_i(t). \tag{1}$$

2.2. Instantaneous Frequency

The Hilbert transform, $v(t)$, of a signal $u(t)$ of the continuous variable t is

$$v(t) = \frac{1}{\pi} \mathrm{P} \int_{-\infty}^{\infty} \frac{u(\eta)}{\eta - t} \, \mathrm{d}\eta \tag{2}$$

where P indicates the Cauchy Principle Value integral.

*The author has been supported under NSF-DMS-9983320.

An instantaneous frequency and amplitude can be obtained for every value of t. Following [4] the instantaneous amplitude is defined as

$$A(t) = \sqrt{u(t)^2 + v(t)^2} \tag{3}$$

and the instantaneous phase can be defined as

$$\varphi(t) = \arctan \frac{v(t)}{u(t)}. \tag{4}$$

Instantaneous frequency, $f(t)$, can be derived from phase by

$$\dot{\varphi}(t) = \omega(t) = 2\pi f(t) \tag{5}$$

which simplifies to

$$\omega(t) = \frac{d}{dt}(\arctan \frac{v(t)}{u(t)}) = \frac{u(t)\dot{v}(t) - v(t)\dot{u}(t)}{u(t)^2 + v(t)^2}. \tag{6}$$

The instantaneous frequency and amplitude serve as an alternative to the power spectrum for describing the frequencies that make up a signal.

The discrete Hilbert transform (DHT), $v(i)$, of a signal $u(i)$ is given by $v(i) = \sum_{m=0}^{N-1} h(i-m)u(m)$ where N is even and

$$h(i) = \frac{2}{N} \sin^2(\frac{\pi i}{2}) \cot(\frac{\pi i}{N}). \tag{7}$$

Using a centered difference approximation to the time derivatives in (6) we can approximate

$$\tilde{f}(i) = \frac{u(i)(v(i+1) - v(i-1)) - v(i)(u(i+1) - u(i-1))}{2t_s(u(i)^2 + v(i)^2)} \tag{8}$$

where t_s is the sampling time. Due to the centered difference operator, the frequency $\tilde{f}(i)$ must be corrected by

$$f(i) = f_s \arcsin(\tilde{f}/f_s) \tag{9}$$

where f_s is the sampling rate.

Applying the Hilbert transform directly to a multicomponent signal provides values of $A(i)$ and $f(i)$ which are unusable for describing the signal. The idea of instantaneous frequency and amplitude does not make sense when a signal consists of multiple components at different frequencies. For this reason, Empirical Mode Decomposition and the Hilbert Transform work very well together by first decomposing a signal into single-frequency components and then finding the instantaneous frequency and amplitude of each component.

2.3. Intermittency

As noted in [5] and [6], intermittency is a major obstacle to the use of EMD on many signals. Intermittency for example occurs in turbulent flow or in any signal that is constantly changing such as speech. In this case we refer to intermittency as a component at a particular time scale either coming into existence or disappearing from a signal entirely. Since EMD locally pulls out the highest frequency component as the current IMF, intermittency in a signal means that the frequency tracked by a particular IMF will jump as the intermittent component begins or ends as in figure 2. The situation when an IMF has components of different frequencies due to intermittency is called **mode mixing**. A solution to mode mixing is proposed in [5] in which a change in the choice of extrema for the envelopes limits the scale over which the sifting process allows a component to pass.

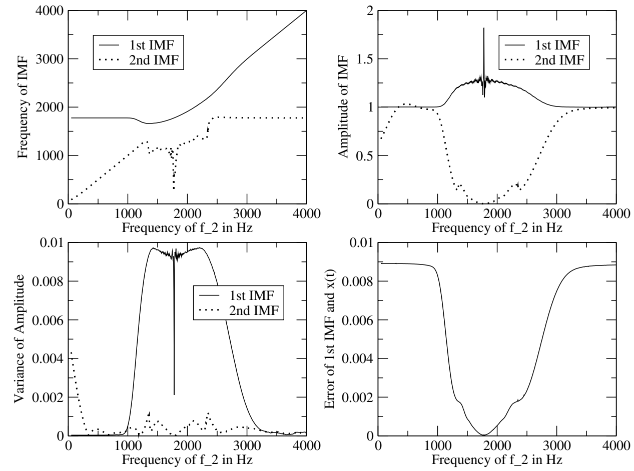

Fig. 1. EMD is performed on the combination of pure tones $x(t) = \sin(2\pi f_1 t) + \sin(2\pi f_2 t)$ where $f_1 = 1776$ Hz and f_2, displayed along the horizontal axes, varies from 50 Hz to 4000 Hz by 6 Hz increments. The top right and top left graphs show the frequencies and amplitudes of the IMFs. The bottom left shows the variance of the amplitude. A greater variance indicates more modulation. The bottom right shows the error between the 1st IMF and the original signal. As f_2 approaches f_1, EMD interprets the entire signal as a single modulated component.

3. PURE TONES AND MODULATION

When considering the performance of EMD on a combination of pure tones, it is important to note that a sum of pure sine waves can often be rewritten as an amplitude modulated (AM) sine wave. The performance of EMD on pairs of pure tones is discussed in [2]. When two sine waves are added together as

$$x(t) = a_1 \sin(2\pi f_1 t) + a_2 \sin(2\pi f_2 t), \tag{10}$$

EMD will decompose the signal with some error due to error in the extrema due to sampling. In [2], it is noted that when $0.5 < f_1/f_2 < 2$ the error between the first IMF and the high frequency component increases dramatically. Calling this error is a bit misleading. Saying that EMD does not perform well because it finds AM tones is as valid as saying that a Discrete Fourier Transform (DFT) does not perform well on a modulated tone because it returns a sum of pure tones. Without a context there is no correct answer as to which interpretation is right. In many problems involving natural phenomena, the modulated interpretation may be the most appropriate, so EMD is more likely to find components that are linked to the underlying process generating the signal.

The question at hand is how EMD interprets combinations of sine waves. In certain cases it is natural to write the sum of two sine waves as an amplitude modulated sine wave. The most basic of trigonometric identities bears this out as $\sin(2\pi xt) + \sin(2\pi yt) = 2\sin(\pi(x+y))\cos(\pi(x-y))$. If $x = 650$ and $y = 600$, then the sum appears as a 625 Hz signal which is amplitude modulated by a 25 Hz sine wave. The AM interpretation is clearly the most natural. If, on the other hand, $x = 50$ and $y = 600$ then the obvious interpretation is to consider the signal to be two separate unmodulated components. EMD interprets both of these signals appropri-

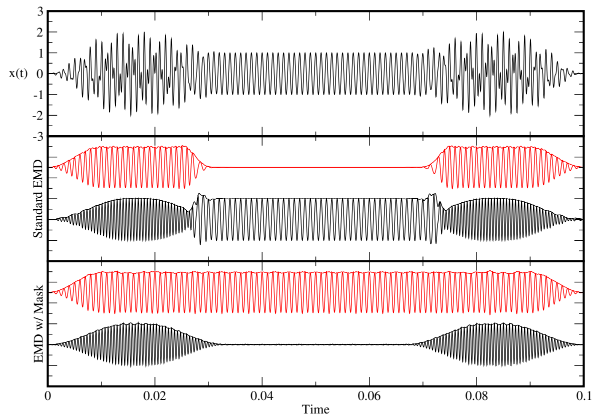

Fig. 2. The signal $x(t)$ consists of a 1776 Hz sine wave that is off between $t = 0.033$ and $t = 0.067$ and a 1000 Hz sine wave. The middle two signals are the IMFs computed with standard EMD. The bottom IMF of the middle pair shows mode mixing as the frequency jumps from 1776 Hz to 1000 Hz when the higher frequency signal is not present. The bottom two signals are IMFs computed with masking signals of 2100 Hz for the first IMF and 1200 Hz for the second IMF. Note how the bottom two IMFs do not show mode mixing. The two IMFs on the bottom are nearly identical to the original two components that form $x(t)$.

ately. At some point, as x varies between 50 and 650, there is a transition between interpreting the signal as two separate tones and as a modulated tone. In order to use EMD, the transition between the two cases must be understood. This transition may appear to be a shortcoming of EMD, but another viewpoint is that EMD may provide a reasonable way to interpret such combinations of sine waves. We examine the case in (10). Figure 1 shows how EMD interprets the sum in (10) with $f_1 = 1776$, $a_1 = a_2 = 1$ and a sampling rate of 44,100 Hz. The second frequency is varied from 50 Hz to 4000 Hz. When f_2 is small or large, EMD interprets the signal as two components. When $0.5 < f_1/f_2 < 2$ the first IMF is a modulated signal that contains the higher frequency component and a portion of the lower frequency component.

4. A SOLUTION TO MODE MIXING

Here we present a novel solution to the problem of mode mixing. The basic idea is to insert a masking signal, in this case a single sine tone, that prevents lower frequency components from being included in the IMF. Since the masking signal is known, it can be removed from the IMF obtained through EMD in the following manner.

1. Construct a masking signal, $s(n)$, from the frequency information of the original data, $x(n)$
2. Perform EMD on $x_+(n) = x(n) + s(n)$ to obtain the IMF $z_+(n)$. Similarly obtain $z_-(n)$ from $x_-(n) = x(n) - s(n)$.
3. Define the IMF as $z(n) = (z_+(n) + z_-(n))/2$.

This algorithm deserves a bit of an explanation. Suppose that the signal $x(n)$ contains intermittent components of two frequencies f_a and f_b with $f_a > f_b$. Wherever f_a is present in the signal,

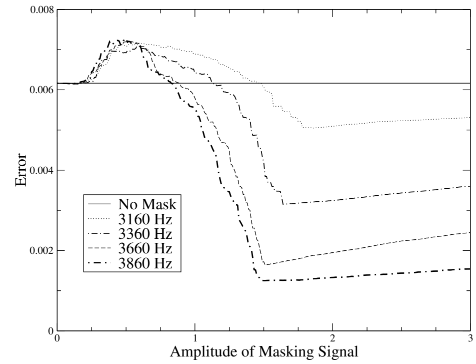

Fig. 3. The error between the first IMF and the actual high frequency component with various masking frequencies. The signals tested are of the form in (10) with $f_1 = 3060$ and $f_2 = 2050$ with unit amplitude. The horizontal axis displays the amplitudes used in the masking signal. The error is minimized around $a_0 = 1.6$.

it will be extracted in z_- and z_+ as long as the masking signal is chosen appropriately. The averaging step will cancel out the positive and negative instances of the masking signal leaving only the portion of the signal at f_a. Where f_a is not present only the masking signal is extracted. In this case $z(t) = (s(t) - s(t))/2 = 0$.

4.1. Choosing a masking signal

In order to pick an appropriate masking signal, the frequency content of a signal must be examined. One possibility is taking a DFT and examining peaks of the spectrum. A difficulty in this approach is that EMD is biased to pick AM signals, so peaks in the spectrum do not necessarily correspond to individual IMFs. The DFT and EMD are very different in their approaches, so coupling the two may be practical but is certainly discordant.

Another approach to constructing a masking signal, that taken in this paper, uses the first unaltered IMF, $y_1(n)$ to describe the highest frequency component of the signal. This IMF may contain mode mixing with two or more frequency bands contained within the signal. After a Hilbert decomposition of the IMF, $y_1(n)$, into $a_1(n)$ and $f_1(n)$, an energy weighted mean

$$\bar{f} = \frac{\sum_{i=1}^{k} a_1(i) f_1^2(i)}{\sum_{i=1}^{k} a_1(i) f_1(i)} \tag{11}$$

of the Hilbert frequencies gives the mean frequency over k samples. An approximation of the frequency of the higher component can be found by finding the energy weighted mean of all points whose frequency is greater than $\bar{f}$. Choosing a masking signal of the form $s(n) = a_0 \sin(2\pi \bar{f} \frac{n}{f_s})$ leads to good performance when each frequency within the signal is separated by at least a factor of 2. The choice of a_0 can affect the performance of the algorithm. Generally the optimal choice of a_m depends on the frequencies and amplitudes of the components, but the factor of 1.6 above the average amplitude of the components is a decent rule of thumb. Figure 3 shows how error changes with a_0. Note the dramatic drop in error around $a_0 = 1.6$ in this case. The best value of a_0 depends on both the amplitudes and frequencies of the components involved. As is usual with EMD, experience must guide the choice of parameters for a particular problem.

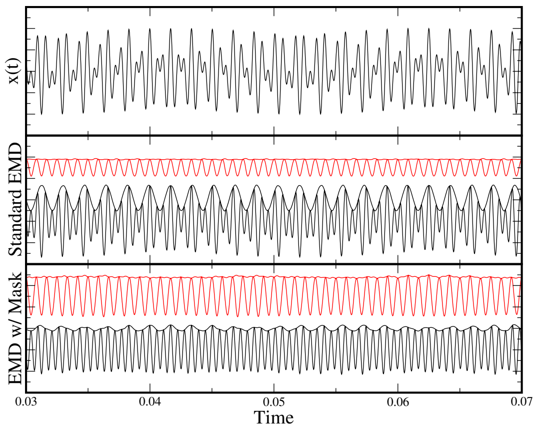

Fig. 4. The signal x(t) as in (10) with $f_1 = 1776$ Hz and $f_2 = 1200$ Hz is in the top graph. The middle graph contains the IMFs computed using standard EMD. Note how a fraction of the lower frequency component combines with the higher frequency component. In the bottom graph a masking signal of 2100 Hz prevents the two signals from combining.

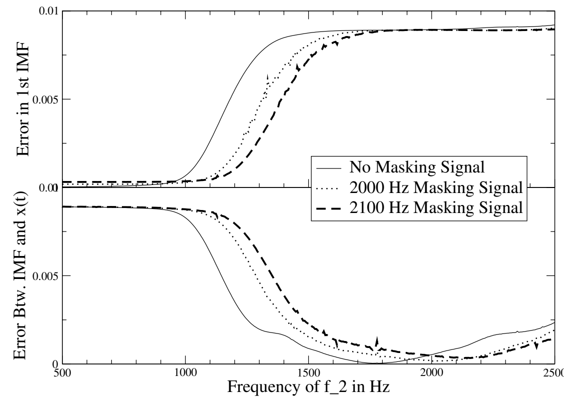

Fig. 5. Results from EMD applied to the two-tone signal $x(t) = \sin(2\pi f_1 t) + \sin(2\pi f_2 t)$ with $f_1 = 1776$ Hz and f_2 varying from 50 to 2500 Hz. The horizontal axis displays the value of f_2. The top graph shows the error between the first IMF and $\sin(2\pi f_1 t)$. The bottom graph shows the error between the first IMF and $x(t)$. While $f_2 < 1000$ Hz, EMD interprets $x(t)$ as two signals. As the error in the bottom graph goes to zero, EMD interprets the signal as a single modulated component. Adding a masking signal increases the maximum value of f_2 for which the two sine waves are separated by EMD from around 1000 Hz to 1200 Hz.

4.2. Using Masking Signals to Separate Components

Whenever the ratio of high frequency to low(next highest) frequency is less than 2, EMD interprets the combination as having some degree of modulation. By choosing a masking frequency higher than the highest frequency component, it is possible to separate two components whose frequencies are within a factor of 2 of each other (Figure 4). The presence of the higher frequency masking signal causes the first component to be included in the first IMF while the second is ignored. Without the masking signal, a portion of the lower frequency component is interpreted as being part of the higher frequency component. Figure 5 shows how masking signals can help to separate signals that are close in amplitude.

5. CONCLUSION

Empirical Mode Decomposition is an algorithm that decomposes signals into Intrinsic Mode Functions which admit an unambiguous definition of instantaneous frequency and amplitude at each sampled time for each component. Because EMD is not defined analytically, experimentation provides the primary means of understanding the algorithm. While the DFT interprets all sums of sine waves as sums, EMD interprets some of these sums as sums and others as AM signals. In many cases this interpretation is elementary, but for the cases in which the interpretation has been ambiguous, EMD provides an answer. An interesting question is whether the EMD interpretation has a natural or experimental basis. For example, does the EMD decomposition of pure tones correspond to psychoacoustic observations of human hearing?

Intermittency is another difficulty of using EMD in many real-world applications. By examining the high frequency content of a signal, we construct a masking signal that is used to prevent mode mixing by adding and then subtracting the masking signal from the original. Unlike methods proposed in [5, 6] there is no hard decision that must be made in eliminating intermittency. In cases when the components are well separated in frequency, this method works as well as could be desired. In certain regions where components are too close together in frequency to be separated using standard EMD, the masking signal technique can be used to distinguish the two components. The decision as to whether two components should be separated depends on the problem at hand. This method provides the researcher an opportunity to separate components of different, but similar, frequencies that was not previously available using EMD.

6. REFERENCES

[1] N. Huang et al., "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-startionary time series analysis," *Proc. R. Soc. Lond. A*, vol. 454, no. 1971, pp. 903–995, March 1998.

[2] G. Rilling, P. Flandrin, and P. Gonçalvès, "On empirical mode decomposition and its algorithms," in *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, NSIP-03*, 2003.

[3] P. Flandrin, G. Rilling, and P. Gonçalvès, "Empirical mode decomposition as a filter bank," *IEEE Sig. Proc. Lett.*, vol. 11, no. 2, pp. 112–114, 2004.

[4] S. Hahn, *Hilbert Transforms in Signal Processing*, Artech House, Boston, 1996.

[5] N. E. Huang, Z Shen, and S. Long, "A NEW VIEW OF NONLINEAR WATER WAVES: The Hilbert Spectrum," *Annu. Rev. Fluid Mech.*, vol. 31, no. 1, pp. 417–457, 1999.

[6] N.E. Huang, M. Wu, S. Long, S. Shen, W. Qu, P. Gloerson, and K. Fan, "A confidence limit for the empirical mode decomposition and hilbert spectral analysis," *Proc. R. Soc. Lond. A*, vol. 459, no. 2037, pp. 2317 – 2345, September 2003.