One or two Frequencies? The empirical Mode Decomposition Answers

Gabriel Rilling, Patrick Flandrin

One or Two Frequencies? The Empirical Mode Decomposition Answers

Gabriel Rilling and Patrick Flandrin*, *Fellow, IEEE*

Abstract

This paper investigates how *Empirical Mode Decomposition* (EMD), a fully data-driven technique recently introduced for decomposing any oscillatory waveform into zero-mean components, behaves in the case of a composite two-tones signal. Essentially two regimes are shown to exist, depending on whether the amplitude ratio of the tones is greater or smaller than unity, and the corresponding resolution properties of EMD turn out to be in good agreement with intuition and physical interpretation. A refined analysis is provided for quantifying the observed behaviours, theoretical claims are supported by numerical experiments, and possible extensions to nonlinear oscillations are briefly outlined.

Index Terms

Resolution, spectral analysis, time-frequency, Empirical Mode Decomposition.

EDICS Category: SSP-SPEC

I. Introduction

One standard issue in spectrum analysis is *resolution*, i.e., the capability of distinguishing between (more or less closely spaced) neighbouring spectral components. At first sight, this question might appear as unambiguous, but a second thought suggests that it is the case only if some prior assumption on—or modelling of—the signal under consideration is given. Indeed, if it is known that the signal $x(t)$ to be analyzed actually consists of two tones of, say, equal amplitudes[1] with frequencies f_1 and f_2, one can write

Manuscript submitted November 13, 2006. G. Rilling and P. Flandrin are with the Physics Department (UMR 5672 CNRS), Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07 France. Tel.: +33 (0)4 72 72 81 60; Fax: +33 (0)4 72 72 80 80; E-mail: {grilling,flandrin}@ens-lyon.fr.

[1] We restrict in this Introduction to such an oversimplified case because situations with unequal amplitudes are computationnaly more complicated, while their interpretation remains essentially unchanged, see Fig. 1. For the same sake of simplicity in the discussion, phase differences are also ignored.

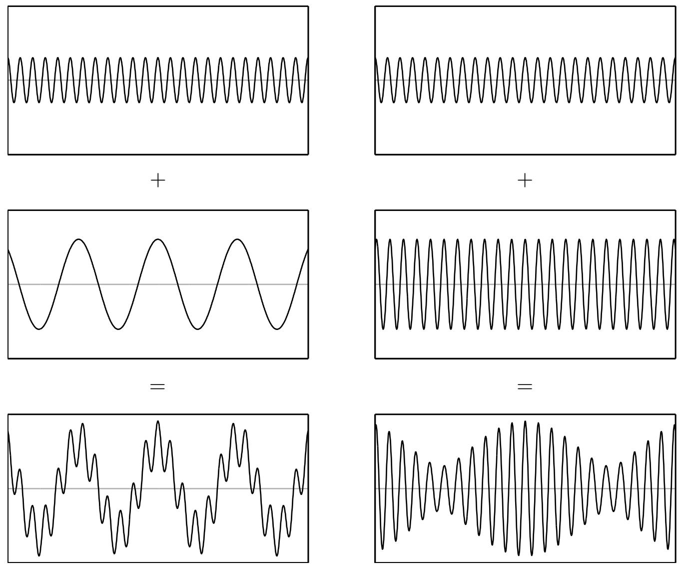

Fig. 1. *Beat effect* — In each column, signals in the bottom row are obtained as the superposition of the waveforms plotted in the top and middle rows. When the frequencies of the two superimposed tones are sufficiently far apart (left column), the two-tones interpretation is meaningful whereas, when they get closer (right column), an interpretation in terms of a single tone modulated in amplitude is clearly favored.

$x(t) = \cos 2\pi f_1 t + \cos 2\pi f_2 t$ and address the question of detecting and estimating f_1 and f_2. However, from a mathematical point of view, one can equally write $x(t) = 2\, \cos \pi(f_1 - f_2)t\, \cos \pi(f_1 + f_2)t$, with an underlying interpretation in terms of a single tone modulated in amplitude rather than of a superposition of two unimodular tones. As it is well-known, such an interpretation is especially relevant when $f_1 \approx f_2$, a prominent example being given by the "beat effect" (see Fig. 1). In this respect, since, at some point, close tones are no longer perceived as such by the human ear but are rather considered as a whole, one can wonder whether a decomposition into tones is a good answer if the aim is to get a representation matched to physics (and/or perception) rather than to mathematics. More generally, if only $x(t)$ is given, there might be no a priori reason to prefer one of the two representations, the effective choice being in some sense driven by the way the signal is processed.

This question is addressed in this paper within the fresh perspective offered by *Empirical Mode Decomposition* (EMD) [4], [5], a relativeley recent technique whose purpose is to adaptively decompose any signal into oscillatory contributions. Since EMD is fully data-driven, not model-based and only defined as the output of an iterative algorithm (see Section II), it is an open question to know what kind of separation can (or cannot) be achieved for two-tones composite signals when using the method. It is worth emphasizing that resolving closely spaced components is not here the ultimate goal, and that poor resolution performance can indeed be accommodated provided that the decomposition is suitably matched to some physically meaningful interpretation. As for earlier publications [2], [3], the aim of the present study is primarily to contribute, within a well-controlled framework, to a better understanding of the possibilities and limitations offered by EMD.

The paper is organized as follows. In Section II, basics of EMD are recalled. After having detailed the two-tones signal model and the performance measure used for quantifying the resolution capabilities of EMD, Section III presents key features of results obtained by numerical simulations. Section IV then offers a theoretical analysis aimed at justifying the reported results, based on a careful study of extrema properties and interpolation schemes. Possible extensions to stochastic and nonlinear situations are briefly mentioned in the Conclusion.

II. EMD BASICS

It is not the purpose of this paper to re-introduce in detail EMD and discuss subtleties about its possible variations and implementations: the interested reader is referred to [4] for the seminal contribution as well as to [1], [9] for implementation issues or [5] for a survey of more recent advances.

In a nutshell, the EMD rationale can be summarized by the motto "signal = fast oscillations superimposed to slow oscillations", with iteration on the slow oscillations considered as a new signal. This shares much with the wavelet philosophy, up to the noticeable difference that the "fast" vs. "slow" distinction is directly controlled by the signal itself, and not by some filtering operations prescribed a priori [2]. More precisely, the decomposition is carried out at the scale of local oscillations, with the "slow oscillations" part obtained as the mean value of an upper and a lower envelope computed as (cubic splines) interpolations between maxima and minima, respectively. Subtracting this component from the original signal, we get what is called an *Intrinsic Mode Function* (IMF). The procedure can then be applied to the slow oscillations part, considered as a new signal to decompose, and successive constitutive components of a signal can therefore be iteratively extracted. In practice, this extraction has to be refined by a *sifting process*, i.e., an inner loop that iterates upon the "fast oscillations" part, until the latter can be considered as zero-mean according to some stopping criterion. Once this is achieved, the fast oscillation is considered as the effective IMF and the corresponding, slow oscillations residual is computed. It follows from this sketchy description of EMD that the key ingredient of the algorithm is what will be referred to in the following as the *sifting operator* $\mathcal{S}(.)$, which acts as

$$(\mathcal{S}x)(t) \triangleq x(t) - m_x(t), \tag{1}$$

where $m_x(t)$ stands for the mean value of the upper and lower envelopes of the analyzed signal $x(t)$, and whose iteration yields the first IMF.

From a physical point of view, an IMF is a zero-mean oscillatory waveform, not necessarily made of sinusoidal functions and possibly modulated in both amplitude and frequency. Going back to the toy example sketched in Fig. 1, the zero-mean criterion makes of the composite signal in the right column a reasonable candidate for being already an admissible IMF, with no need for any further decomposition. This agrees with the beat effect interpretation and contrasts with the situation in the left column, where two zero-mean contributions are clearly visible, making of their identification a meaningful objective. It appears therefore that EMD might be effective in the separation problem, as it has been formulated, and it is the purpose of this paper to switch from intuition and experimental facts to well-supported claims.

III. EXPERIMENTS

In order to understand how EMD decomposes a multicomponent signal into monocomponent ones, one can first remark that, because of the recursive nature of the algorithm, we only have to understand how the first IMF is extracted from the original signal. We will here adopt this perspective, with two-tones signals (in the spirit of Fig. 1) used for the tests. This will allow for analytically tractable analyses, and offer an in-depth elaboration on the preliminary findings reported in [9].

A. Signal model

As far as simulations are concerned, signals are discrete-time in nature and, in the situation we are interested in here, the most general form for a discrete-time two tones signal is:

$$x[n] = a_1 \cos(2\pi f_1 n + \varphi_1) + a_2 \cos(2\pi f_2 n + \varphi_2), n \in \mathbb{Z}.$$

We will however not use such a form with 6 parameters, since only 3 are needed without loss of generality. Concerning first the amplitudes a_1 and a_2, it is obvious that the behaviour of EMD only depends on their ratio $a \triangleq a_2/a_1$. A similar simplification applies as well to the two frequency and phase parameters $\{f_1, f_2\}$ and $\{\varphi_1, \varphi_2\}$ insofar as both frequencies are much smaller than the sampling frequency f_s. As reported in [8], [9], [11], sampling effects may turn the analysis much more complicated when the frequencies of the sinusoids get close to the Nyquist frequency ($f_1, f_2 \gtrsim 0.25 f_s$). Therefore, we will only address the case where $f_1, f_2 \ll f_s$, allowing us to consider that we work with continuous-time signals. In that case, the covariance of EMD with respect to time shifts and dilations makes its behaviour only sensitive to the relative parameters $f \triangleq f_2/f_1$ and $\varphi \triangleq \varphi_2 - \varphi_1$, thus leading to the simpler continuous-time model:

$$x(t; a, f) = \cos 2\pi t + a\cos(2\pi f t + \varphi), t \in \mathbb{R}. \tag{2}$$

As we can moreover restrain ourselves to the case $f \in]0, 1[$, the $\cos 2\pi t$ term will be referred to in the following as the higher frequency component (HF) and the $a\cos(2\pi f t + \varphi)$ term as the lower frequency component (LF).

B. Performance measure

Given the above model, the questions of interest are: (1) “When does EMD retrieve the two individual tones?”, (2) “When does it consider the signal as a single component?” and (3) “When does it do something else?”

In order to address the first question, we can consider the quantity:

$$c_1^{(n)}(a, f, \varphi) \triangleq \frac{\|d_1^{(n)}(t; a, f) - \cos 2\pi t\|_{L^2(T)}}{\|a\cos(2\pi f t + \varphi)\|_{L^2(T)}}, \tag{3}$$

where $d_1^{(n)}(t; a, f) \triangleq (\mathcal{S}^n x(.; a, f))(t)$ stands for the first IMF extracted from $x(t; a, f)$ with exactly n sifting iterations, based on an observation of duration $T \gg 1/f$. When the two components are correctly separated, the fine to coarse nature of the decomposition ensures that the first IMF necessarily matches the HF component $\cos 2\pi t$. Provided this, the second IMF is bound to match the LF component because the first slow oscillations residual from which the second IMF is extracted already is the LF component. Therefore a zero value of (3) indicates a perfect separation of the two components. Finally, the denominator is chosen so that the criterion has a value close to 1 when the two components are badly separated.

C. Results

Fig. 2 summarizes experimental results obtained for $\langle c_1^{(10)}(a, f, \varphi)\rangle_\varphi$, the averaged value over φ of $c_1^{(n)}(a, f, \varphi)$, with $n = 10$ sifting iterations.[2] Examining this figure evidences two rather well separated domains with contrasting behaviours, depending on whether the amplitude ratio is greater or smaller than unity. While it seems rather natural that, for a given amplitude ratio, EMD resolves the two frequencies

[2]The corresponding standard deviation is not shown because it is generally very small, except in some very specific cases involving frequency synchronizations. The number of 10 iterations is arbitrary but its order of magnitude is guided by common practice [5].

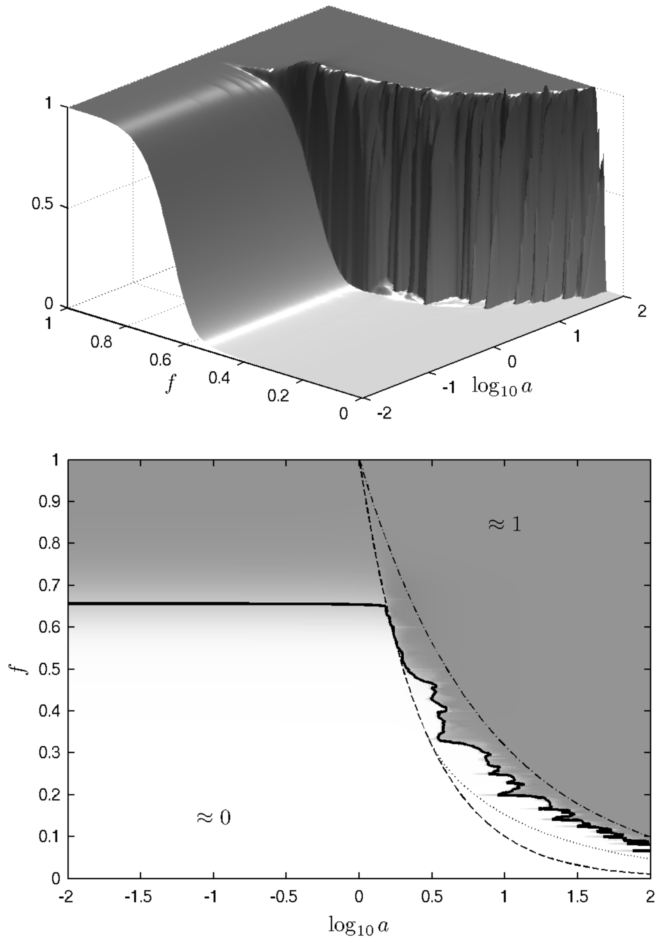

Fig. 2. *Performance measure of separation for two-tones signals* — A 3D version of the averaged criterion (3) with $n = 10$ sifting iterations is plotted in the top diagram. Its 2D projection onto the (a, f)-plane of amplitude and frequency ratios is plotted in the bottom diagram, with the critical curves predicted by theory superimposed as dashed ($af = 1$), dash-dotted ($af^2 = 1$) and dotted ($af \sin(3\pi f/2) = 1$) lines. The black thick line stands for the contour $\langle c_1^{(10)}(a, f, \varphi)\rangle_\varphi = 0.5$.

only when the frequency ratio is below some cutoff, a less usual feature, coming from the highly nonlinear nature of EMD, is that the cutoff frequency also depends on the amplitude ratio, in a non-symmetrical way. What turns out is that this dependence essentially applies when the amplitude of the HF component gets smaller than that of the LF one, and vanishes in the opposite case. Moreover, there is a critical cutoff frequency ratio ($f_c \approx 0.67$ in the present case) above which it is impossible to separate the two components, whatever the amplitude ratio.

Those findings will be given a theoretical justification in the following Sect. IV. In particular, we will show in Sect. IV-B.1.b that the behaviour of EMD is very close to that of a linear filter when the amplitude ratio tends to zero and remains close as long as $a \leq 1$. The transfer function of this equivalent filter will be analytically characterized, and it will be shown that its cutoff frequency only depends on the number of sifting iterations.

IV. Theory

A. About extrema

Extrema play a crucial role in the EMD algorithm, and it is important to discuss precisely what can be known about their locations and/or distributions. There are two asymptotic cases where we know exactly the locations of extrema: if the amplitude of one of the components is infinetly larger than that of the other one ($a \to 0$ or $a \to \infty$), then the locations of the extrema in the sum signal are exactly those of the extrema in the larger component. When the amplitude ratio is finite, locating precisely the extrema becomes tricky, but in almost all situations, we can still obtain the average number of extrema per unit length (or extrema rate) $r_e(a, f)$. This information will in fact serve as a basis for a theory that, while only asymptotically exact, adequately accounts for the behaviour of EMD for almost all frequency and amplitude ratios.

Proposition 1: If $af < 1$, $r_e(a, f) = 2$, i.e., the extrema rate is exactly the same as that of the HF component, whereas, if $af^2 > 1$, $r_e(a, f) = 2f$, i.e., it is exactly that of the LF one.

In order to prove those claims (illustrated in Fig. 3), the first step is to show that the sign of the second derivative of the two-tones signal at its extrema is actually the same as that of the second derivative of the HF component if $af < 1$ and that of the LF component if $af^2 > 1$. To this end, let us assume that $x(t; a, f)$ admits an extremum at $t = t_0$:

$$\partial_t x(t; a, f)|_{t=t_0} \propto \sin 2\pi t_0 + af \sin(2\pi f t_0 + \varphi) = 0. \tag{4}$$

The second derivative of $x(t; a, f)$ is:

$$\partial_t^2 x(t; a, f) \propto \cos 2\pi t + af^2 \cos(2\pi f t + \varphi), \tag{5}$$

and what we want to justify is that

$$|af^2 \cos(2\pi f t_0 + \varphi)| < |\cos 2\pi t_0| \text{ if } af < 1,$$

$$|af^2 \cos(2\pi f t_0 + \varphi)| > |\cos 2\pi t_0| \text{ if } af^2 > 1.$$

Squaring the above equations and replacing $\cos^2 2\pi t_0$ in both of them by its value from (4), we obtain:

$$a^2 f^4 \cos^2(2\pi f t_0 + \varphi) + a^2 f^2 \sin^2(2\pi f t_0 + \varphi) < 1 \text{ if } af < 1,$$
$$a^2 f^4 \cos^2(2\pi f t_0 + \varphi) + a^2 f^2 \sin^2(2\pi f t_0 + \varphi) > 1 \text{ if } af^2 > 1,$$

which hold true since $a^2 f^4 < a^2 f^2 < 1$ if $af < 1$ and $1 < a^2 f^4 < a^2 f^2$ if $af^2 > 1$. Given this result, there can only be one extremum between two successive zero-crossings of the second derivative of the HF (resp. LF) component if $af < 1$ (resp. $af^2 > 1$) because its type (maximum or minimum) is determined by the sign of the second derivative of the HF (resp. LF) component. Thus there can only be one extremum per half-period of the HF (resp. LF) component, hence the extrema rates and the conclusion of the proof.

In the intermediate domain where $af > 1$ and $af^2 < 1$, computing the extrema rate is more involved and a tedious calculation ends up with an expression that operates a transition between the two former domains according to:

$$r_e(a, f) = 2f + \frac{4}{\pi}\left[\sin^{-1}\left(\frac{1}{af}\sqrt{\frac{1 - a^2 f^4}{1 - f^2}}\right) - f \sin^{-1}\sqrt{\frac{1 - a^2 f^4}{1 - f^2}}\right].$$

From the EMD point of view, these informations are already decisive. Indeed, it seems rather natural that EMD can only extract a component if it “sees” extrema that are related to it. Therefore, it seems unlikely to recover the HF component in the $af^2 > 1$ area as the signals extrema are more related to the LF component. On the other hand, recovering the HF component in the $af < 1$ area seems feasible a priori. As extracting the highest frequency component into the first IMF is what we intuitively expect EMD will do, we can think of the $af < 1$ area as a “normal” case while the $af^2 > 1$ area can be viewed as “abnormal”. We will see in the following that quite strange behaviours can indeed be observed in that area. However, the existence of such an “abnormal” domain is far from a complete drawback as it allows in particular to preserve nonlinear periodic waveforms as single IMFs instead of scattering their harmonics over several ones.

The two curves, $af = 1$ and $af^2 = 1$, have been superimposed to the diagram Fig. 2 in order to visualize the link between them and the behaviour of EMD. It appears indeed that the $af^2 = 1$ curve tightly delimits the upper side of the transition area in the right side of the figure. On the other hand, the $af = 1$ curve delimiting the lower side is not as tight as soon as $f < 1/3$. A refinement resulting in a closer theoretical boundary when $f < 1/3$ will be proposed in Sect. IV-C.

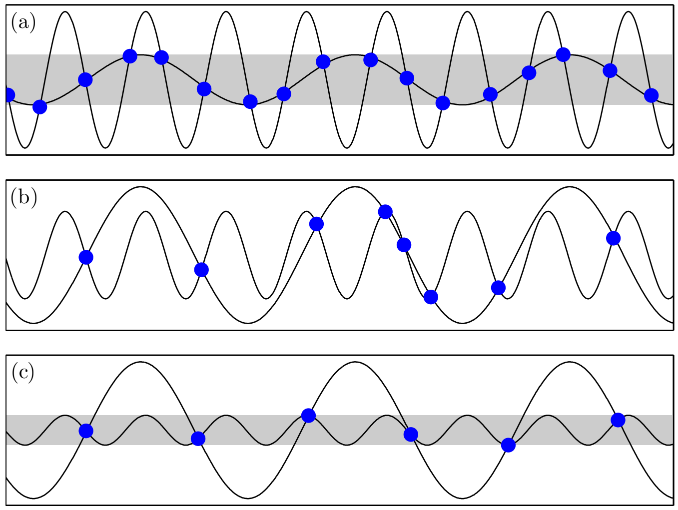

Fig. 3. *Locations of the extrema in the two-tones signal* — In each graph are plotted the derivative of the HF component and the *opposite* of the derivative of the LF component so that each crossing corresponds to an extremum in the composite signal. (a) $af < 1$: as the amplitude of the derivative of the HF component is greater than that of the LF component, there is exactly one crossing between two successive extrema of the HF component; (b) $af > 1$ and $af^2 < 1$: no regular distribution of the crossings can be guaranteed ; (c) $af^2 > 1$: as the maximum slope of the derivative of the LF component is greater than that of the HF component, there is exactly one crossing between two successive extrema of the LF component.

Finally, in order to better support the theory exposed in the next section, some control can be added to the extrema locations. Indeed, if we take into account that the crossings in Fig. 3 (a) can only occur in the shaded band, then it results that each crossing is distant from a zero-crossing of the HF component by at most $1/(2\pi)\sin^{-1}(af)$. In terms of extrema, this exactly means that, when $af < 1$, each extremum is distant from an extremum of the HF component by at most $1/(2\pi)\sin^{-1}(af)$. Likewise, when $af^2 > 1$, each extremum can be shown to be distant from an extremum of the LF component by at most $1/(2\pi f)\sin^{-1}(1/af)$. In both cases, these statements give some quantitative control on the extrema locations supporting the idea that these are close to those of one of the two components as soon as af gets sufficiently smaller than 1 or alternatively af^2 sufficiently larger than 1.

B. A model for EMD when the extrema are equally spaced

Based on the above statements regarding extrema, the model proposed here relies on the only assumption that extrema are located at the exact same places as those of one of the two components (the HF one when $af < 1$ or the LF one when $af^2 > 1$).

1) Case $af < 1$:

a) Derivation of the model: The model assumes that maxima are located at integer time instants $k, k \in \mathbb{Z}$ while minima are located at half-integer ones $k+1/2, k \in \mathbb{Z}$. The extrema being equally spaced, the set of maxima (or minima) can be seen as a specific sampling of the initial signal at the frequency of the HF component, which is 1 in our simplified model. In the Fourier domain, the Dirac combs associated with the maxima and minima samplings are:

$$\text{Ш}_{max/min}(\nu) = \sum_{k\in\mathbb{Z}} (\pm 1)^k \delta(\nu - k), \tag{6}$$

where the $(-1)^k$ term comes from the time shift between the two samplings. Convolving these with the signal in the Fourier domain, we obtain the Fourier representations of the maxima/minima sets (resp. S_{max} and S_{min}):

$$S_{max/min}(\nu) = \left(\text{Ш}_{max/min} * \hat{x}\right)(\nu).$$

If we then use an interpolation scheme that, when the knots are equally spaced, can be expressed in terms of digital filtering (e.g., Shannon, splines,...) with frequency response $I(\nu)$ for unit spaced knots, the envelopes ($e_{max}(t)$ and $e_{min}(t)$) have the Fourier representations:

$$\hat{e}_{max/min}(\nu) = I(\nu)\left(\text{Ш}_{max/min} * \hat{x}\right)(\nu).$$

This implies that

$$\frac{\hat{e}_{min}(\nu) + \hat{e}_{max}(\nu)}{2} = I(\nu)\left(\text{Ш}_{mean} * \hat{x}\right)(\nu),$$

where $\text{Ш}_{mean}(\nu) = \sum_{k\in\mathbb{Z}} \delta(\nu - 2k)$, the two interleaved samplings being replaced by another one with double frequency. It is worth noticing that the samplings (6) introduce aliasing as soon as the signal contains frequencies above the corresponding Nyquist frequency 0.5 (such effects have been first reported in [7]) but, in the present case, the cancellation of the odd indices in the Dirac combs cancels the aliasing effects provided the spectrum of the signal is zero outside of $[-1, 1]$. While this last condition is obviously met in the case of the two-tones signal when $af < 1$, it will not be the case in the following when $af^2 > 1$, and furthermore if we consider nonlinear waveforms.

Finally, we end up with a Fourier representation of the first iteration of the sifting operator given by

$$(\widehat{\mathcal{S}x})(\nu) = \hat{x}(\nu) - I(\nu)\left(\text{Ш}_{mean} * \hat{x}\right)(\nu). \tag{7}$$

As our only assumption was about the locations of extrema, this representation (or a properly dilated one) holds as soon as the extrema locations are nearly equally spaced.

b) Expressions of the IMFs: The signal (2) initially contains four Fourier components at $1, -1, f, -f$ with coefficients $c_{\pm 1} = 1/2$ and $c_{\pm f} = ae^{\pm i\varphi}/2$. After one iteration of the sifting operator, the latters become $c^{(1)}_{\pm 1} = 1/2$ and $c^{(1)}_{\pm f} = (1 - I(f))ae^{\pm i\varphi}/2$, while some new Fourier components appear at $2k \pm f, k \in \mathbb{Z}^*$, because of the extrema sampling. Fortunately, the associated coefficients $c_{2k\pm f} = -I(2k \pm f)ae^{\pm i\varphi}$ are generally very small as $I(\nu)$ is typically close to zero when $|\nu| > 1$. Therefore, those aliased components can generally be neglected, thus implying that the model (7) can be well approximated by a simple *linear filter* with frequency response $1 - I(\nu)$. As $I(\nu)$ typically stays in the range $[0, 1]$ for usual interpolation schemes, the signal after one iteration is then close to a two-tones signal (2), with the same frequency ratio f, a smaller amplitude ratio $a^{(1)} < a$ and the same phase. Thus, the condition $af < 1$ is even better satisfied after one iteration and so the model (7) still holds for all the possible following sifting iterations. This allows us to finally express the first IMF obtained after n iterations in the linear approximation as

$$\begin{aligned} d^{(n)}_1(t; a, f) &= (\mathcal{S}^n x(.; a, f))(t) \\ &= \cos 2\pi t + (1 - I(f))^n \, a \cos(2\pi f t + \varphi), \end{aligned} \tag{8}$$

and consequently the second (and last) IMF as

$$\begin{aligned} d^{(n)}_2(t; a, f) &= x(t; a, f) - d^{(n)}_1(t; a, f) \\ &= (1 - (1 - I(f))^n) \, a \cos(2\pi f t + \varphi). \end{aligned}$$

In the case of cubic spline interpolation, which is by far the most commonly used for EMD, it follows from [12] that the frequency response for unit spaced knots $I(\nu)$ is given by

$$I^{c.s.}(\nu) = \left(\frac{\sin \pi\nu}{\pi\nu}\right)^4 \frac{3}{2 + \cos 2\pi\nu}.$$

Combining this with (8) finally yields a theoretical model for the left side ($af < 1$) of Fig. 2, which is compared to simulations for different numbers of iterations and amplitude ratios in Fig. 4. According to these results, both the model and the simulations point out that EMD performs as a linear filter (high-pass for the first IMF, low-pass for the second one) whose cutoff frequency only depends on the frequency of the HF component and on the number of sifting iterations. It appears moreover that the model remains very close to the simulation results even when af gets close to 1 (i.e., when the model assumption of equally spaced extrema clearly becomes questionable), thus supporting the claim that EMD acts almost as a linear filter over the whole range $af < 1$.

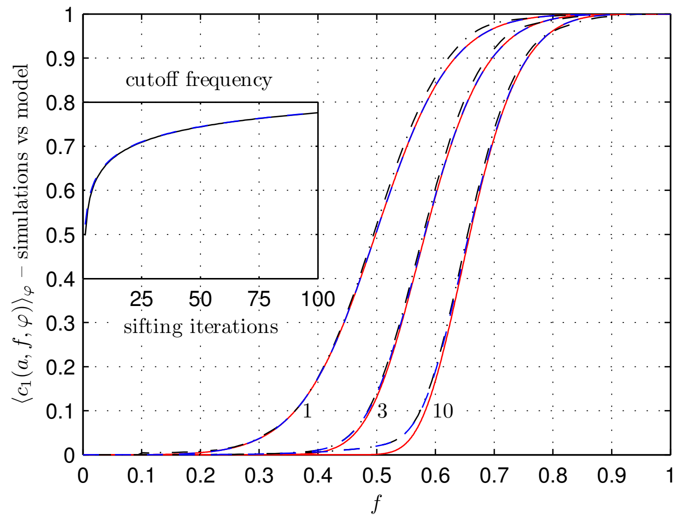

Fig. 4. *Equivalent filter model for EMD* — Experimental results for $\langle c_1^{(n)}(a,f,\varphi)\rangle_\varphi$ (full line curves) are compared to theoretical predictions when $a = 10^{-2}$ (dashed curve) and $a = 1$ (dash-dotted curve), for $n = 1, 3$ and 10 sifting iterations. In the small box are plotted the exact value of the cutoff frequency (full line curve) and its approximation (9) (dashed curve), as a function of the number of sifting iterations.

c) Approximation of the cutoff frequency: A natural definition for the cutoff frequency $f_c^{(n)}$ (where n stands for the number of iterations) is the value for which the response of the EMD equivalent filter is half its maximum:

$$(1 - I(f_c^{(n)}))^n = 1/2.$$

There is unfortunately no analytical solution to this equation in the case of cubic spline interpolation, but a good approximation is given by the following asymptotic formula:

$$f_c^{(n)} = \left(1 + \left(\frac{\ln 2}{n}\right)^{\frac{1}{4}}\right)^{-1} + O\left(n^{-\frac{5}{4}}\right). \tag{9}$$

As pointed out by this expression, the cutoff frequency is a non-decreasing function of the number of iterations that furthermore tends to 1 when n tends to infinity. The increase is however very slow and therefore the cutoff frequency remains significantly lower than 1 for reasonable numbers of iterations: typically $f_c^{(n)} \lesssim 0.75$ when $n \lesssim 100$ (see Fig. 4). Since this cutoff frequency is related to a notion of *relative frequency resolution*, our results agree with the practical observations that suggest a rather poor value for the latter, typically about 0.5. Moreover, the enhancement of that resolution when n is increased, though rather weak, is consistent with empirical observations stating that increasing n increases the number of IMFs and smoothes the envelopes variations. Indeed, both features are consistent with a decrease of each IMF bandwidth, at least locally.[3]

2) Case $af^2 > 1$:

a) Simulations: In the area $af^2 > 1$, we know from the simulation results (Fig. 2) that EMD can never retrieve the two tone components. Thus, the next question of interest here is the second one: does EMD consider the signal as a single modulated component or not? If the answer is 'yes', the first IMF should be the original signal, so that we can address the question by examining the following distance measure (the choice for normalization will become clear later):

$$c_2^{(n)}(a, f, \varphi) \triangleq \frac{\|d_1^{(n)}(t; a, f) - x(t; a, f)\|_{L^2(T)}}{\|\cos 2\pi t\|_{L^2(T)}}. \tag{10}$$

As can be observed in Fig. 5, the behaviour of EMD when $af^2 > 1$ seems to depend mostly on the frequency ratio. More precisely, it appears that the value of $\langle c_2^{(10)}(f_1, f_2, \varphi)\rangle_\varphi$ is either close to one or close to zero depending on whether f is close to $(2k)^{-1}$ or $(2k+1)^{-1}$, $k \in \mathbb{N}^*$. This means that EMD effectively considers the signal as a single component when f is close to $(2k+1)^{-1}$ but does something else when f is close to $(2k)^{-1}$. We will show in the next section that in the latter case, the aliasing effects resulting from the extrema sampling actually give birth to a new lower frequency component.

b) The model: As before, the model assumes that the extrema are equally spaced and therefore the former reasoning applies. We then end up with a formulation which is basically the same as in (7), with a few adjustments:

$$(\widehat{\mathcal{S}x})(\nu) = \hat{x}(\nu) - I(\nu/f)\,(\text{Ш}_{mean} * \hat{x})\,(\nu), \tag{11}$$

with $\text{Ш}_{mean}(\nu) = \sum_{k\in\mathbb{Z}} \delta(\nu - 2kf)e^{i2k\varphi}$.

c) Expressions of the IMFs: If we now apply this model to the two-tones signal, the four Fourier coefficients at $c_{\pm 1} = 1/2$ and $c_{\pm f} = ae^{\pm i\varphi}/2$ become after one sifting iteration $c_{\pm 1} = (1 - I(1/f))/2$ and $c_{\pm f} = ae^{\pm i\varphi}/2$. As in the former case, the component which has its extrema close to those of the multicomponent signal is leaved unchanged by the sifting operator while the other one has its amplitude decreased. However, the decrease is almost nonexistent here because $1/f > 1$ and therefore $I(1/f) \approx 0$. Thus the four initial Fourier components are nearly leaved unchanged after the first iteration. This would be fine if there were no aliased components as in the first case but, unfortunately, there are. Indeed, if $k_f \in \mathbb{N}$ is such that $2k_f - 1 < 1/f < 2k_f + 1$, there are two aliased frequencies $f_a = 2k_f f - 1$ and

[3] It has however to be remarked that, as argued in [10], increasing with no limit the number of sifting iterations would result in IMFs with no amplitude modulation.

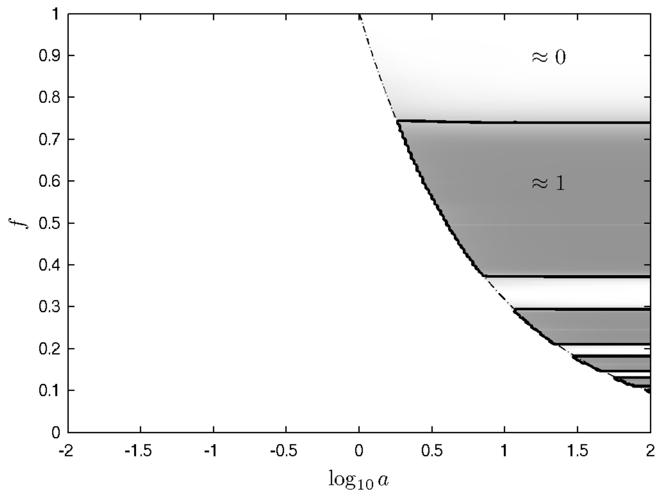

Fig. 5. *Performance measure of separation for two-tones signals* — The averaged value of the criterion (10) is plotted for $n = 10$ sifting iterations. Only the $af^2 > 1$ area is displayed because the criterion is meaningless otherwise. As in Fig. 2, the black thick line stands for the contour $\langle c_2^{(10)}(a, f, \varphi)\rangle_\varphi = 0.5$.

its opposite that both lie in $[-f, f]$ and might therefore not be killed by the interpolation filter $I(\nu/f)$. According to our model, the first IMF is then:

$$d_1^{(1)}(t; a, f) = x(t; a, f) - I\left(f_a/f\right)\cos\left(2\pi f_a t + 2k_f\varphi\right).$$

As this new component has smaller frequency and amplitude than the HF component, we can expect the extrema in the IMF after one iteration to still be near those of the LF component. The aliased component has therefore little impact and can be neglected too. According to this model, the IMF obtained after n iterations is then:

$$d_1^{(n)}(t; a, f) = x(t; a, f) - \lambda_n \cos\left(2\pi f_a t + 2k_f\varphi\right), \tag{12}$$

with $\lambda_n = 1 - \left(1 - I\left(f_a/f\right)\right)^n$ and, consequently,

$$d_2^{(n)}(t; a, f) = \lambda_n \cos\left(2\pi f_a t + 2k_f\varphi\right).$$

Unlike the case $af < 1$, these expressions show that there is no possible linear filtering equivalent for EMD when $af^2 > 1$. As a matter of fact, the behaviour of EMD is rather odd as it just creates a new lower frequency component which is added to the signal to obtain the first IMF, and then compensated in the second one. Moreover, that new frequency is even more annoying when it comes to interpretation as it might be mistaken for an intrinsic time scale of the signal while its relevance is rather questionable.

Indeed, if EMD is followed by instantaneous frequency/amplitude estimation, as in the Hilbert–Huang Transform framework [4], [5], this results in a first IMF with an instantaneous frequency that oscillates around f while that of the second IMF oscillates around f_a. Ultimately, the unfortunate EMD user might uncover a component at f_a, while the frequency of the corresponding intrinsic oscillation is in fact 1!

Quantitatively, the amplitude λ_n of the new frequency component depends on the ratio f_a/f. In a first approximation, $f_a/f \lesssim f_c^{(n)} \Rightarrow \lambda_n \approx 1$ while, on the other hand, $f_a/f \gtrsim f_c^{(n)} \Rightarrow \lambda_n \approx 0$. It results that the new component has either an amplitude close to that of the original at frequency 1 when f is close to $(2k)^{-1}, k \in \mathbb{N}^*$ or an almost zero amplitude when f is close to $(2k+1)^{-1}$. Moreover, as the cutoff frequency $f_c^{(n)}$ increases with n, the width of the bands where $\lambda_n \approx 1$ also increases with n.

C. *Refinements when* $f < 1/3$

1) Extrema may appear during the sifting process: In Sect. IV-A, we showed that the two curves $af = 1$ and $af^2 = 1$ are the boundaries of a transition area for the extrema density. Among these curves, we also noticed that $af = 1$ delimits the lower side of the transition area but rather loosely when $f < 1/3$. In that area, it appears that even if the extrema rate is slightly below 2 ($af \gtrsim 1$), EMD acts as if the extrema were close to those of the HF component. *The explanation for this behaviour is in fact that some extrema may appear when iterating the sifting process*. As a matter of fact, a close look at the two-tones signal when $af > 1$ and $af^2 < 1$ (see Fig. 6) shows that its extrema rate, which is between those of the two tones, is not uniform: the extrema are mostly located around the extrema of the LF component. Besides, the signal also exhibits strong inflections related to extrema pairs of the HF component that clearly would yield extrema pairs in the sum signal if the amplitude ratio was smaller. As the signal exhibits local minima (resp. maxima) around what appears to be the maxima (resp. minima) of the LF component, its lower (resp. upper) envelope and, therefore, their mean roughly follow the shape of the LF component. Subtracting this mean to the signal then naturally reveals the extrema that were hidden as inflections, and it follows that the extrema rate may increase after one sifting iteration to match the rate of the HF component, thus allowing us to use the model (7) for the remaining iterations.

2) A tighter boundary for the transition area: It follows from Proposition 1 that, for a given f, the number of extrema in the two-tones signal that lie in between two successive zero-crossings of the LF component is close to $1/f$ if $a < 1/f$ and exactly 1 if $a > 1/f^2$. However, the threshold value of a below which this number is at least 2 is in fact lower than $1/f^2$ and higher than $1/f$ if $f < 1/3$. It is graphically clear that this threshold $\tilde{a}(f)$ follows from a tangency condition in the worst case situation. Although this condition admits no analytical solution, assuming that the tangency point coincides with

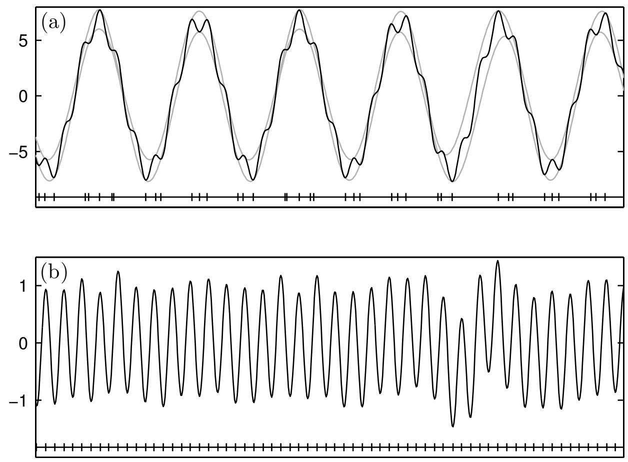

Fig. 6. *Example of a two-tones signal with $af \gtrsim 1$ and $f < 1/3$* — As it can be seen in (a), some extrema of the HF component do not give birth to extrema in the composite signal (identified as marks on the horizontal line at the bottom of the diagram) but only to inflexions surrounded by humps. However, since the lower and upper envelopes (light gray lines) roughly follow the LF trend, their mean is close to the latter and subtracting it to the composite signal uncovers the hidden extrema: this is evidenced in (b), which plots the first IMF after one sifting iteration and the associated distribution of its extrema.

the extremum of the HF component derivative leads to the approximation

$$\tilde{a}(f) \approx (f \sin(3\pi f/2))^{-1}$$

and, hence, to the improved boundary reported on Fig. 2.

D. Summary

Thanks to the previous analysis, we are finally able to answer the three initial questions of interest. A schematic view of the EMD answers to the two-tones separation problem is proposed in Fig. 7, where each area is labelled according to one of the 3 following possibilities: (1) the two components are well separated and correctly identified, (2) their sum is leaved unchanged and considered as a single waveform, (3) EMD does something else, either halfway between (1) and (2), or with the possibility of a decomposition in fake oscillations different from the effective tones.

V. CONCLUSION

This study has considered in detail the way EMD behaves in the simple case of a two-tones signal. A number of experimental findings, well supported by a theoretical analysis, have been reported. This allows for a better understanding of the method and of its relevance in terms of adequation between the

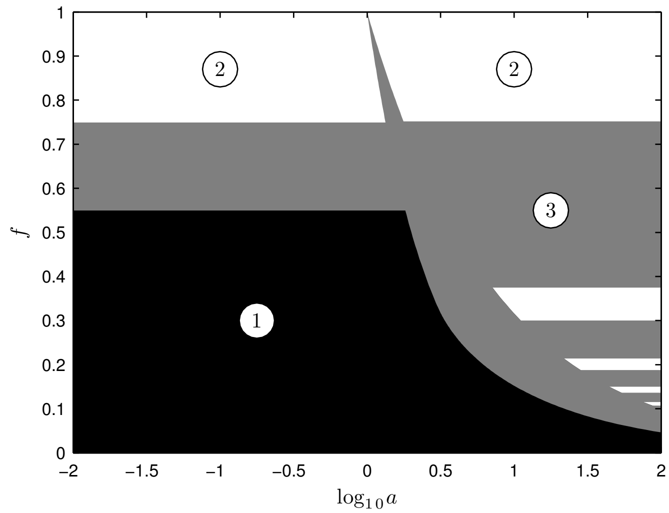

Fig. 7. *Summary of EMD answers to the two-tones separation problem* — Depending on the frequency ratio f and the amplitude ratio a, three domains with different behaviours can be distinguished: (1) the two components are separated and correctly identified, (2) they are considered as a single waveform, (3) EMD does something else.

mathematical decomposition it provides and its physical interpretation. More precisely, one prominent outcome of the study is that EMD permits to address in a fully data-driven way the question whether a given signal is better represented as a sum of two separate, unmodulated tones, or rather as a single, modulated waveform, with an answer that turns out to be in good agreement with intuition (and/or perception).

A limitation of the present study is of course the model on which it is based, which can be thought of as oversimplified and unrealistic. In this respect, it is worth pointing out some possible extensions that would enlarge the perspective. The first one directly follows from the very local nature of EMD, thanks to which conclusions drawn from the study of unmodulated tones still apply to slowly-varying AM-FM situations. A second possible extension is related to Sect. IV-B.1, where it has been shown that the behaviour of EMD can be efficiently described by an equivalent linear filter: a popular situation where such a model has been pushed forward is broadband noise, where it has appeared that EMD acts as quasi-dyadic filter bank [3], [6]. Unlike the two-tones case however, the equivalent filters for broadband noise have only been characterized by numerical experiments, and no theory has been established yet. It is clear however that the two situations are not disconnected, the spectral width of an equivalent filter in the stochastic case being closely related to the ability of distinguishing between neighbouring components in the deterministic case. In that sense, it is not surprising that the transfer function of the equivalent high-pass filter reported in [3], [6] for the first IMF has much to share with the response given in Fig. 4. A more precise approach is of course necessary, and it is under current investigation. Finally, another question of major interest is the possible applicability of the results obtained here to more general cases involving nonlinear oscillations. Such extensions beyond tones are clearly out of the scope of this paper, but it is worth mentioning some encouraging preliminary results in this direction. In fact, when analyzing the separation of two tones, it appeared that it was possible to describe the behaviour of EMD with a simple analytical model provided that the extrema were nearly equally spaced. As the model only requires that last condition to hold, it can also be applied to components different from circular waveforms, provided that the same extrema condition is verified. What turns out is that, whatever the form of the periodic oscillations, there are still two remarkable areas in the frequency-amplitude ratios plane where the behaviour of EMD depends almost only on the frequency ratio. In these areas, where the extrema density is exactly that of one of the two components, the behaviour of EMD can still be fairly approximated by simple, asymptotically exact models derived from (7), the critical frequency ratio depending only on the number of iterations and not on the waveforms.

ACKNOWLEDGMENT

The authors would like to thank Paulo Gonçalvès (LIP-ENS Lyon) for many useful discussions.

REFERENCES

[1] http://perso.ens-lyon.fr/patrick.flandrin/emd.html

[2] P. Flandrin and P. Gonçalvès, "Empirical Mode Decompositions as data-driven wavelet-like expansions," *Int. J. of Wavelets, Multires. and Info. Proc.*, Vol. 2, No. 4, pp. 477–496, 2004.

[3] P. Flandrin, G. Rilling and P. Gonçalvès, "Empirical Mode Decomposition as a filterbank," *IEEE Sig. Proc. Lett.*, Vol. 11, pp. 112–114, 2004.

[4] 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 non-stationary time series analysis," *Proc. Roy. Soc. London A*, Vol. 454, pp. 903–995, 1998.

[5] N.E. Huang and S.S.P. Shen (*eds.*), *Hilbert-Huang Transform and its Applications*, World Scientific, 2005.

[6] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the Empirical Mode Decomposition method," *Proc. Roy. Soc. London A*, Vol. 460, pp. 1597–1611, 2004.

[7] R.N. Meeson, Jr., "HHT sifting and filtering," in [5], pp. 75–105.

[8] G. Rilling and P. Flandrin, "On the influence of sampling on the Empirical Mode Decomposition," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP-06*, Paper SPTM-P5.11, Toulouse (F), 2006.

[9] G. Rilling, P. Flandrin and P. Gonçalvès, "On Empirical Mode Decomposition and its algorithms," in *Proc. IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03*, Grado (I), 2003.

[10] R.C. Sharpley and V. Vatchev, “Analysis of the Intrinsic Mode Functions,” *Constr. Approx.*, Vol. 24, pp. 17–47, 2006.

[11] N. Stevenson, M. Mesbah and B. Boashash, “A sampling limit for the Empirical Mode Decomposition,” in *Proc. Int. Symp. on Signal Proc. and its Appl. ISSPA-05*, pp. 647–650, Sydney (A), 2005.

[12] M. Unser, “Splines: A perfect fit for signal processing,” *IEEE Signal Proc. Mag.*, Vol. 16, No. 6, pp. 22–38, 1999.