ON EMPIRICAL MODE DECOMPOSITION AND ITS ALGORITHMS

*Gabriel Rilling**, *Patrick Flandrin** *and Paulo Gonçalvès***
*Laboratoire de Physique (UMR CNRS 5672), École Normale Supérieure de Lyon
46, allée d'Italie 69364 Lyon Cedex 07, France
{grilling,flandrin}@ens-lyon.fr
**Projet IS2, INRIA Rhône-Alpes, 655, avenue de l'Europe 38330 Montbonnot, France
Paulo.Goncalves@inria.fr

ABSTRACT

Huang's data-driven technique of *Empirical Mode Decomposition* (EMD) is presented, and issues related to its effective implementation are discussed. A number of algorithmic variations, including new stopping criteria and an on-line version of the algorithm, are proposed. Numerical simulations are used for empirically assessing performance elements related to tone identification and separation. The obtained results support an interpretation of the method in terms of adaptive constant-Q filter banks.

1 INTRODUCTION

A new nonlinear technique, referred to as *Empirical Mode Decomposition* (EMD), has recently been pioneered by N.E. Huang *et al.* for adaptively representing nonstationary signals as sums of zero-mean AM-FM components [2]. Although it often proved remarkably effective [1, 2, 5, 6, 8], the technique is faced with the difficulty of being essentially defined by an algorithm, and therefore of not admitting an analytical formulation which would allow for a theoretical analysis and performance evaluation. The purpose of this paper is therefore to contribute experimentally to a better understanding of the method and to propose various improvements upon the original formulation. Some preliminary elements of experimental performance evaluation will also be provided for giving a flavour of the efficiency of the decomposition, as well as of the difficulty of its interpretation.

2 EMD BASICS

The starting point of the Empirical Mode Decomposition (EMD) [2] is to consider oscillations in signals at a very local level. In fact, if we look at the evolution of a signal $x(t)$ between two consecutive extrema (say, two minima occurring at times t_- and t_+), we can heuristically define a (local) high-frequency part $\{d(t), t_- \leq t \leq t_+\}$, or local *detail*, which corresponds to the oscillation terminating at the two minima and passing through the maximum which necessarily exists in between them. For the picture to be complete, one still has to identify the corresponding (local) low-frequency part $m(t)$, or local *trend*, so that we have $x(t) = m(t) + d(t)$ for $t_- \leq t \leq t_+$. Assuming that this is done in some proper way for all the oscillations composing the entire signal, the procedure can then be applied on the residual consisting of all local trends, and constitutive components of a signal can therefore be iteratively extracted.

Given a signal $x(t)$, the effective algorithm of EMD can be summarized as follows [2] (see also `emd.ppt` in [9]):

1. identify all extrema of $x(t)$
2. interpolate between minima (resp. maxima), ending up with some envelope $e_{\min}(t)$ (resp. $e_{\max}(t)$)
3. compute the mean $m(t) = (e_{\min}(t) + e_{\max}(t))/2$
4. extract the detail $d(t) = x(t) - m(t)$
5. iterate on the residual $m(t)$

In practice, the above procedure has to be refined by a *sifting* process [2] which amounts to first iterating steps 1 to 4 upon the detail signal $d(t)$, until this latter can be considered as zero-mean according to some stopping criterion. Once this is achieved, the detail is referred to as an *Intrinsic Mode Function* (IMF), the corresponding residual is computed and step 5 applies. By construction, the number of extrema is decreased when going from one residual to the next, and the whole decomposition is guaranteed to be completed with a finite number of modes.

Modes and residuals have been heuristically introduced on "spectral" arguments, but this must not be considered from a too narrow perspective. First, it is worth stressing the fact that, even in the case of harmonic oscillations, the high vs. low frequency discrimination mentioned above applies only *locally* and corresponds by no way to a pre-determined sub-

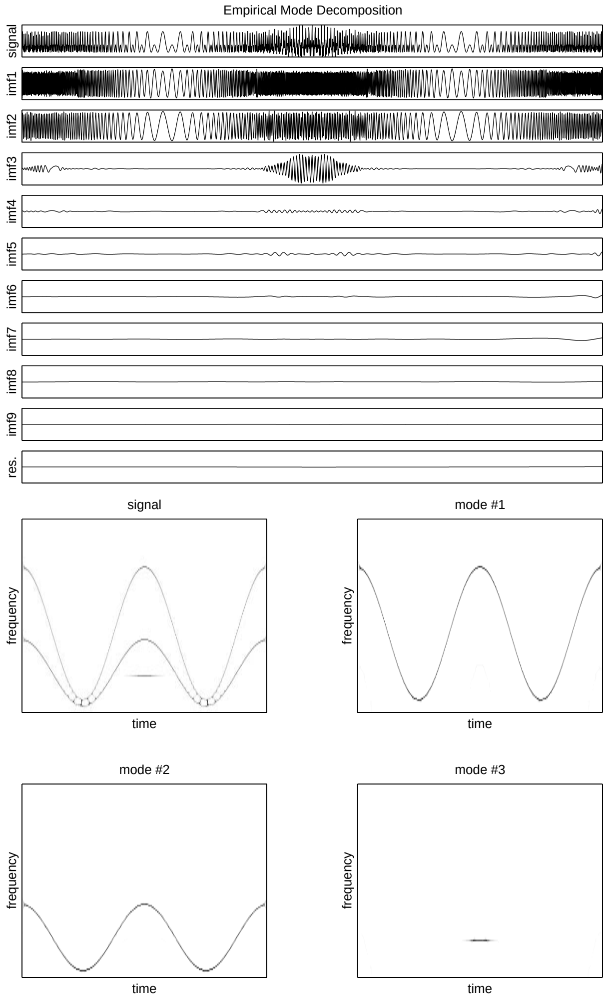

Figure 1: EMD of a 3-component signal. *The analyzed signal (first row of the top diagram) is the sum of 2 sinusoidal FM components and 1 Gaussian wavepacket. The decomposition performed by EMD is given in the 8 IMF's plotted below, the last row corresponding to the final residue. The time-frequency analysis of the total signal (top left of the 4 bottom diagrams) reveals 3 time-frequency signatures which overlap in both time and frequency, thus forbidding the components to be separated by any non-adaptive filtering technique. The time-frequency signatures of the first 3 IMF's extracted by EMD evidence that these modes efficiently capture the 3-component structure of the analyzed signal. (All time-frequency representations are reassigned spectrograms [3, 9].)*

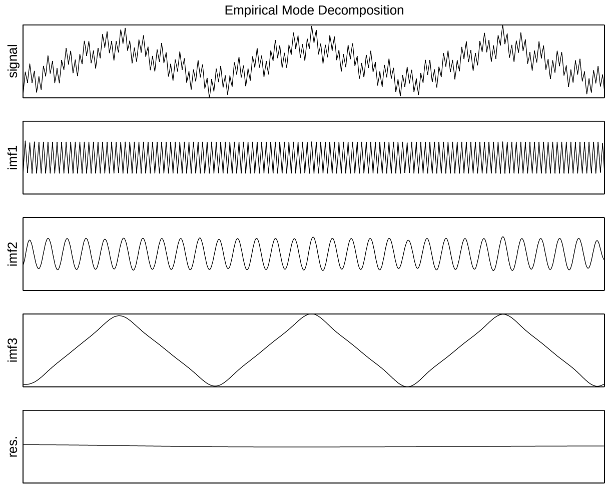

Figure 2: EMD of a 3-component signal — Nonlinear oscillations. *The analyzed signal (first row of the diagram) is the sum of 3 components: a sinusoïd of some medium period T superimposed to 2 triangular waveforms with periods smaller and larger than T. The decomposition performed by EMD is given in the 3 IMF's plotted below, the last row corresponding to the final residue.*

band filtering (as, e.g., in a wavelet transform). Selection of modes rather corresponds to an automatic and adaptive (signal-dependent) time-variant filtering. An example in this direction, where a signal composed of 3 components which significantly overlap in time and frequency is successfully decomposed by the method, is given in Figure 1. This figure has been obtained by running the MATLAB script `emd_fmsin.m`.[1]

Another example (`emd_triang.m`) that puts emphasis on the potentially "non-harmonic" nature of EMD is given in Figure 2. In this case, both linear and nonlinear oscillations (associated respectively with 1 sinusoïd and 2 triangular waveforms) are effectively identified and separated, whereas any "harmonic" analysis (Fourier, wavelets,. . .) would end up in the same context with a much less compact and physically less meaningful decomposition.

3 ALGORITHMIC VARIATIONS

As it has been defined in Section 2, the EMD algorithm depends on a number of options which have to be controlled by the user and which require some expertise. Our purpose here is to make more precise the rationale for such choices, as well as to base upon this analysis variations on the initial formulation.

[1]MATLAB scripts developed for the algorithms and examples described throughout this paper are available as freeware codes [9].

3.1 Sampling, interpolation and border effects

A basic operation in EMD is the estimation of upper and lower "envelopes" as interpolated curves between extrema. The nature of the chosen interpolation plays an important role, and our experiments tend to confirm (in agreement with what is recommended in [2]) that *cubic splines* are to be preferred. Other types of interpolation (linear or polynomial) tend to increase the required number of sifting iterations and to "over-decompose" signals by spreading out their components over adjacent modes.

A second point is that, since the algorithm operates in practice on discrete-time signals, some special attention has to be paid to the fact that extrema must be correctly identified, a pre-requisite which requires a fair amount of *oversampling* (this point will be investigated further in Section 4 below).

Finally, a third issue that has to be taken into account is related to *boundary conditions*, so as to minimize error propagations due to finite observation lengths. To this end, we obtain good results by mirrorizing the extrema close to the edges.

3.2 Stopping criteria for sifting

The extraction of a mode is considered as satisfactory when the *sifting* process is terminated. Two conditions are to be fulfilled in this respect [2]: the first one is that the number of extrema and the number of zero-crossings must differ at most by 1; the second one is that the mean between the upper and lower envelopes must close to zero according to some criterion.

The evaluation of how small is the amplitude of the mean has to be done in comparison with the amplitude of the corresponding mode, but imposing a too low threshold for terminating the iteration process leads to drawbacks similar to the ones mentioned previously (over-iteration leads to over-decomposition). As an improvement to the criteria that have been considered so far [2], we propose in `emd.m` [9] to introduce a new criterion based on 2 thresholds θ_1 and θ_2, aimed at guaranteeing *globally* small fluctuations in the mean while taking into account *locally* large excursions. This amounts to introduce the *mode amplitude* $a(t) := (e_{\max}(t) - e_{\min}(t))/2$ and the *evaluation function* $\sigma(t) := |m(t)/a(t)|$ so that sifting is iterated until $\sigma(t) < \theta_1$ for some prescribed fraction $(1-\alpha)$ of the total duration, while $\sigma(t) < \theta_2$ for the remaining fraction. One can typically set $\alpha \approx 0.05$, $\theta_1 \approx 0.05$ and $\theta_2 \approx 10\,\theta_1$ (default values in `emd.m`).

3.3 Local EMD

In the classical EMD implementation, sifting iterations apply to the *full length* signal, and they are pursued as long as there exists a *local zone* where the mean of the envelopes is not considered as sufficiently small. However, as it has been already mentioned, it turns out that over-iterating on the whole signal for the sake of a better local approximation has the drawback of contaminating other parts of the signal, in particular in uniformizing the amplitude and in over-decomposing it by spreading out its components over adjacent modes. Moreover, the hierarchical and nonlinear nature of the princeps algorithm can by no means guarantee that the EMD of concatenated signals would be the concatenation of individual EMD's.

This observation suggests therefore a first variation upon the initial EMD formulation. This variation, referred to as "local EMD" (`local_emd.m`), introduces an intermediate step in the sifting process: those local zones where the error remains large are identified and isolated, and extra-iterations are applied only to them. This is achieved by introducing a weighting function $w(t)$ such that $w(t) = 1$ on those connected time supports where $\sigma(t) > \theta_1$, with a soft decay to 0 outside those supports. Step 4 of the algorithm described in Section 2 is thus simply replaced by $d(t) = x(t) - w(t)\,m(t)$.

3.4 On-line EMD

A second variation is based on the observation that the sifting step relies on interpolations between extrema, and thus only requires a finite number of them (5 minima and 5 maxima in the case of cubic splines) for being operated at a given point. This suggests that the extraction of a mode could therefore be possible *blockwise*, without the necessary knowledge of the whole signal (or previous residual). This remark paved the road for our development of an EMD algorithm which operates *on-line* and can therefore be applied to data flows (`emd_online.m`).

A pre-requisite for the blockwise extraction of a mode is to apply the same number of sifting steps to all blocks in order to prevent possible discontinuities Since this would require the knowledge of the whole signal, the number of sifting operations is proposed to be fixed a priori, and it proved in fact that a few iterations (less than 10, typically 4) are generally sufficient to extract a meaningful IMF. The effective application of the on-line version of the EMD algorithm that we propose is obtained by means of a sliding window operating on top of the local algorithm described above. The front edge of the window progresses when new data become available, whereas the rear edge progresses by blocks when the stopping criterion is met on a block. Based on this principle, an IMF and its corresponding residual can be computed sequentially. The whole algorithm can therefore be applied to this residual, thus allowing for an extraction of the next mode with some delay.

An example of how the algorithm works on can be appreciated by running `ex_online.m`, an example in which the analyzed signal is the periodization of the 3-component signal used in Figure 1. In this case, the final decomposition obtained on-line on 16000 data points clearly appears as the periodization of the decomposition obtained by decomposing the elementary block of 2000 data points.

Besides the essential usefulness of an on-line algorithm for decomposing data flows, one can also point out its advantage over standard (block) algorithms in terms of computational burden, which quickly becomes very heavy when dealing with long data records.

4 PERFORMANCE ELEMENTS

Since EMD is essentially defined by an algorithm and does not admit an analytical definition, its performance evaluation is difficult and requires extensive simulation experiments. We will report here on two situations: yet elementary, they both point on non-trivial features that are to be known (and better understood) prior applying EMD to real data and trying to interpret the decomposition.

4.1 Tones and sampling

When analyzing a pure tone, EMD is expected to be the identity operator, with only 1 mode (supposed to be identical to the tone), and no residual. Even when keeping apart possible border effects, this happens not to be true because of the unavoidable influence of sampling, which may create jittered extrema when dealing with only a few points per period.

Figure 3 (`emd_sampling.m`) aims at quantifying this phenomenon by plotting, as a function of the tone frequency f, the relative error

$$e(f) = \left(\sum_n (x_f[n] - d_1[n])^2 / \sum_n x_f^2[n] \right)^{1/2}, \quad (1)$$

where $d_1[n]$ stands for the 1st EMD mode extracted from the tone $x_f[n]$ of (normalized) frequency f. It results that such a tone estimation is heavily dependent on f: while the error reaches minima when the tone period is an *even* multiple of the sampling period, we globally observe that $e(f) \leq C\, f^2$.

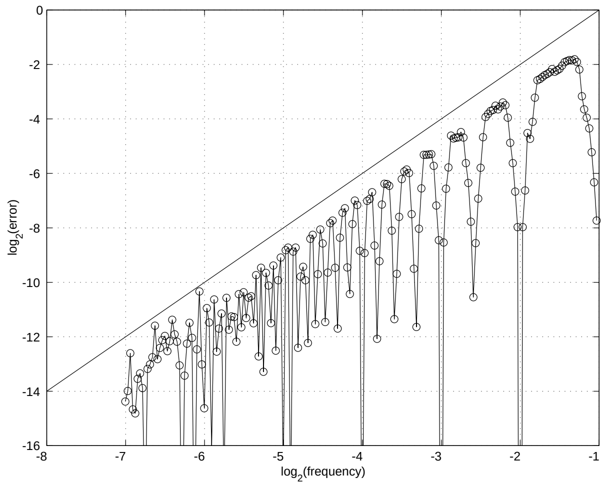

Figure 3: EMD of 1 tone — Estimation error. *Circles correspond to eq.(1) evaluated on 256 points tones. The superimposed straight line has a slope of 2 in this (base 2) log-log plot, thus evidencing that error is upper bounded by a quadratic increase.*

4.2 Tones Separation

In the case of 2 superimposed tones

$$x[n] = a_1 \cos 2\pi f_1 n + a_2 \cos 2\pi f_2 n$$

with $f_2 < f_1 < 1/2$, EMD is expected to extract them via its first 2 modes, although moderate sampling may require more that one mode for extracting f_1, thus locating f_2 in modes of index larger than 2. Errors in the extraction can be quantified via a weighted extension of the criterion (1), with f_1 compared to mode 1, and f_2 to the subsequent mode with smallest error (`emd_separation.m`).

The result is plotted in Figure 4, evidencing an intricate structure with entire domains where tones separation is difficult, in particular when $f_1 > 1/4$. The observed patterns clearly depend on the amplitudes ratio $\rho := a_1/a_2$ but, in a first approximation (and when both f_1 and f_2 are sufficently small so that no sampling issue enters the problem), they all share a common feature: most errors are contained within a triangular domain limited by two lines passing through the origin. In other words, for a given frequency f_1, there exist for each amplitude ratio ρ a fixed $\alpha_\rho < 1$ that defines a "confusion" frequency band $\mathcal{B}(f_1) := [\alpha_\rho f_1, f_1]$ such that f_1 and $f_2 \in \mathcal{B}(f_1)$ cannot be separated. This supports an interpretation of EMD in terms of a *constant-Q filter bank*, in close agreement with the findings reported in [1, 4, 7] in stochastic situations involving broadband noise.

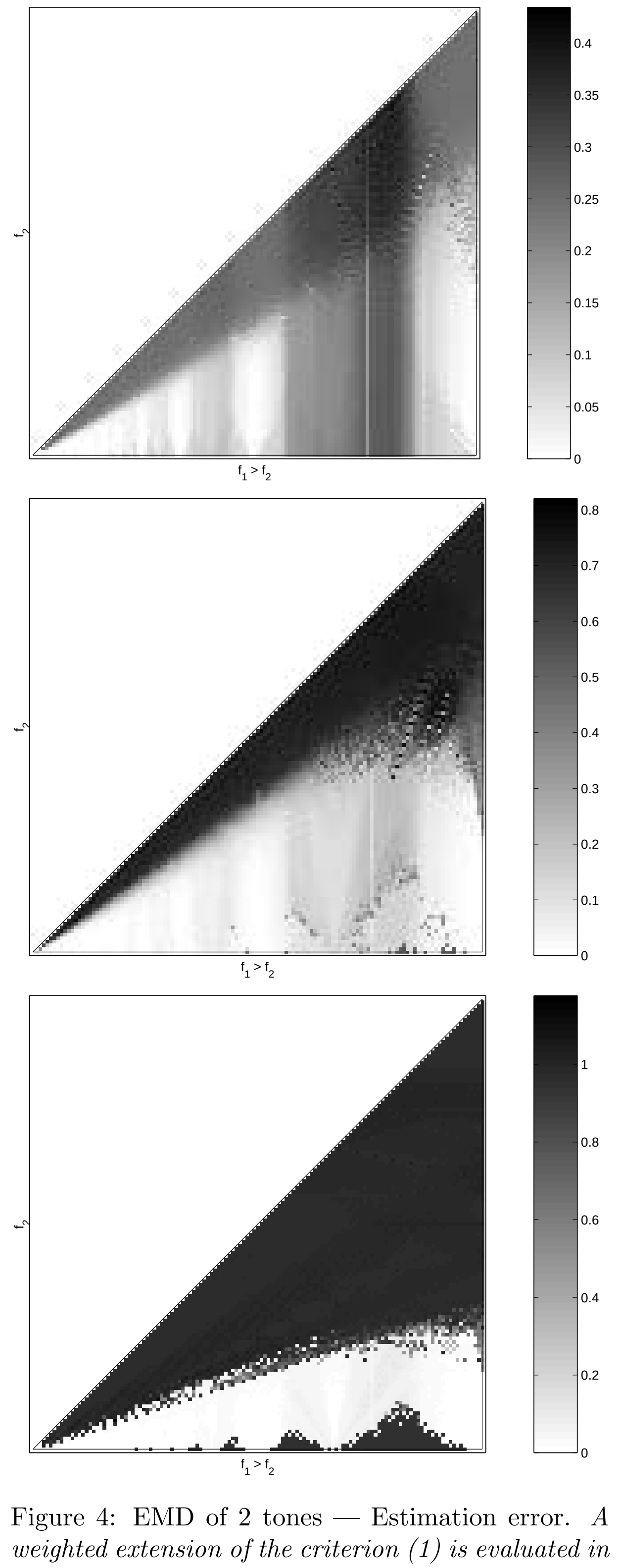

Figure 4: EMD of 2 tones — Estimation error. *A weighted extension of the criterion (1) is evaluated in the lower triangle of the normalized frequency square* $[0, 1/2] \times [0, 1/2]$*, and plotted on a linear gray-level scale (darker pixels correspond to larger errors). Amplitude ratios are set to* $\rho = 4, 1$ *and* $1/4$ *(from top to bottom).*

5 CONCLUSION

EMD is a promising new addition to existing toolboxes for nonstationary and nonlinear signal processing, but it still needs to be better understood. This paper discussed algorithmic issues aimed at more effective implementations of the method, and it proposed some preliminary performance measures.

The results reported here are believed to provide with new insights on EMD and its use, but they are merely of an experimental nature and they clearly call for further studies devoted to more theoretical approaches.

References

[1] K.T. Coughlin and K.K. Tung, "11-year solar cycle in the stratosphere extracted by the empirical mode decomposition method," *Adv. Space Res.*, Nov. 2002 (submitted).

[2] 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.

[3] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, 1999.

[4] P. Flandrin, G. Rilling and P. Gonçalvès, "Empirical Mode Decomposition as a filter bank," *IEEE Sig. Proc. Lett.*, 2003 (in press).

[5] R. Fournier, "Analyse stochastique modale du signal stabilométrique. Application à l'étude de l'équilibre chez l'Homme," Thèse de Doctorat, Univ. Paris XII Val de Marne, 2002.

[6] E.P. Souza Neto, M.A. Custaud, C.J. Cejka, P. Abry, J. Frutoso, C. Gharib and P. Flandrin, "Assessment of cardiovascular autonomic control by the Empirical Mode Decomposition," 4th Int. Workshop on Biosignal Interpretation, Como (I), pp. 123-126, 2002.

[7] 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*, Dec. 2002 (submitted).

[8] Z. Wu, E.K. Schneider, Z.Z. Hu and L. Cao, "The impact of global warming on ENSO variability in climate records," COLA Technical Report, CTR 110, Oct. 2001.

[9] `www.ens-lyon.fr/~flandrin/software.html`