



HAL
open science

3A-EMD: A Generalized Approach for Monovariate and Multivariate EMD.

Julien Fleureau, Amar Kachenoura, Jean-Claude Nunes, Laurent Albera, Lotfi Senhadji

► **To cite this version:**

Julien Fleureau, Amar Kachenoura, Jean-Claude Nunes, Laurent Albera, Lotfi Senhadji. 3A-EMD: A Generalized Approach for Monovariate and Multivariate EMD.. Information Sciences, Signal Processing and their Applications, May 2010, Kuala Lumpur, Malaysia. pp.300 - 303, 10.1109/ISSPA.2010.5605465 . hal-00910730

HAL Id: hal-00910730

<https://hal.archives-ouvertes.fr/hal-00910730>

Submitted on 6 Dec 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

3A-EMD: A GENERALIZED APPROACH FOR MONOVARIATE AND MULTIVARIATE EMD

Julien Fleureau^{1,2}, Amar Kachenoura^{1,2}, Jean-Claude Nunes^{1,2}, Laurent Albera^{1,2}, Lotfi Senhadji^{1,2}

¹Inserm, UMR 642, Rennes, F-35000, France

²Université de Rennes 1, LTSI, Rennes, F-35000, France

For correspondence : Amar Kachenoura

Tel : (33) - 2 23 23 45 52, E-Mail : amar.kachenoura@univ-rennes1.fr

ABSTRACT

EMD is an emerging topic in signal processing research and is applied in various practical fields. Its recent extension to multivariate signals, motivated by the need to jointly analyze multi-channel signals, is an active topic of research. However, all the existing extensions specifically hold either mono-, bi- or tri-variate signals or require multiple projections that complexify the original process. In this communication, a novel EMD approach called 3A-EMD is proposed. It is essentially based on the redefinition of the mean envelope operator and allows, under certain conditions, a straightforward decomposition of monovariate and multivariate signals without any change in the core of the algorithm. A comparative study with classical monovariate and bivariate methods is presented and shows the competitiveness of 3A-EMD. A trivariate decomposition is also given to illustrate the extension of the proposed algorithm to any signal dimension, $D > 2$.

1. INTRODUCTION

Empirical Mode Decomposition (EMD) was originally introduced in the late 1990s to study water surface wave evolution [1]. EMD can be considered as an emerging technique in signal processing with a very important research topics and development in various fields such as biomedical signal analysis [2], Hurst exponent estimation [3], etc. EMD can be viewed as a method which splits a given signal into a set of AM-FM components. The latter components, referred to as the Intrinsic Mode Functions (IMFs), are estimated using an iterative procedure called *sifting process*. Let $\{s(t)\} = \{a_0(t)\}$ be a signal to decompose. Then the sifting process to estimate the n^{th} IMF ($1 \leq n \leq N$) can be summarized as follows:

1. Initialization taking $d_{n,0}(t) = a_{n-1}(t)$ for every t .
2. Computation of the mean envelope $\mathcal{M}(\{d_{n,k}(t)\})$.
3. Extraction of the details $\{d_{n,k+1}(t)\} = \{d_{n,k}(t)\} - \mathcal{M}(\{d_{n,k}(t)\})$.
4. Incrementation of k and go to step 2, repeat until $\{d_{n,k+1}(t)\}$ is designed as an IMF.

In Huang [1], the mean envelope $\mathcal{M}(\{d_{n,k}(t)\})$ is calculated as the half sum of the upper and the lower en-

velopes, which are obtained by interpolation between the local maxima points and the local minima points of $\{d_{n,k}(t)\}$, respectively.

Several studies dealing with EMD have been reported in the last decade: those addressed to improve the Huang's algorithm [4, 5], and those proposing an extension of EMD to bi- or trivariate signals [6, 7, 8] and even recently to multivariate signals [9]. All these latter methods, motivated by the need to jointly analyze multi-channel real signals, are either restricted to a specific signal dimension or are based on multiple projections of the original signal that complexify the original EMD process. The aim of this communication is to propose a new efficient EMD algorithm, named 3A-EMD (Active Angle Averaging EMD), that preserves the original geometrical simplicity of the classical EMD algorithm and extends its performance to multivariate signals defined from \mathbb{R} to \mathbb{R}^D with $D > 2$. The classical EMD algorithm is extended to multivariate signals by redefining the mean envelope operator \mathcal{M} using new characteristic points. A generalized version of the EMD is thus obtained working straightforwardly on monovariate and multivariate signal without any change.

2. A NEW EMD APPROACH: 3A-EMD

The proposed 3A-EMD method aims at providing a simple algorithm working for multivariate signals without any modification. To do so, the signal mean trend (computed by the \mathcal{M} operator) is redefined as the signal going through the *barycenter* of its different *elementary oscillations*. This is performed through three main steps: i) identification of all the elementary oscillations, ii) computation of the barycenters of the identified elementary oscillations, and iii) interpolation between all barycenters to obtain the signal mean trend.

Elementary oscillation An *elementary oscillation* of a given signal with values in \mathbb{R}^D ($D \geq 1$) can be viewed as a subpart of the signal defined between two extrema of oscillation. However, computing the extrema if the output space \mathbb{R}^D is multidimensional ($D > 1$) is not obvious. In this section we propose a new and general way to compute such extrema. Let the signal $\{s(t)\}$ with values in \mathbb{R}^D be

in class \mathcal{C}^1 (or a “regular enough” signal in practical situations). The tangent vector to $\{s(t)\}$ at t' , $\mathbf{T}_s(t')$, is defined by:

$$\mathbf{T}_s(t') = [1, \frac{ds}{dt}(t')]^T \in \mathbb{R}^{D+1} \quad (1)$$

and let's define the quantity, $\alpha_s(t')$, as follows:

$$\alpha_s(t') = \lim_{h \rightarrow 0} \langle \mathbf{T}_s(t' - h), \mathbf{T}_s(t' + h) \rangle \quad (2)$$

which is the inner product of \mathbb{R}^{D+1} (denoted by $\langle \cdot, \cdot \rangle$) between the tangents to $\{s(t)\}$ just before and after t' . In fact, $\alpha_s(t)$ is maximum when both vectors, $\mathbf{T}_s(t-h)$ and $\mathbf{T}_s(t+h)$, are linearly dependent, due to Cauchy-Schwarz inequality, and null when they are orthogonal. This quantity is an indicator of the local angle and thus, of the local oscillation of $\{s(t)\}$ at t' . Due to the continuity of the inner product, we have:

$$\alpha_s(t') = \langle \lim_{h \rightarrow 0} \mathbf{T}_s[t' - h], \lim_{h \rightarrow 0} \mathbf{T}_s[t' + h] \rangle \quad (3)$$

and as $\{s(t)\}$ is in \mathcal{C}^1 , we get:

$$\alpha_s(t') = \|\mathbf{T}_s(t')\|^2 = 1 + \left\| \frac{ds}{dt}(t') \right\|^2 \quad (4)$$

where $\|\cdot\|$ abusively represents the norm of both \mathbb{R}^{D+1} and \mathbb{R}^D . Finally, an oscillation extremum of the signal $\{s(t)\}$ is defined as a local minimum of the function $\{\alpha_s(t')\}$ and the following property holds:

Property: the signal $\{s(t)\}$ has an oscillation extremum in t_o if and only if the function $\{\|\frac{ds}{dt}(t')\|^2\}$ is locally minimum in t_o .

This definition of the extremum does not depend on the output space dimension D , which makes it suitable in multivariate EMD contexts. For $D=1$, the reader could check that the computed extrema using the previous property include the signal scalar extrema (of Huang [1]) but also the saddle points.

Oscillation barycenter Let $P_1 = [t_1, s(t_1)]^T$ and $P_2 = [t_2, s(t_2)]^T$ be two consecutive oscillation extrema. A complete oscillation is defined between P_1 and P_2 and the associated *oscillation barycenter*, $M_{P_1 \rightarrow P_2}$, is obtained by:

$$M_{P_1 \rightarrow P_2} = \left[\frac{t_1 + t_2}{2}, \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} s(t) dt \right]^T \quad (5)$$

The mean trend $\mathcal{M}(\{s(t)\})$ is then computed by interpolating between the barycenters. The \mathcal{M} operator appears with this definition as a non-stationary signal-based averaging operator with a field of application not exclusively restricted to univariate signals. Nevertheless, interpolating directly through all the oscillation barycenters to obtain $\mathcal{M}(\{s(t)\})$, appears in practice to be very sensitive to artefacts introduced by the spline interpolation process and the border effects. Such effects appear very lesser in the classical Huang's approach as if they were “balanced” by the averaging of the min and max envelopes. Similarly,

a robust implementation of the mean envelope computation is obtained for 3A-EMD by averaging two envelopes: a first envelope interpolating the even indexed barycenters (including signal borders) and a second envelope interpolating the odd indexed barycenters (also including signal borders).

Sifting process stop criterion The sifting process termination is handled by means of a classical modified Cauchy-like criterion [4]. The sifting process is thus stopped whenever $\|d_{n,k+1}(t_i) - d_{n,k}(t_i)\| / \|d_{n,k}(t_i)\|$ becomes lesser than a certain threshold ϵ ($1e^{-2}$ in practice) for a ratio of the whole samples of signal equal, at least, to α (90% in practice).

Scope of application Based on the detection of local minima of $\{\|\frac{ds}{dt}(t)\|^2\}$, 3A-EMD raises some assumptions on the candidate signals:

(A1) The signal $\{s(t)\}$ must be regular or sampled enough so that its numerical derivative has a relevant meaning.

(A2) Any signal including one IMF with a piecewise constant derivative norm (and especially purely circular rotating bivariate signals) would not be successfully decomposed by 3A-EMD.

(A3) The energy of the different IMF derivatives of the signal should ideally decrease or, at least, the energy of the n^{th} IMF d_n should not be “too much” bigger than the energy of the following IMFs d_{n+1} . In practice, noisy parts of a signal are often associated to highest frequencies and such condition must not be too restrictive.

Note on computational complexity The computational complexity of 3A-EMD algorithm can be precisely evaluated and compared with some of the reference algorithms of the literature. For any given EMD algorithm, let's call N the number of extracted IMF, K_n the number of sifting iterations performed to extract the n^{th} IMF, and $d_{n,k}$ the signal considered at the k^{th} iterations of the sifting process. Let's also call $M_{n,k}^{Huang}$ the number of extrema detected in $d_{n,k}$ by the univariate Huang's procedure [1], $M_{n,k,p}^{Rilling}$ the number of extrema detected in the p^{th} projection of $d_{n,k}$ by the second geometrical bivariate Rilling's procedure [7] with P projection planes, and $M_{n,k}^{3A}$ the number of barycenters detected in $d_{n,k}$ by 3A-EMD method. The computational complexity of the sifting step necessary to obtain $d_{n,k+1}$ for each of these three methods can be therefore computed for a signal from \mathbb{R} to \mathbb{R}^D with S samples, by the formula given in the table 1. The global complexity for each of those methods is then directly obtained by $C^{method} = \sum_{n=1}^N \sum_{k=1}^{K_n} C_{n,k-1}^{method}$.

3. RESULTS

The objective of this section is twofold: to compare the performance of 3A-EMD algorithm with classical approaches and to illustrate its extensibility to multivariate signals ($D > 2$).

Signals selection The signals used to evaluate the behavior of 3A-EMD algorithm are two univariate (s_{11}, s_{12}),

Method	D	Sifting Complexity
Huang	1	$C_{n,k}^{Huang} = 18S + 15M_{n,k}^{Huang}$
Rilling	2	$C_{n,k}^{Rilling} = S(11P + 2) + 15 \sum_{p=1}^{P/2} M_{n,k,p}^{Rilling}$
3A-EMD	any	$C_{n,k}^{3A} = D(19S + 16M_{n,k}^{3A}) + M_{n,k}^{3A}$

Table 1. Computational complexity of some reference algorithms and of 3A-EMD method.

two bivariate (s_{21} , s_{22}) and one trivariate (s_{31}) signals, defined on the time interval $T = [-1; 2]$ and sampled at $f_s = 10\text{kHz}$:

$$\begin{aligned}
s_{11}(t) &= e^{0.23(1+t)}(0.5\sin(100\pi t) + \sin(50\pi t)) \\
s_{12}(t) &= 2\sin(250\pi t) + e^{0.23(1+t)}\cos(\pi(2.58t + 21.95)^2) + 3\cos(\pi(1.7t + 7.3)^2) + 3t \\
s_{21}(t) &= e^{0.23t-0.46}[0.3\cos(\pi(2.58t + 14.2)^2) + 0.9\sin(\pi(2.58t + 14.2)^2)] \\
&\quad + [0.4\sin(20\pi t) + 0.7\cos(24\pi t)] \\
s_{22}(t) &= [0.3\cos(400\pi t) + 0.8\sin(500\pi t)] \\
&\quad + [4\cos(100\pi t) + 7\sin(100\pi t)] \\
&\quad + e^{0.23t-0.46}[5\cos(\pi(2.24t + 4.47)^2) + 2\sin(\pi(2.24t + 4.47)^2)]
\end{aligned}$$

The trivariate signal s_{31} is given by:

IMF 1	$x(t) = 3.5\sin(500\pi t)$ $y(t) = 2\cos(500\pi t)$ $z(t) = e^{0.23t-0.46}\sin(500\pi t)$
IMF 2	$x(t) = 5e^{0.23(1+t)}\cos(\pi(3.16t + 25.3)^2)$ $y(t) = 2e^{0.23(1+t)}\sin(\pi(3.16t + 25.3)^2)$ $z(t) = 0.5\sin(130\pi t + 1.2)$
IMF 3	$x(t) = 4\sin(60\pi t)$ $y(t) = 2\cos(60\pi t)$ $z(t) = 6\sin(64\pi t + 1.5)$

Note that all these signals are compatible with the assumptions (A1), (A2) and (A3).

Performance criteria Three kinds of performance criterion are used during the simulations.

Criterion 1 A first criterion only takes into account the behavior of the algorithm on a temporal window $[0; 1]$ of $T = [-1; 2]$ where border effects become negligible. Let I_c be the set of associated time indices and \hat{d}_i an estimate of the true i^{th} IMF d_i , then the first criterion is defined by the error $e_c^i(d) = \sum_{k \in I_c} \|\hat{d}_i[k] - d_i[k]\|^2$. For a complete set $\hat{\mathbf{d}} = \{\hat{d}_i\}_{1 \leq i \leq N}$ of estimates of the N IMFs composing the signal to analyse, this error is generalized into $e_c(\hat{\mathbf{d}}) = \sum_{i=1}^N e_c^i(\hat{d}_i)$.

Criterion 2 A second criterion evaluates the ability of the algorithm to minimize border effects. It is represented by e_b^i and e_b and is computed as for criterion 1 but on the set I_b of time indices included in the subinterval $[-1; 0] \cup$

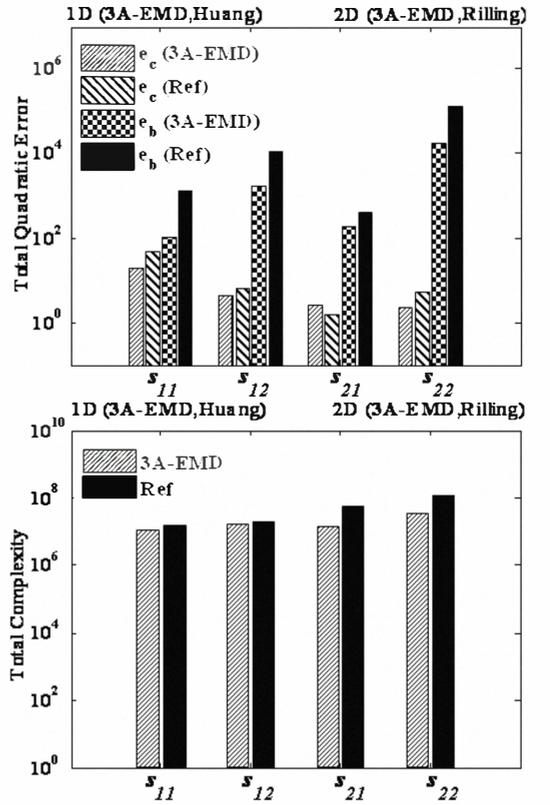


Fig. 1. Comparative study of 3A-EMD versus Huang's (1D) and Rilling's (2D) reference (Ref) methods.

$[1; 2]$ of $T = [-1; 2]$. Indeed, border effects are often critical in practice and their minimization facilitates the exploitation of the IMFs.

Criterion 3 A last criterion evaluates the computational complexity with respect to the formula listed in Table 1.

Comparative study This section compares 3A-EMD performances with those obtained by Huang's algorithm [1] in monovariate case and Rilling's method [7] in bivariate context. More precisely, the second geometric approach proposed in [7] is used as reference for bivariate decompositions with 8 projection planes. In both cases, the sifting process stopping criterion and the border effect management are identical for all the compared methods. The figure 1 presents criteria 1 and 2 on the left, and criterion 3 on the right for both monovariate and bivariate signals. Core performances of 3A-EMD are globally satisfactory and competitive compared with those of the reference methods. Regarding border performances (e_b), 3A-EMD generally offers more efficient results and seems to have a nice behaviour on the signal border. Finally, the computational complexity analysis suggests that, even if no significant differences are visible on monovariate signals, 3A-EMD generally requires less sifting iterations and less computational operations than the Rilling's method.

Trivariate decomposition The figure 2 presents the exact (dashed dark line) and estimated (red line) IMFs of

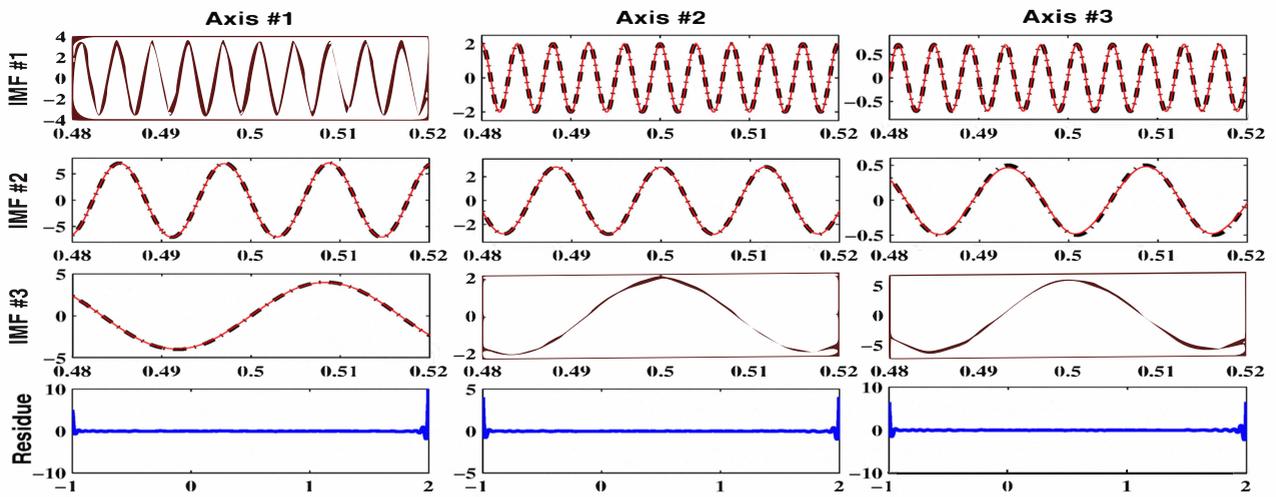


Fig. 2. 3A-EMD trivariate decomposition of the signal s_{31} . Dashed dark line: real IMF / Continuous red line: estimated IMF.

the trivariate ($D > 2$) signal s_{31} projected on the three main axis (from left to right). The three IMFs and the residue are displayed from the top to the bottom of the figure. Note that only a central zoom of the IMFs on the temporal axis has been represented for the sake of clarity. A good behaviour of the algorithm can be observed. Indeed, the three IMFs seem well extracted and the low residue highlights the efficiency of the decomposition on the border. More than quantitative results, this example illustrates the extensibility of 3A-EMD to higher dimension signals without any change in the core of our algorithm. This also shows the relevancy of the use of oscillation barycenters in a general multivariate context.

4. CONCLUSION AND PERSPECTIVES

A new algorithm, called 3A-EMD, based on a novel geometric definition of the mean envelope operator is proposed. The obtained results show that, under certain assumptions on the signals, this alternative definition enables to decompose both monivariate and multivariate signals without any modification in the core of the algorithm. This last point is the one main difference with regard to the existing approaches of the literature. The comparative study also suggests that 3A-EMD seems to offer competitive core and border performances as well as some interesting performances in terms of computational complexity. A restricted scope of application mainly due to the use of the first derivative in the algorithm remains the most important limitation of our approach. However, as pointed before, those restrictions should not be too much restrictive in practical fields. More simulated and real data decompositions should be performed in future works to verify the 3A-EMD interest in practical context.

5. REFERENCES

- [1] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N.-C. Yen, C. Tung, and H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. of the Royal Society A: Mathematical, Physical and Engineering Sciences, London, March 1998, pp. 903-995.
- [2] L. Hualou, L. Qiu-Hua, and J. Chen, "Application of the empirical mode decomposition to the analysis of esophageal manometric data in gastroesophageal reflux disease," IEEE Transactions on Biomedical Engineering, vol. 52, no. 10, pp. 1692-1701, 2005.
- [3] G. Rilling, P. Flandrin, and P. Goncalves, "Empirical mode decomposition, fractional gaussian noise and hurst exponent estimation," Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, Philadelphia, March 2005, vol. 4, pp. 489-492.
- [4] G. Rilling, P. Flandrin, and P. Goncalves, "On empirical mode decomposition and its algorithms," Proc. of IEEE/EURASIP Workshop on Nonlinear Signal and Image Processing, Grado, June 2003.
- [5] E. Delechelle, J. Lemoine, and O. Niang, "Empirical mode decomposition: an analytical approach for sifting process," IEEE Signal Processing Letters, vol. 12, no. 11, pp. 764-767, 2005.
- [6] T. Tanaka and D. Mandic, "Complex empirical mode decomposition," IEEE Signal Processing Letters, vol. 14, no. 2, pp. 101 - 104, 2007.
- [7] G. Rilling, P. Flandrin, P. Goncalves, and J. Lilly, "Bivariate empirical mode decomposition," IEEE Signal Processing Letters, vol. 14, no. 12, pp. 936-939, 2007.
- [8] N. Regman and D. Mandic, "Empirical Mode Decomposition for Trivariate Signals," IEEE Transaction on Signal Processing, vol. 58, no. 3, pp.1059-1068, 2010.
- [9] N. Regman and D. Mandic, "Multivariate empirical mode decomposition," Proc. R. Soc. A, 2009.