A new optimization based approach to the empirical mode decomposition

Basarab Matei and Sylvain Meignen

Abstract - In this paper, an alternative optimization based approach to the empirical mode decomposition (EMD) is proposed. The principle is to build first an approximation of the signal mean envelope, which serves as initial guess for the optimization procedure. We develop several optimization strategies to approximate the mean envelope which compare favorably with the original EMD on AM/FM signals.

Key words and phrases : empirical mode decomposition, adaptive algorithm

Mathematics Subject Classification (2010) : 41,68, 62D99.

1. Introduction

To analyze signals in time and frequency domains at the same time is a challenging issue and many methods exist to this aim: short-time Fourier transform, Wigner-Ville distribution and wavelets [2], [7]. This issue was one of the motivation for the development of the *empirical mode decomposition* (EMD). The EMD expands a given signal into a set of functions defined by the signal itself, the *intrinsic mode functions* (IMFs) computed through the so-called "sifting process" (SP). In the denomination of Huang [5], an IMF is a function that satisfies two conditions: (a) in the whole data set, the number of extrema and the number of zero crossings must either be equal or differ at most by one; and (b) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

In spite of its practical achievements, the EMD technique is essentially an algorithmic approach which lacks a well-established theoretical proof and a direct, systematic optimization of the method. One of the problems posed by the EMD algorithm arises from the interpolation of the extrema, namely undershoots and overshoots. To avoid this situation, many alternative approaches to EMD propose to directly compute an approximation of the signal mean envelope. This knowledge enables to build the mean envelope first and then deduce the mode. In that framework, several approaches have been proposed to find out the intersection points of the signal with

its mean envelope, see [3],[9],[1],[4] and [8]. In [3], the signal mean envelope is a steady state of some partial differential equation which intersects inflection points of the signal. In [1], the mean envelope is a B-spline with knots defined using the signal extrema while, in [4], the mean envelope is a spline intersecting signal local means. In [8], the mean envelope is computed directly through an optimization procedure by imposing a series of linear constraints. In this paper, we first compare these methods to asses of these points of interest for AM/FM signals and we bring about potential improvements to the mean envelope estimation proposed in [4]. We then derive new strategies to approximate the mean envelope using an optimization procedure. The proposed method appears to behave better than the original EMD on AM/FM signals. Let us stress that contrary to the original EMD algorithm where the convergence of the SP is not ensured, we here seek the mean envelope in a finite dimensional subspace which makes the proposed procedure convergent. The outline of our work is as follows. In Section 2, we recall some basic notions about EMD and also present the results of EMD on a sum of AM/FM signals. In Section 3, motivated by the results obtained using the EMD, we propose a different method to compute a mean envelope approximation. We next introduce, in Section 4, a new optimization procedure to approximate the mean envelope. We finally end the paper with some numerical examples on a sum of AM/FM signals.

2. EMD basis and limitations

The EMD decomposes a signal into a sequence of AM/FM components. Let us first briefly recall the principle of the EMD algorithm for a one-dimensional signal x.

1. Initialization (EMD): $r = x$, $k = 1$

2. While r is not monotonic

   (a) Initialization (sifting process): $m_0 = r, i = 0$

   (b) While m_i is not an IMF repeat (sifting process)

      i. Compute the mean envelope $e_i = \frac{1}{2}(u_i + l_i)$ of m_i (where u_i (resp. l_i) denotes the upper (resp. lower) envelope of m_i)

      ii. $m_{i+1} = m_i - e_i$; $i = i + 1$

   (c) $x_k = m_i$, $r = r - x_k$, $k = k + 1$.

When the decomposition is complete, one can write $x = \sum_{k=1}^{K} x_k + r$. Apparently, the modes x_k are computed in an adaptive way. However, the "sifting process" (SP) that computes the IMF is dependent on a stopping criterion

whose choice both influences the number of modes and their shape. Furthermore, in some instances, the SP may not converge. We highlight the limitations of the original EMD (by applying the MATLAB code corresponding to [11]) to a simple AM/FM signal. For the signal x of Figure 3.(A), made of two AM/FM signals, the original EMD (with the default parameters $(0.05, 0.5, 0.05)$, see [11] for details), gives 22 IMFs which are physically not meaningful. A more careful look at the obtained results, shows that the information contained in the modes x_1 and x_2 is spread over more than 10 IMFs. To illustrate this, we depict on Figure 1, seven of these significant modes. The origin of such a behavior is the instability of the SP and this motivates the search for alternative techniques to determine the modes.
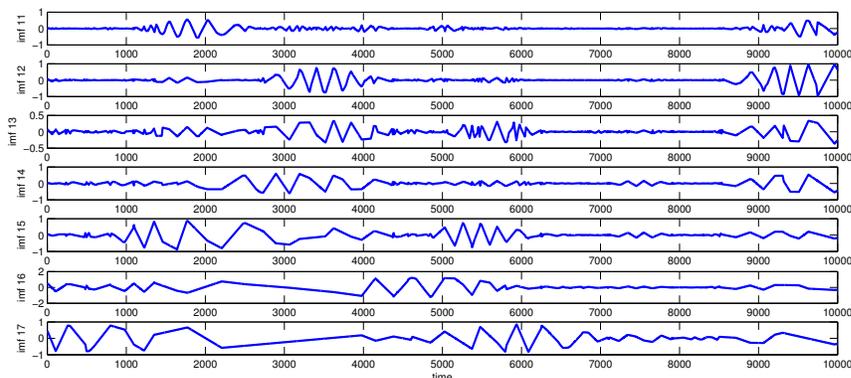

Figure 1: Some of the significant mode obtained for the decomposition of the signal of Figure 3.(A)

3. First approximation of the mean envelope

In this section, we first investigate several methods that were proposed to compute a signal mean envelope approximation in relation with the EMD algorithm. The common idea of these approaches is to seek particular points of the mean envelope which serve as knots. The approximation is then either defined through splines interpolation [4] or piecewise cubic polynomials interpolation [8]. Usually, the knots are approximation of the intersection points of the signal with its mean envelope [4][3], but in some approaches the knots are located at the extrema of the signal [8].

3.1. Determination of points of interest on the mean envelope

Here, we adopt the point of view of [4] or [3], that is we look for the intersection of the signal with its mean envelope, the mean envelope approximation
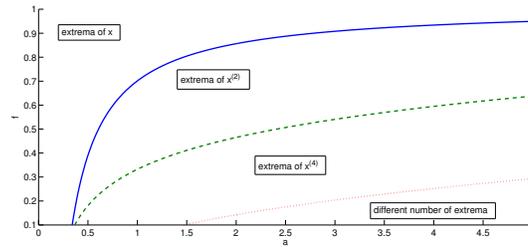
Figure 2. Different regions for the best approximations of θ by $t, t^{(2)}, t^{(4)}$

being defined by spline interpolation. We now propose an improved version of the method developed in [4] which will compare favorably to the existing ones. In what follows, $x(t_j)$ denotes the set of values corresponding to the extrema of x located at t_j. In [4], the mean envelope is assumed to interpolate $(\bar{t}_j, \bar{x}_j)$ defined by:

$$\bar{x}_j = \frac{1}{t_{j+1} - t_j} \int_{t_j}^{t_{j+1}} x(t)dt, \ \bar{t}_j = \int_{t_j}^{t_{j+1}} \frac{t(x(t) - \bar{x}_j)^2 dt}{\int_{t_j}^{t_{j+1}} (x(t) - \bar{x}_j)^2 dt}. \tag{3.1}$$

The choice for $\bar{x}_j$ is motivated by the fact that, if the extrema of the signal corresponded to those of the first mode x_1, putting $x = x_1 + e$ and assuming the integral of a mode is null between its extrema we would obtain: $\bar{x}_j = \frac{1}{t_{j+1} - t_j} \int_{t_j}^{t_{j+1}} e(t)dt$, which is indeed a point of the mean envelope since, from the classical mean theorem for continuous functions, there exist $d \in [t_j, t_{j+1}]$ such that: $\frac{1}{t_{j+1} - t_j} \int_{t_j}^{t_{j+1}} e(t)dt = e(d)$.

However, there exist two strong limitations to that model which are the following: the extrema of x do not, in general, coincide with those of x_1 and the second, the integral of x_1 between its successive extrema does not necessarily vanish (think of frequency modulated signals for instance). In order to discuss the above remarks, we consider the following example:

$$x(t) = \cos(2\pi t) + a \cos(2\pi ft) \quad 0 < f < 1 \ a \in \mathbb{R}_+. \tag{3.2}$$

We first focus on the determination of the locations of the extrema of x_1 given x. For the signal x defined in (3.2), $x_1 = \cos(2\pi t)$. We study, in this particular example, the approximation of the locations of the extrema of x_1 by those of even order derivative of x. As remarked in [6], the extrema of the second order derivative is more likely to give a better approximation of the location of the extrema of x_1 than the extrema of the original signal. However, the proof given in [6] only involves maximal deviation and do not correspond to the general case which we now study. Let us denote by θ the locations of the extrema of x_1 and by $t^{(2m)}$ the locations of the extrema

of $x^{(2m)}$ the $2m$th order derivative of x. Let us recall that $\#X$ denotes the cardinality of the set X. To illustrate our strategy to find the best approximation of the locations θ, we only consider $m = 1$ and $m = 2$ the extension to higher order derivatives being straightforward. By using either $t, t^{(2)}$ or $t^{(4)}$ to approximate θ, we use the following procedure:

- If $\#\theta = \#t = \#t^{(2)} = \#t^{(4)}$, then we compute :

$$\hat{\theta} = \underset{l \in t, t^{(2)}, t^{(4)}}{\operatorname{argmin}} \frac{1}{\#\theta} \sum_j |\theta_j - l_j|$$

- otherwise, if $\#\theta = \#t = \#t^{(2)}$, then we compute :

$$\hat{\theta} = \underset{l \in t, t^{(2)}}{\operatorname{argmin}} \frac{1}{\#\theta} \sum_j |\theta_j - l_j|,$$

  and we proceed similarly for the two other cases ($\#\theta = \#t^{(2)} = \#t^{(4)}$ and $\#\theta = \#t = \#t^{(4)}$).

- otherwise, if $\#\theta = \#t$ then $\hat{\theta} = t$ (and similarly if $\#\theta = \#t^{(2)}$, $\hat{\theta} = t^{(2)}$ or $\#\theta = \#t^{(4)}$, $\hat{\theta} = t^{(4)}$)

- otherwise, the cardinality of any of t, $t^{(2)}$ or $t^{(4)}$ is different from that of θ, and we cannot conclude only using $x^{(2)}$ and $x^{(4)}$ and we need to consider higher order derivatives.

The results of the proposed procedure are displayed on Figure 2, for the signal defined in (3.2). These numerical simulations show that the best way to approximate θ strongly depends on a and f. In general, θ is unknown and may not be equal to t. In that case, the EMD algorithm sometimes manages to compute, from x, a first mode having the correct number of extrema, but the underlying process is unclear. The numerical simulations of Figure 2. show that we can retrieve the number of extrema of x_1 by considering those of higher even order derivatives of x. More precisely, in the above example, to find the number of the extrema of x_1, one considers the smallest order m_0 such that $\#t^{(2m_0)} = \#t^{(2m_0+2)}$, for which we always have $\#t^{(2m_0)} = \# \theta$. In that context, the estimation of θ should be computed using only $t^{(2m)}$, $m \geq m_0$. With this in mind, we will assume $\#t = \#\theta$, of which a particular example is given by the signal x, sum of two AM/FM modulated functions x_1 and x_2, displayed on Figure 3. In this case, the first mode is known to be x_1 while the sought mean envelope is x_2. We then plot the values of x_1 at t, $t^{(2)}$ and $t^{(4)}$ (see Figure 3 (B)). If we assume that the extrema of x_1 are known, we consider the following error measurement: $E(l) = \frac{1}{\#\theta} \sum_j |\theta_j - l_j|$, For the signal of Figure 3.(A), we get $E(t) = 15.83$, $E(t^{(2)}) = 8.34$ and $E(t^{(4)}) = 27.07$. A more careful look at these numerical results shows that
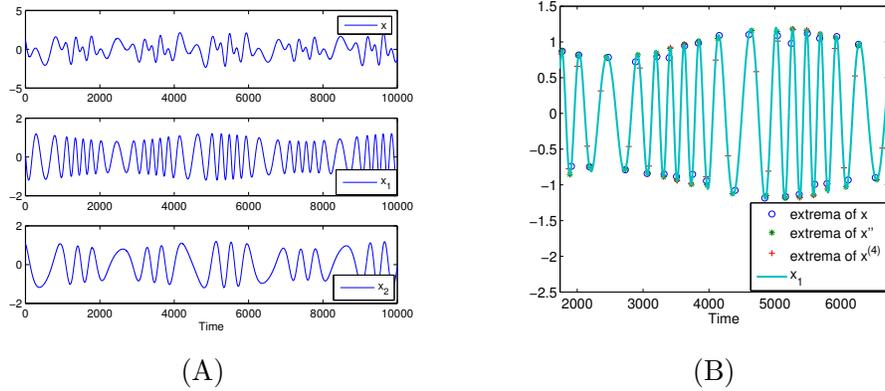
Figure 3: (A): signal x made of two modulated signals x_1 and x_2. (B): signal x_1 and its values at t, $t^{(2)}$ and $t^{(4)}$.

the error using $t^{(4)}$ as an estimator of θ is much larger when the frequency of x_1 is locally lower, which is in accordance with the results displayed on Figure 2. On the contrary, when the frequency of x_1 is locally higher, the estimation is better using the extrema of higher order derivatives (note that such results are obtained because the amplitude modulation is between 0.5 and 1.5 both for x_1 and x_2, see Figure 2). In what follows, $\hat{\theta}$ will denote the locations of the extrema of the even order derivative chosen to approximate θ.

3.2. Initial guess for the mean envelope

The study of the previous subsection suggests that the model given by (3.1) should be replaced by $x_{j,\hat{\theta}}$

$$x_{j,\hat{\theta}} = \frac{1}{\hat{\theta}_{j+1} - \hat{\theta}_j} \int_{\hat{\theta}_j}^{\hat{\theta}_{j+1}} x(t)dt, \ t_{j,\hat{\theta}} = \frac{\int_{\hat{\theta}_j}^{\hat{\theta}_{j+1}} t(x(t) - x_{j,\hat{\theta}})^2 dt}{\int_{\hat{\theta}_j}^{\hat{\theta}_{j+1}} (x(t) - x_{j,\hat{\theta}})^2 dt}.$$

Note that $(t_j(\hat{\theta}), x_{j,\hat{\theta}})$ does not necessarily belong to the graph of x. Therefore, we replace $t_j(\hat{\theta})$ by the abscissa c_j obtained by the mean theorem, namely:

$$\exists c_j \text{ s.t.} \quad \frac{1}{\hat{\theta}_{j+1} - \hat{\theta}_j} \int_{\hat{\theta}_j}^{\hat{\theta}_{j+1}} x(t)dt = x(c_j) = x_{j,\hat{\theta}_j}. \tag{3.3}$$

Then, if one defines the mean envelope as the cubic spline interpolant at $(c, x(c)) := (c_j, x(c_j))_j$, the signal and its approximative mean envelope

coincide at that point. To illustrate these improvements, we now compare the mean envelope computed using $(c, x(c))$ with those proposed in [8], in [3] and in [4] (recalled in (3.1)). For the sake of consistency, let us recall the methods defined in [8] and in [3]. In the latter, the mean envelope approximation intersects the inflection points of the signal, while in [8], the mean envelope approximation satisfies the interpolation conditions that follows. Assume that $x(t_j)$ is a minimum for x and that $x(t_j)$ is an extremum for the sequence $(x(t_{j-2}), x(t_j), x(t_{j+2}))$. Depending on the cases, the shape of the upper (resp. lower) envelope is used to derive that of the lower (resp. upper). Indeed, Let $\tilde{t}_j$ be abscissae of the intersection (when it exists) of straight lines L_1 and L_2 defined by $L_1 : f_1(t) = \frac{x(t_j) - x(t_{j-2})}{t_j - t_{j-2}} t + \frac{x(t_{j-1})(t_j - t_{j-2}) - (x(t_j) - x(t_{j-2}))t_{j-1}}{t_j - t_{j-2}}$ and $L_2 : f_2(t) = \frac{x(t_{j+2}) - x(t_j)}{t_{j+2} - t_j} t + \frac{x(t_{j+1})(t_{j+2} - t_j) - (x(t_{j+2}) - x(t_j))t_{j+1}}{t_{j+2} - t_j}$. If $\tilde{t}_j > t_j$, one imposes : $(1/2)(f_1(t_j) + x(t_j)) = e(t_j)$. Otherwise, one sets $(1/2)(f_2(t_j) + x(t_j)) = e(t_j)$. Note, that when the sequence $(x(t_{j-2}), x(t_j), x(t_{j+2}))$ is monotonic no interpolation condition is imposed on the mean envelope.

We now carry out the comparison of these methods on the modulated sinus functions displayed on Figure 3.(A) The estimation error is computed through the following quantity: $E_f(l, \tilde{x}) = \sqrt{\frac{\sum_j (x_2(l_j) - \tilde{x}_j)^2}{\sum_j (x_2(l_j))^2}}$. By considering that inf denotes the inflection points of x, the results are as follows: $E_f(inf, x(inf)) = 0.2712$, $E_f(\bar{t}, \bar{x}) = 0.3423$, $E(c, x(c)) = 0.1876$, (where $\hat{\theta} = t^{(2)}$). The method developed in [8] leads to significantly worse results. Finally, we assume that the extrema of x_1 are known and we compute (3.3) replacing $\hat{\theta}$ by θ: we obtain a new $\tilde{c}$ such that $E_f(\tilde{c}, x_{j,\theta}) = 0.1246$ which means that $x(\tilde{c}) = x_{j,\theta}$ is not necessarily a point of the sought mean envelope. However, the estimation of the intersection points given by (3.3) is helpful in the following sense: our concern is to define a strategy to find the intersection points of the signal with its mean envelope starting from the initial guess $(c, x(c))$, which we do in the following section.
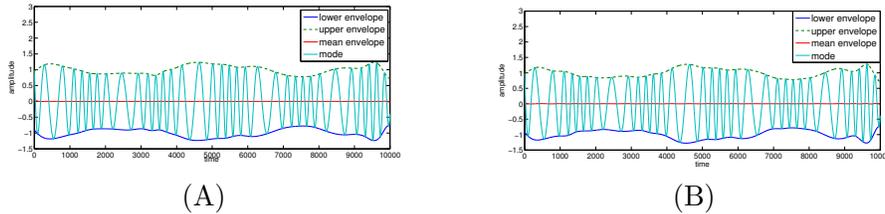

(A) (B)

Figure 4: The signal m^f obtained from the signal x of Figure 3.(A) as well as its upper and lower envelopes, for one interval strategy (A) and for three intervals strategy (B).
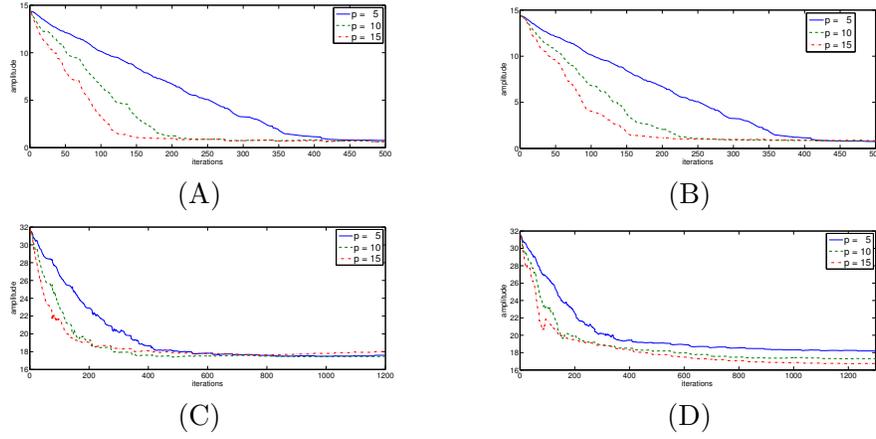
(A)   (B)

(C)   (D)

Figure 5: L^2 norm of m^j with respect to j, for one interval strategy (A) and for three intervals strategy (B). (C) and (D): L^2 error of $x_1 - m^j$ with respect to j, for one interval strategy (C) and for three intervals strategy (D).

4. Optimization procedure for the estimation of the mean envelope

The optimization procedure proposed here computes an approximation of the mean envelope of x. Our approach is based on the fact that, for each c_j, (defined by (3.3)), there exists an intersection point of x and of its mean envelope in the vicinity of the points $(c_j, x(c_j))$. Note that contrary to the approach of [4], where the approximation of the mean envelope serves as initial guess to the SP, we replace the SP by an optimization procedure. In what follows, we will assume that $\#\hat{\theta} = \#\theta$ and that x is sampled at the rate Δt. We denote for any signal x, by $e(x)$ its mean envelope obtained as $\frac{u(x)+l(x)}{2}$, where $u(x)$ (resp. $l(x)$) is the upper (resp. lower) envelope of x, obtained as usual by cubic spline interpolation of its maxima (resp. minima).

4.1. Initialization part

We start from $(c^1, x(c^1)) = (c, x(c))$, which defines the mean envelope e^1 of x through cubic spline interpolation. In relation with the classical EMD algorithm, we notice that e^1 should be such that $e(m^1)$ (with $m_1 = x - e^1$, having extrema located at τ^1), has a L^2 norm which should be close to zero.

Our strategy is then to modify $(c^1, x(c^1))$ to obtain $(c^2, x(c^2))$ leading to e^2 in such a way that the L^2 norm of $e(m^2)$ (with $m^2 = x - e^2$ having extrema τ^2) is lesser than that of $e(m^1)$. To do so, we seek c_j^2 in the interval $[c_j^1 - p\Delta t, c_j^1 + p\Delta t]$ (the choice for p will be discussed later) for each j, following

a certain optimization criterion based on the L^2 norm of $e(m^2)$. Note that the optimal solution lies in a finite dimensional space which ensures the convergence of our procedure. To explore all the possibilities in this finite dimensional space will be computationally prohibitive. Therefore, we decide to move the points $(c_j^1, x(c_j^1))$ sequentially using one of the following two strategies:

- <u>one (resp. three) interval(s) strategy</u>

  - find the interval $[\tau_{j_0}^1, \tau_{j_0+1}^1]$ (resp. $[\tau_{j_0-1}^1, \tau_{j_0+2}^1]$), minimizing locally the L^2 norm of $e(m^1)$.
  - put $c_j^2 = c_j^1, j \neq j_0$ and find $c_{j_0}^2 \in [c_{j_0}^1 - p\Delta t, c_{j_0}^1 + p\Delta t]$, such that the L^2 norm of $e(m^2)$ is minimal over the interval $[\tau_{j_0}^2, \tau_{j_0+1}^2]$ (resp. $[\tau_{j_0-1}^2, \tau_{j_0+2}^2]$).

These two strategies are designed to test the influence of the neighborhood on the computation of the optimal mean envelope.

4.2. Loop part

Once c^2 has been computed the interval indexed by j_0 is marked and we pass on other intervals following the same strategy, never considering the interval indexed by j_0 again. After $\#\theta$ steps, all the intervals have been marked and we obtain a certain set $(c^{\#\theta}, x(c^{\#\theta}))$. Then, we unmark all the intervals and we start again the procedure from the initialization step with $(c, x(c)) = (c^{\#\theta}, x(c^{\#\theta}))$. We run the algorithm until convergence.

5. Numerical results

We give an illustration of the two strategies on the signal x of Figure 3.(A) assuming $\hat{\theta} = t^{(2)}$. The results are shown on Figure 4, where we depict the mode m^f (corresponding to the computed mode at the last iteration) as well as its upper, lower and mean envelope, for both strategies, when $p = 15$. The results show very little differences between the obtained modes. We show on Figure 5.(A) and (B), the decrease of the L^2 norm of the mean envelope of m^j computed with both strategies for $p = 5, 10, 15$ with respect to the iteration. Again, the results show very little differences between the two strategies. On Figure 4.(C) and (D), we check that with the proposed strategies, the error between x_1 and m^j decreases in L^2 norm, again for $p = 5, 10, 15$. We notice that the convergence of m^j to x_1 is improved with the second strategy (three intervals strategy) when p is larger which is not the case with the first strategy. This suggests that to consider a larger interval for the minimization of the L^2 norm of $e(m^j)$ leads to smoother upper and lower envelopes for m^j. Note also that with the second strategy

when p is increased the convergence to x_1 is improved, which is in accordance which what we expected.

6. Conclusion

In this paper we first discussed the approximation of the extrema of the first mode in the EMD method by the extrema of even order derivatives. This study enabled us to build a signal mean envelope approximation, which we then used as initial guess for a new optimization procedure to improve the approximation. This procedure replaces advantageously the SP of the original EMD for AM/FM signals as numerical results show. Future work should involve the improvement of the computational cost which is beyond the scope of present article.

Acknowledgments

References

[1] Q. Chen, N. Huang, S. Riemenschneider and Y. Xu, A B-spline approach for empirical mode decompositions, *Advances in Computational Mathematics*, **24** (2006), 171-195.

[2] L. Cohen, *Time Frequency Analysis*, Prentice Hall, Englewoods Cliffs, 1995.

[3] E. Deléchelle, J. Lemoine and O. Niang, Empirical mode decomposition: an analytical approach for sifting process, *IEEE Signal Processing Letters*, **12**, *11* (2005), 764-767.

[4] H. Hong, X. Wang and Z. Tao, Local mean-based sifting for empirical mode decomposition, *IEEE Signal Processing Letters*, **6**, *10* (2009).

[5] N.E. Huang, Z. Shen, S. R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung and H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non stationary time series analysis, *Proceedings of the Royal Society London A*, pp. 903-1005, 1998.

[6] Y. Kopsinis and S. Mac Laughlin, Investigation and enhancement of the empirical mode decomposition method based on a heuristic search optimization approach, *IEEE Transactions on Signal Processing*, **56**, *1* (2008), 1-13.

[7] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[8] S. Meignen and V. Perrier, A new formulation for empirical mode decomposition based on constrained optimization, *IEEE Signal Processing Letters*, **14**, *12* (2007), 932-935.

[9] S. Peng and W-L. Hwang, Adaptive signal decomposition based on local narrow band signals, *IEEE Transactions on Signal Processing*, **56**, *7* (2008).

[10] G. Rilling and P. Flandrin, One or two Frequencies? The empirical mode decomposition answers, *IEEE Transactions on Signal Processing*, **56**, *1* (2008), 85-95.

[11] G. Rilling, P. Flandrin and P.Goncalvès, *On empirical mode decomposition and its algorithms*, IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, NSIP-03, Grado (I), June 2003.

[12] R.C. Sharpley and V. Vatchev, Analysis of the intrinsic mode functions, *Construct. Approx.*, **24** (2006), 17-47.

Basarab Matei
LAGA, Université Paris XIII
99 Avenue Jean-Baptiste Clément
Villetaneuse - France
E-mail: `matei@math.univ-paris13.fr`

Sylvain Meignen - corresponding author
LJK, Université Joseph Fourier
51 rue des Mathématiques
Grenoble - France
E-mail: `sylvain.meignen@imag.fr`