

INVESTIGATION OF THE EMPIRICAL MODE DECOMPOSITION BASED ON GENETIC ALGORITHM OPTIMIZATION SCHEMES

Yannis Kopsinis, Stephen (Steve) McLaughlin

IDCOM, School of Engineering and Electronics
The University of Edinburgh, King's Buildings, EH9 3JL, Edinburgh

ABSTRACT

Empirical mode decomposition (EMD) has lately received much attention due to the many interesting features that exhibits. However it lacks a strong theoretical basis which would allow a performance analysis and hence the enhancement and optimization of the method in a systematic way. In this paper, an investigation of EMD is attempted in an alternative way: The interpolation points and the piecewise interpolating polynomials for the formation of the upper and lower envelopes of the signal are optimized based on a genetic algorithm framework revealing important characteristics of the method which were previously hidden. As a result, novel directions for both the performance enhancement and the theoretical investigation of the method are unveiling.

Index Terms— Empirical Mode Decomposition, Genetic Algorithm

1. INTRODUCTION

Empirical mode decomposition (EMD) [1] is a relatively new, data-driven adaptive technique for analyzing multicomponent signals into a number of elementary amplitude and frequency modulated (AM/FM) zero mean signals termed intrinsic mode functions (IMFs). The task of EMD is to iteratively reveal locally the slow oscillating part of the signal according to a procedure called sifting which involves the computation of an upper and a lower envelope which *enfold* the signal [2]. Although the validity and the robustness of EMD have been shown in a number of applications [3], [4], it lacks a well established theoretical analysis which would permit a convergence proof and a direct, systematic optimization of the method. Due to the nature of EMD and the obscure way it operates, the so far published modifications of the initially proposed algorithm leading to performance improvement are limited [5], [6].

There are two crucial issues related to EMD namely, the best interpolation method [7] and the optimized positioning of the interpolation points. An answer to the above open problems can lead to a better understanding of the sifting process and subsequently can reveal ways to the performance enhancement of the algorithm. In this paper EMD is examined with the aid of appropriately designed multicomponent signals which allow as to know explicitly the optimum outputs that the EMD should “ideally” provide. The knowledge of the desired EMD outputs allow the use of genetic algorithm searching techniques for the estimate of parameters related to optimized IMF extraction.

This work was performed as part of the BIAS consortium under a grant funded by the EPSRC under their Basic Technology Programme.

2. EMD ALGORITHM

The task of EMD is to decompose a multicomponent signal $x(t)$ into a number of virtually monocomponent IMFs. Each one of them, say the first one $h(t)$, is obtained iteratively through the sifting process. According to this, during the $(n + 1)$ th sifting iteration the temporal IMF estimate $\hat{h}_n(t)$ is getting improved according to the next steps¹ (some of the quantities derived are shown in Fig. 1):

1) Specify some time instances $\tau_u = [\tau_{u,1}, \dots, \tau_{u,M}]$, $\tau_l = [\tau_{l,1}, \dots, \tau_{l,L}]$ called nodes and the corresponding IMF values $\hat{\mathbf{h}}_{u_n} = [\hat{h}_n(\tau_{u,1}), \dots, \hat{h}_n(\tau_{u,M})]$, $\hat{\mathbf{h}}_{l_n} = [\hat{h}_n(\tau_{l,1}), \dots, \hat{h}_n(\tau_{l,L})]$ called interpolation points. These points are utilized in the formation of two envelopes, an upper one and a lower one, which enfold the temporal estimate \hat{h}_n .

2) Interpolate, according to a predetermined scheme, e.g. natural cubic spline interpolation, the points defined in the first step in order to form the upper $I_{\tau_u}(t)$ and the lower envelopes $I_{\tau_l}(t)$ and compute the *mean envelope* $m_{n+1}(t) = (I_{\tau_u}(t) + I_{\tau_l}(t))/2$.

3) Obtain the refined estimate of the IMF as $\hat{h}_{n+1}(t) = \hat{h}_n(t) - m_{n+1}(t)$ and proceed from step 1 again unless a stopping criterion has been fulfilled. In that latter case, set $h(t) = \hat{h}_{n+1}(t)$.

Exactly the same procedure is then applied to the residual signal $x'(t) = x(t) - h(t)$ for the extraction of the next IMF. For this reason the analysis in this paper is focused on the extraction of the first IMF only.

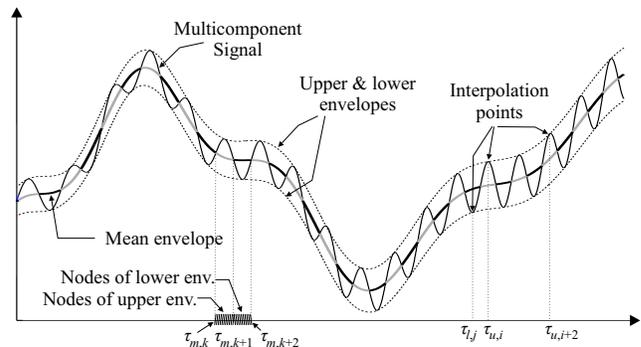


Fig. 1. A multicomponent signal and quantities related to EMD.

Ideally, the final result of each sifting process is the extraction, along the time axis, of the faster oscillations preserving at the same time their amplitude. For example, consider the case where the complex signal consists of N AM-FM modulated monocomponent sig-

¹For the first iteration, $x(t)$ is used as temporal IMF estimate $\hat{h}_{l_1}(t)$.

nals

$$x(t) = \sum_{i=1}^N s_i(t) = \sum_{i=1}^N \alpha_i(t) \cos(2\pi f_i(t)t). \quad (1)$$

In this scenario, EMD should be able to extract, at any time instant, this signal out of N that locally has a highest frequency. Under the assumption of $f_i(t) > f_j(t)$ for any $i < j$ and for all time instances t , the first desired estimated IMF has to be equal to $s_1(t)$. The above assumption will now be adopted since it does not affect the analysis or the findings of this work which are still valid in the case of frequency-overlapped signal components.

Considering that $\hat{h}_1(t) = x(t) - m_1(t)$, $\hat{h}_2(t) = \hat{h}_1(t) - m_2(t)$, \dots , $\hat{h}_n(t) = \hat{h}_{n-1}(t) - m_n(t)$, the estimate of $h(t)$ after n sifting iterations can be written as

$$\hat{h}_n(t) = x(t) - (m_1(t) + \dots + m_n(t)) = x(t) - M_n(t) \quad (2)$$

with $M_n(t)$ defined here as the *total mean envelope*. Equation (2) reformulates the sifting process as an iterative procedure for the estimation of the slow oscillating *local mean* of the multicomponent signal which, as will be seen, is a perspective that offers certain advantages. Assuming that we require the decomposition to be exact then the optimum total mean envelope which leads to the accurate decomposition of the first signal, i.e., $h(t) = x(t) - M_{opt}(t) = s_1(t)$, is given by the sum of the $L - 1$ lower frequency signals $M_{opt}(t) = \sum_{i=2}^N s_i(t)$.

A question that arises with respect to EMD is: Are there optimum interpolation points having specific properties, and if yes, how well are they approximated by the *local extrema*, i.e., the maximum and minimum extrema which are usually adopted in practice? The intersection points, τ_m , between the optimum mean envelope and the signal provide useful information about the possible positions of the optimum interpolation points:

$$\begin{aligned} \tau_m = [\tau_{m,1}, \dots, \tau_{m,Q}] &= \{t : M_{opt}(t) = x(t)\} \\ \Rightarrow \tau_m &= \{t : s_1(t) = 0\}. \end{aligned} \quad (3)$$

The interpolation points of the upper and the lower envelope, should be selected in a way capable of leading to estimated mean envelope equal to $M_{opt}(t)$. At the same time, the upper and the lower envelope should be over and under the mean envelope respectively. Investigating Fig. 1, it can be argued that, each one of the intervals $(\tau_{m,i}, \tau_{m,i+1})$ can only contain interpolation points which belong exclusively to either the upper or the lower envelope depending on whether the residual signal is “over” or “under” $M_{opt}(t)$. In Fig. 1, the intervals that correspond to upper or lower envelopes are shown with gray or black mean envelope respectively. In the next section, the optimum interpolation points will be detected by means of a genetic algorithm (GA) search procedure.

3. GA-BASED EMD OPTIMIZATION: BEST INTERPOLATION POINTS DETECTION

In section 2 it was seen that the use of appropriately designed multicomponent signals allow us to know in advance the intervals along the time axes with nodes which explicitly correspond to either the upper or the lower envelope. This fact will be exploited in order to efficiently detect the best ones among all the possible interpolation points by means of a GA-based search procedure. Note here, that a modified EMD with the GA embedded in it is not being proposed here. The GA is simply being used in order to investigate the way that EMD processes the signals and to reveal directions leading to performance enhancement.

Consider that there are M intervals that are able to contain upper envelope nodes $T_{u,k} = \{(\tau_{m,i}, \tau_{m,i+1}), i : M_{opt}(t) < x(t) \text{ for } t \in (\tau_{m,i}, \tau_{m,i+1})\}$, $1 \leq k \leq M$ and L intervals that are able to contain lower envelope nodes $T_{l,n} = \{(\tau_{m,i}, \tau_{m,i+1}), i : M_{opt}(t) > x(t) \text{ for } t \in (\tau_{m,i}, \tau_{m,i+1})\}$, $1 \leq n \leq L$. Also assume that in each interval there is explicitly one node. It is easy to realize that this assumption in general results in a number of nodes similar or somewhat larger than the one that results from the standard local extrema case. The optimization which is accomplished with the aid of the GA is as follows: Minimize the relative error between the actual mean envelope $M_{opt}(t)$ and the estimated total mean envelope $M_n(t)$ after a preset number of n sifting iterations with respect to the adopted upper and lower envelope node vectors τ_u and τ_l respectively.

The i th chromosome in the adopted GA has the form $\chi_i = [\tilde{\tau}_{u,1}, \tilde{\tau}_{u,2}, \dots, \tilde{\tau}_{u,M}, \tilde{\tau}_{l,1}, \tilde{\tau}_{l,2}, \dots, \tilde{\tau}_{l,L}]$, where $\tilde{\tau}_{,k}$ denotes a random value in the interval $T_{,k}$. The fitness function comprises a number of steps:

- 1) Split the chromosome in the upper and lower envelope node vectors τ_u and τ_l and compute the corresponding interpolation points.
- 2) Perform n times the sifting iteration steps 2 and 3 (see section 2) using in each iteration the same nodes as they were defined by the chromosome under consideration in order to estimate the total mean envelope $M_n(t)$ from (2).
- 3) Compute the chromosome fitness with the following error function applied to the sampled signals:

$$\begin{aligned} F &= \sum_{i=1}^K |M_{opt}(i) - M_n(i)|dt + \\ &\quad \sum_{i: I_{\tau_u}(i) < h_{n-1}(i)} |h_{n-1}(i) - I_{\tau_u}(i)|dt + \\ &\quad \sum_{t: I_{\tau_l}(t) > h_{n-1}(t)} |h_{n-1}(t) - I_{\tau_l}(t)|dt \end{aligned} \quad (4)$$

The first term deals with the main objective of the optimization procedure, i.e. to provide accurate mean envelope estimates, while the second and the third terms guarantee that the interpolation functions will be envelopes which will “tightly include” the processed signal.

The application of the GA optimization scheme presented above will be realized in a multicomponent signal constituted from the sum of the three monocomponent signals shown in Fig. 2. The GA is used in order to detect the upper and lower envelope nodes that optimize the extraction of the first monocomponent signal after the first sifting iteration. The error bars graphs in Fig. 2(a)-(j) show the relative error between the optimum and the estimated mean envelope along the time axis where the dark and the light areas correspond to high and low error values respectively. The error bars are associated with three different sets of interpolation points: The points detected by the GA (a1, b1, c1), the local extrema of the desired monocomponent signal $s_1(t)$ which is about to get extracted (a2, b2, c2), and the local extrema of the multicomponent signal $x(t)$ (a3, b3, c3, c4) which actually correspond to the standard method adopted in practice. Moreover, error bars (a1-a3) and (b1-b3) correspond to 3rd and 5th order natural splines interpolation methods respectively. It can be easily observed, particularly in the higher order spline cases, that the optimized interpolation points estimated by the GA tend to coincide with the extrema of the signal which is about to get extracted. Hereafter, we will refer to the extrema, nodes and interpolation points of the desired signal to be extracted as *desired* and to the extrema, nodes and interpolation points estimated by a GA procedure as *optimized*.

Error bars 2(c1-c4) depict the relative error after 5 sifting operations and 3rd order splines. In (c1-c3) the optimized, the desired

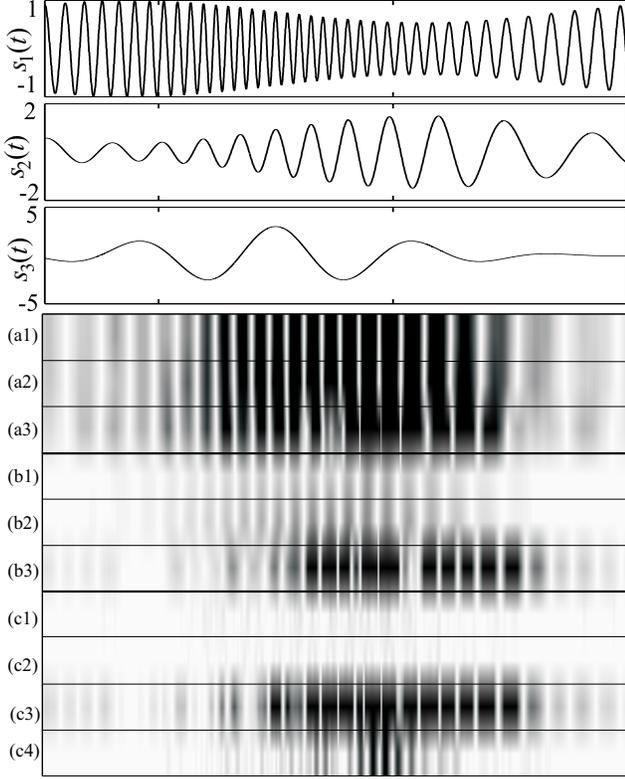


Fig. 2. Relative error between the estimated and the actual mean envelope after the first IMF extraction of a signal consisted of the three monocomponent signals shown on top.

or the interpolation points resulted from the local extrema of the multicomponent signal are set constant in every sifting iteration. In contrast to that, in the case of error bar c4, which corresponds to the standard EMD method, the extrema of the multicomponent signal are recomputed in each sifting iteration. Here is the power of the EMD method, which although adopting interpolation points far from the best ones, generally succeeds in converging to them. When the extrema obtained from the multicomponent signal serve as interpolation points throughout the rest of the sifting iterations, the results are disappointing (see Fig. 2(c4)).

In practice, neither the extrema of the desired signals are known, nor can they be estimated with the aid of a GA. The maxima and minima of the multicomponent signal serve as estimates to the desired interpolation points and it should be expected that the closer the estimated extrema are to the extrema of the higher frequency signal is the better the "extraction" of this signal will be. The development of specific optimized methods for the interpolation points is beyond the scope of this paper and is left for future research.

4. GA-BASED EMD OPTIMIZATION: INVESTIGATION ON IMPROVED INTERPOLATION SCHEMES

Although in Fig. 2 it was observed that incrementing the order of the natural splines leads to an improvement in the performance of EMD this is not a general conclusion and depends on the frequency difference between the fast oscillating monocomponent signal and the rest of them. In this section, the GA-based optimization of the

piecewise polynomials which form the envelopes will be realized while the interpolation points will stay fixed to the desired ones.

For the construction of the envelopes, e.g. the upper one, it is assumed that the corresponding interpolation points $x(\tau_{u,i})$, $1 \leq i \leq L$ are linked with $L - 1$ 4th order polynomial curves² $P_i(t) = a_{u,i}t^4 + b_{u,i}t^3 + c_{u,i}t^2 + d_{u,i}t + e_{u,i}$, $\tau_{u,i} \leq t \leq \tau_{u,i+1}$. Moreover, the polynomials share the following properties

$$P_i(\tau_{u,i}) = x(\tau_{u,i}) \quad (5)$$

$$P_i(\tau_{u,i+1}) = x(\tau_{u,i+1}) \quad (6)$$

$$\frac{dP_i(t)}{dt} \Big|_{t=\tau_{u,i}} = \frac{dP_{i-1}(t)}{dt} \Big|_{t=\tau_{u,i}} \quad (7)$$

The continuity in the first derivative guarantees at least the minimum smoothness at the transitions between the polynomials at the interpolation points.

Each chromosome in the GA has $2 \times (L + M)$ genes $\chi = [D_{u,1}^{(1)}, \dots, D_{u,L}^{(1)}, D_{u,1}^{(2)}, \dots, D_{u,L}^{(2)}, D_{l,1}^{(1)}, \dots, D_{l,M}^{(1)}, D_{l,1}^{(2)}, \dots, D_{l,M}^{(2)}]$ which represent the values of the first and the second derivatives of the polynomials at the interpolation points, i.e.:

$$\frac{dP_i(t)}{dt} \Big|_{t=\tau_{u,i+1}} = D_{u,i+1}^{(1)} \quad (8)$$

$$\frac{d^2P_i(t)}{dt^2} \Big|_{t=\tau_{u,i+1}} = D_{u,i+1}^{(2)} \quad (9)$$

Equations (5)-(9) form $L - 1$ linear systems of equations from which the unknown parameters

$\{a_{u,i}, b_{u,i}, c_{u,i}, d_{u,i}, e_{u,i}\}$ and hence, the upper envelope can be easily computed. Using the last $2 \times M$ genes of the chromosomes, the lower envelope can be computed in a similar way. The mean envelope is computed for all the chromosomes and their fitness is given by $F = \sum_{i=1}^K |M_{opt}^{(1)}(t) - m_1^{(1)}(t)| dt$. The task of GA is to search for the optimum derivatives that the piecewise polynomial should have at the interpolation points. Presumably, the derivative values determines the shape of the envelopes. The above GA-based

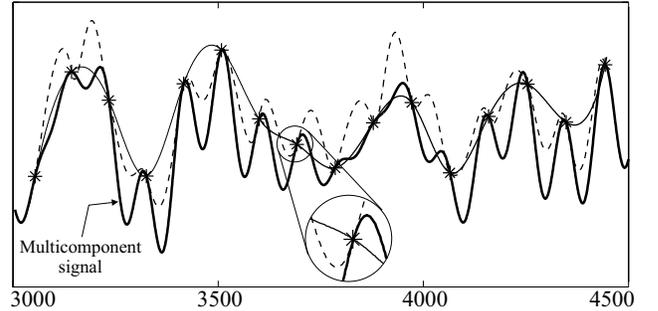


Fig. 3. Upper envelope constructed either by natural cubic spline interpolation (thin curve) or by GA-optimized piecewise polynomial interpolation (dashed curve).

optimization procedure is applied on a portion of the signal depicted in frequency domain on the top of Fig. 4, namely from 3000 to 4500 sec. This signal portion is shown in time domain in Fig. 3 with thick line. In the same figure, the thin line and the dashed line correspond to the upper envelope when it is constructed using natural cubic splines or the polynomials optimized by the GA respectively.

²The fourth order that is adopted here is not restrictive, and the same procedure could be realized with polynomials of higher or lower order.

It is not hard to observe the difference in the behavior of the two envelopes. The spline envelope often crosses the signal, whereas the GA optimized envelope usually stays above it, and more importantly *it tends to be tangential to the signal at the interpolation points* which are shown by stars. In other words, at least the first derivatives of the envelope at the interpolation points coincide with these of the multicomponent signal. The above observation brings directly to mind Hermite piecewise polynomial interpolation [8] which allows the determination of the value of derivatives at selected points.

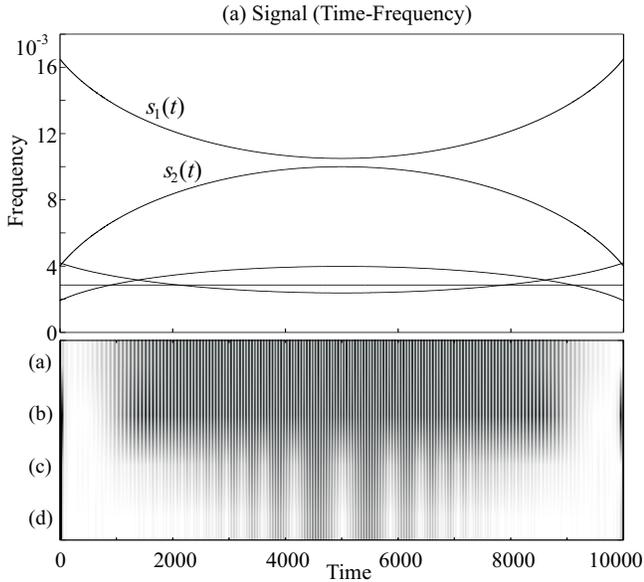


Fig. 4. Performance enhancement achieved with the use of Hermite instead of natural splines interpolation.

The performance improvement that can be achieved with the aid of Hermitian interpolation is shown in Fig. 4. The first two error bars show the relative error when 3rd order and 9th order spline interpolation has been used. The third error bar corresponds to the cubic Hermite interpolation with the first derivatives of the interpolation polynomials at the interpolation points to be set equal to the estimated first derivatives of the multicomponent signal. The fourth error bar corresponds to the 5th order Hermite interpolation where the derivatives higher than the first order are equal between the polynomial pieces in a similar way to the natural splines way. In fact, the improvement is even better as the order of interpolation increases.

The mean square error (MSE) performance³ in the extraction of the highest frequency signal with respect to the number of sifting iterations is depicted in figure 5 for the case of the signal shown in Fig. 4. Sampling frequency issues are not taken into account here [2] and the multicomponent signal has been sampled at least 50 times faster than the Nyquist frequency. In this figure, the 4 first curves correspond to 3rd and 7th order natural spline interpolation with local extrema (1), (2), 3rd and 7th order natural splines and desired extrema (3), (4) and the rest of them correspond to Hermite interpolation. We observe that for the case of desired extrema, the higher the order of the Hermitian interpolation the better the decomposition performance is (see curves (7),(8)). However this trend is not followed in the case of local extrema where the 3rd order Hermitian interpolation performs better than the 7th order one. In general,

³ $MSE = \frac{1}{N} \sum_{n=1}^N (h(n) - s_1(n))^2$, with N the number of samples.

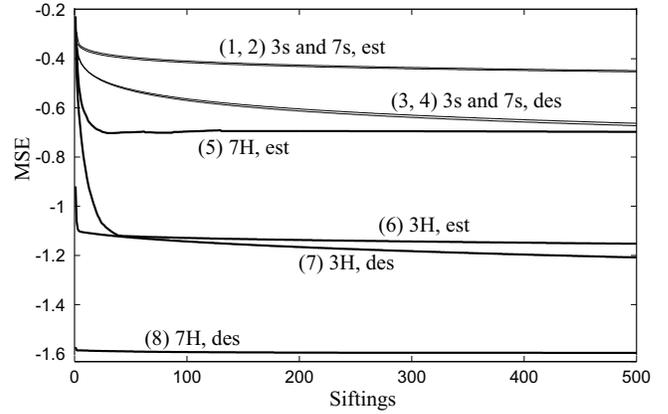


Fig. 5. MSE between the actual and the estimated by EMD first signal component w.r. to the number of sifting iterations.

Hermitian interpolation outperforms the spline interpolation in all the cases examined.

5. CONCLUSION

In this paper, the empirical mode decomposition algorithm was investigated from a novel perspective, that of a genetic algorithm based optimization. The approach facilitated a better understanding of the method and offered significant information about the interpolation points, the envelopes and their optimized settings in several simulation examples. In this manner, a significant performance enhancement of EMD has been achieved. Moreover, the findings regarding the optimized interpolation points and envelopes will be useful in developing a more thorough mathematic analysis of EMD.

6. REFERENCES

- [1] N. E. Huang et. al., "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. A*, vol. 454, pp. 903–995, Mar. 1998.
- [2] G. Rilling and P. Flandrin, "On the influence of sampling on the empirical mode decomposition," in *ICASSP*, 2006.
- [3] H. Liang, Q.-H. Lin, and J. D. Z. Chen, "Application of the empirical mode decomposition to the analysis of esophageal manometric data in gastroesophageal reflux disease," *IEEE Trans. Biomed. Eng.*, vol. 10, pp. 1692–1701, Oct. 2005.
- [4] X. Wang, B. Li, Z. Liu, H. T. Roman, O. L. Russo, K. K. Chin, and K. R. Farmer, "Analysis of partial discharge signal using the Hilbert-Huang transform," *IEEE Trans. Power Delivery*, vol. 21, pp. 1063–1067, July 2006.
- [5] G. Rilling, P. Flandrin, and P. Gonçalvès, "On empirical mode decomposition and its algorithms," in *NSIP workshop*, 2003.
- [6] R. Deering and J. F. Kaiser, "The use of a masking signal to improve empirical mode decomposition," in *ICASSP*, 2005.
- [7] N. E. Huang and S. Shen, *Hilbert-Huang Transform and Its Applications*, World Scientific Publishing Company, first edition, 2005.
- [8] C. Pozrikidis, *Numerical Computation in Science and Engineering*, Oxford University Press, first edition, 1998.