# GRAPH EMPIRICAL MODE DECOMPOSITION

*Nicolas Tremblay, Pierre Borgnat, Patrick Flandrin*

CNRS, Ecole Normale Supérieure de Lyon, Physics Laboratory
Lyon, France

## ABSTRACT

An extension of Empirical Mode Decomposition (EMD) is defined for graph signals. EMD is an algorithm that decomposes a signal in an addition of modes, in a local and data-driven manner. The proposed Graph EMD (GEMD) for graph signals is based on careful considerations on key points of EMD: defining the extrema, interpolation procedure, and the sifting process stopping criterion. Examples of GEMD are shown on the 2D grid and on two examples of sensor networks. Finally the effect of the graph's connectivity on the algorithm's performance is discussed.

***Index Terms*—** Graph signal processing, Empirical Mode Decomposition, Graph interpolation

## 1. INTRODUCTION

Graphs are a useful coding or representation of relations in data for many applications, e.g., neural, sensor, energy, social or biological networks. A *graph signal* is a signal defined on the nodes of a graph, the structure of this graph being either known *a priori* or inferred from proximity or similarity measures between nodes. Fig. 1 shows two examples of such signals, in the context of sensor networks, where nodes are sensors spread out in space. Recently, there has been a substantial effort to adapt classical signal processing tools to graph signals [1, 2] such as the graph wavelet transform [3–6], lifting schemes [7], the windowed Fourier transform [8] or interpolation [9]. We introduce the Graph Empirical Mode Decomposition, an adaptation to graph signals of the now widely used Empirical Mode Decomposition (EMD) [10, 11]. EMD is a data-driven algorithm that aims at locally separating fast from slow oscillations in a signal. As EMD is local and adaptive, it is especially useful when the components of the signal one wants to separate are nonstationary or have overlapping spectra, hence when a simple filtering in the Fourier space fails. For illustrations, we focus on sensor networks, but the method is relevant for any graph. In Sec. 2, the Graph EMD (GEMD) is introduced after recalling the classical EMD (CEMD). In Sec. 3, GEMD is applied to signals on the 2D grid, and on the two examples of Fig. 1. We conclude in Sec. 4.

This work has been partially supported by the French National Research Agency under grant ANR-13-BS03-0002-01.

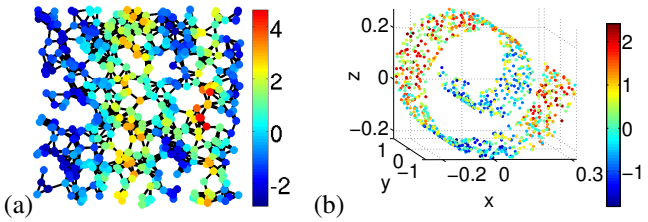


**Fig. 1**. Examples of graph signals in sensor networks, as detailed later on in Sec. 3.2. The values of the graph signals on the nodes is color-coded (as per the colorbars).

## 2. ALGORITHM FOR GRAPH EMD

### 2.1. Classical Empirical Mode Decomposition

Let us recall classical EMD (CEMD). Given a signal $x(t)$, it separates a local "low frequency" component $m_1(t)$ –the trend– from an Intrinsic Mode Function (IMF) $d_1(t)$ which is a local "high frequency" mode having the same number of extrema and zero crossings, and roughly symmetric with respect to zero. By applying the same decomposition to $m_1(t)$ we obtain $m_1(t) = m_2(t) + d_2(t)$, and, recursively:

$$x(t) = m_K(t) + \sum_{k=1}^{K} d_k(t). \qquad (1)$$

The signal is decomposed in IMFs until they are all extracted.

Given $m_i(t)$, this separation of the slow oscillating trend $m_{i+1}(t)$ from the fast oscillating IMF $d_{i+1}(t)$, is done in the EMD algorithm by using the so-called *sifting process* [10]:

1. $s = m_i$. While $s(t)$ does not meet the sifting process stopping criterion, repeat steps 2 to 5:
2. Detect the local extrema of $s(t)$.
3. Interpolate the minima (resp. extrema) to obtain some envelope $e_{min}(t)$ (resp. $e_{max}(t)$).
4. Compute the mean (local trend) $\mu(t) = \frac{e_{min}(t)+e_{max}(t)}{2}$.
5. Subtract it from the signal: $s(t) \leftarrow s(t) - \mu(t)$.
6. Set $d_{i+1}(t) = s(t)$ and $m_{i+1}(t) = m_i(t) - s(t)$.

The most conservative stopping criterion is that the loop stops as soon as $s(t)$ is an IMF. This is usually too strong a constraint and it is relaxed to a stopping criterion yielding approximate IMFs [12].

### 2.2. From CEMD to Graph EMD

Before discussing the elements defining EMD on graphs (extrema, interpolation and stopping criteria), let us study how one creates a graph for data when it is not known beforehand.

#### *2.2.1. Graph creation*

We place ourselves in the context of sensor networks, where the signal has a value at each sensor whose locations in space are known. Let $V$ be the set of the sensors, used as nodes for the graph. Among the options to define edges in the graph supporting the signal, we explore two:

1. a weighted graph parametrized by $\delta$: only pairs of sensors $(i, j)$ at a distance $d_{i,j}$ shorter than $\delta$ are connected by an edge, with weight $w_{i,j} = \exp(-d_{i,j}^2/2\delta^2)$.
2. a binary graph parametrized by $k$: each node is connected to its $k$ nearest neighbors ($k$-NN).

These procedures do not necessarily build connected graphs (typically when $\delta$ or $k$ are too small). To avoid interpolation problems, choose a connected component and add the shortest link connecting it to another component – the component grows larger – and repeat this until the graph is connected.

In other contexts, the graph could be known beforehand, or obtained, e.g., by using statistical similarities as distances. Anyway, we end up with a graph $\mathcal{G} = (V, E)$, where $E$ is the set of edges connecting nodes. Let us note $A$ its adjacency matrix and $D$ the diagonal matrix of degrees.

#### *2.2.2. Definition of local extrema*

For a signal $x$ defined on $V$, a node $i$ is a local maximum (resp. minimum) if, for all its neighbours $k$ in $\mathcal{G}$, $x(i) > x(k)$ (resp. $x(i) < x(k)$). Note that other notions of extremum could be introduced: for instance extremum along one direction only, as it is done for images in "Pseudo-2D" EMD where extrema are along lines or columns only [13]. We explore another definition of extrema in Sec. 3.3.

#### *2.2.3. Interpolation procedure*

There are several ways to interpolate a graph signal. There are for instance global procedures, like the method discussed in [9], where the authors minimize the highest graph Fourier frequency mode necessary to recover the signal on the known nodes. Since this method is based on global Fourier modes, it would not be appropriate to extract modes having some nonstationnarity within the graph, e.g. a chirp. The highest frequency retained would be globally the highest one, and it contradicts the locality of EMD: at local places in the graph where the modes have lower frequency, this interpolation procedure would never extract the mode.

Instead, we rely on an interpolation method formulated through a discrete partial differential equation on the graph. Inspired by Grady et al. [14], interpolation is recast as a Dirichlet problem on the graph. Consider the Laplacian $L = D - A$ of graph $\mathcal{G}$, the signal $s$ on nodes $V$ to be interpolated, and $B$ (resp. $U$) the set of nodes where the signal is known (resp. unknown). Solving the Dirichlet problem comes down to finding $s$ that minimizes $s^\top Ls$ under the constraint $s(b) = s_B(b)$ the known values for $b \in B$. By re-ordering the nodes, one may write $s^\top = [s_B^\top s_U^\top]$ and $L = \begin{bmatrix} L_B & R \\ R^\top & L_U \end{bmatrix}$. Solving the Dirichlet problem boils down to solving the system of linear equations: $L_U s_U = -R^\top s_B$.

#### *2.2.4. Choice of stopping criteria*

With the previous elements, the sifting process is easily modified and we propose a stopping criterion for this sifting process from an energy criterion: stop the loop 2-5 as soon as the energy of the mean $\mu(t)$ (computed at step 4) is lower than the energy of the signal $m_i(t)$ analysed divided by 1000. In all the experiments we have made, this criterion converges. This criteria is reminiscent of choices in CEMD, see [10, 12].

#### *2.2.5. The GEMD algorithm*

Akin to the CEMD, we define the GEMD from its algorithm. Given a set of sensors $V$, a set of measures $\{x_i\}_{i\in V}$, and $K$, the number of modes to be extracted, the algorithm reads:

1. Create the adjacency matrix $A$ for the graph $\mathcal{G}$, here by considering the relative spatial positions (as in 2.2.1).
2. Set $m_0 = x$. Iterating on $i$, extract from $m_i$ the fast mode $d_{i+1}$ and slow trend $m_{i+1}$ following the sifting process of Sec. 2.1 (where $t$ stands now for the indices of nodes) using the extrema, interpolation procedure and stopping criterion described above.
3. Stop and obtain $x(t) = m_K(t) + \sum_{k=1}^{K} d_k(t)$.

Note that we do not discuss here the number of modes to be extracted $K$, it is fixed a priori.

## 3. EXAMPLES AND DISCUSSION

### 3.1. Application and discussion on the 2D grid

Consider the case of $N = 400$ sensors distributed on the 2D grid $20 \times 20$. Instead of discussing it as a regular image, we adopt the point-of-view of graphs. Let us consider, as an example, a superposition of sine waves separated by an angle $\theta$; the signal is the sum of three components:

- a horizontal sine wave of amplitude 2 and frequency 2.
- a sine wave of amplitude 1 and frequency 7, that propagates with an angle $\theta$ with the horizontal.
- a uniform noise of amplitude 0.5.

Fig. 2 shows the results of the GEMD for $\theta = 0, \pi/4$ and $\pi/2$. For the first two cases, the two sines waves are separated as expected (high frequency wave in the first IMF). For orthogonal sine waves ($\theta = \pi/2$), the GEMD does not separate them and the explanation touches the very foundations

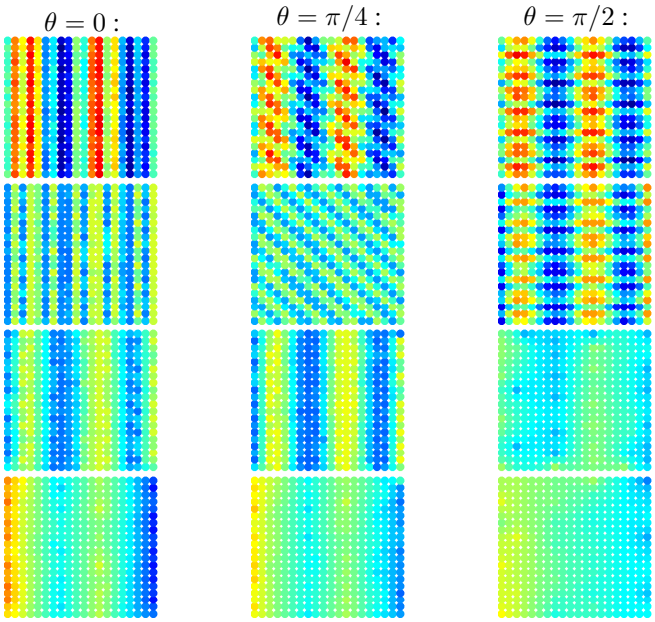


**Fig. 2**. The left (resp. center, right) column represents the result of the GEMD for an angle $\theta = 0$ (resp. $\pi/4$, $\pi/2$) between two sine waves. The first row is the original signal, the second and third rows are the first and second IMF, and the last row is the residue.

of EMD: the definition of extrema in 2.2.2. Fig. 3 displays the first steps of the algorithm. For $\theta = 0$ and $\pi/4$, there are enough extrema to force the envelopes (and thus the mean $\mu$) to follow the low frequency component thereby enabling the separation of components; whereas for $\theta = \pi/2$, there are not enough extrema and they have approximately the same value: the envelopes are flat, $\mu$ has very low energy and the first IMF will contain the whole signal, for no separation. This issue is that this signal is a valid IMF, like it is for EMD in 2D [15], and one should turn to "Pseudo-2D" EMD [13] with another definition of extrema to change that. In fact, the definition of extrema (combined with interpolation and stopping criterion) defines *a posteriori* what is an IMF. Depending on the application, one may change this to force the separation of a component. This issue is discussed in Sec. 3.3.

### 3.2. Two examples of sensor networks and discussion

#### *3.2.1. A sensor network in 2D space*

Consider a network of 512 sensors uniformly distributed on the $1 \times 1$ square. We create a weighted graph from their 2D space positions following 2.2.1 with $\delta = 0.075$. On this graph, we create a signal as the sum of 4 components:

- a sine wave of amplitude 1 and frequency 7, propagating with an angle $\theta = \pi/4$ with the horizontal (Fig. 4a).
- a horizontal linear chirp of amplitude 2 (Fig. 4d).

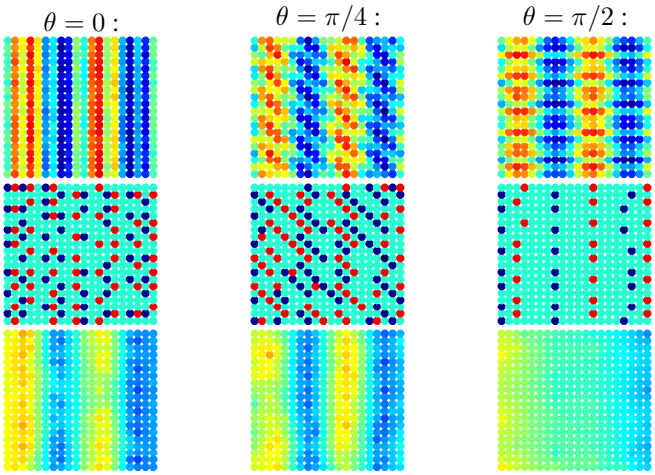


**Fig. 3**. The first steps of the GEMD on the 2D grid for the three angles. The first line is the original signals. The second row shows the extrema (minima in blue and maxima in red). The last row shows the mean $\mu$ of the two envelopes interpolated from these extrema.

- a null signal except for a localized set of nodes of amplitude 2 (Fig. 4g).
- a uniform noise of amplitude 0.5.

The total signal is plotted in Fig. 1a. Results are plotted in the center column of Fig. 4. The first IMF recovers the high frequency sine wave component. The linear chirp is partly in the second IMF and in the residue. The localized signal ends up in the residue. The right column of Fig. 4 shows that a filtering in the graph Fourier space (as defined using the Laplacian [3]) would have failed because of overlap in the Fourier spectra.

#### *3.2.2. A sensor network in 3D space*

Consider a network of $N = 1024$ sensors distributed on a "swiss roll" manifold in 3D space [16] as shown in Fig. 1b. The 3D positions $(X, Y, Z)$ of the sensors on this manifold are computed in 3 steps: i) create $U_1$ and $U_2$, two uniformly random vectors between 0 and 1 of $N$ points; ii) the 3D coordinate vectors are obtained by setting $S_1 = \pi\sqrt{(b^2 - a^2)U_1 + a^2}$, and

$$X = S_1 \cos S_1;\ Y = \pi^2(b^2 - a^2)U_2/2;\ Z = S_1 \sin S_1.$$

Parameter $a\pi$ (resp. $b\pi$) is the starting (resp. ending) angle of the swiss roll. Here, $a = 1$ and $b = 4$. $Y$ is chosen such that the length of the manifold in the $Y$ direction is equal to the total length of the manifold if unrolled: this is to ensure a uniform distribution of sensors; iii) the swiss roll is finally centered and rescaled to fit in the cube of side length 2.

The corresponding $k$-NN binary graph is created following section 2.2.1 with $k = 14$. On this graph, create a signal

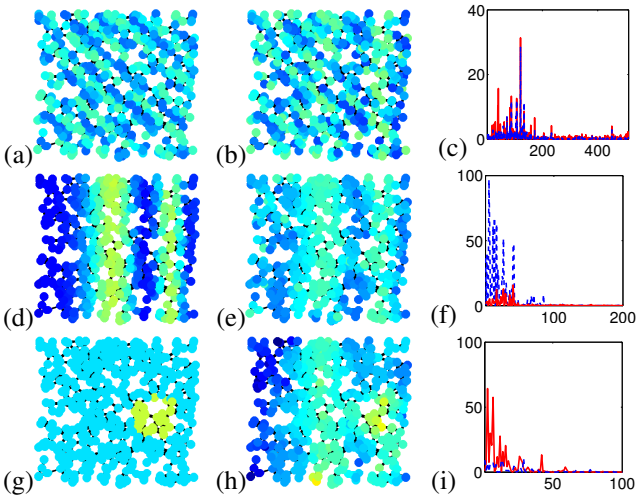


**Fig. 4**. Sensor network of 512 uniformly distributed nodes. Left column: the three components of the original signal. Middle column: the first two IMFs and the residue uncovered by the GEMD. The colormap is the same as in Fig. 1a. Right column: theoretical (resp. uncovered) signals in blue (resp. red) in the graph Fourier domain.

as the sum of 3 components:

- a sine wave in the 3D space: amplitude 1 and frequency 7, that propagates with an angle $\theta = \pi/4$ in the (y,z) plane (Fig.5c),
- a linear chirp along the manifold of amplitude 1 (Fig.5d),
- a uniform noise of amplitude $0.5$.

The total signal is plotted in Fig.5a and Fig.5b for two different view points. The rest of Fig. 5 shows the results of the GEMD: the first (resp. second) IMF in (e) (resp. (f)) recovers the 3D sine wave (c) (resp. the chirp on the manifold (d)). Here again, a simple filtering in the Fourier space would not have separated both signals (see Fig. 5g and h).

Let us now investigate the impact of the construction of the graph from 2.2.1 on the power of recovery of the original components by the GEMD. The recovery rate of the 3D sine wave (resp. the chirp) is measured in terms of its correlation distance with the first (resp. second) IMF. Both methods detailed in section 2.2.1 are investigated and results are plotted in Fig.6. They present a similar behavior: when the connectivity is too low, the method is not sensible to slowly varying signals and recovery fails; when the connectivity is too high, there are too few maxima and the whole signal ends up in the first IMF: recovery fails; there exists an optimal connectivity for which both signals are reasonably uncovered. However, the sensitivity is not too high for $k$; if $k$ is (roughly) between 10 and 20, the recovery appears to be correct.

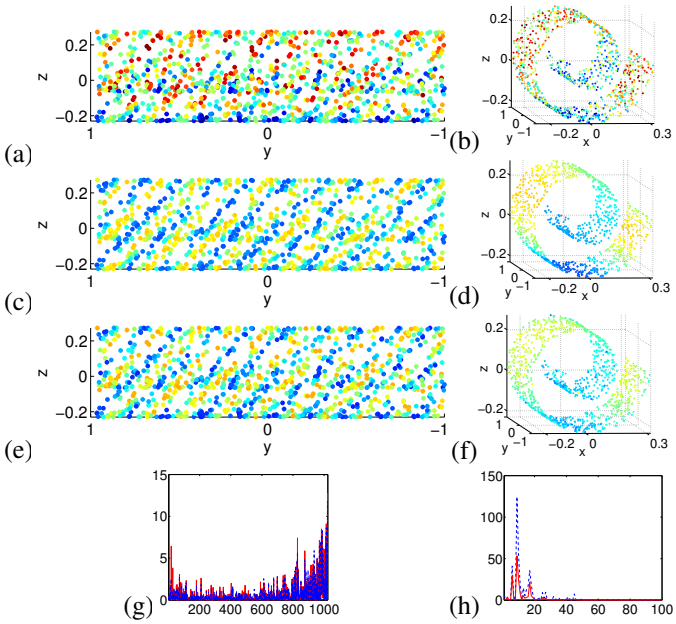


**Fig. 5**. GEMD results on a swiss roll manifold. (a) and (b) are two different views of the original signal composed of a sum of a 3D sine wave (c) and a linear chirp on the manifold (d). (e) and (f) are the first two IMFs. The colormap is the same as in Fig. 1b. (g) and (h) compare the original (dashed blue) and recovered (red) components in the graph Fourier domain.

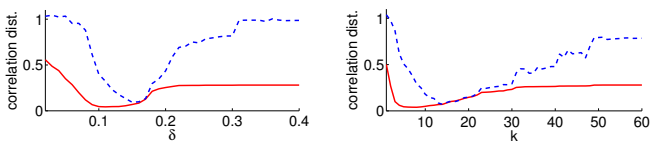


**Fig. 6**. For the swiss roll example, power of recovery of the 3D sine wave (red) and the chirp on the manifold (dashed blue) vs. the connectivity of the graph.

### 3.3. Another definition of local extrema

Suppose a different notion of extremum: a node is a local maximum (resp. minimum) if its value is higher (resp. lower) than a portion of its neighbors, like it would be for maximum along lines or columns on a grid. Here, we only explore this definition for half of the neighbors. In Fig. 7 we compare results obtained with this definition (two right columns) with results previously obtained with Sec. 2.2.2's definition (two left columns) on the 2D grid example with $\theta = \pi/2$. We see how this new definition of extrema increase the number of extrema, thereby enabling the extraction of the fast oscillating mode (first IMF) but pushing the slow oscillating mode into the residue. This observation suggests to look for more elaborate notions of extrema that would take into account the topology of the graph.

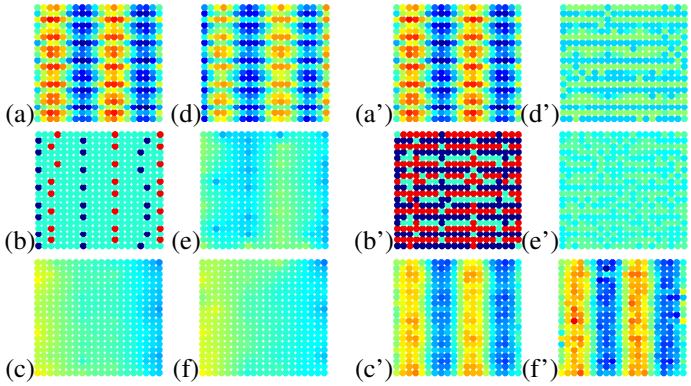


**Fig. 7**. Comparison of results using two different definitions of extrema. The two columns on the left (resp. right) show results obtained with the notion of extrema described in Sec. 2.2.2 (resp. in Sec. 3.3). Fig. a (resp. a') represent the same original signal described in Sec. 3.1 with $\theta = \pi/2$. Fig. b (b') represent the local maxima (in red) and minima (in blue). Fig. c (c') represent the mean of the two envelopes interpolated from these extrema. Fig. d (d') represent the first IMF, Fig. e (e') the second IMF and Fig. f (f') the residue.

## 4. CONCLUSION

A straight-forward adaptation of the classical EMD for graph signals is explored in this communication. The extension of EMD to graph signals opens many degrees of freedom for the key points of EMD: extrema, interpolation, and stopping criterion. In this first communication on the subject, we deliberately chose to use the simplest definitions. We discussed that an additional point is essential to GEMD: the way one chooses to create the graph associated to a given network, more specifically the choice of how connected one chooses to create the graph. In fact, the connectivity has a direct impact on the number of extrema of the signal, therefore a direct impact on the very definition of what is an IMF. Future work will explore this link between connectivity and local extrema.

## 5. REFERENCES

[1] D. Shuman, S. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE SP Mag.*, vol. 30, no. 3, pp. 83–98, 2013.

[2] A. Sandryhaila and J Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 1644–1656, 2013.

[3] D. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *ACHA*, vol. 30, no. 2, pp. 129–150, 2011.

[4] N. Leonardi and D. Van De Ville, "Tight wavelet frames on multislice graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3357–3367, 2013.

[5] R. Coifman and M. Maggioni, "Diffusion wavelets," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 53 – 94, 2006.

[6] S. Narang and A. Ortega, "Perfect reconstruction two-channel wavelet filter banks for graph structured data," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 2786–2799, 2012.

[7] S. Narang and A. Ortega, "Lifting based wavelet transforms on graphs," in *Proc. of APSIPA Annual Summit and Conference (APSIPA ASC)*, 2009.

[8] D. Shuman, B. Ricaud, and P. Vandergheynst, "Vertex-frequency analysis on graphs," *arXiv preprint:1307.5708*, 2013.

[9] S. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *ICASSP*, 2013, pp. 5445–5449.

[10] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[11] D. Mandic, N. Rehman, Z. Wu, and N. Huang, "Empirical mode decomposition-based time-frequency analysis of multivariate signals," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 74–86, 2013.

[12] G. Rilling, F. Flandrin, and P. Gonçalves, "On empirical mode decomposition and its algorithms," in *IEEE-EURASIP Workshop NSIP*, June 2003.

[13] Z. Wua, N.E. Huang, and X. Chan, "The multi-dimensional ensemble empirical mode decomposition method," *Adv. Adapt. Data Anal.*, vol. 1, no. 3, pp. 339–372, 2009.

[14] L. Grady and E. Schwartz, "Anisotropic interpolation on graphs: the combinatorial dirichlet problem," *Technical Report Boston University*, 2003.

[15] J.C Nunes, Y Bouaoune, E Delechelle, O Niang, and Ph Bunel, "Image analysis by bidimensional empirical mode decomposition," *Image and Vision Computing*, vol. 21, no. 12, pp. 1019 – 1026, 2003.

[16] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, pp. 2319–2323, 2000.