

Analyse de signaux AM-FM par approximation parcimonieuse

El Hadji S. DIOP¹, Karl SKRETTING² et Abdel-Ouahab BOUDRAA³

¹Département de Mathématiques, Université de Thiès, Cité Malick Sy, BP 941, Thiès, Sénégal

²Department of Electrical and Computer Engineering, University of Stavanger, 4036 Stavanger, Norway

³IRENav (EA3634), Ecole Navale/Arts-Métiers ParisTech, BCRM Brest, CC 600, 29240 Brest Cedex 9, France
ehsdiop@hotmail.com, karl.skretting@uis.no, boudra@ecole-navale.fr

Résumé – Nous proposons une méthode d’approximation parcimonieuse pour modéliser un signal par des fonctions AM (amplitude modulation)-FM (frequency modulation). La méthode repose sur les modèles AM-FM multicomposantes où chaque composante est obtenue par approximation parcimonieuse résolue par orthogonal matching pursuits, et classée des hautes vers les basses fréquences. L’algorithme proposé s’avère très intéressant pour séparer les contenus fréquentiels de signaux, et se comporte beaucoup mieux que l’EMD (empirical mode decomposition). Comparés à l’EMD, les résultats obtenus montrent une nette amélioration sur la séparation des fréquences, la capacité à bien séparer des fréquences très proches de sinusoides et une robustesse au bruit.

Abstract – The paper presents a sparse coding method that models a signal with amplitude modulation (AM) and frequency modulation (FM) functions. Indeed, the proposed sparse coding approach is based on a multicomponent AM-FM modeling where each monocomponent counterpart is obtained by sparse coding using orthogonal matching pursuits, and sorted from fine to coarse depending on its frequency content. The proposed approach appears to be an efficient tool to properly separate the frequency content of signals, and behaves like the empirical mode decomposition. Results show neat improvements in terms of frequency separation, tone separation capability, and robustness against noise.

1 Introduction

L’EMD [1] a été introduite pour analyser des signaux non stationnaires, et permet de décomposer tout signal comme une somme de composantes appelées IMF (intrinsic mode function) et d’un résidu. Les IMF sont des fonctions AM-FM monocomposantes générées des hautes vers les basses fréquences, et sont extraites dans l’étape du tamisage (SP pour sifting process). Bien qu’étant un outil d’analyse temps-fréquence très intéressant, l’EMD présente beaucoup de limites dues à son origine algorithmique, et même si des améliorations ont été apportées [2, 3, 4] depuis son introduction. L’EMD peut être vue comme une méthode d’approximation par des fonctions AM-FM, où chaque IMF est alors une fonction AM-FM monocomposante. La théorie d’approximation a connu des développements fulgurants durant ces dernières décennies, cela étant dû aux besoins sans cesse croissants tant sur le plan pratique, que sur les aspects théoriques. Parmi les méthodes d’approximations, il y a les solutions parcimonieuses [5] pouvant être obtenues de manière très efficace avec, par exemple, les algorithmes de type matching pursuits [6], least absolute shrinkage and selection operator [8] ou orthogonal MP (OMP) [7]. Des travaux [9, 10, 11] existent sur les fonctions AM-FM/EMD et les méthodes parcimonieuses. Contrairement à ce qui est fait jusque là, nous apportons dans ce travail des contributions à la fois sur les modèles AM-FM et sur l’EMD, en utilisant une approche parcimonieuse. En établissant un lien entre les fonctions

AM-FM et la représentation de dictionnaire, notre objectif est de présenter les modèles AM-FM sous l’angle d’un problème d’approximation parcimonieuse, et de résoudre quelques problèmes majeurs de l’EMD. On considère alors un signal comme la somme de fonctions AM-FM monocomposantes où chacune d’elles est obtenue comme une approximation parcimonieuse.

2 EMD & Opérateur de Teager

Pour tout signal $x(t)$, on peut résumer l’EMD [1], suivant :

1. Trouver les extréma de $x(t)$
2. Interpoler les maxima et les minima de $x(t)$ notés respectivement $E_{max}(t)$ et $E_{min}(t)$
3. Calculer la moyenne locale $m(t)$ de $x(t)$:
$$m(t) = \frac{1}{2}(E_{max}(t) + E_{min}(t)).$$
4. Extraire le détail $d(t) = x(t) - m(t)$

On dit qu’une fonction f est une IMF [1] ssi sa moyenne locale est nulle, et la différence entre le nombre d’extréma et de passage par zéro de f est au plus de un. Le SP constitué par les étapes (1)-(4) permet d’itérer sur détail $d(t)$ jusqu’à ce que sa moyenne locale soit nulle, et obtenir ainsi une IMF. Finalement, l’EMD décompose le signal $x(t)$ sous la forme :

$$x(t) = \sum_{i=1}^M d_i(t) + r(t), \quad (1)$$

où d_i est la $i^{\text{ème}}$ IMF et $r(t)$ est le résidu de la décomposition.

L'opérateur de Teager (OT) a été introduit [12] pour modéliser les signaux de parole. Pour tout signal $x(t)$, on a :

$$\Psi[x(t)] = \left(\frac{dx}{dt}(t) \right)^2 - x(t) \frac{d^2x}{dt^2}(t). \quad (2)$$

Soit $x(t) = a(t) \cos[\phi(t)]$ un signal AM-FM à bande étroite, on suppose que $a(t)$ est une fonction qui varie lentement,

$$\phi(t) = \omega_c t + \omega_m \int_0^t q(u) du + \theta, \quad (3)$$

ω_c est la porteuse de fréquence supposée varier lentement ou constante par morceaux, $|q(u)| \leq 1$, $\omega_m \in [0; \omega_c]$ est le maximum de la déviation de fréquence, et θ est un déphasage constant. Alors, la fréquence instantanée $\omega(t)$ est donnée par :

$$\omega(t) := \phi'(t) = \frac{d}{dt} \phi(t) = \omega_c + \omega_m q(t). \quad (4)$$

L'OT est introduit après dans le domaine du traitement du signal pour approcher avec des erreurs minimales l'enveloppe de l'amplitude ainsi que la fréquence instantanée, dès lors que les hypothèses faites sur $a(t)$ et $\phi(t)$ restent valables, suivant [13] :

$$\hat{a}(t) = \frac{\Psi[x(t)]}{\sqrt{\Psi[x'(t)]}} \text{ and } \hat{\omega}(t) = \sqrt{\frac{\Psi[x'(t)]}{\Psi[x(t)]}}. \quad (5)$$

L'OT pour un signal discret $x[n]$ est défini par :

$\Psi(x[n]) = x[n]^2 - x[n-1]x[n+1]$. En posant :

$G[n] = 1 - \frac{\Psi(x[n] - x[n-1]) + \Psi(x[n+1] - x[n])}{4\Psi(x[n])}$, l'amplitude et la fréquence instantanées sont estimées par :

$$\hat{a}[n] = \sqrt{\frac{\Psi(x[n])}{1 - G[n]^2}} \text{ and } \hat{\omega}[n] = \arccos G[n]. \quad (6)$$

3 Approximation parcimonieuse

Une représentation AM-FM multicomposantes s'écrit :

$$x(t) = \sum_{k=0}^{M-1} a_k(t) \cos[\phi_k(t)] = \sum_{k=0}^{M-1} x_k(t), \quad (7)$$

où $\forall k = 0, \dots, M-1$, $(a_k, \phi_k)_k$ est le domaine de modulation, les a_k, ϕ_k sont des fonctions suffisamment régulières et dénotent respectivement les amplitudes et les phases instantanées; ϕ_k' est la fréquence instantanée. Pour s'assurer de la contrainte de bande étroite, les hypothèses pour (3) et (4) s'appliquent aussi à toutes les fonctions a_k et ϕ_k . Pour l'EMD, les $x_k(t)$ sont les IMF, classées des hautes vers les basses fréquences; chaque IMF étant une fonction AM-FM à bande étroite.

Soit $x[n]$ un signal contenant N échantillons, $D \in \mathbb{C}^{N \times N}$ la matrice de Transformée de Fourier Discrète (TFD) : $D[n, k] = \exp\left(\frac{2i\pi kn}{N}\right)$. D est base orthogonale de \mathbb{C}^N . La TFD de $x[n]$, notée $\bar{x}[k]$ et la TFD inverse sont respectivement données par :

$$\bar{x} = Dx, \text{ and } x = \frac{1}{N} D^* \bar{x}, \quad (8)$$

où D^* désigne la transposée du conjugué complexe de D . En apprentissage de dictionnaires, (8) peut être interprétée comme la représentation du signal x par \bar{x} suivant le dictionnaire D . On aura donc d'autres types de représentations pour différentes formes de D . En domaine continu, les atomes (colonnes) de D sont des fonctions dépendant continûment de t . Le dictionnaire de cosinus : $D = \{\sqrt{2} \cos[(k + \frac{1}{2}) \pi t]\}_{k \in \mathbb{N}}$ est une base orthonormale de $L^2([0; 1])$ et génère la grande famille de représentations AM-FM multicomposantes (7).

3.1 Formulation du problème & Algorithme

Pour $1 \leq p < \infty$, soit $\|\cdot\|$ la norme l^p de $x \in \mathbb{R}^N$: $\|x\|_p = \left(\sum_{n=1}^N |x[n]|^p \right)^{\frac{1}{p}}$. Pour $p = 0$, notons $\|x\|_0$ le nombre d'éléments non nuls [14] de x ; i.e., $\|x\|_0 = \sum_{n=1}^N (x[n] \neq 0)$.

Soit $D = \{d_i\}_i \in \mathbb{R}^{N \times K}$ un dictionnaire, chaque colonne de D s'appelle un atome. Nous souhaitons représenter tout $x = (x[n])_{n=1}^N$ comme une combinaison linéaire des atomes de D , $x \approx \sum_{k=1}^K d_k \alpha_k = D\alpha$ tels que beaucoup d'éléments de $\alpha = (\alpha_k)_k \in \mathbb{R}^K$ soient nuls. De plus, cette approximation devrait être optimale dans le sens où l'erreur de reconstruction $x - D\alpha$ est minimale pour la norme l^2 . Nous formulons donc le problème comme suit :

$$\alpha^* = \underset{\alpha \in \mathbb{R}^K}{\operatorname{argmin}} \frac{1}{2} \|x - D\alpha\|_2^2 \text{ s.t. } \|\alpha\|_p \leq s_p, \quad (9)$$

où s_p représente le degré de parcimonie de α donnée en norme l^p . Avec un multiplicateur de Lagrange, l'équation (9) devient :

$$\alpha^* = \underset{\alpha \in \mathbb{R}^K}{\operatorname{argmin}} \frac{1}{2} \|x - D\alpha\|_2^2 + \lambda \|\alpha\|_p, \quad (10)$$

avec $\lambda > 0$ étant un paramètre dépendant de s_p . Pour $p = 0$, (10) devient :

$$\alpha^* = \underset{\alpha \in \mathbb{R}^K}{\operatorname{argmin}} \frac{1}{2} \|x - D\alpha\|_2^2 + \lambda \|\alpha\|_0. \quad (11)$$

L'équation (11) est non convexe et sa résolution est bien connue comme étant un problème NP-difficile. Différentes méthodes ont été proposées pour la résoudre de manière efficace; il y a par exemple les algorithmes MP ou OMP. En plus de solutions plus parcimonieuses, OMP présente des améliorations notoires par rapport à MP, car à chaque itération, le résidu est projeté orthogonalement sur l'espace engendré par les atomes sélectionnés. En effet, une mise à jour des coefficients de tous les atomes sélectionnés est effectuée dans OMP par projection orthogonale. Ainsi, un atome ne peut pas être sélectionné plus d'une fois, comme dans MP et réduisant ses performances, garantissant également la convergence de OMP [7].

L'EMD nécessite un critère d'arrêt pour sortir du SP, qui est basé sur un calcul à posteriori d'un indice pour vérifier l'orthogonalité entre les IMF et de la décomposition. L'approche proposée repose sur une représentation AM-FM multicomposantes où chaque composante s'obtient par approximation parcimonieuse. Chaque composante étant une fonction AM-FM

à bande étroite, on peut donc calculer les estimations des amplitudes et des fréquences grâce à l'OT. Les composantes sont ensuite classées des hautes vers les basses fréquences comme l'EMD. Les principales étapes sont résumées dans l'algorithme 1.

Algorithm 1: Algorithme de décomposition proposé.

Entrées: x , D et λ .

Sorties : Modes de décomposition $y = (y_i)_{i=1}^M$.

1 Approximation parcimonieuse : calculer avec OMP

$$\alpha^* := \operatorname{argmin}_{\alpha \in \mathbb{R}^K} \frac{1}{2} \|x - D\alpha\|_2^2 + \lambda \|\alpha\|_0.$$

2 Prendre $M = \#\{\alpha_k \neq 0\}$.

3 **foreach** $m = 1$ à M **do**

4 $y_m \leftarrow d_m \alpha_m$

5 Calculer les fréquences $\hat{\omega}_m$ de y_m :

6 **foreach** $k = 1$ à K **do**

7 Calculer $\hat{\omega}_m[k]$ avec (6)

8 **end**

9 **end**

10 Classifier les colonnes de $\hat{\omega} = (\hat{\omega}_i)_{i=1}^M$ par ordre décroissant.

11 Poser $y = (y_i)_{i=1}^M$.

4 Résultats

Désignons par *Mode* les modes obtenus avec l'approche proposée, HF et BF la haute et la basse fréquence du signal, respectivement. On compare HF (resp. BF) exacte à Mode₁ (resp. Mode₂) et à IMF₁ (resp. IMF₂). Le dictionnaire D est constitué d'une famille de cosinus ou bien de sinus et de cosinus. Nous avons testé les deux méthodes d'abord sur le signal défini sur $[0; 3]$ et contenant 769 échantillons : $x_1(t) = 2 \sin(20\pi t) + 3 \sin(2\pi t)$. Les résultats sont présentés sur la Fig. 1. La HF est bien séparée par l'EMD et l'approche proposée, ce qui n'est pas le cas pour le second mode IMF₂, illustrant bien les limites de l'EMD principalement dues aux méfaits de l'interpolation. Considérons le signal défini sur $[0; 10]$ dont l'amplitude varie linéairement et contenant 2561 échantillons : $x_2(t) = \cos(2\pi t) + (2t + 1) \cos(0.6\pi t)$. Les résultats présentés sur la Fig. 2 sont aussi intéressants, mettant aussi bien en évidence les limites de l'EMD (voir IMF₁ et IMF₂). On note aussi une petite atténuation de l'amplitude de Mode₂, ce qui ne nuit en rien à son contenu fréquentiel. Nous avons étudié et comparé avec l'EMD les performances de l'approche par rapport à sa capacité de bien séparer des sinusoides pures dont leurs fréquences sont proches, en considérant le signal défini sur $[0; 6]$: $y(t) = y_1(t) + y_2(t)$, où $y_1(t) = \cos(2\pi t)$ et $y_2(t) = a_k \cos(2\pi f_k t)$. Afin de mieux quantifier les performances, nous utilisons une métrique basée sur la norme L^2 : $M = \sum_{k=1}^2 \|d_k - y_k\|_2 / \|y_k\|_2$, où d_k désigne le k^{ieme} mode de décomposition extrait soit avec l'EMD, soit l'approche proposée. Les résultats sont présentés sur la Fig. 3 pour des valeurs de $a_k \in [10^{-2}; 25]$ et $f_k \in]0; 1]$,

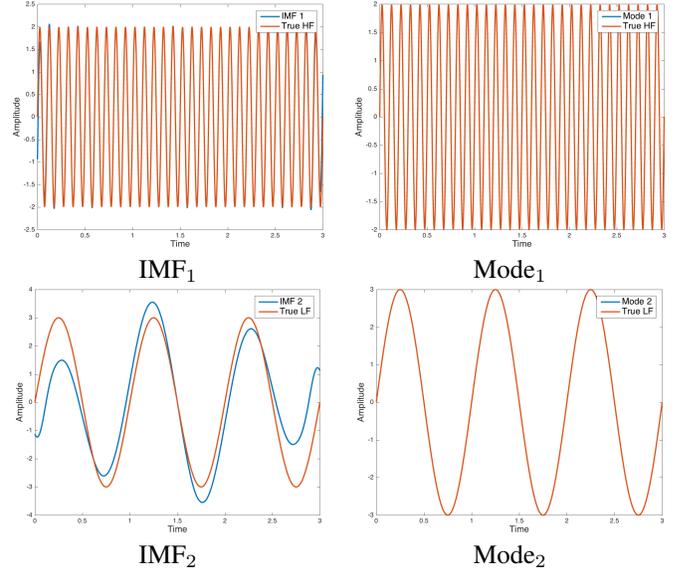


FIGURE 1 – Signal $x_1(t)$: approche proposée et EMD.

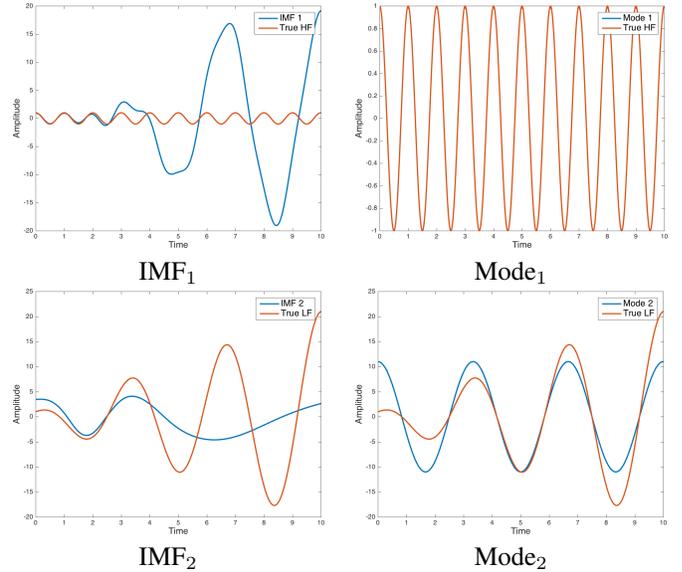


FIGURE 2 – Signal $x_2(t)$: approche proposée et EMD.

montrant bien les améliorations considérables de l'approche proposée ; en effet, la couleur bleue représente les parties où les fréquences sont bien séparées, *i.e.*, $M = 0$. Le dernier signal

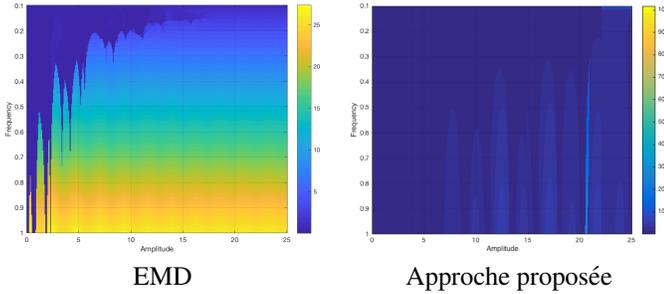


FIGURE 3 – Séparation de fréquences proches.

est non linéaire, défini sur $[0; 6]$ et contenant 769 échantillons : $x_3(t) = 4t \cos\left(\frac{2\pi}{5}(t^2 - 1)\right) + 3 \sin(2\pi(t + 3))$. Les résultats sont présentés sur la Fig. 4 montrant que IMF_2 ne sépare pas bien la LF, contrairement à $Mode_2$. Aussi, comparé à IMF_1 , on note une petite atténuation de l'amplitude de $Mode_1$.

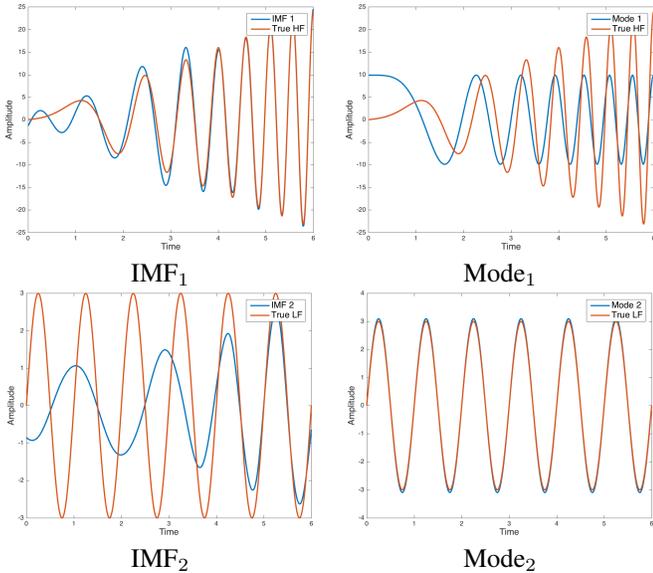


FIGURE 4 – Signal $x_3(t)$: approche proposée et EMD.

5 Conclusion

L'approche AM-FM multicomposantes proposée s'avère être une méthode de séparation de fréquence efficace. L'algorithme proposé montre des performances intéressantes pour bien séparer les contenus fréquentiels de signaux, des fréquences très proches et est robuste au bruit. L'EMD repose sur une interpolation des enveloppes, son absence dans la méthode proposée peut bien expliquer les améliorations. Nous avons cependant noté une atténuation des amplitudes pour certains signaux non stationnaires, même si cela n'affecte pas les fréquences. Les pers-

pectives vont en ce sens en utilisant une méthode d'apprentissage de dictionnaires qui tient en compte les caractéristiques fréquentielles dans le processus d'apprentissage.

Références

- [1] 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," *RS*, vol. 454, pp. 903-995, 1998.
- [2] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms : An empirical mode decomposition-like tool," *ACHA*, vol. 30, no. 2, pp. 243-261, 2011.
- [3] E. H. S. Diop, R. Alexandre, and V. Perrier, "A PDE based and interpolation-free framework for modeling the sifting process in a continuous domain," *Adv. Comput. Math.*, vol. 38, no. 4, pp. 801-835, 2011.
- [4] T. Y. Hou and Z. Shi, "Extracting a shape function for a signal with intra-wave frequency modulation," *RS. A*, vol. 374, no. 2065, pp. , 2016.
- [5] M. Elad, *Sparse and redundant representations*. Springer, New York, 2010.
- [6] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE TSP*, vol. 41, no. 12, pp. 3397-3415, 1993.
- [7] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE TIT*, vol. 53, no. 12, pp. 4655-4666, 2007.
- [8] R. Tibshirani, "Regression shrinkage and selection via the lasso," *JRSS*, vol. 58, no. 1, pp. 267-288, 1996.
- [9] Z. He, Q. Wang, Y. Shen, and M. Sun, "Kernel sparse multitask learning for hyperspectral image classification with empirical mode decomposition and morphological wavelet-based features", *IEEE TGRS*, vol. 52, no. 8, pp. 5150-5163, 2014.
- [10] R. Chen, H. Jia, X. Xie, and W. Gao, "Learning a collaborative multiscale dictionary based on robust empirical mode decomposition," *ArXiv :1704.04422*, 2017.
- [11] X. Meng, A. Jakobsson, X. Li, and Y. Lei, "Estimation of chirp signals with time-varying amplitudes," *Signal Processing*, vol. 147, pp. 1-10, 2018.
- [12] H. M. Teager, "Some observations on oral air flow during phonation", *IEEE TASSP*, vol. ASSP-28, no. 5, pp. 599-601, 1980.
- [13] P. Maragos, J. Kaiser, and T. Quatieri, "On separating amplitude from frequency modulations using energy operators", in *IEEE ICASSP*, 1992.
- [14] S. Chen, C. F. N. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks", *IEEE Transactions on Neural Networks*, vol. 2, no. 2, pp. 302-309, 1991.