FREQUENCY SEPARATION METHOD BASED ON SPARSE CODING

El Hadji S. Diop

Department of Mathematics,
University of Thies, Cite Malick Sy,
BP 941, Thies, Senegal

Karl Skretting

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

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 frequency separation (SCFS) approach is based on a mutli-component 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. SCFS 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.

***Index Terms*—Sparse coding, AM-FM, Frequency separation, Orthogonal matching pursuit.**

I. INTRODUCTION

The empirical mode decomposition (EMD) was introduced for analyzing nonlinear and non-stationary time series [1]. EMD decomposes any signal as a sum of several basic components called intrinsic mode functions (IMFs) and a residual which is the signal trend; IMFs are AM-FM monocomponent signals and generated in a coarse-to-fine basis. The step in which IMFs are extracted is a iterative process called sifting process (SP). Also, it is an interesting time-frequency analysis tool if combined with Hilbert or Teager-Huang transforms [2]. EMD is a data-driven and adaptive method, contrary to wavelet analysis or Cohen class time-frequency representations, for which either a predefined basis function or a kernel is used [3]. EMD has many drawbacks and serious limits for properly working, mainly due to its algorithmic origin and SP step. Efforts have been done as alternatives to overcome such limits and/or provide theoretical framework [4]–[8].

EMD can be seen as an AM-FM approximation method, where each IMFs are narrow band AM-FM functions; AM-FM functions have been widely studied in signal and image processing [9]–[11]. For the last several decades, developing approximation methods have thus been an interesting and stimulating subject of interests, mainly due to increasing demands that can be ranged from practical applications to theoretical frameworks. In mathematical approximations theory, an accurate signal approximation is made as a linear combination of a small number of vectors selected from a carefully constructed basis. Sparse solutions recently appeared [12] as an interesting approximation method that can be efficiently solved with various techniques; for instance, using matching pursuits [13], orthogonal MP (OMP) [14], least absolute shrinkage and selection operator [15], compressed sensing [16], or more recently using Bregman iterations [17].

There exists some works related to AM-FM functions/EMD and sparse coding. In fact, a sparse reconstruction framework was proposed [18] to address the problem of parameters estimation of signals composed of an unknown number of chirps with time-varying amplitude. Using a dictionary refinement strategy, their approach accurately estimated AM and FM parameters of multiple signal components. EMD was used to decompose training signals that belong to different classes into IMFs [19]. Afterwards, a raw dictionary whose atoms are IMFs is trained with the same training signals using matching pursuits, leading to a sparse learned dictionary. A multiscale dictionary learning framework was proposed [20] based on an improved EMD. Basically, this was achieved by looking for the sparsest multiscale representations of images within the largest possible dictionary which is composed of IMFs. EMD was combined with a morphological wavelet transform (MWT) for gaining spectral–spatial image features [21] that can be integrated by the sparse multitask learning.

In this paper, contrary to previous works, contributions are provided on both AM-FM modeling and EMD using a sparse approach. The motivations behind the established link between AM-FM and dictionary representations are to view AM-FM modeling as a sparse approximation problem, and solve major issues in EMD. The proposed SCFS method models a signal as a sum of monocomponent AM-FM functions, each term being approximated by sparse coding. SCFS shows great capabilities for tracking the frequency content of sines and chirps, increases the tone separation capability, and is robust in a noisy environment. Moreover, SCFS eliminates the interpolation process necessary in EMD, which causes many side effects. This work provides also contributions on both the comprehension and well established foundation of

EMD.

II. REVIEW ON EMD & TEAGER OPERATOR

For any signal $x(t)$, EMD can be summarized as [1]:

1) Find $x(t)$ extrema
2) Do linear/spline interpolation of $x(t)$ maxima and $x(t)$ minima, respectively denoted $E_{max}(t)$ and $E_{min}(t)$
3) Compute the local mean $m(t)$ of $x(t)$: $m(t) = \frac{1}{2}(E_{max}(t) + E_{min}(t))$.
4) Extract detail $d(t) = x(t) - m(t)$
5) Iterate on the residual $m(t)$

A function f is an IMF if [1]:

(i) the local mean of f is equal to 0, and
(ii) the number of extrema and zero crossings of f must either be equal or differ at most by one.

The sifting process (SP) is the refinement iterative steps (1)-(4) in which the detail $d(t)$ is iterated until its local mean is equal to 0 in order to get an IMF. The signal $x(t)$ is first decomposed in the main loop: $x(t) = d_1(t) + m_1(t)$. Then, the first residual $m_1(t)$ is decomposed: $m_1(t) = d_2(t) + m_2(t)$, and so on. Finally, EMD decomposes $x(t)$ as:

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

where d_i denotes the so-called i^{th} IMF and $r(x)$ is the residual of the decomposition.

Teager operator (TO) was first introduced [22] for modeling speech signals. For any $x(t)$, TO is defined as:

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

Let $x(t) = a(t)\cos[\phi(t)]$ be a narrow band AM-FM signal, $a(t)$ is assumed to be a slowly varying function,

$$\phi(t) = \omega_c t + \omega_m \int_0^t q(u)\mathrm{d}u + \theta, \tag{3}$$

ω_c denoting the carrier frequency assumed to be slowly varying or piecewise constant, $q(u)$ is a bounded function; *i.e.*, $|q(u)| \leq 1$, $\omega_m \in [0;\ \omega_c]$ is the maximum frequency deviation and θ represents a constant phase offset. Thus, the instantaneous frequency is given as:

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

TO was later introduced in signal processing, and its capability to find good approximations of the amplitude envelope and instanteneous frequency was demonstrated with minimum approximation errors, subject to the previous assumptions on $a(t)$ and $\phi(t)$, as [23]:

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

The discrete TO of a discrete signal $x[n]$ is defined as:

$$\Psi(x[n]) = x[n]^2 - x[n-1]x[n+1]. \tag{6}$$

Let $G[n] = 1 - \dfrac{\Psi(x[n]-x[n-1]) + \Psi(x[n+1]-x[n])}{4\Psi(x[n])}$. The instantaneous amplitude and the instantaneous frequency of $x[n]$ can be estimated as:

$$\hat{a}[n] = \sqrt{\frac{\Psi(x[n])}{1-G[n]^2}} \text{ and } \hat{\omega}[n] = \arccos G[n]. \tag{7}$$

III. PROPOSED SPARSE CODING APPROACH

A multicomponent AM-FM representation of a signal $x(t)$ is generally formulated as:

$$x(t) = \sum_{k=0}^{M-1} a_k(t)\cos[\phi_k(t)] = \sum_{k=0}^{M-1} x_k(t), \tag{8}$$

where $\forall\ k = 0, \cdots, M-1$, $(a_k, \nabla\phi_k)_k$ is the modulation domain, a_k, ϕ_k are smooth functions and respectively denoting the instantaneous amplitude and phase components; ϕ'_k is the instantaneous frequency. To ensure a narrow band constraint, it is common to set the same restrictions for every a_k and ϕ_k as for $a(t)$ and $\phi(t)$ in equations (3) and (4). In EMD, $x_k(t)$ represent IMFs which are sorted depending on their frequency content from fine to coarse, and each IMF is truly a monocomponent narrow band AM-FM signal.

Let $x[n]$ be a signal of N samples, $n \in \{0, 1, \cdots, N-1\}$ and let $\mathrm{D} \in \mathbb{C}^{N\times N}$ be the discrete Fourier transform (DFT) matrix: $D[n,k] = \exp\left(\frac{2i\pi kn}{N}\right)$. D is an orthogonal basis of $\mathbb{C}^N$. The DFT of $x[n]$ denoted $\bar{x}[k]$, and the inverse DFT are respectively given as:

$$\bar{x} = Dx, \text{ and} \tag{9}$$

$$x = \frac{1}{N}D^\star\bar{x}, \tag{10}$$

where $D^\star$ is the complex conjugate transpose of D. In a dictionary learning view, (10) can be interpreted as a representation of the signal x by $\bar{x}$ under the dictionary D. Different forms of D will yield different representations. In continuous time, the atoms (columns) of D will be continuous functions (of t). One example is the cosine dictionary: $D = \left\{\sqrt{2}\cos\left[\left(k+\frac{1}{2}\right)\pi t\right]\right\}_{k\in\mathbb{N}}$ which is an orthonormal basis of $L^2([0;\ 1])$, then it yields the broad family of multicomponent AM-FM representations (8). The motivations behind the established link between multicomponent AM-FM models and dictionary representations are to view AM-FM modeling as a sparse approximation problem, which turns out to resolve main issues of classical EMD.

III-A. Problem statement & Proposed algorithm

For $1 \leq p < \infty$, let $\|\cdot\|$ be the l^p norm of any vector $x \in \mathbb{R}^N$ given as $\|x\|_p = \left(\sum_{n=1}^{N}|x[n]|^p\right)^{\frac{1}{p}}$. Let $D = \{d_i\}_i \in \mathbb{R}^{N\times K}$ be a given dictionary, each column of D is commonly called (dictionary) atom. We aim at representing any given signal $x = (x[n])_{n=1}^{N}$ as a linear combination of the dictionary atoms $x \approx \sum_{k=1}^{K} d_k \alpha_k = D\alpha$, *s.t.* most elements of $\alpha = (\alpha_k)_k \in \mathbb{R}^K$ are zero. In addition, such an approximation should be optimal in the sense that the reconstruction error $x - D\alpha$ is minimal in l^2 norm. Thus, we formulate the problem in the following:

$$\alpha^\star = \underset{\alpha \in \mathbb{R}^K}{\operatorname{argmin}} \frac{1}{2} \|x - D\alpha\|_2^2 \ \textit{s.t.}\ \|\alpha\|_p \le s_p, \qquad (11)$$

where s_p is the sparseness measure on α given in terms of l^p norm. Using a Lagrange multiplier, equation (11) is equivalent to:

$$\alpha^\star = \underset{\alpha \in \mathbb{R}^K}{\operatorname{argmin}} \frac{1}{2} \|x - D\alpha\|_2^2 + \lambda \|\alpha\|_p, \qquad (12)$$

where $\lambda > 0$ is a tuning parameter.

Using common notation, let $\|\alpha\|_0$ denote the number of nonzero elements [24] in α; *i.e.*, $\|\alpha\|_0 = \sum_{k=1}^{K} (\alpha_k \neq 0)$. For $p = 0$, equation (12) yields:

$$\alpha^\star = \underset{\alpha \in \mathbb{R}^K}{\operatorname{argmin}} \frac{1}{2} \|x - D\alpha\|_2^2 + \lambda \|\alpha\|_0. \qquad (13)$$

Equation (13) is nonconvex and well known as a NP-hard problem for finding a solution. Different strategies have been developed to efficiently solve it; for instance, greedy algorithms like MP or OMP. In addition to more sparse solutions, OMP demonstrates improvements on MP as for each iteration the residual is orthogonalized onto the space spanned by the selected atoms. Indeed, an update of the coefficients of all selected dictionary atoms is performed by means of an orthogonal projection of the signal onto such selected atoms. This procedure alleviates the issue of selecting an atom more than once as in MP, which reduces MP's performance. In addition, the way coefficients are selected guarantees OMP convergence [14].

On the other hand, a stopping criterion is required for EMD to quit the SP loop or for rejecting IMFs that are grossly non-orthogonal [25]. The criterion is based on the computation of an a posteriori index of orthogonality in order to check the orthogonality between IMFs and of the overall decomposition. Note that the orthogonality property is not truly reached in EMD [1], since as a Reynolds type decomposition, orthogonality is always approximately met up to the degree of nonlinearity.

The main idea of the proposed SCFS method is based on a multicomponent AM-FM representation of a signal $x(t)$ (8), each monocomponent counterpart $x_k(t)$ is then obtained using sparse coding (13). To mimic EMD, each $x_k(t)$ is sorted according to its frequency content, from fine to coarse. This is done thanks to the approximations of the amplitude and frequency components given by TO (2); each $x_k(t)$ being narrow banded AM-FM functions. SCFS algorithm is presented in Algorithm 1.

Algorithm 1: SCFS algorithm.

Input : x, D and λ.
Output: Decomposition modes $y = (y_i)_{i=1}^{M}$.

Sparse coding: compute using OMP

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

Set $M = \sharp(\alpha_k \neq 0)$.
foreach $m = 1$ *to* M **do**
  $y_m \leftarrow d_m \alpha_m$
  Compute the frequency content $\hat{\omega}_m$ of y_m:
  foreach $k = 1$ *to* K **do**
    Compute $\hat{\omega}_n[k]$ using (7)
  end
end
Sort columns of $\hat{\omega} = (\hat{\omega}_i)_{i=1}^{M}$ in descending order.
Set $y = (y_i)_{i=1}^{M}$ accordingly.

IV. EXPERIMENTS

Decomposition modes obtained with SCFS are referred to as *Mode*. We denote by HF and LF the high and low frequency component, respectively. In all experiments, true HF (*resp.* LF) is compared to Mode_1 (*resp.* Mode_2) and IMF_1 (*resp.* IMF_2); IMFs are obtained from EMD. In our experiments, the dictionary D is either a family of cosines or a family of sines and cosines.

The first tested stationary signal is defined on $[0;\ 3]$, has 769 samples, and is given by: $x_1(t) = 2\sin(20\pi t) + 3\sin(2\pi t)$. Results are shown in Fig. 1. The first mode is well separated by SCFS and EMD as well, but IMF_2 suffers from interpolation issues previously discussed regarding undershoots, overshoots, ...

The second tested signal is defined on $[0;\ 10]$, has a linearly varying amplitude and 2561 samples; it is given as: $x_2(t) = \cos(2\pi t) + (2t + 1)\cos(0.6\pi t)$. Its decomposition is displayed in Fig. 2 showing clearly the limits of EMD (see IMF_1 and IMF_2). Notice some amplitude attenuation in Mode_2, while tracking well the frequency content.

The tone separation capability of EMD and the proposed approach is examined here by considering the following signal $y : [0;\ 6] \mapsto \mathbb{R};\ t \longrightarrow y(t) = y_1(t) + y_2(t)$, where $y_1(t) = \cos(2\pi t)$ and $y_2(t) = a_k \cos(2\pi f_k t)$. To better quantify it, the following L^2-based measure is proposed: $M = \sum_{k=1}^{2} \frac{\|d_k - x_k\|_2}{\|x_k\|_2}$, d_k being the k^{th} decomposition mode obtained either with EMD or SCFS. Results are displayed in Fig. 3 for $a_k \in [10^{-2};\ 25]$ and $f_k \in]0;\ 1]$. The blue area represents regions where frequency contents are well separated; *i.e.*, $M = 0$. The proposed SCFS shows noticeable improvements compared to EMD.

The proposed SCFS is finally tested on the following nonlinear signal defined on $[0;\ 6]$ with 769 samples:

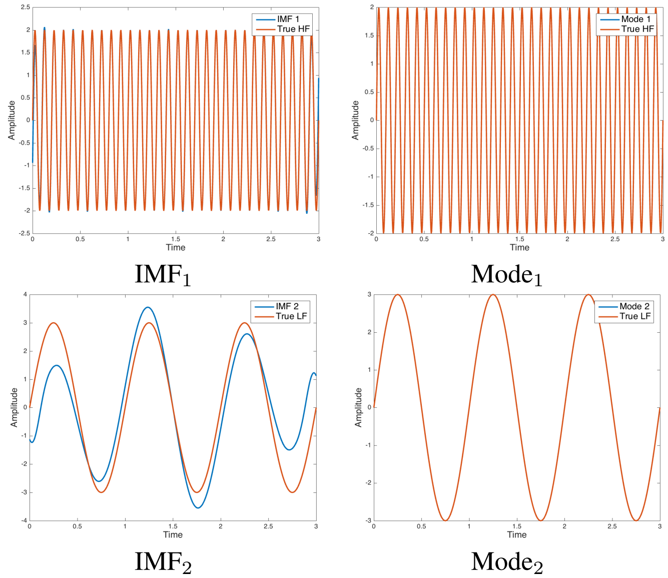

Fig. 1. Decomposition of the signal $x_1(t)$, compared with the true components: SCFS vs EMD.

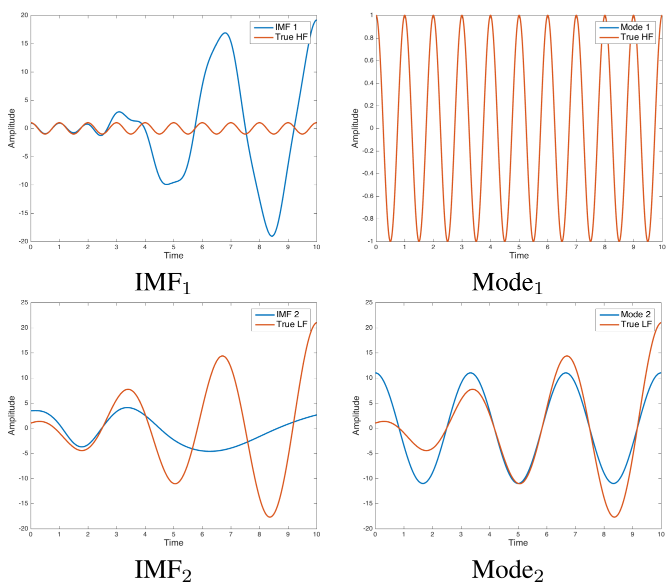

Fig. 2. Decomposition of the signal $x_2(t)$, compared with the true components: SCFS vs EMD.

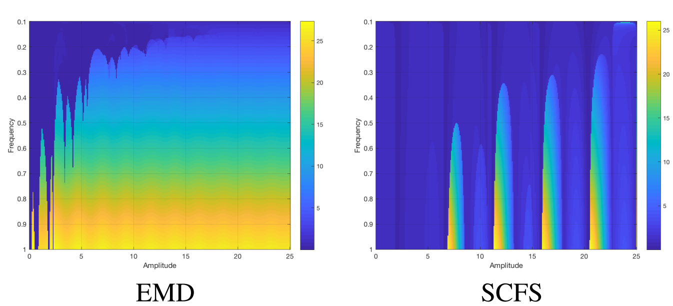

Fig. 3. Tone separation capability: SCFS vs EMD.

$x_3(t) = 4t\cos\left(\frac{2\pi}{5}(t^2-1)\right) + 3\sin(2\pi(t+3))$; its decomposition is shown in Fig. 4. IMF$_2$ fails to correctly track the LF, compared to Mode$_2$. On the other hand, Mode$_1$ suffers from an amplitude attenuation, compared to IMF$_1$.

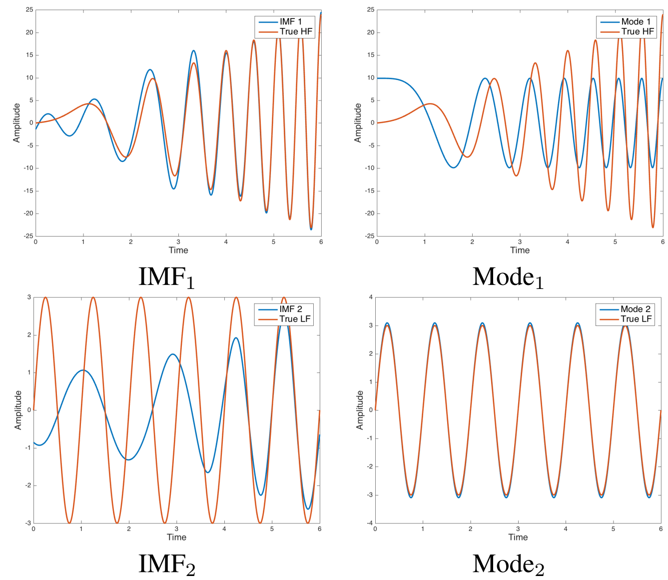

Fig. 4. Decomposition of the signal $x_3(t)$, compared with the true components: SCFS vs EMD.

V. CONCLUSION

The proposed multicomponent AM-FM modeling approach turns out to be an efficient tool to properly separate the frequency content of signals. Based on OMP, the proposed algorithm show great capabilities in frequency separation of linear and nonlinear signals, tone separation, and robustness in a noisy environment. Eliminating the interpolation stage necessary in EMD may explain such improvements. However, some amplitude attenuation is observed in some modes obtained from nonlinear signals, even though this issue does not affect the frequency content of other modes. Ongoing works are on how to resolve that by using a dictionary learning approach, which takes into account the frequency characteristics in the learning process.

VI. REFERENCES

[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] A. Bouchikhi and A. Boudraa, “Multicomponent AM–FM signals analysis based on EMD–B-splines ESA,” *Sig. Proc.*, vol. 92, no. 9, pp. 2214–2228, 2012.

[3] B. Boashash, *Time frequency signal analysis and processing: A comprehensive reference*. The Boulevard,

Langford Lane, Kidlington, Oxford OX5 1GB, UK: Elsevier, 2003.

[4] R. C. Sharpley and V. Vatchev, "Analysis of the Intrinsic Mode Functions," *Constr. Approx.*, vol. 24, pp. 17–47, 2006.

[5] E. H. S. Diop, R. Alexandre, and A. O. Boudraa, "Analysis of intrinsic mode functions: a PDE approach," *IEEE SPL*, vol. 17, no. 4, pp. 398–401, April 2010.

[6] 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.

[7] 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, dec 2011.

[8] T. Y. Hou and Z. Shi, "Extracting a shape function for a signal with intra-wave frequency modulation," *The Royal Society A*, vol. 374, no. 2065, p. 20150194, 2016.

[9] R. McAulay and T. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," *IEEE TASSP*, vol. ASSP-34, no. 4, pp. 744–754, Aug. 1986.

[10] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "Multiple dimensional quasi-eigenfunction approximations and multicomponent am-fm models," *IEEE TIP*, vol. 9, no. 2, pp. 1867–1876, Feb. 2000.

[11] E. H. S. Diop, A. O. Boudraa, and F. Salzenstein, "A joint 2D AM–FM estimation based on higher order Teager–Kaiser energy operators," *Signal, Image and Video Processing*, vol. 5, no. 1, pp. 61–68, Mar. 2011.

[12] M. Elad, *Sparse and redundant representations*. Springer New York, 2010.

[13] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE TSP*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[14] 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.

[15] R. Tibshirani, "Regression shrinkage and selection via the lasso," *JRSS.*, vol. 58, no. 1, pp. 267–288, 1996.

[16] D. L. Donoho, "Compressed sensing," *IEEE TIT*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[17] Y. Isaac, Q. Barthélemy, C. Gouy-Pailler, M. Sebag, and J. Atif, "Multi-dimensional signal approximation with sparse structured priors using split Bregman iterations," *IEE TSP*, vol. 130, pp. 389–402, Jan. 2017.

[18] X. Meng, A. Jakobsson, X. Li, and Y. Lei, "Estimation of chirp signals with time-varying amplitudes," *Signal Processing*, vol. 147, pp. 1–10, June 2018.

[19] M. Kaleem, A. Guergachi, and S. Krishnan, "Empirical mode decomposition based sparse dictionary learning with application to signal classification," in *IEEE DSP/SPE*, Aug. 2013, pp. 18–23.

[20] R. Chen, H. Jia, X. Xie, and W. Gao, "Learning a collaborative multiscale dictionary based on robust empirical mode decomposition," *ArXiv:1704.04422*, 2017.

[21] 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, Aug. 2014.

[22] H. M. Teager, "Some observations on oral air flow during phonation," *IEEE TASSP*, vol. ASSP-28, no. 5, pp. 599–601, Oct. 1980.

[23] P. Maragos, J. Kaiser, and T. Quatieri, "On separating amplitude from frequency modulations using energy operators," in *IEEE ICASSP*, 1992.

[24] 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, Mar. 1991.

[25] N. E. Huang, M.-L. C.Wu, S. R. Long, amuel S. P. Shen, W. Qu, P. Gloersen, and K. L. Fan, "A confidence limit for the empirical mode decomposition and hilbert spectral analysis," *The Royal Society*, vol. 459, no. 2037, pp. 2317–2345, 2003.