NONLINEAR SIGNAL DECOMPOSITION BASED ON BLOCK SPARSE APPROXIMATION

El Hadji S. Diop

NAGIP Group, Department of Mathematics,
University Iba Der Thiam of Thies
Cite Malick Sy, BP 941, Thies, Senegal

Karl Skretting

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

ABSTRACT

We propose here a nonlinear decomposition approach that properly separates the frequency content of signals. To do so, we propose a block-based sparse coding frequency separation (B-SCFS) method with a dictionary constituted of an infinite number of chirp-like atoms so that general signals can be analyzed. Also, a suitable modification of orthogonal matching pursuits (OMP) algorithm is introduced, combined with a block-based procedure for an efficient handling of long signals. Conducted experiments show that the proposed approach noticeably improves on Empirical Mode Decomposition (EMD) and former SCFS as well, which performed well only on a limited class of signals.

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

I. INTRODUCTION

Structuring signal processing algorithms and systems becomes more and more important in order to allow good trade-offs between accuracy and implementation resources. Therefore, various heuristic techniques have been proposed to deal with these trade-offs. In that context, signal representation remains crucial to achieve that, and thus, the need of efficient signal approximation methods for an accurate representation. The main objective in approximation theory is to provide a representation of a signal (or target function) with (ideally a few) basis functions or atoms, which are simpler and easy-to-compute functions. This formulation naturally raised the questions of the trade-offs between the resolution and complexity of the approximation [1]. Approximation theory relies on an accurate approximation of the (target) signal on the basis of its construction thanks to a linear combination of a small number of vectors selected from the dictionary, that is, the collection of all atoms (basis fnctions). In nonlinear approximation, basis functions are not build from a fixed linear space; rather, they depend on the function being approximated (the target) in some way. To find coefficients, various approaches have been proposed [2]–[5]. Here, we are most focused on sparse representations [6], which show great interest in recent years. Sparsity can in fact be exploited to estimate chirp signals [7], [8]. Sparse representation aims at looking for the most compact representation of a signal as a linear combination of atoms in an overcomplete dictionary.

EMD [9] is among nonlinear representation methods, and has been intensively studied since its introduction. Its main goal is to represent a signal in a multicomponent AM (amplitude modulation)-FM (frequency modulation) representation. The basis functions or atoms of EMD are called intrinsic mode functions (IMFs), and are generated in a coarser-to-finer manner on the basis of their frequency contents. Its original algorithmic formulation was for a long time its main criticism, causing a lack of a theoretical background and preventing it for general acceptance in the signal processing community. Remarkable efforts were done so far, and many works proposed a mathematical framework [10]–[14].

SCFS has been recently proposed [15], [16], recasting EMD as a sparse approximation problem and providing a new framework for both its comprehension and formulation. However, some limits actually prevent it to perform well on wider family of signals and real data. The first limit is related to the dictionary which was built as a family of sines and cosines. The second is related to the signal length, SCFS is only efficient on moderate length signals. This work provides two alternatives to improve SCFS. Firstly, an infinite dictionary of chirp atoms is used, and secondly, interpolating between overlapping blocks makes the method appropriate for handling long signals as well.

II. BACKGROUND ON EMD

Given a signal $s(t)$, EMD can be summarized as [9]:

1. Find the extrema of $s(t)$

2. Interpolate the maxima and minima of $s(t)$, denoted respectively $E_{max}(t)$ and $E_{min}(t)$

3. Compute the local mean $m_{loc}(t)$ of $s(t)$:

$$m_{loc}(t) = \frac{1}{2}(E_{max}(t) + E_{min}(t)) \tag{1}$$

4. Extract detail $d(t) = s(t) - m_{loc}(t)$

5. Iterate on the detail.

The refinement iterative steps (1)-(4) constitute the important sifting process where IMFs are extracted by iterating on the residual $d(t)$ until to get a zero local mean (1). Finally, EMD expands $s(t)$ as $s(t) = \sum_{k=1}^{K} s_k(t) + r(t)$, where $s_k(t)$ are IMFs, and $r(t)$ is the residual of the decomposition .

III. SPARSE BLOCK-BASED DECOMPOSITION

A multicomponent AM-FM representation of $s(t)$ gives:

$$s(t) = \sum_{m=0}^{M-1} \gamma_m(t) \cos(\varphi_m(t)) = \sum_{m=0}^{M-1} s_m(t), \tag{2}$$

where the modulation domain is defined as $\{(\gamma_m, \frac{d}{dt}\varphi_m), m = 0, \cdots, M-1\}$. The smooth functions γ_m and φ_m represent respectively the instantaneous amplitude and instantaneous phase components.

III-A. Sparse coding frequency separation method

Let $s = (s[n])_{n=1}^{N}$. Given a full-rank dictionary $D = \{d_i\}_{i=1}^{K}$, one way to proceed is to represent s as:

$$\hat{s} = \sum_{i=1}^{K} d_i \eta_i = D\eta, \tag{3}$$

and most coefficients $\eta_i = 0$ (only M non-zero coefficients). Using (2), each monocomponent $s_k(t)$ can be obtained, using SCFS [15], [16], by solving the sparse coding problem:

$$\eta^\star = \operatorname*{argmin}_{\eta \in \mathbb{R}^K} \frac{1}{2} \|s - D\eta\|_2^2 + \lambda \|\eta\|_p, \tag{4}$$

$\|\cdot\|_p$ is the L^p norm, and $\lambda > 0$ a parameter. An approximation solution of this problem can be found for $p = 0$ by using MP or OMP [17], [18], and for $p = 1$ by using LARS (least angle regression) [19]. Here, we consider the case $p = 0$:

$$\eta^\star = \operatorname*{argmin}_{\eta \in \mathbb{R}^K} \frac{1}{2} \|s - D\eta\|_2^2 + \lambda \|\eta\|_0, \tag{5}$$

$\|\eta\|_0$ denotes the number of nonzero elements in η.

III-B. Parametrized dictionary

In equation (2), the AM, phase and FM functions should be smooth. Thus, for a small segment of the signal, each component can be approximated by a linear function for both the AM and FM functions; therefore, it makes sense to define a dictionary of chirps.

The advanced SCFS makes a sparse approximation of each block using a parametrized dictionary with an infinite number of atoms, where each atom has five parameters:

$$d(t) = (a_0 + a_1 t) \cos(\varphi_0 + \omega t + 0.5\psi t^2), \quad t \in [-1, 1], \tag{6}$$

where a_0 and a_1 are the parameters that define the amplitude function: $a(t) = a_0 + a_1 t$; φ_0, ω and ψ define the phase function: $\varphi(t) = \varphi_0 + \omega t + 0.5\psi t^2$. The frequency function is given as $f(t) = \frac{1}{2\pi}\varphi'(t) = \frac{1}{2\pi}(\omega + \psi t)$. When $\phi(t) = 0$, *i.e.*, $\varphi_0 = \omega = \psi = 0$, then, the atom and amplitude function define a linear trend. For the atom in (6), the coefficient is fixed to 1. A discrete atom of length N is assumed to be centered around $t = 0$, evenly spaced between $t = -1$ and $t = 1$. Thus, $\mathbf{t}$ is a N-length column vector. Let us denote $\circ$ the elementwise multiplication operation. Then, $\varphi(\mathbf{t}) = \varphi_0 + \omega \mathbf{t} + 0.5\psi\ \mathbf{t} \circ \mathbf{t}$ and $\cos(\varphi(\mathbf{t}))$ are column vectors of length N. Actually, each $d(\mathbf{t})$ can be expressed as a linear combination of two basis vectors in two different ways:

$$d(\mathbf{t}) = [\cos\varphi(\mathbf{t}),\ \mathbf{t} \circ \cos\varphi(\mathbf{t})] \cdot [a_0, a_1]^T \quad \text{and} \tag{7}$$

$$d(\mathbf{t}) = \left[\left(\mathbf{1} + \frac{a_1}{a_0}\mathbf{t}\right) \circ \cos(\overline{\varphi}(\mathbf{t})), \left(\mathbf{1} + \frac{a_1}{a_0}\mathbf{t}\right) \circ \sin(\overline{\varphi}(\mathbf{t}))\right] \cdot [a_c, a_s]^T \tag{8}$$

where $\overline{\varphi}(\mathbf{t})$ is $\varphi(\mathbf{t})$ with parameter $\varphi_0 = 0$. For (8), after (new) coefficients $\{a_c, a_s\}$ are found, the parameters $a_0 = \sqrt{a_c^2 + a_s^2}$ and $\varphi_0 = \operatorname{atan2}(-a_s, a_c)$ are updated, $\operatorname{atan2}$ denotes the four-quadrant inverse tangent. The observations above are useful when searching for an atom that best match a given signal; the objective function (often norm of residual), which is to be minimized, can be a function of three arguments $\{\varphi_0, \omega, \psi\}$ or $\{a_1/a_0, \omega, \psi\}$ instead of five parameters. Combining both and allowing a small error may reduce the objective function to two arguments.

III-C. Block processing OMP approach

Standard OMP algorithms cannot be used with a dictionary consisting of all parametric atoms as given by (6), since there is an infinite number of atoms in this dictionary. Therefore, a special approach for OMP has been proposed, consisting in repeating a loop K times, ones for each atom to select. At iteration k, the residual $\mathbf{r}_{k-1} = \mathbf{s} - \tilde{\mathbf{s}}_{k-1}$ is found as the original signal block subtracted from its approximation at step $k-1$ denoted $\tilde{s}_{k-1}$. The approximation is the best possible representation in the space spanned by the previously selected atoms, $\operatorname{Span}\{d_i(\mathbf{t})\}_{i=1}^{k-1}$. Each loop consists of the following two steps:

- **Finding an initial set of parameters**: Fast Fourier transform (FFT) is used on the first $2/3$ of the signal to find instantaneous frequency (IF) at $1/3$ of the signal ($t = -1/3$). Then, FFT is used on the last $2/3$ of the signal to find the IF at $2/3$ of the signal ($t = 1/3$). A window function and zero padding may be used when FFT is done. The mean of the two frequencies gives the initial value for center frequency, and the slope gives the chirpyness. The phase is estimated by projecting the signal onto the sine and cosine parts, and using the amplitude of each part.
- **Optimizing chirp parameters**: We first keep the ratio $\{\frac{a_1}{a_0}\}$ fixed, and optimize over the free variables $\{\omega, \psi\}$ doing some few iterations of the BFGS (Broyden-Fletcher-Goldfarb-Shanno)method [20], [21]. $\{a_0, \phi\}$ depend on the free variables as given by the solution of (8). (8) is solved for $\{a_c, a_s\}$, and the wanted parameters $\{a_0 = \sqrt{a_c^2 + a_s^2},\quad \varphi_0 = \operatorname{atan2}(-a_s, a_c)\}$ can be calculated directly from these. This is done within the objective function;

each time the objective function is executed with (new) values for $\{\omega, \psi\}$. After optimization, $\{\varphi_0, \omega, \psi\}$ are fixed, and then $\{a_0, a_1\}$ are found by solving (7).

III-D. Finding decomposition modes for long signals

The procedure above can be used for signals of 10 to 2000 or even 5000 samples, depending on the signal properties. But usually longer signals (more than 500 samples) should be divided into overlapping windowed blocks. Each block is decomposed into its chirp atoms with the previously described OMP, one of the atoms may be a linear trend (zero frequency). Then, the reconstructed block components can be merged together on the basis on how the blocks initially overlapped each other. This procedure may give an accurate sparse description of both the original signal and its AM-FM components. Steps of the block-based SCFS are presented in Algorithm 1. Its convergence follows since the core of the algorithm is based on OMP which converges [3], [18].

Algorithm 1: Proposed block-based SCFS algorithm.

Input : signal s of length L, number of modes (components) K, number of blocks M, length of each block N, and weights (window) to use on each block w.
Output: Decomposition modes $y = (y_i)_{i=1}^{K}$.

```
Initialization:
Set modes y = (y_i)_{i=1}^K and σ to 0. σ is sum of weights
  and has length L.
Divide s into M overlapping blocks each of length N, i.e.,
find {P_m}_{m=1}^M as index for first element in each block.
foreach block: m = 1, 2, ..., M do
    Add weights w to sum of weights s as follows:
    foreach element: n = 1, 2, ..., N do
        p = P_m + n − 1
        σ[p] = σ[p] + w[n]
    end
end
foreach block: m = 1, 2, ..., M do
    The signal block is s_m = s[P_m : (P_m + N − 1)]
    Use OMP on s_m to find K chirp atoms:
      {d_k, a_{0,k}, a_{1,k}, φ_k, ω_k, ψ_k}_{k=1}^K
    Sort atoms based on ω:
      {i_k}_{k=1}^K; ω_{i_1} < ω_{i_2} < ··· < ω_{i_K}
    foreach atom: j = 1, 2, ..., K do
        Add weighted atom to mode (component) j as:
        foreach element: n = 1, 2, ..., N do
            p = P_m + n − 1
            y_j[p] = y_j[p] + (w[n]/σ[p]) · d_{i_j}[n].
        end
    end
end
```

IV. EXPERIMENTS

Experiments are conducted on two nonlinear AM-FM signals $s_1(t)$ and $s_2(t)$ represented in Fig. 1, as well as their spectrograms. $s_1(t)$ has 769 samples and is composed of two components: a pure sine with a constant amplitude equal to 3 and frequency equal to 11 Hz, and an AM-FM component with linearly varying amplitude and frequency:
$s_1(t) = 3\sin(22\pi(t+3)) + a(t)\cos(\varphi(t)+\varphi_0)$, where $a(t)$ varies linearly in $[1;3]$, its IF also varies linearly in $[5;10]$ with a random phase φ_0.
The signal $s_2(t)$ has 2001 samples and is defined as:
$s_2(t) = s_2^1(t) + s_2^2(t)$, where $s_2^1(t) = a_0\cos(\varphi(t)+\varphi_0)$, $s_2^2(t) = a(t)\cos(\psi(t)+\psi_0)$, $a_0 = 1.2$, the instantaneous frequency (IF) of $s_2^1(t)$ linearly increases in $[50;150]$, $a(t)$ linearly decreases in $[0.3;1.1]$, IF of $s_2^2(t)$ linearly decreases in $[200;100]$, φ_0 and ψ_0 are random phases. The decomposition of $s_1(t)$ using the proposed B-SCFS is obtained by dividing $s_1(t)$ into 10 overlapped blocks of 200 samples using rectangular windows. The different decomposition results using B-SCFS, SCFS and EMD are shown in Fig. 2. This example highlights the weaknesses previously discussed on the limitations of SCFS. For EMD, we only show the first two IMFs, IMF_1 and IMF_2. This example clearly illustrates also the limits of EMD, and exhibits the good behavior of the proposed method.

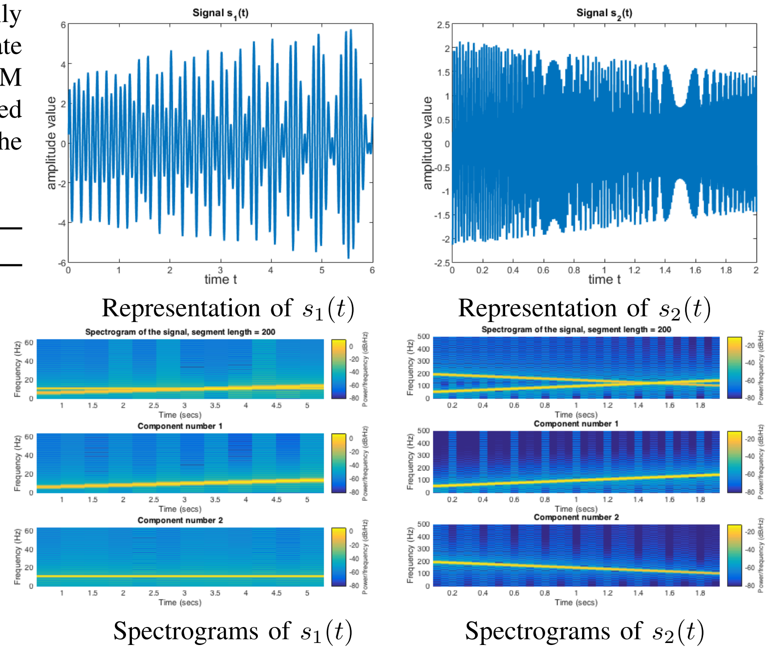

Fig. 1. Representations and spectrograms of $s_1(t)$ and $s_2(t)$.

For $s_2(t)$, we divide it into 26 overlapped blocks of 200 samples using rectangular windows. B-SCFS provides two modes, while EMD provides ten IMFs. B-SCFS and EMD results are displayed in Fig. 3. In addition, we also provide a quantitative evaluation by computing the mean squared errors (MSE) and L^2 norm; results obtained with B-SCFS, SCFS and EMD are compared with the true high frequency (HF) and low frequency (LF) components. We expect the first modes perfectly match the HF and LF, respectively. The lower the errors, the better the method, quantitative results

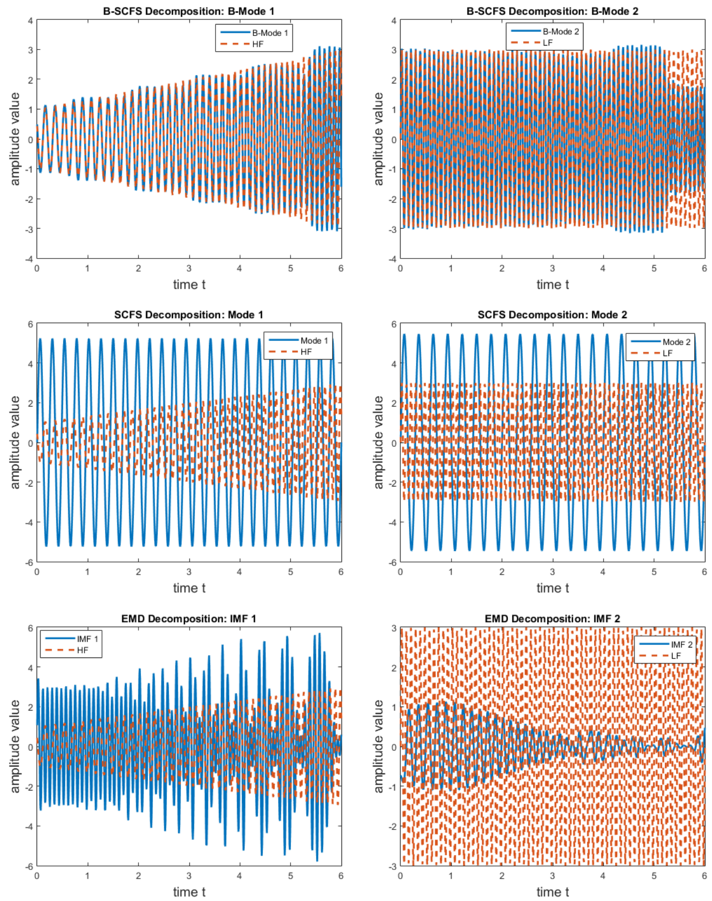

Fig. 2. Decomposition of $s_1(t)$. Top line: B-SCFS. Middle: SCFS. Bottom: EMD.

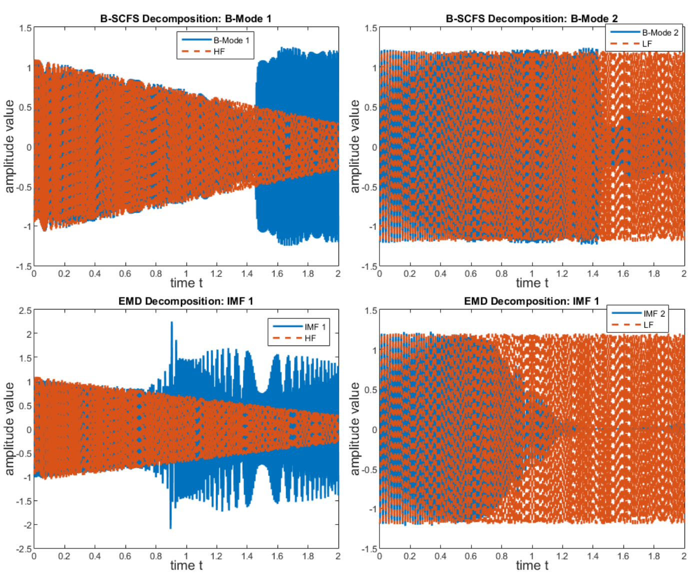

Fig. 3. $s_2(t)$ decomposition. Top: B-SCFS. Bottom: EMD.

are reported in Table I confirming the good behaviors of the proposed B-SCFS.

Table I. Decomposition errors: modes vs. true components.

Metrics	$s_1(t)$			$s_2(t)$		
	B-SCFS	SCFS	EMD	B-SCFS	SCFS	EMD
$\Vert$ HF - $M_1\Vert_2$	14.08	110.8	60.34	20.78	100.8	29.03
MSE(HF, M_1)	0.26	15.97	4.73	0.21	5.07	0.42
$\Vert$ LF - $M_2\Vert_2$	12.77	122.2	60.31	21.2	96.46	28.9
MSE(LF, M_2)	6.21	19.41	4.73	0.22	5.01	0.42
$\Vert s_i - \hat{s}_i\Vert_2$	8.87	164.7	1.82	2.4	139.4	2.67
MSE(s_i, $\hat{s}_i$)	0.1	35.28	0.004	0.003	9.71	0.0036

Finally, the noise robustness of B-SCFS is studied. To proceed, a Gaussian noise is added to $s_1(t)$ by applying different signal-to-noise-ratios (SNRs) from 32 db to -2 db. For these values of SNRs, we compare the B-Mode$_1$ (first B-SCFS mode) and Mode$_1$ (first SCFS mode) with the true uncorrupted HF by computing the root MSE (RMSE) between them, the same is done for the second modes. The plotting of the quantitative results is displayed in Fig. 4. These results should be examined in the meanwhile with the poor decomposition results obtained with SCFS (*cf.* Figure 2, middle line), which is a consequence of the limits of SCFS perviously discussed. For EMD, we obtain over ten IMFs and observe that first IMFs hold the noise; it is also very difficult to identify which IMF is close to HF or LF).

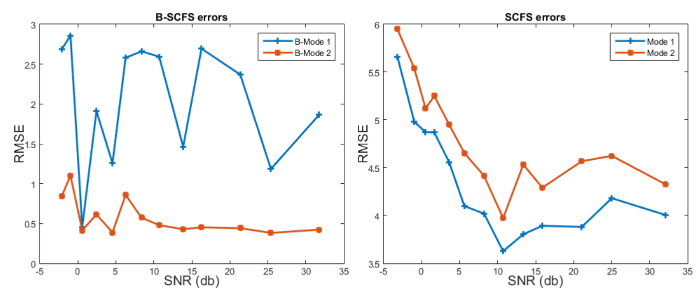

Fig. 4. Decomposition errors of the noisy signal.

V. CONCLUSION

A nonlinear decomposition approach that properly separates the frequency content of signals has been proposed in this work. Indeed, a block-based sparse coding method is proposed with a dictionary composed of an infinite number of chirp-like atoms making a more effective analysis of more general signals. A suitable modification combined with a block-based strategy of orthogonal matching pursuits algorithm is then introduced. Obtained results show noticeably improvements in comparison to the former method and EMD as well.

VI. REFERENCES

[1] R. A. DeVore, “Nonlinear approximation,” *Acta Numerica*, vol. 7, pp. 51–150, Jan. 1998.

[2] T. Blu and M. Unser, “Quantitative Fourier Analysis of Approximation Techniques: Part I- Interpolators and Projectors,” *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2783–2795, Oct. 1999.

[3] Y. C. Pati, R. Razaiifar, and P. S. Krishnaprasad, “Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition,” in *ACSSC*, November 1993, pp. 1–5.

[4] S. Mann and S. Haykin, “The chirplet transform: physical considerations,” *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2745–2761, 1995.

[5] R. Gribonval, “Fast matching pursuit with a multiscale dictionary of gaussian chirps,” *IEEE TSP*, vol. 49, no. 5, pp. 994–1001, may 2001.

[6] M. Elad, *Sparse and Redundant Representations*. Springer New York, 2010.

[7] J. Sward, J. Brynolfsson, A. Jakobsson, and M. Hansson-Sandsten, “Sparse semi-parametric estimation of harmonic chirp signals,” *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1798–1807, apr 2016.

[8] X. Tu, J. Swärd, A. Jakobsson, and F. Li, “Estimating nonlinear chirp modes exploiting sparsity,” *Signal Processing*, vol. 183, p. 107952, jun 2021.

[9] 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,” *The Royal Society*, vol. 454, pp. 903–995, 1998.

[10] 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, Mar. 2011.

[11] 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,” *Advances in Computational Mathematics*, vol. 38, no. 4, pp. 801–835, dec 2011.

[12] H. tieng Wu, “Instantaneous frequency and wave shape functions (I),” *Applied and Computational Harmonic Analysis*, vol. 35, no. 2, pp. 181–199, sep 2013.

[13] P. Tavallali, T. Y. Hou, and Z. Shi, “Extraction of intrawave signals using the sparse time-frequency representation method,” *Multiscale Modeling & Simulation*, vol. 12, no. 4, pp. 1458–1493, jan 2014.

[14] T. Y. Hou and Z. Shi, “Extracting a shape function for a signal with intra-wave frequency modulation,” *The Roy. Soc. A*, vol. 374, no. 2065, mar 2016.

[15] E. H. S. Diop and K. Skretting, “Frequency separation method based on sparse coding,” in *ICASSP*, Brighton, UK, may 2019, pp. 5192–5196.

[16] E. H. S. Diop, K. Skretting, and A.-O. Boudraa, “Multicomponent AM–FM signal analysis based on sparse approximation,” *IET Signal Processing*, vol. 14, no. 1, pp. 32–43, feb 2020.

[17] S. G. Mallat and Z. Zhang, “Matching Pursuits With Time-Frequency Dictionaries,” *IEEE TSP*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[18] J. A. Tropp and A. C. Gilbert, “Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit,” *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[19] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, “Least Angle Regression,” *The Annals of Statistics*, vol. 32, no. 2, pp. 407–451, Apr. 2004.

[20] C. G. Broyden, “The convergence of a class of double-rank minimization algorithms 1. general considerations,” *IMA Journal of Applied Mathematics*, vol. 6, no. 1, pp. 76–90, 1970.

[21] Y.-H. Dai, “Convergence properties of the BFGS algoritm,” *SIAM Journal on Optimization*, vol. 13, no. 3, pp. 693–701, jan 2002.