Signal Processing 183 (2021) 107952

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Estimating nonlinear chirp modes exploiting sparsity

Xiaotong Tu[a], Johan Swärd[c], Andreas Jakobsson[c], Fucai Li[b,*]

[a] School of Informatics, Xiamen University, Xiamen 361005, China
[b] State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China
[c] Centre for Mathematical Sciences, Lund University, Sweden

ARTICLE INFO

Article history:
Received 21 September 2020
Revised 10 December 2020
Accepted 22 December 2020
Available online 12 January 2021

Keywords:
Time-frequency analysis
Fault diagnosis
Nonlinear chirps
Sparse modeling

ABSTRACT

The decomposition of nonlinear chirp modes is a challenging task, typically requiring prior knowledge of the number of modes a signal contains. In this work, we present a greedy nonlinear chirp mode estimation (NCME) technique that forms the used decomposition basis from the signal itself, using an arctangent demodulation technique. The resulting decomposition is formed by considering the residual energy and smoothness of the instantaneous amplitude. We also derive a computationally efficient implementation of the resulting estimator, using the alternating direction method of multipliers (ADMM). Numerical simulations and experimental data analysis illustrate the effectiveness and advantages of the proposed method.

1. Introduction

Time-frequency analysis plays a significant role in analyzing and processing nonlinear signals whose frequency and amplitude vary with time [1,2]. Such non-linear signals occur in a wide range of applications, for example, in human speech [3,4], vibrational signals [5], and in gravitational waves [6]. Given the importance of such signals, numerous approaches have been proposed for estimating the modes detailing this form of signals [7].

Examples include Huang's empirical mode decomposition (EMD), which is commonly used to decompose a signal into a set of intrinsic mode functions [8]. As a data-driven method, the EMD implements a recursive sifting procedure to estimate the modes one by one. Since the EMD is able to separate different modes, this method has received notable attention and has been applied in many fields including finance [9], mechanical engineering [10], and meteorology [11].

Inspired by the EMD, Daubechies described an EMD-like tool called the synchrosqueezing transform (SST) to deal with a non-stationary signal [12]. By the combination of a continuous wavelet transform and a reassignment technique, the SST possesses a precise mathematical definition. The SST has shown its power in a wide range of applications and several extensions have been developed to further improve its performance [13]. Since the original SST can only obtain an unbiased estimation of approximately harmonic components, use of a higher-order phase function was proposed as a generalization [6,14]. At present, although computationally cumbersome, the high-order SST can theoretically be extended to any order of phase functions [6]. However, these SST-based methods provide instantaneous frequency (IF) information in the time-frequency plane in a subjective way. If one wants to extract the accurate value of the IF, further steps are required.

Much effort has focused on the estimation of modes that are well described using spectral lines. For instance, Gills proposed the empirical wavelet transform (EWT) to decompose a signal into subbands by constructing an adaptive wavelet basis [15]. In order to perform spectrum segmentation, this method first determines the spectral peaks of each mode. After that, the corresponding wavelet filter is built to recover each of the modes. However, the EWT heavily relies on the used peak detection algorithm, thereby making it difficult to separate closely spaced modes. Later, the variational mode decomposition (VMD), which uses a recursive algorithm to recover each mode, was proposed by Dragomiretskiy et al [16]. Since the VMD considers a joint optimization algorithm to estimate all of the modes, it can generate a more accurate result for closely spaced modes than the EWT. However, neither the EWT nor the VMD can recover wide-band modes whose frequency ranges overlap.

In the past decade, the use of sparse reconstruction models has been explosively applied in several areas of signal processing and machine learning. In many of these fields, the use of a sparse model has been proved to be applicable in time-frequency analysis and source separation [17]. In [18], a sparse semi-parametric method was proposed that allows for estimating a sum of chirp

* Corresponding author.
E-mail addresses: xttu@xmu.edu.cn (X. Tu), js@maths.lth.se (J. Swärd), aj@maths.lth.se (A. Jakobsson), fcli@sjtu.edu.cn (F. Li).

https://doi.org/10.1016/j.sigpro.2020.107952
0165-1684/

signals without assuming a prior knowledge of the number of modes in the signal. This method is able to approximate nonlinear chirps by dividing the signal into short intervals which are then assumed to be stationary or well modelled as containing linear chirps. In [19], this algorithm was further developed to estimate chirp signals with time-varying amplitudes. As a further alternative, the intrinsic chirp component decomposition was proposed to process a general chirp signal [20]. Generally, these approaches require a dictionary containing the set of considered candidate modes.

In this work, a generalized method, termed the nonlinear chirp mode estimation (NCME), is proposed to model a large part of a signal which contains several components with time-varying amplitudes and frequencies. A sparse model is adopted to extract the modes one by one, by explicitly enforcing the smoothness of the estimated parameters. The estimator is formed in an order-recursive manner, such that no prior knowledge of the number of modes is required. Furthermore, the NCME can directly estimate the instantaneous amplitude (IA) and the IF of the signal, which is accomplished without segmenting the signal into short intervals. The performance of the method is demonstrated using both simulated and real signals. We further introduce a computationally efficient implementation, employing the alternating direction method of multipliers (ADMM). The remainder of this paper is organised as follows:

In the next section, the NCME is proposed together with the assumed signal model, including the introduced efficient implementation. Section III provides a comparision and a validation of the proposed method by analyzing a simulated signal and two real-world signals. Finally, conclusions are described in Section IV.

2. Nonlinear chirp mode estimation

2.1. Signal model

Consider a multicomponent non-stationary signal $y(t)$ that may be well-modelled as

$$y(t) = \sum_{k=1}^{K} y_k(t) = \sum_{k=1}^{K} a_k(t) \cos\left(2\pi \int_0^t f_k(\tau) d\tau + \phi_k\right) \tag{1}$$

where K denotes the number of chirp modes, $a_k(t)$ the instantaneous amplitude of the k^{th} mode, and f_k and ϕ_k the instantaneous frequency and initial phase, respectively. Here, we assume that the instantaneous amplitude $a_k(t)$ and instantaneous frequency f_k are nonnegative numbers.

One may rewrite (1) as

$$y(t) = \sum_{k=1}^{K} u_k(t) \cos\left(2\pi \int_0^t f_k(\tau) d\tau\right) + v_k(t) \sin\left(2\pi \int_0^t f_k(\tau) d\tau\right) \tag{2}$$

with

$$\begin{aligned} u_k(t) &= a_k(t) \cos(2\pi\phi_k) \\ v_k(t) &= -a_k(t) \sin(2\pi\phi_k) \end{aligned} \tag{3}$$

and

$$a_k(t) = \sqrt{u_k{}^2(t) + v_k{}^2(t)} \tag{4}$$

2.2. Algorithm

The NCME algorithm is formed using a matching pursuit formulation such that the modes are estimated in a greedy order-recursive manner, such that the k^{th} component is estimated as

$$\min_{u_k, v_{k,f_k}} \|y(t) - y_k(t)\|_2^2 + \lambda\left(\left\|u''_k(t)\right\|_1 + \left\|v''_k(t)\right\|_1\right) \tag{5}$$

with

$$y_k = u_k \cos\left(2\pi \int_0^t f_k(\tau) d\tau\right) + v_k \sin\left(2\pi \int_0^t f_k(\tau) d\tau\right) \tag{6}$$

where the first term in (5) restricts the residual energy by extracting the desired modes. The remaining two terms are selected to enforce smoothness of the instantaneous amplitude. The user parameter λ allows for a tuning of the relative weighting between the cost function and the desired degree of sparseness of the solution.

For convenience, the minimization in (5) may be expressed as

$$\|\mathbf{y} - \mathbf{A}_k \mathbf{x}_\mathbf{k}\|_2^2 + \lambda \|\mathbf{D}\mathbf{x}_k\|_1 \tag{7}$$

where $\boldsymbol{y} = [y(t_1), \cdots, y(t_N)]^T$

$\boldsymbol{f}_k = [f_k(t_1), \cdots, f_k(t_N)]^T$

$\boldsymbol{u}_k = [u_k(t_1), \cdots, u_k(t_N)]^T$

$\boldsymbol{v}_k = [v_k(t_1), \cdots, v_k(t_N)]^T$

$\boldsymbol{x}_k = \left[\boldsymbol{u}_k^T \ \boldsymbol{v}_k^T\right]^T$

$\boldsymbol{C}_k = diag[\cos(\theta_k(t_1)), \cdots, \cos(\theta_k(t_N))]^T$

$\boldsymbol{S}_k = diag[\sin(\theta_k(t_1)), \cdots, \sin(\theta_k(t_N))]^T$

$\boldsymbol{A}_k = [\boldsymbol{C}_k \ \boldsymbol{S}_k]$

$\theta_k(t) = 2\pi \int_0^t f_k(\tau) d\tau$

$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{H} & 0 \\ 0 & \boldsymbol{H} \end{bmatrix}$

with $\boldsymbol{H} \in \boldsymbol{R}^{(N-2) \times N}$ the second-order difference matrix

$$\boldsymbol{H} = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -2 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 & -2 & 1 \end{bmatrix}$$

The minimization of (7) may thus be expressed as a generalized Lasso problem, which, as shown in the following, may be solved efficiently using an ADMM formulation. This yields an estimate of $\boldsymbol{x}_k = [\boldsymbol{u}_k^T \ \boldsymbol{v}_k^T]^T$ as well as $\boldsymbol{A}_k = [\boldsymbol{C}_k \ \boldsymbol{S}_k]$. According to the arctangent demodulation technique [21], $\boldsymbol{f}_k = [f_k(t_1), \cdots, f_k(t_N)]^T$ may be updated using the estimated $\boldsymbol{x}_k = [\boldsymbol{u}_k^T \ \boldsymbol{v}_k^T]^T$. Thus, the IF increment may be accomplished as

$$\begin{aligned} \Delta\tilde{f}_k^i(t) &= -\frac{1}{2\pi}\frac{d}{dt}\left(\arctan\left(\frac{v_k^i(t)}{u_k^i(t)}\right)\right) \\ &= \frac{v_k^i(t) \cdot \frac{d}{dt}\left(u_k^i(t)\right) - u_k^i(t) \cdot \frac{d}{dt}\left(v_k^i(t)\right)}{2\pi\left(\left(u_k^i(t)\right)^2 + \left(v_k^i(t)\right)^2\right)} \end{aligned} \tag{8}$$

See Appendix A for a proof of (8). Since (8) is computed in discrete-time, it may suffer from numerical errors being propagated in the iterations. To avoid this, one may enforce that the IF increment should be band-limited by forming the update as

$$\min_{\Delta\mathbf{f}_k^i} \left\|\Delta\tilde{\mathbf{f}}_k^i - \Delta\mathbf{f}_k^i\right\|_2^2 + \gamma \left\|\mathbf{H}\Delta\mathbf{f}_k^i\right\|_2^2 \tag{9}$$

Algorithm 1 The NCME algorithm.

1: Initialization: signal $\boldsymbol{y}$; weight parameters λ and γ; stopping criteria ε, δ and maximum number of iterations M; $k=1$ and $\boldsymbol{r}_k=\boldsymbol{y}$.
2: **if** $\|\boldsymbol{r}_k\|_2^2/\|\boldsymbol{y}\|_2^2>\varepsilon$ **then**
3: Set $i=1$, initialize the IF $f_k^1(t)$ and form $\boldsymbol{A}_k^1$ using (7)
4: **if** $\|\boldsymbol{y}_k^{i+1}-\boldsymbol{y}_k^i\|_2^2>\delta\|\boldsymbol{y}_k^i\|_2^2$ *and* $i<M$ **then**
5: Use the Lasso to compute $\boldsymbol{x}_k^i=[\boldsymbol{u}_k^i;\boldsymbol{v}_k^i]$
6: $\boldsymbol{y}_k^i=\boldsymbol{A}_k^i\boldsymbol{x}_k^i$
7: $\Delta\tilde{f}_k^i(t)=\frac{v_k^i(t)\cdot\frac{d}{dt}(u_k^i(t))-u_k^i(t)\cdot\frac{d}{dt}(v_k^i(t))}{2\pi((u_k^i(t))^2+(v_k^i(t))^2)}$
8: $\boldsymbol{f}_k^{i+1}=\boldsymbol{f}_k^i+(\boldsymbol{I}+\frac{1}{\gamma}\boldsymbol{H}^T\boldsymbol{H})^{-1}\Delta\tilde{\boldsymbol{f}}_k^i$
9: Update $\boldsymbol{A}_k^{i+1}$ using (7)
10: $i=i+1$
11: **end if**
12: Set $\boldsymbol{y}_m=\boldsymbol{y}_k^i$, $\boldsymbol{f}_m=\boldsymbol{f}_k^{i+1}$,
and form $\boldsymbol{r}_{k+1}=\boldsymbol{r}_k-\boldsymbol{y}_k$
13: $k=k+1$
14: **end if**
15: Extract signal modes $\{\boldsymbol{y}_k\}$ and IFs $\{\boldsymbol{f}_k\}$

where $\gamma>0$ denotes a weighting parameter, and where $\Delta\tilde{\boldsymbol{f}}_k^i=[\Delta\tilde{f}_k^i(t_1),\cdots,\Delta\tilde{f}_k^i(t_N)]^T$. This is a ridge regression problem, implying that

$$\Delta\mathbf{f}_k^i=\left(\mathbf{I}+\frac{1}{\gamma}\mathbf{H}^T\mathbf{H}\right)^{-1}\Delta\tilde{\mathbf{f}}_k^i \tag{10}$$

where $\boldsymbol{I}$ represents an identity matrix. Finally, the IF can then be updated as

$$\mathbf{f}_k^{i+1}=\mathbf{f}_k^i+\Delta\mathbf{f}_k^i \tag{11}$$

One then employs the estimated IF $\boldsymbol{f}_k^{i+1}$ in order to again update the matrix $\boldsymbol{A}_k^{i+1}$. The algorithm will iteratively update in this manner until the stopping criteria are achieved, such that the estimated IF does no longer change significantly in each iteration or that a maximum number of iterations is reached. After the mode is extracted, the estimated mode is then removed from the original signal $\boldsymbol{y}$. In a similar manner, all the modes are then extracted, one by one, until the residual energy is sufficiently small. Since the proposed method is able to estimate the IFs, as well as the IAs of all the signal modes, the time-frequency representation (TFR) of the signal can be defined as

$$NCME(t,f)=\sum_{k=1}^{K}a_k(t)\delta(f-f_k(t)) \tag{12}$$

where $NCME(t,f)$ represents the TFR of the signal with $\delta(t)$ denoting the Dirac function.

The algorithm is summarized in Algorithm 1.[1] It is suggested that the stopping threshold δ is usually set to 10^{-9}-10^{-7}, and the ε is often set to 10^{-2}- 10^{-1}, and parameter λ as 10^1- 10^4. At present, some schemes have been proposed to update the parameter λ, and find the optimal λ [22,23].

2.3. Efficient implementation

In order to solve the generalized Lasso problem in (7), an efficient implementation using the ADMM framework is proposed.

Generally, the ADMM solves problems on the form [24]

$$\begin{aligned}\min_{\mathbf{x},\mathbf{z}}\quad & \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2+\lambda\|\mathbf{z}\|_1\\ subject\,to\quad & \mathbf{D}\mathbf{x}-\mathbf{z}=\mathbf{0}\end{aligned} \tag{13}$$

where $\boldsymbol{A}$, $\boldsymbol{x}$ here represent $\boldsymbol{A}_k$, $\boldsymbol{x}_k$ in (7), respectively.

[1] The Matlab implementation of the algorithm will be provided online (https://github.com/tormii/NCME) upon publication.

The augmented Lagrangian for this minimization may be written as

$$L(\mathbf{x},\mathbf{z},\mathbf{u})=\frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2+\lambda\|\mathbf{z}\|+\frac{\rho}{2}\|\mathbf{D}\mathbf{x}-\mathbf{z}+\mathbf{u}\|_2^2 \tag{14}$$

where $\rho>0$ denotes the penalty parameter and $\boldsymbol{u}$ the scaled dual variable. The ADMM decomposes the original optimal problem into several sub-problems, which finds the solution to (13) by iteratively minimizing (14) for each variable separately using

$$\mathbf{x}^{i+1}=\arg\min_{\mathbf{x}}\left(\frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2+\frac{\rho}{2}\left\|\mathbf{D}\mathbf{x}-\mathbf{z}^i+\mathbf{u}^i\right\|_2^2\right) \tag{15}$$

$$\mathbf{z}^{i+1}=\arg\min_{\mathbf{z}}\left(\lambda\|\mathbf{z}\|+\frac{\rho}{2}\left\|\mathbf{D}\mathbf{x}^{i+1}-\mathbf{z}+\mathbf{u}^i\right\|_2^2\right) \tag{16}$$

$$\mathbf{u}^{i+1}=\mathbf{D}\mathbf{x}^{i+1}-\mathbf{z}^{i+1}+\mathbf{u}^i \tag{17}$$

Specifically, one may form an analytic solution to (15) by differentiating with respect to $\boldsymbol{x}_k$, yielding

$$\mathbf{x}^{i+1}=\left(\mathbf{A}^T\mathbf{A}+\rho\mathbf{D}^T\mathbf{D}\right)^{-1}\left(\mathbf{A}^T\mathbf{y}+\rho\mathbf{D}^T\left(\mathbf{z}^i-\mathbf{u}^i\right)\right) \tag{18}$$

Though (16) is not differentiable at $\boldsymbol{z}_k=0$, one may solve for $\boldsymbol{z}^{i+1}$ using

$$\mathbf{z}^{i+1}=\begin{cases}\mathbf{D}\mathbf{x}^{i+1}+\mathbf{u}^i-\frac{\lambda}{\rho}\mathbf{D}\mathbf{x}^{i+1}+\mathbf{u}^i\geq\frac{\lambda}{\rho}\\ \mathbf{D}\mathbf{x}^{i+1}+\mathbf{u}^i+\frac{\lambda}{\rho}\mathbf{D}\mathbf{x}^{i+1}+\mathbf{u}^i\leq-\frac{\lambda}{\rho}\\ 0otherwise\end{cases} \tag{19}$$

This formula can more compactly be expressed using the soft threshold operator $S_{\lambda/\rho}(\cdot)$, i.e.,

$$\mathbf{z}^{i+1}=S_{\lambda/\rho}\left(\mathbf{D}\mathbf{x}^{i+1}+\mathbf{u}^i\right) \tag{20}$$

3. Comparison and validation

To investigate the performance of the NCME, we examine both simulated and real signals.

3.1. Simulated signal

We initially consider a simulated signal formed as

$$\begin{aligned}&s(t)=a(t)\cos\left(2\pi\phi(t)\right)\\ &with\begin{cases}a(t)=sawtooth(40\pi t,0.1)+2\\ \phi(t)=0.8+50t+525t^2-300t^3\end{cases}\end{aligned} \tag{21}$$

where $a(t)$ is the instantaneous amplitude and $\phi(t)$ represents the instantaneous phase of the signal. Specifically, the IF of the signal can be expressed as $f(t)=\frac{d\phi}{dt}=-900t^2+1050t+50$. Here, $sawtooth(T,W)$ denotes a function which can generate a modified

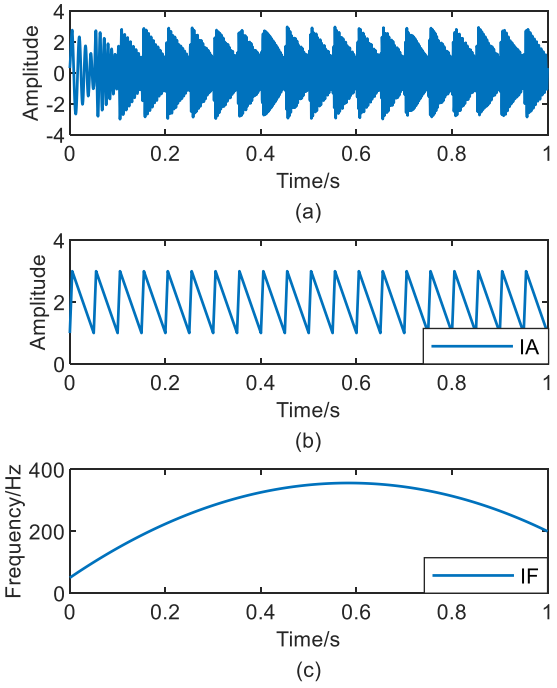

Fig. 1. (a) Waveform of the signal, (b) its instantaneous amplitude, and (c) its instantaneous frequency.

Table 1
EQF Corresponding to different methods.

Methods	NCME	VNCMD
IA	39.3	22.8
IF	45.4	29.5
Mode	41.7	23.4

triangle wave, where W, a scalar parameter between 0 and 1, determines the fraction between 0 and 2π at which the maximum occurs. The function increases from -1 to 1 on the interval 0 to $2\pi W$, then decreases linearly from 1 back to -1 on the interval $2\pi W$ to 2π. The sampling frequency is 2 kHz with a sampling period of 1 s. The time-domain signal, its IA, and IF are shown in Fig. 1.

The proposed method and the variational nonlinear chirp mode decomposition (VNCMD) [7] are used to estimate the modes of the signal. Here, the estimated quality factor (EQF), defined as

$$EQF = 20\log_{10}\frac{\|x\|_2}{\|x-\tilde{x}\|_2} \tag{22}$$

is used to evaluate the accuracy of the algorithm, where x and $\tilde{x}$ represent the theoretical and estimated value, respectively. Thus, a greater value of EQF indicates a better estimation accuracy. The EQF results by the two methods are shown in

Table 1, where the initial IFs of both two methods being chosen as $IF(t) = 300$. As can be seen, the NCME achieves a higher estimation accuracy than VNCMD for both the IA and the IF. Fig. 2 illustrates the differences in the estimation performance of the methods, clearly showing that the NCME estimates well match the theoretical values. Similarly, it may be seen that the VNCMD fails to estimate the true trend of the IA and overfitting occurs especially in the slope change points. The found improvement is due to the fact that the NCME enforce a piecewise linear structure on the IA estimate due to the used l_1 norm, making it more suitable to detail signals with abrupt changing IA. As the VNCMD strives to form a smooth IA estimate, the estimate causes an overfitting

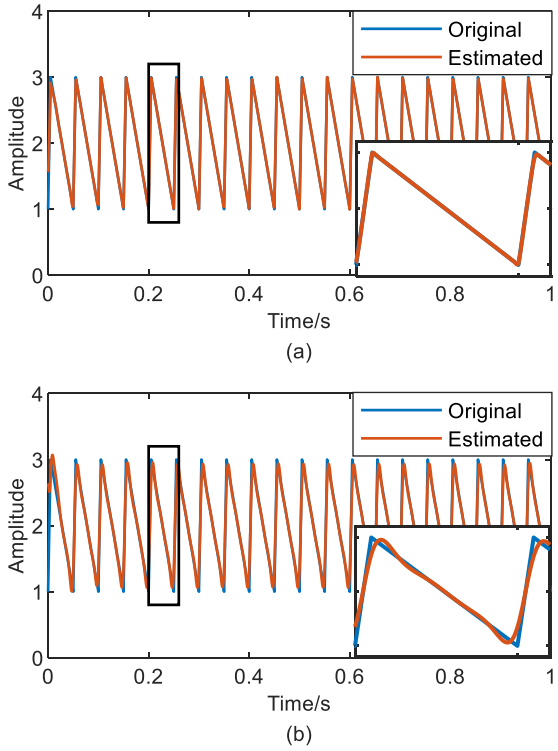

Fig. 2. The estimated IA by differenced methods: (a) NCME and (b) VNCMD.

phenomenon when analyzing the IA of signals with underlying dynamics.

It is worth noting that the convergence property is highly related to the initial IF as well as to the noise level. We proceed to further examine this aspect, using a signal with the initial IF set to $IF_{ini}(t) = f(t) + \Delta f$, where $7\Delta f$ belongs to $\{100, 200, 300, 400, 500\}$. Fig. 3 shows the resulting mode reconstructed accuracy with respect to the iteration number. The faster growth and bigger value of EQF means a better convergence property. The results show that the proposed method possesses a faster convergence speed than that of VNCMD. In general, the NCME requires fewer iterations to converge as compared to the VNCMD estimate, as well as yielding a better estimation accuracy.

In addition, the robustness of the algorithm is tested under different noise levels. In this case, the initial IF is set as $IF_{ini}(t) = f(t) + \Delta f$, where the offset error Δf is 50. A white

Gaussian noise is added to the signal $s(t)$ with the signal to noise ratio (SNR) being defined as -5 dB to 5 dB. At each noise level, 50 Monte Carlo simulations are carried out with the average of these EQF results being shown in Fig. 4. Again, the NCME obtains better performance than the VNCMD in the estimation of the IF. On average, the proposed algorithm achieves roughly a 10 dB improvement in EQF as compared to the VNCMD at every noise level.

3.2. Speech analysis

Both IF and IA estimation are of crucial importance in speech and audio analysis. In this subsection, the proposed method is applied to analyze a speech signal from the TIMIT database of the spoken sentence “Why were you away a year, Roy”. The voice recording is made of a female speaker and recorded at a sampling frequency of 8 kHz. The waveform of the signal and its spectrum are shown in Fig. 5. As can be seen, the signal exhibits a clear nonstationary behavior. Time-frequency (TF) representations may be used to visualize the information content of the nonstationary

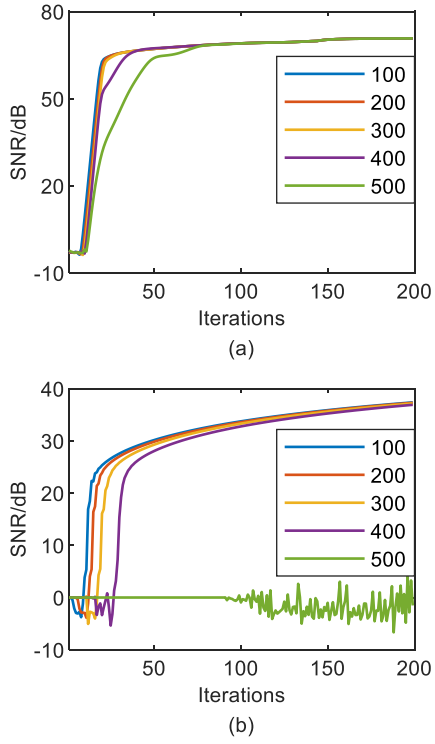

Fig. 3. The mode convergence curves with different initial IFs in the noise-free case for (a) the NCME and (b) the VNCMD.

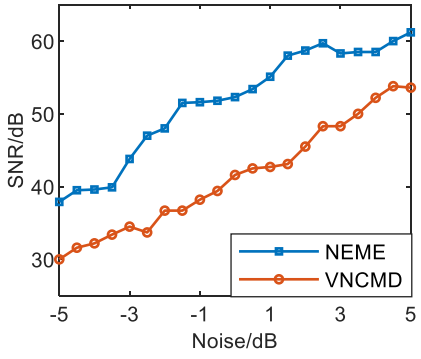

Fig. 4. The EQF of the estimated IF under different noise levels.

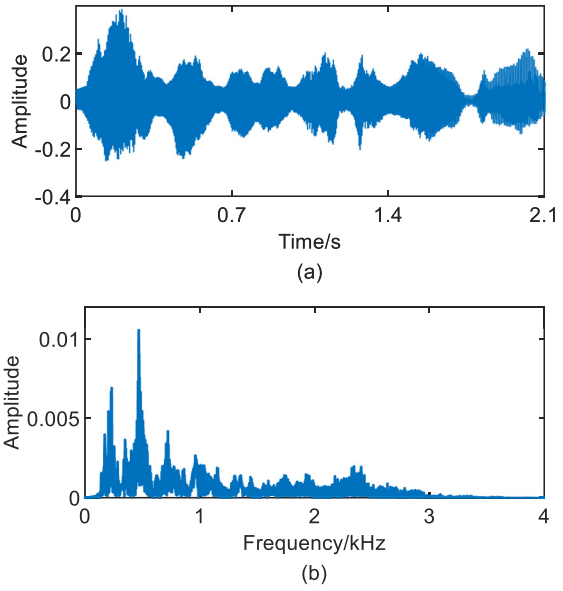

Fig. 5. The speech signal "Why were you away a year, Roy?": (a) time-domain waveform and (b) its spectrum.

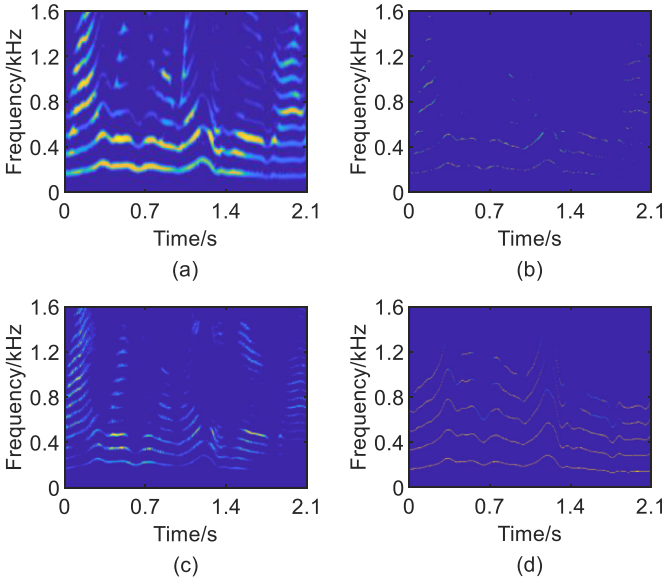

Fig. 6. TFR results generated by (a) STFT, (b) SST, (c) S-method, and (d) NCME.

signal. Here, we examine the estimates of the short-time Fourier transform (STFT), the SST, the S-method [25–27], and the proposed NCME, yielding the TF representations shown in Fig. 6. As may be seen, the STFT, SST, and S-method yield limited frequency information, where only the first and the second harmonics can be detected, thereby failing to accurately represent the weak frequency structure such as the third and fifth overtones.

To further investigate this, the estimated IFs, IAs, and reconstruct modes are presented in Figs. 7 and 8. As shown in Fig. 7 (a), the NCME yields a continuous fundamental frequency estimate. Different from the state-of-the-art methods, the NCME algorithm estimate is formed using the entire signal, without using a segmentation approach. The estimated IAs of the first five harmonics are shown in Fig. 7 (b). As can be seen, large amplitudes appear in harmonics 2, 3, and 5, around 0.4 s, 1.51 s, and 1.88 s, respectively. Since this method can estimate the IA and the IF, the reconstructed mode may be reconstructed using these two parameters as shown in Fig. 8. In summary, the NCME provides an effective way to estimate the IF, the IA, and the mode of a signal even for quite weak components. Note that the initial IF is a crucial issue for the NCME method, even though the proposed method will update the basis based on the initial IF by every iteration. If the initial IF matches well with the real IF of the signal, the number of iterations can be reduced greatly. The initial IF can be obtained from ridge detection algorithms based on TFR [28,29], which have been proved to be an effective tool.

3.3. Rub-impact detection

Condition monitoring is regarded as a potential tool to ensure the operation safety of mechanical systems. Here, the proposed method is tested on a vibration signal and used to identify the rub-impact in a rotor system.

The experimental setup is shown in Fig. 9 [30,31]. A rub rod is utilized to hit the radial surface of the disk to simulate the rub-fault. The eddy current transducers, mounted on the probe base, are applied to acquire the shaft displacement signal using a 2 kHz sampling frequency. The operation speed of the shaft is 85 Hz.

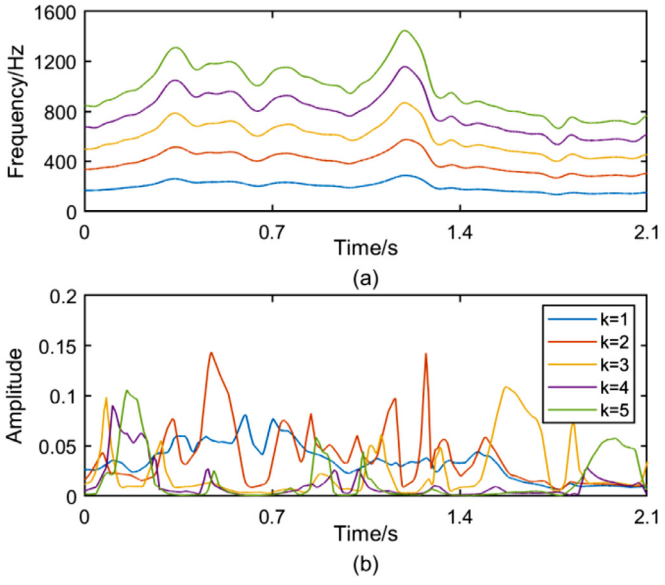

Fig. 7. (a) The estimation of the IFs and (b) the corresponding IAs by NCME.

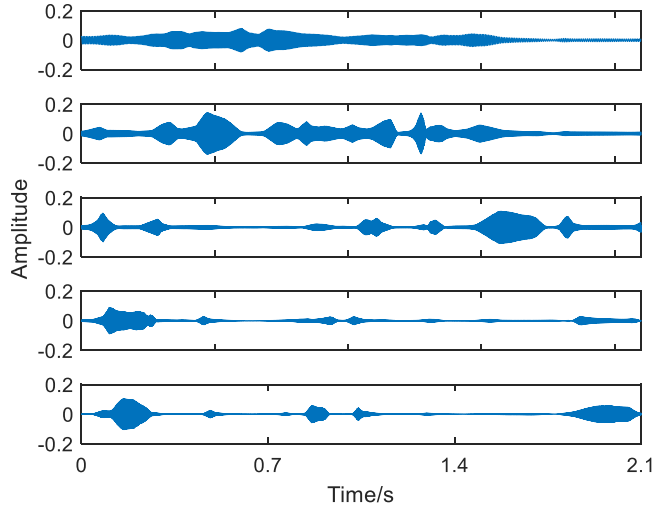

Fig. 8. The estimated modes of the first five components.

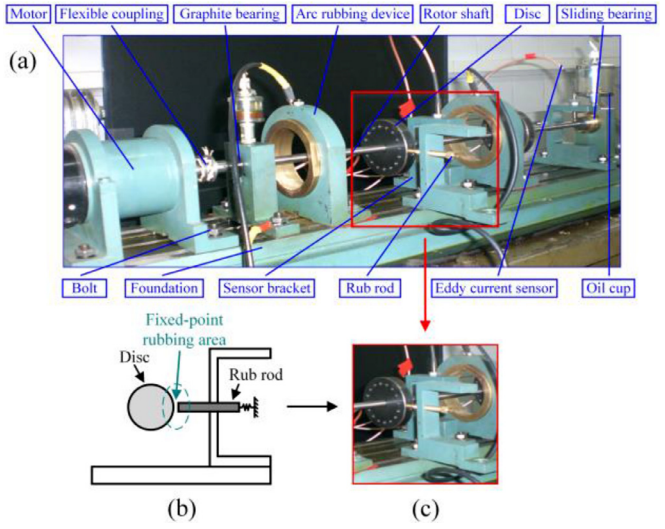

Fig. 9. (a) The test rig of the rotor system, (b) the diagrammatic sketch, and (c) a zoom-in of the rubbing structure.

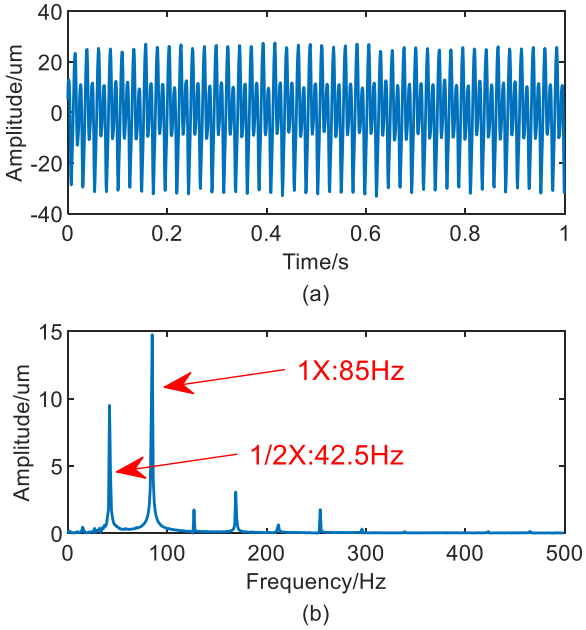

Fig. 10. The vibration signal collected from the test rig. (a) The time waveform and (b) its Fourier spectrum.

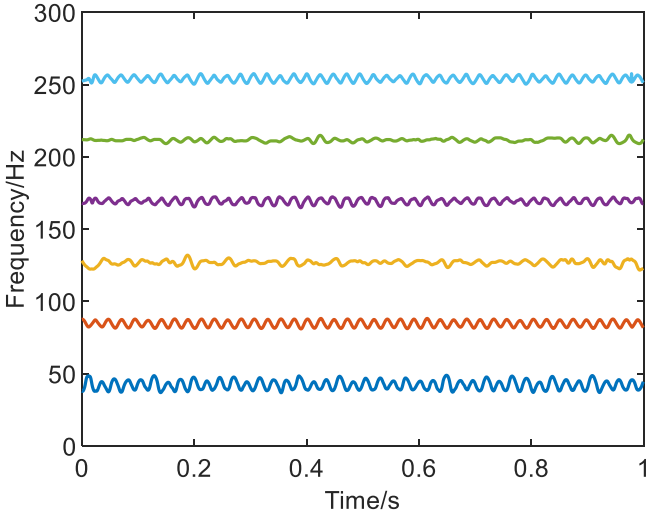

Fig. 11. The estimated IFs.

The raw vibration and its corresponding spectrum are shown in Fig. 10. It can be found that only the operation frequency 85 Hz (1X) and its fractional harmonic 42.5 Hz ($\frac{1}{2}X$) are clearly seen in the spectrum. However, no other obvious evidence indicates the existence of the rub-impact. The NCME is implemented to process the vibration signal and the estimated IFs are presented Fig. 11. The rich fault features can be found in both of the estimated parameters, clearly revealing the oscillation of both the IFs and the IAs. The estimated IFs and the Fourier spectrums of the fundamental frequency (1X) and harmonic (2X) are shown in Fig. 12. It can be seen that the 1X component shows a high degree of fluctuation, and that the oscillation frequency is approximately equal to half the fundamental frequency, demonstrating that this rotor contains a rub-impact fault and that a strong impact occurs in every two rotation periods. Since this method is capable of estimating the time-varying feature of the signal, NCME is an effective tool to identify the fault in the rotor system.

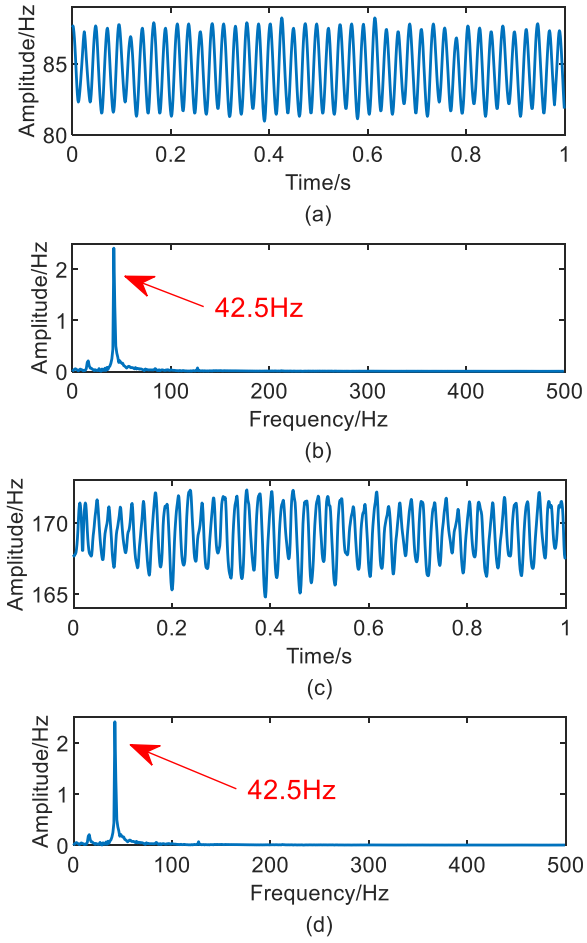

Fig. 12. Estimated IFs and the corresponding spectrums. For (a, b) the fundamental and (c, d) the first overtone.

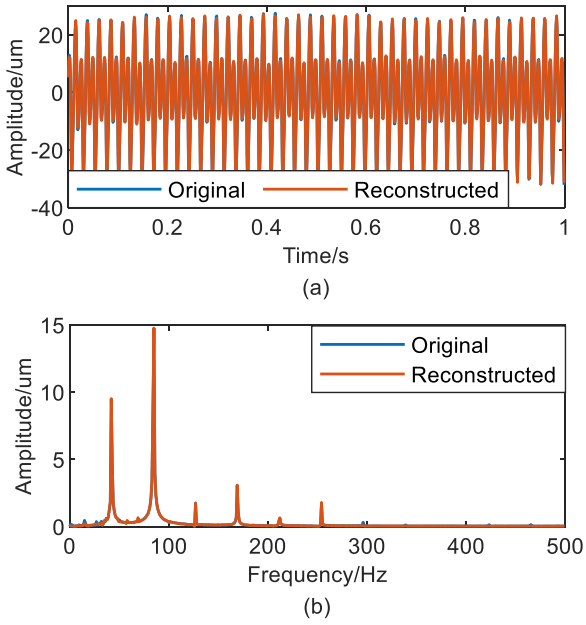

Fig. 13. Reconstructed signal and the corresponding spectrums. (a) waveforms of the reconstructed signal and the original signal, (b) spectrums of reconstructed signal and the original signal.

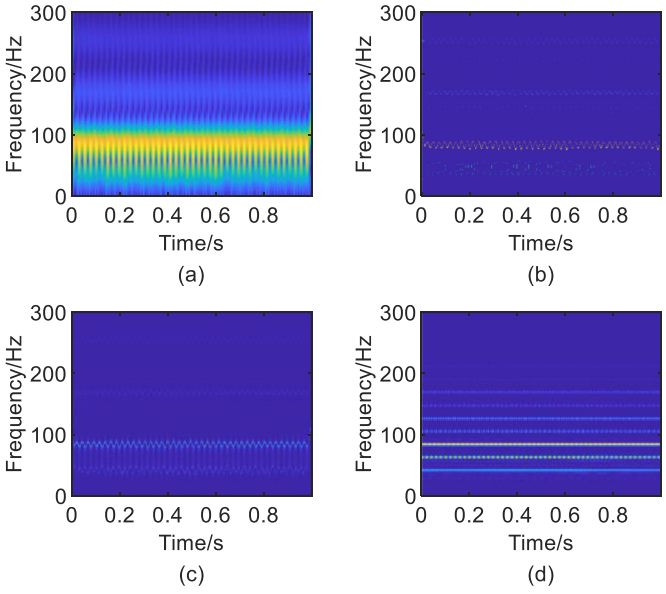

Fig. 14. TFR results generated by (a) STFT, (b) SET, (c) SST and (d) S-method.

We proceed to examine the reconstruction properties of the NCME by estimating the first six modes and calculating their spectrums, as shown in Fig. 13. The results demonstrate that the reconstructed signal is consistent with the original signal, as well as the spectrums.

For the sake of comparison, the STFT, the synchroextracting transform (SET) [32], the SST, and the S-method are also utilized to analyze the vibration signal, with the corresponding TFRs shown in Fig. 14. Due to the uncertainty principle, the STFT cannot simultaneously achieve a reasonable resolution in both time and frequency, as shown in Fig. 14 (a). Thus, the STFT generates a blurred TFR, making it difficult to observe the time-varying feature of the signal. The SET, imposing further assumptions, seems to yield an improved resolution, being able to accurately estimate the time-varying IF of the harmonics 1X and 2X, but it fails to deal with the weak components such as $\frac{1}{2}X$, and $\frac{5}{2}X$, as shown in Fig. 14 (b). The SST is capable of reassigning the energy to the new position and it should meet the requirement of a signal only containing the weak frequency modulation component. Hence, the SST produces a TFR with energy diffusion as shown in Fig. 14 (c). This aliasing representation will increase the challenge to extract the IF from the TF plane. The S-method can produce a sum of the Wigner distribution of individual signal components, avoiding cross-terms. Though S-method can improve the time-frequency resolution of the TFR, it cannot reveal the structure of this kind of signal, as shown in Fig. 14 (d).

4. Conclusion

In this paper, an effective method to decompose a wide-band signal into several nonlinear chirp modes is proposed. Using a greedy matching pursuit style algorithm, we form a recursive algorithm to extract the modes one by one. The resulting estimator can update the basis at each iteration by using the arctangent demodulation technique, being derived from the signal itself. Furthermore, we introduce an efficient implementation using the ADMM, making the algorithm practically useful in applications such as speech analysis and fault diagnosis. Although the NCME is able to separate closely-spaced modes, it fails to accurately decompose the signal with severely crossed modes. In the future, a joint optimization scheme may be considered to estimate all of the modes simultaneously to allows for such case.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Xiaotong Tu: Conceptualization, Methodology, Software, Writing - original draft. **Johan Swärd:** Methodology, Software, Writing - review & editing. **Andreas Jakobsson:** Methodology, Validation, Writing - review & editing. **Fucai Li:** Funding acquisition, Formal analysis, Data curation, Validation.

Acknowledgements

This work was supported in part by the National Science and Technology Major Project under Grant 2018ZX04011001, in part by the Open Project of Jiangsu Key Laboratory of Advanced Food Manufacturing Equipment and Technology under Grant FM-201901.

Appendix A

The proof of the (8) is given in the following. The signal model in (1) can be rewritten as

$$y(t) = \sum_{k=1}^{K} a_k(t) \cos\left(2\pi \int_0^t \left(\tilde{f}_k(\tau) + \Delta\tilde{f}_k(\tau)\right)d\tau + \phi_k\right) \tag{23}$$

where $\tilde{f}_k(\tau)$ is the estimated IF of the k^{th} mode, and $\Delta\tilde{f}_k(\tau)$ denotes the error between the real IF and estimated IF, i.e., $f_k(\tau) = \tilde{f}_k(\tau) + \Delta\tilde{f}_k(\tau)$.

One may rewrite (23) as

$$\begin{aligned} y(t) = \sum_{k=1}^{K} \Bigg\{ & u_k(t) \cos\left(2\pi \int_0^t \tilde{f}_k(\tau)d\tau\right) \\ & + v_k(t) \sin\left(2\pi \int_0^t \tilde{f}_k(\tau)d\tau\right)\Bigg\} \end{aligned} \tag{24}$$

with

$$\begin{cases} u_k(t) = a_k(t)\cos\left(2\pi \int_0^t \Delta\tilde{f}_k(\tau)d\tau + \phi_k\right) \\ v_k(t) = -a_k(t)\sin\left(2\pi \int_0^t \Delta\tilde{f}_k(\tau)d\tau + \phi_k\right) \end{cases} \tag{25}$$

Thus, $-tan((2\pi \int_0^t \Delta\tilde{f}_k(\tau)d\tau + \phi_k))$ can be obtained via $v_k(t)/u_k(t)$. Finally, $\Delta\tilde{f}_k(\tau)$ can be formed as

$$\Delta\tilde{f}_k(\tau) = -\frac{1}{2\pi}\frac{d}{dt}\left(\arctan\left(\frac{v_k^{i+1}(t)}{v_k^{i+1}(t)}\right)\right) \tag{26}$$

Completing the proof.

References

[1] X. Tu, Y. Hu, F. Li, S. Abbas, Z. Liu, W. Bao, Demodulated high-order synchrosqueezing transform with application to machine fault diagnosis, IEEE Trans. Ind. Electron. 66 (4) (2019) 3071–3081.
[2] F. Auger, et al., Time-frequency reassignment and synchrosqueezing: an overview, IEEE Signal Process. Mag. 30 (6) (2013) 32–41.
[3] D. Giacobello, M.G. Christensen, M.N. Murthi, S.H. Jensen, M. Moonen, Sparse linear prediction and its applications to speech processing, IEEE Trans. Audio. Speech. Lang. Processing 20 (5) (2012) 1644–1657.
[4] M.G. Christensen, P. Stoica, A. Jakobsson, S. Holdt Jensen, Multi-pitch estimation, Signal Process. 88 (4) (2008) 972–983.
[5] Z. He, X. Tu, W. Bao, Y. Hu, F. Li, Gaussian-modulated linear group delay model: application to second-order time-reassigned synchrosqueezing transform, Signal Process. 167 (107275) (2020) 1–7.
[6] D. Pham, S. Meignen, High-order synchrosqueezing transform for multicomponent signals analysis - with an application to gravitational-wave signal, IEEE Trans. Signal Process. 65 (12) (2017) 3168–3178.
[7] S. Chen, X. Dong, Z. Peng, W. Zhang, G. Meng, Nonlinear chirp mode decomposition: a variational method, IEEE Trans. Signal Process. 65 (22) (2017) 6024–6037.
[8] N.E. Huang, et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc. London. Ser. A Math. Phys. Eng. Sci. 454 (1971) (1998) 903–995.
[9] G. Oladosu, Identifying the oil price–macroeconomy relationship: an empirical mode decomposition analysis of US data, Energy Policy 37 (12) (2009) 5417–5426.
[10] Q. Gao, C. Duan, H. Fan, Q. Meng, Rotating machine fault diagnosis using empirical mode decomposition, Mech. Syst. Signal Process. 22 (5) (2008) 1072–1081.
[11] T.A. McMahon, A.S. Kiem, M.C. Peel, P.W. Jordan, G.G.S. Pegram, A new approach to stochastically generating six-monthly rainfall sequences based on empirical mode decomposition, J. Hydrometeorol. 9 (6) (Dec. 2008) 1377–1389.
[12] I. Daubechies, J. Lu, H.-T. Wu, Synchrosqueezed wavelet transforms: a tool for empirical mode decomposition, Appl. Comput. Harmon. Anal. 30 (2) (2009) 243–261.
[13] S. Meignen, D.-H. Pham, M.A. Colominas, On the use of short-time fourier transform and synchrosqueezing-based demodulation for the retrieval of the modes of multicomponent signals, Signal Process. 178 (2021) 107760.
[14] X. Tu, Q. Zhang, Z. He, Y. Hu, S. Abbas, F. Li, Generalized horizontal synchrosqueezing transform: algorithm and applications, IEEE Trans. Ind. Electron. (2020) 1.
[15] J. Gilles, Empirical wavelet transform, IEEE Trans. Signal Process. 61 (16) (2013) 3999–4010.
[16] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, IEEE Trans. Signal Process. 62 (3) (2014) 531–544.
[17] S.I. Adalbjörnsson, J. Swärd, A. Jakobsson, Enhancing smoothness in amplitude modulated sparse signals, in: 11th IMA International Conference on Mathematics in Signal Processing, 2016, pp. 1–4.
[18] J. Swärd, et al., Sparse semi-parametric estimation of harmonic chirp signals, IEEE Trans. Signal Process. 64 (7) (2016) 1798–1807.
[19] X. Meng, A. Jakobsson, X. Li, Y. Lei, Estimation of chirp signals with time-varying amplitudes, Signal Process. 147 (2018) 1–10.
[20] S. Chen, Z. Peng, Y. Yang, X. Dong, W. Zhang, Intrinsic chirp component decomposition by using Fourier Series representation, Signal Process. 137 (2017) 319–327.
[21] Q. Lv, et al., High dynamic-range motion imaging based on linearized doppler radar sensor, IEEE Trans. Microw. Theory Tech. 62 (9) (2014) 1837–1846.
[22] S.D. Babacan, R. Molina, A.K. Katsaggelos, Bayesian compressive sensing using Laplace priors, IEEE Trans. Image Process. 19 (1) (2010) 53–63.
[23] Z. Bai, L. Shi, J. Sun, and M.G. Christensen, "Complex sparse signal recovery with adaptive Laplace priors," arXiv Prepr., vol. 2006, no. 16720, pp. 1–13, 2020.
[24] S. Boyd, N. Parikh, E. Chu, Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers, Now Publishers Inc, 2011.
[25] L. Stankovic, A method for time-frequency analysis, IEEE Trans. Signal Process. 42 (1) (1994) 225–229.
[26] F. Li, C. Yang, Y. Xia, X. Ma, T. Zhang, Z. Zhou, An adaptive S-method to analyze Micro-Doppler signals for human activity classification, Sensors 17 (12) (2017) 2769.
[27] Lj. Stanković, T. Thayaparan, M. Daković, Signal decomposition by using the S-method with application to the analysis of HF radar signals in sea-clutter, IEEE Trans. Signal Process. 54 (11) (2006) 4332–4342.
[28] I. Djurović, Lj. Stanković, An algorithm for the Wigner distribution based instantaneous frequency estimation in a high noise environment, Signal Process. 84 (3) (2004) 631–643.
[29] X. Tu, W. Bao, Y. Hu, S. Abbas, F. Li, Parameterized synchrosqueezing transform with application to machine fault diagnosis, IEEE Sens. J. 19 (18) (2019) 8107–8115.
[30] K. Yu, T.R. Lin, H. Ma, H. Li, J. Zeng, A combined polynomial chirplet transform and synchroextracting technique for analyzing nonstationary signals of rotating machinery, IEEE Trans. Instrum. Meas. (2019) 1.
[31] K. Yu, H. Ma, H. Han, J. Zeng, H. Li, X. Li, Second order multi-synchrosqueezing transform for rub-impact detection of rotor systems, Mech. Mach. Theory 140 (2019) 321–349.
[32] G. Yu, M. Yu, C. Xu, Synchroextracting transform, IEEE Trans. Ind. Electron. 64 (10) (2017) 8042–8054.