

# An Empirical Signal Separation Algorithm for Multicomponent Signals Based on Linear Time-Frequency Analysis

Lin Li <sup>1,2,\*</sup>, Haiyan Cai <sup>2</sup>, Qingtang Jiang <sup>2</sup>, Hongbing Ji <sup>1</sup>

<sup>1</sup> School of Electronic Engineering, Xidian University, Xi'an 710071, China

<sup>2</sup> Department of Mathematics and Computer Science, University of Missouri-St. Louis, St. Louis 63121, United States

{lilin, hbji}@xidian.edu.cn, {haiyan\_cai, jiangq}@umsl.edu

**Abstract:** The empirical mode decomposition (EMD) is a powerful tool for non-stationary signal analysis. It has been used successfully for non-stationary signals separation and time-frequency representation. Linear time-frequency analysis (TFA) is another powerful tool for non-stationary signal. Linear TFAs, e.g. short-time Fourier transform (STFT) and wavelet transform (WT), depend linearly upon the signal analysis. In the current paper, we utilize the advantages of EMD and linear TFA to propose a new signal reconstruction method, called the empirical signal separation algorithm. First we represent the signal with STFT or WT. After that, by using an EMD-like procedure, we extract the components in the time-frequency (TF) plane one by one, adaptively and automatically. With the iterations carried out in the sifting process, the proposed method can separate non-stationary multicomponent signals with fast varying frequency components which EMD may not be able to separate. The experiments results demonstrate the efficiency of the proposed method compared to standard EMD, ensemble EMD and synchrosqueezing transform.

**Key words:** time-frequency analysis; empirical mode decomposition; wavelet transform; synchrosqueezing transform; signal separation.

## 1. Introduction

Separation of multicomponent non-stationary signals has broad applications in many engineering fields, such as seismic signal analysis, vibration fault diagnosis, biomedical signal analysis, speech enhancement, sonar signal processing, etc. The EMD algorithm along with the Hilbert spectrum analysis (HSA), introduced by Huang et al. in [1], is an efficient method to decompose and analyze non-stationary signals. The EMD scheme, considered as an adaptive signal analysis, separates a given signal into a number of components, called intrinsic mode functions (IMFs), and then calculates the instantaneous frequency (IF) of each IMF by the HSA.

In recent years, efforts to improve the performance of EMD focus on envelope fitting, including interpolation

---

\*Corresponding author  
Email: lilin@xidian.edu.cn

algorithm and boundary expansion. In [1], the cubic spline function is used to generate the upper envelope and lower envelope by interpolating the local maxima and local minima respectively. An alternative B-spline algorithm for the EMD is introduced in [2]. The blending cubic spline interpolation operator and corresponding error bounds are also derived in [3]. The direct construction algorithm of envelope mean is proposed based on the constrained optimization for narrow-band signals [4]. In addition, [5] proposes an alternative EMD using the filter to calculate the envelope mean. In the iteration process, the boundary expansion error will gradually effect the accuracy of IMFs. Usually the symmetry rule is used to predict the extrema beyond the two boundaries. A slope-based method is introduced to restrain the boundary effect in [6].

The capability of EMD to decompose two sub-signals is discussed in [7], and the influence of sampling on the EMD is discussed in [8]. Moreover, in [9], the EMD is considered as a filter bank, whose sub-band number, band-width and cutoff frequency are adaptive and data-driven. An inverse filter scheme of EMD is proposed in [10], which aims to reduce the scale mixing. The ensemble EMD (EEMD) proposed in [11], is a noise-assisted signal analysis method. EEMD utilizes the advantage of the statistical characteristics of white noise to force the ensemble to exhaust all possible solutions in the sifting process.

On the other hand, time-frequency analysis has been studied for many years, including linear TFA and non-linear TFA (see the overviews in [12], [13], [14], [15] and [16]). Compared with EMD, TFA is of solid theoretical basis. The research tendency of TFA is to enhance the TF resolution and energy concentration [15]. The uncertainty principle [16] makes a tradeoff between temporal and spectral resolutions unavoidable. Many non-linear TFA methods are proposed to overcome this important shortcoming, such as kernel-based time-frequency distributions [15], the time-frequency reassignment [17], etc. However, these non-linear TFA are irreversible, hence unable to reconstruct multicomponent signals. On the contrary, linear TFA are reversible and are able to recover the signal components. The synchrosqueezing transform (SST) based on continuous wavelet transform (CWT), a typical linear TFA, was introduced in [18] and further developed in [19]. The sub-signals are reconstructed by extracting the IF curves in the SST plane one by one. SST works well for single-tone signals, but not for broadband time-varying frequency signals. The generalized SST and instantaneous frequency-embedded SST (IFE-SST) are proposed in [20] and [21] respectively, with both of them changing broadband signals to narrow-band signals and IFE-SST preserving the IFs of the original signals. A multitapered SST is introduced in [22] to enhance the concentration in the TF plane by averaging over random projections with synchrosqueezing. A hybrid EMD-SST scheme is introduced in [3] by applying the modified SST to the IMFs obtained by the EMD. STFT-based SST is discussed in [23] and [24], and STFT-based second order SST is

proposed in [25]. The matching demodulation transform and its synchrosqueezing are discussed in [26]. The linear and synchrosqueezed time-frequency representations are reviewed in [27] in detail.

In this paper, we utilize the advantages of EMD and linear TFA to propose a new signal separation method, called the empirical signal separation algorithm (ESS). In the sifting procedure of EMD, only features in time domain, e.g. maximum and minimum extrema are used to decompose the given signal. So EMD is sensitive to noises and usually results in mode mixing and artifacts. Based on linear TFA (STFT or WT), we use an EMD-like algorithm to extract the signal components in the TF plane one by one. Our algorithm is different from the empirical wavelet transform in [28], which extracts the components by designing an appropriate wavelet filter bank. Compared with the signal reconstruction by SST, our algorithm is adaptive and can be used for multicomponent signals dissatisfying the well-separated conditions in [19]. The remainder of this paper is organized as follows. The linear TFA and SST are reviewed in Section 2. The ESS is introduced in Section 3. We also discuss some properties of ESS in Section 3. The numerical results and comparisons are provided in Section 4.

## 2. Linear Time-Frequency Analysis and Synchrosqueezing Transform

A multicomponent amplitude and frequency-modulated (AM-FM) signal is given by [16],

$$s(t) = \sum_{j=1}^N s_j(t) = \sum_{j=1}^N A_j(t) \cos(2\pi\phi_j(t)), \quad (1)$$

where the IF  $\phi_j'(t)$  and amplitude  $A_j(t)$  of each component  $s_j(t) = A_j(t) \cos(2\pi\phi_j(t))$  are slowly varying compared to  $\phi_j(t)$ , to allow for the extraction and separation of component  $s_j(t)$  from  $s(t)$ .

### 2.1. The Linear Time-Frequency Analysis

Let us start with the short-time Fourier transform (STFT). For a real-valued window function  $u \in L^2(\mathbb{R})$  with unit  $L^2$ -norm, and a signal  $s \in L^2(\mathbb{R})$ , the STFT of  $s(t)$  is defined by,

$$V_s(t, \omega) = \int_{\mathbb{R}} s(\tau) u(\tau - t) e^{-i\omega\tau} d\tau, \quad (2)$$

for  $t, \omega \in \mathbb{R}$ , where  $t$  and  $\omega$  denote time and angular frequency, respectively. Note that  $\omega = 2\pi\xi$ , where  $\xi$  is the frequency.

When the Gaussian function

$$u_\sigma(t) = \frac{1}{2\sigma\sqrt{\pi}} e^{-\left(\frac{t}{2\sigma}\right)^2}, \quad (3)$$

with  $\sigma = 1/\sqrt{4\pi}$  is used as the window function, the STFT is called the Gabor transform [24].

The signal  $s(t)$  can be reconstructed by the inverse STFT, using the following formula,

$$s(\tau) = \iint_{\mathbb{R}^2} V_s(t, \omega) u(\tau - t) e^{i\omega\tau} dt d\omega. \quad (4)$$

And  $s(t)$  can also be reconstructed by

$$s(t) = \frac{1}{u(0)} \int_{\mathbb{R}} V_s(t, \omega) d\omega. \quad (5)$$

Note that  $u(t)$  is required to be non-zero and continuous at 0 (see [23]).

Now we discuss the TF resolution of STFT. The time center of the window function  $u(t)$  is defined by

$$t_u^* = \frac{\int_{-\infty}^{\infty} t |u(t)|^2 dt}{\int_{-\infty}^{\infty} |u(t)|^2 dt}, \quad (6)$$

and the equivalent duration of  $u(t)$  is defined by

$$\Delta_u = \left[ \frac{\int_{-\infty}^{\infty} (t - t_u^*)^2 |u(t)|^2 dt}{\int_{-\infty}^{\infty} |u(t)|^2 dt} \right]^{\frac{1}{2}}. \quad (7)$$

The Fourier transform of  $u(t)$  is denoted by  $\hat{u}(\omega)$ . Then the frequency center and equivalent bandwidth of  $u(t)$  are denoted as  $\omega_{\hat{u}}^*$  and  $\Delta_{\hat{u}}$ , obtained by Eq. (6) and Eq. (7) with  $u$  and  $t$  replaced by  $\hat{u}$  and  $\omega$ , respectively.

From the Heisenberg uncertainty principle, we have  $\Delta_u \Delta_{\hat{u}} \geq \frac{1}{2}$ , and the equality is obtained for Gaussian function only (see details in [16]).  $\Delta_u$  and  $\Delta_{\hat{u}}$  are also called temporal and spectral resolutions, respectively. The uncertainty principle imposes an unavoidable tradeoff between temporal and spectral resolutions.

CWT is another method of linear TF analysis. Like STFT, CWT uses a window to locate a signal to its time and frequency content. However, the window shape of CWT changes continuously at different frequency instant.

Let  $\psi(t) \in L^2(\mathbb{R})$  be a fix function.  $\hat{\psi}(\omega)$  is the Fourier transform of  $\psi(t)$ . If  $\hat{\psi}$  satisfies the conditions,

$$\hat{\psi}(\omega) = 0 \quad \text{for } \omega < 0, \quad (8)$$

$$0 < C_{\psi} = \int_0^{\infty} |\hat{\psi}(\omega)|^2 \frac{d\omega}{\omega} < \infty, \quad (9)$$

then  $\psi(t)$  is called an analytic continuous wavelet (admissible wavelet).

Denote

$$\psi_{(a,b)}(t) = \frac{1}{a} \psi\left(\frac{t-b}{a}\right), \quad (10)$$

where  $a > 0$  is the scale, and  $b \in \mathbb{R}$  is the shift. The Fourier transform of  $\psi_{(a,b)}(t)$  is written as

$$\hat{\psi}_{(a,b)}(\omega) = \hat{\psi}(a\omega)e^{-i\omega b}. \quad (11)$$

The CWT of a signal  $s(t) \in L^2(\mathbb{R})$  is defined by

$$W_s(a,b) = \left\langle s, \psi_{(a,b)} \right\rangle = \int_{-\infty}^{\infty} s(t) \overline{\psi\left(\frac{t-b}{a}\right)} a^{-1} dt. \quad (12)$$

An analytic signal  $s(t)$  can be recovered from  $W_s(a,b)$  by (see e.g. [31], [32] and [33])

$$s(t) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty W_s(a,b) a^{-1} \psi_{(a,b)}(t) db da. \quad (13)$$

For an analytic CW  $\psi(t)$ , if  $C_\psi' = \int_0^\infty \frac{\overline{\hat{\psi}(\eta)}}{\eta} d\eta < +\infty$ , another reconstruction formula can be alternatively used

(see [18] and [19]),

$$s(b) = \frac{1}{C_\psi'} \int_0^\infty W_s(a,b) a^{-1} da. \quad (14)$$

For real signals, the CWT with integral involving  $b$  is invertible too. Let  $s \in L^2(\mathbb{R})$  be a real-valued signal.

Consider the analytic signal with associated  $s$ , which is denoted by

$$s_a(t) = s(t) + i(\mathbb{H}(s))(t), \quad (15)$$

where  $\mathbb{H}(s)$  is the Hilbert transform of  $s$ . The following reconstruction formulae exist in  $L^2(\mathbb{R})$  [19],

$$s(t) = \Re \left\{ \frac{2}{C_\psi} \int_0^\infty \int_{-\infty}^\infty W_s(a,b) a^{-1} \psi_{(a,b)}(t) db da \right\}, \quad (16)$$

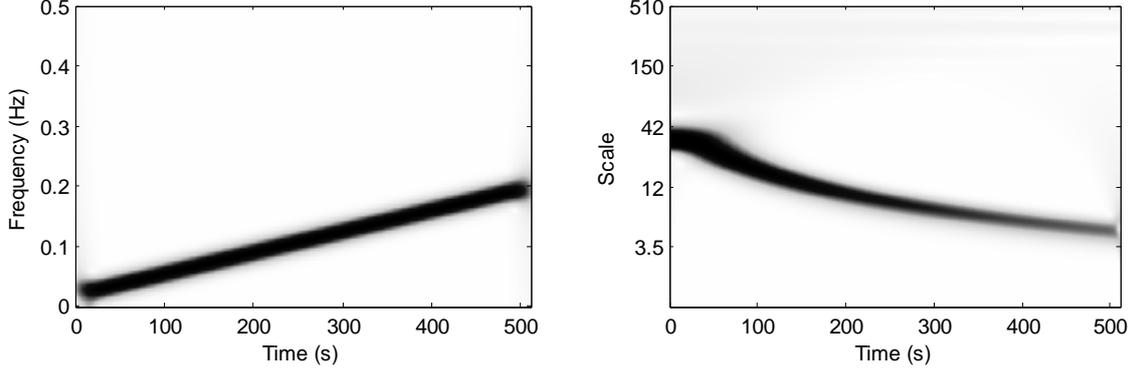
$$s(b) = \Re \left\{ \frac{2}{C_\psi'} \int_0^\infty W_s(a,b) a^{-1} da \right\}. \quad (17)$$

With the definitions in Eq. (6) and Eq. (7), we can also calculate the temporal and spectral resolutions of wavelet

$\psi_{(a,b)}$ , which are given by

$$\Delta_{\psi_{(a,b)}} = \left[ \frac{\int_{-\infty}^{\infty} (t - t_{\psi_{(a,b)}}^*)^2 \left| \frac{1}{a} \psi\left(\frac{t-b}{a}\right) \right|^2 dt}{\int_{-\infty}^{\infty} \left| \frac{1}{a} \psi\left(\frac{t-b}{a}\right) \right|^2 dt} \right]^{\frac{1}{2}} = a \Delta_\psi, \quad (18)$$

$$\Delta_{\hat{\psi}_{(a,b)}} = \left[ \frac{\int_{-\infty}^{\infty} (\omega - \omega_{\hat{\psi}_{(a,b)}}^*)^2 \left| \hat{\psi}(a\omega) e^{-i\omega b} \right|^2 d\omega}{\int_{-\infty}^{\infty} \left| \hat{\psi}(a\omega) e^{-i\omega b} \right|^2 d\omega} \right]^{\frac{1}{2}} = \frac{1}{a} \Delta_{\hat{\psi}}, \quad (19)$$



**Figure 1.** The STFT (left) and CWT (right) of a LFM signal

where  $\Delta_\psi$  and  $\Delta_{\hat{\psi}}$  are the temporal and spectral resolutions of  $\psi$ , respectively,  $t_{\psi(a,b)}^*$  and  $\omega_{\hat{\psi}(a,b)}^*$  are the centers of  $\psi_{(a,b)}$  and  $\hat{\psi}_{(a,b)}$ , respectively. Therefore, we obtain

$$\Delta_{\psi_{(a,b)}} \Delta_{\hat{\psi}_{(a,b)}} = \Delta_\psi \Delta_{\hat{\psi}} \geq \frac{1}{2}. \quad (20)$$

From Eq. (18), Eq. (19) and Eq. (20), we know the localization window of  $\psi_{(a,b)}$  is not fixed when CWT is used. For larger  $a$ , viz. the low frequency band, it provides a narrower frequency window and hence a higher frequency resolution, but a lower time resolution. On the other hand, the opposite is true. Figure1 shows the difference between STFT and CWT. The simulation signal is a linear frequency modulation (LFM) signal, where the frequency changes linearly from 0.02Hz to 0.2Hz.

## 2.2. The Synchrosqueezing Transform

Based on the linear TFA above, the SST is a TF reassignment approach, which aims to sharpen the TF representation while keeping the temporal localization. SST has been proved to be well adapted to multicomponent signals in [19]. However, it is difficult to separate an unknown multicomponent signal to specific components without any prior knowledge and/or appropriate restrictions. Eq. (1) is called the AHM [3], with the restrictions described by

$$\left\{ \begin{array}{ll} A_j \in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R}); & \phi_j \in C^2(\mathbb{R}) \\ \inf_{t \in \mathbb{R}} A_j(t) > c_1; & \sup_{t \in \mathbb{R}} A_j(t) < c_2 \\ \inf_{t \in \mathbb{R}} \phi_j'(t) > c_1; & \sup_{t \in \mathbb{R}} \phi_j'(t) < c_2 \\ |A_j'(t)| \leq \varepsilon \phi_j'(t); & |\phi_j''(t)| \leq \varepsilon \phi_j'(t) \end{array} \right. \quad (21)$$

for all  $t \in \mathbb{R}$ , where  $0 < \varepsilon \ll 1$  and  $0 < \varepsilon < c_1 < c_2 < \infty$ . In addition, the components are said to be well-separated if

$$\left\{ \begin{array}{l} \phi_j'(t) > \phi_{j-1}'(t), \\ \phi_j'(t) - \phi_{j-1}'(t) \geq d(\phi_j'(t) + \phi_{j-1}'(t)), \end{array} \right. \quad (22)$$

for all  $t \in \mathbb{R}$  and some  $0 < d < 1$ . A continuous function  $f(t) = A(t) \cos(2\pi\phi(t))$  with  $A(t)$  and  $\phi(t)$  satisfying conditions (21) and (22) is called an intrinsic mode type (IMT) function in [19].

The CWT-based SST (WSST) works through squeezing the CWT, where the wavelet is required to be admissible (defined in Eq. (9)). The reference IF function (local IF) based on CWT is defined by,

$$\omega_s(a, b) = \frac{-i\partial_b W_s(a, b)}{2\pi W_s(a, b)}. \quad (23)$$

Then the WSST is defined by

$$T_s(\xi, b) = \int_{\{a: |W_s(a, b)| > \gamma\}} W_s(a, b) \delta(\xi - \omega_s(a, b)) a^{-1} da, \quad (24)$$

where  $\gamma > 0$ , which is used to reduction numerical error and noise influence (see the discussions in [18] and [19]).

Similar to the WSST, the STFT-based SST (FSST) works through squeezing the STFT. The local IF is defined by [23]

$$\omega_s(t, \xi) = \Re_e \left\{ \frac{-i\partial_t V_s(t, \xi)}{2\pi V_s(t, \xi)} \right\}. \quad (25)$$

Then the FSST is defined by

$$P_s(t, \xi) = \frac{1}{u(0)} \int_{\{\nu: |V_s(t, \nu)| > \gamma\}} V_s(t, \nu) \delta(\xi - \omega_s(t, \nu)) d\nu. \quad (26)$$

The SST is used not only to sharpen the TF representation, but to separate the multicomponent signal into its constituent modes. For a multicomponent signal in (1), the  $j$ -th component can be reconstructed by WSST and FSST, respectively:

$$s_j(b) = \lim_{\alpha \rightarrow 0} \frac{1}{C_\psi'} \int_{|\xi - \phi_j'(b)| < \rho} T_s(\xi, b) d\xi, \quad (27)$$

$$s_j(t) = \lim_{\alpha \rightarrow 0} \int_{|\xi - \phi_j'(t)| < \rho} P_s(t, \xi) d\xi, \quad (28)$$

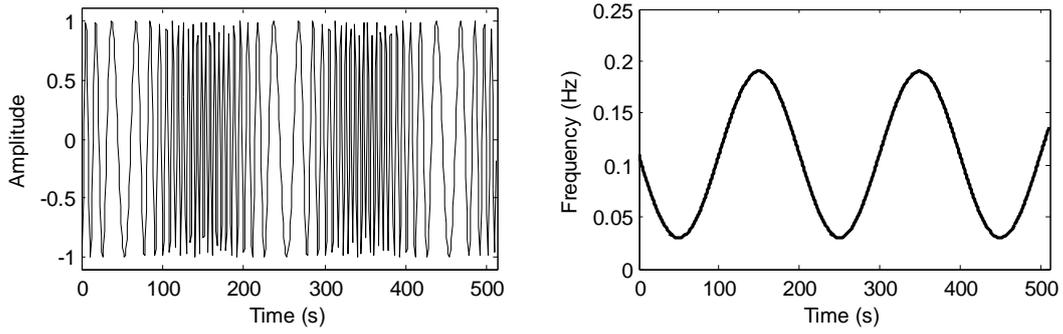
where  $\rho$  is the width of the zone to be summed up around the ridge corresponding to the IF of the  $j$ -th component.

### 3. The Empirical Signal Separation Algorithm

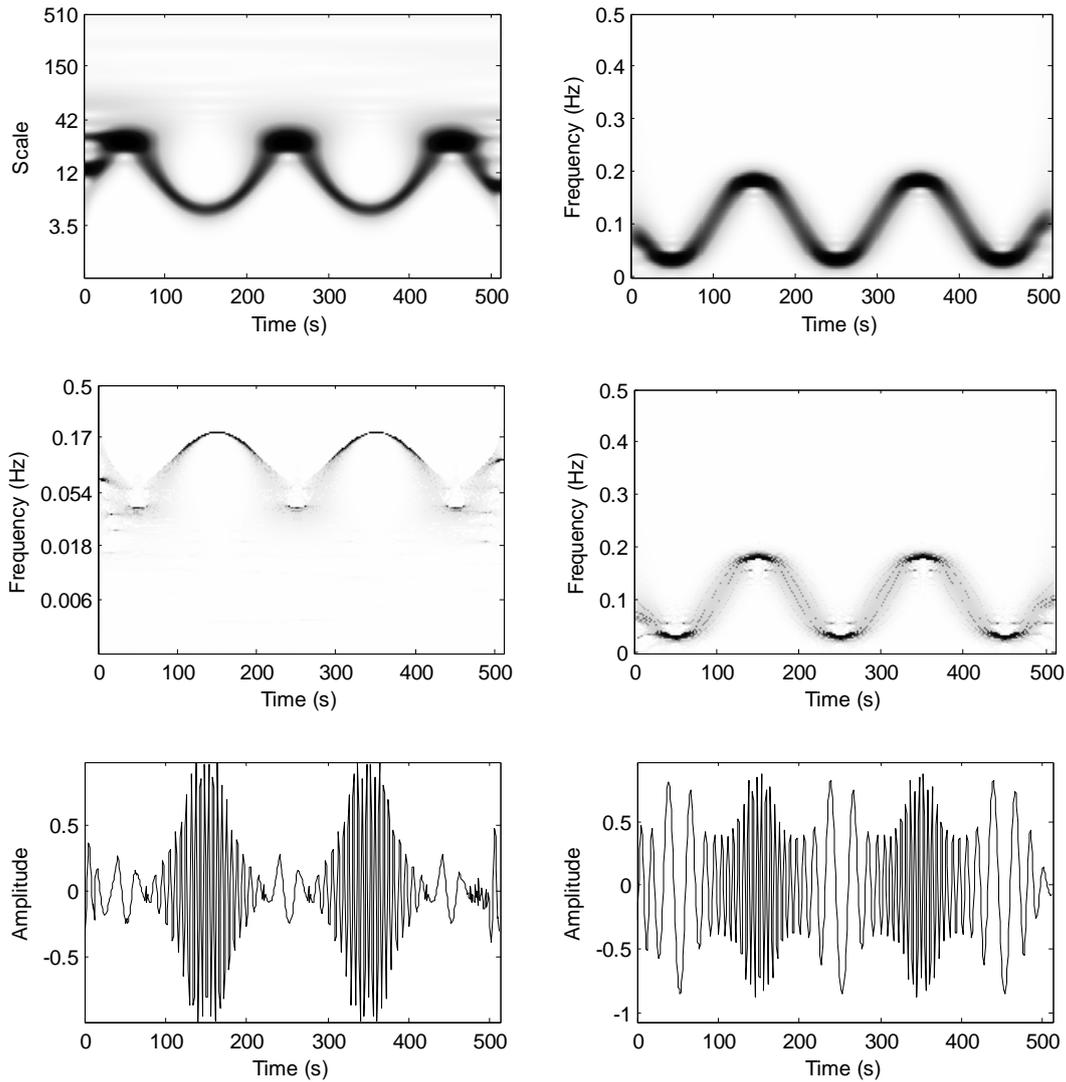
Let us start with the squeezing and recovery of a monocomponent signal by SST. Consider

$$x(t) = \cos \left( 2\pi \left( 0.11t + \frac{0.08}{0.01\pi} \cos(0.01\pi t) \right) + \theta_1 \right), \quad (29)$$

where  $\theta_1$  is the random initial phase.  $x(t)$  is a sinusoidal frequency modulation signal which can be found in radar systems. Figure 2 shows one of the waveform of signal  $x(t)$  and its IF. Note the sampling frequency is 1Hz.



**Figure 2.** Waveform of signal  $x(t)$  (left) and its instantaneous frequency (right).



**Figure 3.** CWT (top left) and STFT (top right) of  $x(t)$ , WSST (middle left) and FSST (middle right) of  $x(t)$ , the waveform reconstructed from WSST (bottom left) and the waveform reconstructed from FSST (bottom right).

Figure 3 shows the SST and the waveform reconstruction results by the Synchrosqueezing Toolbox by E. Brevdo and G. Thakur [19, 29]. The wavelet used is Morlet's wavelet with parameters  $\sigma=1$  and  $\mu=2\pi$  and the number of voices used is 64. The window function for STFT is the standard Gaussian, with parameters mean 0 and standard

deviation 0.12. The threshold  $\gamma$  in both WSST and FSST is equal to  $10^{-3}$ . The width of the zone to be summed up around the ridge for waveform reconstruction is  $\rho = 5$  (the discrete form, unitless), for both WSST and FSST. We find that the SST works well when the IF changes slowly, but not well when the IF changes fast. See WSST and FSST in Figure 3. This can also be found in the recovered waveforms (see the bottom row in Figure 3). Theoretically, for a monocomponent signal, one can increase the width  $\rho$  to improve the recovery performance. However it is not always simply to do in this way to process multicomponent signals since increasing the width  $\rho$  will result in the mixture of components.

### 3.1. Problem with Existing Methods

Consider the IMT function  $s(t) = A(t) \cos(2\pi\phi(t))$ , the CWT of  $s(t)$  is approximated by

$$W_s(a, b) \approx \frac{1}{2} A(b) e^{i2\pi\phi(b)} \overline{\hat{\psi}(2\pi a\phi'(b))}. \quad (30)$$

For a single-tone signal  $s(t) = A \cos(2\pi ct)$ , especially, its CWT is given by

$$W_s(a, b) = \frac{1}{2} A e^{i2\pi bc} \overline{\hat{\psi}(2\pi ac)}. \quad (31)$$

If the center frequency of  $\hat{\psi}(\omega)$  is  $\omega_0$ ,  $W_s(a, b)$  will concentrate around  $a = \omega_0/(2\pi c)$ . The IF function of  $s(t)$  is  $\omega_s(a, b) = c$ . Then by Eq. (24), all points  $(a, b)$  in CWT plane can be "squeezed" to  $(c, b)$  in the WSST plane. However, if there is a significant change of frequency in the signal, the efficiency of synchrosqueezing of IF depends on the instantaneous bandwidth of signal in the range of the analysis wavelet. Looking back at Figure 3, we notice that synchrosqueezing performs well near the extrema of IF of the original signal, where the IF changes slowly.

Note that by increasing the width  $\rho$ , we can enhance the accuracy of the reconstruction waveform continuously in Figure 3. However, the important problem is how to deal with the multicomponent signals. Next, let us consider another numerical example given by

$$\begin{aligned} y(t) &= s_1(t) + s_2(t) + s_3(t) + n(t) \\ &= \cos(0.12\pi t + 10\cos(0.006\pi t) + \theta_1) + \cos(0.22\pi t + 10\cos(0.006\pi t) + \theta_2) + \cos(0.38\pi t + \theta_3) + n(t), \end{aligned} \quad (32)$$

where  $n(t)$  is an additive noise. Figure 4 shows the IFs of the three components in  $y(t)$  and also the waveform with an additive Gaussian white noise. The sampling rate is 1Hz and the signal duration is from 0s to 511s. Observe that  $s_3(t)$  starts at 65s and ends at 384s.

Figure 5 shows the CWT, STFT, WSST and FSST of the multicomponent signal  $y(t)$ , respectively. The corresponding parameters  $\gamma$ ,  $\rho$  of these algorithms are the same as those used in Figure 3. Note that both WSST and

FSST do not work well for the two sinusoidal frequency modulation signals. It seems FSST is better than WSST in Figure 5. But when changing the wavelet parameters in WSST, we can obtain better representation for either the high frequency component  $s_3(t)$  or the low frequency component  $s_1(t)$ . When using SST to extract the components, we need to know exactly the number of components and the width of extraction window first. So it is not adaptive.

Actually, we can use the CWT or STFT to recover a signal too (see Eq. (5) and Eq. (14)). But for multicomponent signal, the recovering algorithm based on CWT or STFT is also not adaptive. [27] indicates that the synchrosqueezed transform cannot improve TF resolution of STFT and CWT in the sense if the components are not reliably represented by STFT or CWT, then they will not be well separated by SST. By comparing the results in Figure 3 and Figure 5, and our other experiments, we find that the CWT and STFT are more stable than WSST and FSST in the whole TF plane.

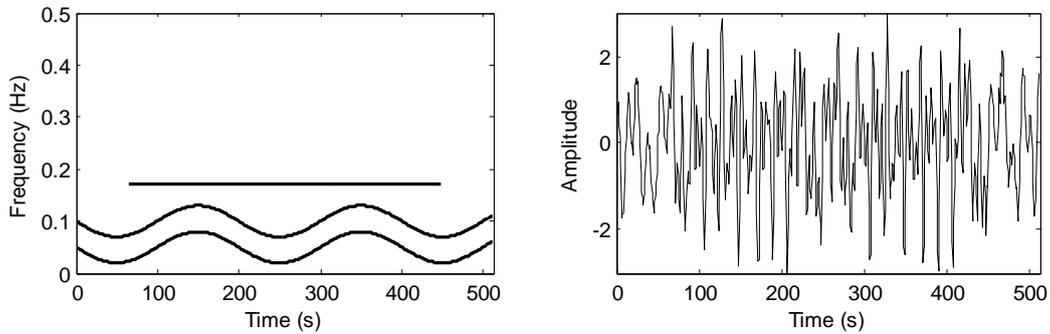


Figure 4. The real IF of  $y(t)$  (left) and the waveform with signal-to-noise ratio (SNR) 15dB (right).

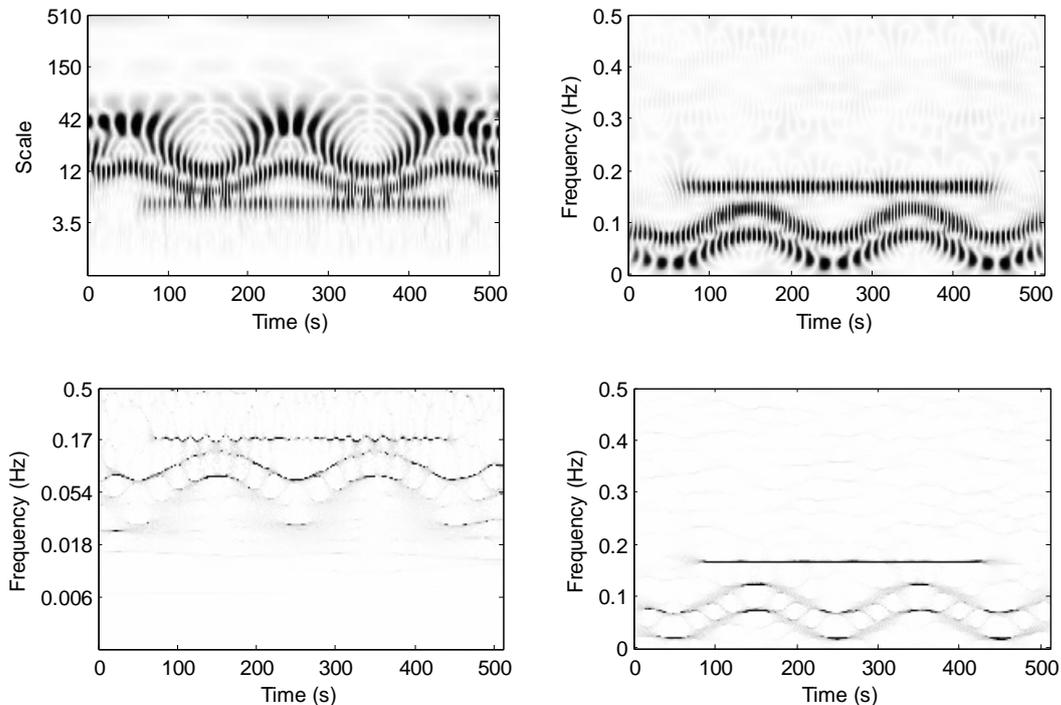


Figure 5. Analysis of multicomponent signal  $y(t)$  : real part of CWT (top left), real part of STFT (top right), WSST (bottom left) and FSST (Bottom right).

### 3.2. Extraction of TF Ridges

Considering the AHM conditions and noise effect, we use the extrema of the real part and imaginary part of STFT or CWT to extract the TF ridges. We take CWT as example in the following analysis. Suppose  $f(t) = s(t) + n(t)$ , where  $n(t)$  is a white noise with mean 0 and variance  $\zeta_n^2$ , and  $s(t) = A(t)\cos(2\pi\phi(t))$  is an IMT function. The CWT of  $f(t)$  is given by

$$\begin{aligned} W_f(a, b) &= W_s(a, b) + W_n(a, b) \\ &\approx \frac{1}{2}A(b)e^{i2\pi\phi(b)}\hat{\psi}(2\pi a\phi'(b)) + \int_{\mathbb{R}} n(t)\frac{1}{a}\overline{\psi\left(\frac{t-b}{a}\right)}dt. \end{aligned} \quad (33)$$

The expectation and variance of  $W_f(a, b)$  are given by

$$E[W_f(a, b)] = W_s(a, b) \approx \frac{1}{2}A(b)e^{i2\pi\phi(b)}\hat{\psi}(2\pi a\phi'(b)), \quad (34)$$

and

$$\text{Var}[W_f(a, b)] \approx \zeta_n^2 C(a), \quad (35)$$

where  $C(a)$  varies with scale  $a$ . See Appendix A for the proofs of Eq. (35) and the following Eq. (36) and Eq. (37).

Suppose  $(a_1, b_1)$  is a point on which  $\Re[W_s(a, b)]$  gains its local extreme value. The signal-to-noise ratio of  $\Re[W_f(a_1, b_1)]$  is defined by

$$r_1(a_1, b_1) = \frac{|\Re[W_s(a_1, b_1)]|^2}{\text{Var}[\Re[W_f(a_1, b_1)]]},$$

and we have

$$r_1(a_1, b_1) = \frac{|A(b_1)\hat{\psi}(2\pi a_1\phi'(b_1))|^2}{\zeta_n^2 C(a_1)}. \quad (36)$$

Analogously, the signal-to-noise ratio of  $W_f(a_1, b_1)$  is

$$r_2(a_1, b_1) = \frac{|W_s(a_1, b_1)|^2}{\text{Var}[W_f(a_1, b_1)]} = \frac{|A(b_1)\hat{\psi}(2\pi a_1\phi'(b_1))|^2}{2\zeta_n^2 C(a_1)} = \frac{1}{2}r_1(a_1, b_1). \quad (37)$$

Note that for imaginary part, the signal-to-noise ratio is the same as Eq. (36) with  $\Re$  replaced by  $\text{Im}$ , provided that  $(a_1, b_1)$  is a local extreme point of  $\text{Im}[W_s(a, b)]$ . Furthermore, the absolute operation is nonlinear, which means that the extracted ridge is not always consistent and unbiased for the IF estimation. After we extract the extrema of the real part and imaginary part first, we use the cubic spline to interpolate the extrema to obtain the ridge of IF.

### 3.3. Effect of the Window/Wavelet Parameters

Unless otherwise noted, the wavelet used in the remainder of this paper is Morlet's wavelet. Morlet's wavelet is composed of a complex exponential multiplied by a Gaussian window, given by

$$\psi(t) = c_{\sigma,\mu} e^{-\frac{1}{2\sigma^2}t^2} \left( e^{i\mu t} - e^{-\frac{1}{2}\mu^2\sigma^2} \right), \quad (38)$$

where  $\mu, \sigma > 0$ ,  $c_{\sigma,\mu}$  is a normalizing constant subject to  $\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1$ . Then we have

$$c_{\sigma,\mu} = \pi^{-\frac{1}{4}} \sigma^{-\frac{1}{2}} / \sqrt{1 + e^{-\mu^2\sigma^2} - 2e^{-\frac{3}{4}\mu^2\sigma^2}}.$$

The Fourier transform of Morlet's wavelet is given by

$$\hat{\psi}(\omega) = c_{\sigma,\mu} \sqrt{2\pi\sigma^2} \left( e^{-\frac{1}{2}\sigma^2(\omega-\mu)^2} - e^{-\frac{1}{2}\sigma^2(\omega+\mu)^2} \right). \quad (39)$$

Note that  $\int_{-\infty}^{\infty} |\hat{\psi}(\omega)|^2 d\omega = 2\pi$ .

The frequency center of  $\psi(t)$  is given by

$$\omega_{\hat{\psi}}^* = \frac{\int_{-\infty}^{\infty} \omega |\hat{\psi}(\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |\hat{\psi}(\omega)|^2 d\omega} = \frac{1 - e^{-\frac{3}{4}\sigma^2\mu^2}}{1 + e^{-\sigma^2\mu^2} - 2e^{-\frac{3}{4}\sigma^2\mu^2}} \mu. \quad (40)$$

The bandwidth of  $\psi(t)$  is given by

$$\Delta_{\hat{\psi}} = \left[ \frac{\int_{-\infty}^{\infty} (\omega - \omega_{\hat{\psi}}^*)^2 |\hat{\psi}(\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |\hat{\psi}(\omega)|^2 d\omega} \right]^{\frac{1}{2}} \approx \left[ \frac{\frac{1}{2\sigma^2} + e^{-\sigma^2\mu^2} \left( \frac{1}{2\sigma^2} + \mu^2 \right) - 2e^{-\frac{3}{4}\sigma^2\mu^2} \left( \frac{1}{2\sigma^2} + \frac{1}{4}\mu^2 \right)}{1 + e^{-\sigma^2\mu^2} - 2e^{-\frac{3}{4}\sigma^2\mu^2}} \right]^{\frac{1}{2}} \approx \frac{1}{\sqrt{2}\sigma}. \quad (41)$$

Usually, we let  $\mu \geq \pi$  and  $\sigma \geq 0.5$ , then  $\omega_{\hat{\psi}}^* \approx \mu$ . Therefore, the center and bandwidth of  $\hat{\psi}(\omega)$  are dependent on  $\mu$  and  $\sigma$ , respectively. In addition, by Eq. (18) and Eq. (19), we have,

$$\begin{cases} \Delta_{\psi(a,b)} = \frac{1}{a} \Delta_{\hat{\psi}} \approx 1/(\sqrt{2}a\sigma), \\ \Delta_{\psi(a,b)} = a \Delta_{\hat{\psi}} \approx a\sigma/\sqrt{2}. \end{cases} \quad (42)$$

By Eq. (31) and Eq. (42), CWT  $W_s(a,b)$  of a single-tone signal  $s(t)$  with Morlet's wavelet will concentrate around  $a = \mu/(2\pi c)$ , with scale support zone

$$\frac{\mu - \frac{1}{2} \frac{1}{\sqrt{2}\sigma}}{2\pi c} \leq a \leq \frac{\mu + \frac{1}{2} \frac{1}{\sqrt{2}\sigma}}{2\pi c}. \quad (43)$$

For a multicomponent signal with pure harmonics  $s(t) = \sum_{k=1}^K A_k \cos(2\pi c_k t)$ , where  $A_k > 0$  and  $c_k > c_{k-1}$ , to separate each component with CWT, the wavelet should satisfy

$$\frac{\mu\sigma + 1/(2\sqrt{2})}{2\pi c_k} \leq \frac{\mu\sigma - 1/(2\sqrt{2})}{2\pi c_{k-1}}, \quad (44)$$

for all  $k > 1$ , which is equivalent to

$$\mu\sigma \geq \frac{1}{2\sqrt{2}} \frac{c_k + c_{k-1}}{c_k - c_{k-1}}. \quad (45)$$

Thus it seems that the bigger value of  $\mu\sigma$  comes better separation of a multicomponent signal in the CWT plane. However, this conclusion may only correct for pure harmonics. To this regard, let us consider a linear frequency modulation signal

$$s(t) = A \cos\left(2\pi\left(ct + \frac{r}{2}t^2\right)\right), \quad (46)$$

with phase  $\phi(t) = ct + \frac{r}{2}t^2$ , instantaneous frequency  $\phi'(t) = c + rt$  and chirp rate  $\phi''(t) = r$ . Then its scale support zone is

$$\frac{\mu - \Delta_h/2}{2\pi(c + rt)} \leq a \leq \frac{\mu + \Delta_h/2}{2\pi(c + rt)}, \quad (47)$$

where  $\Delta_h = \frac{1}{\sqrt{2}} \sqrt{\frac{1}{\sigma^2} + (2\pi r a^2)^2} \sigma^2$ , see the proof in the Appendix B.

Suppose  $c_k > c_{k-1}$  and  $c_k + r_k t > c_{k-1} + r_{k-1} t > 0$ . To separate each component with CWT, the wavelet should satisfy

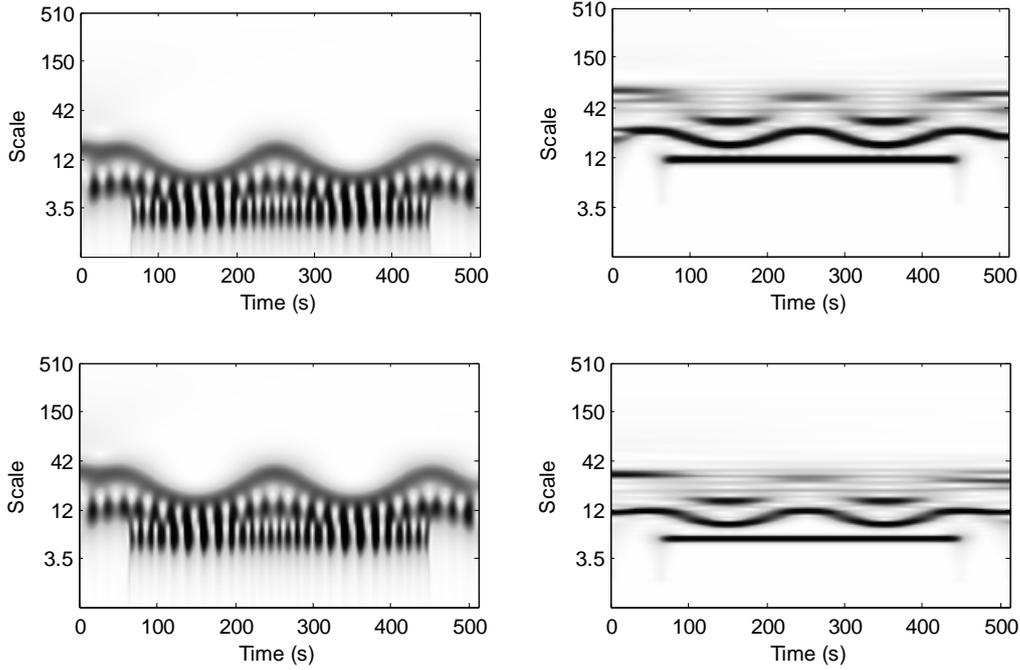
$$\frac{\mu - \Delta_{h_{k-1}}/2}{2\pi(c_{k-1} + r_{k-1}t)} \geq \frac{\mu + \Delta_{h_k}/2}{2\pi(c_k + r_k t)}. \quad (48)$$

Consider  $\Delta_{h_{k-1}}$  and  $\Delta_{h_k}$  on the ridge  $a_{k-1}(t) = \frac{\mu}{2\pi(c_{k-1} + r_{k-1}t)}$  and  $a_k(t) = \frac{\mu}{2\pi(c_k + r_k t)}$ , respectively, we have

$$\frac{\mu\sigma - \frac{1}{2\sqrt{2}} \sqrt{1 + \frac{r_{k-1}^2}{4\pi^2(c_{k-1} + r_{k-1}t)^4} (\mu\sigma)^4}}{2\pi(c_{k-1} + r_{k-1}t)} \geq \frac{\mu\sigma + \frac{1}{2\sqrt{2}} \sqrt{1 + \frac{r_k^2}{4\pi^2(c_k + r_k t)^4} (\mu\sigma)^4}}{2\pi(c_k + r_k t)}. \quad (49)$$

Note that Eq. (49) is a fourth order inequality with respect to  $\mu\sigma$ , where  $0 < \mu\sigma < +\infty$ . So Eq. (49) may hold for some specific intervals on  $\mathbb{R}_+$  of  $\mu\sigma$ .

Therefore, by selecting different the parameters  $\mu$  and  $\sigma$ , we can change the position of the ridge in the CWT plane, and hence change the scale support zone around the ridge. The width of scale support zone depends on both the bandwidth of  $\psi_{(a,b)}$  and the instantaneous bandwidth of the signal to be analyzed. For a monocomponent signal, we



**Figure 6.** CWTs of  $y(t)$  with different parameter:  $\mu = \pi, \sigma = 1$  (top left),  $\mu = 4\pi, \sigma = 1$  (top right),  $\mu = 2\pi, \sigma = 0.5$  (bottom left) and  $\mu = 2\pi, \sigma = 2.5$  (bottom right).

should choose suitable  $\mu$  and  $\sigma$  to make sure the CWT is well concentrated and its ridge keeps away from the upper and lower bound of the CWT plane (finite discrete). For a multicomponent signal, except for the conditions above, we need to separate different components as well as possible.

Take the noise-free multicomponent signal in Figure 4 ( $y(t)$  in Eq. (32) with  $n(t) = 0$ ) for example. Figure 6 shows the CWT with different parameters. When  $\sigma = 1$ ,  $\mu$  increases from  $\pi$  to  $4\pi$ , the distribution moves from the bottom to the top in the CWT plane. On the other hand, when  $\mu = 2\pi$ ,  $\sigma$  increases from 0.5 to 2.5, the distribution holds at the same position but with increasing frequency resolution for the low frequency component. From the left column of Figure 6, we find that the component with lowest frequency is well-represented, but the components with high frequency are not concentrated. On the other hand, from the right column of Figure 6, the two components with high frequency are well-represented, but the component with the lowest frequency is not concentrated. That means we cannot obtain a good representation for both high frequency component and low frequency component simultaneously for CWT and SST.

### 3.4. CWT-based and STFT-based Empirical Signal Separation

Based on the above discussions, we utilize the advantage of EMD to extract the signal components adaptively. The conventional EMD only uses the features in time domain, e.g. maximum and minimum extrema to decompose a signal. Here, we present a new EMD-like method by sifting in the TF plane. The new method is more efficient to separate components close to each other than EMD and SST as demonstrated in the experiment results in Section 4. The proposed

algorithms are called CWT-based empirical signal separation (CWT-ESS) algorithm and STFT-based empirical signal separation (STFT-ESS) algorithm.

In our method, different signal components are extracted by changing the parameters of Morlet's wavelet automatically. Because the higher frequency component always has the ridge with narrower bandwidth when using a suitable window (see Figure 6), namely high energy concentration, so we first extract the high frequency component with maximal amplitude. We use the frequency spectrum to determine the value of parameters adaptively.

Suppose the effective frequency range of the signal is  $[p, q]$  with  $0 < p < q < 1/2$ , and the sampling rate is 1 Hz. Actually, one can use the principle of 3-dB bandwidth to find  $p$  and  $q$ , which are given by

$$\begin{cases} p = \min \left\{ \xi : \arg \left( |\hat{f}(\xi)| > |\hat{f}(\xi_0)|/\sqrt{2} \right) \right\}, \\ q = \max \left\{ \xi : \arg \left( |\hat{f}(\xi)| > |\hat{f}(\xi_0)|/\sqrt{2} \right) \right\}, \end{cases} \quad (50)$$

where  $\hat{f}(\xi)$  is the Fourier transform of  $f(t)$ , and  $\xi_0 = \arg \max_{\xi} |\hat{f}(\xi)|$ . Suppose the signal length is  $N$  ( $N$  is a power of 2, namely  $N = 2^m, m \in \mathbb{Z}_+$ ) and the number of voice is  $n_v$ . We choose a proper  $\mu$  to make the Morlet's wavelet transform of the signal distribute in the center of the CWT plane. Because the domain of the scale is  $\{a_k = 2^{k/n_v}, k = 1, 2, \dots, mn_v\}$  with sample frequency 1Hz, we define the center area (half) of the CWT with the scales  $\{a_k = 2^{k/n_v} : k = mn_v/4, mn_v/4 + 1, \dots, 3mn_v/4\}$  for example. Then by  $a = \mu/(2\pi c)$  in Eq. (43), we calculate

$$\begin{cases} \mu_1 = 2^{3m/4} 2\pi p, \\ \mu_2 = 2^{m/4} 2\pi q. \end{cases} \quad (51)$$

Here to avoid unexpected errors, we set the range of  $\mu$  to be  $[\pi, 5\pi]$ , namely,

$$\mu = \begin{cases} \max \{(\mu_1 + \mu_2)/2, \pi\}, & \text{if } (\mu_1 + \mu_2)/2 < 5\pi; \\ 5\pi, & \text{if } (\mu_1 + \mu_2)/2 \geq 5\pi. \end{cases} \quad (52)$$

By Eq. (45) or Eq. (49), we could determine the range of  $\sigma$  theoretically. But since we have no prior information about the input signal, such as the number of components, instantaneous frequency etc. We choose  $\sigma = 1$  as in [19] for simplicity. Next we calculate the CWT and find a TF ridge. Then obtain a signal mode, an IMF by the sifting process.

Now we present the steps of the CWT-ESS for an amplitude normalized signal  $f(t)$  with detailed algorithm provided in **Algorithm 1**. Figure 7 shows the flow chart of the CWT-ESS algorithm.

For an input signal  $f(t)$ , initialize  $k = 1$  and the residual  $r(t) = f(t)$ .

- **Step 1:** Determine the window/wavelet parameter  $\mu_k$  by (52).

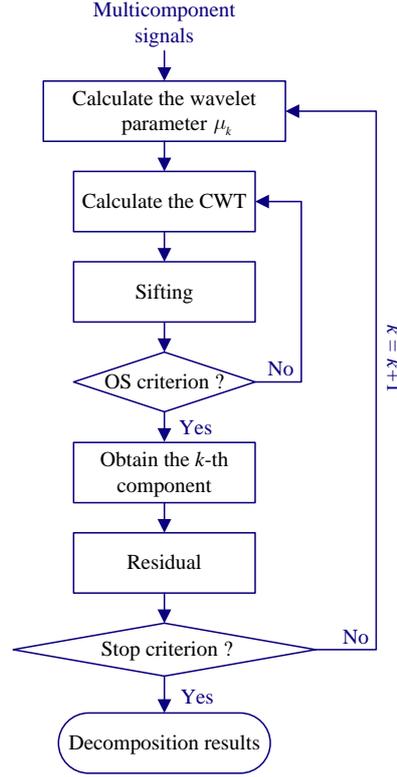


Figure 7. Flow chart of the proposed CWT-ESS algorithm

---

### Algorithm 1 (CWT-ESS)

---

**Input:**  $f(t)$ , let  $k = 1$  and  $r(t) = f(t)$ .

**Start:**

Calculate  $\mu_k$  by Eq. (52).

Let  $l = 1$ ,  $imf_k(t) = 0$ ,  $r_l(t) = r(t)$ , calculate  $W_{r_l}(a, t)$  by Eq. (12).

**Loop A:** While  $\max |W_{r_l}(a, t)| > \gamma_1$ , do (called sifting)

Calculate the ridge  $d_k(t)$ ,

$$c_l(t) = \Re e \left[ 2W_{r_l}(d_k(t), t) / \hat{\psi}(\omega_0) \right], \quad imf_k(t) = imf_k(t) + c_l(t), \quad r_{l+1}(t) = r_l(t) - c_l(t)$$

$$c_{l+1}(t) = \Re e \left[ 2W_{r_{l+1}}(d_k(t), t) / \hat{\psi}(\omega_0) \right].$$

**Loop B:** While  $OS > \gamma_2$ , do

$$l = l + 1,$$

$$c_l(t) = \Re e \left[ 2W_{r_l}(d_k(t), t) / \hat{\psi}(\omega_0) \right], \quad imf_k(t) = imf_k(t) + c_l(t), \quad r_{l+1}(t) = r_l(t) - c_l(t),$$

$$c_{l+1}(t) = \Re e \left[ 2W_{r_{l+1}}(d_k(t), t) / \hat{\psi}(\mu_k) \right].$$

End of **Loop B**

End of **Loop A**

Let  $r(t) = r_l(t)$ ,  $k = k + 1$ , return to **Start**.

---

- **Step 2:** Let  $l = 1$ ,  $imf_k(t) = 0$ ,  $r_l(t) = r(t)$ , calculate the CWT  $W_{r_l}(a, t)$  of  $r_l(t)$  with parameter  $\mu_k$ .
- **Step 3:** Sifting (Loop A in **Algorithm 1**).

(i) Calculate the ridge  $d_k(t)$ . Find an extreme point (with absolute value exceeds a given threshold  $\gamma_1$ ). Then we find local maximum and minimum extrema belong to the same component next to the current point one by one.  $d_k(t)$  is obtained by cubic spline interpolation to these local extrema.

(ii) Obtain the details  $c_l(t)$  and  $c_{l+1}(t)$ . Update  $r_{l+1}(t) = r_l(t) - c_l(t)$  and  $imf_k(t) = imf_k(t) + c_l(t)$ .

$$c_l(t) = \Re e \left[ 2W_{\eta_l} (d_k(t), t) / \hat{\psi}(\omega_0) \right],$$

$$c_{l+1}(t) = \Re e \left[ 2W_{\eta_{l+1}} (d_k(t), t) / \hat{\psi}(\omega_0) \right].$$

(iii) Update  $l = l + 1$ , continue (ii) till  $c_l(t)$  and  $c_{l+1}(t)$  satisfy the OS stop criterion in (54).

(iv) Obtain the  $k$ -th IMF  $imf_k(t)$ .

- **Step 4:** Update  $r(t) = r_l(t)$  and  $k = k + 1$  and repeat Step 1~3, obtain the other IMFs  $imf_k(t)$  and the residual signal  $r(t)$ . ■

In Step 3, we extract the maximum and minimum extrema of the high frequency component in the CWT plane (both real part and the imaginary part), then interpolate the extrema by the cubic spline. We have

$$m_k(t) = \mu_k / d_k(t), \quad (53)$$

where  $m_k(t)$  is the IF of  $imf_k(t)$ .

The stop criterion for **Loop B** is that when  $OS \leq \gamma_2$  ( $\gamma_2 = 1.01$  in this paper), stop iterating [5].

$$OS = \frac{\sum_t |imf_k(t)c_{l+1}(t)|}{\sum_t |(f(t) - imf_k(t))c_{l+1}(t)|}. \quad (54)$$

The stop criterion for **Loop A** is that when  $\max |W_r^l(a, t)| \leq \gamma_1$  ( $\gamma_1 = 0.1$  in this paper).

Note that when the STFT is used, Eq. (53) is replaced by

$$m_k(t) = d_k(t). \quad (55)$$

The detail  $c_l(t)$  in **Algorithm 1** is replaced by

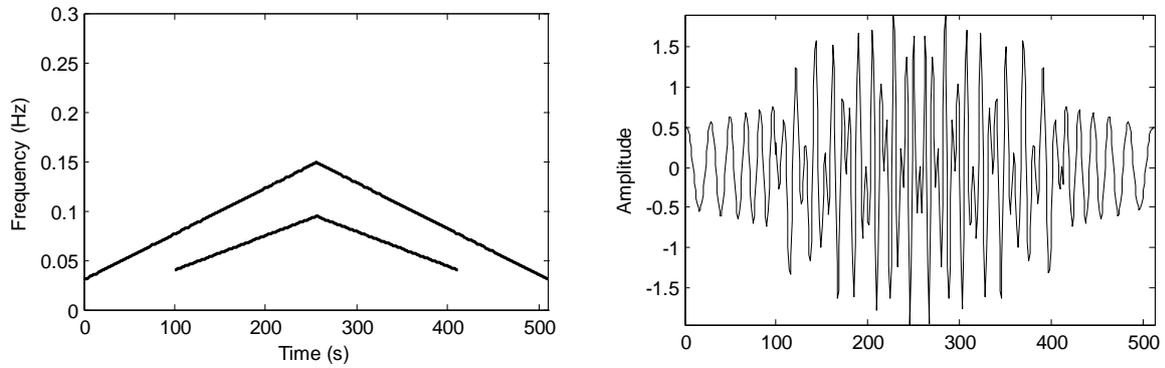
$$c_l(t) = \frac{1}{u(0)} V_f^l (d_k(t), t). \quad (56)$$

Furthermore, because of the equal TF resolution in the whole STFT plane, we simply use a constant window for STFT-ESS, which means we will ignore Step 1.

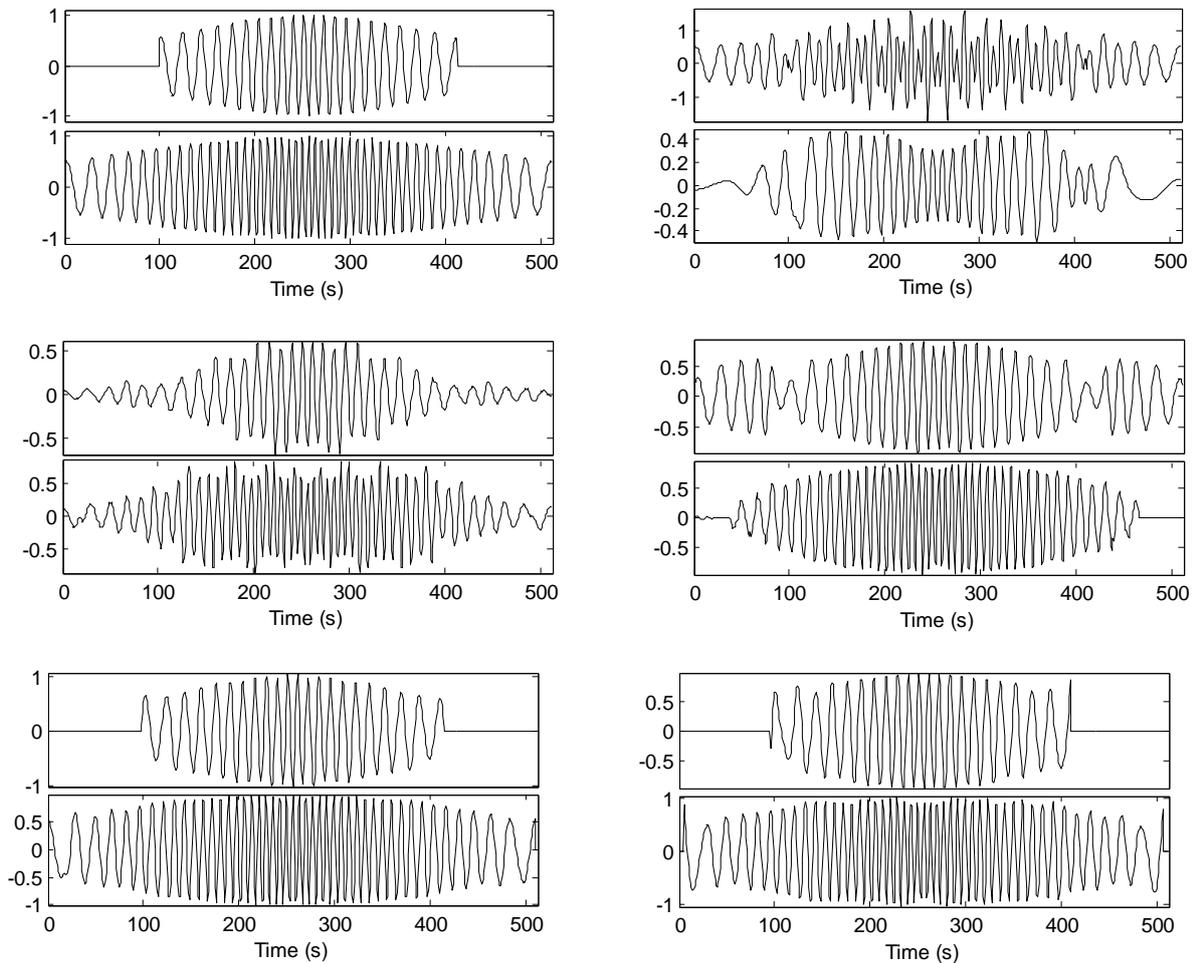
#### 4. Numerical results

In this section, various numerical examples are used to validate the proposed methods. First we deal with the two-component non-stationary signal  $z(t)$  in Figure 8. Component 1 has the low frequency, while Component 2 the

high frequency. Both of their amplitudes are modulated symmetrically by the Gaussian functions, see the original waveforms in Figure 8. And their frequency modulations are symmetrical triangular LFM. The waveforms are continuous.



**Figure 8.** The two-component LFM signal  $z(t)$ : the real IFs (left), the mixed signal (right).



**Figure 9.** The separation results by different methods: two components of the original waveforms (top left), two components obtained by EMD (top right), WSST (middle left), FSST (middle right), CWT-ESS (bottom left) and STFT-ESS (bottom right). In each sub-picture, the top and the bottom correspond to Component 1 and Component 2, respectively.

Figure 9 shows the separation results by different methods. Note that the parameters of WSST and FSST algorithms are the same those in Figure 3 and Figure 5. In addition, the number of components and the width of extraction window are set in advance to be 2 and 5, respectively. We just display the first IMF and second IMF of the decomposition results of EMD by using the code supplied by T. Oberlin [4]. Because of the “mode mixing” (as explained in [7]), EMD cannot separate these two components. For CWT-ESS and STFT-ESS, also we just display the first two modes. By comparing the waveforms of the separation results to the original ones, we can find that our methods are superior to the EMD, WSST and FSST.

We use the relative root mean square error (RMSE) to evaluate the separation performance, which is defined by

$$R_j = \sqrt{\frac{1}{T} \sum_{n=1}^T |s_j(n) - \tilde{s}_j(n)|^2} / \varsigma_{s_j}, \quad (57)$$

where  $\tilde{s}_j(n)$  is the reconstructed  $j$ -th component,  $\varsigma_{s_j}$  is the average power of  $s_j(n)$ ,

$$\varsigma_{s_j} = \sqrt{\frac{1}{T} \sum_{n=1}^T |s_j(n)|^2}, \quad (58)$$

Figure 10 shows the relative RMSE under different signal-to-noise ratios (SNRs) for the two-component triangular LFM signal  $z(t)$ . The additive Gaussian white noises are added to the original signal with SNR from 10 to 20 dB. Under each SNR, we use the Monte-Carlo experiment for 100 runs. Note that the parameters for recovering with WSST and FSST are the same as Figure 9. Because of noises, here we use the EEMD [11] instead of EMD. And for the EEMD, it repeats 200 times for average by adding independent noises with standard deviation 0.1. Obviously, from the results in Figure 10, the algorithms CWT-ESS and STFT-ESS introduced in this paper are superior to EEMD, WSST and FSST.

Next, we move to the three-component signal  $y(t)$  given in Eq. (32), whose IFs and waveform are shown in Figure 4, and WSST and FSST are shown in Figure 5. Because the SST based reconstruction methods are supervised which requires the input of the number of components of the signal and the parameter, the separation performance can be improved a lot by changing the window width manually. However, it is very difficult to get the best window width and component number adaptively and automatically. So, in this experiment, we just compare the separation performance of the unsupervised methods in this paper, i.e. EEMD, CWT-ESS and STFT-ESS. Figure 11 shows the separation errors under different SNRs from 3 to 20 dB. Obviously, the algorithms CWT-ESS and STFT-ESS introduced in this paper are superior to EEMD.

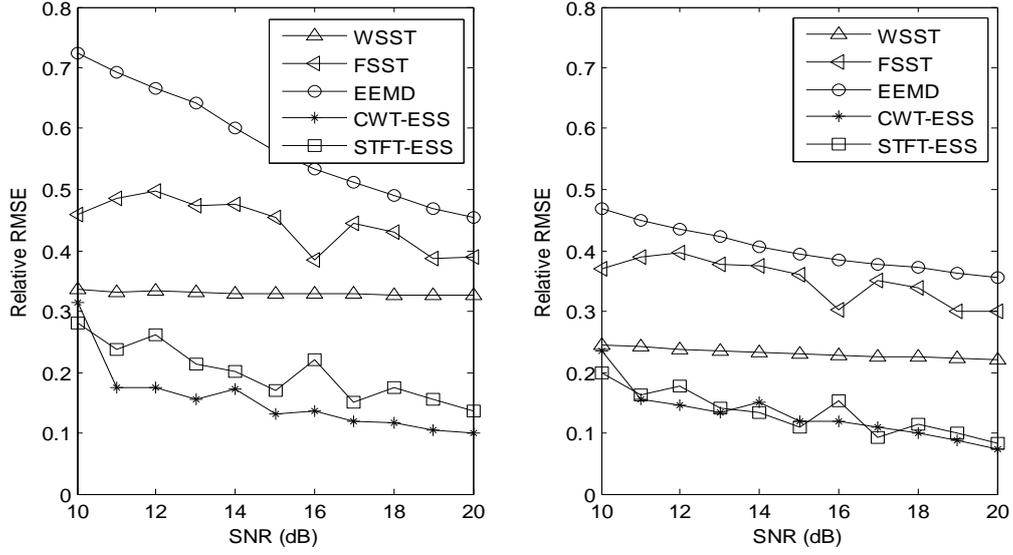


Figure 10. The separation performance of the two-component triangular LFM signal: Component 1 (left) and Component 2 (right).

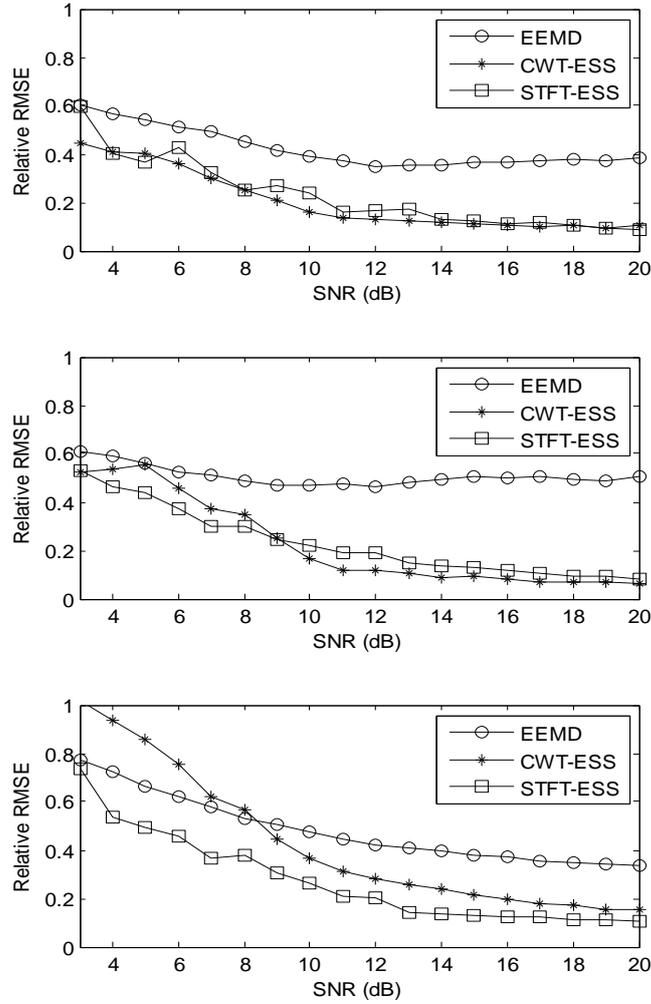
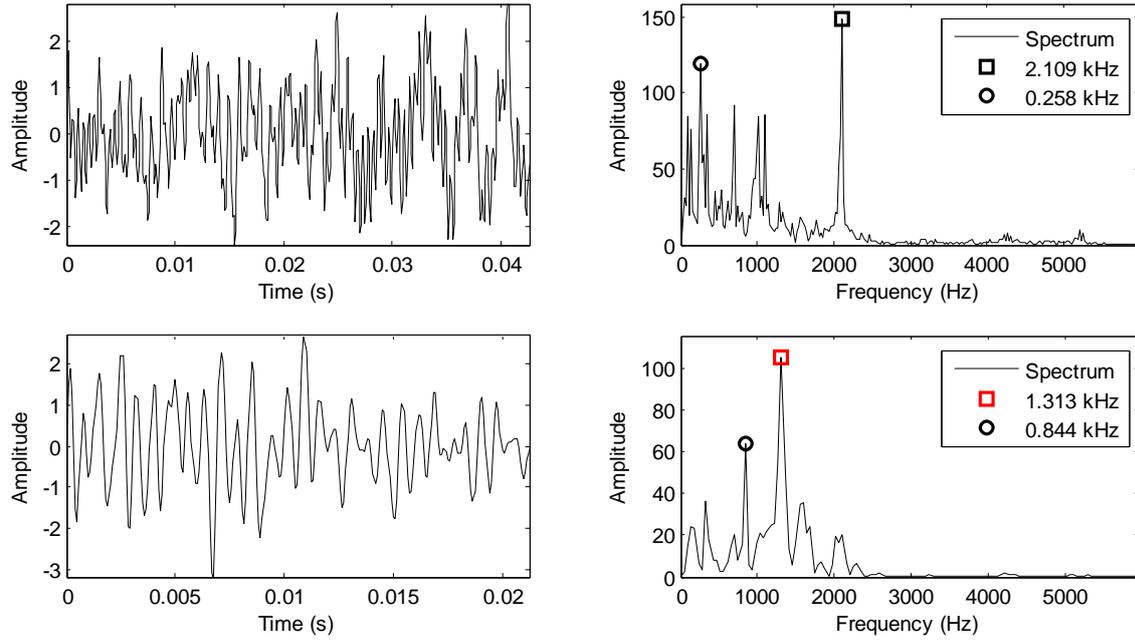


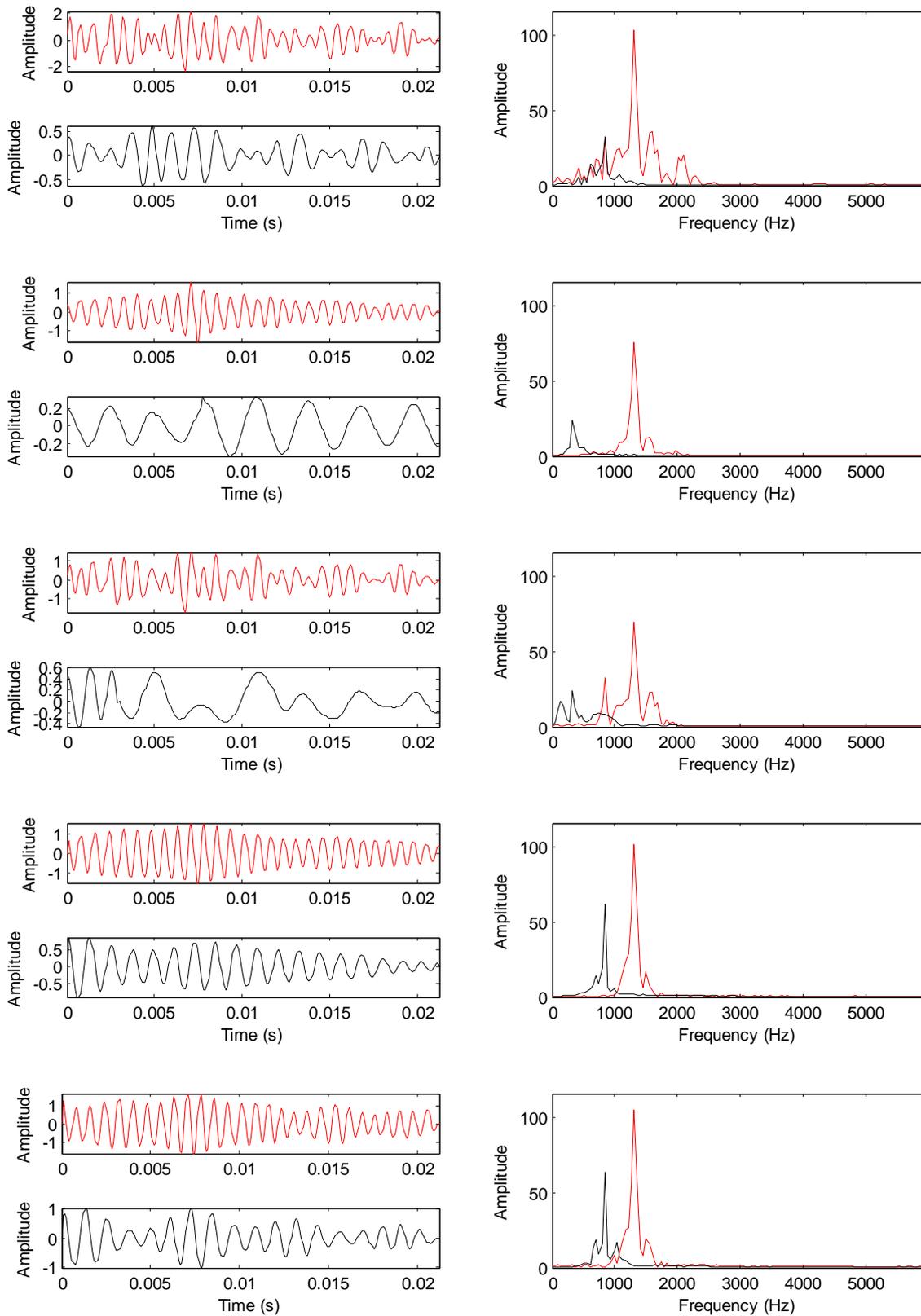
Figure 11. The separation performance of the three-component signal  $y(t)$  in Eq. (32): Component 1 (top), Component 2 (middle) and Component 3 (bottom).



**Figure 12.** The fan end accelerometer data: the waveform of normal bearing (top left), the spectrum of normal bearing (top right), the waveform of fault bearing (bottom left), the spectrum of fault bearing (bottom right).

Finally, we deal with the bearing data from Case Western Reserve University [34] for fault diagnosis (see [35]-[37] for machine fault diagnosis based on the synchrosqueezing method). The bearing data was obtained by using a 2 hp Reliance Electric motor. Vibration data was collected using accelerometers, which were attached to the housing with magnetic bases. In the following experiments, the motor speed is 1730 rpm and the sampling rate is 12 kHz. We use some small pieces of fan end bearing signals to demonstrate the validity of our proposed method. Figure 12 shows the waveforms and spectra of the normal baseline data and fault bearing data. Obverse that the normal bearing signal has one typical frequency (principal component) located at 2.019 kHz. Except for this principal component, there are some complicated vibration frequencies, which are mixed together from about 0 to 1.5 kHz, see the peak at 0.258 kHz.

For the fault bearing signal in Figure 12, the accelerometer for collecting data is placed at 6 o'clock position at the fan end of the motor housing. The motor bearing is seeded with fault using electro-discharge machining. The fault bearing with 0.021 inches fault in diameter is introduced at outer raceway and tested at the speed of 1730 rpm. Note that the spectrum of fault bearing is different from the normal one. The outer raceway faults are stationary. Hence placement of the fault at the fan end has a direct impact on the vibration response of the bearing system. The two typical frequencies are marked at 1.313 kHz and 0.844 kHz. This is the typical feature for bearing fault. Usually we need to find the principal components (intrinsic modes) for bearing fault signal. The principal components are corresponding to the typical features, but reserve the local information of amplitude and phase, which are very useful for fault diagnosis and identification.



**Figure 13.** Separation results of the fault bearing signal by different methods: EMD (first row), WSST (second row), FSST (third row), CWT-ESS (fourth row) and STFT-ESS (last row).

It seems the two components are well-separated in the bottom-right picture of Figure 12. However, because of

background noises and interferences, it is difficult to decompose this signal in frequency domain or TF domain. We use different methods to separate this signal to extract the two principal components, including EMD, WSST, FSST, CWT-ESS and STFT-ESS. Figure 13 shows the separation results with the left column showing the first two modes and the right column showing the corresponding spectra. Note that we use red color for the intrinsic mode with higher frequency, black for the intrinsic mode with lower frequency. Because we have no prior information, the RMSE cannot be applied to evaluate the separation performance. Here we analyze the spectra of the separation results. We say a method can separate this signal well if the spectra of the separation results are just corresponding to the typical frequency components in the bottom-right picture of Figure 12. Observe that EMD cannot extract the intrinsic modes accurately because of the mode mixing. Neither WSST nor FSST can extract the lower frequency component (0.844 kHz) because of interferences. The higher frequency components (1.313 kHz) decomposed by FSST is not accurately too. Both CWT-ESS and STFT-ESS can obtain exactly the two intrinsic modes of the fault bearing signal. Meanwhile, both CWT-ESS and STFT-ESS enhance the energy of the decomposed sub-signals significantly, which means the amplitudes of the spectra in fourth and last row of Figure 13 are close to that in Figure 12.

### **Conclusion**

In this paper, we propose a new EMD-like method to separate multicomponent signal in the STFT or CWT plane. Compared to the SST-based signal separation methods, our method is adaptive and unsupervised. In the sifting process of our method, we adjust the parameters for CWT according to the spectrum of the signal adaptively. The experimental results for synthetic and bearing signals demonstrate that the proposed method separates the models better than EMD and the methods based on standard SST. For multicomponent signals with IFs of components intersecting each other, the methods of how to separate these signals both by our method and the SST will be the subject for further study. The applications of CWT-ESS and STFT-ESS to mechanical signal denoising, machine fault identification and its physics analysis will be our future research.

### **Acknowledgements**

The authors thank anonymous reviewers very much for their valuable comments and suggestions.

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61803294) and Simons Foundation Collaboration Grants for Mathematicians (Grant No. 353185).

### **Appendix**

#### **A. Proofs of Eq. (35)-(37).**

From  $W_f(a, b) = W_s(a, b) + W_n(a, b)$ , we have

$$\begin{aligned}
\text{Var}[W_f(a,b)] &= \text{Var}[W_n(a,b)] \\
&= \text{Var}\left[\int_{\mathbb{R}} n(t+b) \frac{1}{a} \psi\left(\frac{t}{a}\right) dt\right] \\
&\approx \text{Var}\left[\lim_{N \rightarrow \infty} \lim_{\Delta t \rightarrow 0} \sum_{k=-N}^N n(k\Delta t + b) \frac{1}{a} \psi\left(\frac{k\Delta t}{a}\right) \Delta t\right] \\
&= \varsigma_n^2 \lim_{N \rightarrow \infty} \lim_{\Delta t \rightarrow 0} \sum_{k=-N}^N \left| \frac{1}{a} \psi\left(\frac{k\Delta t}{a}\right) \Delta t \right|^2,
\end{aligned} \tag{A.1}$$

where we know  $n(k\Delta t + b)$  is independent with  $n(j\Delta t + b)$ , if  $k \neq j$ . Throughout this paper, we just let sampling rate as 1 Hz, which means  $\Delta t = 1$ , and consider the duration of  $f(t)$  is finite, namely  $k = 0, 1, \dots, N-1$ . Then we have

$$\text{Var}[W_f(a,b)] \approx \varsigma_n^2 \sum_{k=0}^{N-1} \left| \frac{1}{a} \psi\left(\frac{k}{a}\right) \right|^2 = \varsigma_n^2 C(a), \tag{A.2}$$

where  $C(a) = \sum_{k=0}^{N-1} \left| \frac{1}{a} \psi\left(\frac{k}{a}\right) \right|^2$  varies with scale  $a$ .

This shows (35).

Note that  $\text{Var}[W_f(a,b)]$  is irrelevant to the time translation  $b$ . By (36), for real and imaginary parts of CWT, we have

$$\begin{cases} E[\Re[W_f(a,b)]] = \Re[W_s(a,b)], \\ E[\Im[W_f(a,b)]] = \Im[W_s(a,b)]. \end{cases} \tag{A.3}$$

Analogously,

$$\text{Var}[W_f(a,b)] = \text{Var}[\Re[W_f(a,b)]] + \text{Var}[\Im[W_f(a,b)]], \tag{A.4}$$

$$\text{Var}[\Re[W_f(a,b)]] = \text{Var}[\Im[W_f(a,b)]] = \frac{1}{2} \varsigma_n^2 C(a). \tag{A.5}$$

Consider the Morlet's wavelet, where  $\hat{\psi}(\omega)$  is real. When  $\phi(b_1) \in \mathbb{Z}$  or  $(1/2 + \phi(b_1)) \in \mathbb{Z}$ , and  $2\pi a_1 \phi'(b_1) = \omega_0$ , where  $\omega_0 = \arg \max_{\omega} |\hat{\psi}(\omega)|$ , then  $(a_1, b_1)$  is the point with extreme value of the real part  $\Re[W_s(a,b)]$ , with signal to noise ratio

$$r_1(a_1, b_1) = \frac{|\Re[W_s(a_1, b_1)]|^2}{\text{Var}[\Re[W_f(a_1, b_1)]]} = \frac{|A(b_1) \hat{\psi}(2\pi a_1 \phi'(b_1))|^2}{\varsigma_n^2 C(a_1)}. \tag{A.6}$$

The signal-to-noise ratio is the same for the extrema of the imaginary part. Analogously, for  $|W_f(a,b)|$ , the signal to noise ratio is defined by

$$r_2(a_1, b_1) = \frac{|W_s(a_1, b_1)|^2}{\text{Var}[W_f(a_1, b_1)]} = \frac{|A(b_1)\hat{\psi}(2\pi a_1\phi'(b_1))|^2}{2\zeta_n^2 C(a_1)}, \quad (\text{A.7})$$

These show Eq. (36) and Eq. (37).

### B. Proof of Eq. (47).

For  $\psi(t)$  given in (38), when  $\mu\sigma \geq \pi$ ,  $\hat{\psi}(\omega) \approx 0$  for  $\omega \leq 0$ . Thus for  $s(t)$  given by (46), when  $\phi'(t) = c + rt > 0$ , we have

$$\begin{aligned} W_s(a, b) &= \int_{-\infty}^{\infty} s(t) \psi\left(\frac{t-b}{a}\right) \frac{dt}{a} \\ &= \int_{-\infty}^{\infty} s(ax+b) \overline{\psi(x)} dx \\ &= \frac{A}{2} k_0 \int_{-\infty}^{\infty} e^{i2\pi\left(cax+cb+\frac{r}{2}(a^2x^2+2abx+b^2)\right)} e^{-\frac{x^2}{2\sigma^2}} \left( e^{-i\mu x} - e^{-\frac{1}{2}\mu^2\sigma^2} \right) dx \\ &= \frac{A}{2} k_0 \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2} + i\left(-\mu+2\pi ca+2\pi rab\right)x + \pi r a^2 x^2 + 2\pi cb + \pi r b^2} dx - \frac{A}{2} k_0 e^{-\frac{1}{2}\mu^2\sigma^2} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2} + i\left(-0+2\pi ca+2\pi rab\right)x + \pi r a^2 x^2 + 2\pi cb + \pi r b^2} dx \\ &= I(\mu, \sigma, a, b) - e^{-\frac{1}{2}\mu^2\sigma^2} I(0, \sigma, a, b), \end{aligned} \quad (\text{B.1})$$

where  $k_0 = c_{\sigma, \mu}$ . According to the Fourier transform of a linearly-chirped Gaussian pulse [30],

$$\int_{-\infty}^{\infty} e^{-[(g+ih)t^2 - i(\omega-\omega_0)t]} dt = \frac{\sqrt{\pi}}{\sqrt{g+ih}} e^{-\frac{(\omega-\omega_0)^2}{4(g+ih)}} \quad (\text{B.2})$$

we have

$$I(\mu, \sigma, a, b) = \frac{A}{2} k_0 e^{i2\pi cb + i\pi r b^2} \frac{\sqrt{\pi}}{\sqrt{1/(2\sigma^2) - i\pi r a^2}} e^{-\frac{(2\pi ca + 2\pi rab - \mu)^2}{4(1/(2\sigma^2) - i\pi r a^2)}} \quad (\text{B.3})$$

Let  $\omega = 2\pi ca + 2\pi rab$ , and define  $h(\omega)$  by

$$h(\omega) = e^{-\frac{\sigma^2}{2} \frac{1}{1+(2\pi r a^2 \sigma^2)^2} (\omega - \mu)^2} \quad (\text{B.4})$$

Then,

$$I(\mu, \sigma, a, b) = \frac{A}{2} k_0 e^{i2\pi cb + i\pi r b^2} \frac{\sqrt{\pi}}{\sqrt{1/(2\sigma^2) - i\pi r a^2}} e^{-\frac{i\pi r a^2 \sigma^4}{1+(2\pi r a^2 \sigma^2)^2} (\omega - \mu)^2} h(\omega). \quad (\text{B.5})$$

Note that  $e^{-\frac{1}{2}\mu^2\sigma^2} I(0, \sigma, a, b) \approx 0$ , Hence,

$$|W_s(a, b)| \approx |I(\mu, \sigma, a, b)| = \frac{A}{2} k_0 \frac{\sqrt{\pi}}{(1/(2\sigma^2) + \pi r a^2)^{\frac{1}{4}}} h(\omega). \quad (\text{B.6})$$

Therefore, the ridge of  $W_s(a, b)$  is located at  $2\pi ca + 2\pi rab - \mu = 0$ , namely  $a = \frac{\mu}{2\pi(c+rb)} = \frac{\mu}{2\pi\phi'(b)}$ . The bandwidths of  $W_s(a, b)$  and  $h(\omega)$  are the same, which is equal to

$$\Delta_h = \frac{1}{\sqrt{2}} \sqrt{\frac{1}{\sigma^2} + (2\pi ra^2)^2} \sigma^2. \quad (\text{B.7})$$

Hence, the support zone of  $W_s(a, t)$  is  $\frac{\mu - \Delta_h/2}{2\pi\phi'(b)} \leq a \leq \frac{\mu + \Delta_h/2}{2\pi\phi'(b)}$  or  $\frac{\mu - \Delta_h/2}{2\pi(c+rb)} \leq a \leq \frac{\mu + \Delta_h/2}{2\pi(c+rb)}$ .

This shows Eq. (47).

### References

- [1] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C.C. Tung, H.H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceeding of Royal Society A: Mathematical, Physical and Engineering Science* 454 (1998) 903-995.
- [2] Q. Chen, N. Huang, S. Riemenschneider, Y. Xu, "A B-spline approach for empirical mode decompositions," *Advances in Computational Mathematics* 24 (2006) 171-195.
- [3] C.K. Chui, M.D. van der Walt, "Signal analysis via instantaneous frequency estimation of signal components," *International Journal on Geomathematics* 6 (2015) 1-42.
- [4] T. Oberlin, S. Meignen, V. Perrier, "An alternative formulation for the empirical mode decomposition," *IEEE Transactions on Signal Processing* 60 (2012) 2236-2246.
- [5] L. Lin, Y. Wang, H. Zhou, "Iterative filtering as an alternative algorithm for empirical mode decomposition," *Advances in Adaptive Data Analysis* 1 (2009) 543-560.
- [6] F. Wu, L. Qu, "An improved method for restraining the end effect in empirical mode decomposition and its applications to the fault diagnosis of large rotating machinery," *Journal of Sound and Vibration* 314 (2008) 586-602.
- [7] G. Rilling, P. Flandrin, "One or two frequencies? The empirical mode decomposition answers," *IEEE Transactions on Signal Processing* 56 (2008) 85-95.
- [8] G. Rilling, P. Flandrin, "On the influence of sampling on the empirical mode decomposition," *ICASSP 2006, Toulouse, France*, vol. 3, pp. 444-447, May 2006.
- [9] P. Flandrin, G. Rilling, P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Processing Letters* 11 (2004) 112-114.
- [10] L. Li, H. Ji, "Signal feature extraction based on improved EMD method," *Measurement* 42 (2009) 796-803.

- [11] Z. Wu, N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Advances in Adaptive Data Analysis* 1 (2009) 1-41.
- [12] L. Cohen, "Time-frequency distributions-a review," *Proceedings of IEEE* 77 (1979) 941-981.
- [13] S. Qian, D. Chen, "Joint time-frequency analysis," *IEEE Signal Processing Magazine* 16 (1999) 52-67.
- [14] E. Sejdic, I. Djurovic, J. Jiang, "Time-frequency feature representation using energy concentration: An overview for recent advances," *Digital Signal Processing* 19 (2009) 153-183.
- [15] B. Boashash, N. A. Khan, T. Ben-Jabeur, "Time-frequency features for pattern recognition using high-resolution TFDs: A tutorial review," *Digital Signal Processing* 40 (2015) 1-30.
- [16] B. Boashash, "Time-frequency signal analysis and processing: A comprehensive reference," Academic Press, 2015.
- [17] F. Auger, P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Transactions on Signal Processing* 43 (1995) 1068-1089.
- [18] I. Daubechies, S. Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory models," *Wavelet in Medicine and Biology*, pp. 527-546, CRC Press, 1996.
- [19] I. Daubechies, J. Lu, H.-T. Wu, "Synchrosqueezed wavelet transforms: A empirical mode decomposition-like tool," *Applied and Computational Harmonic Analysis* 30 (2011) 243-261.
- [20] C. Li, M. Liang, "A generalized synchrosqueezing transform for enhancing signal time-frequency representation," *Signal Processing* 92 (2012) 2264-2274.
- [21] Q. Jiang, B. W. Suter, "Instantaneous frequency estimation based on synchrosqueezing wavelet transform," *Signal Processing* 138 (2017) 167-181.
- [22] I. Daubechies, Y. Wang, H.-T. Wu, "ConceFT: Concentration of frequency and time via a multitapered synchrosqueezed transform," *Philosophical Transactions of The Royal Society A*, vol. 374, July 2015.
- [23] G. Thakur and H.-T. Wu, "Synchrosqueezing based recovery of instantaneous frequency from nonuniform samples," *SIAM J. Math. Anal.*, 43 (2011) 2078--2095.
- [24] T. Oberlin, S. Meignen, V. Perrier, "The Fourier-based synchrosqueezing transform," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 315-319, May 2014, Florence, Italy.
- [25] T. Oberlin, S. Meignen, V. Perrier, "Second-order synchrosqueezing transform of invertible reassignment? towards ideal time-frequency representations," *IEEE Transactions on Signal Processing* 63 (2015) 1335-1344.
- [26] S. Wang, X. Chen, G. Cai, B. Chen, X. Li, Z. He, "Matching demodulation transform and synchrosqueezing in time-frequency analysis," *IEEE Transactions on Signal Processing* 62 (2014) 69-84.

- [27] D. Iatsenko, P.-V.E. McClintock, A. Stefanovska, "Linear and synchrosqueezed time-frequency representations revisited: Overview, standards of use, resolution, reconstruction, concentration, and algorithms," *Digital Signal Processing* 42 (2015) 1-26.
- [28] J. Gilles, "Empirical wavelet transform," *IEEE Transactions on Signal Processing* 61 (2013) 3999-4010.
- [29] G. Thakur, E. Brevdo, N.-S. Fucar, H.-T. Wu, "The synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications," *Signal Processing* 93 (2013) 1079-1094.
- [30] D.J. Gibson, "Fourier transform of a linearly-chirped Gaussian pulse," 2006, <http://archive.physiker.us/files/physics/ChirpedPulseTransform.pdf>.
- [31] Y. Meyer, *Wavelets and Operators*, vol. 1, Cambridge University Press, 1993.
- [32] I. Daubechies, *Ten lectures on wavelets*, SIAM, CBMS-NSF Regional Conference Series in Applied Mathematics, 1992.
- [33] C. K. Chui, Q. Jiang, *Applied Mathematics-Data Compression, Spectral Methods, Fourier Analysis, Wavelets and Applications*, Atlantis Press, Amsterdam, 2013.
- [34] The Website of the Case Western Reserve University Bearing DATA Center. <http://csegroups.case.edu/bearingdatacenter/pages/12k-drive-end-bearing-fault-data>.
- [35] C. Li and M. Liang, "Time frequency signal analysis for gearbox fault diagnosis using a generalized synchrosqueezing transform," *Mechanical Systems and Signal Processing* 26 (2012) 205-217.
- [36] Z.P. Feng, X.W. Chen, and M. Liang, "Iterative generalized synchrosqueezing transform for fault diagnosis of wind turbine planetary gearbox under nonstationary conditions," *Mechanical Systems and Signal Processing* 52-53 (2015) 360-375.
- [37] S.B. Wang, X.F. Chen, I.W. Selesnick, Y.J. Guo, C.W. Tong and X.W. Zhang, "Matching synchrosqueezing transform: A useful tool for characterizing signals with fast varying instantaneous frequency and application to machine fault diagnosis," *Mechanical Systems and Signal Processing* 100 (2018) 242-288.