Synchrosqueezed wave packet transforms and diffeomorphism based spectral analysis for 1D general mode decompositions

Haizhao Yang

Department of Mathematics, Stanford University

October 2013; revised June 2014

Abstract

This paper develops new theory and algorithms for 1D general mode decompositions. First, we introduce the 1D synchrosqueezed wave packet transform and prove that it is able to estimate instantaneous information of well-separated modes from their superposition accurately. The synchrosqueezed wave packet transform has a better resolution than the synchrosqueezed wavelet transform in the time-frequency domain for separating high frequency modes. Second, we present a new approach based on diffeomorphisms for the spectral analysis of general shape functions. These two methods lead to a framework for general mode decompositions under a weak well-separation condition and a well-different condition. Numerical examples of synthetic and real data are provided to demonstrate the fruitful applications of these methods.

Keywords. Mode decomposition, general shape function, instantaneous, synchrosqueezed wave packet transform, diffeomorphism.

AMS subject classifications: 42A99 and 65T99.

1 Introduction

1.1 Problem statement

In signal processing, analyzing instantaneous properties (e.g., instantaneous frequencies, instantaneous amplitudes and instantaneous phases [4, 23]) of signals has been an important topic for over two decades. In many applications [6, 20, 30, 29, 37, 36], a signal would be a superposition of several components, for example, a complex signal

$$f(t) = \sum_{k=1}^{K} \alpha_k(t) e^{2\pi i N_k \phi_k(t)}, \tag{1}$$

where $\alpha_k(t)$ is the instantaneous amplitude, $2\pi N_k \phi_k(t)$ is the instantaneous phase and $N_k \phi_k'(t)$ is the instantaneous frequency. One wishes to decompose the signal $f(t)$ to obtain each component $\alpha_k(t) e^{2\pi i N_k \phi_k(t)}$ and its corresponding instantaneous properties. This is referred to as the mode decomposition problem.

In spite of considerable successes of analyzing signals by decomposing them in the form (1), a superposition of a few wave-like components belongs to a very limited class of oscillatory patterns. Most of all, decompositions in the form (1) lose important physical

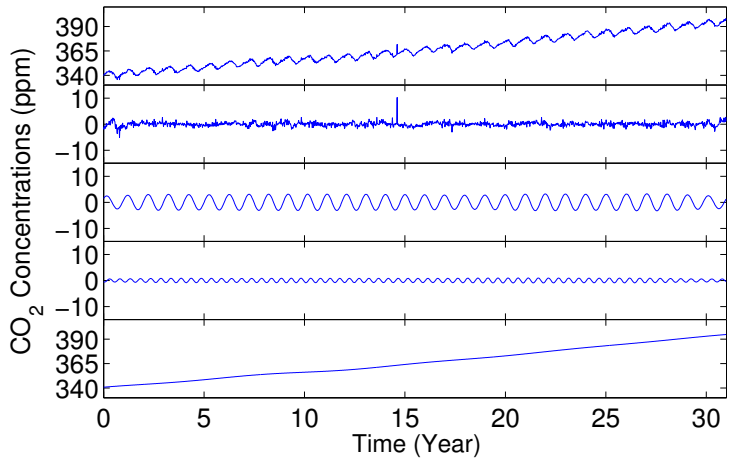

Figure 1: The top signal is the observed CO_2 concentration of recent 31 years (1981-2011) at MLO. Below the original signal are the components provided by the wavelet transform. Only relevant components are separated and presented.

information in some cases as detailed in [30, 35]. To be more concrete, we take the daily atmospheric CO_2 concentration data in [35] as an example. It is observed by National Oceanic and Atmospheric Administration at Mauna Loa (MLO). The method based on wavelet transforms is capable of decomposing data in the form (1), providing one annual cycle, one semiannual cycle and a growing trend (see Figure 1). However, each component alone cannot reflect the true nonlinear evolution pattern: the CO_2 concentration slowly increased in a longer period and quickly decreased in a shorter period. This special pattern is a result of seasonal photosynthetic drawdown and respiratory release of CO_2 by terrestrial ecosystems [35]. Fortunately, such a nonlinear evolution pattern can be recovered by summing up the annual cycle and the semiannual cycle as shown in Figure 2. This motivates the study of a more general decomposition of the form

$$f(t) = \sum_{k=1}^{K} f_k(t) = \sum_{k=1}^{K} \alpha_k(t) s_k(2\pi N_k \phi_k(t)), \tag{2}$$

where $\{s_k(t)\}_{1\le k\le K}$ are 2π-periodic general shape functions. By applying the Fourier expansion of general shape functions, the form (2) is informally similar to the form (1) with infinite terms, i.e.,

$$f(t) = \sum_{k=1}^{K} \alpha_k(t) s_k(2\pi N_k \phi_k(t)) = \sum_{k=1}^{K} \sum_{n=-\infty}^{\infty} \widehat{s_k}(n) \alpha_k(t) e^{2\pi i n N_k \phi_k(t)}. \tag{3}$$

One could combine terms with similar oscillatory patterns in the form (1) to obtain a more efficient and more meaningful decomposition in the form (2). This is the general mode decomposition problem discussed in this paper.

1.2 Synchrosqueezed (SS) time-frequency analysis

A powerful tool for mode decomposition problem is the synchrosqueezed (SS) time-frequency analysis consisting of a linear time-frequency analysis tool and a synchrosqueezing technique [6, 28, 21, 27]. Synchrosqueezed wavelet transforms (SSWT), first proposed in [7] by

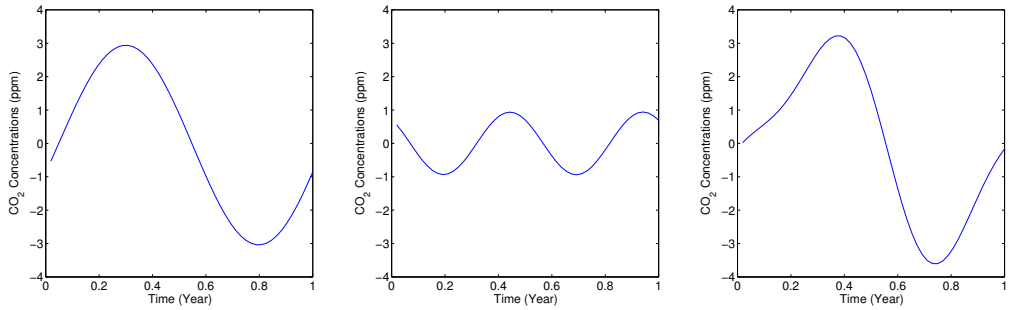

Figure 2: Wave shapes of relevant components provided by wavelet transform. Left: Annual wave shape. Middle: Semiannual wave shape. Right: Summation of the annual wave shape and semiannual wave shape.

Daubechies et al., can accurately decompose a class of superpositions of wave-like components and estimate their instantaneous frequencies, as proved rigorously in [6]. Following this research line, a synchrosqueezed short-time Fourier transform (SSSTFT) and a generalized synchrosqueezing transform have been proposed in [28] and [21], respectively. Stability properties of these synchrosqueezing approaches have been studied in [27]. Recently, [3] adapts SS transforms to strongly modulated components by adding an extra term to the instantaneous frequency estimate function to measure the stronger nonlinearity using some reassignment techniques [1, 2]. This improves the estimation of instantaneous frequencies while keeping the advantage of allowing for simple reconstruction. With these newly developed theories, these synchrosqueezing transforms have been applied to analyze signals in the form (1) in many applications successfully [5, 13, 27, 31, 33].

In the analysis of existing SS transforms [6, 28], the well-separation condition for a class of superpositions of intrinsic mode type functions is a key requirement. Let us take the SSWT as an example. Since it is of significance to study the relation between the magnitudes of instantaneous frequencies and the accuracy of instantaneous frequency estimates, N and N_k are introduced in the following definitions.

Definition 1.1. *(Intrinsic mode type function for the SSWT). A continuous function $f : \mathbb{R} \to \mathbb{C}$, $f \in L^\infty(\mathbb{R})$ is said to be intrinsic-mode-type (IMT) with accuracy $\epsilon > 0$ if $f(t) = a(t)e^{2\pi i N\phi(t)}$ with $a(t)$ and $\phi(t)$ having the following properties:*

$$a \in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R}), \quad \phi \in C^2(\mathbb{R})$$
$$\inf_{t\in\mathbb{R}} \phi'(t) > 0, \quad \sup_{t\in\mathbb{R}} \phi'(t) < \infty, \quad \sup_{t\in\mathbb{R}} |\phi''(t)| < \infty,$$
$$|a'(t)| \le \epsilon |N\phi'(t)|, |\phi''(t)| \le \epsilon |\phi'(t)|, \quad \forall t \in \mathbb{R}.$$

Definition 1.2. *(Superposition of well-separated intrinsic mode functions for the SSWT). A function $f : \mathbb{R} \to \mathbb{C}$ is said to be a superposition of well-separated intrinsic mode functions, up to accuracy ϵ, and with separation $\triangle$, if there exists a finite K, such that*

$$f(t) = \sum_{k=1}^{K} f_k(t) = \sum_{k=1}^{K} a_k(t) e^{2\pi i N_k \phi_k(t)},$$

where all the f_k are IMT, and where moreover their respective phase functions ϕ_k satisfy

$$N_k\phi_k'(t) > N_{k-1}\phi_{k-1}'(t) \quad \text{and} \quad |N_k\phi_k'(t) - N_{k-1}\phi_{k-1}'(t)| \ge \triangle \left[N_k\phi_k'(t) + N_{k-1}\phi_{k-1}'(t)\right], \quad \forall t \in \mathbb{R}.$$

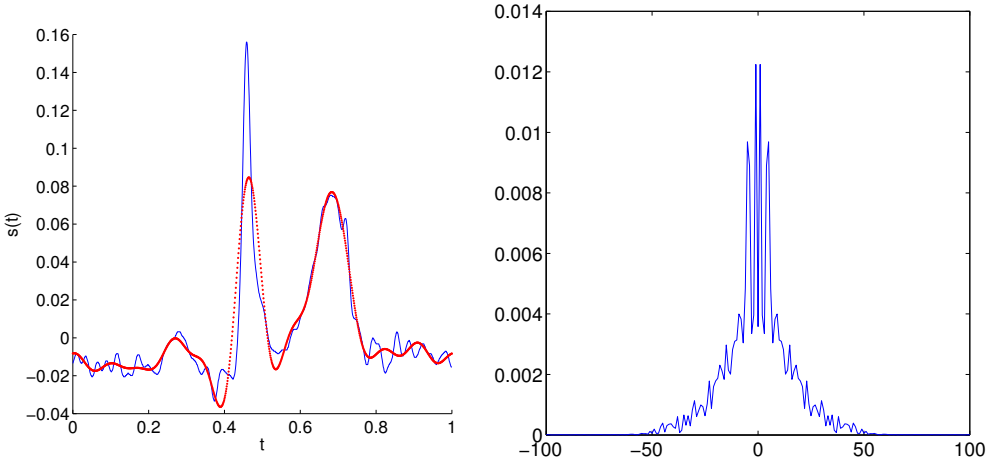

Figure 3: Left: The solid line plots a spike shape function $s(t)$ of a real ECG signal and the dotted line is its band-limited approximation $\sum_{|n|\leq 10} \widehat{s}(n)e^{2\pi int}$. The sum of a few Fourier expansion terms cannot approximate the shape function accurately. Most importantly, the highest peak, the key quantity called R-peak in [11], is smoothed and is hardly distinguished. Right: The Fourier power spectrum $|\widehat{s}(\xi)|$ of $s(t)$ is plotted. The energy in the Fourier domain is spreading widely.

In [6], it is proved that the SSWT can estimate instantaneous frequencies of well-separated intrinsic mode functions from their superposition, using a mother wavelet supported in $[1-d, 1+d]$, with $d < \triangle/(1+\triangle)$. The well-separation condition can be essentially referred to as the condition that the instantaneous frequencies $N_k\phi_k'(t)$ are not crossing over the support of the same wavelet in the time-frequency domain.

For the general mode decomposition problem, a straightforward question would be whether the SS time-frequency representation can extract general modes $\{\alpha_k(t)s_k(2\pi N_k\phi_k(t))\}_{1\leq k\leq K}$, identify general shape functions $\{s_k(t)\}_{1\leq k\leq K}$ and estimate instantaneous properties. Recently, [30] shows that the SSWT can be used to solve the general mode decomposition problem for a superposition of well-separated general modes with analytic wave shape functions $s_k(t)$ sufficiently close to the exponential function e^{it}, i.e., a few terms of the Fourier expansion of $s_k(t)$ are sufficient to approximate $s_k(t)$ well. However, this class of band-limited wave shape functions in [30] is still restrictive in some situations, e.g., spike signals in neurons have shape functions with a wide Fourier band as shown in Figure 3. The SSWT would not be suitable to address the general mode decomposition problem in these circumstances because the well-separation condition for the SSWT is impractical for the following two reasons.

1. The superposition of two nearby Fourier expansion terms $\widehat{s_k}(n)\alpha_k(t)e^{2\pi inN_k\phi_k(t)}$ and $\widehat{s_k}(n+1)\alpha_k(t)e^{2\pi i(n+1)N_k\phi_k(t)}$ are not well-separated when n is large (see Figure 4 for an example), due to the low resolution of wavelet transforms in the high frequency part of the time-frequency domain.

2. For two different instantaneous frequencies $N_k\phi_k'(t)$ and $N_j\phi_j'(t)$, their multiples may have crossover frequencies with high probability.

One possible idea to address the first problem might be to apply the SSSTFT in [28]. The SSSTFT has a much weaker requirement for the well-separation condition and it seems

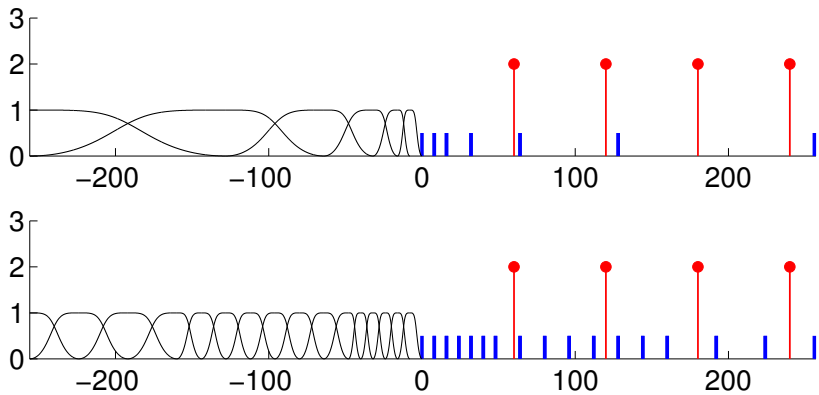

Figure 4: Top: Wavelet tiling and bump functions in the high frequency Fourier domain. The mother wavelet has a support of size $d = 1$. Each short budding in the positive part indicates the center of one bump function, while the bump functions in the negative part are plotted. The dots denote the support of the Fourier transform of $\sum_{n=1}^{4} e^{2\pi i n N t}$, where $N = 60$. The well-separation condition for the SSWT is not satisfied, because the wavelet transform of these modes are overlapping in the time-frequency domain. Bottom: Wave packet tiling and bump functions with $d = 1$. The well-separation condition holds.

to have a much better resolution to distinguish intrinsic mode type functions with high frequencies. However, the SSSTFT may not be suitable to estimate instantaneous frequencies $nN_k\phi_k'(t)$ accurately when nN_k is large. First of all, the resolution parameter $\alpha \geq O(nN_k)$ in [28] is large, resulting in a large error bound of instantaneous frequency estimates. Second, when nN_k is large, $nN_k\phi_k''(t)$ is large. This may lead to an almost zero SSSTFT of the component $\widehat{s_k}(n)\alpha_k(t)e^{2\pi i n N_k \phi_k(t)}$, because the threshold parameter γ in SSSTFT is large. In this case, it is difficult to estimate $nN_k\phi_k'(t)$ and to recover $\widehat{s_k}(n)\alpha_k(t)e^{2\pi i n N_k \phi_k(t)}$. Another idea is to apply the improved SS operator based on time-frequency reassignment in [3], which might lead to a smaller α and a smaller γ in the improved SSSTFT. However, rigorous analysis of the improvement is still under development.

The above problems motivate the design of 1D synchrosqueezed wave packet transforms (SSWPT) and a diffeomorphism based spectral analysis method (DSA). First, the SSWPT has a better resolution to distinguish high frequency harmonic modes than the SSWT and provides more accurate instantaneous frequency estimate than the SSSTFT. Hence, the SSWPT is a good alternative to solve the general mode decomposition problem of the form (2). Second, under the weak well-separation condition that all the general components have at least one term in their Fourier expansions well-separated in the time-frequency domain, the SSWPT can estimate the instantaneous frequency information of each general component. Third, the DSA method can overcome the resolution problem and the crossover frequency problem in existing synchrosqueezing methods once sufficient information is provided by the SSWPT. The DSA method consists of diffeomorphisms and a short-time Fourier transform (in practice, the Fourier transform is applied if $f(t)$ is defined only in a bounded interval). It is capable of decomposing a wide class of general superpositions accurately.

1.3 Related work

There are three other research lines to address the mode decomposition problems of the form (1). The first one is the empirical mode decomposition (EMD) method initialized by Huang et al. in [18] and refined in [19]. To improve the noise resistance of the EMD methods, some variants have been proposed in [17, 34]. It has been shown that the EMD methods can decompose some signals into more general components of the form (2) instead of the form (1) in some cases (see Figure 5 left) in [35]. However, this advantage is not stable and consistent as illustrated in Figure 5 right. It is worth more effort to understand the EMD methods on general mode decompositions.

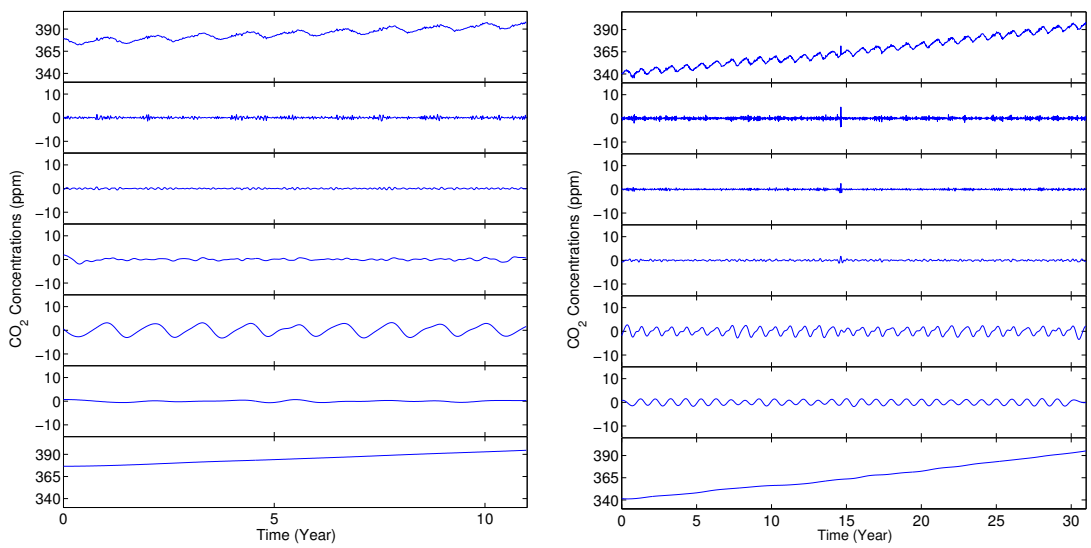

Figure 5: Left: Decomposition results of recent 11 years (2001-2011) CO_2 concentration at MLO by EEMD. Right: The results of recent 31 years (1981-2011). The first row of each table shows the original data and the components provided by the EEMD method in [34] are plotted in an order of decreasing frequencies below the original data. Left: The EEMD method provides a single general component. Right: The EEMD method provides two annual cycles, contrary to the result of 11 years data.

Hou and Shi proposed several nonlinear optimization schemes to decompose signals. The first model in [15] is based on nonlinear TV^3 minimization, which is computationally costly. The second model [16] proposed a nonlinear matching pursuit model based on sparse representations of signals in a data-driven time-frequency dictionary, which has a fast algorithm for periodic data. Under some sparsity assumptions, the analysis of convergence for the latter scheme has been recently studied in [14].

The third method is the empirical wavelet transform recently proposed in [9, 10] by Gilles, Tran and Osher, which empirically builds a wavelet filter bank according to the energy distribution of a given signal in the Fourier domain so as to obtain an adaptive time-frequency representation.

The rest of this paper is organized as follows. In Section 2, the 1D SSWPT and the DSA method are briefly introduced by providing a simple example. In Section 3, main theorems for 1D SSWPT to solve the general mode decomposition problem are presented. In Section 4, the DSA method is theoretically analyzed. In Section 5, some synthetic and

real examples are provided to demonstrate the efficiency of the above two methods. Finally, we conclude with some discussions of future work in Section 6.

2 Implementation of proposed methods

2.1 1D SSWPT

In what follows, we briefly introduce the 1D SSWPT based on the 2D SSWPT in [36]. Let $w(t)$ be a mother wave packet in the Schwartz class and the Fourier transform $\widehat{w}(\xi)$ is a real-valued, non-negative, smooth function with a support equal to $(-d, d)$ determined by a parameter $d \leq 1$. We can use $w(t)$ to define a family of wave packets through scaling, modulation, and translation, controlled by a geometric parameter s.

Definition 2.1. *Given the mother wave packet $w(t)$ and the parameter $s \in (1/2, 1)$, the family of wave packets $\{w_{ab}(t) : |a| \geq 1, b \in \mathbb{R}\}$ is defined as*

$$w_{ab}(t) = |a|^{s/2} w(|a|^s (t-b)) e^{2\pi i (t-b)a},$$

or equivalently, in the Fourier domain as

$$\widehat{w_{ab}}(\xi) = |a|^{-s/2} e^{-2\pi i b \xi} \widehat{w}(|a|^{-s}(\xi - a)).$$

Notice that if s were equal to 1, these functions would be qualitatively similar to the standard wavelets. On the other hand, if s were equal to $1/2$, we would obtain the wave atoms defined in [8]. But $s \in (1/2, 1)$ is essential as we shall see in the main theorems.

The instantaneous frequency of the low frequency part is not well defined as discussed in [23]. For this reason, it is enough to consider the wave packets with $|a| \geq 1$. The high frequency modes can be identified and extracted independently of the low frequency part so that the low frequency part can be recovered by removing high frequency modes.

Definition 2.2. *The 1D wave packet transform of a function $f(t)$ is a function*

$$W_f(a, b) = \langle w_{ab}, f \rangle = \int \overline{w_{ab}(t)} f(t) dt \tag{4}$$

for $|a| \geq 1, b \in \mathbb{R}$.

For $f \in L^2(\mathbb{R})$, if the Fourier transform $\widehat{f}(\xi)$ vanishes for $|\xi| < 1$, the L^2 norms of $W_f(a, b)$ and $f(t)$ are equivalent, i.e., $\exists c_1$ and c_2 such that $0 < c_1 < c_2 < \infty$ and

$$c_1 \int |f(t)|^2 dt \leq \int |W_f(a, b)|^2 da db \leq c_2 \int |f(t)|^2 dt. \tag{5}$$

Definition 2.3. *Instantaneous frequency information function:*

Let $f \in L^\infty(\mathbb{R})$. The instantaneous frequency information function of f is defined by

$$v_f(a, b) = \begin{cases} \frac{\partial_b W_f(a,b)}{2\pi i W_f(a,b)}, & \text{for } |W_f(a, b)| > 0; \\ \infty, & \text{otherwise.} \end{cases} \tag{6}$$

It will be proved that, for a class of wave-like functions $f(t) = \alpha(t) e^{2\pi i N \phi(t)}$, $v_f(a, b)$ precisely approximates $N\phi'(b)$ independently of a as long as $W_f(p, b) \neq 0$. Hence, if we squeeze the coefficients $W_f(a, b)$ together based upon the same instantaneous frequency information function $v_f(a, b)$, then we would obtain a sharpened time-frequency representation of $f(t)$. This motivates the definition of the SS energy distribution as follows.

Definition 2.4. *Given $f(t)$, $W_f(a,b)$, and $v_f(a,b)$, the SS energy distribution $T_f(v,b)$ is defined by*

$$T_f(v,b) = \int_{\mathbb{R}} |W_f(a,b)|^2 \delta(\Re v_f(a,b) - v) da \tag{7}$$

for $v, b \in \mathbb{R}$.

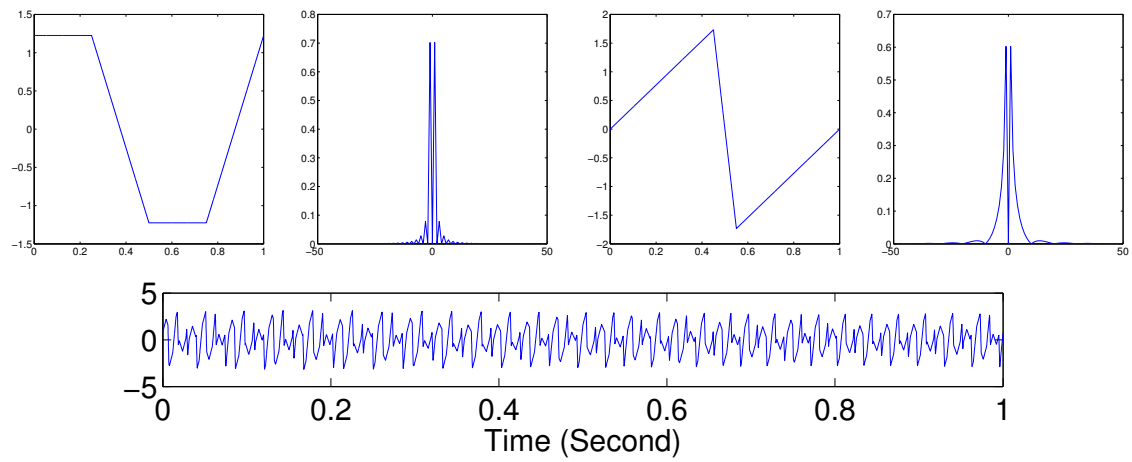

Figure 6: Top left: The general shape function $s_1(t)$ and its spectral energy $|\widehat{s_1}(\xi)|$. Top right: The general shape function $s_2(t)$ and its spectral energy $|\widehat{s_2}(\xi)|$. Bottom: A superposition of general modes generated by using $s_1(t)$ and $s_2(t)$.

For a multi-component signal $f(t)$, the SS energy of each component will concentrate around its corresponding instantaneous frequency. Hence, the SSWPT can provide information about their instantaneous frequencies.

With the definition of the SSWPT above, it is ready to explain how it solves the general mode decomposition problem of the form (2) under well-separation condition. We denote this method as the GMDWP method for short. Let us consider the **Example** 1:

$$f_1(t) = \alpha_1(t) s_1(2\pi N_1 \phi_1(t)) = (1 + 0.05\sin(4\pi x)) s_1\left(120\pi(x + 0.01\sin(2\pi x))\right)$$

and

$$f_2(t) = \alpha_2(t) s_2(2\pi N_2 \phi_2(t)) = (1 + 0.1\sin(2\pi x)) s_1\left(180\pi(x + 0.01\cos(2\pi x))\right),$$

where $s_1(t)$ and $s_2(t)$ are periodic general shape functions defined in $[0,1]$ as shown in Figure 6. Let $f(t) = f_1(t) + f_2(t)$ (see Figure 6 bottom) and we try to recover $f_1(t)$ and $f_2(t)$ from $f(t)$.

Step 1: Apply the SSWPT on $f(t)$ to compute the SS energy distribution $T_f(v,b)$. The essential support of $T_f(v,b)$ (the support above a certain level) is separated into essentially disjoint sets $\{S_{kn}\}$ as shown in Figure 7 left. Each set S_{kn} corresponds to one term in the Fourier expansion of one general mode, say, $\widehat{s_k}(n)\alpha_k(t)e^{2\pi i n N_k \phi_k(t)}$.

Step 2: As we shall see in Theorem 3.4, those points in S_{kn} are concentrating around $nN_k\phi_k'(t)$. This is also illustrated by Figure 7 left. Since each set S_{kn} is well separated from other sets, by applying the clustering method in [36], one can identify each set S_{kn}.

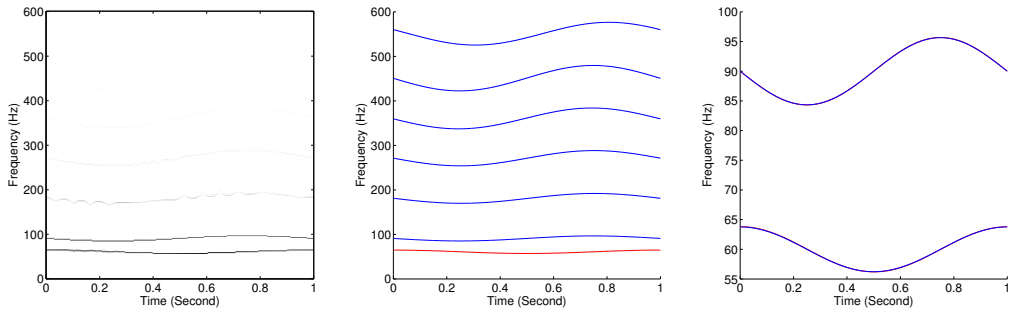

Figure 7: Left: The SS energy distribution of $f(t)$. The Fourier expansion terms with strong energy are well-separated. Middle: The instantaneous frequency estimates $\psi_{kn}(b)$ of $nN_k\phi_k'(b)$ and the result of curve classification as indicated by different colors. Right: The red curves are the instantaneous frequency estimates $\tilde{\psi}_k$ and the blue curves are the real instantaneous frequencies $N_k\phi_k'(b)$.

Step 3: For each point $(v,b) \in S_{kn}$, $v \approx nN_k\phi_k'(b)$. This motivates the definition of a weighted mean of the positions of the points in S_{kn} as

$$\psi_{kn}(b) = \frac{\sum_{(v,b)\in S_{kn}} |T_f(v,b)|v}{\sum_{(v,b)\in S_{kn}} |T_f(v,b)|}.$$

$\psi_{kn}(b)$ provides an accurate estimate of the instantaneous frequency $nN_k\phi_k'(b)$ as shown in Figure 7 middle. In the presence of noise, $\psi_{kn}(b)$ will be disturbed by noise. Hence, some low-pass filter or smoothing process could be applied to $\psi_{kn}(b)$.

Step 4: Apply the curve classification Algorithm 3.7 to identify $\{\psi_{kn}\}_{|n|\geq 1}$ for each k. In the case of Example 1, there are two general modes. Hence, this step gives ψ_{11} and $\{\psi_{2n}\}_{n=1}^{6}$. ψ_{11} is plotted in red and $\{\psi_{2n}\}_{n=1}^{6}$ are plotted in blue in Figure 7 middle.

Step 5: Since $\psi_{kn} \approx nN_k\phi_k'$, one arrives at a function $\tilde{\psi}_k$ as an estimation of the instantaneous frequency $N_k\phi_k'$ of $\alpha_k(t)s_k(2\pi N_k\phi_k(t))$ by applying the method in Theorem 3.9. For Example 1, this step gives $\tilde{\psi}_1$ and $\tilde{\psi}_2$ and they are shown in Figure 7 right in red.

Step 6: Suppose $U_k = \bigcup_{n=-\infty}^{\infty} S_{kn}$, then each general mode can be recovered by

$$f_k(t) = \alpha_k(t)s_k(2\pi N_k\phi_k(t)) = \int_{v_f(a,b)\in U_k} \tilde{w}_{ab}(t)W_f(a,b)dadb,$$

where the set of functions $\{\tilde{w}_{ab}(t), |a| \in [1,\infty), b \in \mathbb{R}\}$ is the dual frame of $\{w_{ab}(x), |a| \in [1,\infty), b \in \mathbb{R}\}$. This step can reconstruct two components $f_1(t)$ and $f_2(t)$ for Example 1 using essential supports of $T_f(v,b)$. These two reconstructed components are shown in Figure 8.

Step 7: Since each Fourier expansion term can be recovered by

$$\widehat{s_k}(n)\alpha_k(t)e^{2\pi i n N_k\phi_k(t)} = \int_{v_f(a,b)\in S_{kn}} \tilde{w}_{ab}(t)W_f(a,b)dadb,$$

one can recover the instantaneous amplitude up to a constant factor by

$$\alpha_k(t) \approx \tilde{\alpha}_k(t) = \sqrt{\sum_n |\widehat{s_k}(n)\alpha_k(t)e^{2\pi i n N_k\phi_k(t)}|^2}.$$

This step gives two instantaneous amplitude estimates $\tilde{\alpha}_1(t)$ and $\tilde{\alpha}_2(t)$ for Example 1. As shown in Figure 8 left, the normalized $\tilde{\alpha}_k(t)$ can estimate the normalized $\alpha_k(t)$ accurately. Hence, the general shape function $s_k(t)$ can be recovered by

$$s_k(t) \approx \frac{f_k((N_k\phi_k)^{-1}(\frac{t}{2\pi}))}{\tilde{\alpha}_k((N_k\phi_k)^{-1}(\frac{t}{2\pi}))},$$

up to a constant factor.

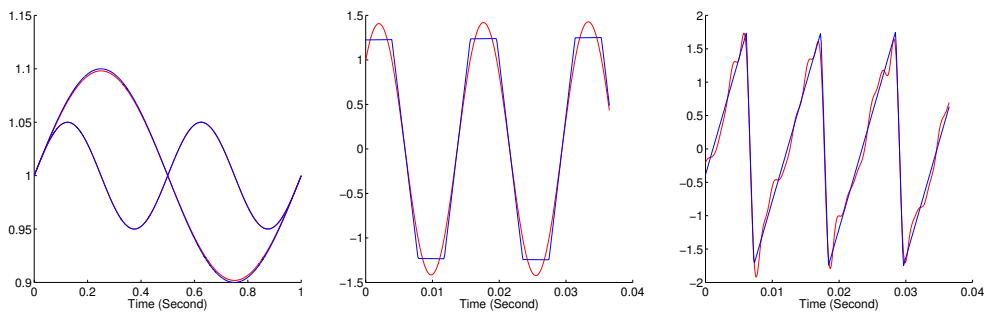

Figure 8: Blue: Real signals. Red: Reconstructed results. Left: Normalized instantaneous amplitude estimates $\tilde{\alpha}_k(t)$ in red and normalized instantaneous amplitudes $\alpha(t)$ in blue. Middle and right: The real general modes and the recovered general modes provided by the GMDWP method.

As we can see in the above example, the GMDWP method can provide accurate estimates of instantaneous frequencies and instantaneous amplitudes from the well-separated essential supports of $T_f(v, b)$. However, the reconstructed general modes are not satisfactory (see Figure 8 middle and right). As Figure 9 shows, considering only the essential supports would ignore Fourier expansion terms with weak energy, the information of which is indispensable to reconstruct exact general modes. This desires the diffeomorphism based spectral analysis for exact reconstructions of general modes.

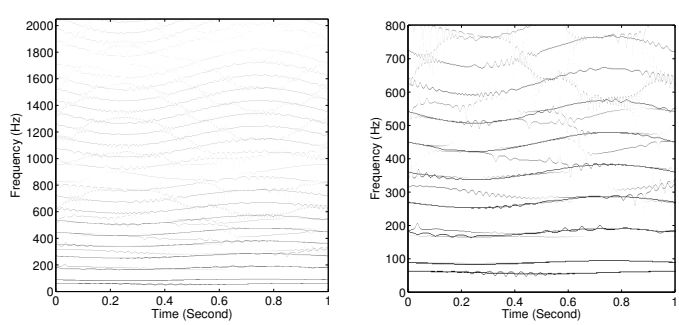

Figure 9: Left: $\log_{10}(T_f(v, b))$ in the visible time-frequency domain. Right: $\log_{10}(T_f(v, b))$ in the low frequency part of the time-frequency domain. Some Fourier expansion terms with weak energy are interfering other terms.

2.2 Diffeomorphism based spectral analysis (DSA)

As discussed previously, the well-separation condition, which assumes that the instantaneous frequencies $nN_k\phi_k'(t)$ of nonzero Fourier expansion terms (i.e. $\widehat{s_k}(n) \neq 0$) are well-separated from each other, is not practical. However, it is reasonable to assume that each general mode has at least one Fourier expansion term $\widehat{s_k}(n_k)\alpha_k(t)e^{2\pi i n_k N_k \phi_k(t)}$ with $n_k N_k \phi_k'(t)$ well-separated from other instantaneous frequencies, so that the SSWPT can estimate $n_k N_k \phi_k'(t)$ accurately. This is referred to as the weak well-separation condition. In what follows, the DSA is introduced to identify all the nonzero Fourier expansion terms using the pre-estimated instantaneous frequencies $\{n_k N_k \phi_k'(t)\}_{k=1}^K$ and the instantaneous amplitudes $\{|\widehat{s_k}(n_k)|\alpha_k(t)\}_{k=1}^K$ provided by the GMDWP method.

Let us assume the signal of interest is defined in $[0, 1]$. Notice that the smooth function $\phi_k(t)$ has the interpretations of a warping in each general mode via a diffeomorphism $\phi_k : \mathbb{R} \to \mathbb{R}$. With the instantaneous frequencies $\{n_k N_k \phi_k'(t)\}_{k=1}^K$ available, we can therefore define the instantaneous phase profile by

$$p_k(t) = \frac{1}{m_k} \int_0^t n_k N_k \phi_k'(x) dx,$$

where $m_k = \frac{1}{2}\left(\max_t n_k N_k \phi_k'(t) + \min_t n_k N_k \phi_k'(t)\right)$. Because $p_k(t)$ is a smooth monotonous function, we can define the inverse-warping profile in $[0, 1]$ by

$$\begin{aligned} h_k(t) &= \frac{f \circ p_k^{-1}(t)}{|\widehat{s_k}(n_k)|\alpha_k \circ p_k^{-1}(t)} \\ &= \sum_{n=-\infty}^{\infty} \frac{\widehat{s_k}(n)}{|\widehat{s_k}(n_k)|} e^{2\pi i(\frac{n m_k}{n_k} t + n N_k \phi_k(0))} \\ &\quad + \sum_{j \neq k} \sum_{n=-\infty}^{\infty} \frac{\widehat{s_j}(n)}{|\widehat{s_k}(n_k)|} \frac{\alpha_j \circ p_k^{-1}(t)}{\alpha_k \circ p_k^{-1}(t)} e^{2\pi i n N_j \phi_j \circ p_k^{-1}(t)}. \end{aligned}$$

If the diffeomorphisms $\phi_k : \mathbb{R} \to \mathbb{R}$ are well different and the phases $2\pi N_k \phi_k(t)$ are sufficiently steep in $[0, 1]$, which will be clarified later, the Fourier transform of each inverse-warping profile $\widehat{h_k}(\xi)$ will have sheer peaks at $\xi = \frac{n m_k}{n_k}$ and will be relative small elsewhere. This motivates the design of the DSA method as follows.

Step 1: Input: A signal $f(t)$, its instantaneous phase profiles $\{p_k(t)\}_{k=1}^K$ and instantaneous amplitudes $\{|\widehat{s_k}(n_k)|\alpha_k(t)\}_{k=1}^K$.

Step 2: Initialize: Set up the initial residual $r(t) = f(t)$ and the tolerance ϵ. Let $f_k(t) = 0$ be the initial guess of the kth general mode and denote $S_k = \emptyset$ as the initial guess of the spectrum information of the kth general shape function s_k for $k = 1, \ldots, K$.

Step 3: For $k = 1, \ldots, K$, compute the inverse-warping profiles in $[0, 1]$ by

$$h_k(t) = \frac{r \circ p_k^{-1}(t)}{|\widehat{s_k}(n_k)|\alpha_k \circ p_k^{-1}(t)}.$$

Step 4: Apply the discrete Fourier transform on $h_k(t)$ in $[0, 1]$ to obtain $\widehat{h_k}(\xi)$ for $k = 1, \ldots, K$ and solve the following optimization problem,

$$(\tau, j) = \arg\max_{(\xi,k)} |\widehat{h_k}(\xi)|.$$

Then $\tau \approx \frac{nm_j}{n_j}$ for some n such that $\widehat{s_j}(n) \neq 0$.

Step 5: Let $g(t) = e^{2\pi i \tau t}$. Warp the harmonic $g(t)$ with the jth instantaneous phase profile $p_j(t)$ and multiply the warped harmonic by the jth instantaneous amplitude $|\widehat{s_j}(n_j)|\alpha_j(t)$ to obtain

$$\begin{aligned} |\widehat{s_j}(n_j)|\alpha_j(t) g \circ p_j(t) &\approx |\widehat{s_j}(n_j)|\alpha_j(t) e^{2\pi i \frac{nm_j}{n_j} p_j(t)} \\ &= |\widehat{s_j}(n_j)| e^{-2\pi i n N_j \phi_j(0)} \alpha_j(t) e^{2\pi i n N_j \phi_j(t)}. \end{aligned}$$

Step 6: Solve the L^2 minimization problem for a complex factor $\beta \in \mathbb{C}$ such that

$$\beta = \arg\min_{\beta \in \mathbb{C}} \|r(t) - \beta |\widehat{s_j}(n_j)|\alpha_j(t) g \circ p_j(t)\|_{L^2}.$$

Then

$$\beta |\widehat{s_j}(n_j)|\alpha_j(t) g \circ p_j(t) \approx \widehat{s_j}(n)\alpha_j(t) e^{2\pi i n N_j \phi_j(t)},$$

which implies

$$|\beta| = \frac{|\widehat{s_j}(n)|}{|\widehat{s_j}(n_j)|}.$$

Step 7: Update: Compute the new residual

$$r(t) = r(t) - \beta |\widehat{s_j}(n_j)|\alpha_j(t) g \circ p_j(t).$$

Update the jth recovered general mode

$$f_j(t) = f_j(t) + \beta |\widehat{s_j}(n_j)|\alpha_j(t) g \circ p_j(t),$$

and the jth spectrum information set

$$S_j = S_j \cup \{(\tau, |\beta|)\}.$$

Step 8: If $\|r(t)\|_{L^2} > \epsilon$, repeat step 3-7. Otherwise, stop iterating and export the general mode estimates f_k and the spectrum information S_k for $k = 1, \ldots, K$.

Notice that for each pair $(\tau, |\beta|) \in S_k$, $(\tau, |\beta|) \approx (\frac{nm_k}{n_k}, \frac{|\widehat{s_k}(n)|}{|\widehat{s_k}(n_k)|})$ for some n such that $\widehat{s_k}(n) \neq 0$. For each k, let

$$d_k(\xi) = \begin{cases} |\beta|, & \text{for } \xi = \tau \text{ such that } (\tau, \beta) \in S_k \\ 0, & \textit{otherwise}. \end{cases}$$

Then

$$|\widehat{s_k}(n)| \approx |\widehat{s_k}(n_k)| d_k(\frac{m_k n}{n_k}) \Rightarrow d_k(\xi) \approx \frac{1}{|\widehat{s_k}(n_k)|} \left| \widehat{s_k}(\frac{n_k}{m_k}\xi) \right|.$$

Hence, $d_k(\xi)$ is an approximation of the spectral energy $|\widehat{s_k}(\xi)|$ up to a constant factor and a scaling.

The DSA method proposed above can take into account all the Fourier expansion terms, even if there are weak energy terms and crossover frequencies. Let us consider the **Example** 1 again. The reconstructed general modes recovered by the DSA method shown in Figure 10 are exactly the desired general modes.

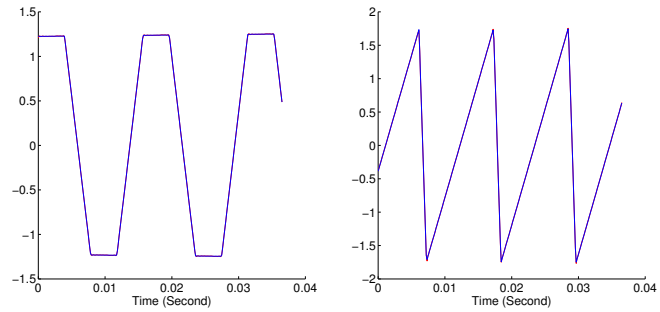

Figure 10: Blue: Real signals. Red: Reconstructed results. Two recovered general modes provided by the DSA method.

3 Analysis of the 1D SSWPT

In this section, we provide rigorous analysis of the 1D SSWPT for the general mode decomposition problem following the model in [36].

3.1 General mode decomposition problems

Definition 3.1. *General shape functions:*

The general shape function class $\mathcal{S}_M$ consists of 2π-periodic functions $s(t)$ in the Wiener Algebra with a unit $L^2([-\pi,\pi])$-norm and a L^∞-norm bounded by M satisfying the following spectral conditions:

1. *The Fourier series of $s(t)$ is uniformly convergent;*
2. *$\sum_{n=-\infty}^{\infty} |\widehat{s}(n)| \leq M$ and $\widehat{s}(0) = 0$;*
3. *Let Λ be the set of integers $\{|n| : \widehat{s}(n) \neq 0\}$. The greatest common divisor $\gcd(s)$ of all the elements in Λ is 1.*

In fact, if $\gcd(s) > 1$, then the general mode $s(2\pi N\phi(t))$ can be considered as a more oscillatory mode $\tilde{s}(2\pi \gcd(s) N\phi(t))$ with $\gcd(\tilde{s}) = 1$ and the Fourier coefficients $\widehat{\tilde{s}}(n) = \widehat{s}(\gcd(s)n)$. The requirement that $\widehat{s}(0) = 0$ and s has a unite $L^2([-\pi,\pi])$-norm is to normalize the general shape function.

Definition 3.2. *A function $f(t) = \alpha(t)s(2\pi N\phi(t))$ is a general intrinsic mode type function (GIMT) of type (M,N), if $s(t) \in \mathcal{S}_M$ and $\alpha(t)$ and $\phi(t)$ satisfy the conditions below.*

$$\alpha(t) \in C^\infty, \quad |\alpha'| \leq M, \quad 1/M \leq \alpha \leq M$$
$$\phi(t) \in C^\infty, \quad 1/M \leq |\phi'| \leq M, \quad |\phi''| \leq M.$$

Definition 3.3. *A function $f(t)$ is a well-separated general superposition of type (M,N,K,s), if*

$$f(t) = \sum_{k=1}^{K} f_k(t),$$

where each $f_k(t) = \alpha_k(t)s_k(2\pi N_k\phi_k(t))$ is a GIMT of type (M,N_k) such that $N_k \geq N$ and the phase functions satisfy the separation condition: for any pair (a,b), there exists at most

one pair (n,k) *such that* $\widehat{s_k}(n) \neq 0$ *and that*

$$|a|^{-s}|a - nN_k\phi_k'(b)| < d.$$

We denote by $GF(M,N,K,s)$ *the set of all such functions.*

Notice that the mode decomposition problem of the form (1) is a special case of the general mode decomposition problem of the form (2). We therefore only analyze the 1D SSWPT for the general mode decomposition problem. Besides, the components are not necessarily defined in the whole domain $\mathbb{R}$, because the SS transforms are localized so as to capture the non-linear, non-stationary features of signals as illustrated in [6, 36]. For the sake of convenience, we omit the discussion of the data segments.

3.2 Instantaneous frequency estimates

The estimates of the instantaneous frequencies $\{N_k\phi_k'(t)\}_{k=1}^K$ rely on Theorem 3.4, Algorithm 3.7, and Theorem 3.9 below.

Theorem 3.4. *For a function* $f(t)$ *and* $\varepsilon > 0$*, we define*

$$R_\varepsilon = \{(a,b) : |W_f(a,b)| \geq |a|^{-s/2}\sqrt{\varepsilon}\}$$

and

$$Z_{n,k} = \{(a,b) : |a - nN_k\phi_k'(b)| \leq d|a|^s\}$$

for $1 \leq k \leq K$ *and* $|n| \geq 1$*. For fixed* M *and* K*, for any* $\varepsilon > 0$*, there exists a constant* $N_0(M,K,s,\varepsilon) > 0$ *such that for any* $N > N_0(M,K,s,\varepsilon)$ *and* $f(t) \in GF(M,N,K,s)$ *the following statements hold.*

(i) $\{Z_{n,k} : 1 \leq k \leq K, \widehat{s_k}(n) \neq 0\}$ *are disjoint and* $R_\varepsilon \subset \bigcup_{1\leq k\leq K}\bigcup_{\widehat{s_k}(n)\neq 0} Z_{n,k}$*;*

(ii) For any $(a,b) \in R_\varepsilon \cap Z_{n,k}$*,*

$$\frac{|v_f(a,b) - nN_k\phi_k'(b)|}{|nN_k\phi_k'(b)|} \lesssim \sqrt{\varepsilon}.$$

The proof of Theorem 3.4 relies on two lemmas as follows to estimate the asymptotic behavior of $W_f(a,b)$ and $\partial_b W_f(a,b)$ as N going to infinity. In what follows, when we write $O(\cdot)$, $\lesssim$, or $\gtrsim$, the implicit constants may depend on M and K.

Lemma 3.5. *Suppose* $\Omega_a = \{(k,n) : a \in [\frac{nN_k}{2M}, 2MnN_k]\}$*. Under the assumption of Theorem 3.4, for any* $\varepsilon > 0$*, we have*

$$W_f(a,b) = |a|^{-s/2}\left(\sum_{(k,n)\in\Omega_a} \widehat{s_k}(n)\alpha_k(b)e^{2\pi i n N_k\phi_k(b)}\widehat{w}\left(\left(a - nN_k\phi_k'(b)\right)|a|^{-s}\right) + O(\varepsilon)\right),$$

when N *is sufficiently large.*

Proof. Without loss of generality, we can simply assume $N_k = N$ for all k and only prove the case for $a > 1$. Because $w(t)$ decays rapidly, the wave packet transform $W_f(a,b)$ is

well defined. By the uniform convergence of the Fourier series of $s_k(t)$ and the change of variables, we have

$$\begin{aligned} W_f(a,b) &= \int_{\mathbb{R}} \sum_{k=1}^{K} \alpha_k(t) s_k(2\pi N\phi_k(t)) a^{s/2} w(a^s(t-b)) e^{-2\pi i(t-b)a} dt \\ &= a^{-s/2} \sum_{k=1}^{K} \sum_{|n|\geq 1} \widehat{s_k}(n) \int_{\mathbb{R}} \alpha_k(a^{-s}x+b) w(x) e^{2\pi i(nN\phi_k(a^{-s}x+b)-a^{1-s}x)} dx. \end{aligned}$$

Let us estimate $I_{kn} = \widehat{s_k}(n) \int_{\mathbb{R}} \alpha_k(a^{-s}x+b) w(x) e^{2\pi i(nN\phi_k(a^{-s}x+b)-a^{1-s}x)} dx$. Let $h(x) = \widehat{s_k}(n)\alpha_k(a^{-s}x+b)w(x)$ and $g(x) = 2\pi(nN\phi_k(a^{-s}x+b) - a^{1-s}x)$, then

$$I_{kn} = \int_{\mathbb{R}} h(x) e^{ig(x)} dx,$$

and

$$g'(x) = 2\pi a^{-s}(nN\phi_k'(a^{-s}x+b) - a).$$

If $a < \frac{nN}{2M}$, then $|g'(x)| \gtrsim a^{-s}nN \gtrsim (nN)^{1-s}$. If $a > 2MnN$, then $|g'(x)| \gtrsim a^{1-s} \gtrsim (nN)^{1-s}$. So, if $a \notin [\frac{nN}{2M}, 2MnN]$, then $|g'(x)| \gtrsim (nN)^{1-s}$. For real smooth functions $g(x)$, we define the differential operator

$$L = \frac{1}{i}\frac{\partial_x}{g'}.$$

Because $h(x)$ decays sufficiently fast at infinity, we perform integration by parts r times to get

$$\int_{\mathbb{R}} h e^{ig} dx = \int_{\mathbb{R}} h(L^r e^{ig}) dx = \int_{\mathbb{R}} ((L^*)^r h)\, e^{ig} dx,$$

where L^* is the adjoint of L. A few algebraic calculation shows that L^* contributes a factor of order $\frac{1}{|g'|} \lesssim \frac{1}{(nN)^{1-s}}$ if $a \notin [\frac{nN}{2M}, 2MnN]$, and we therefore have

$$|I_{kn}| = \left| \int_{\mathbb{R}} e^{ig}\, ((L^*)^r h)\, dx \right| \lesssim |\widehat{s_k}(n)|(nN)^{-(1-s)r} \lesssim |\widehat{s_k}(n)|\varepsilon.$$

Since $s < 1$ and $\sum_{n=-\infty}^{\infty} |\widehat{s_k}(n)| \leq M$, if $N \gtrsim \varepsilon^{\frac{-1}{(1-s)r}}$, then

$$a^{-s/2} \sum_{(k,n)\notin\Omega_a} I_{kn} \lesssim a^{-s/2} \sum_{(k,n)\notin\Omega_a} |\widehat{s_k}(n)| O(\varepsilon) \lesssim a^{-s/2} O(\varepsilon). \tag{8}$$

Now let us estimate I_{kn} when $a \in [\frac{nN}{2M}, 2MnN]$. Recall that

$$I_{kn} = \widehat{s_k}(n) \int_{\mathbb{R}} \alpha_k(a^{-s}x+b) w(x) e^{2\pi i(nN\phi_k(a^{-s}x+b)-a^{1-s}x)} dx.$$

By Taylor expansion,

$$\alpha_k(a^{-s}x+b) = \alpha_k(b) + \alpha_k'(b^*)a^{-s}x$$

and

$$\phi_k(a^{-s}x+b) = \phi_k(b) + \phi_k'(b)a^{-s}x + \frac{1}{2}\phi_k''(b^{**})a^{-2s}x^2$$

for some b^* and b^{**}. Notice that, if $N \gtrsim \varepsilon^{-1/s}$, then

$$\begin{aligned} & |I_{kn} - \widehat{s_k}(n)\alpha_k(b)\int_{\mathbb{R}} w(x)e^{2\pi i(nN\phi_k(a^{-s}x+b)-a^{1-s}x)}dx| \\ \lesssim\ & |\widehat{s_k}(n)|\alpha_k'(b^*)a^{-s}\int_{\mathbb{R}}|x||w(x)|dx \\ \lesssim\ & |\widehat{s_k}(n)|O(\varepsilon). \end{aligned}$$

This implies that

$$I_{kn} = \left(\widehat{s_k}(n)\alpha_k(b)\int_{\mathbb{R}} w(x)e^{2\pi i(nN\phi_k(a^{-s}x+b)-a^{1-s}x)}dx + |\widehat{s_k}(n)|O(\varepsilon)\right)$$

for $a \in [\frac{nN}{2M}, 2MnN]$ and $N \gtrsim \varepsilon^{-1/s}$. Since $|e^{ix}-1| \le |x|$, if $N \gtrsim \varepsilon^{-1/(2s-1)}$, then we have

$$\begin{aligned} & |I_{kn} - \widehat{s_k}(n)\alpha_k(b)\int_{\mathbb{R}} w(x)e^{2\pi i(nN\phi_k(b)+nN\phi_k'(b)a^{-s}x-a^{1-s}x)}dx| \\ \lesssim\ & |\widehat{s_k}(n)|\left(O(\varepsilon) + \left|\alpha_k(b)\int_{\mathbb{R}} w(x)e^{2\pi i(nN\phi_k(b)+nN\phi_k'(b)a^{-s}x-a^{1-s}x)}\left(e^{2\pi inN\frac{1}{2}\phi_k''(b^{**})a^{-2s}x^2}-1\right)dx\right|\right) \\ \lesssim\ & |\widehat{s_k}(n)|\left(O(\varepsilon) + nNa^{-2s}\int_{\mathbb{R}} x^2|w(x)|dx\right) \\ \lesssim\ & |\widehat{s_k}(n)|O(\varepsilon). \end{aligned}$$

Hence, it holds that

$$I_{kn} = \left(\widehat{s_k}(n)\alpha_k(b)e^{2\pi inN\phi_k(b)}\widehat{w}\left((a-nN\phi_k'(b))a^{-s}\right) + |\widehat{s_k}(n)|O(\varepsilon)\right), \tag{9}$$

if $a \in [\frac{nN}{2M}, 2MnN]$ and $N \gtrsim \max\{\varepsilon^{-1/s}, \varepsilon^{-1/(2s-1)}\} = \varepsilon^{-1/(2s-1)}$.

In sum, by (8) and (9), we arrive at

$$\begin{aligned} W_f(a,b) &= a^{-s/2}\left(\sum_{(k,n)\in\Omega_a} I_{kn} + \sum_{(k,n)\notin\Omega_a} I_{kn}\right) \\ &= |a|^{-s/2}\left(\sum_{(k,n)\in\Omega_a}\widehat{s_k}(n)\alpha_k(b)e^{2\pi inN\phi_k(b)}\widehat{w}\left(\left(a-nN\phi_k'(b)\right)|a|^{-s}\right) + O(\varepsilon)\right), \end{aligned}$$

if $N \gtrsim \max\{\varepsilon^{\frac{-1}{(1-s)r}}, \varepsilon^{\frac{-1}{2s-1}}\}$.

Similar argument can prove the above conclusion for $a < -1$ and it is simple to generalize it for different N_k to complete the proof. □

The next lemma is to estimate $\partial_b W_f(a,b)$ when $\Omega_a = \{(k,n) : a \in [\frac{nN_k}{2M}, 2MnN_k]\}$ is not empty, i.e., when $W_f(a,b)$ is relevant.

Lemma 3.6. *Suppose $\Omega_a = \{(k,n) : a \in [\frac{nN_k}{2M}, 2MnN_k]\}$ is not empty. Under the assumption of Theorem 3.4, for any $\varepsilon > 0$, we have*

$$\begin{aligned} & \partial_b W_f(a,b) \\ =\ & |a|^{-s/2}\left(\sum_{(k,n)\in\Omega_a} 2\pi inN_k\widehat{s_k}(n)\alpha_k(b)\phi_k'(b)e^{2\pi inN_k\phi_k(b)}\widehat{w}\left(\left(a-nN_k\phi_k'(b)\right)|a|^{-s}\right) + |a|O(\varepsilon)\right), \end{aligned}$$

when N is sufficiently large.

Proof. Similar to the proof of Lemma 3.5, we can assume $N_k = N$ for all k and only need to prove the case when $a > 1$. By the definition of the wave packet transform, we have

$$\begin{aligned}\partial_b W_f(a,b) &= \sum_{k=1}^{K} 2\pi i a^{1+s/2} \int_{\mathbb{R}} \alpha_k(t) s_k(2\pi N \phi_k(t)) w(a^s(t-b)) e^{-2\pi i (t-b)a} dt \\ &\quad - \sum_{k=1}^{K} a^{3s/2} \int_{\mathbb{R}} \alpha_k(t) s_k(2\pi N \phi_k(t)) w'(a^s(t-b)) e^{-2\pi i (t-b)a} dt.\end{aligned}$$

Denote the first term by T_1 and the second term by T_2. By a similar discussion in the proof of Lemma 3.5, we have the following asymptotic estimates when N is sufficiently large.

$$\begin{aligned}T_2 &= -a^{s/2} \sum_{k=1}^{K} \sum_{|n|\geq 1} \widehat{s_k}(n) \int_{\mathbb{R}} \alpha_k(a^{-s}x+b) w'(x) e^{2\pi i (nN\phi_k(a^{-s}x+b) - a^{1-s}x)} dx \\ &= -a^{s/2} \sum_{(k,n)\in\Omega_a} \int_{\mathbb{R}} \widehat{s_k}(n) \alpha_k(a^{-s}x+b) w'(x) e^{2\pi i (nN\phi_k(a^{-s}x+b) - a^{1-s}x)} dx + a^{s/2} O(\varepsilon) \\ &= a^{s/2} \sum_{(k,n)\in\Omega_a} \int_{\mathbb{R}} \widehat{s_k}(n) w(x) \alpha_k(a^{-s}x+b) e^{2\pi i (nN\phi_k(a^{-s}x+b) - a^{1-s}x)} \\ &\quad \left(2\pi i n N \phi_k'(a^{-s}x+b) a^{-s} - 2\pi i a^{1-s}\right) dx + a^{-s/2} \sum_{(k,n)\in\Omega_a} \int_{\mathbb{R}} \widehat{s_k}(n) w(x) \\ &\quad \alpha_k'(a^{-s}x+b) e^{2\pi i (nN\phi_k(a^{-s}x+b) - a^{1-s})} dx + a^{s/2} O(\varepsilon) \\ &= a^{-s/2} \sum_{(k,n)\in\Omega_a} 2\pi i n N \int_{\mathbb{R}} \widehat{s_k}(n) \phi_k'(a^{-s}x+b) \alpha_k(a^{-s}x+b) w(x) e^{2\pi i (nN\phi_k(a^{-s}x+b) - a^{1-s}x)} dx \\ &\quad - a^{1-s/2} \sum_{(k,n)\in\Omega_a} 2\pi i \int_{\mathbb{R}} \widehat{s_k}(n) w(x) \alpha_k(a^{-s}x+b) e^{2\pi i (nN\phi_k(a^{-s}x+b) - a^{1-s}x)} dx \\ &\quad + a^{-s/2} O(1) + a^{s/2} O(\varepsilon) \\ &= a^{-s/2} \sum_{(k,n)\in\Omega_a} 2\pi i n N \left(\widehat{s_k}(n) \phi_k'(b) \alpha_k(b) e^{2\pi i n N \phi_k(b)} \widehat{w}(a^{-s}(a - nN\phi_k'(b))) + |\widehat{s_k}(n)| O(\varepsilon) \right) \\ &\quad - a^{1+s/2} \sum_{(k,n)\in\Omega_a} \widehat{s_k}(n) 2\pi i \int_{\mathbb{R}} \alpha_k(t) w(a^s(t-b)) e^{2\pi i (nN\phi_k(t) - (t-b)a)} dt \\ &\quad + a^{-s/2} O(1) + a^{s/2} O(\varepsilon),\end{aligned}$$

if $N \gtrsim \max\{\varepsilon^{\frac{-1}{(1-s)r}}, \varepsilon^{\frac{-1}{2s-1}}\}$. The third equality holds by integration by parts and the last equality holds by changing variables. Notice that

$$\begin{aligned}T_1 &= a^{1+s/2} \sum_{(k,n)\in\Omega_a} \widehat{s_k}(n) 2\pi i \int_{\mathbb{R}} \alpha_k(t) w(a^s(t-b)) e^{2\pi i (nN\phi_k(t) - (t-b)a)} dt \\ &\quad + \sum_{(k,n)\notin\Omega_a} 2\pi i a^{1-s/2} \widehat{s_k}(n) \int_{\mathbb{R}} \alpha_k(a^{-s}x+b) w(x) e^{2\pi i (nN\phi_k(a^{-s}x+b) - a^{1-s}x)} dx \\ &= a^{1+s/2} \sum_{(k,n)\in\Omega_a} \widehat{s_k}(n) 2\pi i \int_{\mathbb{R}} \alpha_k(t) w(a^s(t-b)) e^{2\pi i (nN\phi_k(t) - (t-b)a)} dt + a^{1-s/2} O(\varepsilon),\end{aligned}$$

if $N \gtrsim \varepsilon^{\frac{-1}{(1-s)r}}$ for any $r >= 1$. Hence $T_1 + T_2$ results in

$$\begin{aligned}
&\partial_b W_f(a,b) \\
= \;& a^{-s/2} \sum_{(k,n)\in\Omega_a} 2\pi i n N \left(\widehat{s_k}(n)\phi_k'(b)\alpha_k(b) e^{2\pi i n N \phi_k(b)} \widehat{w}(a^{-s}(a - nN\phi_k'(b))) + |\widehat{s_k}(n)| O(\varepsilon) \right) \\
&+ a^{-s/2} O(1) + a^{s/2} O(\varepsilon) + a^{1-s/2} O(\varepsilon) \\
= \;& |a|^{-s/2} \left(\sum_{(k,n)\in\Omega_a} 2\pi i n N \widehat{s_k}(n)\alpha_k(b)\phi_k'(b) e^{2\pi i n N \phi_k(b)} \widehat{w}\left(\left(a - nN\phi_k'(b)\right)|a|^{-s} \right) + |a| O(\varepsilon) \right),
\end{aligned}$$

if N is sufficiently large. So, the Lemma 3.6 is proved. □

We are now ready to prove Theorem 3.4 with Lemma 3.5 and Lemma 3.6.

Proof. Let us first consider (i). The well-separation condition implies that $\{Z_{n,k} : 1 \le k \le K, \widehat{s_k}(n) \neq 0\}$ are disjoint. Let (a,b) be a point in R_ε, then $|W_f(a,b)| \ge a^{-s/2}\sqrt{\varepsilon}$, which means that Ω_a is not empty and $\exists (k,n) \in \Omega_a$ such that $\widehat{w}((a - nN_k\phi_k'(b))a^{-s}) \neq 0$. Because the support of $\widehat{w}(\xi)$ is $(-d,d)$, we know $|a - nN_k\phi_k'(b)| \le a^s d$, i.e., $(a,b) \in Z_{n,k}$. Hence, $R_\varepsilon \subset \bigcup_{1\le k\le K} \bigcup_{\widehat{s_k}(n)\neq 0} Z_{n,k}$.

To show (ii), let us recall that $v_f(a,b)$ is defined as

$$v_f(a,b) = \frac{\partial_b W_f(a,b)}{2\pi i W_f(a,b)},$$

for $W_f(a,b) \neq 0$. If $(a,b) \in R_\varepsilon \bigcap Z_{n,k}$, then by Lemma 3.5

$$\begin{aligned}
W_f(a,b) &= |a|^{-s/2} \left(\sum_{(k,n)\in\Omega_a} \widehat{s_k}(n)\alpha_k(b) e^{2\pi i n N_k \phi_k(b)} \widehat{w}\left(\left(a - nN_k\phi_k'(b)\right)|a|^{-s} \right) + O(\varepsilon) \right) \\
&= |a|^{-s/2} \left(\widehat{s_k}(n)\alpha_k(b) e^{2\pi i n N_k \phi_k(b)} \widehat{w}\left(\left(a - nN_k\phi_k'(b)\right)|a|^{-s} \right) + O(\varepsilon) \right),
\end{aligned}$$

as the other terms drop out, since $\{Z_{n,k}\}$ are disjoint. Similarly, by Lemma 3.6

$$\begin{aligned}
&\partial_b W_f(a,b) \\
= \;& |a|^{-s/2} \left(2\pi i n N_k \widehat{s_k}(n)\alpha_k(b)\phi_k'(b) e^{2\pi i n N_k \phi_k(b)} \widehat{w}\left(\left(a - nN_k\phi_k'(b)\right)|a|^{-s} \right) + |a| O(\varepsilon) \right).
\end{aligned}$$

Let g denote the term $\widehat{s_k}(n)\alpha_k(b) e^{2\pi i n N_k \phi_k(b)} \widehat{w}\left(\left(a - nN_k\phi_k'(b)\right)|a|^{-s} \right)$, then

$$\begin{aligned}
v_f(a,b) &= \frac{nN_k\phi_k'(b) g + |a| O(\varepsilon)}{g + O(\varepsilon)} \\
&= \frac{nN_k\phi_k'(b)(g + O(\varepsilon))}{g + O(\varepsilon)},
\end{aligned}$$

since $a \in [\frac{nN_k}{2M}, 2MnN_k]$. Because $|W_f(a,b)| \ge a^{-s/2}\sqrt{\varepsilon}$ for $(a,b) \in R_\varepsilon$, then $|g| \gtrsim \sqrt{\varepsilon}$. Therefore

$$\frac{|v_f(a,b) - nN_k\phi_k'(b)|}{|nN_k\phi_k'(b)|} \lesssim \left| \frac{O(\varepsilon)}{g + O(\varepsilon)} \right| \lesssim \sqrt{\varepsilon}.$$

□

Theorem 3.4 shows that the instantaneous frequency information function $v_f(a,b)$ can estimate $nN_k\phi_k'(t)$ accurately for a class of superpositions of general mode functions if their phases are sufficiently steep. This guarantees the well concentration of the SS energy distribution $T_f(v,b)$ around $nN_k\phi_k'(t)$. Hence, $\psi_{kn}(b)$ defined in the introduction of the GMDWP method is an accurate estimate of $nN_k\phi_k'(b)$. Next, a curve classification method and an instantaneous frequency identification method are introduced below.

Let us reindex the functions $\{\psi_{kn} : 1 \le k \le K, |n| \ge 1\}$ by $\{\psi_j\}_{1\le j\le L}$. Our goal is to obtain K index sets $\{\Lambda_k\}_{1\le k\le K}$ such that $\{\psi_{kn}\}_{|n|\ge 1} = \{\psi_j\}_{j\in\Lambda_k}$. Because $\frac{\psi_{kn_1}}{\psi_{kn_2}} \approx \frac{n_1}{n_2}$, $\{\psi_{kn}\}_{|n|\ge 1}$ can approximately be considered as a set in a one dimensional point set in a high dimensional space. Hence, the curve classification can be considered as a subspace clustering problem studied in [25, 26]. This motivates the following method to classify $\{\psi_j\}_{1\le j\le L}$. This method is similar to the method in [25] for subspace clustering.

Algorithm 3.7. *Curve classification of* $\{\psi_j\}_{1\le j\le L}$

1: *For each pair* (k,j) *with* $k \ne j$, *let* $m_k = \|\psi_k\|_{L^\infty}$ *and* $m_j = \|\psi_j\|_{L^\infty}$. *Compute the linear regression of* $\frac{\psi_k m_j}{\psi_j m_k}$ *and its residual* r_{kj}. *Let* R *be the* $L\times L$ *matrix such that* $R_{kj} = r_{kj}$.

2: *Set up a variance parameter* σ^2 *and define a Gaussian function* $g(x) = e^{-\frac{x^2}{2\sigma^2}}$.

3: *Form the affinity graph* G *with nodes representing the* L *instantaneous frequencies and edge weights given by* $g(R_{kj}) + g(R_{jk})$.

4: *Compute the eighenvalues of the normalized Laplacian of* G *and sort them in descending order* $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_L$. *Let*

$$K = L - \underset{i=1,\cdots,L-1}{\arg\max}\,(\sigma_i - \sigma_{i+1}).$$

5: *Apply the spectral clustering method proposed in [22] and used in [36] to the affinity graph using* K *as the number of curve class to divide the index set* $\{1,\ldots,L\}$ *into* K *subsets* $\{\Lambda_k\}_{1\le k\le K}$.

6: *Associate each index set* Λ_k *with a class of curves* $\{\psi_j\}_{j\in\Lambda_k}$.

If $\{\phi_k'(t)\}_{1\le k\le K}$ are not very similar, then the residual of the linear regression of $\frac{\phi_k'(t)}{\phi_j'(t)}$ is large for $k \ne j$. By setting up a proper parameter σ, Algorithm 3.7 can classify $\{\psi_{kn} : 1 \le k \le K, |n| \ge 1\}$ accurately with high probability. In the case in which instantaneous frequencies are disturbed by noise, robust subspace clustering techniques in [26] can be applied.

Algorithm 3.7 results in K classes of curves $\{\psi_j\}_{j\in\Lambda_k} = \{\psi_{kn}\}_{|n|\ge 1} \approx \{nN_k\phi_k'\}_{|n|\ge 1}$ for $1 \le k \le K$. The theorems below show how to estimate the instantaneous frequency $N_k\phi_k'(t)$ of the general mode $\alpha_k(t)s_k(2\pi N_k\phi_k(t))$.

Theorem 3.8. *Suppose* $\alpha(t)s(2\pi N_0\phi(t))$ *is a general intrinsic mode function of type* (M_0, N_0) *and* $s(t)$ *has some nonzero Fourier coefficients* $\{\widehat{s}(n_i)\}_{1\le i\le N}$ *such that* $\widehat{s}(n) = 0$ *for* $n < |n_1|$ *and* $\gcd(|n_1|,\ldots,|n_N|) = 1$. *If we know* $\psi_i(t) = n_i N_0\phi'(t)$ *for* $1 \le i \le N$ *and*

$$n_0 = \min\left\{n : 1 \le n \le M, \frac{n\psi_i(t)}{\psi_1(t)} \text{ is a constant integer for } 2 \le i \le N\right\}, \tag{10}$$

where $M = M_0 \min|\psi_1(t)|$, *then the instantaneous frequency* $N_0\phi'(t)$ *of* $\alpha(t)s(2\pi N_0\phi(t))$ *is* $\frac{|\psi_1(t)|}{n_0}$.

Proof.

$$\begin{aligned}
&\frac{n\psi_i(t)}{\psi_1(t)} \text{ is a constant integer for } 2 \le i \le N \Rightarrow \frac{nn_i}{n_1} \text{ is an integer for } 2 \le i \le N \\
\Rightarrow\ & \frac{n_1}{\gcd(n,n_1)} \Big| n_i \text{ for } 2 \le i \le N \Rightarrow \frac{n_1}{\gcd(n,n_1)} \Big| \gcd(n_2,\ldots,n_N) \\
\Rightarrow\ & \gcd(n,n_1) = n_1 \Rightarrow n = k|n_1| \text{ for some integer } k \ge 1.
\end{aligned}$$

Hence, $|n_1| = \min\left\{n : 1 \le n \le M, \frac{n\psi_i(t)}{\psi_1(t)} \text{ is a constant integer for } 2 \le i \le N\right\} = n_0$, which implies $N_0\phi'(t) = \frac{|\psi_1(t)|}{n_0}$. □

Determining whether $\frac{n\psi_i(t)}{\psi_1(t)}$ is a constant integer is not practical unless the instantaneous frequencies are exactly recovered. This motivates the design of the following method.

Theorem 3.9. *Suppose the same condition of Theorem 3.8 holds and n_0 is the solution of the following minimization problem,*

$$n_0 = \min\left(\underset{1\le n\le M}{\arg\min}\left\{\frac{1}{N-1}\sum_{i=2}^{N}\|\frac{n\psi_i(t)}{\psi_1(t)} - \lfloor\frac{n\psi_i(t)}{\psi_1(t)} + 0.5\rfloor\|_{L^2}^2\right\}\right), \tag{11}$$

where $M = M_0 \min|\psi_1(t)|$. Then the instantaneous frequency $N_0\phi'(t)$ of $s(2\pi N_0\phi(t))$ is $\frac{|\psi_1(t)|}{n_0}$.

Proof. Let $f(n) = \frac{1}{N-1}\sum_{i=2}^{N}\|\frac{n\psi_i(t)}{\psi_1(t)} - \lfloor\frac{n\psi_i(t)}{\psi_1(t)} + 0.5\rfloor\|_{L^2}^2$. Notice that

$$\frac{n\psi_i(t)}{\psi_1(t)} - \lfloor\frac{n\psi_i(t)}{\psi_1(t)} + 0.5\rfloor = \frac{nn_i}{n_1} - \lfloor\frac{nn_i}{n_1} + 0.5\rfloor.$$

If $n = k|n_1|$ for some integer $k \ge 1$, then $f(n) = 0$.

Otherwise, then by the proof of Theorem 3.8, $\exists i$, $2 \le i \le N$ such that $\frac{nn_i}{n_1}$ is not an integer. Then

$$\|\frac{n\psi_i(t)}{\psi_1(t)} - \lfloor\frac{n\psi_i(t)}{\psi_1(t)} + 0.5\rfloor\|_{L^2}^2 > 0,$$

which implies $f(n) > 0$. So, $|n_1| = \min\left(\underset{1\le n\le M}{\arg\min} f(n)\right)$. Therefore, $N_0\phi'(t) = \frac{|\psi_1(t)|}{n_0}$. □

Notice that the function $g(x) = |x - \lfloor x + 0.5\rfloor|$ is absolutely continuous. If $\psi_i(t) \approx n_i N_0\phi'(t)$ is sufficiently accurate, the conclusion of Theorem 3.9 is still true.

For each k, Theorem 3.9 estimates the instantaneous frequency of the general mode $\alpha_k(t)s_k(2\pi N_k\phi_k(t))$ using the result of Theorem 3.4 and Algorithm 3.7. This completes the estimates of the instantaneous frequencies of the general modes from a superposition of the form (2).

3.3 The analysis of spectral resolution

In Theorem 3.4, the lower bound $s > 1/2$ ensures that the wave packets is sufficiently localized in space so that it can reflect the second order properties of the phase functions precisely. The upper bound $s < 1$ enables the SSWPT to detect a more general class of shape functions compared to the wave shape functions defined in [30]. An intuitive

reason for this more general result is that the supports of wave packets in the Fourier domain are more localized than those of wavelets, resulting in a better resolution for mode decompositions. In what follows, the single scale resolution and the multiscale resolution of SS transforms will be defined and studied.

Definition 3.10. *The single scale resolution at a level N of a SS transform is $\frac{1}{N\lambda_0} - \frac{1}{N}$, where $\lambda_0 \in (0,1)$ is the critical number such that $\forall \lambda \in (0, \lambda_0)$, the SS transform is able to distinguish two modes $f_1(t) = e^{2\pi i N t}$ and $f_2(t) = e^{2\pi i \lambda N t}$ from their superposition $f(t) = f_1(t) + f_2(t)$.*

The single scale resolution analysis is related to the beating phenomenon of the EMD method in [24]. In [32], the authors have proved a conclusion which is equivalent to the fact that the single scale resolution at the level 1 of the SSWT with a mother wavelet supported in an interval of size $2d$ is $\frac{2d}{1-d}$. As we shall prove below, the SSWPT has a higher single scale resolution than the SSWT and a smaller geometric scale parameter s benefits a higher resolution. This means that the SSWPT has a better ability to distinguish two harmonics with close frequencies.

Recall that the wave packet transform is controlled by the geometric parameter s and the parameter d for the size of the support of the mother wave packet in the Fourier domain. Consider two complex harmonics $f_1(t) = e^{2\pi i N t}$, $f_2(t) = e^{2\pi i \lambda N t}$, and their superposition $f(t) = f_1(t) + f_2(t)$. Then the wave packet transform of f is

$$W_f(a,b) = |a|^{-s/2} e^{2\pi i N b} \widehat{w}(|a|^{-s}(a-N)) + |a|^{-s/2} e^{2\pi i N \lambda b} \widehat{w}(|a|^{-s}(a - N\lambda)),$$

and the instantaneous frequency information function is

$$v_f(a,b) = \frac{N\left(e^{2\pi i N b}\widehat{w}(|a|^{-s}(a-N)) + \lambda e^{2\pi i N\lambda b}\widehat{w}(|a|^{-s}(a-N\lambda))\right)}{e^{2\pi i N b}\widehat{w}(|a|^{-s}(a-N)) + e^{2\pi i N\lambda b}\widehat{w}(|a|^{-s}(a-N\lambda))}.$$

The necessary and sufficient condition of an exact decomposition by the SSWPT is $v_f(a,b) = N$ in $Z_{f_1} = \{(a,b) : W_{f_1}(a,b) \neq 0\}$ and $v_f(a,b) = N\lambda$ in $Z_{f_2} = \{(a,b) : W_{f_2}(a,b) \neq 0\}$. This is equivalent to say Z_{f_1} and Z_{f_2} are disjoint. Since

$$W_{f_1}(a,b) = |a|^{-s/2} e^{2\pi i N b} \widehat{w}(|a|^{-s}(a-N)),$$

$$W_{f_2}(a,b) = |a|^{-s/2} e^{2\pi i N \lambda b} \widehat{w}(|a|^{-s}(a-N\lambda)),$$

and the support of $\widehat{w}(\xi)$ is $(-d,d)$, the condition for Z_{f_1} and Z_{f_2} being disjoint is that the supports of the wave packets $\widehat{w_{N\lambda b}}$ and $\widehat{w_{Nb}}$ are disjoint for all b, i.e., $\widehat{w}(|N\lambda|^{-s}(x-N\lambda))$ and $\widehat{w}(|N|^{-s}(x-N))$ have non-overlapping supports. So, the critical number λ_0 of the SSWPT with a geometric scaling parameter s is the solution of the following equations:

$$\begin{cases} N - a_1 &= d|a_1|^s, \\ a_2 - N\lambda_0 &= d|a_2|^s, \\ a_1 &= a_2, \end{cases}$$

where N and d are known. When $s = 1$, the solution is $\lambda_0 = \frac{1-d}{1+d}$, which accords with the result in [32]. Let $a_1 = a_2 = a$, then we have $\lambda_0 = \frac{2a-N}{N}$. Notice that λ_0 is increasing, when s is decreasing. Therefore, a smaller s benefits a higher single scale resolution, especially for high frequency signals.

For the mode decomposition problems of the form (1), the single scale resolution is enough to quantify the resolution of a certain SS transform. However, for the general mode

decomposition problems of the form (2), each general mode $\alpha_k(t)s_k(2\pi N_k\phi_k(t))$ would result in multiple instantaneous frequencies $nN_k\phi_k'(t)$, i.e., a superposition of infinitely many terms $\widehat{s_k}(n)\alpha_k(t)e^{2\pi i n N_k\phi_k(t)}$. It is important to know how many multiple instantaneous frequencies can be identified by the SS transform.

Definition 3.11. *The multiscale resolution at a level N of a certain SS transform is $\frac{1}{Nk} - \frac{1}{N(k+1)}$, where k is the largest number of the components such that the SS transform can distinguish all the components in*

$$f(t) = \sum_{n=1}^{k} e^{2\pi i n N t}.$$

By definition, a SS transform with a smaller multiscale resolution can distinguish more Fourier expansion terms $\widehat{s_k}(n)\alpha_k(t)e^{2\pi i n N_k\phi_k(t)}$, so that one would obtain a better recovery of $\alpha_k(t)s_k(2\pi N_k\phi_k(t))$ by combining these recovered Fourier expansion terms.

For a superposition

$$f(t) = \sum_{n=1}^{K} e^{2\pi i n N t} \tag{12}$$

with K arbitrarily large, the wave packet transform is

$$W_f(a,b) = \sum_{n=1}^{K} |a|^{-s/2} e^{2\pi i n N b}\widehat{w}\left(|a|^{-2}(a-nN)\right)$$

and the instantaneous frequency information function is

$$v_f(a,b) = \frac{N\sum_{n=1}^{K} n e^{2\pi i n N b}\widehat{w}\left(|a|^{-s}(a-nN)\right)}{\sum_{n=1}^{K} e^{2\pi i n N b}\widehat{w}\left(|a|^{-s}(a-nN)\right)}.$$

Each term $\widehat{w}\left(|a|^{-s}(a-nN)\right)$ is supported in $Z_n = \{a : |a-nN| < d|a|^s\}$ centered at $a = nN$. Similar to previous discussions, an exact recovery of the nth term is equivalent to $Z_{n-1}\bigcap Z_n = \emptyset$ and $Z_n\bigcap Z_{n+1} = \emptyset$. Since the size of the interval Z_n is monotonously increasing as n increases and the space between their centers is fixed, one should identify n_0, the greatest n such that $Z_n\bigcap Z_{n+1} = \emptyset$, i.e., $n_0 = \max\{n \in \mathbb{Z}^+ : nN < a_0\}$, where $a_0 = \left(\frac{N}{2d}\right)^{1/s}$ is a solution of the following equations,

$$\begin{cases} (n+1)N - a &= d|a|^s, \\ a - nN &= d|a|^s. \end{cases}$$

Hence, $n_0 = \lfloor \frac{N^{\frac{1}{s}-1}}{(2d)^{1/s}} \rfloor$. Therefore, the multiscale resolution at the level N of the SSWPT with a scaling parameter s and a mother wave packet supported in $(-d,d)$ is

$$\frac{1}{(n_0-1)N} - \frac{1}{n_0 N} = \frac{1}{N(\lfloor \frac{N^{\frac{1}{s}-1}}{(2d)^{1/s}} \rfloor - 1)\lfloor \frac{N^{\frac{1}{s}-1}}{(2d)^{1/s}} \rfloor} \approx O(N^{1-2/s}).$$

Notice that the SSWT (i.e., $s = 1$) can only distinguish $O(1)$ terms in (12). This limits its application to general mode decompositions of the form (2). However, the SSWPT is able to identify $O(N^{\frac{1}{s}-1})$ terms exactly in (12). This motivates its application to the general mode decomposition problems of the form (2). Concrete examples will be presented to support this argument in Section 5.

4 Theory for the DSA

The analysis of the DSA method essentially consists of two main theorems. Theorem 4.2 proves that the SSWPT is able to provide accurate input instantaneous frequency estimates if the weak well-separation condition defined below holds. Then Theorem 4.5 proves that Step 4, the key idea of the DSA method, can provide precise spectral analysis for general shape functions, if their corresponding phase functions are well-different and steep enough. We omit the proof of the other steps in the DSA method to save space.

Definition 4.1. *A function $f(t)$ is a weak well-separated general superposition of type (M, N, K, s) if*

$$f(t) = \sum_{k=1}^{K} f_k(t)$$

where each $f_k(t) = \alpha_k(t) s_k(2\pi N_k \phi_k(t))$ is a GIMT of type (M, N_k) such that $N_k \geq N$ and the phase functions satisfy the following weak well-separation conditions.

1. *Suppose*
$$Z_{nk} = \left\{(a, b) : |a - nN_k \phi_k'(b)| \leq d|a|^s\right\}.$$
For each k, there exists n_k such that $\widehat{s_k}(n_k) \neq 0$ and $Z_{n_k k} \cap Z_{nj} = \emptyset$ for all pairs $(n, j) \neq (n_k, k)$ and $\widehat{s_j}(n) \neq 0$.

2. *$\exists K_0 < \infty$ such that $\forall a \in \mathbb{R}$ and $\forall b \in \mathbb{R}$ there exists at most K_0 pairs of (n, k) such that $(a, b) \in Z_{nk}$.*

We denote by $wGF(M, N, K_0, K, s)$ the set of all such functions. Note that $wGF(M, N, 1, K, s) = GF(M, N, K, s)$.

In what follows, when we write $O(\cdot)$, $\lesssim$, or $\gtrsim$, the implicit constants may depend on M, K and K_0.

Theorem 4.2. *For a function $f(t)$ and $\varepsilon > 0$, we define*

$$R_\varepsilon = \{(a, b) : |W_f(a, b)| \geq |a|^{-s/2}\sqrt{\varepsilon}\}$$

and

$$Z_{n,k} = \{(a, b) : |a - nN_k \phi_k'(b)| \leq d|a|^s\}$$

for $1 \leq k \leq K$ and $|n| \geq 1$. For fixed M, K_0, K and $\forall \varepsilon > 0$, there exists a constant $N_0(M, K_0, K, s, \varepsilon) > 0$ such that $\forall N > N_0$ and $f(t) \in wGF(M, N, K_0, K, s)$ the following statements hold.

(i) For each j, there exists n_j such that $\widehat{s_j}(n_j) \neq 0$ and $Z_{n_j j} \cap Z_{nk} = \emptyset$ for all pairs $(n, k) \neq (n_j, j)$ and $\widehat{s_k}(n) \neq 0$;

(ii) For any $(a, b) \in R_\varepsilon \cap Z_{n_j,j}$,

$$\frac{|v_f(a, b) - n_j N_j \phi_j'(b)|}{|n_j N_j \phi_j'(b)|} \lesssim \sqrt{\varepsilon}.$$

(iii) For each j, let

$$l_{n_j}(b) = \min\left\{a : (a,b) \in R_\epsilon \cap Z_{n_j j}\right\}, \quad u_{n_j}(b) = \max\left\{a : (a,b) \in R_\epsilon \cup Z_{n_j j}\right\}.$$

Suppose $v_f(a,b) \neq \infty$. If $a \leq l_{n_j}(b)$, then $v_f(a,b) \leq l_{n_j}(b)(1+O(\sqrt{\epsilon}))$. If $a \geq u_{n_j}(b)$, then $v_f(a,b) \geq u_{n_j}(b)(1-O(\sqrt{\epsilon}))$.

Proof. The weak well-separation condition implies (i). (ii) is true by the same argument of Theorem 3.4 (ii). We only need to prove (iii). Let us recall that

$$\Omega_a = \{(k,n) : a \in [\frac{nN_k}{2M}, 2MnN_k]\}.$$

By Lemma 3.5

$$W_f(a,b) \quad = \quad |a|^{-s/2}\left(\sum_{(k,n)\in\Omega_a} \widehat{s_k}(n)\alpha_k(b)e^{2\pi i n N_k \phi_k(b)}\widehat{w}\left(\left(a - nN_k\phi_k'(b)\right)|a|^{-s}\right) + O(\varepsilon)\right),$$

as the other terms drop out. Similarly, by Lemma 3.6

$$\begin{aligned} &\partial_b W_f(a,b) \\ = \quad & |a|^{-s/2}\left(\sum_{(k,n)\in\Omega_a} 2\pi i n N_k \widehat{s_k}(n)\alpha_k(b)\phi_k'(b)e^{2\pi i n N_k \phi_k(b)}\widehat{w}\left(\left(a - nN_k\phi_k'(b)\right)|a|^{-s}\right) + |a|O(\varepsilon)\right). \end{aligned}$$

Let $g_{n,k}$ denote the term $\widehat{s_k}(n)\alpha_k(b)e^{2\pi i n N_k \phi_k(b)}\widehat{w}\left((a - nN_k\phi_k'(b))|a|^{-s}\right)$, then

$$\begin{aligned} \frac{v_f(a,b)-a}{a} &= \frac{\partial_b W_f(a,b) - 2\pi i a W_f(a,b)}{2\pi i a W_f(a,b)} \\ &= \frac{\sum_{(k,n)\in\Omega_a}(nN_k\phi_k'(b) - a)g_{n,k} + |a|O(\varepsilon)}{a|a|^{s/2}W_f(a,b)}, \end{aligned}$$

since $a \in [\frac{nN_k}{2M}, 2MnN_k]$ for $(k,n) \in \Omega_a$. Because $f(t) \in wGF(M,N,K_0,K,s)$, the number of $g_{n,k}$ not vanishing is at most K_0. Because $|W_f(a,b)| \geq |a|^{-s/2}\sqrt{\varepsilon}$ for $(a,b) \in R_\varepsilon$, $|a - nN_k\phi_k'(b)| \leq |a|^s d$ for $g_{n,k}$ not vanishing, and $|g_{n,k}| \lesssim 1$, then

$$\left|\frac{v_f(a,b)-a}{a}\right| \lesssim \frac{K_0|a|^s d + |a|O(\epsilon)}{|a|\sqrt{\epsilon}} \lesssim \sqrt{\epsilon},$$

if $|a| \gtrsim N \gtrsim \epsilon^{\frac{-1}{1-s}}$. Therefore, if $a \leq l_{n_j}(b)$, then

$$v_f(a,b) \leq a(1+O(\sqrt{\epsilon})) \leq l_{n_j}(b)(1+O(\sqrt{\epsilon}))$$

for N sufficiently large. If $a \geq u_{n_j}(b)$, then

$$v_f(a,b) \geq a(1-O(\sqrt{\epsilon})) \geq u_{n_j}(b)(1-O(\sqrt{\epsilon}))$$

for N large enough. □

Theorem 4.2 (ii) shows that $v_f(a,b)$ can estimate instantaneous frequencies $\{n_jN_j\phi_j'(t)\}_{j=1}^K$ accurately so that the energy of $\widehat{s_j}(n_j)\alpha_j(t)e^{2\pi i n_j N_j \phi_j(t)}$ is squeezed to sharpened areas around $n_jN_j\phi_j'(t)$. Theorem 4.2 (iii) implies that the SS energy distribution $T_f(a,b)$ has well-separated and sharp supports around $\{n_jN_j\phi_j'(t)\}_{j=1}^K$, each of which only corresponds to $\widehat{s_j}(n_j)\alpha_j(t)e^{2\pi i n_j N_j \phi_j(t)}$. This guarantees the accurate estimate of $n_jN_j\phi_j'(t)$ and the precise extraction of $\widehat{s_j}(n_j)\alpha_j(t)e^{2\pi i n_j N_j \phi_j(t)}$.

Next, Theorem 4.5 below shows that the DSA method with exact estimates of the instantaneous frequencies is able to provide accurate spectral analysis of the general shape functions, if the phase functions are well-different and steep sufficiently.

Since $f(t)$ is defined in $\mathbb{R}$ with non-vanishing amplitudes, we consider the following short-time Fourier transform with real-valued, non-negative and smooth window function $w_1(t)$ compactly supported in $(-1,1)$ such that $|\widehat{w_1}|$ has a sheer peak around the origin and rapidly decays elsewhere.

Definition 4.3. *Given the window function $w_1(t)$ and a parameter $T>1$, the short-time Fourier transform of a function $f(t)$ with a parameter T is a function*

$$\mathcal{F}_T(f)(a,b)=\int_{\mathbb{R}} f(t)w_T(t-b)e^{-2\pi i a t}dt$$

for $a,b\in\mathbb{R}$, where $w_T(t)=w_1(t/T)$ and $\mathcal{F}_T$ denote the short-time Fourier transform operator with the parameter T.

Definition 4.4. *For $M>0$ and $K>0$, the phase functions $\{\phi_k(t)\}_{1\le k\le K}$ are well-different of type (M,K) at $b\in\mathbb{R}$, if they satisfy the following conditions.*

1. *For any $T>0$, the number of extrema of $\phi_k\circ\phi_j^{-1}(t)$ in $(b-T,b+T)$ is at most TM for $k\neq j$.*
2. *For any $T>0$ there exists $\eta_0>0$, $\eta_1>0$ and $N_0(M,K,T,b)$ such that $\forall a\in(\frac{1}{2M^2},2M^2)$ and $\forall N>N_0(M,K,T,b)$*

$$\lambda^*\left(\left\{t:\left|\partial_t\left(\phi_k(\phi_j^{-1}(t))\right)-a\right|\le\frac{1}{N^{1-\eta_0}}\right\}\cap\{t:b-T\le t\le b+T\}\right)\lesssim O(\frac{1}{N^{\eta_1}})$$

for $k\neq j$, where $\lambda^*(\cdot)$ denotes the Lebesgue measure and $\lesssim$ means the implicit constant may depend on M, K, T and b.

The definition of well-different phase functions is crucial to general mode decompositions. The difference of phase functions is the key feature for grouping the Fourier expansion terms of the general modes. If two phase functions are similar, their corresponding general modes would have similar evolution patterns. It is reasonable to combine them as one general mode. On the other hand, the well-difference of phase functions guarantees that the key idea of the DSA method can provide accurate spectral information of general shape functions, as proved in the following theorem.

Theorem 4.5. *Suppose $f(t)=\sum_{k=1}^K f_k(t)$, where $f_k(t)=\alpha_k(t)s_k(2\pi N_k\phi_k(t))$ is a GIMT of type (M,N_k) with $N_k\ge N$ and the phase functions $\{\phi_k(t)\}_{1\le k\le K}$ are well-different of type (M,K) at b. Let $s_0=\max_{(k,n)}|\widehat{s_k}(n)|$. Define*

$$h_k(t)=\frac{f\circ\phi_k^{-1}(t)}{\alpha_k\circ\phi_k^{-1}(t)}$$

for $1 \le k \le K$. *For fixed* M, K, b, s_0 *and* $\delta > 0$, $\exists T_0(M, K, s_0, \delta, b)$, $\forall T > T_0$, $\exists N_0(M, K, s_0, T, b) > 0$ *such that* $\forall N > N_0$ *the solution of the following optimization problem*

$$(a_0, k_0) = \arg\max_{(a,k)} |\mathcal{F}_T(h_k)(a, b)|$$

satisfies $|a_0 - nN_{k_0}| < \delta$ *for some* n *such that* $\widehat{s_{k_0}}(n) \neq 0$.

In what follows, when we write $O(\cdot)$, $\lesssim$, or $\gtrsim$, the implicit constants may depend on M, K, T and b.

Proof. Notice that

$$h_k(t) = \frac{f \circ \phi_k^{-1}(t)}{\alpha_k \circ \phi_k^{-1}(t)} = \sum_{n=-\infty}^{\infty} \widehat{s_k}(n) e^{2\pi i n N_k t} + \sum_{j \neq k} \sum_{n=-\infty}^{\infty} \widehat{s_j}(n) \frac{\alpha_j \circ \phi_k^{-1}(t)}{\alpha_k \circ \phi_k^{-1}(t)} e^{2\pi i n N_j \phi_j \circ \phi_k^{-1}(t)},$$

then

$$\begin{aligned} \mathcal{F}_T(h_k)(a, b) &= \sum_{n=-\infty}^{\infty} \widehat{s_k}(n) \int_{\mathbb{R}} w_T(t-b) e^{2\pi i (nN_k - a)t} dt \\ &\quad + \sum_{j \neq k} \sum_{n=-\infty}^{\infty} \widehat{s_j}(n) \int_{\mathbb{R}} \frac{\alpha_j \circ \phi_k^{-1}(t)}{\alpha_k \circ \phi_k^{-1}(t)} w_T(t-b) e^{2\pi i (nN_j \phi_j \circ \phi_k^{-1}(t) - at)} dt \end{aligned}$$

by the uniform convergence of the Fourier series of $s_k(t)$. The first part of $\mathcal{F}_T(h_k)(a, b)$ is

$$\begin{aligned} I_1(a, k) &= \sum_{n=-\infty}^{\infty} \widehat{s_k}(n) \int_{\mathbb{R}} w_T(t-b) e^{2\pi i (nN_k - a)t} dt \\ &= \sum_{n=-\infty}^{\infty} T \widehat{s_k}(n) e^{2\pi i b (nN_k - a)} \int_{\mathbb{R}} w_1(x) e^{2\pi i T (nN_k - a)x} dx \\ &= \sum_{n=-\infty}^{\infty} T \widehat{s_k}(n) e^{2\pi i b (nN_k - a)} \widehat{w_1}\left(T(a - nN_k)\right). \end{aligned}$$

Hence, $\exists T_0(M, K, s_0, \delta, b)$ such that, if $T > T_0$, then $|I_1(a, k)|$ has well-separated sheer energy peaks at $a = nN_k$ of order $T \left|\widehat{s_k(n)}\right|$ and $|I_1(a, k)| < \frac{T s_0}{3}$ if $|a - nN_k| \ge \delta$ for all n. The estimate of the second part

$$I_2(a, k) = \sum_{j \neq k} \sum_{n=-\infty}^{\infty} \widehat{s_j}(n) \int_{\mathbb{R}} \frac{\alpha_j \circ \phi_k^{-1}(t)}{\alpha_k \circ \phi_k^{-1}(t)} w_T(t-b) e^{2\pi i (nN_j \phi_j \circ \phi_k^{-1}(t) - at)} dt$$

relies on the estimate of each term

$$I_{jn} = \widehat{s_j}(n) \int_{\mathbb{R}} \frac{\alpha_j \circ \phi_k^{-1}(t)}{\alpha_k \circ \phi_k^{-1}(t)} w_T(t-b) e^{2\pi i (nN_j \phi_j \circ \phi_k^{-1}(t) - at)} dt.$$

Notice that $\frac{\alpha_j \circ \phi_k^{-1}(t)}{\alpha_k \circ \phi_k^{-1}(t)} w_T(t-b)$ and $2\pi(nN_j \phi_j \circ \phi_k^{-1}(t) - at)$ are real smooth functions and $w_T(t-b)$ has a compact support in $(b - T, b + T)$. If $\partial_t(nN_j \phi_j \circ \phi_k^{-1}(t) - at) \neq 0$ in $(b - T, b + T)$, a similar argument of the integration by parts in Lemma 3.5 shows that

$$|I_{jn}| \lesssim |\widehat{s_j}(n)| \frac{1}{\left|nN_j \partial_t(\phi_j \circ \phi_k^{-1})(t) - a\right|}.$$

Therefore, the order of $|I_{jn}|$ is determined by points t such that $\left|nN_j\partial_t(\phi_j\circ\phi_k^{-1})(t)-a\right|$ is vanishing or relatively small.

If $a\notin(\frac{nN_j}{2M^2},2nN_jM^2)$, then by the fact that $\partial_t(\phi_j\circ\phi_k^{-1})(t)\in[\frac{1}{M^2},M^2]$, we have $\left|nN_j\partial_t(\phi_j\circ\phi_k^{-1})(t)-a\right|\gtrsim nN_j$, which implies

$$|I_{jn}|\lesssim\frac{|\widehat{s_j}(n)|}{nN_j}\lesssim\frac{1}{N}. \tag{13}$$

If $a\in(\frac{nN_j}{2M^2},2nN_jM^2)$, then $\frac{a}{nN_j}\in(\frac{1}{2M^2},2M^2)$. Let

$$A=\left\{t:\left|\partial_t\left(\phi_j\circ\phi_k^{-1}(t)\right)-\frac{a}{nN_j}\right|\le\frac{1}{(nN_j)^{1-\eta_0}}\right\}\cap\{t:b-T\le t\le b+T\}.$$

Because the phase functions are well-different of type (M,K) at b, for fixed T there exists $\eta_0>0$, $\eta_1>0$ and $N_1(M,K,T,b)$ such that for $\frac{a}{nN_j}\in(\frac{1}{2M^2},2M^2)$ and $nN_j>N_1(M,K,T,b)$, we have $\lambda^*(A)\lesssim O(\frac{1}{(nN_j)^{\eta_1}})$. This gives

$$\left|\widehat{s_j}(n)\int_A\frac{\alpha_j\circ\phi_k^{-1}(t)}{\alpha_k\circ\phi_k^{-1}(t)}w_T(t-b)e^{2\pi i(nN_j\phi_j\circ\phi_k^{-1}(t)-at)}dt\right|\lesssim O(\frac{|\widehat{s_j}(n)|}{(nN_j)^{\eta_1}}).$$

By the definition of well-difference of type (M,K), $(\mathbb{R}\setminus A)\cap(b-T,b+T)$ is a union of at most $O(TM)$ intervals. Hence, similar to the method of stationary phase, we have

$$\left|\widehat{s_j}(n)\int_{\mathbb{R}\setminus A}\frac{\alpha_j\circ\phi_k^{-1}(t)}{\alpha_k\circ\phi_k^{-1}(t)}w_T(t-b)e^{2\pi i(nN_j\phi_j\circ\phi_k^{-1}(t)-at)}dt\right|\lesssim O(\frac{|\widehat{s_j}(n)|}{(nN_j)^{\eta_0}}).$$

In sum,

$$\begin{aligned}|I_{jn}| &\le \left|\widehat{s_j}(n)\int_{\mathbb{R}\setminus A}\frac{\alpha_j\circ\phi_k^{-1}(t)}{\alpha_k\circ\phi_k^{-1}(t)}w_T(t-b)e^{2\pi i(nN_j\phi_j\circ\phi_k^{-1}(t)-at)}dt\right|\\ &\quad+\left|\widehat{s_j}(n)\int_{A}\frac{\alpha_j\circ\phi_k^{-1}(t)}{\alpha_k\circ\phi_k^{-1}(t)}w_T(t-b)e^{2\pi i(nN_j\phi_j\circ\phi_k^{-1}(t)-at)}dt\right|\\ &\lesssim O(\frac{|\widehat{s_j}(n)|}{(nN_j)^{\eta_1}})+O(\frac{|\widehat{s_j}(n)|}{(nN_j)^{\eta_0}}).\end{aligned}$$

Recall that $N_k\ge N$ and $\sum\limits_{n=-\infty}^{\infty}|\widehat{s_k}(n)|\le M$ for $1\le k\le K$. So, if $N>N_1(M,K,T,b)$

$$|I_2(a,k)|\lesssim\sum_{j\ne k}\sum_{n=-\infty}^{\infty}\left(O(\frac{|\widehat{s_j}(n)|}{(nN_j)^{\eta_1}})+O(\frac{|\widehat{s_j}(n)|}{(nN_j)^{\eta_0}})\right)\lesssim O(\frac{(K-1)M}{N^{\eta}})\lesssim O(\frac{1}{N^{\eta}}), \tag{14}$$

where $\eta=\min\{\eta_0,\eta_1\}$.

By (13) and (14), $\exists N_0=\max\left\{N_1(M,K,T,b),\left(\frac{3}{Ts_0}\right)^{1/\eta},\frac{3}{Ts_0}\right\}$ such that $\forall N>N_0$, we have $|I_2(a,k)|<\frac{Ts_0}{3}$.

Let Ξ_k be the index set $\{n:\widehat{s_k}(n)\ne0\}$ and $(\tilde{n},\tilde{k})=\arg\max\limits_{(n,k)}|\widehat{s_k}(n)|$. Now suppose $N>N_0$. Let $|\mathcal{F}_T(h_k)(a,b)|$ take the maximum value at the pair (a_0,k_0). If there is no $n\in\Xi_{k_0}$ such that $|a_0-nN_{k_0}|<\delta$, then

$$|\mathcal{F}_T(h_{k_0})(a_0,b)|\le|I_1(a_0,k_0)|+|I_2(a_0,k_0)|<\frac{2Ts_0}{3}.$$

However, for the pair $(\tilde{n}, \tilde{k})$, we have

$$\left|\mathcal{F}_T(h_{\tilde{k}})(\tilde{n}, b)\right| \geq \left|I_1(\tilde{n}, \tilde{k})\right| - \left|I_2(\tilde{n}, \tilde{k})\right| > Ts_0 - \frac{Ts_0}{3} > \frac{2Ts_0}{3}.$$

This conflicts with the fact that $|\mathcal{F}_T(h_k)(a, b)|$ takes the maximum value less than $\frac{2Ts_0}{3}$ at the pair (a_0, k_0). Hence, there exists $n \in \Xi_{k_0}$ satisfying that $|a_0 - nN_{k_0}| < \delta$. This completes the proof. □

In practice, the signal $f(t)$ is defined in a bounded interval, e.g., $[0, 1]$ without loss of generality. Applying the Fourier transform on $f(t)$ in $[0, 1]$ is equivalent to applying the short-time Fourier transform on $f(t)$ with a rectangle window function centered at $t = \frac{1}{2}$. In this sense, Theorem 4.5 implies that the DSA method can accurately decompose $f(t)$ into GIMTs $\{\alpha_k(t)s_k(2\pi N_k\phi_k(t))\}_{k=1}^K$ and analyzes the spectra of general shape functions $\{\alpha_k(t)\}_{k=1}^K$ by extracting the Fourier expansion terms $\widehat{s_k}(n)\alpha_k(t)e^{2\pi i n N_k \phi_k(t)}$ one by one from the one with highest energy.

5 Numerical examples

In this section, some numerical examples of synthetic and real data are provided to demonstrate the proposed properties of 1D SSWPT and the efficiency of the GMDWP method and the DSA method in fruitful applications. In all of these examples, the 1D SSWPT is implemented using a fast algorithm similar to the one in [8, 36] and the complexity is $O(L\log(L))$, where L is the number of sample points of a given signal. The mother wave packet $w(t)$ is constructed using the same method in [8] with a support parameter $d = 1$. The threshold parameter in the main theorems is $\epsilon = 10^{-6}$ and the scaling parameter s is equal to $2/3$. For the purpose of convenience, the synthetic data is defined in $[0, 1]$ and the number of samples is between 2^{13} and 2^{15}.

5.1 The comparison of multiscale resolutions

Let us start by repeating the comparison of the resolutions of the SSWPT and the SSWT, since it is a fundamental issue in general mode decomposition problems. As we shall see in the following two examples, the SSWT would mix up high frequency terms and would provide misleading high frequency information. However, the SSWPT can relieve much of this trouble. In the following two examples, the mother wavelet and the mother wave packet have the same size of supports $d = 1$ in the Fourier domain.

Example 2: Let us consider

$$f(t) = e^{200\pi i(x+5x^2)} + \sum_{n=1}^{20} e^{2\pi i n N(x+0.005*\sin(2\pi x))}$$

where $N = 100$. Figure 11 left shows the real instantaneous frequencies of all the oscillatory terms in $f(t)$. The SSWPT of $f(t)$ shown in Figure 11 middle agrees with Theorem 3.4 and 4.2 that the essential support of the SS energy distribution $T_f(v, b)$ is concentrating around isolated instantaneous frequencies, even with crossover frequencies on the scene. Recall that the SSWPT is able to identify $O(N^{1/s-1}) = O(10)$ components when $s = 2/3$. The number of clearly identified components is 20 according with the multiscale resolution analysis of the SSWPT. However, the SSWT can only identify $O(1)$ components

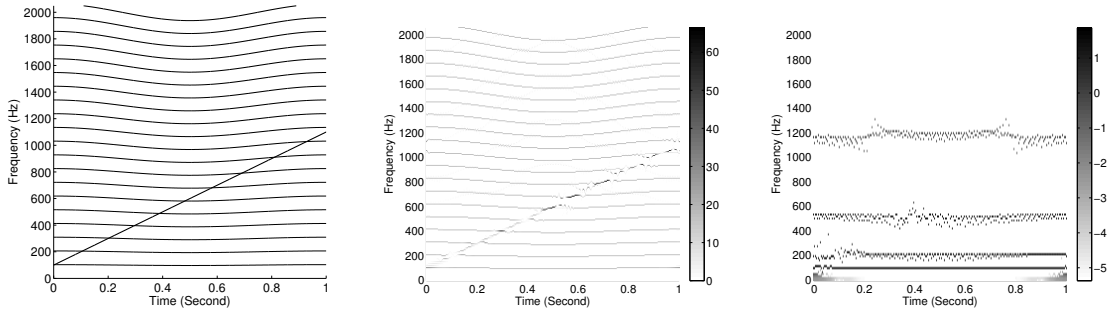

Figure 11: Left: Real instantaneous frequencies of Fourier expansion terms of $f(t)$ in Example 2. Middle: The SS energy distribution of $T_f(v,b)$ provided by SSWPT. Right: The logarithm of the SS energy distribution $\log_{10}(T_f(v,b))$ provided by SSWT.

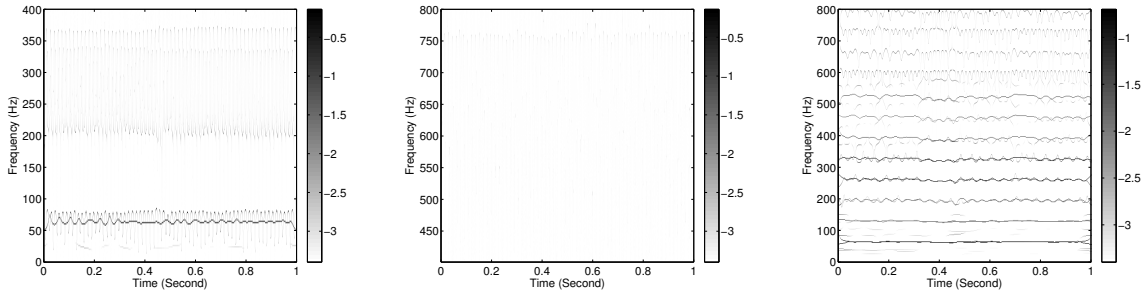

Figure 12: Left: The logarithm of the SS energy distribution $\log_{10}(T_f(v,b))$ of the ECG signal in the low frequency part using the SSWT. Middle: $\log_{10}(T_f(v,b))$ of the same ECG signal in the high frequency part using the SSWT. Right: $\log_{10}(T_f(v,b))$ of the same ECG signal provided by the SSWPT.

as shown in Subsection 3.3. This explains why the SSWT of $f(t)$ is misleading as shown in Figure 11 right. The SSWT mixes up the high frequency terms and cannot reflect the true instantaneous information of signals. Nevertheless, the high frequency information is crucial to precise reconstructions of general shape functions.

Example 3: In [6], the SSWT has been shown the ability to estimate the instantaneous frequencies of real ECG signals accurately. Figure 12 left and middle show the SS energy distribution of the SSWT of the real ECG signal in [6]. The component with the lowest frequency can reflect the instantaneous frequency of the ECG signal according with the conclusion of [6]. However, the synchrosqueezed result in high frequency part is blurry and useless due to severe interferences between high frequency components. This limits the application of the SSWT to recover spike shape functions. Fortunately, as the Figure 12 right shows, the SSWPT can identify $O(N^{1/s-1}) = O(10)$ components, which are the Fourier expansion terms of the general mode in a spike shape. This encourages the application of the SSWPT for general mode decompositions.

5.2 General mode decompositions and the robustness

As we have seen in Example 1, the GMDWP method and the DSA method can exactly recover general modes. In what follows, we would study the robustness against noise and present some more examples of general shape functions. The shapes of general modes are determined by all the Fourier expansion terms, including those weak energy terms which have been concealed by noise. We will show the recovered results in noisy cases and then present an example about denoising according to the feature of recovered modes. The noise used here is a Gaussian random noise $n(t)$ with zero mean and variance σ^2. To quantify the influence of the noise on each general mode, we introduce the following Signal-to-Noise Ratio (SNR)

$$\mathrm{SNR}[dB] = \min\left\{10\log_{10}\left(\frac{\|f_i\|_{L^2}}{\sigma^2}\right), 1 \le i \le K\right\},$$

where $\{f_i\}_{i=1}^K$ are the general modes contained in the original signal $f(t)$.

Example 1: Let us revisit Example 1 in Figure 7 and add study its noisy case,

$$f(t) = \alpha_1(t)s_1(2\pi N_1\phi_1(t)) + \alpha_2(t)s_2(2\pi N_2\phi_2(t)) + n(t).$$

Figure 13 shows two superpositions with different noise levels. As the reconstructed results show in Figure 14 and Figure 15 left, the instantaneous frequencies are accurately estimated, even if the signal is disturbed by severe noise. The essential feature of the general modes are recovered. When the noise is overwhelming the general modes, additional denoising procedure is application dependent, as we will show in the next example.

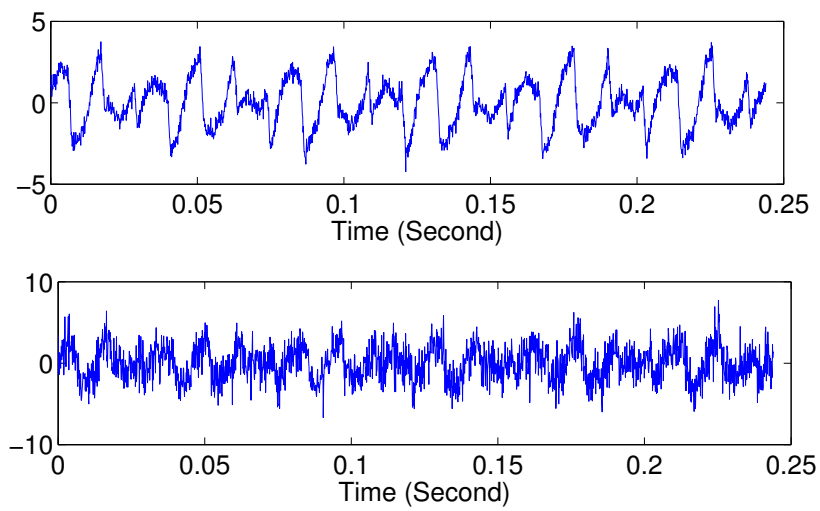

Figure 13: Noisy signals of Example 1 and their SNRs are 6 and -3, respectively.

As a comparison, the EEMD method in [34] is applied to the noisy data with $SNR = 6$ in this example. We set up the ratio of the standard deviation of the added noise and that of the original data as 0.2. The ensemble number is 50 and the expected number of modes is 5. This method is able to provide several modes in different frequency scales. However, they are not those modes expected.

Example 4**:** Combining the proposed methods with some post processing techniques can detect general shape functions in a wider class than the one defined in Definition 3.1.

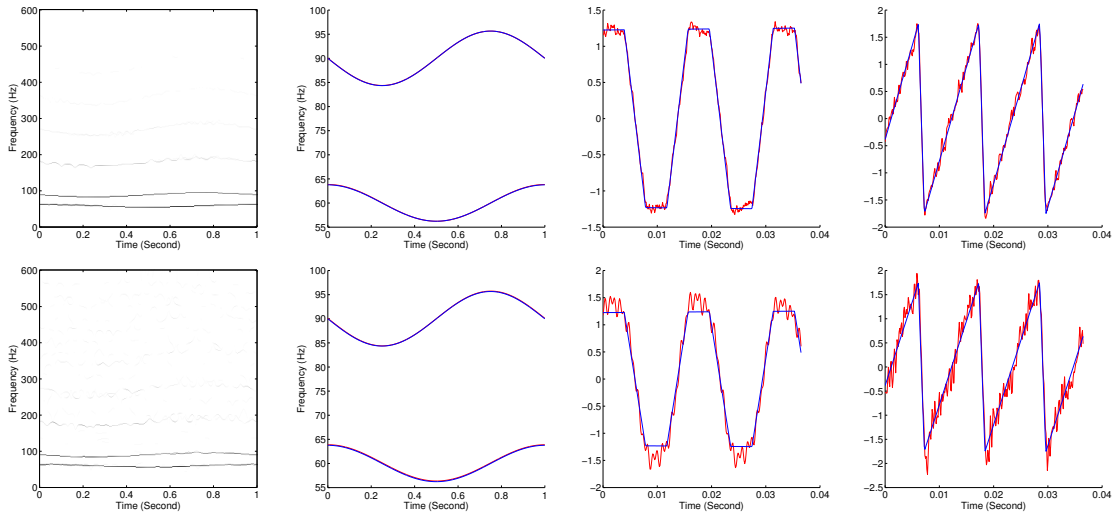

Figure 14: Noisy Example 1. Top: SNR = 6. Bottom: SNR = −3. Left: The SS energy distributions of signals. Middle left: The real instantaneous frequencies (blue) and the estimated instantaneous frequencies (red). Middle right and right: Recovered general modes.

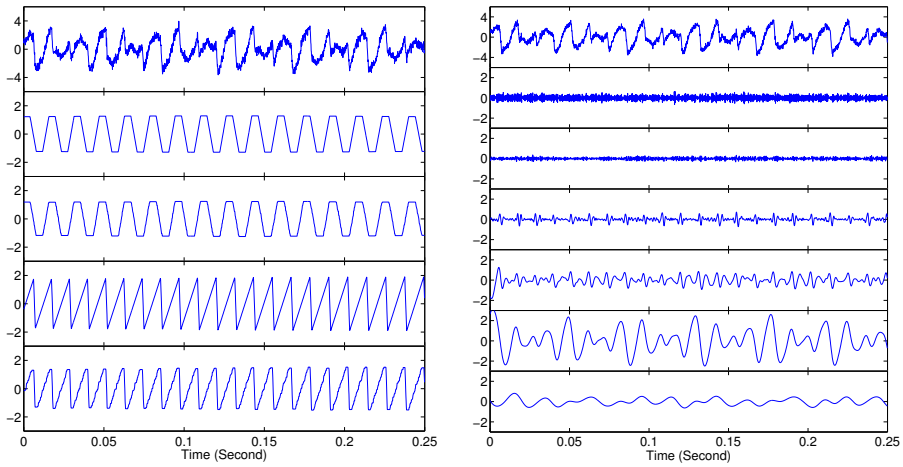

Figure 15: Left: Noisy signals of Example 1 with SNR = 6. The first row: original data. The second and the third row: the first noiseless general mode and its recovered result. The fourth and the fifth row: The second noiseless general mode and its recovered result. Right: The first row: original data. Rows below the first one: identified modes provided by the EEMD method.

For example, we test piecewise constant shape functions s_3 and s_4 as shown in Figure 16. A noisy superposition of general modes is generated as follows.

$$f(t) = \alpha_3(t)s_3(2\pi N_3\phi_3(t)) + \alpha_4(t)s_4(2\pi N_4\phi_4(t)) + n(t),$$

where $\alpha_3(t) = 1 + 0.4\sin(4\pi t)$, $\alpha_4(t) = 1 - 0.3\sin(2\pi t)$, $N_3 = 120$, $N_4 = 185$, $\phi_3(t) = t + 0.005\sin(2\pi t)$, and $\phi_4(t) = t + 0.01\cos(4\pi t)$. In this example, the SSWPT is applied

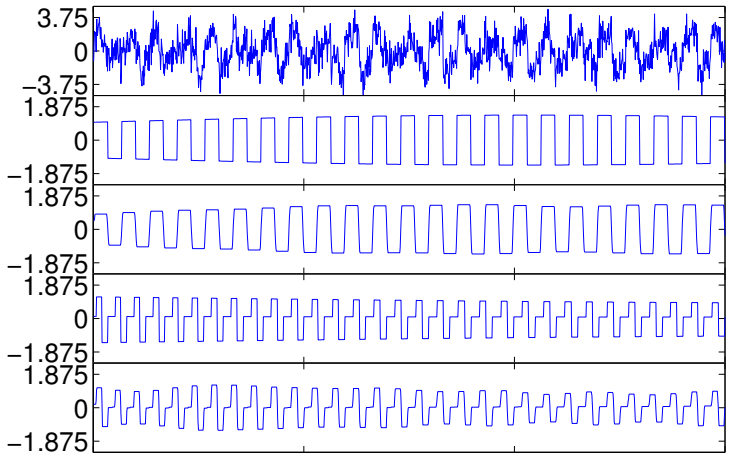

Figure 16: The first row: A noisy superposition with SNR = 0. The second and the third row: The first noiseless general mode and its recovered result. The fourth and the fifth row: The second noiseless general mode and its recovered result.

to estimate the instantaneous information first and then the DSA method is applied to decompose $f(t)$ into two general modes. Finally, a TV norm minimization is applied to obtain the final results shown in Figure 16. The DSA method is able to detect the basic feature of these general modes and the post processing TV norm minimization helps to reduce the noise.

5.3 Real applications

Example 5: In the first example of real applications, we study ECG signals. Two real ECG general shape functions $s_5(t)$ and $s_6(t)$ (see Figure 17) are cut out from real ECG signals in [12] and [30] . A noisy superposition with SNR = 0 is generated as follows:

$$f(t) = \alpha_5(t)s_5(2\pi N_5\phi_5(t)) + \alpha_6(t)s_6(2\pi N_6\phi_6(t)) + n(t),$$

where $\alpha_5(t) = 1 + 0.05\sin(2\pi t)$, $\alpha_6(t) = 1 + 0.05\cos(2\pi t)$, $N_5 = 150$, $N_6 = 220$, $\phi_5(t) = t + 0.006\sin(2\pi t)$, and $\phi_6(t) = t + 0.006\cos(2\pi t)$. As shown in Figure 17, the SS energy distribution is well concentrated around the real instantaneous frequencies and the instantaneous frequencies are accurately estimated. Most importantly, the main spikes of real ECG shape functions are precisely recovered, even if the SNR is small. The decomposition results are plotted in Figure 18 left.

As a comparison, the EEMD method with the same parameters in previous examples is also applied to the same data. As shown in Figure 18 right, the EEMD method cannot provide usefull mode decomposition results.

Example 6: Let us revisit the example shown in Figure 1 in the introduction. The original data $f_0(t)$ has a slowly growing trend linear in time. Suppose $f_r(t)$ is the linear regression of $f_0(t)$ and let $f(t) = f_0(t) - f_r(t)$. The SS energy distribution of $f(t)$ shown in Figure 19 left has three essential supports corresponding to three wave-like components. By weighting the locations of theses supports, we obtain the instantaneous frequency estimates of each component as shown in Figure 19. According to the evolutive pattern of the intrinsic frequencies, there are only two general modes contained in the superposition. The curve

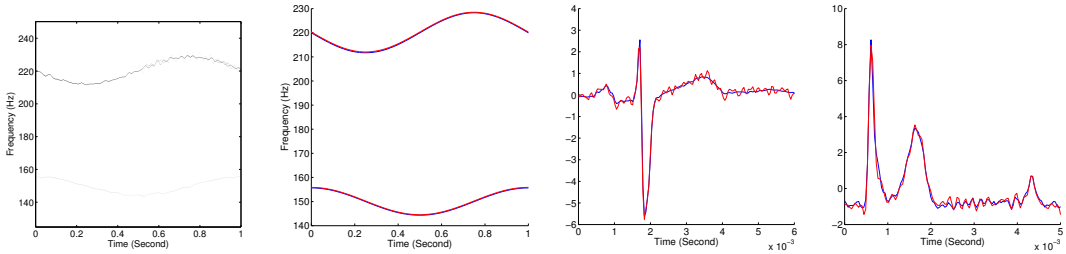

Figure 17: Left: The SS energy distribution of Example 5. Middle left: Real instantaneous frequencies (red) and instantaneous frequency estimates (blue). Middle right: Real ECG shape function $s_5(t)$ (blue) and its estimate (red). Right: Real ECG shape function $s_6(t)$ (blue) and its estimate (red).

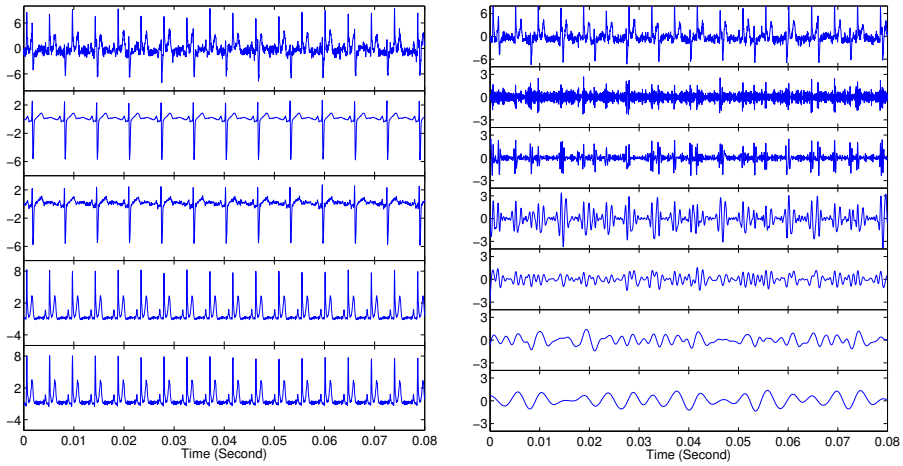

Figure 18: Left: The first row is a noisy superposition of two synthetic ECG signals with SNR = 0. The second and the third row: the first noiseless ECG component and its recovered result. The fourth and the fifth row: the second noiseless ECG component and its recovered result. Right: The first row is the same superposition of two ECG signals. Rows below the first one are identified modes provided by the EEMD method.

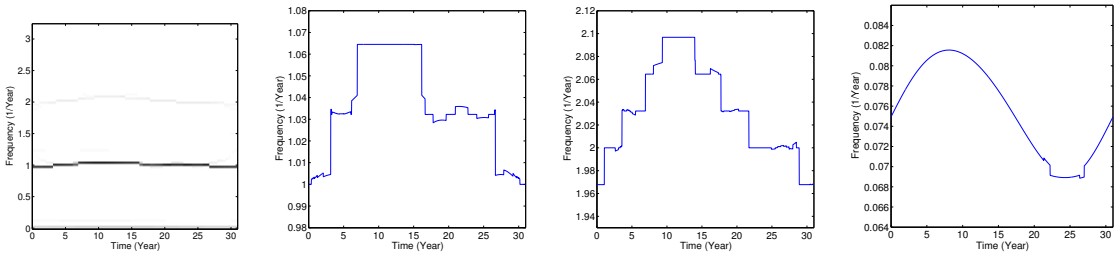

Figure 19: Left: $T_f(a,b)$ of the 31 years CO_2 concentration data. Middle left: The instantaneous frequency of the annual cycle. Middle right: The instantaneous frequency of the semiannual cycle. Right: the instantaneous frequency of the low frequency cycle. The curve classification algorithm groups the annual and semiannual cycle together.

classification step in the GMDWP method automatically groups the annual estimate and the semiannual estimate together. Hence, the decomposition result of the GMDWP method contains a general mode which is the sum of the annual cycle and the semiannual cycle shown in Figure 20. Because of the low frequency of the third term, it is reasonable to combine it with $f_r(t)$ to obtain a slowly varying growing trend shown in Figure 20.

6 Conclusion

This paper proposed the 1D synchrosqueezed wave packet transform (SSWPT) and the diffeomorphism based spectral analysis (DSA) method to solve the general mode decomposition problem under a weak well-separation condition and a well-different condition. These algorithms combine several ideas, namely (i) time-frequency transforms with better resolutions, (ii) a curve classification method, (iii) instantaneous information estimates, (iv) diffeomorphisms and (v) the (short time) Fourier transform.

In particular, item (i) allows us to estimate the instantaneous information of high frequency Fourier expansion terms of general modes; items (ii) and (iii) classify the extracted Fourier expansion terms and provide the instantaneous information of general modes for the DSA method; item (iv) linearizes the phase functions of general modes so that item (v) is able to identify the spectra of general shape functions. As a consequence of the above steps, the general modes are reconstructed by adding up their Fourier expansion terms one-by-one.

There are many future directions for the general mode decomposition problem. The most important work is to estimate the instantaneous information of each general mode without any well-separation condition. Given the instantaneous information, the DSA method is able to decompose the general superposition accurately. Another work of importance is the rigorous noise analysis of these methods. Although numerical results have shown robustness against random Gaussian noise, theoretical analysis is still an open problem. The robustness properties of synchrosqueezed wavelet transforms have been analyzed in [27] recently. Similar conclusions may be true for the methods proposed in this paper. It is also of interest to study other kinds of noise and to explore the effects of noise on the reconstruction. Finally, it would be appealing to weaken the well-different condition of phase functions in Theorem 4.5 and to classify the class of well-different phase functions.

Acknowledgments. H.Y. was partially supported by NSF grant CDI-1027952. H.Y.

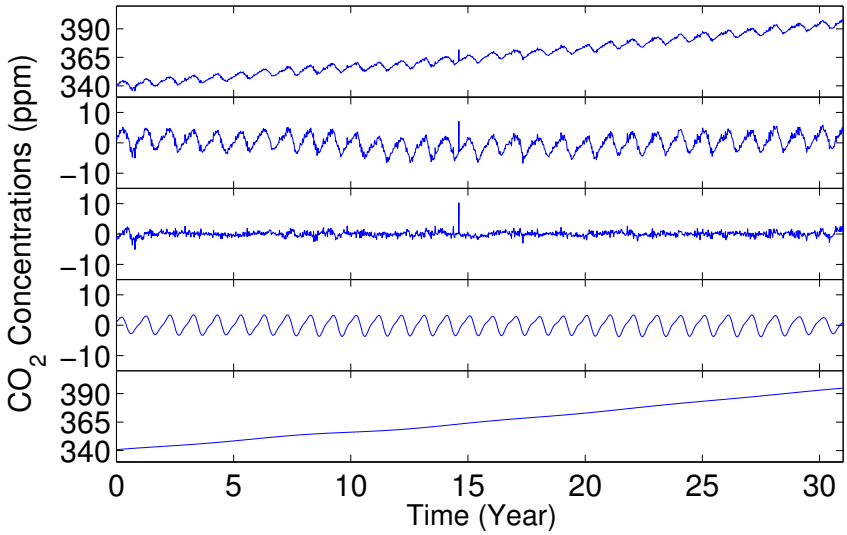

Figure 20: Decomposition results of the 31 years CO_2 concentration data provided by the GMDWP method. Top: The original data $f_0(t)$. Second row: The signal $f(t) = f_0(t) - f_r(t)$. Third row: The remaining noise term. Fourth row: The annual general mode. Last row: The slowly growing mode, which is the sum of $f_r(t)$ and the low frequency component.

thanks Lexing Ying for comments on the manuscript.

References

[1] F. Auger, E. Chassande-Mottin, and P. Flandrin. Making reassignment adjustable: The levenberg-marquardt approach. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 3889–3892, March 2012.

[2] F. Auger and P. Flandrin. Improving the readability of time-frequency and time-scale representations by the reassignment method. *Signal Processing, IEEE Transactions on*, 43(5):1068–1089, May 1995.

[3] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu. Time-frequency reassignment and synchrosqueezing: An overview. *Signal Processing Magazine, IEEE*, 30(6):32–41, Nov 2013.

[4] B. Boashash and S. Member. Estimating and interpreting the instantaneous frequency of a signal. In *Proceedings of the IEEE*, pages 520–538, 1992.

[5] Y.-C. Chen, M.-Y. Cheng, and H.-T. Wu. Non-parametric and adaptive modelling of dynamic periodicity and trend with heteroscedastic and dependent errors. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, pages n/a–n/a, 2013.

[6] I. Daubechies, J. Lu, and H.-T. Wu. Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool. *Appl. Comput. Harmon. Anal.*, 30(2):243–261, 2011.

[7] I. Daubechies and S. Maes. A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models. In *Wavelets in Medicine and Biology*, pages 527–546. CRC Press, 1996.

[8] L. Demanet and L. Ying. Wave atoms and sparsity of oscillatory patterns. *Appl. Comput. Harmon. Anal.*, 23(3):368–387, 2007.

[9] J. Gilles. Empirical wavelet transform. *Signal Processing, IEEE Transactions on*, 61(16):3999–4010, 2013.

[10] J. Gilles, G. Tran, and S. Osher. 2d empirical transforms. wavelets, ridgelets, and curvelets revisited. *SIAM Journal on Imaging Sciences*, 7(1):157–186, 2014.

[11] A. Goldberger. *Clinical Electrocardiography: A Simplified Approach.* Mosby-Elsevier, 7th edition, 2006.

[12] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000 (June 13). Circulation Electronic Pages: http://circ.ahajournals.org/cgi/content/full/101/23/e215 PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

[13] R. H. Herrera, J.-B. Tary, and M. van der Baan. Time-frequency representation of microseismic signals using the synchrosqueezing transform. *CoRR*, abs/1301.1295, 2013.

[14] T. Hou, Z. Shi, and P. Tavallali. Convergence of a data-driven time-frequency analysis method. *arXiv:1303.7048 [math.NA]*, 2013.

[15] T. Y. Hou and Z. Shi. Adaptive data analysis via sparse time-frequency representation. *Adv. Adapt. Data Anal.*, 3(1-2):1–28, 2011.

[16] T. Y. Hou and Z. Shi. Data-driven timefrequency analysis. *Applied and Computational Harmonic Analysis*, 35(2):284 – 308, 2013.

[17] T. Y. Hou, M. P. Yan, and Z. Wu. A variant of the EMD method for multi-scale data. *Adv. Adapt. Data Anal.*, 1(4):483–516, 2009.

[18] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 454(1971):903–995, 1998.

[19] N. E. Huang, Z. Wu, S. R. Long, K. C. Arnold, X. Chen, and K. Blank. On instantaneous frequency. *Adv. Adapt. Data Anal.*, 1(2):177–229, 2009.

[20] W. Huang, Z. Shen, N. E. Huang, and Y. C. Fung. Engineering analysis of biological variables: An example of blood pressure over 1 day. *Proc. Natl. Acad. Sci.*, 95, 1998.

[21] C. Li and M. Liang. A generalized synchrosqueezing transform for enhancing signal timefrequency representation. *Signal Processing*, 92(9):2264 – 2274, 2012.

[22] A. Y. Ng, M. I. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. *Neural Information Processing Systems*, 14, 2001.

[23] B. Picinbono. On instantaneous amplitude and phase of signals. *IEEE Trans. Signal Processing*, pages 552–560, 1997.

[24] G. Rilling and P. Flandrin. One or two frequencies? The empirical mode decomposition answers. *IEEE Trans. Signal Process.*, 56(1):85–95, 2008.

[25] M. Soltanolkotabi and E. J. Cands. A geometric analysis of subspace clustering with outliers. *CoRR*, abs/1112.4258, 2011.

[26] M. Soltanolkotabi, E. Elhamifar, and E. J. Candès. Robust subspace clustering. *CoRR*, abs/1301.2603, 2013.

[27] G. Thakur, E. Brevdo, N. S. Fučkar, and H.-T. Wu. The synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new paleoclimate applications. *Signal Processing*, 93(5):1079–1094, 2013.

[28] G. Thakur and H.-T. Wu. Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples. *SIAM J. Math. Analysis*, 43(5):2078–2095, 2011.

[29] A. D. Veltcheva. Wave and group transformation by a hilbert spectrum. *Coastal Engineering Journal*, 44(4), 2002.

[30] H.-T. Wu. Instantaneous frequency and wave shape functions (i). *Applied and Computational Harmonic Analysis*, 35(2):181 – 199, 2013.

[31] H.-T. Wu, Y.-H. Chan, Y.-T. Lin, and Y.-H. Yeh. Using synchrosqueezing transform to discover breathing dynamics from {ECG} signals. *Applied and Computational Harmonic Analysis*, 36(2):354 – 359, 2014.

[32] H.-T. Wu, P. Flandrin, and I. Daubechies. One or two frequencies? The synchrosqueezing answers. *Adv. Adapt. Data Anal.*, 3(1-2):29–39, 2011.

[33] H.-T. Wu, S.-S. Hseu, M.-Y. Bien, Y. R. Kou, and I. Daubechies. Evaluating physiological dynamics via synchrosqueezing: Prediction of ventilator weaning. *Biomedical Engineering, IEEE Transactions on*, 61(3):736–744, March 2014.

[34] Z. Wu and N. E. Huang. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 01(01):1, 2009.

[35] Z. Wu, N. E. Huang, and X. Chen. Some considerations on physical analysis of data. *Advances in Adaptive Data Analysis*, 3(1-2):95–113, 2011.

[36] H. Yang and L. Ying. Synchrosqueezed wave packet transform for 2d mode decomposition. *SIAM Journal of Imaging Science*, 2013.

[37] H. Yang and L. Ying. Synchrosqueezed curvelet transform for 2d mode decomposition. *SIAM Journal on Mathematical Analysis*, to appear.