The Fourier-based Synchrosqueezing Transform

Thomas Oberlin, Sylvain Meignen, Valérie Perrier

THE FOURIER-BASED SYNCHROSQUEEZING TRANSFORM

T. Oberlin, S. Meignen and V. Perrier

Laboratoire Jean Kuntzmann, University of Grenoble and CNRS
51 rue des Mathmatiques - BP 53 38041 Grenoble cedex 09, France

ABSTRACT

The short-time Fourier transform (STFT) and continuous wavelet transform (CWT) are intensively used to analyze and process multicomponent signals, ie superpositions of modulated waves. The synchrosqueezing is a post-processing method which circumvents the uncertainty relations, inherent to these linear transforms, by reassigning the coefficients in scale or frequency. Originally introduced in the setting of the continuous wavelet transform, it provides a sharp, concentrated representation, while remaining invertible. This technique received a renewed interest with the recent publication of an approximation result, which provides guarantees for the decomposition of a multicomponent signal. This paper adapts the formulation of the synchrosqueezing to the STFT, and states a similar theoretical result. The emphasis is put on the differences with the CWT-based synchrosqueezing, and all the content is illustrated through numerical experiments.

***Index Terms*—** multicomponent signals, short-time Fourier transform, ridge analysis, synchrosqueezing, reassignment

1. INTRODUCTION

Linear time-frequency and time-scale analysis are standard tools for the study of nonstationnary signals, or deterministic signals with varying frequency content. In particular, multicomponent signals, ie superpositions of amplitude- and frequency modulated waves (AM–FM), are accurately analyzed by the short-time Fourier transform (STFT) [1] or the continuous wavelet transform (CWT) [2]. It is well known that both transforms for such signals draw strips in the time-frequency (TF) or time-scale (TS) plane, centered at the *ridges* corresponding to the instantaneous frequencies [3].

The Synchrosqueezing transform, introduced in [4], is a kind of reassignment method [5] that aims to sharpen a time-scale representation, while remaining invertible. This method received a renewed interest in 2011 with the publication of a strong theoretical result concerning the decomposition of multicomponent signals with Synchrosqueezing [6]. A wide range of applications have been developped since, for instance in [7, 8, 9]. In this work, we will focus on the theoretical fundations of the method.

The SST was indeed originally introduced in the context of wavelet analysis, and it has not been entirely adapted to the STFT. For example, [10] proposes a SST-like decomposition based on the STFT, that comes with an approximation result. But this transform does not allow the reconstruction of the modes, which is yet one of the main characterisstics of the original SST. More problematic, the assumptions on the modes are global, whereas the STFT and the CWT are local transforms. A more natural extension has been proposed and used for example in [11], but without any theoretical considerations.

This paper addresses this issue by defining properly and Fourier-based SST (FSST), and providing an approximation theorem similar to [6], but adapted to the case of the STFT. Besides, a substantial part of this paper will be devoted to a comparison between wavelets and STFT, both on a theoretical and a numerical perspective. To this end, we start by defining the STFT and CWT, then we introduce the FSST and state the approximation result, underlining the differences with the CWT-based SST. Finally, these differences are illustrated and demonstrated though numerical experiments on synthetic multicomponent signals.

2. SHORT-TIME FOURIER TRANSFORM AND MULTICOMPONENT SIGNALS

In the paper we denote by $\hat{f}(\nu)$ the Fourier transform of function f with the following normalization:

$$\hat{f}(\nu) = \int_{\mathbb{R}} f(x) e^{2i\pi\nu x} \, dx. \tag{1}$$

The short-time Fourier transform (STFT) is a local version, obtained through the use of a sliding window g:

$$V_f(\eta, t) = \int_{\mathbb{R}} f(\tau) g(\tau - t) e^{-2i\pi\eta(\tau - t)} \, d\tau. \tag{2}$$

The representation of $|V_f(\eta, t)|^2$ in the time-frequency plane is called the *spectrogram* of signal f. Let us also recall the wavelet transform W_f, which uses an admissible wavelet $\psi \in L^2(\mathbb{R})$ (satisfying $0 < C_\psi = \int_0^\infty |\hat{\psi}(\xi)|^2 \frac{d\xi}{\xi} < \infty$), and is defined for anyt time t and scale $a > 0$ by:

$$W_f(a, t) = \frac{1}{a} \int_{\mathbb{R}} f(\tau) \psi\left(\frac{\tau - t}{a}\right)^* d\tau. \tag{3}$$

We now want to study the STFT of a multicomponent signal of the form

$$f(t)=\sum_{k=1}^{K} f_k(t)=\sum_{k=1}^{K} A_k(t)e^{2i\pi\phi(t)}. \qquad (4)$$

If we assume slow variations on the instantaneous amplitudes A_k and frequencies ϕ_k', we can write the following approximation, wich amounts to approximate f by a sum of pure waves:

$$f(\tau)\approx\sum_{k=1}^{K} A_k(t)e^{2i\pi[\phi(t)+\phi'(t)(\tau-t)]}. \qquad (5)$$

The corresponding approximation for the STFT then writes:

$$V_f(\eta,t)\approx\sum_{k=1}^{K} f_k(t)\hat{g}(\eta-\phi_k'(t)). \qquad (6)$$

This shows that the representation of a multicomponent signal in the time-frequency plane is concentrated around the so-called *ridges*, defined by $\eta=\phi_k'(t)$. If the frequencies ϕ_k' are separated enough compared to the support of $\hat{g}$, each mode occupies a distinct domain of the TF plane, allowing their detection, separation and reconstruction.

3. FOURIER-BASED SYNCHROSQUEEZING

The aim of the synchrosqueezing is twofold : first, to provide a concentrated representation of multicomponent signals in the time-frequency plane; second, to give a decomposition method, that allows to separate and demodulate the different modes. A major theoretical result, originally stated in [6] in the wavelet context, shows its usefulness for separated low-modulated multicomponent signals. This sections defines the STFT-based synchrosqueezing (FSST), and then extends the approximation result of [6] to the case of the STFT. We will pay a particular attention to the differences between the FSST and the wavelet-based synchrosqueezing (WSST).

3.1. Motivation, definition

Starting from the TFCT V_f, the SST moves the coefficients $V_f(\eta,t)$ according to the map $(\eta,t)\mapsto(\hat{\omega}_f(\eta,t),t)$, where $\hat{\omega}_f$ is the *local instantaneous frequency* defined by

$$\hat{\omega}_f(\eta,t)=\frac{1}{2\pi}\partial_t\arg V_f(\eta,t). \qquad (7)$$

This operator is simply the instantaneous frequency of the signal at time t, filtered at frequency η. We will see that it is indeed a good local approximation of the instantaneous frequencies $\phi_k'(t)$. The second key ingredient of the synchrosqueezing is the following “vertical” reconstruction formula, which stands in $L^2(\mathbb{R})$ provided that the window g is continuous and does not vanish at 0:

$$f(t)=\frac{1}{g(0)}\int_{\mathbb{R}} V_f(\eta,t)\,d\eta. \qquad (8)$$

The synchrosqueezing consists in resticting the integration domain in equation (8) to the interval where $\hat{\omega}_f(\eta,t)=\omega$, by formally writing:

$$T_f(\omega,t)=\frac{1}{g(0)}\int_{\mathbb{R}} V_f(\eta,t)\delta\left(\omega-\hat{\omega}_f(\eta,t)\right)\,d\eta. \qquad (9)$$

The next section defines more mathematically the SST, and extends the approximation theorem of [6].

3.2. An approximation result

Definition 3.1. *Let $\varepsilon>0$ and $\Delta\in(0,1)$. The set $\mathcal{B}_{\Delta,\varepsilon}$ of multicomponent signals with modulation ε and separation Δ is the set of all signals $f(t)=\sum_{k=1}^{K} f_k(t)$ where*

- *the $f_k(t)=A_k(t)e^{2i\pi\phi_k(t)}$ satisfy: $A_k\in\mathcal{C}^1(\mathbb{R})\bigcap L^\infty(\mathbb{R})$, $\phi_k\in\mathcal{C}^2(\mathbb{R})$, $\sup_t\phi_k'(t)<\infty$ and for all t, $A_k(t)>0$, $\phi_k'(t)>0$, $|A_k'(t)|\le\varepsilon$ and $|\phi_k''(t)|\le\varepsilon$.*
- *the f_k are separated with resolution Δ, ie for all $k\in\{1,\cdots,K-1\}$ and all t,*

$$\phi_{k+1}'(t)-\phi_k'(t)>2\Delta. \qquad (10)$$

Definition 3.2. *Let $\rho\in\mathcal{D}(\mathbb{R})$ be such that $\int\rho=1$, and consider $\gamma,\delta>0$, and $f\in\mathcal{B}_{\Delta,\varepsilon}$, the STFT-based synchrosqueezing (SST) of f with threshold γ and accuracy δ is defined by:*

$$T_f^{\delta,\gamma}(\omega,t)=\frac{1}{g(0)}\int_{|V_f(\eta,t)|>\gamma} V_f(\eta,t)\frac{1}{\delta}\rho\left(\frac{\omega-\hat{\omega}_f(\eta,t)}{\delta}\right)d\eta. \qquad (11)$$

If we make δ and γ tend to zero, we formally obtain the usual signal processing definition (9).

Theorem 3.1. *Consider $f\in\mathcal{B}_{\Delta,\varepsilon}$, and $\nu\in(0,\frac{1}{2})$. Let $g\in\mathcal{S}(\mathbb{R})$ be such that $\operatorname{supp} g\in[-\Delta,\Delta]$, and $\rho\in\mathcal{D}(\mathbb{R})$ satisfying $\int\rho=1$. Then, if ε is small enough, the following holds:*

- *$|V_f(\eta,t)|>\varepsilon^\nu$ only when there exists $k\in\{1\cdots,K\}$ such that $(\eta,t)\in Z_k:=\{(\eta,t)\,/\,|\eta-\phi_k'(t)|<\Delta\}$.*
- *For all $k\in\{1\cdots,K\}$ and all $(\eta,t)\in Z_k$ such that $|V_f(\eta,t)|>\varepsilon^\nu$, we have*

$$|\hat{\omega}_f(\eta,t)-\phi_k'(t)|\le\varepsilon^\nu. \qquad (12)$$

- *For all $k\in\{1\cdots,K\}$, there exists a constant C such that for all $t\in\mathbb{R}$,*

$$\left|\lim_{\delta\to 0}\left(\int_{|\omega-\phi_k'(t)|<\varepsilon^\nu} T_f^{\delta,\gamma}(\omega,t)\,d\omega\right)-f_k(t)\right|\le C\varepsilon^\nu. \qquad (13)$$

This theorem gives a strong approximation result, since is ensures that the non-zero coefficients of the SST are localized around the ridges, and that a reconstruction of the modes is easily obtained from the concentrated representation. The proof of this result is very similar to [6], and will not be detailed here. We will just give the main steps leading to the results, referring the reader to [12], appendix A for a complete and detailed proof.

Sketch of the proof. The proof starts by approximating V_f and $\hat{\omega}_f$ using a Taylor expansion. One obtains:

$$\left| V_f(\eta,t) - \sum_{k=1}^{K} f_k(t)\hat{g}(\eta - \phi_k'(t)) \right| \leq \varepsilon\Gamma_1(t), \tag{14}$$

where $\Gamma_1(t) = KI_1 + \pi I_2 \sum_{k=1}^{K} A_k(t)$, $I_n = \int_{\mathbb{R}} |x|^n |g(x)|\, dx$.

$$\begin{aligned} &\left| \partial_t V_f(\eta,t) - 2i\pi \sum_{k=1}^{K} f_k(t)\phi_k'(t)\hat{g}(\eta - \phi_k'(t)) \right| \\ &\qquad\qquad \leq \varepsilon\left(\Gamma_2(t) + 2\pi|\eta|\Gamma_1\right), \end{aligned} \tag{15}$$

where $\Gamma_2(t) = KI_1' + \pi I_2' \sum_{k=1}^{K} A_k(t)$, $I_n' = \int_{\mathbb{R}} |x|^n |g'(x)|\, dx$. Then we show that, if ε satisfies

$$\varepsilon \leq \Gamma_1(t)^{\frac{-1}{1-\nu}}, \tag{16}$$

then for any $1 \leq k \leq K$ and (η, t) such that $|\eta - \phi_k'(t)| < \Delta$ and $|V_f(\eta,t)| > \varepsilon^\nu$,

$$|\hat{\omega}_f(\eta,t) - \phi_k'(t)| \leq \left[(2\phi_k'(t) + \Delta)\Gamma_1(t) + \frac{1}{2\pi}\Gamma_2(t) \right] \varepsilon^{1-\nu}. \tag{17}$$

The end of the proof needs the following conditions on ε:

$$\varepsilon^\nu < \Delta \text{ and } \varepsilon < \left[(2\phi_k'(t) + \Delta)\Gamma_1(t) + \frac{1}{2\pi}\Gamma_2(t) \right]^{\frac{-1}{1-2\nu}}, \tag{18}$$

and uses mainly the Fubini and dominated convergence theorems. □

3.3. Relation with wavelet-based Synchrosqueezing

We will not define the wavelet-based synchrosqueezing here, but only underline the differences. They concern two different aspects :

- The assumptions on the modulation for the wavelet-based SST depends on the instantaneous frequency, ie they writes $|A_k'(t)| \leq \varepsilon\phi_k'(t)$ and $|\phi_k''(t)| \leq \varepsilon\phi_k'(t)$.
- The frequency separation between the components is logarithmic, and writes $\frac{\phi_{k+1}'(t) - \phi_k'(t)}{\phi_{k+1}'(t) + \phi_k'(t)} > \Delta$, where the wavelet is supposed to satisfy $\operatorname{supp}\hat{\psi} \subset [1-\Delta, 1+\Delta]$.

4. NUMERICAL RESULTS

The following numerical experiments use the classical logarithmic discretization of the scales of the CWT. We will use the Gaussian window and the complex Morlet wavelet, depending on a parameter σ and defined in the Fourier domain by:

$$\hat{g}(\nu) = \sigma^{\frac{1}{2}} e^{-\pi\sigma^2\nu^2} \text{ and } \hat{\psi}(\nu) = \sigma^{\frac{1}{2}} e^{-\pi\sigma^2(1-\nu)^2}.$$

The Matlab code used to create all the figures of the paper can be downloaded from `http://www-ljk.imag.fr/membres/Thomas.Oberlin/ic14.tar.gz`.

4.1. Limiting modulations

Let us start by determining which kind of signals are adapted to which method, either wavelets or the STFT. The question is : considering a fixed $\varepsilon > 0$, what are the signals that satisfies the assumptions? The pure waves obviously satisfy the assumptions for either the FSST or the WSST, and any ε. To simplify, we will first consider only single modes without AM, which write $h(t) = e^{2i\pi\phi(t)}$. We are particularly interesting in the strongest possible modulation, ie the phases ϕ satisfying $|\phi''(t)| = \varepsilon$ for the FSST, or $|\phi''(t)| = \varepsilon\phi'(t)$ for the WSST. One easily sees that the first kind of modes are *linear chirp*, ie they have a quadratic phase ϕ with $\phi'' = \varepsilon$, whereas the second ones are *exponential chirp* whose phase writes $h(t) = e^{2i\pi Ce^{\varepsilon t}}$.

To illustrate this, we draw on Figure 1 both transforms for a linear chirp with phase $\phi(t) = 10t + 100t^2$. It is clear that the quality of the representation given by the FSST remains constant along the time. For the WSST however, the representation is very concentrated around $t = 1$, but it is of poor quality for low t. This is because the quality of the representation depends on $\phi'(t)$. Then, we show on Figure 2 the same

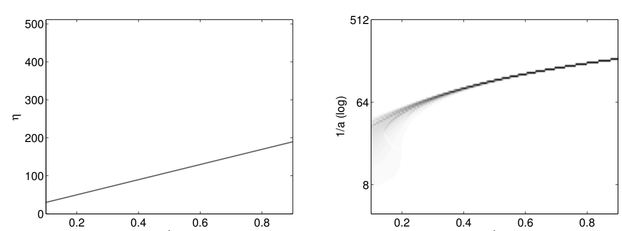

Fig. 1. Comparison of the methods on a linear chirp with a constant frequency modulation ϕ''. Left: the FSST. Right: the WSST.

test, but for an exponential chirp whose phase is $\phi(t) = 10e^{3t}$. As we expected, the FSST provides a sharp representation for low t, but does not manage to handle the high frequency modulation $\phi''(t)$ for higher times. Interestingly, the WSST does not provide a representation with constant quality, but seems to be more concentrated for high t. Actually, the result in [6] shows that the error for this kind of signal remains globally constant. But since the scales are discretized in a logarithmic way, the representation is sharper for high frequencies.

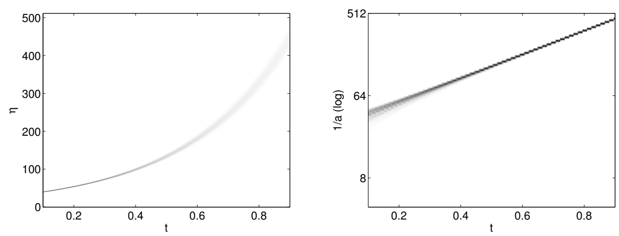

Fig. 2. Comparison of the methods on an exponential chirp with constant ratio ϕ''/ϕ'. Left: the FSST. Right: the WSST.

In order to obtain a representation with constant quality, the mode must satisfy a constant ration ϕ''/ϕ'^2. To check this numerically, we present the same test for such a signal, called an hyperbolic chirp, whose phase is $\phi(t) = -50 * \log(1.02 - t)$. One makes sure easily that the corresponding WSST remains sharp whatever the instantaneous frequency.

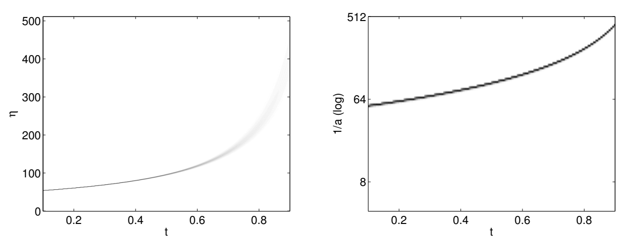

Fig. 3. Comparison of the methods on an hyperbolic chirp with constant ratio ϕ''/ϕ'^2. Left: the FSST. Right: the WSST.

4.2. separation vs localization

Let us here discuss the role of parameter σ, ie the size of the window or wavelet. From theorem 3.1 and the separation condition (10), it is clear that $\operatorname{supp} g$ should be small enough, which requires a parameter σ sufficiently high. Nevertheless, when detailing the error term in the constant C, one shows that it depends on the positive moments $\int_{\mathbb{R}} x^n |g(x)|\, dx$, thus σ should be small enough to ensure a good reassignment step. This phenomenon is illustrated on Figure 4, where one displays the FSST for a sum of polynomial chirps, using two different sizes of window g. For the small value $\sigma =$ (left), one gets a sharp representation but interferences, and the contrary is observed for a high σ. The same kind of phenomenon is observed on Figure 5, but in the wavelet case. This shows that σ must be chosen carefully to achieve a tradeoff between localization and separation. When one has only little a priori information on the signal, a convenient way to chose between the Fourier- and wavelet-based synchrosqueezing is the frequency range of the signal. The STFT can indeed handle a

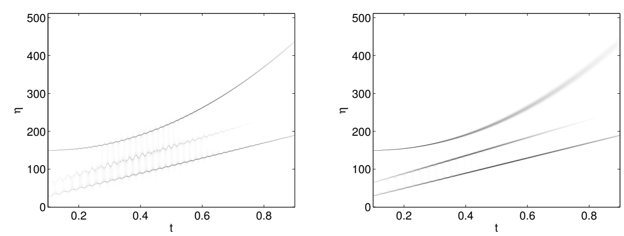

Fig. 4. FSST of s sum of polynomial chirps, using two different windows : $\sigma = 0.02$ (left) and $\sigma = 0.04$ (right).

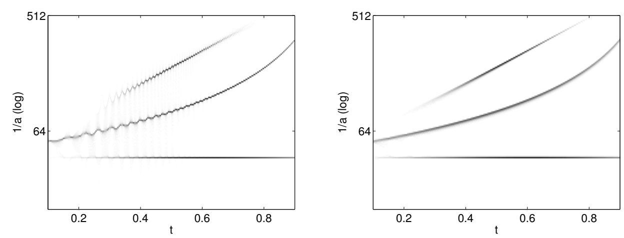

Fig. 5. WSST of s sum of highly modulated chirps, using two different windows : $\sigma = 2$ (left) and $\sigma = 5$ (right).

wide range of modulations at low frequency, and the WSST will behave satisfactorily at high frequencies in most cases.

5. CONCLUSION

This paper gave a natural extension of the Synchrosqueezing transform in the setting of the short-time Fourier transform, extending the approximation result shown in [6]. We insisted on the main differences between the wavelet-based and the Fourier-based SST, namely the frequency separation and the low-modulations assumptions. This allowed us to determine which class of modulations are analyzed best with wich transform. We showed that the FSST is adapted to linear or sublinear modulations, whereas the WSST can deal with exponential and hyperbolic chirps. This results have been illustrated and confirmed through numerical examples, where we also put the emphasis on the size of the window or wavelet.

We also showed that a tradeoff is needed between localization and separation, that can not be satisfactorily solved for some highly modulated multicomponent signals containing close instantaneous frequencies. In this regard, we should attempt in future works to extend the synchrosqueezing to strong modulations, that are already handled in the original reassignment method, as partially done by [13, 14]. Another interesting perspective to keep in mind is to use adaptive window's sizes $\sigma(t)$, as done for instance in [15, 16].

6. REFERENCES

[1] D. Gabor, "Theory of communication. part 1: The analysis of information," *Electrical Engineers-Part III: Radio and Communication Engineering, Journal of the Institution of*, vol. 93, no. 26, pp. 429–441, 1946.

[2] A. Grossmann and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *SIAM journal on mathematical analysis*, vol. 15, no. 4, pp. 723–736, 1984.

[3] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrsani, "Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies," *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 644–664, 1992.

[4] I. Daubechies and S. Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," *Wavelets in Medicine and Biology*, pp. 527–546, 1996.

[5] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," vol. 43, no. 5, pp. 1068–1089, 1995.

[6] I. Daubechies, J. Lu, and H-T. Wu, "Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool," *Applied and computational harmonic analysis*, vol. 30, no. 2, pp. 243–261, 2011.

[7] 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*, 2012.

[8] C. Franco, P-Y. Guméry, N. Vuillerme, A. Fleury, and J. Fontecave-Jallon, "Synchrosqueezing to investigate cardio-respiratory interactions within simulated volumetric signals," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*. IEEE, 2012, pp. 939–943.

[9] D. Iatsenko, A. Bernjak, T. Stankovski, Y. Shiogai, P.J. Owen-Lynch, P.B.M. Clarkson, P.V.E. McClintock, and A. Stefanovska, "Evolution of cardiorespiratory interactions with age," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 371, no. 1997, 2013.

[10] G. Thakur and H-T. Wu, "Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples," *SIAM Journal on Mathematical Analysis*, vol. 43, no. 5, pp. 2078–2095, 2011.

[11] Francois Auger, Patrick Flandrin, Y Lin, Stephen McLaughlin, Sylvain Meignen, Thomas Oberlin, and H Wu, "Time-frequency reassignment and synchrosqueezing: An overview," *Signal Processing Magazine, IEEE*, vol. 30, no. 6, pp. 32–41, 2013.

[12] T. Oberlin, *Analyse de signaux multicomposantes*, Ph.D. thesis, Université de Grenoble, 2013, `http://www-ljk.imag.fr/membres/Thomas.Oberlin/these_TO_15oct.pdf`.

[13] C. Li and M. Liang, "Time–frequency signal analysis for gearbox fault diagnosis using a generalized synchrosqueezing transform," *Mechanical Systems and Signal Processing*, vol. 26, pp. 205–217, 2012.

[14] S. Wang, X. Chen, G. Cai, B. Chen, X. Li, and Z. He, "Matching demodulation transform and synchrosqueezing in time-frequency analysis," 2013.

[15] L.J. Stankovic and V. Katkovnik, "Algorithm for the instantaneous frequency estimation using time-frequency distributions with adaptive window width," *Signal Processing Letters, IEEE*, vol. 5, no. 9, pp. 224–227, 1998.

[16] J. Xiao and P. Flandrin, "Multitaper time-frequency reassignment for nonstationary spectrum estimation and chirp enhancement," *Signal Processing, IEEE Transactions on*, vol. 55, no. 6, pp. 2851–2860, 2007.