A New Adaptive Technique for Multicomponent Signals Reassignment Based on Synchrosqueezing Transform

N. Laurent
University of Grenoble and CNRS
Laboratoire Jean Kuntzmann
38 041 Grenoble cedex 09, France

S. Meignen
University of Grenoble and CNRS
Laboratoire Jean Kuntzmann
38 041 Grenoble cedex 09, France

***Abstract*—In this paper, we tackle the issue of making adaptive the reassignment of multicomponent signals using synchrosqueezing transforms. Indeed, depending on the frequency modulation of the modes making up these signals, the reassignment results with synchrosqueezing transforms are significantly different depending on the so-called order of the transform. In this paper, we introduce a new technique to choose locally the best order of the synchrosqueezing transform to represent a multicomponent signal, and study its behavior in terms of mode reconstruction on both simulated and real signals.**

***Index Terms*—Synchrosqueezing transforms, short-time Fourier transform, mode reconstruction.**

I. Introduction

Linear Time-frequency representations (LTFRs) are widely used to analyze *multicomponent signals* (MCSs) since they enable the estimation of the *instantaneous frequencies* (IFs) of their constituent modes as well as their reconstruction. One of the most commonly used LTFR is the *short-time Fourier transform* (STFT) [1] whose performance in terms of IF estimation and mode reconstruction are dependent on the choice of an analysis window [2]. To circumvent this limitation of STFT a reassignment technique, called *Fourier-based synchrosqueezing transform* (FSST), was developed in [3], while alternative techniques aiming at making the STFT more adaptive were proposed in [4], [5].

FSST is however well designed only to reassign MCSs made of modes with small frequency modulation. Therefore new approaches were developed to take into account this aspect, through the so-called *second-order synchrosqueezing transform* (FSST2) [6], [7], and to deal with the reassignment of modes with fast oscillating phase [8], through the so-called *higher order synchrosqueezing transforms* (FSSTNs). The main problem associated with all these FSSTs is that they are designed for specific types of modes and, if the latter depart from this ideal situation, the reassignment process becomes rapidly inaccurate. Note that another source of inaccuracy is that the IF estimates used in FSSTs are very sensitive to noise.

Our goal in this paper is to find a way to choose locally in the TF plane the best order for the synchrosqueezing transform to represent the modes of a MCS. The proposed criterion is based on the analysis of the energy on the ridges of FSSTs of different orders. More precisely, we hypothesize that the FSST leading the largest coefficient magnitudes on a ridge is the most relevant to represent the associated mode. From this criterion, we build a new adaptive FSST which proves to outperform FSST with a fixed order for the reconstruction of both simulated and real MCSs.

The layout of the paper is as follows: in Section II, we introduce the notation we use throughout the paper as well as the basics on FSSTs. Then, we introduce our new adaptive FSST in Section III, and conclude the paper with different simulations showing the relevance of the proposed approach.

II. Definition and Notation

In this section, we introduce a series of definitions and notation used throughout the paper. Considering a signal $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and a real window $g \in L^\infty(\mathbb{R}) \cap L^2(\mathbb{R})$, the (modified) *Short-Time Fourier Transform* (STFT) is defined as:

$$V_f^g(t,\eta) = \int_{\mathbb{R}} f(\tau)g(\tau - t)e^{-2i\pi(\tau-t)\eta}d\tau. \quad (1)$$

STFT enables to perform the TF analysis of *multicomponent signals* (MCS) containing P modes defined by:

$$f(t) = \sum_{p=1}^{P} f_p(t), \quad (2)$$

in which each mode $f_p(t) = A_p(t)e^{2i\pi\phi_p(t)}$ where $A_p(t)$ is the *instantaneous amplitude* (IA) and $\phi_p(t)$ the *instantaneous phase* (IP). We assume $A_p > 0$ and is varying slowly. For such a signal, its *ideal TFR* (ITF) is given by

$$TI(t,\eta) = \sum_{p=1}^{P} A_p(t)\delta(\eta - \phi_p'(t)). \quad (3)$$

FSST aims at reassigning STFT through the following formula [3]

$$T_f^g(t,\eta) = \int_{\mathbb{R}} V_f^g(t,\xi)\delta(\eta - \widehat{\omega}_f(t,\xi))d\xi, \quad (4)$$

where $\widehat{\omega}_f(t,\xi)$ is an IF estimate of the signal at point (t,ξ) based on a local first order polynomial approximation of the phase of the modes, and is computed through

$$\widehat{\omega}_f(t,\xi) = \Re\left\{\frac{1}{2i\pi}\frac{\partial_t V_f^g(t,\xi)}{V_f^g(t,\xi)}\right\}. \quad (5)$$

Note that FSST can be extended to the discrete time and frequency setting [9], through the definition of

$$T_f^g[n,k] \approx T_f^g(\frac{n}{L}, k\frac{L}{M}), \quad (6)$$

where L is length of the signal, M the number of frequency bins and $\frac{L}{M}$ the frequency resolution. Since FSST is based on IF estimate $\widehat{\omega}_f$, its relevance is restricted to pure harmonics. To cope with this issue, the second order [7], [10] and then higher orders [8], [11] FSSTs were proposed. These were defined by $T_f^{g,N}(t,\eta)$, obtained by replacing $\widehat{\omega}_f$ in (4) by another IF estimate $\widehat{\omega}_f^{[N]}$ built assuming local phase polynomial oscillations of order N [8]. Then provided $g(0) \neq 0$, f can be approximated by

$$f(t) \approx \frac{1}{g(0)}\int_{\mathbb{R}} T_f^{g,N}(t,\eta)d\eta. \quad (7)$$

Similarly to what was done for FSST, it is possible to define $T_f^{g,N}[n,k] \approx T_f^{g,N}(\frac{n}{L}, k\frac{L}{M})$ and to reconstruct f at time $\frac{n}{L}$, using a sum instead of an integral.

In the case of a monocomponent signal and when N matches the order of its phase, the magnitude of FSSTN normalized by $g(0)$ is the ITF. This, in practice, means that all the information of the signal should be concentrated on the FSSTN ridge corresponding to the global maximum of the magnitude of FSSTN along the frequency axis at each time. Therefore, to measure the quality of the reassignment process associated with FSSTs the Earth mover's distance (EMD) is often used [12]. EMD is a sliced (fixed time) Wasserstein distance aimed at comparing probability distributions, which has already been used in the TF context for instance in [13], [7]. However, since the true ITF is unknown one cannot use EMD to determine the best order of FSST to represent a MCS. Furthermore, for a MCS, this order depends both on time and frequency since it has to adapt to the modes present at a given TF point. In the following section, we are going to explain how to make a relevant choice for N locally in the TF plane.

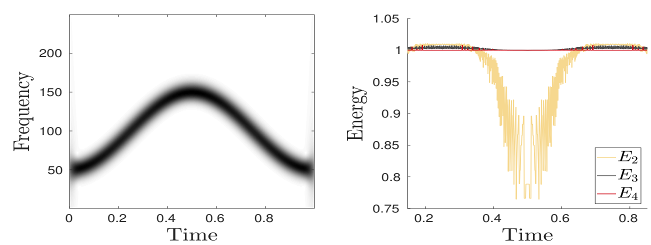
III. Adaptive Synchrosqueezing Transform Based on the Energy on the Ridges

In this section, we first investigate how the magnitude of FSSTN on its ridges and normalized by $g(0)$ are good estimates of the IA of the modes of a MCS. Note that these are exact estimates when the former fit into the model used to design FSSTN. From now on, FSST1 denotes the original FSST.

In that context, let ψ_N^p be the p^{th} ridge, namely corresponding to the p^{th} mode, computed as the local maxima of the magnitude of $T_f^{g,N}[n,k]$ for each time index n when k varies. To build the set of P ridges we use the classical approach developed in [14], [15] and used in [7], [16] to name a few. The normalized reassigned transform is defined, for time index n, by $\frac{1}{g(0)}|T_f^{g,N}[n,\psi_N^p[n]]|$ and is an estimate of $A_p(\frac{n}{L})$. Note that, due to frequency resolution constraints, STFT in the vicinity of a ridge associated with a mode may not be reassigned with FFSTN in a single frequency bin even if a mode fits perfectly into the model used to design FSSTN. Indeed, the IF of a mode can cross several frequency bins between two successive time indices. To circumvent such a drawback, we introduce two other FSSTNs denoted by $T_{f,1/2}^{g,N}$ and $T_{f,-1/2}^{g,N}$ corresponding to FSSTN but reassigned in a set of frequency bins shifted by half a bin upward or downward, respectively. Then we define the energy on the ridge as follows:

$$E_N^p[n] = \frac{\max\limits_{q=-1/2,0,1/2}\left(|T_{f,q}^{g,N}[n,\psi_N^p[n]]|\right)}{g(0)}, \quad (8)$$

in which $T_{f,0}^{g,N} = T_f^{g,N}$. Doing so, we favor the representation corresponding to the most concentrated energy, namely such that the IF crosses only one frequency bin between two successive time indices.



(a) (b)

Fig. 1. (a): magnitude of the STFT of a mode with an oscillating phase; (b): E_N for $2 \leq N \leq 4$ for the signal whose STFT is depicted in (a)

Our goal is then to investigate the quality of this amplitude estimate in the absence of noise. We first carry out simulations on a pure harmonic mode with constant amplitude, for which no significant differences arise between the estimates when N varies. For a linear chirp, we then note that there are no visible differences when N is larger or equal to 2, because for these N FSSTN handles well this type of frequency modulation, but as expected FSST1 behaves significantly worse. Now, if the signal corresponds to the STFT mangitude of Fig. 1 (a), in which $A = 1$ and the frequency is sinusoidal, the local linear chirp model used in FSST2 is no longer accurate at location where the frequency exhibits some curvature (around $t = 0.5$ s in that case), and thus the amplitude estimate E_2 is worse than E_3 or E_4 which exhibit very small differences (p is omitted because we consider a single mode). We also notice that, in the absence of noise, E_N is smaller than the mode amplitude when the reassignment process does not work well. This behavior is expected since, in that case, some part of the energy of the mode is not reassigned onto the ridge. In practice, we numerically notice that E_N, whatever N, is smaller than A, provided there are no time interference. Indeed, these interference can cause some overshoots in amplitude estimation with E_N. To illustrate this, let us consider the mode with amplitude $A = 1$ and corresponding to STFT magnitude of Fig. 2 (a), for which some time interference arise in the first half of the signal. For this reason, the signal no longer fits into the ideal model for $N = 4$, and this results in some overshoots in the estimate E_4, i.e $E_4[n] > A$ for some n (see Fig. 2 (d)). In the second part of the signal, the time interference are far less important (the sinusoidal frequency being multiplied by a dumping function), limiting the overshoots. Now, if one considers E_2 and E_3, depicted in Fig. 1 (b) and (c), the overshoot phenomenon is less important, but the overall amplitude estimate is a lot worse.

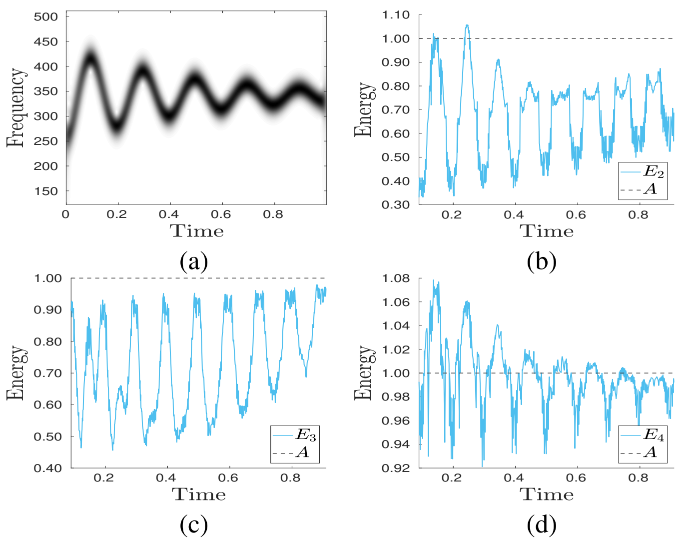


Fig. 2. (a): STFT magnitude of a mode with dumped sinusoidal frequency and amplitude $A = 1$; (b),(c),(d): E_2,E_3 and E_4 for the mode corresponding to (a), respectively.

We are now interested in assessing the overshoot phenomenon when a mode is affected by a complex white Gaussian noise ε such that $\tilde{f} = f + \varepsilon$. Though E_N is no longer an exact estimate of the mode amplitude, it should still give us a good indication of the quality of the underlying reassignment process. We investigate whether the presence of noise can cause some overshoots as time interference do. To do so, we consider noisy versions of the modes corresponding to STFT magnitudes of Fig. 1 (a) and 2 (a) (input SNR equals 10 dB, such an amount of noise ensures the ridge is always detectable). The upper 95^{th} percentiles of the random variables $(E_N[n])_n$ displayed in Fig. 3 (a) and (b) are always lower than the true amplitude A of the mode, meaning that the noise reassignment does not create any overshoots on the ridge. We also note that at this noise level, matching the order of FSST to the type of signal seems irrelevant, since FSST2 leads to the most accurate estimate of the amplitude.

Based on the previous simulations we now explain how to choose the order in FSSTN both locally and in an adaptive way. To this end, we propose to consider the very simple criterion which consists of selecting the order N for which E_N^p is the maximum for the p^{th} mode. Formally, considering

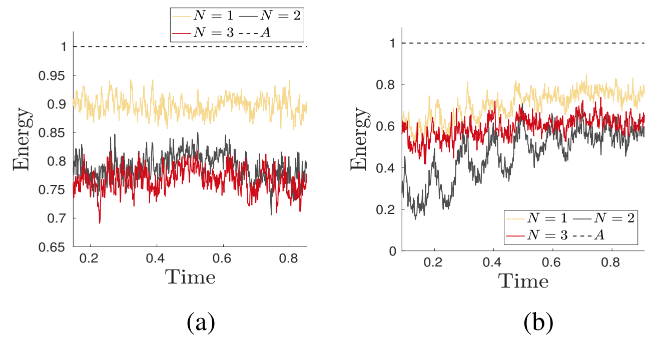


Fig. 3. (a),(b): upper 95^{th} percentiles of $(E_N[n])_n$, when $2 \leq N \leq 4$ for the signal of Fig. 1 (a) (resp Fig. 2 (a)). The number of noise realizations is 100.

the orders up to N_0, this corresponds to

$$\mathcal{N}^p[n] = \underset{N\in 1,\cdots,N_0}{\arg\max}\ E_N^p[n]. \quad (9)$$

With that formalism, at each time index n, the p^{th} mode is associated with a certain order which may vary with time. Then, to confirm that the order given by $\mathcal{N}^p$ at time index n is relevant, we propose to reconstruct $f_p(\frac{n}{L})$ by means of the formula:

$$f_p(\frac{n}{L}) \approx R_{f_p}^{\mathcal{N}^p}[n] := \frac{1}{g(0)} T_{\tilde{f},q[n]}^{g,\mathcal{N}^p[n]}[n, \psi_{\mathcal{N}^p[n]}[n]]. \quad (10)$$

in which $q[n]$ is either $-\frac{1}{2}, 0$ or $\frac{1}{2}$ according to (8).

IV. Results

Following formula (10), one defines the reconstruction of f_p from the coefficients on the ridge of FSSTN as:

$$f_p(\frac{n}{L}) \approx R_{f_p}^{N}[n] := \frac{1}{g(0)} T_{\tilde{f}}^{g,N}[n, \psi_N[n]]. \quad (11)$$

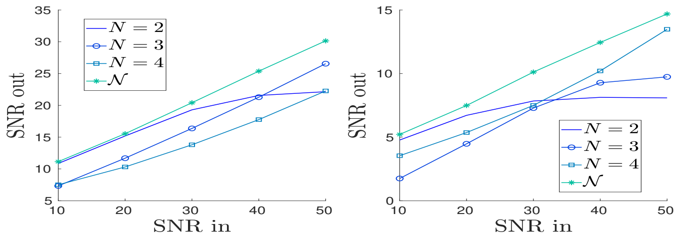
To compare the quality of mode reconstruction with the different techniques we introduce

$$\mathrm{SNR}(F_p, E - F_p) = 20\log_{10}\left(\frac{\|F_p\|_2}{\|E - F_p\|_2}\right), \quad (12)$$

with $F_p[n] = f_p(\frac{n}{L})$, E being an estimate of F_p, and in which $\|.\|_2$ denotes the l_2 norm.

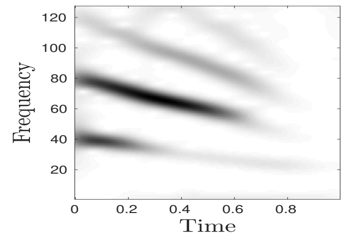
A. Monocomponent Case

We first study the reconstruction of the monocomponent signals of Fig. 1 (a) and 2 (a), and display the results in Fig. 4 (a) and (b). At high SNR and comparing FSSTs with fixed orders, we first notice that higher order FSSTs lead to better reconstruction for both signals. Then when the noise level increases, FSST2 behaves better than FSST3 and FSST4, meaning it is not recommended to use high order polynomial approximation of the phase of a mode even at a moderate noise level. We also notice that the new technique we propose, i.e. $R_f^{\mathcal{N}}$, leads to a better reconstruction than the other tested techniques, proving the relevance for mode reconstruction of adapting locally the order of the synchrosqueezing transform as defined in (9).



(a) (b)

Fig. 4. (a): output SNR corresponding to SNR$(f, R_f^N - f)$ for $N = 2, \cdots, 4$ or to SNR$(f, R_f^{\mathcal{N}} - f)$ for the signal of Fig. 1 (a); (b): same as (a) but for the signal of Fig. 2 (a).



(a)

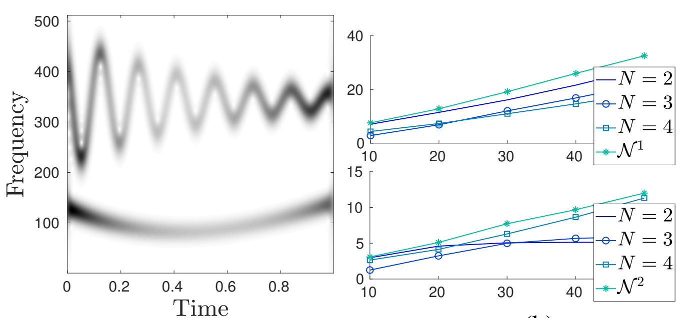
Order	SNR (dB)
1	3.4361
2	9.4786
3	8.5328
4	7.7444
$\mathcal{N}_a$	10.7182

(b)

Fig. 6. (a): STFT modulus of a bat echolocation call; (b): SNR corresponding to signal reconstruction from the coefficients on the ridges of the 3 most energetic detected modes.

B. Multicomponent Case

To give an illustration for the MCS case, we study the signal of Fig. 5 (a) which differs from the monocomponent case in that the amplitudes of the modes are not constant. In that case, we again notice that mode reconstruction is more accurate when one uses the technique we propose rather than FSSTN (see the reconstruction results for the two modes of Fig. 5 (a) displayed in Fig. 5 (b)).



(a) (b)

Fig. 5. (a): modulus of the STFT of the sum of two modes with varying amplitude; (b): mode reconstruction results as a function of input SNR.

C. Real Signals

We now investigate how the proposed adaptive technique improves the quality of reconstruction on a bat echolocation call whose STFT magnitude is depicted in Fig. 6 (a). In our study, we only look for the three most energetic modes in ridge detection and put $R_f^{\mathcal{N}_a} := \sum_{p=1}^{3} R_{f_p}^{\mathcal{N}^p}$, $\mathcal{N}_a$ meaning we consider an adaptive N in the MCS case. Similarly to $R_f^{\mathcal{N}_a}$, we respectively denote signal reconstruction from the three detected modes for, $1 \leq N \leq 4$, by R_f^N. We then compare $R_f^{\mathcal{N}_a}$ and R_f^N, for different N, to f in terms of SNR, and we again notice that the first technique is much better than the others (see the Table of Fig. 6 (b)).

V. Conclusion

In this paper, we have proposed a novel technique to assess locally in the time-frequency plane the best order in synchrosqueezing transforms to represent multicomponent signals. Our approach to find this order is based on the analysis of the coefficient magnitude on the ridges of synchrosqueezing transforms, and proves to be more relevant for mode reconstruction than synchrosqueezing transforms based on a fixed order. The main hypothesis used to design this new adaptive method is that the magnitude of the normalized synchrosqueezing transforms on their ridges estimate the modes amplitudes while remaining lower. Though this appears to be numerically true in most cases, a deeper mathematical analysis is still needed to support this claim. This will be done in a near future.

References

[1] L. Stankovic, M. Dakovic, and V. Ivanovic, "Performance of spectrogram as IF estimator," *Electronics Letters*, vol. 37, no. 12, pp. 797–799, 2001.

[2] P. Flandrin, *Time-frequency/time-scale analysis*. Academic Press, 1998, vol. 10.

[3] T. Oberlin, S. Meignen, and V. Perrier, "The Fourier-based synchrosqueezing transform," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 315–319.

[4] L. Li, H. Cai, H. Han, Q. Jiang, and H. Ji, "Adaptive short-time Fourier transform and synchrosqueezing transform for non-stationary signal separation," *Signal Processing*, vol. 166, p. 107231, 2020.

[5] L. Li, H. Cai, and Q. Jiang, "Adaptive synchrosqueezing transform with a time-varying parameter for non-stationary signal separation," *Applied and Computational Harmonic Analysis*, 2019.

[6] T. Oberlin, S. Meignen, and V. Perrier, "Second-order synchrosqueezing transform or invertible reassignment? Towards ideal time-frequency representations," *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1335–1344, March 2015.

[7] R. Behera, S. Meignen, and T. Oberlin, "Theoretical analysis of the second-order synchrosqueezing transform," *Applied and Computational Harmonic Analysis*, vol. 45, no. 2, pp. 379–404, 2018.

[8] D. H. Pham and S. Meignen, "High-order synchrosqueezing transform for multicomponent signals analysis-with an application to gravitational-wave signal." *IEEE Trans. Signal Processing*, vol. 65, no. 12, pp. 3168–3178, 2017.

[9] N. Laurent and S. Meignen, "A novel time-frequency technique for mode retrieval based on linear chirp approximation," *IEEE Signal Processing Letters*, vol. 27, pp. 935–339, 2020.

[10] T. Oberlin and S. Meignen, "The second-order wavelet synchrosqueezing transform," in *42th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2017.

[11] S. Meignen, T. Oberlin, and D.-H. Pham, "Synchrosqueezing transforms: From low-to high-frequency modulations and perspectives," *Comptes Rendus Physique*, vol. 20, no. 5, pp. 449–460, 2019.

[12] O. Pele and M. Werman, "Fast and robust earth movers distances," *2009 IEEE 12th International Conference on Computer Vision*, Sep 2009.

[13] I. Daubechies, Y. G. Wang, and H.-T. Wu, "ConceFT: concentration of frequency and time via a multitapered synchrosqueezed transform," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, Mar 2016.

[14] R. Carmona, W. Hwang, and B. Torresani, "Characterization of signals by the ridges of their wavelet transforms," *IEEE Transactions on Signal Processing*, vol. 45, no. 10, pp. 2586–2590, Oct 1997.
[15] ——, "Multiridge detection and time-frequency reconstruction," *IEEE Transactions on Signal Processing*, vol. 47, no. 2, pp. 480–492, Feb 1999.
[16] S. Meignen and D.-H. Pham, "Retrieval of the modes of multicomponent signals from downsampled short-time Fourier transform," *IEEE Transactions on Signal Processing*, vol. 66, no. 23, pp. 6204–6215, 2018.