

THE FOURIER-BASED SYNCHROSQUEEZING TRANSFORM

T. Oberlin, S. Meignen and V. Perrier

Laboratoire Jean Kuntzmann, University of Grenoble-Alpes, UMR 5224 CNRS
51 rue des Mathématiques - BP 53, 38041 Grenoble cedex 09, France

ABSTRACT

The short-time Fourier transform (STFT) and the continuous wavelet transform (CWT) are extensively used to analyze and process multicomponent signals, *i.e.* superpositions of modulated waves. The synchrosqueezing is a post-processing method which circumvents the uncertainty relation inherent to these linear transforms, by reassigning the coefficients in scale or frequency. Originally introduced in the setting of the CWT, it provides a sharp, concentrated representation, while remaining invertible. This technique received a renewed interest with the recent publication of an approximation result related to the application of the synchrosqueezing to multicomponent signals. In the current paper, we adapt the formulation of the synchrosqueezing to the STFT and state a similar theoretical result to that obtained in the CWT framework. The emphasis is put on the differences with the CWT-based synchrosqueezing with numerical experiments illustrating our statements.

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 nonstationary signals or deterministic signals with varying frequency content. In particular, multicomponent signals, *i.e.* superpositions of amplitude- and frequency modulated waves (AM–FM), are accurately analyzed with STFT [1] or CWT [2]. It is well known that both transforms for such signals draw stripes in the time-frequency (TF) or time-scale (TS) planes, centered on *ridges* corresponding to the instantaneous frequencies of the modes making up the signal [3].

The Synchrosqueezing transform (SST), introduced in [4], is a kind of reassignment method [5] that aims to sharpen a TS representation while remaining invertible. This method recently received a renewed interest with the publication of approximation results on the analysis of multicomponent signals with CWT-based SST [6]. A wide range of applications have been developed since, for instance in [7, 8, 9]. In this

work, we focus on the theoretical foundations of the method but for the STFT-based SST.

Indeed, the SST, originally introduced in the context of CWT, has not entirely been adapted to the STFT. For example, in [10] a SST-like decomposition based on the STFT is proposed that enables the derivation of some approximation results. However this transform does not allow for modes reconstruction, yet one of the main characteristics of the original SST. More problematic are the global assumptions made on the modes whereas the STFT and the CWT are local transforms. A more natural extension has been proposed in [11] but without any theoretical considerations.

This paper addresses this issue by defining properly a STFT-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 the comparison between CWT and STFT, both on a theoretical and a numerical perspective. To this end, we first define STFT and CWT, then introduce FSST and state the approximation result, underlining the differences with the CWT-based SST. Finally, these differences are illustrated and demonstrated through numerical experiments on synthetic multicomponent signals.

2. SHORT-TIME FOURIER TRANSFORM AND MULTICOMPONENT SIGNALS

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

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

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

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

The representation of $|V_f(\eta, t)|^2$ in the TF plane is called the *spectrogram* of signal f . Let us also recall that the CWT, W_f , 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 any time t and scale

The authors acknowledge the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-13-BS03-0002-01 (ASTRES).

$a > 0$ by:

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

We now want to study the STFT of a multicomponent AM-FM signal of the form:

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

If we assume slow variations on the instantaneous amplitudes A_k and frequencies ϕ'_k , we can write the following approximation in the vicinity of a fixed time t_0 , which amounts to approximate f by a sum of pure waves:

$$f(t) \approx \sum_{k=1}^K A_k(t_0) e^{2i\pi[\phi_k(t_0) + \phi'_k(t_0)(t-t_0)]}. \quad (5)$$

The corresponding approximation for the STFT then reads (changing t_0 by a generic t):

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

This shows that the representation of a multicomponent signal in the TF plane is concentrated around so-called *ridges*, defined by $\eta = \phi'_k(t)$. If frequencies ϕ'_k are enough separated when k varies, each mode occupies a distinct domain of the TF plane, allowing for their detection, separation and reconstruction.

3. FOURIER-BASED SYNCHROSQUEEZING

The aim of the SST is twofold: to provide a concentrated representation of multicomponent signals in the TF plane, and a decomposition method that enables to separate and demodulate the different modes. A theoretical result, originally stated in [6] in the CWT context, shows its usefulness for well separated low-modulated multicomponent signals. This section defines the STFT-based synchrosqueezing (FSST), and then extends the approximation result of [6] to that framework. We will pay a particular attention to the differences between FSST and CWT-based synchrosqueezing (WSST).

3.1. Motivation, definition

Starting from the STFT V_f , the FSST 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

$$\hat{\omega}_f(\eta, t) = \frac{1}{2\pi} \partial_t \arg V_f(\eta, t) = \mathcal{R}e \left(\frac{1}{2i\pi} \frac{\partial_t V_f(\eta, t)}{V_f(\eta, t)} \right) \quad (7)$$

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 SST is the following ‘‘vertical’’ reconstruction formula, which stands in $L^2(\mathbb{R})$ provided the window g is continuous and does not vanish at 0:

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

This enables to define the FSST which consists in restricting the integration domain in (8) to the interval where $\hat{\omega}_f(\eta, t) = \omega$, by formally writing:

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

The next section defines the FSST more mathematically 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

- $f_k(t) = A_k(t) e^{2i\pi\phi_k(t)}$ satisfies: $A_k \in \mathcal{C}^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$, $\phi_k \in \mathcal{C}^2(\mathbb{R})$, $\sup_t \phi'_k(t) < \infty$ and for all t , $A_k(t) > 0$, $\phi'_k(t) > 0$, $|A'_k(t)| \leq \varepsilon$ and $|\phi''_k(t)| \leq \varepsilon$.
- the f_k s are separated with resolution Δ , ie for all $k \in \{1, \dots, K-1\}$ and all t ,

$$\phi'_{k+1}(t) - \phi'_k(t) > 2\Delta. \quad (10)$$

Definition 3.2. Let ρ be in $\mathcal{D}(\mathbb{R})$ the space of smooth compactly supported function such that $\int \rho = 1$, and set $\gamma, \delta > 0$. The FSST of $f \in \mathcal{B}_{\Delta, \varepsilon}$, with threshold γ and accuracy δ is defined by:

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

If δ and γ tend to zero, one formally obtains the usual definition in signal processing (9).

Theorem 3.1. Consider $f \in \mathcal{B}_{\Delta, \varepsilon}$, and $\nu \in (0, \frac{1}{2})$. Let g be in $\mathcal{S}(\mathbb{R})$, the space of smooth rapidly decreasing functions, such that $\text{supp } \hat{g} \in [-\Delta, \Delta]$. Pick $\rho \in \mathcal{D}(\mathbb{R})$ satisfying $\int \rho = 1$. Then, provided ε is small enough, the following holds:

- $|V_f(\eta, t)| > \varepsilon^\nu$ only when there exists $k \in \{1 \dots, K\}$ such that $(\eta, t) \in Z_k := \{(\eta, t) / |\eta - \phi'_k(t)| < \Delta\}$.
- For all $k \in \{1 \dots, K\}$ and all $(\eta, t) \in Z_k$ such that $|V_f(\eta, t)| > \varepsilon^\nu$, one has

$$|\hat{\omega}_f(\eta, t) - \phi'_k(t)| \leq \varepsilon^\nu. \quad (12)$$

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

$$\left| \lim_{\delta \rightarrow 0} \left(\int_{|\omega - \phi'_k(t)| < \varepsilon^\nu} T_f^{\delta, \varepsilon^\nu}(\omega, t) d\omega \right) - f_k(t) \right| \leq C\varepsilon^\nu. \quad (13)$$

This theorem gives a strong approximation result since it ensures that the non-zero coefficients of the FSST are localized around the ridges and that mode reconstruction is easily achieved from this concentrated representation. The proof of this result, following the same lines as those of [6], is not detailed here and we just recall the main steps, referring the reader to [12], appendix A, for a complete and detailed proof.

Sketch of the proof. The proof starts with approximating V_f and $\hat{\omega}_f$ using Taylor expansions of the phases. One first obtains

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

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

$$\left| \partial_t V_f(\eta, t) - 2i\pi \sum_{k=1}^K f_k(t) \phi'_k(t) \hat{g}(\eta - \phi'_k(t)) \right| \leq \varepsilon (\Gamma_2(t) + 2\pi|\eta| \Gamma_1(t)), \quad (15)$$

with $\Gamma_2(t) = KI'_1 + \pi I'_2 \sum_{k=1}^K A_k(t)$, $I'_n = \int_{\mathbb{R}} |x|^n |g'(x)| dx$. Then one shows that, if ε satisfies

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

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

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

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

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

and uses mainly Fubini and dominated convergence theorems to finally establish (12) and (13). \square

3.3. Relation with CWT-based Synchrosqueezing

We now underline the differences with the WSST regarding the assumptions made on the multicomponent signal. These involve two different aspects:

- The assumptions on the modulation for the WSST depends on the instantaneous frequency, *i.e.* $|A'_k(t)| \leq \varepsilon \phi'_k(t)$ and $|\phi''_k(t)| \leq \varepsilon \phi'_k(t)$,

- The frequency separation between the components is logarithmic, and reads $\frac{\phi'_{k+1}(t) - \phi'_k(t)}{\phi'_{k+1}(t) + \phi'_k(t)} > \Delta$, where the wavelet is supposed to satisfy $\text{supp } \hat{\psi} \subset [1 - \Delta, 1 + \Delta]$,

which should be compared with assumptions given in definition 3.1.

4. NUMERICAL RESULTS

The following numerical experiments illustrate the behavior of FSST and WSST on multicomponent signals where the modulations obey different laws. We will use the Gaussian window (resp. complex Morlet wavelet) to define the STFT (resp. CWT), which *numerically* satisfy the assumptions of compact support and admissibility condition, and which respectively depend on a parameter σ as follows:

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

Note that FSST (resp. WSST) will be represented on a linear (resp. logarithmic) scale. The Matlab code used to create all these figures can be downloaded from <http://www-1.jk.imag.fr/membres/Thomas.Oberlin/ic14.tar.gz>.

4.1. Limiting modulations

Let us start with determining what type of signals, for a fixed $\varepsilon > 0$, fulfill the assumptions required for FSST or WSST. To simplify, let us first consider a single mode without AM: $h(t) = e^{2i\pi\phi(t)}$. We are particularly interested in the strongest possible modulation, *i.e.* a phase ϕ s.t. $|\phi''(t)| = \varepsilon$ for the FSST, or $|\phi''(t)| = \varepsilon\phi'(t)$ for the WSST. One easily sees that the first kind of modes are *linear chirps*, having a quadratic phase ϕ with $\phi'' = \varepsilon$, whereas the second ones are *exponential chirps* like $h(t) = e^{2i\pi F_0 e^{\varepsilon t}}$.

We illustrate this on Figure 1 by drawing both transforms for a linear chirp with phase $\phi(t) = 10t + 100t^2$. It is clear that the quality of the representation given by FSST is invariant with time. For WSST however, the representation is very concentrated around $t = 1$ but of poor quality at lower t , the quality of the representation being dependent on $\phi'(t)$. Figure

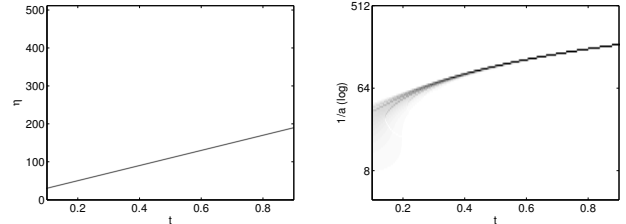


Fig. 1. Comparison of the methods on a linear chirp with a constant frequency modulation ϕ'' . Left: FSST. Right: WSST.

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

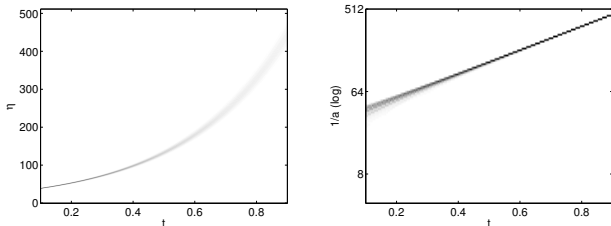


Fig. 2. Comparison of the methods on an exponential chirp with constant ratio ϕ''/ϕ' . Left: the FSST. Right: the WSST.

In order to obtain a representation with constant quality, the mode must satisfy a constant ratio ϕ''/ϕ'^2 . Note that this phenomenon has been thoroughly discussed in [13], ch. 4. To check this numerically, we present on Figure 3 the same test for such a signal, called an hyperbolic chirp, whose phase is $\phi(t) = -50 * \log(1.02 - t)$. One easily makes sure that the corresponding WSST remains sharp whatever the instantaneous frequency.

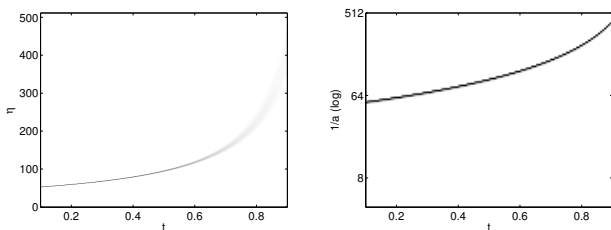


Fig. 3. Comparison of the methods on an hyperbolic chirp with constant ratio ϕ''/ϕ'^2 . Left: the FSST. Right: the WSST.

4.2. Separation vs localization

Let us here discuss the role of parameter σ , *i.e.* the size of the window or wavelet. From theorem 3.1 and separation condition (10), it is clear that $\text{supp } \hat{g}$ should be small enough, which requires σ sufficiently large. Nevertheless, by detailing the error term in the constant C in (13), one can show that it depends on $\int_{\mathbb{R}} x^n |g(x)| dx$, thus σ should be small enough to ensure a good reconstruction step. This phenomenon is illustrated on Figure 4, where one displays FSST for a sum of polynomial chirps, using two different window sizes. For σ

small (left), one gets a sharp representation with some interference, and the contrary for σ large. The same phenomenon

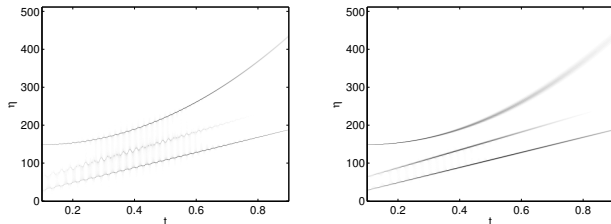


Fig. 4. FSST of a sum of polynomial chirps, using two different windows : $\sigma = 0.02$ (left) and $\sigma = 0.04$ (right).

is observed on Figure 5, but in the wavelet case. This shows that σ must be chosen carefully to achieve a trade-off between localization and separation. When one has only little a priori

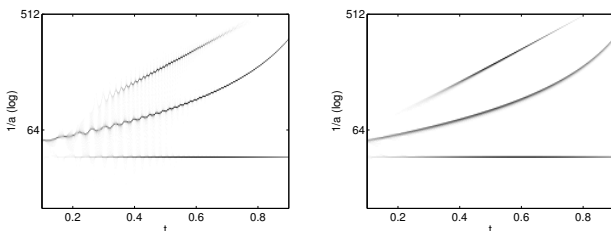


Fig. 5. WSST of a sum of highly modulated chirps, using two different windows : $\sigma = 2$ (left) and $\sigma = 5$ (right).

information on the signal, a convenient way to choose between FSST and WSST is the frequency range of the signal. The FSST can indeed handle a wide range of modulations at low frequency while the WSST behaves satisfactorily at high frequencies in most cases.

5. CONCLUSION

In this paper, we proposed a natural extension of the synchrosqueezing transform to the STFT setting, leading to an approximation result similar to the one shown in [6]. We emphasized the differences between WSST and FSST regarding frequency separation and low-modulation assumptions. This allowed us to determine which class of modulations are analyzed best with each transform, *i.e.* linear or sub-linear modulations should be dealt with FSST and exponential and hyperbolic chirps with WSST. Since a trade-off is needed between time localization and frequency separation, highly modulated multicomponent signals containing close instantaneous frequencies cannot be satisfactorily dealt with. We should therefore attempt in future works to improve the SST behavior in such contexts, while developing new applications.

6. REFERENCES

- [1] D. Gabor, "Theory of communication. part 1: The analysis of information," *Journal of I.E.E.*, 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. Torrèsani, "Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies," *IEEE Transactions on Information Theory*, 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 *Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*. 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] F. Auger, P. Flandrin, Y. Lin, S. McLaughlin, S. Meignen, T. 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_ThomasOberlin.pdf.
- [13] S. Mallat, *A wavelet tour of signal processing*, Academic press, 1999.