Making Synchrosqueezing Locally Adaptive in The Time-Frequency Plane

Marcelo Colominas, Sylvain Meignen

▶ **To cite this version:**

Marcelo Colominas, Sylvain Meignen. Making Synchrosqueezing Locally Adaptive in The Time-Frequency Plane. International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2023, Jun 2023, Rhodes Island, Greece. pp.1-5, 10.1109/ICASSP49357.2023.10097003 . hal-04218338

HAL Id: hal-04218338
https://hal.science/hal-04218338v1

Submitted on 26 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MAKING SYNCHROSQUEEZING LOCALLY ADAPTIVE IN THE TIME-FREQUENCY PLANE

Marcelo A. Colominas[a] *and Sylvain Meignen*[b]

[a] Institute for Research and Development in Bioengineering and Bioinformatics (IBB), CONICET, and Faculty of Engineering, Universidad Nacional de Entre Ríos (UNER), Argentina
[b] Jean Kuntzmann Laboratory, University Grenoble Alpes and CNRS 5225, France

ABSTRACT

In this work, we explore the problem of making synchrosqueezing transform adaptive. To deal with multicomponent signals, existing methods use a fixed order N, thus assuming all the modes making up the signal have order N polynomial phases. To go beyond this limitation, we introduce a new criterion based on the concentration of the representation to locally choose the best order for the synchrosqueezing transform. We study the performance of the proposed approach in terms of the error with respect to the ideal time-frequency representation, both on synthetic and real signals.

Index Terms— Synchrosqueezing transforms, nonstationary signals, time-frequency, multicomponent signals

1. INTRODUCTION

Nonstationary signals are encountered in many scientific fields, and are often represented as the superimposition of time-varying amplitude and frequency modes. A common way to analyze these *multicomponent signals* (MCSs), is to use linear *time-frequency representations* (TFRs), which enable to estimate the *instantaneous frequencies* (IFs) present in the signal. Among them, the probably most commonly used is the so-called *short-time Fourier transform* (STFT), which uses an analysis window (hence the short-time) to perform a local frequency analysis. The presence of this window, however, limits the resolution of the representation, introducing an uncertainty principle [1, 2]. To overcome this limitation, reassignment methods were introduced [3, 4]. A special form of reassignment, that allows for signal reconstruction, called synchrosqueezing transform [5, 6], was originally developed for the continuous wavelet transform, and then extended to the STFT to obtain the *Fourier-based synchrosqueezing transform* (FSST) [7]. Reconstruction is important for tasks such as *time-frequency* (TF) filtering [8, 9].

This work was supported by Agencia I+D+i (Argentina) via grant PICT-2020-01808, by UNER (Argentina) via grant PID 6224, the STIC AmSud ASPMLM-Voice grant, and also by the ANR ASCETE project (France) with grant number ANR-19- CE48-0001-01. macolominas@conicet.gov.ar; sylvain.meignen@univ-grenoble-alpes.fr.

FSST in its original formulation is, however, known to perform well only on MCSs made of modes with small frequency modulations. To overcome this limitation, the *second-order synchrosqueezing transform* (FSST2) was first introduced [10], and then the *high-order synchrosqueezing transform* (FSSTN) [11]. The idea behind FSSTN (considering the original FSST as FSST1) is to assume the phases of the modes can locally be approximated by polynomials of order N. Departing from this ideal situation, and also the presence of noise, introduce inaccuracies that harm the final result. Moreover, the presence of different modes with different modulations challenges the approach that uses a fixed N.

Our goal in this paper is to investigate a way to adaptively and locally choose the best order for the synchrosqueezing transform ; and we propose to do that by maximizing the concentration of the TFR using a greedy approach time by time. The rest of the paper is organized as follows. In Sec. 2 we introduce the notation and basic definitions we use throughout the paper. Our proposal, along with a step-by-step algorithm, is presented in Sec. 3. Sec. 4 proposes some numerical results on both synthetic and real data, illustrating the advantages of our approach over those based on fixed order FSSTs, and Sec. 5 concludes the paper.

2. NOTATION AND BASIC DEFINITIONS

We introduce in this section the notation and basic definitions that will be used throughout the paper. Given a signal $x \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and a real and even window $g \in L^\infty(\mathbb{R}) \cap L^2(\mathbb{R})$, we define the modified short-time Fourier transform (STFT) as

$$F_x^g(t,f) = \int_{-\infty}^{+\infty} x(u) g(u-t) e^{-i2\pi f(t-u)} du, \qquad (1)$$

where $t \in \mathbb{R}$ and $f \in \mathbb{R}$ are interpreted as time and frequency respectively. The STFT offers a TF analysis of MCSs made of a superimposition of P modes:

$$x(t) = \sum_{p=1}^{P} A_p(t) e^{i2\pi\phi_p(t)}, \qquad (2)$$

where $A_p(t)$ and $\phi_p(t)$ are, respectively, the instantaneous amplitude and phase of the p-th mode. We assume $A_p > 0$ and slowly varying, and $\phi'_p > 0$ is called the instantaneous frequency (IF) of the p-th mode. For this type of signal, the ideal TFR reads

$$IT(t,f) = \sum_{p=1}^{P} A_p(t)\delta(f - \phi'_p(t)), \tag{3}$$

where $\delta(\cdot)$ is the Dirac delta distribution.

FSSTN "sharpens" the STFT by vertically reassigning the coefficients using the following formula :

$$T_x^{g,N}(t,f) = \frac{1}{g(0)} \int_\Gamma F_x^g(t,v)\delta(f - \tilde{\omega}_x^{[N]}(t,v))dv, \tag{4}$$

such that the new TFR is "closer" to the ideal one. Here, $\tilde{\omega}_x^{[N]}(t,f)$ is the N-th order IF estimation built assuming the phase of the modes of x locally behaves as a polynomial of order N [11], and $\Gamma = \{(t,f)/|F_x^g(t,f)| > 3\gamma\}$, where γ is estimated as $\tilde{\gamma} = \sqrt{2}\text{median}(|\Re\{F_x^g(t,f)\}|)/0.6745$ [12, 13, 14]. These definitions can be directly extended to the discrete TF setting, by defining

$$F_x^g[n,k] \approx F_x^g\left(\frac{n}{L}, k\frac{L}{M}\right), \tag{5}$$

where L is the length of the signal, M the number of frequency bins and L/M the frequency resolution. In this context, the discrete FSSTN reads

$$T_x^{g,N}[n,k] = \frac{1}{g(0)} \sum_{q=0}^{M-1} F_x^g[n,q]\chi[k - \lfloor\tilde{\omega}_x^{[N]}[n,q]\rceil], \tag{6}$$

where $\chi[c] = 1$ if $c = 0$ and 0 otherwise. Definition (6) leads to the following implementation for synchrosqueezing:

$$T_x^{g,N}[n, \lfloor\tilde{\omega}_x^{[N]}[n,k]\rceil] \leftarrow T_x^{g,N}[n, \lfloor\tilde{\omega}_x^{[N]}[n,k]\rceil] + F_x^g[n,k], \tag{7}$$

for $n = 0, 1, \ldots, L-1$ and $k = 0, 1, \ldots, M-1$. This means that each coefficient $F_x^g[n,k]$ is reassigned according to $\tilde{\omega}_x^{[N]}[n,k]$.

3. LOCALLY ADAPTING THE SYNCHROSQUEEZING ORDER

The stability and robustness to noise of IF estimation with SST was studied in [15]. Regarding the use of FSSTN for that purpose, one can identify at least three sources of errors: (i) modes with phases that locally depart from a polynomial of order N; (ii) ubiquitous noise, which alters the computation of the reassignment operators (the higher the order, the more affected they are by noise); and (iii) the distance to the ridge corresponding to a mode (ideally to the IF), since $|F_x^g[n,k]|$ decreases when one moves away from that ridge the impact of noise is greater.

Previous efforts have tried to adapt the order N of the synchrosqueezing transform. In [16], for instance, the authors defined $N = \mathfrak{F}[n,p]$, i.e. as a function of the discrete time instant n and of the mode p. Though, with such an approach, N is no longer fixed, it requires the knowledge of the TF ridges associated with each mode, which is not relevant when the signal contains crossing modes, for instance.

Considering that, for a given n, we propose to define $N = \mathfrak{G}[n,k]$, i.e. as a function of both discrete time and frequency, and no longer make any reference to TF ridges in the definition of $\mathfrak{G}$. More precisely, we first calculate for an n, N_0 such that:

$$N_0 = \underset{1\leq N\leq 4}{\text{argmin}}\, RE_{1D}(T_x^{g,N}[n,:]), \tag{8}$$

with

$$RE_{1D}(v) = \frac{1}{1-\alpha}\log\left(\sum_k \frac{v[k]^\alpha}{(\sum_a v[a])^\alpha}\right). \tag{9}$$

where v stands for a generic vector. By finding N_0 minimizing the Rényi entropy as described in (8), one obtains the order of the FSST that maximizes the TF concentration. Though there exist many different measures to study TF concentration, to consider the minimal Rényi entropy is a common choice for that purpose [17]. Then, still considering a fixed n, one orders the coefficients of the spectrogram at that time instant in descending order, and starting from the coefficient with the highest magnitude, one considers the order of FSST used to reassign the corresponding TF point that leads to the smaller Rényi entropy, the reassignment order of the other spectrogram points at n being left unchanged. Then, the same approach is applied to all the sorted spectrogram points at time n, and finally one makes the time index vary. The algorithm is summarized as follows:

Algorithm 1 Adaptive Synchrosqueezing

Input: $F_x^g[n,k]$, $T_x^{g,N}[n,k]$ and $\tilde{\omega}_x^{[N]}[n,k]$ for $N \in \mathcal{N} = \{1,\ldots,4\}$.
for $n = 0, \ldots, L-1$ **do**
  Set $N_0 = \arg\min_{N\in\mathcal{N}} RE_{1D}(T_x^{g,N}[n,:])$.
  Set $T_x^{g,A}[n,:] = T_x^{g,N_0}[n,:]$.
  Sort $|F_x^g[n,:]|$ defining the vector m s. t. $|F_x^g[n,m[:]]|$ is decreasing.
  for $i = 0, \ldots, L-1$ **do**
    Undo the reassignment: $T_x^{g,A}[n, \lfloor\tilde{\omega}_x^{[N_0]}[n,m[i]]\rceil] \leftarrow T_x^{g,A}[n, \lfloor\tilde{\omega}_x^{[N_0]}[n,m[i]]\rceil] - F_x^g[n,m[i]]$.
    for $P \in \mathcal{N}$ **do**
      Define $T_x^{g,A(P)}[n,:] = T_x^{g,A}[n,:]$.
      Compute

$$T_x^{g,A(P)}[n, \lfloor\tilde{\omega}_x^{[P]}[n,m[i]]\rceil] \leftarrow T_x^{g,A(P)}[n, \lfloor\tilde{\omega}_x^{[P]}[n,m[i]]\rceil] + F_x^g[n,m[i]].$$

11: **end for**
12: Set order $N^* = \arg\min_{P \in \mathcal{N}} RE_{1D}(T_x^{g,A(P)}[n,:])$.
13: Define $T_x^{g,A}[n,:] = T_x^{g,A(N^*)}[n,:]$.
14: **end for**
15: **end for**
16: **Output:** $T_x^{g,A}[n,k]$.

In terms of complexity, the algorithm needs the knowledge of 4 FFSTN. Then, at each TF point one computes 4 Rényi entropies of vectors of length M, which is the most important computational cost of the algorithm.

4. NUMERICAL RESULTS

In this section, we offer some numerical results on both synthetic and real signals. In order to assess the quality of the obtained TFRs, we use two measures. For the error with respect to the ideal TFR (according to Eq. (3)), we use the Earth mover's distance (EMD), also called the optimal transport distance [18]. EMD is computed as a sliced Wasserstein distance in order to compare probability distributions, and it was already used in the TF context [19, 20]. For the overall TF concentration, we use the 2D Rényi entropy

$$RE_{2D} = \frac{1}{1-\alpha} \log \left(\sum_n \sum_k \frac{R_x[n,k]^\alpha}{(\sum_a \sum_b R_x[a,b])^\alpha} \right), \quad (10)$$

where $R_x[n,k]$ stands for a generic TFR. A lower value of this measure would indicate a more concentrated representation. In the figures, we call our proposal 'FSSTa'.

4.1. Monocomponent signal

As a first example, we study a sinusoidal chirp of the form

$$x(t) = \cos\left(2\pi\left(256t - 100e^{-t}(10\pi\sin(10\pi t) - \cos(10\pi t))/1 + 100\pi^2\right)\right), \quad (11)$$

for $t \in [0,1]$ and sampled at 1024 Hz. For all TFRs, we used 512 frequency bins. The results can be appreciated in Fig. 1, where we used $N_{max} = 4$. Zoomed-in versions of FSST3, FSST4, and FSSTa evidence that our proposal achieves a better representation, more concentrated and closer to the ideal TFR. Most of the inaccuracies present in FSST3 and FSST4 are absent in FSSTa. Errors and 2D Rényi entropies are displayed for 3 different SNRs (error bars with means and standard deviations). It is clear that our new proposal outperforms the fixed order FSSTN both in concentration and error, in the range of tested SNRs.

To better illustrate this, we show the optimum orders selected by our algorithm for a SNR of 30 dB and 50 dB. We can observe here the action of noise. For a low level of noise (SNR equal to 50 dB), the algorithm usually selects $N = 4$.

For moderate noise (30 dB) we observe that the algorithm adapts the order to the local modulation. Indeed, for those

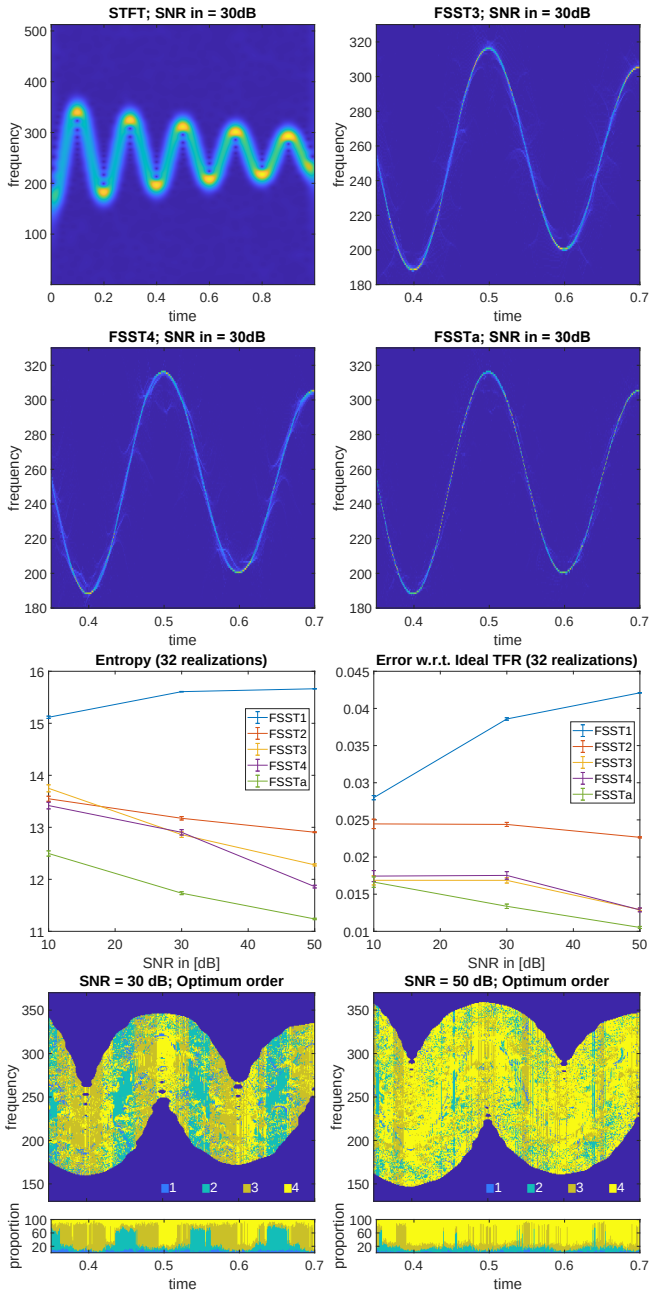

Fig. 1. Results for the monocomponent signal defined by Eq. (11). First row: modulus of the STFT and zoomed in version of the modulus of FSST3. Second row: zoomed-in versions of the moduli of FSST4 and FSSTa. Third row: 2D Rényi entropy and errors for the different methods, and for 3 SNRs. Fourth row: optimum orders for each TF point for which $|F_x^g(t,f)| > 3\tilde{\gamma}$, and proportions for each time instant.

portion where the mode oscillates as a linear chirp (ascending and descending parts of the sinusoidal) the algorithm gives preponderance to $N = 2$ (suitable for linear chirps). On the

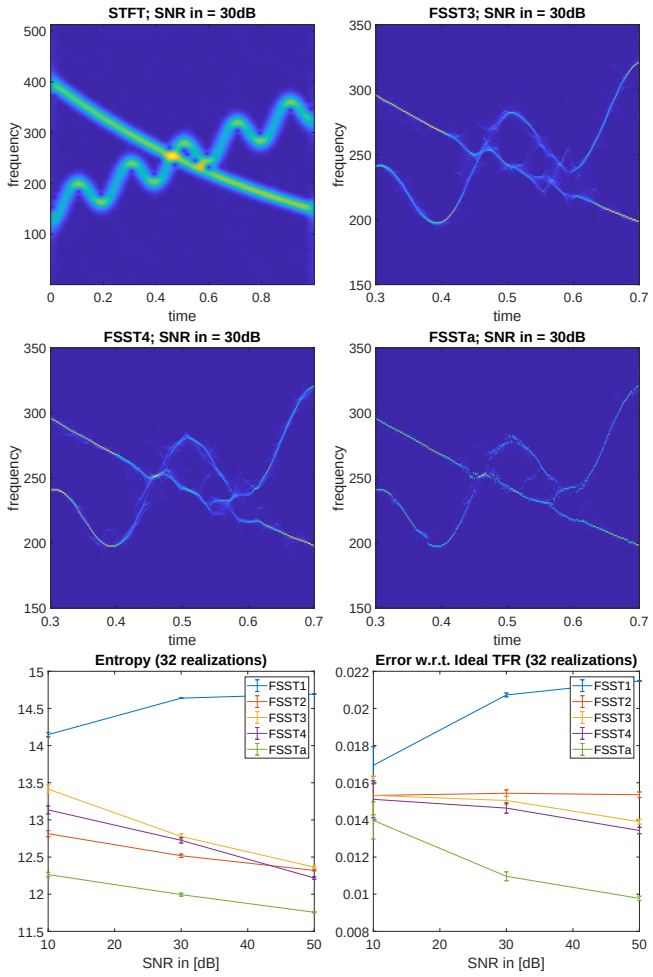

Fig. 2. Results for the multicomponent signal of Eq. (12).

other hand, for those portions where the sinusoidal frequency presents local extrema, the algorithm favors $N = 3$ and $N = 4$.

4.2. Multicomponent signal with crossing modes

As a second challenging example we study

$$y(t) = \cos(2\pi(100t^2 + 150t - \sin(10\pi t))) + \cos(2\pi(-400e^{-t})), \quad (12)$$

for $t \in [0, 1]$ and sampled at 1024 Hz. As before, we used 512 frequency bins and $N_{max} = 4$. The results are presented in Fig. 2. For this example, FSSTN (with $N = 1, \ldots, 4$) show undesirable behavior since the entropies associated with FSSTN with different orders "cross" each other when the noise level varies: while at 10 dB FSST2 has the lowest entropy, at 50 dB it is FSST4. Our proposal, on the other hand, achieves better results for all cases. The concentration of FSSTa is clearly the best, and it can be observed how the

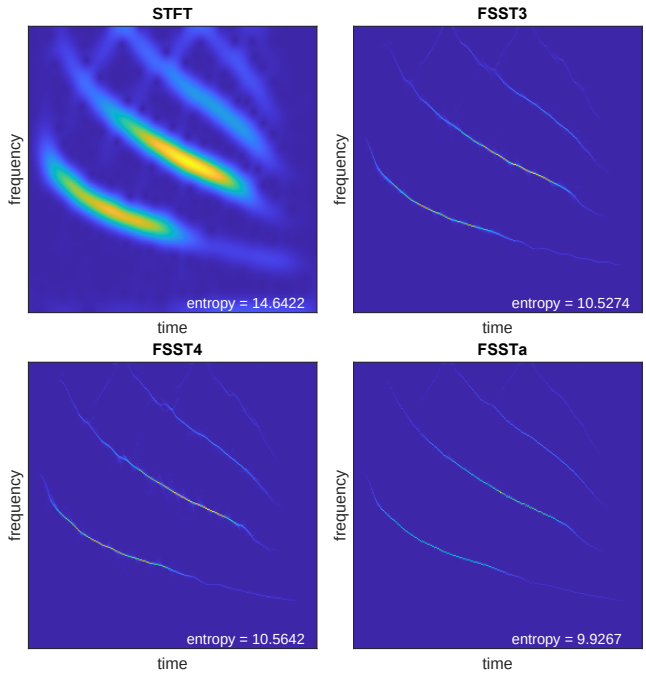

Fig. 3. Results for the bat signal.

IF of the frequency-decreasing mode is much better reconstructed. The "hairy" effect of FSSTN (that tends to create false ridges) is evidently alleviated in FSSTa.

4.3. Real signal

To deal with a real-world signal, we revisit the classical example of the bat echolocation call. The results are in Fig. 3, where we compute the 2D Rényi entropy for all four TFRs. Once again, FSSTa corresponds to the more concentrated TFR, which can be confirmed not only by its lowest entropy but also by visually inspecting the representations, where one can observe that FSSTa does not present the "blurring" present in FSST3 and FSST4.

5. CONCLUSIONS

In this work, we proposed a novel criterion to locally choose, in the time-frequency plane, the best order in the synchrosqueezing transform in terms of concentration and error with respect to the ideal TFR. Our proposal is based on the maximization of the concentration by means of a greedy algorithm that works time by time. We hypothesize the highest amplitude coefficients should be reassigned first, since those are dominant in terms of concentration and error. This approach resulted in better results than existing fixed order FSSTs. Future works will be related to deepen the mathematical aspects of this problem.

6. REFERENCES

[1] Leon Cohen, *Time-frequency analysis*, vol. 778, Prentice hall New Jersey, 1995.

[2] Patrick Flandrin, *Time-frequency/time-scale analysis*, Academic press, 1998.

[3] François Auger and Patrick Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Transactions on signal processing*, vol. 43, no. 5, pp. 1068–1089, 1995.

[4] Kunihiko Kodera, Roger Gendrin, and Claude Villedary, "Analysis of time-varying signals with small bt values," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, no. 1, pp. 64–76, 1978.

[5] Ingrid Daubechies and Stephane Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," *Wavelets in medicine and biology*, pp. 527–546, 1996.

[6] Ingrid Daubechies, Jianfeng Lu, and Hau-Tieng Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Applied and computational harmonic analysis*, vol. 30, no. 2, pp. 243–261, 2011.

[7] Thomas Oberlin, Sylvain Meignen, and Valérie Perrier, "The Fourier-based synchrosqueezing transform," in *2014 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2014, pp. 315–319.

[8] Sylvain Meignen, Thomas Oberlin, Philippe Depalle, Patrick Flandrin, and Stephen McLaughlin, "Adaptive multimode signal reconstruction from time–frequency representations," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, pp. 20150205, 2016.

[9] Marcelo A Colominas, Sylvain Meignen, and Duong-Hung Pham, "Time-frequency filtering based on model fitting in the time-frequency plane," *IEEE Signal Processing Letters*, vol. 26, no. 5, pp. 660–664, 2019.

[10] Thomas Oberlin, Sylvain Meignen, and Valérie 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, 2015.

[11] Duong-Hung Pham and Sylvain Meignen, "High-order synchrosqueezing transform for multicomponent signals analysis - with an application to gravitational-wave signal," *IEEE Transactions on Signal Processing*, vol. 65, no. 12, pp. 3168–3178, 2017.

[12] David L Donoho and Jain M Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[13] Stéphane Mallat, *A wavelet tour of signal processing: the sparse way*, Academic Press, 2008.

[14] Duong-Hung Pham and Sylvain Meignen, "A novel thresholding technique for the denoising of multicomponent signals," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 4004–4008.

[15] Matt Sourisseau, Hau-Tieng Wu, and Zhou Zhou, "Inference of synchrosqueezing transform–toward a unified statistical analysis of nonlinear-type time-frequency analysis," *arXiv preprint arXiv:1904.09534*, 2019.

[16] N Laurent and S Meignen, "A new adaptive technique for multicomponent signals reassignment based on synchrosqueezing transform," in *30th European Signal Processing Conference EUSIPCO 2022*, 2022.

[17] Richard G Baraniuk, Patrick Flandrin, Augustus JEM Janssen, and Olivier JJ Michel, "Measuring time-frequency information content using the Rényi entropies," *IEEE Transactions on Information theory*, vol. 47, no. 4, pp. 1391–1409, 2001.

[18] Ofir Pele and Michael Werman, "Fast and robust earth mover's distances," in *2009 IEEE 12th international conference on computer vision*. IEEE, 2009, pp. 460–467.

[19] Ingrid Daubechies, Yi Wang, and Hau-tieng 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, pp. 20150193, 2016.

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