A Novel Adaptive Approach to Mode Decomposition for Multicomponent Signals

Duong-Hung Pham, and Sylvain Meignen.

***Abstract*—This letter addresses the problem of the detection and estimation of the modes of a multicomponent signal using the reassignment framework. More precisely, we propose a new algorithm to estimate the ridges representing the time-frequency (TF) signatures of the modes based on the local orientation of the reassignment vector (RV), and use them to define the so-called "basins of attraction" enabling modes' retrieval. Compared with previous approaches, this new technique not only enable reconstruction of AM/FM modes but also Dirac impulses, which is of great interest in many practical situations. Numerical experiments conducted with both synthetic and real data illustrate the effectiveness of the technique.**

***Index Terms*—Time-frequency, reassignment, synchrosqueezing, AM/FM, multicomponent signals.**

I. Introduction

In the signal processing community, multicomponent signals (MCS), defined as a superposition of amplitude- and frequency-modulated (AM-FM) modes, have received a considerable attention [1], [2]. Indeed, they enable to very accurately represent a large class of signals arising from audio recordings (music, speech), meteorology, structural stability analysis [3]–[5], or medical data (electrocardiogram, thoracic and abdominal movement signals) [6], [7]. Linear techniques, as for instance the short-time Fourier transform (STFT), are commonly used to characterize them in the TF plane. However, an inherent limitation of such methods, known as the "uncertainty principle", stipulates that one cannot localize a signal with arbitrary precision both in time and frequency. Many efforts have been made to cope with this issue and one of them, called the *reassignment method* (RM), received a considerable attention [8]. RM however does not allow for mode reconstruction, contrary to another phase-based technique called *synchrosqueezing transform* (SST), introduced in [9]. Unfortunately, this technique cannot deal with MCS containing modes with strong frequency modulation and irregular amplitudes or Dirac impulses. Regarding the frequency modulation, a novel technique, called *second order synchrosqueezing* (VSST), was developed in [10] and further theoretically studied in [11]. Nevertheless, these techniques assuming the MCS is made of a fixed number of AM/FM modes, cannot deal with vanishing modes or Dirac impulses.

One key ingredient for mode reconstruction is the estimation of its TF signature, the knowledge of which enables definition of plenty different techniques other than those based on RM [12], [13]. The type of modes sought conditions the technique for TF signatures estimation: some approaches concentrate on ridge detection for AM/FM modes [14], [15], while some others, using the properties of the *reassignment vector* (RV), can handle a wider class of TF signatures, the constraints on the modes being less stringent [16] [17]–[19]. In this latter case, the estimated TF signatures are then used to define *basins of attraction* (BAs) for the modes enabling their reconstruction. However, the just mentioned approaches based on RV fail to assess the TF signature associated with a noisy Dirac impulse. To improve the behavior of techniques based on RV on that type of modes while preserving their main characteristics on AM/FM modes is the aim of the present paper.

To do so, after having introduced some useful definitions in Section II, we recall that of RV and existing approaches based on the latter for TF signature estimation in Section III. Then, we introduce, in Section IV, our new TF signature estimator based on the local orientation of RV, and then show how to perform mode reconstruction using BAs. Finally, numerical simulations in Section V demonstrate the improvement brought by the proposed new technique both on a complex simulated MCS and a real signal.

II. Basic Definitions

For a given discrete signal $f \in L^2(\mathbb{R})$, *STFT* corresponds to:

$$V_f^g(t,\omega) = \int_{\mathbb{R}} f(u) g(u-t) e^{-i2\pi\omega(u-t)}\, du, \tag{1}$$

where the window g is assumed to be real-valued. The spectrogram is then defined as $|V_f^g(t,\omega)|^2$. When $f = \delta_{t_0}$ is a Dirac distribution at t_0, the above definition can be extended using the duality product in the space of distributions as follows:

$$V_f^g(t,\omega) = \langle \delta_{t_0}, g(.-t) e^{-i2\pi\omega(.-t)} \rangle = g(t_0-t) e^{-i2\pi\omega(t_0-t)}. \tag{2}$$

Note that, in such a case, the amplitude of STFT is the same whatever ω. In the sequel, we study MCSs defined as a superposition of modes:

$$f(t) = \sum_{k=1}^{K} f_k(t) \quad \text{with } f_k(t) = A_k(t) e^{i2\pi\phi_k(t)} \text{or } f_k = A_k \delta_{t_k} \tag{3}$$

for some finite $K \in \mathbb{N}$, a priori unknown. $A_k(t)$ and $\phi_k(t)$ are respectively instantaneous amplitude (IA) and phase (IP), $\phi_k'(t)$ being referred to as the instantaneous frequency (IF) of

D-H Pham and S. Meignen are with the Jean Kuntzmann Laboratory, University of Grenoble-Alpes, and CNRS, Grenoble 38041, France (email:duong-hung.pham@imag.fr and sylvain.meignen@imag.fr).

mode f_k at time t. Note also that, in our context, $A_k(t)$ needs not be continuous. Also, we define the *TF signature* of the first type of mode as $(t, \phi_k'(t))$ while for the Dirac impulse it corresponds to (t_k, ω).

III. RV Based TF Signature Estimation

A. RV Definition and Illustrations

The key idea of the reassignment method (RM) [8] is to map a TF representation, as for instance the spectrogram, to the location corresponding to the TF signature of the nearest mode. This corresponds to the centroid of the distribution:

$$\begin{aligned}\hat{\tau}_f(t,\omega) &= t - \frac{1}{2\pi}\partial_\omega \left\{\arg(V_f^g(t,\omega))\right\}\\ \hat{\omega}_f(t,\omega) &= \frac{1}{2\pi}\partial_t \left\{\arg(V_f^g(t,\omega))\right\},\end{aligned} \tag{4}$$

where $\arg(Z)$ is the argument of complex number Z. These can be computed through [8]:

$$\begin{aligned}\hat{\tau}_f(t,\omega) &= t + \Re\left\{\frac{V_f^{tg}(t,\omega)}{V_f^g(t,\omega)}\right\}\\ \hat{\omega}_f(t,\omega) &= \omega - \frac{1}{2\pi}\Im\left\{\frac{V_f^{g'}(t,\omega)}{V_f^g(t,\omega)}\right\},\end{aligned} \tag{5}$$

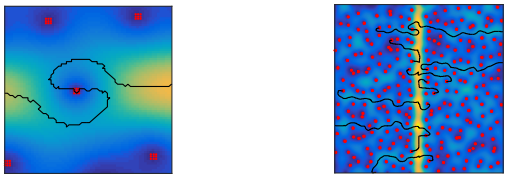
where $V_f^{tg}, V_f^{g'}$ are respectively STFTs of f computed with windows $t \mapsto tg(t), g'(t)$ and $\Re\{Z\}$ (*resp.* $\Im\{Z\}$) is the real (*resp.* imaginary) part of the complex number Z. With this in mind, RV is then defined as [16]:

$$RV(t,\omega) = \left(\ \hat{\tau}_f(t,\omega) - t,\ \hat{\omega}_f(t,\omega) - \omega\ \right). \tag{6}$$

As an illustration, it is easy to see that for $f = \delta_{t_0}$, $RV(t,\omega) = (t_0 - t, 0)$: it has a component only along the time axis. Conversely, for a purely harmonic mode, $f(t) = e^{i2\pi\omega_0 t}$, one has $V_f^g(t,\omega) = \hat{g}(\omega-\omega_0)e^{2i\pi\omega_0 t}$, where $\hat{g}$ is the Fourier transform of g, and thus $\dfrac{V_f^{tg}(t,\omega)}{V_f^g(t,\omega)} = \dfrac{\widehat{tg}(\omega-\omega_0)}{\hat{g}(\omega-\omega_0)}$ which is an imaginary complex number when g is even. Similarly, one has $\frac{V_f^{g'}(t,\omega)}{V_f^g(t,\omega)} = \frac{\widehat{g'}(\omega-\omega_0)}{\hat{g}(\omega-\omega_0)} = 2i\pi(\omega-\omega_0)$, so that, $RV(t,\omega) = (0, \omega_0 - \omega)$: it has a component only along the frequency axis. Another simple case is that of a constant amplitude linear chirp $f(u) = A(t)e^{2i\pi(\phi(t)+(u-t)\phi'(t)+\frac{(u-t)^2}{2}\phi''(t))}$ (t fixed), whose STFT reads: $V_f^g(t,\omega) = \widehat{f(t)g(u)e^{i\pi\phi''(t)u^2}}(\omega-\phi'(t))$. When $g(t) = e^{-\sigma\pi t^2}$, the following two relations can be easily proven:

$$\begin{aligned}V_f^{tg}(t,\omega) &= \frac{1}{-2\pi\sigma + 2i\pi\phi''(t)}\widehat{(g(u)e^{i\pi\phi''(t)u^2})'}(\omega-\phi'(t))\\ &= \frac{i(\omega-\phi'(t))}{-\sigma + i\phi''(t)}V_f^g(t,\omega)\\ V_f^{g'}(t,\omega) &= -2\pi\sigma V_f^{tg}(t,\omega),\end{aligned}$$

leading to $RV(t,\omega) = \frac{(\omega-\phi'(t))}{\sqrt{\sigma^2+\phi''(t)^2}}(-\phi''(t), \sigma^2)$. The IF of the mode being a straight line whose orientation is given by vector $(1, \phi''(t))$, RV is orthogonal to the ridge corresponding to the TF signature only if $\sigma = 1$ (the window is unitary in L^2). More generally it points to that ridge following the direction $(-\phi''(t), \sigma^2)$.



(a) (b)

Fig. 1. (a): a close-up of a zero of the spectrogram and its corresponding contour computed with method M_1; (b): STFT of a noisy Dirac impulse (SNR = 0 dB) along with the first 10 contours computed with method M_1.

B. Definitions of Contour Points

This section reviews two existing approaches to define *contour points*, corresponding to the TF signature, based on the projection of RV in some specific direction. Indeed, we have just seen that RV points to the ridge associated with the TF signature of a mode, which means that, when crossing a ridge, RV undergoes a strong variation in its orientation. To determine the location of these sudden orientation changes, a first strategy was developed in [16], and consisted in projecting RV in a specific direction, given by an angle θ, and then in determining the location of the sign change of the projection. Thus, *contour points* (CPs) were defined as the zeros of $\langle RV(t,\omega), v_\theta\rangle$, where v_θ is the unit vector in the direction θ, and $\langle .,. \rangle$ denotes the inner product. Note that the direction of projection θ being fixed a priori, the technique does not adapt well to the determination of CPs corresponding to varying orientations. To deal with this problem, an improved technique to compute CPs was proposed in [17], [18]. It first consisted in remarking that, due to the discrete nature of the studied signals, RV should be viewed as a displacement on a grid not a vector of with real coordinates. Indeed, if the signal is supposed to be defined on $0, \cdots, M-1$, the STFT is evaluated at frequencies $\frac{p}{N}$, $p = 0, \cdots, N-1$ (N is the number of frequency bins), so the grid is indexed by (k,p), k denoting one time instant,. By rounding to the nearest integers both in time and frequency the coordinates of RV, one obtains RV_r and then defines a new set of CPs by projecting this vector as follows:

$$\alpha(k,p) := \langle RV_r(k, \frac{p}{N}), v_{\theta_r(k,\frac{p}{N})\mod\pi}\rangle = 0 \tag{7}$$

with $\theta_r(k, \frac{p}{N})$ the argument of $RV_r(k, \frac{p}{N})$ (we consider $(\theta_r(k, \frac{p}{N}) \mod \pi) \in [0, \pi[$). This alternative allows to define a new type of CPs that no longer depend on a fixed angle θ. However, this technique, called M_1 in the sequel, suffers from some serious limitations [19]. Firstly, special structures are created in the vicinity of the zeros of the spectrogram since the mod π computation induces $\alpha(k,p)$ to be zero on horizontal TF lines crossing the zeros. Secondly, it is not capable of detecting vertical ridges, still because of the mod π factor, which generates numerical instabilities. All these phenomena are respectively illustrated in Figure 1 (a) and (b) (the CPs

are practically chained by considering level zero contours of $\alpha(k,p)$ using *contourc* MATLAB function).

IV. New Adaptive Determination of Contours Points and Basins of Attraction

A. Contour Estimation Based on Local RV Orientation

The applicability of the just recalled approaches based on the projection of RV_r to compute CPs is hindered by the fact that the orientation of RV_r, in the vicinity of the TF signature of a mode, fluctuates, and all the more so that the noise level increases. This section introduces a new adaptive algorithm that uses a criterion based on a local rather than punctual orientation of RV_r to define a direction of projection. This results in a much more robust estimation of the TF signature of modes like Dirac impulses, even at high noise level, while maintaining a good behavior for AM/FM modes.

We first investigate the impact of viewing RV as a displacement on a grid not a vector of with real coordinates. To do so, we depict the distributions of the argument of RV or RV_r (both taken modulo π), respectively in Figures 2 (a), (b), when the signal is a white Gaussian noise. We remark that the argument of RV (modulo π) is almost uniformly distributed in all directions while RV_r (modulo π) clearly favors four directions: $0, \pi/4, \pi/2$ and $3\pi/4$. Despite these four orientations are not informative if one considers the whole TF plane, we now see that they are features enabling definition of new local direction of projection for RV_r, called *local projection angles* (LPAs), These help improve the performance of the estimator of TF signature based on the projection of the reassignment vector.

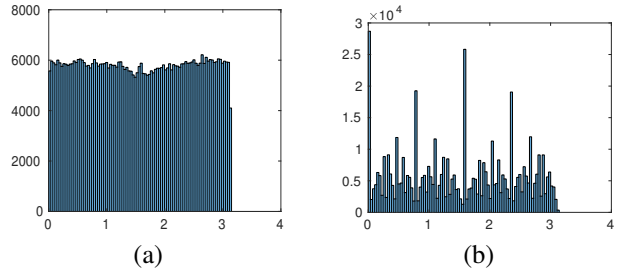


Fig. 2. (a): Histogram of the argument of RV modulo π, for a white Gaussian noise; (b): Same as (a) but with RV having its coordinates rounded to the nearest integers both in time and frequency (RV_r), prior to histogram computation.

We define, at each grid point (p,k), LPA as the most frequent value of θ_r in a square neighborhood of size $(2T_s+1)^2$ centered at the point of study. We then project RV_r on the direction given by LPA and define the new CPs as the zeros of the projection, as summarized in the following algorithm (the *mode* function returns the most frequent value in an array):

We are going to show that the direction of projection is stabilized by using the proposed local estimation. T_s, controlling the size of the neighborhood, should have a great impact on

Algorithm 1 LPA Algorithm

Input: RV_r
$\theta_r := mod\,(arg(RV_r), \pi)$, $[M,N] := size(RV_r)$
for $(k,p) \in \{0,\cdots,M-1\} \times \{0,\cdots,N-1\}$ **do**
  $tmp = \theta_r(max(0,k-T_s) : min(M-1,k+T_s),$
  $\qquad max(0,p-T_s) : min(N-1,p+T_s))$
  $lpa(k,p/N) = mode(tmp)$
  $\alpha(k,p) := \langle RV_r(k,p/N), v_{lpa(k,p/N)} \rangle$
Define CPs as the zeros of α

CPs computation and will be further studied in Section V (M_1 corresponds to $T_s = 0$).

B. Determination of Basins of Attraction Using RV and Mode Reconstruction

Using the just estimated modes' TF signatures, the *basin of attraction* (BA) associated with a mode, is defined as in [18], [19]. Once the BAs are computed, each corresponding mode f_i is reconstructed through:

$$f_i(t) = \frac{1}{g(0)} \int_{(t,\omega)\in\mathcal{B}_i} V_f^g(t,\omega) d\omega. \tag{8}$$

where $\mathcal{B}_i \subset \mathbb{R}^2$ is the BA associated with mode i.

V. Numerical Experiments

This section investigates the properties of the proposed algorithm for mode TF signature identification, signal denoising and mode reconstruction.

A. Numerical Results

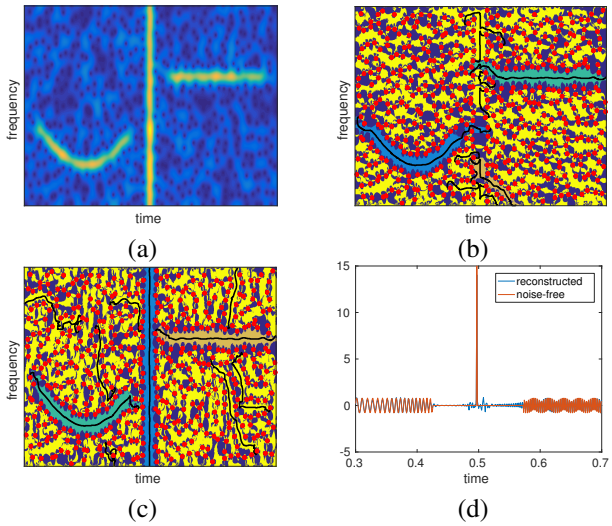


Fig. 3. (a): spectrogram of the simulated signal (SNR 0 dB); (b): BAs associated with the first 10 contours computed by method M_1; (c): same as (a) but computed with LPA algorithm (with $T_s = 30$); (d): reconstructed signal based on the coefficients contained in the three most energetic BAs depicted in (c) along with the original noise-free signal.

Let us first consider a simulated MCS composed of three components: a Dirac impulse, a cosine chirp, and a purely harmonic mode. This signal is then contaminated by an additive white Gaussian noise (Signal-to-Noise Ratio (SNR) 0 dB, MCS is sampled at a rate $N = 1024$Hz on $[0, 1]$ and STFT is computed with the Gaussian window $\sigma = 1$). We first display in Figure 3 (a) the spectrogram of the signal. Then, we depict respectively in Figures 3 (b) and (c) the basins of attraction along with the first 10 contours computed with method M_1 and the algorithm based on LPA. It is clear that the former cannot detect the Dirac impulse, whereas the latter manages to capture the TF structures associated with the three modes. Finally, we illustrate in Figure 3 (d) the reconstruction of the signal by selecting the coefficients associated with the three most energetic BAs (in cyan, orange and blue for decreasing energy order) displayed in Figure 3 (c). The output SNR after reconstruction are 9.6, 12.5 and 10.5 dB for the cosine chirp, Dirac impulse, and purely harmonic mode respectively, meaning that the algorithm not only estimates the modes but also performs some kind of denoising.

B. Sensitivity to Parameter T_s of LPA Algorithm

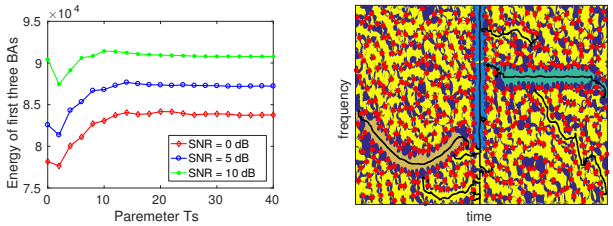


Fig. 4. (a): Energy contained in the three most energetic BAs computed by LPA algorithm on Fig.3 (a) for different values of T_s and noise levels; (b): BAs and the first 10 contours computed by LPA algorithm with $T_s = 6$.

The issue we now discuss is how to choose an appropriate parameter T_s for a specific signal so as to compute CPs efficiently. The measure we use is the energy contained in the first K most energetic BAs with respect to T_s:

$$E_f(T_s) = \sum_{i=1}^{K} \sum_{(t,\omega)\in \mathcal{BA}_i^{T_s}} |V_f^g(t,\omega)|^2, \tag{9}$$

where $\mathcal{BA}_i^{T_s}$ is the ith BA. The larger the quantity E_f, the better the computation of CPs (provided K is meaningful for the studied signal). In Figure 4 (a), we display $E_f(T_s)$ for $K = 3$, for the MCS of Figure 3 (a), and at three different noise levels (SNR = 0, 5 and 10 dB). We remark that E_f fluctuates when T_s is small whatever the noise level and then stagnates when some particular value for T_s is reached. The reason for such a behavior is that when T_s is small and for the Dirac impulse, the neighborhood is too small to enable the determination of a stable direction of projection. As a result, only part of the contour associated with the Dirac impulse is taken into account in the first three contours resulting in a lower E_f. As an illustration, BAs associated with the first 10 contours when $T_s = 6$ and $T_s = 30$ are shown in Figure 4 (b) and Figure 3 (c)), respectively.

C. Application to Real Signal

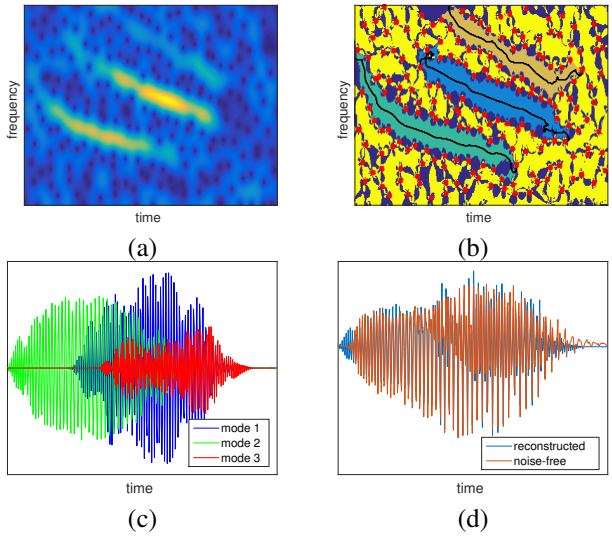


Fig. 5. (a): the spectrogram; (b): BAs associated with the first 3 contours computed with LPA algorithm; (c): three reconstructed modes based on the coefficients contained in the three most energetic BAs depicted in (b); (d): reconstructed signal along with the original noise-free signal.

We now illustrate our new technique on a bat echolocation signal, made of 400 samples recorded at 143 Hz, to which a white Gaussian noise is added such that the input SNR equals 5.0 dB. The spectrogram of the noisy bat signal is displayed in Figure 5 (a). Then, it can be seen from Figure 5 (b) that the BAs corresponding to the three main components of the echolocation signal are well estimated by LPA algorithm, enabling the reconstruction of the three detected modes (Figure 5 (c)). Finally, we compare the total resulting signal with the original noise-free signal. The output SNR of the final reconstruction is 10.9 dB, which confirms the potential interest of our new technique for the denoising of real MCS.

VI. Conclusion

In this letter, we have introduced a new technique to estimate the TF signatures or contours of the modes of multicomponent signals by projecting the reassignment vector along its local orientation. We then defined basins of attraction as the set of coefficients associated with these contours used the latter to reconstruct the modes. The technique proves to be efficient to reconstruct non AM/FM modes like Dirac impulses or discontinuous modes even at high noise level and can be profitably used to denoise real multicomponent signal as for instance bat echolocation calls.

References

[1] K. Kodera, R. Gendrin, and C. 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, Feb 1978.

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

[3] M. Costa, A. A. Priplata, L. A. Lipsitz, Z. Wu, N. E. Huang, A. L. Goldberger, and C.-K. Peng, "Noise and poise: Enhancement of postural complexity in the elderly with a stochastic-resonance-based therapy," *Europhysics Letters (EPL)*, vol. 77, no. 6, p. 68008, Mar 2007.

[4] D. A. Cummings, R. A. Irizarry, N. E. Huang, T. P. Endy, A. Nisalak, K. Ungchusak, and D. S. Burke, "Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand," *Nature*, vol. 427, no. 6972, pp. 344 –347, Jan 2004.
[5] N. E. Huang and Z. Wu, "A review on Hilbert-huang transform: Method and its applications to geophysical studies," *Reviews of Geophysics*, vol. 46, no. 2, Jun 2008.
[6] Y. Y. Lin, H.-T. Wu, C. A. Hsu, P. C. Huang, Y. H. Huang, and Y. L. Lo, "Sleep apnea detection based on thoracic and abdominal movement signals of wearable piezo-electric bands," *IEEE Journal of Biomedical and Health Informatics*, 2016.
[7] C. L. Herry, M. Frasch, A. J. Seely, and H.-T. Wu, "Heart beat classification from single-lead ecg using the synchrosqueezing transform," *Physiological Measurement*, vol. 38, no. 2, pp. 171–187, 2017.
[8] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE J SP*, vol. 43, no. 5, pp. 1068–1089, 1995.
[9] 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.
[10] 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.
[11] R. Behera, S. Meignen, and T. Oberlin, "Theoretical analysis of the second-order synchrosqueezing transform," *Applied and Computational Harmonic Analysis*, Nov 2016.
[12] S. Meignen, T. Oberlin, and S. McLaughlin, "A new algorithm for multicomponent signals analysis based on synchrosqueezing: With an application to signal sampling and denoising," *IEEE J SP*, vol. 60, no. 11, pp. 5787–5798, 2012.
[13] T. Oberlin, S. Meignen, and S. Mclaughlin, "A novel time-frequency technique for multicomponent signal denoising," in *Proceedings of 21st European Signal Processing Conference (EUSIPCO-13)*, 2013, pp. 1–5.
[14] N. Delprat, B. Escudie, P. Guillemain, R. Kronland, P. Tchamitchian, and B. Torresani, "Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies," *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 644–664, 1992.
[15] D. E. Newland, "Ridge and phase identification in the frequency analysis of transient signals by harmonic wavelets," *Journal of Vibration and Acoustics*, vol. 121, no. 2, pp. 149–155, 1999.
[16] Y. Lim, B. Shinn, and T. Gardner, "Sparse contour representations of sound," *Signal Processing Letters, IEEE*, vol. 19, no. 10, pp. 684–687, Oct 2012.
[17] S. Meignen, T. Oberlin, P. Depalle, P. Flandrin, and S. McLaughlin, "Adaptive multimode signal reconstruction from time–frequency representations," *Phil. Trans. R. Soc. A*, vol. 374, no. 2065, p. 20150205, 2016.
[18] S. Meignen, T. Gardner, and T. Oberlin, "Time-frequency ridge analysis based on reassignment vector," in *Proceedings of the 23st European Signal Processing Conference (EUSIPCO-15)*. EURASIP, 2015.
[19] S. Meignen, T. Oberlin, and S. McLaughlin, "Fully adaptive mode decomposition from time-frequency ridges," in *The 42nd IEEE International Conference on Acoustics, Speech and Signal Processing*. ICASSP2017, 2017.