



HAL
open science

On Demodulation, Ridge Detection and Synchrosqueezing for Multicomponent Signals

Sylvain Meignen, Duong-Hung Pham, Steve Mclaughlin

► **To cite this version:**

Sylvain Meignen, Duong-Hung Pham, Steve Mclaughlin. On Demodulation, Ridge Detection and Synchrosqueezing for Multicomponent Signals. IEEE Transactions on Signal Processing, 2017, 65 (8), pp.2093-2103. 10.1109/TSP.2017.2656838 . hal-01586414

HAL Id: hal-01586414

<https://hal.science/hal-01586414v1>

Submitted on 12 Sep 2017

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.

On Demodulation, Ridge Detection and Synchronsqueezing for Multicomponent Signals

Sylvain Meignen Duong-Hung Pham and Steve McLaughlin, FIEEE

Sylvain Meignen
LJK,
700 Avenue Centrale, Campus
de Saint Martin d'Hères,
38401 Domaine Universitaire
de Saint-Martin-d'Hères,
France
Tel 0033 4 57 42 17 66
sylvain.meignen@imag.fr

Duong-Hung Pham
LJK,
700 Avenue Centrale, Campus
de Saint Martin d'Hères,
38401 Domaine Universitaire
de Saint-Martin-d'Hères,
France
Tel 0033 4 57 42 18 33
duong-hung.pham@imag.fr

Steve McLaughlin
School of Engineering and
Physical Sciences,
Heriot-Watt University,
Edinburgh Scotland, UK
Tel 0044 131 451 4127
s.mclaughlin@hw.ac.uk

Abstract

In this paper, we present a novel technique for the retrieval of the modes of a multicomponent signal using a time-frequency (TF) representation of the signal. Our approach is based on a novel ridge extraction method that takes into account the fact that the TF representation is both discrete in time and frequency, followed by a demodulation procedure. Numerical results show the benefits of the proposed approach for mode reconstruction in comparison to similar techniques that do not make use of demodulation. Furthermore, numerical investigations show that the proposed approach sharpens the TF representation on which it is built.

Index Terms

time-frequency, reassignment, synchronsqueezing, AM/FM demodulation, multicomponent signals.

EDICS Category: DSP-TFSR, DSP-TRSF

The authors acknowledge the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-13-BS03-0002-01 (ASTRES) and of the EPSRC via grant EP/J015180/1..

On Demodulation, Ridge Detection and Synchrosqueezing for Multicomponent Signals

I. INTRODUCTION

There are many physical systems which generate complex signals that are often modeled as a sum of amplitude and frequency-modulated (AM–FM) waves. These signals are generally referred to as multicomponent signals (MCSs). In many situations, it is often desirable and necessary to decompose these MCSs into their individual components. As a result of their computational simplicity and efficiency, linear time-frequency (TF) transforms such as the short-time Fourier transform (STFT) and the continuous wavelet transform (CWT) have received considerable attention to this end, over the last 40 years. The STFT of a MCS determines *ridges* in the TF plane which, once detected, allow for the reconstruction of the different components based on the TF representation evaluated on these ridges [1]. More recently, it has been shown in [2][3] that local frequency integration can improve the robustness to noise of the reconstruction. An alternative approach to the reconstruction of the modes, based on the information computed on the ridge, is known as the *synchrosqueezing transform* (SST) originally introduced in [4], and theoretically studied in [5]. In essence, the technique consists of enhancing the time-scale (TS) representation, given by a wavelet transform, by reassigning the coefficients using an estimate of the instantaneous frequency (IF). Such a technique is easily transferable to the TF representation given by the STFT [6][7]. However, in their original formulations, these techniques are not well suited for signals made of non purely harmonic modes, therefore an extension of SST was recently proposed to better deal with this case, and is known as *second order synchrosqueezing transform* [8].

In this present paper, we introduce a novel demodulation technique, built on the second order synchrosqueezing transform that leads to an even sharper representation along with better mode reconstruction results. A previous attempt which involved demodulating the signal before applying SST was presented in [9]. In that paper, the demodulation was based on the computation of the phase of the analytic signal associated with the MCS. However, it is well known that this phase cannot be related to the instantaneous frequency (IF) of the modes which the MCS consists of. Indeed, to demodulate a MCS requires estimates of the IF of the modes. In relation to this issue, in [10], an estimate of the IF of a mono-component signal was computed using local frequency extrema of the spectrogram. An iterative procedure was proposed to accurately estimate the IFs of the modes, but mode reconstruction was not discussed.

In this paper, we first introduce a novel ridge estimation technique, based on the second-order SST method introduced in [8], followed by demodulation and finally mode reconstruction. The benefits of using such a procedure is an improvement in the sharpness of the TF representation obtained and also in the mode reconstruction performance. The paper is structured as follows, after this brief introduction, we recall the basics of STFT-based SST (Section II). Then, we focus on the practical implementation of ridge estimation in Section III, and move on to the definition of the demodulation procedure followed by the reconstruction algorithm based on the demodulated signal (Section IV). Numerical examples showing the relevance of the proposed approach on both simulated and real data conclude the paper.

II. BACKGROUND TO FOURIER SYNCHROSQUEEZING TRANSFORM

Prior to starting, it is useful to define our notation, and remind readers of the basic elements of STFT based SST (FSST) and of the approach to FSST which is better adapted to deal with modulated modes.

A. Basic Definitions and Notation

Let f be a function in $L^1(\mathbb{R})$, the space of integrable functions, we denote by \hat{f} the Fourier transform of f , defined using the following normalization:

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

Taking a window g in the Schwartz class, the space of smooth functions with fast decaying derivatives of any order, the (modified) STFT of f is defined by

$$V_f^g(\eta, t) = \int_{\mathbb{R}} f(\tau)g^*(\tau - t)e^{-i2\pi\eta(\tau - t)} d\tau, \quad (2)$$

where $g^*(t)$ is the complex conjugate of $g(t)$.

In what follows, we investigate the retrieval of the components f_k of a MCS f defined by:

$$f(t) = \sum_{k=1}^K A_k(t)e^{2i\pi\phi_k(t)} = \sum_{k=1}^K f_k(t), \quad (3)$$

where $A_k(t) > 0$ and $\phi_k'(t) > 0$. The signal f is completely defined by its so-called ideal TF representation as follows:

$$TI_f(\omega, t) = \sum_{k=1}^K A_k(t)\delta(\omega - \phi_k'(t)). \quad (4)$$

B. The Basics of FSST

The aim of STFT-based SST (FSST) [6][7] is to retrieve the ideal TF representation of f from its STFT, based on estimation of the IF at time t and frequency η :

$$\hat{\omega}_f(\eta, t) = \frac{1}{2\pi} \partial_t \arg V_f^g(\eta, t) = \eta - \mathcal{I}m \left\{ \frac{1}{2\pi} \frac{V_f^{g'}(\eta, t)}{V_f^g(\eta, t)} \right\}, \quad (5)$$

where $\mathcal{I}m \{X\}$ denotes the imaginary part of complex number X . The principle of FSST is to reassign the complex coefficients $V_f^g(\eta, t)$ according to the following map $(\eta, t) \mapsto (\hat{\omega}_f(\eta, t), t)$, by means of the *synchrosqueezing operator*:

$$T_f^g(\omega, t) = \frac{1}{g^*(0)} \int_{\{\eta, |V_f^g(\eta, t)| \geq \gamma\}} V_f^g(\eta, t) \delta(\omega - \hat{\omega}_f(\eta, t)) d\eta, \quad (6)$$

where δ is the Dirac distribution. Knowing ϕ'_k , the k th mode can then be retrieved by considering:

$$f_k(t) \approx \int_{|\omega - \phi'_k(t)| < d} T_f^g(\omega, t) d\omega. \quad (7)$$

Essentially, FSST reassigns the information in the TF plane and then makes use of this sharpened representation to recover the modes. Previous theoretical investigations [8][6] have highlighted a set of signals on which the performance of FSST can be evaluated. These are defined as follows:

Definition II.1. Let $\varepsilon > 0$. We define the set $\mathcal{B}_{\Delta, \varepsilon}$ of MCS where

- for all k , f_k 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 , i.e. for all $k \in \{1, \dots, K-1\}$ and for all t ,

$$\phi'_{k+1}(t) - \phi'_k(t) > 2\Delta. \quad (8)$$

The synchrosqueezing operator with threshold $\gamma > 0$ and accuracy parameter $\lambda > 0$ is then defined, using a function $\rho \in \mathcal{D}(\mathbb{R})$, in the space of compactly supported smooth functions, such that $\int_{\mathbb{R}} \rho(x) dx = 1$, as:

$$T_f^{\lambda, \gamma}(\omega, t) = \frac{1}{g(0)} \int_{\{\eta, |V_f^g(\eta, t)| > \gamma\}} V_f^g(\eta, t) \frac{1}{\lambda} \rho\left(\frac{\omega - \hat{\omega}_f(\eta, t)}{\lambda}\right) d\eta. \quad (9)$$

This definition allows us to state the main approximation result of FSST [6]:

Theorem II.1. Let $f \in \mathcal{B}_{\Delta, \varepsilon}$, $\tilde{\varepsilon}^3 = \varepsilon$, and g be a window in the Schwartz class, such that \hat{g} is compactly supported in $[-\Delta, \Delta]$. Then, if ε is small enough,

- $|V_f^g(\eta, t)| > \tilde{\varepsilon}$ only if there exists $k \in \{1, \dots, K\}$ s. t. $(\eta, t) \in Z_k := \{(\eta, t), \text{ s.t. } |\eta - \phi'_k(t)| < \Delta\}$.

- For all $k \in \{1, \dots, K\}$ and all pair $(\eta, t) \in Z_k$, s. t. $|V_f^g(\eta, t)| > \tilde{\varepsilon}$, one has

$$|\hat{\omega}_f(\eta, t) - \phi'_k(t)| \leq \tilde{\varepsilon}. \quad (10)$$

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

$$\left| \lim_{\lambda \rightarrow 0} \left(\int_{|\omega - \phi'_k(t)| < \tilde{\varepsilon}} T_f^{\lambda, \tilde{\varepsilon}}(\omega, t) d\omega \right) - f_k(t) \right| \leq C\tilde{\varepsilon}. \quad (11)$$

A detailed proof is available in [6].

C. Second-Order Synchrosqueezing

One of the limitations of FSST is that it does not allow for the reconstruction of modes subject to significant frequency modulation. A method was recently reported to deal with mode modulation in the FSST context via *second order synchrosqueezing* (VSST) [8][11]. This technique uses a second order approximation of the phase of the modes in the definition of the synchrosqueezing operator. VSST is based on a new complex estimate of the second order derivative of the phase of f , defined as follows:

$$\tilde{q}_f(\eta, t) = \frac{\partial_t(\partial_t V_f^g(\eta, t)/V_f^g(\eta, t))}{2i\pi - \partial_t(\partial_\eta V_f^g(\eta, t)/V_f^g(\eta, t))}, \quad (12)$$

which is computable by means of five different STFTs:

$$\tilde{q}_f(\eta, t) = \frac{1}{2i\pi} \frac{V_f^{g''}(\eta, t)V_f^g(\eta, t) - (V_f^{g'}(\eta, t))^2}{V_f^{tg}(\eta, t)V_f^{g'}(\eta, t) - V_f^{tg'}(\eta, t)V_f^g(\eta, t)}. \quad (13)$$

Now, introducing $\tilde{\omega}_f(\eta, t) = \frac{\partial_t V_f^g(\eta, t)}{2i\pi V_f^g(\eta, t)}$, and $\tilde{t}_f(\eta, t) = t - \frac{\partial_\eta V_f^g(\eta, t)}{2i\pi V_f^g(\eta, t)}$, enables the definition of a new IF estimate as [11]:

$$\hat{\omega}_f^{(2)}(\eta, t) = \begin{cases} \mathcal{Re} \{ \tilde{\omega}_f(\eta, t) + \tilde{q}_f(\eta, t)(t - \tilde{t}_f(\eta, t)) \} & \text{if } \partial_t \tilde{t}_f(\eta, t) \neq 0 \\ \hat{\omega}_f(\eta, t) & \text{otherwise,} \end{cases} \quad (14)$$

where $\mathcal{Re}\{X\}$ denotes the real part of X . It is worth noting here that $\hat{\omega}_f(\eta, t) = \mathcal{Re}\{\tilde{\omega}_f(\eta, t)\}$. It can be easily shown that, when $f(t) = A(t)e^{2i\pi\phi(t)}$ is a linear chirp with Gaussian amplitude, $\hat{\omega}_f^{(2)}(\eta, t) = \phi'(t)$. For a more general mode with Gaussian amplitude, when the IF is estimated by $\hat{\omega}_f^{(2)}(\eta, t)$, the estimation error involves only derivatives of the phase whose orders are larger than 3.

VSST then consists of replacing $\hat{\omega}_f$ by $\hat{\omega}_f^{(2)}$ in FSST:

$$TV_f^g(\omega, t) = \frac{1}{g^*(0)} \int_{\{\eta, |V_f^g(\eta, t)| \geq \gamma\}} V_f^g(\eta, t) \delta(\omega - \hat{\omega}_f^{(2)}(\eta, t)) d\eta. \quad (15)$$

The reconstruction of the mode f_k is subsequently performed by means of the following formula:

$$f_k(t) \approx \int_{|\omega - \phi'_k(t)| < d} TV_f^g(\omega, t) d\omega. \quad (16)$$

D. On the Computation of the STFT and Synchrosqueezed Transforms

In many practical situations, the signal f is of finite length, typically defined on the interval $[0, T]$, and discretized into $f(\frac{nT}{N})_{n=0, \dots, N-1}$. In what follows and without loss of generality, N is assumed to be a power of 2 to ease the presentation. Assuming g is supported on $[-\frac{LT}{N}, \frac{LT}{N}]$, with $L < N/2$ the STFT is then computed as follows:

$$\begin{aligned} V_f^g(\eta, t) &= \int_{\mathbb{R}} f(t + \tau)g(\tau)e^{-2i\pi\tau\eta}d\tau = \int_{-\frac{LT}{N}}^{\frac{LT}{N}} f(t + \tau)g(\tau)e^{-2i\pi\tau\eta}d\tau \\ &\approx \frac{T}{N} \sum_{n=-L}^L f(t + \frac{nT}{N})g(\frac{nT}{N})e^{-i2\pi\frac{nT}{N}\eta}, \end{aligned} \quad (17)$$

from which we infer that:

$$\begin{aligned} V_f^g(\frac{p}{T}, \frac{qT}{N}) &\approx \frac{T}{N} \sum_{n=-L}^L f(\frac{(q+n)T}{N})g(\frac{nT}{N})e^{-i2\pi\frac{np}{N}} \\ &= \frac{Te^{\frac{2i\pi pL}{N}}}{N} \sum_{n=0}^{2L} f(\frac{(q+(n-L))T}{N})g(\frac{(n-L)T}{N})e^{-i2\pi\frac{np}{N}} \end{aligned} \quad (18)$$

$$:= e^{\frac{2i\pi pL}{N}} \sum_{n=0}^{2L} S(q, n)e^{-i2\pi\frac{np}{N}}, \quad (19)$$

where the last sum is computed by means of an FFT. It is common to extend, for each q , the sequence $(S(q, n))_n$ into a sequence of size $N_f > N$ by adding $N_f - N$ zeros to it. This operation is known as *zero-padding*. By doing so, one obtains an increased frequency resolution in the TF grid but not of the time resolution, since:

$$\begin{aligned} V_f^g(\frac{N}{N_f}\frac{p}{T}, \frac{qT}{N}) &\approx \frac{T}{N} \sum_{n=-L}^L f(\frac{(q+n)T}{N})g(\frac{nT}{N})e^{-i2\pi\frac{np}{N_f}} \\ &= \frac{Te^{\frac{2i\pi pL}{N_f}}}{N} \sum_{n=0}^{2L} f(\frac{(q+(n-L))T}{N})g(\frac{(n-L)T}{N})e^{-i2\pi\frac{np}{N_f}} \\ &= e^{\frac{2i\pi pL}{N_f}} \sum_{n=0}^{2L} S(q, n)e^{-i2\pi\frac{np}{N_f}}. \end{aligned} \quad (20)$$

Since $V_f^g(\frac{N}{N_f}\frac{p}{T}, \frac{qT}{N})$ is approximated by means of an FFT, only the first half of the frequency set is meaningful. That is, V_f^g is approximated on the TF grid $\{0, \frac{T}{N}, \dots, \frac{(N-1)T}{N}\} \times \{0, \frac{N}{N_f}\frac{1}{T}, \dots, \frac{N}{N_f}\frac{N_f/2-1}{T}\}$.

It is worth noting here that the synchrosqueezed TF representations T_f^g or TV_f^g correspond to reassigned versions of the STFT on the discrete time-frequency grid defined above. The computation of T_f^g from V_f^g can then be carried out as explained in Algorithm 1 (putting $t_q = \frac{qT}{N}$) [8].

Algorithm 1

for $q = 0$ to $N - 1$ **do**

for $k = 0$ to $N_f/2 - 1$ **do**

$T_f^g(\frac{N}{N_f} \frac{k}{T}, t_q) := 0$

for $p = 0$ to $N_f/2 - 1$ **do**

 Compute $\hat{\omega}_f(\frac{N}{N_f} \frac{p}{T}, t_q) = \frac{N}{N_f} \frac{p}{T} - \mathcal{I}m \left\{ \frac{1}{2\pi} \frac{V_f^{g'}(\frac{N}{N_f} \frac{p}{T}, t_q)}{V_f^g(\frac{N}{N_f} \frac{p}{T}, t_q)} \right\}$.

 Put $k = \text{round}(T \frac{N_f}{N} \hat{\omega}_f(\frac{N}{N_f} \frac{p}{T}, t_q))$.

 Reassign $|V_f^g(\frac{N}{N_f} \frac{p}{T}, t_q)| > \gamma$ as follows:

$$T_f^g(\frac{N}{N_f} \frac{k}{T}, t_q) = T_f^g(\frac{N}{N_f} \frac{p}{T}, t_q) + \frac{1}{g^*(0)} V_f^g(\frac{N}{N_f} \frac{p}{T}, t_q).$$

The same algorithm is applied to get TV_f^g from V_f^g , replacing $\hat{\omega}_f$ by $\hat{\omega}_f^{(2)}$. The role of zero-padding is going to be further investigated in the sequel.

III. RIDGE ESTIMATION

Any mode reconstruction techniques based on the synchrosqueezing transform requires an estimate of the ridges $(t, \phi'_k(t))$ (mode reconstruction being then either based on formula (7) or (16), depending on the type of TF representation used). In that context, we are going to introduce a classical ridge detector that is usually applied to the spectrogram, and then investigate whether to perform the ridge detection on the reassigned transform is profitable. The influence of all of the different parameters on the accuracy of ridge estimation both in noiseless and noisy contexts will also be studied.

A. Algorithm for Ridge Extraction

To compute an estimate of the ridge $(t, \phi'_k(t))$, assuming knowledge of the number of modes K , we can use the same algorithm as described in [5] or [12], and which was originally proposed in [13]. This computes a local minimum of the functional

$$E_f(\psi_1, \dots, \psi_K) = \sum_{k=1}^K - \int_{\mathbb{R}} |TF_f(\psi_k(t), t)|^2 dt + \int_{\mathbb{R}} \lambda \psi'_k(t)^2 dt + \beta \psi''_k(t)^2 dt, \quad (21)$$

where TF_f is one of the TF representations given by V_f^g , T_f^g or TV_f^g . However, as presented, equation (21) does not offer any algorithmic means to compute the ridges. Inspired by the above minimization problem, we derive Algorithm 2, for that purpose.

Algorithm 2

Pick $q \in \{0, \dots, N-1\}$

for $k = 1$ to K **do**

1. $t_q = \frac{qT}{N}$.

2. Define $p_{k,q} = \underset{l}{\operatorname{argmax}} |TF_f(\frac{N}{N_f} \frac{l}{T}, t_q)|$.

3. Define $I_{k,q} = \{\max(0, p_{k,q} - \frac{N_f}{N} T \Delta), \dots, \min(p_{k,q} + \frac{N_f}{N} T \Delta, N_f/2 - 1)\}$.

4. Define $p_{k,q-1} = \underset{l \in I_{k,q}}{\operatorname{argmax}} |TF_f(\frac{N}{N_f} \frac{l}{T}, t_{q-1})|$.

5. Define $p_{k,q+1} = \underset{p \in I_{k,q}}{\operatorname{argmax}} |TF_f(\frac{N}{N_f} \frac{p}{T}, t_{q+1})|^2 - \lambda(p - p_{k,q})^2 - \beta(p - 2p_{k,q} + p_{k,q-1})^2$

6. Iterate forward in time.

7. Iterate steps 2-6 backward from time t_q .

8. $(\psi_k(t_q))_q = (\frac{N}{N_f} \frac{p_{k,q}}{T})_q$

9. $TF_f = TF_f \setminus \bigcup_q [\psi_k(t_q) - \Delta, \psi_k(t_q) + \Delta]$

To improve the robustness of the procedure, several random initializations are required, leading to the detection of many different ridge sets $(\psi_k)_{k=1, \dots, K}$, and the one retained as the output corresponds to the one maximizing

$$\sum_{k=1}^K \sum_{n=0}^{N-1} |TF_f(\frac{N}{N_f} \frac{p_{k,n}}{T}, t_n)|^2 - \lambda \sum_{n=1}^{N-1} (p_{k,n} - p_{k,n-1})^2 - \beta \sum_{n=1}^{N-2} (p_{k,n+1} - 2p_{k,n} + p_{k,n-1})^2.$$

Note that, at the end of the procedure, the estimated ridges need to be resorted according to increasing IF.

B. Influence of Zero-Padding on Ridge Estimation

To start the discussion on the influence of zero-padding on ridge estimation, we recall, for the case of a mono-component signal, the following estimate of $\phi'(t_n)$:

$$\Psi(t_n) = \underset{\eta}{\operatorname{argmax}} |V_f^g(\eta, t_n)|^2, \quad (22)$$

which was studied in [14], for a noisy version of the signal $f(t) = A(t)e^{2i\pi\phi(t)}$, i.e. $\tilde{f}(t_n) = f(t_n) + \epsilon(t_n)$, where ϵ is a Gaussian white noise with variance σ_ϵ^2 . Selecting

$$\Delta\Psi(t_n) := \Psi(t_n) - \phi'(t_n), \quad (23)$$

and assuming g is the Gaussian window, $g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$, it was proven in [14] that:

$$\begin{aligned} Bias\{\Delta\Psi(t_n)\} &\sim_{N \rightarrow +\infty} \sum_{k=1}^{+\infty} \frac{2\pi\phi^{(2k+1)}(t_n)\sigma^{2k}}{2^k k!} \\ Var\{\Delta\Psi(t_n)\} &\sim_{N \rightarrow +\infty} \frac{\sigma_\epsilon^2}{8\sqrt{\pi}|A(t_n)|^2} \left[1 + \frac{\sigma_\epsilon^2 T}{2N\sqrt{\pi}\sigma|A(t_n)|^2} \right] \frac{T}{N\sigma^3}. \end{aligned} \quad (24)$$

These results are interesting but they do not consider the fact that $|V_f^g(\eta, t_n)|$ is only available on a discrete frequency grid, since it is computed using an FFT. More precisely, (22) actually corresponds to the ridge detector we would like to study (when λ and β are null), assuming a continuous frequency representation. To illustrate the impact of the discrete grid associated with frequency resolution, remembering that, as already noted, TF representations are evaluated at frequencies $\left(\frac{N}{N_f} \frac{p}{T}\right)_{p=0, \dots, N_f/2-1}$, we investigate the quality of IF estimate depending on this discretization, i.e. the choice for N_f . This can be quantified by measuring the mean square error (MSE) between the estimated ridge and the ground truth:

$$MSE(\psi) = \sqrt{\frac{1}{N-1} \sum_{n=0}^{N-1} (\phi'(t_n) - \psi(t_n))^2}, \quad (25)$$

when the frequency resolution varies. To have a better understanding of what is at work in this ridge detection, we not only investigate the influence of zero-padding but also of the noise-level. Since, the study of a linear chirp is somewhat limiting, we extend the analysis to three different types of mono-component signals whose STFT are displayed in Figure 1, first row (they correspond to a linear chirp, a polynomial chirp and a mode with sinusoidal phase).

The results, displayed in Figure 1 D, show that, in a noise-free context, when STFT or VSST are used for ridge detection, MSEs are the same for the linear chirp, which corresponds to the fact that the coefficients are reassigned to a maximum of the STFT with VSST (this method being based on an exact IF estimate for linear chirps). For the other two signals, the detector based on STFT behaves a little bit better than VSST, but not significantly so. In contrast, since FSST is based on an inaccurate IF estimate (even for the linear chirp), the results in terms of ridge estimation are significantly worse when the former is used as TF representation. For this reason, we do not consider it in the simulations which follow. Finally, we remark, that in the noise-free context for the linear and polynomial chirps of Figure 1 A and B, the MSE error when using STFT or VSST decreases when the frequency resolution across the sampling grid is increased. However, this is no longer true with the signal of Figure 1 C. In such a case, since the signal modulation is important, there is no staircase effect even at a low frequency resolution such as $N_f = N$. The conclusion of this study is that the frequency resolution, for the purpose of ridge estimation, has to be tuned depending on the signal modulation: a small modulation requires a higher

frequency resolution.

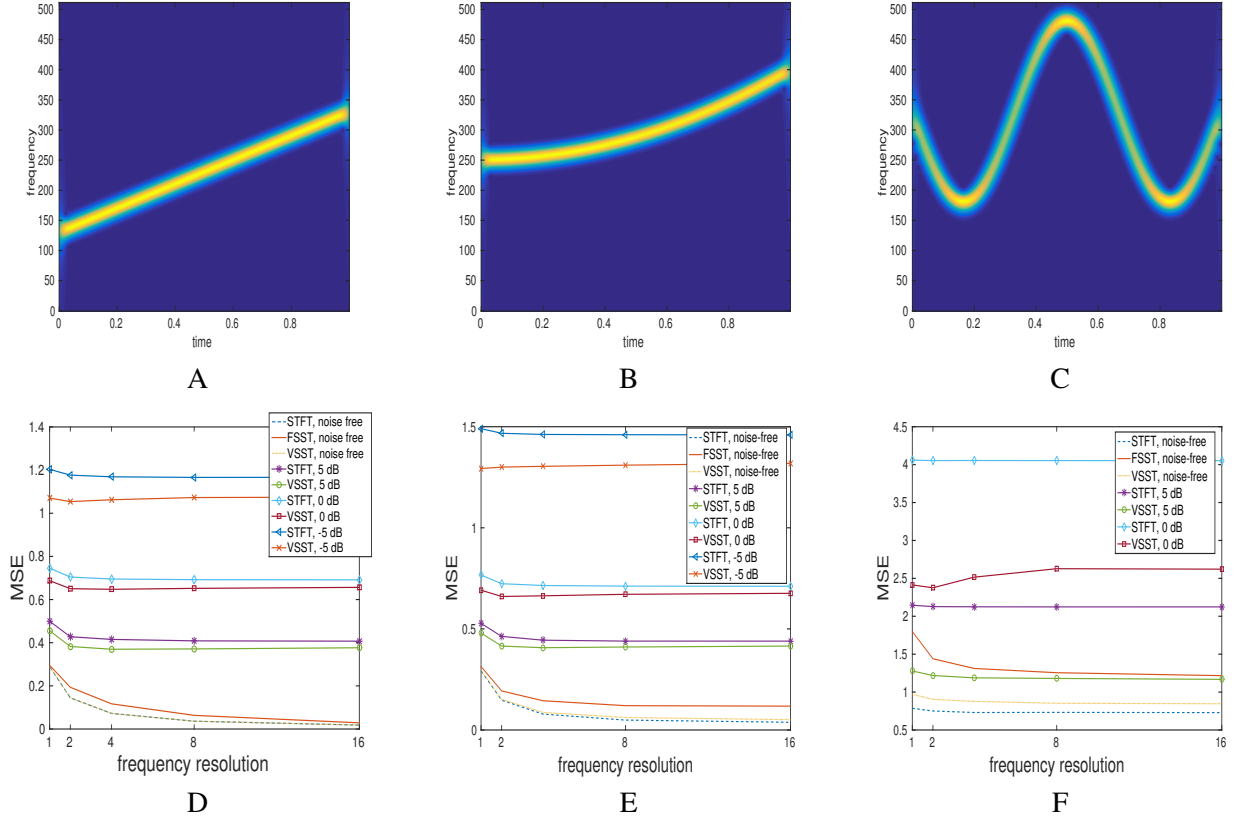


Figure 1. A: STFT of a linear chirp; B: STFT of a polynomial chirp; C: STFT of a mode with sinusoidal phase; D: Computation of the mean square error associated with the ridge detection for the linear chirp displayed in A, for various frequency resolution (k in abscissa means $N_f = kN$), different TF representations and noise level; E: same as D but for the polynomial chirp displayed in B; F: same as D but for the mode with sinusoidal phase displayed in C.

Now, we would like to understand what happens in noisy situations, therefore we perform the ridge detection on the linear and polynomial chirps and also on the mode with sinusoidal phase but with an SNR equal to 5, 0 or -5 dB. The results are depicted in Figure 1 D to F (for the latter type of signals, and whatever the TF representation used, the ridge detector does not perform well at -5 dB, therefore the results are not depicted). It is clear from Figures 1 D and E that, while a finer frequency resolution, associated with a larger N_f , leads to a more accurate IF estimate in the noise-free case, N_f has a much smaller impact on the quality of the estimation in a noisy context. Furthermore, the quality of the estimate provided by applying the ridge detector to VSST rather than to STFT is always better: the ridge detection operates on a much sharper TF representation which appears to be less sensitive to noise. Finally, we note that, from these simulations, $N_f = 8N$ is a good choice for frequency resolution for ridge detection purpose.

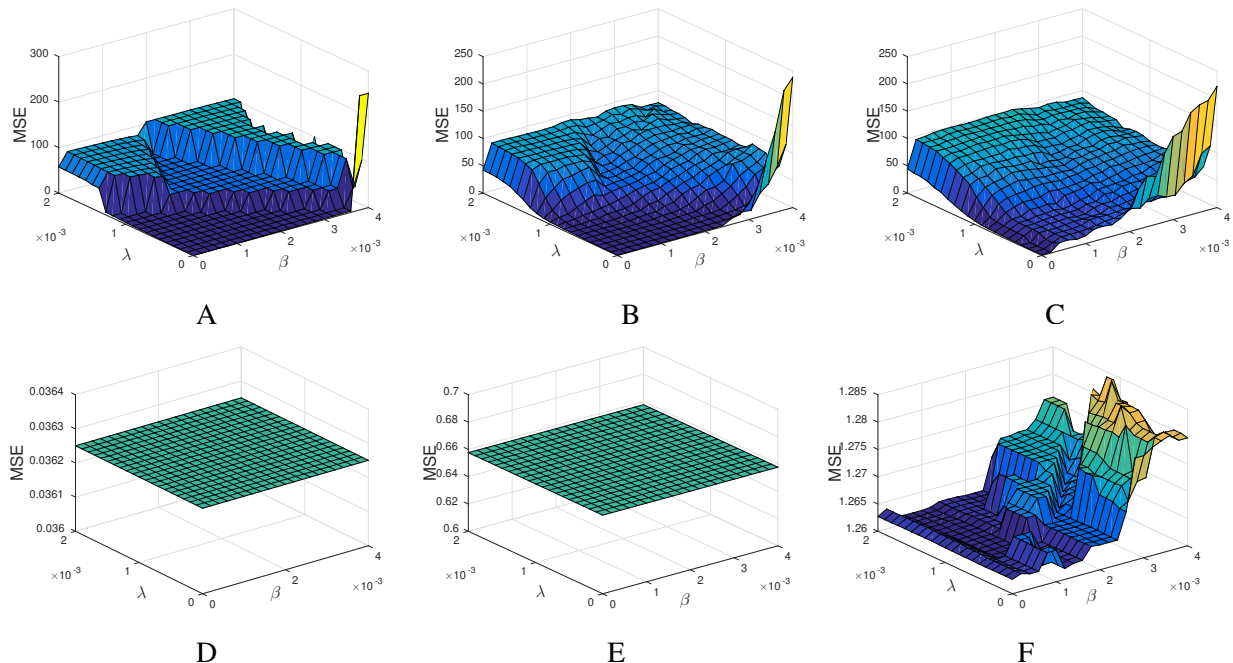


Figure 2. A: MSE corresponding to the ridge estimation for the linear chirp of Figure 1 A with V_f^g as TF representation (noise-free case, $N_f = 8N$); B: Same as A but at a 0 dB noise level; C: Same as A but at a -5 dB noise level; D: MSE corresponding to the ridge estimation for the linear chirp of Figure 1 A with TV_f^g as TF representation (noise-free case, $N_f = 8N$); E: Same as D but at a 0 dB noise level; H: Same as D but at a -5 dB noise level.

C. Influence of Regularization Parameters

Taking into account the study carried out in the previous section, the ridge detector applied either to STFT or VSST both lead to good results when no regularization is used, even though, as illustrated in Figure 1 (second row), to perform ridge detection on VSST rather than STFT is always better in noisy situations.

We now study the behavior of the ridge detector applied to STFT or VSST when regularization terms vary, both in the noise-free and noisy cases. To do so, we consider the same linear chirp as previously either in the noise-free, 0 dB or -5 dB cases. We remark that the ridge detector is much more sensitive to regularization parameters when applied to STFT rather than VSST (see Figure 2): the reassignment technique enables a more robust ridge detection even at high noise level, because it corresponds to a sharper TF representation. Finally, note that, the regularization parameters do not offer any improvement in terms of the accuracy of the ridge estimation, which argues against using them, (the simulations shown in Figure 2 were carried out for $N_f = 8N$, but the same results could be derived for any reasonable value of N_f). It is important to note here that the same conclusions would hold if the simulations were carried out on the polynomial chirp or on the mode with sinusoidal phase, as soon the algorithm detects

the ridge.

IV. DEMODULATION ALGORITHM AND MODE RECONSTRUCTION

Once a ridge is detected using an appropriate N_f to avoid the staircase effect mentioned above, we compute a demodulation operator for each mode which is going to be subsequently used to extract the corresponding demodulated mode. Inverting the demodulation operator, we will finally obtain the desired mode. The modes are extracted in a sequential fashion, i.e. one at a time, a commonly used technique often referred to as the *peeling method* in the literature [15][16].

It is worth noting here that, in most cases, and in contrast to our approach, when demodulation problems are considered, it is often assumed that knowledge of a phase function $v(t)$ is available and, this is then used to compute the so-called short time generalized Fourier transform (STGFT). Indeed, the STGFT corresponds to:

$$V_f^{g,v}(\eta, t) = \int_{\mathbb{R}} f(t)g^*(t - \tau)e^{-2i\pi v(t)}e^{-2i\pi\eta(t-\tau)}dt. \quad (26)$$

This kind of approach has also been used in [17][18] and ridge detection can be viewed as a way to estimate this phase function. Attempts have also been made to estimate the ridges using parametric models [10]. As will be explained later, our approach is fully non-parametric.

A. Definition of Demodulation Operator

Based on the ridge estimate defined above, we introduce the demodulation algorithm for a mono-component signal $f(t) = A(t)e^{2i\pi\phi(t)}$, for which we assume the IF estimate $\psi(t)$ is computed. For the case of a linear chirp, i.e. $\phi(t) = at + bt^2$, $\psi(t)$ approximates $a + 2bt$. So, by multiplying $f(t)$ by $e^{-2i\pi(\psi(t)t/2)}$, and if the IMF estimation is accurate, we should obtain a demodulated signal f_D with constant frequency $a/2$. However, it is worth remarking that this demodulation procedure is only well suited to a linear chirp, because it removes only second order terms. Therefore, to demodulate a more general mode $f(t) = A(t)e^{i2\pi\phi(t)}$, the following demodulation operator $e^{-i2\pi(\int_0^t \psi(x)dx - \psi_0 t)}$, where ψ_0 is some positive constant frequency, is a better choice since, no assumption is made about ϕ . Indeed, by considering the signal

$$f_D(t) = f(t)e^{-i2\pi(\int_0^t \psi(x)dx - \psi_0 t)}, \quad (27)$$

one should get a signal with constant frequency ψ_0 . An illustration of this is shown in Figure 3, for three different types of mode, where ψ_0 is equal to 100 Hz. In that figure, we display the VSST of the considered modes in the noise-free (resp. 0 dB) case, in the first (resp. second) row. In the bottom row of

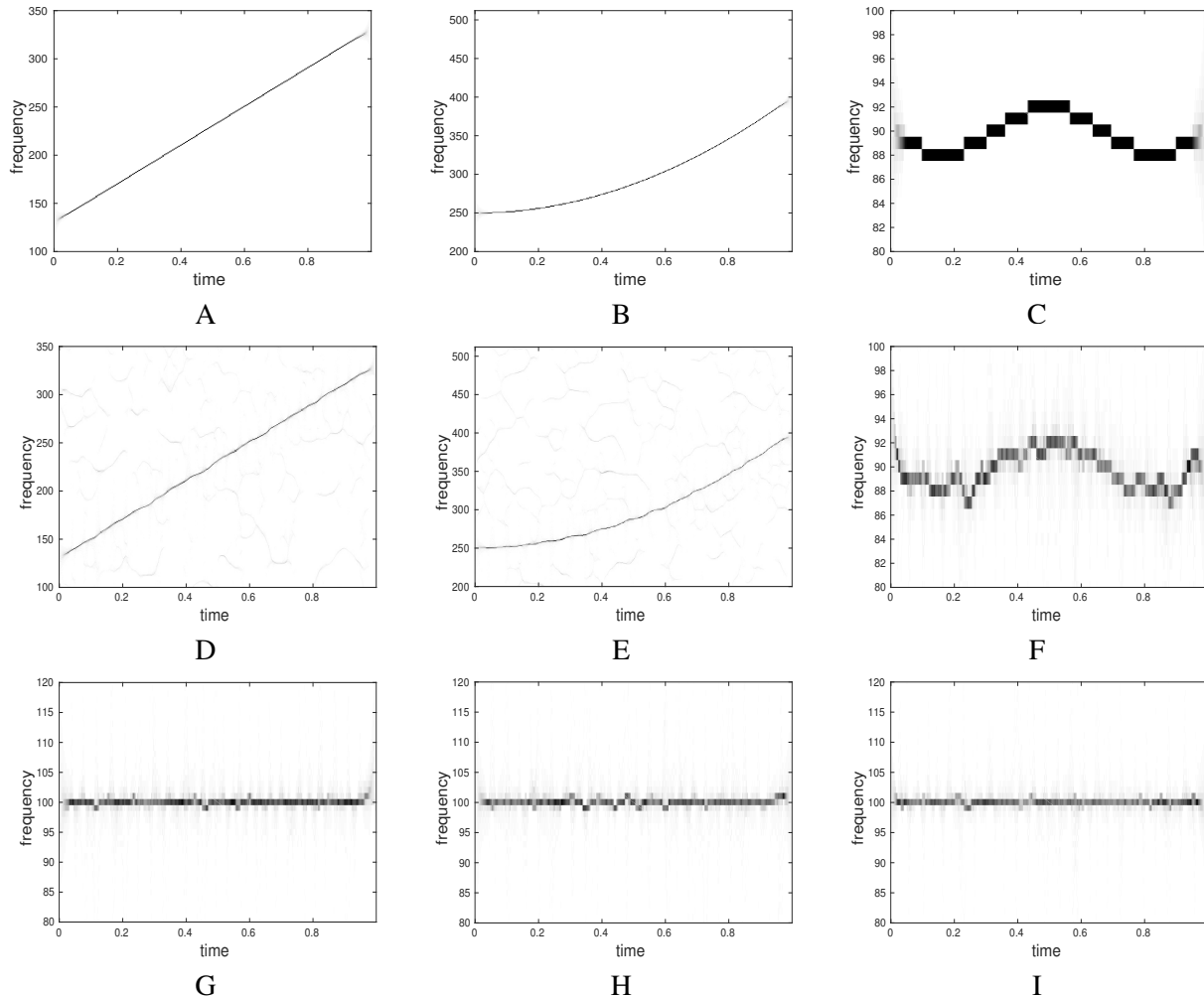


Figure 3. A: VSST of a linear chirp (noise-free case); B: VSST of a polynomial chirp (noise-free case); C: VSST of a mode with sinusoidal phase (noise-free case); D: VSST of a linear chirp (noise level 0 dB case); E: VSST of a polynomial chirp (0 dB case); F: VSST of a mode with sinusoidal phase (0 dB case); G: demodulated signal D; H: demodulated signal E; I: demodulated signal F.

that figure, we display the VSSTs of the demodulated signals associated with the three modes represented in the second row (N_f being taken equal to $8N$ in the ridge detection). Despite the high noise level, the demodulation performs well.

Now, let us consider how this procedure works in the multicomponent case. We will illustrate this by adopting a signal consisting of the three different modes, displayed in Figure 4 A. Then by applying Algorithm 2 to the VSST computed with $N_f = 8N$, we obtain the estimates (ψ_1, ψ_2, ψ_3) , which are subsequently used to compute three demodulated signals, as follows:

$$f_{D,k}(t) = f(t)e^{-i2\pi(\int_0^t \psi_k(x)dx - \psi_0 t)}, \quad k = 1, 2, 3. \quad (28)$$

The VSST of the three signals $(f_{D,k})_{k=1,2,3}$ are shown in Figures 4 B to D, where the SNR in the original signal equals 0 dB. It is worth noting that, in $f_{D,k}$, only the k th mode is demodulated.

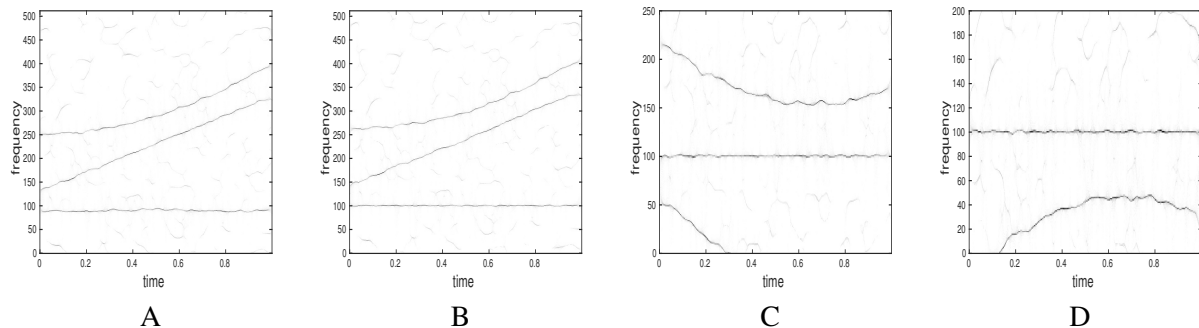


Figure 4. A: Three modes signal (noise level 0 dB); B: VSST of $f_{D,1}$; C: VSST of $f_{D,2}$; D: VSST of $f_{D,3}$.

B. Algorithm for Mode Extraction Based on Demodulation

The previous section has provided us with a means to demodulate any of the modes of the signal f , the number, K , of which is assumed to be known. With that in mind, the algorithm for mode extraction can then be summarized as follows:

Algorithm 3

Compute the ridge estimates (ψ_1, \dots, ψ_K) with Algorithm 2 applied to VSST.

for $k = 1$ to K **do**

1. Compute $f_{D,k}(t) = f(t)e^{-i2\pi(\int_0^t \psi_k(x)dx - \psi_0 t)}$.
 2. From $TV_{f_{D,k}}$, extract the ridge $\psi_{D,k}$ corresponding to mode k of $f_{D,k}$, by considering single ridge detection in the frequency range $[\psi_0 - \Delta, \psi_0 + \Delta]$.
 3. Reconstruct the k th mode of $f_{D,k}$ and then multiply it by the inverse of demodulation operator to recover f_k : $f_k(t) \approx \left(\int_{|\omega - \psi_{D,k}(t)| < \Delta} TV_{f_{D,k}}(\omega, t) d\omega \right) e^{i2\pi(\int_0^t \psi_k(x)dx - \psi_0 t)}$.
-

Note here that the TF representation used to compute the ridge of $f_{D,k}$ and then mode k could alternatively be T_f^g since the mode sought is demodulated, there is no need to take into account the modulation at this stage. Indeed, the k th mode of signal $f_{D,k}$ should be a purely harmonic signal at frequency ψ_0 (see Figures 4 B to D for illustrations). Furthermore, while it is important to fix the frequency resolution parameter N_f according to mode modulation for ridge estimation, to compute $TV_{f_{D,k}}$, $N_f = N$ is used because the mode k , extracted at step 3 of Algorithm 3, is a purely harmonic one.

V. EVALUATION OF THE PERFORMANCE OF THE RECONSTRUCTION ALGORITHM

Before we assess the reconstruction technique proposed, we discuss how an optimal window length, (which is crucial in all TF representations), might be determined. The emphasis is placed on the difficulty of estimating this window length in a noisy context.

A. Automatic Window Length Determination

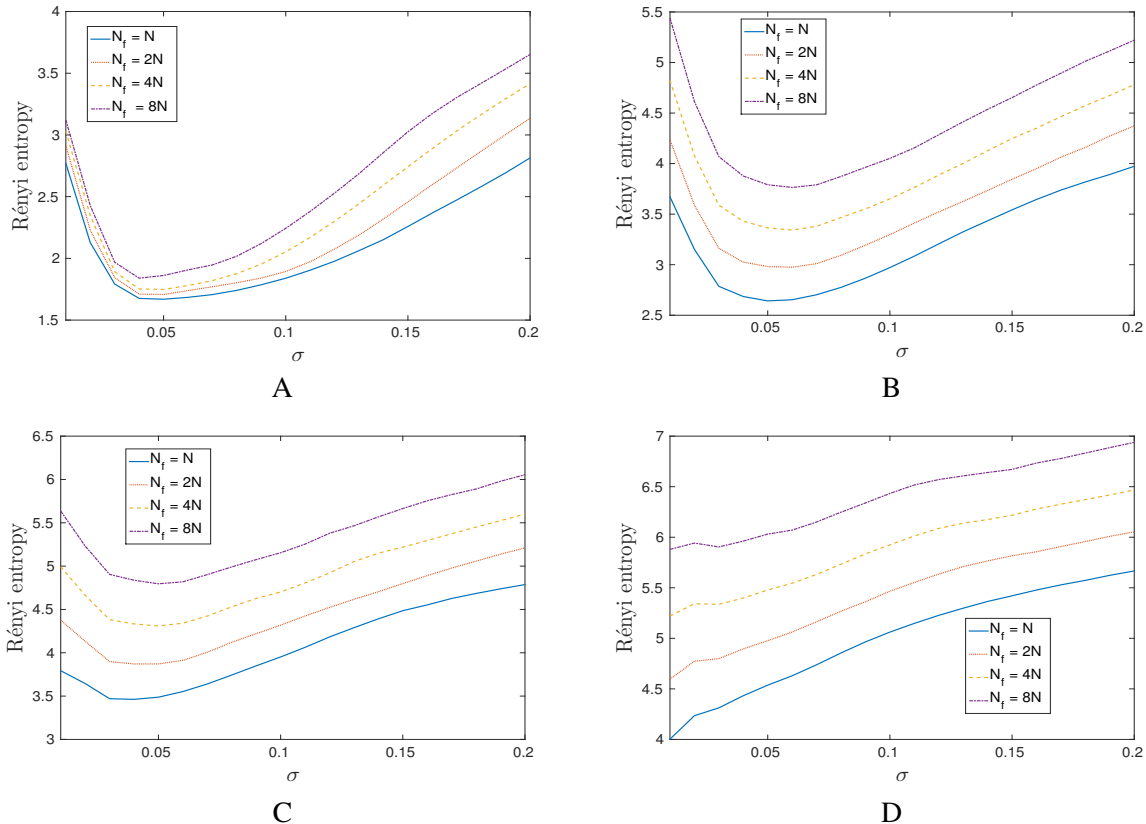


Figure 5. A: Rényi entropy associated with the VSST of Figure 4 A (but with no noise) as a function of parameter σ , and with various frequency resolutions; B: same as A but with a SNR equal to 5 dB; C: same as A but with a SNR equal to 0 dB; D: same as A but with a SNR equal to -5 dB

To determine an optimal window length, we consider that g is Gaussian, i.e. $g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$, so that its length is controlled by parameter σ . In our framework, we seek the value of that parameter leading to the most concentrated representation. Following [19][20], this concentration can be measured on the VSST by means of the Shannon entropy:

$$H_S(\sigma) = - \sum_{p=0}^{N_f/2-1} \sum_{n=0}^{N-1} \frac{|TV_f^g(\frac{N}{N_f} \frac{p}{T}, t_n)|}{\|TV_f^g\|_1} \log_2 \left(\frac{|TV_f^g(\frac{N}{N_f} \frac{p}{T}, t_n)|}{\|TV_f^g\|_1} \right), \quad (29)$$

or the Rényi entropy:

$$H_R(\sigma) = \frac{1}{1-\alpha} \log_2 \left(\sum_{p=0}^{N_f/2-1} \sum_{n=0}^{N-1} \left(\frac{|TV_f^g(\frac{N}{N_f} \frac{p}{T}, t_n)|}{\|TV_f^g\|_1} \right)^\alpha \right), \quad (30)$$

whose behaviors are reported to be very similar [19], and in which $\|TV_f^g\|_1 = \sum_{p=0}^{N_f/2-1} \sum_{n=0}^{N-1} |TV_f^g(\frac{N}{N_f} \frac{p}{T}, t_n)|$. To determine σ in this manner is particularly relevant, but only when the noise level is relatively low. Indeed, looking at Figures 5 A to C, representing the Rényi entropy (with α equal to 3) of the VSST of the signal f displayed in Figure 4 A, one notices that it exhibits a local minimum at a specific value for σ at noise level lower than 0 dB, and the optimal value is relatively stable for these cases. Note also that, in such a case, the result is not dependent on the frequency resolution. We note here that the Rényi entropy is computed by considering that the second order reassignment operator reassigns only the coefficients $V_f^g(\eta, t)$ such that $|\mathcal{R}e \{V_f^g(\eta, t)\}| > \gamma_1$ or $|\mathcal{I}m \{V_f^g(\eta, t)\}| > \gamma_2$, where γ_1 (resp. γ_2) is the standard deviation of $\mathcal{R}e \{V_f^g(\eta, t)\}$ (resp. $\mathcal{I}m \{V_f^g(\eta, t)\}$). The choice for such a threshold is motivated by the fact that the STFT of a zero mean white Gaussian noise is also a zero mean Gaussian process.

However, the technique based on Rényi entropy to determine the optimal σ no longer works in a very noisy context, see Figure 5 D, in particular because it does not take into account the number K of modes. Since Algorithm 3 performs better when the ridge detection is efficient, it is natural to define the optimal value σ as the one that concentrates the most the information on the K detected ridges. This could be measured by introducing the following quantity:

$$E_R(\sigma) = \frac{\sum_{k=1}^K \sum_{n=0}^{N-1} |TV_f^g(\psi_k(t_n), t_n)|^2}{\sum_{p=0}^{N_f/2-1} \sum_{n=0}^{N-1} |TV_f^g(\frac{N}{N_f} \frac{p}{T}, t_n)|^2}, \quad (31)$$

where E_R stands for "energy on the ridge", bearing in mind the dependence on σ is contained in g . E_R actually corresponds to the proportion of the total energy located on the ridges. We depict $E_R(\sigma)$ for the same signal as before, for different noise levels, and for $N_f = N$ in Figure 6. We notice, first, that the optimal value is close to that given by the Rényi entropy for noise levels lower than 0 dB, and that the information located on the ridge becomes less and less significant as the noise level increases. What is interesting with this technique is that, in contrast to the Rényi entropy, it offers us a means to find a relevant σ at noise levels as high as -5 dB. Such a technique will thus be used to determine the optimal σ in very noisy situations. Finally, note that similar results can be obtained by considering different values of N_f .

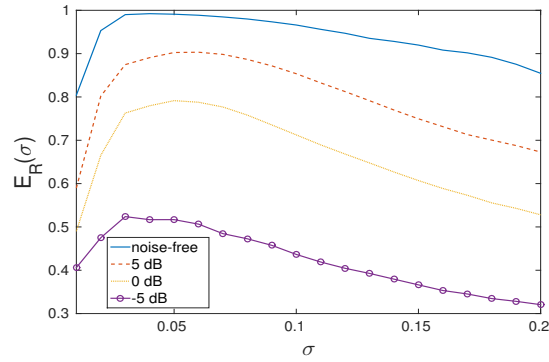


Figure 6. Computation of E_R , corresponding to the proportion of the energy contained in VSST computed on the ridges, with respect to σ , and in either the noise-free, 5 dB, 0 dB or -5 dB cases

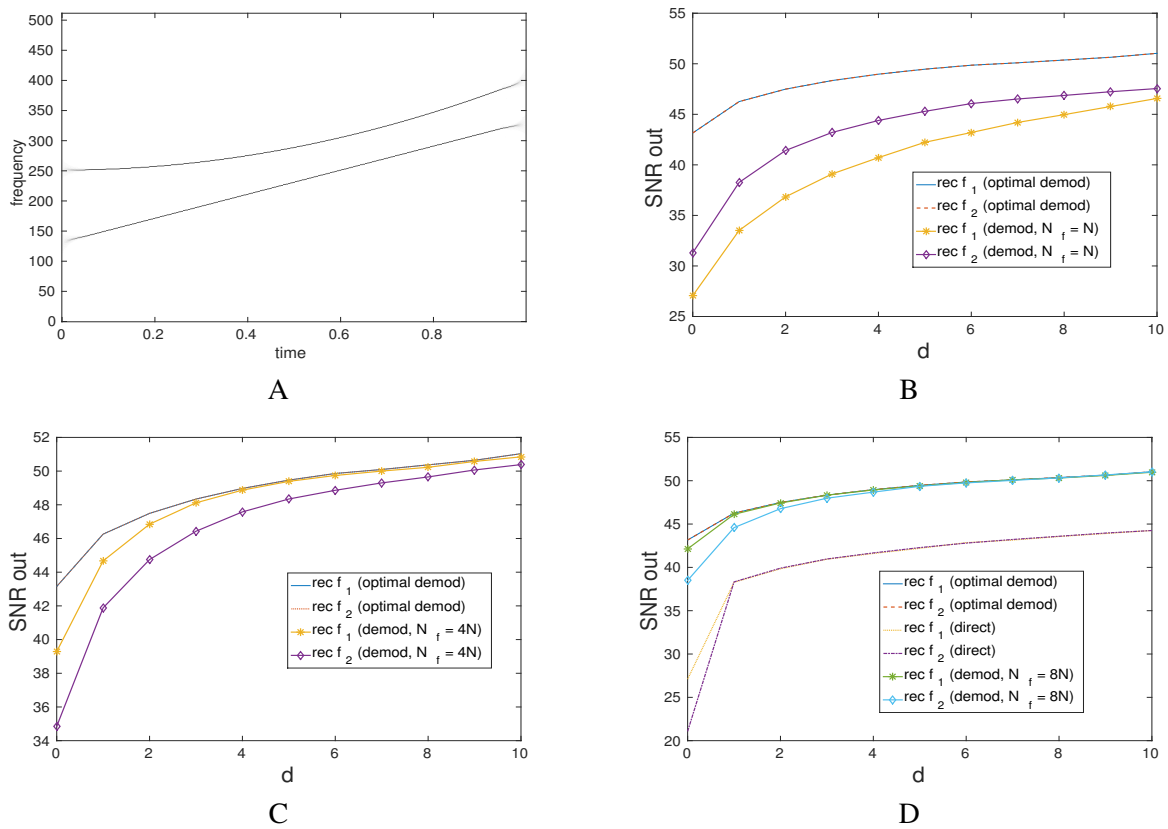


Figure 7. A: VSST of a two mode signal; B: mode reconstruction ("rec f_i " corresponds to reconstructed mode f_i) when ridges used in Algorithm 3 computed with $N_f = N$, along with the reconstruction when the IFs of the mode are assumed to be known in Algorithm 3 ("optimal demod" in the figure); C: Same as B, but when $N_f = 4N$ in the ridge computation; D: Same as B, but when $N_f = 8N$ in the ridge computation

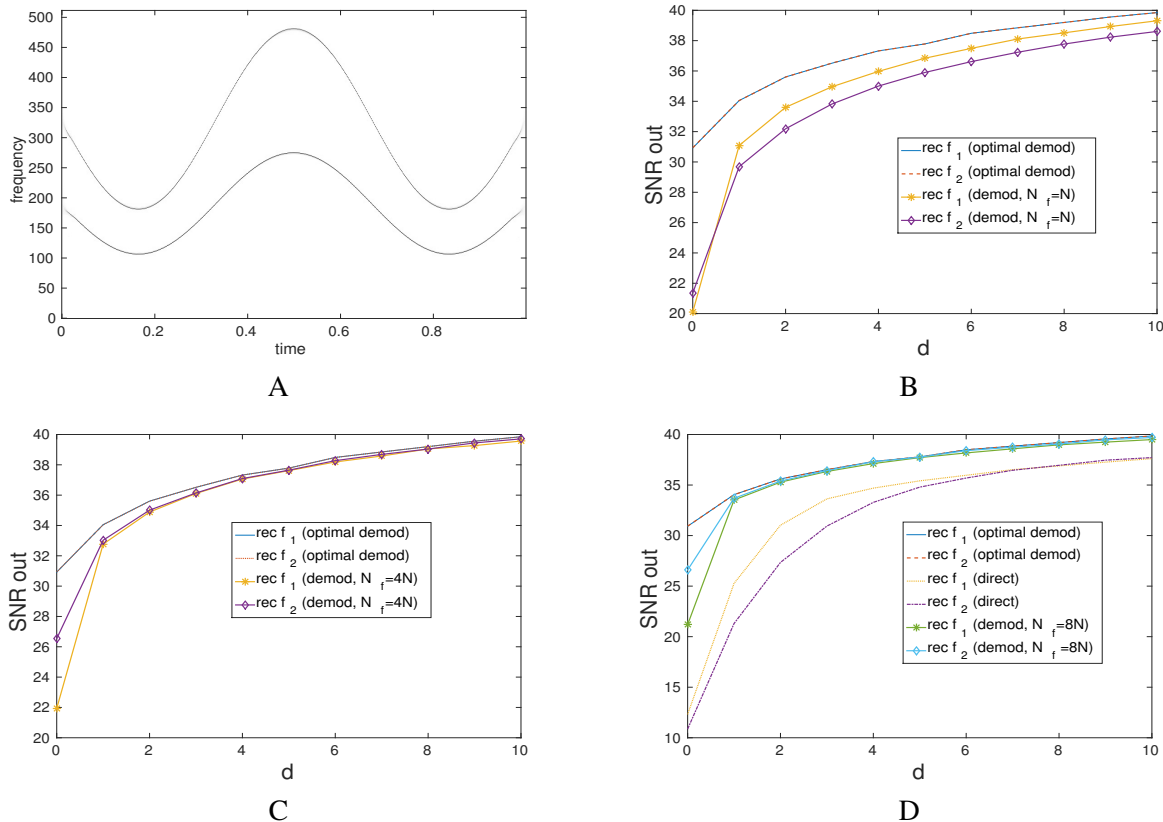


Figure 8. A: VSST of a two mode signal; B: mode reconstruction ("rec f_i " corresponds to reconstructed mode f_i) when ridges used in Algorithm 3 are computed with $N_f = N$, along with the reconstruction when the IFs of the mode are assumed to be known in Algorithm 3 ("optimal demod" in the figure); C: Same as B, but when $N_f = 4N$ in the ridge computation; D: Same as B, but when $N_f = 8N$ in the ridge computation

B. Reconstruction Procedure: Noise-free Case

In this section, we illustrate the improvement offered by Algorithm 3 in terms of the quality of the reconstructed modes in the noise free case to enable the impact of parameter selection to be considered. Our test signals are displayed in Figures 7 A and 8 A. The window used to build the TF representation is Gaussian and its length is optimized as explained in the previous subsection. Then, ridge detection and mode reconstruction are performed using a small value for γ in the definition of the reassignment operator (typically $\gamma = 10^{-3}$), since in such a case, all the non zero coefficients are related to the signal.

Since the ridge computation is influenced by the frequency resolution, we investigate the impact of N_f used in ridge computation on mode reconstruction. Also, to show that to use demodulation results in a more compact TF representation than by the original VSST method, the role of d , used both in reconstruction formula (16) and Algorithm 3, is investigated. To assess how the ridge detection impacts mode reconstruction, we also compute the mode reconstruction assuming the IFs of the modes are known.

We study two types of signals which are depicted in Figure 7 A and 8 A. The results for the first type are depicted in Figures 7 B to D and represent the output SNR defined, for mode i , as $20 \log_{10}(\frac{\|f_i\|_2}{\|f_i - \hat{f}_i\|_2})$, where the norm is the l_2 norm and \hat{f}_i is the i th mode reconstructed using Algorithm 3, and when the frequency resolution used in the ridge detection varies (in the different figures we use the term "demod"). In each case, we also display the reconstruction results using the true IFs of the modes in Algorithm 3 (in the figures we use the term "optimal demod"). We note that, as expected, for $N_f = 8N$ the results are very close to those obtained assuming knowledge of the IFs of the modes, as illustrated in Figure 7 D. We also display the results obtained by reconstructing the modes directly using formula (16): whatever the value of d the reconstruction is better when using Algorithm 3. Also, since the signal studied is slightly modulated, to choose a sufficiently large N_f for ridge estimation is crucial. Similar conclusions can be drawn from the study of the signal whose VSST is displayed in Figure 8 A: first, to increase N_f clearly improves the reconstruction results, and, then, when $N_f = 8N$, the results are close to those that would be obtained if the IFs of the modes were known. Also, we again remark that the results are far better than direct reconstruction. In this case however, and since the modes are more modulated than those of Figure 7 A, the impact of N_f on ridge computation and then mode reconstruction is less important.

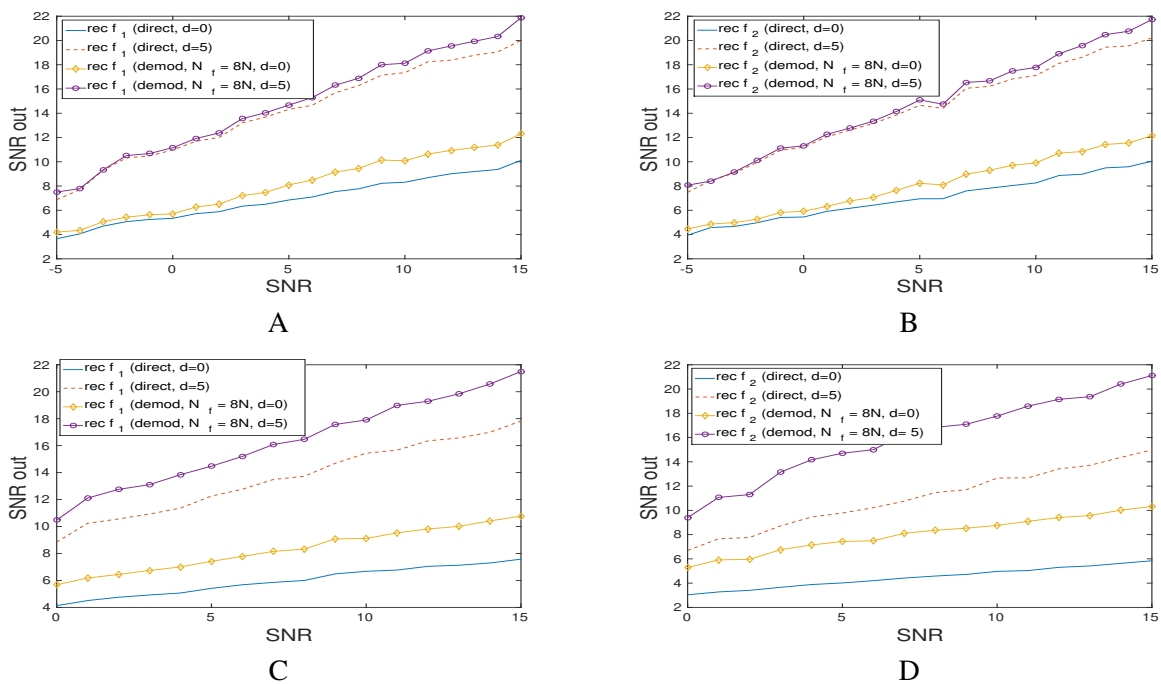


Figure 9. A: SNR after reconstruction for mode f_1 of the signal whose VSST is depicted in Figure 7 A using either the direct reconstruction (direct) or Algorithm 3 (demod), and for $d = 0$ or $d = 5$ in both cases; B: Same as A but for mode f_2 of the same signal; C: SNR after reconstruction for mode f_1 of the signal whose VSST is depicted in Figure 8 A using either the direct reconstruction (direct) or Algorithm 3 (demod), and for $d = 0$ or $d = 5$ in both cases; of the same signal; D: Same as C but for mode f_2 of that signal

C. Reconstruction Procedure: Noisy Case

In this section, we investigate the sensitivity to noise of our new method for mode retrieval, considering again the two types of signals displayed in Figures 7 A and 8 A. From the study of the noise-free case, ridge computation leads to good reconstruction when $N_f = 8N$, so we retain this value in the simulations that follow. Again we use a Gaussian window with the optimal window length σ computed as before, and with threshold $\gamma = 10^{-3}$.

The results displayed in the first row of Figure 9, represent the output SNR associated with mode reconstruction, when $d = 0$ or $d = 5$, respectively for the first and second mode of Figure 7 A, with respect to global input SNR. We note the following based on these observations: whatever the noise level, the mode reconstruction is improved by using Algorithm 3 rather than direct reconstruction; the discrepancy in terms of reconstruction performance between the two types of techniques increases when the noise level is decreasing; the gain of demodulating first is not that important because VSST is optimized for linear chirps.

Switching to the study of the signal of Figure 8 A, the benefit of using the demodulation procedure is much clearer: when a mode is very different from a linear chirp, the demodulation procedure greatly improve the reconstruction results. Finally, we remark that, as in the noise-free case, the parameter d plays a crucial role in the quality of the reconstruction, and that by demodulating first, we obtain a more concentrated representation since, for a given d , the reconstruction is always better using demodulation than without.

D. Application to Real Data: VSST versus EMD and limitations

Here we consider the reconstruction of a bat echolocation signal whose VSST is shown in Figure 10 A. Assuming the number of modes is three, which is consistent with the representation in the aforementioned figure (there is actually a fourth mode but due to aliasing effect we do not take it into account), we perform ridge extraction on the VSST (the extracted ridges are also depicted in the figure) and then compute the different modes by either using VSST or by demodulating first. In this regard, we study the influence of parameter d and frequency resolution on the reconstruction.

Since the signal studied is real, and as Algorithm 3 applies to complex signal, we first consider the Hilbert transform of the signal before applying that algorithm. Then, the length of the signal N not being a power of 2, we use $N_1 = 2^{\lfloor \log_2(N) \rfloor + 1}$, and its multiples, to define the different frequency resolutions subsequently used in the ridge detection. As previously, we investigate the impact of the frequency resolution used in the ridge estimation, on signal reconstruction. For this purpose, we compute the output SNR associated with the reconstruction of the signal by summing the first three modes. The

results depicted in Figure 10 B, again show the benefit of demodulating first the signal compared to direct computation, the improvement brought by using a higher frequency resolution being much less obvious than in controlled situations such as those studied before. In spite of this reconstruction results are satisfactory, some information is lost when considering the reconstructed signal obtained using only the first three modes. This problem arises because for real-world signals, such as the bat signal considered here, the number of modes is not constant over time: i.e. some modes vanish but the ridge estimation assumes that the modes will persist throughout the data record. This is a failing of many methods and is topic of current research. For comparison purposes we depict the VSST of the first three intrinsic mode functions (IMF) obtained with the *empirical mode decomposition* (EMD) [21], which is an alternative technique to extract the modes of a multicomponent signal, on Figure 11 A, B and C. From these figures, we note that the modes obtained are not related to the TF content of the signal depicted in Figure 10 A since the TF structure based on modes corresponding to TF ridges is completely broken.

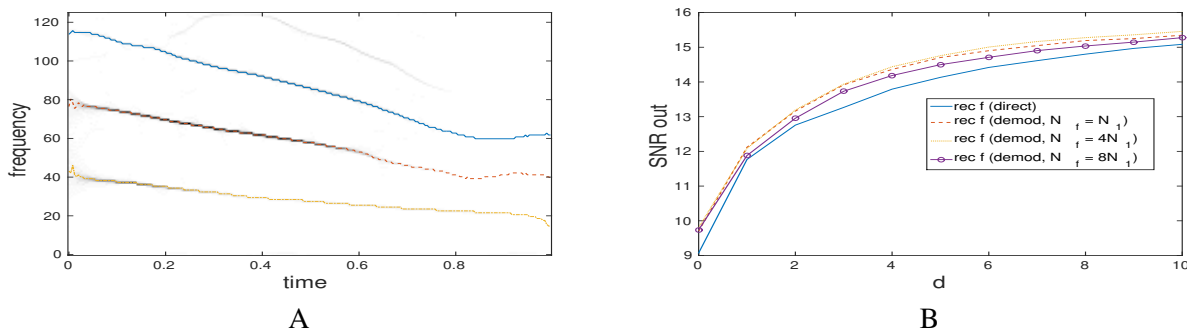


Figure 10. A: VSST of a bat echolocation call along with the corresponding ridges, B: reconstructed signal based on VSST and assuming the number of modes equals 3

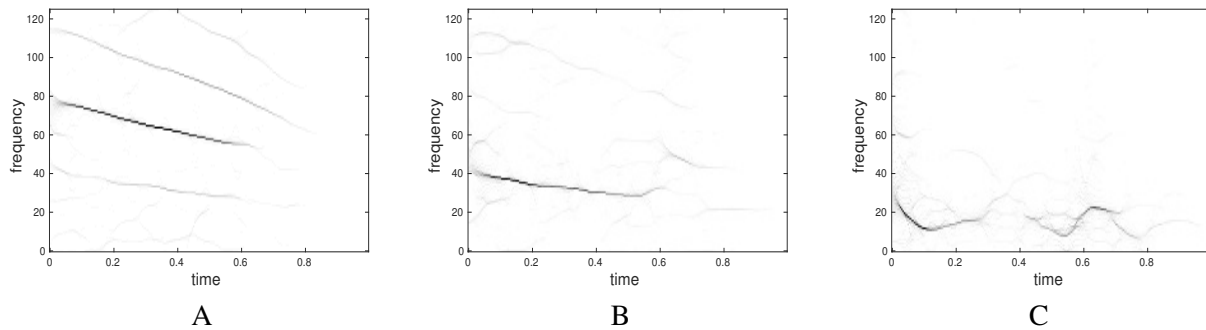


Figure 11. A: VSST of the first IMF obtained corresponding to signal whose VSST is depicted on Figure 10 A; B: second IMF; C: third IMF

VI. CONCLUSIONS

In this paper, we have introduced a new algorithm for the retrieval of the modes of a multicomponent signal from the study of some time-frequency representations. It is based on a novel technique for ridge estimation followed by a demodulation procedure. By using an appropriate frequency resolution, it is possible to compensate for the discretization of the frequencies induced by the use of FFTs in the computation of the TF representations, and thus obtained reliable IF estimates. The simulation carried out on test signals show that, by demodulating the signal first using these IF estimates, the associated time-frequency representation is sharpened and that the accuracy of the reconstruction is much better than when direct reconstruction is performed, both in noiseless and noisy situations. Simulations performed on real signals where the number of modes may vary with time however show the limits of signal reconstruction based on ridge extraction, and is a challenging issue we will address in the near future.

ACKNOWLEDGMENTS

The authors would like to thank anonymous reviewers for their suggestions that enabled to greatly improve the readability of the manuscript.

REFERENCES

- [1] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrèsani, “Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies,” *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 644–664, 1992.
- [2] 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,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 11, pp. 5787–5798, nov. 2012.
- [3] —, “Multicomponent signal denoising with synchrosqueezing,” in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, aug. 2012, pp. 660–663.
- [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] I. Daubechies, J. Lu, and H. Wu, “Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool,” *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 243–261, 2011.
- [6] T. Oberlin, S. Meignen, and V. Perrier, “The Fourier based synchrosqueezing transform,” in *Proceedings of ICASSP*, 2014.
- [7] H.-T. Wu, “Adaptive analysis of complex data sets,” PhD thesis, Princeton University, 2011.
- [8] T. Oberlin, S. Meignen, and P. Perrier, “Second-order synchrosqueezing transform of invertible reassignment? towards ideal time-frequency representation,” *Signal Processing, IEEE Transactions on*, vol. 63, no. 3, pp. 1335–1344, 2015.
- [9] C. Lia and M. Liang, “A generalized synchrosqueezing transform for enhancing signal time-frequency representation,” *Signal Processing*, vol. 92, no. 9, pp. 2264–2274, 2012.
- [10] S. Wang, X. Chen, G. Cai, B. Chen, X. Li, and Z. He, “Matching demodulation transform and synchrosqueezing in time-frequency analysis,” *Signal Processing, IEEE Transactions on*, vol. 62, no. 1, pp. 69–84, 2014.

- [11] R. Behera, S. Meignen, and T. Oberlin, “Theoretical analysis of second-order synchrosqueezing transform,” *Applied and Computational Harmonic Analysis*, doi:10.1016/j.acha.2016.11.001, 2016.
- [12] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu, “An overview of time-frequency reassignment and synchrosqueezing,” *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 32–41, 2013.
- [13] R. Carmona, W. Hwang, and B. Torresani, “Characterization of signals by the ridges of their wavelet transforms,” *Signal Processing, IEEE Transactions on*, vol. 45, no. 10, pp. 2586–2590, 1997.
- [14] L. Stankovic, M. Dakovic, and V. Ivanovic, “Performance of spectrogram as if estimator,” *Electron. Lett.*, vol. 37, no. 12, pp. 797–799, 2001.
- [15] B. Barkhat and B. Boashash, “A high-resolution quadratic time-frequency distribution for multicomponent signal analysis,” *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2232–2239, 2001.
- [16] M. Z. Ikram and G. Zhou, “Estimation of multicomponent polynomial phase signals of mixed orders,” *Signal Processing*, vol. 81, no. 11, pp. 2293 – 2308, 2001.
- [17] S. Olhede and A. Walden, “A generalized demodulation approach to time-frequency projections for multicomponent signals,” *Proc. R. Soc. A.*, vol. 461, pp. 2159–2179, 2005.
- [18] Z. Yu, Y. Sun, and W. Jin, “A novel generalized demodulation approach for multi-component signals,” *Signal Processing*, vol. 118, pp. 188–202, 2016.
- [19] R. Baraniuk, P. Flandrin, A. Janssen, and O. Michel, “Measuring time-frequency information content using the renyi entropies,” *IEEE Transactions on Information theory*, vol. 47, no. 4, pp. 1391–1409, 2001.
- [20] L. Stanković, “A measure of some time–frequency distributions concentration,” *Signal Processing*, vol. 81, no. 3, pp. 621–631, 2001.
- [21] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, and H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.