Contributions to the Analysis of Multicomponent Signals Synchrosqueezing and Associated Methods

Duong-Hung Pham

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Mathématiques appliquées**

Arrêté ministériel : 7 août 2006

Présentée par

PHAM Duong Hung

Thèse dirigée par **Sylvain Meignen**

préparée au sein du **Laboratoire Jean Kuntzmann**
et de **l'École Doctorale Mathématiques, Sciences et Technologies de l'Information, Informatique**

Contributions to the Analysis of Multicomponent Signals

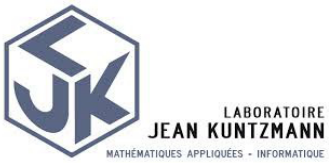
Synchrosqueezing and Associated Methods

Thèse soutenue publiquement le **17 Septembre 2018**,
devant le jury composé de:

M. Patrick Flandrin
Directeur de recherche, École Normale Supérieure de Lyon, Président
M. François Auger
Professeur, Université de Nantes, Rapporteur
M. Roland Badeau
Maître de conférence, HdR, Telecom ParisTech, Rapporteur
M. Jérôme Mars
Professeur, Grenoble-INP, Examinateur
Mme. Valérie Perrier
Professeur, Grenoble-INP, Examinatrice
M. Sylvain Meignen
Maître de conférence, HdR, Grenoble-INP, Directeur de thèse

Contributions to the Analysis of Multicomponent Signals

Synchrosqueezing and Associated Methods



PHAM Duong Hung

Laboratoire Jean Kuntzmann
Université Grenoble Alpes

This dissertation is submitted for the degree of
Doctor of Philosophy

September 2018

I would like to dedicate this dissertation to my loving parents.

Remerciements

This thesis is the result of a three-year working experience as PhD candidate at the Laboratoire Jean Kuntzmann, Université Grenoble Alpes, Grenoble, France. This work was supported by a fellowship from Programme de Bourses d'Excellence de l'Ambassade de France au Vietnam in collaboration with the Laboratoire Jean Kuntzmann.

First and foremost, I would like to express my deepest gratitude to my great advisor, Sylvain Meignen. Over the course of countless research meetings, he has taught me how to take a loose idea or result and push it to its limits by discovering variations and generalizations, implications and applications. He has also taught me to think of the big picture and present my work accordingly. In addition, Sylvain's advice has extended beyond research to teaching, career planning, and even personal issues; and he has always been willing to take extra time out of his busy schedule when important or unusual issues came up. Meanwhile, he worked behind the scenes to make sure that I did not have to worry about funding or adminstrative problems, and could focus on research. Finally, and perhaps most importantly, Sylvain has been endlessly motivating and encouraging and brings his infective energy and enthusiasm to everything he does.

Besides my advisor, I would like to thank Prof. Patrick Flandrin (président); Prof. Jérôme Mars and Prof. Valérie Perrier (examinateurs) for serving as my committee members despite their busy schedules, especially Prof. François Auger and Prof. Roland Badeau (rapporteurs) for their valuable and insightful feedback on a preliminary version of this dissertation.

Working at the Laboratoire Jean Kuntzmann has taught me much about how my work relates and applies to practice, and I am grateful for all the interactions that I have had with the great people there, especially Ph.D students for their sympathy and friendship.

Finally, I owe special gratitude to my family for continuous and unconditional support to my parents, for giving birth to me at the first place and supporting me spiritually throughout my life, and am sad to always be living so far away from them. Perhaps science will eventually make the distance less significant.

Abstract

Many physical signals including audio (music, speech), medical data (electrocardiogram (ECG), phonocardiogram (PCG)), marine mammals or gravitational-waves can be accurately modeled as a superposition of amplitude and frequency-modulated waves (AM-FM modes), called multicomponent signals (MCSs). Time-frequency (TF) analysis plays a central role in characterizing such signals and in that framework, numerous methods have been proposed over the last decade. However, these methods suffer from an intrinsic limitation known as the uncertainty principle. In this regard, reassignment method (RM) was developed with the purpose of sharpening TF representations (TFRs) given respectively by the short-time Fourier transform (STFT) or the continuous wavelet transform (CWT). Unfortunately, it did not allow for mode reconstruction, in opposition to its recent variant known as synchrosqueezing transforms (SST). Nevertheless, many critical problems associated with the latter still remain to be addressed such as the weak frequency modulation condition, the mode retrieval of an MCS from its downsampled STFT or the TF signature estimation of irregular and discontinuous signals. This dissertation mainly deals with such problems in order to provide more powerful and accurate invertible TF methods for analyzing MCSs.

This dissertation gives six valuable contributions. The first one introduces a second-order extension of wavelet-based SST along with a discussion on its theoretical analysis and practical implementation. The second one puts forward a generalization of existing STFT-based synchrosqueezing techniques known as the high-order STFT-based SST (FSSTn) that enables to better handle a wide range of MCSs. The third one proposes a new technique established on the second-order STFT-based SST (FSST2) and demodulation procedure, called demodulation-FSST2-based technique (DSST2), enabling a better performance of mode reconstruction. The fourth contribution is that of a novel approach allowing for the retrieval of modes of an MCS from its downsampled STFT. The fifth one presents an improved method developed in the reassignment framework, called adaptive contour representation computation (ACRC), for an efficient estimation of TF signatures of a larger class of MCSs. The last contribution is that of a joint analysis of ACRC with non-negative matrix factorization (NMF) to enable an effective denoising of phonocardiogram (PCG) signals.

Keywords: time-frequency or time-scale analysis, uncertainty principle, AM-FM signals, reassignment, synchrosqueezing.

Résumé

De nombreux signaux physiques incluant des signaux audio (musique, parole), médicaux (electrocardiogramme (ECG), phonocardiogramme (PCG)), de mammifères marins ou d'ondes gravitationnelles peuvent être modélisés comme une superposition d'ondes modulées en amplitude et en fréquence (modes AM-FM), appelés signaux multicomposantes (SMCs). L'analyse temps-fréquence (TF) joue un rôle central pour la caractérisation de tels signaux et, dans ce cadre, diverses méthodes ont été développées au cours de la dernière décennie. Néanmoins, ces méthodes souffrent d'une limitation intrinsèque appelée le principe d'incertitude. Dans ce contexte, la méthode de réallocation (MR) a été développée visant à améliorer les représentations TF (RTFs) données respectivement par la transformée de Fourier à court terme (TFCT) et la transformée en ondelette continue (TOC), en les concentrant autour des lignes de crête correspondant aux fréquences instantanées. Malheureusement, elle ne permet pas de reconstruction des modes, contrairement à sa variante récente connue sous le nom de transformée synchrosqueezée (TSS). Toutefois, de nombreux problèmes associés à cette dernière restent encore à traiter tels que le traitement des fortes modulations en fréquence, la reconstruction des modes d'un SMC à partir de sa TFCT sous-échantillonnée or l'estimation des signatures TF de modes irréguliers et discontinus. Cette thèse traite principalement de tels problèmes afin de construire des nouvelles méthodes TF inversibles plus puissantes et précises pour l'analyse des SMCs.

Cette thèse offre six nouvelles contributions précieuses. La première contribution introduit une extension de TSS d'ordre deux appliqué à la TOC ainsi qu'une discussion sur son analyse théorique et sa mise en œuvre pratique. La seconde contribution propose une généralisation des techniques de synchrosqueezing construites sur la TFCT, connue sous le nom de transformée synchrosqueezée d'ordre supérieur (FTSSn), qui permet de mieux traiter une large gamme de SMCs. La troisième contribution propose une nouvelle technique utilisant sur la transformée synchrosqueezée appliquée à la TFCT de second ordre (FTSS2) et une procédure de démodulation, appelée DTSS2, conduisant à une meilleure performance de la reconstruction des modes. La quatrième contribution est celle d'une nouvelle approche permettant la récupération des modes d'un SMC à partir de sa TFCT sous-échantillonnée. La cinquième contribution présente une technique améliorée, appelée calcul de représentation des contours adaptatifs (CRCA), utilisée pour une estimation efficace des signatures TF d'une plus grande classe de SMCs. La dernière contribution est celle d'une analyse conjointe entre l'CRCA et la factorisation matricielle non-négative (FMN) pour un débruitage performant des signaux phonocardiogrammes (PCG).

Mot clés: Analyse temps-fréquence or temps-échelle, principe d'incertitude, signaux AM-FM, réallocation, synchrosqueezing.

List of my publications

Journal Papers

[1] S. Meignen and D.-H. Pham, "Retrieval of the Modes of Multicomponent Signals from Downsampled Short-Time Fourier Transform," *IEEE Transactions on Signal Processing*, 2018. Accepted.

[2] D. H. Pham, S. Meignen, N. Dia, J. Fontecave-Jallon, and B. Rivet, "Phonocardiogram signal denoising based on non-negative matrix factorization and adaptive contour representation computation." *IEEE Signal Processing Letters*, vol. 25, pp. 1475–1479, October 2018.

[3] D.-H. Pham and S. Meignen, "Second-order synchrosqueezing transform: The wavelet case, comparisons and applications, " *hal-01586372*, 2018. Submitted.

[4] D.-H. Pham and S. Meignen, "An adaptive computation of contour representations for mode decomposition," *IEEE Signal Processing Letters*, vol. 24, pp. 1596–1600, November 2017.

[5] D.-H. Pham and S. Meignen, "High-order synchrosqueezing transform for multicomponent signals analysis - with an application to gravitational-wave signal," *IEEE Transac tions on Signal Processing*, vol. 65, pp. 3168–3178, June 2017.

[6] S. Meignen, D.-H. Pham, and S. McLaughlin, "On demodulation, ridge detection and synchrosqueezing for multicomponent signals," *IEEE Transactions on Signal Processing*, vol. 65, no. 8, pp. 2093–2103, 2017.

Conferences & Workshops

[7] D.-H. Pham and S. Meignen, "A novel thresholding technique for the denoising of multicomponent signals," in *IEEE ICASSP, Calgary, Alberta, Canada*, 2018.

[8] D.-H. Pham, "Poster for signal processing meets deep learning," in *IEEE-EURASIP Summer School, Capri, Italy*, 2017.

[9] D.-H. Pham, "Analysis, synthesis and transformations by reassignment, emd and synchrosqueezing," in *Workshop ASTRES, Grenoble*, 2016.

Table of contents

List of figures

List of tables

Nomenclature

Mathematical Symbols

$\mathbb{N}$, $\mathbb{Z}$	set of natural numbers and integers
$\mathbb{R}$, $\mathbb{C}$	set of real and complex numbers
$\mathbb{R}_+$	set of non-negative real numbers
i	unit imaginary number $i^2 = -1$
$\lfloor z \rfloor$	floor function (the greatest integer less than or equal to z)
$\Re\mathrm{e}\{z\}$, $\Im\mathrm{m}\{z\}$	real, imaginary parts of a complex number z
$\arg\{z\}$, $\lvert z \rvert$	argument, modulus of a complex number z
$\overline{z}$	complex conjugate of z
$\langle z \rangle$	mean value of z
$\mathrm{supp}(f)$	support of function f
δ	Dirac function
$\mathbf{1}_I$	indicator function on interval I
p.v.$\{.\}$	Cauchy principal value
sgn	sign function
Var	Variance

Spaces and Norms

$\lVert . \rVert_p$	l_p-norm. For a function f: $\lVert f \rVert_p = \left(\int \lvert f(x) \rvert^p \mathrm{d}x\right)^{1/p} < +\infty$ and for a vector $x \in \mathbb{C}^n$: $\lVert x \rVert_p = \left(\sum_{i=1}^n \lvert x_i \rvert^p\right)^{1/p}$
$\lVert . \rVert_\infty$	sup-norm: $\lVert f \rVert_\infty = \sup_{\mathbb{R}} \lvert f(x) \rvert$ and $\lVert x \rVert_\infty = \max_{1 \le i \le n} \lvert x_i \rvert$
$L^p(\mathbb{R})$	space of p-integrable functions $\lVert f \rVert_{L^p} = \left(\int \lvert f(x) \rvert^p \mathrm{d}x\right)^{1/p} < +\infty$, for $1 \le p < +\infty$
$L^\infty(\mathbb{R})$	space of essentially bounded measurable functions: $\lVert f \rVert_\infty = \inf\{C \ge 0 : \lvert f(x) \rvert \le C$ for almost every $x\}$
$\langle \cdot, \cdot \rangle_X$	scalar product in Hilbert space X. In $L^2(\mathbb{R})$, $\langle f, g \rangle = \int_{\mathbb{R}^2} f(x)\overline{g(x)}\mathrm{d}x$

$\mathcal{C}^k(\mathbb{R})$	space of continuously differentiable functions of order k on $\mathbb{R}$
$C_c^k(\mathbb{R})$	space of continuous and compactly supported functions on $\mathbb{R}$
$\mathcal{D}(\mathbb{R})$	space of compactly supported smooth functions $\mathcal{C}^\infty(\mathbb{R})$ on $\mathbb{R}$
$\mathcal{S}(\mathbb{R})$	the Schwartz space consisting of smooth functions whose derivatives of any order decay at infinity faster than any polynomial

Transforms and Operators

$\mathcal{F}(f), \hat{f}$	Fourier Transform, $\hat{f}(\eta) = \int_{\mathbb{R}} f(t)e^{-i2\pi\eta t}dt$
$\mathcal{F}^{-1}(f), \check{f}$	inverse Fourier Transform, $\check{f}(t) = \int_{\mathbb{R}} \hat{f}(\eta)e^{i2\pi\eta t}d\eta$
$\mathcal{H}\{.\}$	Hilbert Transform.
$\bigtriangledown f$	gradient of function f
$\partial_x f$	partial derivative of function x with respect to variable t
$\frac{df}{dt}$	derivative of function f with variable t

Acronyms / Abbreviations

ACRC	Adaptive Contour Representation Computation
AM	Amplitude Modulation
ARMA	AutoRegressive Moving Average
AS	Analytic Signal
ASTRES	Analysis, Synthesis and Transformations by Reassignment, EMD and Synchrosqueezing
BA	Basin of Attraction
CP	Contour Point
CVD	Cardiovascular Disease
DFSST2	Demodulation-FSST2-based Technique
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
EMD	Earth Mover's Distance
EMD	Empirical Mode Decomposition
FFT	Fast Fourier Transform
FM	Frequency Modulation
FR	Frequency Representation

FSST2	Second-order Fourier-based Synchrosqueezing Transform
FSST	Fourier-based Synchrosqueezing Transform
FSSTn	Nth-order STFT-based SynchroSqueezing Transform
FT	Fourier Transform
HT	Hard Thresholding
IA	Instantaneous Amplitude
IDFT	Inverse Discrete Fourier Transform
IF	Instantaneous Frequency
IFT	Inverse Fourier Transform
ITFR	Ideal Time Frequency Representation
IP	Instantaneous Phase
LIGO	Laser Interferometer Gravitational-Wave Observatory
LPA	Local Projection Angle
AM-FM	Amplitude Modulation-Frequency Modulation
MCS	Multicomponent Signal
MSE	Mean Square Error
NMF	Non-negative Matrix Factorization
PCG	Phonocardiogram
RM	Reassignment Method
RMSE	Root Mean Squared Error
RV	Reassignment Vector
SNR	Signal-to-Noise Ratio
SSR-HT	Shifted-Symmetrized-Regularized Hard-Thresholding
SST	Synchrosqueezing Transform
STFT	Short Time Fourier Transform
STGFT	Short Time Generalized Fourier Transform
TFR	Time Frequency Representation
TF	Time Frequency
TR	Time Representation
UP	Uncertainty Principle
WVD	Wigner-Ville Distribution

WSST2	Second-order Wavelet-based Synchrosqueezing Transform
WSST	Wavelet-based Synchrosqueezing Transform
WT	Wavelet Transform

Chapter 1

Introduction

1.1 Context and Questions of Research

Signal processing is considered as a key factor in a vast array of applications including audio, speech recognition, radar, sonar, geophysical exploration, medical biology, physics or systems related to communication, processing or retrieval of information. The fundamental goal in such applications is to provide underlying information on the specific problems for the purpose of making decision (extraction, detection, estimation, classification, recognition of patterns, synthesis or morphing) [12–15]. One of the most powerful methods developed over the past century to respond to this goal is probably *time-frequency (TF) analysis*, pioneered by Alfréd Haar [16] and Dennis Gabor [17]. The major motivation stems from the fact that time domain or frequency domain (*Fourier transform* (FT)) representation alone is not sufficient to describe *non-stationary signals* whose frequency content changes with time. Such signals are indeed very commonly encountered in real-world situations such as musical notes [13, 14], marine mammals [18], radars and sonars [19], gravitational-wave [20, 21], or biomedicine [22–24]. One conjectures that these limitations can be effectively overcome by *time-frequency representations* (TFR) that study the signal jointly as a two-dimensional object of time and frequency rather than separately.

One thus has to seek mathematical transformations that enable to map the analyzed signal into its TF domain. This then brings up the question: which *generalized FT* establishes this mapping? The primary answers are the most basic forms of linear TF methods, *short-time Fourier transform* (STFT) and *wavelet transform* (WT). However, one is faced with a fundamental limitation of such methods, known as the *uncertainty principle* (UP), which stipulates that one cannot localize a signal with an arbitrary precision both in time and frequency. Many attempts have been made to cleverly circumvent this trade-off, among which the most celebrated *Wigner-Ville distribution* (WVD). Actually, it belongs to quadratic TF methods in which one can interpret a signal as energy density. In addition to WVD, the latter consists of other members, as for instance, *Cohen class*, a special case of which

being squared STFT (spectrogram), and *affine class* illustrated by squared WT (scalogram) [12, 25, 13]. Although offering a dramatic improvement in terms of TF resolution, WVD inevitably produces spurious cross-term interference hampering the TF readability of the so-called *multicomponent signals* (MCSs) [26, 27, 12]. Such a signal model defined as the sum of AM-FM components or modes will be extensively studied in this thesis. In fact, it was successfully used for modeling a variety of physical phenomena, as for example, speech processing [28], biomedical signal [29–31], or fingerprint [32, 33]. The most significant attempt to reduce the interference generated by WVD while retaining many useful properties is the use of a smoothing operation. However, this operation yields the negative effect of spreading out the signal in the TF plane as a consequence of the UP [12, 25]. A considerable effort has been put in to address this issue, among which a general methodology to sharpen TFRs, called *reassignment method* (RM), was proposed. The concept of RM dates back to Kodera et *al.* in the 1970s, in a somehow restricted framework [34, 35], and then further developed in [36], as a post-processing technique. It, however, faced with the limitation that the reassigned transform is no longer invertible.

In the context of audio signal analysis [37], Daubechies and Maes proposed another phase-based technique, called *synchrosqueezing transform* (SST), whose theoretical analysis followed in [38]. Its purpose is relatively similar to that of RM, i.e. to sharpen the time-scale (TS) representation given by CWT, with the additional advantage of allowing for the mode retrieval. Using the principle of *wavelet-based SST* (WSST), Thakur and Wu proposed an extension of SST to the TFR given by STFT that is called *STFT-based SST* (FSST) [39]. It was then proven to be robust to small bounded perturbations and noise [40], while other efforts were put on exploring the bidimensional case, as for instance by using the monogenic synchrosqueezed wavelet transform [41], developing other types of TFRs as the synchrosqueezed wavelet packet transform [42, 43], or multi-taper approaches as in the ConceFT technique [44]. In spite of all these advances, the applicability of SST is somewhat hindered by the requirement of weak frequency modulation hypothesis for the modes constituting the signal. In contrast, most real signals are made up of very strongly modulated AM-FM modes, as for instance chirps involved in radar [45], speech processing [46], or gravitational waves [20, 21]. Many different approaches and methods have been applied in an effort to deal with this disadvantage, among which a recent adaptation of FSST to the context of strongly modulated modes, known as the *second-order synchrosqueezing transform* (FSST2), was proposed in [47, 48], and further mathematically analyzed in [49]. Unfortunately, the aforementioned technique was proven to only provide an ideal invertible TFR for linear chirps with Gaussian modulated amplitudes, which is still restrictive when processing other more general signals. Arising from these studies, the *first two* research questions need to be answered. The first one relates to a lack of a counterpart of FSST2 established in the wavelet context while the second one puts the focus on the improvement of the existent synchrosqueezing techniques to enable to handle a wider class of MCSs.

Besides, performing signal demodulation before applying SST was of high interest for taking into account the weak frequency modulation condition [50, 51]. However, such a demodulation required the phase computation of the analytic signal (AS) associated with the MCS. It is well-known that the latter cannot be related to the instantaneous frequencies (IF) of the modes which the MCS consists of, which are considered as a key ingredient in the demodulation of MCS. In relation to this issue, in [51], 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, but mode retrieval was not discussed. Therefore, the *third* research question arises of how to include the demodulation procedure into the synchrosqueezing framework more appropriately and efficiently to enable the handling of the above drawbacks.

Moreover, a very important issue related to the use of SST is that while the latter is sharper than the TFR it is based on, mode reconstruction cannot be carried out from SST downsampled in time. On the contrary, such a reconstruction can be achieved from a downsampled STFT: in some sense, SST, by sharpening the frequency representation, annihilates the time redundancy of the TFR. Then, the *fourth* research question arises of how to use the latter correctly so that the reconstruction of the modes of a noisy MCS can be performed with a better accuracy and using fewer STFT coefficients than the one associated with SST.

Finally, another critical problem of SST is that it assumes the instantaneous characteristics (i.e. instantaneous amplitude, frequency, and phase) of the modes to be regular and continuous. This is in contrast to many practical signals including marine mammals [52], damped tones [53], musical sounds [54, 55], or thermoacoustic vibration [56] that contain many highly irregular or discontinuous modes, known as Dirac impulses or vanishing modes. Many different attempts have been made to tackle this problem, among which the methods, built on the properties of reassignment vector (RV), draw the most attention [57–59]. However, such methods cannot handle signals containing noisy Dirac impulses, which is also the main objective of the *fifth* research question needed to be addressed. Additionally, another serious problem related to these RV-based methods is that they do not efficiently handle signals containing highly energetic noises which are very common in *phonocardiogram* (PCG) signals [59, 5]. Indeed, in the case of such signals, the RV-based methods only yield contours in order of descending, but their nature (signal or noise) cannot be identified. In another direction, an appealing method for denoising PCG signals that relied upon their analysis accompanied by the one associated with an *electrocardiogram* (ECG) recorded simultaneously, was developed in [22]. Such a method comprises three main steps: first, one decomposes the spectrograms of PGC and ECG using *non-negative matrix factorization* (NMF); secondly, one computes the cross-correlation between the so-called *activation functions* derived from the above decomposition; thirdly, one thresholds this cross-correlation enabling an efficient noise elimination. However, the threshold is fixed a priori making the method not adaptive. Therefore, the *sixth* research

question arises as to whether a joint analysis of the RV-based approaches and NMF technique is profitable for an adaptive denoising for PCG signals and how to perform this analysis.

1.2 Outline of This Dissertation

This dissertation is devoted to a large extent to the SSTs and associated methods for the analysis of MCSs. The main objective is to bring powerful and efficient solutions to respond respectively to all the six research questions highlighted above. For a better understanding, we succinctly summarize such solutions as in the following outline.

In Chapter 2, we provide a concise overview of previous research done in the analysis of MCSs using TFRs. We first recall the major reason for using TFRs for signal description, followed by some useful notation, concepts and definitions. Then, we recall several typical TFRs either belonging to linear or quadratic classes: STFT, WT, spectrogram, WVD, Cohen class, scalogram, and affine class. We then point out an intrinsic limitation associated with such methods known as the uncertainty principle (UP). Next, we introduce the RM enabling to cope with this issue. We then show an inherent limitation associated with RM which states that it does not allow for the mode reconstruction, and finally introduce its invertible variant SST.

In Chapter 3, we put forward a novel second-order extension of WSST: second-order wavelet-based SST (WSST2), for which we conduct a mathematical analysis and discuss its practical implementation. This work has been just done in [3]. Then, we propose an improvement of existing STFT-based SSTs by computing more accurate estimates of the IFs of the modes making up the signal, using higher-order approximations both for the amplitude and phase. This enables a perfect concentration and reconstruction for a wider variety of AM-FM modes than what was possible up to now with synchrosqueezing techniques. This work was published in [5].

In Chapter 4, we provide two new contributions to the analysis of MCSs developed in the STFT framework. First, we introduce a technique established on the second-order STFT-based SST (FSST2) and demodulation procedure, called demodulation-FSST2-based technique (DFSST2). This results in an even sharper TFR as well as a better mode reconstruction than the ones obtained by only FSST2. This work was published in [6]. Next, we present a novel method for the retrieval of modes of a noisy MCS carried out from its downsampled STFT using the key property of STFT: time redundancy. This work has been completed in [1].

In Chapter 5, we first introduce an improved technique, called adaptive contour representation computation (ACRC) that relies on the use of the properties of the reassignment vector (RV) to efficiently compute TF signatures of MCSs. In fact, existing RV-based techniques cannot assess the latter associated with a noisy Dirac impulse. Instead, the improved technique uses a local averaging rather than a punctual orientation of RV to define the direction of projection. This results in an effective estimation of that impulse, as well as AM-FM

modes. Also, the resultant TF estimates enable the definition of particular corresponding TF regions called basins of attractions (BAs) subsequently used for mode retrieval. This work was published in [4]. Then, we present a novel method based on a joint analysis between ACRC and non-negative matrix factorization (NMF). Such a method is called NMF-ACRC that enables for a better performance in denoising phonocardiogram (PCG) signals. This work has been just carried out in [2].

In Chapter 6, we summarize the major contributions and results of this dissertation and give some directions for future research.

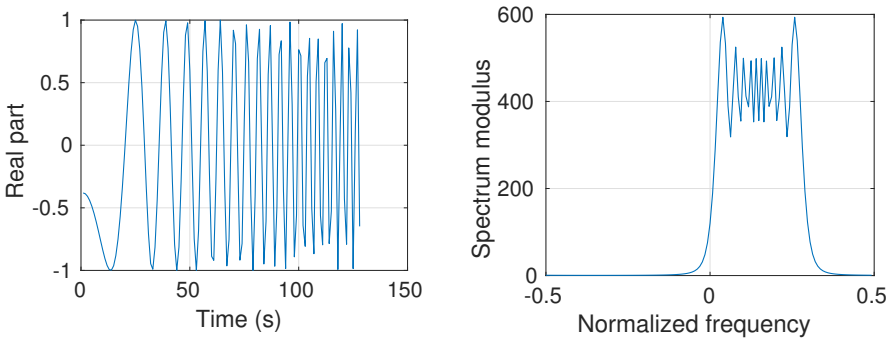
Chapter 2

Background and state of the art

This chapter provides a brief state of the art on the analysis of multicomponent signals (MCSs) from their time-frequency (TF) and time-scale (TS) representations. These tools were initially proposed respectively in [16] and [17] and historically developed as a key factor for solving many decision problems [13, 25, 60]. We first recall, in Section 2.1, the motivation for developing TF representations (TFRs) to describe non-stationary signals, followed by some useful notation and definitions in Section 2.2. Then, we introduce, in Section 2.3, TFRs either belonging to *linear* or *quadratic* classes: for the first one, the focus is put on the *short time Fourier transform* (STFT) and *wavelet transform* (WT), while for the second, we review some important examples: spectrogram, Wigner-Ville distribution (WVD), Cohen class, scalogram, and affine class. We then discuss, in Section 2.4, an intrinsic limitation of such methods known as the *uncertainty principle* (UP) which states that a signal cannot be arbitrarily well concentrated in both time and frequency. To deal with this issue, we recall, in Section 2.5, the principle of the so-called *reassignment method* (RM), which enables the concentration of the energy distribution to improve the readability of TFR. However, RM does not allow for the mode retrieval, in contrast to its variant called *synchrosqueezing transform* (SST) that we present in Section 2.6. Finally, we draw a conclusion of this chapter in Section 2.7.

2.1 From Time and Frequency Representation to Time-Frequency Representation

Time representation (TR) and frequency representation (FR) are the two most classical and widespread methods for signal representations [61, 62, 25]. The former shows how the amplitude of a signal changes with time, whereas the latter, obtained by Fourier Transform (FT) (see Section 2.2.1 for FT definition of a signal), uncovers its spectral content. However, neither of them can deal with signals having time-varying frequency content, known as *non-stationary* signals, as found in a wide range of applications, such as radar, sonar, telecommunications, [13, 12], meteorology, structural stability analysis [63–65], or biomedicine [23, 24]. As an



(a) (b)

Fig. 2.1 Time and frequency representations of a linear FM signal with duration 128 seconds: (a): time domain; (b): frequency domain: energy spectrum.

illustration, Figure 2.1 displays TR and FR of a linear *frequency-modulated* (FM) signal with constant amplitude. From these figures, it is difficult to tell the type of modulation, i.e. linear, parabolic or hyperbolic, and the energy spectrum (in normalized frequency)[1] only shows what frequencies are present, but not their time localization.

In order to tackle these issues, one generally combines TR and FR into a single representation, called *time-frequency representation* (TFR), which shows how the spectral content of a signal evolves with time and proves to be an ideal tool to dissect, analyze and interpret non-stationary signals [25]. A variety of TFRs were proposed and developed in the literature and mainly categorized as linear or quadratic methods [12, 25, 66, 67]. Linear methods, which notably include STFT and WT, focus on making FT time-dependent, while quadratic methods, as for instance the spectrogram, scalogram, and WVD, focus on the associated spectral density. All of these methods are respectively discussed in the following section, but before that, we present several key notation and definitions that will be used in the sequel.

2.2 Basis Notation and Definitions

2.2.1 Fourier Transform (FT)

The *Fourier Transform* (FT) is a mathematical transformation used to map a function from time to frequency domains. More specifically, one has:

Definition 2.1. *The Fourier Transform (FT)* $\hat{f}$ *of a function* $f \in L^1(\mathbb{R})$ *is defined by:*

$$\hat{f}(\eta) = \mathcal{F}(f(\eta)) = \int_{\mathbb{R}} f(t) e^{-i2\pi\eta t} dt. \tag{2.1}$$

[1] Ratio between the frequency in Hertz η and the sampling frequency η_s, with respect to the Shannon sampling theorem.

If $\hat{f}$ is also integrable, f can be reconstructed via the inverse FT (IFT).

$$f(t) = \int_{\mathbb{R}} \hat{f}(\eta) e^{i2\pi\eta t} d\eta. \tag{2.2}$$

Finally, FT can be extended to $L^2(\mathbb{R})$ through Parseval-Plancherel theorem:

Proposition 2.2. *If $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then its FT $\hat{f}$ is $L^2(\mathbb{R})$, and the function $f \mapsto \hat{f}$ is an isometry with respect to L^2 norm: $\|f\|_2 = \|\hat{f}\|_2$.*

2.2.2 Multicomponent Signal (MCS)

Several types of signal models used to characterize the properties of a given signal were proposed in the literature and generally dichotomized into two classes: deterministic and stochastic models [68, 13]. In this dissertation, we make extensive use of a deterministic model of the so-called *multicomponent signal* (MCS), defined as the superposition of AM-FM components or modes [26, 27, 12].

Definition 2.3. *An MCS can be defined as:*

$$f(t) = \sum_{k=1}^{K} f_k(t) \quad \text{with} \quad f_k(t) = A_k(t) e^{i2\pi\phi_k(t)}, \tag{2.3}$$

for some finite $K \in \mathbb{N}$, $A_k(t)$ and $\phi_k(t)$ are respectively the instantaneous amplitude (IA) and instantaneous phase (IP) functions of the k^{th} mode satisfying: $A_k(t) > 0, \phi_k'(t) > 0$ and $\phi_{k+1}'(t) > \phi_k'(t)$ for all t, where $\phi_k'(t)$ is referred to as the instantaneous frequency (IF) of mode f_k at time t.

Remark 2.4. In what follows, three types of modes extensively used are mathematically defined as.

Definition 2.5. A mode or a chirp is said to be *linear* if the IF varies exactly linearly with time: $\phi_k'(t) = \alpha_0 + \alpha_1 t$ with $\alpha_0, \alpha_1 \in \mathbb{R}$ being called the starting frequency (at time $t = 0$), and the rate of frequency change, respectively.

Definition 2.6. A mode or a chirp is said to be *sinusoidal* if the IF varies exactly sinusoidally with time: $\phi_k'(t) = \alpha_0 + \sin(2\pi\alpha_1 t)$ with $\alpha_0, \alpha_1 \in \mathbb{R}$ being called the starting frequency (at time $t = 0$) and the rate of sinusoidal change in frequency, respectively.

Definition 2.7. A mode or a chirp is said to be *exponential* if the IF varies exactly exponentially with time: $\phi_k'(t) = \alpha_0 + \sin(2\pi\alpha_1 t)$ with $\alpha_0, \alpha_1 \in \mathbb{R}$ being called the starting frequency (at time $t = 0$) and the rate of exponential change in frequency, respectively.

Remark 2.8. It is also noteworthy that as we will see, the modes $A_k(t)e^{i2\pi\phi_k(t)}$ are assumed to be continuous and regular in the first three chapters, while the method *adaptive contour*

representation computation (ACRC), presented in the last chapter, works on highly irregular or discontinuous modes.

the spectrogram is

Definition 2.9. *Such a signal is fully described by its ideal TFR (ITFR) defined by:*

$$\mathrm{ITF}_{\varsigma}(t,\eta) = \sum_{k=1}^{K} A_k(t)^{\varsigma}\delta(\eta - \phi_k'(t)), \tag{2.4}$$

where $\varsigma = 1$ or $\varsigma = 2$ respectively corresponds to a linear or quadratic TFR and δ denotes the Dirac distribution.

2.2.3 Analytic Signal (AS)

The underlying rationale for using analytic signal (AS) lies in the fact that the negative frequency components of the spectrum of the real-valued function f are superfluous, known as the Hermitian symmetry. To be more specific, a signal is real if and only if $\hat{f}(-\eta) = \overline{\hat{f}(\eta)}$, which means that the negative frequency components can be discarded without loss of information. Therefore, this leads to the use of the analytic associate of a given real-valued signal rather than the signal itself, which is expressed as the following definition [17, 12].

Definition 2.10. *The analytic signal (AS) f_a associated with a real signal f is defined by:*

$$f_a(t) = f(t) + i\mathcal{H}\{f(t)\}, \tag{2.5}$$

where $\mathcal{H}(f)$ is the Hilbert transform of f defined in the time domain as:

$$\mathcal{H}\{f(t)\} = \frac{1}{\pi}\mathrm{p.v.}\left\{\int_{\mathbb{R}} \frac{f(\tau)}{t-\tau}d\tau\right\} = \frac{1}{\pi}\lim_{\delta\to 0}\int_{|\tau|\geq\delta}\frac{f(\tau)}{t-\tau}d\tau. \tag{2.6}$$

This definition has simple interpretation: in f_a, the negative frequency components of f have been removed, the strictly positive ones have been doubled, and the DC[2] component is kept unchanged: $\hat{f}_a(\eta) = \hat{f}(\eta) + \mathrm{sgn}(\eta)\hat{f}(\eta)$, where $\mathrm{sgn}(\eta)$ is the *sign* function defined as follows.

$$\mathrm{sgn}(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

[2] spectral component at zero frequency

Let us go back to the above AM-FM signal. The Bedrosian theorem states that if $A(t)$ is with low frequency and $\cos(2\pi\phi(t))$ is with high frequency, then one obtains [69, 70]:

$$\mathcal{H}\{A(t)\cos(2\pi\phi(t))\} = A(t)\mathcal{H}\{\cos(2\pi\phi(t))\}. \tag{2.7}$$

Finally, AS brings various benefits to signal analysis. For example, it enables to define instantaneous frequency (IF) and group delay (GD) that we present in the following subsection or to avoid interference generated by interaction between positive and negative frequency components when using quadratic TFRs (to be discussed in detail in Section 2.3.4).

2.2.4 Instantaneous Frequency (IF) and Group Delay (GD)

The IF of f is defined as the local rate oscillation of the signal at time instant t. It enables a full description of time-varying spectral content of the signal [13, 71].

Definition 2.11. *The IF of a real signal f is defined as the derivative of the instantaneous phase (IP) of its associated AS f_a:*

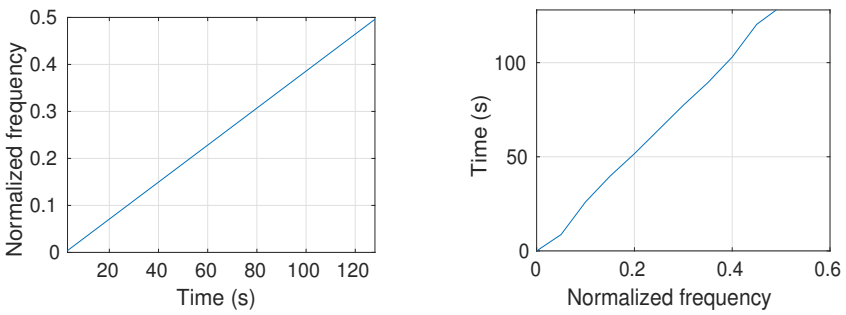
$$\eta_i(t) = \frac{1}{2\pi}\frac{d}{dt}arg\{f_a(t)\}, \tag{2.8}$$

where $arg\{z\}$ denotes the argument of complex number z, and $\frac{df_a}{dt}$ is the time derivative of f_a.

A dual way for the description of the time-varying spectral content is the group delay (GD). It is considered as a local time operator and defined as follows.

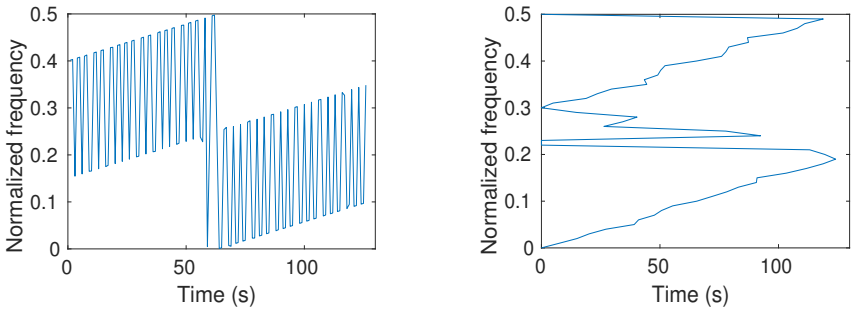
Definition 2.12. *Given a signal f, its GD is defined by:*

$$t_d(\eta) = -\frac{1}{2\pi}\frac{d}{d\eta}arg\{\hat{f}_a(\eta)\}. \tag{2.9}$$



(a) (b)

Fig. 2.2 IF and GD of a linear FM signal: (a): IF; (b): GD



(a) (b)

Fig. 2.3 IF and GD of an MCS: (a): IF; (b): GD

In Figure 2.2, we display the IF and GD of a linear FM signal. We remark that they clearly offer ideal descriptions of the time localization of the spectral content of this *mono-component* signal. Unfortunately, they fail to represent MCSs [13] as illustrated in Figure 2.3, in which the IF and GD of an MCS as a superposition of two linear FM modes are displayed. It is clear that they do not show two different frequencies at each time instant t.

To address this limitation, one generally describes TF structures of an MCS by surfaces defined on the TF plane. Mathematically, this corresponds to joint functions of time t and frequency η, known as *time-frequency representations* (TFRs), some of which are discussed in more details in the following section.

2.3 Some Typical Linear and Quadratic TFRs

Having established the useful notation and concepts, this section reviews several linear and quadratic TFRs, in particular, the motivation for their use, definitions, and properties. A linear TFR is generally obtained by taking inner products of the analyzed signal with a predefined family of templates (called *atoms*) generated by one or a few basic templates [71, 38]. Linear TFRs all satisfy the superposition principle, i.e. the TFR of an MCS equals the sum of the TFRs of its components. This linearity is a desirable property for many practical applications [72, 61, 13], but quadratic TFRs are more commonly used when one wants to interpret a representation as an energy distribution or instantaneous power spectrum. Indeed, as we will see later, a quadratic TFR allows for an approximative description of the energy density of a signal [71, 12]. For a better understanding of all these methods, we discuss respectively, in Sections 2.3.1 and 2.3.2, the short time Fourier transform (STFT) and wavelet transform (WT), while some fundamental examples of quadratic TFRs are reviewed: spectrogram, Wigner-Ville distribution (WVD) and Cohen class, and then scalogram and affine class in the rest of this section.

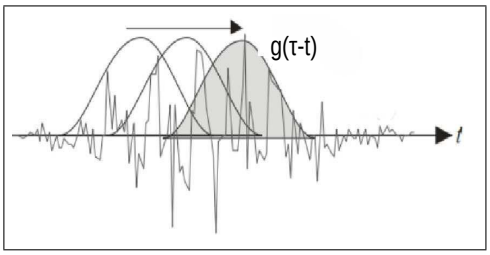


Fig. 2.4 STFT of a signal obtained by computing FT of a block while sliding a time analysis window g

2.3.1 Short Time Fourier Transform (STFT)

STFT is the most widely used method to study non-stationary signals because of its simplicity and powerfulness. The key principle is to break up the signal into several small time blocks that are disjoint or overlapping and then perform FT on each block. This is done by multiplying the analyzed signal with a family of templates obtained by time-translating and frequency-modulating a window function. The principle is explicitly presented in Figure 2.4.

A. Continuous STFT

Definition 2.13. *Given a signal $f \in L^1(\mathbb{R})$, and $g \in L^2(\mathbb{R})$, the (modified) STFT of f is defined by:*

$$V_f^g(t,\eta) = \int_{\mathbb{R}} f(\tau)\overline{g(\tau - t)}e^{-2i\pi\eta(\tau - t)}d\tau, \tag{2.10}$$

where $\overline{g}$ is the complex conjugate of g.

Remark 2.14. This definition is different from the conventional one by the modulation factor $e^{i2\pi\eta t}$, which is added for the sake of consistency with the continuous wavelet transform (CWT) that we will study later.

B. STFT Properties

We here recalled some of nice properties of STFT [13, 73]. First, it can be expressed in terms of signal and window Fourier spectra.

$$V_f^g(t,\eta) = \int_{\mathbb{R}} \hat{f}(\xi)\overline{\hat{g}(\xi - \eta)}e^{i2\pi\xi t}d\xi, \tag{2.11}$$

Secondly, STFT is an invertible linear transform which is expressed as the following theorem.

Theorem 2.15. *Given $f \in L^2(\mathbb{R})$, then $V_f^g \in L^2(\mathbb{R})$ and can be inverted through:*

$$f(\tau) = \frac{1}{\|g\|^2}\int_{\mathbb{R}}\int_{\mathbb{R}} V_f^g(t,\eta)g(\tau - t)e^{i2\pi\eta(\tau - t)}dtd\eta, \tag{2.12}$$

and Parseval-Plancherel equality holds:

$$\|f\|_2^2 = \int_{\mathbb{R}} |f(\tau)|^2 d\tau = \int_{\mathbb{R}} \int_{\mathbb{R}} |V_f^g(t,\eta)|^2 dt d\eta. \tag{2.13}$$

Remark 2.16. This formula also indicates that a signal f can be decomposed into a weighted sum of elementary waveforms:

$$g_{t,\eta}(\tau) = g(\tau - t)e^{i2\pi\eta(\tau - t)}. \tag{2.14}$$

It is referred to as a *building block* or *atom*. Then, STFT is written as: $V_f^g(t,\eta) = \langle f, g_{t,\eta} \rangle$. This corresponds to transformation group of translation in both time and frequency, called *Weyl-Heisenberg* group.

Thirdly, another formula allowing for signal reconstruction from STFT, that we will use later in the STFT-based synchrosqueezing context, are introduced as follows.

Proposition 2.17. *Given $g \in L^2(\mathbb{R})$ and $\overline{g(0)} \neq 0$, then any $f \in L^2(\mathbb{R})$ can be reconstructed through:*

$$f(t) = \frac{1}{\overline{g(0)}} \int_{\mathbb{R}} V_f^g(t,\eta) d\eta. \tag{2.15}$$

Remark 2.18. This is called *STFT Morlet formula* whose proof can be found in [74]. If f is analytic, i.e. $\eta \leq 0$ then $\hat{f}(\eta) = 0$, the integral in (2.15) only takes place on $\mathbb{R}_+$. Further, we can get the same reconstruction formula as above if $f, \hat{f} \in L^1(\mathbb{R})$.

Finally, it should be noted that the analysis window g plays a vital role in the TFR analysis, especially its length [25, 73]. In fact, its choice is associated with a trade-off between time and frequency resolutions, known as the *uncertainty principle* (UP). This issue will be further discussed in Section 2.4.

2.3.2 Wavelet Transform (WT)

The original idea of wavelet transform (WT) emanates from the fact that STFT is associated with a fixed time and frequency resolution because the length of the analysis window remains unchanged, while many signals in practice need more flexibility [12, 71]. Indeed, WT leads to good time resolution for high-frequency events and good frequency resolution for low-frequency events. In order to better understand this transform, we first recall continuous wavelet transform (CWT) computed with real wavelets, then CWT with complex wavelets, and finally some commonly used wavelets.

A. CWT

The principle of CWT is to decompose a signal f over a family of window templates of variable lengths obtained by translating and rescaling (i.e. dilation and contraction) a window

function, called *mother wavelet* [25, 73]:

$$\psi_{t,a}(\tau) = \frac{1}{a}\psi\left(\frac{\tau - t}{a}\right), \tag{2.16}$$

where $a > 0$ represents the scale parameter ($a > 1$ dilates ψ and $a < 1$ contracts ψ), and is inversely proportional to a frequency η, while t is a shifting parameter.

Definition 2.19. *A real wavelet $\psi \in L^2(\mathbb{R})$ is called* admissible *if it satisfies:*

$$0 < C_\psi = \int_{\mathbb{R}_+} \left|\hat{\psi}(\xi)\right|^2 \frac{d\xi}{|\xi|} < \infty. \tag{2.17}$$

This is also called *admissibility condition*. Then, CWT is defined as follows:

Definition 2.20. *Given $f \in L^2(\mathbb{R})$ and a real admissible wavelet ψ, its CWT is defined for $(t, a) \in \mathbb{R} \times \mathbb{R}_+^*$ as:*

$$W_f^\psi(t, a) = \langle f, \psi_{t,a}\rangle = \frac{1}{a}\int_{\mathbb{R}} f(\tau)\overline{\psi\left(\frac{\tau - t}{a}\right)}d\tau. \tag{2.18}$$

By definition, CWT is more a time-scale (TS) than a TF representation. However, for wavelets which are well localized around a non-zero frequency η_0 at scale $a = 1$, a TF interpretation is possible thanks to $\eta = \eta_0/a$.

B. CWT Properties

Several important CWT properties are here presented [25]. The first one is the reconstruction formula from CWT [75, 73].

Theorem 2.21. *If ψ is a real admissible wavelet, $f \in L^2(\mathbb{R})$ can be inverted by:*

$$f(\tau) = \frac{1}{C_\psi}\int_{\mathbb{R}+^*}\int_{\mathbb{R}} W_f^\psi(t, a)\frac{1}{a}\psi(\frac{\tau - t}{a})\frac{dadt}{a}, \tag{2.19}$$

and Parseval-Plancherel equality holds:

$$\|f\|_2^2 = \int_{\mathbb{R}} |f(t)|^2 dt = \frac{1}{C_\psi}\int_{\mathbb{R}+^*}\int_{\mathbb{R}} |W_f^\psi(t, a)|^2 \frac{dadt}{a}. \tag{2.20}$$

Moreover, WT is covariant by translation in time and scale. Finally, like STFT, since it is a windowed transform, it is associated with a trade-off between TF resolutions as a result of the uncertainty principle (UP) [25, 13]. This issue will be further discussed in Section 2.4.

C. CWT with complex wavelets

One of the most critical issues related to the use of CWT is the wavelet choice. In practice, one often chooses a wavelet such that its FT $\hat{\psi}$ is real, regular and centered at frequency η_0. However, an analytic wavelet ψ is much more useful because it can separate the phase and amplitude information of a signal enabling a measurement of the time evolution of frequency transients [73]. In particular, we recall two reconstruction formulae associated with an analytic wavelet, widely used in the wavelet-based synchrosqueezing context, in the following propositions.

Proposition 2.22. *$f \in L^2(\mathbb{R})$ is real and ψ an analytic admissible wavelet ($\eta < 0 \Rightarrow \hat{\psi} = 0$), then one has $W_f^{\psi} = \frac{1}{2} W_{f_a}^{\psi}$ and particularly [73]:*

$$f(\tau) = \frac{2}{C_\psi} \Re\mathrm{e}\left[\int_{\mathbb{R}_+^*} \int_{\mathbb{R}} W_f^{\psi}(t,a) \frac{1}{a} \psi\left(\frac{\tau - t}{a}\right) \frac{dtda}{a} \right], \tag{2.21}$$

and

$$\|f\|_2^2 = \frac{2}{C_\psi} \int_{\mathbb{R}_+^*} \int_{\mathbb{R}} |W_f^{\psi}(t,a)|^2 \frac{dtda}{a}, \tag{2.22}$$

where $\Re\mathrm{e}\{z\}$ denotes the real part of complex number z.

Proposition 2.23. *Let $\psi \in L^2(\mathbb{R})$ such that $0 < C'_\psi = \int_{\mathbb{R}_+} \overline{\hat{\psi}(\xi)} \frac{d\xi}{\xi} < +\infty$, then for any $f \in L^2(\mathbb{R})$ and analytic, one has:*

$$f(t) = \frac{1}{C'_\psi} \int_{\mathbb{R}_+^*} W_f^{\psi}(t,a) \frac{da}{a}. \tag{2.23}$$

The latter is referred to as WT Morlet formula whose proof can be found in [73, 74].

D. Some commonly used mother wavelets

Various wavelets were proposed in the literature, three of which are recalled here [73, 76, 74].

- The most widely used class is that of complex *Morlet wavelets* defined in the frequency domain as:

$$\hat{\psi}_{\sigma,\eta_0}(\eta) = a_{\sigma,\eta_0} e^{-\pi\sigma^2(\eta-\eta_0)^2}, \tag{2.24}$$

  where a_{σ,η_0} is a constant, σ the bandwidth and η_0 the central frequency. Having optimal TF support, they have been extensively used in signal analysis. However, they are not analytic since a Gaussian function is not null at 0.

- The second class is that of *Bump wavelets* defined as:

$$\hat{\psi}_{\mu,\sigma}(\eta) = a_{\mu,\sigma} e^{1 - \frac{1}{1 - \frac{\eta-\mu}{\sigma}}} \mathbf{1}_{[\mu-\sigma,\mu+\sigma]} \tag{2.25}$$

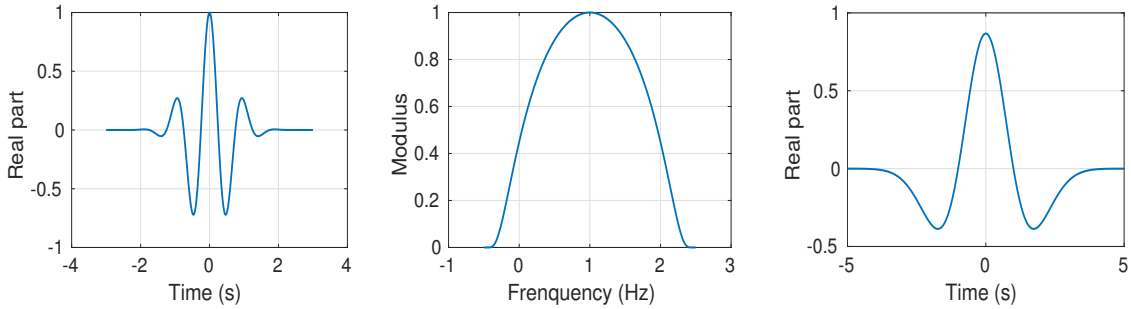
where $a_{\mu,\sigma}$ is a normalizing constant, $\mu > \sigma > 0$, and $\mathbf{1}_I$ the indicator function of I. Note that $\hat{\psi}_{\mu,\sigma} \in \mathbb{C}^\infty$ supported on $[\mu - \sigma, \mu + \sigma]$ and admits central frequency $\eta_0 = \mu$.

- The third class is that of generalized *Morse wavelets* defined as follows:

$$\hat{\psi}_{\beta,\gamma}(\eta) = U(\eta) a_{\beta,\gamma} \eta^\beta e^{-\eta^\gamma}, \tag{2.26}$$

where $a_{\beta,\gamma}$ is a normalizing constant. Their central frequencies are $\eta_0 = (\beta/\gamma)^{\frac{1}{\gamma}}$. For instance, Mexican hat wavelet, defined as the second derivative of the Gaussian function, is included in this class.

As an illustration, we display, in Figure 2.5, some commonly used wavelets, related to the three just mentioned above.



(a) (b) (c)

Fig. 2.5 (a): Morlet Wavelet; (b): Bump wavelet; (c): Mexican hat wavelet.

2.3.3 The Spectrogram

The spectrogram is a good illustration of a quadratic TFR and is defined as the squared modulus of STFT. It, in fact, enables an estimation of the time-varying spectral content of the signal by studying its energy concentration.

Definition 2.24. *Given $f \in L^1(\mathbb{R})$ and $g \in L^2(\mathbb{R})$, the spectrogram is defined by:*

$$|V_f^g(t,\eta)|^2 = \left| \int_{\mathbb{R}} f(\tau) \overline{g(\tau - t)} e^{-2i\pi\eta(\tau - t)} d\tau \right|^2. \tag{2.27}$$

In addition to its quadratic nature, the spectrogram also satisfies time- and frequency-covariant property. Moreover, it is generally interpreted as a measure of the energy of the signal contained in the TF domain since it satisfies the global energy conservation property, namely assuming g is with unit energy [10]:

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |V_f^g(t,\eta)|^2 dt d\eta = \int_{\mathbb{R}} |f(t)|^2 dt. \tag{2.28}$$

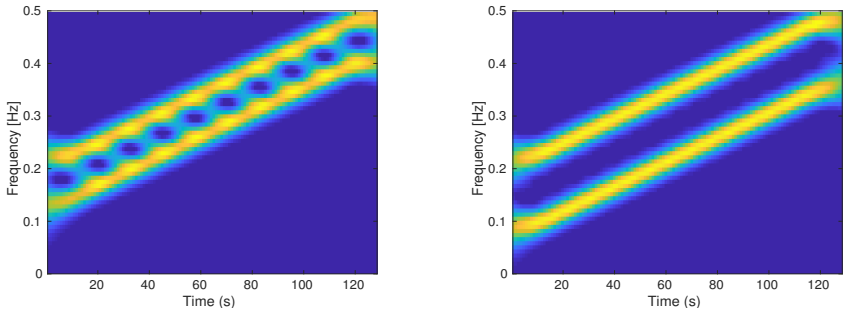


Fig. 2.6 Illustration of interference structure in the spectrogram of two linear parallel chirps: (a) with cross-terms; (b) without cross-terms (using the Toolbox developed in [10]).

Due to this quadratic nature, the spectrogram is also subjected to the so-called cross-term interference that result from oscillating contributions located midway between any two interacting components of an MCS. To be more specific, considering an MCS $f = f_1 + f_2$, one obtains $|V_f^g|^2 = |V_{f_1}^g|^2 + |V_{f_2}^g|^2 + 2\Re\mathrm{e}\left\{V_{f_1}^g \overline{V_{f_2}^g}\right\}$, where $|V_{f_1}^g|^2$ and $|V_{f_2}^g|^2$ are called auto-terms, while $2\Re\mathrm{e}\left\{V_{f_1}^g \overline{V_{f_2}^g}\right\}$ is a cross-term. It is worth mentioning here that if these two components are too close in the TF plane, the cross interference will disturb the readability of TFR and, otherwise they will be essentially zero. To illustrate these phenomena, we display, in Figure 2.6, spectrograms of an MCS composed of two linear parallel chirps computed with the same Gaussian window but different separation level.

Finally, the spectrogram is a non-invertible transform, namely, it does not allow for signal reconstruction because the squaring operation destroys the phase information required for that purpose. Nevertheless, it is used in many practical applications, as for example, fundamental frequency measurement, cross-synthesis, spectral envelope extraction, sinusoidal modeling and FFT audio filter banks [77–79]. More details on issues, properties, and application of the spectrogram can be found in [36, 71, 25].

2.3.4 Wigner-Ville Distribution (WVD) and Cohen Class

In contrast to the spectrogram, WVD is an adaptive quadratic TF method that does not make use of any analysis window. Instead, it uses $g(t) = f_-(t) = f(-t)$, the time-reversed version of the analyzed signal, which leads to [25, 67]:

$$V_f^{f_-}(t,\eta) = \mathcal{W}_f(t/2, \eta/2)/2, \tag{2.29}$$

where

$$\mathcal{W}_f(t,\eta) = \int_{\mathbb{R}} f(t+\tau/2)\overline{f(t-\tau/2)}e^{-i2\pi\eta\tau}d\tau. \tag{2.30}$$

This expression is referred to as WVD. Figure 2.7 depicts the WVD of the two linear parallel chirps displayed in Figure 2.6 (b). Compared with the spectrogram, it clearly shows a more concentrated picture but cross-term still exist [12].

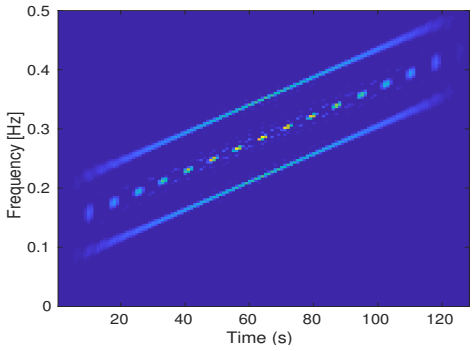


Fig. 2.7 WVD of two linear parallel chirps (using the Toolbox developed in [10])

While WVD enjoys many useful properties: time or frequency-shift invariance, perfect localization of linear chirps, etc., it still suffers from cross-term interference. Numerous techniques were proposed to deal with this issue, among which using analytic signal (AS) rather than real one is the first simple one [12, 25, 13].

Smoothed versions of WVD also bring a huge benefit in terms of interference attenuation. They have a general form:

$$\mathrm{Co}_f(t,\eta;\Pi) = \int_{\mathbb{R}}\int_{\mathbb{R}} \Pi(s-t,\xi-\eta)\mathcal{W}_f(s,\xi)dsd\xi. \tag{2.31}$$

where Π is smoothing kernel. They satisfy: time and frequency shift invariance and belong to the so-called *Cohen class*. Note that WVD and spectrogram are members of this class when $\Pi(t,\eta)$ equals $\delta(t)\delta(\eta)$ and $\mathcal{W}_g(t,\eta)$, respectively. However, the smoothing operation inevitably entails a loss of TF concentration, which leads to a trade-off between TF concentration and interference attenuation. The most widely used method to address this issue is *reassignment method* (RM), which will be discussed in more details in Section 2.5. Other properties and applications of Cohen class can be found in [25, 13, 73].

2.3.5 The Scalogram and Affine Class

Similar to the spectrogram, ones define the *scalogram* as the squared modulus of CWT of a signal.

Definition 2.25. *Given $f \in L^2(\mathbb{R})$ and a real admissible wavelet ψ, its scalogram is defined as the squared modulus of CWT:*

$$|W_f^{\psi}(t,a)|^2 = \left|\frac{1}{a}\int_{\mathbb{R}} f(\tau)\overline{\psi\left(\frac{\tau-t}{a}\right)}d\tau\right|^2. \tag{2.32}$$

It is also a quadratic TF energy distribution but defined in the TS plane. It can be viewed as the counterpart of Cohen class in the wavelet framework, called the *affine class*, which contains all quadratic TF representations covariant by translation in time and dilation. More details on issues, properties, and application of the scalogram, as well as on the affine class are available in [73, 13, 25].

2.4 Uncertainty Principle (UP)

One of the most important issues associated with the use of TFR is the *uncertainty principle* (UP), which states that one cannot localize a signal with arbitrary precision both in time and frequency [12, 13, 67]. The UP expresses a fundamental relation between the standard deviation of a function in time σ_t and the standard deviation of its FT in frequency σ_η, which corresponds to the following theorem.

Theorem 2.26. *Given $f \in L^2(\mathbb{R})$, one defines its mean time and frequency positions:*

$$\langle t \rangle = \frac{1}{\|f\|_2^2} \int_{\mathbb{R}} t|f(t)|^2 dt \quad \text{and} \quad \langle \eta \rangle = \frac{1}{\|f\|_2^2} \int_{\mathbb{R}} \eta|\hat{f}(\eta)|^2 d\eta, \tag{2.33}$$

and the standard deviations along time and frequency:

$$\sigma_t^2 = \frac{1}{\|f\|_2^2} \int_{\mathbb{R}} (t - \langle t \rangle)^2|f(t)|^2 dt \quad \text{and} \quad \sigma_\eta^2 = \frac{1}{\|f\|_2^2} \int_{\mathbb{R}} (\eta - \langle \eta \rangle)^2|\hat{f}(\eta)|^2 d\eta. \tag{2.34}$$

Then, one has:

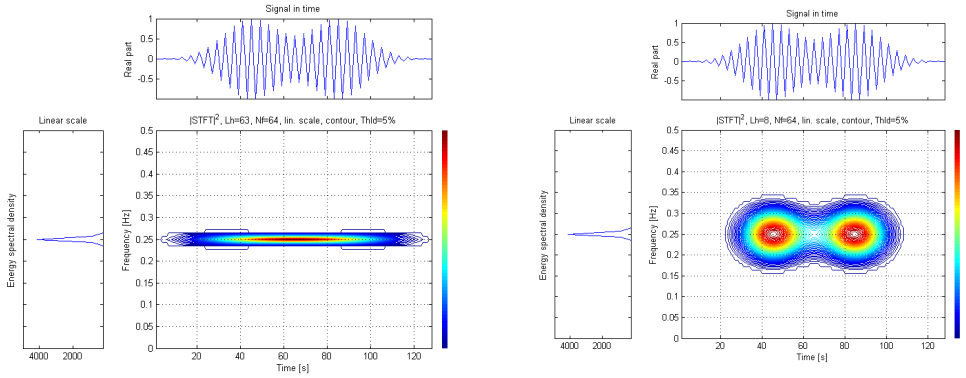
$$\sigma_t \sigma_\eta \geq \frac{1}{2}. \tag{2.35}$$

This expression, known as *Heisenberg-Gabor inequality*, has a simple interpretation: one cannot have a signal for which both σ_t and σ_η are arbitrarily small. The lower bound $\sigma_t \sigma_\eta = \frac{1}{2}$ is reached for Gaussian signals [13]:

$$f(t) = C \exp[(-\beta(t - \langle t \rangle)^2) + i2\pi\langle \eta \rangle (t - \langle t \rangle)], \tag{2.36}$$

where $C \in \mathbb{R}$ and $\beta \in \mathbb{R}_+$.

To illustrate the UP, we consider an MCS composed of two AM Gaussian modes with different time shifts whose STFTs are computed with the Gaussian window with different lengths. The results are depicted in Figure 2.8. In Figure 2.8 (a), the frequency resolution is very good, but it is almost impossible to give any time variation. In contrast, in Figure 2.8 (b), the frequency resolution is coarser, but the time resolution is sufficiently good to lead to some information on the change in the behavior of the signal with the time.



(a) (b)

Fig. 2.8 Two Gaussian atoms analyzed by STFT using a Gaussian window g of different lengths: (a): 128 samples; (b): 17 samples (using the Toolbox developed in [10]).

2.5 Reassignment Method (RM)

The primary rationale for developing RM stems from the fact that the smoothing operation, mentioned in Definition (2.31), smears the signal in the TF plane as a result of the uncertainty principle. For instance, we can see, in Figure 2.6, each AM-FM mode makes up a band of non-zero coefficients with a non-negligible width in the TF plane. A lot of effort has been put in over the years to tackle this issue, among which a widely used technique is the *reassignment method* (RM). It was first introduced in [34, 35], in a somehow restricted framework, and then further developed in [36] as a post-processing technique for signal analysis. Indeed, it focuses the energy of a TFR towards ideal IF curves to lead to a more concentrated TF picture, thus improving the readability of TFRs. Unfortunately, it suffers from an inherent limitation which is its non-invertibility, namely, it does not allow for signal reconstruction. For a better understanding of this method, we introduce, in Section 2.5.1, RM applied to the spectrogram and to Cohen class in Section 2.5.2, and finally to the scalogram in Section 2.5.3.

2.5.1 RM for the Spectrogram

The key principle of RM, when applied to an MCS, is to map each value of a TFR, as for instance the spectrogram, to the location corresponding to the IF curve of the nearest mode [36]. To be more specific, the spectrogram can be equivalently written as: $\int_{\mathbb{R}}\int_{\mathbb{R}} \mathcal{W}_f(s,\xi)\mathcal{W}_g(s-t,\xi-\eta)dsd\xi$, for which the smoothing kernel $\mathcal{W}_g$ enables a reduction of interference, but at the expense of smearing the distribution. To tackle this issue, one considers local energy distribution $\mathcal{W}_f(s,\xi)\mathcal{W}_g(s-t,\xi-\eta)$ as a mass distribution. Then, making an analogy with mechanical physics, one assigns the total mass of the object to its centroid. This is exactly what RM does: it moves each value of the spectrogram computed at a point (t,η) to the *centroid* $(\widehat{\tau}_f(t,\eta),\widehat{\omega}_f(t,\eta))$ of the energy distribution.

Definition 2.27. *Given $f \in L^2(\mathbb{R})$, the centroid (called also reassignment operators) of the spectrogram are defined as:*

$$\widehat{\omega}_f(t,\eta) = \frac{\int_{\mathbb{R}}\int_{\mathbb{R}} \xi \mathcal{W}_f(s,\xi)\mathcal{W}_g(s-t,\xi-\eta)dsd\xi}{\int_{\mathbb{R}}\int_{\mathbb{R}} \mathcal{W}_f(s,\xi)\mathcal{W}_g(s-t,\xi-\eta)dsd\xi}, \tag{2.37}$$

$$\widehat{\tau}_f(t,\eta) = \frac{\int_{\mathbb{R}}\int_{\mathbb{R}} s\mathcal{W}_f(s,\xi)\mathcal{W}_g(s-t,\xi-\eta)dsd\xi}{\int_{\mathbb{R}}\int_{\mathbb{R}} \mathcal{W}_f(s,\xi)\mathcal{W}_g(s-t,\xi-\eta)dsd\xi}. \tag{2.38}$$

Note that the centroid uses phase information of STFT, and can be rewritten as follows:

Proposition 2.28. Given $f \in L^2(\mathbb{R})$, the reassignment operators of the spectrogram are defined for all (t,η) such that $V_f^g(t,\eta) \neq 0$ by:

$$\widehat{\omega}_f(t,\eta) = \frac{1}{2\pi}\partial_t \left\{\arg(V_f^g(t,\eta))\right\}, \tag{2.39}$$

$$\widehat{\tau}_f(t,\eta) = t - \frac{1}{2\pi}\partial_\eta \left\{\arg(V_f^g(t,\eta))\right\}, \tag{2.40}$$

where ∂_x denotes the partial derivative of a function with respect to variable x.

Note also that in practice one uses a more efficient procedure [74]:

Proposition 2.29. *For a signal $f \in L^2(\mathbb{R})$, the expressions $\widehat{\omega}_f(t,\eta)$, $\widehat{\tau}_f(t,\eta)$ can be written for wherever $V_f^g(t,\eta) \neq 0$ by:*

$$\widehat{\omega}_f(t,\eta) = \eta - \frac{1}{2\pi}\Im\mathrm{m}\left\{\frac{V_f^{g'}(t,\eta)}{V_f^g(t,\eta)}\right\}, \tag{2.41}$$

$$\widehat{\tau}_f(t,\eta) = t + \Re\mathrm{e}\left\{\frac{V_f^{tg}(t,\eta)}{V_f^g(t,\eta)}\right\}, \tag{2.42}$$

where $V_f^{tg}, V_f^{g'}$ are respectively STFTs of f computed with windows $t \mapsto tg(t), g'(t)$ and $\Im\mathrm{m}\{z\}$ denotes the imaginary part of complex number z.

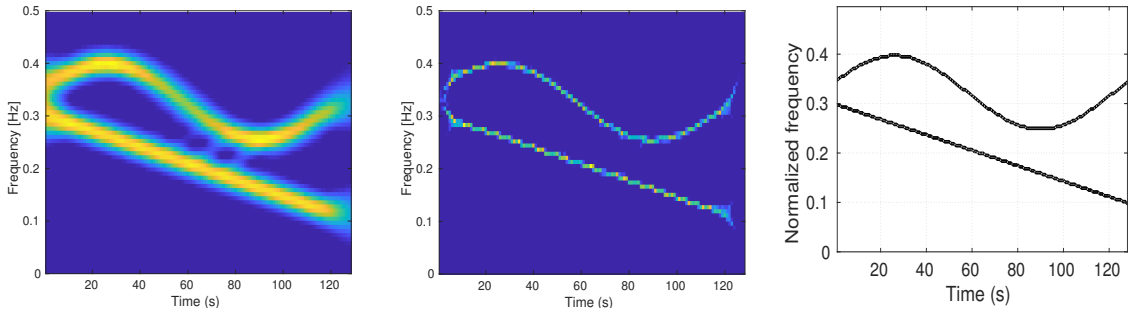
Then, the definition of the reassigned spectrogram is derived as follows.

Definition 2.30. *Given $f \in L^2(\mathbb{R})$, the reassigned spectrogram is defined by:*

$$\breve{\Theta}_f^g(\tau,\omega) = \int_{\mathbb{R}}\int_{\mathbb{R}} |V_f^g(t,\eta)|^2 \delta(\tau - \widehat{\tau}_f(t,\eta))\delta(\omega - \widehat{\omega}_f(t,\eta))dtd\eta, \tag{2.43}$$

where δ denotes the Dirac distribution.

In order to illustrate this method, reassigned spectrogram of an MCS made of a sinusoidal FM and a linear FM chirps is shown in Figure 2.9 (b). It is also compared with its STFT and ITF in Figures (a) and (c), respectively. Obviously, TF concentration given by RM is



(a) (b) (c)

Fig. 2.9 Spectrogram reassignment on an MCS (using the Toolbox developed in [10]): (a) spectrogram; (b): reassigned spectrogram, (c): ideal TFR.

significantly improved compared with the spectrogram: the two components are much more localized.

Moreover, the reassigned spectrogram is no longer bilinear and invertible, but it still satisfies time and frequency shifts invariance, energy conservation provided g is with unit energy and non-negative property. In particular, it gives a perfect localization for linear chirps, which is expressed by the following theorem.

Theorem 2.31. *For a linear chirp $f(t) = Ae^{i2\pi\phi(t)}$, i.e. $A > 0$ and ϕ is a second-order polynomial, one has: $\hat{\omega}_f(t,\eta) = \phi'(\hat{\tau}_f(t,\eta))$.*

The proof of this theorem is available in [36, 74].

Finally, a recent extension of the classical RM based on the Levenberg-Marquardt algorithm enables to adjust the energy localization in the TF plane through a damping parameter μ [80, 81]. The new reassignment operators are defined as follows.

Definition 2.32. *Given $f \in L^2(\mathbb{R})$, the reassignment operators of the spectrogram are defined for all (t,η) such that $V_f^g(t,\eta) \neq 0$ by:*

$$\begin{pmatrix} \widehat{\omega}_f^{\mu}(t,\eta) \\ \widehat{\tau}_f^{\mu}(t,\eta) \end{pmatrix} = \begin{pmatrix} \eta \\ t \end{pmatrix} - \left(\bigtriangledown^t \boldsymbol{R}_f^{\mu}(t,\eta) + \mu \boldsymbol{I}_2 \right)^{-1} \boldsymbol{R}_f^{\mu}(t,\eta), \tag{2.44}$$

where $\boldsymbol{R}_f^{\mu}(t,\eta) = \begin{pmatrix} \eta - \widehat{\omega}_f(t,\eta) \\ t - \widehat{\tau}_f(t,\eta) \end{pmatrix}$ *and* $\bigtriangledown^t \boldsymbol{R}_f^{\mu}(t,\eta) = \begin{pmatrix} \dfrac{\partial \boldsymbol{R}_f^{\mu}(t,\eta)}{\partial \eta} & \dfrac{\partial \boldsymbol{R}_f^{\mu}(t,\eta)}{\partial t} \end{pmatrix}$, *with $\boldsymbol{I}_2$ is the 2×2 identity matrix.*

Then, the Levenberg-Marquardt reassigned spectrogram is defined by replacing $(\widehat{t}_f^{\mu},\ \widehat{\omega}_f^{\mu})$ by $(\widehat{t}_f^{\mu},\ \widehat{\omega}_f^{\mu})$ in (2.43).

2.5.2 RM for Cohen Class

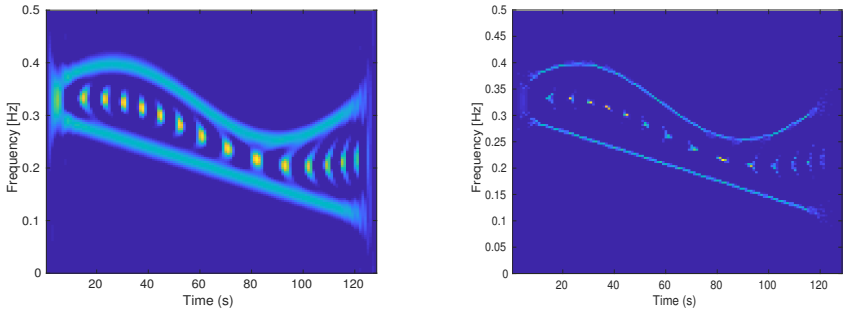
RM is extended to other distributions of Cohen class by replacing WVD $\mathcal{W}_g(t,\eta)$ by an arbitrary kernel $\Pi(t,\eta)$ in formulae (2.37) and (2.43) [36]:

$$\widehat{c\omega}_f(t,\eta) = \frac{\int_{\mathbb{R}}\int_{\mathbb{R}} \xi \mathcal{W}_f(s,\xi)\Pi_g(s-t,\xi-\eta)dsd\xi}{\int_{\mathbb{R}}\int_{\mathbb{R}} \mathcal{W}_f(s,\xi)\Pi_g(s-t,\xi-\eta)dsd\xi}, \tag{2.45}$$

$$\widehat{c\tau}_f(t,\eta) = \frac{\int_{\mathbb{R}}\int_{\mathbb{R}} s\mathcal{W}_f(s,\xi)\Pi_g(s-t,\xi-\eta)dsd\xi}{\int_{\mathbb{R}}\int_{\mathbb{R}} \mathcal{W}_f(s,\xi)\Pi_g(s-t,\xi-\eta)dsd\xi}. \tag{2.46}$$

$$\check{\mathcal{C}}o_f(\tau,\omega) = \int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{Co}_f(t,\eta)\delta(\tau-\widehat{c\tau}_f(t,\eta))\delta(\omega-\widehat{c\omega}_f(t,\eta))dtd\eta. \tag{2.47}$$

Figures 2.10 (a) and (b) respectively illustrate WVD and its reassigned version of the two-component signal presented in Section 2.5.1. Like the reassigned spectrogram, the reassigned WVD generates also a highly concentrated TFR, but there are a few cross-terms still appearing.



(a) (b)

Fig. 2.10 RM for a member of Cohen class (using the Toolbox developed in [10]): (a): WVD; (b): reassigned WVD.

2.5.3 RM for the Scalogram

Similarly, one can define RM for the scalogram as follows [36, 25].

Definition 2.33. *Given $f \in L^2(\mathbb{R})$ and an admissible wavelet ψ, the reassignment operators for the scalogram are defined for all (t,a) such that $W_f^{\psi}(t,a) \neq 0$ by:*

$$\widehat{\omega}_f(t,a) = \Re\mathrm{e}\left\{\frac{1}{i2\pi}\frac{\partial_t W_f^{\psi}(t,a)}{W_f^{\psi}(t,a)}\right\}, \tag{2.48}$$

$$\widehat{\tau}_f(t,a) = \Re\mathrm{e}\left\{\frac{\int_{\mathbb{R}} \tau f(\tau)\frac{1}{a}\overline{\psi(\frac{\tau-t}{a})}d\tau}{W_f^{\psi}(t,a)}\right\}. \tag{2.49}$$

In what follows, we use the variable notation (t, η) and (t, a) to make the distinction between frequency and scale. Further, these operators are practically computed in the Fourier domain by means of only three CWTs as follows:

Proposition 2.34. *Given $f \in L^2(\mathbb{R})$ and an admissible wavelet ψ, the expressions $\widehat{\omega}_f(t,a)$, $\widehat{\tau}_f(t,a)$ can be written wherever $W_f^{\psi}(t,a) \neq 0$ as:*

$$\widehat{\omega}_f(t,a) = -\frac{1}{2\pi a}\Im\mathrm{m}\left\{\frac{W_f^{\psi'}(t,a)}{W_f^{\psi}(t,a)}\right\}, \tag{2.50}$$

$$\widehat{\tau}_f(t,a) = t + a\Re\mathrm{e}\left\{\frac{W_f^{t\psi}(t,a)}{W_f^{\psi}(t,a)}\right\}, \tag{2.51}$$

where $t \mapsto W_f^{\psi'}, W_f^{t\psi}$ are respectively CWTs of f computed with wavelets $\psi', t\psi$.

Then, the definition of the reassigned scalogram is derived as follows.

Definition 2.35. *Given $f \in L^2(\mathbb{R})$ and an admissible wavelet ψ, the reassigned scalogram is defined by:*

$$\breve{\Xi}_f^g(\tau,\omega) = \int_{\mathbb{R}+^*}\int_{\mathbb{R}} |W_f^{\psi}(t,a)|^2 \delta(\tau - \widehat{\tau}_f(t,a))\delta(\omega - \widehat{\omega}_f(t,a))dtda. \tag{2.52}$$

Finally, the reassigned scalogram also localizes perfectly linear chirps.

Theorem 2.36. *For a linear chirp $f(t) = Ae^{i2\pi\phi(t)}$ and if ψ is such that $t \mapsto \psi(t), \psi'(t), t\psi(t)$ are admissible wavelets. Then, one has: $\hat{\omega}_f(t,a) = \phi'(\hat{\tau}_f(t,a))$.*

The proof of this theorem is available in [36].

2.6 Synchrosqueezing Transform (SST)

As mentioned previously, RM proves to be a simple and powerful tool to enhance the readability of TFRs, it, however, does not allow for mode reconstruction, as opposed to another phase-based reassigned technique called *synchrosqueezing transform* (SST). Such a transform was firstly proposed in the wavelet case [37, 38] and then extended to the STFT one [39, 66, 74]. It actually corresponds to a nonlinear operator that sharpens the TFR of a signal, combines the localization and sparsity properties of RM with the invertibility property of linear TFRs, and is robust to a variety of signal perturbations [82, 76, 40]. To better understand this technique, we recall the original wavelet-based SST (WSST) in Section 2.6.1. Then, we move to an extension of WSST: STFT-based SST (FSST) in Section 2.6.2, in which its numerical implementation, associated ridge extraction technique, and numerical simulations are discussed. Finally, we recall, in Section 2.6.3, an adaptation of FSST, known as the second-order FSST (FSST2), to better handle signals with strong frequency modulation.

2.6.1 Wavelet-based Synchrosqueezing Transform (WSST)

A. WSST Principle

The principle of WSST is to sharpen the *blurred* representation given by CWT using the IF estimate at time t and scale a mentioned in Definition 2.33, namely:

$$\widehat{\omega}_f(t,a) = \Re\mathrm{e}\left\{\frac{1}{i2\pi}\frac{\partial_t W_f^\psi(t,a)}{W_f^\psi(t,a)}\right\}. \tag{2.53}$$

In that framework, $W_f^\psi(t,a)$ is only reassigned along the *scale* axis to a new position $(t,\widehat{\omega}_f(t,a))$ using the synchrosqueezing operator defining WSST, as follows:

$$S_f^{\psi,\gamma}(t,\omega) = \int_{|W_f^\psi(t,a)|>\gamma} W_f^\psi(t,a)\delta\left(\omega-\widehat{\omega}_f(t,a)\right)\frac{da}{a}, \tag{2.54}$$

with γ some threshold parameter.

Because the reassignment is only carried out along the scale direction, WSST preserves the *causality property*, thus making the k^{th} mode approximately reconstructed by integrating $S_f^{\psi,\gamma}(t,\omega)$ in the vicinity of the corresponding ridge $(t,\frac{1}{\phi_k'(t)})$. This leads to the associated reconstruction formula:

$$f_k(t) \approx \frac{1}{C_\psi'}\int_{\{\omega,|\omega-\varphi_k(t)|<d\}} S_f^{\psi,\gamma}(t,\omega)d\omega, \tag{2.55}$$

where C_ψ' was defined in Proposition 2.23, $\varphi_k(t)$ is an estimate of $\phi_k'(t)$, and d, that we call *compensation parameter*, enables to compensate for both the inaccurate approximation $\varphi_k(t)$ of $\phi_k'(t)$ and the error made by estimating IF by means of $\widehat{\omega}_f(t,a)$. It is worthy of note here that approximation $\varphi_k(t)$ must be computed before retrieving mode f_k. For that purpose, a commonly used technique relying upon ridge extraction assuming $S_f^{\psi,\gamma}$ and K are known. This technique was initially proposed by Carmona et *al.* [83] and then successfully used in the synchrosqueezing context in [38, 66]. Indeed, it relies upon the minimization of the following *energy functional*:

$$E_f(\varphi_1,\cdots,\varphi_K) = \sum_{k=1}^{K} -\int_{\mathbb{R}} |S_f^{\psi,\gamma}(t,\varphi_k(t))|^2 dt + \int_{\mathbb{R}} \lambda\varphi_k'(t)^2 + \beta\varphi_k''(t)^2 dt, \tag{2.56}$$

where λ and β are regularization parameters chosen such that the trade-off between smoothness of φ_k and energy is maximized. In practice, this energy functional is hard to implement because of its non-convexity, and one has to find tricks to avoid local minima as much as possible, as for example by using a simulated annealing algorithm proposed in [83]. A recent algorithm, first introduced in [40], enables the determination of the ridge associated

with the corresponding mode thanks to a forward-backward approach applied for different initializations. Finally, the influence of the regularization parameters in this algorithm will be analyzed and discussed in more details in Section 4.1.1 of Chapter 4. This work was published recently in [6].

B. WSST Mathematical Framework

WSST is supported by a solid mathematical framework [38] which we now recall. Let us first define the class of chirp-like functions on which one builds the WSST theory:

Definition 2.37. *Let $\varepsilon > 0$ and $c > 0$. The set $\mathcal{A}_{c,\varepsilon}$ of MCSs with modulation ε and separation c corresponds to signals defined in Section 2.2.2 with f_k satisfying:*

$$\begin{aligned} &A_k \in \mathbb{C}^1(\mathbb{R}) \cap L^\infty(\mathbb{R}), \phi_k \in \mathbb{C}^2(\mathbb{R}), \\ &\inf_{t\in\mathbb{R}} \phi_k'(t) > 0, \sup_{t\in\mathbb{R}} \phi_k'(t) < \infty, A_k(t) > 0, \\ &|A_k'(t)| \leq \epsilon\phi_k'(t), |\phi_k''(t)| \leq \epsilon\phi_k'(t) \quad \forall t \in \mathbb{R}. \end{aligned} \tag{2.57}$$

Further, the f_ks are separated with resolution c, i.e., for all $k \in \{1, \ldots, K-1\}$ and all t

$$\phi_{k+1}'(t) - \phi_k'(t) \geq c(\phi_{k+1}'(t) + \phi_k'(t)). \tag{2.58}$$

In what follows, Δ denotes a real number in $]0, \frac{c}{c+1}[$.

Definition 2.38. *Let h be a positive L^1-normed window belonging to $C_c^\infty(\mathbb{R})$ and consider $\gamma, \lambda > 0$. The wavelet-based synchrosqueezing transform of f (WSST) with threshold γ and accuracy λ is defined by:*

$$S_f^{\psi,\lambda,\gamma}(t,\omega) = \int_{|W_f^\psi(t,a)|>\gamma} W_f^\psi(t,a) \frac{1}{\lambda} h\left(\frac{\omega - \widehat{\omega}_f(t,a)}{\lambda}\right) \frac{da}{a}. \tag{2.59}$$

If $\lambda \to 0$, then $S_f^{\psi,\lambda,\gamma}(t,\omega)$ tends, in the sense of distribution, to some value which we formally write as in (2.54).

Theorem 2.39. *Consider $f \in \mathcal{A}_{c,\varepsilon}$, set $\tilde{\varepsilon} = \varepsilon^{\frac{1}{3}}$ and let ψ be a wavelet such that $supp(\widehat{\psi}) \subset [1-\triangle, 1+\triangle]$. Then, provided ε is sufficiently small, the following hold:*

(a) $|W_f^\psi(t,a)| > \tilde{\varepsilon}$ only when, there exists $k \in \{1, \ldots, K\}$, such that $(t,a) \in Z_k := \{(t,a),\ s.t.\ |a\phi_k'(t) - 1| < \Delta\}$.

(b) For each $k \in \{1, \ldots, K\}$ and all $(t,a) \in Z_k$ for which holds $|W_f^\psi(t,a)| > \tilde{\varepsilon}$, one has:

$$|\widehat{\omega}_f(t,a) - \phi_k'(t)| \leq \tilde{\varepsilon}. \tag{2.60}$$

(c) Moreover, for each $k \in \{1, \dots, K\}$*, there exists a constant* $D_{W,1}$ *such that for any* $t \in \mathbb{R}$

$$\left| \lim_{\lambda \to 0} \left(\frac{1}{C'_\psi} \int_{|\omega - \phi'_k(t)| < \tilde{\varepsilon}} S_f^{\psi,\lambda,\tilde{\varepsilon}}(t,\omega) d\omega \right) - f_k(t) \right| \leq D_{W,1} \tilde{\varepsilon}. \tag{2.61}$$

The proof of this theorem is available in [38] and gives a strong approximation result for the class $\mathcal{A}_{c,\epsilon}$, since it ensures that the non-zero coefficients of WSST are localized around the curves $(t, \frac{1}{\phi'_k(t)})$ in the TS space, and that the reconstruction of the modes is easily obtained from the concentrated representation. However, it is more useful for real-time applications [84] to extend this result to *non-compact wavelet* in the Fourier domain. This development has recently been made in [3] and will be presented in Section 3.1.1 of Chapter 3. Finally, numerical implementation, simulations and applications of WSST are well presented in [38, 40, 85].

2.6.2 STFT-based Synchrosqueezing Transform (FSST)

Originally proposed in the wavelet case, SST was extended similarly to the STFT context, known as *STFT-based synchrosqueezing transform* (FSST) in a somehow restrictive framework, in [39, 85]. A more general formulation was proposed in [66] and mathematically analyzed in [49], all of which we now recall.

A. FSST Mathematical Framework

In the FSST framework, the IF estimate of signal f at time t and frequency η, introduced in Definition 2.28, reads:

$$\begin{aligned} \widehat{\omega}_f(t,\eta) &= \frac{1}{2\pi} \partial_t \left\{ \arg(V_f^g(t,\eta)) \right\}, \\ &= \eta - \frac{1}{2\pi} \Im\mathrm{m} \left\{ \frac{V_f^{g'}(t,\eta)}{V_f^g(t,\eta)} \right\}. \end{aligned} \tag{2.62}$$

Then, one defines the class of chirp-like functions on which one builds the FSST theory.

Definition 2.40. *Let* $\epsilon > 0$ *and* $\Delta > 0$*. The set* $\mathcal{B}_{\Delta,\epsilon}$ *of MCSs with modulation* ϵ *and separation* Δ *is the set of all MCSs defined in Section 2.2.2 satisfying:*

$$\begin{aligned} &A_k \in \mathbb{C}^1(\mathbb{R}) \cap L^\infty(\mathbb{R}), \phi_k \in \mathbb{C}^2(\mathbb{R}), \\ &\sup_{t \in \mathbb{R}} \phi'_k(t) < \infty, \phi'_k(t) > 0, A_k(t) > 0, \quad \forall t \\ &|A'_k(t)| \leq \epsilon, |\phi''_k(t)| \leq \epsilon \quad \forall t \in \mathbb{R}. \end{aligned} \tag{2.63}$$

Further, the f_ks are separated with resolution Δ, i.e., for all $k \in \{1, ..., K-1\}$ and all t,

$$\phi'_{k+1}(t) - \phi'_k(t) \geq 2\Delta. \tag{2.64}$$

Definition 2.41. *Let h be a L^1-normalized positive function belonging to $C_c^\infty(\mathbb{R})$, and pick $\gamma, \lambda > 0$. The STFT-based synchrosqueezing (FSST) of f with threshold γ and accuracy λ is defined by:*

$$T_f^{g,\lambda,\gamma}(t,\omega) = \int_{|V_f^g(t,\eta)|>\gamma} V_f^g(t,\eta) \frac{1}{\lambda} h\left(\frac{\omega - \widehat{\omega}_f(t,\eta)}{\lambda}\right) d\eta. \tag{2.65}$$

If $\lambda \to 0$, then $T_f^{g,\lambda,\gamma}(t,\omega)$ tends, in the sense of distribution, to some value which we formally write as follows.

$$T_f^{g,\gamma}(t,\omega) = \int_{|V_f^g(t,\eta)|>\gamma} V_f^g(t,\eta)\delta(\omega - \widehat{\omega}_f(t,\eta))d\eta. \tag{2.66}$$

called FSST in the sequel.

Theorem 2.42. *Consider $f \in \mathcal{B}_{\Delta,\epsilon}$ and put $\tilde{\epsilon} = \epsilon^{1/3}$. Let $g \in \mathcal{S}(\mathbb{R})$, be such that supp$(\widehat{g}) \subset [-\Delta, \Delta]$. Then, if ϵ is small enough, the following hold:*

(a) $|V_f^g(t,\eta)| > \tilde{\epsilon}$ only when there exists $k \in \{1, ..., K\}$ such that $(t,\eta) \in Z_k := \{(t,\eta),\ s.t.\ |\eta - \phi'_k(t)| < \Delta\}$.

(b) For all $k \in \{1, ..., K\}$ and all $(t,\eta) \in Z_k$ such that $|V_f^g(t,\eta)| > \tilde{\epsilon}$, ones has

$$|\widehat{\omega}_f(t,\eta) - \phi'_k(t)| \leq \tilde{\epsilon}. \tag{2.67}$$

(c) For all $k \in \{1, ..., K\}$, there exists a constant $D_{F,1}$ such that for all $t \in \mathbb{R}$,

$$\left| \lim_{\lambda \to 0} \left(\frac{1}{g(0)} \int_{|\omega - \phi'_k(t)|<\tilde{\epsilon}} T_f^{g,\lambda,\tilde{\epsilon}}(t,\omega) d\omega \right) - f_k(t) \right| \leq D_{F,1}\tilde{\epsilon}. \tag{2.68}$$

This theorem is very similar to that in the wavelet case, but some differences appear mainly on the modulation of the mode and the separation condition (see Definition 2.40). It shows that the synchrosqueezing operator $T_f^{g,\gamma}$ is concentrated in narrow bands around curves $(t, \phi'_k(t))$ in the TF plane and the modes f_ks can be reconstructed from $T_f^{g,\gamma}$ with a reasonably high accuracy. The proof of this theorem is presented in [38]. Note that similar to the wavelet case, an extension of this theorem to a non-compactly supported window in the Fourier domain was also introduced in [49].

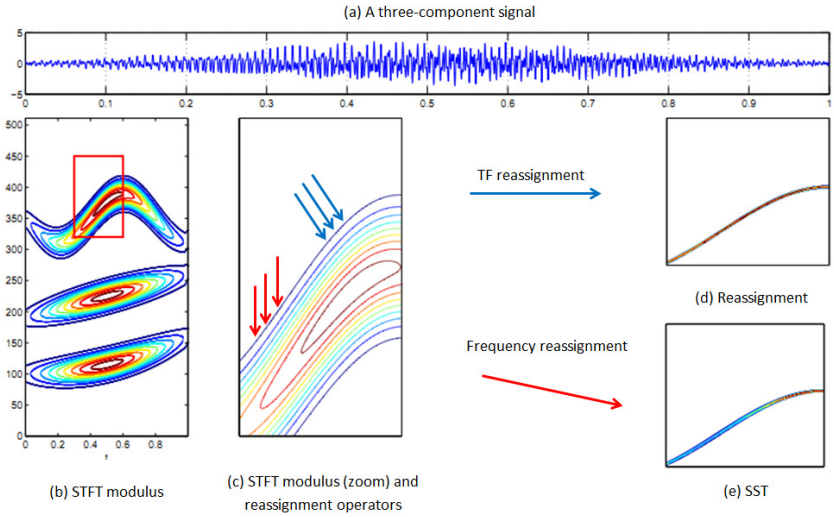


Fig. 2.11 FSST and RM of an MCS (using Gaussian window of size $\sigma_F = 0.03$): (a): time representation; (b): STFT modulus; (c): a small patch extracted from (b); (d): RM; (e): FSST.

B. Numerical Illustrations of FSST

This section provides several numerical examples to illustrate the principle of FSST. In a nutshell, the FSST of a signal is generally conducted in three main steps: (1) compute $T_f^{g,\gamma}$ using formula (2.66); (2) detect and extract the ridges corresponding to modes from $T_f^{g,\gamma}$ using formula (2.56) (just replacing $T_f^{g,\gamma}$ by $S_f^{\psi,\gamma}$ in that formula); (3) retrieve mode f_k using detected ridges φ_k and $T_f^{g,\gamma}$ through:

$$f_k(t) \approx \frac{1}{g(0)} \int_{\{\omega, |\omega - \varphi_k(t)| < d\}} T_f^{g,\gamma}(t,\omega) d\omega, \tag{2.69}$$

where $\varphi_k(t)$ is the estimate of $\phi_k'(t)$ given by the ridge detector, and d is the compensation parameter.

Let us first consider a complex simulated MCS (f) composed of three different kinds of AM-FM modes: a sinusoidal chirp f_1, a linear FM chirp f_2 and a second-order FM chirp f_3 all with the same Gaussian modulated amplitude. In our simulations, f is uniformly sampled over interval $[0, 1]$ with a sampling rate M $= 1024$ Hz. In addition, STFT is computed with L^1-normalized Gaussian window $g(t, \sigma_F) = \sigma_F^{-1} e^{-\pi \frac{t^2}{\sigma_F^2}}$ with $\sigma_F = 0.03$. In Figure 2.11, the principle of FSST and RM carried out on this MCS are clearly shown: FSST moves the STFT coefficients only along the frequency axis, while those reassigned by RM along both time and frequency. In comparison with RM, FSST gives a less concentrated result wherever the frequency modulation $\phi_k''(t)$ is large, and otherwise behaves similarly. Moreover, we illustrate, in Figure 2.12, the procedure of the mode retrieval with FSST on the same MCS as above.

These results show that all the three modes are separated and reconstructed with a relatively high accuracy.

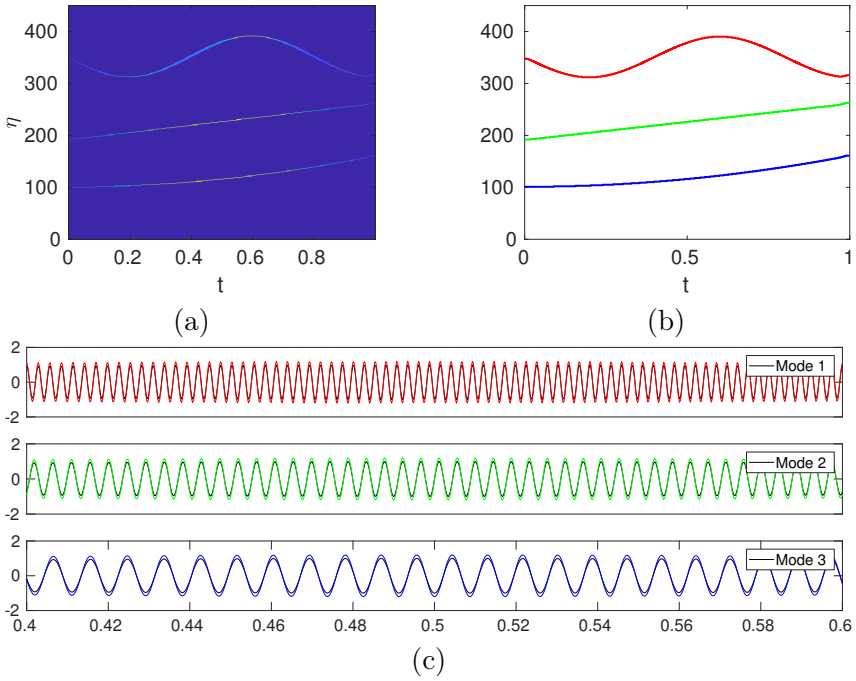


Fig. 2.12 Modes' retrieval using FSST: (a): FSST modulus; (b): ridge extraction; (c): three reconstructed modes (only a part of reconstructed modes is displayed together with the original ones (black lines)).

Furthermore, we illustrate the FSST behavior on a noisy MCS by adding the above signal with a white Gaussian noise $f_\zeta(t) = f(t) + \zeta(t)$, where $\zeta(t)$ is a complex white Gaussian process with variance $\mathrm{Var}\left(\Re\mathrm{e}\left\{\zeta(t)\right\}\right) = \mathrm{Var}\left(\Im\mathrm{m}\left\{\zeta(t)\right\}\right) = \sigma_\zeta^2$. The noise level is measured by the *Signal-to-Noise Ratio* (SNR):

$$\mathrm{SNR}_{\mathrm{intput}}[\mathrm{dB}] = 20\log_{10}\frac{\|f\|_2}{\|f_\zeta - f\|_2}, \tag{2.70}$$

and $\zeta(t)$ is the white Gaussian noise added, and $\|.\|_2$ is the l_2 norm. In Figure 2.13, we display the procedure of FSST on a noisy MCS with input SNR 10dB. It is clear that FSST produces a quite focused TFR from the smeared STFT. Moreover, the reconstruction of the modes is performed by selecting the FSST coefficients only on the detected ridges (i.e. $d = 0$). The reconstruction quality of the techniques is measured by the output SNR defined as:

$$\mathrm{SNR}_{\mathrm{output}}[\mathrm{dB}] = 20\log_{10}\left(\frac{\|f_i\|_2}{\|f_i - f_{r,i}\|_2}\right), \tag{2.71}$$

$f_{r,i}$ is the ith reconstructed mode. In our simulations, the output SNR of each mode is about 15dB, meaning that FSST not only estimates the modes but also performs some kind of

denoising. Other properties and comparisons of FSST with other methods are introduced in [40, 74, 49].

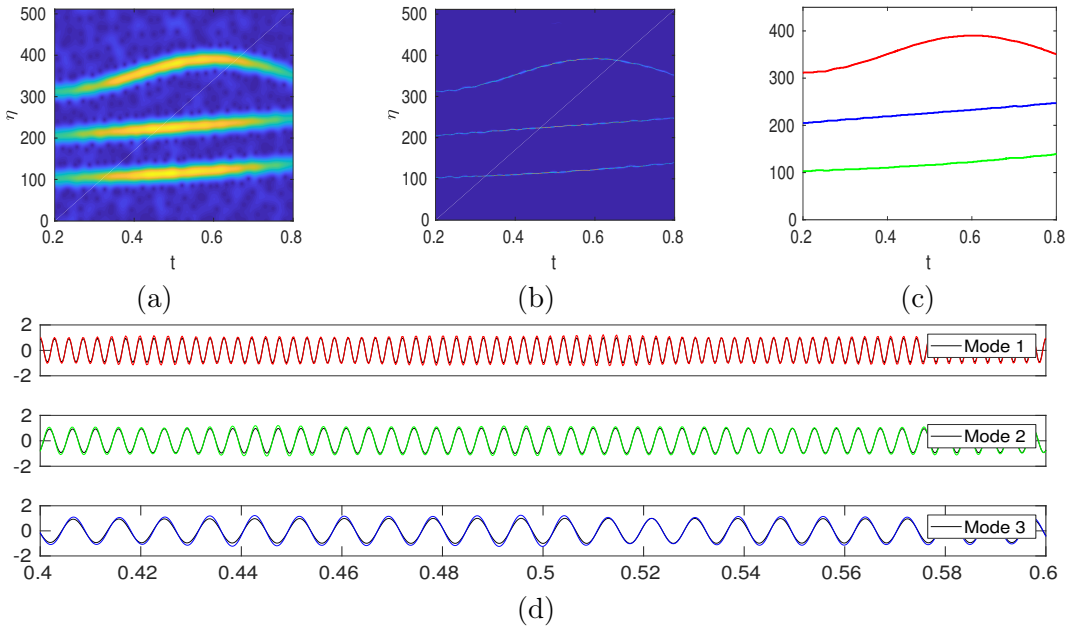


Fig. 2.13 FSST procedure of a noisy MCS: (a): STFT; (b): FSST; (b): ridge extraction; (d): three reconstructed modes (only a part of reconstructed modes is displayed along with the original noisy ones (black lines)).

2.6.3 Second Order STFT-based Synchrosqueezing Transform (FSST2)

A. FSST2 Principle

Although FSST proves to be an interesting solution for enhancing TFR while allowing the mode retrieval, it cannot deal with the modes having significant FMs ϕ'' which are frequently met in real-world applications [53, 46, 20, 45, 21]. In this section, we recall *second-order FSST* (FSST2), known as an adaptation of FSST to the strongly modulated modes [66, 47]. Let us first define the *second-order local modulation operator* $\widetilde{q}_{t,f}(t,\eta)$ allowing for a more accurate IF estimate, which is then used to define an improved synchrosqueezing operator $T^{g,\gamma}_{2,f}$.

Proposition 2.43. *Given a signal $f \in L^2(\mathbb{R})$, the complex reassignment operators $\widetilde{\omega}_f(t,\eta)$ and $\widetilde{\tau}_f(t,\eta)$ are respectively defined for any (t,η) s.t. $V^g_f(t,\eta) \neq 0$ as:*

$$\begin{aligned}\widetilde{\omega}_f(t,\eta) &= \frac{\partial_t V^g_f(t,\eta)}{2i\pi V^g_f(t,\eta)},\\ \widetilde{\tau}_f(t,\eta) &= t - \frac{\partial_\eta V^g_f(t,\eta)}{2i\pi V^g_f(t,\eta)}.\end{aligned} \tag{2.72}$$

Then, the second-order local complex modulation operator $\widetilde{q}_{t,f}(t,\eta)$ *is defined by:*

$$\widetilde{q}_{t,f}(t,\eta) = \frac{\partial_t \widetilde{\omega}_f(t,\eta)}{\partial_t \widetilde{\tau}_f(t,\eta)} \qquad \text{whenever } \partial_t \widetilde{\tau}_f(t,\eta) \neq 0. \tag{2.73}$$

In what follows, the subscript t (*resp.* η) denotes the obtained function using the derivatives with respect to the variabl t (*resp.* η).

In that case, the definition of the improved IF estimate associated with the TF representation given by STFT is derived as:

Definition 2.44. Let $f \in L^2(\mathbb{R})$, the second-order local complex IF estimate of f is defined as:

$$\widetilde{\omega}^{[2]}_{t,f}(t,\eta) = \begin{cases} \widetilde{\omega}_f(t,\eta) + \widetilde{q}_{t,f}(t,\eta)(t - \widetilde{\tau}_f(t,\eta)) & \text{if } \partial_t \widetilde{\tau}_f \neq 0 \\ \widetilde{\omega}_f(t,\eta) & \text{otherwise.} \end{cases}$$

Then, its real part $\widehat{\omega}^{[2]}_{t,f}(t,\eta) = \Re\mathrm{e}\{\widetilde{\omega}^{[2]}_{t,f}(t,\eta)\}$ is the desired IF estimate.

It was demonstrated in [48] that $\Re\mathrm{e}\{\widetilde{q}_{t,f}(t,\eta)\} = \phi''(t)$ when f is a Gaussian modulated linear chirp, i.e. $f(t) = A(t)e^{i2\pi\phi(t)}$ where both $\log(A(t))$ and $\phi(t)$ are quadratic. Also, $\widehat{\omega}^{[2]}_{t,f}(t,\eta)$ is an exact estimate of $\phi'(t)$ for this kind of signals. Furthermore, $\widetilde{\omega}_f(t,\eta)$, $\widetilde{\tau}_f(t,\eta)$ and $\widetilde{q}_{t,f}(t,\eta)$ can be computed by means of only five STFTs as follows:

Proposition 2.45. *For a signal* $f \in L^2(\mathbb{R})$*, the expressions* $\widetilde{\omega}_f$*,* $\widetilde{\tau}_f$ *and* $\widetilde{q}_{t,f}$ *can be written as:*

$$\widetilde{\omega}_f = \eta - \frac{1}{i2\pi}\frac{V_f^{g'}}{V_f^g} \tag{2.74}$$

$$\widetilde{\tau}_f = t + \frac{V_f^{tg}}{V_f^g} \tag{2.75}$$

$$\widetilde{q}_{t,f} = \frac{1}{i2\pi}\frac{V_f^{g''}V_f^g - \left(V_f^{g'}\right)^2}{V_f^{tg}V_f^{g'} - V_f^{tg'}V_f^g}, \tag{2.76}$$

where V_f^g *denotes* $V_f^g(t,\eta)$ *for the sake of simplicity and* $V_f^{g'}, V_f^{tg}, V_f^{g''}, V_f^{tg'}$ *are respectively STFTs of* f *computed with windows* $t \mapsto g'(t), tg(t), g''(t)$ *and* $tg'(t)$*.*

The second-order FSST (FSST2) is then defined by replacing $\widehat{\omega}_f(t,\eta)$ by $\widehat{\omega}^{[2]}_{t,f}(t,\eta)$ in (2.66):

$$T^{g,\gamma}_{2,f}(t,\omega) = \int_{\{\eta,|V_f^g(t,\eta)|>\gamma\}} V_f^g(t,\eta)\delta\left(\omega - \widehat{\omega}^{[2]}_{t,f}(t,\eta)\right)d\eta. \tag{2.77}$$

Mode f_k is finally retrieved by replacing $T^{g,\gamma}_f(t,\omega)$ by $T^{g,\gamma}_{2,f}(t,\omega)$ in (2.69) with a different normalization as follows.

$$f_k(t) \approx \frac{1}{D'_{g,k}} \int_{\{\omega,|\omega-\varphi_k(t)|<d\}} T^{g,\gamma}_{2,f}(t,\omega)d\omega, \tag{2.78}$$

where $D'_{g,k} = \int_{\mathbb{R}} \overline{\mathcal{F}\{g(\tau)e^{-i\pi\phi''_k(t)\tau^2}\}(\eta)}d\eta$ is a normalization constant that depends on mode k and is slightly different from $\overline{g(0)}$ used in the case of non-negligible FM. In practice, $\phi''_k(t)$ should be estimated by $\widetilde{q}_{t,f}(t,\eta)$ before computing this normalization constant. To implement the FSST2 in practice, the same procedure as that introduced in Section 2.6.2 FSST are applied using $T^{g,\gamma}_{2,f}(t,\eta)$ instead of $T^{g,\gamma}_{f}(t,\eta)$. Finally, numerical simulations to illustrate the efficiency of FSST2 and its comparison with other methods are well presented in [66, 48, 49, 86] and also provided in the following chapters.

B. Mathematical Foundations of FSST2

Now, we recall the mathematical framework of FSST2 in the non-compact support case, which was introduced in [49], by first defining a class of chirp-like functions $\mathcal{B}^{[2]}_{\Delta,\epsilon}$, which is larger than $\mathcal{B}_{\Delta,\epsilon}$ use in FSST case.

Definition 2.46. *Let $\epsilon > 0, \Delta > 0$. The set $\mathcal{B}^{[2]}_{\Delta,\epsilon}$ of MCSs with second-order modulation ϵ and separation Δ corresponds to the signals defined in Section 2.2.2 satisfying:*

(a) f_k is such that A_k and ϕ_k satisfy the following conditions:

$$A_k(t) \in L^\infty(\mathbb{R}) \cap C^2(\mathbb{R}),\ \phi_k(t) \in C^3(\mathbb{R}), \phi'_k(t),\ \phi''_k(t),\ \phi'''_k(t) \in L^\infty(\mathbb{R}),$$
$$A_k(t) > 0,\ \inf_{t\in\mathbb{R}} \phi'_k(t) > 0,\ \sup_{t\in\mathbb{R}} \phi'_k(t) < \infty, |A'_k(t)| \leq \epsilon,\ |A''_k(t)| \leq \epsilon,\ \text{and } |\phi'''_k(t)| \leq \epsilon.$$

(b) the f_ks are separated with resolution Δ, i.e., for all $k \in \{1, ..., K-1\}$ and all t,

$$\phi'_{k+1}(t) - \phi'_k(t) \geq 2\Delta. \tag{2.79}$$

Now, let us define the second-order FSST as follows.

Definition 2.47. *Let h be a L^1-normalized positive function belonging to $C^\infty_c(\mathbb{R})$, and pick $\gamma, \lambda > 0$. The second-order FSST (FSST2) of f with threshold γ and accuracy λ is defined by:*

$$T^{g,\lambda,\gamma}_{2,f}(t,\omega) = \int_{|V^g_f(t,\eta)|>\gamma} V^g_f(t,\eta)\frac{1}{\lambda}h\left(\frac{\omega - \widehat{\omega}^{[2]}_f(t,\eta)}{\lambda}\right)d\eta. \tag{2.80}$$

If $\lambda \to 0$, then $T^{g,\lambda,\gamma}_{2,f}(t,\omega)$ tends, in the sense of distribution, to some value which we formally write as in (2.77).

The approximation theorem of FSST2, in which we prove that $\hat{\omega}^{[2]}_f(t,\eta)$ is an accurate IF estimate for a function $f \in \mathcal{B}^{[2]}_{\Delta,\epsilon}$, is now recalled:

Theorem 2.48. *Consider* $f \in \mathcal{B}^{[2]}_{\Delta,\epsilon}$, *and put* $\tilde{\epsilon} = \epsilon^{1/6}$. *Let* g *be a window satisfying, all* $k = 1, \cdots, K$ *and* $r \in \{0, 1, 2\}$, $\left|\mathcal{F}\{\tau^r g(\tau) e^{-i\pi\phi_k''(t)\tau^2}\}(\eta)\right| \leq K_r \epsilon$ *when* $|\eta| \geq \Delta$ *and,* $\int_{|\eta|>\Delta} \left|\mathcal{F}\{g(\tau) e^{-i\pi\phi_k''(t)\tau^2}\}(\eta)\right| d\eta \leq K_3 \tilde{\epsilon}$, *for some constants* K_r *and* K_3. *Then, provided* ϵ *is sufficiently small, the following hold:*

(a) $|V_f^g(t,\eta)| > \tilde{\epsilon}$ *only when there exists* $k \in \{1, ..., K\}$ *such that* $(t,\eta) \in Z_k := \{(t,\eta),\ s.t.\ |\eta - \phi_k'(t)| < \Delta\}$.

(b) For all $k \in \{1, ..., K\}$ *and all* $(t,\eta) \in Z_k$ *such that* $|V_f^g(t,\eta)| > \tilde{\epsilon}$ *and* $\left|\partial_t \tilde{t}_f(t,\eta)\right| > \tilde{\epsilon}$, *one has:*

$$\left|\widehat{\omega}^{[2]}_{t,f}(t,\eta) - \phi_k'(t)\right| \leq \tilde{\epsilon}. \tag{2.81}$$

(c) For all $k \in \{1, ..., K\}$, *there exists a constant* $D_{F,2}$ *such that for all* $t \in \mathbb{R}$,

$$\left|\lim_{\lambda\to 0}\left(\frac{1}{D'_{g,k}}\int_{|\omega-\phi_k'(t)|<\tilde{\epsilon}} T^{g,\lambda,\tilde{\epsilon}}_{2,f}(t,\omega)d\omega\right) - f_k(t)\right| \leq D_{F,2}\tilde{\epsilon}. \tag{2.82}$$

with $D'_{g,k}$ *being defined in (2.78).*

The proof of this theorem is available in [49].

2.7 Conclusions

In this chapter, we have conducted a succinct review on TFRs for the analysis of MCSs. We first introduced the motivation why a TFR is preferred to a time or frequency representation alone in many applications whose signals have time-varying nature or multiple components. Then, we recalled some basic notation, definitions, and several principal linear and quadratic TF methods: short time Fourier transform (STFT), wavelet transform (WT), spectrogram, Wigner Ville distribution and scalogram. On that matter, we also pointed out the main limitation associated with a TFR known as the uncertainty principle preventing any accurate temporal localization of a frequency. To tackle such a limitation, RM aiming at sharpening TFR was then developed, but it did not allow for signal reconstruction. In this regard, we recalled its variant, known as synchrosqueezing transform which has the additional advantage of allowing the mode retrieval. The latter was also developed in both wavelet and STFT frameworks. Unfortunately, it was shown that these techniques cannot deal with signals containing modes with strong frequency modulation. In order to handle this case, we presented the second-order synchrosqueezing transform (FSST2) in the STFT context, an adaptation of FSST to the strong frequency modulated context, in which the local instantaneous frequency estimate ($\widehat{\omega}_f(t,\eta)$) is replaced by a more accurate one. However, it only perfectly adapts to the signals that can be locally well approximated by linear chirps, which is still limited when

processing other general types of signals. In this regard, we put forward, in Chapter 3.2, a generalization of FSST and FSST2 to the so-called *high-order STFT-based synchrosqueezing transforms* (FSSTn) based on a more accurate instantaneous frequency estimate that enables to circumvent this limitation. This work was published in [5]. Before going into the details of this method, we introduce, in Chapter 3.1, the counterpart of FSST2 in the wavelet framework, known as the *second-order wavelet-based synchrosqueezing transform* (WSST2) accompanied by its rigorous mathematical analysis, all of which are available in [3].

Chapter 3

Contributions to Synchrosqueezing Transforms for Multicomponent Signals Analysis

Contents

In Chapter 2, we introduced three effective synchrosqueezing transforms (WSST, FSST, FSST2) for analyzing multicomponent signals (MCSs), along with their mathematical foundations. However, these techniques were proven to only provide an ideal invertible TFR for either pure harmonic or linear chirps with Gaussian modulated amplitudes, which is still restrictive. Indeed, most real-world signals made of very strongly modulated AM-FM modes, as for instance chirps involved in radar [45], speech processing [46], or gravitational waves [20, 21], need more accurate and powerful techniques. With this in mind, we propose to improve existing STFT-based SSTs by computing more accurate estimates of the instantaneous frequencies (IFs) of the modes making up the signal, using higher-order approximations both for the amplitude and phase. This results in perfect concentration and reconstruction for a wider variety of AM-FM modes than what was possible up to now with synchrosqueezing techniques [5]. Before characterizing this new approach, we put forward the counterpart of FSST2 in the wavelet framework a build-up on a second-order extension of WSST: the so-called *second-order wavelet-based synchrosqueezing transform* (WSST2), for which we develop a mathematical analysis and discuss its practical implementation [3].

This chapter is structured as follows: we first revisit WSST assuming the analysis wavelet is not compactly supported in the Fourier domain in Section 3.1.1. This helps derive, in Section 3.1.2, the second-order wavelet-based synchrosqueezing transform (WSST2) and subsequently

prove approximations results generalizing those related to WSST. In Section 3.1.3, we detail the practical implementation of WSST2 and perform several numerical experiments on simulated signals to compare the latter with other synchrosqueezing techniques in Section 3.1.4. Then, we introduce, in Section 3.2.1, a different definition of the second-order FSST (FSST2) compared with the one in Section 2.6.3. In that framework, we present the proposed generalization: the so-called *higher-order STFT-based synchrosqueezing transform* (FSSTn), in Section 3.2.2. Then, numerical simulations are performed on synthetic signals to demonstrate the efficiency of FSSTn compared with the reassignment method (RM) and other synchrosqueezing techniques, in Section 3.2.3. Finally, numerical experiments conducted on gravitational-wave signals illustrate the interest of WSST2 and FSSTn in comparison with the other synchrosqueezing techniques.

3.1 Second-order Wavelet-based Synchrosqueezing Transform

3.1.1 Wavelet-based Synchrosqueezing Transform with Non-compactly Supported Wavelet

The wavelet-based synchrosqueezing transform (WSST) using a compactly supported wavelet in the Fourier domain is supported by a solid mathematical framework, which was recalled in Section 2.6.1. Now, we define the class of chirp-like signals on which we build the theory for the non-compact support case.

Definition 3.1. *Let $\varepsilon > 0$ and $c > 0$. The set $\mathcal{A}_{c,\varepsilon}$ of MCSs with modulation ε and separation c corresponds to signals defined in Section 2.2.2 with f_k satisfying:*

$$\begin{gathered} A_k \in C^1(\mathbb{R}) \cap L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}), \phi_k \in C^2(\mathbb{R}), \\ \inf_{t\in\mathbb{R}} \phi_k'(t) > 0, \sup_{t\in\mathbb{R}} \phi_k'(t) < \infty, M = \max_k \left(\sup_{t\in\mathbb{R}} \phi_k'(t)\right), \\ A_k(t) > 0, |A_k'(t)| \leq \varepsilon\phi_k'(t) \leq \varepsilon M, |\phi_k''(t)| \leq \varepsilon\phi_k'(t) \leq \varepsilon M, \quad \forall t \in \mathbb{R}. \end{gathered}$$

Further, the f_ks are separated with resolution c, i.e., for all $k \in \{1, \ldots, K-1\}$ and all t

$$\phi_{k+1}'(t) - \phi_k'(t) \geq c(\phi_{k+1}'(t) + \phi_k'(t)). \tag{3.1}$$

Remark 3.2. Note that the conditions defining this class are slightly different from those introduced in Definition 2.37 to enables a theoretical analysis associated with the non-compactly supported wavelet. In what follows Δ denotes a real number in $]0, \frac{c}{c+1}[$.

Definition 3.3. *Let h be a positive L^1-normed window belonging to $C_c^\infty(\mathbb{R})$, and consider $\gamma, \lambda > 0$. The wavelet-based synchrosqueezing transform of f (WSST) with threshold γ and*

accuracy λ *is defined by:*

$$S_f^{\psi,\lambda,\gamma}(t,\omega) := \int_{|W_f^\psi(t,a)|>\gamma} W_f^\psi(t,a)\frac{1}{\lambda}h\left(\frac{\omega-\widehat{\omega}_f(t,a)}{\lambda}\right)\frac{da}{a}. \tag{3.2}$$

If $\lambda \to 0$*, then* $S_f^{\psi,\lambda,\gamma}(t,\omega)$ *tends, in the sense of distribution, to some value which we formally write as in Equation (2.54).*

Theorem 3.4. *Consider* $f \in \mathcal{A}_{c,\varepsilon}$*, set* $\tilde{\varepsilon} = \varepsilon^{\frac{1}{3}}$ *and let* ψ *be a non-compactly supported wavelet in the Fourier domain satisfying:* $\left|\hat{\psi}(\eta)\right| \leq N_0\varepsilon$ *when* $|\eta-1|>\Delta$*, and* $\int_{|\eta-1|>\Delta}\left|\hat{\psi}(\eta)\right|\frac{d\eta}{\eta} \leq N_1\tilde{\varepsilon}$*, for some constants* N_0 *and* N_1*.*

Assuming $(t,a) \in \mathbb{E} = \mathbb{R}\times\left[0, \frac{1+\Delta}{\inf_{t\in\mathbb{R}}\phi_1'(t)}\right]$*, then, provided* ε *is sufficiently small, the following hold:*

(a) $|W_f^\psi(t,a)| > \tilde{\varepsilon}$ *only when, there exists* $k \in \{1,\dots,K\}$*, such that* $(t,a) \in Z_k := \{(t,a),\ s.t.\ |a\phi_k'(t)-1|<\Delta\}$*.*

(b) For each $k \in \{1,\dots,K\}$ *and all* $(t,a) \in Z_k$ *for which holds* $|W_f^\psi(t,a)| > \tilde{\varepsilon}$*, one has:*

$$|\widehat{\omega}_f(t,a) - \phi_k'(t)| \leq \tilde{\varepsilon}. \tag{3.3}$$

(c) Moreover, for each $k \in \{1,\dots,K\}$*, there exists a constant* D_1 *such that for any* $t \in \mathbb{R}$

$$\left|\lim_{\lambda\to 0}\left(\frac{1}{C_\psi'}\int_{|\omega-\phi_k'(t)|<\tilde{\varepsilon}} S_f^{\psi,\lambda,\tilde{\varepsilon}}(t,\omega)d\omega\right) - f_k(t)\right| \leq D_1\tilde{\varepsilon}, \tag{3.4}$$

where C_ψ' *was defined in Proposition 2.23.*

The proof of this theorem is available in Appendix A.1.

3.1.2 Second-order Wavelet-based Synchrosqueezing Transform (WSST2)

As highlighted above, the inherent limitation associated with WSST is that it is only efficient to enhance TFRs of MCS composed of slightly perturbed harmonic modes. In order to overcome this drawback, a recent extension of WSST was introduced based on a more accurate IF estimate that is then used to define an improved synchrosqueezing operator, called the *second-order wavelet-based synchrosqueezing transform* (WSST2) [87], and our goal is to develop its mathematical study.

A. WSST2 Principle

Let us first define a *second-order local modulation operator* subsequently used to compute the new IF estimate. This modulation operator corresponds to the ratio of the first-order derivatives, with respect to t, of the reassignment operators, as explained in the following:

Proposition 3.5. *Given a signal $f \in L^2(\mathbb{R})$, the complex reassignment operators $\widetilde{\omega}_f(t,a)$ and $\widetilde{\tau}_f(t,a)$ are respectively defined for any (t,a) s.t. $W_f^\psi(t,a) \neq 0$ as:*

$$\widetilde{\omega}_f(t,a) = \frac{1}{i2\pi}\frac{\partial_t W_f^\psi(t,a)}{W_f^\psi(t,a)} \tag{3.5}$$

$$\widetilde{\tau}_f(t,a) = \frac{\int_{\mathbb{R}} \tau f(\tau)\frac{1}{a}\overline{\psi(\frac{\tau-t}{a})}d\tau}{W_f^\psi(t,a)} = t + a\frac{W_f^{t\psi}(t,a)}{W_f^\psi(t,a)}, \tag{3.6}$$

which are now defined provided $t\psi$ and ψ' are in $L^1(\mathbb{R})$. Then, the second-order local complex modulation operator $\widetilde{q}_{t,f}(t,a)$ is defined by:

$$\widetilde{q}_{t,f}(t,a) = \frac{\partial_t\widetilde{\omega}_f(t,a)}{\partial_t\widetilde{\tau}_f(t,a)}, \qquad \text{whenever } \partial_t\widetilde{\tau}_f(t,a) \neq 0. \tag{3.7}$$

Remark 3.6. It is interesting to note that using partial derivatives with respect to a instead of t, one can obtain a new second-order local modulation operator $\widetilde{q}_{a,f}(t,a) = \dfrac{\partial_a\widetilde{\omega}_f(t,a)}{\partial_a\widetilde{\tau}_f(t,a)}$, whose properties are exactly the same as those of $\widetilde{q}_{t,f}(t,a)$.

In this regard, the definition of the improved IF estimate associated with the TFR given by CWT is derived as:

Definition 3.7. *Let $f \in L^2(\mathbb{R})$, the second-order local complex IF estimate of f is defined as:*

$$\widetilde{\omega}^{[2]}_{t,f}(t,a) = \begin{cases} \widetilde{\omega}_f(t,a) + \widetilde{q}_{t,f}(t,a)(t-\widetilde{\tau}_f(t,a)) & \text{if } \partial_t\widetilde{\tau}_f(t,a) \neq 0 \\ \widetilde{\omega}_f(t,a) & \text{otherwise.} \end{cases} \tag{3.8}$$

Then, its real part $\widehat{\omega}^{[2]}_{t,f}(t,a) = \Re\mathrm{e}\{\widetilde{\omega}^{[2]}_{t,f}(t,a)\}$ is the desired IF estimate.

Similar to the FSST2 case, $\Re\mathrm{e}\{\widetilde{q}_{t,f}(t,a)\} = \phi''(t)$ when f is a Gaussian modulated linear chirp. Further, $\widehat{\omega}^{[2]}_{t,f}(t,a)$ is an exact estimate of $\phi'(t)$ for this kind of signals. Furthermore, $\widetilde{\omega}_f(t,a)$ and $\widetilde{q}_{t,f}(t,a)$ can be computed by means of five CWTs as follows:

Proposition 3.8. *Let $f \in L^2(\mathbb{R})$, $\widetilde{\omega}_f(t,a)$ and $\widetilde{q}_{t,f}(t,a)$ can be written as:*

$$\widetilde{\omega}_f(t,a) = -\frac{1}{i2\pi a}\frac{W_f^{\psi'}(t,a)}{W_f^\psi(t,a)} \tag{3.9}$$

$$\widetilde{q}_{t,f}(t,a) = \frac{1}{i2\pi a^2} \frac{W_f^{\psi''}(t,a)W_f^{\psi}(t,a) - W_f^{\psi'}(t,a)^2}{W_f^{t\psi}(t,a)W_f^{\psi'}(t,a) - W_f^{t\psi'}(t,a)W_f^{\psi}(t,a)}, \tag{3.10}$$

where $t \mapsto W^{\psi'}, W^{t\psi}, W^{\psi''}, W^{t\psi'}$ *are respectively CWTs of* f *computed with wavelets* $\psi', t\psi, \psi'', t\psi'$ *all in* $L^1(\mathbb{R})$.

Proof. These expressions are easily derived using $\partial_t^p W_f^{\psi}(t,a) = \left(-\frac{1}{a}\right)^p W_f^{\psi^{(p)}}(t,a)$. □

The second-order WSST (WSST2) is then defined by replacing $\widehat{\omega}_f(t,a)$ by $\widehat{\omega}_{t,f}^{[2]}(t,a)$ in (2.54):

$$S_{2,f}^{\psi,\gamma}(t,\omega) := \int_{|W_f^{\psi}(t,a)|>\gamma} W_f^{\psi}(t,a)\delta\left(\omega - \widehat{\omega}_{t,f}^{[2]}(t,a)\right)\frac{da}{a}, \tag{3.11}$$

and f_k is finally retrieved by replacing $S_f^{\psi,\gamma}(t,\omega)$ by $S_{2,f}^{\psi,\gamma}(t,\omega)$ in (2.55) with a different normalization as follows:

$$f_k(t) \approx \frac{1}{C'_{\psi,k}} \int_{\{\omega,|\omega-\varphi_k(t)|<d\}} S_f^{\psi,\gamma}(t,\omega)d\omega, \tag{3.12}$$

where $C'_{\psi,k} = \int_0^{\infty} \overline{\mathcal{F}\{\psi(\tau)e^{-i\pi\frac{\phi_k''(t)}{\phi_k'^2(t)}\eta^2\tau^2}\}(\eta)}\frac{d\eta}{\eta}$ depends on mode k and is definitely different from C'_{ψ} as soon as the modulation is non zero. The reason behind this normalization is explained in the proof of the following Theorem 3.11. Before detailing numerical implementation of WSST2, the following section introduces a study of WSST2 from the angle of mathematical analysis.

B. Mathematical Foundations for WSST2

This section begins with the definition of another class of chirp-like functions, larger than $\mathcal{A}_{c,\varepsilon}$ and that can be successfully dealt with WSST2:

Definition 3.9. *Let* $\varepsilon > 0$. *The set* $\mathcal{A}_{c,\varepsilon}^{[2]}$ *of multicomponent signals with second-order modulation* ε *and separation* c *corresponds to the signals defined in Section 2.2.2 satisfying:*

(a) f_k *is such that* A_k *and* ϕ_k *satisfy the following conditions:*

$$\begin{gathered}
A_k(t) \in C^2(\mathbb{R}) \cap L^1(\mathbb{R}) \cap L^{\infty}(\mathbb{R}),\ \phi_k(t) \in C^3(\mathbb{R}),\\
\phi_k'(t),\ \phi_k''(t),\ \phi_k'''(t) \in L^{\infty}(\mathbb{R}),\\
A_k(t) > 0,\ \inf_{t\in\mathbb{R}} \phi_k'(t) > 0,\ \sup_{t\in\mathbb{R}} \phi_k'(t) < \infty,\quad M = \max_k \left(\sup_{t\in\mathbb{R}} \phi_k'(t)\right),\\
|A_k'(t)| \le \varepsilon\phi_k'(t) \le \varepsilon M, |A_k''(t)| \le \varepsilon\phi_k'(t) \le \varepsilon M,\\
\text{and } |\phi_k'''(t)| \le \varepsilon\phi_k'(t) \le \varepsilon M \quad \forall t \in \mathbb{R}.
\end{gathered}$$

(b) the ϕ_ks satisfy the following separation condition

$$\phi'_{k+1}(t) - \phi'_k(t) \geq c(\phi'_{k+1}(t) + \phi'_k(t)), \forall t \in \mathbb{R}\ ,\ \forall k \in \{1, \dots, K-1\}.$$

Note that Δ is some value in $]0, \frac{c}{c+1}[$. Now, let us define WSST2 as follows:

Definition 3.10. *Let h be a positive L^1-normed window belonging to $C_c^\infty(\mathbb{R})$, and consider $\gamma, \lambda > 0$, the WSST2 of f with threshold γ and accuracy λ is defined by:*

$$S_{2,f}^{\lambda,\gamma}(t,\omega) = \int_{|W_f^\psi(t,a)|>\gamma} W_f^\psi(t,a) \frac{1}{\lambda} h\left(\frac{\omega - \widehat{\omega}_{t,f}^{[2]}(t,a)}{\lambda}\right) \frac{da}{a}. \tag{3.13}$$

In Section 3.1.1, we showed that, for functions $f \in \mathcal{A}_{c,\varepsilon}$, a good IF estimate was given by $\widehat{\omega}_f(t,a)$ and the approximation theorem followed. Here, to assess the approximation property of WSST2 we have just introduced, we consider $f \in \mathcal{A}_{c,\varepsilon}^{[2]}$ for which we are going to prove that $\widehat{\omega}_{t,f}^{[2]}(t,a)$ is a good IF estimate. The approximation theorem is as follows:

Theorem 3.11. *Consider $f \in \mathcal{A}_{c,\tilde{\varepsilon}}^{[2]}$, set $\tilde{\varepsilon} = \varepsilon^{1/6}$. Let ψ be a wavelet satisfying, for all $k = 1, \cdots, K$, $r \in \{0,1\}$ and $p \in \{0,1\}$, $\left|\mathcal{F}\{\tau^r \psi^{(p)}(\tau) e^{-i\pi \frac{\phi_k''(t)}{\phi_k'^2(t)}\eta^2\tau^2}\}(\eta)\right| \leq N_{r,p}\varepsilon$ when $|\eta - 1| > \Delta$, and $\int_{|\eta-1|>\Delta} \left|\mathcal{F}\{\psi(\tau) e^{-i\pi \frac{\phi_k''(t)}{\phi_k'^2(t)}\eta^2\tau^2}\}(\eta)\right| \frac{d\eta}{\eta} \leq N_2\tilde{\varepsilon}$, for some constants $N_{r,p}$ and N_2.*

Assuming $(t,a) \in \mathbb{E}$ defined in Theorem 3.4, then, provided ε is sufficiently small, the following hold:

(a) $|W_f^\psi(t,a)| > \tilde{\varepsilon}$ on $\mathbb{E}$ only when, there exists $k \in \{1, \dots, K\}$, such that $(t,a) \in Z_k := \{(t,a),\ s.t.\ |a\phi'_k(t) - 1| < \Delta\}$.

(b) For each $k \in \{1, \dots, K\}$ and for all $(t,a) \in Z_k$, for which hold $|W_f^\psi(t,a)| > \tilde{\varepsilon}$ and $|\partial_t \widetilde{\tau}_f(t,a)| > \tilde{\varepsilon}$, one has

$$|\widehat{\omega}_{t,f}^{[2]}(t,a) - \phi'_k(t)| \leq \tilde{\varepsilon}. \tag{3.14}$$

(c) Moreover, for each $k \in \{1, \dots, K\}$, there exists a constant $D_{W,2}$ such that

$$\left|\left(\lim_{\lambda\to 0} \frac{1}{C'_{\psi,k}} \int_{|\omega - \phi'_k(t)|<\tilde{\varepsilon}} S_{2,f}^{\lambda,\tilde{\varepsilon}}(t,\omega) d\omega\right) - f_k(t)\right| \leq D_{W,2}\tilde{\varepsilon}, \tag{3.15}$$

where $C'_{\psi,k}$ is defined in (3.12).

The proof of Theorem 3.11 is available in Appendix A.2.

3.1.3 Numerical Implementation of WSST2

This section details the numerical implementation of WSST2. Let consider a signal f being assumed to be defined on $[0,1]$ and then uniformly discretized at time $t_m = m/n$ with $m = 0,\dots,n-1$ and $n = 2^L$, $L \in \mathbb{N}$. First, we discretize W_f^ψ at $(m/n, a_j)$, where $a_j = \frac{2^{j/n_v}}{n}$, $j = 0,\dots,Ln_v$ with the *voice number* n_v being a user-defined parameter controlling the number of scales ($n_v = 32$ or 64 in practice). The discrete wavelet transform (DWT) of f is computed in the Fourier domain as follows:

$$W_f^\psi(t_m, a_j) \approx W_{\mathbf{d},f}^{\hat{\psi}}(m,j) := \left(\mathcal{F}_{\mathbf{d}}^{-1}\left(\left(\mathcal{F}_{\mathbf{d}}(f) \odot \overline{\hat{\psi}_{j,.}}\right)\right)\right)_m, \tag{3.16}$$

where $\mathcal{F}_{\mathbf{d}}(f)$ (*resp.* $\mathcal{F}_{\mathbf{d}}^{-1}$) denotes the standard (*resp.* inverse) discrete Fourier transform (DFT) (*resp.* IDFT), $\odot$ the elementwise multiplication, $\hat{\psi}_{j,q} = \hat{\psi}(a_j q)$ with $q = 0,\dots,n-1$, and the subscript $\mathbf{d}$ stands for *discrete* value.

With this in mind, we compute the complex estimate of the second-order modulation operator $\widetilde{q}_{t,f}$ defined as in (3.10), as follows:

$$\widetilde{q}_{\mathbf{d},t,f}(m,j) = \frac{i2\pi\left(W_{\mathbf{d},f}^{\hat{\psi}}(m,j)W_{\mathbf{d},f}^{\xi^2\hat{\psi}}(m,j) - \left(W_{\mathbf{d},f}^{\xi\hat{\psi}}(m,j)\right)^2\right)}{a_j^2\left[\left(W_{\mathbf{d},f}^{\hat{\psi}}(m,j)\right)^2 + W_{\mathbf{d},f}^{\hat{\psi}}(m,j)W_{\mathbf{d},f}^{\xi\hat{\psi}'}(m,j) - W_{\mathbf{d},f}^{\hat{\psi}'}(m,j)W_{\mathbf{d},f}^{\xi\hat{\psi}}(m,j)\right]},$$

where $W_{\mathbf{d},f}^{\xi^2\hat{\psi}}$, $W_{\mathbf{d},f}^{\xi\hat{\psi}}$, $W_{\mathbf{d},f}^{\xi\hat{\psi}'}$, $W_{\mathbf{d},f}^{\hat{\psi}'}$ denote respectively DWTs of f computed using the wavelets $\xi \mapsto \xi^2\hat{\psi}$, $\xi \mapsto \xi\hat{\psi}$, $\xi \mapsto \xi\hat{\psi}'$, and $\xi \mapsto \hat{\psi}'$. For instance $(\xi\hat{\psi})_{j,q} = (a_j q)\hat{\psi}(a_j q)$.

Introducing $\widetilde{\omega}_{\mathbf{d},f}(m,j) = \dfrac{W_{\mathbf{d},f}^{\xi\hat{\psi}}(m,j)}{a_j W_{\mathbf{d},f}^{\hat{\psi}}(m,j)}$ and $\widetilde{\tau}_{\mathbf{d},f}(m,j) = t + \dfrac{a_j}{i2\pi}\dfrac{W_{\mathbf{d},f}^{\hat{\psi}'}(m,j)}{W_{\mathbf{d},f}^{\hat{\psi}}(m,j)}$, enables the definition of a discrete version of a second-order complex IF estimate of f:

$$\widetilde{\omega}_{\mathbf{d},f}^{[2]}(m,j) = \begin{cases} \widetilde{\omega}_{\mathbf{d},f}(m,j) + \widetilde{q}_{\mathbf{d},t,f}(m,j)(t - \widetilde{\tau}_{\mathbf{d},f}(m,j)) & \text{if } \partial_t\widetilde{\tau}_{\mathbf{d},f}(m,j) \neq 0 \\ \widetilde{\omega}_{\mathbf{d},f}(m,j) & \text{otherwise,} \end{cases}$$

where $\partial_t\widetilde{\tau}_{\mathbf{d},f}(m,j) = \dfrac{\left(W_{\mathbf{d},f}^{\hat{\psi}}(m,j)\right)^2 + W_{\mathbf{d},f}^{\hat{\psi}}(m,j)W_{\mathbf{d},f}^{\xi\hat{\psi}'}(m,j) - W_{\mathbf{d},f}^{\hat{\psi}'}(m,j)W_{\mathbf{d},f}^{\xi\hat{\psi}}(m,j)}{W_{\mathbf{d},f}^{\hat{\psi}}(m,j)^2}$. We then take the real part $\widehat{\omega}_{\mathbf{d},f}^{[2]}(m,j) = \Re\mathrm{e}\{\widetilde{\omega}_{\mathbf{d},f}^{[2]}(m,j)\}$, which leads to the desired discreted IF estimate.

We now show how to compute WSST2. First, we highlight how the frequency domain is split when performing second-order synchrosqueezing transform. First we remark that each scale a_j is the inverse of frequency $f_j = 1/a_j = 2^{-j/n_v}n$. Putting $f_{Ln_v} = 0$ and $f_{-1} = +\infty$, we define frequency bins corresponding to the wavelet representation as $\mathcal{W}_j = \left[\frac{f_{j+1}+f_j}{2}, \frac{f_j+f_{j-1}}{2}\right[$, where $0 \leq j \leq Ln_v - 1$. With this in mind, the second-order synchrosqueezing operator is

implemented as follows:

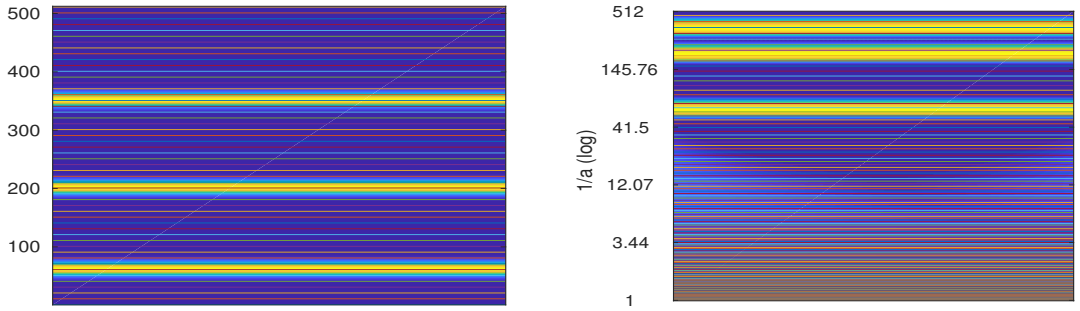
$$S^{\psi,\gamma}_{\mathbf{d},2,f}(m,f_j)=\sum_{\mathbb{G}_{\mathbf{d}}(j)} W^{\hat{\psi}}_{\mathbf{d},f}(m,l)\frac{\log(2)}{n_v}, \tag{3.17}$$

where $\mathbb{G}_{\mathbf{d}}(j)=\left\{0\leq l\leq Ln_v-1 \text{ s.t. } \widehat{\omega}^{[2]}_{\mathbf{d},f}(m,l)\in\mathcal{W}_j \text{ and } |W^{\hat{\psi}}_{\mathbf{d},f}(m,l)|>\gamma\right\}$.

Finally, each mode f_k is retrieved by summing $S^{\psi,\gamma}_{\mathbf{d},2,f}$ along the frequency axis in the vicinity of the kth mode. More precisely, one has, for each t_m,

$$f_k(m/n)\approx\frac{1}{C'_{\mathbf{d},\psi,k}}\sum_{l\in\Upsilon_k(m)} S^{\psi,\gamma}_{\mathbf{d},2,f}(m,\omega_l), \tag{3.18}$$

where $\Upsilon_k(m)$ is a set of indices corresponding to a small frequency band located around the ridge curve of kth mode, which is selected by ridge extraction method [83, 5]. Note also that $C'_{\mathbf{d},\psi,k}$ is a discrete approximation of $C'_{\psi,k}$ computed as in (3.12): $C'_{\psi,k}=\int_0^\infty \overline{\mathcal{F}\{\psi(\tau)e^{-i\pi\frac{\phi''_k(t)}{\phi'^2_k(t)}\eta^2\tau^2}\}(\eta)}\frac{d\eta}{\eta}$, in which $\phi''_k(t)$ and $\phi'_k(t)$ should be first estimated by $\widetilde{q}_{\mathbf{d},t,f}$ and $\widetilde{\omega}^{[2]}_{\mathbf{d},f}$, respectively.



(a) (b)

Fig. 3.1 (a), (b): modulus of STFT and WT of three constant frequency modes f_1, f_2, and f_3, with lines corresponding to the frequencies where the transforms are actually computed (linear scale for STFT and exponential one for WT).

Remark 3.12. It is important to remark here that the set $\Upsilon_k(m)$ is computed via ridge extraction on the wavelet representation so that the accuracy of the reconstruction of the mode depends on the frequency band the mode leaves in. Indeed the size of $\mathcal{W}_{j-1}$ is $2^{1/n_v}$ the size of $\mathcal{W}_j$, meaning the accuracy of the set $\Upsilon_k(m)$ depends on the frequency of mode k. On the contrary, this is not the case when considering a synchrosqueezing operator based on STFT, for which one uses a uniform sampling of the frequency axis. With STFT, the length of the frequency bins is 1, while the length of $\mathcal{W}_j$ is smaller than 1 for large j and much bigger for small ones. This is illustrated in Figure 3.1, where we consider three constant frequency modes: on that figure, we draw a line at each frequency used in the computation of CWT and STFT (along with the transforms themselves). We notice that while the error associated with

IF estimation by ridge extraction is bounded by 1 for STFT, it depends on the frequency for CWT (the average IF estimation error associated with ridge extraction ranges from 0.55, 1.74 and 3.34 for f_1, f_2 and f_3 respectively).

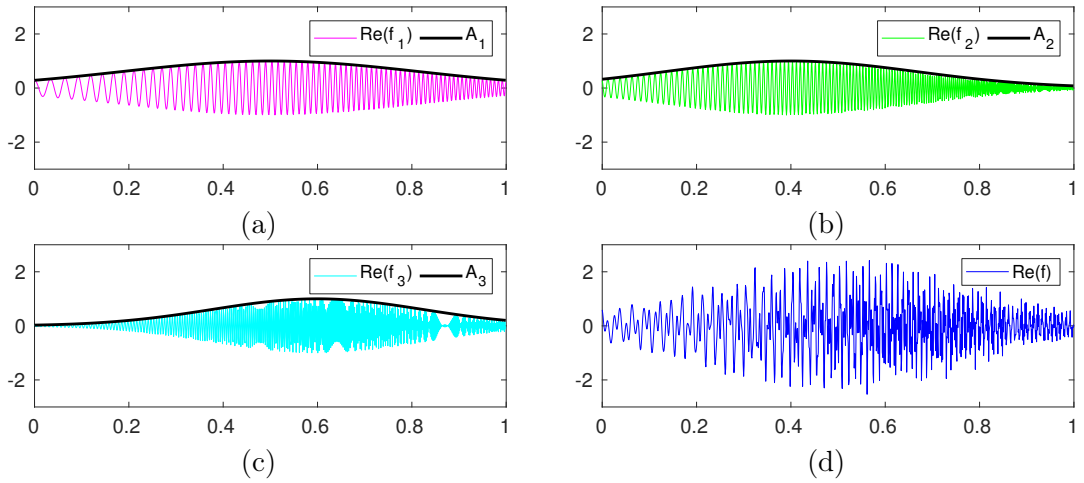


Fig. 3.2 (a), (b) and (c): real part of f_1, f_2, and f_3 respectively with Gaussian modulated amplitudes A_1, A_2 and A_3 superimposed; (d): real part of f.

3.1.4 Numerical Analysis of the Behavior of WSST2 and Comparisons

In this section, we provide numerical experiments to demonstrate the efficiency of WSST2 compared with other synchrosqueezing transforms including WSST, FSST, and FSST2 as introduced in the previous chapter. More precisely, we carry out a comparison in terms of concentration and accuracy of the TFRs obtained. For that purpose, we start with considering a complex simulated MCS (f) composed of three components: a linear chirp (f_1), a hyperbolic chirp (f_2) and an exponential chirp (f_3) with Gaussian modulated amplitudes, whose instantaneous frequencies are respectively linear ($\phi''(t) \propto cst$), hyperbolic ($\phi''(t) \propto \phi'(t)^2$) and exponential ($\phi''(t) \propto \phi'(t)$). Note also that f_1 behaves locally as a Gaussian modulated linear chirp that is mathematically proved to be perfectly handled by both FSST2 and WSST2, while the other two components contain strong nonlinear frequency modulations.

In our simulations, f is uniformly sampled over time interval $[0, 1]$ with a sampling rate $M = 1024$ Hz. In Figures 3.2 (a), (b) and (c), we display respectively the real part of the three components along with their amplitudes, and, in Figure 3.2 (d), the real part of the whole signal. Moreover, the CWT (*resp.* STFT) of f is then computed with the L^1- normalized complex Morlet wavelet (*resp.* Gaussian window):

$$\psi(t, \sigma_W) = \frac{1}{\sqrt{\sigma_W}} e^{-\pi \frac{t^2}{\sigma_W^2}} e^{i2\pi t} \text{ and } g(t, \sigma_F) = \sigma_F^{-\frac{1}{2}} e^{-\pi \frac{t^2}{\sigma_F^2}},$$

where σ_W and σ_F are optimal values in some sense as explained hereafter. An arbitrary threshold $\gamma = \gamma_0 = 0.001$ is also set for noise-free signals (the results obtained being relatively insensitive to that threshold). In addition, the wavelet-based (*resp.* STFT-based) synchrosqueezing transforms are represented on a logarithmic (*resp.* linear) scale. The Matlab codes for synchrosqueezing transforms and the scripts leading to all figures of this chapter can be found https://github.com/phamduonghung/WSST2.

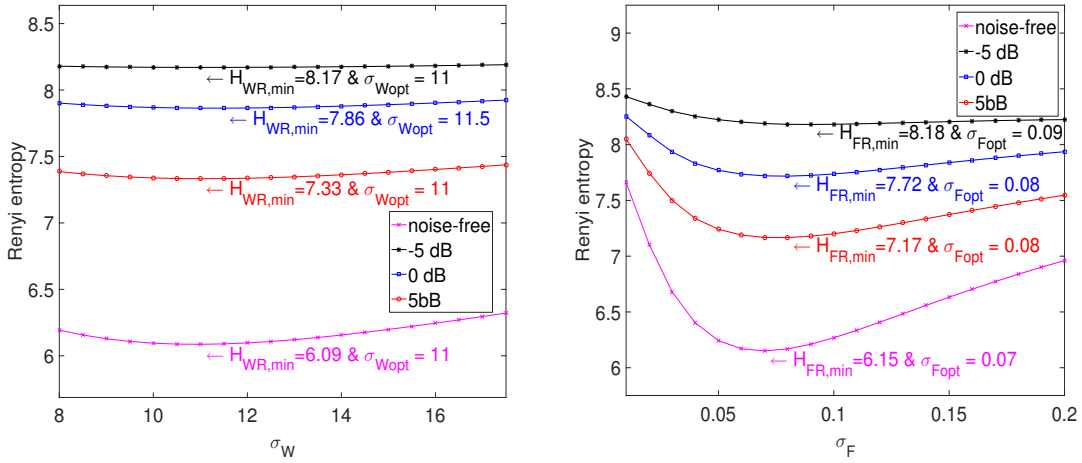


Fig. 3.3 Evolution of Rényi entropies with respect to σ_W and σ_F either in the noise-free, 5 dB, 0 dB or -5 dB cases for: (a) CWT; (b) STFT.

A. Optimal Wavelet and Window Length Determination

One of the well-known issues regarding the use of CWT (*resp.* STFT) is the choice of an appropriate length for Morlet wavelet (*resp.* Gaussian window) to allow for a good trade-off between time and frequency localization. In the synchrosqueezing context, this choice has also a strong impact on the accuracy of mode reconstruction: to use an inappropriate window may lead to the failure of ridge extraction and then of mode retrieval. To tackle this issue, a widely used approach is to measure the concentration of CWT (*resp.* STFT) which then allows us to pick the *optimal* window length as the one associated with the most concentrated representation. For that purpose, a relevant work is [88, 89], in which the concentration of the CWT (*resp.* STFT) is measured by means of Rényi entropy:

$$H_{WR}(\sigma_W) = \frac{1}{1-\alpha}\log_2\left(\frac{\int\int_{\mathbb{R}^2}|W_f^{\psi}(t,a)|^{\alpha}dadt}{\int\int_{\mathbb{R}^2}|W_f^{\psi}(t,a)|dadt}\right), \tag{3.19}$$

$$H_{FR}(\sigma_F) = \frac{1}{1-\alpha}\log_2\left(\frac{\int\int_{\mathbb{R}^2}|V_f^{g}(t,\eta)|^{\alpha}d\eta dt}{\int\int_{\mathbb{R}^2}|V_f^{g}(t,\eta)|d\eta dt}\right), \tag{3.20}$$

where integer orders $\alpha > 2$ is recommended. The larger the Rényi entropy, the less concentrated the CWT (*resp.* STFT). The optimal wavelet (*resp.* window) length parameter is thus determined as: $\sigma_{Wopt} = \arg\min_{\sigma}(H_{WR}(\sigma_W))$ (*resp.* $\sigma_{Fopt} = \arg\min_{\sigma}(H_{FR}(\sigma_F))$). Note also that, as we will see, this optimal determination procedure allows for a relatively fair comparison between the tested STFT-based and wavelet-based methods.

In Figure 3.3, we display the evolution of Rényi entropies ($\alpha = 3$) with respect to σ of the CWT and STFT for the signal f introduced above at different noise levels (noise-free, -5, 0, 5 dB). We notice that it exhibits a local minimum at a specific value for σ_W or σ_F and the corresponding optimal value is relatively stable with the noise levels.

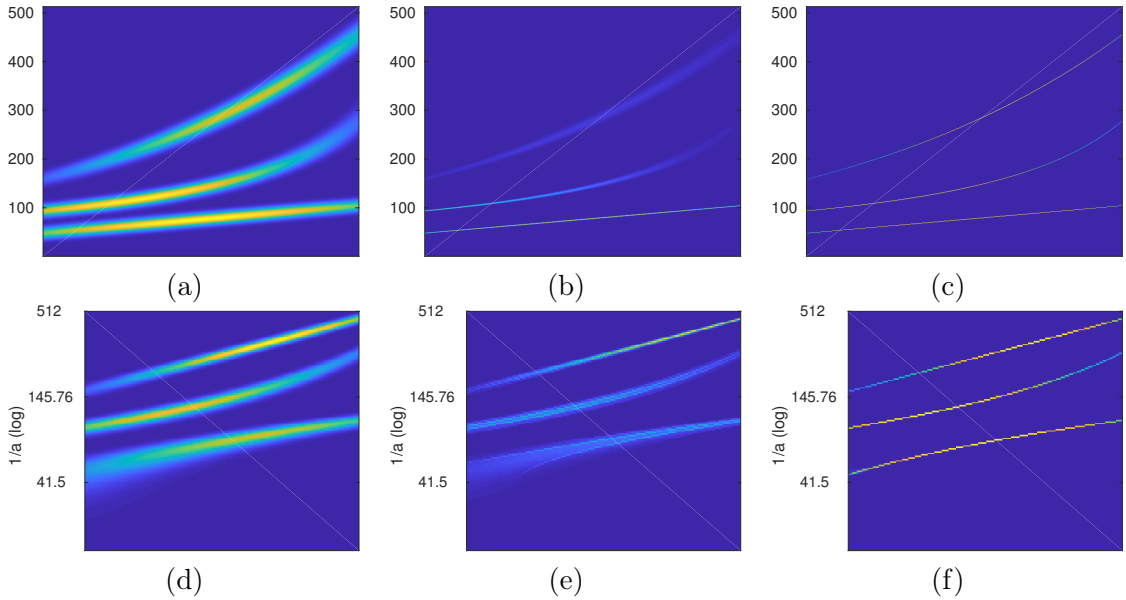


Fig. 3.4 *First row*, (a): modulus of the STFT of f; (b): FSST; (c): FSST2 ; *Second row*, (d): modulus of the CWT of f; (e): WSST; (f) WSST2. Threshold $\gamma_0 = 0.001$.

Having determined the optimal σ_{Wopt} and σ_{Fopt}, we display, still in the noise-free context, the STFT and CWT of f in the first column of Figure 3.4. Then, on the other two columns of this figure, the reassigned versions of CWT and STFT given by the aforementioned synchrosqueezing transforms are depicted. Analyzing these figures, we first remark that, as expected, FSST leads to a relatively sharp TFR for the linear chirp f_1, that looks similar to the ones given by WSST2 and FSST2, and much better than that corresponding to WSST. It is worth mentioning here that this worse representation is related to the scale discretization and not to the quality of the IF estimate as will be shown later. We shall also remark that FSST fails to reassign correctly the STFT of f_2 and f_3 where their frequency modulations are non-negligible. In contrast, the reassigned representations of f_2 and f_3 provided by WSST are much more concentrated at these locations. Moreover, it is also of interest to remark that the quality of the representation corresponding to WSST seems not to depend on the scale for f_2 contrary to what happens with f_3. Finally, for f_2 and f_3, both WSST2 and FSST2

seem to behave very similarly when considering either of the three studied modes and result in compact TFRs.

For a better understanding of the performance improvements brought by the use of WSST2 over other studied methods, the following section introduces a quantitative comparison of all these techniques from the angle of energy concentration of TFRs, and then a measure of their accuracy by means of the Earth mover's distance (EMD). But, before detailing them, we first study the stability of IF estimation with FSST and WSST on a linear chirp and then switch to that of WSST on a hyperbolic chirp.

B. Stability of IF Estimation with FSST and WSST on a Linear Chirp

In order to explain the different behaviors of STFT and WSST when applied to a linear chirp. Given a linear chirp $h_c(t) = Ae^{2i\pi\phi(t)}$, and if $|\eta - \phi^{'}(t)| < \Delta$ and STFT is performed with $g_{\sigma_F}(t, \sigma_F)$, one has:

$$|\widehat{\omega}_{h_c}(t,\eta) - \phi^{'}(t)| \leq \quad \Delta \left|1 - \frac{1}{1+\sigma_F^4 \phi^{''}(t)^2}\right| \leq \Delta, \tag{3.21}$$

which means that this IF estimation is stable for all t.

Similarly, in the wavelet case, we study the quality of the estimate $\widehat{\omega}_{h_c}(t,a)$, for which we have the following result:

Theorem 3.13. *If one performs the decomposition with the Morlet wavelet $\psi(t, \sigma_W)$ and if $|\frac{1}{a} - \phi^{'}(t)| \leq \Delta$ then:*

$$|\widehat{\omega}_{h_c}(t,a) - \phi^{'}(t)| \quad \leq \Delta, \tag{3.22}$$

meaning the IF estimation is stable for all t.

The proof of this theorem is available in Appendix A.3.

C. Stability of IF Estimation with WSST on a Hyperbolic Chirp

In this section, we show that WSST is well behaved when applied to an hyperbolic chirp, as suggested by Figure 3.4 (e). We show, in the following theorem, the stability of IF estimation with $\widehat{\omega}_f(t,a)$, when a Cauchy wavelet is used for the decomposition (the proof for the Morlet wavelet still needs to be carried out).

Theorem 3.14. *Let f be an hyperbolic chirp defined by $f(t) = t^{i\alpha}$ for $0 < t$ and any α in $\mathbb{R}$, and consider the estimate $\widehat{\omega}_f(t,a)$ computed using the Cauchy wavelet with parameter 1, whose definition of order β, a strictly positive real,*

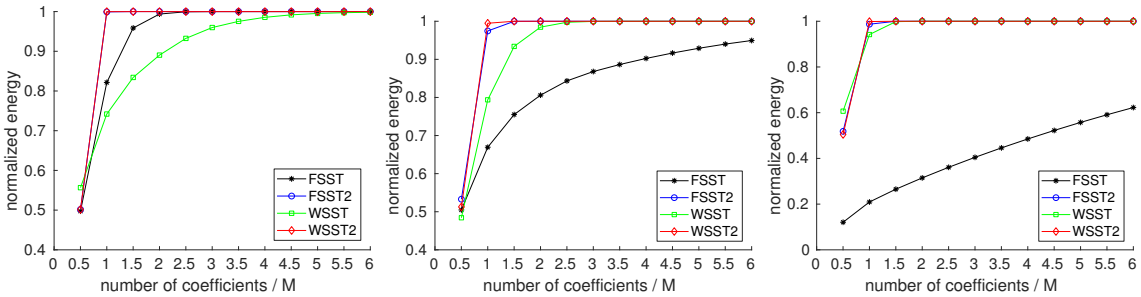
$$g_\beta(t) = \Gamma(\beta+1)(1 - i2\pi t)^{-(1+\beta)}, \tag{3.23}$$

where for any complex z with strictly positive real part, $\Gamma(z) = \int_0^{+\infty} t^{z-1}e^{-t}dt$. Then, if $|\frac{1}{a} - \phi^{'}(t)| \leq \Delta$, one has:

$$|\widehat{\omega}_f(t,a) - \phi^{'}(t)| \quad \leq \Delta, \tag{3.24}$$

meaning the IF estimation is stable for all t.

The proof of this theorem is available in Appendix A.4.



(a) (b) (c)

Fig. 3.5 (a) Normalized energy as a function of the number of sorted TF coefficients for f_1. Abscissa corresponds to the number of coefficients over the size M of the signal; (b): same as (a) but for f_2 ; (c): same as (a) but for f_3. Threshold $\gamma_0 = 0.001$.

D. Evaluation of TF Concentration

The TF concentration is one of the outstanding features used for evaluating the performance of the different TF techniques. In order to quantify this, an appealing method first introduced in [48] is used in this chapter. The main goal of such a method is to measure the energy concentration by considering the proportion of the latter contained in the first nonzero coefficients associated with the highest amplitudes, which we call *normalized energy*. When computed on a mono-component signal, the faster it increases towards 1 with the number of coefficients involved, the more concentrated the TFR. In Figure 3.5 (a), we depict the normalized energy corresponding to the reassignment of the STFT of f_1 using different techniques, with respect to the number of coefficients kept divided by the length of f_1 (which corresponds to the sampling rate M in our case). Since we consider only one mode, a good representation has to have its energy mostly contained in the first M coefficients, which correspond to abscissa 1 in the graph of Figure 3.5 (a) Not surprisingly, the energy of f_1 is perfectly localized when using either WSST2 or FSST2, since they require only one coefficient per time instant to recover the signal energy, while WSST and FSST need more coefficients (5 and 2 respectively). The same computations carried out on f_2 and f_3 show that WSST2 still better performs than the other three methods, especially WSST or FSST.

In order to further challenge the different TF reassignment techniques in the presence of noise, we consider a noisy signal $f_\zeta(t) = f(t) + \zeta(t)$, where $\zeta(t)$ is a complex white Gaussian process with variance $\mathrm{Var}\left(\Re\left\{\zeta(t)\right\}\right) = \mathrm{Var}\left(\Im\left\{\zeta(t)\right\}\right) = \sigma_\zeta^2$. Note also that, in this noisy context, one of the well-known issue regarding the use of SST is the choice of an appropriate threshold γ on $W_f^\psi(t,a)$ or $V_f^g(t,\eta)$ in the definition of the synchrosqueezing operator to allow for signal denoising and a fair comparison between the different tested methods. Here, we propose to use one of the the most commonly used techniques for signal denoising: hard thresholding (HT) [90, 73]. Such a technique enables adaptive determination of the threshold γ as a function of the noise level. Indeed, it exploits the linearity of CWT and the fact that, for a fixed scale a, one has:

$$\mathrm{Var}\left(\Re\left\{W_\zeta^\psi(t,a)\right\}\right) = \frac{1}{a}\sigma_\zeta^2\left\|\psi\right\|_2^2 \text{ and } \mathrm{Var}\left(\Im\left\{W_\zeta^\psi(t,a)\right\}\right) = \frac{1}{a}\sigma_\zeta^2\left\|\psi\right\|_2^2,$$

where Var stands for the variance. From this, we deduce that $\left|W_\zeta^\psi(t,a)\right|^2 \Big/ \frac{1}{a}\sigma_\zeta^2\left\|\psi\right\|_2^2$ is χ_2 distributed with 2 degrees of freedom. If $\left|W_\zeta^\psi(t,a)\right|^2$ is larger than $9 \times \frac{1}{a}\sigma_\zeta^2\left\|\psi\right\|_2^2$, the probability of false alarm is less than 1%. Therefore, assuming the variance of the noise σ_ζ^2 is known, HT for CWT means to threshold coefficients as:

$$\overline{W}_{f_\zeta}^\psi(t,a) = \begin{cases} W_{f_\zeta}^\psi(t,a), & \text{if } \left|W_{f_\zeta}^\psi(t,a)\right| \geq \gamma_W = 3\sigma_\zeta\frac{1}{\sqrt{a}}\left\|\psi\right\|_2 \\ 0 & \text{otherwise,} \end{cases} \tag{3.25}$$

The same arguments applied to the STFT case leads to:

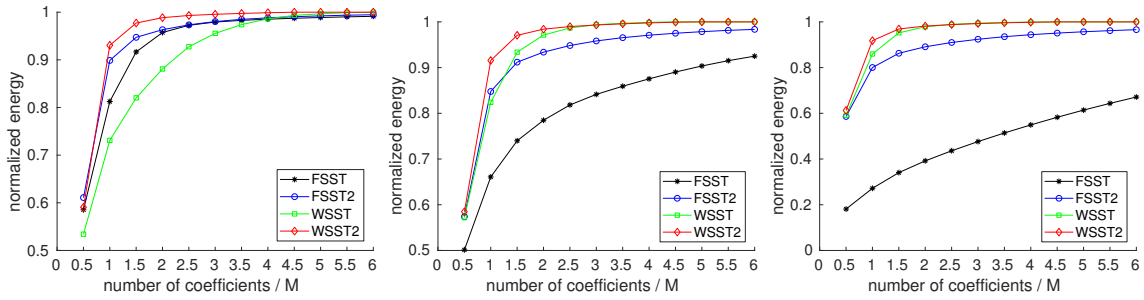
$$\overline{V}_{f_\zeta}^g(t,\eta) = \begin{cases} V_{f_\zeta}^g(t,\eta), & \text{if } \left|V_{f_\zeta}^g(t,\eta)\right| \geq \gamma_F = 3\sigma_\zeta\|g\|_2 \\ 0 & \text{otherwise,} \end{cases} \tag{3.26}$$

which enables to guarantee an efficient noise removal. In real-life applications, threshold levels γ_W and γ_F are unknown and need to be estimated. For example, by considering the *median absolute deviation* approach [91], a robust estimator for wavelet decomposition was proposed in [90]:

$$\hat{\gamma}_W = \frac{\underset{a}{\mathrm{median}}\left|\Re\left\{W_{f_\zeta}^\psi(t,a)\right\}\right|}{0.6745} \text{ and } \hat{\gamma}_F = \frac{\underset{\eta}{\mathrm{median}}\left|\Re\left\{V_{f_\zeta}^g(t,\eta)\right\}\right|}{0.6745},$$

where median represents the median of the coefficients.

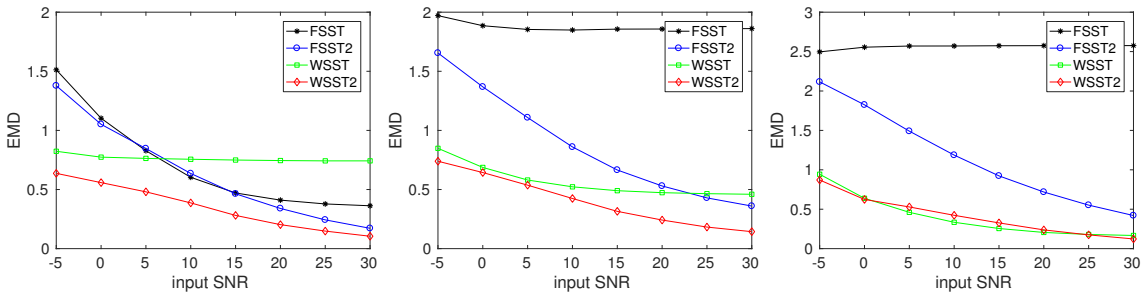
Then, we carry out the same numerical experiments regarding energy concentration as in the noise-free case, each mode being this time contaminated by a white Gaussian noise (SNR = 0 dB). The results displayed in Figure 3.6 exhibit a slightly slower growth of the normalized energy since the coefficients corresponding to noise, that the above technique cannot completely eliminate, are spread out over the whole TF or TS planes. However, the normalized energy is still more concentrated when using WSST2 than the other methods, even



(a) (b) (c)

Fig. 3.6 (a) Normalized energy as a function of the number of sorted TF coefficients for noisy f_1 (SNR $= 0$ dB); (b): same as (a) but for noisy f_2 (SNR $= 0$ dB); (c): same as (a) but for noisy f_3 (SNR $= 0$ dB).

for mode f_1 when compared with FSST2. These facts clearly show that the representation provided by the former technique is the most concentrated, even in heavy noise situations.



(a) (b) (c)

Fig. 3.7 (a): EMD corresponding to different TFRs of f_1 given by the synchrosqueezing transforms; (b): same as (a) but for f_2; (c): same as (a) but for f_3.

Although quite informative, the method based on normalized energy does not deliver any insight into the location accuracy of the reassigned transforms. The latter can alternatively be quantified by measuring the dissimilarity between the resultant TFR and the corresponding ideal one by means of the *Earth mover's distance* (EMD), a procedure already used in the synchrosqueezing context in [44]. The EMD is a sliced Wasserstein distance, commonly used in optimal transport, which gives the amount of *work* needed to deform one probability distribution to another. The principle of EMD in the synchrosqueezing context is to compute the 1D EMD between the resultant TFR and the corresponding ideal one, for each individual time t, and then take the average overall t to define the global EMD. Mathematically, one has the following definition of EMD [92, 44].

Definition 3.15. *Given two probability distributions ϖ_1 and ϖ_2 on a metric space (Γ, ϱ) and an optimization over $\mathcal{P}(\varpi_1, \varpi_1)$, the set of all probability measures on $\Gamma \times \Gamma$ with ϖ_1 and ϖ_2*

being considered as marginals, the standard EMD is defined by:

$$d_{EMD}(\varpi_1, \varpi_2) = \int_{\vartheta \in \mathcal{P}(\varpi_1, \varpi_1)} \varrho(x, y) d\vartheta(x, y). \tag{3.27}$$

In particular, if $\Gamma \in \mathbb{R}$ *and* ϱ *is the canonical Euclidean distance* $\varrho(x, y) = |x - y|$*, then the standard EMD reduces to 1D EMD as:*

$$d_{1DEMD}(\varpi_1, \varpi_2) = \int_{\Gamma} |f_{\varpi_1}(x) - f_{\varpi_2}(x)| dx, \tag{3.28}$$

where $f_{\varpi_i}(x) = \int_{-\infty}^{x} d\varpi_i$ *with* $i = 1, 2$.

A smaller EMD means a better TFR concentration to the ground truth and fewer noise fluctuations. Note also that since EMD is defined for probability distributions, the TFRs is first normalized such that their integral equal 1 for t. In Figures 3.7 (a), (b) and (c), we display, respectively for the three modes already tested, the evolution of EMD with respect to the noise level, for TFRs given either by WSST, WSST2, FSST or FSST2. For linear chirp f_1, WSST2 always achieves the best performance of the TF concentration to the ground truth whatever the input SNR, even compared with FSST2. Moving to f_2 and f_3, WSST2 performs similarly to WSST at the high noise level and is more accurate at the low noise level, while it consistently outperforms FSST and FSST2. These results confirm the interest of using WSST2 on many different types of signals, even in the presence of heavy noise.

Table 3.1 Accuracy of mode retrieval in the noise-free case

	FSST	**FSST2**	**WSST**	**WSST2**
Mode f_1	8.77	29.4	4.57	**50.3**
Mode f_2	2.64	18.4	8.12	**22.9**
Mode f_3	0.646	18.1	14.6	**30.4**
MCS f	3.18	20.0	7.11	**27.2**

Table 3.2 Accuracy of mode retrieval in the noisy case, (SNR = 0 dB)

	FSST	**FSST2**	**WSST**	**WSST2**
Mode f_1	4.46	6.01	3.76	**6.65**
Mode f_2	2.62	4.84	4.58	**6.00**
Mode f_3	0.46	3.07	3.06	**3.41**
MCS f	2.38	4.59	3.80	**5.34**

E. Evaluation of Mode Reconstruction Performance

As discussed above, the variants of second-order SST (WSST2) proposed in this chapter leading to significantly better TFRs, this should translate into better performance in terms of mode reconstruction. Let us first briefly recall that f_k is retrieved from the TFR of f given by the WSST2 (other SSTs have the same mode retrieval procedure) through the formula (3.18):

$$f_k(m/n) \approx \frac{1}{C'_{d,\psi,k}} S^{\psi,\gamma}_{d,2,f}(m, \varphi_k(m/n)), \tag{3.29}$$

where $\varphi_k(m/n)$ is the estimate of $\phi'_k(m/n)$ given by the ridge detector (see Section 2.6.1 for details on such a technique). This formula also indicates that we only use the information on the ridge to reconstruct the mode (i.e. $d = 0$). In Table 3.1, we display the output SNRs for modes f_1, f_2, f_3 and also for f, using either FSST, FSST2, WSST or WSST2 for mode reconstruction. Further, we carry out the same experiments, but each mode is embedded in a white Gaussian noise at a noise level: SNR = 0 dB. The resultant accuracies for such a reconstruction are displayed in Table 3.2. From these results, we can see that the improvement brought by using WSST2 is clear and consistent with the previous study of the accuracy of the proposed TFRs. In particular, WSST2 always provide better results than FSST2 in both noise-free and noisy cases although the mathematical analysis shows that they have a similar behavior. The reason of this phenomena is that the discretization of scale in the wavelet case is finer the in low-frequency range and coarser in the high-frequency range, whereas that of frequency in the STFT case is unchanged, all mentioned in Section 3.1.3. As a result, TFRs given by wavelet-based transforms are generally more accurate than the ones associated with STFT-based transforms, which leads undoubtedly to a great impact on the accuracy of the reconstruction results.

3.2 High-order STFT-based Synchrosqueezing Transforms

As highlighted in Section 2.6.3, FSST2 is proven capable of handling a class of chirp signals that are larger than the original FSST does. Actually, the former is obtained through the definition of the improved synchrosqueezing operator $T^{g,\gamma}_{2,f}(t,\eta)$ using the second-order local modulation operator $\widetilde{q}_{t,f}(t,\eta)$, which corresponds to the ratio of the first-order derivatives, with respect to t, of the reassignment operators. However, by using partial derivatives with respect to η instead of t, a new second-order local modulation operator $\widetilde{q}_{\eta,f}(t,\eta)$ showing the same properties as those of $\widetilde{q}_{t,f}(t,\eta)$ can also be obtained as follows.

3.2.1 New Definition of Second Order STFT-based SST (FSST2)

Let first introduce the new definition of the second-order local complex modulation operator $\widetilde{q}_{\eta,f}(t,\eta)$.

Definition 3.16. *Given a signal $f \in L^2(\mathbb{R})$, the second-order local complex modulation operator $\widetilde{q}_{\eta,f}$ is defined by:*

$$\widetilde{q}_{\eta,f}(t,\eta) = \frac{\partial_\eta \widetilde{\omega}_f(t,\eta)}{\partial_\eta \widetilde{\tau}_f(t,\eta)} \qquad \text{whenever } \partial_\eta \widetilde{\tau}_f(t,\eta) \neq 0, \tag{3.30}$$

where $\widetilde{\omega}_f(t,\eta)$ and $\widetilde{\tau}_f(t,\eta)$ are respectively defined in Proposition 2.43.

Note that a more practical formula for this operator is also obtained as:

Proposition 3.17. *The second-order modulation operator $\widetilde{q}_{\eta,f}(t,\eta)$ can be computed through:*

$$\widetilde{q}_{\eta,f} = \frac{1}{i2\pi} \frac{\left(V_f^g\right)^2 + V_f^g V_f^{tg'} - V_f^{g'} V_f^{tg}}{\left(V_f^{tg}\right)^2 - V_f^g V_f^{t^2 g}}, \tag{3.31}$$

where $V_f^{t^2 g}$ is the STFT of the signal f computed with window $t \mapsto t^2 g(t)$.

Proof. By considering the partial derivatives of $\widetilde{\omega}_f(t,\eta)$ and $\widetilde{\tau}_f(t,\eta)$ with respect to η in the expressions given in Proposition 2.45, and then using the formula $\partial_\eta V_f^g(t,\eta) = -i2\pi V_f^{tg}(t,\eta)$, the expression for $\widetilde{q}_{\eta,f}$ follows. □

Then, a new definition of the improved IF estimate having the same properties as $\widetilde{\omega}^{[2]}_{t,f}(t,\eta)$ is derived as follows:

Definition 3.18. *Let $f \in L^2(\mathbb{R})$, the second-order local complex IF estimate of signal f is defined by:*

$$\widetilde{\omega}^{[2]}_{\eta,f}(t,\eta) = \begin{cases} \widetilde{\omega}_f(t,\eta) + \widetilde{q}_{\eta,f}(t,\eta)(t - \widetilde{\tau}_f(t,\eta)) & \text{if } \partial_\eta \widetilde{\tau}_f(t,\eta) \neq 0 \\ \widetilde{\omega}_f(t,\eta) & \text{otherwise.} \end{cases}$$

Then, its real part $\widehat{\omega}^{[2]}_{\eta,f}(t,\eta) = \Re e\{\widetilde{\omega}^{[2]}_{\eta,f}(t,\eta)\}$ is the desired IF estimate.

The next proposition shows that it perfectly estimates the IF of a Gaussian modulated linear chirp.

Proposition 3.19. *If $f(t) = A(t)e^{i2\pi\phi(t)}$ is a Gaussian modulated linear chirp, then $\widehat{\omega}^{[2]}_{\eta,f}(t,\eta) = \phi'(t)$.*

Proof. Let us consider a mode $f(\tau) = A(\tau)e^{i2\pi\phi(\tau)}$ where $\log(A(\tau))$ and $\phi(\tau)$ are quadratic functions described by:

$$\log(A(\tau)) = \sum_{k=0}^{2} \frac{\alpha_k}{k!}\tau^k \quad \text{and} \quad \phi(\tau) = \sum_{k=0}^{2} \frac{\beta_k}{k!}\tau^k,$$

with $\alpha_k, \beta_k \in \mathbb{R}$. The STFT of this mode with any window g, at time t and frequency η, can be written as:

$$\begin{aligned} V_f^g(t,\eta) &= \int_{\mathbb{R}} f(\tau+t)g(\tau)e^{-i2\pi\eta\tau}d\tau \\ &= \int_{\mathbb{R}} \exp\left(\sum_{k=0}^{2}\frac{1}{k!}\left(\alpha_k + i2\pi\beta_k\right)(\tau+t)^k\right) g(\tau)e^{-i2\pi\eta\tau}d\tau. \end{aligned}$$

By taking the partial derivative of $V_f^g(t,\eta)$ with respect t, and then dividing by $i2\pi V_f^g(t,\eta)$, the local IF estimate $\widetilde{\omega}_f(t,\eta)$ defined in Proposition 2.43 can be obtained for $V_f^g(t,\eta) \neq 0$:

$$\widetilde{\omega}_f(t,\eta) = \sum_{k=1}^{2}\left(\frac{1}{i2\pi}\alpha_k + \beta_k\right)t^{k-1} + \left(\frac{1}{i2\pi}\alpha_2 + \beta_2\right)\frac{V_f^{tg}(t,\eta)}{V_f^g(t,\eta)}. \tag{3.32}$$

Then, taking the partial derivative of (3.32) with respect to η and recalling from Proposition 2.45 that $\dfrac{V_f^{tg}(t,\eta)}{V_f^g(t,\eta)} = \widetilde{\tau}_f(t,\eta) - t$, we get the following expression:

$$\partial_\eta\widetilde{\omega}_f(t,\eta) = \left(\frac{1}{i2\pi}\alpha_2 + \beta_2\right)\partial_\eta\widetilde{\tau}_f(t,\eta). \tag{3.33}$$

Setting $\widetilde{q}_{\eta,f}(t,\eta) = \dfrac{\partial_\eta\widetilde{\omega}_f(t,\eta)}{\partial_\eta\widetilde{\tau}_f(t,\eta)}$ assuming $\partial_\eta\widetilde{\tau}_f(t,\eta) \neq 0$ and noting that $\beta_2 = \phi''(t)$, from (3.32) and (3.33), we also have the following result:

$$\begin{aligned} \phi'(t) &= \beta_1 + \beta_2 t \\ &= \Re\mathrm{e}\left\{\widetilde{\omega}_f(t,\eta) - \left(\frac{1}{i2\pi}\alpha_2 + \beta_2\right)\left(\widetilde{\tau}_f(t,\eta) - t\right)\right\} \\ &= \Re\mathrm{e}\left\{\widetilde{\omega}_f(t,\eta) + \widetilde{q}_{\eta,f}(t,\eta)(t - \widetilde{\tau}_f(t,\eta))\right\}. \end{aligned} \tag{3.34}$$

Putting $\widetilde{\omega}^{[2]}_{\eta,f}(t,\eta) = \widetilde{\omega}_f(t,\eta) + \widetilde{q}_{\eta,f}(t,\eta)(t - \widetilde{\tau}_f(t,\eta))$, it follows that $\phi'(t) = \Re\mathrm{e}\left\{\widetilde{\omega}^{[2]}_{\eta,f}(t,\eta)\right\}$, which ends the proof. □

Finally, a new definition of second-order FSST (FSST2) is derived by simply replacing $\widehat{\omega}^{[2]}_{t,f}(t,\eta)$ by $\widehat{\omega}^{[2]}_{\eta,f}(t,\eta)$ in (2.77), which then allows for the retrieval of mode f_k, as in (2.78).

3.2.2 Higher Order Synchrosqueezing Transform

Despite FSST2 definitely improves the concentration of TFR, it is only demonstrated to work well on perturbed linear chirps with Gaussian modulated amplitudes. In order to handle signals containing more general types of AM-FM modes having non-negligible $\phi_k^{(n)}(t)$ for

$n \geq 3$, we are going to define new synchrosqueezing operators, based on approximation orders higher than three for both amplitude and phase [5].

A. Nth-order IF Estimate

The new IF estimate we define here relies on high-order Taylor expansions of the amplitude and phase of a mode. To do so, let us first consider a mode defined as in the following:

Definition 3.20. *Given a mode $f(\tau) = A(\tau)e^{i2\pi\phi(\tau)}$ in $L^2(\mathbb{R})$ with $A(\tau)$ (resp. $\phi(\tau)$) equal to its L^{th}-order (resp. N^{th}-order) Taylor expansion for τ close to t:*

$$\log(A(\tau)) = \sum_{k=0}^{L} \frac{[\log(A)]^{(k)}(t)}{k!} (\tau - t)^k$$
$$\phi(\tau) = \sum_{k=0}^{N} \frac{\phi^{(k)}(t)}{k!} (\tau - t)^k$$

where $Z^{(k)}(t)$ denotes the k^{th} derivative of Z evaluated at t.

A mode f defined as above, with $L \leq N$, can be written as:

$$f(\tau) = \exp\left(\sum_{k=0}^{N} \frac{1}{k!} \left([\log(A)]^{(k)}(t) + i2\pi\phi^{(k)}(t)\right)(\tau - t)^k\right),$$

since $[\log(A)]^{(k)}(t) = 0$ if $L + 1 \leq k \leq N$. Consequently, the STFT of this mode at time t and frequency η reads:

$$\begin{aligned} V_f^g(t,\eta) &= \int_{\mathbb{R}} f(\tau + t)g(\tau)e^{-i2\pi\eta\tau}d\tau \\ &= \int_{\mathbb{R}} \exp\left(\sum_{k=0}^{N} \frac{1}{k!} \left([\log(A)]^{(k)}(t) + i2\pi\phi^{(k)}(t)\right)\tau^k\right) g(\tau)e^{-i2\pi\eta\tau}d\tau. \end{aligned}$$

By taking the partial derivative of $V_f^g(t,\eta)$ with respect to t and then dividing by $i2\pi V_f^g(t,\eta)$, the local IF estimate $\widetilde{\omega}_f(t,\eta)$ defined in Proposition 2.43 can be written when $V_f^g(t,\eta) \neq 0$ as:

$$\begin{aligned} \widetilde{\omega}_f(t,\eta) &= \sum_{k=1}^{N} r_k(t) \frac{V_f^{t^{k-1}g}(t,\eta)}{V_f^g(t,\eta)} \\ &= \frac{1}{i2\pi}[\log(A)]'(t) + \phi'(t) + \sum_{k=2}^{N} r_k(t) \frac{V_f^{t^{k-1}g}(t,\eta)}{V_f^g(t,\eta)}, \end{aligned} \tag{3.35}$$

where $r_k(t)$ are functions of t defined for $k = 1, \dots, N$ as:

$$r_k(t) = \frac{1}{(k-1)!}\left(\frac{1}{i2\pi}[\log(A)]^{(k)}(t) + \phi^{(k)}(t)\right).$$

It is clear from (3.35) that, since $A(t)$ and $\phi(t)$ are real expressions, $\Re\mathrm{e}\{\widetilde{\omega}_f(t,\eta)\} = \phi'(t)$ does not hold when the sum on the right hand side of (3.35) has a non-zero real part. As in the case of the Gaussian modulated linear chirp introduced before, to get the exact IF estimate for the studied signal, one needs to subtract $\Re\mathrm{e}\left\{\sum_{k=2}^{N} r_k(t)\frac{V_f^{t^{k-1}g}(t,\eta)}{V_f^g(t,\eta)}\right\}$ to $\Re\mathrm{e}\{\widetilde{\omega}_f(t,\eta)\}$, for which $r_k(t)$, for all $k = 2, \dots, N$, must be estimated.

For that purpose, inspired by our study of the Gaussian-modulated linear chirp, we derive a frequency modulation operator $\widetilde{q}_{\eta,f}^{[k,N]}$, equal to $r_k(t)$ when f satisfies Definition 3.20, obtained by differentiating different STFTs with respect to η, as explained hereafter. Note that we choose to differentiate with respect to η rather than t because it leads to much simpler expressions, mainly as a result of the following formulae:

$$\begin{aligned}\partial_t V_f^g(t,\eta) &= i2\pi\eta V_f^g(t,\eta) - V_f^{g'}(t,\eta)\\ \partial_\eta V_f^g(t,\eta) &= -i2\pi V_f^{tg}(t,\eta).\end{aligned} \tag{3.36}$$

The different modulation operators $\widetilde{q}_{\eta,f}^{[k,N]}$ for $k = 2, \dots, N$ can then be derived recursively, as explained in the next proposition:

Proposition 3.21. *Given a mode $f \in L^2(\mathbb{R})$ that satisfies Definition 3.20 with $L \leq N$, the $N-1$ local modulation operators $\widetilde{q}_{\eta,f}^{[k,N]}$ such that $\Re\mathrm{e}\left\{\widetilde{q}_{\eta,f}^{[k,N]}(t,\eta)\right\} = \frac{\phi^{(k)}(t)}{(k-1)!}$, $k = 2, \dots, N$, can be determined by:*

$$\begin{aligned}&\widetilde{q}_{\eta,f}^{[N,N]}(t,\eta) = y_N(t,\eta) \text{ and}\\ &\widetilde{q}_{\eta,f}^{[j,N]}(t,\eta) = y_j(t,\eta) - \sum_{k=j+1}^{N} x_{k,j}(t,\eta)\widetilde{q}_{\eta,f}^{[k,N]}(t,\eta) \qquad \text{for } j = N-1, N-2, \dots, 2,\end{aligned}$$

where $y_j(t,\eta)$ and $x_{k,j}(t,\eta)$ are defined as follows. For any (t,η) s.t. $V_f^g(t,\eta) \neq 0$ and $\partial_\eta x_{j,j-1}(t,\eta) \neq 0$, we put:

$$\begin{aligned}&\text{for } k = 1 \dots N, \quad y_1(t,\eta) = \widetilde{\omega}_f(t,\eta) \text{ and } x_{k,1}(t,\eta) = \frac{V_f^{t^{k-1}g}(t,\eta)}{V_f^g(t,\eta)},\\ &\text{and then for } j = 2 \dots N \text{ and } k = j \dots N,\\ &y_j(t,\eta) = \frac{\partial_\eta y_{j-1}(t,\eta)}{\partial_\eta x_{j,j-1}(t,\eta)} \text{ and } x_{k,j}(t,\eta) = \frac{\partial_\eta x_{k,j-1}(t,\eta)}{\partial_\eta x_{j,j-1}(t,\eta)}.\end{aligned}$$

The proof of Proposition 3.21 is given in Appendix B.1. Then, the definition of the N$^{\text{th}}$-order IF estimate follows:

Definition 3.22. *Let $f \in L^2(\mathbb{R})$, the N^{th}-order local complex IF estimate $\widetilde{\omega}_{\eta,f}^{[N]}$ is defined by:*

$$\widetilde{\omega}_{\eta,f}^{[N]}(t,\eta) = \begin{cases} \widetilde{\omega}_f(t,\eta) + \sum_{k=2}^{N} \widetilde{q}_{\eta,f}^{[k,N]}(t,\eta)\left(-x_{k,1}(t,\eta)\right), & \text{if } V_f^g(t,\eta) \neq 0 \\ & \text{and } \partial_\eta x_{j,j-1}(t,\eta) \neq 0 \text{ for } j = 2 \dots N. \\ \widetilde{\omega}_f(t,\eta) & \text{otherwise.} \end{cases}$$

Then, its real part $\widehat{\omega}_{\eta,f}^{[N]}(t,\eta) = \Re e\{\widetilde{\omega}_{\eta,f}^{[N]}(t,\eta)\}$ is the desired IF estimate.

For this estimate, we have the following approximation result:

Proposition 3.23. *Given a mode $f \in L^2(\mathbb{R})$ that satisfies Definition 3.20 with $L \leq N$, then $\phi'(t) = \widehat{\omega}_{\eta,f}^{[N]}(t,\eta)$.*

Proof. From (3.35), we have:

$$\begin{aligned} \phi'(t) &= \Re\mathrm{e}\left\{\widetilde{\omega}_f(t,\eta) + \sum_{k=2}^{N} r_k(t)\left(-\frac{V_f^{t^{k-1}g}(t,\eta)}{V_f^g(t,\eta)}\right)\right\} \\ &= \Re\mathrm{e}\left\{\widetilde{\omega}_f(t,\eta) + \sum_{k=2}^{N} r_k(t)\left(-x_{k,1}(t,\eta)\right)\right\} \\ &= \Re\mathrm{e}\left\{\widetilde{\omega}_f(t,\eta) + \sum_{k=2}^{N} \widetilde{q}_{\eta,f}^{[k,N]}(t,\eta)\left(-x_{k,1}(t,\eta)\right)\right\}. \end{aligned} \tag{3.37}$$

Let us put $\widetilde{\omega}_{\eta,f}^{[N]}(t,\eta) = \widetilde{\omega}_f(t,\eta) + \sum_{k=2}^{N} \widetilde{q}_{\eta,f}^{[k,N]}(t,\eta)\left(-x_{k,1}(t,\eta)\right)$, we obtain $\phi'(t) = \widehat{\omega}_{\eta,f}^{[N]}(t,\eta)$, which ends the proof. □

B. Efficient Computation of Modulation Operators

The local modulation operators $\widetilde{q}_{\eta,f}^{[k,N]}$ defined in Proposition 3.21 should not be computed by approximating partial derivatives by means of discrete differentiation since this would generate numerical instability especially in the presence of noise. Therefore, to address this issue, we remark that these modulation operators can instead be computed analytically as functions of different STFTs. This is illustrated for $N = 4$ through the following proposition:

Proposition 3.24. *Let $f \in L^2(\mathbb{R})$, the modulation operators $\widetilde{q}_{\eta,f}^{[k,N]}$ for $N = 4$ and $k = 2, 3, 4$ can be expressed as:*

$$\widetilde{q}_{\eta,f}^{[4,4]} = G_4\left(V_f^{t^{0\dots6}g}, V_f^{t^{0\dots3}g'}\right),$$
$$\widetilde{q}_{\eta,f}^{[3,4]} = G_3\left(V_f^{t^{0\dots4}g}, V_f^{t^{0\dots2}g'}\right) - \widetilde{q}_{\eta,f}^{[4,4]} G_{3,4}\left(V_f^{t^{0\dots5}g}\right),$$
$$\widetilde{q}_{\eta,f}^{[2,4]} = G_2\left(V_f^{t^{0\dots2}g}, V_f^{t^{0\dots1}g'}\right) - \widetilde{q}_{\eta,f}^{[3,4]} G_{2,3}\left(V_f^{t^{0\dots3}g}\right) - \widetilde{q}_{\eta,f}^{[4,4]} G_{2,4}\left(V_f^{t^{0\dots4}g}\right),$$

where $G_k\left(V_f^{t^{0\dots m}g}, V_f^{t^{0\dots n}g'}\right)$ *is a function of* $V_f^{t^l g}$ *for* $l = 0, \dots, m$ *and* $V_f^{t^l g'}$ *for* $l = 0, \dots, n$ *while* $G_{k,j}\left(V_f^{t^{0\dots m}g}\right)$ *is associated with coefficient* $\widetilde{q}_{\eta,f}^{[j,N]}$ *in the computation of* $\widetilde{q}_{\eta,f}^{[k,N]}$ *for* $k \neq j$.

Also, we recall that the fourth-order IF estimate can be written as:

$$\widetilde{\omega}_{\eta,f}^{[4]}(t,\eta) = \widetilde{\omega}_f(t,\eta) + \widetilde{q}_{\eta,f}^{[2,4]}(t,\eta)\left(-x_{2,1}(t,\eta)\right) + \widetilde{q}_{\eta,f}^{[3,4]}(t,\eta)\left(-x_{3,1}(t,\eta)\right) + \widetilde{q}_{\eta,f}^{[4,4]}(t,\eta)\left(-x_{4,1}(t,\eta)\right).$$

The proof of Proposition 3.24 is available in Appendix B.2 where explicit forms for G_k and $G_{k,j}$ are given.

Remark 3.25. We first note that when $N = 2$, i.e. by neglecting $\widetilde{q}_{\eta,f}^{[3,4]}$ and $\widetilde{q}_{\eta,f}^{[4,4]}$ corresponding to orders 3 and 4, the second-order IF estimate $\widetilde{\omega}_{\eta,f}^{[2]}(t,\eta)$ defined in Proposition 3.18 is found again. Secondly, it is clear that the number of STFTs used to compute $\widetilde{q}_{\eta,f}^{[4,4]}$ is 11, namely $V_f^{t^l g}$ for $l = 0, \dots, 6$, and $V_f^{t^l g'}$ for $l = 0, \dots, 3$. Finally, generalizing the procedure detailed in the proof of Proposition 3.24 to any N, one obtains that $\widetilde{q}_{\eta,f}^{[N,N]}$ can be computed by means of $3N - 1$ STFTs, namely $V_f^{t^l g}$ for $l = 0, \dots, 2N-2$, and $V_f^{t^l g'}$ for $l = 0, \dots, N-1$.

C. Nth-order STFT-based SST (FSSTn)

As for FSST2, the N^{th}-order FSST (FSSTn) is defined by replacing $\widehat{\omega}_f(t,\eta)$ by $\widehat{\omega}_{\eta,f}^{[N]}(t,\eta)$ in (2.66):

Definition 3.26. *Given* $f \in L^2(\mathbb{R})$ *and a real number* $\gamma > 0$*, one defines the FSSTn operator with threshold* γ *as:*

$$T_{N,f}^{g,\gamma}(t,\omega) = \int_{|V_f^g(t,\eta)|>\gamma} V_f^g(t,\eta)\,\delta\left(\omega - \widehat{\omega}_{\eta,f}^{[N]}(t,\eta)\right) d\eta. \tag{3.38}$$

Finally, the modes of the MCS can be reconstructed by replacing $T_f^{g,\gamma}(t,\omega)$ by $T_{N,f}^{g,\gamma}(t,\omega)$ in (2.69) with a normalization $D'_{N,g,k} = \int_{\mathbb{R}} \overline{\mathcal{F}\{g(\tau)e^{-i\pi\phi_k^{(N)}(t)\tau^N}\}(\eta)}d\eta$ instead of $\overline{g(0)}$, which depends on the non-zero modulation of each mode k. In practice, $\phi_k^{(N)}(t)$ should be estimated by $\widetilde{q}_{\eta,f}^{[k,N]}(t,\eta)$ before computing this normalization. Regarding numerical implementation of FSSTn, the same procedures as introduced in Section 2.6.2 for FSST are used to compute $T_{N,f}^{g,\gamma}(t,\eta)$ from $V_f^g(t,\eta)$ by replacing $\widehat{\omega}_f(t,\eta)$ by $\widehat{\omega}_{\eta,f}^{[N]}(t,\eta)$, and to estimate the associated ridges for mode reconstruction.

3.2.3 Numerical Analysis of the Behavior of STFT-based SSTs and Comparisons

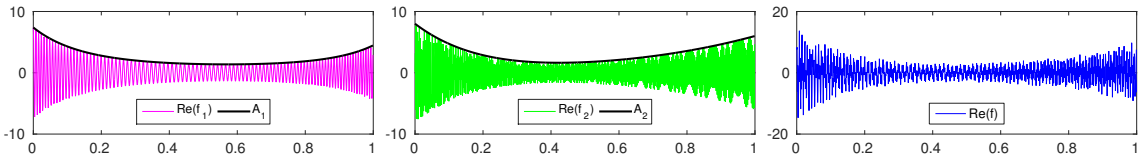
This section presents numerical investigations to illustrate the improvements brought by our new technique in comparison with the reassignment method (RM) or existing STFT-based SSTs (FSST and FSST2) on both simulated and real signals. The steps of these investigations are followed similarly to those given by WSST2 case in Section 3.1.4. Let us first consider a simulated MCS composed of two AM-FM components:

$$f(t) = f_1(t) + f_2(t) = A_1(t)e^{i2\pi\phi_1(t)} + A_2(t)e^{i2\pi\phi_2(t)},$$

with $A_k(t)$ and $\phi_k(t)$ defined on $[0,1]$, for $k = 1, 2$, by:

$$A_1(t) = \exp\left(2(1-t)^3 + t^4\right),\ A_2(t) = 1 + 5t^2 + 7(1-t)^6 \text{ and}$$
$$\phi_1(t) = 50t + 30t^3 - 20(1-t)^4,\ \phi_2(t) = 340t - 2\exp\left(-2(t-0.2)\right)\sin\left(14\pi(t-0.2)\right).$$

Note that f_1 is a polynomial chirp that satisfies Definition 3.20 with $L = N = 4$, while f_2 is a damped-sine function containing very strong nonlinear sinusoidal frequency modulations and high-order polynomial amplitude modulations. In our simulations, f is sampled at a rate $M = 1024$ Hz on $[0,1]$. The Matlab codes for synchrosqueezing transforms and the scripts leading to all figures of this chapter can be found https://github.com/phamduonghung/FSSTn. In Figures 3.8 (a) and (b), we display the real part of f_1 and f_2 together with their amplitudes, and, in Figure 3.8 (c), the real part of f.



(a) (b) (c)

Fig. 3.8 (a) and (b): real part of f_1 and f_2 with A_1 and A_2 superimposed; (c): real part of f.

The STFT of f is then computed with the L^1-normalized Gaussian window $g(t,\sigma_F) = \sigma_F^{-1}e^{-\pi\frac{t^2}{\sigma_F^2}}$, where σ_F is the optimal value that is determined by means of Rényi entropy technique introduced in Section 3.1.4.A. In Figure 3.9, we display the evolution of Rényi entropy ($\alpha = 3$) with respect to σ_F for the signal f introduced above for different noise levels (noise-free, -5, 0, 5 dB). This leads to an optimal value in each case (σ_{Fopt}), relatively stable with the noise level.

Having determined the optimal σ_F, we display, in the noise-free context, the STFT of f on the left of Figure 3.10. Then, on the right of this figure, close-ups of the STFT itself are depicted, along with reassigned versions of STFT either given by the RM or FSST and variants.

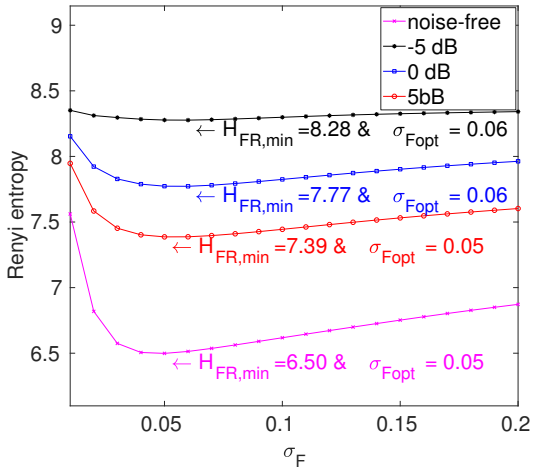


Fig. 3.9 Evolution of Rényi entropies ($H_{F,R}$) with respect to σ_F either in the noise-free, 5 dB, 0 dB or -5 dB cases.

Analyzing these close-ups, we remark that, as expected, FSST2 leads to a relatively sharp TFR for f_1, very similar to the one given by RM and much better than the one corresponding to FSST. However, all these methods fail to reassign the STFT of f_2 correctly, especially where the IF of that mode has a non-negligible curvature $\phi_2'''(t)$. In contrast, the TF reassignment of the STFT of f_2 provided by FSST3 or FSST4 is much sharper at these locations. Looking at what happens for mode f_1 also tells us that, FSST3 and FSST4 seem to behave very similarly to FSST2 or RM in terms of the sharpness of the representation. However, as we shall see later, the accuracy of the representation is dramatically improved by using one of the former two methods. Finally, note that since f_1 obeys Definition 3.20, the IF estimate used in FSST4 is exact for that mode which results in a perfect reassignment of the STFT.

In what follows, we introduce a quantitative comparison of all these techniques, for both noise-free and noisy cases, using the energy concentration and a measure of their accuracy by means of the Earth mover's distance (EMD) as in 3.1.4.D.

A. Evaluation of TF Concentration

Firstly, we depict, in Figure 3.11 (a), the normalized energy corresponding to the reassignment of the STFT of f_1 using the different techniques, with respect to the number of coefficients kept divided by the length of f_1. From this figure, it is hard to figure out the benefits of using FFST3 or FSST4 rather than the other two methods. The only thing one can check is that the energy is perfectly localized with FSST4 because f_1 obeys Definition 3.20. The results of the same computation obtained for mode f_2 are displayed in Figure 3.11 (b), showing that the normalized energy is much more concentrated using FSST4 than the other methods and that FSST3 also outperforms FSST2 and RM.

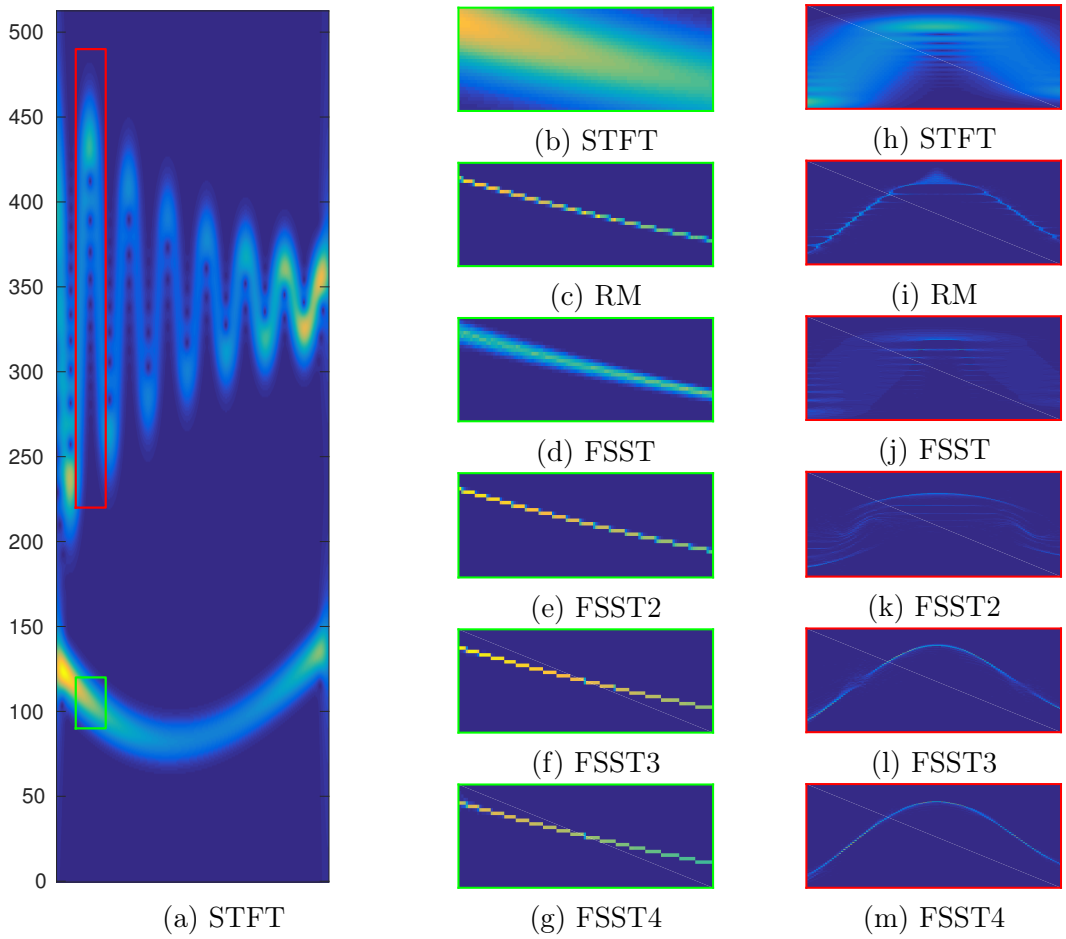


Fig. 3.10 (a): modulus of the STFT of f; (b): STFT of a small TF patch corresponding to mode f_1 (delimited by green segments) extracted from (a); (c) RM performed on the STFT shown in (b); from (d) to (g), same as (c) but using respectively FSST, FSST2, FSST3, FSST4; the same as (b) to (g) but for f_2.

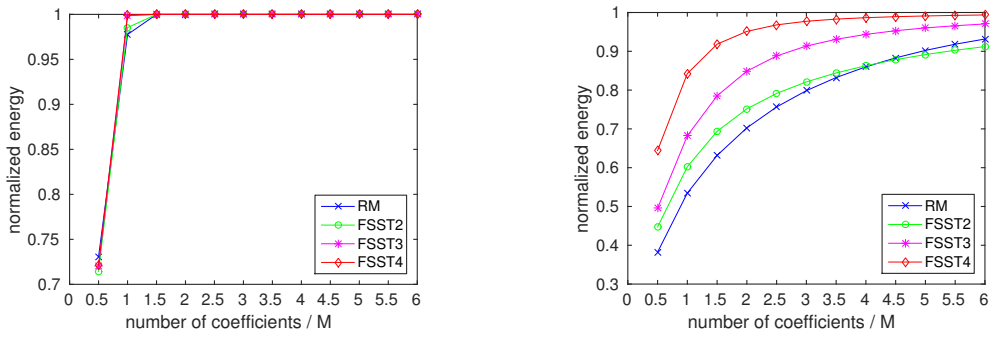


Fig. 3.11 (a) Normalized energy as a function of the number of sorted associated coefficients for f_1; (b): same as (a) but for f_2.

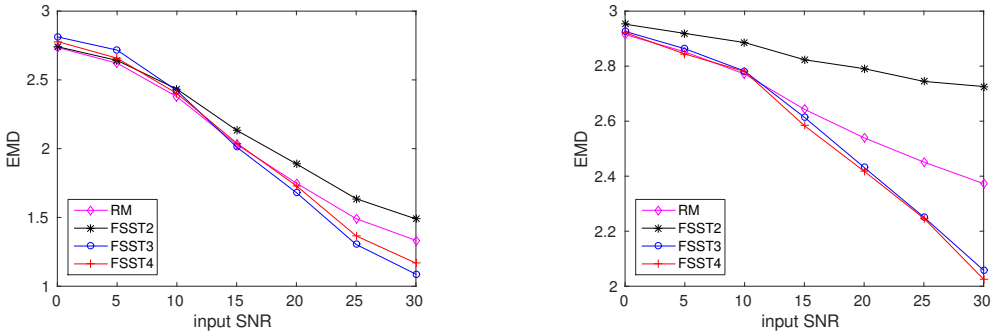


Fig. 3.12 (a): EMD corresponding to different TFRs of f_1 either given by RM, FSST2, FSST3 or FSST4; (b): same as (a) but for f_2.

Moreover, we study the performance of the TFRs in the presence of noise by HT technique as introduced in 3.1.4.D. Then, we apply the Earth mover's distance (EMD) on noisy TFRs to measure the accuracy of the reassigned transforms of the noisy signal. In Figures 3.12 (a) and (b), we display, respectively for f_1 and f_2, the evolution of EMD with respect to the noise level, for TFRs given either by FSST2, FSST3, FSST4 or RM. This study tells us that, at low noise level and for mode f_1, FSST3 and FSST4 are more accurate than the other studied methods. Note that this is something that could not be derived from the previous study on the normalized energy. The same investigations but for mode f_2 confirms the interest of using FSST3 or FSST4 to reassign the STFT of a mode with IF exhibiting the strong curvature. Note that the benefits of using the proposed new methods remain important even at the high noise level.

B. Evaluation of Mode Reconstruction Performance

As discussed above, using mode reconstruction is a good way to assess the performance of TFRs. Let us first briefly recall the procedure to retrieve f_k from the TFR of f given by the FSST of order N:

$$f_k(t) \approx \int_{\{\omega, |\omega - \varphi_k(t)| < d\}} T^{g,\gamma}_{N,f}(t,\omega) d\omega. \tag{3.39}$$

Note that $\varphi_k(t)$ is estimate of $\phi'_k(t)$ given by the ridge detector (see Section 2.6.1 for details on such a technique), and d is *compensation parameter* used to compensate for the inaccuracy of this estimation and also for the errors caused by approximating the IF by $\widehat{\omega}^{[N]}_{\eta,f}(t,\eta)$.

We first analyze the performance of the reconstruction procedure by considering the information on the ridge only, i.e. we take $d = 0$, and measuring output SNRs. In Table 3.3, we display this output SNR for modes f_1, f_2 and also for f, using either FSST2, FSST3 or FSST4 for mode reconstruction. The improvement brought by using FSST3 and FSST4 is clear and coherent with the previous study of the accuracy of the proposed new TFRs.

Table 3.3 Performance of mode reconstruction in the noise-free case

	FSST2	**FSST3**	**FSST4**
Mode f_1	17.8	25.7	**28.8**
Mode f_2	1.73	3.62	**6.87**
MCS f	3.57	5.57	**8.82**

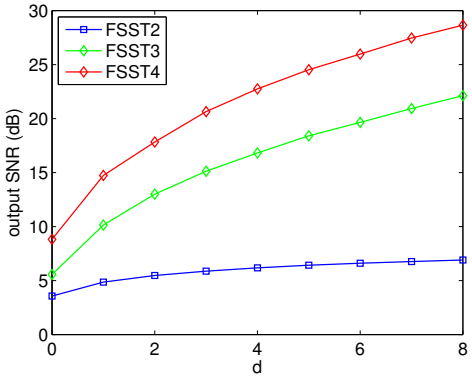


Fig. 3.13 Reconstruction accuracy measured in SNR, with respect to d, of the noise-free signal.

Parameter d also measures how well the TFR is concentrated around the detected ridges: if the former is well concentrated, even if one uses a small d, the reconstruction results should be satisfactory. In order to measure this, we display in Figure 3.13, the output SNR corresponding to the reconstruction of f when d varies, and when the TFR used for mode reconstruction is either FSST2, FSST3, and FSST4. From this figure and for all tested methods, it is clear that a larger d means a more accurate reconstruction of the signal. Nevertheless, the accuracy of the reconstruction using FSST2 seems to stagnate when some critical value for d is reached, which is not the case with the other two methods: the parameter d can only partly compensate for the inaccuracy of IF estimation. For that very reason, it is crucial to use the most accurate estimate as possible which again pleads in favor of FSST3 and FSST4.

3.3 Application of Synchrosqueezing Transforms to Gravitational-wave Signals

This section investigates the applicability of synchrosqueezing techniques including FSST2, our new ones WSST2 and FSST4 to the analysis of a transient gravitational-wave signal. The first observation of such a signal was made in September 2015 and announced by the (LIGO) and Virgo collaborations in February 2016, and was named **GW150914** [21]. It was detected by the LIGO detector in Hanford Washington that closely matches the predictions of Albert Einstein about general relativity for a gravitational wave emanating from the inward spiral

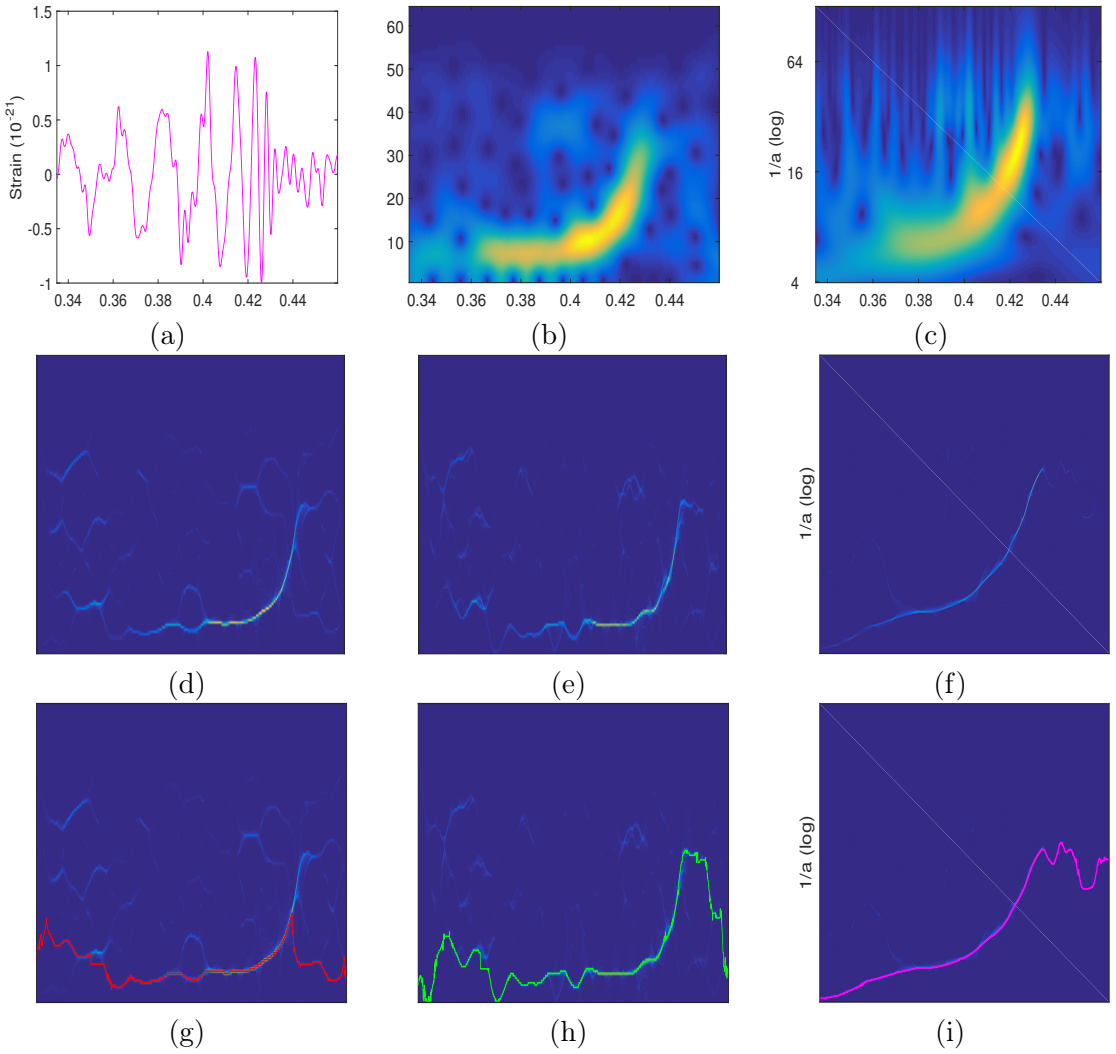


Fig. 3.14 Illustration of the TFRs of the gravitational-wave event **GW150914**, (a): observed Hanford signal; (b): STFT; (c): CWT; (d): FSST2; (e): FSST4; (f): WSST2; (g): the ridge estimated from FSST2 displayed in (d); (h): same as (g) but on FSST4; (i): same as (g) but on WSST2.

and merger of a pair of black holes and the subsequent *ringdown* of the single resulting black hole. In our simulations, we use a Gaussian window and the Morlet wavelet with respective optimal lengths computed by the Rényi entropy technique: $\sigma_F = 0.05$ and $\sigma_W = 1$.

We first display the gravitational-wave strain observed by the LIGO Hanford and the modulus of its STFT and CWT in the top-row panel of Figure 3.14, and then the reassigned transforms corresponding to FSST2, FSST4 and WSST2 in the middle-row panel of Figure 3.14, respectively. The sharpened representations provided by these synchrosqueezing techniques make the TF information more easily interpretable: as a matter of fact, the gravitational-wave signal consists of only one mode sweeping sharply upwards. However, the difference of the

performance of these reassigned transforms cannot be seen at this point. Moving on to the bottom line of Figures 3.14, we perform ridge detection on each of the TFRs given by FSST2, FSST4 and WSST2. We remark that compared with the others, FSST4 enables a better ridge detection of the three stages of the collision of two black-holes; especially the *ring down* one, which commences when the IF of the mode starts to decrease. This is associated with a sudden variation of the curvature of its IF which is better taken into account by FSST4. This thus demonstrates the interest of this new technique in practical applications.

3.4 Conclusion

In this chapter, we have introduced two important contributions to the synchrosqueezing techniques for the analysis of multicomponent signals. The first one is a novel second-order wavelet-based synchrosqueezing transform (WSST2) for analyzing signals made of strongly frequency-modulated modes based on the continuous wavelet transform. It simply consists of a refinement of the instantaneous frequency estimate, computed using a second-order expansion of the phase. After having revisited the case of first-order synchrosqueezing (WSST) while releasing the hypothesis of a wavelet compactly supported in the frequency domain, we demonstrated a novel approximation theorem involving the proposed new synchrosqueezing transform applied to multicomponent signals made of strongly modulated modes. In this regard, we put forward a novel reconstruction technique for the modes. Numerical experiments showed the benefits of taking into account frequency modulation for both representation and reconstruction purposes, and also the better performance of the second-order reassignment based on wavelet compared to that based on STFT. The second contribution of this chapter is the generalization of the existing STFT-based synchrosqueezing transforms by defining new synchrosqueezing operators based on the high-order amplitude and phase approximations. Such a generalization allows us to better handle a wide variety of multicomponent signals containing very strongly modulated AM-FM modes. The interest of the proposed new technique was also demonstrated through numerical experiments on simulated signals. Indeed, it successfully produces a TF picture more concentrated than other methods based on synchrosqueezing or reassignment, while allowing for a better invertibility of the TFR. Finally, a set of numerical simulations on the gravitational-wave signal was given to demonstrate the interest of the new generation technique, FSSTn.

Chapter 4

Denoising and Mode Reconstruction of Multicomponent Signals

Contents

In this chapter, we put forward two new contributions to the analysis of MCSs developed in the short time Fourier transform (STFT) framework. The first one is a new technique built on the second-order STFT-based synchrosqueezing transform (FSST2) and demodulation procedure, that we coin DFSST2 [6]. It enables an even sharper representation and a better mode reconstruction than with FSST2. In fact, a previous attempt consisting in signal demodulation before applying synchrosqueezing transform (SST) was presented in [50, 51], in which the demodulation relies upon the computation of the phase of the associated analytic signal (AS). However, it is well known that this phase cannot be related to the instantaneous frequency (IF) of the modes. In this regard, an iterative approach based on the local frequency extrema of the spectrogram to accurately estimate the IFs was proposed in [51], unfortunately, mode reconstruction was not discussed.

The second contribution lies in the fact that although SST and SST-based methods prove to be effective for enhancing the TFRs they are based on, none of which allows for mode retrieval of MCSs from their SST downsampled in time. Using the time redundancy of STFT we here show that the reconstruction of the modes of a noisy MCS can be performed with a better accuracy and fewer STFT coefficients than when STFT-based SSTs is considered. This work has been recently accomplished in [1].

For a better understanding of these contributions, we first introduce, in Section 4.1.1, some new analyses on the practical implementation of *ridge estimation* carried out on STFT-based SSTs (i.e FSSTn). Then, we move on to the definition of the demodulation procedure followed by the proposed reconstruction algorithm, called DFSST2, in Section 4.1.2. Numerical examples show the relevance of this new technique on both simulated and real data in Section 4.1.3. Then, we recall, in Section 4.2.1, the different signal reconstruction techniques from shifted downsampled STFT and then study how to use these for signal denoising, in Section 4.2.2, assuming the noise is Gaussian white. Exploiting these findings, we introduce, in Section 4.2.3, a novel algorithm, based on downsampled STFT, to retrieve the modes of an MCS. A comparison of this new algorithm with two synchrosqueezing techniques: DFSST2 and FSST2 computed from a full STFT (i.e. without downsampling), on both simulated and real data, in Section 4.2.4, concludes this chapter.

4.1 On Ridge Detection, Demodulation and Synchrosqueezing

Before delineating the *demodulation-FSST2-based technique* (DFSST2), the following section presents some new analyses on the practical implementation of ridge estimation.

4.1.1 Some New Analyses on Ridge Estimation

Any mode reconstruction techniques based on the STFT-based SSTs (FSSTn) requires an estimate of the ridges $(t, \phi_k'(t))$ and for that purpose, we introduced one of the most commonly used ridge detectors in Chapter 2. In this subsection, we detail this detector and then investigate the influence of all of the different parameters on the accuracy of ridge estimation in both noiseless and noisy contexts. To do so, we first show how to numerically implement the STFT and FSSTn.

A. On the Computation of the STFT, FSSTn and zero-padding

Let us consider a signal f being of finite length, typically defined on the interval $[0, T]$, and discretized into $f(\frac{nT}{N})_{n=0,\cdots,N-1}$. Assuming g is supported on $[-\frac{LT}{N}, \frac{LT}{N}]$, with $L < N/2$, STFT is then computed as follows:

$$\begin{aligned} V_f^g(t,\eta) &= \int_{\mathbb{R}} f(t+\tau)g(\tau)e^{-2i\pi\tau\eta}d\tau \\ &= \int_{-\frac{LT}{N}}^{\frac{LT}{N}} f(t+\tau)g(\tau)e^{-2i\pi\tau\eta}d\tau \\ &\approx \frac{T}{N}\sum_{n=-L}^{L} f(t+\frac{nT}{N})g(\frac{nT}{N})e^{-i2\pi\frac{nT}{N}\eta}, \end{aligned} \tag{4.1}$$

from which we infer that:

$$\begin{aligned} V_f^g\left(\frac{qT}{N}, \frac{p}{T}\right) &\approx \frac{T}{N}\sum_{n=-L}^{L} f\left(\frac{(q+n)T}{N}\right) g\left(\frac{nT}{N}\right) e^{-i2\pi\frac{np}{N}} \\ &= \frac{Te^{\frac{2i\pi pL}{N}}}{N}\sum_{n=0}^{2L} f\left(\frac{(q+(n-L))T}{N}\right) g\left(\frac{(n-L)T}{N}\right) e^{-i2\pi\frac{np}{N}} \\ &= e^{\frac{2i\pi pL}{N}}\sum_{n=0}^{2L}\Gamma(q,n)e^{-i2\pi\frac{np}{N}} := V_{\mathbf{d},f}^g\left(\frac{qT}{N}, \frac{p}{T}\right), \end{aligned} \tag{4.2}$$

where the subscript **d** stands for *discrete* value and the last sum is computed by means of an FFT. It is common to extend, for each q, the sequence $(\Gamma(q,n))_n$ 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:

$$\begin{aligned} V_f^g\left(\frac{qT}{N}, \frac{N}{N_f}\frac{p}{T}\right) &\approx \frac{T}{N}\sum_{n=-L}^{L} f\left(\frac{(q+n)T}{N}\right) g\left(\frac{nT}{N}\right) e^{-i2\pi\frac{np}{N_f}} \\ &= \frac{Te^{\frac{2i\pi pL}{N_f}}}{N}\sum_{n=0}^{2L} f\left(\frac{(q+(n-L))T}{N}\right) g\left(\frac{(n-L)T}{N}\right) e^{-i2\pi\frac{np}{N_f}} \\ &= e^{\frac{2i\pi pL}{N_f}}\sum_{n=0}^{2L}\Gamma(q,n)e^{-i2\pi\frac{np}{N_f}} := V_{\mathbf{d},f}^g\left(\frac{qT}{N}, \frac{N}{N_f}\frac{p}{T}\right). \end{aligned} \tag{4.3}$$

Since $V_{\mathbf{d},f}^g\left(\frac{qT}{N}, \frac{N}{N_f}\frac{p}{T}\right)$ is computed by means of an FFT, only the first half of the frequency set is meaningful. That is, $V_{\mathbf{d},f}^g$ is computed on the TF grid $\left\{0, \frac{T}{N}, \cdots, \frac{(N-1)T}{N}\right\} \times \left\{0, \frac{N}{N_f}\frac{1}{T}, \cdots, \frac{N}{N_f}\frac{N_f/2-1}{T}\right\}$.

The computation of FSST $T_{\mathbf{d},f}^{g,\gamma}$ from $V_{\mathbf{d},f}^g$ can then be performed as explained in Algorithm 1 (putting $t_q = \frac{qT}{N}$) [48].

Algorithm 1 STFT-based SSTs

for $q = 0$ to $N-1$ **do**
  for $k = 0$ to $N_f/2-1$ **do**
    $T_{\mathbf{d},f}^{g,\gamma}(t_q, \frac{N}{N_f}\frac{k}{T}) := 0$
  for $p = 0$ to $N_f/2-1$ **do**
    Compute $\widehat{\omega}_{\mathbf{d},f}(t_q, \frac{N}{N_f}\frac{p}{T}) = \frac{N}{N_f}\frac{p}{T} - \Im\mathrm{m}\left\{\frac{1}{2\pi}\frac{V_{\mathbf{d},f}^{g'}(t_q, \frac{N}{N_f}\frac{p}{T})}{V_{\mathbf{d},f}^{g}(t_q, \frac{N}{N_f}\frac{p}{T})}\right\}$.
    Put $k = \mathrm{round}(T\frac{N_f}{N}\widehat{\omega}_{\mathbf{d},f}(t_q, \frac{N}{N_f}\frac{p}{T}))$.
    Reassign $|V_{\mathbf{d},f}^g(t_q, \frac{N}{N_f}\frac{p}{T})| > \gamma$ as follows:

$$T_{\mathbf{d},f}^{g,\gamma}(t_q, \frac{N}{N_f}\frac{k}{T}) = T_{\mathbf{d},f}^{g,\gamma}(t_q, \frac{N}{N_f}\frac{k}{T}) + V_{\mathbf{d},f}^g(t_q, \frac{N}{N_f}\frac{p}{T}).$$

The same algorithm is applied to get other synchrosqueezing transforms FSSTn $T_{\mathbf{d},\mathbf{N},f}^{g,\gamma}$ from V_f^g, replacing $\widehat{\omega}_{\mathbf{d},f}$ by $\widehat{\omega}_{\mathbf{d},\mathbf{N},t,f}^{[2]}$, where $\mathbf{N}$ denotes order of synchrosqueezing transforms and we only study $\mathbf{N} = 1$ or $\mathbf{N} = 2$ in this chapter. The role of zero-padding is going to be further investigated in the sequel.

B. Ridge Extraction for FSSTn

As mentioned precisely, to compute an estimate of the ridge $(t, \phi'_k(t))$ assuming knowledge of the number of modes K, one needs compute a local minimum of the functional as follows:

$$E_f(\varphi_1, \cdots, \varphi_K) \quad = \quad \sum_{k=1}^{K} - \int_{\mathbb{R}} |T^{g,\gamma}_{\mathbf{N},f}(t, \varphi_k(t))|^2 dt + \int_{\mathbb{R}} \lambda \varphi'_k(t)^2 dt + \beta \varphi''_k(t)^2 dt, \quad (4.4)$$

where λ, β are the regularization parameters. However, as presented, formula (4.4) does not offer any algorithmic means to compute the ridges. Furthermore, $T^{g,\gamma}_{\mathbf{N},f}$ is approximated on a TF grid to get $T^{g,\gamma}_{\mathbf{d},\mathbf{N},f}$. Inspired by the above minimization problem, we derive Algorithm 2, for that purpose.

Algorithm 2 Ridge Extraction on FSSTn

Pick $q \in \{0, \cdots, N-1\}$
for $k = 1$ to K **do**
1. $t_q = \frac{qT}{N}$.
2. Define $p_{q,k} = \underset{l}{\operatorname{argmax}} |T^{g,\gamma}_{\mathbf{d},\mathbf{N},f}(t_q, \frac{N}{N_f}\frac{l}{T})|$.
3. Define $I_{q,k} = \{\max(0, p_{q,k} - \frac{N_f}{N} T\Delta), \cdots, \min(p_{q,k} + \frac{N_f}{N} T\Delta, N_f/2 - 1)\}$.
4. Define $p_{q-1,k} = \underset{l \in I_{q,k}}{\operatorname{argmax}} |T^{g,\gamma}_{\mathbf{d},\mathbf{N},f}(t_{q-1}, \frac{N}{N_f}\frac{l}{T})|$.
5. Define $p_{q+1,k} = \underset{p \in I_{q,k}}{\operatorname{argmax}} |T^{g,\gamma}_{\mathbf{d},\mathbf{N},f}(t_{q+1}, \frac{N}{N_f}\frac{p}{T})|^2 - \lambda(p - p_{q,k})^2 - \beta(p - 2p_{q,k} + p_{q-1,k})^2$
6. Iterate forward in time.
7. Iterate steps 2-6 backward from time t_q.
8. $(\varphi_k(t_q))_q = (\frac{N}{N_f}\frac{p_{q,k}}{T})_q$
9. $T^{g,\gamma}_{\mathbf{d},\mathbf{N},f} = T^{g,\gamma}_{\mathbf{d},\mathbf{N},f} \setminus \bigcup_q [\varphi_k(t_q) - \Delta, \varphi_k(t_q) + \Delta]$

In this algorithm, the ridges are extracted one by one as follows. One first picks a random time and finds the local maximum of the modulus of $T^{g,\gamma}_{\mathbf{d},\mathbf{N},f}$ at that time (steps 1 and 2, this point is supposed to be on a ridge), then, since mode modulation is bounded by $\frac{N_f}{N} T\Delta$, one finds neighboring points on the ridge using the fact that the maximum displacement of a point on a ridge between two successive locations is $\pm\frac{N_f}{N} T\Delta$ (which means $\pm\frac{N_f}{N} T\Delta \times N$ in terms of indices, steps 3–5), and finally iterates backward and forward (steps 6 and 7). One then removes the detected ridge and corresponding TF neighborhood (steps 8 and 9, the ridges are associated with K sufficiently separated local maxima at each time instant), and proceeds with the remaining ridges.

Note that to improve the robustness of the procedure, several random initializations are required, leading to the detection of many different ridge sets $(\varphi_k)_{k=1,\cdots,K}$, and the one

retained as the output corresponds to the one maximizing:

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

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

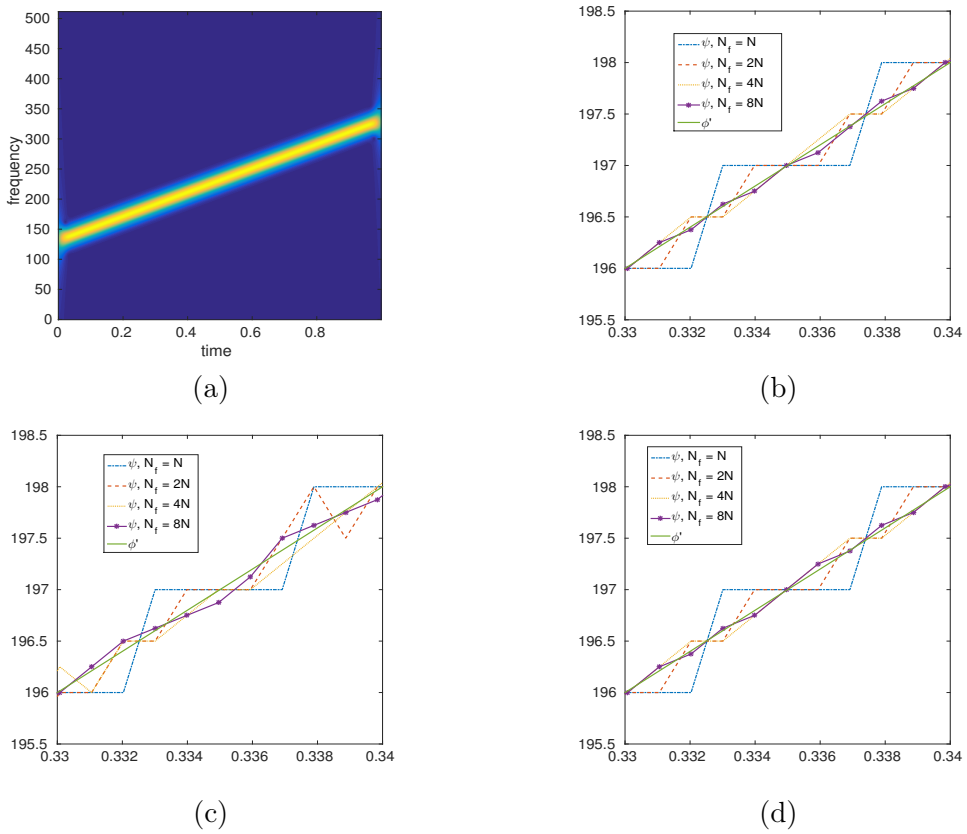


Fig. 4.1 (a): STFT of a linear chirp; (b): Zoom on the ridges estimated for the signal whose STFT is displayed in (a), using the ridge detector based on STFT with no regularization parameters and for various frequency resolution (corresponding to different values for N_f); (c): Similar to (b) but for the ridge detector based on FSST; (d): Similar to (b) but for the ridge detector based on FSST2.

C. 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 $\phi'(t_n)$:

$$\varphi(t_n)=\operatorname*{argmax}_{\eta}|V^g_f(t_n,\eta)|^2, \tag{4.5}$$

which was studied in [93], for a noisy version of the signal $f(t) = A(t)e^{2i\pi\phi(t)}$, i.e. $f_\zeta(t_n) = f(t_n) + \zeta(t_n)$, where ζ is a white Gaussian noise with variance σ_ζ^2. Selecting $\Delta\varphi(t_n) = \varphi(t_n) - \phi'(t_n)$ 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 [93] that:

$$\begin{aligned} Bias\{\Delta\varphi(t_n)\} &\sim_{N\to+\infty} \sum_{k=1}^{+\infty} \frac{2\pi\phi^{(2k+1)}(t_n)\sigma^{2k}}{2^k k!} \\ Var\{\Delta\varphi(t_n)\} &\sim_{N\to+\infty} \frac{\sigma_\zeta^2}{8\sqrt{\pi}|A(t_n)|^2}\left[1 + \frac{\sigma_\zeta^2 T}{2N\sqrt{\pi}\sigma|A(t_n)|^2}\right]\frac{T}{N\sigma^3}. \end{aligned} \tag{4.6}$$

These results are interesting but do not consider the fact that $|V_f^g(t_n,\eta)|$ is only available on a discrete frequency grid, since it is computed using an FFT. More precisely, formula (4.5) 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, we perform ridge detection on the linear chirp whose STFT is depicted in Figure 4.1 (a) (with $N_f = N$ for that plot, setting λ and β to zero). To study such a simple case is very essential to fully understand the behavior of the ridge detector introduced above.

Ridge estimates are shown in Figures 4.1 (b) to (d), where the detection is performed for various frequency resolutions (associated with different values for N_f) and for different TF representations. What is striking is the stair-case effect visible in each of the Figures 4.1 (b) to (d), when the frequency resolution of the sampling grid is insufficiently fine. This arises because, as already noted, TF representations are evaluated at frequencies $\left(\frac{N}{N_f}\frac{p}{T}\right)_{p=0,\cdots,N_f/2-1}$, and the IF estimate reflects this discretization. This can be quantified by measuring the *mean square error* (MSE) between the estimated ridge and the ground truth:

$$MSE(\varphi) = \sqrt{\frac{1}{N-1}\sum_{n=0}^{N-1}(\phi'(t_n) - \varphi(t_n))^2}, \tag{4.7}$$

when the frequency resolution varies. In order to better understand 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 4.2, first row (they correspond to a linear chirp, a polynomial chirp and a mode with sinusoidal phase).

The results, displayed in Figure 4.2 (d), show that, in a noise-free context, when STFT, FSST or FSST2 is 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 FSST2 (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 FSST2, but not

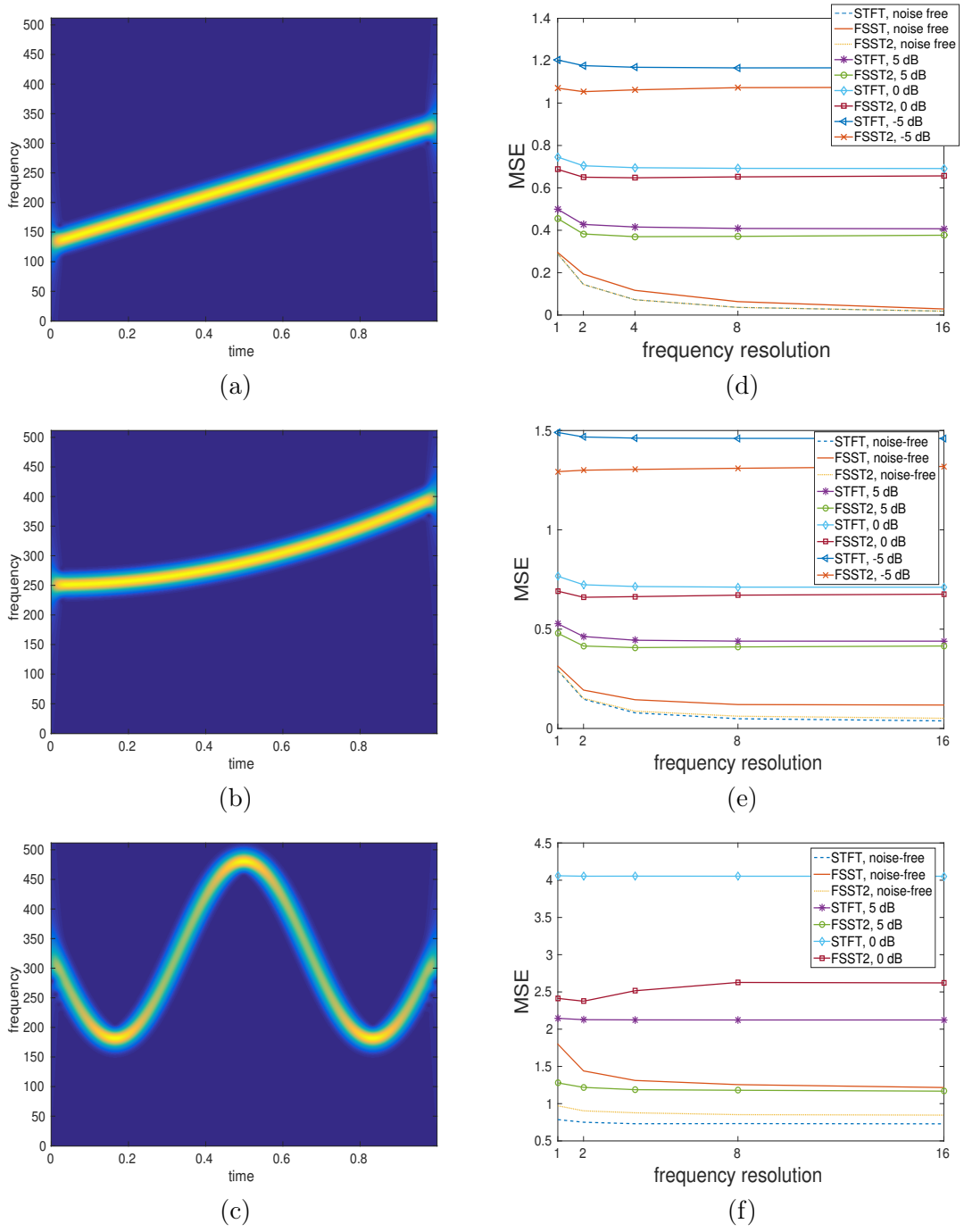


Fig. 4.2 Influence of zero-padding on different types of chirps with different noise levels: (a): STFT of a linear chirp; (b): STFT of a polynomial chirp; (c): STFT of a mode with sinusoidal phase; (d): MSE associated with the ridge detection for the linear chirp displayed in (a), for various frequency resolution (k in abscissa means $N_f = kN$), different TFRs 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).

significantly so. In contrast, since FSST relies 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 TFR. For this reason, we do not consider it in the simulations that follow. Finally, we remark, that in the noise-free context for the linear and polynomial chirps of Figure 4.2 (a) and (b), MSE when using STFT or FSST2 decreases when the frequency resolution across the sampling grid is increased. However, this is no longer true with the signal of Figure 4.2 (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.

Now, we would like to understand what happens in noisy situations, therefore we perform the ridge detection on the linear, sinusoidal and polynomial chirps with an SNR equal to 5, 0 or -5 dB. The results are depicted in Figures 4.2 from (d) to (f) (for the latter type of signals, and whatever the TFR used, the ridge detector does not perform well at -5 dB, therefore the results are not depicted). It is clear from Figures 4.2 (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 FSST2 rather than to STFT is always better: the ridge detection operates on a much sharper TFR 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.

D. Influence of Regularization Parameters

Taking into account the study carried out in the previous section, the ridge detector applied to either STFT or FSST2 both lead to good results when no regularization is used, even though, as illustrated in Figure 4.2 (second row), to perform ridge detection on FSST2 rather than STFT is always better in noisy situations.

We now study the behavior of the ridge detector applied to STFT or FSST2 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 to FSST2 (see Figure 4.3): the reassignment technique enables a more robust ridge detection even at high noise level because it corresponds to a sharper TFR. 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 are shown in Figure 4.3 were done 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

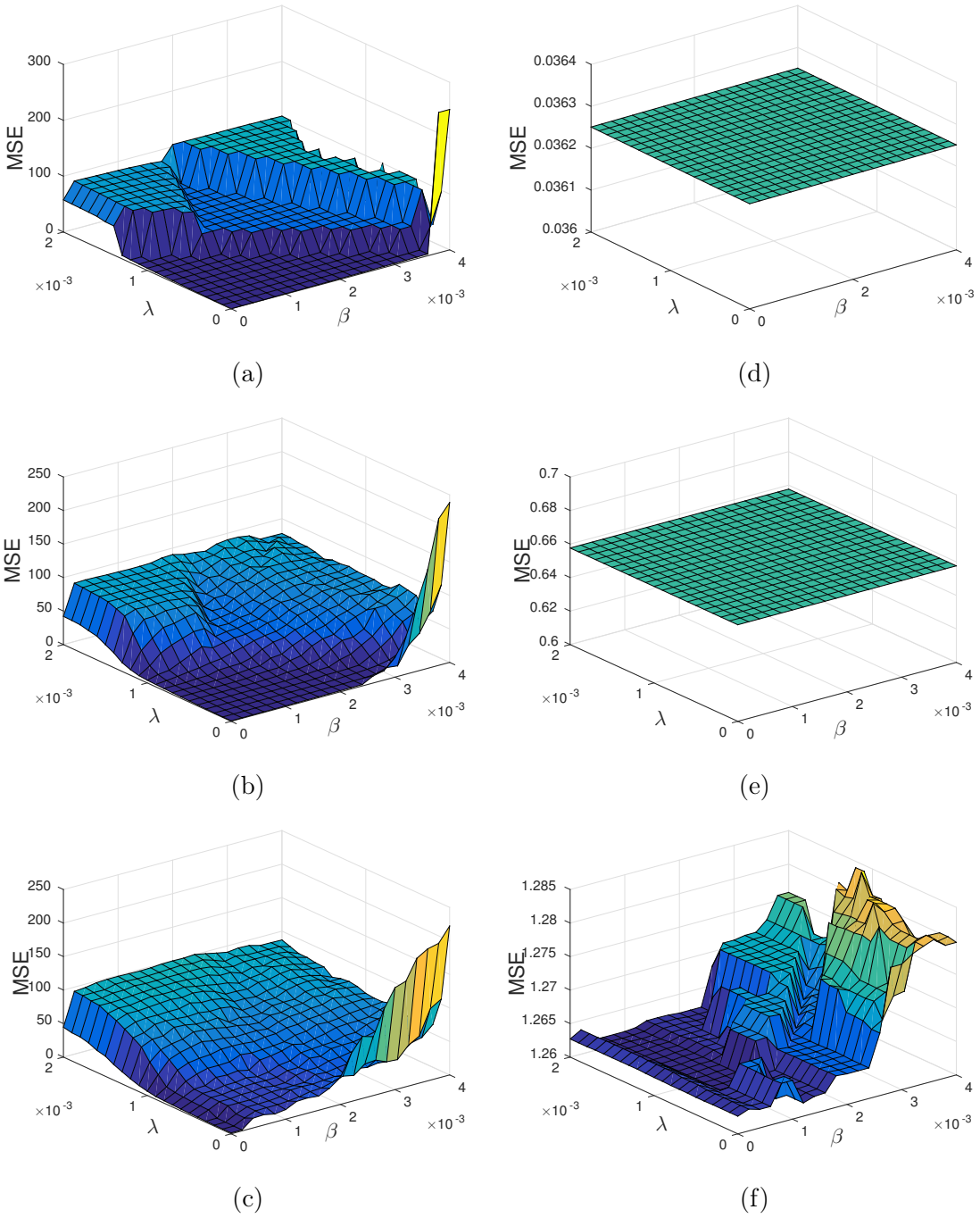


Fig. 4.3 Influence of regularization parameters on different types of chirps with different noise levels: (a): MSE corresponding to the ridge estimation for the linear chirp of Figure 4.2 (a) with V_f^g as TFR (noise-free case, $N_f = 8N$); (b): same as (a) but at a 0 dB noise level; (c): same as (a) but at a -5 dB noise level; (d): MSE corresponding to the ridge estimation for the linear chirp of Figure 4.2 (a) with $T_{2,f}^{g,\gamma}$ as TFR (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.

were conducted on the polynomial or sinusoidal chirps, as soon the algorithm detects the ridge.

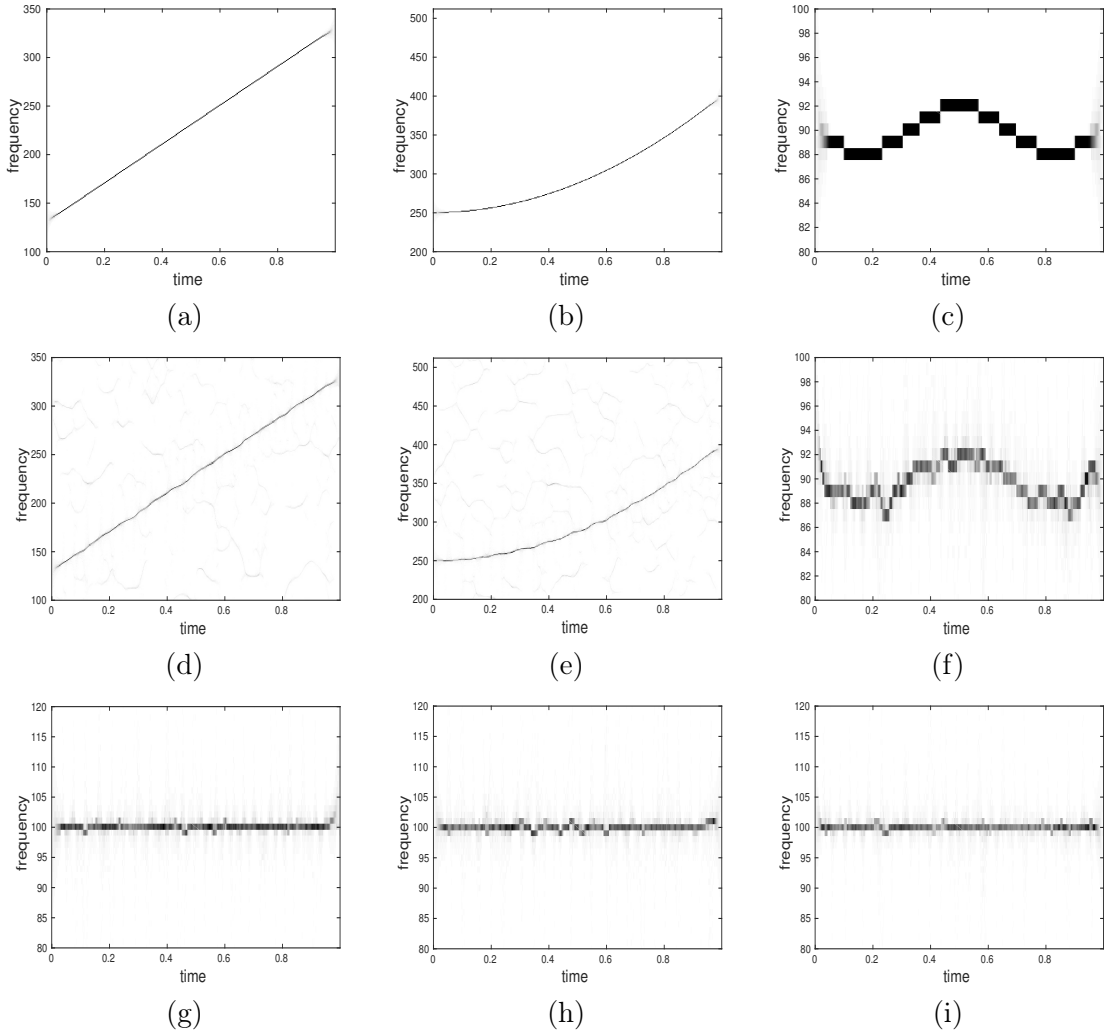


Fig. 4.4 Demodulation procedure on mono-component signals: (a): FSST2 of a linear chirp (noise-free case); (b): FSST2 of a polynomial chirp (noise-free case); (c): FSST2 of a mode with sinusoidal phase (noise-free case); (d): FSST2 of a linear chirp (noise level 0 dB case); (e): FSST2 of a polynomial chirp (0 dB case); (f): FSST2 of a mode with sinusoidal phase (0 dB case); (g): demodulated signal (d); (h): demodulated signal (e); (i): demodulated signal (f).

4.1.2 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 subsequently used to extract the corresponding demodulated mode. Inverting the demodulation operator, we 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 [94, 95].

It is noteworthy 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:

$$V_f^{g,v}(t,\eta) \quad = \quad \int_{\mathbb{R}} f(t)\overline{g(t-\tau)}e^{-2i\pi v(t)}e^{-2i\pi\eta(t-\tau)}dt. \tag{4.8}$$

This kind of approach has also been used in [96, 97] 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 [51]. 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 $\varphi(t)$ is computed. For the case of a linear chirp, i.e. $\phi(t) = at + bt^2$, $\varphi(t)$ approximates $a + 2bt$. So, by multiplying $f(t)$ by $e^{-2i\pi(\varphi(t)t/2)}$, and if the IF 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 $e^{-i2\pi(\int_0^t \varphi(x)dx - \varphi_0 t)}$, where φ_0 is some positive constant frequency, is a better choice since, no assumption is made about ϕ. Indeed, by considering the signal:

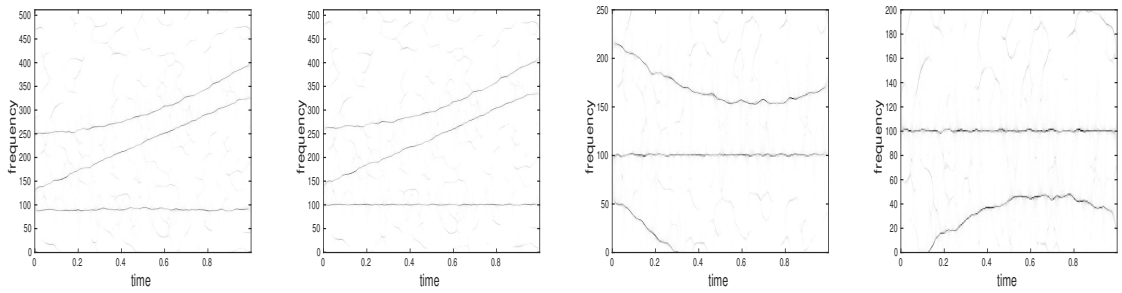
$$f_D(t) = f(t)e^{-i2\pi(\int_0^t \varphi(x)dx - \varphi_0 t)}, \tag{4.9}$$

one should get a signal with constant frequency φ_0.

An illustration of this is shown in Figure 4.4, for three different types of mode, where φ_0 is equal to 100 Hz. In that figure, we display the FSST2 of the considered modes in the noise-free (*resp.* 0 dB) case, in the first (*resp.* second) row. In the bottom row of that figure, we display the FSST2 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 MCS case. We will illustrate this by adopting a signal consisting of the three different modes, displayed in Figure 4.5 (a). Then by applying Algorithm 2 to the FSST2 computed with $N_f = 8N$, we obtain the estimates $(\varphi_1, \varphi_2, \varphi_3)$, which are subsequently used to compute three demodulated signals, as follows:

$$f_{D,k}(t) = f(t)e^{-i2\pi(\int_0^t \varphi_k(x)dx - \varphi_0 t)}, \quad k = 1, 2, 3. \tag{4.10}$$



(a) (b) (c) (d)

Fig. 4.5 Demodulation procedure on an MCS (a): three modes signal (noise level 0 dB); (b): FSST2 of $f_{D,1}$; (c): FSST2 of $f_{D,2}$; (d): FSST2 of $f_{D,3}$.

The FSST2 of the three signals $(f_{D,k})_{k=1,2,3}$ are shown in Figures 4.5 (b) to (d), where the SNR in the original signal equals 0 dB. It is interesting to note that, in $f_{D,k}$, only the kth mode is demodulated.

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 new algorithm for mode extraction that we coin *demodulation-FSST2-based technique* (DFSST2), can then be summarized as follows:

Algorithm 3 *Demodulation-FSST2-based Technique* (DFSST2)

Compute the ridge estimates $(\varphi_1, \cdots, \varphi_K)$ with Algorithm 2 applied to FSST2.
for $k = 1$ to K **do**

1. Compute $f_{D,k}(t) = f(t)e^{-i2\pi(\int_0^t \varphi_k(x)dx - \varphi_0 t)}$.
2. From $T^{g,\gamma}_{2,f_{D,k}}$, extract the ridge $\varphi_{D,k}$ corresponding to mode k of $f_{D,k}$, by considering single ridge detection in the frequency range $[\varphi_0 - \Delta, \varphi_0 + \Delta]$.
3. Reconstruct the kth mode of $f_{D,k}$ and then multiply it by the inverse of demodulation operator to recover f_k: $f_k(t) \approx \left(\int_{|\omega - \varphi_{D,k}(t)|<d} T^{g,\gamma}_{2,f_{D,k}}(t,\omega)d\omega\right) e^{i2\pi(\int_0^t \varphi_k(x)dx - \varphi_0 t)}$.

Note here that the TFR used to compute the ridge of $f_{D,k}$ and then mode k could alternatively be $T^{g,\gamma}_f$ since the mode sought is demodulated, there is no need to take into account the modulation at this stage. Indeed, the kth mode of signal $f_{D,k}$ should be a purely harmonic signal at frequency φ_0 (see Figures 4.5 (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 $T^{g,\gamma}_{2,f_{D,k}}$, $N_f = N$ is used because the mode k, extracted at step 3 of Algorithm 3, is a purely harmonic one.

4.1.3 Evaluation of the Performance of DFSST2

This section introduces some numerical experiments to demonstrate the superiority offered by the foregoing technique DFSST2 over the original FSST2 on both simulated (noise-free and noisy) and real signals in term of the quality of the reconstructed modes. For that purpose, let us consider test noise-free signals displayed in Figures 4.6 (a) and 4.7 (a). The window used to build the TFRs is Gaussian whose length is optimized by means of Rényi entropy technique introduced in Section 3.1.4.A. Note also that an arbitrary threshold $\gamma = \gamma_0 = 0.001$ is used for noise-free cases, while the hard-thresholding (HT) technique is applied to noisy cases.

Moreover, since the ridge computation is influenced by the frequency resolution, we investigate the impact of N_f used in ridge computation on mode reconstruction. To assess how the ridge detection impacts mode reconstruction, we also compute the mode reconstruction assuming the IFs of the modes are known. Finally, it was shown in Section 3.1.4.E. that the *compensation parameter* d plays a vital role in mode reconstruction, which is shown again in this section.

A. Noise-free Signal Case

The results for the noise-free signals are depicted from Figures 4.6 (b) to (d), in which we represent the output SNR when using DFSST2, and when the frequency resolution used in the ridge detection varies. 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 DFSST2"). We first remark that, as expected, for $N_f = 8N$ the results given by DFSST2 are very close to the ones obtained by optimal DFSST2, as illustrated in Figure 4.6 (d). We also display the results obtained by reconstructing the modes using directly FSST2: whatever the value of d the reconstruction is better when using DFSST2. Additionally, 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 FSST2 is displayed in Figure 4.7 (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 DFSST2 provides far better results than those given by direct FSST2. In this case, however, and since the modes are more modulated than those of Figure 4.6 (a), the impact of N_f on ridge computation and then mode reconstruction is less important.

B. Noisy Signal Case

In order to investigate the sensitivity to noise of DFSST2, we consider again the previous types of signals. Also, ridge computation leads to good reconstruction when $N_f = 8N$ in the noise-free case, so this value is retained in the simulations that follow.

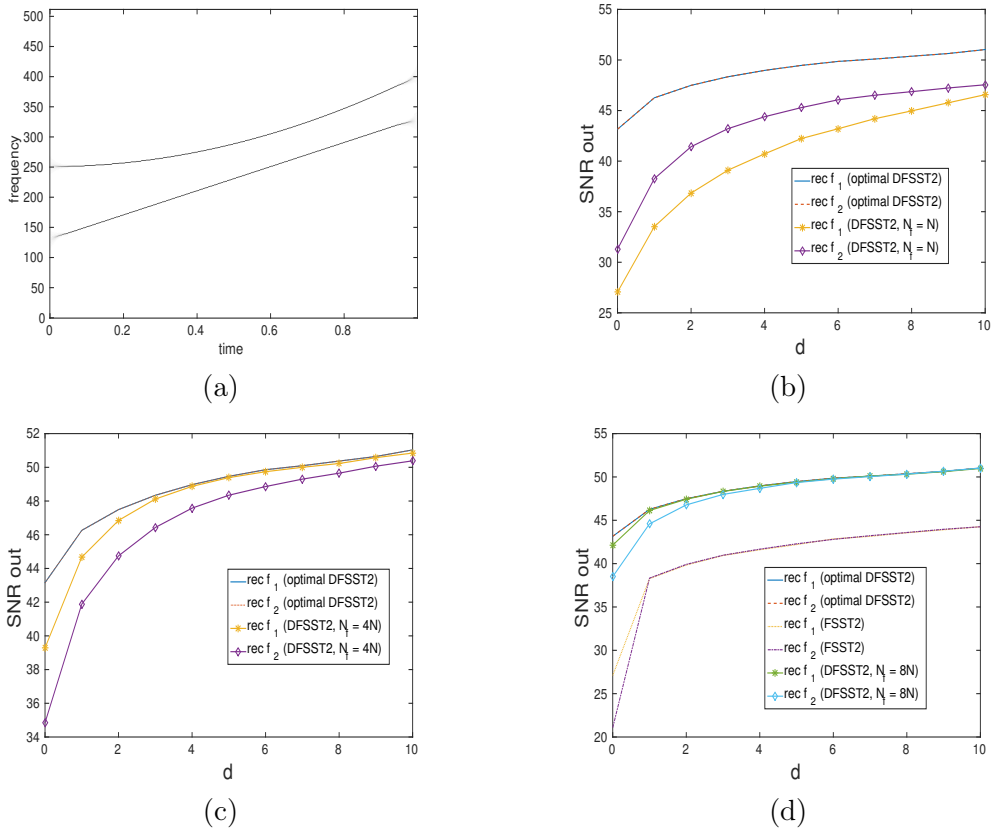


Fig. 4.6 DFSST2 application to a noise-free two-component signal: (a): its FSST2; (b): mode reconstruction ("rec f_i" corresponds to reconstructed mode f_i) when using DFSST2 computed with $N_f = N$, together with the reconstruction when the IFs of the mode are assumed to be known ("optimal DFSST2" in 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.

The results displayed in the first row of Figure 4.8, represent the output SNR associated with mode reconstruction when $d = 0$ or $d = 5$, respectively for the first and second mode of Figure 4.6 (a), with respect to global input SNR. We note that the following based on these observations: whatever the noise level, the mode reconstruction is improved by using DFSST2 rather than direct FSST2; the discrepancy in terms of reconstruction performance between the two types of techniques increases when the noise level is decreasing; the gain of DFSST2 is not that important because FSST2 is optimized for linear chirps.

Switching to the study of the signal of Figure 4.7 (a), the benefit of using the demodulation procedure is much clearer: when a mode is very different from a linear chirp, DFSST2 greatly improves 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 using DFSST2, we obtain a more concentrated representation since, for a given d, the reconstruction is always better using demodulation than without.

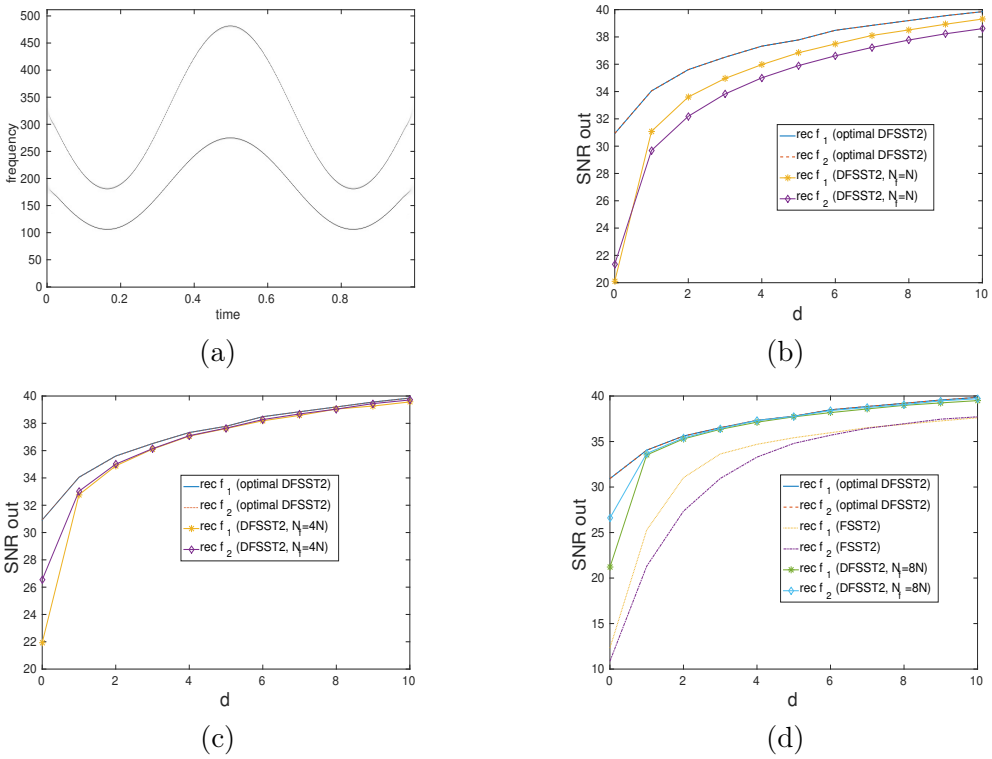


Fig. 4.7 DFSST2 application to another noise-free two-component signal: (a): its FSST2; (b): mode reconstruction ("rec f_i"corresponds to reconstructed mode f_i) when using DFSST2 computed with $N_f = N$, along with the reconstruction when the IFs of the modes are assumed to be known ("optimal DFSST2" 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.

C. Application of DFSST2 to Real Data and limitations

Here we consider the reconstruction of a bat echolocation signal whose FSST2 is shown in Figure 4.9 (a). Assuming the number of modes is 3, 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 FSST2 (the extracted ridges are also depicted in the figure) and then compute the different modes by either using FSST2 or DFSST2. In this regard, we study the influence of parameter d and frequency resolution on the reconstruction.

Since the signal studied is real, and as DFSST2 applies to a 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 $N_1 = 2^{\lfloor log_2(N) \rfloor + 1}$ with $\lfloor . \rfloor$ being the floor function, 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

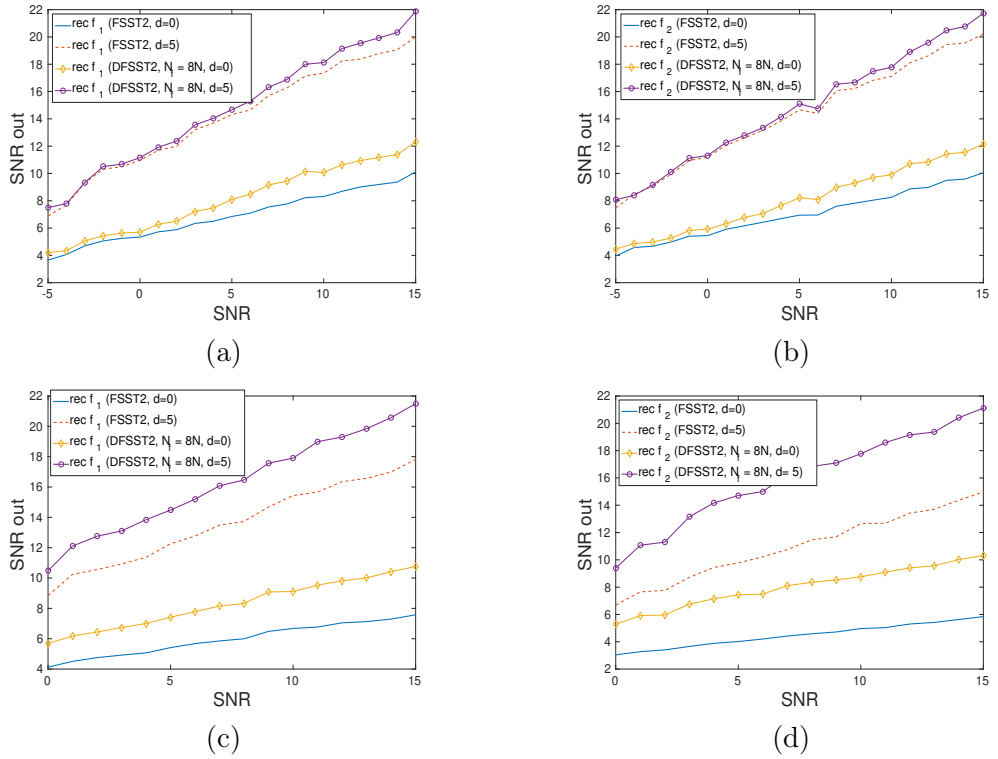


Fig. 4.8 Illustration of reconstruction procedure for a noisy signal case: (a): SNR after reconstruction for mode f_1 of the signal whose FSST2 is depicted in Figure 4.6 (a) using either FSST2 or DFSST2, and for $d = 0$ or $d = 5$ in both cases; (b): same as (a) but for mode f_2 of the same signal; (c): SNR after reconstruction for mode f_1 of the signal whose FSST2 is depicted in Figure 4.7 (a) using either FSST2 or DFSST2, 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.

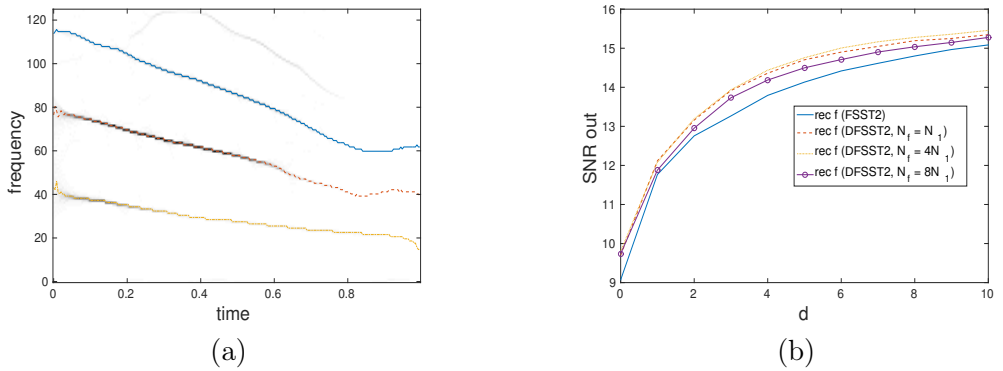


Fig. 4.9 Illustration of DFSST2 on a real signal: (a): FSST2 of a bat echolocation call along with the corresponding ridges, (b): reconstructed signal based on either DFSST2 or FSST2 and assuming the number of modes equals 3.

output SNR associated with the reconstruction of the signal by summing the first three

modes. The results depicted in Figure 4.9 (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. In the next chapter, we propose an interesting method, called *adaptive contour representation computation*, which enables to tackle this problem.

4.2 On Downsampled Short-Time Fourier Transform (STFT)

This second section investigates the denoising and mode reconstruction of MCSs from their downsampled STFT. This work has been recently done in [1]. First, we concisely recall three signal reconstruction techniques based on STFTs in the discrete setting and then show how to use them for signal denoising.

4.2.1 Signal Reconstruction from Downsampled STFT

Let first consider a signal f in $l_2(\mathbb{Z})$, and a discrete real window g also in $l_2(\mathbb{Z})$, the STFT downsampled by a factor of R (a positive integer) and shifted by an integer parameter p ($0 \leq p < R$) is defined for each ω by:

$$V_f^g(mR+p,\omega) = \sum_{n\in\mathbb{Z}} f[n]g[n-mR-p]e^{-i2\pi\omega(n-mR-p)}. \tag{4.11}$$

If one assumes the signal is of length L and supported on $\{0,\cdots,L-1\}$, the filter g supported on $\{-M,\cdots,M\}$ such that $L_g := 2M+1 \leq N \leq L/2$, where N is the number of frequency bins, one may write:

$$\begin{aligned} V_f^g(mR+p,\frac{k}{N}) &= \sum_{n\in\mathbb{Z}} f[n]g[n-mR-p]e^{-i2\pi\frac{k(n-mR-p)}{N}} \\ &= \sum_{n=-M}^{M} f[mR+p+n]g[n]e^{-i2\pi\frac{kn}{N}} \\ &= \sum_{n=0}^{2M} f[mR+p+n-M]g[n-M]e^{-i2\pi\frac{k(n-M)}{N}}. \end{aligned}$$

Since g is null on $\{M+1,\cdots,N-1-M\}$, STFT can be rewritten as:

$$V_f^g(mR+p,\frac{k}{N})e^{-i2\pi\frac{kM}{N}} = \sum_{n=0}^{N-1} f[mR+p+n-M]g[n-M]e^{-i2\pi\frac{kn}{N}}.$$

Using the properties of the *discrete Fourier transform* (DFT), one obtains, for any $n \in \{0,\cdots,N-1\}$:

$$f[mR+p+n-M]g[n-M] = \frac{1}{N}\sum_{k=0}^{N-1} V_f^g(mR+p,\frac{k}{N})e^{i2\pi\frac{k(n-M)}{N}}. \tag{4.12}$$

When g is non-zero on $\{-p,\cdots,R-1-p\}$, one gets, for any $q \in \{0,\cdots,R-1\}$:

$$f[mR+q] = \frac{1}{g[q-p]N}\sum_{k=0}^{N-1} V_f^g(mR+p,\frac{k}{N})e^{i2\pi\frac{k(q-p)}{N}}. \tag{4.13}$$

This reconstruction technique will be referred to as the *first reconstruction technique* (RT1) in the sequel. Since p varies in $\{0,\cdots,R-1\}$, if one wants to use the same filter g whatever p, the former has to be non-zero on $\{-R+1,\cdots,R-1\}$.

Alternatively, using (4.12) and assuming $\sum_{m\in\mathbb{Z}} g[n-mR]^2 = 1$ for all n, we may write:

$$f[n] = \sum_{k=0}^{N-1}\sum_{m\in\mathbb{Z}} V_f^g(mR+p,\frac{k}{N})g[n-mR-p]\frac{e^{i2\pi\frac{k(n-mR-P)}{N}}}{N}. \tag{4.14}$$

The proof of (4.14) is available in Appendix C.1. Furthermore, when such a filter is used, one also has energy conservation, namely:

$$\sum_{n\in\mathbb{Z}} f[n]^2 = \sum_{m\in\mathbb{Z}}\frac{1}{N}\sum_{k=0}^{N-1}|V_f^g(mR+p,\frac{k}{N})|^2, \tag{4.15}$$

the proof being also available in Appendix C.2. Note that the reconstruction can still be achieved by making the appropriate renormalization as follows:

$$f[n] = \frac{\sum_{k=0}^{N-1}\sum_{m\in\mathbb{Z}} V_f^g(mR+p,\frac{k}{N})g[n-mR-p]\frac{e^{i2\pi\frac{k(n-mR-p)}{N}}}{N}}{\sum_{m\in\mathbb{Z}} g[n-mR-p]^2}, \tag{4.16}$$

provided $\sum_{m\in\mathbb{Z}} g[n-mR]^2 \neq 0$ for all n. This type of reconstruction technique is denoted *second reconstruction technique* (RT2) in the sequel.

Remark 4.1. It is important to remark that such a reconstruction formula was already studied in [98] but for infinite signals and assuming a continuous frequency. Here, the signal is finite and the frequencies are discretized. Also, in our context, the number of frequency bins is much smaller than the signal length for the method to be tractable for long signals.

Alternatively, assuming this time $\sum_{m\in\mathbb{Z}} g[n-mR] = 1$ for all n, one similarly has:

$$f[n] = \sum_{k=0}^{N-1} \sum_{m\in\mathbb{Z}} V_f^g(mR+p, \frac{k}{N}) \frac{e^{i2\pi \frac{k(n-mR-p)}{N}}}{N}, \tag{4.17}$$

the proof being available in Appendix C.3. Note that, as soon as $\sum_{m\in\mathbb{Z}} g[n-mR] \neq 0$ for all n, one can still reconstruct f through:

$$f[n] = \frac{\sum_{k=0}^{N-1} \sum_{m\in\mathbb{Z}} V_f^g(mR+p, \frac{k}{N}) \frac{e^{i2\pi \frac{k(n-mR-p)}{N}}}{N}}{\sum_{m\in\mathbb{Z}} g[n-mR-p]}. \tag{4.18}$$

This reconstruction formula is referred to as *third reconstruction technique* (RT3) in the sequel.

Remark 4.2. Typical filters g that can be used in RT2 and RT3 are those satisfying the *constant overlap add* property [99] with hop-size R ($COLA(R)$): $\sum_{m\in\mathbb{Z}} g[n-mR] = 1$, for all n.

For example, the rectangular window defined on $\{0, \cdots, R-1\}$ is $COLA(R)$. Other popular examples of $COLA(R)$ windows is the Bartlett window: $w[n] = w_r[n]\left[1 - \frac{|n|}{R}\right]$, for $n \in \{-R, \cdots, R\}$ and where w_r is the rectangular window on $\{-R, \cdots, R\}$, the Hann window defined, for $n \in \{-R, \cdots, R\}$, by: $w[n] = w_r[n]\left[\frac{1}{2} + \frac{1}{2}\cos(\frac{\pi n}{R})\right]$, or the Hamming window, $w[n] = w_r[n]\left[0.54 + (1-0.54)\cos(\frac{\pi n}{R})\right]$, for this time $n \in \{-R, \cdots, R-1\}$, and putting $w[R] = 0$. One has, however, to bear in mind that $COLA(R)$ condition is only sufficient for signal reconstruction from downsampled STFT.

4.2.2 Signal Denoising from Downsampled STFT

To remove noise from a signal, we propose to use an adaptation of hard-thresholding (HT) technique introduced in Section 3.1.4.D. to the downsampled STFT context. Assuming the variance of the Gaussian noise σ_ζ^2 is known, HT for downsampled STFT is defined as:

$$\overline{V}_{f_\zeta}^g(m, \frac{k}{N}) = \begin{cases} V_{f_\zeta}^g(m, \frac{k}{N}), & \text{if } |V_{f_\zeta}^g(m, \frac{k}{N})| \geq \gamma_F = 3\sigma_\zeta \|g\|_2 \\ 0 & \text{otherwise,} \end{cases} \tag{4.19}$$

where m denotes time index. Then, denoised signals are obtained using one of the three reconstruction formulae replacing V_f^g by $\overline{V}_{f_\zeta}^g$. For the sake of simplicity, we will also denote these techniques by RT1, RT2 and RT3, and f_r a reconstructed signal using one of these methods.

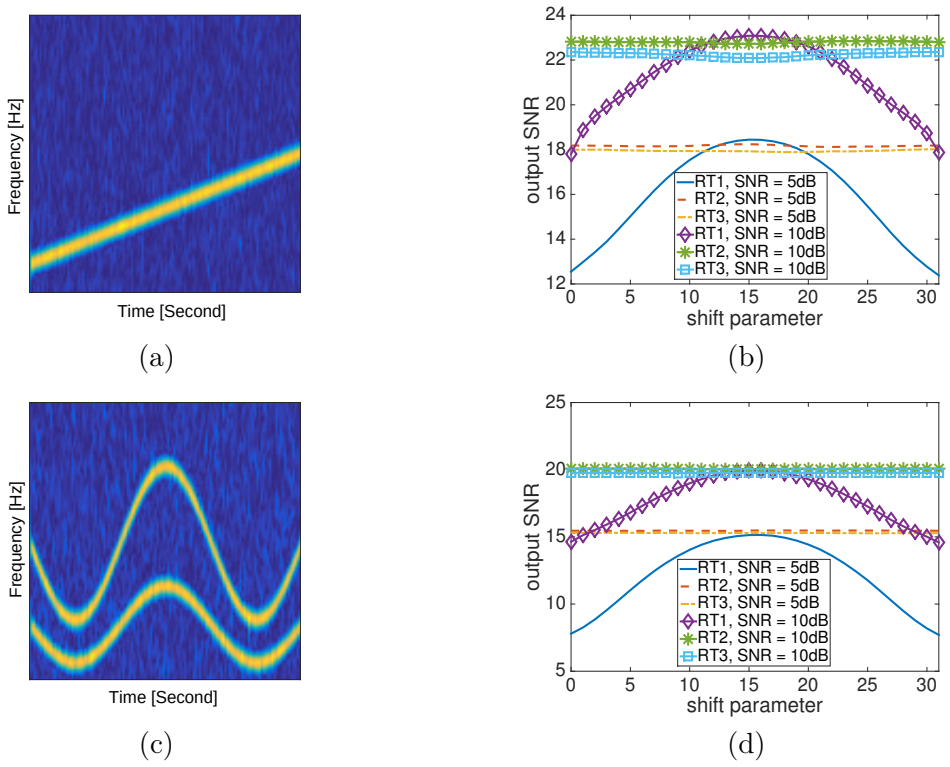


Fig. 4.10 (a): STFT of a linear chirp; (b): output SNR corresponding to the reconstruction of signal associated with STFT displayed in (a), when the shift parameter varies, for different input SNR and for RT1, RT2, and RT3; (c): STFT of an MCS; (d): output SNR corresponding to the reconstruction of signal associated with STFT displayed in (c), when the shift parameter varies, for different input SNRs and for RT1, RT2, and RT3.

In order to assess the quality of the reconstruction procedure using a shifted downsampled STFT, we remark that a good reconstruction technique should be very slightly sensitive to the shift parameter p. Pointing out that the higher the SNR the lower the l_2 error $\|f - f_r\|_2$, we may write, for RT1:

$$\sum_{q=0}^{R-1} |f[mR+q] - f_r[mR+q]|^2$$
$$= \sum_{q=0}^{R-1} \frac{1}{g[q-p]^2 N^2} \Big| \sum_{k=0}^{N-1} \left(V_f^g(mR+p, \frac{k}{N}) - \overline{V}_{f_\zeta}^g(mR+p, \frac{k}{N}) \right) e^{i2\pi \frac{k(q-p)}{N}} \Big|^2.$$

Defining $K(p,q) := |\sum_{k=0}^{N-1}\left(V_f^g(mR+p,\frac{k}{N}) - \overline{V}_{f_\zeta}^g(mR+p,\frac{k}{N})\right)e^{i2\pi\frac{k(q-p)}{N}}|^2$, and assuming it is independent of p and q, then the l_2 error is minimal when $\sum_{q=0}^{R-1}\frac{1}{g[q-p]^2}$ is minimal for some p. For low-pass, positive, and symmetric windows (like Hamming or Bartlett windows, for instance), this corresponds to considering the central taps of filter g, i.e. $p = \lfloor R/2 \rfloor$. Regarding method RT2, energy conservation enables us to write:

$$\sum_{q\in\mathbb{Z}}|f[q]-f_r[q]|^2 = \sum_{m\in\mathbb{Z}}\frac{1}{N}\sum_{k=0}^{N-1}|V_f^g(mR+p,\frac{k}{N}) - \overline{V}_{f_\zeta}^g(mR+p,\frac{k}{N})|^2, \tag{4.20}$$

which is independent of p, provided $K(p,p)$ is independent of p.

To illustrate this, we consider a linear chirp defined for t in $\{0,...,L-1\}$ by $f(t) = e^{2i\pi(0.05t+(0.10/L)t^2)}$, with $L = 4096$. The maximal reduced frequency is 0.25, and the Shannon-Nyquist sampling theory tells us that one needs at least $L/2$ samples to have perfect reconstruction, meaning the sampling period is half the critical one (i.e. the signal is oversampled by a factor of 2). Then, we compute the STFT of the signal using the Hamming window with length $L_g = 63$ and consider a number $N = 256$ of frequency bins, and $R = 32$ (the largest downsampling factor compatible with the three reconstruction techniques since $M = 32$). For $p = 0$, we obtain the representation of Figure 4.10 (a), corresponding to a 256×128 matrix and, changing p does not change the matrix size. So, if one compares with the critical sampling, requiring 2048 values for f, the downsampling factor actually equals 16. Then, for different noise levels, shift parameter in $\{0,\cdots,R-1\}$ and input SNRs, we display in Figure 4.10 (b), the output SNR corresponding to the reconstruction of the linear chirp using either RT1, RT2, or RT3.

For RT1, having numerically checked that $K(p,q)$ is almost independent of p and q, we notice that, as expected, best reconstruction results are obtained when $p = \lfloor R/2 \rfloor$. On the contrary, we check that RT2 is insensitive to p, as a consequence of energy conservation. The behavior of RT3 is very similar to that of RT2 but is more difficult to analyze. As we are interested in reconstructing the signal from its downsampled STFT for any given shift parameter, RT1 should not be used. However, such a reconstruction should be borne in mind since it is used in the synchrosqueezing techniques discussed at the end of this paper.

It is also worth mentioning that when the SNR is equal to 10 dB, and no downsampling is applied, the output SNRs for the linear chirp and with either of the three methods RT1, RT2, and RT3 are 23.39, 23.27 and 22.87 dB respectively: looking at Figure 4.10 (b), we notice that the denoising performance is not sensitive to downsampling factor since similar results are obtained using downsampled STFT in RT2, RT3 and also for RT1, provided the shift parameter is optimally chosen for the latter.

This behavior is also observable when applying the reconstruction techniques to a more general signal, as the MCS whose downsampled STFT is depicted in Figure 4.10 (c), for which

the output SNR is displayed in Figure 4.10 (d). In this example, the reduced frequencies are $\frac{400}{L} + \frac{30}{L}\cos(3\pi\frac{t}{L})$ and $\frac{1000}{L} + \frac{60}{L}\cos(3\pi\frac{t}{L})$ with $L = 4096$, meaning that the maximal reduced frequency equals 0.38: the signal is over-sampled by a factor less than 2 (so with $R = 32$, the downsampling factor compared with the critical Nyquist sampling is even bigger than in the previous example). As previously, to use downsampled STFT instead of the full transform does not result in a significant performance loss. In the following section, we are going to exploit these findings to propose a novel algorithm for the reconstruction and denoising of the modes of an MCS.

4.2.3 Mode Recontruction from Downsampled STFT of Noisy MCS

This section presents a new technique derived from the HT strategy introduced in the previous section for denoising and mode reconstruction of an MCS from its downsampled STFT. Bear in mind that mode reconstruction in the TF plane is carried out by integrating STFT along the frequency axis and in the vicinity of detected ridges approximating the IFs ($\phi_l's$) of the modes. Therefore, we first explain how ridges are detected and then detail how mode retrieval is performed.

Algorithm 4 Ridge extraction from downsampled STFT

Pick $p \in \{0, \cdots, R-1\}$
Define $Z_{p,R} = \{mR + p, m \in \lfloor \frac{L-1}{R} \rfloor\}$
$\Delta_f = R \times M \times N$
for $l = 1$ to K **do**
1. Pick $t_m = mR + p \in Z_{p,R}$.
2. $p_{m,l} := \underset{k}{\operatorname{argmax}} |V^g_{f_\zeta}(t_m, \frac{k}{N})|$.
3. $I_{m,l} := [p_{m,l} - \Delta_f, p_{m,l} + \Delta_f] \bigcap \{0, ..., N-1\}$.
4. $p_{m-1,l} := \underset{k \in I_{m,l}}{\operatorname{argmax}} |V^g_{f_\zeta}(t_{m-1}, \frac{k}{N})|$, iterate backward.
5. $p_{m+1,l} := \underset{k \in I_{m,l}}{\operatorname{argmax}} |V^g_{f_\zeta}(t_{m+1}, \frac{k}{N})|$, iterate forward.
6. For $t_m \in Z_{p,R}$, set $V^g_{f_\zeta}(t_m, k/N) := 0$, for $k \in [p_{m,l} - \Delta N, p_{m,l} + \Delta N]$

Sort $(p_{m,l})_l$ in increasing order to obtain $(\tilde{p}_{m,l})_l$
$(\varphi_l[t_m])_m = (\frac{\tilde{p}_{m,l}}{N})_m$

A. Ridge Extraction from Noisy Downsampled STFT

In order to compute an estimate ($\varphi_l[n]$s) of the ridges $(n, \phi_l'[n])$ in the downsampled STFT context, we propose an adaption of Algorithm 2 (see Section 4.1.1.B.) in which the downsampling factor R taking into account. As mentioned in 4.1.1.C. that to use regularization terms in the energy functional results in less accurate ridge estimation. Therefore, we consider the

following IF estimate:

$$\varphi[mR+p] = \underset{0\leq k\leq N-1}{\operatorname{argmax}} |V^g_{f_\zeta}(mR+p, \frac{k}{N})|, \tag{4.21}$$

making m vary, where f_ζ will be a noisy version of f. From this, a procedure used to extract the ridges of an MCS from its downsampled STFT is summarized in Algorithm 4.

B. Mode Reconstruction

Based on RT2 and HT technique for downsampled STFT introduced in Section 4.2.2, we define a novel technique for the retrieval of the modes of an MCS (to consider RT3 for mode reconstruction instead would not change the results that much). Let us focus on the reconstruction of the l-th mode, assuming φ_l is known. For that purpose, we recall that a first-order approximation of STFT close to the lth ridge and in the continuous time setting, reads [47]:

$$V^g_{f_\zeta,c}(t,\omega) \approx A_l(t)\hat{g}(\omega - \phi_l'(t)), \tag{4.22}$$

If the window g is the Gaussian window $\sigma^{-1/2}e^{-\pi\frac{x^2}{\sigma^2}}$ and considering the approximation of ϕ_l' by φ_l, and adapting (4.22) to our discrete time and frequency settings, we get at time $t_m := mR+p$:

$$V^g_{f_\zeta}(t_m, \frac{k}{N}) \approx |V^g_{f_\zeta}(t_m, \varphi_l[t_m])|e^{-\pi\sigma^2(\frac{k}{N}-\varphi_l[t_m])^2},$$

for $\frac{k}{N}$ sufficiently close to $\varphi_l[t_m]$. If, as previously, the noise variance is σ_ζ^2, the coefficients associated with mode l one selects at each time t_m, following the procedure detailed in Section 4.2.2, correspond to the indices k satisfying:

$$|V^g_{f_\zeta}(t_m, \varphi_l[t_m])|e^{-\pi\sigma^2(\frac{k}{N}-\varphi_l[t_m])^2} \geq \gamma_F = 3\sigma_\zeta\|g\|_2,$$

namely the indices inside the interval $J_{l,m} := [N\varphi_l[t_m] - \Gamma_{l,m}, N\varphi_l[t_m] + \Gamma_{l,m}]$, with

$$\Gamma_{l,m} = \frac{N}{\sigma}\sqrt{-\frac{1}{\pi}\log\left(\frac{3\sigma_\zeta\|g\|_2}{|V^g_{f_\zeta}(t_m,\varphi_k[t_m])|}\right)}.$$

Knowing interval $J_{l,m}$, mode f_l is reconstructed through:

$$f_{r,l}[n] := \sum_{m\in\mathbb{Z}}\sum_{k\in J_{l,m}} V^g_{f_\zeta}(t_m, \frac{k}{N})g[n-t_m]\frac{e^{i2\pi\frac{k(n-t_m)}{N}}}{N}. \tag{4.23}$$

This method will be denoted by M_1 in the sequel. Doing so, one selects the portion of the signal above the noise level in the vicinity of each ridge. However, such a simple expression

for $J_{l,m}$ exists only because one assumes the window is Gaussian. Also, the width of the interval $J_{l,m}$ only depends on the magnitude of STFT at local maxima location, but not on the frequency modulation of the corresponding modes. Therefore, an alternative to compute the frequency interval for mode reconstruction is to define $\eta_{l,m}^{[1]}$ and $\eta_{l,m}^{[2]}$:

$$\eta_{l,m}^{[1]} = \underset{k}{\operatorname{argmax}} \left\{ \frac{k}{N} < \varphi_l[t_m], |V_{f_\zeta}^g(t_m, \frac{k}{N})| < 3\sigma_\zeta \|g\|_2 \right\}$$
$$\eta_{l,m}^{[2]} = \underset{k}{\operatorname{argmin}} \left\{ \frac{k}{N} > \varphi_l[t_m], |V_{f_\zeta}^g(t_m, \frac{k}{N})| < 3\sigma_\zeta \|g\|_2 \right\},$$

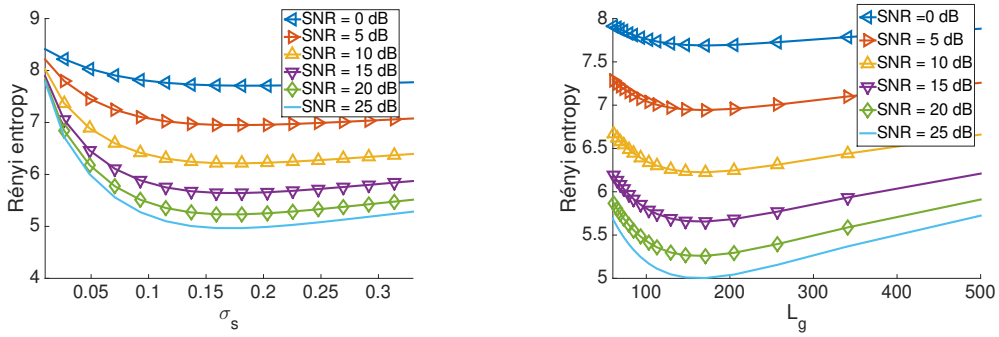
and then set :

$$J_{l,m} := \left[N\varphi_l[t_m] - \eta_{l,m}^{[1]}, N\varphi_l[t_m] + \eta_{l,m}^{[2]} \right]. \tag{4.24}$$

In such a case, the interval is no longer symmetric with respect to the maximum amplitude location and adapts to frequency modulation. This method will be denoted by M_2 in the sequel. The comparison between methods M_1 and M_2 will be conducted in the following section.

4.2.4 Numerical Applications

In this section, we first investigate the quality of mode reconstruction with respect to the type of reconstruction method, i.e. M_1 or M_2, type of filter, downsampling factor and shift parameter. Then, we compare the proposed mode reconstruction methods with other ones based on STFT-based SSTs. In this regard, a particular emphasis is put on the amount of information needed to perform mode reconstruction.

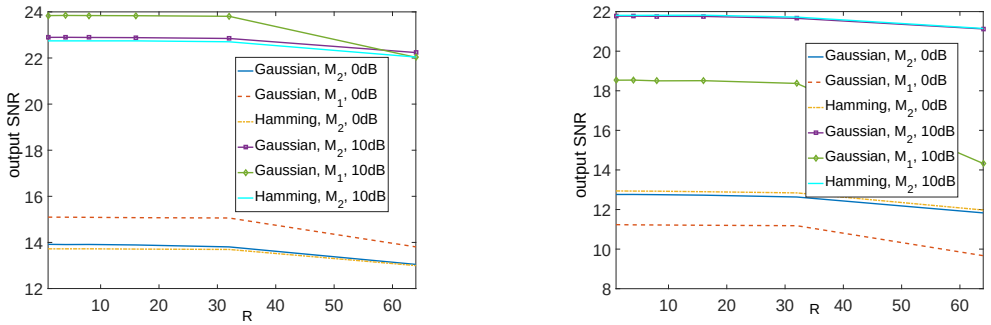


(a) (b)

Fig. 4.11 (a): Computation of Rényi entropy with respect to σ when the Gaussian window is used to compute STFT, for different noise level ; (b): same as (a) but when the Hamming window is used instead

In our simulations, we use both the Gaussian and Hamming windows whose optimal lengths σ and L_g are determined by the Rényi entropy technique introduced in Section 3.1.4.A.

In Figure 4.11 (a) and (b), we display the Rényi entropy of the STFT of the signal associated with Figure 4.10 (c), for various input SNRs, and with respect to the filter length parameter, for the Gaussian and Hamming windows respectively. Looking at these figures, the minimum entropy clearly puts forward an optimal window length which varies very little with the noise level. In these simulations, we took $N = 512$ and then the optimal window parameter σ equals 0.15 and the optimal L_g equals 161 for the Hamming window.



(a) (b)

Fig. 4.12 (a): output SNR associated with the reconstruction of mode f_1 of the signal associated with Figure 4.10 (c), for different Rs, different optimized filters, for either reconstruction technique M_1 or M_2, and when the input SNR is 0 dB or 10 dB; (b): same as (a) but for mode f_2

A. Sensitivity of Mode Reconstruction to Filter Choice, Downsampling Factor, and Shift Parameter

We here study the sensitivity of mode reconstruction techniques M_1 and M_2, to the choice of downsampling factor, shift parameter, and filter g. Our experiments are again based on noisy versions of the signal associated with Figure 4.10 (c), with $N = 512$, and the optimal filters (Gaussian and Hamming) being chosen as explained previously. Taken into account the optimal length associated with each of these filters, the largest downsampling factor one can consider is $R = 128$, but to take $R \leq 64$ is sufficiently informative on the behaviors of the proposed techniques. In order to measure the performance of the different techniques, we compute the output SNRs associated with the reconstruction of each mode, i.e. $SNR(f_l, f_{r,l})$, where the output SNR is defined in (2.71) and $f_{r,l}$ in (4.23).

As for the whole signal reconstruction procedure studied in Section 4.2.2, the output SNR, when considering mode reconstruction, is almost insensitive to the shift parameter. Therefore, for the sake of simplicity, we omit this parameter in the simulations and consider $p = 0$. The results depicted in Figure 4.12 (a), for an input SNR equal to 0 dB or 10dB, show that the quality of reconstruction for f_1 worsens when the downsampling factor increases, whatever the type of reconstruction method used, but the performance loss is not significant at least

for M_2. On that example, M_1 behaves slightly better than M_2 but the discrepancy between the results provided by the two methods remains within a range of 1 dB. Finally, to use the Hamming or the Gaussian window makes very little difference.

If one carries out the same study on mode f_2, which is much more modulated than f_1, the importance of taking into account the modulation in the reconstruction procedure transpires. The gain obtained by using M_2 rather than M_1 is at least 2 dB at the high noise level, and the benefits of using the former technique are even greater when the noise level is low. Based on these results, we will use M_2 in the comparison with SSTs that follows.

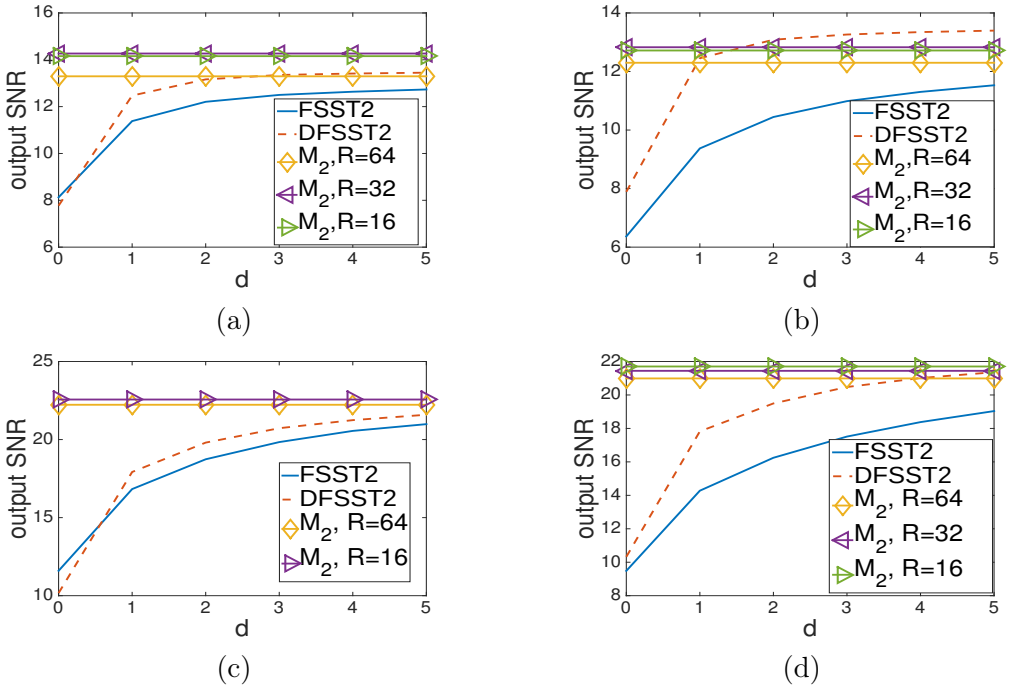


Fig. 4.13 (a): output SNR associated with the reconstruction of mode f_1 of the signal associated with Figure 4.10 (c), using different FSST2 techniques and with respect to parameter d, or M_2 with optimized Hamming filter and different downsampling factors R (SNR 0 dB); (b): same as (a) but for mode f_2; (c): same as (a) but with SNR 10 dB; (d): same as (b) but with SNR 10 dB.

B. Comparison with SST Methods

This section presents a comparison in terms of the quality of reconstructed modes between M_2 and SST methods including FSST2 and DFSST2 previously introduced. At first glance, since the information is reassigned by two latter techniques in the vicinity of the ridge associated with one mode, for a given time, mode reconstruction involves much fewer coefficients than with the former. Nevertheless, we have just proven that we could still perform mode reconstruction from a downsampled version of STFT which is not the case with FSST2 and DFSST2. So,

even though FSST2 leads to a much sharper representation than STFT, if one is interested in signal reconstruction using the smallest amount of information as possible, we are going to show that the reconstruction based on STFT performs better than FSST2 and DFSST2.

To quantify this, we compute the mode reconstruction performance for the signal associated with Figure 4.10 (c). We investigate the behavior of FSST2 and DFSST2 with respect to parameter d, along with mode reconstruction by M_2, using the optimized Hamming filter and when the downsampling parameter R varies. The output SNRs are displayed in Figures 4.13 from (a) to (d), and correspond to input SNRs 0 or 10 dB (the results are averaged over 5 realizations). As the reconstruction with M_2 does not depend on d, in these figures, the mode reconstruction results with this technique are displayed as constant functions. As already noticed, M_2 is not much sensitive to the downsampling factor when the latter goes from $R = 16$ to $R = 64$, for both f_1 and f_2.

Analyzing the results regarding FSST2 and DFSST2, we remark that the DFSST2 results in better mode reconstruction than when using FSST2. Now comparing M_2 with them, M_2 leads better results when only the information on the ridge is considered for FSST2 ($d = 0$), but when d increases, typically $d = 5$, mode reconstruction performed with DFSST2 seems to perform similarly to M_2 in most cases.

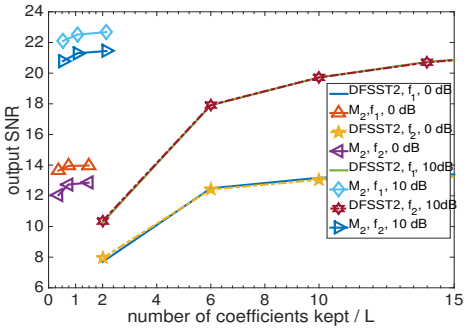


Fig. 4.14 Output SNR associated with the reconstruction of the signal associated with Figure 4.10 (c) with respect to the number of coefficients kept divided by the signal length (input SNR 0 or 10 dB)

However, there is one key issue hidden in the just computed output SNRs, which is the number of coefficients used to perform the reconstruction with each method. Indeed, for a fair comparison, we compute the number of coefficients required by the different reconstruction procedures. If the reconstruction procedure is performed on DFSST2 or FSST2, the total number of coefficients to be stored is $N_{FSST2} = K(2d+1)L$. On the contrary, the number of coefficients used by M_2 corresponds to:

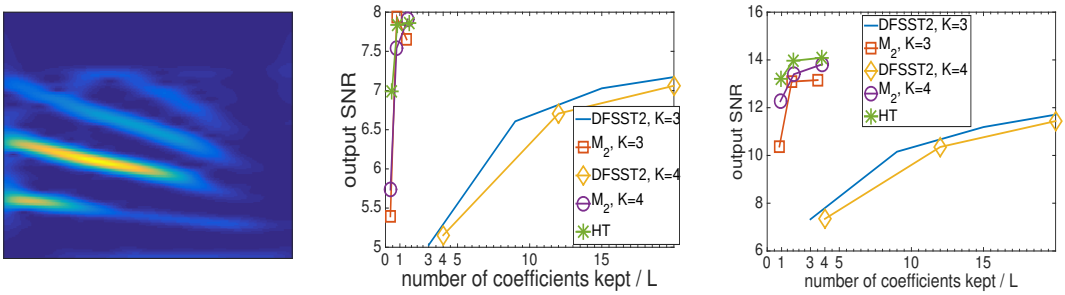
$$N_{M_2} = \sum_{l=1}^{K} \sum_{m} \# \left\{ J_{l,m} \bigcap \mathbb{Z} \right\}, \tag{4.25}$$

where $J_{l,m}$ is the interval defined in (4.24) and $\#\{X\}$ the cardinal of set X. Note that the range for m depends on the downsampling factor R. In Figure 4.14, we depict for input SNRs equal to 0 or 10 dB, the output SNRs associated with mode reconstruction with respect to the number of coefficients kept, i.e. N_{FSST2} or N_{M_2}, divided by signal length L. Doing so, we highlight the very different nature of the mode reconstruction techniques based on DFSST2 and M_2. Indeed, whatever the type of modes considered, for a given output SNR, mode retrieval is performed using much less information using the latter technique than the former. But more than that: the latter technique, while denoising the modes, can also compress the whole signal since the number of coefficients kept over the signal length L is lesser than one when $R \geq 64$. On the contrary, FSST2 based denoising techniques always require more than L coefficients. In this example, the amount of information needed to get the same output SNR is about 15 times larger with the technique associated with DFSST2 than with M_2. Note finally that for the former technique the results are not sensitive to the modulation (f_1 reconstruction versus that of f_2), which is expected because this algorithm demodulates each mode before proceeding with reconstruction.

C. Application to Real Data

In this section, we investigate the behavior of our new method for the denoising of a bat echolocation call, whose noise-free STFT is depicted in Figure 4.15 (a). As one does not know the ground truth in terms of the modes to be extracted, one cannot compute output SNR for each mode; the only thing one can do is to compute a global output SNR. In this regard, we are going to check that we do not lose much information by considering mode extraction as in M_2. Another difference between this signal and the one associated with Figure 4.10 (c) is that the modes only last for a certain amount of time, which makes ridge extraction more complicated. For denoising purpose, we thus investigate three different strategies: M_2, DFSST2, and HT.

We again analyze the quality of denoising with respect to the number of coefficients kept over the length of the signal, the results depicted in Figure 4.15 (b) and (c), respectively correspond to a noise level of 0 and 10 dB. The number K of modes is set to 3 or 4 in M_2 and also in DFSST2. We first remark that while the latter technique is sensitive to the number of modes (N_{FSST2} is linear with respect to K), M_2 is not. Indeed, if one considers an irrelevant mode number with M_2, the corresponding coefficients will be mostly below the noise level and will thus not be considered. Another interesting fact is that, in such a context, to extract the mode for denoising purpose may be useless, because the coefficients above the noise threshold are almost all located in the vicinity of the different ridges: given a downsampling factor, the number of selected coefficients is almost the same with M_2 or HT. Finally, we also remark that when the downsampling factor is large, the denoising performance becomes



(a) (b) (c)

Fig. 4.15 (a): STFT of a bat echolocation call, (b): output SNR associated with the reconstruction of the signal associated with (a) with respect to the number of coefficients kept divided by the signal length (input SNR 0 dB), the sensitivity to the number K of modes is also tested; (c): same as (b) but with an input SNR of 10 dB.

worse when extracting the ridges in the denoising process: if one wants to denoise the signal while compressing it, in such a situation, ridge extraction should be precluded.

4.3 Conclusion

In this chapter, we have introduced two contributions to the problem of denoising and mode reconstruction of multicomponent signals. First, a new algorithm for the retrieval of the modes from the study of some TFRs is presented. It relies upon 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 FFT in the computation of the TFRs and thus obtained reliable IF estimates. The simulation conducted on test signals show that, by demodulating the signal first using these IF estimates, the associated TFR is sharpened and that the accuracy of the reconstruction is much better than when direct reconstruction is performed, both in noiseless and noisy situations. Numerical simulations conducted on real signals where the number of modes may vary with time, however, show the limits of signal reconstruction based on ridge extraction. To address this issue, the following chapter proposes a new technique called *adaptive contour representation computation* based on properties of the reassignment vector (RV) which enables to estimate any time-frequency signatures of multicomponent signals. This work was published in [4].

Secondly, we introduced a new technique for the denoising and mode reconstruction of multicomponent signals from their downsampled STFT that enables to deal with the practical use of synchrosqueezing transform. We first recalled different signal reconstruction procedure using downsampled versions of short-time Fourier transform. From this, we have derived a classical hard thresholding strategy for signal denoising which we have then adapted to the

denoising of multicomponent signals and mode extraction. We then showed that the proposed technique compares favorably, on simulated and real data, with the most recent denoising ones based on synchrosqueezing in terms of the number of coefficients used in the reconstruction procedure. Recently, we have proposed a technique called *Shifted-Symmetrized-Regularized Hard-Thresholding* (SSR-HT) which enables an improvement of denoising behavior compared with the original HT [7].

Chapter 5

Adaptive Contour Representation Computation & its Application with Non-negative Matrix Factorization to Cardiac Signal Denoising

Contents

In previous chapters, we introduced a family of synchrosqueezing transforms (SSTs) not only enhancing time-frequency representations (TFRs), but also allowing for mode retrieval for a wide class of multicomponent signals (MCSs). Nevertheless, as highlighted above, because of the assumption on the MCSs being made of a fixed number of AM-FM modes, they cannot

handle *vanishing modes* or *Dirac impulses* that are widely encountered in many practical situations, as for instance, marine mammals [52], damped tones [53], musical sounds [54, 55], or thermoacoustic vibration [56]. Substantial methods were proposed to circumvent this problem, among which the ones based on using the properties of the *reassignment vector* (RV) to estimate TF signatures of MCSs have recently gained a renewed interest [57, 59, 58]. In contrast to the conventional methods which work directly with the ridge detection for AM-FM modes [100, 83, 101], the RV-based ones consider these modes as particular TF regions, called *basins of attractions* (BAs) subsequently used for mode reconstruction. However, the latter still fail to assess the TF signatures associated with a noisy Dirac impulse. To tackle this disadvantage, we introduce an improved version of the RV-based techniques based on using a local averaging rather than a punctual orientation of RV to lead to a more efficient computation of the direction of projection. As we will see, such an improved technique, coined *adaptive contour representation computation* (ACRC), enables to not only handle that type of impulse but also preserve its main characteristics for AM-FM modes. This work is the first contribution of this chapter [4].

The second contribution is a novel method for denoising *phonocardiogram* (PCG) signals based on a joint analysis between the just proposed technique ACRC and *non-negative matrix factorization* (NMF). Indeed, as we will see, this method, called *NMF-ACRC*, can avail the strengths of each technique to deal with their intrinsic drawbacks: the former does not work well on signals containing highly energetic noises that are regularly seen in PCGs [59], while the latter depends on a priori threshold that limits its adaptivity. This work has been recently carried out in [2].

This chapter is organized as follows: after having recalled, in Section 5.1.1, the basics associated with RV-based approaches, we detail, in Section 5.1.2, the new technique ACRC and then show how to perform mode reconstruction using the associated BAs. The numerical simulations of Section 5.1.3 demonstrate the improvement brought by ACRC on both a complex simulated MCS and a real data. Then, we introduce, in Section 5.2.2, the new method NMF-ACRC for PCG denoising. In Section 5.2.3, numerical simulations conducted on SiSEC 2016 illustrate the performance of the latter.

5.1 New Adaptive Contour Representation Computation (ACRC)

Before going into details of the new technique ACRC [4], the following section succinctly reviews the basics that are relevant to the RV-based methods.

5.1.1 RV-Based methods for TF Signature Estimation

A. RV Definition and Illustrations

Let first us recall the centroid $(\hat{\tau}_f(t,\eta), \hat{\omega}_f(t,\eta))$ of the spectrogram in the reassignment framework introduced in Section 2.5.1, whose coordinates are computed by:

$$\widehat{\tau}_f(t,\eta) = t + \Re\mathrm{e}\left\{\frac{V_f^{tg}(t,\eta)}{V_f^{g}(t,\eta)}\right\}, \quad \text{and} \quad \widehat{\omega}_f(t,\eta) = \eta - \frac{1}{2\pi}\Im\mathrm{m}\left\{\frac{V_f^{g'}(t,\eta)}{V_f^{g}(t,\eta)}\right\}. \tag{5.1}$$

With this in mind, the definition of the *reassignment vector* (RV) can be derived as follows:

Definition 5.1. *RV at time t and frequency η is defined by [57]:*

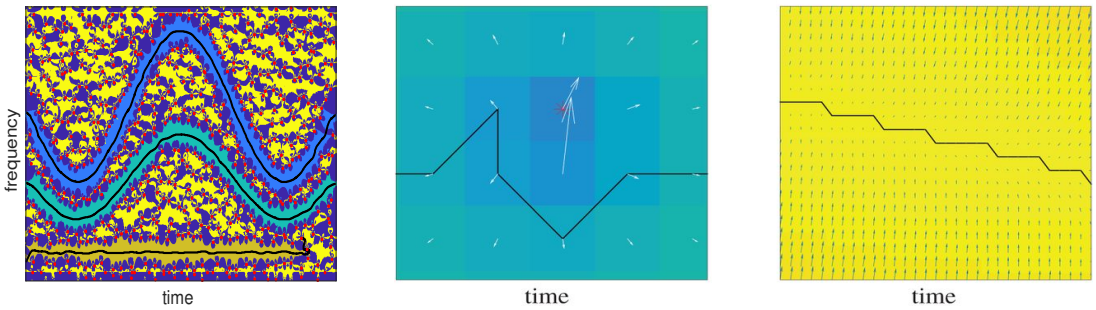
$$RV(t,\eta) = \begin{pmatrix} \hat{\tau}_f(t,\eta) - t \\ \hat{\omega}_f(t,\eta) - \eta \end{pmatrix}, \tag{5.2}$$

where the window g is assumed to be real-valued.

As an illustration, considering a Dirac distribution at t_0: $f = \delta_{t_0}$, one easily obtain: $RV(t,\eta) = (t_0 - t, 0)$: it has a component only along the time axis. Conversely, for a purely harmonic mode, $f(t) = e^{i2\pi\eta_0 t}$, one has $V_f^g(t,\eta) = \hat{g}(\eta-\eta_0)e^{2i\pi\eta_0 t}$ and thus $\dfrac{V_f^{tg}(t,\eta)}{V_f^{g}(t,\eta)} = \frac{\widehat{tg}(\eta-\eta_0)}{\hat{g}(\eta-\eta_0)}$ which is an imaginary complex number when g is even. Similarly, one has $\frac{V_f^{g'}(t,\eta)}{V_f^{g}(t,\eta)} = \frac{\widehat{g'}(\eta-\eta_0)}{\hat{g}(\eta-\eta_0)} = 2i\pi(\eta-\eta_0)$, so that, $RV(t,\eta) = (0, \eta_0 - \eta)$: it has a component only along the frequency axis. Another simple case is the one of a constant amplitude linear chirp, whose STFT reads [48]: $V_f^g(t,\eta) = f(t)g(u)\widehat{e^{i\pi\phi''(t)u^2}}(\eta - \phi'(t))$. When $g(t) = e^{-\sigma\pi t^2}$, the following two relations can be easily proven:

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

leading to $RV(t,\eta) = \frac{(\eta-\phi'(t))}{\sqrt{\sigma^2+\phi''(t)^2}}(-\phi''(t), \sigma^2)$. 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) (c)

Fig. 5.1 (a): STFT modulus of a three-component MCS with the zeros (red points) and three ridges; (b): a close-up of RV close to a zero (red asterisk): white arrows represent the RV while the nearby contour is depicted in black; (c): behavior of RV close to a ridge: blue arrows represent the RV while the ridge is plotted in black.

B. Existing RV-based Methods and Limitations

One of the major challenging problems in the TF analysis is how to detect ridge points and connect them together to construct smooth contours corresponding to TF signatures [100]. Here, we briefly recall a ridge detector that was initially proposed in [57] and then improved in [58, 59]. Such a detector relies upon the properties of RV in the vicinity of the ridge associated with the TF signatures: when crossing a ridge, RV undergoes a sharp variation in its orientations. As an illustration, let us consider a simulated MCS made of a constant and two sinusoidal chirps whose spectrogram is depicted in Figure 5.1 (a). Then, we display respectively, in Figures 5.1 (b) and (c), close-ups of RV close to a zero and to a ridge, which enables to illustrate the properties of RVs: on each side of a ridge, RVs have opposite directions. To determine the changes of the location of these sudden orientation, a first strategy was developed in [57] 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 the projected vector, i.e.

$$\langle RV(t,\eta), v_\theta \rangle = 0, \tag{5.3}$$

where v_θ is the unit vector in the direction θ, and $\langle .,. \rangle$ the inner product on $\mathbb{R}^2$. However, because the direction of projection θ is fixed a priori, the technique did 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 [58, 59] that consisted in modifying the definition of (5.3) as follows.

$$\alpha(t,\eta) := \langle RV(t,\eta), v_{\theta(t,\eta) \bmod \pi} \rangle = 0, \tag{5.4}$$

with $\theta(t,\eta)$ the argument of $RV(t,\eta)$ and where we consider $\theta(t,\eta) \bmod \pi \in [0,\pi[$. It is of interest to note that when $RV(t,\eta) \in [0,\pi]$ (*resp.* $[-\pi,0]$), $\langle RV(t,\eta), v_{\theta(t,\eta) \bmod \pi} \rangle$ equals 1 (*resp.* -1). This alternative, called M_1 in the sequel, enables the definition of a new type of CPs that no longer depends on a fixed angle θ. In Figure 5.1 (a), we display three detected ridges of an MCS using M_1.

(a) (b)

Fig. 5.2 (a): a close-up of a zero of the spectrogram and its corresponding contour computed with M_1; (b): STFT of a noisy Dirac impulse (SNR = 0 dB) together with the first 10 contours computed with M_1, which clearly point out the failure of the method in detecting the vertical ridge.

However, this technique suffers from some serious limitations [102]. Firstly, special structures are created in the vicinity of the zeros of the spectrogram since the mod π computation induces $\alpha(t,\eta)$ 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 results in numerical instabilities. All these phenomena are respectively illustrated in Figure 5.2 (a) and (b). In the following subsection, we develop a new adaptive method, called *adaptive contour representation computation* (ACRC), to enable to effectively overcome such limitations. Finally, it is worth of note that the CPs are practically chained by considering level zero contours of $\alpha(t,\eta)$ using *contourc* MATLAB function. Then, one segments the resulting contours with respect to the zeros of the spectrogram and subsequently sorts them into descending energy order corresponding to a series of decreasing energy ridges of modes.

5.1.2 New Adaptive Contour Representation Computation (ACRC) and Basins of Attraction (BA)

The applicability of the just recalled approaches based on the projection of RV to compute CPs is hindered by the fact that the orientation of RV, 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 technique ACRC that uses a criterion based on a local rather than a punctual orientation of RV to define a direction of projection. More precisely, the direction of projection for each RV is defined by considering the average over a squared neighborhood of RV centered at the point of study instead of considering only one single grid point as in [58, 59]. This

results in a much more robust estimation of the TF signature of modes like Dirac impulses, even at the high noise level, while maintaining a good behavior for AM/FM modes. Before going into the details of ACRC, we first investigate the impact of viewing RV as a displacement on a grid and not a vector with real coordinates that we call TF discretization effect.

A. Effect of TF discretization on RV orientation

Let us consider the definition of RV in (5.2) in the discrete setting, we remark that in practice, due to the discrete nature of the studied signals, RV is associated with a displacement on a grid, not a vector with real coordinates. Indeed, the signal is supposed to be defined on $0, \cdots, M-1$, then 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 a time instant. By rounding off to the nearest integers both in time and frequency the coordinates of RV, one obtains RV_r.

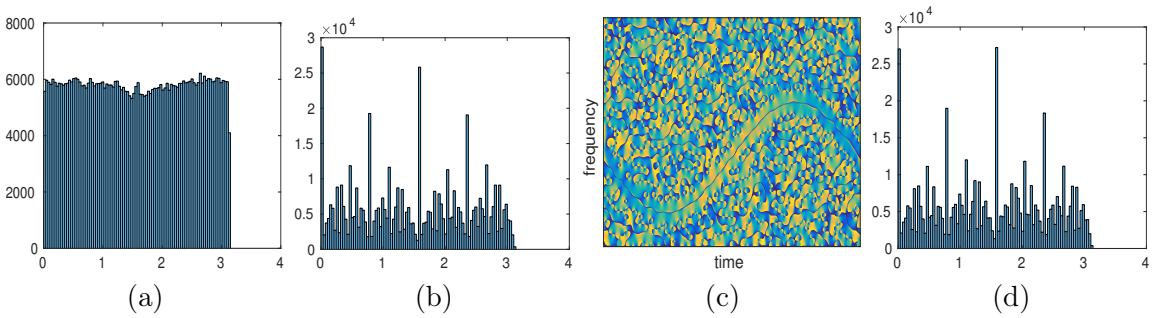


Fig. 5.3 (a): histogram of the argument of RV modulo π, for a white Gaussian noise; (b): same as (a) but with RV having its coordinates rounded to the nearest integers both in time and frequency (called RV_r in the paper), prior to histogram computation; (c): Argument of RV_r for a mode with sinusoidal frequency with some white Gaussian noise added (SNR 0dB); (d): the corresponding histogram of (c).

Then, in order to illustrate the effect of TF discretization on RV orientation, we depict the distributions of the argument of RV or RV_r (both taken modulo π), respectively in Figures 5.3 (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 $3\pi/4$. Let us now consider a signal with a sinusoidal frequency contaminated by the above white Gaussian noise (SNR 0dB), for which the argument modulo π of RV_r and the corresponding histogram are respectively displayed in Figures 5.3 (c) and (d). Comparing between Figures 5.3 (b) and (d), we notice that the presence of a mode with sinusoidal frequency does not alter the distribution of the angles. Despite these four orientations are not informative if one considers the whole TF plane, we are going to see in the following subsection that they are features enabling definition of the

new local direction of projection for RV_r, so-called *local projection angles* (LPAs). This helps improve the performance of the estimator of TF signatures based on the projection of RV.

B. ACRC Algorithm

To determine LPAs, we first consider the argument modulo π of RV_r, which we denote $\theta_r(k, \frac{p}{N})$ and recall that it belongs to $[0, \pi[$. Then, at each grid point (p, k), a local projection angle (LPA) is defined as the most frequent value of θ_r in a squared neighborhood of size $(2T_s + 1)^2$ centered at the point of study. We then project RV_r onto the direction given by this just computed LPA and define the new CPs as the zeros of the projection. The whole procedure called ACRC is detailed in the following algorithm (the *mode* function returns the most frequent value in an array):

Algorithm 5 ACRC Algorithm

Input: RV_r
$\theta_r := mod\,(arg(RV_r), \pi), \quad [M, N] := size(RV_r)$
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 and then chain them using *contourc* MATLAB function.

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 CPs computation and will be further studied in Section 5.1.3 (M_1 corresponds to $T_s = 0$).

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

Having determined the ridges (or contours) associated with the modes making up the signal using ACRC algorithm, we define the *basin of attraction* (BA) associated with a ridge, i.e. the set of coefficients associated with a given contour as in [58]. Since RV points to a ridge in its vicinity, we determine the BA of a given ridge as the set of points such that RV points to that ridge. However, because the localization property of RV is only valid for linear chirps, and also because of the presence of noise, RV does not point exactly to a ridge. Therefore, it is proposed in [58] to associate with a given coefficient (t, η) the closest ridge to point $(\hat{\tau}_f(t, \eta), \hat{\omega}_f(t, \eta))$. Once the BAs are computed, each corresponding mode f_i is reconstructed through:

$$f_i(t) = \frac{1}{g(0)} \int_{(t,\eta)\in\mathcal{B}_i} V_f^g(t, \eta) d\eta, \tag{5.5}$$

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

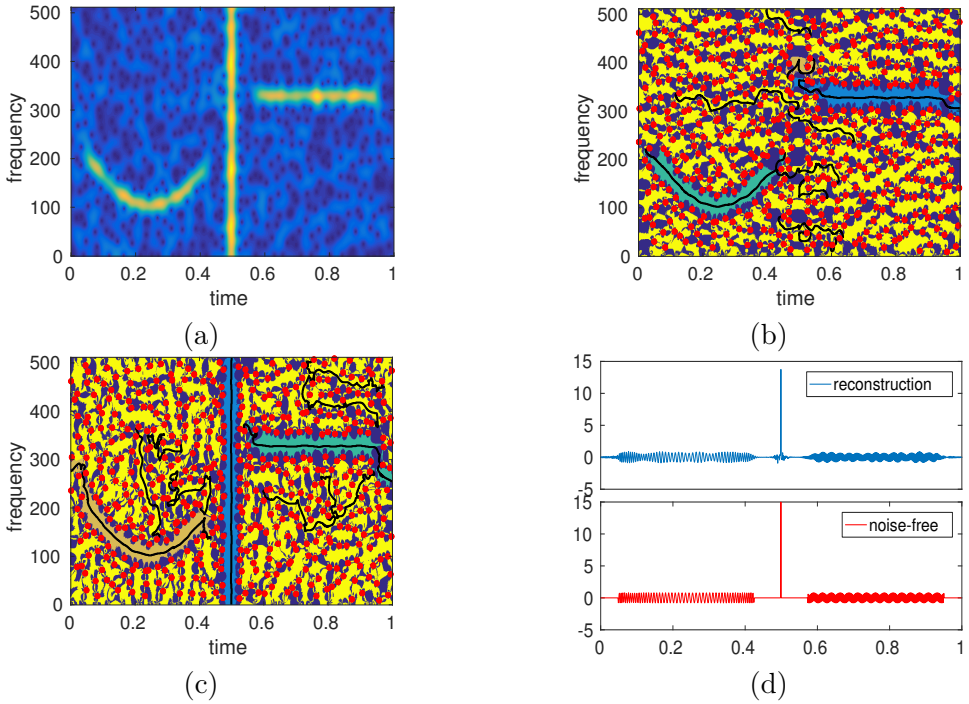


Fig. 5.4 (a): spectrogram of the simulated signal (SNR 0 dB); (b): BAs associated with the first 10 contours computed by method M_1; (c): same as (a) but computed with ACRC algorithm (with $T_s = 30$); (d): reconstructed signal based on the coefficients contained in the three most energetic BAs depicted in (c) accompanied by the original noise-free signal.

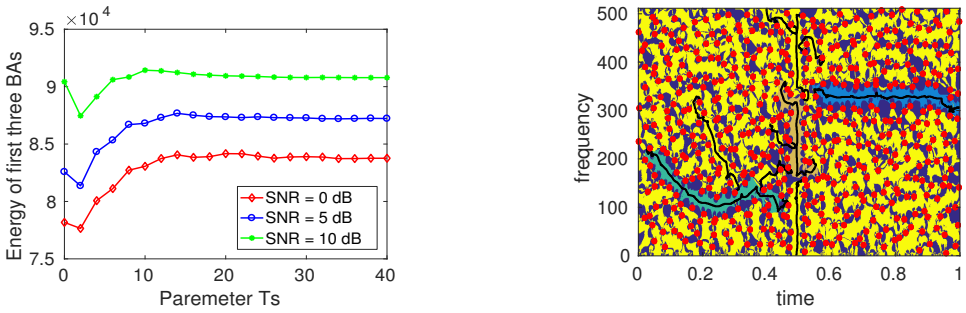
5.1.3 Numerical Experiments

This section investigates the properties of the proposed algorithm for mode TF signature identification, signal denoising, and mode reconstruction. Indeed, as we will see, the proposed algorithm enables a fully adaptive denoising and mode reconstruction.

A. Numerical Results

Let us first consider a simulated MCS made 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, the signal is sampled at a rate $N = 1024$Hz on $[0, 1]$ and its STFT is computed with the Gaussian window $\sigma = 1$). We first display in Figure 5.4 (a) the spectrogram of the signal and then in Figures 5.4 (b) and (c) the basins of attraction along with the first 10 contours computed with method M_1, and ACRC algorithm, respectively. 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 5.4 (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 5.4 (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, especially without requiring any threshold.



(a) (b)

Fig. 5.5 (a): Energy contained in the three most energetic BAs computed by ACRC algorithm on Fig.5.4 (a) for different values of T_s and noise levels; (b): BAs and the first 10 contours computed by ACRC algorithm with $T_s = 6$ and noise level at SNR = 0dB.

B. Sensitivity to Parameter T_s of ACRC Algorithm

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, namely:

$$E_f(T_s) = \sum_{i=1}^{K} \sum_{(t,\eta)\in\mathcal{B}_i^{T_s}} |V_f^g(t,\eta)|^2, \tag{5.6}$$

where we know that $\mathcal{B}_i^{T_s}$ 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 5.5 (a), we display $E_f(T_s)$ for $K = 3$, for the MCS of Figure 5.4 (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 5.5 (b) and Figure 5.4 (c)), respectively.

C. Application to Real Signal

We now illustrate the proposed technique ACRC 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

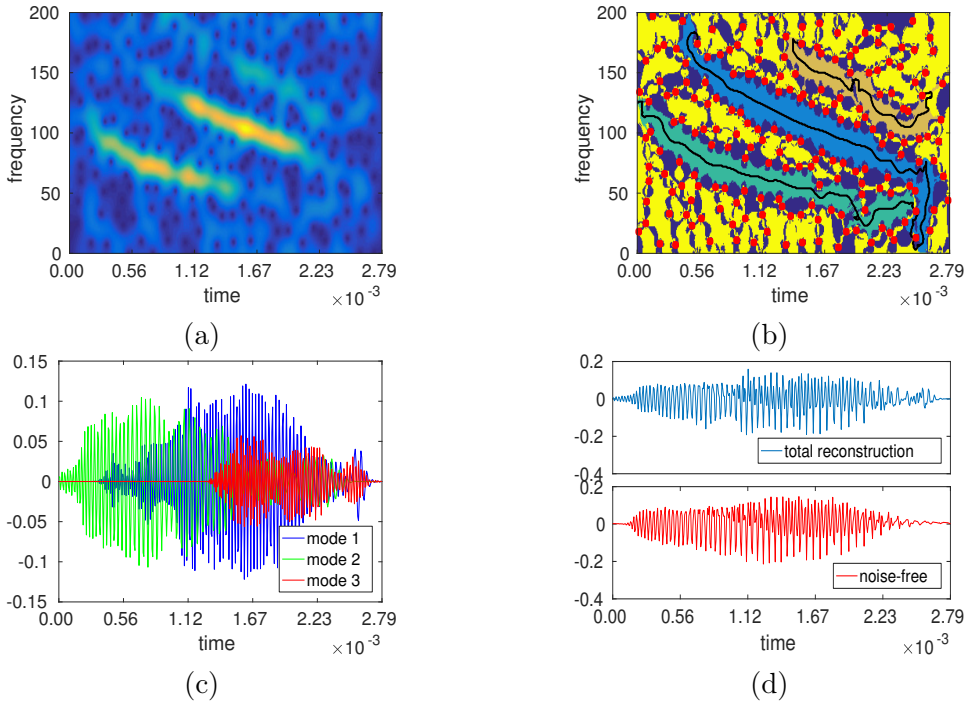
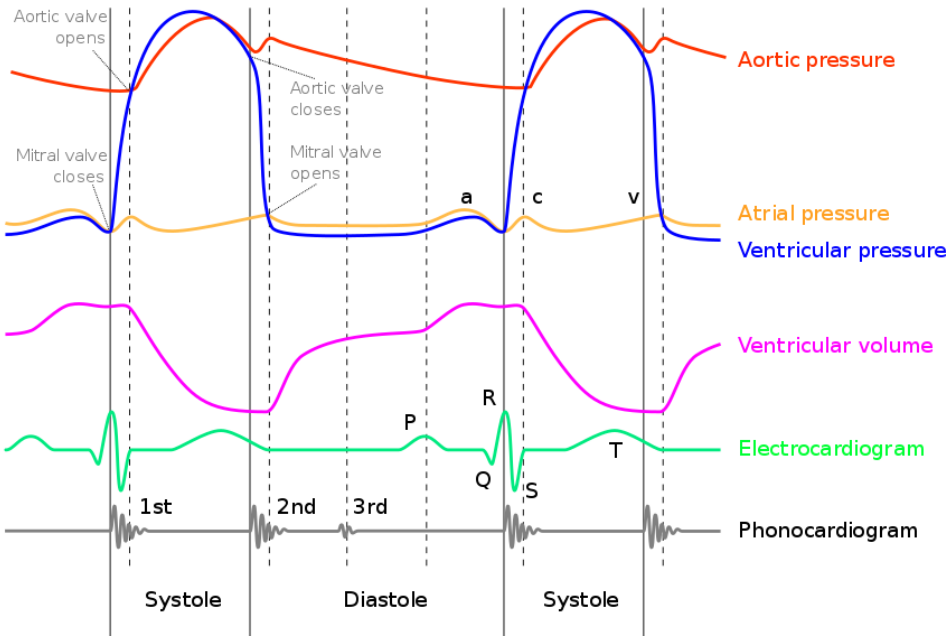


Fig. 5.6 (a): the spectrogram; (b): BAs associated with the first 3 contours computed with ACRC algorithm; (c): three reconstructed modes based on the coefficients contained in the three most energetic BAs depicted in (b); (d): reconstructed signal followed by the original noise-free signal.

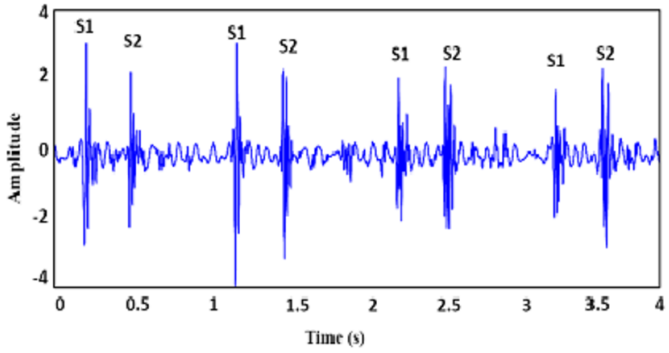
SNR equals 5.0 dB. The spectrogram of the noisy bat signal is displayed in Figure 5.6 (a). Then, it can be seen from Figure 5.6 (b) that the BAs corresponding to the three main components of the echolocation signal are well estimated by ACRC algorithm, enabling the reconstruction of the three detected modes (Figure 5.6 (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 a real MCS.

5.2 A Joint Analysis of NMF and ACRC to Cardiac Signal Denoising

Having introduced ACRC to estimate TF signatures of a wide class of MCSs, this second section puts forward a new denoising method based on a joint analysis of ACRC with NMF to enable to denoise phonocardiogram (PCG) signals. For a better understanding, let us first present the context and motivation of this study.



(a)



(b)

Fig. 5.7 (a) The Wiggers diagram to illustrate the cardiac cycle events together with *phonocardiogram* (PCG) signal and electrocardiogram (ECG) trace [11]; (b) two major audible S1 and S2 of a normal PCG signal.

5.2.1 Context and Motivation

Cardiovascular diseases (CVDs) are one of the world's largest public health problems, causing the death of nearly 18 million people annually, accounting for one-third of deaths worldwide [103]. Many efforts have been made to cope with this issue, among which CVDs diagnosis method based on cardiac auscultation for early-stage detection of heart abnormalities is the most commonly used. Cardiac auscultation is a non-invasive and low-cost technique defined as the listening of the heart sounds using an acoustic stethoscope. In fact, such sounds are the result of the mechanical vibrations controlled by two sets of valves: AV-valves (mitral and tricuspid) between the atria and ventricles, and semilunar valves (aortic and pulmonary)

between the ventricles and arteries. Also, during a normal cardiac cycle consisting of two periods, systole and diastole, they can be categorized into seven main components: S1, S2, S3, S4, murmurs, clicks, and snaps, among which the first two ones are heard more clearly. More precisely, S1, (*resp.* S2,) is associated with the closing of atrioventricular (*resp.* aortic and pulmonary) valves corresponding to the beginning of ventricular systole (*resp.* diastole) [104]. With the development of electronic stethoscope, the graphical recording of a heart sound, called *phonocardiogram* (PCG), can also be displayed on a digital computer and then analyzed to provide more insightful information on heart condition [105]. As an illustration, we display, in Figure 5.7 (a), the events and details of the cardiac cycle with *phonocardiogram* (PCG) signal and electrocardiogram (ECG) trace. It is referred to as the Wiggers diagram [11]. Further, we show, in Figure 5.7 (b), the two major audible components S1 and S2 through a normal PCG.

Unfortunately, PCGs are severely contaminated by many different types of noises including subject movement, subject speech, ambient sources, stethoscope movement, and lung sounds, making their analysis quite difficult [106]. Various methods for PCG denoising are reported in the literature. A class of commonly used techniques based on wavelet thresholding and variants were developed [107–109], or approaches using empirical mode decomposition (EMD) were introduced in [110–112]. However, these methods are all based on the analysis of the PCG on its own, while, for denoising purpose, it may be of interest to use an extra non-invasive recording, as for instance *electrocardiogram* (ECG). This idea was investigated in [22], where the analysis of simultaneous PCG and ECG was carried out by decomposing their respective spectrogram using *non-negative matrix factorization* (NMF). In that paper, denoising was performed by computing the cross-correlation between so-called *activation functions* associated with PCG and ECG, and then by thresholding this quantity to enable performant noise removal. Unfortunately, the choice of the threshold needed to be made a priori, thus limiting the adaptivity of the proposed technique. In addition, the method, known as *adaptive contour representation computation* (ACRC), just introduced in Section 5.1, enables an adaptive mode decomposition and denoising of any complex MCS. Unfortunately, it does not work well on signals containing highly energetic noises, as for instance, PCG signals. Therefore, in what follows, we investigate how the joint analysis of noisy PCG signals with NMF and ACRC for denoising purpose.

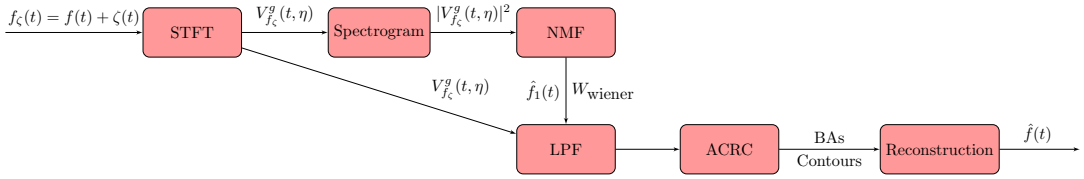


Fig. 5.8 Block diagram of the proposed NMF-ACRC denoising method.

5.2.2 Proposed Method

This section introduces a new algorithm *NMF-ACRC* for PCG denoising that relies on successively combining NMF with ACRC. As we will see, this combination will enable to take advantage of the strengths of each technique, thus resulting in a significant performance improvement. The explicit block diagram of the proposed algorithm is depicted in Figure 5.8, in which the two main steps, i.e. NMF and ACRC, are discussed in details in the following subsections. Before detailing them, the following subsection presents the signal database used extensively in the sequel.

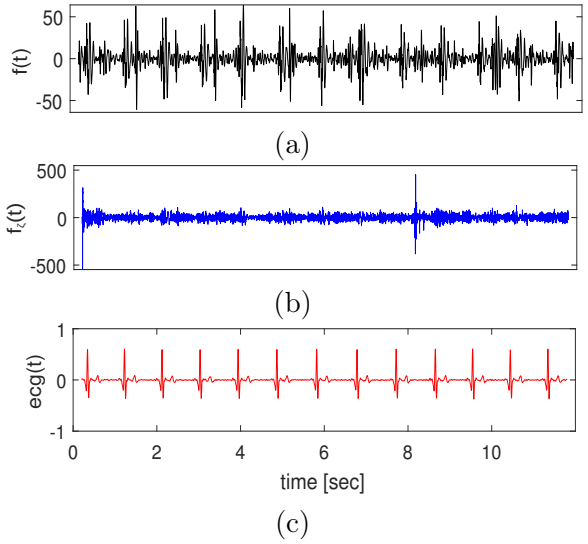


Fig. 5.9 (a): noise-free PCG; (b): noisy PCG; (c): synchronous ECG.

A. Database

The simulations are exploiting a database already used during the sixth community-based Signal Separation Evaluation Campaign (SiSEC 2016), during which PCGs were recorded with a cardiac microphone MLT210 on three healthy volunteers, while ECGs were simultaneously acquired by PowerLab instrument. Such signals, sampled at 1KHz, were then passed through a 15 to 300Hz band-pass filter. In total, sixteen such synchronous PCGs and ECGs were recorded, lasting from ten seconds to more than a minute. In a second time, PCGs were artificially contaminated by different real interference (radio, cough, pseudo-periodic breathing noise, etc.). As an illustration, we display, in Figure 5.9, a noise-free PCG ($f(t)$), its noisy version ($f_\zeta(t) = f(t) + \zeta(t)$), and synchronous ECG ($ecg(t)$).

B. NMF: Filtering out High-energy Noises

The key idea of NMF technique stems from the fact that information of a single natural phenomenon can be acquired using different devices, called *modalities*, associated with different

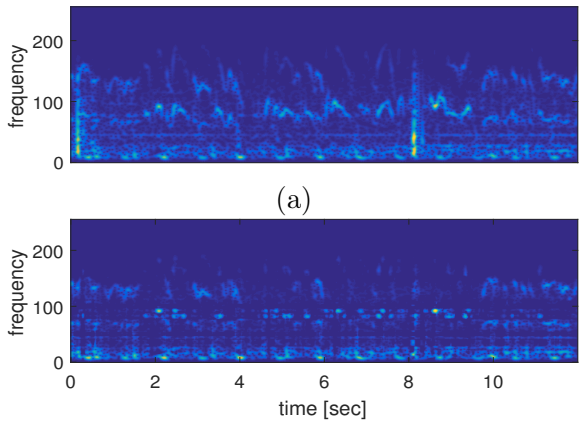


Fig. 5.10 Illustration of the application of NMF algorithm to PCG displayed in Figure 5.9 (b): (a) noisy spectrogram; (b) denoised spectrogram.

datasets [113, 114]. For instance, ECG corresponds to the recording of the electrical activity of the heart, while PCG to heart sounds. Modalities have similarities with one another, among which the most relevant is definitely quasi-periodicity, i.e. activation and inactivation periods are almost the same. In this context, the use of *non-negative matrix factorization* (NMF) enables to reflect this property by putting forward similar parameters, called *shared factors* [115]. In a nutshell, NMF approximates a m by n matrix V with non-negative elements by the product of two non-negative matrices W and H both with non-negative elements and with respective sizes $m \times k$ and $k \times n$: $V \approx WH$, where k, much smaller than m or n, is called the number of estimated components [116, 117]. Note that this factorization is not exact; WH is a lower-rank approximation to V. The factors W and H are chosen to minimize the root mean squared error (RMSE):

$$\min_{W,H} ||V - WH||_2, \quad \text{sujet to} \quad W \geq 0, \; H \geq 0. \tag{5.7}$$

Note also that H is the shared factor which varies little across all the recordings of multimodal datasets, while W is the unshared factor. For physiological signals, NMF applied to their spectrograms are of particular interest for identifying signals exhibiting similar temporal behaviors [118]. In this regard, a denoising procedure enabling the elimination of high-energy noises from noisy PCGs using synchronous ECGs was proposed in [22], and summarized hereafter. First, one computes the spectrograms of noisy PCG and synchronous ECG, respectively denoted V_{f_ζ} and V_{ecg} (in our simulation a Gaussian window $g(t) = \sigma^{-1} e^{-\pi \frac{t^2}{\sigma^2}}$ is used). Then, NMF using an alternating least-squares algorithm [119] is applied to decompose the spectrogram of the noisy PCG into $k = 12$ (practically chosen number) components, i.e. $V_{f_\zeta}^{m,n}(t,\eta) = W_{f_\zeta}^{m,12}(\eta) H_{f_\zeta}^{12,n}(t)$, and one component for ECG spectrogram

$V_{ecg}^{m,n}(t,\eta) = W_{ecg}^{m,1}(\eta)H_{ecg}^{1,n}(t)$ (the size of the matrices added as superscript to emphasize the size of matrix H varies depending on the type of signal). One then computes the cross-correlation between $H_{f_\zeta}^{k,\cdot}(t)$ and $H_{ecg}^{1,\cdot}(t)$, for each k. When it is larger than some a prior threshold γ, the corresponding $H_{f_\zeta}^{k,\cdot}(t)$ is associated with PCG, and noise otherwise. This enables the determination of an estimation of the spectrogram of the PCG and of the noise as follows: $V_{\hat{f}_1}(t,\eta) = W_{f_\zeta}^{m,12}(\eta)H_{\hat{f}_1}^{12,n}(t)$ and $V_{\hat{\zeta}_1}(t,\eta) = W_{f_\zeta}^{m,12}(\eta)H_{\hat{\zeta}_1}^{12,n}(t)$ where $H_{\hat{f}_1}^{12,n}(t)$ and $H_{\hat{\zeta}_1}^{12,n}(t)$ are activation components associated with estimated signal and noise, respectively, and satisfying: $H_{\hat{f}_1}^{12,n}(t) + H_{\hat{\zeta}_1}^{12,n}(t) = H_{f_\zeta}^{12,n}(t)$. In that context, the Wiener filter used for high-energy noise removal is derived as follows [120, 121]:

$$W_{\text{wiener}} = \frac{V_{\hat{f}_1}(t,\eta)}{V_{\hat{f}_1}(t,\eta) + V_{\hat{\zeta}_1}(t,\eta)}. \tag{5.8}$$

In Figure 5.10 (a), we first depict the spectrogram of the noisy PCG displayed in Figure 5.9 (b) in which two high-energy impulse noises appear at the beginning (at time 0.2 seconds) and in the middle (at time 8.1 seconds). Then, we show, in Figure 5.10 (b), the denoised spectrogram using the algorithm just described and based on NMF. It is clear that the latter removes the two high-energy impulse noises. Although NMF proves to be an efficient solution for PCG denoising, it can still be improved using ACRC as a post-processing step as we are going to show. A brief summary of how we apply ACRC in PCG denoising context is the subject of the following section.

C. ACRC: Components Estimation and Signal Retrieval

After having removed high-energy noises using NMF, the ACRC technique is applied on a low-pass filtered version of the spectrogram. It is worth remembering that the outputs of ACRC are contours corresponding to the TF signatures of the relevant components, which is then used to define the *basins of attraction* (BAs) associated with each contour enabling us to proceed with signal retrieval.

Before performing ACRC, a low-pass filter (LPF) with a cutoff frequency 80Hz is used to enable not only a removal of the impact of high-frequency noises but also a significant reduction of the computational cost of ACRC. Also, we notice that the average number of contours detected is 3.5 contours/second, and due to the quasi-periodicity of PCG that can be reflected by ACRC, the total number of contours is close to $3.5N/1000$, N being the signal length. We display, in Figure 5.11 (a) and (b), the basins of attraction along with the first contours $3.5N/1000$ (in red) computed with ACRC on the non-filtered and filtered STFTs associated with the signals whose spectrograms are displayed in Figure 5.10 (a) and (b), respectively. From Figure 5.11 (a), it is clear that ACRC computes contours associated with the two high-energy impulse noises mentioned above, while, after noise removal using the NMF algorithm, the computed contours are apparently more relevant. Finally, we illustrate

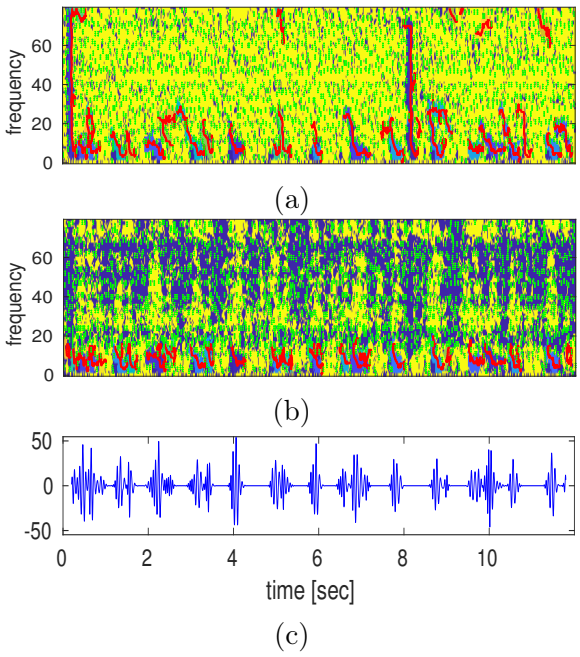


Fig. 5.11 Illustration of the application of ACRC and LPF with a cutoff frequency 80Hz: (a): on the signal displayed in Figure 5.9 (b); (b): on the denoised signal whose spectrogram is displayed in Figure 5.10 (c); (c) reconstructed PCG.

in Figure 5.11 (c) the reconstruction of PCG signal by selecting STFT coefficients associated with BAs displayed in Figure 5.11 (b). In the following section, we introduce how to assess the performance of the proposed technique in comparison with the state-of-art methods.

5.2.3 Results and Discussion

This section investigates the performance of the proposed denoising algorithm NMF-ACRC by comparing it with *empirical mode decomposition* (EMD) [122], *non-negative matrix factorization* (NMF), and *adaptive contour representation computation* (ACRC) on the database of PCG signals introduced in Section 5.2.2.A.. Before going into the details, we introduce, in the following subsection, a quantitative measurement for the performance evaluation of the techniques. The Matlab code implementing the method and the scripts generating all the figures are available at github.com/phamduonghung/EUSIPCO2018.

A. Evaluation Criteria

To measure the performance of the techniques, we use *BSS Eval Toolbox* [123], which is based on the decomposition of each estimated signal into a number of contributions associated with the target signal, interference of the unwanted sources and artifacts. More precisely, the following three evaluation criteria corresponding to the three following energy ratios expressed in decibels (dB) are used:

- Signal-to-Distortion Ratio (SDR): measures globally the level of all error terms.
- Signal-to-Interference Ratio (SIR): estimates the level of interference from all the other interfering sources.
- Signal-to-Artifacts Ratio (SAR): estimates the level of algorithmic artifacts and the linearity separation of the algorithm.

Having defined these criteria, we assess the performance of the different methods by computing the gain between each evaluation criterion computed on the denoised (output) and initial noisy (input) signals. For instance, SDR gain = output SDR-input SDR.

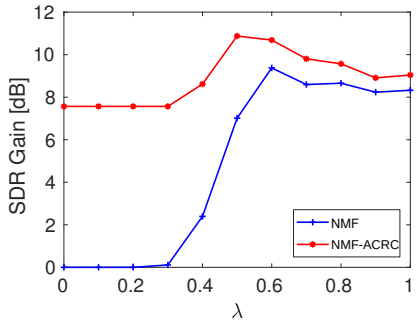


Fig. 5.12 SDR gain for different cross-correlation threshold values γ when applying two techniques NMF and NMF-ACRC to the noisy PCG signal displayed in Figure 5.9 (b).

B. Sensitivity to Cross-Correlation Threshold of NMF-ACRC

As mentioned above, one of the main drawbacks of NMF is that the choice for threshold γ used to separate the noise and signal components strongly influences the technique performance. To illustrate how NMF-ACRC manages to alleviate this dependence on γ, we display, in Figure 5.12, SDR gains obtained when using either NMF or NMF-ACRC algorithms, for different γ and for the signal depicted in Figure 5.9 (b). It is clear that the SDR gain obtained with NMF is null when γ is small, meaning, in that case, PCG is not denoised at all. Furthermore, the SDR gain increases and then stagnates when some particular value for γ is reached. In contrast, SDR gain obtained when using NMF-ACRC is bigger and also more stable with respect to parameter γ, especially when γ is small. These results confirm NMF-ACRC is only slightly sensitive to the choice of threshold γ, which thus makes it more adaptive than NMF.

C. Comparison of Denoising Performance

In this section, we compare NMF-ACRC with the three state-of-art techniques including EMD, NMF, and ACRC on the database of the sixteen real noisy PCGs of the SiSEC database. We display the distribution of the different gains using the boxplot representation [124], in which the central line indicates the median, the bottom and top edges of the box indicating the 25th

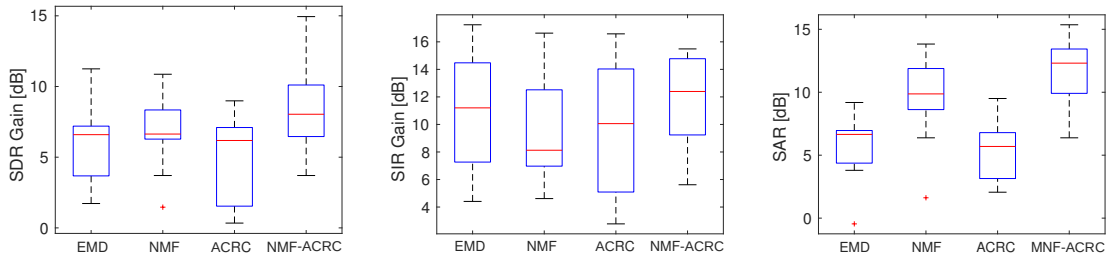


Fig. 5.13 Denoising performance (expressed in dB) of the four techniques on the database of real noisy PCG signals: (a) SDR gain; (b) SIR gain; (c) SAR.

and 75th percentiles, respectively. In Figure 5.13, we depict the denoising results associated with each of the tested techniques in terms of SDR, SIR and SAR gains. At first glance, we remark that NMF-ACRC produces better results than the other studied methods for all three evaluation criteria. More precisely, in Figure 5.13 (a), SDR median gain obtained using NMF-ACRC is 7.1 dB, while it equals 6.6, 6.6, and 6.2 dB with EMD, NMF, and ACRC, respectively. Moving to Figure 5.13 (b), NMF-ACRC leads to an SIR median gain of 12.4 dB bigger than the one obtained with the other three methods (12.2, 8.12, and 10.0 dB for EMD, NMF, and ACRC, respectively). Finally, Figure 5.13 (c) shows that SAR median gain with NMF-ACRC is 12.3 dB, while those of EMD, NMF, and ACRC are much smaller: 6.6, 8.7, 5.7 dB, respectively. In addition, the analysis of the standard deviation of the obtained results shows that NMF-ACRC behaves similarly to the other methods in that respect. All in all, for PCG denoising, these results plead in favor of mixing NMF with ACRC as is done by NMF-ACRC.

5.3 Conclusion

In this chapter, we have first introduced a fully adaptive technique ACRC 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 and used the former 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 bat echolocation call.

Then, we have presented another application of ACRC for PCG denoising by successively combining it with NMF. As a result, we managed to circumvent the essential limitations of NMF and ACRC techniques, which result in a significant improvement in terms of denoising performance. In a nutshell, NMF was first used to remove the high-energy noises from the initial noisy PCG signal, while ACRC was carried out as a post-processing step to estimate TF signatures of relevant components. The resultant TF signatures were subsequently used for

signal reconstruction. Numerical experiments demonstrated the effectiveness of the proposed approach on a database of real PCG signals.

Chapter 6

Conclusions and Future Research

Contents

This dissertation was set out to investigate various aspects of the synchrosqueezing and associated methods in both theoretical analyses and numerical experiments. In what follows, we highlight the main research contributions of this dissertation, as well as discuss some directions for future work.

6.1 Main Contributions

After a concise introduction for the context and questions of research in Chapter 1, we gave a brief overview of the works accomplished with regard to TFR methods for the analysis of MCSs in Chapter 2. Then, all the major results carried out in the next three chapters are summarized as follows.

- **The second-order wavelet-based SST (WSST2)** derived from CWT of an MCS was first presented to deal with signals made of components or modes with strong

frequency modulation. It relied on the recalculation of the IF estimate using a second-order expansion of the phase. Also, we readdressed the theoretical analysis of the first-order synchrosqueezing (WSST) as proposed in [38] but computed with a non-compact wavelet in the Fourier domain. In this regard, a novel approximation theorem for the new technique WSST2 using that type of wavelet was then developed. Finally, numerical implementation and simulations emphasizing the differences between WSST2 and existent synchrosqueezing techniques (WSST, FSST and FSST2) were introduced.

- **A generalization of the STFT-based SSTs** was put forward by using the redefinition of IF estimates using high-order amplitude and phase expansions. Such a technique was proved not only to achieve a highly concentrated TFR for a wide variety of MCSs but also reconstruct their modes with a high accuracy. Numerical investigation on a synthetic signal showed the efficiency of this new approach.

- **An appealing application** on a real gravitational-wave signal of the synchrosqueezing techniques (both wavelet and STFT) was also given. In particular, the fourth-order SST (FSST4) enabled the accurate detection of three stages of the collision of two black holes including the inspiral, merger and ringdown. This confirms the interest of the SST generalization on a real-world situation.

- **A novel demodulation-FSST2-based (DFSST2) technique** was presented for the retrieval of the modes of an MCS derived from the study of its TFR given by FSST2. It relied on a novel ridge extraction method, that took into account the fact that the TF representation is both discrete in time and frequency, followed by a demodulation procedure. Numerical results showed a better behavior of DFSST2 for mode reconstruction over similar techniques without using demodulation.

- **A new approach for the reconstruction and denoising** of MCSs from their downsampled STFTs was proposed. More precisely, we first recalled signal reconstruction techniques based on STFTs and how to use them for signal denoising. We then investigated how to adapt this new approach to the context of MCSs. Numerical experiments using the proposed approach and other synchrosqueezing techniques (FSST2 and DFSST2) were given to demonstrate the interest of the former.

- **A new technique called adaptive contour representation computation (ACRC)** for an adaptive estimation of TF signatures corresponding to the IFs of the components of MCSs was introduced. It relied on the estimation of the local orientation of the reassignment vector (RV). The resultant IF estimates enabled the segmentation of the TF plane into BAs that was subsequently used to reconstruct the modes. Numerical simulation showed that compared with previous approaches, this new method fully adapted to any TF complex structures, even those associated with noisy Dirac impulses or vanishing modes, thus resulting in better reconstruction performance.

- **A novel approach (called NMF-ACRC) for PCG denoising** based on an adaptive combination of two different techniques NMF and ACRC was put forward. NMF was first used to filter out high-energy noises thanks to the multimodality between PCG and synchronous ECG signals. Then ACRC was performed on a low-pass filtered version of the obtained signal to identify the relevant TF components subsequently used for signal retrieval. This new method was assessed on noisy PCG signals, already studied during the SiSEC 2016 evaluation campaign, and showed a significant improvement in terms of denoising performance compared with other state-of-the-art methods.

6.2 Future Research

A variety of synchrosqueezing and associated methods were proposed and developed for the analysis of MCSs in this dissertation. Much efforts should be pursued in the future in some specific directions as mentioned hereafter.

As highlighted in Chapter 3, the second-order wavelet-based SST (WSST2) provided better practical results than the second-order STFT-based SST (FSST2) even though they were theoretically expected to perform similarly. However, the former was less accurate than the generalization of the latter (FSST4) when applied to the gravitational-wave signals. Therefore, it would be interesting to make the same generalization synchrosqueezing in the wavelet case and then compare them with those based on STFT in terms of both TF concentration and mode reconstruction performance. Then, we should devote ourselves to the theoretical analysis of the behavior of all these techniques when applied to noisy signals, as was done in [40, 125] for the original WSST. Next, the impact of noise on synchrosqueezing operators still needs to be further studied, in particular in heavy noise situations. From a practical point of view, different techniques have been developed to handle the noise in the SST context, among which a very promising one, inspired by multitaper approaches [126], is based on averaging the SSTs obtained with different kernels [44]. Also, it would be of interest to study the behavior of the transforms when the type of noise is non-Gaussian, as for instance, Poisson or autoregressive-moving-average noises (ARMA) [44]. In addition, we should apply the proposed synchrosqueezing techniques to a more extensive set of practical signals as was done in [40, 127–131]. Finally, the fact that in this dissertation, the synchrosqueezing techniques were only compared with each other or with the standard reassignment (RM). Therefore, it would be interesting to make comparisons between the former and other state-of-the-art methods, as for instance, dictionary approaches based on chirplet path pursuit [20, 132].

As also mentioned in Chapter 3, the Gaussian window length has a strong influence on the performance of the SSTs. For a fair comparison of tested TF methods, the optimal value of such a length was determined by the technique based on Rényi entropy performing on the TFRs given by STFT, but not by SST. This is because the Rényi-based technique applied

on the latter no longer works in a very noisy context as emphasized in [6]. Thus, a detailed study of such a phenomenon should be conducted in the near future.

Furthermore, it is well known that ridge extraction is an essential step prior to mode reconstruction in the synchrosqueezing context. We introduced, in Chapter 4, a new method for that purpose carrying out on the TFRs of MCSs given by the reassigned (SSTs) or initial (STFT and CWT) transforms. However, it does not always behave better on the TFRs given by the former as those associated with the latter. The reason behind this fact is that the reassignment sometimes produces time irregularities, which cannot be taken into account by the ridge extraction technique. Therefore, it would be of interest to develop a ridge extraction technique better adapted to the reassigned context. Several related works would be useful for that purpose, as for example, a ridge extraction method using image processing technique of active contours or snakes [133]. Moreover, as mentioned in Chapter 4, the technique DFSST2 for mode reconstruction was built on the FSST2, thus it could be interesting to extend it to other synchrosqueezing techniques based on both STFT and wavelet. Regarding the technique for mode retrieval of MCSs from their downsampled STFT, future work should involve its generalization to non-Gaussian noises and its application to a wider range of large real signals.

In Chapter 5, we presented an improved technique ARCR for the adaptive estimation of TF signatures of any complex MCSs, but several issues arising need to be addressed in the future. First, the fact that the factorization in (5.7) in practice uses an iterative method starting with random initial values for W and H. However, since the RMSEs have local minima, repeated factorizations may yield different W and H. The question then arises as to whether there would be a way in which one can obtain the best factors W and H enabling optimal results. Secondly, as mentioned in Section 5.2.2.B., the number of estimated components is experimentally determined ($k = 12$ for SiSEC2016 database), which requires an automatic optimal determination procedure. Thirdly, we should devote ourselves to the application of the proposed method to larger real databases of highly energetic noise signals.

Finally, the ASTRES toolbox comprising a set of Matlab functions has recently introduced as in [134]. It provides a number of efficient numerical tools for the purpose of "Analysis, Synthesis and Transformations by Reassignment, EMD and Synchrosqueezing" (ASTRES) of non-stationary MCSs. It also can be viewed as an extension of the renowned Time-Frequency ToolBox (TFTB) [10]. Thus, one future work should be focused is to integrate the contributions developed in this dissertation into the ASTRES toolbox that will allow them to be more diffused and used in the scientific community.

Appendix A

Proofs of Section 3.1: Second-order Wavelet-based Synchrosqueezing Transform

A.1 Proof of Theorem 3.4 on page 39

The main steps for the proof of Theorem 3.4 are detailed hereafter. First, we introduce the following proposition that is useful to prove item (a) of Theorem 3.4.

Proposition A.1. *For any* $(t,a) \in \mathbb{R} \times \mathbb{R}^+$*, one has:*

$$\left| W_f^{\psi}(t,a) - \sum_{k=1}^{K} f_k(t) \overline{\hat{\psi}(a\phi_k'(t))} \right| \leq \varepsilon E_0(t,a), \tag{A.1}$$

where $E_p(t,a) = aMKJ_{1,p} + \pi a^2 M J_{2,p} \sum_{k=1}^{K} A_k(t)$ *and* $J_{n,p} = \int_{\mathbb{R}} |u|^n |\psi^{(p)}(u)| du$.

Proof. For each $k \in \{1, \ldots, K\}$, a zeroth order Taylor expansion of the amplitude and first order expansion of the phase of f_k leads to:

$$\begin{aligned} f_k(\tau) &= A_k(\tau) e^{i2\pi\phi_k(\tau)} \\ &= A_k(t) e^{i2\pi[\phi_k(t) + \phi_k'(t)(\tau - t)]} + (A_k(\tau) - A_k(t)) e^{i2\pi\phi_k(\tau)} \\ &\qquad + A_k(t)[e^{i2\pi[\phi_k(t) + \phi_k'(t)(\tau-t) + \int_t^\tau \phi_k''(x)(\tau - x)dx]} - e^{i2\pi[\phi_k(t) + \phi_k'(t)(\tau-t)]}] \\ &= f_{k,1}(\tau) + f_{k,2}(\tau) + f_{k,3}(\tau). \end{aligned}$$

Then, for any (t,a), the first term can be written as:

$$W_{f_{k,1}}^{\psi}(t,a) = \frac{1}{a} A_k(t) e^{i2\pi\phi_k(t)} \int_{\mathbb{R}} e^{i2\pi\phi_k'(t)(\tau - t)} \overline{\psi\left(\frac{\tau - t}{a}\right)} d\tau = f_k(t) \overline{\widehat{\psi}(a\phi_k'(t))}.$$

The second term is bounded by:

$$\begin{aligned}\left|W^{\psi}_{f_{k,2}}(t,a)\right| &\leq \frac{1}{a}\int_{\mathbb{R}} |A_k(\tau)-A_k(t)|\left|\psi\left(\frac{\tau-t}{a}\right)\right| d\tau \\ &\leq \frac{\varepsilon M}{a}\int_{\mathbb{R}} |\tau-t|\left|\psi\left(\frac{\tau-t}{a}\right)\right| d\tau = \varepsilon a M J_{1,0}.\end{aligned}$$

and the third term by:

$$\begin{aligned}\left|W^{\psi}_{f_{k,3}}(t,a)\right| &\leq \frac{2\pi A_k(t)}{a}\int_{\mathbb{R}}\left(\int_t^{\tau} |\phi_k''(u)||(\tau-u)|du\right)\left|\psi\left(\frac{\tau-t}{a}\right)\right| d\tau \\ &\leq \frac{\varepsilon\pi M A_k(t)}{a}\int_{\mathbb{R}} |\tau-t|^2\left|\psi\left(\frac{\tau-t}{a}\right)\right| d\tau = \varepsilon\pi a^2 M J_{2,0} A_k(t).\end{aligned}$$

Writing $|W^{\psi}_f - \sum_{k=1}^{K} W^{\psi}_{f_{k,1}}| = |\sum_{k=1}^{K}(W^{\psi}_{f_{k,2}} + W^{\psi}_{f_{k,3}})|$, we obtain the desired result. □

Now we can prove item (a) of Theorem 3.4: since $E_0(t,a)$ is bounded on $\mathbb{E}$, we can consider:

$$\tilde{\varepsilon} \leq \frac{1}{\sqrt{2}}\min\left(\|E_0(t,a)\|^{-\frac{1}{2}}_{\infty,\mathbb{E}}, \left\|N_0\sum_{k=1}^{K} A_k(t)\right\|^{-\frac{1}{2}}_{\infty,\mathbb{E}}\right) \tag{A.2}$$

where $\|z(t,a)\|_{\infty,X} = \sup_{(t,a)\in X}|z(t,a)|$. For $(t,a) \in \mathbb{E} \setminus \bigcup_{l=1}^{K} Z_l$, we immediately get $|W^{\psi}_f(t,a)| \leq \tilde{\varepsilon}$. Thus, if $|W^{\psi}_f(t,a)| > \tilde{\varepsilon}$, there is at least one k such that $(t,a) \in Z_k$. Furthermore, because of the separation condition on the modes, one can easily show the Z_ks are disjoint sets, so k is unique.

Remark A.2. Note that $E_0(t,a)$ is uniformly bounded for $(t,a) \in \mathbb{E}$ because a is lower than α. In the seminal paper of Daubechies [38], this constraint on a was missing.

Let us now detail the proof of item (b) of Theorem 3.4. Writing Proposition A.1 with wavelet ψ' we get for any $(t,a) \in \mathbb{R}\times\mathbb{R}^+$:

$$\left|W^{\psi'}_f(t,a) + \sum_{k=1}^{K} f_k(t) 2i\pi a\phi_k'(t)\overline{\hat{\psi}(a\phi_k'(t))}\right| \leq \varepsilon E_1(t,a),$$

with $E_1(t,a)$ being defined at the end of Proposition A.1. Thus, if $(t,a) \in Z_k$, one gets:

$$\left|W^{\psi'}_f(t,a) + f_k(t) 2i\pi a\phi_k'(t)\overline{\hat{\psi}(a\phi_k'(t))}\right| \leq \varepsilon\left(2\pi N_0 a\sum_{l\neq k}\phi_l'(t)A_l(t) + E_1(t,a)\right).$$

Note that since $\partial_t W_f^\psi(t,a) = -\frac{1}{a} W_f^{\psi'}(t,a)$, one can right for $(t,a) \in Z_k$ satisfying $|W_f^\psi(t,a)| > \tilde{\varepsilon}$:

$$\begin{aligned}
\left|\widehat{\omega}_f(t,a) - \phi_k'(t)\right| &= \left| \Re\mathrm{e}\left\{ \frac{1}{2\pi} \frac{\partial_t W_f^\psi(t,a)}{W_f^\psi(t,a)} - \phi_k'(t) \right\} \right| \\
&= \left| \Re\mathrm{e}\left\{ \frac{1}{i2\pi a} \frac{W_f^{\psi'}(t,a) + i2\pi a \phi_k'(t) W_f^\psi(t,a)}{W_f^\psi(t,a)} \right\} \right| \\
&\leq \left| \frac{1}{i2\pi a} \frac{W_f^{\psi'}(t,a) + 2i\pi a \phi_k'(t) f_k(t) \overline{\hat{\psi}(a\phi_k'(t))}}{W_f^\psi(t,a)} \right| \\
&\qquad + \left| \frac{\phi_k'(t) W_f^\psi(t,a) - \phi_k'(t) f_k(t) \overline{\hat{\psi}(a\phi_k'(t))}}{W_f^\psi(t,a)} \right| \\
&\leq \tilde{\varepsilon}^2 \left(N_0 \sum_{l \neq k} (\phi_l'(t) + \phi_k'(t)) A_l(t) + \frac{E_1(t,a)}{2\pi a} + \phi_k'(t) E_0(t,a) \right).
\end{aligned}$$

By putting, $B_k(t,a) = N_0 \sum_{l\neq k} (\phi_l'(t) + \phi_k'(t)) A_l(t) + \frac{E_1(t,a)}{2\pi a} + \phi_k'(t) E_0(t,a)$, and remarking it is bounded on $\mathbb{E}$, we may choose :

$$\tilde{\varepsilon} \leq \min_k \| B_k(t,a) \|_{\infty,\mathbb{E}}^{-1}, \tag{A.3}$$

so that for $(t,a) \in Z_k$ such that $|W_f^\psi(t,a)| > \tilde{\varepsilon}$, we immediately get:

$$\left|\widehat{\omega}_f(t,a) - \phi_k'(t)\right| \leq \tilde{\varepsilon}. \tag{A.4}$$

Let us now introduce the following lemma, which is useful to prove item (c) of Theorem 3.4.

Lemma A.3. *Suppose that both (A.2) and (A.3) are satisfied, and that the following condition is also verified:*

$$\varepsilon \leq 1/8c^3 (\phi_1'(t) + \phi_2'(t))^3. \tag{A.5}$$

Consider the following sets:

$$\begin{aligned}
\mathbb{X} &= \{a \ s.t. \ |W_f^\psi(t,a)| > \widetilde{\varepsilon} \ and \ |\widehat{\omega}_f(t,a) - \phi_k'(t)| \leq \widetilde{\varepsilon}\}, \\
\mathbb{Y} &= \{a \ s.t. \ |W_f^\psi(t,a)| > \widetilde{\varepsilon} \ and \ |a\phi_k'(t) - 1| < \Delta\},
\end{aligned}$$

then $\mathbb{X} = \mathbb{Y}$.

The proof of Lemma A.3 is available in available in [38] .

Coming back to the proof of item (c) of Theorem 3.4, let $t \in \mathbb{R}$ and note that $W_f^\psi(t,a) \in L^\infty(\mathbb{X})$. Then, since $a > 0$ on $\mathbb{X}$, $\frac{1}{a}W_f^\psi(t,a) \in L^1(\mathbb{X})$, and, thus, using the same type of technique as in [38] (Estimate 3.9), one gets:

$$\begin{aligned}
&\left|\lim_{\lambda\to 0}\left(\frac{1}{C'_\psi}\int_{|\omega-\phi'_k(t)|<\widetilde{\varepsilon}} S_f^{\psi,\lambda,\widetilde{\varepsilon}}(t,\omega)d\omega\right) - f_k(t)\right| = \left|\frac{1}{C'_\psi}\int_{\mathbb{X}} W_f^\psi(t,a)\frac{da}{a} - f_k(t)\right| \\
&\leq \left|\frac{1}{C'_\psi}\int_{|a\phi'_k(t)-1|<\Delta} W_f^\psi(t,a)\frac{da}{a} - f_k(t)\right| + \left|\frac{1}{C'_\psi}\int_{|W_f^\psi(t,a)|\leq\widetilde{\varepsilon}\cap|a\phi'_k(t)-1|<\Delta} W_f^\psi(t,a)\,\frac{da}{a}\right| \\
&\leq \frac{1}{\left|C'_\psi\right|}\left[\int_{|a\phi'_k(t)-1|<\Delta}\left|W_f^\psi(t,a) - f_k(t)\overline{\hat{\psi}(a\phi'_k(t))}\right|\frac{da}{a} + \frac{A_k(t)}{\left|C'_\psi\right|}\int_{|a\phi'_k(t)-1|\geq\Delta}\left|\hat{\psi}(a\phi'_k(t))\right|\frac{da}{a} + \right. \\
&\qquad\qquad \left. \widetilde{\varepsilon}\log\left(\frac{1+\Delta}{1-\Delta}\right)\right] \\
&\leq \frac{1}{\left|C'_\psi\right|}\left[\int_{|a\phi'_k(t)-1|<\Delta}\varepsilon(E_0(t,a) + N_0\sum_{l\neq k}A_l(t))\frac{da}{a} + A_k(t)N_1\widetilde{\varepsilon} + \widetilde{\varepsilon}\log\left(\frac{1+\Delta}{1-\Delta}\right)\right] \\
&\leq \widetilde{\varepsilon}\frac{1}{\left|C'_\psi\right|}\left[\|A_k\|_\infty N_1 + 2\log\left(\frac{1+\Delta}{1-\Delta}\right)\right] \leq D_1\widetilde{\varepsilon},
\end{aligned}$$

which ends up proving the theorem.

A.2 Proof of Theorem 3.11 on page 42

Theorem 3.11 is a generalization of Theorem 3.4, so the proof of the former is in principle similar to that of the latter. Proposition A.1 generalizes into:

Proposition A.4. *For any $k \in \{1,\dots,K\}$, any $r \in \{0,1\}$ and $p \in \{0,1\}$, and $(t,a) \in \mathbb{R}\times\mathbb{R}^+$, one has:*

$$\left|W_f^{\tau^r\psi^{(p)}(\tau)}(t,a) - \sum_{k=1}^{K} f_k(t)\overline{\mathcal{F}\{\tau^r\psi^{(p)}(\tau)e^{-i\pi\phi''_k(t)a^2\tau^2}\}(a\phi'_k(t))}\right| \leq \varepsilon E_{r,p}(t,a), \tag{A.6}$$

with $E_{r,p}(t,a) = a^{r+1}MKJ_{r+1,p} + \frac{\pi}{3}a^{r+3}MJ_{r+3,p}\sum_{k=1}^{K}A_k(t)$.

Proof. Following the same steps as the proof of Proposition A.1, but using a zeroth order Taylor expansion of the amplitude and second-order expansion of the phase of f_k, one has:

$$\begin{aligned}
f_k(\tau) &= A_k(\tau)e^{i2\pi\phi_k(\tau)} \\
&= A_k(t)e^{i2\pi[\phi_k(t)+\phi_k'(t)(\tau-t)+\frac{1}{2}\phi_k''(t)(\tau-t)^2]} + (A_k(\tau)-A_k(t))e^{i2\pi\phi_k(\tau)} \\
&+ A_k(t)\left[e^{i2\pi[\phi_k(t)+\phi_k'(t)(\tau-t)+\frac{1}{2}\phi_k''(t)(\tau-t)^2+\frac{1}{2}\int_t^\tau \phi_k'''(x)(\tau-x)^2dx]} - e^{i2\pi[\phi_k(t)+\phi_k'(t)(\tau-t)+\frac{1}{2}\phi_k''(t)(\tau-t)^2]}\right] \\
&= f_{k,1}(\tau) + f_{k,2}(\tau) + f_{k,3}(\tau).
\end{aligned}$$

Then, for any $(t,a) \in \mathbb{R}\times\mathbb{R}^+$, one has:

$$W_{f_{k,1}}^{\tau^r\psi^{(p)}(\tau)}(t,a) = f_k(t)\overline{\mathcal{F}\{\tau^r\psi^{(p)}(\tau)e^{-i\pi\phi_k''(t)a^2\tau^2}\}(a\phi_k'(t))},$$

$$\left|W_{f_{k,2}}^{\tau^r\psi^{(p)}(\tau)}(t,a)\right| \leq \varepsilon a^{r+1}MJ_{r+1,p},$$

and

$$|W_{f_{k,3}}^{\tau^r\psi^{(p)}(\tau)}(t,a)| \leq \varepsilon\frac{\pi}{3}a^{r+3}MA_k(t)J_{r+3,p},$$

from which one easily gets the inequality (A.6). □

Item (a) follows from this proposition remarking that if $(t,a) \in \mathbb{E}\setminus\bigcup_{l=1}^{K} Z_l$:

$$|W_f^\psi(t,a)| \leq \varepsilon(E_{0,0}(t,a) + N_{0,0}\sum_{k=1}^{K} A_k(t)) \leq \tilde{\varepsilon},$$

when $\tilde{\varepsilon}$ is sufficiently small, i.e.:

$$\tilde{\varepsilon} \leq \frac{1}{\sqrt{2}}\min\left(\| E_{0,0}(t,a)\|_{\infty,\mathbb{E}}^{-\frac{1}{2}}, \left\| N_{0,0}\sum_{k=1}^{K} A_k(t)\right\|_{\infty,\mathbb{E}}^{-\frac{1}{2}}\right) \tag{A.7}$$

because $E_{0,0}(t,a)$ is bounded on $\mathbb{E}$.

Now, to prove item (b) of Theorem 3.11, we remark that Proposition A.4 rewrites for any $(t,a) \in \mathbb{R}\times\mathbb{R}^+$:

$$\left|W_f^{\tau^r\psi^{(p)}(\tau)}(t,a) - \sum_{k=1}^{K} W_{f_{k,1}}^{\tau^r\psi^{(p)}(\tau)}(t,a)\right| \leq \varepsilon E_{r,p}(t,a), \tag{A.8}$$

which rewrites when $r=0$ using an integration by parts:

$$\left|W_f^{\psi^{(p)}}(t,a) + 2i\pi a\sum_{k=1}^{K}\left(a\phi_k''(t)W_{f_{k,1}}^{\tau\psi^{(p-1)}(\tau)}(t,a) + \phi_k'(t)W_{f_{k,1}}^{\psi^{(p-1)}}(t,a)\right)\right| \leq \varepsilon E_{0,p}(t,a). \tag{A.9}$$

From equation (A.8), we deduce that if $(t,a) \in Z_k$,

$$\left|W_f^{\tau^r \psi^{(p)}(\tau)}(t,a) - W_{f_{k,1}}^{\tau^r \psi^{(p)}(\tau)}(t,a)\right| \leq \varepsilon \left(E_{r,p}(t,a) + \sum_{l \neq k} A_l(t) N_{r,p} \right). \tag{A.10}$$

Proposition A.5. *For any $(t,a) \in Z_k$ such that $|W_f^{\psi}(t,a)| > \tilde{\varepsilon}$ and $|\partial_t \widetilde{\tau}_f(t,a)| > \tilde{\varepsilon}$ one has:*

$$|\widetilde{q}_{t,f}(t,a) - \phi_k''(t)| \leq \tilde{\varepsilon}. \tag{A.11}$$

Proof. For any $(t,a) \in Z_k$ one has, using (A.9) and (A.10):

$$\begin{aligned}
&|\phi_k''(t) - \widetilde{q}_{t,f}(t,a)| \\
&= \left| \frac{1}{2\pi a^2} \frac{W_f^{\psi'}\left[W_f^{\psi'} + i2\pi a(a\phi_k''(t)W_f^{t\psi} + \phi_k'(t)W_f^{\psi})\right] - W_f^{\psi}\left[W_f^{\psi''} + i2\pi a(a\phi_k''(t)W_f^{t\psi'} + \phi_k'(t)W_f^{\psi'})\right]}{W_f^{t\psi}W_f^{\psi'} - W_f^{t\psi'}W_f^{\psi}} \right| \\
&\leq \left| \frac{1}{2\pi a^2} \frac{W_f^{\psi'}\left[W_f^{\psi'} + i2\pi a(a\phi_k''(t)W_{f_{k,1}}^{t\psi} + \phi_k'(t)W_{f_{k,1}}^{\psi})\right] - W_f^{\psi}\left[W_f^{\psi''} + i2\pi a(a\phi_k''(t)W_{f_{k,1}}^{t\psi'} + \phi_k'(t)W_{f_{k,1}}^{\psi'})\right]}{W_f^{t\psi}W_f^{\psi'} - W_f^{t\psi'}W_f^{\psi}} \right| \\
&+ \frac{1}{a} \frac{\left|a\phi_k''(t)W_f^{\psi'}(W_f^{t\psi} - W_{f_{k,1}}^{t\psi})\right| + \phi_k'(t)\left|W_f^{\psi} - W_{f_{k,1}}^{\psi}\right| + \left|a\phi_k''(t)W_f^{\psi}(W_f^{t\psi'} - W_{f_{k,1}}^{t\psi'})\right| + \phi_k'(t)\left|W_f^{\psi'} - W_{f_{k,1}}^{\psi'}\right|}{\left|W_f^{t\psi}W_f^{\psi'} - W_f^{t\psi'}W_f^{\psi}\right|} \\
&= \left| \frac{1}{2\pi a^2} \frac{W_f^{\psi'}(W_f^{\psi'} - W_{f_{k,1}}^{\psi'}) - W_f^{\psi}(W_f^{\psi''} - W_{f_{k,1}}^{\psi''})}{W_f^{t\psi}W_f^{\psi'} - W_f^{t\psi'}W_f^{\psi}} \right| \\
&+ \frac{1}{a} \frac{\left|a\phi_k''(t)W_f^{\psi'}(W_f^{t\psi} - W_{f_{k,1}}^{t\psi})\right| + \phi_k'(t)\left|W_f^{\psi} - W_{f_{k,1}}^{\psi}\right| + \left|a\phi_k''(t)W_f^{\psi}(W_f^{t\psi'} - W_{f_{k,1}}^{t\psi'})\right| + \phi_k'(t)\left|W_f^{\psi'} - W_{f_{k,1}}^{\psi'}\right|}{\left|W_f^{t\psi}W_f^{\psi'} - W_f^{t\psi'}W_f^{\psi}\right|} \\
&\leq \frac{\varepsilon\left(\left(\left|\frac{W_f^{\psi'}}{2\pi a^2}\right| + \left|\frac{\phi_k'(t)}{a}\right|\right)E_{0,1}(t,a) + \left|\frac{W_f^{\psi}}{2\pi a^2}\right|E_{0,2}(t,a)| + \left|\phi_k''(t)W_f^{\psi'}\right|E_{1,0}(t,a) + \frac{\phi_k'(t)}{a}E_{0,0}(t,a) + |\phi_k''(t)|\,E_{1,1}(t,a)\right)}{\left|W_f^{t\psi}W_f^{\psi'} - W_f^{t\psi'}W_f^{\psi}\right|} \\
&+ \varepsilon\frac{\sum_{l\neq k} A_l(t)\left(\left(\left|\frac{W_f^{\psi'}}{2\pi a^2}\right| + \left|\frac{\phi_k'(t)}{a}\right|\right)N_{0,1} + \left|\frac{W_f^{\psi}}{2\pi a^2}\right|N_{0,2}| + \left|\phi_k''(t)W_f^{\psi'}\right|N_{1,0} + \frac{\phi_k'(t)}{a}N_{0,0} + |\phi_k''(t)|\,N_{1,1}\right)}{\left|W_f^{t\psi}W_f^{\psi'} - W_f^{t\psi'}W_f^{\psi}\right|} \\
&\leq \frac{\varepsilon\left(\left(\left|\frac{W_f^{\psi'}}{2\pi a^2}\right| + \left|\frac{\phi_k'(t)}{a}\right|\right)E_{0,1}(t,a) + \left|\frac{W_f^{\psi}}{2\pi a^2}\right|E_{0,2}(t,a)| + \left|\phi_k''(t)W_f^{\psi'}\right|E_{1,0}(t,a) + \frac{\phi_k'(t)}{a}E_{0,0}(t,a) + |\phi_k''(t)|\,E_{1,1}(t,a)\right)}{\widetilde{\varepsilon}^3} \\
&+ \varepsilon\frac{\sum_{l\neq k} A_l(t)\left(\left(\left|\frac{W_f^{\psi'}}{2\pi a^2}\right| + \left|\frac{\phi_k'(t)}{a}\right|\right)N_{0,1} + \left|\frac{W_f^{\psi}}{2\pi a^2}\right|N_{0,2}| + \left|\phi_k''(t)W_f^{\psi'}\right|N_{1,0} + \frac{\phi_k'(t)}{a}N_{0,0} + |\phi_k''(t)|\,N_{1,1}\right)}{\widetilde{\varepsilon}^3} \leq \widetilde{\varepsilon},
\end{aligned}$$

if $\widetilde{\varepsilon}$ is sufficiently small, the last inequality being obtained by remarking the numerator is bounded on $\mathbb{E}$. Note also, that $|\phi_k''(t) - \widetilde{q}_{t,f}(t,a)|$ is of the order of $\widetilde{\varepsilon}^3$ if ε is sufficiently small. □

Proof of item (b): according to definition of $\widetilde{\omega}_{t,f}^{[2]}(t,a)$ in (3.8), one has:

$$\widetilde{\omega}_{t,f}^{[2]}(t,a) = \widetilde{\omega}_f(t,a) + \widetilde{q}_{t,f}(t,a)(t - \widetilde{\tau}_f(t,a)).$$

It follows that for $(t,a) \in Z_k$, such that $|W_f^{\psi}(t,a)| > \tilde{\varepsilon}$ and $|\partial_t \widetilde{\tau}_f(t,a)| > \tilde{\varepsilon}$

$$\begin{aligned}
&\left|\widetilde{\omega}_{t,f}^{[2]}(t,a) - \phi_k'(t)\right| \\
&\quad = \left|\frac{1}{i2\pi a}\frac{W_f^{\psi'}(t,a) + i2\pi a\phi_k'(t)W_f^{\psi}(t,a) + i2\pi a^2\phi_k''(t)W_f^{t\psi}(t,a)}{W_f^{\psi}(t,a)}\right| + \left|a\frac{(\widetilde{q}_{t,f}(t,a) - \phi_k''(t))\,W_f^{t\psi}(t,a)}{W_f^{\psi}(t,a)}\right| \\
&\leq \frac{\frac{1}{2\pi a}\left|W_f^{\psi'}(t,a) - W_{f_{k,1}}^{\psi'}(t,a)\right| + \phi_k'(t)\left|W_f^{\psi}(t,a) - W_{f_{k,1}}^{\psi}(t,a)\right| + a\left|\phi_k''(t)\right|\left|W_f^{t\psi(t)}(t,a) - W_{f_{k,1}}^{t\psi(t)}(t,a)\right|}{W_f^{\psi}(t,a)} \\
&\qquad\qquad + \left|a\frac{(\widetilde{q}_{t,f}(t,a) - \phi_k''(t))\,W_f^{t\psi}(t,a)}{W_f^{\psi}(t,a)}\right| \\
&\leq \widetilde{\varepsilon}^5\left(\frac{1}{2\pi a}(E_{0,1} + \sum_{l\neq k} A_l(t)N_{0,1})\right) + \phi_k'(t)\left(E_{0,0} + \sum_{l\neq k} A_l(t)N_{0,0}\right) + a\left|\phi_k''(t)\right|\left(E_{0,1} + \sum_{l\neq k} A_l(t)N_{1,1}\right) \\
&\qquad\qquad + \left|a\frac{(\widetilde{q}_{t,f}(t,a) - \phi_k''(t))\,W_f^{t\psi}(t,a)}{\widetilde{\varepsilon}}\right| \leq \tilde{\varepsilon}
\end{aligned}$$

when $\tilde{\varepsilon}$ is sufficiently small.

Proof of item (c): It is exactly the same as in the weak modulation case (item (c) of Theorem 3.4), except that we use, at the very end of the proof, the following hypotheses:

$$\left|\lim_{\lambda\to 0}\left(\frac{1}{C'_{\psi,k}}\int_{|\omega-\phi'_k(t)|<\tilde{\varepsilon}} S^{\psi,\lambda,\tilde{\varepsilon}}_{2,f}(t,\omega)d\omega\right)-f_k(t)\right| = \left|\frac{1}{C'_{\psi,k}}\int_{\mathbb{X}} W^{\psi}_f(t,a)\frac{da}{a}-f_k(t)\right|$$

$$\leq \left[\left|\frac{1}{C'_{\psi,k}}\int_{|a\phi'_k(t)-1|<\Delta} W^{\psi}_f(t,a)\frac{da}{a}-f_k(t)\right|+\left|\frac{1}{C'_{\psi,k}}\int_{|W^{\psi}_f(t,a)|\leq\widetilde{\varepsilon}\cap|a\phi'_k(t)-1|<\Delta} W^{\psi}_f(t,a)\,\frac{da}{a}\right|\right]$$

$$\leq \frac{1}{\left|C'_{\psi,k}\right|}\left[\int_{|a\phi'_k(t)-1|<\Delta}\left|W^{\psi}_f(t,a)-f_k(t)\overline{\mathcal{F}\{\psi(\tau)e^{-i\pi\phi''_k(t)a^2\tau^2}\}(a\phi'_k(t))}\right|\frac{da}{a}+\right.$$

$$\left. A_k(t)\int_{|a\phi'_k(t)-1|\geq\Delta}\left|\overline{\mathcal{F}\{\psi(\tau)e^{-i\pi\phi''_k(t)a^2\tau^2}\}(a\phi'_k(t))}\right|\frac{da}{a}+\widetilde{\varepsilon}\log\left(\frac{1+\Delta}{1-\Delta}\right)\right]$$

$$\leq \frac{1}{\left|C'_{\psi,k}\right|}\left[\varepsilon\int_{|a\phi'_k(t)-1|<\Delta}\left(E_{0,0}(t,a)+N_{0,0}\sum_{l\neq k}A_l(t)\right)\frac{da}{a}++\widetilde{\varepsilon}(1+A_k(t)N_2)\log\left(\frac{1+\Delta}{1-\Delta}\right)\right]\leq D_2\widetilde{\varepsilon}.$$

A.3 Proofs of Theorem 3.13 on page 48

In such a case, one has, using a second order Taylor expansion of the phase of h_c:

$$\begin{aligned}
W^{\psi}_{h_c}(t,a) &= \int_{\mathbb{R}} Ae^{2i\pi\phi(\tau)}\sigma_W^{-1}e^{-\frac{\pi}{\sigma_W^2}\left(\frac{\tau-t}{a}\right)^2}e^{-2i\pi\left(\frac{\tau-t}{a}\right)}d\tau\\
&= h_c(t)a\sigma_W^{-1}\mathcal{F}\left\{e^{-\pi\left[\frac{1}{\sigma_W^2}-ia^2\phi''(t)\right]u^2}\right\}(1-a\phi'(t))\\
&= h_c(t)a\sigma_W^{-1}\left(\frac{1}{\sigma_W^2}-ia^2\phi''(t)\right)^{-\frac{1}{2}}e^{\left[\frac{-\pi\sigma_W^2(1-a\phi'(t))^2}{1-i\sigma_W^2a^2\phi''(t)}\right]}.
\end{aligned}$$

With that expression, one can compute the estimate, bearing in mind that $\phi''(t)$ is constant and $\Im\mathrm{m}(z)$ is the imaginary part of complex number z:

$$\begin{aligned}
\widehat{\omega}_{h_c}(t,a) &= \Re\mathrm{e}\left\{\frac{1}{i2\pi}\frac{\partial_t W^{\psi}_{h_c}(t,a)}{W^{\psi}_{h_c}(a,t)}\right\}=\frac{1}{2\pi}\Im\mathrm{m}\left\{\frac{\partial_t W^{\psi}_{h_c}(t,a)}{W^{\psi}_{h_c}(t,a)}\right\}\\
&= \frac{1}{2\pi}\Im\mathrm{m}\left\{\frac{h_c{}'(t)}{h_c(t)}+\frac{2\pi\sigma_W^2a\phi''(t)(1-a\phi'(t))}{1-i\sigma_W^2a^2\phi''(t)}\right\}\\
&= \frac{1}{2\pi}\Im\mathrm{m}\left\{2i\pi\phi'(t)\right\}+\left(\frac{\sigma_W^2a\phi''(t)(1-a\phi'(t))}{1+\sigma_W^4a^4\phi''(t)^2}(\sigma_W^2a^2\phi''(t))\right).
\end{aligned}$$

From this one derives:

$$|\widehat{\omega}_{h_c}(t,a) - \phi^{'}(t)| \quad = \quad \left| \frac{\sigma_W^4 a^4 \phi^{''}(t)^2 \left(\frac{1}{a} - \phi^{'}(t)\right)}{1 + \sigma_W^4 a^4 \phi^{''}(t)^2} \right|.$$

If $|\frac{1}{a} - \phi^{'}(t)| \leq \Delta$ then $|\widehat{\omega}_{h_c}(t,a) - \phi^{'}(t)| \leq \Delta \ \left|1 - \frac{1}{1+\sigma_W^4 a^4 \phi^{''}(t)^2}\right| \leq \Delta.$

A.4 Proofs of Theorem 3.14 on page 48

Note first that when β is an integer, g_β admits the following Fourier transform $\hat{g}_\beta(\eta) = \eta^\beta e^{-\eta} H(\eta)$ with H the Heaviside function. So, even if $\hat{g}_\beta$ is not compactly supported, it has a fast decay (the behavior of $\hat{g}_\beta$ being similar for non integer β).

First, let us first consider $f(t) = t^\alpha H(t)$, for any $\alpha > -1$ and let us compute its wavelet transform with the Cauchy wavelet with $\beta > \alpha$:

$$\begin{aligned}
W_f^{g_\beta}(t,a) &= \frac{1}{a} \int_0^\infty \tau^\alpha \overline{g_\beta \left(\frac{\tau - t}{a}\right)} d\tau \\
&= a^\beta \Gamma(\beta+1) \int_0^\infty \tau^\alpha \frac{1}{(a + i2\pi(\tau - t))^{-(1+\beta)}} d\tau \\
&= a^\beta \Gamma(\beta+1) \int_0^\infty \tau^\alpha \left(\frac{-i}{(-2\pi t - ia) + 2\pi\tau}\right)^{1+\beta} d\tau \\
&= a^\beta \Gamma(\beta+1) \int_0^\infty \tau^\alpha \left(\frac{-i}{z + 2\pi\tau}\right)^{1+\beta} d\tau,
\end{aligned}$$

with $z = -2\pi t - ia$, so that the wavelet transform can be viewed as analytic function in z.

Now, assuming z is a positive integer, one can rewrite the wavelet transform, making the appropriate change of variable, as:

$$\begin{aligned}
W_f^{g_\beta}(t,a) &= (2\pi)^{-\alpha-1} a^\beta \Gamma(\beta+1)(-i)^{1+\beta} z^{\alpha-\beta} \int_0^\infty \tau^\alpha \left(\frac{1}{1+\tau}\right)^{1+\beta} d\tau \\
&= (2\pi)^{-\alpha-1} a^\beta \Gamma(\beta+1)(-i)^{1+\beta} z^{\alpha-\beta} B(\alpha+1, \beta-\alpha) \\
&= (2\pi)^{-\alpha-1} a^\beta \Gamma(\beta+1) B(\alpha+1, \beta-\alpha) e^{-i\pi\frac{1+\beta}{2}} z^{\alpha-\beta},
\end{aligned}$$

where B is the beta function. The expression is then also true when z is complex using the analytic continuation theorem.

Now using the analytic continuation theorem with variable α we have for the hyperbolic chirp $f(t) = t^{i\alpha} H(t)$, $\alpha \in \mathbb{R}$:

$$\begin{aligned}
W_f^{g_\beta}(t,a) &= (2\pi)^{-i\alpha-1} a^\beta \Gamma(\beta+1) B(i\alpha+1, \beta-i\alpha) e^{-i\pi\frac{1+\beta}{2}} z^{i\alpha-\beta}, \\
&= (2\pi)^{-i\alpha-1} a^\beta \Gamma(\beta+1) B(i\alpha+1, \beta-i\alpha) e^{-i\pi\frac{1+\beta}{2}} (-2\pi t - ia)^{i\alpha-\beta}.
\end{aligned}$$

Note that this expression is valid for any positive β, regardless of α

Using that expression for the wavelet transform, we may then write:

$$\widehat{\omega}_f(t,a) \;=\; \frac{1}{2\pi}\mathcal{I}\left[\frac{\partial_t W_f^{g_\beta}(t,a)}{W_f^{g_\beta}(t,a)}\right] = \mathcal{I}\left[-(i\alpha-\beta)(2\pi t+ia)^{-1}\right] = \frac{\alpha 2\pi t + a\beta}{a^2+4\pi^2t^2}$$

Thus one can deduce: $\left|\widehat{\omega}_f(t,a) - \frac{\alpha}{2\pi t}\right| = |\widehat{\omega}_f(a,t) - \phi'(t)| = \left|\frac{\frac{\beta}{a}-\frac{\alpha}{2\pi t}}{1+4\pi^2(\frac{t}{a})^2}\right|$.
Finally, if $|\frac{1}{a} - \phi^{'}(t)| \leq \Delta$ and $\beta = 1$: $|\widehat{\omega}_f(t,a) - \phi'(t)| \leq \frac{\Delta}{1+4\pi^2(\frac{t}{a})^2} \leq \Delta$.

Appendix B

Proofs of Section 3.2: High-order STFT-based Synchrosqueezing Transforms

B.1 The proof of Proposition 3.21 on page 57

Proof. First of all, we rewrite the expression (3.35) under matrix form:

$$\widetilde{\omega}_f(t,\eta) = \mathbf{X}_N(t,\eta) \cdot \mathbf{R}_N(t)^T$$

where $\mathbf{Z}^T$ is the transpose of matrix $\mathbf{Z}$ and the two row vectors $\mathbf{X}, \mathbf{R}$ defined as:

$$\mathbf{X}_N(t,\eta) = [1 \quad x_{2,1}(t,\eta) \quad ... \quad x_{N,1}(t,\eta)]$$
$$\mathbf{R}_N(t) = [r_1(t) \quad r_2(t) \quad ... \quad r_N(t)]$$

Let us denote $y_1 = \mathbf{X}_N \cdot \mathbf{R}_N^T$, we may thus write:

$$y_1 = \begin{bmatrix} x_{1,1} & x_{2,1} & x_{3,1} & \dots & x_{N,1} \end{bmatrix} \mathbf{R}_N^T. \tag{B.1}$$

It is noteworthy that $\Re\mathrm{e}\{r_1(t)\} = \phi'(t)$. To get r_k, we build up a system of N equations with variables r_k for $k = 1, \dots, N$ from (B.1) using the following procedure. By computing the partial derivatives of (B.1) with respect to η and using notation $y_2 = \dfrac{\partial_\eta y_1}{\partial_\eta x_{2,1}}$ and $x_{k,2} = \dfrac{\partial_\eta x_{k,1}}{\partial_\eta x_{2,1}}$, the second equation can be obtained:

$$y_2 = \begin{bmatrix} 0 & 1 & x_{3,2} & \dots & x_{N,2} \end{bmatrix} \mathbf{R}_N^T.$$

Doing the same thing iteratively, we can get the jth equation:

$$y_j = \begin{bmatrix} 0 & 0 & \dots & 1 & \dots & x_{N,j} \end{bmatrix} \mathbf{R}_N^T.$$

Combining all these equations, the desired system of equations can be deduced:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{N-1} \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & x_{2,1} & x_{3,1} & \dots & x_{N,1} \\ 0 & 1 & x_{3,2} & \dots & x_{N,2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & x_{N,N-1} \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_{N-1} \\ r_N \end{bmatrix}$$

or

$$\Big[\mathbf{Y}_N\Big] = \Big[\mathbf{MX}\Big] \Big[\mathbf{R}_N\Big]^T. \tag{B.2}$$

Since **MX** is a upper triangular matrix with nonzero diagonal coefficients, we use back-substitution algorithm to get r_k for $k = 1, \dots, N$ as follows:

$$r_N(t) = y_N(t,\eta) \text{ and } r_j(t) = y_j(t,\eta) - \sum_{k=j+1}^{N} x_{k,j}(t,\eta) r_k(t) \quad \text{for } j = N-1, N-2, \dots, 1.$$

As a result, $\widetilde{q}_{\eta,f}^{[k,N]}(t,\eta) = r_k(t)$ for $k = 2, \dots, N$.

From (3.35), we get: $\Re\mathrm{e}\left\{\widetilde{q}_{\eta,f}^{[k,N]}(t,\eta)\right\} = \dfrac{\phi^{(k)}(t)}{(k-1)!}$ for $k = 2, \dots, N$, which finishes the proof. □

B.2 Proof of the Proposition 3.24 on page 58

Proof. By using $\partial_\eta V_f^{t^{k-1}g} = -i2\pi V_f^{t^k g}$ for $k \in \mathbb{N}$ and defining $X_{k,j} = V_f^g V_f^{t^k g} - V_f^{t^{j-1}g} V_f^{t^{k-j+1}g}$, we get the following formula:

$$\partial_\eta X_{k,j} = -i2\pi \left(X_{k+1,j} + X_{k+1,j} - X_{k+1,2}\right).$$

Thus, the upper triangular part of matrix **MX** defined in (B.2) with $N = 4$ can be obtained as follows:

$$\begin{aligned}
x_{k,1} &= \frac{V_f^{t^{k-1}g}}{V_f^g} && \text{for } k = 1 \dots 4, \\
x_{k,2} &= \frac{\partial_\eta x_{k,1}}{\partial_\eta x_{2,1}} = \frac{V_f^g V_f^{t^k g} - V_f^{tg} V_f^{t^{k-1}g}}{V_f^g V_f^{t^2 g} - \left(V_f^{tg}\right)^2} = \frac{X_{k,2}}{X_{2,2}} && \text{for } k = 2 \dots 4, \\
x_{k,3} &= \frac{\partial_\eta x_{k,2}}{\partial_\eta x_{3,2}} = \frac{X_{k+1,3} X_{2,2} - X_{k,2} X_{3,3}}{X_{4,3} X_{2,2} - X_{3,2} X_{3,3}} && \text{for } k = 3 \dots 4, \\
x_{k,4} &= 1 && \text{for } k = 4
\end{aligned}$$

Also, the elements of vector $\mathbf{Y}$ are obtained by:

$$
\begin{aligned}
y_1 &= \widetilde{\omega}_f = \eta - \frac{1}{i2\pi}\frac{V_f^{g'}}{V_f^g},\\
y_2 &= \frac{\partial_\eta y_1}{\partial_\eta x_{2,1}}\\
&= \frac{1}{i2\pi}\frac{\left(V_f^g\right)^2 + V_f^g V_f^{tg'} - V_f^{tg}V_f^{g'}}{V_f^g V_f^{t^2 g} - \left(V_f^{tg}\right)^2} = \frac{W_2}{X_{2,2}},\ \text{with}\ \ W_2 = \frac{1}{i2\pi}\left[\left(V_f^g\right)^2 + V_f^g V_f^{tg'} - V_f^{tg}V_f^{g'}\right].\\
y_3 &= \frac{\partial_\eta y_2}{\partial_\eta x_{3,2}} = \frac{W_3 X_{2,2} - W_2 X_{3,3}}{X_{4,3}X_{2,2} - X_{3,2}X_{3,3}},\ \text{with}\ \ W_3 = \partial_\eta W_2.\\
y_4 &= \frac{\partial_\eta y_3}{\partial_\eta x_{4,4}} = \frac{\begin{array}{r}(X_{4,3}X_{2,2} - X_{3,2}X_{3,3})\,W_4\\ -\,(W_3X_{2,2} - W_2X_{3,3})\,(X_{5,4} + X_{5,3} - X_{5,2})\\ +\,(W_3X_{3,2} - W_2X_{4,3})\,(X_{4,4} + X_{4,3} - X_{4,2})\end{array}}{\begin{array}{r}(X_{4,3}X_{2,2} - X_{3,2}X_{3,3})\,(X_{6,4} + X_{6,3} - X_{6,2})\\ -\,(X_{5,3}X_{2,2} - X_{4,2}X_{3,3})\,(X_{5,4} + X_{5,3} - X_{5,2})\\ +\,(X_{5,3}X_{3,2} - X_{4,2}X_{4,3})\,(X_{4,4} + X_{4,3} - X_{4,2})\end{array}},\ \text{with}\ \ W_4 = \partial_\eta W_3.
\end{aligned}
$$

With the help of back-substitution algorithm, the modulation operators read:

$$
\begin{aligned}
\widetilde{q}_{\eta,f}^{[4,4]} &= \frac{\begin{array}{r}(X_{4,3}X_{2,2} - X_{3,2}X_{3,3})\,W_4\\ -\,(W_3X_{2,2} - W_2X_{3,3})\,(X_{5,4} + X_{5,3} - X_{5,2})\\ +\,(W_3X_{3,2} - W_2X_{4,3})\,(X_{4,4} + X_{4,3} - X_{4,2})\end{array}}{\begin{array}{r}(X_{4,3}X_{2,2} - X_{3,2}X_{3,3})\,(X_{6,4} + X_{6,3} - X_{6,2})\\ -\,(X_{5,3}X_{2,2} - X_{4,2}X_{3,3})\,(X_{5,4} + X_{5,3} - X_{5,2})\\ +\,(X_{5,3}X_{3,2} - X_{4,2}X_{4,3})\,(X_{4,4} + X_{4,3} - X_{4,2})\end{array}}\\
\widetilde{q}_{\eta,f}^{[3,4]} &= \frac{W_3X_{2,2} - W_2X_{3,3}}{X_{4,3}X_{2,2} - X_{3,2}X_{3,3}} - \widetilde{q}_{\eta,f}^{[4,4]}\frac{X_{5,3}X_{2,2} - X_{4,2}X_{3,3}}{X_{4,3}X_{2,2} - X_{3,2}X_{3,3}}\\
\widetilde{q}_{\eta,f}^{[2,4]} &= \frac{W_2}{X_{2,2}} - \widetilde{q}_{\eta,f}^{[3,4]}\frac{X_{3,2}}{X_{2,2}} - \widetilde{q}_{\eta,f}^{[4,4]}\frac{X_{4,2}}{X_{2,2}}.
\end{aligned}
$$

Finally, we complete the proof of this proposition by using notation G_k and $G_{j,k}$ to rewrite the above expressions. □

Appendix C

Proof of Section 4.2: On Downsampled STFT

C.1 The proof of formula (4.14) on page 85

If g is such that, for any n, $\sum_{m\in\mathbb{Z}} g[n-mR]^2 = 1$, from (4.12) and putting $q = n+mR+p-M$, one has:

$$f[q]g[q-mR-p] = \frac{1}{N}\sum_{k=0}^{N-1} V_f^g(mR+p, \frac{k}{N})e^{i2\pi\frac{k(q-mR-p)}{N}}.$$

From this we deduce:

$$f[q]\sum_{m\in\mathbb{Z}} g[q-mR-p]^2 = \sum_{m\in\mathbb{Z}}\frac{1}{N}\sum_{k=0}^{N-1} V_f^g(mR+p, \frac{k}{N})e^{i2\pi\frac{k(q-mR-p)}{N}} g[q-mR-p]$$

$$\Leftrightarrow f[q] = \frac{\sum_{m\in\mathbb{Z}}\frac{1}{N}\sum_{k=0}^{N-1} V_f^g(mR+p, \frac{k}{N})e^{i2\pi\frac{k(q-mR-p)}{N}} g[q-mR-p]}{\sum_{m\in\mathbb{Z}} g[q-mR-p]^2}.$$

If $\sum_{m\in\mathbb{Z}} g[n-mR]^2 = 1$, we get (4.14), and if it is only non-zero for any n, (4.16) follows.

C.2 The proof of formula (4.15) on page 85

Regarding energy conservation, from (4.12) we may write:

$$\sum_{m\in\mathbb{Z}}\sum_{q=mR+p-M}^{N-1+mR+p-M} f[q]^2 g[q-mR-p]^2 = \sum_{m\in\mathbb{Z}}\sum_{q=mR+p-M}^{N-1+mR+p-M} |\frac{1}{N}\sum_{k=0}^{N-1} V_f^g(mR+p, \frac{k}{N})e^{i2\pi\frac{k(q-mR-p)}{N}}|^2$$

$$\Leftrightarrow \sum_{m\in\mathbb{Z}}\sum_{q\in\mathbb{Z}} f[q]^2 g[q-mR-p]^2 = \sum_{m\in\mathbb{Z}} \frac{1}{N}\sum_{k=0}^{N-1} |V_f^g(mR+p,\frac{k}{N})|^2$$

$$\Leftrightarrow \sum_{q\in\mathbb{Z}} f[q]^2 = \sum_{m\in\mathbb{Z}} \frac{1}{N}\sum_{k=0}^{N-1} |V_f^g(mR+p,\frac{k}{N})|^2.$$

C.3 The proof of formula (4.17) on page 86

The proof of (4.17) is similar to that of (4.14). Indeed, we have:

$$f[q]\sum_{m\in\mathbb{Z}} g[q-mR-p] = \sum_{m\in\mathbb{Z}} \frac{1}{N}\sum_{k=0}^{N-1} V_f^g(mR+p,\frac{k}{N}) e^{i2\pi\frac{k(q-mR-p)}{N}}$$

$$\Leftrightarrow f[q] = \frac{\sum_{m\in\mathbb{Z}} \frac{1}{N}\sum_{k=0}^{N-1} V_f^g(mR+p,\frac{k}{N}) e^{i2\pi\frac{k(q-mR-p)}{N}}}{\sum_{m\in\mathbb{Z}} g[q-mR-p]}.$$

If $\sum_{m\in\mathbb{Z}} g[n-mR] = 1$ for any n, then one obtains (4.17) and otherwise if the sum is only non-zero (4.18) follows.

References

[1] S. Meignen and D.-H. Pham, "Retrieval of the modes of multicomponent signals from downsampled short-time fourier transform," *IEEE Transactions on Signal Processing*, 2018. Accepted.

[2] D. H. Pham, S. Meignen, N. Dia, J. Fontecave-Jallon, and B. Rivet, "Phonocardiogram signal denoising based on non-negative matrix factorization and adaptive contour representation computation," *IEEE Signal Processing Letters*, vol. 25, pp. 1475–1479, October 2018.

[3] D.-H. Pham and S. Meignen, "Second-order synchrosqueezing transform: The wavelet case, comparisons and applications," *hal-01586372*, 2018. Submitted.

[4] D.-H. Pham and S. Meignen, "An adaptive computation of contour representations for mode decomposition," *IEEE Signal Processing Letters*, vol. 24, pp. 1596–1600, November 2017.

[5] D.-H. Pham and S. Meignen, "High-order synchrosqueezing transform for multicomponent signals analysis - with an application to gravitational-wave signal," *IEEE Transactions on Signal Processing*, vol. 65, pp. 3168–3178, June 2017.

[6] S. Meignen, D.-H. Pham, and S. McLaughlin, "On demodulation, ridge detection and synchrosqueezing for multicomponent signals," *IEEE Transactions on Signal Processing*, vol. 65, no. 8, pp. 2093–2103, 2017.

[7] D.-H. Pham and S. Meignen, "A novel thresholding technique for the denoising of multicomponent signals," in *IEEE ICASSP, Calgary, Alberta, Canada*, 2018.

[8] D.-H. Pham, "Poster for signal processing meets deep learning," in *IEEE-EURASIP Summer School, Capri, Italy*, 2017.

[9] D.-H. Pham, "Analysis, synthesis and transformations by reassignment, emd and synchrosqueezing," in *Workshop ASTRES,Grenoble*, 2016.

[10] F. Auger, P. Flandrin, P. Gonçalvès, and O. Lemoine, *The Time-Frequency Toolbox (TFTB) - For Use with MATLAB*. CNRS (France) and Rice University (USA), 1996.

[11] J. R. Mitchell and J.-J. Wang, "Expanding application of the wiggers diagram to teach cardiovascular physiology," *Advances in physiology education*, vol. 38, no. 2, pp. 170–175, 2014.

[12] L. Cohen, *Time-frequency Analysis: Theory and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1995.

[13] B. Boashash, *Time frequency signal analysis and processing - A comprehensive reference*. Gulf Professional Publishing, 2003.

[14] S. V. Vaseghi, *Advanced digital signal processing and noise reduction.* John Wiley & Sons, 2008.

[15] E. Sejdić, I. Djurović, and J. Jiang, "Time–frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Processing*, vol. 19, no. 1, pp. 153–183, 2009.

[16] A. Haar, "Zur theorie der orthogonalen funktionensysteme," *Mathematische Annalen*, vol. 69, no. 3, pp. 331–371, 1910.

[17] D. Gabor, "Theory of communication. part 1: The analysis of information," *Electrical Engineers - Part III: Radio and Communication Engineering, Journal of the Institution of*, vol. 93, pp. 429–441, November 1946.

[18] W. W. Au and M. C. Hastings, *Principles of marine bioacoustics.* Springer, 2008.

[19] A. W. Rihaczek, "Principles of high-resolution radar, artech house radar library," *Artech House, Norwood, MA, USA*, 1996.

[20] E. J. Candes, P. R. Charlton, and H. Helgason, "Detecting highly oscillatory signals by chirplet path pursuit," *Applied and Computational Harmonic Analysis*, vol. 24, no. 1, pp. 14–40, 2008.

[21] B. P. Abbott and al., "Observation of gravitational waves from a binary black hole merger," *Phys. Rev. Lett.*, vol. 116, 2016.

[22] N. Dia, J. Fontecave-Jallon, P.-Y. Guméry, and B. Rivet, "Application de la factorisation non-négative des matrices (nmf) pour le débruitage des signaux phonocardiographiques," *XXVIeme Colloque GRETSI à Juan-Les-Pins*, Sep 2017.

[23] 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.

[24] 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.

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

[26] L. Cohen and C. Lee, "Instantaneous frequency, its standard deviation and multicomponent signals," in *Advanced Algorithms and Architectures for Signal Processing III*, vol. 975, pp. 186–209, International Society for Optics and Photonics, 1988.

[27] L. Cohen, "What is a multicomponent signal?," in *Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on*, vol. 5, pp. 113–116, IEEE, 1992.

[28] D. Dimitriadis, P. Maragos, and A. Potamianos, "Robust am-fm features for speech recognition," *IEEE signal processing letters*, vol. 12, no. 9, pp. 621–624, 2005.

[29] V. Murray, C. Agurto, S. Barriga, M. S. Pattichis, and P. Soliz, "Real-time diabetic retinopathy patient screening using multiscale am-fm methods," in *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pp. 525–528, IEEE, 2012.

[30] S. Washimkar and S. Chede, "Application of fcm clustering on am-fm to detect mri disease progression for multiple sclerosis," in *Computational Intelligence and Communication Networks (CICN), 2015 International Conference on*, pp. 283–287, IEEE, 2015.

[31] S. Taran, V. Bajaj, and D. Sharma, "Teo separated am-fm components for identification of apnea eeg signals," in *Signal and Image Processing (ICSIP), 2017 IEEE 2nd International Conference on*, pp. 391–395, IEEE, 2017.

[32] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "The multicomponent am-fm image representation," *IEEE Transactions on Image Processing*, vol. 5, no. 6, pp. 1094–1100, 1996.

[33] M. S. Pattichis, G. Panayi, A. C. Bovik, and S.-P. Hsu, "Fingerprint classification using an am-fm model," *IEEE Transactions on Image Processing*, vol. 10, no. 6, pp. 951–954, 2001.

[34] K. Kodera, C. D. Villedary, and R. Gendrin, "A new method for the numerical analysis of non-stationary signals," *Physics of the Earth and Planetary Interiors*, vol. 12, no. 23, pp. 142 – 150, 1976.

[35] 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, pp. 64–76, Feb 1978.

[36] F. Auger and P. 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.

[37] 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.

[38] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 243–261, 2011.

[39] G. Thakur and H.-T. Wu, "Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples," *SIAM Journal on Mathematical Analysis*, vol. 43, p. 2078–2095, Jan 2011.

[40] G. Thakur, E. Brevdo, N. S. FučKar, and H.-T. Wu, "The synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications," *Signal Process*, vol. 93, pp. 1079–1094, May 2013.

[41] M. Clausel, T. Oberlin, and V. Perrier, "The monogenic synchrosqueezed wavelet transform: a tool for the decomposition/demodulation of am-fm images," *Applied and Computational Harmonic Analysis*, vol. 39, pp. 450–486, Nov 2015.

[42] H. Yang and L. Ying, "Synchrosqueezed wave packet transform for 2D mode decomposition," *SIAM Journal on Imaging Sciences*, vol. 6, pp. 1979–2009, Jan 2013.

[43] H. Yang, "Synchrosqueezed wave packet transforms and diffeomorphism based spectral analysis for 1d general mode decompositions," *Applied and Computational Harmonic Analysis*, vol. 39, pp. 33–66, July 2015.

[44] I. Daubechies, Y. Wang, and H.-T. Wu, "Conceft: concentration of frequency and time via a multitapered synchrosqueezed transform," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, Mar 2016.

[45] M. Skolnik, *Radar Handbook*. McGraw-Hill Education, 2008.

[46] J. W. Pitton, L. E. Atlas, and P. J. Loughlin, "Applications of positive time-frequency distributions to speech processing," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 4, pp. 554–566, 1994.

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

[48] 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, pp. 1335–1344, March 2015.

[49] R. Behera, S. Meignen, and T. Oberlin, "Theoretical analysis of the second-order synchrosqueezing transform," *Applied and Computational Harmonic Analysis*, Nov 2016.

[50] C. Li and M. Liang, "A generalized synchrosqueezing transform for enhancing signal time–frequency representation," *Signal Processing*, vol. 92, p. 2264–2274, Sep 2012.

[51] S. Wang, X. Chen, G. Cai, B. Chen, X. Li, and Z. He, "Matching demodulation transform and synchrosqueezing transform in time-frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp. 69–84, 2014.

[52] D. Groutage, J. Schempp, and L. Cohen, "Characterization and analysis of marine mammal sounds using time-frequency and time-prony techniques," in *OCEANS'94.'Oceans Engineering for Today's Technology and Tomorrow's Preservation.'Proceedings*, vol. 1, pp. I–253, IEEE, 1994.

[53] X. Rodet, "Time—domain formant—wave—function synthesis," in *Spoken Language Generation and Understanding*, pp. 429–441, Springer, 1980.

[54] T. F. Quatieri and R. J. McAulay, "Audio signal processing based on sinusoidal analysis/synthesis," in *Applications of digital signal processing to audio and acoustics*, pp. 343–416, Springer, 2002.

[55] N. H. Fletcher and T. D. Rossing, *The physics of musical instruments*. Springer Science & Business Media, 2012.

[56] S. L. Garrett, J. A. Smith, R. W. Smith, B. J. Heidrich, and M. D. Heibel, "Fission-powered in-core thermoacoustic sensor," *Applied Physics Letters*, vol. 108, no. 14, p. 144102, 2016.

[57] Y. Lim, B. Shinn-Cunningham, and T. Gardner, "Sparse contour representations of sound," *IEEE Signal Processing Letters*, vol. 19, pp. 684–687, Oct 2012.

[58] 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.

[59] 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.

[60] H. Fan, Y. Jiang, and G. Kuang, "Target detection of non-stationary radar signal and riemannian geometry," in *Geometric Science of Information*, pp. 821–828, Springer, 2013.

[61] J. G. Proakis, *Digital signal processing: principles algorithms and applications*. Pearson Education India, 2001.

[62] A. Vretblad, *Fourier analysis and its applications*, vol. 223. Springer Science & Business Media, 2006.

[63] 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, p. 68008, Mar 2007.

[64] 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, pp. 344 –347, Jan 2004.

[65] N. E. Huang and Z. Wu, "A review on Hilbert-huang transform: Method and its applications to geophysical studies," *Reviews of Geophysics*, vol. 46, Jun 2008.

[66] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu, "Time-frequency reassignment and synchrosqueezing: An overview," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 32–41, 2013.

[67] F. Hlawatsch and F. Auger, *Time-Frequency Analysis: Concepts and Methods*. Wiley-ISTE, 2008.

[68] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. i. fundamentals," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 520–538, 1992.

[69] E. Bedrosian, "A product theorem for hilbert transforms," *Proceedings of the IEEE*, vol. 51, no. 5, pp. 868–869, 1963.

[70] G. Xu, X. Wang, X. Xu, L. Zhou, and L. Shao, "Time-varying bandpass filter based on assisted signals for am-fm signal separation: A revisit," *Journal of Signal and Information Processing*, vol. 4, no. 03, p. 229, 2013.

[71] Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE M SP*, vol. 9, no. 2, pp. 21–67, 1992.

[72] S. W. Smith *et al.*, "The scientist and engineer's guide to digital signal processing," 1997.

[73] S. Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 3rd ed., 2008.

[74] T. Oberlin, *Analyse de signaux multicomposantes : contributions à la décomposition modale Empirique, aux représentations temps-fréquence et au Synchrosqueezing*. PhD thesis, Mathématiques Appliquées de l'Université de Grenoble, Novembre 2013.

[75] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1992.

[76] S. Meignen, T. Oberlin, and S. McLaughlin, "Multicomponent signal denoising with synchrosqueezing," in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, pp. 660–663, 2012.

[77] J. O. Smith, *Spectral audio signal processing.* W3K publishing, 2011.

[78] B. Yang, "A study of inverse short-time fourier transform," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pp. 3541–3544, 2008.

[79] H. Nawab and T. Quatieri, "Short-time fourier transform," *In Chapter 6 of Advanced Topics in Signal Processing, J. S. Lim, and A. V. Oppenheim (Eds), Prentice Hall, New Yersey*, pp. 289–337, 1988.

[80] F. Auger, E. Chassande-Mottin, and P. Flandrin, "Making reassignment adjustable: The levenberg-marquardt approach," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 3889–3892, IEEE, 2012.

[81] D. Fourer, F. Auger, and P. Flandrin, "Recursive versions of the levenberg-marquardt reassigned spectrogram and of the synchrosqueezed stft," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pp. 4880–4884, IEEE, 2016.

[82] 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 Transactions on Signal Processing*, vol. 60, no. 11, pp. 5787–5798, 2012.

[83] R. Carmona, W. Hwang, and B. Torresani, "Characterization of signals by the ridges of their wavelet transforms," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2586–2590, Oct 1997.

[84] C. K. Chui, Y.-T. Lin, and H.-T. Wu, "Real-time dynamics acquisition from irregular samples — with application to anesthesia evaluation," *Analysis and Applications*, vol. 14, p. 537–590, Jul 2016.

[85] H.-T. Wu, *Adaptive analysis of complex data sets.* PhD thesis, 2012.

[86] D. Fourer, F. Auger, K. Czarnecki, S. Meignen, and P. Flandrin, "Chirp rate and instantaneous frequency estimation: application to recursive vertical synchrosqueezing," *IEEE Signal Processing Letters*, no. 99, 2017.

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

[88] R. G. Baraniuk, P. Flandrin, A. J. Janssen, and O. J. Michel, "Measuring time-frequency information content using the rényi entropies," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1391–1409, 2001.

[89] L. Stanković, "A measure of some time–frequency distributions concentration," *Signal Processing*, vol. 81, no. 3, pp. 621–631, 2001.

[90] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[91] L. Sachs, *Applied statistics: a handbook of techniques*. Springer Science & Business Media, 2012.

[92] C. Villani, "Topics in optimal transportation, volume 58 of grad," *Stud. Math. AMS, Providence*, vol. 33, 2003.

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

[94] B. Barkat and B. Boashash, "A high-resolution quadratic time-frequency distribution for multicomponent signals analysis," *IEEE Transactions on Signal Processing*, vol. 49, no. 10, pp. 2232–2239, 2001.

[95] M. Z. Ikram and G. T. Zhou, "Estimation of multicomponent polynomial phase signals of mixed orders," *Signal Processing*, vol. 81, no. 11, pp. 2293–2308, 2001.

[96] S. Olhede and A. Walden, "A generalized demodulation approach to time-frequency projections for multicomponent signals," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 461, pp. 2159–2179, The Royal Society, 2005.

[97] Z. Yu, Y. Sun, and W. Jin, "A novel generalized demodulation approach for multicomponent signals," *Signal Processing*, vol. 118, pp. 188–202, 2016.

[98] D. Griffin and J. Lim, "Signal estimation from modified short-time fourier transform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 2, pp. 236–243, 1984.

[99] J. Allen, "Short term spectral analysis, synthesis, and modification by discrete fourier transform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 3, pp. 235–238, 1977.

[100] 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.

[101] 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.

[102] S. Meignen, T. Oberlin, and S. McLaughlin, "Fully adaptive mode decomposition from time-frequency ridges," in *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, pp. 3884–3888, IEEE, 2017.

[103] W. H. Organization, "Global status report on noncommunicable diseases 2017," tech. rep., World Health Organization, 2017.

[104] H. Koymen, B.-K. Altay, and Y. Z. Ider, "A study of prosthetic heart valve sounds," *IEEE transactions on biomedical engineering*, no. 11, pp. 853–863, 1987.

[105] J. Singh and R.-S. Anand, "Computer aided analysis of phonocardiogram," *Journal of Medical Engineering & Technology*, vol. 31, pp. 319–323, jan 2007.

[106] D. Gradolewski and G. Redlarski, "Wavelet-based denoising method for real phonocardiography signal recorded by mobile devices in noisy environment," *Computers in Biology and Medicine*, vol. 52, pp. 119–129, Sep 2014.

[107] S.-R. Messer, J. Agzarian, and D. Abbott, "Optimal wavelet denoising for phonocardiograms," *Microelectronics Journal*, vol. 32, pp. 931–941, Dec 2001.

[108] P. Varady, "Wavelet-based adaptive denoising of phonocardiographic records," in *Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2001.

[109] A.-K. Kumar and G. Saha, "A discrete wavelet packet transform based approach for selection of artifact-free phonocardiogram segment," in *Emerging Applications of Information Technology (EAIT), 2014 Fourth International Conference of*, pp. 83–88, IEEE, 2014.

[110] O. Beya, B. Jalil, E. Fauvet, and O. Laligant, "Empirical modal decomposition applied to cardiac signals analysis," in *Wavelet Applications in Industrial Processing VII* (F. Truchetet and O. Laligant, eds.), SPIE, feb 2010.

[111] A. Gavrovska, M. Slavkovic, I. Reljin, and B. Reljin, "Application of wavelet and emd-based denoising to phonocardiograms," in *Signals, Circuits and Systems (ISSCS), 2013 International Symposium on*, pp. 1–4, IEEE, 2013.

[112] O. Beya, E. Fauvet, and O. Laligant, "Eda, approche non linéaire de débruitage des signaux cardiaques," in *CORESA*, 2013.

[113] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: An overview of methods, challenges, and prospects," *Proceedings of the IEEE*, vol. 103, pp. 1449–1477, sep 2015.

[114] F. Sedighin, M. Babaie-Zadeh, B. Rivet, and C. Jutten, "Multimodal soft nonnegative matrix co-factorization for convolutive source separation," *IEEE Transactions on Signal Processing*, vol. 65, pp. 3179–3190, jun 2017.

[115] B. Rivet, M. Duda, A. Guerin-Dugue, C. Jutten, and P. Comon, "Multimodal approach to estimate the ocular movements during EEG recordings: A coupled tensor factorization method," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, aug 2015.

[116] Y. Bao, H. Fang, and J. Zhang, "Topicmf: Simultaneously exploiting ratings and reviews for recommendation.," in *AAAI*, vol. 14, pp. 2–8, 2014.

[117] S. Sra and I. S. Dhillon, "Generalized nonnegative matrix approximations with bregman divergences," in *Advances in neural information processing systems*, pp. 283–290, 2006.

[118] M. Niegowski and M. Zivanovic, "Ecg-emg separation by using enhanced non-negative matrix factorization," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pp. 4212–4215, IEEE, 2014.

[119] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Computational statistics & data analysis*, vol. 52, no. 1, pp. 155–173, 2007.

[120] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, "Optimal (wiener) filtering with the fft," *Numerical recepies in pascal: The art of scientific computing*, pp. 459–462, 1986.

[121] R. G. Brown and P. Y. Hwang, "Introduction to random signals and applied kalman filtering: with matlab exercises and solutions," *Introduction to random signals and applied Kalman filtering: with MATLAB exercises and solutions, by Brown, Robert Grover.; Hwang, Patrick YC New York: Wiley, c1997.*, 1997.

[122] G. Rilling, P. Flandrin, P. Goncalves, *et al.*, "On empirical mode decomposition and its algorithms," in *IEEE-EURASIP workshop on nonlinear signal and image processing*, vol. 3, pp. 8–11, NSIP-03, Grado (I), 2003.

[123] E. Vincent, R. Gribonval, and C. Févotte, "Performance measurement in blind audio source separation," *IEEE transactions on audio, speech, and language processing*, vol. 14, no. 4, pp. 1462–1469, 2006.

[124] R. McGill, J. W. Tukey, and W. A. Larsen, "Variations of box plots," *The American Statistician*, vol. 32, no. 1, pp. 12–16, 1978.

[125] H. Yang, "Statistical analysis of synchrosqueezed transforms," *Applied and Computational Harmonic Analysis*, jan 2017.

[126] J. Xiao and P. Flandrin, "Multitaper time-frequency reassignment for nonstationary spectrum estimation and chirp enhancement," *IEEE Transactions on Signal Processing*, vol. 55, pp. 2851–2860, June 2007.

[127] H.-T. Wu, Y.-H. Chan, Y.-T. Lin, and Y.-H. Yeh, "Using synchrosqueezing transform to discover breathing dynamics from ecg signals," *Applied and Computational Harmonic Analysis*, vol. 36, no. 2, pp. 354–359, 2014.

[128] H.-T. Wu, S.-S. Hseu, M.-Y. Bien, Y. R. Kou, and I. Daubechies, "Evaluating physiological dynamics via synchrosqueezing: Prediction of ventilator weaning," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 3, pp. 736–744, 2014.

[129] Y. Hu, X. Tu, F. Li, and G. Meng, "Joint high-order synchrosqueezing transform and multi-taper empirical wavelet transform for fault diagnosis of wind turbine planetary gearbox under nonstationary conditions," *Sensors*, vol. 18, no. 1, p. 150, 2018.

[130] Z. Cai, M. Zhang, and Y. Liu, "Sea-surface weak target detection scheme using a cultural algorithm aided time-frequency fusion strategy," *IET Radar, Sonar & Navigation*, 2018.

[131] W. Liu, S. Cao, Z. Wang, K. Jiang, Q. Zhang, and Y. Chen, "A novel approach for seismic time-frequency analysis based on high-order synchrosqueezing transform," *IEEE Geoscience and Remote Sensing Letters*, 2018.

[132] E. Chassande-Mottin and A. Pai, "Best chirplet chain: Near-optimal detection of gravitational wave chirps," *Physical review D*, vol. 73, no. 4, p. 042003, 2006.

[133] J. Terrien, C. Marque, and G. Germain, "Ridge extraction from the time–frequency representation (tfr) of signals based on an image processing approach: application to the analysis of uterine electromyogram ar tfr," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 5, pp. 1496–1503, 2008.

[134] D. Fourer, J. Harmouche, J. Schmitt, T. Oberlin, S. Meignen, F. Auger, and P. Flandrin, "The astres toolbox for mode extraction of non-stationary multicomponent signals," in *2017 25th European Signal Processing Conference (EUSIPCO)*, pp. 1130–1134, Aug 2017.