

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN  
INGEGNERIA ELETTRONICA, TELECOMUNICAZIONI E  
TECNOLOGIE DELL'INFORMAZIONE

Ciclo 31

**Settore Concorsuale: Telecomunicazioni**

**Settore Scientifico Disciplinare: Telecomunicazioni**

Time-Frequency Signal Analysis and Adaptive Instantaneous Frequency Estimation

**Presentata da: Abdoush, Yazan**

**Coordinatore Dottorato**

**Supervisore**

**Costanzo, Alessandra**

**Corazza, Giovanni Emanuele**

**Esame finale anno 2019**



# Time-Frequency Signal Analysis and Adaptive Instantaneous Frequency Estimation

by

Yazan Abdoush

Submitted to the Department of Electrical, Electronic and Information Engineering  
(DEI)

on March 13, 2019, in partial fulfillment of the  
requirements for the degree of

Ph.D. in Electronics, Telecommunications and Information Technologies Engineering

## Abstract

Most of the human-made and physical signals have nonstationary spectra that evolve rapidly with time. To study and characterize such signals, the classic time-domain and frequency-domain representations are inadequate, since they do not provide joint time and frequency information; meaning that, they are signal representations in which the time and frequency variables are mutually exclusive. Time-frequency (TF) signal analysis (TFSA) concerns the processing of signals with time-varying spectral content. It allows for the construction of a signal representation in which the time and frequency variables are not averaged with respect to each other, but rather present together.

This doctoral thesis has two main points of focus: TFSA based on a linear TF transform with progressive frequency-dependent resolution in the TF domain, known in the literature as the *S-transform* (ST), and designing adaptive methods for instantaneous frequency (IF) estimation, which is a fundamental concept in TFSA with numerous practical applications. The main original contributions are:

- Modifications in the existing discrete definitions for implementing and inverting the ST to ensure exact invertibility and eliminate artifacts in the synthesized signal (Chapter 2).
- Derivation of an algorithm for least-squares signal synthesis from a modified discrete ST (Chapter 2).
- Formulation of a computationally efficient, fully discrete, and exactly invertible ST with a controllable TF sampling scheme, providing frequency resolution that can be varied and made as high as required. This transform is denoted by the controlled converge discrete ST (CC-DST), and it is tested in various applications using real-world data (Chapter 2).
- Accuracy analysis of IF estimation based on a family of linear TF transforms that use Gaussian observation windows to localize the Fourier oscillatory kernel

with arbitrarily defined standard deviations, and derivation of closed-form easily interpreted expressions for the bias and the variance of the estimation error.

- Design of adaptive (i.e., signal-dependent) methods for IF estimation based on linear and quadratic TF representations (Chapter 3).

The first three topics are addressed in Chapter 2, while the last two in Chapter 3. Chapter 1 provides a brief introduction to TFSA in terms of theory and algorithms for the general reader. In Chapter 4, we discuss briefly three promising applications in which TFASP has been proven to bring significant improvements over conventional approaches, which are: jamming suppression for the global navigation satellite systems, automatic signal abnormality detection, and array signal processing for non-stationary signals, in order to give the reader further perspective.

Thesis Supervisor: Giovanni Emanuele Corazza  
Title: Full Professor

## Acknowledgments

Foremost, I would like to thank my family for their unlimited support, love, and encouragement. My sincere appreciation goes to my supervisor, Prof. Giovanni Emanuele Corazza, for giving me the trust to keep working on my research and explore new ideas; his patience, professional help, and the encouragement he provided greatly eased the work on this thesis. I thank Jose Garcia-Molina and Prof. Alessandro Vanelli Coralli for facilitating my internship with the European Space Research and Technology Center (ESTEC), The Netherlands. I am so grateful to the University of Bologna and the European Union's Erasmus Mundus program for their scholarship, which made it possible for me to pursue the master's and Ph.D. studies here in Italy. Last but not least, I would like to thank my good friends Ali Shbeeb, Elias Mouty, Gyimóthi László, Mohannad O. Zughayer, and Nicolas Massloub for their support, generosity, and advice, which aided me to continue this work and overcome difficult times in the last three years.



# Contents

<b>1</b>	<b>Introduction to time-frequency analysis</b>	<b>21</b>
1.1	Introduction . . . . .	22
1.2	Windowed Fourier transform . . . . .	22
1.3	Cohen's class of TFRs . . . . .	24
1.3.1	The Wigner distribution . . . . .	24
1.3.2	Quadratic TFRs . . . . .	26
1.3.3	Cross and inner artifacts . . . . .	28
1.3.4	Role of the kernel . . . . .	30
1.3.5	Generalized Wigner distribution . . . . .	32
1.4	Reassignment method . . . . .	34
1.5	Discrete implementation . . . . .	35
1.6	Objective assessment of TFRs . . . . .	37
1.7	Other advances . . . . .	38
<b>2</b>	<b>The S-transform</b>	<b>41</b>
2.1	Background theory . . . . .	42
2.1.1	Inverting the ST . . . . .	44
2.2	Discrete ST and exact inevitability . . . . .	45
2.3	Least-squares signal synthesis . . . . .	47
2.3.1	Information redundancy . . . . .	51
2.3.2	Examples . . . . .	53
2.4	Controlled-coverage discrete S-transform . . . . .	56
2.4.1	Limitations of the uniformly sampled ST . . . . .	56

2.4.2	The ST as a filter bank . . . . .	58
2.4.3	Controlled spectral partitioning scheme . . . . .	61
2.4.4	Spectral windowing . . . . .	65
2.4.5	DFT-based implementation . . . . .	69
2.4.6	Exact reconstruction . . . . .	72
2.4.7	Quality-factor and complexity . . . . .	74
2.4.8	Application: IF estimation . . . . .	78
2.4.9	Application: time-scaling audio signal . . . . .	82
2.4.10	Application: denoising ECG signal . . . . .	84
2.4.11	Application: visual representation . . . . .	85
<b>3</b>	<b>Adaptive methods for instantaneous frequency estimation</b>	<b>91</b>
3.1	Background theory . . . . .	92
3.1.1	Analytic signals . . . . .	92
3.1.2	Bandwidth-duration product . . . . .	93
3.1.3	The instantaneous frequency . . . . .	94
3.2	Adaptive IF estimation . . . . .	95
3.2.1	Adaptive design of TFRs . . . . .	95
3.2.2	Accuracy analysis of IF estimation based on linear TFRs . . . . .	97
3.2.3	Three-step method for adaptive IF estimation . . . . .	106
3.2.4	Performance evaluation . . . . .	113
3.3	Adaptive IF estimation based on QTFRs . . . . .	123
3.3.1	Introduction . . . . .	123
3.3.2	Background theory . . . . .	124
3.3.3	IF estimation based on the ICI rule . . . . .	126
3.3.4	Motivation for the proposed approach . . . . .	126
3.3.5	Selecting a PTFR . . . . .	128
3.3.6	Derivative approximation by the CWT . . . . .	130
3.3.7	Numerical evaluation . . . . .	132
3.3.8	Simulation results . . . . .	133

<b>4</b>	<b>Applications of time-frequency representations</b>	<b>137</b>
4.1	Jamming mitigation . . . . .	138
4.1.1	Introduction . . . . .	138
4.1.2	Jamming mitigation . . . . .	139
4.2	Abnormalities detection . . . . .	142
4.2.1	TF-based features for classifying nonstationary signals . . . . .	143
4.3	Multisensor TF signal processing . . . . .	144
4.3.1	Extension to multisensor TFRs . . . . .	144
<b>5</b>	<b>Conclusion</b>	<b>147</b>
<b>A</b>	<b>Proof of Lemma 1</b>	<b>151</b>
A.1	Notation . . . . .	151
A.2	Noise properties . . . . .	152
A.3	Proof . . . . .	152
<b>B</b>	<b>Proof of Lemma 2</b>	<b>155</b>
B.1	Proof . . . . .	155



# List of Figures

1-1	Comparison between STFT and ST of signal with two sinusoidal components. The ST provides progressive TF resolution: the higher spectral component is projected onto the TF domain with lower frequency resolution. . . . .	24
1-2	Artifacts in the WD. Cross-terms are present in (a) because the signal is composed of two linear components, and inner-terms are present in (b) because the signal IF law is nonlinear . . . . .	30
1-3	Comparison between selected TFRs using a sinusoidal FM signal. . .	33
1-4	Comparison between spectrogram and reassigned spectrogram using a sinusoidal FM signal. . . . .	35
2-1	Reconstruction errors (in absolute value) resulting from transforming a signal to the TF domain then back-transforming it to the time domain through different forward and backward ST algorithms. . . . .	48
2-2	Examples of TF filtering through the ST. Magnitude of STs of two test signals in (a) and (c). Corresponding binary masks in (b) and (d). . .	54
2-3	Filter-based representation of one voice provided by the DST. $\bar{w}_{k_0}[n]$ represents the utilized window up-converted by the exponential modulator $\exp(i2\pi nk_0/N)$ . . . . .	58
2-4	Progression of the segment size $L_c(\nu, j_0)$ (in logarithmic scale) as a function of $\nu$ . Increasing $\gamma$ leads to faster growth in the segment size. . . . .	62
2-5	Prototype of the proposed asymmetric raised-cosine window. . . . .	66

2-6	Comparison between asymmetric raised-cosine window with roll-off factor equal to 0.5, rectangular window, and truncated Gaussian window in the frequency and the time domain. Frequency responses in (a). Absolute values of time responses in (b). . . . .	68
2-7	Spectral coverage provided by the CC-DST with four different parameter settings. (a) $N_0 = 16$ , $\gamma = 2$ , $\alpha = 0$ (in this case, CC-DST samples the TF domain dyadically). (b) $N_0 = 16$ , $\gamma = 2$ , $\alpha = 0.8$ . (c) $N_0 = 16$ , $\gamma = 1.3$ , $\alpha = 0.8$ (these values provide finer spectral decomposition compared with previous cases). (d) $N_0 = 64$ , $\gamma = 1.0001$ , $\alpha = 0.8$ (CC-DST in this case acts as STFT). . . . .	71
2-8	Variations of the Q-factor of the CC-DST. By reducing $\gamma$ , $Q(\iota)$ increases, but, on the other hand, larger $\gamma$ leads to smaller variations in the Q-factor. . . . .	74
2-9	Comparison between efficiently sampled DSTs and the original uniformly sampled DST in terms of computation time. Computation speed of FFT is added as a reference point. The largest signal processed by original DST was of size $2^{14}$ samples, due to memory limitations. The CC-DST is slightly slower than the one-to-one DOST, but both have computation speeds comparable to those of FFT and show similar growth rate in computation time. This experiment was run on an Intel Core i-5 platform having 6 GB RAM and MATLAB was used for implementation. . . . .	76
2-10	Comparison between several DSTs in terms of IF estimation. Figure shows the NMSE (in dB) of the IF estimates of a sinusoidally FM signal as a function of SNR (in dB). Starting from SNR = 5 dB, CC-DST ( $N_0 = 64$ , $\gamma = 1.1$ , $\alpha = 0.8$ ) provided the most accurate IF estimates. . . . .	77

2-11	IF estimation of a sinusoidally FM signal using the CC-DST with a moving average filter of various sizes. The filter aims at reducing the noise power in the local phase estimates. Figure shows the NMSE of the IF estimates as a function of SNR. Results demonstrate that a filter with a moderate smoothing length should be used. . . . .	78
2-12	Performance of IF estimators based on selected DSTs. Results illustrate the impact of the period of sinusoidal frequency modulation, controlled by the parameter $T$ , on the accuracy of the IF estimates. Increasing $T$ leads to shorter IF periods. Figure shows the NMSE of the IF estimates as a function of $T$ at SNR = 10 dB using log-scales for clarity. At low $T$ , the CC-DST ( $N_0 = 64$ , $\gamma = 1.1$ , $\alpha = 0.8$ ) returns accurate IF estimates, while at large $T$ , the ADSTs are superior. . . . .	81
2-13	Denoising ECG signal via hard thresholding using selected DSTs. Figure shows results of RMSE with different thresholds. CC-DST ( $N_0 = 64$ , $\gamma = 1.1$ , $\alpha = 0.8$ ) is the best performing non-adaptive DST. The ADST provides better denoising performance but requires far more computation time than the CC-DST. . . . .	84
2-14	Denoising ECG signal contaminated by AWGN at SNR=10 dB using the CC-DST and hard thresholding. Noisy ECG signal (without DC component) in (a) and the enhanced signal in (b). . . . .	86
2-15	TFRs of ECG signal using the CC-DST. After interpolating the CC-DST voices in time, the resulting matrix is interpolated in frequency to create $N \times N$ TFR. Interpolation using the nearest neighbor method in (a) and spline in (b). . . . .	87

2-16 TFRs of a real-world bat echolocation sound signal using: (a) the original DST, (b) ADST [1], (c) ADST [2], (d) DOST [3], (e) DST of Brown *et al.* [4], (f) the CC-DST ( $N_0 = 8$ ,  $\gamma = 1.1$ ,  $\alpha = 0.8$ ) with the nearest neighbor interpolation method, and (g) the CC-DST with spline interpolation. All the TFRs are given with the same MATLAB colormap. The CC-DST is superior to all the non-adaptive DSTs. Parameters of the CC-DST were selected based on visual inspection. . . . . 88

2-17 TFRs of a synthetic signal with crossing IF components using: (a) original DST, (b) ADST [1], (c) ADST [2], (d) DOST [3], (e) DST of Brown *et al.* [4], (f) the CC-DST ( $N_0 = 16$ ,  $\gamma = 1.15$ ,  $\alpha = 0.8$ ) with the nearest neighbor interpolation method, and (g) the CC-DST with spline interpolation algorithm. All the TFRs are given with the same MATLAB colormap. Parameters of the CC-DST were selected based on visual inspection. The TFRs obtained with the CC-DST feature improved EC compared with those based on the DOST and the DST of Brown *et al.*. . . . . 89

3-1 Flowchart of procedure used to construct adaptive TFRs based on linear TF transforms belonging to (3.17). . . . . 104

3-2 Theoretical and experimental performance of IF estimators based on STFT and ST using FM signals belonging to three different classes. In upper subplots, STFT with  $\sigma$  equal to 16, 16, 10, respectively, is used. In the last subplots, we use ST with  $\sigma(f)$  equal to  $1/(0.1 + 0.6|f|)$ ,  $1/(0.1 + |f|)$ ,  $1/(0.1 + |f|)$ , respectively. . . . . 105

3-3	Comparison between TFRs based on ECMs with that based on the proposed regular rate. Analyzed signal is given in (3.51) and corrupted by AWGN (SNR equal to 10 dB). TFRs are based on STFT the width of its observation window is set based on (a) ECM of Stanković [5], (b) ECM of Pei and Huang [6], and (c) proposed regular rate. TFRs based on ECMs fail to concentrate enough energy along sinusoidal IF component. . . . .	107
3-4	TFRs of synthetic signal composed of two components given in (3.51) and corrupted by AWGN (SNR equal to 12 dB). Respective parameters of TFRs are adjusted based on visual inspection as provided in Table 3.1. . . . .	114
3-5	TFRs of synthetic signal of 4 components given in (3.52) corrupted by AWGN (SNR equal to 12 dB). Respective parameters of TFRs are optimized based on visual inspection as provided in Table 3.1. . . . .	115
3-6	Comparison between selected TFRs in terms of the NMSE of IF estimation. Figure shows the NMSE (in decibels) as function of SNR (in decibels). We normalize both the true and estimated IFs by the sampling frequency before computing the NMSE. Results of Test Signal A in (a), and results of Test Signal B in (b). TFRs of Test Signal A and Test Signal B are depicted in Fig. 3-4 and Fig. 3-5, respectively. Results of the MDD with Test Signal A are not shown, because the MDD does not resolve the sinusoidally modulated component. . . . .	117
3-7	Comparison between selected TFRs in terms of IF estimation. Analyzed signal is real-world EEG seizure signal corrupted by AWGN at SNR equal to 15 dB. Figure shows TFRs and corresponding IF estimates extracted using an algorithm based on image processing. . . . .	118

3-8	TFRs of real-world bat echolocation sound corrupted by AWGN (SNR equal to 20 dB). Respective parameters of TFRs are provided in Table 3.1. Parameters of EMBD and ADTFD are adjusted based on the results in [7], while those of the SM and the Re-Spec are based on visual inspection. Proposed methods do not require parameter optimization and are fully automated . . . . .	119
3-9	Variations of $L_1(\rho)$ and $L_2(\rho)$ (in dB). Function $L_2(\rho)$ increases above the stationary point at $\rho = 1$ much slower compared with $L_1(\rho)$ . Note, for example, $L_1(\rho = 2)$ is more than 5 times larger than $L_2(\rho = 2)$ . . .	127
3-10	Example of computing the second-order derivative of a noisy sine waveform (SNR=15 dB). (a) Derivative computed by the conventional finite-difference method. (b) Spectral response amplitude of the filter $\Theta_s(t)$ given by the second-order derivative of a Gaussian and SG second-order differential filter (note the high side lobes of SG filter). Parameters of the filters are set such that they have the same cutoff frequency. Derivative computed by: SG differential filter in (c), and the CWT in (d). Most accurate result is provided by CWT. . . . .	129
3-11	Flowchart of proposed method for adaptive IF estimation based on QTFR. . . . .	131
3-12	Comparison between IF estimation methods in terms of the NMSE using two test signals. All the considered methods are implemented using the WD and the SM. . . . .	133
3-13	IF estimation for Signal A. Proposed method is compared with the modified ICI-based algorithm. First and second columns from left: SNR = 11 dB and the WD is used for TFR. Third and fourth columns from left: SNR = 0 dB and the SM is used for TFR. First row from top: estimated IF (solid line) and true IF (red dashed line). Second row: absolute error. Third row from top: the window width. . . . .	134

3-14	IF estimation for Signal B. Proposed method is compared with the modified ICI-based algorithm. First and second columns from left: SNR = 11 dB and the WD is used for TFR. Third and fourth columns from left: SNR = 0 dB and the SM is used for TFR. First row from top: estimated IF (solid line) and true IF (red dashed line). Second row: absolute error. Third row from top: the window width. . . . .	135
4-1	Sample of jammers for sale online ( <a href="https://www.jammer-store.com/">https://www.jammer-store.com/</a> ).140	
4-2	Block diagram of abnormality detection methodology using TF-based approach . . . . .	143



# List of Tables

1.1	Definitions of selected popular Doppler-lag kernels. . . . .	31
2.1	$MSE_{TF}$ and $MSE_T$ of two signals synthesized from modified STs through selected inverting algorithms. . . . .	56
3.1	Parameters used to construct TFRs of test signals. . . . .	110
3.2	Coefficients of Test Signal B given in (3.52). . . . .	113
3.3	ECM of selected TFRs of a real-world bat echolocation sound. . . . .	121
3.4	Comparison between selected TFRs in terms of computation time. . . . .	122
3.5	Comparison between adaptive IF estimation methods in terms of computation time. SNR is set to 11 dB, and the WD is used as TFR. The results are given in seconds, and the experiments were run on Windows-7 Intel Core i-5 platform with 6 GB RAM. Proposed method is the most computationally efficient. . . . .	136



# Chapter 1

## Introduction to time-frequency analysis

## 1.1 Introduction

The spectral analysis of signals is a crucial task in countless scientific applications. As well known, the Fourier transform (FT) is the conventional tool for decomposing a signal into individual spectral components by projecting it onto sinusoidal basis functions. The Fourier analysis allows for effective feature extraction from signals whose spectra can be regarded stationary, or slowly changing, during the processing time. However, most of the real physical signals have nonstationary spectra, meaning that they cannot be completely characterized in the time domain or frequency domain alone: the time domain lacks any frequency information, while, on the other hand, the frequency domain does not provide the necessary time description of varying spectra. Therefore, nonstationary signals are best represented in the time-frequency (TF) plane where the time and frequency variables are present together without being mutually exclusive. The signal TF representation (TFR) should reveal the spectral variations in time; that is, the constant-time cross-section of a TFR should depict the frequencies that emerge at the respective time instant. By doing so, a TFR may provide comprehensive information about nonstationary signals including the time variations, the frequency variations, and the number of components; it enables separation of the signal components from each other, and effective time-varying filtering

## 1.2 Windowed Fourier transform

The simplest approach to include a sort of time description within the Fourier analysis is by localizing the oscillatory Fourier basis in the time domain through a windowing function. Namely, at a certain time instant  $t$ , we may multiply the signal by a window centered at  $t$  and perform the conventional FT on the windowed signal to obtain a localized spectrum. The resulting transform is called the short-time Fourier transform (STFT) and is given according to

$$\text{STFT}_x(t, f) = \int_{-\infty}^{\infty} x(\tau)w(t - \tau) \exp(-i2\pi f\tau) d\tau, \quad (1.1)$$

where  $x(\tau)$  is the signal under analysis, and  $w(\tau)$  is a localizing window, which is usually a real-valued even function. To represent the energy distribution in the TF plane, we may use the squared magnitude of the STFT, which is called the spectrogram:

$$\begin{aligned} \text{SPEC}_x(t, f) &= |\text{STFT}_x(t, f)|^2 \\ &= \left| \int_{-\infty}^{\infty} x(\tau)w(t - \tau) \exp(-i2\pi f\tau) d\tau \right|^2. \end{aligned} \quad (1.2)$$

The concept behind the STFT is simple, and it can be realized in the discrete case efficiently through the Fast FT (FFT) algorithm. It localizes the signal amplitude and phase information independently, and it is easily invertible; therefore, it has been always a classic choice whenever a simple straightforward TF analysis is needed [8, 9]. However, the main limitation of the STFT lies in its use of a fixed observation window, which imposes an unavoidable trade-off between time and frequency resolution. That is, if a large window is used, the frequency resolution is sufficient, since in this case, the FT works upon a large signal, hence providing precise frequency information. If a short window is used, on the other hand, the time resolution is improved and distinguishing fast temporal events becomes possible, but at the expense of smeared spectral resolution.

The S-transform (ST) was originally proposed by Stockwell *et al.* in [10] to improve on the STFT by localizing the signal components in the TF plane with frequency-dependent resolution. This objective is achieved by replacing the fixed observation window in (1.2) with one whose width reduces with increasing frequency. The rationale behind the ST is that, low spectral contents represent slowly varying signal components, hence it is reasonable to sacrifice the time resolution for better frequency resolution. On the contrary, high frequencies represent components with rapid variations in the time domain, where accurate time information can be considered more valuable. A general description of the ST may be given as [4, 11]

$$\text{ST}_x(t, f) = \int_{-\infty}^{\infty} x(\tau)w(t - \tau, f) \exp(-i2\pi f\tau) d\tau, \quad (1.3)$$

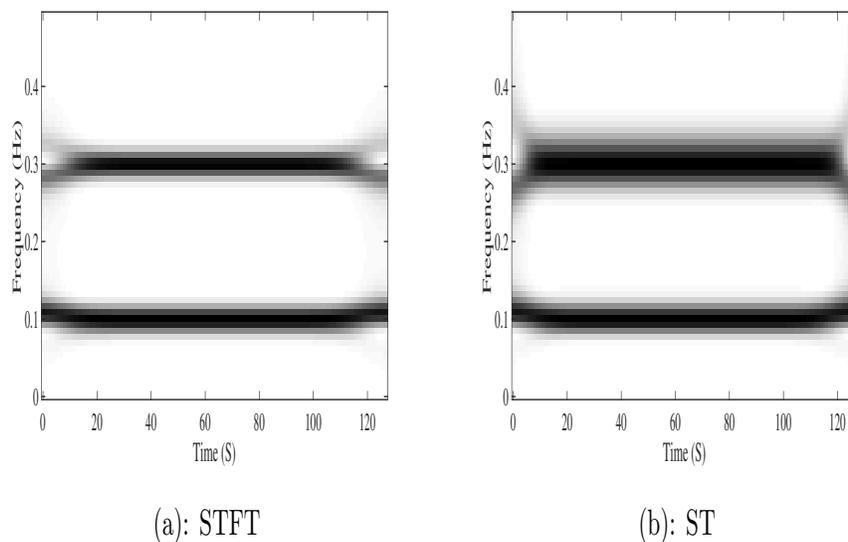


Figure 1-1: Comparison between STFT and ST of signal with two sinusoidal components. The ST provides progressive TF resolution: the higher spectral component is projected onto the TF domain with lower frequency resolution.

in which  $w(t, f)$  is a frequency-dependent observation window. A comparison between the TFRs provided by the spectrogram and the ST (as squared magnitude) is illustrated in Fig. 1-1 using a signal composed of two sinusoids.

## 1.3 Cohen's class of TFRs

### 1.3.1 The Wigner distribution

Ideal TFRs would concentrate energy exactly at the signal instantaneous spectrum. Assuming that the analyzed signal is monocomponent frequency modulated (FM) with unitary amplitude and an arbitrary differentiable phase  $\phi(t)$  defined as:

$$z(t) = \exp(i\phi(t)), \tag{1.4}$$

the ideal TFR, which would be concentrated exactly along the trajectory of the signal instantaneous frequency (IF), can be given by the following form [12]:

$$\text{TFR}_z(t, f) = \delta(f - f_i(t)), \quad (1.5)$$

where  $\delta(t)$  is a the Delta function, and  $f_i(t)$  is the signal IF. The concept of the IF is addressed in detail in Chapter 3; here, we will just present the mathematical definition [13, 14]:

$$f_i(t) = \frac{1}{2\pi} \frac{d\phi}{dt}(t) \equiv \frac{\phi'(t)}{2\pi}. \quad (1.6)$$

Clearly, the IF represents the instantaneous rate of change of the signal phase function. The time derivative of the phase function can be written as

$$\phi'(t) = \lim_{\tau \rightarrow 0} \frac{\phi(t + \tau/2) - \phi(t - \tau/2)}{\tau}. \quad (1.7)$$

If (1.7) is approximated by

$$\phi'(t) \approx \frac{\phi(t + \tau/2) - \phi(t - \tau/2)}{\tau} \quad (1.8)$$

and substituted into (1.5),  $\text{TFR}_z(t, f)$  can be given by [15]

$$\begin{aligned} \text{TFR}_z(t, f) &= \text{FT}_{\tau \rightarrow f} \{ \exp(i\phi(t + \tau/2)) \exp(-i\phi(t - \tau/2)) \} \\ &= \text{FT}_{\tau \rightarrow f} \{ z(t + \tau/2) z^*(t - \tau/2) \} \\ &= \text{FT}_{\tau \rightarrow f} \{ K_z(t, \tau) \}, \end{aligned} \quad (1.9)$$

where  $\text{FT}\{\cdot\}$  is the FT operator. The 2-D function  $K(t, \tau)$  is called the signal *instantaneous autocorrelation function* (IAF) [16] and is defined for a general signal  $z(t)$  according to

$$K_z(t, \tau) = z(t + \tau/2) z^*(t - \tau/2). \quad (1.10)$$

The TFR given by (1.9) is the Wigner distribution<sup>1</sup>(WD) [17], denoted by  $WD_z(t, f)$ . Since the approximation (1.8) is exact when the IF law of the analyzed signal is a linear function of time, we may conclude that the WD is indeed the ideal TFR for monocomponent LFM signals [18, 19]. The FT of the WD with respect to the time variable  $t$  results in another 2-D function called the *spectral autocorrelation function*(SAF):

$$\begin{aligned} k_z(\nu, f) &= \text{FT}_{t \rightarrow \nu} \{WD_z(t, f)\} \\ &= Z(f + \nu/2)Z^*(f - \nu/2), \end{aligned} \tag{1.11}$$

where  $Z(f)$  is the FT of the analyzed signal  $z(t)$ , and  $\tau$  and  $\nu$  are variables called lag and Doppler, respectively. According to the previous expression, an equivalent form of the WD in (1.9) may be defined as

$$WD_z(t, f) = \text{IFT}_{\nu \rightarrow t} \{k_z(\nu, f)\}, \tag{1.12}$$

where  $\text{IFT}\{\cdot\}$  is the inverse FT operator.

### 1.3.2 Quadratic TFRs

A general form of TFRs that are quadratic in signal can be derived from the WD by applying a smoothing kernel according to [15, 17, 16, 20]:

$$\text{TFR}_z(t, f) = \gamma(t, f) \underset{t}{*} \underset{f}{*} WD_z(t, f), \tag{1.13}$$

where  $\gamma(t, f)$  is a 2-D kernel that acts upon the WD as a 2-D TF filter. The IFT of the kernel with respect to frequency is a time-lag function called the *time-lag kernel*,

---

<sup>1</sup>If the analyzed signal is analytic, this TFR is also called the Wigner-Ville distribution.

denoted by  $G(t, \tau)$ , using which the TFR in (1.13) can be also defined as

$$\begin{aligned} \text{TFR}_z(t, f) &= \text{FT}_{\tau \rightarrow f} \left\{ G(t, \tau) *_{t} K_z(t, \tau) \right\} \\ &= \text{FT}_{\tau \rightarrow f} \{ R_z(t, \tau) \}. \end{aligned} \quad (1.14)$$

Note that for the WD, we have that  $R_z(t, \tau) = K_z(t, \tau)$ . We can now introduce the FT of the IAF with respect to the time variable  $t$ :

$$\begin{aligned} A_z(\nu, \tau) &= \text{FT}_{t \rightarrow \nu} \{ K_z(t, \tau) \} \\ &= \text{IFT}_{f \rightarrow \tau} \{ k_z(\nu, f) \}, \end{aligned} \quad (1.15)$$

which is called the *ambiguity function* (AF). Now, by applying the convolution theorem to (1.14), we obtain

$$\text{TFR}_z(t, f) = \text{IFT}_{\nu \rightarrow t} \left\{ \text{FT}_{\tau \rightarrow f} \{ g(\nu, \tau) A_z(\nu, \tau) \} \right\}, \quad (1.16)$$

in which

$$g(\nu, \tau) = \text{FT}_{t \rightarrow \nu} \{ G(t, \tau) \} \quad (1.17)$$

is called the *Doppler-lag kernel*. One more form of  $\text{TFR}_z(t, f)$  can be also derived from (1.16) and (1.15) according to

$$\text{TFR}_z(t, f) = \text{IFT}_{\nu \rightarrow t} \left\{ \chi(\nu, f) *_{f} k_z(\nu, f) \right\}, \quad (1.18)$$

where

$$\chi(\nu, f) = \text{FT}_{\tau \rightarrow f} \{ g(\nu, \tau) \} \quad (1.19)$$

is called the *Doppler-frequency kernel*.

To sum up, the quadratic TFRs (QTFRs) can be constructed through 4 different domains, which are:

### Time-Frequency construction

$$\text{TFR}_z(t, f) = \gamma(t, f) \underset{t}{*} \underset{f}{*} \text{WD}_z(t, f); \quad (1.20)$$

### Doppler-lag construction

$$\begin{aligned} \text{TFR}_z(t, f) &= \text{IFT}_{\nu \rightarrow t} \left\{ \text{FT}_{\tau \rightarrow f} \{g(\nu, \tau) A_z(\nu, \tau)\} \right\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\nu, \tau) A_z(\nu, \tau) \exp(i2\pi(\nu t - f\tau)) d\tau d\nu; \end{aligned} \quad (1.21)$$

### Doppler-frequency construction

$$\begin{aligned} \text{TFR}_z(t, f) &= \text{IFT}_{\nu \rightarrow t} \left\{ \chi(\nu, f) \underset{f}{*} k_z(\nu, f) \right\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \chi(\nu, f - \eta) Z(\eta + \nu/2) Z^*(\eta - \nu/2) \exp(i2\pi\nu t) d\eta d\nu; \end{aligned} \quad (1.22)$$

### Time-lag construction

$$\begin{aligned} \text{TFR}_z(t, f) &= \text{FT}_{\tau \rightarrow f} \left\{ G(t, \tau) \underset{t}{*} K_z(t, \tau) \right\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(t - \xi, \tau) z(\xi + \tau/2) z^*(\xi - \tau/2) \exp(-i2\pi\tau f) d\xi d\tau. \end{aligned} \quad (1.23)$$

This family of QTFRs is called Cohen's family [16].

### 1.3.3 Cross and inner artifacts

The WD and all other QTFRs, which are essentially smoothed versions of the WD according to (1.20), are bilinear in the analyzed signal. As a result of bilinearity, they are inevitably accompanied by undesirable spurious artifacts that represent the interactions between different signal components.

The first sort of artifacts arises if the signal IF law is nonlinear. To explain this, consider a signal  $z(t) = \exp(i\phi(t))$ , the WD of this signal can be written as [21]

$$\text{WD}_z(t, f) = \delta(t - f_i(t)) * \text{FT} \left\{ \exp \left( i \sum_{n=2}^{\infty} \phi^{(n)}(t) \tau \frac{K_n}{n!} \right) \right\}, \quad (1.24)$$

where

$$K_n = \sum_{j=0}^{n-1} \left(\frac{\tau}{2}\right)^{n-1-j} \left(\frac{-\tau}{2}\right)^j. \quad (1.25)$$

The above equation makes clear that if the signal is LFM, meaning that the IF law is a linear function of time, then the term between parentheses in (1.24) is zero. In other words, the WD is the ideal TFR for LFM signals, as mentioned before. However, if the signal phase is of order higher than second, the disturbing term in (1.24) is not zero and, as a result, spurious artifacts in the WD arise; they are called *inner-terms*. The same conclusions apply to any QTFR.

If the signal is composed of multiple components, the QTFRs will contain spurious features, called *cross-terms*. For simplicity, let us assume that the analyzed signal is composed of two components, that is  $z(t) = z_1(t) + z_2(t)$ , then the IAF of  $z(t)$  is given by

$$K_z(t, \tau) = K_{z_1}(t, \tau) + K_{z_2}(t, \tau) + K_{z_1 z_2}(t, \tau) + K_{z_2 z_1}(t, \tau), \quad (1.26)$$

where

$$K_{z_i z_j}(t, \tau) = z_i(t + \tau/2) z_j^*(t - \tau/2). \quad (1.27)$$

As a result of (1.26), the WD of  $z(t)$  is given by

$$\text{WD}_z(t, f) = \text{WD}_{z_1}(t, f) + \text{WD}_{z_2}(t, f) + 2\Re\{\text{WD}_{z_1 z_2}(t, f)\}. \quad (1.28)$$

$\text{WD}_{z_1}$  and  $\text{WD}_{z_2}$  are called *auto-terms*. This result can be generalized to any member of Cohen's class of TFRs. The problem of cross and inner artifacts is illustrated on the WD in Fig. 1-2. In (a), we use a signal composed of two LFM chips, where cross-terms can be clearly observed oscillating between the signal components, while in (b), the signal is nonlinearly modulated and inner artifacts are clearly manifested in the TFR.

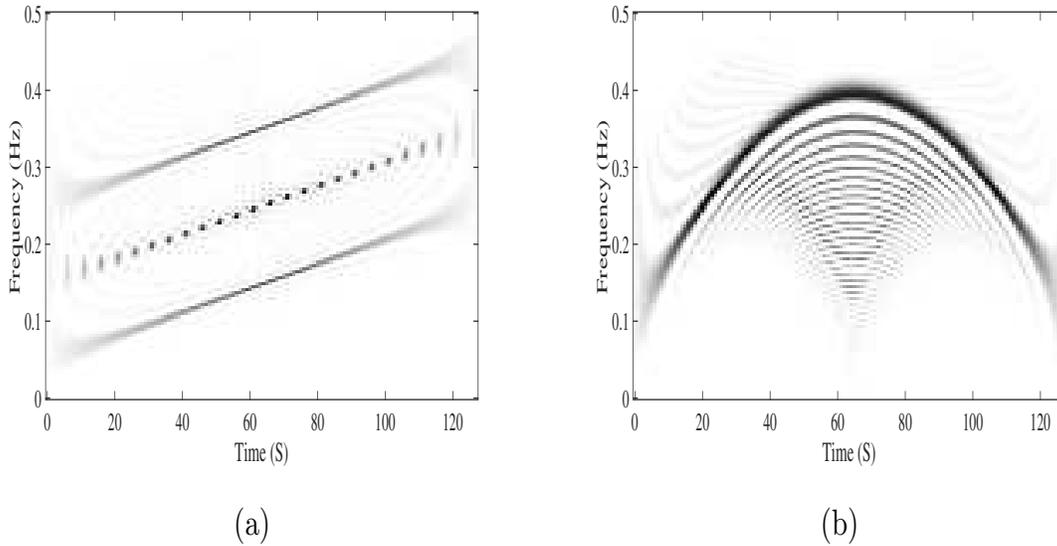


Figure 1-2: Artifacts in the WD. Cross-terms are present in (a) because the signal is composed of two linear components, and inner-terms are present in (b) because the signal IF law is nonlinear

### 1.3.4 Role of the kernel

The presence of inner- and cross-terms represent the main limitation of the QTFRs. When the signal is composed of multiple nonlinear components, these interfering terms make interpretation of a TFR an exhausting and sometimes impossible task. The shapes, locations, and other properties of the inner- and cross-terms have been studied extensively (e.g. [21]). Recently, Boashash and Ouelha in [22] derived exact analytical formulas giving the locations of cross-terms in the ambiguity domain of multicomponent piecewise LFM signals. Generally, it was shown that the cross-terms are oscillatory in nature along their major axis, meaning that they have high-pass characteristics. The auto-terms, on the other hand, have slowly changing amplitudes along their major axis. Since the QTFRs are smoothed versions of the WD, by selecting a proper signal kernel  $\gamma(t, f)$  (i.e., one that matches the signal structure), the impact of cross- and inner-terms can be alleviated [15, 23, 24]. These kernels can be designed using basic filter design principles with the goal to eliminate the interfering artifacts, while at the same time preserving the useful signal components as much as possible. Alternatively, one may characterize the location and behavior of the

QTFR	Doppler-lag kernel
	$g(\nu, \tau)$
WD	1
Spectrogram	$\text{FT}_{t \rightarrow \nu} \{w(t + \tau/2)w(t - \tau/2)\}$
Modified B (MB) [15]	$\frac{ \Gamma(\beta + j\nu\pi) ^2}{\Gamma^2(\beta)}$
Extended Modified B (EMB) [26]	$\frac{ \Gamma(\beta + j\nu\pi) ^2}{\Gamma^2(\beta)} \frac{ \Gamma(\alpha + j\nu\pi) ^2}{\Gamma^2(\alpha)}$
pseudo WD (PWD) [27]	$w(\tau)$
Choi-Williams (CW) [28]	$\exp(-\nu^2\tau^2/a)$
Born-Jordan (BJ) [29]	$\text{sinc}(2a\nu\tau)$

Table 1.1: Definitions of selected popular Doppler-lag kernels.

inner- and cross-terms in the ambiguity domain  $(\nu, \tau)$ , and through a careful design of the Doppler-lag kernel  $g(\nu, \tau)$ , the undesirable terms can be filtered out. In the ambiguity domain, it was shown that the auto-terms are concentrated close to or pass through the origin as opposed to the cross-terms, which are placed far from the origin [22, 21, 25]. Filtering out the undesired cross-terms is perhaps easier and commonly done in the ambiguity domain because the filtering operation is multiplicative.

Boashash showed that the properties of a QTFR are related to those of the kernel [15], implying that the process of designing a QTFR with certain desirable properties boils down to designing the kernel that produces such properties. The kernel  $g(\nu, \tau)$  could be:

- non-separable;
- separable  $g(\nu, \tau) = g_1(\nu)g_2(\tau)$ ;
- Doppler-independent  $g(\nu, \tau) = g_1(\nu)$ ;
- lag-independent  $g(\nu, \tau) = g_2(\tau)$ .

The advantage of using separable kernels lies in the ability to smooth the WD in time and frequency almost independently. Some of the popular kernels are listed in table 1.1. In the same context, we should also mention the compact support

kernel proposed in [30] and its extended separable version presented in [23], which are designed to vanish outside a given range in the ambiguity domain.

Stanković in [31] introduced an interesting TFR, usually called the S-method (SM). This TFR can be considered as a mixture between the spectrogram and the WD, combining the advantages of both; it is defined by

$$\text{SM}_z(t, f) = \int_{-\infty}^{\infty} g(\theta) \text{STFT}_z(t, f + \theta) \text{STFT}_z^*(t, f - \theta) d\theta, \quad (1.29)$$

where  $g(\theta)$  is a short frequency window. By properly choosing the length of  $g(\theta)$ , we can preserve the high resolution of the WD, while at the same time having an immunity to cross-terms likewise the spectrogram. Further, the SM can be computed using the STFT, which makes it more computationally efficient than other QTFRs.

Although designing QTFRs that have certain mathematical properties, such as non-negativity and preservation of time and frequency marginals (see [16]) could be desired for some applications, we see that most of the research in the field of TF signal processing aimed at designing TFRs with high resolution. That is, the main goal in practice is having high energy concentration along the signal IF law with a significant reduction in the amplitudes of the inner- and cross-terms.

### 1.3.5 Generalized Wigner distribution

A generalization to the WD, called the L-WD, was presented by Stanković in [32, 12] to reduce the impact of high-order derivatives of the signal phase function on the WD; it is defined by

$$\text{LWD}_z(t, f) = \int_{-\infty}^{\infty} z^L(t + \tau/2) z^{*L}(t - \tau/2) \exp(-i2\pi\tau f) d\tau, \quad (1.30)$$

where  $L$  is an integer greater than zero. The LWD is simply a scaled WD of the analyzed signal raised to the  $L$ th power. If we assume that the analyzed signal is  $z(t) = \exp(i\phi(t))$ , then by expanding the phase function  $\phi(t)$  into Taylor series

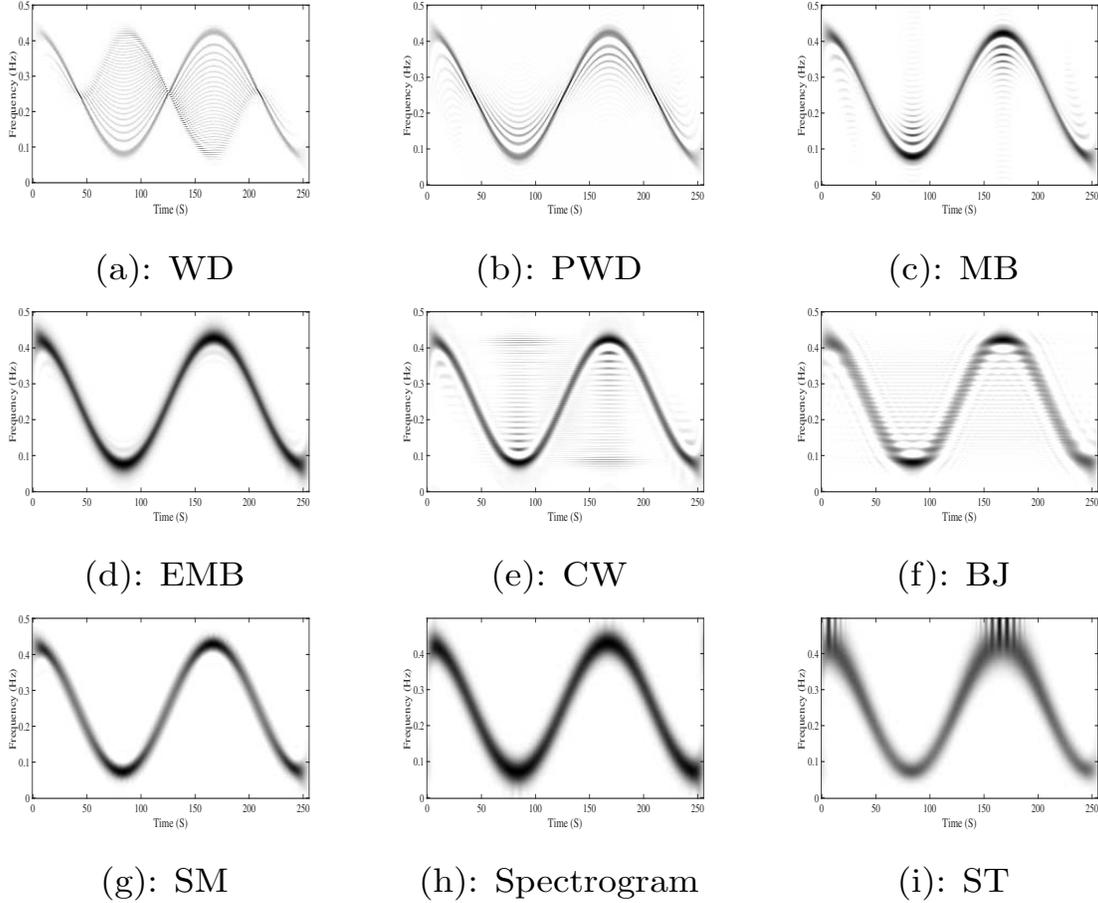


Figure 1-3: Comparison between selected TFRs using a sinusoidal FM signal.

around  $t$ , the LWD can be given in the following form:

$$\text{LWD}(t, f) = \delta(f - f_i(t)) * \text{FT} \left\{ \exp \left( 2i \sum_{n=3,5,\dots}^{\infty} \phi^{(n)}(t) \frac{\tau^n}{2^n L^{n-1} n!} \right) \right\}. \quad (1.31)$$

The above equation is a generalization to (1.24); it shows that the ideal TFR, which locates energy in the TF plane exactly at the signal IF, is convolved with higher-order terms of the derivatives of  $\phi(t)$ , which causes a spread of energy around the IF trajectory. As is the case with the WD, only odd terms of order higher than third exist, but different from the WD, in the LWD, these disturbing terms are divided by a positive factor  $L^{n-1}$ , meaning that their impact can be significantly reduced by increasing the distribution order  $L$ . Stanković applied the same idea to Cohen's class of TFRs in [32]. He also presented a simple and efficient implementation method to

reduce cross-terms, which, if not dealt with properly, can be more emphatic in the LWD. In a similar research stream, TFRs with *complex-time* signal arguments were introduced [33, 34, 35] to improve the energy concentration in the TF plane for signals with rapid variations in their IF laws.

Several TFRs of a signal with sinusoidally modulated frequency is shown in Fig. 1-3.

## 1.4 Reassignment method

Improving the localization of the signal components in the TF plane and attenuating the misleading interference terms was addressed from a different perspective through the *reassignment method*; it was originally proposed for the spectrogram in [36] and substantially developed by Auger [37] and applied with other members of Cohen's class. In short, this method reassigns each value of a TFR to a different location. Consider a TFR, which, as we saw before, results by convolving the WD with a kernel  $\gamma(t, f)$  according to

$$\text{TFR}_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{WD}_z(\tau, \nu) \gamma(t - \tau, f - \nu) d\tau d\nu. \quad (1.32)$$

The above equation shows that  $\gamma(t - \tau, f - \nu)$  delimits a TF region at the vicinity of the  $(t, f)$  point, inside which, the WD values are weighted by the kernel and the average is assigned at the point  $(t, f)$ . The basic principle of the reassignment method is to move this average from the point  $(t, f)$  and reassign it at the center of gravity of its composing distribution  $(t', f')$  given by

$$\begin{aligned} t' &= \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tau \gamma(t - \tau, f - \nu) \text{WD}_z(\tau, \nu) d\tau d\nu}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma(t - \tau, f - \nu) \text{WD}_z(\tau, \nu) d\tau d\nu}, \\ f' &= \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \nu \gamma(t - \tau, f - \nu) \text{WD}_z(\tau, \nu) d\tau d\nu}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma(t - \tau, f - \nu) \text{WD}_z(\tau, \nu) d\tau d\nu}. \end{aligned} \quad (1.33)$$

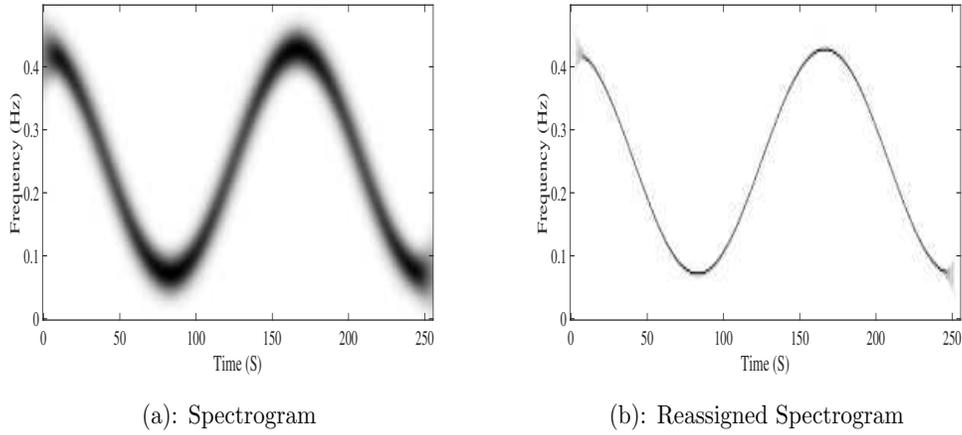


Figure 1-4: Comparison between spectrogram and reassigned spectrogram using a sinusoidal FM signal.

Given these centers of gravities, the reassigned TFR, denoted  $\text{RTFR}_z(t, f)$ , is obtained as

$$\text{RTFR}_z(t', f') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{TFR}_z(\tau, \nu) \delta(t' - \tau) \delta(f' - \nu) d\tau d\nu. \quad (1.34)$$

Through the reassignment method, excellent energy concentration can be obtained, but the method was shown to be sensitive to noise. A spectrogram and a reassigned spectrogram of a signal with sinusoidal FM are shown in Fig. 1-4.

## 1.5 Discrete implementation

We explained previously that any TFR of Cohn's class can be expressed as a convolution between the WD and a kernel. Also, the transformation to the TF domain can be via different domains: lag-Doppler, time-lag, or Doppler-frequency. Sampling the kernel in any of the previous domains is an easy task, provided that the exact definition exists (which is the case with all the kernels previously mentioned), and because the kernel is usually time and frequency bandlimited, no aliasing problems arise. Therefore, the main challenge is sampling the WD, or more specifically, sampling the IAF  $K_z(t, \tau)$ . Perhaps the most popular time-lag sampling scheme is the one proposed by Claasen and Mecklenbrauker [17], where  $K_z(t, \tau)$  is sampled in time with a sampling frequency equal to  $1/T$  and in lag with a sampling frequency equal

to  $1/(2T)$ . Let  $z[n]$  be a discrete complex-valued time series defined as

$$z[n] = \begin{cases} z_a[n], & N \leq 2N - 1 \\ 0, & 0 \leq N - 1, \end{cases} \quad (1.35)$$

where  $z_a[n]$  is the analytic associate of the input real-valued signal, the discretization of (1.10) results in

$$K_z[n, m] = z[n + m]z^*[n - m]. \quad (1.36)$$

Hence, the discrete WD can be defined as

$$\begin{aligned} \text{WD}_z[n, k] &= \sum_{m=0}^{N-1} z[n + m]z^*[n - m] \exp\left(-i\pi \frac{mk}{N}\right) \\ &= \text{DFT}_{m \rightarrow k} \{K_z[n, m]\}. \end{aligned} \quad (1.37)$$

Now, using (1.36) and (1.37), we can numerically realize any of the definitions given in (1.20)-(1.23), keeping in mind that the convolution operation can be implemented through the DFT. We remark that the use of the analytic signal for the discrete realization of the WD, as well as any QTFR, is necessary to avoid aliasing problems [38]; otherwise, the signal should be sampled at double the Nyquist rate. The time-lag domain is often used for the discrete-time implementation of TFRs, due to the computational efficiency of this realization, according to the following steps [39]:

- Generate  $K[n, m]$  using (1.36);
- Obtain the time-lag function  $R_z[n, m]$  as follows:

$$R_z[n, m] = \text{IDFT}_{l \rightarrow n} \left\{ \text{DFT}_{n \rightarrow l} \{k_z[n, m]\} g[l, m] \right\}, \quad (1.38)$$

where  $g[l, m]$  is a discrete Doppler-lag kernel;

- Form the discrete TFD according to

$$\text{TFR}_z[n, k] = \text{DFT}_{m \rightarrow k} \{R_z[n, m]\}. \quad (1.39)$$

In the above equations,  $m$ ,  $l$ , and  $k$  are indicative of lag, Doppler, and frequency, respectively. We should also mention that other definitions of discrete TFRs that satisfy some desirable mathematical properties can be found in [39, 40]. Algorithms that minimize both the computation and memory loads required to realize discrete TFRs are developed by Toole and Boashash in [41].

## 1.6 Objective assessment of TFRs

Comparing TFRs on the basis of visual inspection can be difficult and somehow subjective. The need to objectively compare and assess the quality of TFRs motivated the development of objective quantitative measures [42]. More importantly, when meaningful quantitative measures are available, they can be used as criteria of goodness to automate the selection of proper kernel parameters without interference from a user. Different measures have been proposed; they are reviewed by Stanković in [5]. Perhaps the most important measures are those that measure the energy concentration in the TF domain. In this regard, Jones and Parks [43] proposed an energy concentration measure (ECM) based on the fourth power of the  $L_4$  norm of a TFR divided by its  $L_2$  norm:

$$\text{ECM}_1 = \frac{\sum_n \sum_k \text{TFR}_z^4[n, k]}{(\sum_n \sum_k \text{TFR}_z^2[n, k])^2}. \quad (1.40)$$

A high value of the above ratio indicates that the TFR is highly concentrated. It was shown, however, that this measure favors the selection of a TFR in which one component is highly concentrated at the expense of other poorly localized components [5]. To solve this problem, Stanković in [5] proposed a simpler measure based on the support region of a TFR:

$$\text{ECM}_2(p) = \left[ \sum_n \sum_k |\text{NTFR}_z[n, k]|^{1/p} \right]^p, \quad (1.41)$$

where  $\text{NTFR}_z[n, k]$  is a normalized TFR such that  $\sum_n \sum_k \text{NTFR}_z[n, k] = 1$ , and  $p$  is a positive parameter larger than 1 ( $p = 2$  was shown to provide good results). In the same context, the Renyi entropy (RE) [44] is a quantitative measure sensitive to the number of signal components and their time durations [45]; it is defined by

$$\text{RE}(\alpha) = \frac{1}{1 - \alpha} \log_2 \left[ \sum_n \sum_k \text{TFR}_z^\alpha[n, k] \right], \quad (1.42)$$

where  $\alpha > 0$ . For even  $\alpha$ , cross-terms will contribute positively to the above measure, leading to false conclusion that the concentration improves, whereas for odd  $\alpha$ , the measure is insensitive to zero-mean cross-terms. Therefore, normalized Renyi entropy (NRE) was introduced in [46]. Also, Boashash and Sucic in [42] proposed a resolution performance measure suitable to design TFRs of multicomponent signals with high resolution, but it requires careful measuring of some attributes of a given TFR, which may be difficult to obtain in some situations.

## 1.7 Other advances

Rotating the TF plane to be aligned with the principal axis of the analyzed signal was shown to improve the energy concentration [47]. The rationale behind this approach is that in several cases, the signal under analysis has its energy concentrated along specific directions in the TF plane (or the ambiguity domain). In this context, a WD with Radon transform, called Radon WD (RWD), was introduced for detection and classification of multicomponent LFM signals in noisy environments [25, 48]. In a similar stream of research, the local polynomial Fourier transform (LPFT) was proposed to estimate both the IF and its derivatives [49, 50]. More, recently, a multi-directional rotating kernel was introduced by Boashash and Ouelha [22] for signals that can be modeled as multicomponent piecewise LFM chirps having their energy distributed along multiple directions in the TF plane.

Obtaining TFRs with high clarity is of primary interest to allow for correct extraction of relevant signal features [26]. Unfortunately, TFRs inherently compromise

between cross-term suppression and auto-term resolution. To improve the clarity and readability of TFRs, image processing techniques were successfully applied by some researchers. The resolution of the spectrogram is improved through a de-blurring mechanism in [51], while in a different approach, image de-noising algorithms including wavelet-based denoising [52], singular value decomposition (SVD) [53], and morphological processing [54] showed promising results in some cases. Khan and Sandsten [55] proposed an image processing technique to smooth a given TFR with a kernel that adapts its direction at each location in the TFR to remain aligned along its dominant ridges [24]. This technique is effective in eliminating the cross-terms without blurring the auto-terms; it is useful particularly to resolve close signal components, but it is highly expensive in terms of computations.



# Chapter 2

## The S-transform

## 2.1 Background theory

The S-transform (ST) [10] is a linear time-frequency representation (TFR) with hybrid characteristics from the short-time Fourier transform (STFT) and the wavelet transform (WT) [56]. Like the classic STFT [57], the ST provides a TF representation by projecting the analyzed signal onto local TF functions derived from the Fourier basis by a simple windowing operation. Therefore, the ST independently localizes the real and imaginary parts of the spectrum and provides globally referenced local phase information, thereby preserving the intuitive meaning of phase given by the Fourier transform (FT). Unlike the STFT, however, and similarly to the WT, the ST enables a multi-resolution analysis by using an observation window the temporal width of which reduces with increasing frequency. In doing so, the ST captures local variations in the global spectrum with progressive trade-off between time and frequency resolution in accordance with Heisenberg’s uncertainty principle<sup>1</sup>. Furthermore, the analytic Fourier-similar basis functions of the ST makes certain tasks relevant to oscillatory signal processing, such as the instantaneous frequency (IF) estimation, easier to handle compared with the WT. In addition, the progressive resolution renders a TF representation that is more consistent than that presented by a constant-bandwidth transform, like the STFT. The previous desirable features have made the ST a popular TF tool in various fields of study, such as medicine [58, 59] seismology [60], and power engineering [61, 62], just to name a few.

The ST of a continuous-time signal  $x(t)$  can be given according to

$$\text{ST}_x(t, f) = \int_{-\infty}^{+\infty} x(\tau)w(t - \tau, \sigma(f)) \exp(-i2\pi f\tau) d\tau, \quad (2.1)$$

where  $w(t, \sigma(f))$  is a Gaussian analysis window with unitary area and standard deviation  $\sigma(f)$ , which is frequency dependent. The Gaussian window is a classic choice to perform TF analysis, since it features the minimum time-bandwidth product [15];

---

<sup>1</sup>Heisenberg’s uncertainty principle states that there exists a fundamental limit to the precision with which certain pairs of physical properties can be known.

this window is defined as

$$w(t, \sigma(f)) = \frac{1}{\sqrt{2\pi}\sigma(f)} \exp\left(\frac{-t^2}{2\sigma^2(f)}\right). \quad (2.2)$$

In the original paper [10], the standard deviation of the observation window is given as  $\sigma(f) = 1/|f|$ . In subsequent works, other forms of  $\sigma(f)$  are used to meet different requirements depending on the signal under analysis, such as:  $k/|f|$ , and  $(k_0 + k_1f)/|f|$  [63, 64]. In general, any form of  $\sigma(f)$  can be adopted, provided that it is a decreasing function of frequency to achieve progressive trade-off between time and frequency resolution, which is useful when analyzing some nonstationary signals [65]. We wish to mention that STs with windows other than Gaussian were proposed, and indeed, improving the transform analysis window attracted much research interest. A bi-Gaussian window [66], a window based on the  $t$ -student distribution [67], a frequency-dependent Kaiser window [68], a compact support kernel [69], and a frequency-dependent window having hyperbolic asymmetry [11] were utilized for different applications—the STs employing these windows admit the general definition given in (2.1). Note that the Fourier oscillatory exponential kernel in (2.1) remains stationary and, unlike the window, does not translate in time, which results in absolutely referenced phase information. Meaning that the local phase information returned by the ST can be interpreted exactly as those provided by the FT. Overall, the ST can be regarded as a bridge between the STFT and the WT, maintaining the appealing features of both.

Through the convolution theorem, the FT of the analyzed signal  $X(f)$  together with the FT of the localizing window  $W(f, \sigma(f))$  can be exploited to arrive at a form equivalent to (2.1) according to

$$\begin{aligned} \text{ST}_x(t, f) &= [x(t) \exp(-i2\pi ft)] * w(t, \sigma(f)) = \text{IFT} \{X(f + \nu)W(\nu, \sigma(f))\} \\ &= \int_{-\infty}^{+\infty} X(f + \nu)W(\nu, \sigma(f)) \exp(i2\pi\nu t) d\nu, \end{aligned} \quad (2.3)$$

where  $\text{IFT}\{\cdot\}$  is the inverse FT (IFT) operator.

### 2.1.1 Inverting the ST

Linearity of the ST naturally leads to exact inevitability. Under the assumption that the analysis window satisfies

$$\int_{-\infty}^{\infty} w(t, \sigma(f)) dt = 1, \quad (2.4)$$

averaging the ST over time results in the FT of the analyzed signal according to

$$\begin{aligned} \int_{-\infty}^{\infty} \text{ST}_x(t, f) dt &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x(\tau) w(t - \tau, \sigma(f)) \exp(-i2\pi f\tau) d\tau dt \\ &= \int_{-\infty}^{\infty} x(\tau) \exp(-i2\pi f\tau) \left[ \int_{-\infty}^{\infty} w(t - \tau, \sigma(f)) dt \right] d\tau \\ &= \int_{-\infty}^{\infty} x(\tau) \exp(-i2\pi f\tau) d\tau \\ &= X(f). \end{aligned} \quad (2.5)$$

Equation (2.5) establishes a direct relationship between the ST and the FT, which permits a fast and simple retrieval of signals from their ST representations. The inverse ST described by (2.5) is usually referred to in the literature as the frequency inverse ST (FIST). Another inverting method that avoids averaging the ST over time is presented in [70] according to:

$$\tilde{x}(t) = \int_{-\infty}^{\infty} \text{ST}_x(t, f) c(f) \exp(i2\pi ft) df, \quad (2.6)$$

where  $c(f)$  is given by  $c(f) = \sigma(f)\sqrt{2\pi}$ . For a Gaussian analysis window with a standard deviation given by  $\sigma(f) = k/|f|$ , Simon *et al.* in [71] proved that the inverse ST given by (2.6) is an approximation. Eq. (2.6) can be written according to

$$\begin{aligned} \tilde{x}(t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(\tau) \exp\left(\frac{-(t - \tau)^2}{2\sigma^2(f)}\right) \exp(i2\pi f(t - \tau)) d\tau df \\ &= \int_{-\infty}^{\infty} x(\tau) \left[ \int_{-\infty}^{\infty} \exp\left(\frac{-(t - \tau)^2}{2\sigma^2(f)}\right) \exp(i2\pi f(t - \tau)) df \right] d\tau \\ &= x(t) * h(t), \end{aligned} \quad (2.7)$$

where

$$h(t) = \int_{-\infty}^{\infty} \exp\left(\frac{-t^2}{2\sigma^2(f)}\right) \exp(i2\pi ft) df. \quad (2.8)$$

Hence, for the inverse to be exact,  $h(t)$  should be equal to the Dirac function, which is not exactly satisfied, but rather approximately. Following this consideration, the exact input signal can be retrieved from  $\tilde{x}(t)$  according to [72]

$$x(t) = \text{IFT} \left\{ \frac{\tilde{X}(f)}{H(f)} \right\}, \quad (2.9)$$

in which  $\tilde{X}(f)$  and  $H(f)$  are the FTs of  $\tilde{x}(t)$  and  $h(t)$ , respectively. Inverting the ST via (2.6) is referred to as the time inverse ST (TIST).

## 2.2 Discrete ST and exact inevitability

Let  $x[nT]$ ,  $n = 0, 1, \dots, N-1$  denote a discrete time series corresponding to  $x(t)$  with a time sampling interval  $T$ . The discrete FT (DFT) is given by [10]

$$X[k] = \sum_{n=0}^{N-1} x[n] \exp\left(-i2\pi k \frac{n}{N}\right), \quad (2.10)$$

where  $n = 0, 1, \dots, N-1$ ,  $x[n] \equiv x[nT]$ ,  $X[k] \equiv X[k/N]$ , and  $k = -N/2, \dots, N/2-1$  is frequency index. Using (2.3), the ST of a discrete time series  $x[n]$  is given by

$$\text{DST}_x[n, k] = \sum_{m=0}^{N-1} x[m] w_k[n-m] \exp\left(-i2\pi \frac{mk}{N}\right), \quad (2.11)$$

where  $w_k[n]$  denotes the discrete-time version of the analysis window  $w(t, \sigma(f))$ . When the frequency variable is fixed, the 1-dimensional (1-D) function defined by  $\text{DST}_x[n, k]$  is called a *voice*; it represents the time evolution of the spectral content associated with the analyzed frequency. Each voice is a signal consisting of  $N$  samples, meaning that the DST defined in (2.11) is a  $N \times N$  TF representation, which samples the continuous ST with a uniform sampling scheme. An equivalent form of

(2.11) is given according to

$$\begin{aligned} \text{DST}_x[n, k] &= \sum_{\nu=-N/2}^{N/2-1} X[\nu + k] W_k[\nu] \exp\left(i2\pi \frac{\nu n}{N}\right) \\ &= \text{IDFT}\{X[\nu + k] W_k[\nu]\}, \end{aligned} \quad (2.12)$$

where  $\text{IDFT}\{\cdot\}$  is the inverse DFT (IDFT) operator, and  $X[\nu]$  and  $W_k[\nu]$  are the DFTs of  $x[n]$  and  $w_k[n]$ , respectively. The discretized version of the FIST is given by

$$x[n] = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} \sum_{n'=0}^{N-1} \text{DST}_x[n', k] \exp\left(i2\pi \frac{nk}{N}\right), \quad (2.13)$$

while the discretized associate of the TIST is defined as

$$\tilde{x}[n] = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} \text{DST}_x[n, k] c[k] \exp\left(i2\pi \frac{nk}{N}\right), \quad (2.14)$$

in which  $c[k] = \sum_{n=0}^{N-1} (w_k[n]/w_k[N/2]) = 1/w_k[N/2]$ . Substituting (2.12) into the equation above, we obtain

$$\tilde{x}[n] = x[n] \circledast h[n], \quad (2.15)$$

where

$$h[n] = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} c[k] w_k[n] \exp\left(i2\pi \frac{nk}{N}\right). \quad (2.16)$$

We note that in several works, the fact that the FT of a Gaussian function is also Gaussian is exploited when implementing the DST through (2.12), that is  $W_k[\nu]$  is substituted with  $\exp(-2\pi^2 \nu^2 \sigma^2 [k])$ . While it is true that the continuous FT of a Gaussian function is also Gaussian, this is not necessarily true anymore in the discrete case. Neglecting this fact leads to artifacts in the signal reconstructed via the FIST, as shown in [71]. The implementation in (2.12) in which  $W_k[\nu]$  is given by the DFT of  $w_k[n]$  and not by a direct discretization of the respective Gaussian function in the frequency domain ensures exact invertibility via (2.13). Similarly,  $c[k]$  in (2.14) is defined using the actual discrete summation of the window weights; this is because

the Gaussian windowing function used to realize the DST is in fact truncated to have the same size as the input time series, meaning that  $1/w_k[N/2]$  is slightly less than  $\sqrt{2\pi}\sigma[k]$ . Following the previous consideration, we can recover the input signal exactly from  $\tilde{x}[n]$  by

$$x[n] = \text{IDFT} \left\{ \frac{\tilde{X}[k]}{H[k]} \right\}, \quad (2.17)$$

where  $H[k]$  is the DFT of  $h[n]$ . The previous simple modifications eliminate the reconstruction errors irrespective of the methods used to invert the DST, given of course that the TFR is not modified before the signal is synthesized. Fig. 2-1 shows the reconstruction errors (in squared magnitude) between a LFM chirp and signals retrieved from its DST representation after mixing between different forward and backward algorithms; namely, we use the time domain-based implementation (TST) in (2.11) or the frequency domain-based implementation (FST) in (2.12), and the FIST in (2.13) or the TIST with reconstruction error correction according to (2.17). We see clearly from Fig. 2-1 that the synthesized signals differ from the original signal no more than what would be expected from machine precision round-off errors, regardless of the implementation or reconstruction algorithms.

## 2.3 Least-squares signal synthesis

One reason behind the popularity of linear TFRs, such as the ST, is that they enable flexible and efficient TF filtering through the analysis modification-synthesis (AMS) method in which the TFR of the signal under analysis is computed, a filtering mask is imposed, and then the filtered signal is synthesized from the modified representation. The higher weights of the mask localize regions in the TF domain which are expected to be the desired signal components, while the lower weights attenuate undesired components. Filtering with a filter  $F[n, k]$ , then back-transforming the DST, leads to

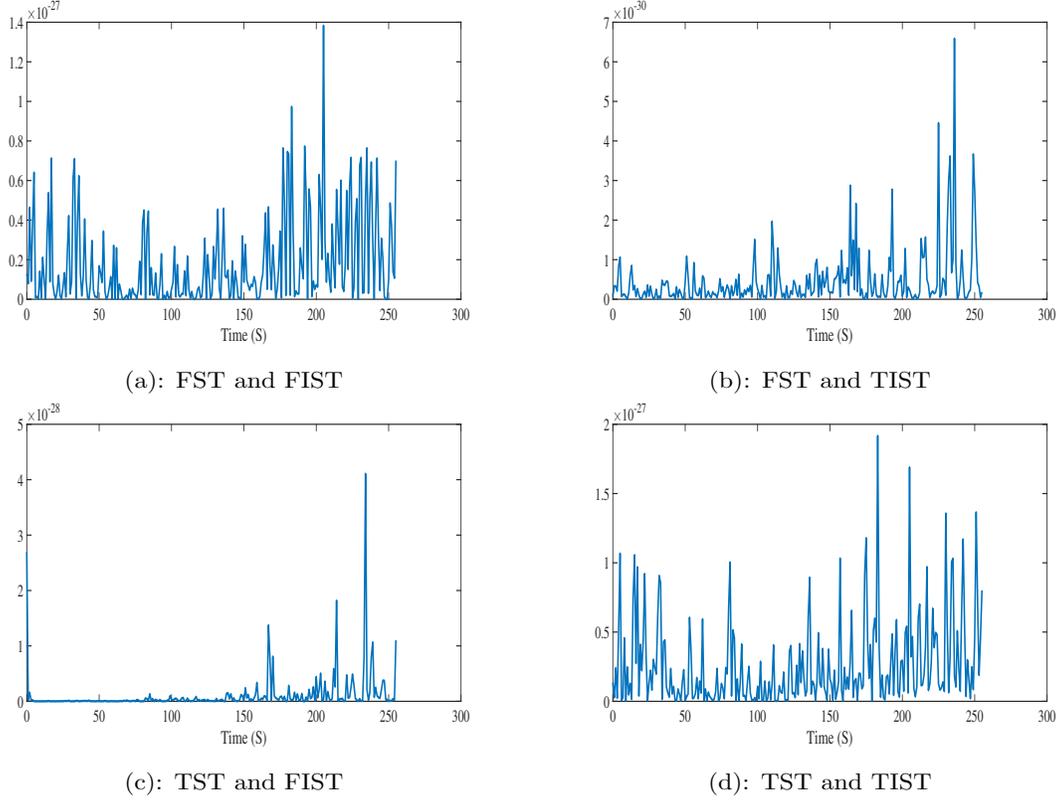


Figure 2-1: Reconstruction errors (in absolute value) resulting from transforming a signal to the TF domain then back-transforming it to the time domain through different forward and backward ST algorithms.

the following signals:

$$x_{F_1}[n] = \text{IDFT} \left\{ \sum_{n_0=0}^{N-1} \text{DST}_x[n_0, k] F[n_0, k] \right\} \quad (2.18)$$

$$x_{F_2}[n] = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} \text{DST}_x[n, k] F[n, k] c[k] \exp \left( i2\pi \frac{nk}{N} \right). \quad (2.19)$$

While the signals synthesized using the FIST and the TIST are equivalent and coincide with the input time series to machine precision (assuming that the modifications previously introduced to eliminate artifacts are adopted), these two inverses differ in their results if the DST is to be modified before constructing a filtered version of the input time series, producing different artifacts. Simon *et al.* in [71] and Pei and Wang in [72] studied the performance of the FIST and the TIST in filtering. These

works demonstrated that the FIST does not translate the time localization imposed by the filter  $F[n, k]$  directly into the time domain because of the averaging performed in time. The TIST, on the other hand, by-passes the need to averaging the ST over time, but it causes frequency leakage when filtering. Overall, it was concluded that the TIST may often be preferred if one is interested in time-domain post-processing. Conversely, the FIST can be more advantageous whenever frequency separation is given higher priority.

A common problem with the AMS method is that in most cases, the modified representation is not valid in the sense that no signal has the desired TFR. Therefore, Pei and Wang [73] presented a new inverse DST to construct a signal whose DST is the closest to a desired representation in the least-squared-error (LSE) sense; we recall this inverse briefly here. Let  $\mathbf{T}$  be a  $N^2 \times N$  matrix representing the transformation matrix of the DST given by (2.12) [or the one given by (2.11)], and let  $\mathbf{x}$  be a  $N \times 1$  vector and  $\mathbf{S}$  be a  $N^2 \times 1$  vector representing the discrete signal and its unfolded ST, respectively. Now, we can model the ST of  $x[n]$  as the matrix multiplication  $\mathbf{S} = \mathbf{T} \mathbf{x}$ . Accordingly, the vector  $\mathbf{x}$  can be recovered using

$$\mathbf{x} = \mathbf{T}^+ \mathbf{S}, \quad (2.20)$$

where  $\mathbf{T}^+ = (\mathbf{T}^* \mathbf{T})^{-1} \mathbf{T}^*$ , and  $\mathbf{T}^*$  is the conjugate transpose of the matrix  $\mathbf{T}$ . From linear algebra, it is well known that if the vector  $\mathbf{S}$  is modified, then the reconstruction given in (2.20) provides a vector whose DST is as close as possible to the desired one in the LSE sense. Obviously, synthesizing a signal based on (2.20) requires computational complexity of order  $O(N^3)$ , and if the matrix  $\mathbf{T}^+$  is stored,  $N^3$  memory elements should be allocated. To improve on complexity, we present now the mathematical derivation of the LSE inverse DST [74]. Let  $\text{DST}_z[n, k]$  be a generic modified DST; we may define a distance measure between  $\text{DST}_z[n, k]$  and another representation of

a generic signal  $y[n]$  denoted as  $\text{DST}_y[n, k]$  as the following:

$$D \{ \text{DST}_z[n, k] - \text{DST}_y[n, k] \} = \sum_{n=0}^{N-1} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} | \text{DST}_z[n, k] - \text{DST}_y[n, k] |^2. \quad (2.21)$$

Using (2.12), the DST can be written as

$$\begin{aligned} \text{DST}_y[n, k] &= \exp\left(-i2\pi k \frac{n}{N}\right) \sum_{\nu=k-\frac{N}{2}}^{k+\frac{N}{2}-1} Y[\nu] W_k[\nu - k] \exp\left(i2\pi n \frac{\nu}{N}\right) \\ &= \exp\left(-i2\pi k \frac{n}{N}\right) \text{IDFT} \{ Y[\nu] W_k[\nu - k] \}, \end{aligned} \quad (2.22)$$

which results in

$$D \{ \text{DST}_z[n, k] - \text{DST}_y[n, k] \} = \sum_{n=0}^{N-1} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} | \text{IDFT} \{ Z[\nu] W_k[\nu - k] - Y[\nu] W_k[\nu - k] \} |^2. \quad (2.23)$$

Now, by Parseval's theorem, (2.21) can be written equivalently as

$$D \{ \text{DST}_z[n, k] - \text{DST}_y[n, k] \} = \frac{1}{N} \sum_{\nu=-\frac{N}{2}}^{\frac{N}{2}-1} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} | Z_W[\nu, k] - Y[\nu] W_k[\nu - k] |^2, \quad (2.24)$$

where  $Z_W[\nu, k] = Z[\nu] W_k[\nu - k]$ . Accordingly,  $Y[\nu]$  that minimizes the distance measure in (2.24) is the one that minimizes the following expression:

$$\sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} [ Z_W[\nu, k] - Y[\nu] W_k[\nu - k] ]^2 \quad (2.25)$$

at each  $\nu$ . To determine  $Y[\nu]$ , we set the gradient of (2.25) to zero and solve in  $Y[\nu]$ , we obtain

$$Y[\nu] = \frac{\sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} Z_W[\nu, k] W_k[\nu - k]}{\sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} W_k^2[\nu - k]}. \quad (2.26)$$

The above formula defines the DFT of the signal whose DST is as close as possible to  $\text{DST}_z[n, k]$  in the LSE sense. Consequently, computation of the discrete LSE inverse DST can be given according to Algorithm 1, which has computational complexity of order  $O(N^2 \log N)$ .

---

**Algorithm 1** The LSE discrete inverse DST

---

- 1: **for**  $k = -N/2$  to  $N/2 - 1$  **do**
- 2:     compute the FFT of the following signal:

$$Z_W[\nu + k, k] = \text{FFT}_{n \rightarrow \nu} \{ \text{DST}_z[n, k] \}$$

- 3:     compute  $Y[\nu]$  according to (2.26)
- 4: **end for**
- 5: compute the Inverse FFT of  $Y[\nu]$ :

$$y[n] = \text{IFFT}_{\nu \rightarrow n} \{ Y[\nu] \}$$


---

### 2.3.1 Information redundancy

The DST of a finite-energy sequence of size  $N$  is a linear transformation from the space  $L^2(\mathbb{C}^N)$  to the space  $L^2(\mathbb{C}^{N \times N})$ . The transformation matrix  $\mathbf{T}$  is of size  $N^2 \times N$  with a rank equal to  $N$ . Accordingly, the DST is a one-to-one transform but not a cover of  $L^2(\mathbb{C}^{N \times N})$ . Each row of  $\mathbf{T}$  is a vector of size  $1 \times N$  representing the Fourier kernel windowed by a Gaussian function according to:

$$\bar{w}_i = \left[ \frac{1}{\sqrt{2\pi\sigma[k]}} \exp\left(\frac{-(n-m)^2}{2\sigma^2[k]}\right) \exp\left(-i2\pi m \frac{k}{N}\right) \right]_{m=0, \dots, N-1}. \quad (2.27)$$

The matrix  $\mathbf{T}$  can be written now in the following form:

$$\mathbf{T} = \begin{bmatrix} \bar{w}_1 \\ \bar{w}_2 \\ \vdots \\ \bar{w}_{N^2} \end{bmatrix}. \quad (2.28)$$

Since the rank of  $\mathbf{T}$  is  $N$ , recovering the vector  $\mathbf{x}$  by using (2.20) is an overdetermined problem:  $\mathbf{x}$  can be recovered using only  $N$  elements of the vector  $\mathbf{S}$  corresponding to any  $N$  independent rows of the matrix  $\mathbf{T}$ . To determine these rows, the rank-revealing QR factorization (RRQRF) method or the singular value decomposition (SVD) can be employed. However, because the size of  $\mathbf{T}$  grows fast with  $N$ , the previous algorithms are very expensive in terms of computation. Alternatively, we propose a fast method that searches for the *semi-orthogonal* row vectors of  $\mathbf{T}$ , which we define as those having the minimum dot products. Since each of these vectors has a Gaussian envelope, as evident from (2.27), it is possible to define an effective time width (ETW), denoted by  $B_t[k]$  and given in time samples, to refer to the significant part of the Gaussian envelope. This width is set according to an arbitrary attenuation in the envelope amplitude equal to  $10 \log_{10}(\alpha)$  dB, that is  $B_t[k] = 2 \left\lceil \sqrt{2 \log(\alpha) \sigma^2[k]} \right\rceil$ , where  $\alpha$  is a dimensionless positive parameter to control the value of  $B_t[k]$ . Let us consider now two rows of  $\mathbf{T}$  associated to the same frequency bin (say  $k_0$ ) and to different time samples (say  $n_i$  and  $n_j$ ), for their dot product to be negligible, their Gaussian envelopes should not overlap within the relative ETWs; hence, they should be shifted in time by at least  $|n_i - n_j| = B_t[k_0] + 1$  samples. According to the previous relation, for each frequency bin, a set of semi-orthogonal vectors can be determined. Similarly, at each time sample, a set of semi-orthogonal vectors can be identified starting from the following dot product:

$$\bar{w}_i \cdot \bar{w}_j = \frac{1}{N} (\bar{W}_i \cdot \bar{W}_j), \quad (2.29)$$

where  $\bar{W}_i$  and  $\bar{W}_j$  are the DFTs (over  $m$ ) of  $\bar{w}_i$  and  $\bar{w}_j$ , respectively. The DFT of the vector in (2.27) is given by

$$\bar{W}_i = \left[ \exp(-2\pi^2(k + \nu)^2 \sigma^2[k]) \exp\left(-i2\pi n \frac{k + \nu}{N}\right) \right]_{\nu=-N/2, \dots, N/2-1}. \quad (2.30)$$

It is plain to see that  $\bar{W}_i$  has a Gaussian envelope with a standard deviation  $1/(2\pi\sigma[k])$ , thus, and as for  $\bar{w}_i$ , an effective frequency width (EFW) given in frequency samples

can be defined as  $B_f[k] = 2N \left\lfloor \sqrt{\log(\alpha) / (\pi\sigma[k])} \right\rfloor$  in order to isolate the significant part of this envelop. Now, any two vectors associated to different frequency bins (say  $k_i$  and  $k_j$ ) are regarded semi-orthogonal if their DFTs have non-overlapping EFWs, which is satisfied when the spectral distance between their envelopes is at least  $(B_f[k_i] + B_f[k_j])/2 + 1$  frequency samples. By means of the procedure described above, the semi-orthogonal row vectors of  $\mathbf{T}$  can be easily identified. The parameter  $\alpha$  controls the number of these vectors; by increasing it, the dot products between the vectors reduce and also the possibility that some of them are linear dependent. But on the other hand, exaggeration with the value assigned to  $\alpha$  makes the number of obtained vectors less than the rank of  $\mathbf{T}$ . The suggested working methodology is to try some values around  $\alpha = 2$ , then the one that provides a number of vectors equal to or slightly larger than  $N$  (i.e., few linearly dependent vectors) is selected. The remaining vectors are discarded from  $\mathbf{T}$ , and the resulting matrix is denoted by  $\mathbf{T}_f$ . It is important to note that this matrix is signal-independent, hence constructed only once. Let  $\mathbf{S}_f$  be the ST matrix whose elements are those corresponding to the rows of  $\mathbf{T}_f$ , then the vector  $\mathbf{x}$  can be synthesized from  $\mathbf{S}_f$  according to

$$\mathbf{x} = \mathbf{T}_f^+ \mathbf{S}_f. \quad (2.31)$$

## 2.3.2 Examples

### Example 1: signal separation

Consider a signal composed of seven Gaussian components according to

$$x(t) = \sum_{n=1}^7 \exp \left[ -\pi \left( \frac{t - t_n}{T} \right)^2 \right] \exp (i2\pi f_n (t - t_n)) \quad (2.32)$$

where  $T = 12$ ,  $t_1 = t_2 = 64$ ,  $t_3 = t_4 = 176$ ,  $t_5 = t_6 = t_7 = 120$ ,  $f_1 = f_3 = 0.17$ , and  $f_2 = f_4 = 0.35$ ,  $f_5 = 0.225$ ,  $f_6 = 0.4$  and  $f_7 = 0.07$ . The signal is sampled with unitary sampling frequency and the number of signal samples is  $N = 256$ . The task is to separate the Gaussian atom associated to  $t_1$  and  $f_1$  from the composite signal  $x(t)$ .

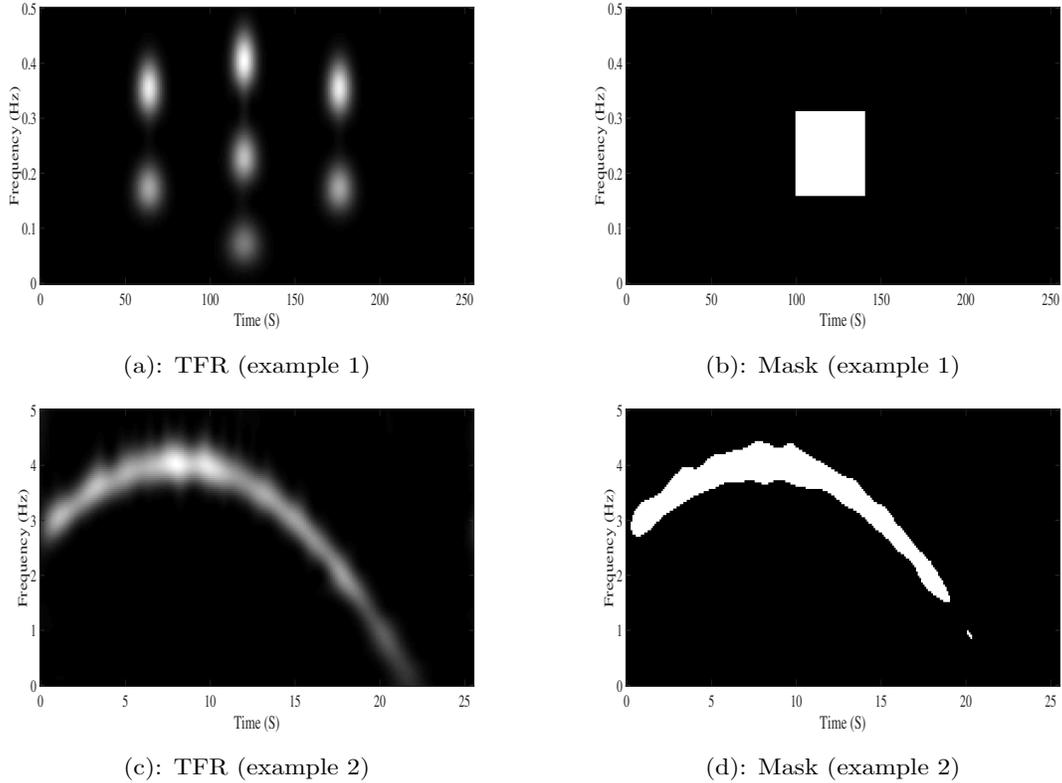


Figure 2-2: Examples of TF filtering through the ST. Magnitude of STs of two test signals in (a) and (c). Corresponding binary masks in (b) and (d).

For this purpose, the DST of  $x(t)$  is first computed using a Gaussian observation window with a standard deviation defined as  $\sigma[k] = 1/(0.1 + 0.3|k|/N)$ . Then, a binary mask is designed to pass the desired part of the representation and nullify the undesired one; finally, a filtered signal is retrieved from the masked representation through a designated inverting method. The performance of filtering is quantified in terms of the TF-MSE ( $\text{MSE}_{\text{TF}}$ ), which is defined as

$$\text{MSE}_{\text{TF}} = \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} |\text{DST}_0[n, k] - \text{DST}_{x_f}[n, k]|^2, \quad (2.33)$$

with  $\text{DST}_0[n, k]$  and  $\text{DST}_{x_f}[n, k]$  representing the masked representation and the DST of a signal  $x_f[n]$  synthesized using the modified representation, respectively. We use also in the evaluation the time-domain MSE computed between the desired time series

$x_0[n]$  and the one obtained by filtering  $x_f[n]$ :

$$\text{MSE}_T = \frac{1}{N} \sum_{n=0}^{N-1} |x_0[n] - x_f[n]|^2. \quad (2.34)$$

The magnitude of the DST of the analyzed signal and the binary filter are displayed in (a) and (b) of Fig. 2-2.

### Example 2: noise reduction

In this example we consider the following chirp:

$$x(t) = \exp \left[ i \left( 8\pi t - \frac{\pi(t-8)^3}{72} \right) \right]. \quad (2.35)$$

The sampling frequency is 10 Hz and the number of samples is  $N = 256$ . Additive white Gaussian noise (AWGN) is added to the noiseless signal at signal-to-noise ratio (SNR) equal to 10 dB. The DST of the analyzed signal is computed using a Gaussian observation window with the same standard deviation as in the previous example. The DST of the noisy signal is filtered with a binary mask, then a filtered signal is synthesized from the masked TFR. The criterion used to define the pass-region of the binary filter is:  $|\text{DST}_{x_n}[n, k]| / \max(|\text{DST}_{x_n}[n, k]|) > 0.6$ , where  $\text{DST}_{x_n}[n, k]$  is the DST of the noisy signal. The quality of denoising is quantified by using  $\text{MSE}_{\text{TF}}$  and  $\text{MSE}_T$ , defined in the previous example. The magnitudes of  $\text{DST}_{x_n}[n, k]$  and the binary filter are displayed in (c) and (d) of Fig. 2-2.

### Results

The results of  $\text{MSE}_{\text{TF}}$  and  $\text{MSE}_T$  by using different inverting methods are reported in Table 2.1. These results show that the LSE discrete IST given by Algorithm 1 provides the minimum errors, outperforming all the other methods, at least for the two considered examples. Also, by increasing the size of the vector  $\mathbf{S}_f$  (i.e., by reducing  $\alpha$ ), the errors of the discrete IST in (2.31) decrease. This result is explained by the fact that the signal is synthesized from a filtered TFR, thus by incorporating more

Algorithm	Example 1		Example 2	
	MSE <sub>TF</sub>	MSE <sub>T</sub>	MSE <sub>TF</sub>	MSE <sub>T</sub>
FIST (2.13)	$29 \times 10^{-6}$	$8.6 \times 10^{-5}$	0.012	0.28
TIST (2.14)	$26 \times 10^{-6}$	$14 \times 10^{-5}$	0.01	0.32
LSE IST (Algorithm 1)	<b><math>24 \times 10^{-6}</math></b>	<b><math>5.7 \times 10^{-5}</math></b>	<b>0.008</b>	<b>0.24</b>
IST in (2.31) ( $\alpha = 2$ )	$60 \times 10^{-6}$	$73 \times 10^{-5}$	0.017	0.43
IST in (2.31) ( $\alpha = 1.5$ )	$27 \times 10^{-6}$	$6 \times 10^{-5}$	0.01	0.3

Table 2.1: MSE<sub>TF</sub> and MSE<sub>T</sub> of two signals synthesized from modified STs through selected inverting algorithms.

samples into the vector  $\mathbf{S}_f$ , a signal closer to the theoretical one in the LSE sense is obtained. With  $\alpha = 2$ , the vector  $\mathbf{S}_f$  has 283 elements (roughly equal to  $N$ ), while with  $\alpha = 1.5$  the size of this vector increases to 485 elements (less than  $2N$ ). Note that with  $\alpha = 1.5$ , the errors of the IST in (2.31) are close to or less than those of the FIST and the TIST. The locations of the DST coefficients that constitute the vector  $\mathbf{S}_f$  are independent from the analyzed signal, depending only on the employed set of windows (recall that the matrix  $\mathbf{T}$  is signal-independent). These locations can be stored and used with other signals.

## 2.4 Controlled-coverage discrete S-transform

### 2.4.1 Limitations of the uniformly sampled ST

We explained in the previous sections how the ST uniquely combines the appealing characteristics of the STFT and the WT. However, sampling the continuous ST with a uniform TF lattice according to (2.12) or (2.11) leads to a discrete TFR with huge amount of redundant information, and, furthermore, it obliges the analysis of as many frequencies as time slots, which is a major inconvenience when dealing with large datasets, requiring computation of  $N \times N$  TF coefficients. Unlike the STFT and the WT, developing efficient sampling schemes for the ST is a topic that has not been explored enough and needs more research in our opinion. In this context, Stockwell in [3] and Brown *et al.* in [4] presented dyadically sampled one-to-one DSTs. The core idea is based on partitioning the signal spectrum into non-overlapping

subbands whose bandwidths double from each subband to the next one. The time-domain signals carrying these subbands are sampled at the Nyquist rate, meaning that the rate is also doubled from one subband signal to the next one. To divide the spectrum this way, non-overlapping boxcar functions are used in [3], whereas in [4], truncated non-overlapping Gaussian windows are employed for their improved time localization. Over-redundancy of the TF information provided by the ST is also recognized in [75]. The authors therein proposed a harmonic sampling scheme tailored for power applications. Because of the rigid manner in which they partition the frequency domain, all the previous transforms are characterized by having low spectral resolution (i.e., they have a low Q-factor, as explained later), which limits their effectiveness in processing signals of clear oscillatory behavior, such as audio, frequency-modulated, and some biomedical signals, as is the case with the dyadic WT. To process such signals, we are indeed interested in a transform that allows for a finer frequency analysis, where the spectral resolution increases more gradually. To the best of our knowledge, such a DST has not been developed yet.

Our main contribution in the coming sections is the construction of a DST with a controllable covering of the TF domain. By means of the proposed transform, the frequency domain is divided into subbands through asymmetric raised-cosine windows whose bandwidths grow progressively but in a fully controllable manner. That is, two parameters are introduced in the developed transform: one defines the size of the narrowest frequency window, while the second controls the relative frequency spacing between two successive windows. The time domain is sampled non-uniformly in accordance with the Nyquist Theorem, meaning that each subband channel has its own data rate, which is equivalent to the double-sided bandwidth of the respective partitioning window. With this level of control on its sampling scheme, the presented DST can be made a dyadic transform, and by choosing particular parameter settings, it can provide spectral resolution as high as desired. This transform requires low computational resources and produces a modest number of TF coefficients, comparable to the number of signal samples. Moreover, the windows by which the frequency domain is partitioned have unitary composite amplitude, making the proposed transform easily

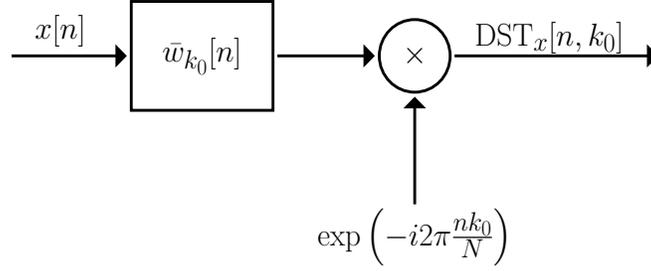


Figure 2-3: Filter-based representation of one voice provided by the DST.  $\bar{w}_{k_0}[n]$  represents the utilized window up-converted by the exponential modulator  $\exp(i2\pi nk_0/N)$ .

and exactly invertible through a simple method, which resembles the overlap addition (OLA) algorithm used for inverting the STFT.

## 2.4.2 The ST as a filter bank

Fixing the frequency index to  $k_0$ , the voice in (2.11) may be expressed using the circular convolution operator according to:

$$\begin{aligned} \text{DST}_x[n, k_0] &= \sum_{m=0}^{N-1} x[m]w_{k_0}[n-m] \exp\left[i2\pi\frac{(n-m)k_0}{N}\right] \exp\left(-i2\pi\frac{nk_0}{N}\right) \\ &= \left[x[n] \otimes w_{k_0}[n] \exp\left(i2\pi\frac{nk_0}{N}\right)\right] \exp\left(-i2\pi\frac{nk_0}{N}\right). \end{aligned} \quad (2.36)$$

The complex exponential modulator  $\exp(i2\pi nk_0/N)$  up-converts the spectral response of the analysis window, centering it at the frequency  $k_0$ , to act upon the analyzed signal as a passband filter. The result of this filtering operation is multiplied by a phase factor, which performs the necessary phase corrections, returning absolutely referenced phase information. Hence, the DST in (2.11) may be looked upon as a bank of digital filters in which each channel isolates one out of  $N$  uniformly spaced spectral subbands, and the output of each filter is phase-corrected, as shown in Fig. 2-3. The fact that these filters are nonidentical, but have bandwidths dependent on the analyzed frequencies, provides non-uniform TF resolution, distinguishing this interpretation from that of the STFT. This system of passband filters covers the signal spectrum excessively, where the number of output channels (or voices) is equal

to the signal size  $N$ . Moreover, the rates of the voices are identical and equal to the input rate, despite the fact that they have different bandwidths, in a clear departure from the uncertainty principle. With this uniform sampling lattice, the DST provides over-redundant TF information and requires expensive computational power, which overshadows its other attractive features and poses the main barrier to more widespread use of this transform.

Besides identifying the source of information redundancy, the previous filter-bank interpretation creates a roadmap to designing efficient sampling schemes for the ST. Namely, to address the aforementioned shortcomings, the filter bank should split the Fourier analysis into subbands with bandwidths that increase gradually from one voice to the next one in order to retain progressive TF resolution. These subbands should cover the signal spectrum entirely to guarantee that the transform is exactly invertible, but should not do so excessively to limit the information redundancy, hence also the computational complexity. At the outputs of these filters, the signal rates should not be identical, but rather proportional to the corresponding bandwidths, where a lower bound on the rate is imposed by the Nyquist Theorem. According to the previous considerations, a one-to-one DST, as an example, can be constructed through a filter bank whose channels are bandlimited with non-overlapping bandwidths and the output of each is critically sampled at the Nyquist rate, and if the bandwidth doubles as the channel center frequency increases, then the DST will sample the TF domain dyadically.

Stockwell in [3] develops a discrete orthonormal ST (DOST) with a dyadic sampling scheme. This transform is presented as a projection of the input time series onto basis functions constructed by linear combinations of the Fourier basis in non-overlapping frequency bands whose sizes double from one band to the next one. We shall show explicitly how this DOST results from the filter bank interpretation previously presented. If we let  $w_{k_0}[n]$  in (2.36) be a filter with a rectangular spectral response such that  $W_{k_0}[\nu] = \sqrt{\beta}$  for  $\nu \in [-\beta_0/2, \beta_0/2 - 1]$  and zero otherwise, it

follows that:

$$w_{k_0}[n] \exp\left(i2\pi \frac{nk_0}{N}\right) = \sum_{\nu=k_0-\beta_0/2}^{k_0+\beta_0/2-1} \frac{1}{\sqrt{\beta_0}} \exp\left(i2\pi \frac{\nu n}{N}\right). \quad (2.37)$$

Based on (2.36), the voice  $k_0$  can be given by:

$$\text{DST}_x[n, k_0] = \sum_{m=0}^{N-1} x[m] \sum_{\nu=k_0-\beta_0/2}^{k_0+\beta_0/2-1} \frac{1}{\sqrt{\beta_0}} \exp\left(i2\pi \frac{\nu}{N}(n-m)\right) \exp\left(-i2\pi \frac{nk_0}{N}\right). \quad (2.38)$$

Since the voice in (2.38) is bandlimited to  $[-\beta_0/2, \beta_0/2 - 1]$  in the frequency domain, it can be decimated in time by  $N/\beta_0$  without causing aliasing, equating the data rate with the Nyquist rate, which results in:

$$\text{DST}_{\downarrow x}[n, k_0] = \sum_{m=0}^{N-1} x[m] \sum_{\nu=k_0-\beta_0/2}^{k_0+\beta_0/2-1} \frac{1}{\sqrt{\beta_0}} \exp\left(i2\pi \nu \left(\frac{n}{\beta_0} - \frac{m}{N}\right)\right) \exp\left(-i2\pi \frac{nk_0}{\beta_0}\right). \quad (2.39)$$

Stockwell divides the frequency domain into non-overlapping subbands using the formulas in [3, (7-10)] by which the center frequency of a subband and its bandwidth are related by:  $k_0/\beta_0 = 3/2$ . Inserting this relation between  $k_0$  and  $\beta_0$  into (2.39), we see that  $\text{DST}_{\downarrow x}[n, k_0]$  is the inner products between the input time series and basis functions defined as:

$$S_{[n, k_0, \beta_0]}[m] = \frac{1}{\sqrt{\beta_0}} \exp(i\pi n) \sum_{\nu=k_0-\beta_0/2}^{k_0+\beta_0/2-1} \exp\left(i2\pi \nu \left(\frac{m}{N} - \frac{n}{\beta_0}\right)\right), \quad (2.40)$$

where  $1/\sqrt{\beta_0}$  is a normalization factor to insure orthonormality. The functions in (2.40) are identical to those presented in [3, (15)]. We conclude that Stockwell DOST is indeed a special configuration of the previously presented filter bank by using rectangular filters that divide the signal spectrum into non-overlapping bandlimited subbands in a dyadic manner, and at the output of each filter, the signal is down-sampled such that it is at the Nyquist rate.

Computing the DOST in (2.39) through inner products is less efficient compared

with an implementation that makes use of the Fast FT (FFT) algorithm. Namely, when  $W_{k_0}[\nu]$  is strictly bandlimited to a bandwidth  $\beta_0$ , the voice  $k_0$  can be given according to

$$\text{DST}_{\downarrow x}[n, k_0] = \text{IDFT} \left\{ X_{W_{k_0}}[\nu] \right\}, \quad (2.41)$$

where  $\text{IDFT}\{\cdot\}$  is the inverse DFT operator, usually implemented by the FFT algorithm, and  $X_{W_{k_0}}[\nu]$  is a slice of the spectrum centered at the frequency  $k_0$  bandlimited to  $[-\beta_0/2, \beta_0/2 - 1]$  and windowed by  $W_{k_0}[\nu]$ , that is  $X_{W_{k_0}}[\nu] = X[\nu + k_0]W_{k_0}[\nu]$ , where  $\nu \in [-\beta_0/2, \beta_0/2 - 1]$ . The voice obtained using (2.41) is at the Nyquist rate, but it can be easily interpolated in time by widening the spectral portion  $X_{W_{k_0}}[\nu]$  through zero-padding before performing the IDFT operation. It is clear now that the time-domain data rate of the DST can be increased easily, and in the extreme case, each voice can be fully interpolated in time by zero-padding the corresponding spectral subband until it has the same length as the input sequence. Because the windows proposed by Stockwell for partitioning the spectrum are rectangular, they have compact supports in frequency, but, via the uncertainty principle, they are poorly localized in time (they oscillate as sinc functions). In [4], the spectrum is partitioned in the same way as with the DOST but by using truncated Gaussian functions instead of boxcar windows in order to improve the time localization of the transform basis functions, resulting in another one-to-one DST.

### 2.4.3 Controlled spectral partitioning scheme

Because of the rigid and uncontrollable manner by which it decomposes the spectrum, the dyadic DST provides a TF representation with poor frequency resolution (it has a low Q-factor, as explained later), which makes it of limited effectiveness for processing signals with clear oscillatory behavior. Indeed, in many practical applications, we require a TF transform by which the spectral decomposition is controllable, thus can be made denser compared with what the dyadic DST provides.

For the sake of argument, consider the positive semi-axis of a generic continuous variable. We may use a real parameter  $\gamma > 1$  to segment this domain indefinitely

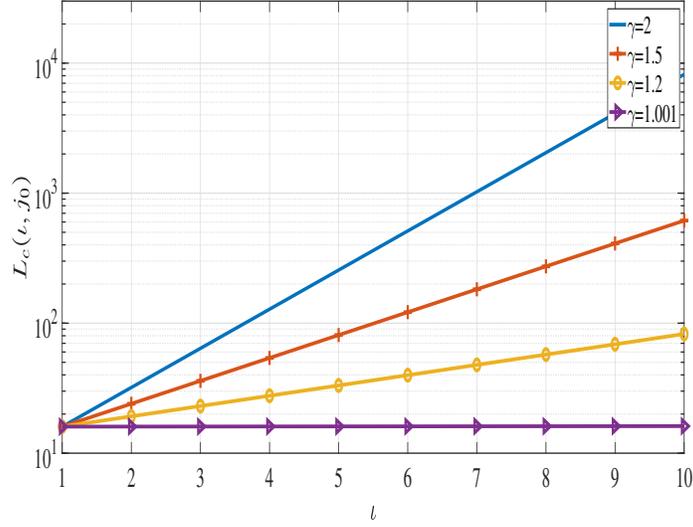


Figure 2-4: Progression of the segment size  $L_c(\iota, j_0)$  (in logarithmic scale) as a function of  $\iota$ . Increasing  $\gamma$  leads to faster growth in the segment size.

into partitions of increasing sizes as follows. We start by defining for each segment a positive quantity, we call it identity, defined as  $j_\iota = j_0 + \iota - 1$ , where  $j_0$  is a positive constant and the parameter  $\iota$  is indicative of the segment index. That is, the identity of the first segment is  $j_0$ , the identity of the second segment is  $j_0 + 1$ , and so on and so forth. We delimit these partitions in the continuous domain such that the left- and right-hand edges of a generic partition are at the coordinates  $D_L(\iota, j_0) = \gamma^{j_\iota - 1}$  and  $D_R(\iota, j_0) = \gamma^{j_\iota}$ , respectively. It follows from the last two relations that the width of a generic segment is  $L(\iota, j_0) = \gamma^{j_\iota - 1}(\gamma - 1)$  and its midpoint is at the coordinate  $D_C(j_\iota) = \gamma^{j_\iota - 1}(\gamma + 1)/2$  (note that both are dependent on the segment identity). Clearly, the width  $L(\iota, j_0)$  grows exponentially as  $\iota$  increases, where the speed of growth is governed by the parameter  $\gamma$ . With a large  $\gamma$ , segment  $\iota$  will have a size significantly wider than that of the segment  $\iota - 1$ , while as  $\gamma$  approaches 1, the difference in size becomes smaller. For example, when  $\gamma = 2$ , the width doubles from one partition to the next one, and by using a smaller  $\gamma$ , we achieve a finer segmentation of the domain, as illustrated in Fig. 2-4. Furthermore, the parameter  $j_0$  controls the size of the first segment (whose identity is  $j_0$ ). If we ask that the first segment has a particular size, say  $L_0$ , we set  $j_0 = 1 + \log_\gamma(L_0/(\gamma - 1))$ . Note that

every two successive partitions overlap only at their common border, meaning that  $D_R(\iota, j_0) = D_L(\iota + 1, j_0)$ .

The previous scheme offers features that are particularly desirable for decomposing the spectrum by the DST. Namely, the midpoint of a partition  $D_C(\iota, j_0)$  may be looked upon as the center frequency of a voice, where the width  $L(\iota, j_0)$  is seen as the bandwidth of this voice. To make this scheme return frequency indices, we shall modify the previous formulas by restricting them to produce integer values as follows. The signal spectrum is partitioned into subbands, each of which is assigned a unique positive quantity called identity, which we now define it as

$$j_\iota = j_0 + |\iota| - 1. \quad (2.42)$$

As before,  $\gamma$  is a parameter larger than 1 to control the frequency spacing between the subbands, and  $\iota$  is a parameter used for indexing these subbands. The value of  $\iota$  is positive when the corresponding subband spans positive frequencies, and vice versa, as explained later. The parameter  $j_0$  is a positive constant, determined such that the first subband contains a predefined number of frequency samples  $N_0$ :

$$j_0 = 1 + \log_\gamma \left( \frac{N_0 - 1}{\gamma - 1} \right). \quad (2.43)$$

Let us assume for now that  $\iota > 0$ , considering only the positive side of the spectrum. The frequency indices corresponding to the lower and upper edges of a generic subband are respectively given by

$$\begin{aligned} \nu_L(\iota, j_0) &= \text{Round}(\gamma^{j_\iota-1}) - \nu_0, & \iota > 0, \\ \nu_R(\iota, j_0) &= \text{Round}(\gamma^{j_\iota}) - \nu_0, & \iota > 0, \end{aligned}$$

with  $\nu_0$  being defined as  $\nu_0 = \text{Round}(\gamma^{j_0-1})$ . The previous formulas are unnumbered, as more general forms will be given later. Note that  $\nu_L(\iota, j_0)$  and  $\nu_R(\iota, j_0)$  are modified forms of  $D_L(\iota, j_0)$  and  $D_R(\iota, j_0)$ , respectively, to ensure that the lower edge of the first subband is aligned with the frequency zero. The center frequency of a subband is

given as

$$\nu_C(\iota, j_0) = \text{Round} \left( \frac{\nu_L(\iota, j_0) + \nu_R(\iota, j_0)}{2} \right). \quad (2.44)$$

The bandwidth of a subband is defined as the difference between its upper and lower edges:

$$\beta(\iota, j_0) = \nu_R(\iota, j_0) - \nu_L(\iota, j_0). \quad (2.45)$$

Note that the number of frequency samples covered by a subband is

$$N_\beta(\iota, j_0) = \beta(\iota, j_0) + 1. \quad (2.46)$$

The number of subbands should ensure a full coverage of the signal spectrum; this number can be derived as follows. Let us assume that the upper edge of the last subband is aligned with the highest frequency of the analyzed signal. Then, for this subband, whose identity is denoted herein by  $j_{max}$ , we have that:  $\gamma^{j_{max}} - \nu_0 = N/2$ . From this relation, it follows that:  $j_{max} = \log_\gamma(N/2 + \nu_0)$ . However, depending on the value of  $\gamma$  and the number of samples  $N$ , it might not be possible to achieve a spectral decomposition satisfying the previous relation. Therefore, it suffices to formulate a condition requiring the signal highest frequency to fall within the last subband. It is straightforward to see that this condition is always satisfied if the identity of the last subband is given by  $j_0 + \text{Round}(j_{max} - j_0) + 1$ , which implies, using (2.43), that the positive spectrum is partitioned into subbands the number of which is given by

$$J = \text{Round} \left\{ \log_\gamma \left[ \frac{\gamma - 1}{N_0 - 1} \left( \frac{N}{2} + \nu_0 \right) \right] \right\} + 1. \quad (2.47)$$

Because of the implicit periodicity in the DFT-domain, one can see that the last subband overlaps into the next DFT-period, causing what is known as *self-aliasing*, which is a familiar phenomenon in discrete-transform analysis. Once the positive spectrum is partitioned, the negative part is segmented likewise into subbands that mirror those spanning the positive spectrum. That is, for a subband centered at a

generic frequency  $\nu_C(\iota, j_0)$  covering the band  $[\nu_L(\iota, j_0), \nu_R(\iota, j_0)]$ , there is a counterpart centered at  $-\nu_C(\iota, j_0)$  spanning the frequencies  $[-\nu_R(\iota, j_0), -\nu_L(\iota, j_0)]$ . To ensure the foregoing, these subbands are indexed by  $\iota \in \{-J, \dots, -1, 1, \dots, J\}$ , where the quantities  $\nu_L(\iota, j_0)$ , and  $\nu_R(\iota, j_0)$  can be given now general definitions valid for positive and negative values of  $\iota$ , according to

$$\nu_L(\iota, j_0) = \text{sgn}(\iota) [\text{Round}(\gamma^{j_\iota+o}) - \nu_0], \quad (2.48)$$

$$\nu_R(\iota, j_0) = \text{sgn}(\iota) [\text{Round}(\gamma^{j_\iota-o-1}) - \nu_0], \quad (2.49)$$

where  $\text{sgn}(\cdot)$  is the signum function, and

$$o = \begin{cases} 0, & \iota \in \{-J, \dots, -1\} \\ -1, & \iota \in \{1, \dots, J\}. \end{cases} \quad (2.50)$$

Note that (2.44)-(2.46) are valid for positive and negative values of  $\iota$ . The parameters  $j_0$  and  $\iota$  characterize a subband completely, since they determine its center, lowest, and highest frequencies by application of (2.44), (2.48), (2.49). Overall, the signal spectrum is divided into  $2J$  subbands, which progressively increase in size starting from a predefined bandwidth in a manner tuned by  $\gamma$ .

#### 2.4.4 Spectral windowing

The scheme explained above divides the signal spectrum into segments, where contiguous subbands overlap only at their common edge. These frequency slices determine the support regions of the filters that constitute the DST filter bank. A weighting function (or a window) should be applied to each segment in order to shape the frequency response of the corresponding filter. The DOST of Stockwell, for example, uses rectangular windows, while, alternatively, truncated Gaussian functions are employed in [16]. The adopted weighting function defines the spectral response of the filter, and thus also its impulse response and selectivity in the time domain. Here, we propose to use asymmetric raised-cosine windows as apodization functions, defined

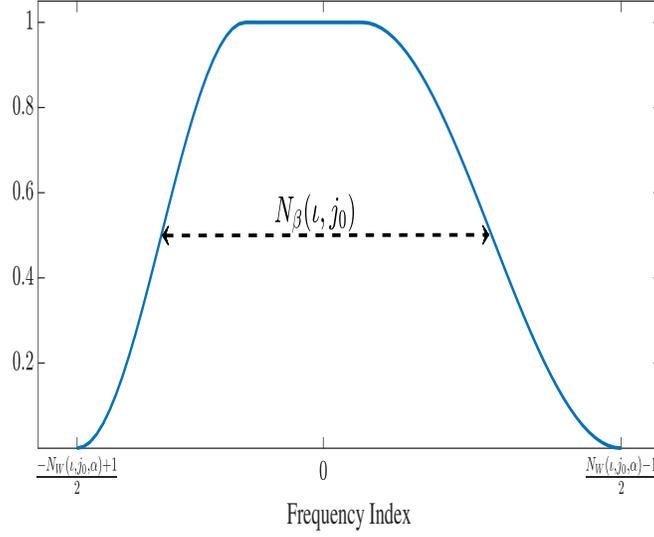


Figure 2-5: Prototype of the proposed asymmetric raised-cosine window.

by the following piecewise function:

$$W_{[l, j_0, \alpha]}[\nu] = \begin{cases} 0.5 \left[ 1 + \cos \left( \frac{\pi}{2\alpha} (t_1(\nu) + \alpha - 1) \right) \right], & \frac{-N_\beta(l, j_0) - T_1(l, j_0, \alpha)}{2} + 1 \leq \nu \leq \frac{-N_\beta(l, j_0) + T_1(l, j_0, \alpha)}{2} \\ 1, & \frac{-N_\beta(l, j_0) + T_1(l, j_0, \alpha)}{2} < \nu < \frac{N_\beta(l, j_0) - T_2(l, j_0, \alpha)}{2} \\ 0.5 \left[ 1 + \cos \left( \frac{\pi}{2\alpha} (t_2(\nu) + \alpha - 1) \right) \right], & \frac{N_\beta(l, j_0) - T_2(l, j_0, \alpha)}{2} \leq \nu \leq \frac{N_\beta(l, j_0) + T_2(l, j_0, \alpha)}{2} - 1, \end{cases} \quad (2.51)$$

where  $N_\beta(l, j_0)$ , given in (2.46), is assumed to be an odd number in the above definition, and  $\alpha$  is the roll-off factor, which is assigned a value between 0 and 1.  $T_1(l, j_0, \alpha)$  and  $T_2(l, j_0, \alpha)$  are the widths (in frequency samples) of transition bands, and  $t_1(\nu)$  and  $t_2(\nu)$  are auxiliary variables defined according to:

$$t_1(\nu) = 1 + \frac{\alpha(1 - N_\beta(l, j_0) - 2\nu)}{T_1(l, j_0, \alpha) - 1}, \quad \frac{-N_\beta(l, j_0) - T_1(l, j_0, \alpha)}{2} + 1 \leq \nu \leq \frac{-N_\beta(l, j_0) + T_1(l, j_0, \alpha)}{2}, \quad (2.52)$$

$$t_2(\nu) = 1 + \frac{\alpha(1 - N_\beta(l, j_0) + 2\nu)}{T_2(l, j_0, \alpha) - 1}, \quad \frac{N_\beta(l, j_0) - T_2(l, j_0, \alpha)}{2} \leq \nu \leq \frac{N_\beta(l, j_0) + T_2(l, j_0, \alpha)}{2} - 1. \quad (2.53)$$

The transition bands are defined as:

$$T_1(\iota, j_0, \alpha) = \begin{cases} T_R(\iota, j_0, \alpha), & \iota \in \{-J, \dots, -1\} \\ T_L(\iota, j_0, \alpha), & \iota \in \{1, \dots, J\}, \end{cases} \quad (2.54)$$

$$T_2(\iota, j_0, \alpha) = \begin{cases} T_L(\iota, j_0, \alpha), & \iota \in \{-J, \dots, -1\} \\ T_R(\iota, j_0, \alpha), & \iota \in \{1, \dots, J\}, \end{cases} \quad (2.55)$$

where

$$T_R(\iota, j_0, \alpha) = \begin{cases} T(\iota, j_0, \alpha), & T(\iota, j_0, \alpha) > 1 \\ 0, & \text{otherwise} \end{cases} \quad (2.56)$$

$$T_L(\iota, j_0, \alpha) = \begin{cases} T_R(\iota, j_0, \alpha), & |\iota| = 1 \\ T_R(|\iota| - 1, j_0, \alpha), & \text{otherwise} \end{cases} \quad (2.57)$$

$$T(\iota, j_0, \alpha) = 2\text{Round} \left[ \frac{\alpha}{2} (N_\beta(\iota, j_0) - 1) \right] + 1. \quad (2.58)$$

Note that our definitions of  $T_R(\iota, j_0, \alpha)$  and  $T_L(\iota, j_0, \alpha)$  ensure that both are odd numbers so as to equate the window amplitude with 0.5 at  $\nu = \pm \frac{1}{2} (N_\beta(\iota, j_0) - 1)$ , while  $T_1(\iota, j_0, \alpha)$  and  $T_2(\iota, j_0, \alpha)$  are defined as in (2.54) and (2.55), respectively, because we require that the negative side of the spectrum be segmented by windows that mirror their counterparts on the positive side. The roll-off factor controls the lengths of transition bands, hence measuring the bandwidth occupied by  $W_{[\iota, j_0, \alpha]}[\nu]$  beyond the fundamental width determined by the spectral partitioning scheme previously presented. For example, when  $\alpha = 0$ , the proposed window transforms into an ideal *brick-wall* filter, which has a sinc-shaped time-domain response with a very slow decay. Therefore, the presence of transition bands is indeed important to improve the time-domain localization of the transform basis functions. It is important to notice that, unlike  $T_R(\iota, j_0, \alpha)$ , the length  $T_L(\iota, j_0, \alpha)$  is defined as a percentage of the bandwidth of the subband prior to  $\iota$ , making all the frequency windows asymmetric, except for the first window and its counterpart on the other side of the spectrum (i.e.,

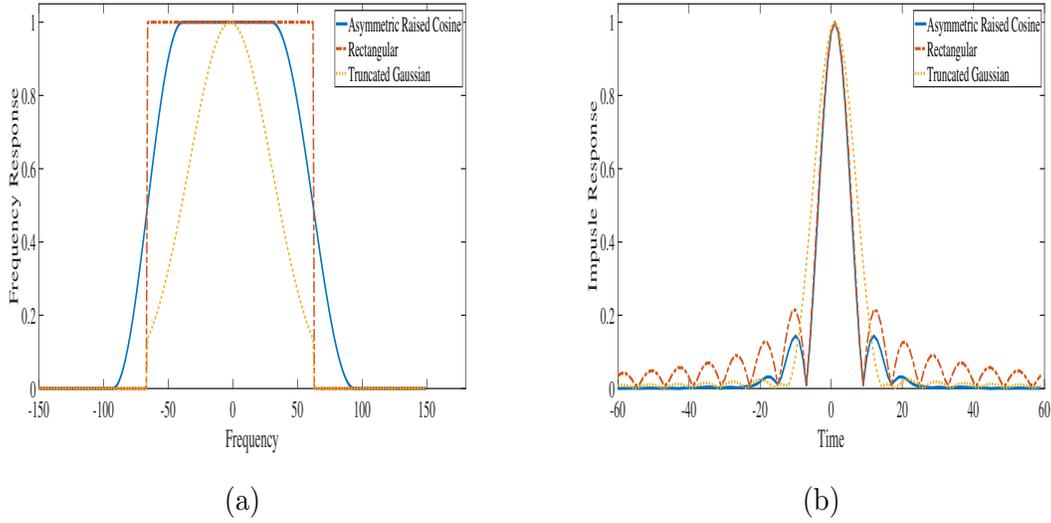


Figure 2-6: Comparison between asymmetric raised-cosine window with roll-off factor equal to 0.5, rectangular window, and truncated Gaussian window in the frequency and the time domain. Frequency responses in (a). Absolute values of time responses in (b).

for  $\iota = \pm 1$ ). Therefore, based on the proposed scheme for spectral segmentation, it is always true that  $T_L(\iota, j_0, \alpha) \leq T_R(\iota, j_0, \alpha)$ , hence with larger  $\gamma$ , the asymmetry of the window becomes more evident. A prototype of the proposed window is shown in Fig. 2-5. Due to the presence of transition bands, the size of  $W_{[\iota, j_0, \alpha]}[\nu]$  elongates beyond  $N_\beta(\iota, j_0)$ . Note that  $N_\beta(\iota, j_0)$  determines the number of frequency samples in the corresponding window at which it is true that  $W_{[\iota, j_0, \alpha]}[\nu] \geq 0.5$ . Consequently, one can see that the windowed subbands overlap, providing an over-complete spectral coverage. In Fig. 2-6, we compare an asymmetric raised-cosine window with a roll-off factor equal to 0.5 with a rectangular and truncated Gaussian windows. The spectral responses of the rectangular and Gaussian windows are confined within the band at which the amplitude of the asymmetric raised-cosine window equal to or larger than 0.5 to resemble the windows used with the dyadic one-to-one DSTs in [3, 4]. Observing the impulse responses, we find that even with a moderate value of the roll-off factor, the proposed window has a fast decay and is better localized in the time domain compared with the other two. To account for the presence of transition bands, the center frequency and the boundaries of a windowed subband should generalize

those presented in (2.44), (2.48), (2.49), respectively, according to

$$\nu_C(\iota, j_0, \alpha) = \text{Round} \left( \frac{\nu_L(\iota, j_0, \alpha) + \nu_R(\iota, j_0, \alpha)}{2} \right), \quad (2.59)$$

$$\nu_L(\iota, j_0, \alpha) = \text{sgn}(\iota) [\text{Round}(\gamma^{j_0+o}) - \nu_0] - \text{floor} \left( \frac{T_1(\iota, j_0, \alpha)}{2} \right), \quad (2.60)$$

$$\nu_R(\iota, j_0, \alpha) = \text{sgn}(\iota) [\text{Round}(\gamma^{j_0-o-1}) - \nu_0] + \text{floor} \left( \frac{T_2(\iota, j_0, \alpha)}{2} \right), \quad (2.61)$$

where we recall (2.50), (2.54), (2.55). The number of frequency samples covered by a generic subband is now given by

$$N_W(\iota, j_0, \alpha) = N_\beta(\iota, j_0) + \text{floor} \left( \frac{T_L(\iota, j_0, \alpha)}{2} \right) + \text{floor} \left( \frac{T_R(\iota, j_0, \alpha)}{2} \right). \quad (2.62)$$

## 2.4.5 DFT-based implementation

The previously explained partitioning scheme offers controlled coverage of the signal spectrum. Therefore, henceforth, the DST that employs this scheme is referred to by the controlled-coverage DST (CC-DST). Adopting the notation introduced before, the CC-DST separates the input signal into  $2J$  channels, each one carrying a voice defined similarly to (2.41) according to:

$$\text{CC-DST}_x[n; \iota, j_0, \alpha] = \frac{N_W(\iota, j_0, \alpha)}{N} \text{IDFT} \left\{ X_{W_{[\iota, j_0, \alpha]}}[\nu] \right\}, \quad (2.63)$$

in which  $N_W(\iota, j_0, \alpha)/N$  is a normalization factor, and  $X_{W_{[\iota, j_0, \alpha]}}[\nu]$  is the spectral segment confined within the band  $[\nu_L(\iota, j_0, \alpha), \nu_R(\iota, j_0, \alpha)]$  and weighted by  $W_{[\iota, j_0, \alpha]}[\nu]$ ; but here, we should differentiate whether the previous band contains an even or odd number of frequency samples. To explain this, let us consider a complex sinusoid  $h[n] = \cos(2\pi fn + \phi)$  with a frequency  $f$  that falls within the passband of a generic subband  $\iota$ , meaning that we can write  $f = \nu_C(\iota, j_0, \alpha) + \Delta_f$ . To confirm our understanding of the CC-DST as a local image of the DFT, we expect to obtain (assuming

no phase-wrapping problems):

$$\begin{aligned} \text{CC-DST}_x[n; \iota, j_0, \alpha] &= 0.5 \exp \left[ i \left( 2\pi \Delta_f \frac{nN}{N_W(\iota, j_0, \alpha)} + \phi \right) \right] \\ \text{CC-DST}_x[n; -\iota, j_0, \alpha] &= 0.5 \exp \left[ -i \left( 2\pi \Delta_f \frac{nN}{N_W(\iota, j_0, \alpha)} + \phi \right) \right]. \end{aligned}$$

To have this result,  $X_{W_{[\iota, j_0, \alpha]}}[\nu]$  should preserve the complex conjugate symmetry characterizing the DFT, and thus be defined as following:

$$X_{W_{[\iota, j_0, \alpha]}}[\nu] = \begin{cases} X[\nu + \nu_C(\iota, j_0, \alpha)] W_{[\iota, j_0, \alpha]}[\nu], & \iota > 0 \text{ or } N_W(\iota, j_0, \alpha) \\ & \text{is odd} \\ \text{Shift}_R \left\{ X[\nu + \nu_C(\iota, j_0, \alpha) + 1] W_{[\iota, j_0, \alpha]}[\nu], 1 \right\}, & \text{otherwise,} \end{cases} \quad (2.64)$$

where in the above definition, the frequency index belongs to the following range:

$$\nu \in \begin{cases} \left[ \frac{-N_W(\iota, j_0, \alpha) + 1}{2}, \frac{N_W(\iota, j_0, \alpha) - 1}{2} \right], & N_W(\iota, j_0, \alpha) \text{ is odd} \\ \left[ \frac{-N_W(\iota, j_0, \alpha)}{2}, \frac{N_W(\iota, j_0, \alpha)}{2} - 1 \right], & N_W(\iota, j_0, \alpha) \text{ is even} \end{cases} \quad (2.65)$$

and the operator  $\text{Shift}_R \{ \cdot, m \}$  shifts a generic sequence circularly by  $m$  locations to the right. Hence, in simpler words, the sequence  $X_{W_{[\iota, j_0, \alpha]}}[\nu]$  is a weighted slice of  $X[\nu]$  confined within the band  $[\nu_L(\iota, j_0), \nu_R(\iota, j_0)]$  and applied to it a one-point right circular shift in case  $\iota$  is negative and  $N_w(\iota, j_0, \alpha)$  is even. The parameters  $j_0$ ,  $\gamma$ , and  $\alpha$  give full control on the spectral partitioning scheme and determine, therefore, the lattice by which the CC-DST samples the TF domain. For example, if we ask that the subbands do not overlap, we set  $\alpha = 0$ , and if we additionally require that the bandwidth of one subband be double that of the one prior to it, we select  $\gamma = 2$ . The narrowest bandwidth is specified through the parameter  $j_0$ . If we set  $\gamma = 2$ ,  $\alpha = 0$ , and  $j_0 = 1$ , the CC-DST covers the TF domain as does the DOST of Stockwell. Reducing  $\gamma$  allows for a finer segmentation of the spectrum and, as a result, a higher frequency resolution is achieved. When  $\gamma \rightarrow 1$ , the voices returned by the CC-DST will have, roughly, equal bandwidths (because of rounding, the bandwidths might

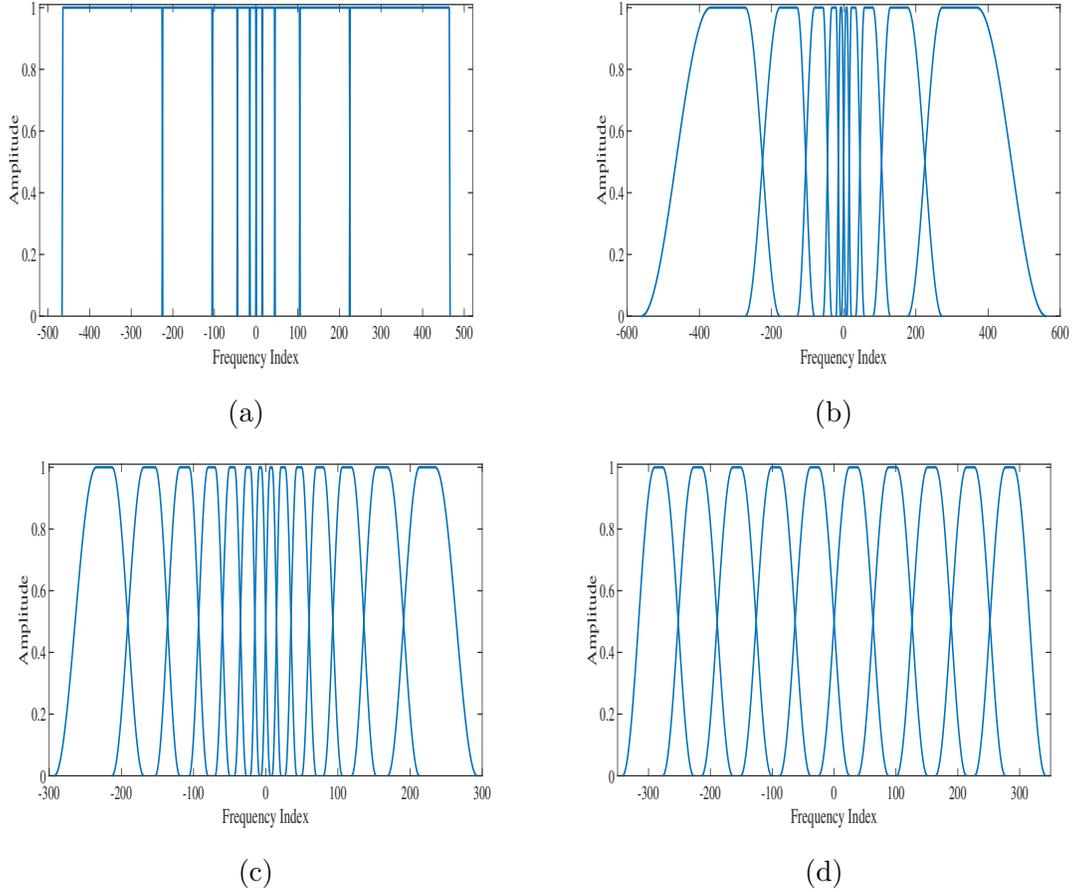


Figure 2-7: Spectral coverage provided by the CC-DST with four different parameter settings. (a)  $N_0 = 16, \gamma = 2, \alpha = 0$  (in this case, CC-DST samples the TF domain dyadically). (b)  $N_0 = 16, \gamma = 2, \alpha = 0.8$ . (c)  $N_0 = 16, \gamma = 1.3, \alpha = 0.8$  (these values provide finer spectral decomposition compared with previous cases). (d)  $N_0 = 64, \gamma = 1.0001, \alpha = 0.8$  (CC-DST in this case acts as STFT).

differ in a few frequency samples), which transforms the CC-DST into a STFT with the conventional raised-cosine filter. Since  $\gamma$  must be larger than 1, we may select  $\gamma = 1 + 1/N$  to meet the previous purpose. The IDFT-based implementation in (2.63) ties up the data rate of each voice with the bandwidth of the corresponding spectral segment via the Nyquist Theorem. That is, as the partitioning window is exactly bandlimited, its bandwidth defines the minimum density of time samples required to represent the respective voice without aliasing. It is simple, however, to change this data rate and interpolate a voice information through zero-padding  $X_{W_{[\ell, j_0, \alpha]}}[\nu]$  before performing the IDFT operation, as previously mentioned. The flexibility of

the CC-DST is illustrated in Fig. 2-7, which depicts the spectral coverage provided by this transform with different parameter settings.

### 2.4.6 Exact reconstruction

The linear relationship between the DST and the DFT implies that the DST is exactly invertible, which makes it is a valuable tool not only for TF analysis but also synthesis. As we established before, a voice of the DST is essentially the IDFT of a weighted slice of the spectrum. Consequently, provided that the weighting windows have nonzero values, by taking the DFTs of the voices then dividing out the corresponding windows, the spectral portions that are used in the DST analysis can be exactly recovered. These segments should be reconnected carefully to synthesize back the spectrum. If the voices are computed using the entire spectrum, as is the case with the original uniformly sampled DST, the reconstruction method is further simplified as one voice is enough to retrieve back the spectrum. As for the CC-DST, the inverting process is straightforward: all that is required to synthesize the signal spectrum is taking the DFT of each voice, placing the resulting sequence within the corresponding band  $[\nu_L(\iota, j_0, \alpha), \nu_R(\iota, j_0, \alpha)]$ , then adding up these frames together. From (2.63), we see that  $X_{W_{[\iota, j_0, \alpha]}}[\nu]$  results by taking the DFT of  $\text{CC-DST}_x[n; \iota, j_0, \alpha]$ . Once the result is available, we can construct a signal defined as

$$\tilde{X}_{[\iota, j_0, \alpha]}[\nu] = \begin{cases} X_{W_{[\iota, j_0, \alpha]}}[\nu - \nu_C(\iota, j_0, \alpha)], & \nu \in [\nu_L(\iota, j_0, \alpha), \nu_R(\iota, j_0, \alpha)] \text{ and} \\ & \{N_w(\iota, j_0, \alpha) \text{ is odd or } \iota > 0\} \\ Y_{W_{[\iota, j_0, \alpha]}}[\nu - \nu_C(\iota, j_0, \alpha) - 1], & \nu \in [\nu_L(\iota, j_0, \alpha), \nu_R(\iota, j_0, \alpha)] \text{ and} \\ & N_w(\iota, j_0, \alpha) \text{ is even and } \iota < 0 \\ 0, & \text{otherwise,} \end{cases} \quad (2.66)$$

where  $Y_{W_{[\iota, j_0, \alpha]}}[\nu] = \text{Shift}_R\{X_{W_{[\iota, j_0, \alpha]}}[\nu], -1\}$  (i.e., circularly shifting  $X_{W_{[\iota, j_0, \alpha]}}[\nu]$  by 1 position to the left). Clearly, the signal  $\tilde{X}_{[\iota, j_0, \alpha]}[\nu]$  can be given equivalently according

to the following formulation:  $\tilde{X}_{[\ell, j_0, \alpha]}[\nu] = X[\nu]\tilde{W}_{[\ell, j_0, \alpha]}[\nu]$ , where

$$\tilde{W}_{[\ell, j_0, \alpha]}[\nu] = \begin{cases} W_{[\ell, j_0, \alpha]}[\nu - \nu_C(\ell, j_0, \alpha)], & \nu \in [\nu_L(\ell, j_0, \alpha), \nu_R(\ell, j_0, \alpha)] \text{ and} \\ & \{N_w(\ell, j_0, \alpha) \text{ is odd or } \ell > 0\} \\ W_{[\ell, j_0, \alpha]}[\nu - \nu_C(\ell, j_0, \alpha) - 1], & \nu \in [\nu_L(\ell, j_0, \alpha), \nu_R(\ell, j_0, \alpha)] \text{ and} \\ & N_w(\ell, j_0, \alpha) \text{ is even and } \ell < 0 \\ 0, & \text{otherwise.} \end{cases} \quad (2.67)$$

Now, if we sum up  $\tilde{X}_{[\ell, j_0, \alpha]}[\nu]$  for each  $\ell$  at which the CC-DST was computed, we get

$$\begin{aligned} Y[\nu] &= \sum_{\ell} \tilde{X}_{[\ell, j_0, \alpha]}[\nu] = \sum_{\ell} X[\nu]\tilde{W}_{[\ell, j_0, \alpha]}[\nu] \\ &= X[\nu] \sum_{\ell} \tilde{W}_{[\ell, j_0, \alpha]}[\nu]. \end{aligned} \quad (2.68)$$

This equation makes clear that, for  $Y[\nu]$  to be identical to  $X[\nu]$  within one DFT-period, the sum of all the overlapping windows  $\tilde{W}_{[\ell, j_0, \alpha]}[\nu]$  should add up to unity at the frequencies  $-N/2 \leq \nu \leq N/2 - 1$ . If this condition is satisfied, the synthesis procedure becomes a matter of summing up overlapping sections, each of which results by taking the DFT of one of the CC-DST voices. The proposed asymmetric raised-cosine windows are constructed in a way so as to satisfy the previous condition. Namely, these frequency windows are unitary in their passbands and null in their stopbands, but they have asymmetric transition bands, which are cosine-tapered sections. For every two successive windows, the right-hand transition band of the former and the left-hand transition band of the later are identical in length and sum up to unity. The procedure described above is the *frequency-domain counterpart* of the OLA method [57], which is widely used for inverting the STFT. The difference between the two is that, with the OLA method, overlapping time-domain frames are used to reconstruct the input signal, while with the one presented here, the spectrum is retrieved by summing up overlapping sections and then the time-domain signal is obtained by taking

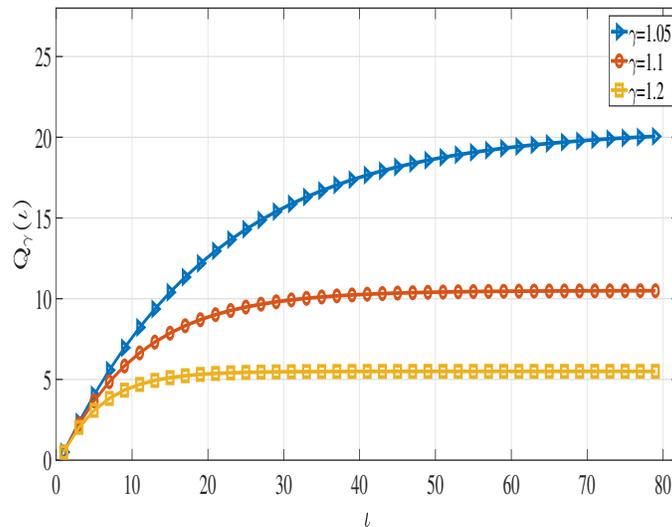


Figure 2-8: Variations of the Q-factor of the CC-DST. By reducing  $\gamma$ ,  $Q(\ell)$  increases, but, on the other hand, larger  $\gamma$  leads to smaller variations in the Q-factor.

the IDFT of the resulting sequence.

### 2.4.7 Quality-factor and complexity

The main motivation behind developing the CC-DST is to have a low-complexity DST with high spectral resolution. The Q-factor (i.e., quality factor) of a function, which is defined as the ratio of its center frequency to its bandwidth, describes quantitatively the resonance behavior of this function. A TF transform is said to be *constant-Q* if all of its voices have the same Q-factor [76]. In the original DST, the employed frequency window is a Gaussian function with a standard deviation that is proportional to frequency. Therefore, it is a constant-Q transform, as opposed to the STFT, which is a constant-bandwidth transform. For the dyadic DSTs presented by Stockwell in [3] and Brown *et al.* in [4], the ratio of the center frequency of a voice to its bandwidth is constant and equal to  $3/2$  [3, (7-10)]. Accordingly, these two transforms are constant-Q, but they both have a low and uncontrollable Q-factor, which makes them of limited effectiveness for processing oscillatory signals. Let us now compute the Q-factor of the proposed CC-DST. As we previously explained, this transform splits the spectrum into  $2J$  subbands, each is weighted by an asymmetric raised-cosine window whose

frequency support is  $[\nu_L(\iota, j_0, \alpha), \nu_R(\iota, j_0, \alpha)]$ . We define for this window an effective bandwidth  $\text{BW}(\iota, j_0)$  as the full width at half maximum (FWHM), meaning that it is the difference between the two frequencies at which the amplitude of the window is 0.5. Clearly, the effective bandwidth of a window (or a voice) is equal to the width of the band  $[\text{Round}(\gamma^{j_0-1}) - \nu_0, \text{Round}(\gamma^{j_\iota}) - \nu_0]$  (assuming  $\iota$  is positive). To simplify the calculations, in what follows, we forgo the rounding function. Hence, the effective bandwidth of a voice is given by

$$\text{BW}(\iota, j_0) = \gamma^{j_\iota-1}(\gamma - 1). \quad (2.69)$$

As for the resonance frequency, we define it as the midpoint of the previous band:

$$\text{RF}(\iota, j_0) = \frac{1}{2}\gamma^{j_\iota-1}(\gamma + 1) - \gamma^{j_0-1}. \quad (2.70)$$

The Q-factor of a CC-DST voice can be given now according to

$$\text{Q}_\gamma(\iota) = \frac{\text{RF}(\iota, j_0)}{\text{BW}(\iota, j_0)} = \frac{\gamma^{|\iota|-1}(\gamma + 1) - 2}{2\gamma^{|\iota|-1}(\gamma - 1)}. \quad (2.71)$$

This formula reveals two important properties of the CC-DST. First, the Q-factor can be tuned by  $\gamma$ , and the smaller is  $\gamma$ , the larger is  $\text{Q}_\gamma(\iota)$ . Second, the Q-factor is not constant, but rather increases with  $\iota$ , which implies that the high-frequency voices have larger Q-factors than those that resonate at lower frequencies. However, as the center frequency of the voice increases, the value of  $\gamma^{|\iota|-1}(\gamma + 1)$  in the numerator of  $\text{Q}_\gamma(\iota)$  becomes much larger than 2, which implies that the growth of  $\text{Q}_\gamma(\iota)$  at higher-frequency voices (i.e., those with larger  $\iota$ ) is small. That is, after some voices, the CC-DST becomes *approximately* a constant-Q transform. Note that when choosing a value for  $\gamma$ , there is always a trade-off between having a large Q-factor, which requires a small  $\gamma$ , and having a transform with rapidly stabilizing Q-factor, which is achieved through a large  $\gamma$ , but at the expense of lowering the value of  $\text{Q}_\gamma(\iota)$ , as illustrated in Fig. 2-8.

The CC-DST is implemented based on the DFT, which is usually realized through

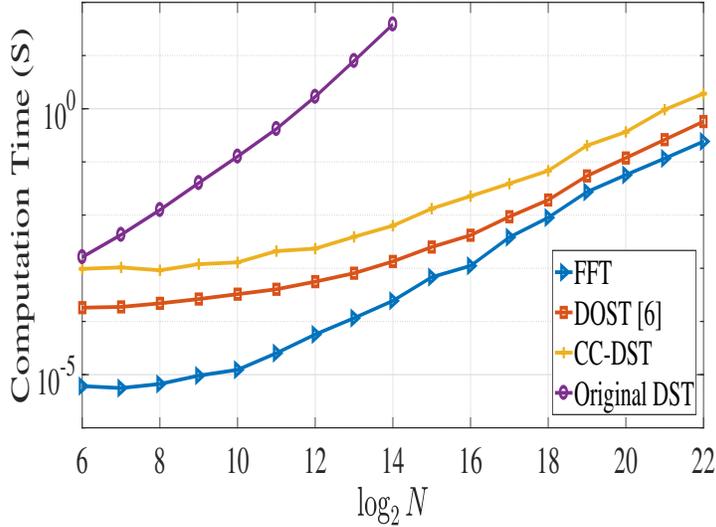


Figure 2-9: Comparison between efficiently sampled DSTs and the original uniformly sampled DST in terms of computation time. Computation speed of FFT is added as a reference point. The largest signal processed by original DST was of size  $2^{14}$  samples, due to memory limitations. The CC-DST is slightly slower than the one-to-one DOST, but both have computation speeds comparable to those of FFT and show similar growth rate in computation time. This experiment was run on an Intel Core i-5 platform having 6 GB RAM and MATLAB was used for implementation.

the FFT algorithm. Particularly, this transform requires  $2J$  FFTs, each of which is of length equal to the corresponding  $N_w(\ell, j_0, \alpha)$ , which is not a power of two in most cases, meaning that the FFT-based implementation might be suboptimal. Nonetheless, the CC-DST works fast and is used in the next section to analyze signals of large numbers of samples. It is also worthy of noting that by using the chirp Z-transform algorithm, the DFT can be implemented with a computational cost equivalent to that in which the number of signal samples is a power of 2. Using straightforward computations, the number of TF coefficients returned by the CC-DST can be approximated by:

$$R(\gamma, j_0, \alpha, J) \approx \text{Round} \{ [2 + \alpha(1 + 1/\gamma)] [J + \gamma^{j_0-1}(\gamma^J - 1)] \}. \quad (2.72)$$

As one might expect, the number of coefficients increases with the roll-off factor  $\alpha$ , and even with  $\alpha = 0$  (i.e., brick-wall filters), this number exceeds the number of signal samples  $N$ , since the spectral coverage of the CC-DST extends beyond one

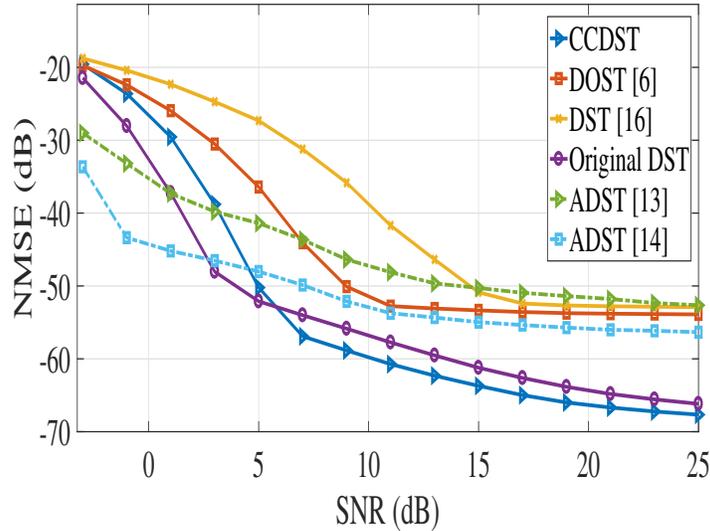


Figure 2-10: Comparison between several DSTs in terms of IF estimation. Figure shows the NMSE (in dB) of the IF estimates of a sinusoidally FM signal as a function of SNR (in dB). Starting from SNR = 5 dB, CC-DST ( $N_0 = 64$ ,  $\gamma = 1.1$ ,  $\alpha = 0.8$ ) provided the most accurate IF estimates.

period in the DFT-domain because of the self-aliasing. We note that the phenomenon of self-aliasing can be eliminated by truncating the last frequency window and its counterpart on the negative spectrum so that they do not overlap into contiguous DFT-periods. Generally, the CC-DST produces a modest number of TF coefficients. Actual computation times for the CC-DST, the DOST, the original DST, and the FFT algorithm are compared in Fig. 2-9. Due to memory limitations, the largest signal processed by the original DST was of size  $2^{14}$  samples. The DST of Brown *et al.* was not included in the benchmark comparison for brevity, since it requires computation times similar to those of the DOST. Typical values for the CC-DST parameters were selected:  $N_0 = N/16$ ,  $\gamma = 1.2$ , and  $\alpha = 0.8$ . The depicted results demonstrate the numerical efficiency of the CC-DST. This transform requires slightly more computation times with respect to the one-to-one DOST, and, furthermore, the computation time shows a growth rate similar to that of the DOST and the FFT, which is  $O(N \log_2 N)$ .

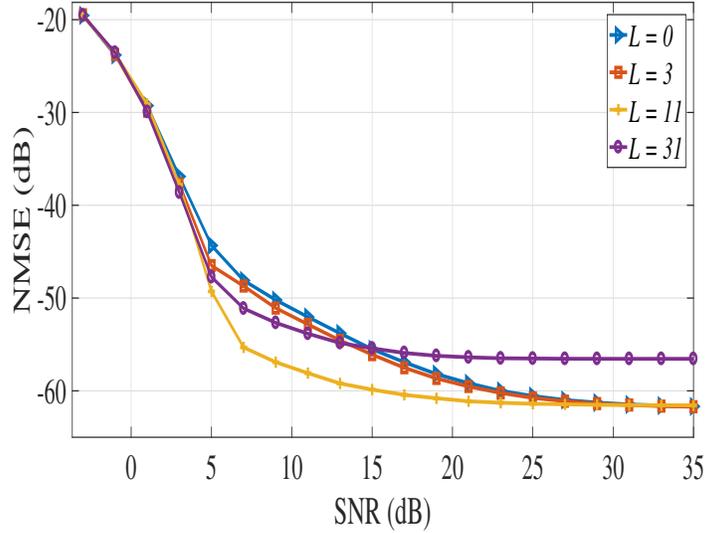


Figure 2-11: IF estimation of a sinusoidally FM signal using the CC-DST with a moving average filter of various sizes. The filter aims at reducing the noise power in the local phase estimates. Figure shows the NMSE of the IF estimates as a function of SNR. Results demonstrate that a filter with a moderate smoothing length should be used.

### 2.4.8 Application: IF estimation

The instantaneous frequency (IF) estimation is a fundamental application of the TFRs [13, 14]. The concept of the IF is explained in depth in the next chapter; here, the aim is just to illustrate how the CC-DST can be used for IF estimation. Let  $x(t) = A(t)\exp(i\phi(t))$  be the analytic associate of a continuous-time real-valued monocomponent signal whose time-varying amplitude and instantaneous phase (IP) are denoted by  $A(t)$  and  $\phi(t)$ , respectively. The IF of  $x(t)$  is defined as  $\text{IF}(t) = \phi'(t)/(2\pi)$ . Thanks to its absolutely referenced phase information, the CC-DST extends the concept of the IF, formulating a definition of the *channel* IF (CIF) [3]. That is, the CIF is simply the IF of a channel (or a voice) isolated by the CC-DST. Let us consider one of the  $2J$  voices provided by the CC-DST, say  $\iota_0$ , the returned local amplitude and phase information enables one to write the voice  $\iota_0$  in the following form:

$$\text{CC-DST}_x[n; \iota_0, j_0, \alpha] = A[n; \iota_0, j_0, \alpha] \exp \left[ i\phi[n; \iota_0, j_0, \alpha] \right]. \quad (2.73)$$

We may use the local phase information in (2.73) to define the CIF according to

$$\text{CIF}[n; \iota_0, j_0, \alpha] = \frac{1}{N} \nu_C(\iota_0, j_0, \alpha) + \frac{1}{2\pi} \text{deriv}\{\text{unwrap}\{\phi[n; \iota_0, j_0, \alpha]\}\}, \quad (2.74)$$

in which  $\text{deriv}\{\cdot\}$  is a function that computes the numerical derivative, and  $\text{unwrap}\{\cdot\}$  is a phase unwrapping function. Note that the CIF in (2.74) is given normalized to the sampling frequency. As for estimating the signal IF, an effective and widely used method for nonparametric IF estimation is based on the maxima positions of a TFR. This method, however, cannot be used directly with the CC-DST, since this transform does not provide a uniformly sampled TFR (i.e., the CC-DST cannot be put in a matrix form). The maxima-based method, nonetheless, can be adapted to work with the CC-DST as follows. First, the voices are interpolated in time so that they all have the same rate as the analyzed signal. In other words, the data rate of a generic voice  $\iota_0$  is raised by  $N/N_W(\iota_0, j_0, \alpha)$ . Second, at a given time instant  $n_0$ , we identify the voice at which the CC-DST peaks in power according to:

$$\iota_{\max} = \arg \max_{\iota} |\text{CC-DST}[n_0; \iota, j_0, \alpha]|^2. \quad (2.75)$$

The IF falls within the subband  $\iota_{\max}$ . Now, to account for the deviation of the IF from the center frequency of the voice  $\iota_{\max}$ , at the considered time instant, the IF is estimated as the CIF of the voice  $\iota_{\max}$ :

$$\text{IF}[n_0] = \text{CIF}[n_0; \iota_{\max}, j_0, \alpha]. \quad (2.76)$$

If the signal is not contaminated by noise, the derivative in (2.74) can be computed simply through the finite-difference method. In the presence of noise, however, the phase samples should be smoothed by a filter before applying the finite-difference method in order to suppress the noise power. We remark that numerical differentiation of a noisy process is known to be an ill-posed problem, meaning that without a smoothing filter for noise suppression, the accuracy of the IF estimate would be poor. To illustrate applicability of the proposed method, let us consider an analytic

frequency modulated (FM) signal with a unitary amplitude and a sinusoidal IF law corrupted by complex-valued additive white Gaussian noise (AWGN). The signal IF law is defined as

$$\text{IF}[n] = 0.175 + 0.125 \cos \left[ \frac{2\pi T(n - 511)}{N} - \frac{\pi}{2} \right], \quad 0 \leq n < 1024, \quad (2.77)$$

where  $T$  is a parameter to control the period of the sinusoidal frequency modulation, set here to 1. The signal is analyzed using the CC-DST with  $N_0 = 64$ ,  $\gamma = 1.1$ , and  $\alpha = 0.8$ , then the IF is estimated through the previously explained algorithm. The phase samples are smoothed using an average moving filter of length  $L = 11$  before differentiating the result through the conventional finite-difference method, as explained before. Results of the normalized mean squared error (NMSE) of the IF estimates are depicted in Fig. 2-10 as a function of the signal-to-noise ratio (SNR). The statistical data are obtained by 200 realizations. For comparative purposes, we provide the results of the DOST of Stockwell, the DST of Brown *et al.*, and the original DST. Moreover, we include in the benchmark compression two adaptive DSTs (ADSTs) optimized based on energy concentration (EC) measures proposed by Sejdić *et al.* in [1] and Pei and Wang in [2]. Note that the DSTs of Stockwell and Brown *et al.* are used to estimate the IF following the method proposed here, while with the original DST and the ADSTs, the IF is estimated based on the maxima positions of the computed TFR. The results in Fig. 2-10 show that at moderate and high SNRs (i.e.,  $\text{SNR} > 5$  dB), the CC-DST returns the most accurate IF estimates, at least for this signal and with the selected set of parameters, outperforming the original DST and the ADST. However, at lower SNRs, the uniformly sampled DSTs, particularly the ADST of [1], appear to provide more accurate IF estimates, but, of course, they require much more involved computations. This example clearly demonstrates the improvement achieved by the CC-DST over the DSTs of Stockwell and Brown *et al.* in IF estimation accuracy; it also shows that the proposed method for IF estimation requires moderate to high SNRs to be effective.

The length of the smoothing filter  $L$  is an important parameter of the proposed

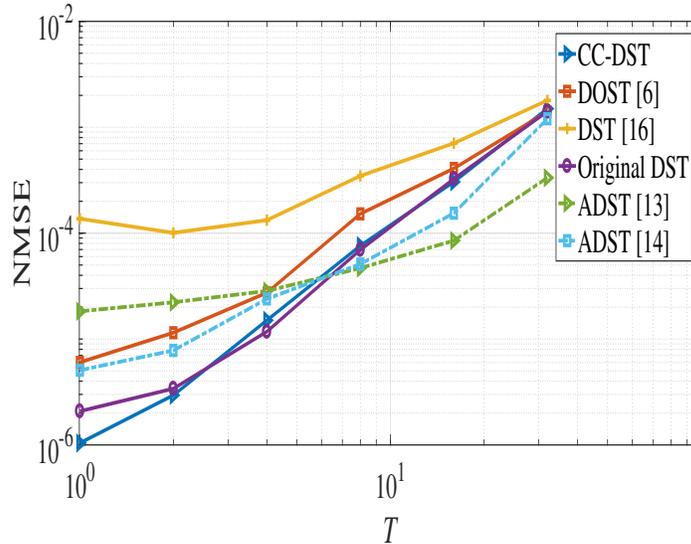


Figure 2-12: Performance of IF estimators based on selected DSTs. Results illustrate the impact of the period of sinusoidal frequency modulation, controlled by the parameter  $T$ , on the accuracy of the IF estimates. Increasing  $T$  leads to shorter IF periods. Figure shows the NMSE of the IF estimates as a function of  $T$  at SNR = 10 dB using log-scales for clarity. At low  $T$ , the CC-DST ( $N_0 = 64$ ,  $\gamma = 1.1$ ,  $\alpha = 0.8$ ) returns accurate IF estimates, while at large  $T$ , the ADSTs are superior.

algorithm for IF estimation. As previously mentioned, the smoothing filter aims at reducing the noise power in the local phase samples before computing the derivative in (2.74). Although larger  $L$  might lead to higher reduction in the noise power, it causes distortion in the differentiated signal, due to loss of derivative details. More importantly, the use of overlapped frequency windows to construct the CC-DST implies that the errors in the phase samples are correlated, hence effectiveness of the smoothing filter is limited. To illustrate the impact of the parameter  $L$  on the accuracy of the IF estimate obtained through the CC-DST, we display in Fig. 2-11 the NMSE of the IF estimates of a sinusoidally FM signal, defined in (2.77) with  $T = 1.5$ , as a function of the SNR using a small ( $L = 3$ ), medium ( $L = 11$ ), and large value of  $L$  ( $L = 31$ ), and also with no smoothing (i.e.,  $L = 0$ ). The results show that starting from a SNR equal to 5 dB, the use of an average moving filter of medium size  $L = 11$  significantly improves the accuracy of the IF estimate compared with filters of shorter or larger lengths. However, at high SNRs (i.e., SNR > 25 dB), smoothing the phase

samples with a short- or moderate- length filter does not result in reduction of the NMSE, whereas excessive smoothing worsens the accuracy of the IF estimate. This example confirms that the smoothing size should be moderate.

In the last experiment, we evaluate the impact of the parameter  $T$  on the accuracy of the IF estimate. In Fig. 2-12, we depict the NMSE of the IF estimates, obtained through the same STs used in the previous experiment, as a function of the parameter  $T$ , fixing the SNR to 10 dB. The results show that reducing the period of the sinusoidal frequency modulation (i.e., increasing  $T$ ) leads to increase in the NMSE. At low values of  $T$ , the CC-DST, with the selected set of parameters ( $N_0 = 64$ ,  $\gamma = 1.1$ ,  $\alpha = 0.8$ ) returns more accurate IF estimates compared with the other considered forms of the ST, whereas at high  $T$ , the ADSTs, particularly the ADST of Sejdić *et al.*, are superior in terms of IF estimation, thanks to their adaptive adjustment of the observation window. It is important to point out that the results of the CC-DST are dependent on the used set of parameters; hence, all the comparisons made here are relative.

### 2.4.9 Application: time-scaling audio signal

Time-scaling audio signals is one of those applications that require a tool for signal analysis as well as synthesis, like the CC-DST. The goal is to slow down or speed up the temporal evolution of a given audio signal without altering its harmonic content. At each time instant, the analyzed audio signal can be modeled as the sum of sinusoids, each of which has a time-varying amplitude and an IP according to [77]:

$$f(t) = \sum_{l=0}^I A_l(t) \exp(i\phi_l(t)). \quad (2.78)$$

Ideally, scaling  $f(t)$  in time by a constant modification factor  $\delta$  would give:

$$f_\delta(t) = \sum_{l=0}^I A_l(\delta t) \exp \left[ 2\pi i \int_0^t \text{IF}_l(\delta\tau) d\tau \right]. \quad (2.79)$$

Through the CC-DST, one can expand the input signal into channels, each carrying a portion of the spectrum. If the spectral segmentation is dense enough, the

channels will have slowly varying spectra, and thus, each subband channel can be regarded as a single harmonic component. By modifying the phase and amplitude of each of these components according to (2.79), then collapsing the resulting CC-DST into a time-domain signal, we can obtain a time-scaled replica of the input audio signal. This application shows the necessity for a DST with high frequency resolution in order to scale the audio signal with high fidelity. Employing a dyadic DST for this application is ineffective because of its poor spectral resolution. Traditionally, time-scaling a recorded sound is performed through the STFT-based phase vocoder, which has found many commercial applications. This technique, however, introduces a well-known perceptual artifact known as *phasiness* (or *loss of presence*), due to which, the time-scaled audio signal is perceived as if the source is much farther from the microphone compared with original recording [77]. Moreover, being a constant-bandwidth transform, the STFT is not consistent with the human auditory system, which is rather modeled as a constant-Q device [76]. These two aspects constitute an incentive for using the CC-DST in this application in place of the STFT. As an example, we time-scaled a three-second audio signal sampled at 16 kHz using the CC-DST with  $N_0 = 128$ ,  $\gamma = 1.06$ , and  $\alpha = 0.8$ . For comparison, we applied the DSTs of Stockwell and Brown *et al.*, the STFT, and the complex WT (CWT) proposed in [78] for the same task. Because of their low spectral resolution, it is unsurprising that the dyadic one-to-one DSTs perform poorly in this application. The STFT-based phase vocoder<sup>2</sup> is fast and performs well, but the resulting audio signal suffers from clear phasiness. The CC-DST with the previous set of parameters time-scales the audio signal with high fidelity without producing phasiness. This phenomenon is also eliminated by using the CWT<sup>3</sup>, but this transform requires about 5 times the time required by the CC-DST to time-scale the considered signal, and, furthermore, it introduces some audible distortion, which is not present when the CC-DST is used. Because the considered signal is of a large number of samples, processing it with a uniformly sampled  $N \times N$  DST is impractical on a regular workstation.

---

<sup>2</sup>MATLAB code is available at <http://www.ee.columbia.edu/~dpwe/resources/matlab/pvoc/>

<sup>3</sup>MATLAB code is available at <http://web.itu.edu.tr/ibayram/DtRadwt/>

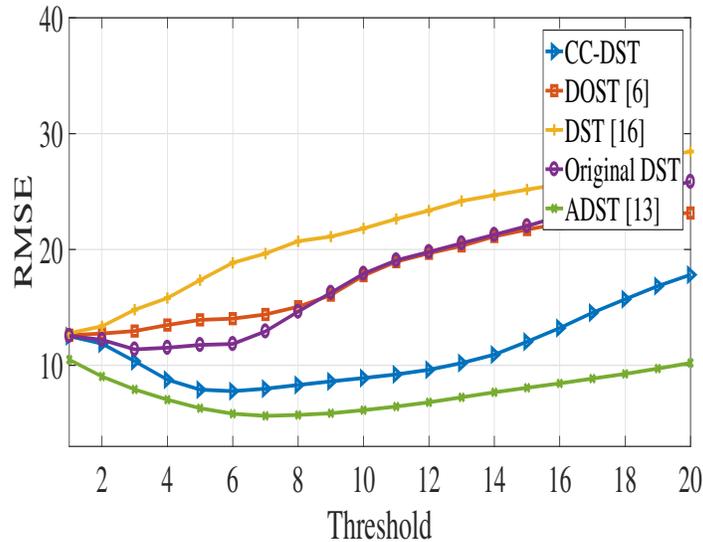


Figure 2-13: Denoising ECG signal via hard thresholding using selected DSTs. Figure shows results of RMSE with different thresholds. CC-DST ( $N_0 = 64$ ,  $\gamma = 1.1$ ,  $\alpha = 0.8$ ) is the best performing non-adaptive DST. The ADST provides better denoising performance but requires far more computation time than the CC-DST.

#### 2.4.10 Application: denoising ECG signal

Electrocardiogram (ECG) is a noninvasive and widely used technique for diagnosing cardiovascular diseases; it provides important information about the structure and function of the heart. During acquisition and transmission, ECG recordings are exposed to different sources of noise, which may hinder the correct analysis of these signals, leading to unreliable and possibly wrong diagnosis [79]. Although other methods exist, TF-based techniques have been proven highly effective for denoising ECG signals. Particularly, methods that employ the original DST were shown to outperform others based on WTs [79]. One clear limitation when using the ST, however, is the high computational requirements, which is further magnified by the fact that the ECG signals might be provided as long-duration records. To evaluate the proposed CC-DST, we perform a simple signal denoising via hard thresholding on a ten-second ECG recording (without the DC component) sampled at 250 Hz and available publicly from Physionet database<sup>4</sup>. AWGN is added to the ECG signal at SNR equal to

<sup>4</sup><http://physionet.org>

10 dB to resemble poor channel conditions, then the noisy signal is analyzed with the CC-DST. TF coefficients with amplitudes that do not exceed a predefined threshold are set to zero, then an enhanced version of the ECG signal is obtained by inverting the modified transform. The parameters used with the CC-DST are  $N_0 = 64$ ,  $\beta = 1.1$ , and  $\alpha = 0.8$ , which have been chosen based on visual evaluation of the resulting ECG signal. For the sake of comparison, the previous steps are repeated using the DOST of Stockwell, the DST of Brown *et al.*, the original  $N \times N$  DST, and the EC-based ADST presented in [1]. The denoising performance is quantitatively evaluated based on the root MSE (RMSE) between the noise-free signal and the enhanced one. Using 100 noise realizations and various values of the threshold, we obtain the curves in Fig 2-13. The results demonstrate that the CC-DST outperforms all the non-adaptive DSTs in this application, at least with the selected set of parameters, but it is slightly surpassed by the ADST. However, the computation speed of the CC-DST is a lot faster than that of the ADST, hence the CC-DST might be preferred in practice. Even though the original DST provides far more redundant TF information compared with the employed CC-DST, requiring much more computational resources, it is less effective for this application. The noisy ECG signal and the enhanced one using the CC-DST with a threshold equal to 6 are shown in Fig. 2-14.

### 2.4.11 Application: visual representation

In several applications, a TF transform is computed to obtain a visual 2-D representation (or image) of a signal of interest in the TF domain. In medical signal processing applications, for example, TFRs are frequently used for signal classification through either visual analysis, or, alternatively, this task can be automated using a machine learning approach [23, 80]. The CC-DST, likewise the DOST, is a discrete non-uniformly sampled TF transform, meaning that it cannot be represented in a matrix form, hence not directly comparable with the original or other uniformly sampled DSTs on a pixel-by-pixel basis. Nonetheless, an  $N \times N$  point image (or TFR) can be formed using the CC-DST through interpolation over time and frequency. As previously explained, a voice of the CC-DST can be interpolated in time by a FFT-based

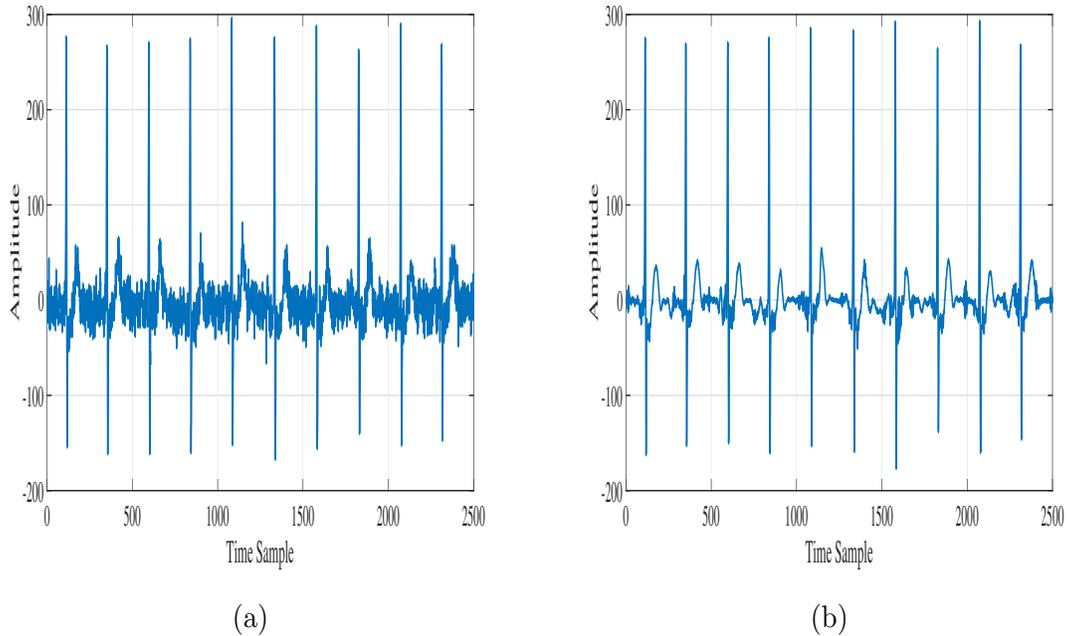


Figure 2-14: Denoising ECG signal contaminated by AWGN at SNR=10 dB using the CC-DST and hard thresholding. Noisy ECG signal (without DC component) in (a) and the enhanced signal in (b).

method: zero padding the voice spectrum, then transforming the result back to the time domain to produce  $N$  equally spaced points. Thence, the resulting matrix, of size  $2J \times N$ , can be interpolated in frequency, forming an  $N \times N$  point TFR, using a linear or nonlinear interpolation method, depending on the available computational and storage resources. We remark that to construct a visual TFR, only the amplitudes of the CC-DST samples are used for interpolation, while the phase samples are discarded to reduce the number of computations. It should be noted that the continuous 2-D approximation obtained through interpolation does not contain any extra information over the CC-DST. An example of constructing a TFR using the CC-DST is illustrated using the ECG signal used in the previous application in Fig. 2-15.

Let us now examine selected DSTs in terms of readability of their TFRs. We will use in this example a real-world bat echolocation signal of size 400 samples taken from [15], and compare the CC-DST with the original DST, the DSTs of Stockwell

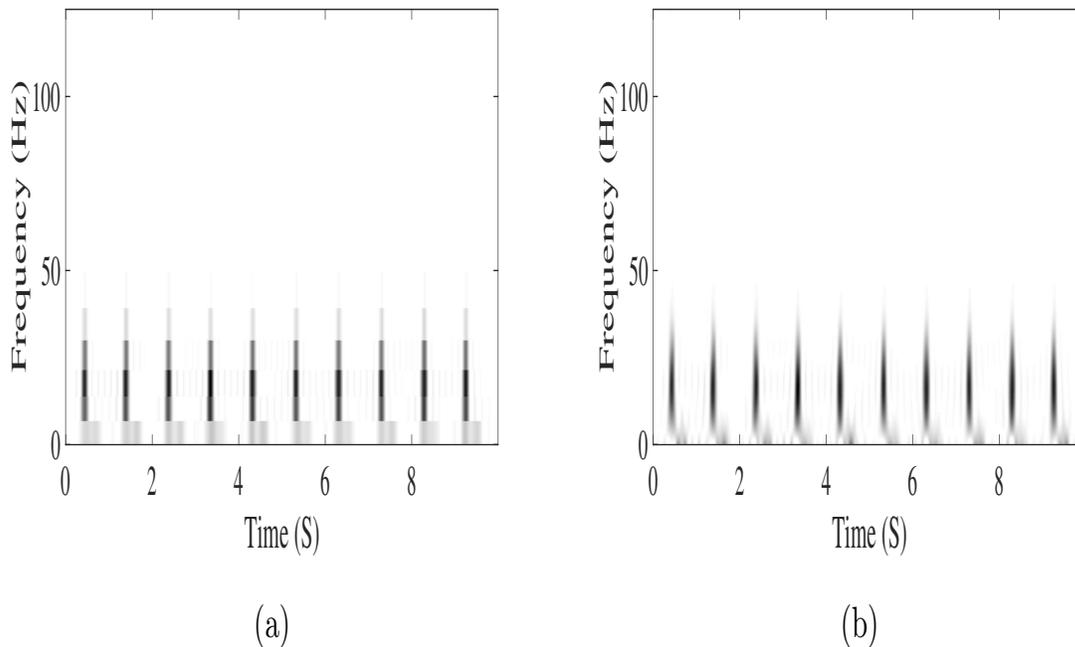


Figure 2-15: TFRs of ECG signal using the CC-DST. After interpolating the CC-DST voices in time, the resulting matrix is interpolated in frequency to create  $N \times N$  TFR. Interpolation using the nearest neighbor method in (a) and spline in (b).

and Brown *et al.*, and the ADSTs developed in [1, 2]. Observing the TFRs in Fig. 2-16, it is plain to see that the original DST provides a TFR with smeared and poorly localized energy. The DSTs of Stockwell and Brown *et al.* are very poor: the signal components are completely blurred, and no useful information can be extracted from the TFRs. The proposed CC-DST, with proper parameter settings, was able to achieve a very good TFR, clearly outperforming the original DST. As for the interpolation methods used with the CC-DST, spline results in a smoother TFR compared with the nearest neighbor method but, of course, at the expense of heavier computational requirements. The ADSTs provided sharper TFRs with impressive EC. These algorithms, however, vary the observation window of the DST iteratively for each frequency to select the most suitable window, making them of prohibitive computational complexity for applications that involve processing large datasets.

As a second example, we consider a multicomponent synthetic signal with fast frequency variation and crossing components. The signal is composed of two compo-

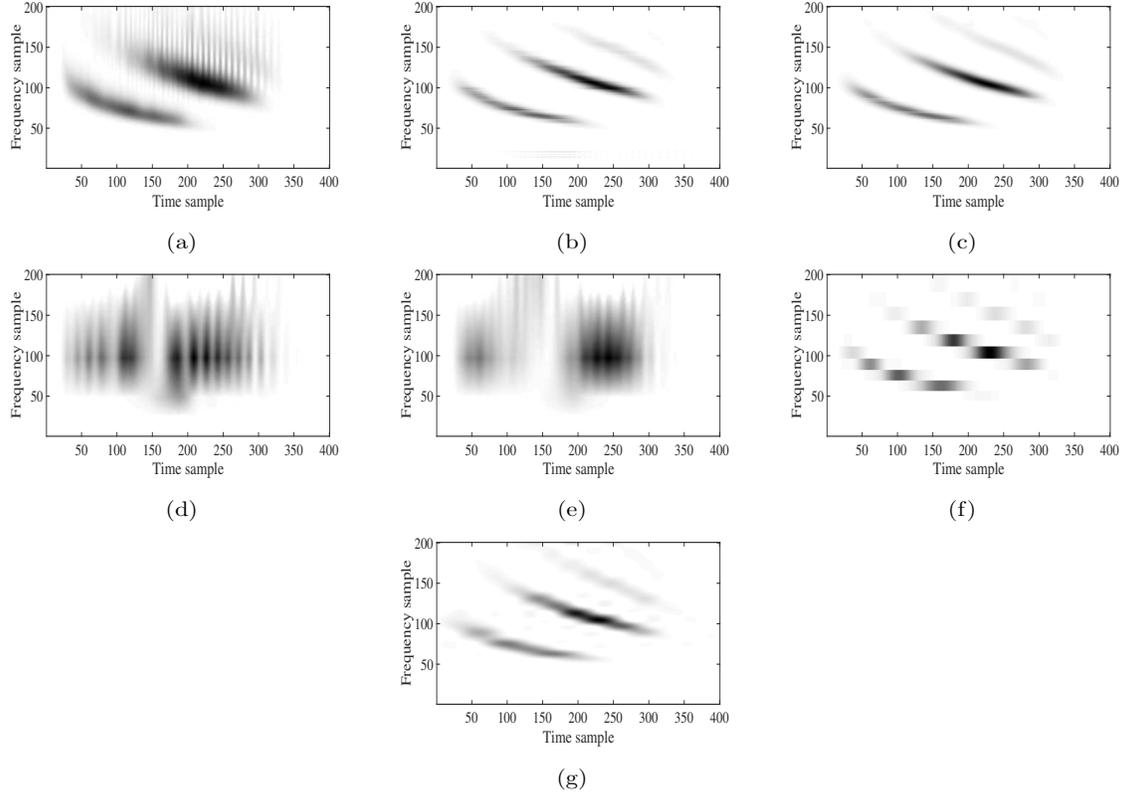


Figure 2-16: TFRs of a real-world bat echolocation sound signal using: (a) the original DST, (b) ADST [1], (c) ADST [2], (d) DOST [3], (e) DST of Brown *et al.* [4], (f) the CC-DST ( $N_0 = 8$ ,  $\gamma = 1.1$ ,  $\alpha = 0.8$ ) with the nearest neighbor interpolation method, and (g) the CC-DST with spline interpolation. All the TFRs are given with the same MATLAB colormap. The CC-DST is superior to all the non-adaptive DSTs. Parameters of the CC-DST were selected based on visual inspection.

nents: a linear chirp and a hyperbolic component, defined as

$$x(t) = \cos(75t + 20\pi \log(50t + 1)) + \cos(100\pi t + 200\pi t^2). \quad (2.80)$$

The above signal is sampled between  $0 \leq t < 1$  with a sampling frequency equal to 1024 Hz. The signal is analyzed using the DSTs used in the previous example, and the obtained TFRs are displayed in Fig. 2-17. We see from the figure that the original DST concentrates high energy along the hyperbolic component but with notably decreased resolution for the linear chirp, hence failing to compromise fairly between the signal components. The ADSTs, namely the DSTs of [1] and [2], improves the EC at middle and high frequencies, but with poorer resolution at low frequencies.

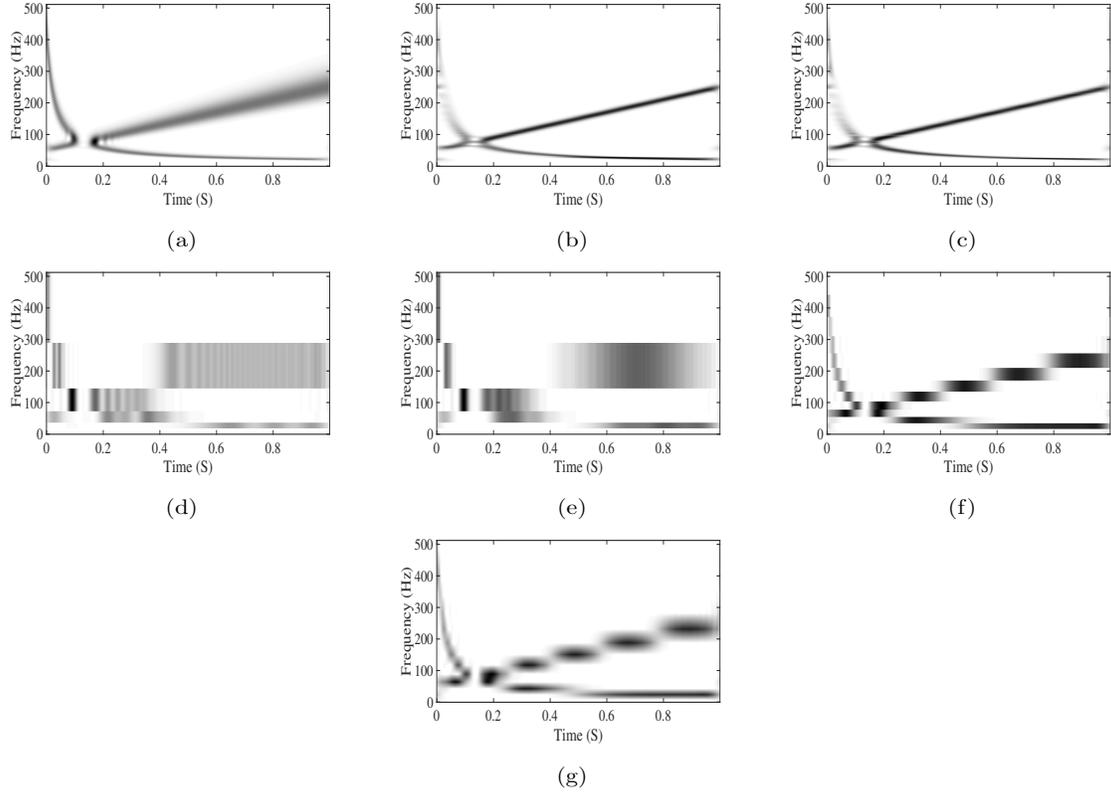


Figure 2-17: TFRs of a synthetic signal with crossing IF components using: (a) original DST, (b) ADST [1], (c) ADST [2], (d) DOST [3], (e) DST of Brown *et al.* [4], (f) the CC-DST ( $N_0 = 16$ ,  $\gamma = 1.15$ ,  $\alpha = 0.8$ ) with the nearest neighbor interpolation method, and (g) the CC-DST with spline interpolation algorithm. All the TFRs are given with the same MATLAB colormap. Parameters of the CC-DST were selected based on visual inspection. The TFRs obtained with the CC-DST feature improved EC compared with those based on the DOST and the DST of Brown *et al.*.

As for the non-uniformly sampled DSTs, the proposed CC-DST provides the best TFR in terms of EC; the IF components are not blurred, rather easily distinguishable. Furthermore, the TF resolution at low frequencies is somehow better compared with the ADSTs. The previous two examples demonstrate visually that the CC-DST significantly improves the quality of the signal TFR compared with other forms of efficiently sampled STs. This result is attributed to the CC-DST parameters, which allow for controlling the CC-DST coverage of the TF domain. These examples also show that the more ample sampling of the TF domain provided by the CC-DST compared with orthogonal DSTs is desirable in many situations as it allows one to build up a smoother TFR with improved readability, without exaggerative computa-

tional requirements. Of course, it should not be expected that the CC-DST, or other forms of efficiently sampled STs, would render TFRs smoother than those obtained with a *fully redundant*  $N \times N$  DST, such as optimized DSTs based on EC measures. These TF transforms, however, are effective when processing short-duration signals, but for large datasets, the computational efficiency of the CC-DST becomes of higher priority.

## Chapter 3

# Adaptive methods for instantaneous frequency estimation

## 3.1 Background theory

### 3.1.1 Analytic signals

A signal  $z(t)$  is said to be analytic if it contains no negative frequencies, hence satisfying the following condition:

$$Z(f) = 0 \quad \text{if } f < 0, \quad (3.1)$$

where  $Z(f)$  is the Fourier transform (FT) of  $z(t)$ . Note that a real signal does not satisfy the previous condition, because it is well known that for a real signal  $s(t)$ , we have that

$$S(f) = S^*(-f). \quad (3.2)$$

The analytic associate of a real signal can be constructed as

$$z(t) = s(t) + is_0(t) \quad (3.3)$$

in which  $s(t)$  and  $s_0(t)$  are real signals whose FTs  $S(f)$  and  $S_0(f)$ , respectively, satisfy

$$S_0(f) = (-i \operatorname{sgn}(f))S(f), \quad (3.4)$$

where  $\operatorname{sgn}(\cdot)$  is the signum function. Now, we can think of a linear operator that takes the real signal  $s(t)$  as an input to produce at the output the signal  $s_0(t)$ ; this operator is the Hilbert transform. That is, the Hilbert transform of a real signal  $s(t)$  whose FT is  $S(f)$  is defined as

$$\mathcal{H}\{s(t)\} = \text{IFT}\{(-i \operatorname{sgn}(f))S(f)\}, \quad (3.5)$$

where  $\text{IFT}\{\cdot\}$  is the inverse FT (IFT) operator. The analytic signal given in (3.3) can be rewritten as

$$z(t) = s(t) + i\mathcal{H}\{s(t)\}. \quad (3.6)$$

It can be shown that for a signal defined as

$$s(t) = A(t) \cos(\phi(t)) \quad (3.7)$$

whose real-valued instantaneous amplitude  $A(t)$  varies sufficiently slower compared to its instantaneous phase  $\phi(t)$  such that there is no overlap between the FTs of  $A(t)$  and  $\cos(\phi(t))$ , we have that

$$\mathcal{H}\{s(t)\} = A(t) \sin(\phi(t)), \quad (3.8)$$

which implies that the analytic associate of  $s(t)$  under the previous condition is given as

$$z(t) = A(t) \exp(i\phi(t)). \quad (3.9)$$

### 3.1.2 Bandwidth-duration product

In practice, all records of observations are finite, meaning that the spectral analysis ever done on measured data is always restricted to the time duration of observation. Also, signals are processed with devices of finite bandwidths. As well known from the Fourier analysis, the bandwidth of a time-limited signal is infinite. On the other hand, signals whose bandwidths are finite have infinite durations. For a signal  $s(t)$  whose FT is denoted by  $S(f)$ , we may define an effective bandwidth  $B_{\text{eff}}$  and an effective time duration  $T_{\text{eff}}$  according to

$$B_{\text{eff}} = \left[ \frac{1}{E_s} \int_{-\infty}^{\infty} f^2 |S(f)|^2 df \right]^{\frac{1}{2}}, \quad T_{\text{eff}} = \left[ \frac{1}{E_s} \int_{-\infty}^{\infty} t^2 |s(t)|^2 dt \right]^{\frac{1}{2}}, \quad (3.10)$$

where  $E_s$  is the total energy of the signal, given by

$$E_s = \int_{-\infty}^{\infty} |S(f)|^2 df. \quad (3.11)$$

Note that  $B_{\text{eff}}^2$  is the second moment of  $|S(f)|^2$  with respect to frequency, and, similarly,  $T_{\text{eff}}^2$  is the second moment of  $|s(t)|^2$  with respect to time. The definitions  $B_{\text{eff}}$

and  $T_{\text{eff}}$  may be used to obtain finite bandwidth and duration. Now, the bandwidth-duration product  $B_{\text{eff}}T_{\text{eff}}$  can be used to measure the temporal and spectral spreads of the signal. As noted by Boashash [15], this product can be thought of as a representative of the number of samples at the sampling rate  $B_{\text{eff}}$  from which the signal can be reconstructed, and thus  $B_{\text{eff}}T_{\text{eff}}$  is a measure of the *information richness* of the signal. It can be proved that any signal satisfies the following relation:

$$B_{\text{eff}}T_{\text{eff}} \geq \frac{1}{4\pi}, \quad (3.12)$$

where the equality holds only for the case of a Gaussian pulse.

### 3.1.3 The instantaneous frequency

Most of the physical signals have spectra that evolve rapidly with time. Therefore, analyzing and characterizing nonstationary signals in the time or frequency domain alone is inadequate, and the signal analyst would naturally seek representing such signals in the 2-dimensional (2-D) time-frequency (TF) plane. The fact that the spectra of nonstationary signals change with time implicates that for the concept of frequency to be meaningful, it should also include time description, thus be defined with respect to time. To clarify this point, consider a complex-valued sinusoid with a constant amplitude  $A$ , frequency  $f$ , and phase shift  $\phi_0$  described as:

$$s(t) = A \exp [i(2\pi ft + \phi_0)]. \quad (3.13)$$

Since its frequency is constant, for the signal  $s(t)$ , the concept of frequency can be completely described by the FT: the FT of  $s(t)$  at the frequency  $f$  returns the amplitude  $A$  and the phase constant  $\phi_0$ . The phase function of  $s(t)$  is linear in time

$$\phi(t) = 2\pi ft + \phi_0, \quad (3.14)$$

hence, its derivative with respect to time is constant and can be used to define the signal frequency. Accordingly, the frequency  $f$  can be interpreted as the constant

rate of change of the phase function in (3.14), and the signal  $s(t)$  can be thought of as a rotating vector in the complex plane making  $f$  cycles per second. Now, our understanding of the frequency as the rate of change of the phase can be generalized to signals whose phase law is an arbitrary differentiable function of time, but in this case, we should define the signal frequency as the instantaneous rate of change of the signal phase. That is, for a signal defined as

$$z(t) = A \exp(i\phi(t)), \quad (3.15)$$

where  $\phi(t)$  represents the instantaneous phase, the previous interpretation of the signal frequency naturally suggests a definition of the *instantaneous frequency* (IF) according to:

$$f_i(t) = \frac{1}{2\pi} \frac{d\phi}{dt}(t). \quad (3.16)$$

The signal  $z(t)$  can be considered as the analytic associate of the real-valued signal  $s(t)$ , whose amplitude changes significantly slower compared to its instantaneous phase. At a given time instant  $t_0$ ,  $f_i(t_0)$  can be considered as the frequency of a sinusoid defined by  $A \exp(i2\pi f_i(t_0)t)$  that fits locally the signal  $z(t)$ . It should be emphasized that defining the IF according to (3.16) is meaningful only for the case of a monocomponent signal. That is, if the signal is composed of different components, each of which is given as in (3.15) and has its own IF, then (3.16) results in a single value that generally does not correspond to any of the IFs of the signal components.

## 3.2 Adaptive IF estimation

### 3.2.1 Adaptive design of TFRs

The IF is a fundamental concept in TF analysis. A simple and very well-known method for IF estimation is based on the maxima position of a TF representation (TFR) [13], [14]. Developing TFRs that reveal accurately the signal IF law often requires a signal-dependent design, since no TFR can be optimized for all signals [15],

[20], as is the case with all fixed mappings. Therefore, numerous research focused on developing adaptive TFRs whose respective parameter(s) can be adjusted according to the signal at hand, sometimes without interference by a user. In this context, energy concentration measures (ECMs) have been widely used to define criteria for automating the selection of *good* TFRs, where the distribution of energy is highly concentrated along the signal components [43, 5, 1, 2]. The multi-view approach for adaptive TFRs is based on constructing multiple TFRs, then combining them according to a predetermined criterion. One example of this approach is the multi-view adaptive fractional spectrogram (AFS) [81]. Further, it has been observed that many signals of interest have their energy distributed along specific directions in the TF domain. Therefore, some researchers developed adaptive TFRs that account for the directions of energy concentration. This stream of research includes the multi-directional distribution (MDD), proposed in [22] for signals approximated as piecewise linear frequency modulated (LFM) chirps, and the locally adaptive directional TFD (ADTFD) [24], [82], which is based on optimized directional filtering. Improving localization of the signal components in the TF domain while at the same time attenuating the interfering cross-terms can be tackled by the reassignment method [37], which, as described in Chapter 1, reassigns each point in the TF domain to the center of gravity of a region located in its vicinity.

Since the accuracy of IF estimation is one of the most important performance criteria in TF analysis, perhaps a more rigorous approach to obtain adaptive TFRs can be built around their statistical performance as IF estimators. Evaluating the accuracy of IF estimation based on TFRs is addressed in several works [83], [65], where it has been demonstrated that the estimation accuracy depends on the width of the used observation window, and optimizing this width requires unknown information about the IF derivatives. To tackle this limitation, a method that approximates the optimal width based on the intersection of confidence intervals (ICI) rule was proposed in [84]; it requires implementation of multiple TFRs, each with a different window size, adding up to a significant computational burden. Aiming at reducing the computations required to obtain linear adaptive TFRs, a three-step adaptation procedure was

introduced in [85] and later improved by Pei and Huang in [6]. It consists of three steps: a preliminary TFR (PTFR) is first constructed and the peaks of its dominant ridges are used to obtain a coarse estimate the IF trajectory; an operator to compute the first derivative of the IF [here is called the IF rate (IFR)] is then employed; and the IFR is finally utilized to adjust the width of the observation window of a final TFR. Neither in [85] nor in [6], however, the criterion used for the optimal width selection is based on the statistical performance of the IF estimator. Namely, the authors did not consider signals contaminated by noise, and, furthermore, the optimal window size was not derived for the general case of a multicomponent signal, whose TFR represents more than one time-varying ridge. The study presented here improves on the previous works by deriving the optimal window width of linear TFRs using novel general formulas that describe the bias and the variance of the IF estimation error in noisy environments. We will arrive at these closed-form expressions in the next section by considering a general form of a TF transform that uses Gaussian windowing functions and the Fourier kernel, assuming that the standard deviation of the window is an arbitrary function of time and/or frequency. Besides being interesting from a theoretical point of view, our results will be incorporated into the previously mentioned adaptation procedure to develop two fully automated linear TFRs. The first TFR employs a time-adaptive window to minimize the sum of the mean squared errors (MSEs) of all the IF estimates at each time instant, while in the second TFR, the window is TF-adaptive, minimizing the estimation MSE at each location in the TF domain.

### 3.2.2 Accuracy analysis of IF estimation based on linear TFRs

#### Terminology

The general short Fourier transform (GSFT) encompasses a family of linear transforms that employ Gaussian observation windows to localize the oscillatory Fourier

kernel. This transform is formulated as<sup>1</sup>

$$\text{GSFT}_x(t, f) = \int_{-\infty}^{+\infty} x(\tau)w(t - \tau, \sigma(t, f)) \exp(-i2\pi f\tau) d\tau, \quad (3.17)$$

where  $x(t)$  is a signal under analysis, and  $w(t, \sigma(t, f))$  is a Gaussian window (usually of unitary area), whose standard deviation is an arbitrary function of time and frequency, denoted by  $\sigma(t, f)$ . As mentioned before, the Gaussian window is a classic choice for TF analysis because it features the minimum time-bandwidth product [15]. If the window width<sup>2</sup> is independent from time and frequency [i.e.,  $\sigma(t, f)$  is constant], (3.17) describes the STFT, whereas if this width is a function of only frequency [i.e.,  $\sigma(t, f) = \sigma(f)$ ], then the transform in (3.17) is an ST. We remark that the ST was originally presented in [10] with a specific form of the window standard deviation that is  $1/|f|$ . Various forms of  $\sigma(f)$  were later presented in the literature (e.g., [11]) to meet specific requirements. The adaptive STFT (ASTFT) is also a member of the family in (3.17), it employs a Gaussian window the standard deviation of which varies in time [i.e.,  $\sigma(t, f) = \sigma(t)$ ]. Besides the aforementioned cases, the window width may change in both time and frequency, hence making the representation adaptive over the TF domain.

## Estimation model

Let  $x(t)$  be an analytic frequency modulated (FM) signal whose amplitude  $A$  varies significantly slower than its phase  $\phi(t)$ ; this signal is given by

$$x(t) = A \exp(i\phi(t)). \quad (3.18)$$

The IF is defined as in (3.16), where we assume that  $\phi(t)$  is an arbitrary smooth differentiable function of time with bounded derivatives, as in [83], [65]. The GSFT

---

<sup>1</sup>We wish to mention that the formalism in (3.17) was firstly introduced in [4] under the name of “general Fourier family”. We chose to slightly modify this name to emphasize that the interest here is in Fourier-like transforms in which  $\sigma(t, f)$  is much shorter than the signal duration.

<sup>2</sup>The term “width” is used loosely to refer to the standard deviation of the Gaussian observation window.

of  $x(t)$  is given by

$$\text{GSFT}_x(t, f) = A \int_{-\infty}^{+\infty} w(t - \tau, \sigma(t, f)) \exp(i\theta_f(\tau)), \quad (3.19)$$

where  $\theta_f(\tau) = \phi(\tau) - 2\pi f\tau$ , and the Gaussian window is defined as

$$w(t, \sigma(t, f)) = \frac{1}{\sqrt{2\pi}\sigma(t, f)} \exp\left(\frac{-t^2}{2\sigma^2(t, f)}\right). \quad (3.20)$$

While in practice we deal with discrete signals, continuous transforms are used in this analysis to obtain easily interpreted closed-form results. By expanding  $\theta_f(\tau)$  into Taylor series around a generic time instant  $t$ , neglecting the terms of order higher than second, we obtain

$$\theta_f(\tau) \approx a_0 + a_1(\tau - t) + a_2(\tau - t)^2, \quad (3.21)$$

where

$$\begin{aligned} a_0 &= \theta_f(t); \\ a_1 &= \partial\theta_f/\partial\tau(t); \\ a_2 &= 1/2 \partial^2\theta_f/\partial\tau^2(t). \end{aligned}$$

Note that (3.21) is exact and not an approximation for sinusoidal and LFM signals, whereas it contains an approximation error when the phase law is of higher order. After substituting (3.20) and (3.21) into (3.19), the resultant integral can be resolved giving

$$\text{GSFT}_x(t, f) = \frac{A \exp\left[ia_0 - \frac{ia_1^2\sigma^2(t, f)}{2i+4a_2^2\sigma^2(t, f)}\right]}{\sqrt{1 - 2ia_2\sigma^2(t, f)}}. \quad (3.22)$$

The same integral was evaluated in [6] for the specific case of LFM signal. It can be easily shown that the result presented therein is a special case of the more general expression in (3.22). We consider the signal as contaminated by complex-valued AWGN, denoted by  $n_0(t)$ , with i.i.d. real and imaginary parts. Due to linearity, the

GSFT of a noisy signal  $x_n(t) = x(t) + n_0(t)$  is given by:

$$\text{GSFT}_{x_n}(t, f) = \text{GSFT}_x(t, f) + \text{GSFT}_{n_0}(t, f). \quad (3.23)$$

At each time instant, the IF estimate coincides with the frequency at which a TFR concentrates most of the energy according to [13], [83], [65]:

$$\hat{f}_i(t) = \arg \max_f |\text{GSFT}_{x_n}(t, f)|^2. \quad (3.24)$$

Equivalently,  $\hat{f}_i(t)$  is the frequency corresponding to

$$\frac{\partial |\text{GSFT}_{x_n}|^2}{\partial f} (t, \hat{f}_i(t)) = 0. \quad (3.25)$$

$|\text{GSFT}_{x_n}(t, f)|^2$  can be written as a summation of two terms:  $S_x(t, f) + S_n(t, f)$  in which  $S_x(t, f)$  represents the noiseless part of  $|\text{GSFT}_{x_n}(t, f)|^2$  and  $S_n(t, f)$  is the part affected by noise according to

$$\begin{aligned} S_n(t, f) = & |\text{GSFT}_{n_0}(t, f)|^2 + 2\Re(\text{GSFT}_x(t, f)) \Re(\text{GSFT}_{n_0}(t, f)) + 2\Im(\text{GSFT}_x(t, f)) \\ & \times \Im(\text{GSFT}_{n_0}(t, f)) \end{aligned} \quad (3.26)$$

with  $\Re(\cdot)$  and  $\Im(\cdot)$  representing the real and imaginary parts, respectively. When the useful signal dominates noise in terms of power, we can expand  $|\text{GSFT}_{x_n}(t, f)|^2$  at a generic time instant  $t$  into Taylor series up to the second-order term around the true IF to obtain

$$\begin{aligned} |\text{GSFT}_{x_n}(t, f)|^2 = & |\text{GSFT}_{x_n}(t, f_i(t))|^2 + (f - f_i(t)) \frac{\partial |\text{GSFT}_{x_n}|^2}{\partial f} (t, f_i(t)) \\ & + \frac{1}{2} (f - f_i(t))^2 \frac{\partial^2 |\text{GSFT}_{x_n}|^2}{\partial f^2} (t, f_i(t)). \end{aligned} \quad (3.27)$$

By taking the first-order frequency derivative of the right- and left-hand sides of (3.27), and after recalling (3.23) and (3.25), we arrive at

$$\hat{f}_i(t) = f_i(t) - \frac{\frac{\partial S_x}{\partial f}(t, f_i(t)) + \frac{\partial S_n}{\partial f}(t, f_i(t))}{\frac{\partial^2(S_x + S_n)}{\partial f^2}(t, f_i(t))}. \quad (3.28)$$

Assuming that  $S_x(t, f_i(t)) \gg S_n(t, f_i(t))$ , we can write (3.28) at moderate to high signal-to-noise ratios (SNRs) as

$$\hat{f}_i(t) = f_i(t) - \frac{\frac{\partial S_x}{\partial f}(t, f_i(t)) + \frac{\partial S_n}{\partial f}(t, f_i(t))}{\frac{\partial^2 S_x}{\partial f^2}(t, f_i(t))}. \quad (3.29)$$

Consequently, the estimate  $\hat{f}_i(t)$  contains the following error:

$$\xi(t, f_i(t)) = \frac{\frac{\partial S_x}{\partial f}(t, f_i(t)) + \frac{\partial S_n}{\partial f}(t, f_i(t))}{\frac{\partial^2 S_x}{\partial f^2}(t, f_i(t))}. \quad (3.30)$$

### Estimation bias

The estimation bias is by definition the expected value of the estimation error  $\xi(t, f_i(t))$ , which is equal to

$$\bar{\xi}(t, f_i(t)) = \frac{\frac{\partial S_x}{\partial f}(t, f_i(t)) + E \left\{ \frac{\partial S_n}{\partial f}(t, f_i(t)) \right\}}{\frac{\partial^2 S_x}{\partial f^2}(t, f_i(t))}. \quad (3.31)$$

**Lemma 1.** *For a signal defined as in (3.18) and contaminated by AWGN, whose double-sided power spectral density is denoted by  $N_0$ , we have the following identity:*

$$E \left[ \frac{\partial S_n}{\partial f}(t, f_i(t)) \right] = -\frac{N_0 \frac{\partial \sigma}{\partial f}(t, f_i(t))}{2\sqrt{\pi}\sigma^2(t, f_i(t))}. \quad (3.32)$$

Proof of Lemma 1 is provided in Appendix A.

**Proposition 1.** *The bias of the IF estimation error based on a transform of the form*

in (3.17) is given by

$$\bar{\xi}(t, f_i(t)) = \frac{k_1^2 \sigma'_i (k_1^3 / C_0 + 4\sqrt{\pi} k_2 \sigma_i)}{4\sqrt{\pi} (4\pi^2 k_1^2 \sigma_i^4 - 3k_3 \sigma_i'^2 + k_1^2 k_2 \sigma_i \sigma_i'')}, \quad (3.33)$$

where

$$k_1 = \sqrt{1 + (2\pi f'_i(t) \sigma_i^2)^2};$$

$$k_2 = k_1^2 - 1;$$

$$k_3 = k_1^4 - 3k_1^2 + 2;$$

$$f'_i(t) = \partial f_i / \partial t(t);$$

$$\sigma_i = \sigma(t, f_i(t));$$

$$\sigma'_i = \partial \sigma / \partial f(t, f_i(t));$$

$$\sigma''_i = \partial^2 \sigma / \partial f^2(t, f_i(t));$$

$C_0$  is the signal-to-noise-spectral-density ratio:  $A^2/N_0$ .

*Proof.* From (3.22), the squared magnitude of  $\text{GSFT}_x(t, f)$  (i.e.,  $S_x(t, f)$ ) can be given by

$$S_x(t, f) = \frac{A^2 \exp \left[ \frac{-a_1^2 \sigma^2(t, f)}{1 + (2a_2 \sigma^2(t, f))^2} \right]}{\sqrt{1 + (2a_2 \sigma^2(t, f))^2}}. \quad (3.34)$$

In (3.31), the first and second derivatives of  $S_x(t, f)$  with respect to frequency at  $f = f_i(t)$  can be computed using (3.34), by substituting their values therein, together with the result of Lemma 1, we prove Proposition 1.  $\square$

It is plain to see that the bias depends on the first derivative of the IF (i.e, the IFR), and also on the values of the first- and second-order frequency derivatives of  $\sigma(t, f)$  at the IF.

**Remark.** Any transform of the form in (3.17) with a frequency-dependent window provides a biased IF estimate, unless it is designed such that  $\sigma'_i = 0$ .

This remark follows directly from (3.33). Consequently, while the STFT and the ASTFT provide unbiased IF estimation, this fact does not hold for the ST whose IF estimate is essentially biased.

## Estimation variance

The estimation variance is defined as

$$\sigma_\xi^2(t, f_i(t)) = E \left[ (\xi(t, f_i(t)) - \bar{\xi}(t, f_i(t)))^2 \right]. \quad (3.35)$$

Using (3.29) and (3.31), the variance at a moderate to high SNR can be written as

$$\sigma_\xi^2(f_i) = \frac{E \left[ \left( \frac{\partial S_n}{\partial f}(t, f_i(t)) \right)^2 \right] - E^2 \left[ \frac{\partial S_n}{\partial f}(t, f_i(t)) \right]}{\left( \frac{\partial^2 S_n}{\partial f^2}(t, f_i(t)) \right)^2}. \quad (3.36)$$

**Lemma 2.** *For a signal defined as in (3.18) and contaminated by complex-valued AWGN, we have the following identity:*

$$E \left[ \left( \frac{\partial S_n}{\partial f}(t, f_i(t)) \right)^2 \right] = \frac{8\pi^2 k_1^2 \sigma_i^4 + (11k_1^2 - 8)\sigma_i'^2}{4\sqrt{\pi} k_1^3 \sigma_i^3 C_0} + \frac{\pi}{C_0^2} + \frac{(11\sqrt{2} + 12\sqrt{\pi}\sigma_i)\sigma_i'^2}{16\pi^{\frac{3}{2}} \sigma_i^5 C_0^2}. \quad (3.37)$$

Proof of Lemma 2 is provided in Appendix B.

**Proposition 2.** *The variance of the IF estimation error based on a transform of the form in (3.17) is given by*

$$\sigma_\xi^2(t, f_i(t)) = \frac{k_1^7 \sigma_i (8\pi^2 k_1^2 \sigma_i^4 + k_4 \sigma_i'^2)}{16\sqrt{\pi} C_0 (4k_1^2 \pi^2 \sigma_i^4 + k_5 \sigma_i'^2 + k_6 \sigma_i \sigma_i'')^2}, \quad (3.38)$$

where

$$k_4 = 11k_1^2 - 8;$$

$$k_5 = -3k_2(k_2 - 1);$$

$$k_6 = k_1^2 k_2.$$

*Proof.* The second-order derivative in the denominator of (3.36) can be computed using (3.34); by substituting its value therein, together with the results of Lemma 1

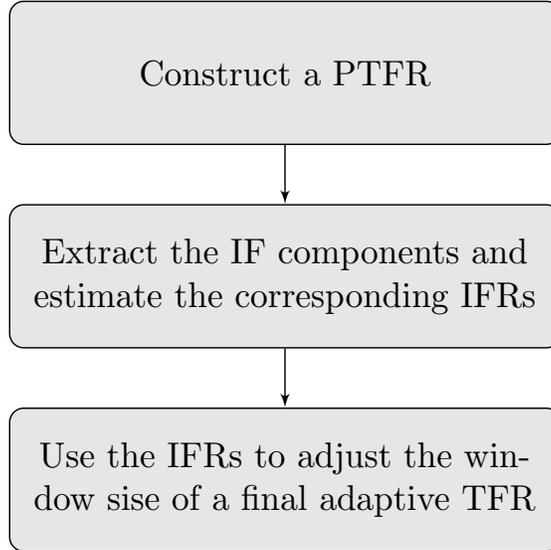


Figure 3-1: Flowchart of procedure used to construct adaptive TFRs based on linear TF transforms belonging to (3.17).

and Lemma 2, and after discarding the terms divided by  $C_0^2$  under the assumption of high SNR, we prove Proposition 2.  $\square$

As the bias, the estimation variance is signal-dependent, because it is a function of the IF derivatives. It is worth noting that if the truncated Taylor series incorporates more terms, then higher-order time derivatives of the IF and higher-order frequency derivatives of  $\sigma(t, f)$  enter the expressions of the bias and the variance, unless the analyzed signal is a sinusoid or a LFM chirp.

**Corollary 1.** *The IF estimate obtained through a transform of the form in (3.17) with a frequency-independent window attains the least MSE if*

$$\sigma(t) = \frac{\left(\frac{3}{7}\right)^{\frac{1}{4}}}{\sqrt{2\pi |f'_i(t)|}}. \quad (3.39)$$

*Proof.* When the window is fixed in frequency, the IF estimate is unbiased (see Remark), and thus the estimation MSE, defined as the sum of the squared bias and the variance, is identical to the variance in (3.38). We replace  $\sigma'_i$  and  $\sigma''_i$  in (3.38) with zeros, and  $\sigma_i$  with  $\sigma_i(t)$  to emphasize that the standard deviation is constant over frequency and changes only in time. By taking the first derivative of the resultant

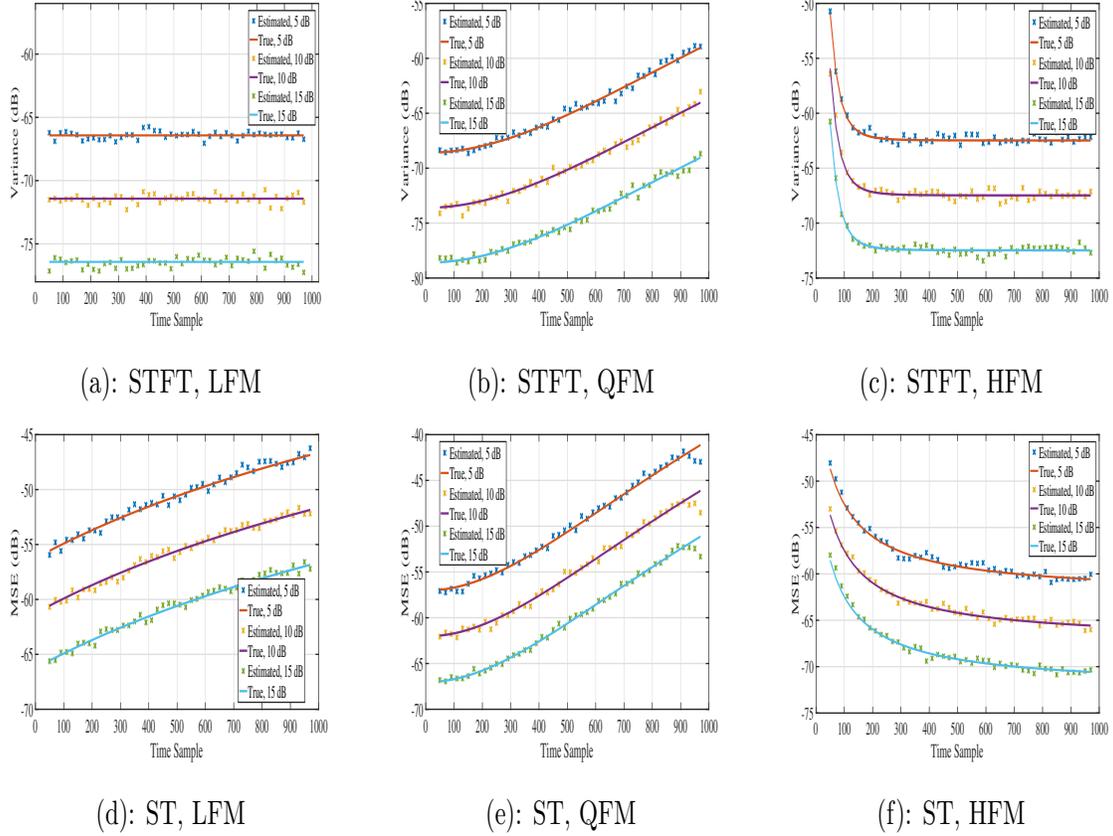


Figure 3-2: Theoretical and experimental performance of IF estimators based on STFT and ST using FM signals belonging to three different classes. In upper subplots, STFT with  $\sigma$  equal to 16, 16, 10, respectively, is used. In the last subplots, we use ST with  $\sigma(f)$  equal to  $1/(0.1 + 0.6|f|)$ ,  $1/(0.1 + |f|)$ ,  $1/(0.1 + |f|)$ , respectively.

equation with respect to  $\sigma(t)$ , we obtain

$$\frac{\partial \sigma_{\xi}^2}{\partial \sigma(t)}(t, f_i(t)) = \frac{k_1^3 [7(2\pi f_i'(t)\sigma^2(t))^2 - 3]}{32\sigma^4(t)\pi^{\frac{5}{2}} C_0}. \quad (3.40)$$

Also, it can be easily shown that

$$\frac{\partial^2 \sigma_{\xi}^2}{\partial \sigma^2(t)}(t, f_i(t)) > 0 \quad \forall \sigma(t) > 0. \quad (3.41)$$

Therefore, to find the value of  $\sigma(t)$  at which the estimation variance is minimum, we put (3.40) to zero and solve in  $\sigma(t)$ ; one positive real solution can be found, which is given by (3.39). □

Deriving the optimal window width of linear TF transforms with frequency-independent windows in absence of noise is straightforward. Indeed, for such transforms in a noiseless environment, it is immediate to see that  $S_x(t, f)$  in (3.34) is a Gaussian function centered around the IF. Hence, the optimal choice of  $\sigma(t)$  shall minimize the variance of  $S_x(t, f)$  to maximize the energy concentration around the IF. The expression of the standard deviation in this case is slightly different than that given by (3.39) and can be readily found to be

$$\sigma(t) = \frac{1}{\sqrt{2\pi |f'_i(t)|}}. \quad (3.42)$$

In [6], the above relation was evaluated explicitly for noiseless LFM signals.

To confirm the validity of the previously derived expressions, we run Monte-Carlo simulations to estimate the IFs of FM signals belonging to 3 different classes contaminated by AWGN using the STFT and the ST. The signals used in the experiments are: a LFM chirp with the following IF law:  $f_{iL}(t) = 0.1 + (0.3/1024)t$ , a quadratic FM (QFM) signal with the following IF law:  $f_{iQ}(t) = 0.05 + (0.4/1024^2)t^2$ , and a hyperbolic FM (HFM) signal whose IF law is given by:  $f_{iH}(t) = 459/[32(t + 255/8)]$ . The results of the numerical analysis along with the theoretical values are depicted in Fig. 3-2. The upper three subplots depict the numerical and theoretical variance of the IF estimate obtained through the STFT, while in the last three subplots, we show the experimental and theoretical MSE of the IF estimate of the ST. The statistical data are obtained by 300 trials for 3 SNR values (5, 10, and 15 dB). A good agreement between the theoretical results and the statistical results can be observed.

### 3.2.3 Three-step method for adaptive IF estimation

The previous section introduces general formulas that relate the bias and the variance of the IF estimation error to the observation window and the smoothness of the IF itself. Therefore, accurate IF estimation requires a data-driven method to select the optimal, or at least a reasonable, window width. To this end, our work recalls and extends a low-complexity adaptation procedure presented in [85], [6]; it consists of

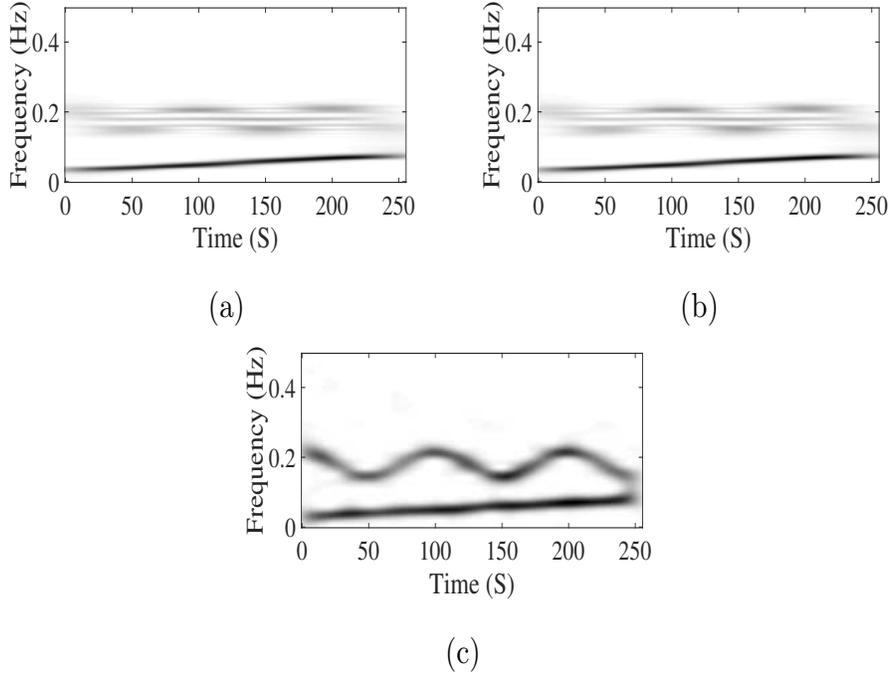


Figure 3-3: Comparison between TFRs based on ECMs with that based on the proposed regular rate. Analyzed signal is given in (3.51) and corrupted by AWGN (SNR equal to 10 dB). TFRs are based on STFT the width of its observation window is set based on (a) ECM of Stanković [5], (b) ECM of Pei and Huang [6], and (c) proposed regular rate. TFRs based on ECMs fail to concentrate enough energy along sinusoidal IF component.

three steps, discussed in this section and summarized by the flowchart in Fig. 3-1. Based on this adaptation procedure, we propose two novel linear TFRs suitable for IF estimation of noisy multicomponent signals, relying on the closed-form expressions of the estimator bias and variance derived in the previous section. The first TFR employs a time-adaptive window that minimizes the sum of the MSEs of the IF estimates at each time instant, whereas in the second TFR, the window width is TF-adaptive, minimizing the estimation MSE at each location in the TF domain.

For the sake of practicality, we shall restrict ourselves now to the discrete version of the GSFT, which we define as<sup>3</sup>

$$\text{GSFT}_x[n, m] = \sum_{l=0}^{N-1} x[l]w[n-l, \sigma[n, m]] \exp\left(-i2\pi m \frac{l}{N}\right), \quad (3.43)$$

<sup>3</sup>The convention that the sampling frequency is normalized to unity is adopted.

where  $x[n]$  denotes a discrete time series of length  $N$  corresponding to  $x(nT_s)$ , with sampling interval  $T_s$  and  $m = -N/2, \dots, N/2 - 1$  representing frequency range index.

### Preliminary TFR

Optimizing the window width of any TFR of the form in (3.43) requires knowledge about the IF itself. Therefore, the first step of the adaptation procedure is to construct a PTFR, which provides a rough estimate of the IF trajectory. While small estimation errors are tolerable at this stage [6], serious ones may propagate to the next stages, degrading the accuracy of the final IF estimate. For this reason, Pei and Huang in [6] employed a STFT whose window width is optimized for the signal at hand according to an ECM defined in [6, (16)]. Despite the good performance shown therein, after examining this ECM with different multicomponent signals, we observed that in noisy environments, the measure favors slow-varying components to the detriment of those that change faster. In other words, it tends to favor the selection of TFRs where only quasi-stationary components are extremely highly concentrated. Here, as a PTFR, we propose a STFT the standard deviation of its Gaussian observation window is set based on so-called *regular rate*. We define this rate to be the ratio between the effective bandwidth  $B_{\text{eff}}$  and the effective time duration  $T_{\text{eff}}$  of the signal under analysis according to

$$\nu_0 = \frac{B_{\text{eff}}}{T_{\text{eff}}} = \left[ \frac{1}{N} \frac{\sum_{k=-N/2}^{N/2-1} (k - k_0)^2 |X[k]|^2}{\sum_{n=0}^{N-1} (n - n_0)^2 |x[n]|^2} \right]^{\frac{1}{2}}, \quad (3.44)$$

where  $X[k]$  is the discrete FT (DFT) of  $x[n]$ , and

$$k_0 = \frac{\sum_{k=-N/2}^{N/2-1} k |X[k]|^2}{\sum_{k=-N/2}^{N/2-1} |X[k]|^2}, \quad n_0 = \frac{\sum_{n=0}^{N-1} n |x[n]|^2}{\sum_{n=0}^{N-1} |x[n]|^2}. \quad (3.45)$$

If we think of the signal normalized power spectral density (i.e., normalized to unity) as a probability mass function, then  $B_{\text{eff}}^2$  would be its second central moment with respect to frequency and  $B_{\text{eff}}$  is its standard deviation, which is a quantity indicative of the signal bandwidth [15]. We may think of  $T_{\text{eff}}$  in the same previous manner as a measure of the signal temporal width. The regular rate  $\nu_0$ , therefore, may be

interpreted as the rate of a linear chirp that sweeps a frequency range of  $B_{\text{eff}}$  during  $T_{\text{eff}}$ . We make use of the regular rate to define a *global* constant rate of change for the signal IF law, which, when inserted into (3.39), can be used to set the standard deviation of the fixed observation window of a STFT, used as a PTFR, according to  $\sigma = (3/7)^{1/4}/\sqrt{2\pi\nu_0}$ . A comparison between the proposed PTFR and TFRs optimized using the ECMs of [5] and [6] is provided in Fig. 3-3. The results show that the ECM-based TFRs fail to concentrate enough energy along a sinusoidal IF component, whereas the TFR whose window width is set based on the regular rate reveals all signal components, in addition to being more computationally efficient, because, unlike in the ECM-based approach, it does not require computation of multiple TFRs to select the optimal one.

### IFR estimation

In the second step of the adaptation procedure, we obtain a rough estimate of the IFR. To this end, the peaks of the dominant ridges of a PTFR are used to estimate the IF trajectory, then an operator that computes the first-order time derivative is employed to retrieve an estimate of the IFR. If the signal is multicomponent, an algorithm to extract separately all components is needed before computing the derivative of each extracted IF ridge. Here, IF components separation is performed through an algorithm based on image processing presented in [86]. It consists of a transformation of a TFR into a binary image followed by a process called component linking, which extracts linked crests of a TFR based on a neighborhood-connectivity criterion if their time duration is longer than a predetermined threshold. This algorithm does not require prior information about the number of components, but it assumes that the IF ridges are separable in a TFR, which is a characteristic shared by many real-world signals, including most physiological signals [87].

As for derivative computations, in [85], the conventional difference operator is used. However, it is well known that differentiating a noisy discrete signal by means of this method is highly sensitive to noise. Namely, the SNR of the differentiated signal is almost always much lower than that of the original signal. Therefore, the

Signal	SM		EMBD		Re-Spec	ADTFD		
Test Signal A	$N_w = 55$	$L = 12$	$\alpha = 0.12$	$\beta = 0.12$	$N_w = 55$	$a = 4$	$b = 9$	$F_w = 81$
Test Signal B, real-world EEG	$N_w = 85$	$L = 6$	$\alpha = 0.075$	$\beta = 0.5$	$N_w = 85$	$a = 3$	$b = 15$	$F_w = 101$
Real-world bat sound	$N_w = 65$	$L = 10$	$\alpha = 0.19$	$\beta = 0.47$	$N_w = 65$	$a = 3$	$b = 15$	$F_w = 101$

Table 3.1: Parameters used to construct TFRs of test signals.

difference operator was replaced in [6] by the principal component analysis (PCA), which performs eigenvalue decomposition on the  $2K$  estimates surrounding  $\hat{f}_i[n]$  in order to approximate  $f'_i[n]$ . Our approach here for computing the first derivative is simpler. First, the estimates  $\hat{f}_i[n]$  are smoothed according to

$$\tilde{\hat{f}}_i[n] = \frac{1}{2K+1} \sum_{n-K}^{n+K} \hat{f}_i[n], \quad n = K, \dots, N-1-K, \quad (3.46)$$

where  $K$  is a positive integer. The above smoother is a simple moving average filter, which aims at increasing the SNR of the differentiated signal [88]. After smoothing, the conventional difference operator is applied to estimate the IFR. The larger is  $K$ , the greater is the noise reduction, but also the larger is the attenuation in the derivative amplitude. Furthermore, filtering the IF estimates according to (3.46) helps remove possible outliers before the derivative is calculated, but the accuracy of this process is limited, due to existing correlation between errors at consecutive samples in the TFR [14]. Therefore,  $K$  should be chosen with caution. The smoother width is set here to 11 (i.e.,  $K = 5$ ) and three passes are used to smooth the coarse IF estimates.

### Window width optimization

Once the IF components are extracted and the respective IFRs are estimated, the window width can be optimized accordingly. At this stage, there are two possibilities for adapting the width of the TFR observation window: univariate and multivariate width optimization. In the first possibility, the width of the window may be defined as a function of time (i.e. ASTFT) or frequency (i.e., ST). This approach to width optimization boils down to searching for the width that achieves the best compromise

for all signal components at each time instant or frequency. Therefore, it requires a meaningful cost function (e.g., MSE) that is minimum if the window width is optimal. One clear advantage of this approach is its limited complexity, as both the ASTFT and the ST can be implemented efficiently by means of FFTs. On the other hand, the observation window can be set to vary in both time and frequency. This second option entails more flexibility by allowing for width adjustment at each location in the TF domain; however, it requires higher computational complexity because the GSFT in (3.43) cannot be realized by using FFTs when  $\sigma[n, m]$  is dependent on both  $n$  and  $m$ .

### Univariate width optimization

We propose an ASTFT the width of its observation window is optimized at each time instant to minimize the sum of the MSEs of the IF estimates of all signal components. Since the ASTFT yields a TFR that is generally free from cross-terms, it can be approximated by superimposing the TFRs corresponding to all the signal monocomponents (see [83] for an argument). In this context, the formulas derived before can be extended to the case of multicomponent signals whose IF ridges are separable in the TF domain. Namely, using an ASTFT, unbiased IF estimates are obtained, and the total MSE can be derived from (3.38) to give:

$$\text{MSE}_{\text{tot}}[n, \sigma[n]] = \sum_{l=1}^{L_n} \frac{\left[ (2\pi f'_{il}[n]\sigma^2[n])^2 + 1 \right]^{\frac{5}{2}}}{32\pi^{\frac{5}{2}}\sigma^3[n] C_0}, \quad (3.47)$$

where  $L_n$  is the number of signal components at the time instant  $nTs$ , and  $f'_{il}[n]$  stands for the IFR of a single IF component  $l$ . We define the optimal window width as that minimizing  $\text{MSE}_{\text{tot}}[n, \sigma[n]]$ , which may be approximated by

$$\sigma_{\text{opt}}[n] \approx \arg \min_{\sigma[n]} \sum_{l=1}^{L_n} \frac{\left[ (2\pi \hat{f}'_{il}[n]\sigma^2[n])^2 + 1 \right]^{\frac{5}{2}}}{\sigma^3[n]}. \quad (3.48)$$

Note that  $\text{MSE}_{\text{tot}}[n, \sigma[n]]$  is the sum of functions each of which is strictly convex and has a global minimum at  $\sigma_l[n]$ , which is returned by (3.39). Therefore, there also exists a global minimum to  $\text{MSE}_{\text{tot}}[n, \sigma[n]]$ , and it belongs to the following band:

$$\sigma_{\text{opt}}[n] \in \left[ \min \{ \sigma_1[n], \dots, \sigma_{L_n}[n] \}, \max \{ \sigma_1[n], \dots, \sigma_{L_n}[n] \} \right]. \quad (3.49)$$

A good approximation of  $\sigma_{\text{opt}}[n]$  can be found by executing the following two steps: first, the values of  $\sigma_l[n]$  corresponding to the minimum and maximum estimated IFRs are computed; thence, between these two limits, candidates for the optimal width are tested based on the objective function in (3.48) from a small set (a set of 10 elements is used in the numerical examples of the next section). We refer to the resultant TFR as the optimized ASTFT (OASTFT).

Optimizing the width of a frequency-dependent window within an ST to minimize the sum of the MSEs of the IF estimates requires information about the first- and second-order frequency derivatives of the window width [recall that (3.33) and (3.38) are dependent on both  $\sigma'_i$  and  $\sigma''_i$ ]. These derivatives are unknown, since the width function itself is unknown and to be designed, which makes this optimization approach impractical; for this reason, we will not proceed with it.

### **Multivariate width optimization**

The OASTFT presented before compromises meaningfully between all signal components in the TF domain, but to further improve adaptivity, the observation window could vary in both time and frequency. If the IF ridges are separable in the TF domain, the window width of the GSFT can be adjusted for each component independently. At this stage, coarse estimates of the IFs and the corresponding IFRs are available, hence the question is how to set the window standard deviation at the on- and off-ridge points. We proved before that the standard deviation given in (3.39) is optimal (in the MSE sense) in case its first- and second-order derivatives with respect to frequency are null at the IF. Therefore, for each on-ridge point of an IF component,

$i$	1	2	3	4
$b_i$	1.2	1	0.9	0.8
$c_i$	-0.0033	-0.0033	-0.009	-0.014
$d_i$	1.6	3.2	6.4	9.6

Table 3.2: Coefficients of Test Signal B given in (3.52).

we define a spectral interval  $\Delta_f$  within which the window width is fixed according to

$$\sigma[n, m] = \frac{\left(\frac{3}{7}\right)^{\frac{1}{4}}}{\sqrt{2\pi \left| \hat{f}'_{il}[n] \right|}}, \quad 0 \leq |m - \hat{f}_{il}[n]| \leq \frac{\Delta_f}{2}, \quad (3.50)$$

whereas outside this interval, the width is interpolated using 2-D linear interpolation as in [6]. The previous interpolation guarantees that the first- and second-order frequency derivatives of  $\sigma[n, m]$  are zeros around  $\hat{f}_{il}[n]$ , and thus the selected width is optimal in the MSE sense. Furthermore, due to noise, the coarse IF estimate is likely to be shifted from its true value, thus it is preferable to assign the width in (3.50) not only to  $\hat{f}_{il}[n]$  but also to the frequencies in its vicinity. Experimentally, we found that a width of few frequency bins yields an accurate IF estimate ( $\Delta_f$  is set to 8 in the numerical examples of the next section). The resultant transform is denoted by the optimized GSFT (OGSFT).

### 3.2.4 Performance evaluation

In this section, the proposed methods are evaluated and compared with popular state-of-the-art algorithms. The comparison includes the methods proposed by Pei and Huang in [6], denoted here by the chirp rate-based ASTFT (CR-ASTFT), the chirp rate-based ST (CR-ST), and the chirp rate-based GSFT (CR-GSFT), and also two of the most promising QTFRs reported in the literature: the S-method (SM) [31], and the extended modified B-distribution (EMBD) [23]. Furthermore, four adaptive TFRs are considered: the MDD [22], the AFS [81], the ADTFD [24], and an adaptive ST based on energy concentration optimization [2], denoted by EC-ST. For completeness, a reassigned spectrogram (Re-Spec) is also included in the benchmark comparison

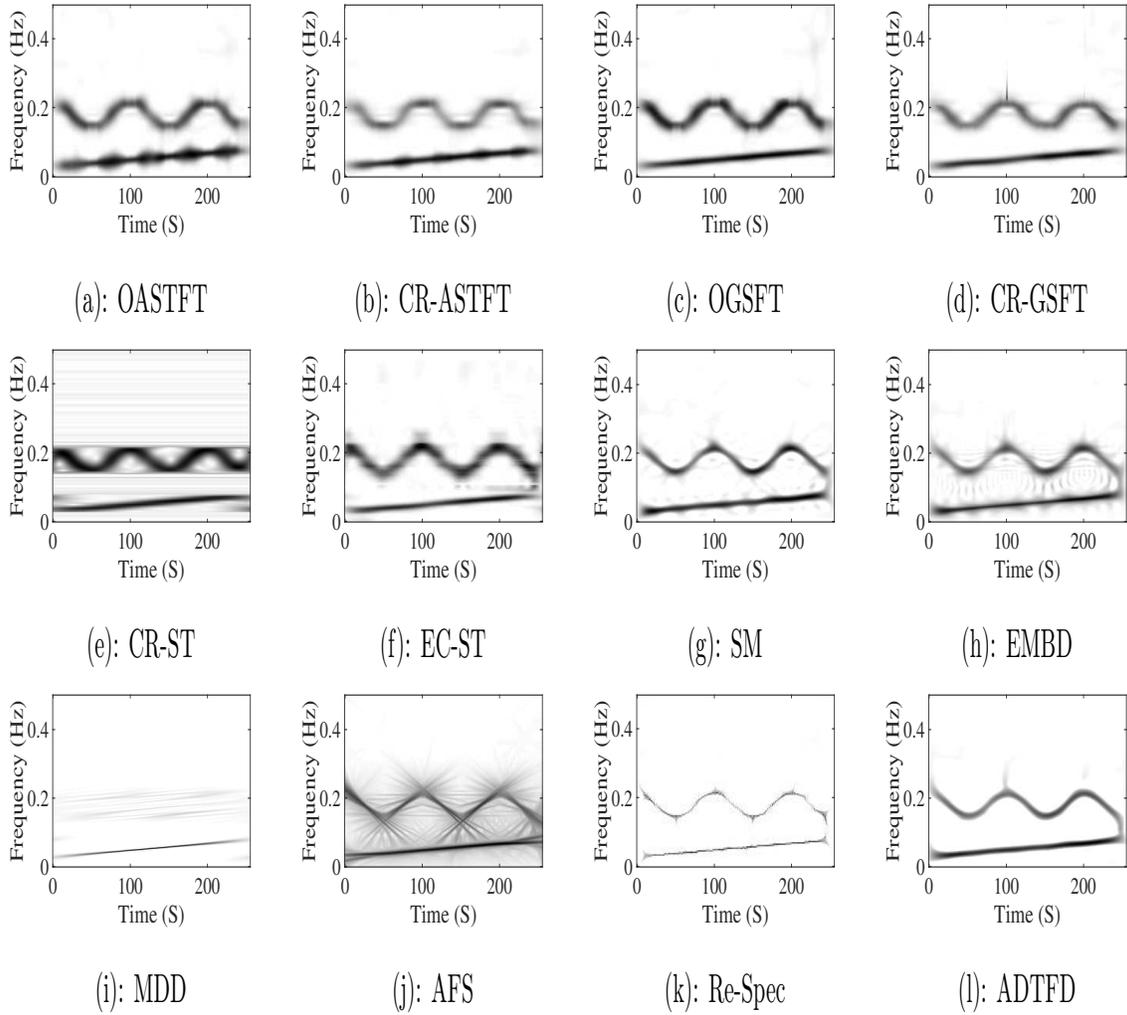


Figure 3-4: TFRs of synthetic signal composed of two components given in (3.51) and corrupted by AWGN (SNR equal to 12 dB). Respective parameters of TFRs are adjusted based on visual inspection as provided in Table 3.1.

representing the reassignment method. The algorithms of [6] require a PTFR; for the sake of a fair and meaningful comparison, they are realized here using the proposed PTFR based on the regular rate.

### Evaluation on synthetic signals

IF estimation of two multicomponent test signals is carried out in AWGN environments characterized by SNR ranging from 0 to 30 dB with a 2-dB step. The normalized MSE (NMSE) of IF estimation over a large number of Monte-Carlo simulations

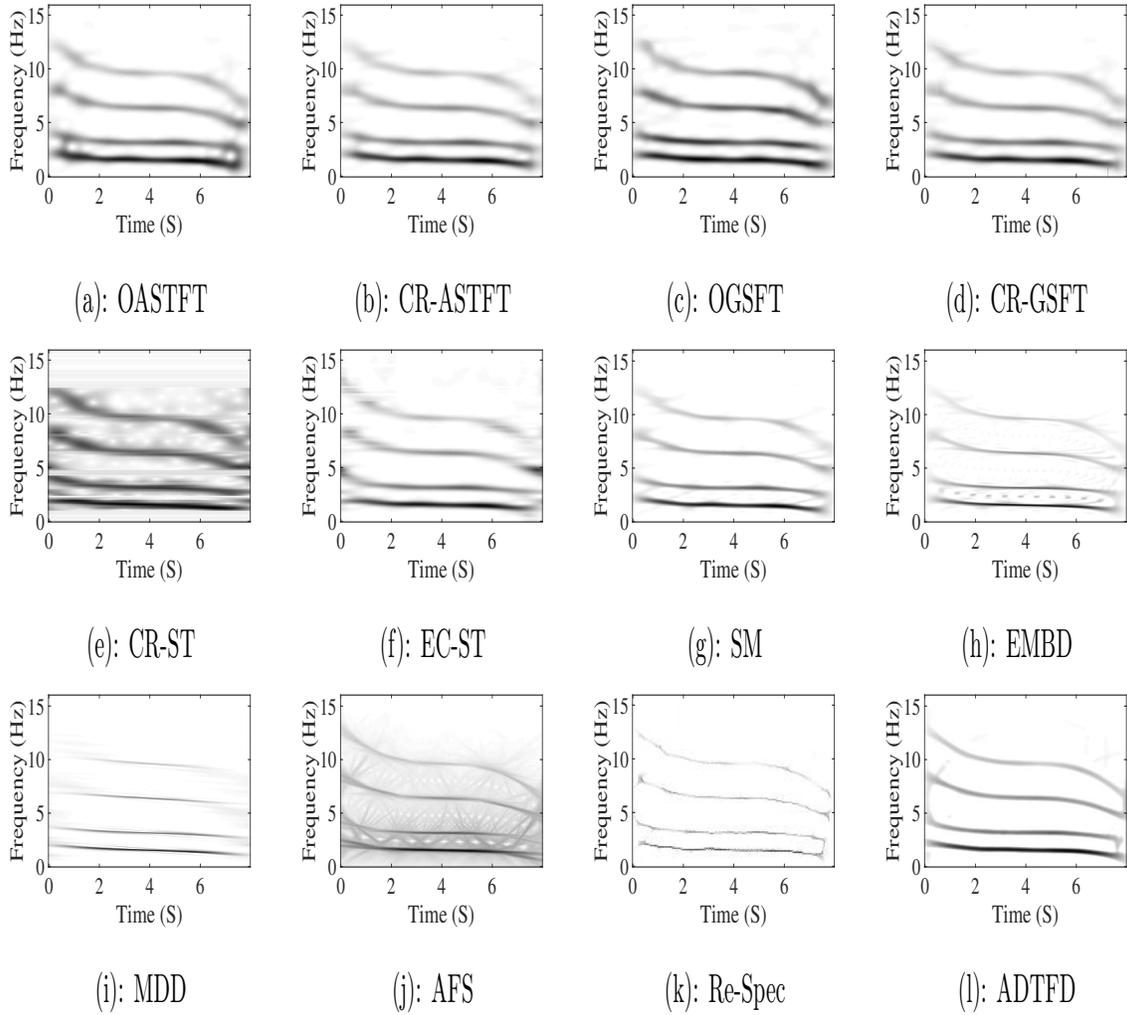


Figure 3-5: TFRs of synthetic signal of 4 components given in (3.52) corrupted by AWGN (SNR equal to 12 dB). Respective parameters of TFRs are optimized based on visual inspection as provided in Table 3.1.

(here is set to 100) is chosen to quantitatively evaluate the precision of the selected methods. As mentioned before, the employed algorithm for IF components separation and tracking is presented in [86]; in our simulation, IF components whose durations are less than 20 samples will be discarded. Moreover, to reduce the quantization error in the final IF estimate, we interpolate the cross-section of a TFR corresponding to a given time instant between the two frequency samples surrounding its peak using a parabolic function.

## Test Signal A

This synthetic signal is composed of two components: a LFM chirp, and a sinusoidal FM component; it is defined as

$$x_A(t) = \cos [2\pi t (a_1 + a_2 t)] + \cos (2\pi a_3 t + 4 \sin(2\pi a_4 t)), \quad (3.51)$$

where  $a_1 = 0.03$ ,  $a_2 = 0.025/256$ ,  $a_3 = 0.18$ ,  $a_4 = 0.01$ ,  $0 \leq t < 256$  seconds, and the sampling frequency is  $f_s = 1$  Hz. TFRs of Test Signal A with AWGN at SNR equal to 12 dB are shown in Fig. 3-4 using the same colormap. The parameters used to construct these TFRs are reported in Table 3.1; they were optimized based on visual inspection. Results of the NMSE are shown in Fig. 3-6.

## Test Signal B

Test Signal B is of four components with significant variations in their relative amplitudes and FM laws in order to resemble real EEG data; it is defined as:

$$x_B(t) = \sum_{i=1}^4 b_i \cos [2\pi (c_i(t-4)^4 + d_i t)], \quad 0 \leq t < 8, \quad (3.52)$$

where the coefficients  $b_i$ ,  $c_i$  and  $d_i$  are given in Table 3.2, and the sampling frequency is  $f_s = 32$  Hz. TFRs of Test Signal B with AWGN at SNR equal to 12 dB are shown in Fig. 3-5 using the same colormap. Parameters of the considered TFRs are optimized based on visual inspection and tabulated in Table 3.1. Results of the NMSE are shown in Fig. 3-6.

## Discussion of results

From the TFRs in Fig. 3-4, we may observe the following. The proposed OASTFT compromises more fairly between the two signal components in terms of energy distribution in the TF domain compared with the CR-ASTFT. The EC-ST appears to provide a *cleaner* TFR than that achieved by the CR-ST. The SM seems to be more effective in attenuating the spurious cross-terms in the TF domain while keeping good

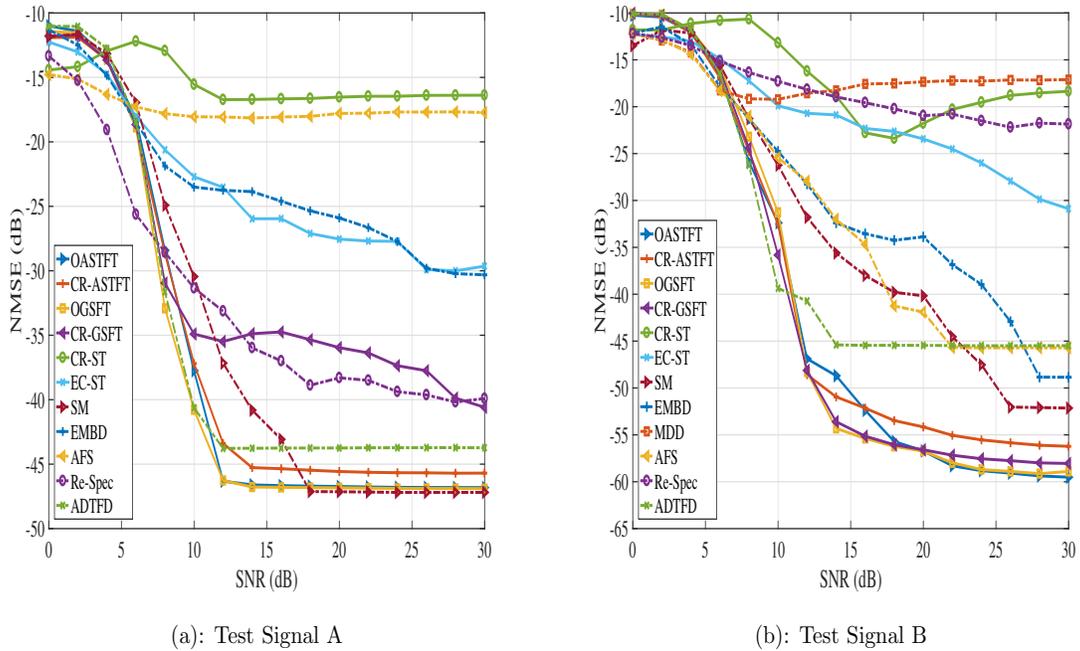


Figure 3-6: Comparison between selected TFRs in terms of the NMSE of IF estimation. Figure shows the NMSE (in decibels) as function of SNR (in decibels). We normalize both the true and estimated IFs by the sampling frequency before computing the NMSE. Results of Test Signal A in (a), and results of Test Signal B in (b). TFRs of Test Signal A and Test Signal B are depicted in Fig. 3-4 and Fig. 3-5, respectively. Results of the MDD with Test Signal A are not shown, because the MDD does not resolve the sinusoidally modulated component.

energy concentration along the IF ridges compared with the EMBD, where these artifacts are more evident. The MDD fails to resolve the sinusoidal FM component, because it is a TFR suitable for signals that can be modeled based on piecewise LFM chirps, whereas the AFS tends to *linearize* this FM segment and introduces some fake artifacts. Perhaps the best visual representations are provided by the Re-Spec and the ADTFD, but those obtained by the OGSFT and the SM are also *sharp* and easily readable; they feature high energy concentration along the IF ridges with almost complete cross-terms suppression.

Moving now to the TFRs of Test Signal B in Fig. 3-5, as far as the ST-based methods are concerned, we see again that the TFR achieved by means of the EC-ST is much cleaner than that provided by the CR-ST. The ASTFT-based TFRs, namely the OASTFT and the CR-ASTFT, seem to yield TFRs somehow close to

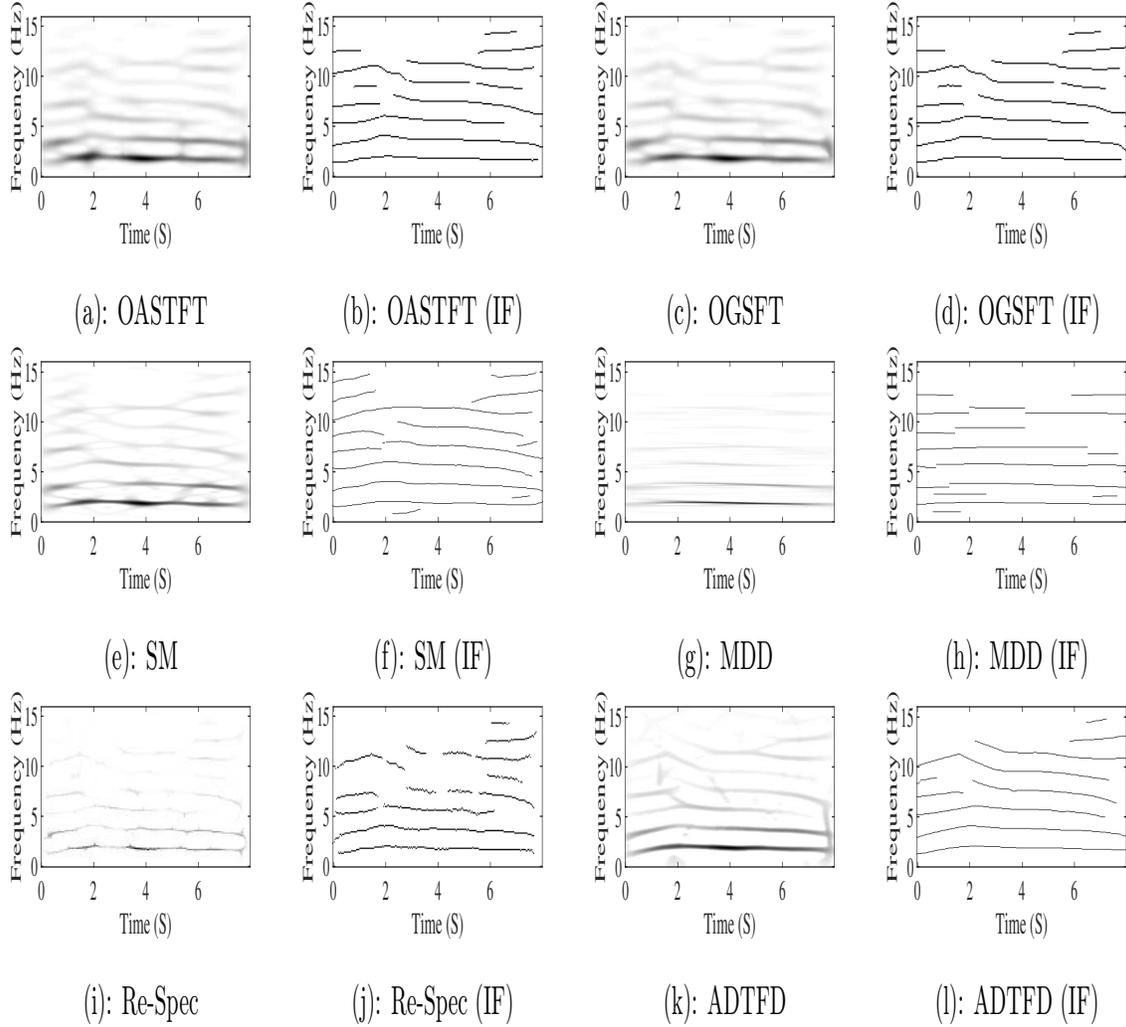


Figure 3-7: Comparison between selected TFRs in terms of IF estimation. Analyzed signal is real-world EEG seizure signal corrupted by AWGN at SNR equal to 15 dB. Figure shows TFRs and corresponding IF estimates extracted using an algorithm based on image processing.

those provided by the OGSFT and the CR-GSFT, respectively, which is a result justified by the fact that the differences in the IFR between the signal components are not as significant as in the previous example. Among the linear TF transforms, the OGSFT achieves the sharpest TFR. The SM, once again, appears to outperform the EMBD in terms of the quality of the visual representation. The linearization effect of the MDD is clear with this test signal, where the four IF ridges appear as linear functions in the TF domain, while the AFS produces notable artifacts, yet the IF components can be visually identified. The TFR returned by the ADTFD (with

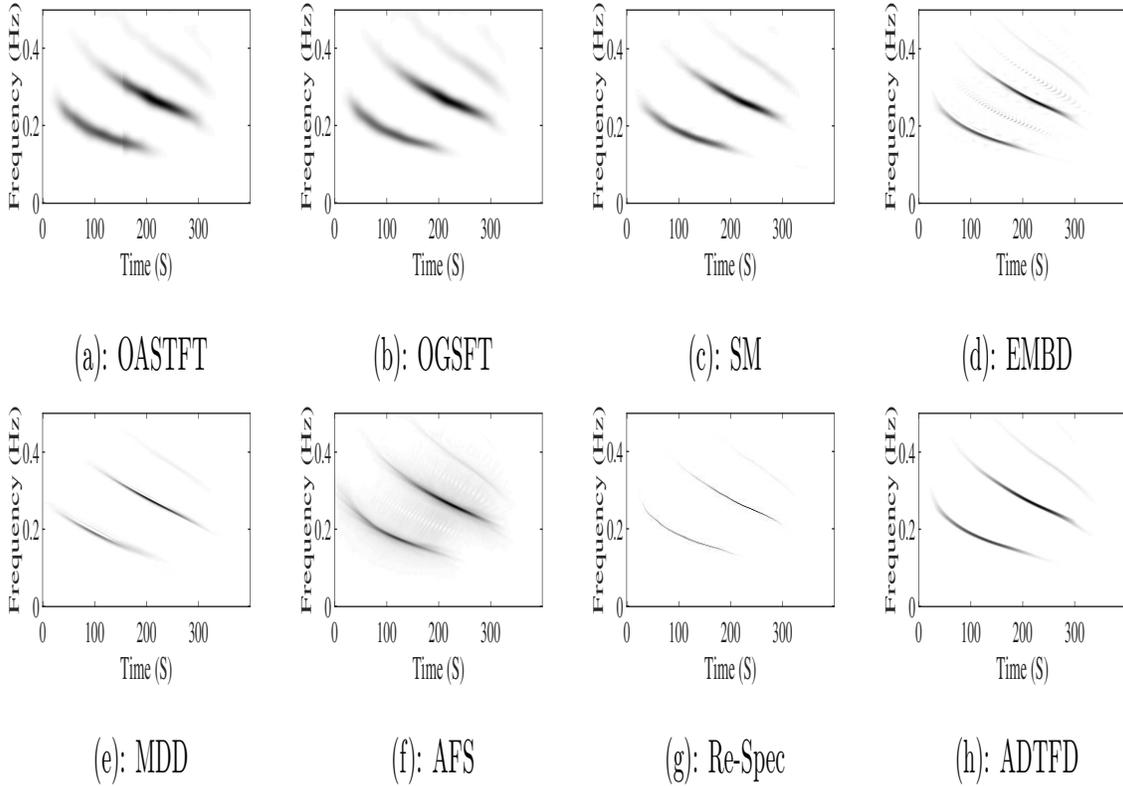


Figure 3-8: TFRs of real-world bat echolocation sound corrupted by AWGN (SNR equal to 20 dB). Respective parameters of TFRs are provided in Table 3.1. Parameters of EMBD and ADTFD are adjusted based on the results in [7], while those of the SM and the Re-Spec are based on visual inspection. Proposed methods do not require parameter optimization and are fully automated

optimized parameter setting) is perhaps the most impressive.

As one can clearly see, evaluating a TFR based on visual inspection could be hard and somehow subjective. This work considers the accuracy of IF estimation as the primary criterion of goodness. Observing the NMSE curves in Fig. 3-6, the following conclusions can be drawn.

1. The proposed methods outperform all the other considered algorithms in IF estimation starting from about 10 dB of SNR, at least for these two examples. This result is explained by the facts that the OASTFT and the OGSFT are linear, hence returning TFRs almost free from cross-terms, and their IF estimates are approximately unbiased.
2. The OASTFT and the OGSFT provide more accurate IF estimates compared

with their counterparts presented in [6], thanks to their statistically optimized window sizes. Note that, unlike the OGSFT, the performance of the CR-GSFT degrades significantly with Signal A compared with Signal B, indicating the sensitivity of the 2-D interpolation used in [6] to the IF ridges in the TF domain.

3. The ST-based methods, namely the CR-ST and the EC-ST, are significantly less accurate in IF estimation than those that adapt the TFR window in time or time and frequency. Indeed, it turns out that the CR-ST is sensitive to errors in the coarse IF estimates, which may require some sort of post-processing.
4. Generally, the OASTFT and the OGSFT seem to perform closely at SNR larger than 10 dB, although the latter is more adaptive and renders a TFR that is, in general, more impressive in terms of readability and energy concentration compared with the former.
5. The SM and the ADTFD seem to be superior to the other methods based on QTFRs in terms of IF estimation; their estimates, nonetheless, are less accurate than those returned by the proposed methods for the considered test signals. Their good performance, moreover, is dependent on proper setting of their respective parameters. Note also that the accuracy of the IF estimates of the AFS and those of the Re-Spec differ significantly between Signal A and Signal B.

### **Evaluation on real signals**

Here, the proposed methods are evaluated using real-world data. Although, as mentioned before, this work considers the accuracy of IF estimation to be the primary evaluation criterion, this criterion cannot be used with real signals, since the true IF is not exactly known.

### **Real-world EEG seizure signal**

We consider a newborn EEG seizure signal of duration 8 seconds acquired at sampling frequency of 32 Hz taken from [80]. AWGN is added to this signal at SNR equal to

Method	OASTFT	OGSFT	SM	EMBD	MDD	AFS	Re-Spec	ADTFD
EMC	$1.8 \times 10^4$	$1.6 \times 10^4$	$2.2 \times 10^4$	$3.2 \times 10^4$	$1.5 \times 10^4$	$8.4 \times 10^4$	$1.6 \times 10^3$	$1 \times 10^4$

Table 3.3: ECM of selected TFRs of a real-world bat echolocation sound.

15 dB, and the noisy EEG signal is analyzed using the OASTFT, the OGSFT, the SM, the MDD, the Re-Spec, and the ADTFD. Figure 3-7 shows the resultant TFRs with the corresponding IF ridges estimated using the connected component linking algorithm. Observing Fig. 3-7, we see that the IF trajectories estimated by the OASTFT and the OGSFT are almost identical and free from spurious components. The SM and the MDD concentrate high energy along the IF components in the TF domain, but they also produce some cross-terms, which can be seen in Fig. 3-7(f) and Fig. 3-7(h), respectively. The Re-Spec does not estimate accurately the IF trajectories of the weaker signal components, as can be seen in Fig. 3-7(j). The ADTFD provides a TFR with very high energy concentration and almost no cross-terms, but the previous example, which uses a synthetic signal, shows that the IF estimates of the ADTFD are less accurate than those provided by the OASTFT and the OGSFT. This example demonstrates that the proposed methods do not yield TFRs with the highest energy concentration, but their IF estimates are very accurate. More importantly, these methods are completely *automatic* in the sense that they do not require interference by a user to properly set any parameter, while the good performance of other algorithms is conditioned by appropriate parameter tuning, which might be tedious when performed manually or requires an optimization algorithm, as proposed in [7] and [82] for the ADTFD.

### Bat echolocation sound

To further illustrate the performance of the developed methods on real-world data, we consider now a bat echolocation sound signal of 400 samples taken from [15]. To analyze this signal, we employ the OASTFT and the OGSFT, representing TFRs based on linear transforms, and for comparative purposes, the SM, the EMBD, the MDD, the AFS, the Re-Spec, and the ADTFD are also considered (see the parameters used to implement these TFRs in Table 3.1). TFRs of the analyzed signal with AWGN

Method	OASTFT	OGSFT	SM	EMBD	MDD	AFS	Re-Spec	ADTFD
time (s)	0.46	2.9	0.09	0.24	0.18	1.5	0.21	9.3

Table 3.4: Comparison between selected TFRs in terms of computation time.

set at SNR equal to 20 dB are displayed in Fig. 3-8 using the same colormap. The accuracy of IF estimation cannot be used to objectively quantify the performance of a TFR of a real-world signal; therefore, we use instead Stanković’s ECM [5]. A low value of this measure indicates a TFR with high energy concentration. Table 3.3 reports results of the ECM for the selected TFRs. They show that the Re-Spec has the minimum ECM, followed by the ADTFD, thus implying maximum energy concentration, which is also confirmed by visual inspection. Note that the proposed methods achieve lower ECMs than the EMBD, which gives rise to notable cross-terms, and also the SM.

A comparison between the previous methods in terms of computation time (averaged over 50 realizations) is provided in Table 3.4. The experiments were run on Windows-7 Intel Core i-5 platform with 6 GB RAM. The results confirm that the OASTFT is more computationally efficient than the OGSFT. They also show that the SM is the fastest algorithm, but it is nonadaptive, while the ADTFD is the most computationally demanding, since it locally optimizes the direction of a smoothing kernel. It is important to note that the results reported in Table 3.4 do not account for the time and numerical complexity required to properly tune, or possibly optimize, the parameters of the TFRs given in Table 3.1, which can be significantly demanding. On the other hand, and as mentioned before, the proposed methods are completely automated. It is useful to mention that in this example, around 96% of the computation time of the OASTFT and 15% of that of the OGSFT is spent in the IF extraction and tracking algorithm.

## 3.3 Adaptive IF estimation based on QTFRs

### 3.3.1 Introduction

The focus in this section is on adaptive IF estimation through QTFRs. As previously explained, when accurate signal models are not available, a simple yet effective method for nonparametric IF estimation is established based on the maxima position of a TFR [15, 20, 16]. Identified sources of error are: 1. bias, originating from the IF high-order derivatives; 2. small noise because of which the TFR local maxima remain within corresponding auto-terms; and 3. large noise, causing the local maxima to be located outside the signal components. The first two sources of error are analyzed in [83], whereas the last is considered in [89]. Analysis of the error caused by small disturbances demonstrates that the estimation bias and variance are monotonically increasing and decreasing functions of the TFR window size, respectively. For the Wigner distribution (WD), the window size that resolves the bias-variance trade-off in the least MSE sense was derived in [84], but was deemed of impractical value for being a function of the IF second-order derivative, which is, of course, unknown, since the IF itself is to be estimated. An algorithm for adaptive IF estimation which does not require *a priori* knowledge of the IF derivatives was presented in [84, 90] based on the ICI rule. In short, this method tests adaptively candidates for the window length from a predefined discrete set, and for each size, a confidence interval, within which the true IF exists with a certain probability, is defined using the asymptotic formula of the estimation variance. An estimate of the optimal width is determined as the largest width in the set for which two successive intervals have at least one point in common. The ICI-based method was proven to provide good estimation accuracy in working environments characterized by moderate to high SNRs; its performance was analyzed in [91] and procedures to properly adjust its parameters were presented. It was subsequently applied with other TFRs, including higher-order ones, in [92, 93, 94]. A modification of this method was proposed in [95], where the amount of overlap between the current and previous confidence intervals is tested to improve the accuracy of estimation. Djurović and Stanković in [96] refined the method behavior

in high-noise environments by excluding narrow windows that produce large errors and applying a median filter to the estimates returned by middle-size windows to remove possible outliers.

Despite the fact that the ICI-based method is of central importance in adaptive IF estimation of monocomponent signals, as accentuated in [14], the following limitations remain. First, since the method nature is based on statistical hypothesis testing, the IF estimation accuracy depends, obviously, on the cardinality of the set of windows used by the ICI rule. Second, if we make use of a larger set of window widths, we may improve on accuracy but at the expense of higher computational complexity, compromising the method practicality. Tackling the previous shortcomings, in this section, we take a different approach to adaptive IF estimation. Essentially, we argue that the closed-form formula which relates the optimal width (in the MSE sense) to the IF second-order derivative can be used directly as a criterion for width selection once a reasonable estimate of the second-order derivative is available. To this end, we first construct a preliminary PTFR, which is a WD or one based on the SM [31], and use the peaks of its dominant ridges as a rough estimate of the IF trajectory. Thence, the continuous wavelet transform (CWT) is introduced to accommodate an estimate of the IF second-order derivative. Although numerical differentiation of noisy process is known to be a classical ill-posed problem, we will show that, by a proper selection of a wavelet mother, the CWT combines smoothing with high-order differentiation, providing a good approximation of derivatives.

### 3.3.2 Background theory

Let  $y(n) = x(n) + \epsilon(n)$  be a noisy signal, where  $x(n) \equiv x(nT)$  is a sampled version of the continuous-time and noise-free signal  $x(t) = A \exp(j\phi(t))$  with  $T$  being the sampling interval.  $\epsilon(n) \equiv \epsilon(nT)$  is a complex-valued AWGN of variance  $\sigma_c^2$ . The signal phase is  $\phi(t)$  and  $A$  is a constant real-valued amplitude. By definition, the signal IF is the first-order derivative of its phase [13, 14, 15]:  $\omega(n) \triangleq \phi'(t)|_{t=nT}$ . Here, the IF is estimated based on the maxima position of the WD according to:

$\hat{\omega}(n) = \arg \max_{\omega} \text{WD}_y(n, \omega)$ , where  $\text{WD}_y(n, \omega)$  denotes the WD of  $y(n)$  which is given by [84]:

$$\text{WD}_y(n, \omega) = \sum_{k=-h/2}^{h/2-1} w_h(k) y(n+k) y^*(n-k) \exp(-j2\omega k) \quad (3.53)$$

in which  $w_h(k) \equiv w_h(kT)$  is a real-valued symmetric window of length  $h$ . The estimation error, defined as  $\Delta\hat{\omega}(n) = \hat{\omega}(n) - \omega(n)$ , is analyzed in [84]; therein, it is demonstrated that the error has the following bias and variance, respectively:

$$\text{Bias}(n, h) \cong c_1 h^2 \omega''(n), \quad \text{Var}(h) = \sigma^2(h) \cong c_2 \frac{\sigma_\epsilon^2}{A^2} \frac{1}{h^3}, \quad (3.54)$$

where  $c_1$  and  $c_2$  are window-dependent parameters. The contradictory behavior between the bias and the variance with respect to the length  $h$  is made clear by (3.54), and the fact that the estimation bias is dependent on the IF itself establishes that any reasonable choice of the length  $h$  should be data-driven. Optimizing  $h$  by minimizing the estimation MSE, defined as the sum of the squared bias and the variance, results in [90]:

$$h_{opt}(n) = \left[ \frac{3c_2\sigma_\epsilon^2/A^2}{(2c_1\omega''(n))^2} \right]^{1/7}. \quad (3.55)$$

If the signal is highly contaminated by noise, the previous analysis loses its validity. In such a case, it was demonstrated in [89] that the estimation error exhibits an impulsive nature, and the WD peaks in any location outside the auto-term position with the following probability:

$$P_e(h) = 1 - \frac{1}{\sqrt{2\pi}\sigma_{\text{WD}}} \int_{-\infty}^{+\infty} \left( 1 - 0.5 \operatorname{erfc} \left( \frac{z}{\sqrt{2}\sigma_{\text{WD}}} \right) \right)^{h-1} \exp \left( \frac{-(z - A_{\text{WD}})^2}{2\sigma_{\text{WD}}^2} \right) dz, \quad (3.56)$$

where  $A_{\text{WD}} = E_0 A^2$  and  $\sigma_{\text{WD}}^2 = E_1 \sigma_\epsilon^2 (2A^2 + \sigma_\epsilon^2)$  with

$$E_0 = \sum_{k=-h/2}^{h/2-1} w_h(k), \quad E_1 = \sum_{k=-h/2}^{h/2-1} w_h^2(k). \quad (3.57)$$

### 3.3.3 IF estimation based on the ICI rule

The fact that the estimation bias and variance are dependent on the IF itself establishes that any reasonable choice of the length  $h$  should be data-driven. Katkovnik and Stanković in [84] presented an algorithm for adaptive IF estimation using the WD that does not require prior knowledge of the bias; we briefly recall it here. Let  $H = \{h_1 < h_2 < \dots < h_j\}$  be a finite-length set of increasing window lengths. A confidence interval associated with a length  $h_j$  is defined given by  $D_{h_j}(n) = [L_{h_j}(n), U_{h_j}(n)]$  where  $L_{h_j}(n) = \hat{\omega}_{h_j}(n) - K\sigma(h_j)$ ,  $U_{h_j}(n) = \hat{\omega}_{h_j}(n) + K\sigma(h_j)$ , and with  $\hat{\omega}_{h_j}$  and  $\sigma(h_j)$  indicating the IF estimate, obtained using a window of length  $h_j$ , and its standard deviation, respectively, and  $K$  is a constant. At each time instant  $nT$ , the ICI-based method takes window width adaptively, approximating the optimal width  $h_{opt}(n)$  by the largest member of  $H$  for which the following condition is still satisfied:

$$D_{h_j}(n) \cap D_{h_{j-1}}(n-1) \neq \emptyset. \quad (3.58)$$

For computation of  $\sigma(h_j)$ , the quantities  $A$  and  $\sigma_\epsilon$  are usually estimated through [97, (23-24)]. Accuracy of the IF estimate returned by the ICI-based method shows some improvement by introducing an additional condition, requiring the quantity  $|D_{h_j}(n) \cap D_{h_{j-1}}(n)|/|D_{h_j}(n)|$  to be larger than or equal to a predefined threshold, as shown in [95]. Applicability of the ICI-based method was extended to high-noise environments in [96]. The authors essentially proposed to exclude the members of  $H$  that are prone to high noise, identifying them as those for which  $P_e(h) < P_1$ . Also, the IF estimates obtained by windows whose widths satisfy  $P_1 \leq P_e(h) \leq P_2$  are smoothed using a median filter to remove possible outliers.  $P_1$  and  $P_2$  are predefined probabilities. For the other windows in  $H$ , the IF estimates are left unchanged.

### 3.3.4 Motivation for the proposed approach

Essentially, we try to obtain an estimate of the IF second-order derivative, which, once available, can be used directly to adjust the window size by a direct application of (3.55). The main argument in support of this alternative approach comes from

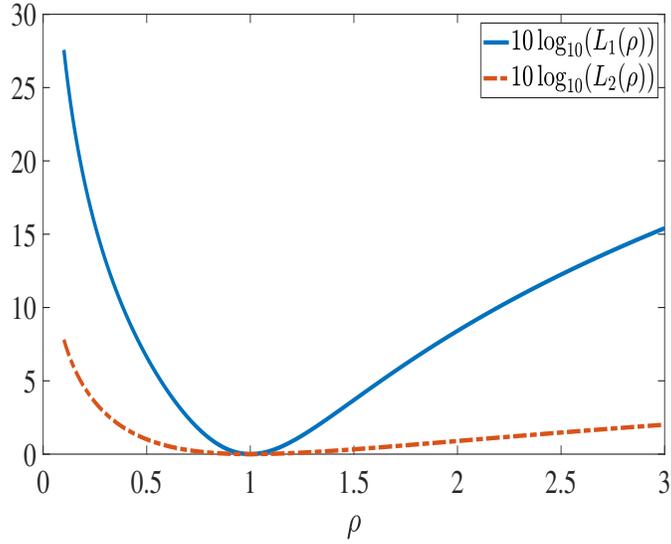


Figure 3-9: Variations of  $L_1(\rho)$  and  $L_2(\rho)$  (in dB). Function  $L_2(\rho)$  increases above the stationary point at  $\rho = 1$  much slower compared with  $L_1(\rho)$ . Note, for example,  $L_1(\rho = 2)$  is more than 5 times larger than  $L_2(\rho = 2)$ .

observing that, by letting a reasonable estimate of the IF second-order derivative control the width selection according to (3.55), variations of the MSE around its minimal value become significantly slower compared with the case in which the MSE changes directly as a function of  $h$ . To clarify this point, we begin by writing the MSE in two forms. In the first one, the bias and the variance are given by the set of equations (3.54), resulting in:

$$\text{MSE}_1(n, h) = (c_1 \omega''(n) h^2)^2 + \frac{c_2 \sigma_\epsilon^2 / A^2}{h^3}. \quad (3.59)$$

While in the second formulation,  $h$  is substituted by

$$\left[ (3c_2 \sigma_\epsilon^2 / A^2) / (2c_1 \widehat{\omega}''(n))^2 \right]^{1/7}$$

with  $\widehat{\omega}''(n)$  denoting an estimate of the IF second-order derivative:

$$\text{MSE}_2(\widehat{\omega}''(n)) = \left[ 3(\omega''(n))^2 + 4(\widehat{\omega}''(n))^2 \right] \left[ \frac{c(\sigma_\epsilon^2 / A^2)^4}{(\widehat{\omega}''(n))^8} \right]^{1/7}, \quad (3.60)$$

where  $c = (c_1^3 c_2^2)^2 / 108$ . At a given time instant, the minimum MSE, denoted by  $\text{MSE}_{min}(n)$ , results from (3.59) by the substitution  $h = h_{opt}(n)$ , or from (3.60) by replacing  $\widehat{\omega}''(n)$  with  $\omega''(n)$ . Now, to understand how the functions in (3.59) and (3.60) varies around  $\text{MSE}_{min}(n)$  as a result of  $h(n)$  and  $\widehat{\omega}''(n)$  deviating from their respective optimal values, we start by defining the following two quantities:  $L_1(\rho) = \text{MSE}_1(n, \rho h_{opt}(n)) / \text{MSE}_{min}(n)$  and  $L_2(\rho) = \text{MSE}_2(\rho \omega''(n)) / \text{MSE}_{min}(n)$ , where the deviations of  $h(n)$  and  $\widehat{\omega}''(n)$  from  $h_{opt}(n)$  and  $\omega''(n)$ , respectively, are described by a real parameter  $\rho$ . It can be shown that the ratios  $L_1(\rho)$  and  $L_2(\rho)$  change as functions of  $\rho$  according to:  $L_1(\rho) = (4 + 3\rho^7) / (7\rho^3)$  and  $L_2(\rho) = (3 + 4\rho^2) / (7\rho^{8/7})$ . These expressions, in turn, make clear that, as  $\rho$  shifts away from 1,  $L_1(n)$  and  $L_2(n)$  increase above their stationary point (at  $\rho = 1$ ) with a rate being significantly larger for  $L_1(n)$  than it is for  $L_2(n)$ , as illustrated in Fig. 3-9. We conclude that, the IF estimation approach in which an estimate of the IF second-order derivative controls the size of the TFR observation window using (3.55) can tolerate error in the derivative estimation larger than the error an approach based on adaptive length selection from a discrete scheme (like the ICI method) can endure in its search for the most suitable length. The evident flatness of  $L_2(\rho)$  around its stationary point indicates that the estimate  $\widehat{\omega}''(n)$  does not have to be highly accurate in order to achieve a low MSE, but rather *reasonable*. To obtain such an estimate, the method proposed in this communication constructs first a PTFR and uses its dominant ridges as a rough estimate of the IF trajectory. Then, the IF second-order derivative is estimated by the CWT. This estimate is finally inserted in (3.55) to adjust the window size of an improved TFR, whose local peaks provide a more accurate estimation of the IF.

### 3.3.5 Selecting a PTFR

The PTFR is aimed to provide a rough estimate of the IF trajectory, which, once available, is used by the CWT for derivative calculations. We stress again that this preliminary estimation does not have to be highly accurate, since, as explained before, an IF estimate with small MSE can be obtained even with an error in the derivative estimate, provided that this error is not very large. Indeed, it turns out that only

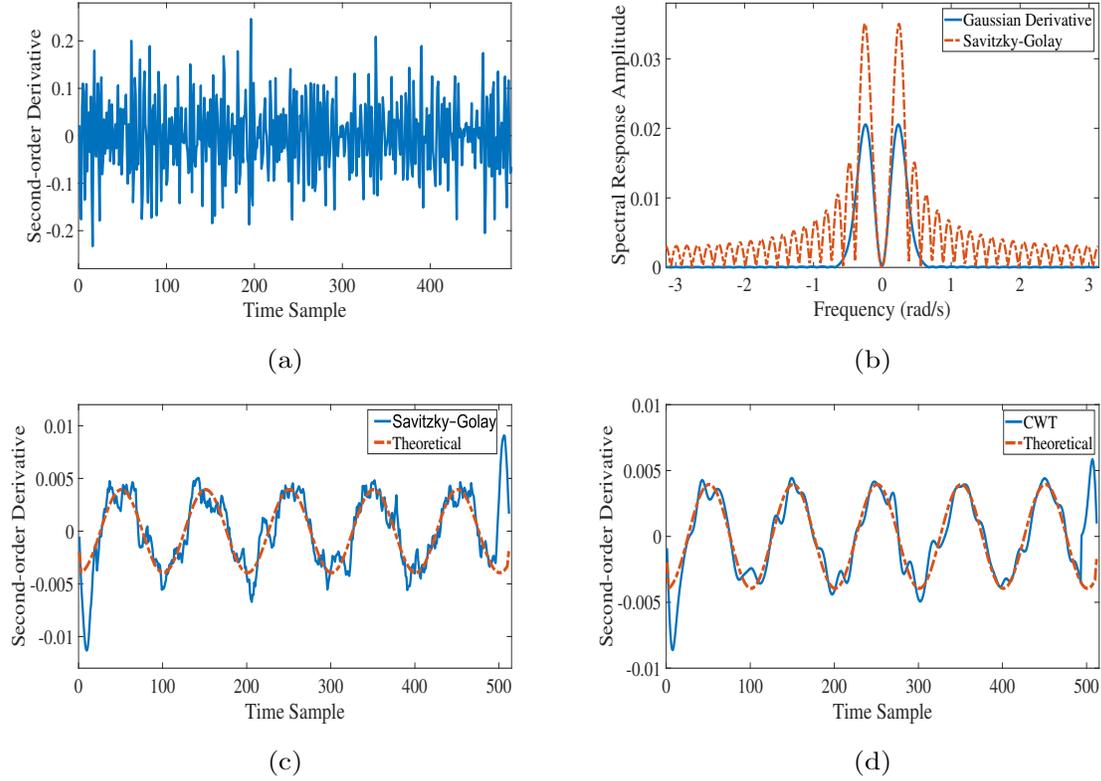


Figure 3-10: Example of computing the second-order derivative of a noisy sine waveform (SNR=15 dB). (a) Derivative computed by the conventional finite-difference method. (b) Spectral response amplitude of the filter  $\Theta_s(t)$  given by the second-order derivative of a Gaussian and SG second-order differential filter (note the high side lobes of SG filter). Parameters of the filters are set such that they have the same cutoff frequency. Derivative computed by: SG differential filter in (c), and the CWT in (d). Most accurate result is provided by CWT.

large errors, which dislocate the PTFR peaks far from the corresponding auto-terms, impair the quality of the final IF estimates. To avoid the occurrence of such errors, based on an estimate of  $A$  and  $\sigma_\epsilon$  obtained following [84, (23, 24)], the PTFR window size, denoted henceforth by  $h_{int}$ , is determined as the smallest in a predefined discrete set of window widths such that the probability of large error is smaller than a threshold:  $P_e(h_{int}) < P_{th}$ . By satisfying this requirement on  $P_e(h_{int})$ , large errors occur with a small probability so as to allow for the smoothing offered by the CWT to effectively alleviate their impact on derivative calculations. We remark that the reasons behind selecting the narrowest window among those that satisfy the previous accuracy requirement are: 1. to avoid IF estimation with large bias, because large bi-

asedness in the initial IF estimate is more detrimental to derivative calculations than variance, whose impact can be better mitigated by smoothing; 2. abrupt changes in the IF trajectory cannot be revealed by a wide observation window, leading to serious errors in derivative estimation. In the numerical realization,  $P_{th}$  is set to 0.1% at high values of SNR ( $\text{SNR} \geq 5$  dB), while at lower SNRs ( $\text{SNR} < 5$  dB), to avoid selection of very large windows, which are prone to inner-artifacts,  $P_{th}$  is increased to 1/3 (similar probability thresholds were used in [96]). For signals with highly nonlinear IF laws in high-noise environments, the WD is contaminated by aggravating interfering inner-terms, which constitute in this case the dominant source of estimation errors. As a possible remedy to inner-terms, we may apply the proposed algorithm with the SM, which is a reduced-interference TFR containing the appealing features of the WD and the STFT. We note that the SM was also used for adaptive IF estimation in [96]. The proposed method is summarized by the flowchart in Fig. 3-11.

### 3.3.6 Derivative approximation by the CWT

Despite being of fundamental importance in various fields of study, numerical differentiation (ND) is known to be an ill-posed problem, meaning that small perturbations lead to large error in the approximate solution. In this context, several works have reported the potential offered by the CWT to derivative calculation [98, 99], which originates from its intrinsic capability to combine smoothing with differentiation, as briefly explained in what follows.

The CWT of a continuous-time signal  $f(t)$  is defined by

$$\text{CWT}_f(t, s; \psi) = \int_{-\infty}^{+\infty} f(u) \frac{1}{\sqrt{s}} \psi \left( \frac{u-t}{s} \right) du, \quad (3.61)$$

where the analyzing wavelet (assumed real in the previous definition) results from a mother wavelet  $\psi(u)$  by dilation and translation through  $s$  and  $t$ , respectively. The transform in (3.61) can be written equivalently using the convolution operator according to:  $\text{CWT}_f(t, s; \psi) = 1/\sqrt{s} f(t) * \psi(-t/s)$ . For a wavelet characterized by  $m$  vanishing moments, there exists a function  $\theta(t)$  whose  $m$ th-order derivative relates

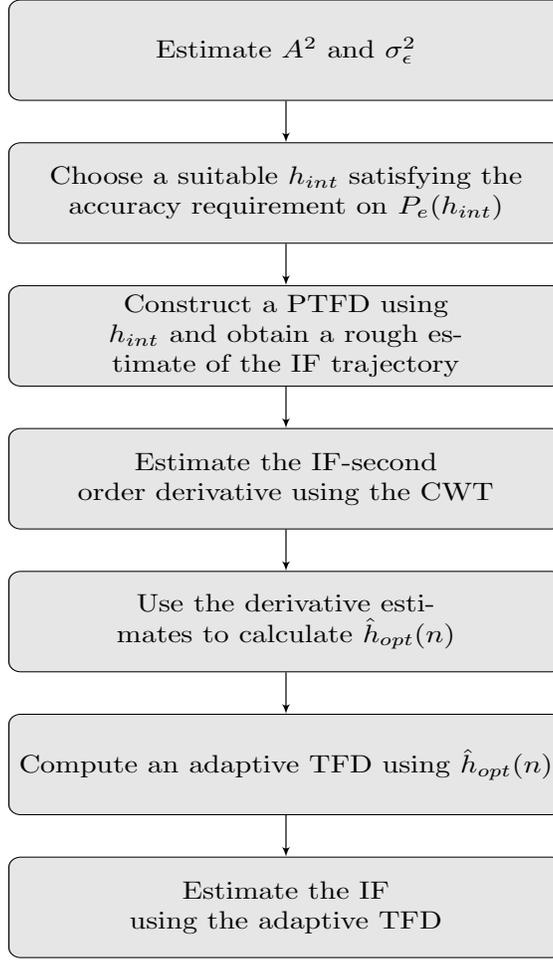


Figure 3-11: Flowchart of proposed method for adaptive IF estimation based on QTFR.

to this wavelet by [100]:

$$\int_{-\infty}^{+\infty} t^i \psi(t) dt = 0, \quad i = 0, \dots, m-1 \iff \psi(t) = (-1)^m \frac{d^m}{dt^m} \theta(t). \quad (3.62)$$

It was also proven that  $\theta(t)$  satisfies the conditions of a smoothing function by having a fast decay and nonzero integral, which establishes that  $\theta(t)$  behaves as a lowpass filter in the spectral domain. By the theory of convolution, the CWT of  $f(t)$  can be written now as:

$$\text{CWT}_f(t, s; \psi) = s^{m+1/2} \frac{d^m}{dt^m} \left[ f(t) * \frac{1}{s} \theta \left( \frac{-t}{s} \right) \right]. \quad (3.63)$$

Rearranging (3.63) gives:

$$\frac{\text{CWT}_f(t, s; \psi)}{s^{m+1/2}} = f(t) * \frac{d^m}{dt^m} \left[ \frac{1}{s} \theta \left( \frac{-t}{s} \right) \right] = f(t) * \Theta_s(t). \quad (3.64)$$

Equations (3.63, 3.64) clearly demonstrate that the CWT through a mother wavelet with  $m$  vanishing moments is intrinsically equivalent to the  $m$ th-order derivative of a lowpass-filtered version of the analyzed signal (up to a constant factor). We remark that the lowpass behavior exhibited by  $\Theta_s(t)$  is certainly desirable if the signal to be differentiated is noisy. This behavior is controlled by the dilation parameter, which sets the filter cutoff frequency. With increasing  $s$ , the spectral response of  $\Theta_s(t)$  becomes narrower, suppressing a larger amount of noise power. On the other hand, excessive smoothing leads to loss of important derivative details. Therefore, a trade-off should be adopted when selecting a suitable value for  $s$ , but, in general, the sole constraint on the chosen mother wavelet is given in (3.62). Note also that the  $m$ th-order derivative can be computed by one transform procedure. For the purpose of this work, we selected the second-order derivative of a Gaussian as mother wavelet, namely:  $\theta(t) = 1/\sqrt{\pi} \exp(-t^2)$ . Example of computing the second-order derivative of a noisy sine waveform with a comparison using Savitzky-Golay (SG) filter [101] is illustrated in Fig. 3-10.

### 3.3.7 Numerical evaluation

In this section, the proposed IF estimation method is evaluated and compared with that based on the ICI rule [84] and its modified version described in [96] with the extra condition of [95]. The performance is quantified by means of the NMSE of the IF estimates over a large number of Monte-Carlo simulations (here is set to 100); in each iteration, AWGN is added to a noiseless signal at SNR ( $\text{SNR} = A^2/\sigma_\epsilon^2$ ) ranging from 0 to 25 dB with a 2-dB step. For numerical evaluation, we use two signals,

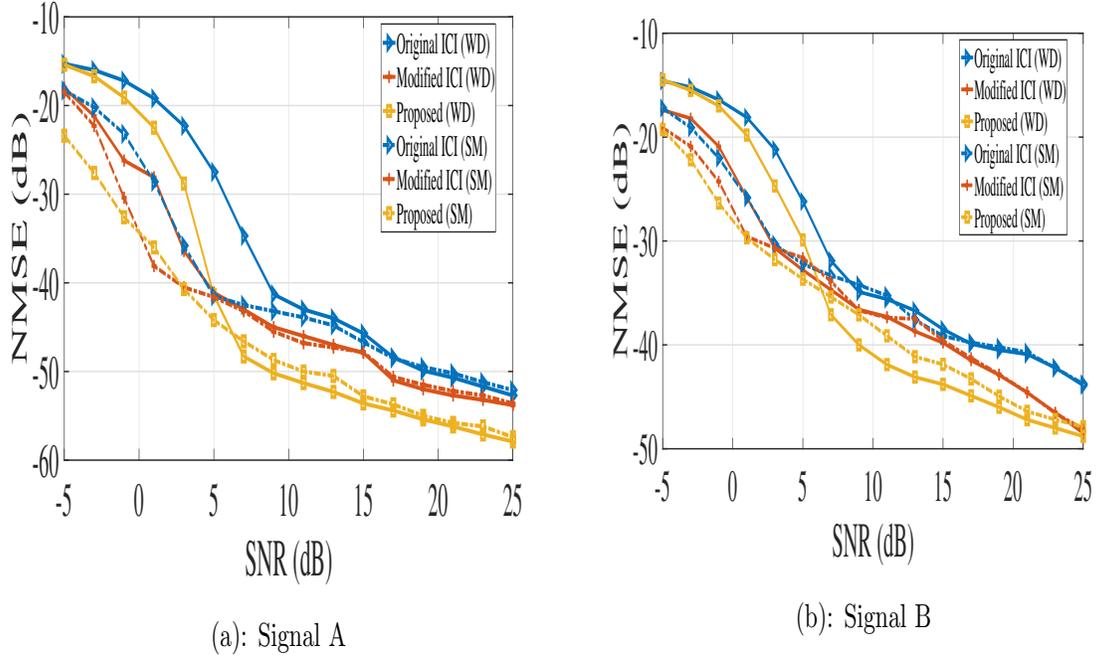


Figure 3-12: Comparison between IF estimation methods in terms of the NMSE using two test signals. All the considered methods are implemented using the WD and the SM.

Signal A and Signal B, with IF laws,  $\omega_A(n)$  and  $\omega_B(n)$ , given according to:

$$\omega_A(n) = \begin{cases} \frac{2.58\pi}{3}, & \text{for } n \in \{0, 127\} \\ \frac{1.5\pi}{3} + \frac{\pi}{3} \cos(4\pi(n_1 - 1)), & \text{for } n \in \{128, 255\} \\ \frac{1.5\pi}{3} + \frac{\pi}{3} \cos(2\pi(n_1 - 2)), & \text{for } n \in \{256, 383\} \\ \frac{2.5\pi}{3} + (0.5\pi - \frac{2.5\pi}{3})(n_1 - 3), & \text{for } n \in \{384, 511\} \end{cases} \quad (3.65)$$

$$\omega_B(n) = 0.5\pi + 0.4\pi \left[ \cos(4\pi n_2) - \frac{1}{3} \cos(12\pi n_2) + \frac{1}{3} \cos(20\pi n_2) \right], \quad n \in \{0, 255\}, \quad (3.66)$$

where  $n_1 = n/128$ ,  $n_2 = (n - 128)/256$ . The amplitude of both signals is set to unity.

### 3.3.8 Simulation results

In all experiments, the proposed method and the ICI-based algorithms are implemented using the WD and the SM as well. Results of the NMSE (in decibels) of

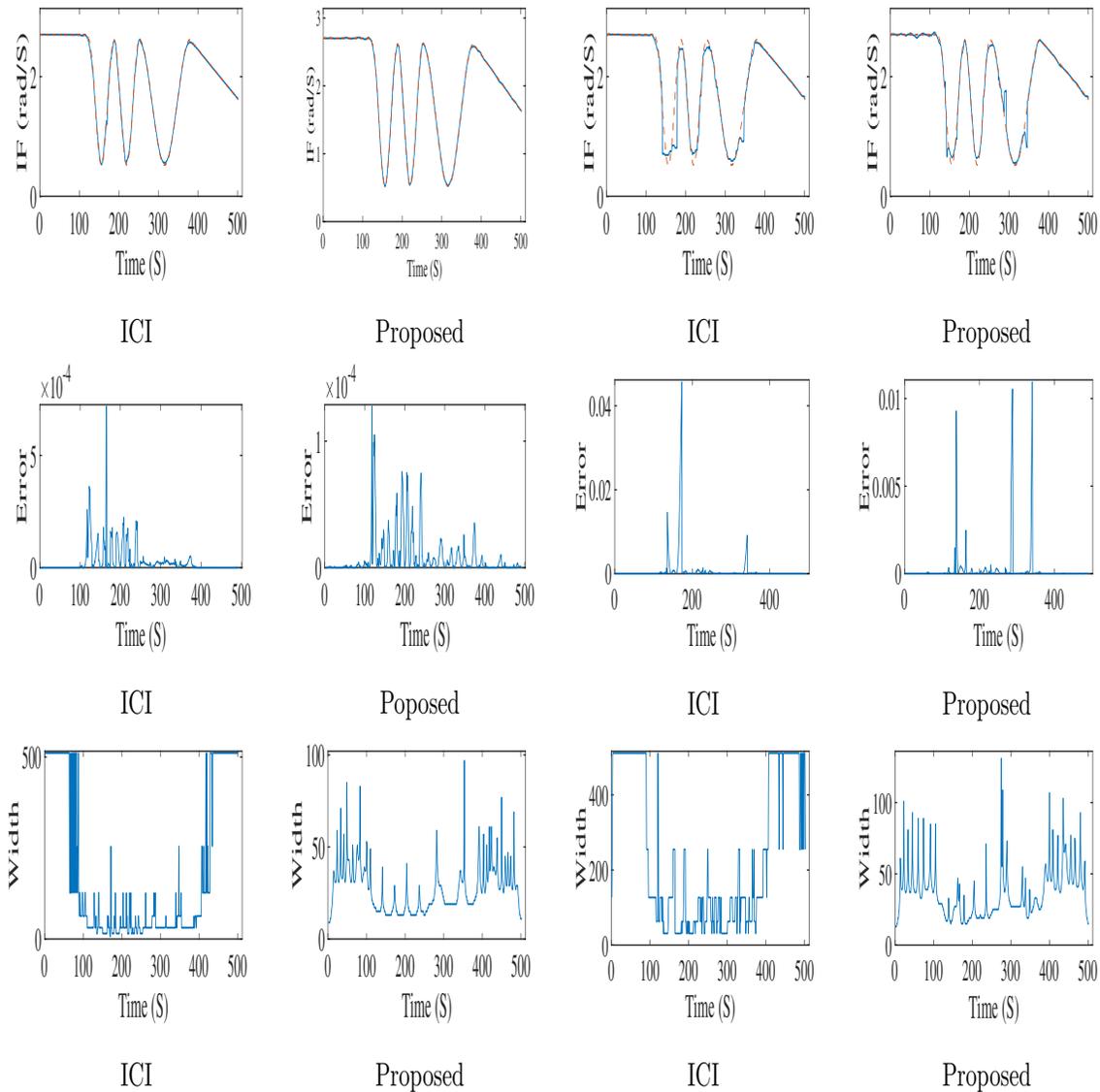


Figure 3-13: IF estimation for Signal A. Proposed method is compared with the modified ICI-based algorithm. First and second columns from left: SNR = 11 dB and the WD is used for TFR. Third and fourth columns from left: SNR = 0 dB and the SM is used for TFR. First row from top: estimated IF (solid line) and true IF (red dashed line). Second row: absolute error. Third row from top: the window width.

IF estimation are shown in Fig. 3-12 as functions of SNR. The following conclusions can be drawn from the experimental results. 1. The proposed method using the WD outperforms the ICI-based algorithms at SNRs larger than 5 dB, whereas at lower SNRs, the modified ICI method provides more accurate IF estimates. The performance shortcoming in high-noise environments is due to emphatic inner-inter-

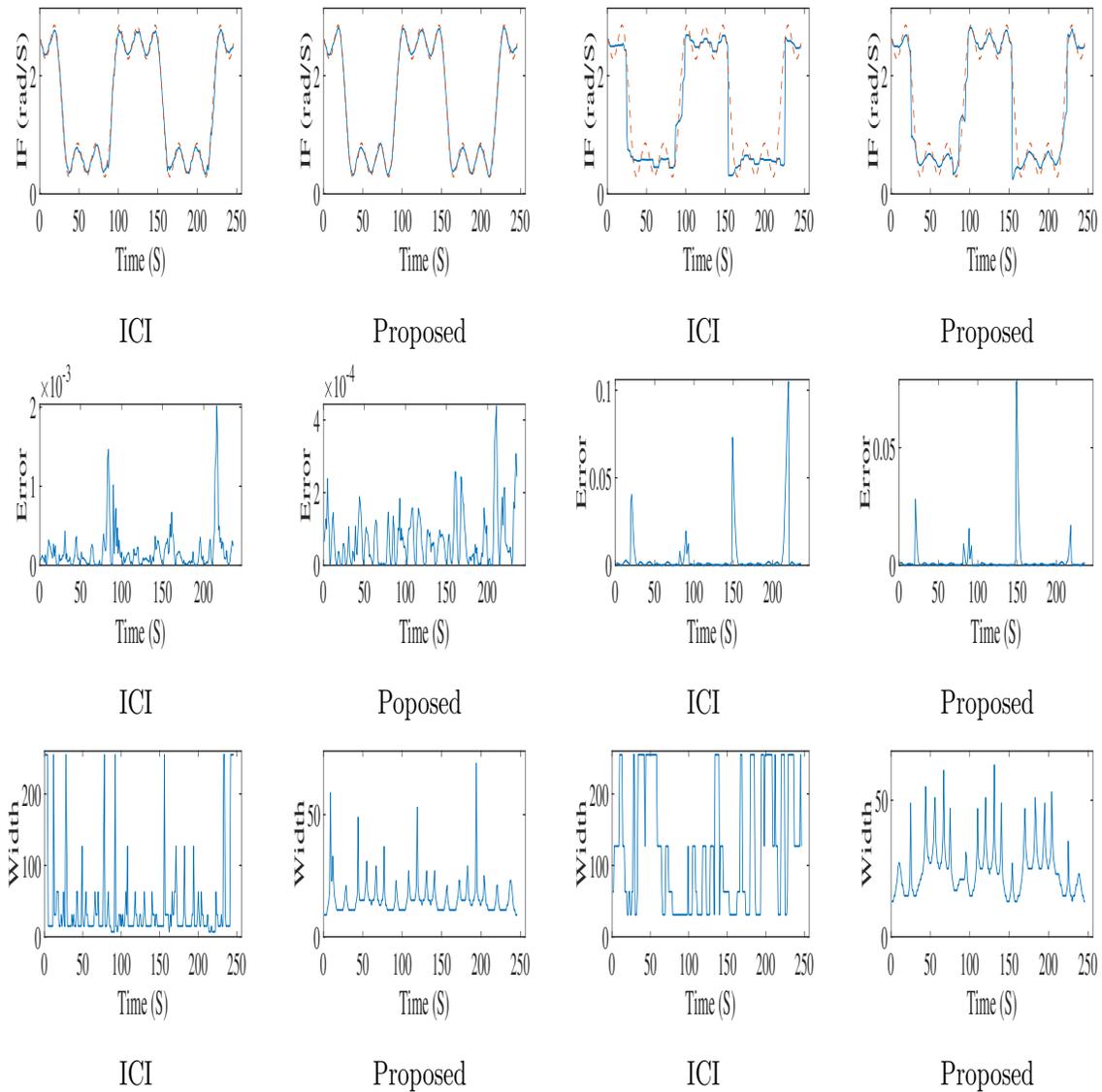


Figure 3-14: IF estimation for Signal B. Proposed method is compared with the modified ICI-based algorithm. First and second columns from left: SNR = 11 dB and the WD is used for TFR. Third and fourth columns from left: SNR = 0 dB and the SM is used for TFR. First row from top: estimated IF (solid line) and true IF (red dashed line). Second row: absolute error. Third row from top: the window width.

ference presented in the WD, which impairs the validity of (3.55). 2. When the TFRs are implemented using the SM instead of the WD for inner-terms suppression, the proposed method regains its effectiveness, improving on the ICI-based algorithms in terms of the accuracy of IF estimates. Note that the SM has also led to significant improvements in the results of the ICI-based methods at SNRs below 5 dB. 3. At

Signal	Original ICI	Modified ICI	Proposed method
Signal A ( $N = 512$ )	2.09	4.59	0.55
Signal B ( $N = 256$ )	0.81	1.94	0.34

Table 3.5: Comparison between adaptive IF estimation methods in terms of computation time. SNR is set to 11 dB, and the WD is used as TFR. The results are given in seconds, and the experiments were run on Windows-7 Intel Core i-5 platform with 6 GB RAM. Proposed method is the most computationally efficient.

SNRs larger than 5 dB, where the inner-terms are less prominent, the accuracy of the proposed method estimates is slightly degraded when the SM is used instead of the WD (this degradation is more evident with Signal B). Nonetheless, these estimates are still more accurate than those of the ICI-based algorithms. Therefore, the suggestion is to use the SM only at SNRs less than 5 dB; otherwise, the WD is preferred. We illustrate in Fig. 3-13 a single trial of IF estimation using Signal A at two values of SNR (0 and 11 dB). Fig. 3-14 replicates the results in Fig. 3-13 for Signal B.

Another advantage of the proposed method is reduced computational complexity, since only two TFRs are implemented (PTFR and final TFR), each of which requires computational complexity of order  $O(N^2 \log_2 N)$ , besides the CWT for derivative calculation, which can be realized with  $O(N)$  computational complexity. With the ICI-based algorithm, on the other hand, various TFRs should be constructed, depending on the size of the used set of widths. A comparison between the developed method and the ICI-based algorithms in terms of computation time is provided in Table 3.5. Results demonstrate the computational efficiency of the proposed method with respect to its counterparts.

# Chapter 4

## Applications of time-frequency representations

## 4.1 Jamming mitigation

### 4.1.1 Introduction

The global navigation satellite systems (GNSS) deliver precise position and time information with global coverage. Accurate and worldwide positioning is enabling an ever-growing number of applications, such as location-based services, agriculture, surveying, synchronization, just to name a few. The over-reliance of vital applications and systems on the GNSS poses a significant concern about the vulnerability of these satellite-based systems. Indeed, the susceptibility to interference is arguably the major threat to reliability and availability of the GNSS [104]. Most of the GNSS are based on the direct sequence spread spectrum (DS-SS), where the navigation bits are modulated by a faster PRN code, which spreads the signal spectrum over a bandwidth much larger than that imposed by the Nyquist Theorem. This modulation technique grants the GNSS inherent and limited immunity against interference; however, these systems are not completely interference-proof. Namely, the GNSS signals have extremely low power levels (i.e., about -160 dBW) once they arrive on the Earth's surface, thus they are likely to be overpowered by any source of interference in the surroundings, causing a large deterioration in the carrier-to-noise-density ratio. The sources of interference may be categorized according to their purpose as unintentional (e.g. solar activities, malfunctions of equipment, etc.) and intentional. Intentional interference is unauthorized jamming attempts that consist in brute-force emission of powerful interfering signals, which can be generated with basic hardware and no specific software with the aim of causing denial of GNSS services over a certain area. In this context, the prevalence of low-cost jamming devices has been reported [105], increasing awareness about susceptibility of the GNSS to interference. These devices are commercialized as Personal Privacy Devices (PPD), and although their usage is illegal, they are gaining popularity among the public due to privacy concerns related to the GNSS localization services. According to experimental surveys [106, 107], the vast majority of jammers on the market are arrays of transmitters (see sample in Fig. 4-1) that broadcast a wideband interference with periodic sawtooth-like modulations

in order to maximize the ratio between the average energy and the peak power. In the last decade, low-cost portable jammers have become more and more widespread; outages of the GNSS operation due to such devices have been experienced and reported. In late 2009, the ground-base augmentation system (GBAS) of the Newark Liberty International Airport, United States, was suffering from daily breaks, due to uninformed usage of an in-car jammer. In April 2012, in Kent, England, the police arrested members of a criminal gang responsible for the theft of some 150 high-value vehicles using jammers to disable the in-car tracking systems. In November 2013, in Australia, a Melbourne newspapers reported that more than 100 cabs in the city were suspected of using GPS jammers in order to fool the fleet management [108].

#### **4.1.2 Jamming mitigation**

The consequences of jamming onto the GNSS range from degradation of the accuracy and integrity of the navigation messages to disruption of the availability and continuity of the system. We refer the reader to [109, 110, 111], where the impact of jamming attempts onto GNSS receivers is investigated. These studies show that jammers can impair both the acquisition and tracking performance of consumer-grade receivers in a range up to 9 km, approximately. Therefore, equipping GNSS receivers with anti-jamming modules is a crucial upgrade to guarantee the reliability and availability of these systems. The demand for cost-effective implementations has motivated research on digital-signal-processing techniques for the excision of received jamming waveforms in order to avoid additional hardware, such as expensive antenna arrays or inertial measurement units. The majority of jamming mitigation methods process the raw I/Q samples at the output of the ADC/AGC loop in a domain where the powerful interference exhibits distinguishable characteristics; depending on the domain of processing, these methods may be classified as in the following.



Figure 4-1: Sample of jammers for sale online (<https://www.jammer-store.com/>).

### **Time-domain techniques**

Time-domain techniques include adaptive notch filtering through FIR or IIR filters [112]. This approach is effective as long as the modulation waveform of the jamming signal is slowly changing in time. Within this category falls also the pulse blanking (PB) method, which was suggested in [113] for narrowband GNSS receivers. This low-complexity mitigation algorithm capitalizes on the fact that when the bandwidth of the jamming signal is much wider than that of the receiver front-end, the jamming signal will resemble a periodic sequence of wideband pulses.

### **Frequency-domain techniques**

These techniques analyze the GNSS signal in the Fourier domain [114]; they are effective only when a small number of spectral components are contaminated by interference, which makes them useful against narrowband jamming signals.

### **Time-frequency techniques**

The jamming signal may be synthesized from a time-frequency representation (TFR), then subtracted from the input signal [115]. Alternatively, a weighting filter can be used to attenuate the TF components corresponding to the interfering signal in a TFR, and then a *cleaner* signal can be retrieved by inverting the modified TFR

[116]. The previous rejection methods are nonparametric, meaning that they do not rely on any a priori models or knowledge about the jamming signal, while other TF-based methods assume that a parametric model is known [117]. The most challenging aspect shared by all anti-jamming modules based on TF analysis is the demanding computational complexity, especially considering that the mitigation unit is supposed to run in nearly real time. The ultimate goal is to develop an anti-jamming technique that is indeed a practical add-on for GNSS receivers. Furthermore, an important point regarding equipping the GNSS receivers with a jamming excision unit is that the quantizer of the receiver ADC should have sufficient resolution. That is, low-end mass-market receivers equipped with 1/2-bit ADCs are prone to saturation in the presence of powerful interference, and their short dynamic range frustrates the effectiveness of any interference mitigation technique based on digital signal processing [108].

Aiming at providing the GNSS receivers with a low-complexity interference excision unit, the authors of [118] developed a mitigation technique based on an adaptive S-transform (ST), called the frequency-adaptive ST (FAST), which can be summarized by the following steps. 1. In a preliminary detection stage, the frequency bins at which the power spectral density exceeds a predefined level are identified; 2. computation of the ST is restricted to the voices corresponding to the frequency bins detected in the previous stage; 3. a TF detection stage determines the timings of the TF components affected by interference; 4. once detected in time and frequency, the interference components are blanked; 5. thence, the TF components in the masked ST are time-averaged to recover the corresponding bins of the Fourier spectrum with reduced interference power. A block diagram of the previous method is shown in [118, Fig. 2]. Obviously, when the bandwidth of the interfering signal is wide and comparable to that of the receiver front-end, as in the case of a wideband noise-like interference, the previous method will not provide a significant saving in terms of computations, since the spectrum of the GNSS signal will be entirely wiped out by interference.

Overall, it is the author's opinion that the TF analysis and signal processing has

the potential to provide very effective solutions for the problem of jamming mitigation and jamming localization [119]; the main challenge, however, is the computational complexity, which is a very important aspect often overlooked, unfortunately. Developing nonparametric algorithms able to provide effective low-complexity TF-based solutions for the problem of jamming suppression requires more research in our view, and the available methods are not mature enough to be translated into commercial solutions.

## 4.2 Abnormalities detection

TF signal and image processing is an effective and widely used technique for abnormality detection and classification of physiological signals, because they exploit the nonstationary and multi-component characteristics of such signals. The early and automated detection of abnormalities in physiological signals avoids the laborious intervention from experts, hence bringing significant improvements to the medical practice. TF-based features extracted from a TFR have been shown to effectively characterize newborn electroencephalogram (EEG)<sup>1</sup> seizures, which are harmful to the brain if inadequately treated. The TF-based approach was proven to outperform the conventional time-only or frequency-only classification approaches [26, 80]. Generally, TF-based feature detection and classification are applicable to all types of nonstationary signals regardless of their nature or origin, and can be exploited to build decision-based systems in various fields of study.

A general TF-based scheme for abnormalities detection is shown in Fig. 4-2; it includes [120]: 1. TF signal analysis; 2. features extraction to characterize different abnormalities; 3. classification of these features and assigning them to different types. The first step may entail multichannel analysis (EEG data, for example, often come from multiple electrodes on the patient's scalp), which usually offers a much more accurate model, but at the expense of higher computational requirements.

---

<sup>1</sup>EEG is a representative signal containing information about the electrical activity generated by the cerebral cortex nerve cells [120].

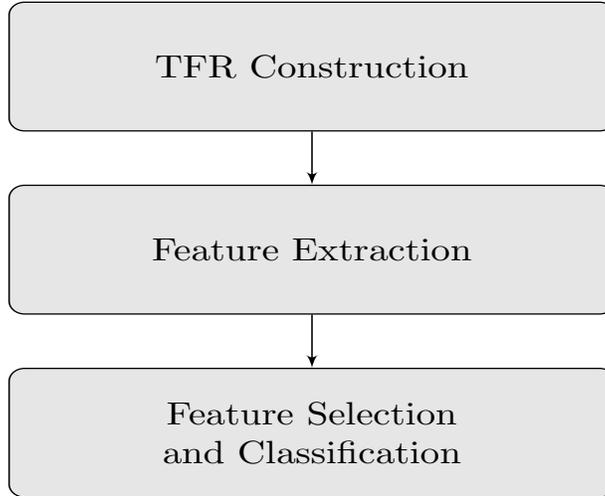


Figure 4-2: Block diagram of abnormality detection methodology using TF-based approach

#### 4.2.1 TF-based features for classifying nonstationary signals

Discriminatory features pertaining to nonstationary signals can be extracted from TFRs. To provide accurate features for detection, the selected TFR needs to have high resolution and low interference [23]. Extraction of relevant features is the key stage of any schemes for pattern recognition and classification. Once extracted, these features may be used for automated classification using a machine learning approach [80] for example. We may differentiate between two sorts of features as explained below.

##### Signal related features

The following features, which are directly related to the signal parameters, have been defined: 1. IF-based features: the mean and deviation of the IF; 2. singular value decomposition (SVD)-based features: the singular values and vectors of the TFR matrix have been proven useful in characterizing abnormalities [120]; 3. TF complexity: a measure that uses both SVD and Shannon entropy; 4. energy concentration measure: this measure is discussed in Chapter 1; 5. TF flux: to measure the rate of change of the signal energy in the TF domain; 6. subbands energies: features representative of the energy concentrated in the subbands where most of the energy is concentrated in

the TFR.

### **Image related features**

These features can be thought of as visual descriptors extracted using TFRs treated as images using image-processing techniques. Morphometric features can be extracted from TFRs after transforming them into binary images, such as area, perimeter, compactness, coordinates of the centroid, and convex hull, which can be used then for classification [121, 120, 80].

## **4.3 Multisensor TF signal processing**

In several practical applications, the data under analysis are collected from an array of measurement sensors, such as in radar, sonar, audio, and medical applications, in order to take advantage of any spatial diversity that might be present in the available data. In many real-world situations, the spectral characteristics of the signals acquired by the deployed sensors are nonstationary and vary with time. Therefore, combining array signal processing with TF analysis may provide a significant advantage in processing signals acquired by multichannel systems. The multisensor TF signal processing is concerned with representing signals in a 3-dimensional (3-D) domain that is space, time, and frequency [122].

### **4.3.1 Extension to multisensor TFRs**

We consider a non-stationary zero-mean real signal vector of  $m$  elements given by  $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ , and vector  $\mathbf{z}(t)$  representing the analytic associate of  $\mathbf{x}(t)$ , which is obtained by Hilbert-transforming each element of  $\mathbf{x}(t)$ . Here, we extend the formulations of TFRs given in Chapter 1 to the multisensor case to form

the so-called *multisensor* TFR (MTFR) for an analytic vector  $\mathbf{z}(t)$  according to [122]:

$$\begin{aligned} \text{MTFR}_{zz}(t, f) &= \begin{bmatrix} \text{TFR}_{z_1 z_1}(t, f) & \text{TFR}_{z_1 z_2}(t, f) & \dots & \text{TFR}_{z_1 z_m}(t, f) \\ \text{TFR}_{z_2 z_1}(t, f) & \text{TFR}_{z_2 z_2}(t, f) & \dots & \text{TFR}_{z_2 z_m}(t, f) \\ \vdots & \vdots & \ddots & \vdots \\ \text{TFR}_{z_m z_1}(t, f) & \text{TFR}_{z_m z_2}(t, f) & \dots & \text{TFR}_{z_m z_m}(t, f) \end{bmatrix} \\ &= \text{FT}_{\tau \rightarrow f} \left\{ G(t, \tau) *_{t} \mathbf{K}_{zz}(t, \tau) \right\}, \end{aligned} \quad (4.1)$$

where  $G(t, \tau)$  is a time-lag kernel and  $\mathbf{K}_{zz}(t, \tau)$  is the instantaneous correlation matrix (ICM) defined by

$$\begin{aligned} \mathbf{K}_{zz}(t, \tau) &= \begin{bmatrix} K_{z_1 z_1}(t, \tau) & K_{z_1 z_2}(t, \tau) & \dots & K_{z_1 z_m}(t, \tau) \\ K_{z_2 z_1}(t, \tau) & K_{z_2 z_2}(t, \tau) & \dots & K_{z_2 z_m}(t, \tau) \\ \vdots & \vdots & \ddots & \vdots \\ K_{z_m z_1}(t, \tau) & K_{z_m z_2}(t, \tau) & \dots & K_{z_m z_m}(t, \tau) \end{bmatrix} \\ &= \mathbf{z}(t + \tau/2) \mathbf{z}^H(t - \tau/2). \end{aligned} \quad (4.2)$$

Note that each element of (4.2) is given by (1.27). The diagonal elements of (4.1) are called *auto-TFRs* and those that are off-diagonal are called *cross-TFRs*. Therefore, in addition to the conventional cross-terms associated with auto-TFRs, in the MTFRs, we have cross-terms generated by the cross-TFRs, which represent the interaction between components belonging to different source signals.

The multisensor Winger distribution (MWD) of an analytic vector  $\mathbf{z}(t)$  is defined as:

$$\text{MWD}_{zz}(t, f) = \text{FT}_{\tau \rightarrow f} \{ \mathbf{K}_{zz}(t, \tau) \}. \quad (4.3)$$

The MWD is the core MTFR and from which an infinite number of representations can be generated through smoothing kernels, as is the case with the monochannel TFRs. Therefore, a MTFR can be defined as

$$\text{MTFR}_{zz}(t, f) = \gamma(t, f) *_{t, f} \text{MWD}_{zz}(t, f). \quad (4.4)$$

In a similar manner, we can extend (1.22), (1.23) to the multisensor case, constructing quadratic MTFRs (QMTFRs). That is, we can define the *spatial ambiguity function* (SAF) according to

$$\mathbf{A}_{zz}(\nu, \tau) = \begin{bmatrix} A_{z_1 z_2}(\nu, \tau) & A_{z_1 z_2}(\nu, \tau) & \dots & A_{z_1 z_m}(\nu, \tau) \\ A_{z_2 z_1}(\nu, \tau) & A_{z_2 z_2}(\nu, \tau) & \dots & A_{z_2 z_m}(\nu, \tau) \\ \vdots & \vdots & \ddots & \vdots \\ A_{z_m z_1}(\nu, \tau) & A_{z_m z_2}(\nu, \tau) & \dots & A_{z_m z_m}(\nu, \tau) \end{bmatrix}, \quad (4.5)$$

where the elements of the  $\mathbf{A}_{zz}(\nu, \tau)$  matrix are given by  $A_{z_i z_j}(\nu, \tau) = \text{IFT}_{t \rightarrow \nu} \{K_{z_i z_j}(t, \tau)\}$ . We may define a MTFR according to

$$\text{MTFR}_{zz}(t, f) = \text{IFT}_{\nu \rightarrow t} \left\{ \text{FT}_{\tau \rightarrow f} \{g(\nu, \tau) \mathbf{A}_{zz}(\nu, \tau)\} \right\}, \quad (4.6)$$

in which  $g(\nu, \tau)$  is a Doppler-lag kernel. Applying the kernels introduced in Chapter 1, we may now extend any 2-D TFR to the multisensor (i.e., multichannel) case.

In [122], multisensor TF signal processing is applied for blind source separation (BSS), direction of arrival (DOA) estimation, causality analysis, and analysis of EEG signals, where the authors demonstrated improved results with respect to conventional approaches. Therefore, it is expected that there is a lot of potential for multisensor/multichannel TF signal processing in a wide range of fields; one possible and interesting application could be multisensor-based jamming localization.

# Chapter 5

## Conclusion

The focus of this PhD dissertation was on time-frequency (TF) analysis and adaptive instantaneous frequency (IF) estimation. We started our study with the S-transform (ST) in Chapter 2, a linear TF transform with hybrid characteristics from the short-time Fourier transform (STFT) and the wavelet transform (WT), and with absolutely reference local phase information. We demonstrated that simple modifications in the existing definitions of the discrete ST (DST) can ensure exact signal recovery. The DST should be realized as in (2.12), using the discrete Fourier transform (DFT) of the observation window, to enable exact signal synthesis through the frequency inverse ST (FIST) (2.13), whereas the time inverse ST (TIST) in (2.17) eliminates any artifacts in the reconstructed signal. For TF filtering applications, the previous inverting methods produce different results. If the quality of filtering is to be evaluated by the similarity between the targeted DST and that of the reconstructed signal, we derived a novel algorithm for least squares signal synthesis from a modified DST (2.26). It requires the same computational complexity as the forward DST, and its effectiveness was demonstrated using numerical examples, where it outperformed other already available inverse ST (IST) methods.

The uniform sampling scheme of the original DST is inconsistent with Heisenberg's uncertainty principle and produces exaggerated amount of redundant TF information. This limitation poses the main barrier to more widespread use of the ST for applications that involve processing large datasets (e.g., audio and medical signal processing). The already existing dyadic DSTs sample the TF domain in a rigid uncontrollable manner, which makes them of limited effectiveness for processing signals of clear oscillatory behavior. Indeed, we showed that these transforms fail to provide appropriate results in various applications. Tackling the previous problems, we developed a DST with controlled TF sampling scheme, denoted controlled coverage DST (CC-DST). The CC-DST allows for finer frequency resolution, and thanks to its controllable coverage of the TF domain, it is effective in processing a wider class of signals, including those with no oscillatory behavior as well as signals arising from vibration phenomena. The CC-DST is exactly invertible through a simple method, because it employs asymmetric raised-cosine windows for spectral decom-

position, and it is numerically efficient with computational complexity close to that of the fast Fourier transform (FFT). We demonstrated its performance in different applications using real-world data. It is expected, therefore, that the CC-DST will be viable in many practical applications, allowing for a wider usage of efficient TF tools to process arbitrarily large datasets. Future research directions could involve developing methods for automating the proper selection of the CC-DST parameters based on the signal under analysis, and combining the CC-DST with compressive sensing (CS) and sparse reconstruction techniques to efficiently process signals suffering from missing observations.

The second part of this work (Chapter 3) was dedicated to designing adaptive methods for instantaneous frequency (IF) estimation based on TF representations (TFRs). We presented first a statistical analysis to derive the bias and the variance of the error of the IF estimate obtained through a linear TF transform that uses a Gaussian observation window with an arbitrarily defined standard deviation. Results of the analysis were then incorporated into a low-complexity three-step adaptation procedure to develop two fully automated linear TFRs. The first TFR is an adaptive STFT (ASTFT) with an observation window that minimizes the sum of the mean squared errors (MSEs) of the IF estimates at each time instant. While in the second TFR, the width of the observation window is adaptive over time and frequency, minimizing the MSE of the IF estimate at each location in the TF domain. The proposed algorithms were carefully evaluated and compared with state-of-the-art methods using synthetic as well as real-world signals. Our results demonstrated that, when dealing with noisy multicomponent signals whose IF ridges are well separated in the TF domain at moderate to high signal-to-noise ratios (SNRs), the developed TFRs may outperform many popular and more sophisticated methods, including those that are signal-adaptive, in terms of IF estimation. Furthermore, the presented algorithms are fully automated, thus avoiding a procedure for parameter optimization, whereas the good performance of most of their already existing counterparts necessitates an appropriate parameter tuning. On the other hand, the presented methods are not the best-performing TFRs in terms of energy contraction, meaning that other tech-

niques, such as the reassignment method and the adaptive directional TF distribution (ADTFD), may render TFRs that are sharper with improved readability. Overall, the derived methods provide attractive solutions for various applications in which the accuracy of IF estimation is of primary interest. Furthermore, the fact that these algorithms are based on linear TF transforms could be an additional reason for preferring them to their counterparts based on quadratic TFRs (QTFRs) in applications involving signal analysis and synthesis.

Then, the three-step adaptation procedure was applied with QTFRs. An estimate of the IF second-order derivative, provided by the continuous WT (CWT), was used to approximate the optimal window width of a wigner distribution (WD). At moderate to high SNRs, numerical results demonstrated superiority of the developed method over those based on the intersection-of-confidence-intervals (ICI) rule. Also, at lower SNRs and with application of the S-method (SM) instead of the WD for suppression of inner-terms, the proposed method provided more accurate IF estimates. Furthermore, the computational complexity of the developed method is significantly less than that of the ICI-based algorithm. Future research could involve 1. applying the presented method with high-order TFRs, 2. extending it for IF estimation of multicomponent signals using an algorithm for component extraction and tracking, and comparing it with the quasi-maximum-likelihood (QML) [103] estimator for polynomial phase signals.

Overall, the presence of cross-terms, which may surpass the auto-terms in terms of power, constitutes the dominant source of IF estimation error, especially at low SNRs. Therefore, the proposed adaptive QTFR is recommended for IF estimation of monocomponent signals, while when the analyzed signal is of multiple components, the presented methods based on linear TFRs are preferred.

# Appendix A

## Proof of Lemma 1

### A.1 Notation

For the sake of brevity and clarity, we will use the following notation:

$$n_{0R}(t) = \Re(n_0(t))$$

$$n_{0I}(t) = \Im(n_0(t))$$

$$W(t, \tau, f, \sigma) = \frac{1}{\sqrt{2\pi}\sigma(t, f)} \exp\left(\frac{-(t - \tau)^2}{2\sigma^2(t, f)}\right) e^{-i2\pi f\tau}$$

$$W^*(t, \tau, f, \sigma) = \frac{1}{\sqrt{2\pi}\sigma(t, f)} \exp\left(\frac{-(t - \tau)^2}{2\sigma^2(t, f)}\right) e^{i2\pi f\tau}$$

$$W_R(t, \tau, f, \sigma) = \frac{1}{\sqrt{2\pi}\sigma(t, f)} \exp\left(\frac{-(t - \tau)^2}{2\sigma^2(t, f)}\right) \cos(-i2\pi f\tau)$$

$$W_I(t, \tau, f, \sigma) = \frac{1}{\sqrt{2\pi}\sigma(t, f)} \exp\left(\frac{-(t - \tau)^2}{2\sigma^2(t, f)}\right) \sin(-i2\pi f\tau)$$

$$W'(t, \tau, f, \sigma) = \frac{\partial W}{\partial f}(t, \tau, f, \sigma)$$

$$W^{*'}(t, \tau, f, \sigma) = \frac{\partial W^*}{\partial f}(t, \tau, f, \sigma)$$

$$W'_R(t, \tau, f, \sigma) = \frac{\partial W_R}{\partial f}(t, \tau, f, \sigma)$$

$$W'_I(t, \tau, f, \sigma) = \frac{\partial W_I}{\partial f}(t, \tau, f, \sigma)$$

$$\text{GSFT}_{y_R}(t, f) = \Re(\text{GSFT}_y(t, f))$$

$$\text{GSFT}_{y_I}(t, f) = \Im(\text{GSFT}_y(t, f))$$

$$\text{GSFT}_y^*(t, f) = (\text{GSFT}_y(t, f))^*$$

$$\text{GSFT}_y^{*'}(t, f) = \frac{\partial \text{GSFT}_y^*}{\partial f}(t, f)$$

$$\text{GSFT}'_{y_R}(t, f) = \frac{\partial \text{GSFT}_{y_R}}{\partial f}(t, f)$$

$$\text{GSFT}'_{y_I}(t, f) = \frac{\partial \text{GSFT}_{y_I}}{\partial f}(t, f)$$

## A.2 Noise properties

The noise  $n_0(t)$  is complex-valued AWGN with i.i.d. real and imaginary parts, thus satisfying the following properties [18]:

$$E[n_0(t_1)n_0(t_2)] = 0, \quad (\text{A.1})$$

$$E[n_0(t_1)n_0^*(t_2)] = N_0\delta(t_2 - t_1), \quad (\text{A.2})$$

$$E[n_0(t_1)n_0^*(t_2)n_0(t_3)] = 0, \quad (\text{A.3})$$

$$E[n_0(t_1)n_0^*(t_2)n_0(t_3)n_0^*(t_4)] = N_0^2[\delta(t_2 - t_1)\delta(t_4 - t_3) + \delta(t_4 - t_1)\delta(t_3 - t_2)]. \quad (\text{A.4})$$

## A.3 Proof

*Proof.* We start from

$$\begin{aligned} E \left[ \frac{\partial (\text{GSFT}_{x_R} \text{GSFT}_{n_0R})}{\partial f}(t, f) \right] &= \text{GSFT}'_{x_R}(t, f) E[\text{GSFT}_{n_0R}(t, f)] \\ &+ E[\text{GSFT}'_{n_0R}(t, f)] \text{GSFT}_{x_R}(t, f), \end{aligned} \quad (\text{A.5})$$

which results in<sup>1</sup>

$$\begin{aligned} E [\text{GSFT}_{n_{0R}}(t, f)] &= \int E [n_{0R}(\tau)] W_R(t, \tau, f, \sigma) d\tau - \int E [n_{0I}(\tau)] W_I(t, \tau, f, \sigma) d\tau \\ &= 0. \end{aligned} \tag{A.6}$$

Similarly, we can prove that

$$\begin{aligned} E [\text{GSFT}'_{n_{0R}}(t, f)] &= \int E [n_{0R}(\tau)] W'_R(t, \tau, f, \sigma) d\tau - \int E [n_{0I}(\tau)] \\ &\quad \times W'_I(t, \tau, f, \sigma) d\tau = 0. \end{aligned} \tag{A.7}$$

By substituting (A.6) and (A.7) into (A.5), we get that that

$$E \left[ \frac{\partial (\text{GSFT}_{xR} \text{GSFT}_{n_{0R}})}{\partial f}(t, f) \right] = 0. \tag{A.8}$$

Following the same previous steps, we prove that

$$E \left[ \frac{\partial (\text{GSFT}_{xI} \text{GSFT}_{n_{0I}})}{\partial f}(t, f) \right] = 0. \tag{A.9}$$

By the results given in (A.8) and (A.9), we arrive at

$$E \left\{ \frac{\partial S_n}{\partial f}(t, f_i(t)) \right\} = E \left[ \frac{\partial |\text{GSFT}_{n_0}|^2}{\partial f}(t, f_i(t)) \right]. \tag{A.10}$$

$$\begin{aligned} E \left[ \frac{\partial |\text{GSFT}_{n_0}|^2}{\partial f}(t, f) \right] &= E [\text{GSFT}'_{n_0}(t, f) \text{GSFT}_{n_0}^*(t, f)] + E [\text{GSFT}_{n_0}'^*(t, f) \\ &\quad \times \text{GSFT}_{n_0}(t, f)], \end{aligned} \tag{A.11}$$

---

<sup>1</sup>  $f \equiv \int_{-\infty}^{\infty}$  unless otherwise stated.

which results in

$$\begin{aligned}
& E [\text{GSFT}'_{n_0}(t, f) \text{GSFT}^*_{n_0}(t, f)] \\
&= \int \int E [n_0(\tau_1)n_0^*(\tau_2)] W'(t, \tau_1, f, \sigma)W^*(t, \tau_2, f, \sigma) d\tau_1 d\tau_2 \quad (\text{A.12}) \\
&= N_0 \int W'(t, \tau, f, \sigma)W^*(t, \tau, f, \sigma) d\tau.
\end{aligned}$$

Similarly,

$$E [\text{GSFT}^*_{n_0}(t, f) \text{GSFT}_{n_0}(t, f)] = N_0 \int W^{*'}(t, \tau, f, \sigma)W(t, \tau, f, \sigma) d\tau. \quad (\text{A.13})$$

The integrals (A.12) and (A.13) can be computed easily, by substituting their values into (A.11), and after recalling (A.10), we prove Lemma 1.  $\square$

# Appendix B

## Proof of Lemma 2

### B.1 Proof

*Proof.* We will use the same notation provided in Section A.1, and the properties of the complexed-valued AWGN are given in Section A.2.

By taking the first-order frequency derivative of (3.26) at  $f = f_i(t)$ , squaring the result, then taking its expected value, we obtain

$$\begin{aligned}
E \left[ \left( \frac{\partial S_n}{\partial f}(t, f_i(t)) \right)^2 \right] &= E \left\{ \left( \frac{\partial |\text{GSFT}_{n_0}|^2}{\partial f}(t, f_i(t)) \right)^2 \right\} + 4E \left\{ \left[ \frac{\partial (\text{GSFT}_{x_R} \text{GSFT}_{n_0 R})}{\partial f}(t, f_i(t)) \right]^2 \right\} \\
&+ 4E \left\{ \left[ \frac{\partial (\text{GSFT}_{x_I} \text{GSFT}_{n_0 I})}{\partial f}(t, f_i(t)) \right]^2 \right\} + 8E \left\{ \frac{\partial (\text{GSFT}_{x_R} \text{GSFT}_{n_0 R})}{\partial f}(t, f_i(t)) \right. \\
&\times \left. \frac{\partial (\text{GSFT}_{x_I} \text{GSFT}_{n_0 I})}{\partial f}(t, f_i(t)) \right\} + 4E \left\{ \frac{\partial |\text{GSFT}_{n_0}|^2}{\partial f}(t, f_i(t)) \frac{\partial (\text{GSFT}_{x_R} \text{GSFT}_{n_0 R})}{\partial f}(t, f_i(t)) \right\} \\
&+ 4E \left\{ \frac{\partial |\text{GSFT}_{n_0}|^2}{\partial f}(t, f_i(t)) \frac{\partial (\text{GSFT}_{x_I} \text{GSFT}_{n_0 I})}{\partial f}(t, f_i(t)) \right\}.
\end{aligned} \tag{B.1}$$

In what follows, all the noise-dependent terms in (B.1) will be evaluated<sup>1</sup>.

---

<sup>1</sup>  $f \equiv \int_{-\infty}^{+\infty}$  unless otherwise stated.

**Term 1:**

$$\begin{aligned}
E \left[ \left( \frac{\partial |\text{GSFT}_{n_0}|^2}{\partial f}(t, f) \right)^2 \right] &= E \left\{ \underbrace{[\text{GSFT}'_{n_0}(t, f) \text{GSFT}^*_{n_0}(t, f)]^2}_{I_1} \right\} \\
&+ E \left\{ \underbrace{[\text{GSFT}^*{}'_{n_0}(t, f) \text{GSFT}_{n_0}(t, f)]^2}_{I_2} \right\} \\
&+ 2E \left[ \underbrace{\text{GSFT}_{n_0}(t, f) \text{GSFT}^*_{n_0}(t, f) \text{GSFT}'_{n_0}(t, f) \text{GSFT}^*{}'_{n_0}(t, f)}_{I_3} \right].
\end{aligned} \tag{B.2}$$

$$\begin{aligned}
E[I_1] &= \int \int \int \int E [n_0(\tau_1) n_0^*(\tau_2) n_0(\tau_3) n_0^*(\tau_4)] W'(t, \tau_1, f, \sigma) \\
&\quad \times W^*(t, \tau_2, f, \sigma) W'(t, \tau_3, f, \sigma) W^*(t, \tau_4, f, \sigma) d\tau_1 d\tau_2 d\tau_3 d\tau_4.
\end{aligned}$$

By (A.4),  $E[I_1]$  is given by the sum of the following integrals:  $E[I_1] = I_{1,1} + I_{1,2} + I_{1,3}$  where

$$\begin{aligned}
I_{1,1} &= N_0^2 \left[ \int W'(t, \tau, f, \sigma) W^*(t, \tau, f, \sigma) d\tau \right]^2, \\
I_{1,2} &= I_{1,1}, \\
I_{1,3} &= 2N_0^2 \int [W'(t, \tau, f, \sigma) W^*(t, \tau, f, \sigma)]^2 d\tau.
\end{aligned}$$

Following the same previous steps, we prove that  $E[I_2] = I_{2,1} + I_{2,2} + I_{2,3}$  where

$$\begin{aligned}
I_{2,1} &= N_0^2 \left[ \int W(t, \tau, f, \sigma) W^{*'}(t, \tau, f, \sigma) d\tau \right]^2, \\
I_{2,2} &= I_{2,1}, \\
I_{2,3} &= 2N_0^2 \int [W(t, \tau, f, \sigma) W^{*'}(t, \tau, f, \sigma)]^2 d\tau.
\end{aligned}$$

and  $E[I_3] = I_{3,1} + I_{3,2} + I_{3,3}$  where

$$\begin{aligned}
I_{3,1} &= N_0^2 \int W'(t, \tau, f, \sigma) W^*(t, \tau, f, \sigma) d\tau \int W^{*'}(t, \tau, f, \sigma) W(t, \tau, f, \sigma) d\tau, \\
I_{3,2} &= N_0^2 \int W'(t, \tau, f, \sigma) W^{*'}(t, \tau, f, \sigma) d\tau \int W^*(t, \tau, f, \sigma) W(t, \tau, f, \sigma) d\tau,
\end{aligned}$$

$$I_{3,3} = 2N_0^2 \int W(t, \tau, f, \sigma) W'(t, \tau, f, \sigma) W^*(t, \tau, f, \sigma) W^{*'}(t, \tau, f, \sigma) d\tau.$$

The previous integrals can be computed easily, by substituting their values into (A.10), we obtain

$$E \left[ \left( \frac{\partial |\text{GSFT}_{n_0}|^2}{\partial f}(t, f) \right)^2 \right] = N_0^2 \left[ \frac{(12\sqrt{\pi}\sigma(t, f) + 11\sqrt{2}) \sigma'^2(t, f)}{16\pi^{\frac{3}{2}}\sigma^5(t, f)} + \pi \right].$$

## Term 2

$$\begin{aligned} & E \left\{ \left[ \frac{\partial (\text{GSFT}_{x_R} \text{GSFT}_{n_0 R})}{\partial f}(t, f) \right]^2 \right\} = \text{GSFT}'^2_{x_R}(t, f) \\ & \times E [\text{GSFT}^2_{n_0 R}(t, f)] + E [\text{GSFT}'^2_{n_0 R}(t, f)] \\ & \times \text{GSFT}^2_{x_R}(t, f) + 2E [\text{GSFT}_{n_0 R}(t, f) \text{GSFT}'_{n_0 R}(t, f)] \\ & \times \text{GSFT}_{x_R}(t, f) \text{GSFT}'_{x_R}(t, f). \end{aligned} \quad (\text{B.3})$$

$$\begin{aligned} \text{GSFT}^2_{n_0 R}(t, f) &= \int \int [n_{0R}(\tau_1) W_R(t, \tau_1, f, \sigma) - n_{0I}(\tau_1) W_I(t, \tau_1, f, \sigma)] \\ & \times [n_{0R}(\tau_2) W_R(t, \tau_2, f, \sigma) - n_{0I}(\tau_2) W_I(t, \tau_2, f, \sigma)] \\ & d\tau_1 d\tau_2. \end{aligned}$$

By expanding the above expression then taking the expected value of the result, we obtain

$$\begin{aligned} E [\text{GSFT}^2_{n_0 R}(t, f)] &= \frac{N_0}{2} \int W_R^2(t, \tau, f, \sigma) + W_I^2(t, \tau, f, \sigma) d\tau \\ &= \frac{N_0}{4\sqrt{\pi}\sigma(t, f)}. \end{aligned} \quad (\text{B.4})$$

Similarly, we may prove that

$$E [\text{GSFT}^2_{n_0 I}(t, f)] = \frac{N_0}{4\sqrt{\pi}\sigma(t, f)}.$$

Following the same previous steps, we obtain:

$$\begin{aligned}
E \left[ \text{GSFT}'^2_{n_0R}(t, f) \right] &= E \left[ \text{GSFT}'^2_{n_0I}(t, f) \right] = \frac{N_0}{2} \int W_R'^2(t, \tau, f, \sigma) + W_I'^2(t, \tau, f, \sigma) d\tau \\
&= \frac{N_0 \left[ 8\pi^2 \sigma^2(t, f) (\sigma^2(t, f) + 2t^2) + 3\sigma'^2(t, f) \right]}{16\sqrt{\pi}\sigma^3(t, f)},
\end{aligned} \tag{B.5}$$

and

$$\begin{aligned}
E \left[ \text{GSFT}_{n_0R}(t, f) \text{GSFT}'_{n_0R}(t, f) \right] &= E \left[ \text{GSFT}_{n_0I}(t, f) \text{GSFT}'_{n_0I}(t, f) \right] \\
&= \frac{N_0}{2} \int W_R(t, \tau, f, \sigma) W_R'(t, \tau, f, \sigma) + W_I(t, \tau, f, \sigma) \\
&\quad \times W_I'(t, \tau, f, \sigma) d\tau = -\frac{N_0 \sigma'(t, f)}{8\sqrt{\pi}\sigma^2(t, f)}.
\end{aligned} \tag{B.6}$$

### Term 3

$$\begin{aligned}
&E \left[ \frac{\partial \left( \text{GSFT}_{xR} \text{GSFT}_{n_0R} \right)}{\partial f}(t, f) \frac{\partial \left( \text{GSFT}_{xI} \text{GSFT}_{n_0I} \right)}{\partial f}(t, f) \right] = \\
&\text{GSFT}'_{xR}(t, f) \text{GSFT}'_{xI}(t, f) E \left[ \text{GSFT}_{n_0R}(t, f) \text{GSFT}_{n_0I}(t, f) \right] \\
&+ \text{GSFT}'_{xR}(t, f) \text{GSFT}_{xI}(t, f) E \left[ \text{GSFT}'_{n_0I}(t, f) \text{GSFT}_{n_0R}(t, f) \right] \\
&+ \text{GSFT}'_{xI}(t, f) \text{GSFT}_{xR}(t, f) E \left[ \text{GSFT}'_{n_0R}(t, f) \text{GSFT}_{n_0I}(t, f) \right] \\
&+ \text{GSFT}_{xR}(t, f) \text{GSFT}_{xI}(t, f) E \left[ \text{GSFT}'_{n_0R}(t, f) \text{GSFT}'_{n_0I}(t, f) \right].
\end{aligned} \tag{B.7}$$

It can be easily shown that:

$$\begin{aligned}
E \left[ \text{GSFT}_{n_0R}(t, f) \text{GSFT}_{n_0I}(t, f) \right] &= 0, \\
E \left[ \text{GSFT}'_{n_0R}(t, f) \text{GSFT}'_{n_0I}(t, f) \right] &= 0,
\end{aligned} \tag{B.8}$$

$$\begin{aligned}
& E \left[ \text{GSFT}'_{n_{0I}}(t, f) \text{GSFT}_{n_{0R}}(t, f) \right] = \\
& \frac{N_0}{2} \int W_R(t, \tau, f, \sigma) W_I'(t, \tau, f, \sigma) - W_I(t, \tau, f, \sigma) \\
& \quad \times W_R'(t, \tau, f, \sigma) d\tau \\
& = -\frac{N_0 \sqrt{\pi} t}{2\sigma(t, f)},
\end{aligned} \tag{B.9}$$

and

$$\begin{aligned}
& E \left[ \text{GSFT}'_{n_{0R}}(t, f) \text{GSFT}_{n_{0I}}(t, f) \right] = \\
& = \frac{N_0}{2} \int W_R'(t, \tau, f, \sigma) W_I(t, \tau, f, \sigma) - W_R(t, \tau, f, \sigma) \\
& \quad \times W_I'(t, \tau, f, \sigma) d\tau. \\
& = \frac{N_0 \sqrt{\pi} t}{2\sigma(t, f)}.
\end{aligned} \tag{B.10}$$

#### Term 4

$$\begin{aligned}
& E \left\{ \frac{\partial |\text{GSFT}_{n_0}|^2}{\partial f}(t, f) \frac{\partial (\text{GSFT}_{xR} \text{GSFT}_{n_{0R}})}{\partial f}(t, f) \right\} = \\
& 2\text{GSFT}_{xR}(t, f) E \left[ \text{GSFT}_{n_{0R}}(t, f) \text{GSFT}'_{n_{0R}}{}^2(t, f) \right] \\
& + 2\text{GSFT}'_{xR}(t, f) E \left[ \text{GSFT}_{n_{0R}}^2(t, f) \text{GSFT}'_{n_{0R}}(t, f) \right] \\
& + 2\text{GSFT}_{xR}(t, f) E \left[ \text{GSFT}'_{n_{0R}}(t, f) \text{GSFT}_{n_{0I}}(t, f) \right. \\
& \quad \times \text{GSFT}'_{n_{0I}}(t, f) \left. \right] + 2\text{GSFT}'_{xR}(t, f) E \left[ \text{GSFT}_{n_{0R}}(t, f) \right. \\
& \quad \times \text{GSFT}_{n_{0I}}(t, f) \text{GSFT}'_{n_{0I}}(t, f) \left. \right].
\end{aligned}$$

Using the AWGN property given in (A.3), we may prove that

$$\begin{aligned}
& E \left[ \text{GSFT}_{n_{0R}}(t, f) \text{GSFT}'_{n_{0R}}{}^2(t, f) \right] = 0, \\
& E \left[ \text{GSFT}_{n_{0R}}^2(t, f) \text{GSFT}'_{n_{0R}}(t, f) \right] = 0, \\
& E \left[ \text{GSFT}'_{n_{0R}}(t, f) \text{GSFT}_{n_{0I}}(t, f) \text{GSFT}'_{n_{0I}}(t, f) \right] = 0, \\
& E \left[ \text{GSFT}_{n_{0R}}(t, f) \text{GSFT}_{n_{0I}}(t, f) \text{GSFT}'_{n_{0I}}(t, f) \right] = 0,
\end{aligned}$$

which results in

$$E \left\{ \frac{\partial |\text{GSFT}_{n_0}|^2}{\partial f}(t, f) \frac{\partial \left( \text{GSFT}_{x_R} \text{GSFT}_{n_0 R} \right)}{\partial f}(t, f) \right\} = 0. \quad (\text{B.11})$$

Similarly, we may prove that:

$$E \left\{ \frac{\partial |\text{GSFT}_{n_0}|^2}{\partial f}(t, f) \frac{\partial \left( \text{GSFT}_{x_I} \text{GSFT}_{n_0 I} \right)}{\partial f}(t, f) \right\} = 0. \quad (\text{B.12})$$

At this point, all the noise-dependent terms in (B.1) are evaluated. To complete the computation, the real and imaginary parts of  $\text{GSFT}_x(t, f_i(t))$ , together with their first-order frequency derivatives should evaluate and substituted into (B.1). By complex expanding (7),  $\text{GSFT}_{x_R}(t, f_i(t))$  and  $\text{GSFT}_{x_I}(t, f_i(t))$  can be evaluated giving the following results:

$$\begin{aligned} \text{GSFT}_{x_R}(t, f_i(t)) &= \frac{A \cos(a_0 - \theta_0)}{\sqrt[4]{1 + [2\pi f'_i(t)\sigma^2(t, f_i(t))]^2}} \\ \text{GSFT}_{x_I}(t, f_i(t)) &= \frac{A \sin(a_0 - \theta_0)}{\sqrt[4]{1 + [2\pi f'_i(t)\sigma^2(t, f_i(t))]^2}} \end{aligned}$$

where

$$\theta_0 = \frac{1}{2} \tan^{-1} [-2\pi f'_i(t)\sigma^2(t, f_i(t))].$$

The derivatives  $\text{GSFT}'_{x_R}(t, f_i(t))$  and  $\text{GSFT}'_{x_I}(t, f_i(t))$  can also be computed using (7), by substituting their values into (B.1) and after some algebraic simplifications we prove Lemma 2.  $\square$

# Bibliography

- [1] E. Sejdic, I. Djurović, and J. Jiang, “A window width optimized S-transform,” *EURASIP Journal on Advances in Signal Processing*, vol. 2008, p. 59, 2008.
- [2] S. C. Pei and P. W. Wang, “Energy concentration enhancement using window width optimization in S transform,” in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 4106–4109.
- [3] R. G. Stockwell, “A basis for efficient representation of the S-transform,” *Digital Signal Process.*, vol. 17, no. 1, 2007.
- [4] R. A. Brown, M. L. Lauzon, and R. Frayne, “A general description of linear time-frequency transforms and formulation of a fast, invertible transform that samples the continuous S-transform spectrum nonredundantly,” *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 281–290, Jan. 2010.
- [5] L. Stanković, “A measure of some time-frequency distributions concentration,” *Signal Process.*, vol. 81, no. 3, pp. 621–631, 2001.
- [6] P. C. Pei and S. G. Huang, “STFT with adaptive window width based on the chirp rate,” *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4065–4080, 2012.
- [7] M. A. Awal and B. Boashash, “An automatic fast optimization of quadratic time-frequency distribution using the hybrid genetic algorithm,” *Signal Process.*, vol. 131, pp. 134–142, 2017.
- [8] J. Allen, “Short term spectral analysis, synthesis, and modification by discrete Fourier transform,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 25, no. 3, pp. 235–238, Jun 1977.
- [9] M. Portnoff, “Implementation of the digital phase vocoder using the fast Fourier transform,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 3, pp. 243–248, 1976.
- [10] R. G. Stockwell, L. Mansinha, and R. P. Lowe, “Localization of the complex spectrum: the S transform,” *IEEE Trans. on Signal Process.*, vol. 44, no. 4, pp. 998–1001, Apr. 1996.

- [11] C. R. Pinnegar and L. Mansinha, "The S-transform with windows of arbitrary and varying shape," *Geophysics*, vol. 68, no. 1, pp. 381–385, 2003.
- [12] L. Stankovic and S. Stankovic, "An analysis of instantaneous frequency representation using time-frequency distributions-generalized Wigner distribution," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 549–552, 1995.
- [13] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. ii. algorithms and applications," *Proc. IEEE*, vol. 80, no. 4, pp. 540–568, 1992.
- [14] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, and M. Daković, "Instantaneous frequency in time-frequency analysis: Enhanced concepts and performance of estimation algorithms," *Digital Signal Process.*, vol. 35, pp. 1–13, 2014.
- [15] B. Boashash, *Time Frequency Signal Analysis and Processing: A Comprehensive Reference*, 2nd ed. London, UK: Elsevier, 2003.
- [16] L. Cohen, "Time-frequency distributions-a review," *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, Jul. 1989.
- [17] T. Claasen and W. F. G. Mecklenbrauker, "The Wigner distribution-a tool for time-frequency signal analysis," *Philips J. Research*, vol. 35, pp. 276–350, 1980.
- [18] K. M. Wong and Q. Jin, "Estimation of the time-varying frequency of a signal: the Cramer-Rao bound and the application of Wigner distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 3, pp. 519–536, Mar 1990.
- [19] P. Rao and F. Taylor, "Estimation of instantaneous frequency using the discrete Wigner distribution," *Electron. Lett.*, vol. 26, no. 4, pp. 246–248, 1990.
- [20] E. Sejdić, I. Djurović, and J. Jiang, "Time-frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Process.*, vol. 19, no. 1, pp. 153–183, 2009.
- [21] L. Stankovic and S. Stankovic, "An analysis of instantaneous frequency representation using time-frequency distributions-generalized Wigner distribution," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 549–552, 1995.
- [22] B. Boashash and S. Ouelha, "An improved design of high-resolution quadratic time–frequency distributions for the analysis of nonstationary multicomponent signals using directional compact kernels," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2701–2713, 2017.
- [23] B. Boashash, N. A. Khan, and T. Ben-Jabeur, "Time–frequency features for pattern recognition using high-resolution TFDs: A tutorial review," *Digital Signal Process.*, vol. 40, pp. 1–30, 2015.

- [24] N. A. Khan and B. Boashash, "Multi-component instantaneous frequency estimation using locally adaptive directional time frequency distributions," *Int. J. Adapt. Control Signal Process.*, vol. 30, no. 3, pp. 429–442, 2016.
- [25] B. Ristic and B. Boashash, "Kernel design for time-frequency signal analysis using the Radon transform," *IEEE Trans. Signal Process.*, vol. 41, no. 5, pp. 1996–2008, 1993.
- [26] B. Boashash, G. Azemi, and J. M. O'Toole, "Time-frequency processing of non-stationary signals: Advanced TFD design to aid diagnosis with highlights from medical applications," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 108–119, Nov 2013.
- [27] B. Boashash and P. Black, "An efficient real-time implementation of the Wigner-Ville distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 11, pp. 1611–1618, 1987.
- [28] H. I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 6, pp. 862–871, 1989.
- [29] M. Born and P. Jordan, "Zur quantenmechanik," *Zeitschrift für Physik*, vol. 34, no. 1, pp. 858–888, 1925.
- [30] M. Abed, A. Belouchrani, M. Cheriet, and B. Boashash, "Time-frequency distributions based on compact support kernels: Properties and performance evaluation," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2814–2827, June 2012.
- [31] L. Stankovic, "A method for time-frequency analysis," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 225–229, 1994.
- [32] —, "A method for improved distribution concentration in the time-frequency analysis of multicomponent signals using the L-Wigner distribution," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1262–1268, 1995.
- [33] —, "Time-frequency distributions with complex argument," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 475–486, 2002.
- [34] M. Morelande, B. Senadji, and B. Boashash, "Complex-lag polynomial Wigner-Ville distribution," in *TENCON'97. IEEE Region 10 Annual Conference. Speech and Image Technologies for Computing and Telecommunications., Proceedings of IEEE*, vol. 1. IEEE, 1997, pp. 43–46.
- [35] A. Omidvarnia, G. Azemi, J. M. O'Toole, and B. Boashash, "Robust estimation of highly-varying nonlinear instantaneous frequency of monocomponent signals using a lower-order complex-time distribution," *Signal Process.*, vol. 93, no. 11, pp. 3251–3260, 2013.

- [36] K. Kodera, R. Gendrin, and C. d. Villedary, "Analysis of time-varying signals with small bt values," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 26, no. 1, pp. 64–76, 1978.
- [37] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, 1995.
- [38] B. Boashash, "Note on the use of the Wigner distribution for time-frequency signal analysis," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 9, pp. 1518–1521, 1988.
- [39] J. M. O'Toole, M. Mesbah, and B. Boashash, "Algorithms for discrete quadratic time–frequency distributions," *WSEAS Trans. Signal Processing, invited paper*, vol. 4, no. 5, pp. 320–329, 2008.
- [40] J. O'Toole, M. Mesbah, and B. Boashash, "A discrete time and frequency Wigner–Ville distribution: properties and implementation," in *Proc. Int. Conf. on Digital Signal Processing and Comm. Systems*, vol. 400, 2005, pp. 19–21.
- [41] J. O. Toole and B. Boashash, "Fast and memory-efficient algorithms for computing quadratic time–frequency distributions," *Appl. Comput. Harmon. Anal.*, vol. 35, no. 2, pp. 350–358, 2013.
- [42] B. Boashash and V. Sucic, "Resolution measure criteria for the objective assessment of the performance of quadratic time-frequency distributions," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1253–1263, 2003.
- [43] D. L. Jones and T. W. Parks, "A high resolution data-adaptive time-frequency representation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 12, pp. 2127–2135, 1990.
- [44] P. Flandrin, R. G. Baraniuk, and O. Michel, "Time-frequency complexity and information," in *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, vol. 3. IEEE, 1994, pp. III–329.
- [45] V. Sucic, N. Saulig, and B. Boashash, "Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency rényi entropy," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, no. 1, p. 125, 2011.
- [46] T.-H. Sang and W. J. Williams, "Renyi information and signal-dependent optimal kernel design," in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, vol. 2. IEEE, 1995, pp. 997–1000.
- [47] M. J. Bastiaans, T. Alieva, and L. Stankovic, "On rotated time-frequency kernels," *IEEE Signal Process. Lett.*, vol. 9, no. 11, pp. 378–381, 2002.

- [48] J. C. Wood and D. T. Barry, "Radon transformation of time-frequency distributions for analysis of multicomponent signals," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3166–3177, 1994.
- [49] V. Katkovnik, "Local polynomial periodograms for signals with the time-varying frequency and amplitude," in *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*, vol. 3. IEEE, 1996, pp. 1399–1402.
- [50] X. Li, G. Bi, S. Stankovic, and A. M. Zoubir, "Local polynomial Fourier transform: A review on recent developments and applications," *Signal Process.*, vol. 91, no. 6, pp. 1370–1393, 2011.
- [51] I. Shafi, J. Ahmad, S. I. Shah, and F. M. Kashif, "Computing deblurred time-frequency distributions using artificial neural networks," *Circuits, Systems & Signal Processing*, vol. 27, no. 3, pp. 277–294, 2008.
- [52] R. G. Baraniuk, "Wavelet soft-thresholding of time-frequency representations," in *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, vol. 1. IEEE, 1994, pp. 71–74.
- [53] H. Hassanpour, "A time–frequency approach for noise reduction," *Digital Signal Process.*, vol. 18, no. 5, pp. 728–738, 2008.
- [54] S. Gómez, V. Naranjo, and R. Miralles, "Removing interference components in time–frequency representations using morphological operators," *J. Visual Commun. Image Represent.*, vol. 22, no. 5, pp. 401–410, 2011.
- [55] N. A. Khan and M. Sandsten, "Time–frequency image enhancement based on interference suppression in Wigner–Ville distribution," *Signal Process.*, vol. 127, pp. 80–85, 2016.
- [56] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep 1990.
- [57] J. B. Allen and L. R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis," *Proc. IEEE*, vol. 65, no. 11, pp. 1558–1564, Nov. 1977.
- [58] R. Brown, M. Zlatescu, A. Sijben, G. Roldan, J. Easaw, P. Forsyth, I. Parney, R. Sevick, E. Yan, D. Demetrick *et al.*, "The use of magnetic resonance imaging to noninvasively detect genetic signatures in oligodendroglioma," *Clin. Cancer Res.*, vol. 14, no. 8, pp. 2357–2362, 2008.
- [59] G. Livanos, N. Ranganathan, and J. Jiang, "Heart sound analysis using the S transform," in *Computers in Cardiology 2000*. IEEE, 2000, pp. 587–590.
- [60] M. Schimmel and J. Gallart, "Frequency-dependent phase coherence for noise suppression in seismic array data," *Journal of Geophysical Research: Solid Earth*, vol. 112, no. B4, 2007.

- [61] S. Mishra, C. N. Bhende, and B. K. Panigrahi, "Detection and classification of power quality disturbances using S-transform and probabilistic neural network," *IEEE Trans. Power Del.*, vol. 23, no. 1, pp. 280–287, 2008.
- [62] P. K. Dash, B. K. Panigrahi, and G. Panda, "Power quality analysis using S-transform," *IEEE Trans. Power Del.*, vol. 18, no. 2, pp. 406–411, April 2003.
- [63] L. Mansinha, R. Stockwell, and R. Lowe, "Pattern analysis with two-dimensional spectral localisation: Applications of two-dimensional s transforms," *Physica A*, vol. 239, no. 1-3, pp. 286–295, 1997.
- [64] S. Assous, A. Humeau, M. Tartas, P. Abraham, and J.-P. L'Huillier, "S-transform applied to laser doppler flowmetry reactive hyperemia signals," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 6, pp. 1032–1037, June 2006.
- [65] E. Sejdic, L. Stanković, M. Daković, and J. Jiang, "Instantaneous frequency estimation using the S-transform," *IEEE Signal Process. Lett.*, vol. 15, pp. 309–312, 2008.
- [66] C. R. Pinnegar and L. Mansinha, "The bi-gaussian S-transform," *SIAM J. Sci. Comput.*, vol. 24, no. 5, pp. 1678–1692, 2003.
- [67] K. Kazemi, M. Amirian, and M. J. Dehghani, "The s-transform using a new window to improve frequency and time resolutions," *Signal, image and Video processing*, vol. 8, no. 3, pp. 533–541, 2014.
- [68] E. Sejdic, I. Djurovic, and J. Jiang, "S-transform with frequency dependent kaiser window," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3. IEEE, 2007, pp. III–1165.
- [69] Z. Zidelmal, H. Hamil, A. Moukadem, A. Amirou, and D. Ould-Abdeslam, "S-transform based on compact support kernel," *Digital Signal Process.*, vol. 62, pp. 137–149, 2017.
- [70] M. Schimmel and J. Gallart, "The inverse S-transform in filters with time-frequency localization," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4417–4422, Nov 2005.
- [71] C. Simon, S. Ventosa, M. Schimmel, A. Heldring, J. J. Danobeitia, J. Gallart, and A. Manuel, "The S-transform and its inverses: Side effects of discretizing and filtering," *IEEE Trans. on Signal Process.*, vol. 55, no. 10, pp. 4928–4937, Oct. 2007.
- [72] S. C. Pei and P. W. Wang, "Novel inverse S transform with equalization filter," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3858–3868, Oct 2009.
- [73] —, "New discrete inverse s transform with least square error in filtering," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 2989–2992.

- [74] Y. Abdoush, G. Pojani, and G. E. Corazza, “Least-squares signal synthesis from modified S-transform,” in *2018 IEEE Statistical Signal Processing Workshop (SSP)*, June 2018, pp. 423–427.
- [75] M. Biswal and P. K. Dash, “Detection and characterization of multiple power quality disturbances with a fast S-transform and decision tree based classifier,” *Digital Signal Process.*, vol. 23, no. 4, pp. 1071–1083, 2013.
- [76] J. Youngberg, “Rate/pitch modification of speech using the constant-Q transform,” in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP’79.*, vol. 4. IEEE, 1979, pp. 748–751.
- [77] J. Laroche and M. Dolson, “Improved phase vocoder time-scale modification of audio,” *IEEE Trans. Speech Audio Process.*, vol. 7, no. 3, pp. 323–332, 1999.
- [78] I. Bayram and I. W. Selesnick, “A dual-tree rational-dilation complex wavelet transform,” *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6251–6256, 2011.
- [79] S. Ari, M. K. Das, and A. Chacko, “ECG signal enhancement using S-transform,” *Comput. Biol. Med.*, vol. 43, no. 6, pp. 649–660, 2013.
- [80] B. Boashash and S. Ouelha, “Automatic signal abnormality detection using time-frequency features and machine learning: A newborn EEG seizure case study,” *Knowledge-Based Systems*, vol. 106, pp. 38–50, 2016.
- [81] N. A. Khan and B. Boashash, “Instantaneous frequency estimation of multi-component nonstationary signals using multiview time-frequency distributions based on the adaptive fractional spectrogram,” *IEEE Signal Process. Lett.*, vol. 20, no. 2, pp. 157–160, 2013.
- [82] M. Mohammadi, A. A. Pouyan, and N. A. Khan, “A highly adaptive directional time–frequency distribution,” *Signal, Image and Video Processing*, vol. 10, no. 7, pp. 1369–1376, 2016.
- [83] V. N. Ivanović, M. Daković, and L. Stanković, “Performance of quadratic time-frequency distributions as instantaneous frequency estimators,” *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 77–89, Jan 2003.
- [84] V. Katkovnik and L. Stanković, “Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length,” *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2315–2325, Sep 1998.
- [85] J. Zhong and Y. Huang, “Time-frequency representation based on an adaptive short-time Fourier transform,” *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5118–5128, 2010.
- [86] L. Rankine, M. Mesbah, and B. Boashash, “IF estimation for multicomponent signals using image processing techniques in the time–frequency domain,” *Signal Process.*, vol. 87, no. 6, pp. 1234–1250, 2007.

- [87] S. Dong, G. Azemi, B. Lingwood, P. B. Colditz, and B. Boashash, "Performance evaluation of multi-component instantaneous frequency estimation techniques for heart rate variability analysis," in *Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on*. IEEE, 2012, pp. 1211–1216.
- [88] P. O'Shea, "On refining polynomial phase signal parameter estimates," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 3, pp. 978–987, July 2010.
- [89] I. Djurovic and L. Stankoic, "Influence of high noise on the instantaneous frequency estimation using quadratic time-frequency distributions," *IEEE Signal Process. Lett.*, vol. 7, no. 11, pp. 317–319, 2000.
- [90] L. J. Stankovic and V. Katkovnik, "Algorithm for the instantaneous frequency estimation using time-frequency distributions with adaptive window width," in *Signal Processing Conference (EUSIPCO 1998), 9th European*. IEEE, 1998, pp. 1–4.
- [91] L. Stankovic, "Performance analysis of the adaptive algorithm for bias-to-variance tradeoff," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1228–1234, 2004.
- [92] V. Katkovnik and L. Stanković, "Periodogram with varying and data-driven window length," *Signal Process.*, vol. 67, no. 3, pp. 345–358, 1998.
- [93] L. Stankovic and V. Katkovnik, "Instantaneous frequency estimation using higher order L-Wigner distributions with data-driven order and window length," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 302–311, 2000.
- [94] Z. M. Hussain and B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent FM signals using quadratic time-frequency distributions," *IEEE Trans. Signal Process.*, vol. 50, no. 8, pp. 1866–1876, 2002.
- [95] J. Lerga and V. Sucic, "Nonlinear IF estimation based on the pseudo WVD adapted using the improved sliding pairwise ICI rule," *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 953–956, 2009.
- [96] I. Djurovic and L. Stankovic, "Modification of the ICI rule-based IF estimator for high noise environments," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2655–2661, 2004.
- [97] V. Katkovnik and L. Stanković, "Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length," *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2315–2325, Sep 1998.
- [98] A. Gentile and A. Messina, "On the continuous wavelet transforms applied to discrete vibrational data for detecting open cracks in damaged beams," *Int. J. Solids Struct.*, vol. 40, no. 2, pp. 295–315, 2003.

- [99] X. Shao and C. Ma, "A general approach to derivative calculation using wavelet transform," *Chemom. Intell. Lab. Syst.*, vol. 69, no. 1-2, pp. 157–165, 2003.
- [100] S. Mallat, *A wavelet tour of signal processing*. Elsevier, 1999.
- [101] R. W. Schafer, "What is a Savitzky-Golay filter?[lecture notes]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 111–117, 2011.
- [102] J. Lerga, V. Sucic, and B. Boashash, "An efficient algorithm for instantaneous frequency estimation of nonstationary multicomponent signals in low SNR," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, no. 1, p. 725189, 2011.
- [103] I. Djurovic and L. Stankovic, "Quasi-maximum-likelihood estimator of polynomial phase signals," *IET Signal Proc.*, vol. 8, no. 4, pp. 347–359, 2014.
- [104] J. A. Volpe, "Vulnerability assessment of the transportation infrastructure relying on the global positioning system," <http://www.navcen.uscg.gov/>, 2001.
- [105] S. Pullen and G. X. Gao, "GNSS jamming in the name of privacy-potential threat to GPS aviation," *Inside GNSS*, vol. 7, no. 2, pp. 34–43, 2012.
- [106] T. Kraus, R. Bauernfeind, and B. Eissfeller, "Survey of in-car jammers-analysis and modeling of the RF signals and IF samples (suitable for active signal cancellation)," *Proceedings of ION GNSS 2011*, pp. 20–23, 2011.
- [107] R. H. Mitch, R. C. Dougherty, M. L. Psiaki, S. P. Powell, B. W. O’Hanlon, J. A. Bhatti, and T. E. Humphreys, "Signal characteristics of civil GPS jammers," in *Proceedings of ION GNSS*, vol. 2011. Citeseer, 2011, pp. 20–23.
- [108] G. Pojani, "Interference mitigation and localization based on time-frequency analysis for navigation satellite systems," Ph.D. dissertation, University of Bologna, 2018.
- [109] D. Borio, C. O’Driscoll, and J. Fortuny, "Jammer impact on Galileo and GPS receivers," in *Localization and GNSS (ICL-GNSS), 2013 International Conference on*. IEEE, 2013, pp. 1–6.
- [110] A. T. Balaei, A. G. Dempster, and D. Akos, "Quantization degradation of GNSS signal quality in the presence of CW RFI," in *Spread Spectrum Techniques and Applications, 2008 IEEE 10th International Symposium on*. IEEE, 2008, pp. 48–52.
- [111] H. Kuusniemi, E. Airos, M. Z. H. Bhuiyan, and T. Kröger, "GNSS jammers: how vulnerable are consumer grade satellite navigation receivers?" *European Journal of Navigation*, vol. 10, no. 2, pp. 14–21, 2012.
- [112] Y.-R. Chien, "Design of GPS anti-jamming systems using adaptive notch filters," *IEEE Syst. J.*, vol. 9, no. 2, pp. 451–460, 2015.

- [113] D. Borio, “Swept GNSS jamming mitigation through pulse blanking,” in *Navigation Conference (ENC), 2016 European*. IEEE, 2016, pp. 1–8.
- [114] P. T. Capozza, B. J. Holland, T. M. Hopkinson, and R. L. Landrau, “A single-chip narrow-band frequency-domain excisor for a global positioning system (GPS) receiver,” *IEEE J. Solid-State Circuits*, vol. 35, no. 3, pp. 401–411, 2000.
- [115] S. Savasta, L. L. Presti, and M. Rao, “Interference mitigation in GNSS receivers by a time-frequency approach,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 1, pp. 415–438, 2013.
- [116] X. Ouyang and M. G. Amin, “Short-time fourier transform receiver for nonstationary interference excision in direct sequence spread spectrum communications,” *IEEE Trans. Signal Process.*, vol. 49, no. 4, pp. 851–863, 2001.
- [117] S. Barbarossa and A. Scaglione, “Adaptive time-varying cancellation of wide-band interferences in spread-spectrum communications based on time-frequency distributions,” *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 957–965, 1999.
- [118] Y. Abdoush, G. Pojani, M. Bartolucci, and G. E. Corazza, “Time-frequency interference rejection based on the S-transform for GNSS applications,” in *2017 IEEE Int. Conf. on Commun. (ICC)*, May 2017.
- [119] G. Pojani, Y. Abdoush, M. Bartolucci, J. A. Garcia-Molina, and G. E. Corazza, “Multiple jammer localization and characterization based on time-frequency analysis,” in *2018 9th ESA Workshop on Satellite Navigation Technologies (Navitec 2018) European Workshop on GNSS Signals and Signal Process.*, Accepted.
- [120] B. Boashash, L. Boubchir, and G. Azemi, “A methodology for time-frequency image processing applied to the classification of non-stationary multichannel signals using instantaneous frequency descriptors with application to newborn EEG signals,” *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, p. 117, 2012.
- [121] —, “Improving the classification of newborn eeg time-frequency representations using a combined time-frequency signal and image approach,” in *Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on*. IEEE, 2012, pp. 280–285.
- [122] B. Boashash and A. Aïssa-El-Bey, “Robust multisensor time–frequency signal processing: A tutorial review with illustrations of performance enhancement in selected application areas,” *Digital Signal Process.*, vol. 77, pp. 153–186, 2018.