Time-Frequency Distributions: Approaches for Incomplete Non-Stationary Signals

Yen Thi Hong Nguyen

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

University of Leeds

Department of Electronics and Electrical Engineering

March 9, 2018

Declaration

This work in this thesis is based on research carried out at the Institute of Integrated Information Systems, School of Electronic and Electrical Engineering, Leeds University, UK. The candidate confirms that no part of this thesis has been submitted elsewhere for any other degree or qualification and it is all her own work except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. Most materials contained in the chapters of this thesis have been previously published in research articles, which are written by the author of this work (Yen Thi Hong Yen), who appears as lead (first) author in all of them. The research has been supervised and guided by Dr Mounir Ghogho, Dr Des McLernon, Dr Moeness Amin and Dr Syed Ali Raza Zaidi, and they appear as co-authors on these articles.

Chapter 4 covers the work of the two articles in the international conferences as follows:

- Nguyen, Y. T., Amin, M. G., Ghogho, M., & McLernon, D. (2015, May). Local sparse reconstructions of doppler frequency using chirp atoms. In Radar Conference (RadarCon), 2015 IEEE (pp. 1280-1284). IEEE.
- Nguyen, Y. T., Amin, M. G., Ghogho, M., & McLernon, D. (2015, May). Time-frequency signature sparse reconstruction using chirp dictionary. In Compressive Sensing IV (Vol. 9484, p. 94840H). International Society for Optics and Photonics. *Compressive Sensing*

IV. Vol. 9484. Society of Photo-optical Instrumentation Engineers, 2015.

Chapter 5 covers the content of the two following articles appeared in the international conferences:

- Nguyen, Y., McLernon, D., & Ghogho, M. (2016, September). Simplified chirp dictionary for time-frequency signature sparse reconstruction of radar returns. In Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa), 2016 4th International Workshop on (pp. 178-182). IEEE.
- Nguyen, Y. T., McLernon, D., Ghogho, M., & Zaidi, S. A. R. (2017, January). Sparse reconstruction of time-frequency representation using the fractional Fourier transform. In Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom), International Conference on (pp. 16-20). IEEE.

This thesis is dedicated to my family for their immense love and support.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my two supervisors Dr. Des McLernon and Professor Mounir Ghogho for the continuous support of my Ph.D study and related research, and also for their patience, motivation, and immense knowledge. Their guidance helped me through both my research and the writing of this thesis. Without their assistance, this work would have not been possible. I also want to show gratefulness to Dr. Des McLernon for his help and care in my life, in both good and bad times.

Secondly, I would like to express my deep gratitude to Professor Moeness Amin and all the staff in the university of Villanova, USA for their support and supervision. They have provided me with both the foundations and the cutting-edge knowledge of time-frequency analysis. Professor Moeness Amin oriented me towards a specific research topic and this contributed significantly to my final results. It was a significant breakthrough in my PhD work to meet and collaborate with them in the USA.

I thank my fellow lab-mates in room 3.62, School of Electronic and Electrical, for the stimulating discussions, for the sleepless nights when we were working together and for all the fun we have had in the last four years. I would like to specifically thank Asma, Ali, Mohanad, Edmond, Tuan, Miaomiao and Naveed for all the precious things we have shared.

Last but not the least, I would like to thank my family: my parents, my sister and my husband for supporting me both spiritually throughout writing this thesis and in my life in general.

Abstract

There are many sources of waveforms or signals existing around us. They can be natural phenomena such as sound, light and invisible like electromagnetic fields, voltage, etc. Getting an insight into these waveforms helps explain the mysteries surrounding our world and the signal spectral analysis (i.e. the Fourier transform) is one of the most significant approaches to analyze a signal. Nevertheless, Fourier analysis cannot provide a time-dependent spectrum description for spectrum-varying signals-non-stationary signal. In these cases, time-frequency distributions are employed instead of the traditional Fourier transform. There have been a variety of methods proposed to obtain the time-frequency representations (TFRs) such as the spectrogram or the Wigner-Ville distribution. The time-frequency distributions (TFDs), indeed, offer us a better signal interpretation in a two-dimensional time-frequency plane, which the Fourier transform fails to give. Nevertheless, in the case of incomplete data, the time-frequency displays are obscured by artifacts, and become highly noisy. Therefore, signal time-frequency features are hardly extracted, and cannot be used for further data processing. In this thesis, we propose two methods to deal with compressed observations. The first one applies compressive sensing with a novel chirp dictionary. This method assumes any windowed signal can be approximated by a sum of chirps, and then performs sparse reconstruction from windowed data in the time domain. A few improvements in computational complexity are also included. In the second method, fixed kernel as well as adaptive optimal kernels are used. This work is also based on the assumption that any windowed signal can be approximately represented by a sum of chirps. Since any chirp 's auto-terms only occupy a certain area in the ambiguity domain, the kernel can be designed in a way to remove

the other regions where auto-terms do not reside. In this manner, not only cross-terms but also missing samples' artifact are mitigated significantly. The two proposed approaches bring about a better performance in the time-frequency signature estimations of the signals, which are simulated with both synthetic and real signals. Notice that in this thesis, we only consider the non-stationary signals with frequency changing slowly with time. It is because the signals with rapidly varying frequency are not sparse in time-frequency domain and then the compressive sensing techniques or sparse reconstructions could not be applied. Also, the data with random missing samples are obtained by randomly choosing the samples' positions and replacing these samples with zeros.

Contents

List of Acronyms

EM	Electromagnetic
LOS	Light of sight
FM	Frequency modulated
ECG	Electrocardiograph
ALCM	Rotating air-launched cruise missile
STFT	Short-time Fourier transform
WVD	Wigner-Ville distribution
TFD	Time-frequency distribution
TF	Time-frequency
CS	Compressive sensing
DTFT	Discrete time Fourier transform
DFT	Discrete Fourier transform
DCT	Discrete Cosine transform
FRFT	Fractional Fourier transform
EMD	Empirical mode decomposition
IMP	Intrinsic mode function
SNR	Signal to noise ratio
IAF	Instantaneous autocorrelation function
AF	Ambiguity function
FT	Fourier transform
RIP	Restricted isometry property
OMP	Orthogonal matching pursuit
POMP	Prune orthogonal matching pursuit
MMV	Multiple measurement vector
IF	Instantaneous frequency
RID	Reduced interference distribution
RGK	Radially-Gaussian kernel
AOK	Adaptive optimal kernel
STAF	Short-time ambiguity function
SMV	Single measurement vector
TFSR	Time-frequency signal representation
QTFD	Quadratic time-frequency distribution
2D	Two-dimensional

List of Symbols

t	Continuous time variable
n	Discrete time variable
f	Continuous frequency variable
k	Discrete frequency
f_R	Received frequency
f_{TX}	Transmit frequency
f_D	Doppler frequency
v	Radial velocity
c	Light propagation speed
$H\{...\}$	Hilbert transform
$z(t)$	Analytic signal
ω	Continuous angular frequency
ω'	Continuous angular Doppler frequency
$\mathbb{F}$	Fourier transform
$\mathbb{F}^{-1}$	Inverse Fourier transform
$\mathbb{F}_{2D}^{-1}$	Two-dimensional inverse Fourier transform
$\otimes$	Convolution
W	Wigner Ville distribution
$s(n)$ or $s(t)$	Signal
S	Fourier transform of signal s
τ	Continuous lag variable
b	Discrete lag variable
p	Discrete Doppler variable
SNR	Signal to noise ratio
N	Signal length
$D(n,k)$	Cohen's class distribution
$C(p,b)$	Kernel function
IAF	Instantaneous autocorrelation function
$\|...\|_0$	Norm 0
$\|...\|_1$	Norm 1
$\|...\|_2$	Norm 2 or Euclidean norm
$\text{supp}(x)$	Support of x
E	Number of signal components
$\mathbb{R}$	A set of real number

$\mathbb{C}$	A set of complex number
M	Number of measurements
$\mathbf{\Psi}$	Basic dictionary
$\mathbf{\Phi}$	Measurement matrix
v	Noise vector
$\mathbb{S}$	A set of observed time instants
$\lvert\mathbb{S}\rvert$	Cardinality of the observed time instants set $\mathbb{S}$
$m(n)$	Observation data
$M(n)$	Observation mask
$miss(n)$	Missing data
$Miss(n)$	Missing data mask
$S(n)$	Full data mask
$S(k)$	Discrete Fourier transform of $s(n)$
$\mathbf{N}_M$	Position of measurements
$AF(p,b)$	Discrete ambiguity function
$TF(n,k)$	Time-frequency distribution
F_s	Sampling frequency
$F_{\max}$	Maximum frequency
m	The window index
$\mathbf{S}_m$	The signal vector
N_w	The window length
T_w	The window time
ϵ	The noise level
$F_e(t)$	The continuous-time instantaneous frequency
T	Total observation interval
α	Chirp-rate
β	Initial frequency
x	Continuous fractional variable
u	Discrete fractional variable
ϕ	Angle rotation
$(\mathbb{F}s)(x)$ or $S^{\phi}(x)$	Fractional Fourier transform operator associated with angle ϕ
sl	Scale parameter
a	The order of the FRFT
x_0, x_1	The scale coordinates for time and frequency
μ	The time-bandwidth product
Δt	Time interval
Δf	Frequency bandwidth
Δx	Scaled variable interval
Q	Dimension of the dictionary
ξ	Concentration level
$R_{ss}(b,n)$	Instantaneous ambiguity function

η	The shift between consecutive windows
$h(n)$	Window function
$\mathbf{\Psi}_c$	Compact dictionary
σ	Standard deviation of the Gaussian kernel
vo	Kernel volume
δf	Frequency resolution
δt	Time resolution
δb	Lag resolution
r	Continuous radius variable
q	Discrete aspect angle
g	Discrete radius variable

Chapter 1

Introduction

1.1 Motivation

There are many sources of waveforms or signals existing around us. They can be natural phenomena such as sound, light and electromagnetic fields, etc. Getting an insight into these waveforms helps us understand the mysteries of our world.

Signal spectral analysis (i.e., the Fourier transform) is one of the most important approaches. The mathematics behind it was invented by Fourier and the Fourier distribution is one of the great innovations of mathematics and science [5]. It helped Fraunhofer to find out and catalogue spectral lines, serving as a finger print of substance composition at element and molecule level in Bunsen-Kirchhoff 's work. It is applied largely in radar, for example, detecting direction and velocity of moving objects by measuring the Doppler shift. We know that a radar transmits an electromagnetic (EM) signal to an object and receives a returned wave from it. If the target is moving, the received frequency will be shifted from the original frequency, and this is known as the Doppler effect [6, 7]. The backscatter frequency is expressed as:

$$f_R = f_{TX} + f_D = f_{TX} - f_{TX}\frac{2v}{c},$$

where f_R, f_{TX}, f_D are the received frequency, transmitted frequency and Doppler shift, v is the radial velocity of the target along the light of sight (LOS) of the radar and c is the propagation speed of EM waves. The object velocity v is defined to be negative if it moves towards the radar and positive if it moves away from the radar. So if the target moves toward the observer, the received frequency is

increased compared with that of the transmitted wave. This explains why a light source moving toward an observer appears bluer and while moving away the the observer, the light becomes more red. By measuring the frequency of the reflected signal, both direction and speed of the object are determined.

It can be seen that the Fourier transform plays a very important role in signal processing. Nevertheless, Fourier analysis cannot provide time-dependent spectrum description. It can display all spectral components contained in the data, but it cannot show when they are actually present. For example, while musical notes are written to indicate the changing of frequency with time, from the lowest one, called the fundamental, to the overtones, frequency analysis cannot facilitate such interpretation. The magnitude spectrum many exhibit hundreds of peaks in the audible frequency range, and the relative heights of those peaks may tell us about the tonality of the music, but not the timing of the notes. In order to get both temporal and spectral information, joint time-frequency (TF) analysis has been proposed, where signals are expanded in two dimensions, i.e., time and frequency. Non-stationary signals, like music, sinusoidal frequency modulated (FM) signals, chirp signals and micro-Doppler radar returns, etc., can reveal their properties via the TF distribution (TFD). TF analysis has many applications. Recently, there has been a huge interest in using radar to detect human activity "through the wall" by analyzing the micro-Doppler frequency from the radar returns. This technique has been used in disaster aid, medical care, and defence. The micro-Doppler estimation also helps us to determine the kinematic properties of an object. For example, measurements of the surface vibration of the vehicle could assist us in detecting the type of vehicle, such as a tank with a gas turbine engine or a bus with a diesel engine [5]. So, we can say the micro-Doppler can serve as the movement's signature.

The following are two examples to illustrate the advantages that time-frequency analysis brings about over the separate time and frequency displays. In the first one, electrocardiograph (ECG) data is used. The data is obtained from the MIMIC II database [8]. The heart beat is a very essential heath description. Analyzing and classifying ECG in the time-frequency domain gives an accuracy up to 99%, which outperforms normal spectral analysis [9]. Three different ECG segments are displayed in Fig. 1.1-1.3. These ECGs are abnormal due to action artifacts and the differences in waveforms can be seen in Fig. 1.1. However, it is hard to extract

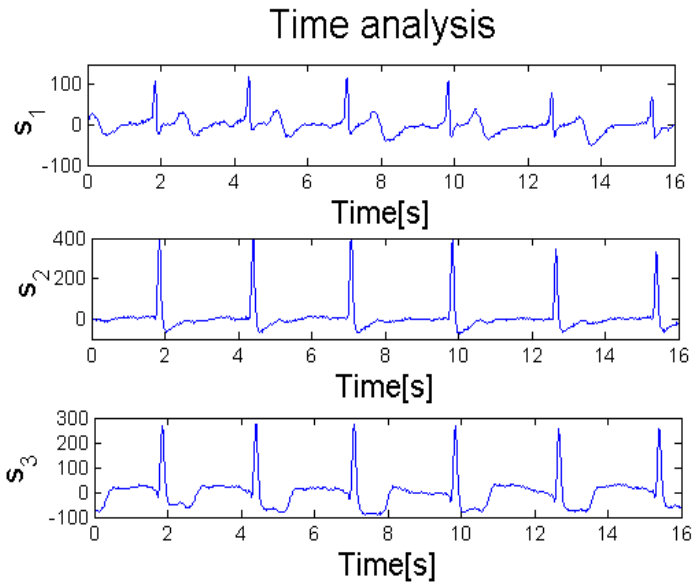


Figure 1.1: Three ECG waveforms in the time domain.

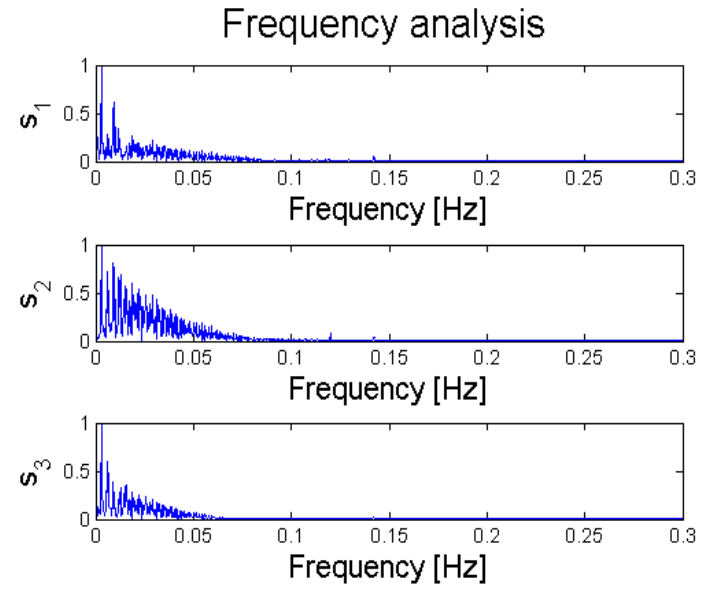


Figure 1.2: Spectral analysis of the three ECG waveforms in Fig. 1.1.

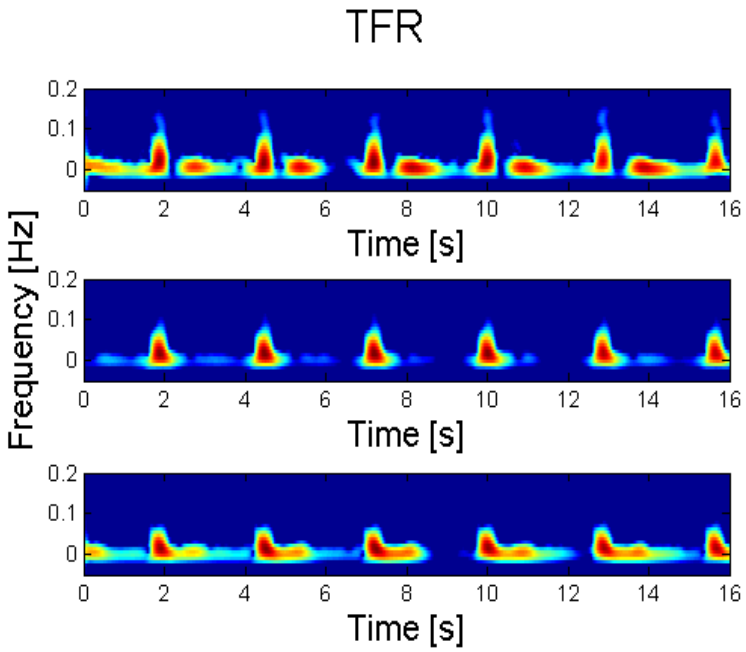


Figure 1.3: TF analysis of the three ECG waveforms in Fig. 1.1.

the features of each ECG to implement classification. The frequency domain does not reveal the signal signature as clearly as in the joint time-frequency domain. So Fig. 1.3 obviously displays that the first signal is composed of high frequency content as well as background components at strong magnitude. The peak frequency of the third waveform is the lowest. In the second example, the micro-Doppler frequency is discussed. It is defined as the backscattered spectral shift due to micro motion, i.e., oscillatory motion of an object or structural components of the object in addition to the bulk motion. The source of micro-motions may be a rotating propeller of a fixed-wing aircraft, a rotating antenna, a walking person with swinging arms and legs, etc. [5]. This frequency modulation on the carrier frequency of a radar transmitted signal can be deployed as a target signature for identification, classification and recognition. Fig. 1.4-1.6 show the temporal, frequency and time-frequency analysis of a signal reflected from a rotating air-launched cruise missile (ALCM) provided in the simulation software in [10]. The radar transmits 8192 pulses with a pulse repetition interval of $67\mu s$ during a period of 0.55 seconds to cover the total target's rotation angle of 360^0. The micro-Doppler features of the

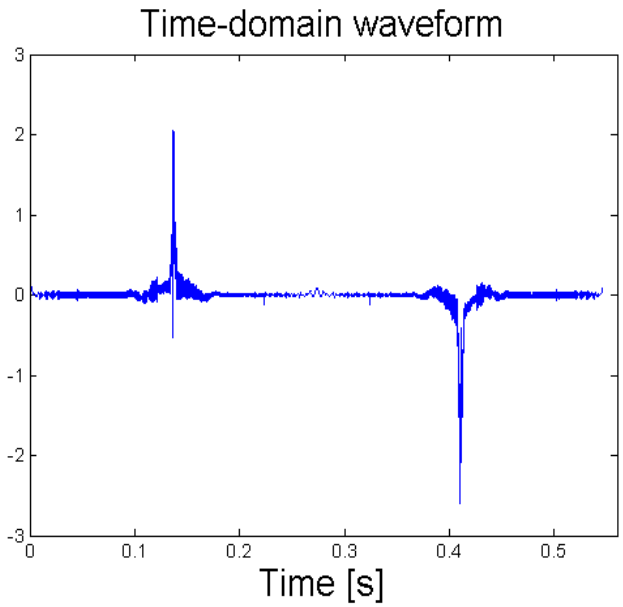


Figure 1.4: Time domain reflected ALCM signal.

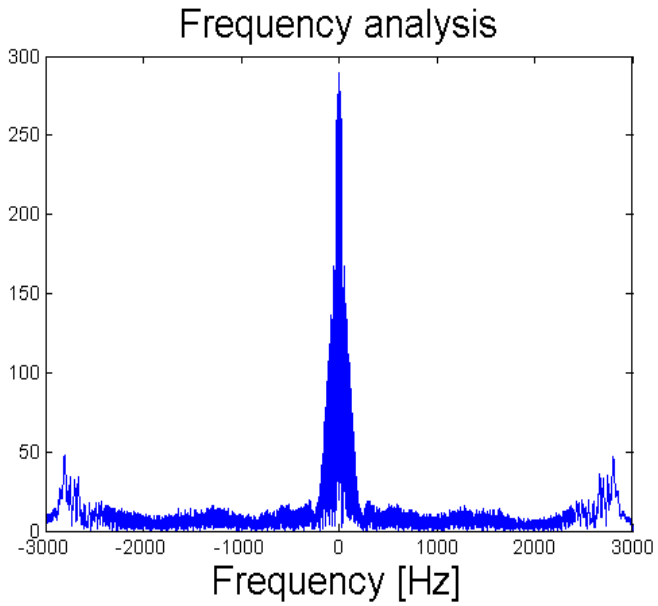


Figure 1.5: Spectral analysis of reflected ALCM signal.

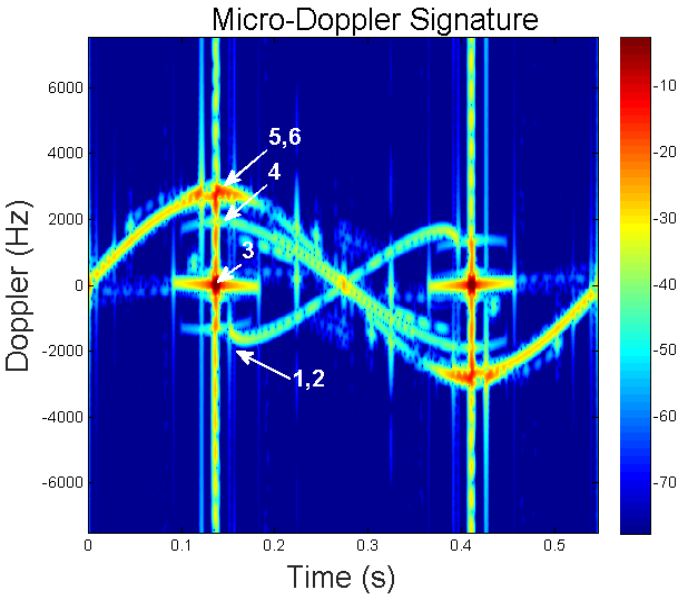


Figure 1.6: The micro-Doppler signature of a simulated rotating ALCM (1: Head tip, 2: Head joint, 3: Wing joint, 4: Engine intake, 5: Tail fin, tail plane, 6: tail tip).

rotating ALCM are revealed more clearly in the joint time-frequency analysis. It can be seen from Fig. 1.6 that the Doppler shift from the wing-joint is almost zero as it locates at the middle of the ALCM and thus the distance between it and the radar is nearly unchanged. In contrast, the tail and its structures have a Doppler shift in the form of a sinusoid due to a dramatical change in distance when the ALCM rotates. As a result, this causes large micro-Doppler shift. The magnitude of the frequencies displayed in the Fig. 1.6 is determined by the angular velocities. They are maximum when the missile is at a 90^0 or 270^0 aspect to the radar.

1.2 Literature Overview

In 1948, Dennis Gabor, a Hungarian Nobel laureate, proposed the first algorithm on TF analysis of an arbitrary signal [11]. He basically applies a short Gaussian window on the signal, and implements the Fourier transform to ascertain the frequency components in the signal segment. The Gaussian window is used because it obtains the minimum product of time and frequency resolution.

The spectrogram is a widely used method to display the time-varying spectral density of non-stationary signals. The spectrogram is calculated by using the short-time Fourier transform (STFT) and then the absolute magnitude is squared to obtain the energy representation [5]. The STFT performs the Fourier transform on a short-time window rather than taking a Fourier transform on the whole signal. The resolution of the STFT depends on the window size. There is a trade-off between the time and the frequency resolution. The larger the window length, the better the frequency resolution, but the poorer the time resolution becomes. The Gabor transform indeed belongs to the STFT with the Gaussian window.

Later on, a better TF resolution method is proposed, i.e., the Wigner-Ville distribution (WVD). It is basically the Fourier transform of the signal bilinear product over a lag variable. Its drawback is that if the signal contains more than one component, its WVD will contain cross-terms that occur halfway between each pair of auto-terms. The magnitude of this interference could be twice as large as the auto-terms. To mitigate the cross-terms, filtered WVDs have been suggested. They apply kernels to reduce large interferences at the expense of slightly reduced TF resolution. There are many kernels such as Choi-Williams, Margrnau-Hill, Born-Jordan, etc. They all belong to the Cohen's class.

Other high-resolution TFDs are the adaptive Gabor representation and the TFD series [12]. They decompose a signal into a family of basis functions, such as the Gabor function, which is well localized in both the time and frequency domains and is adaptive to match the local behaviour of the analyzed signal.

1.3 Challenges and Approaches to Obtain Reliable TFDs with Compressed Data

It can be seen that TF analysis offers us better signal representations, and it can be deployed in many applications from military to medical, disaster aid, etc. Nevertheless, in the case of incomplete data, the time-frequency displays are obscured by artifacts, becoming highly noisy. Therefore, signal TF features are hardly extracted, and cannot be used for the further steps of data processing. The TF representations of both ECG signals and the micro-Doppler shift from the ALCM with 50%

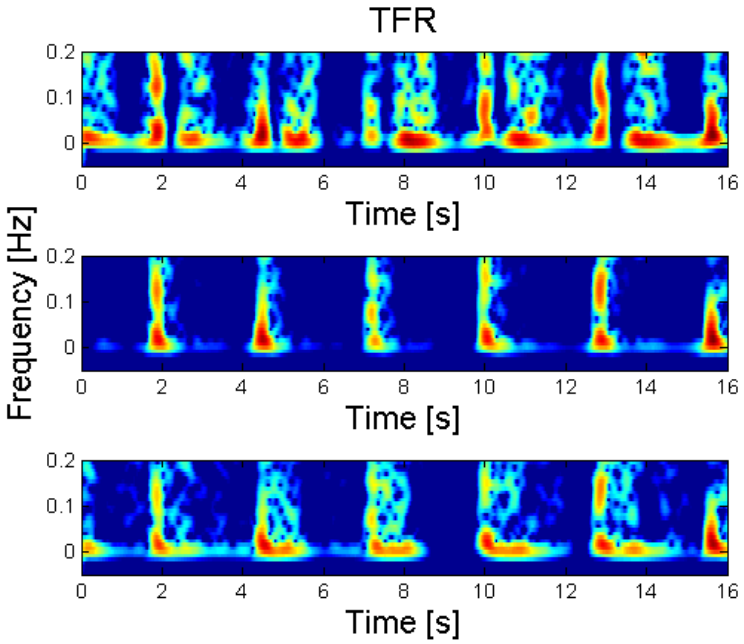


Figure 1.7: Time-frequency analysis of an ECG signal with 50% data missing.

data missing are plotted in Fig. 1.7 and Fig. 1.8. We can observe that the missing data effects are extremely severe. Noise-like effect clutters all the space, and eclipses desired information. In modern life where big data is processed everyday, signals could be partially cut to reduce the burden on hardware and to save time. Signals can also be degraded by excessive noise, which can be filtered out. Thus, time-frequency analysis approaches which are robust to missing samples are of significance. This thesis introduces two approaches which can combat missing data. The first one applies a sparse reconstruction with a novel chirp dictionary. The second method introduces new fixed and adaptive kernels which can effectively mitigate the cross-terms as well as the missing data's artifacts.

1.4 Thesis Outline and Contribution

This PhD thesis describes the research carried out on the reconstruction of a TF signatures of non-stationary signals, such as micro-Doppler radar returns and ECG, etc., especially when the signals are incomplete or randomly sampled. Missing data

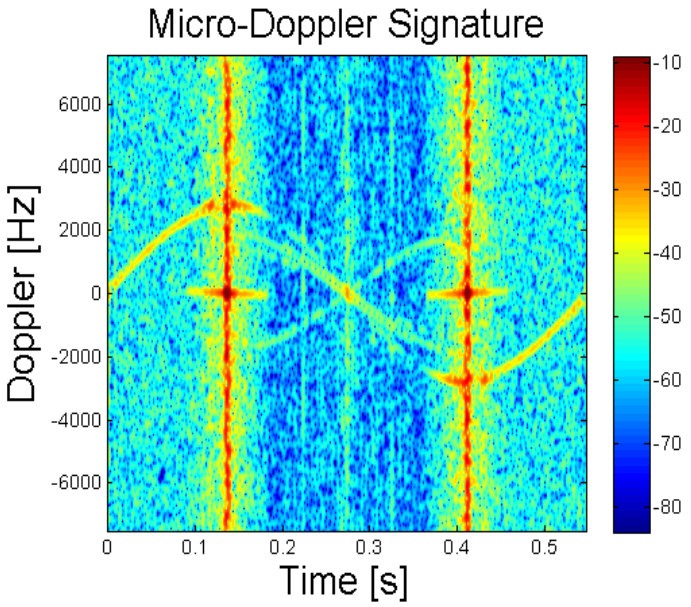


Figure 1.8: The micro-Doppler signature of a simulated rotating ALCM with 50% data missing.

causes artifacts spreading all over the ambiguity and TF domain, which clutter the signal components and hide the pertinent signal structure including the instantaneous frequencies. Traditional methods like STFT, WVD and Cohen's reduced interference class all fail to give accurate TF estimations of the signal. This thesis introduces two methods which ensure good instantaneous frequency approximation even in the case of compressed observations. The first method applies compressive sensing with a novel chirp dictionary. A few improvements in computational complexity are also included. Notice that in this part, we only consider the non-stationary signals with frequency changing slowly with time. It is because the signals with rapidly varying frequency are not sparse in time-frequency domain and then the compressive sensing techniques or sparse reconstructions could not be applied. In the second method, fixed as well as adaptive optimal reduced interference kernels are used. Different from conventional kernels, our proposed kernel can partially combat missing sample effects. Throughout this thesis, we extensively use the concepts in conventional methods such as STFT, ambiguity domain and WVD. Hence, for that reason, the next chapter is devoted to traditional TF techniques, where all important concepts are explained. Chapter 3 considers the effects of missing

data and compressive sensing basics. Chapter 4 is devoted to the chirp dictionary approach, where we introduce two different ways to build the atom set, discussing the signal model as well as solving the sparse problem for TF signature estimation. Chapter 5 considers two methods for reducing the complexity of the chirp dictionary approach. The fixed and signal adaptive reduced interference kernel designs are included in chapter 6. Chapter 7 concludes the thesis and gives further research directions.

Chapter 2: Conventional Time-Frequency Analysis

Time-frequency distributions (TFDs) concern the analysis and processing of signals with time-varying frequency content or non-stationary signals. Such signals are best represented by TFDs because they show how the energy of the signal is distributed over the two-dimensional time-frequency space instead of only one (time or frequency). This chapter presents the key concepts of conventional TFDs such as the Hilbert transform, the short time Fourier transform, the Wigner Ville distribution, the fixed reduced interference kernels which belong to Cohen's class and the adaptive radial Gaussian kernel.

Chapter 3: Sparse Time-Frequency Distribution Fundamentals

As most non-stationary signals are sparse in TF domain, compressive sensing (CS) has been applied in TFDs to give better TF estimations with full and incomplete data. These methods are called spare TFDs. In this chapter, we will study closely the motivation for using CS (especially using CS in TFDs), the fundamentals of CS techniques and some recent TFDs approaches using CS.

Chapter 4: Sparse Reconstruction of Time-Frequency Signature using Chirp Dictionary

This chapter includes our proposed sparse TFD. This method performs sparse reconstruction from windowed data in the time domain with a novel chirp dictionary. In many situations, the non-stationary signal frequency law is more properly approximated by piece-wise second order polynomials than fixed frequency sinusoids. Therefore, the chirp dictionary, instead of the sinusoidal dictionary, is better suited

for sparse reconstruction problems dealing with FM signals. The chirp dictionary is built in two ways. The first includes all possible chirps which can appear in any signal segment. The second is also composed of all chirps, but they are formed from sinusoids which are rotated in all eligible angles by the fractional Fourier transform. Although the dictionary construction procedure is different, the two ways actually lead to the same results. The purpose of presenting the second method is to introduce an alternative way to build the chirp dictionary. Its theory is also the foundation for other applications in the following chapters.

The work of this chapter has led to the publication of two articles in international conferences.

- Nguyen, Yen TH, et al. "Local sparse reconstructions of doppler frequency using chirp atoms." *Radar Conference*, 2015 IEEE. IEEE, 2015.
- Nguyen, Yen TH, et al. "Time-frequency signature sparse reconstruction using chirp dictionary." *Compressive Sensing IV*. Vol. 9484. Society of Photo-optical Instrumentation Engineers, 2015.

Chapter 5: Simplified Chirp Dictionary

The chirp dictionary approach has been proven to provide more reliable TF estimation compared with the sinusoid atom method. The chirp approach, nevertheless, deploys a very large dimension measurement dictionary. Since there are two parameters to be estimated (i.e. the chirp rate and the initial frequency), the dictionary dimension can be equal to the square of the dimension when using the sinusoid atom. This very large atom set leads to a much heavier computation burden and a longer calculation time. Therefore, in order to obtain good TF estimation at low computational complexity, chirp dictionary simplification methods are needed. In this chapter, we introduce two approaches which reduce the chirp dictionary dimension and give a low calculation load. In the first approach, we estimate the chirp rate through the DTFT of the bilinear product at a certain time lag. The initial frequency is solved in the time domain, with a lower dimensional dictionary than the computationally complex full chirp atom. In the second approach, the fractional Fourier transform (FRFT) is used to obtain an initial frequency for each chirp-rate.

This leads to a much simplified chirp atom set. The work of this chapter has led to the publication of two articles in international conferences.

- Nguyen, Yen TH, Des McLernon, and Mounir Ghogho. "Simplified chirp dictionary for time-frequency signature sparse reconstruction of radar returns." *Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, 2016 4th International Workshop on. IEEE, 2016.

- Nguyen, Yen TH, et al. "Sparse reconstruction of time-frequency representation using the fractional Fourier transform." *Recent Advances in Signal Processing, Telecommunications and Computing (SigTelCom)*, International Conference on. IEEE, 2017.

Chapter 6: Reduced Interference Chirp-based Time-Frequency Distribution for Limited Data

In this chapter, we introduce novel fixed and signal-dependent kernels in the ambiguity domain, which can efficiently remove cross-term interference and partially combat missing sample effects without using compressive sensing techniques. These kernels are applied on windowed signals to facilitate online implementation, or processing long signals. According to [13, 14], any non-stationary signal segment can be approximated by a sum of chirps. Additionally, the chirps' auto-terms always reside in only half of ambiguity domain and do not cover the Doppler axis. By removing the areas where the auto-terms do not lie, part of interference and artifacts are mitigated. Moreover, the analysis of the artifacts' distribution shows that the artifact always appear along the Doppler axis. By removing the region along the Doppler axis, our chirp-based kernels give good TFRs in the case of incomplete data.

The technical contributions of this chapter has been written into a journal paper and is waiting for submission.

Chapter 2

Conventional Time-Frequency Analysis

2.1 Introduction

Time-frequency distribution (TFD) concerns the analysis and processing of signals with time-varying frequency content, or non-stationary signals. Such signals are best represented by TFDs because they show how the energy of the signals is distributed over the two-dimensional time-frequency space instead of only one (time or frequency). This chapter presents the key concepts of conventional TFDs such as the Hilbert transform, the short time Fourier transform, the Wigner Ville distribution, the fixed reduced interference kernels which belong to Cohen's class and the adaptive radial Gaussian kernel.

2.2 Instantaneous Frequency Analysis

2.2.1 Hilbert transform

For a real signal $s(t)$, the analytic signal or its associated complex signal $z(t)$ is defined by [5]:

$$z(t) = s(t) + jH\{s(t)\} = a(t)\exp(j\varphi(t)), \tag{2.1}$$

where $a(t)$ and $\varphi(t)$ are the time-varying amplitude and phase of the analytic signal, and $H\{.\}$ is the Hilbert transform of the signal, which is expressed as:

$$H\{s(t)\} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{s(\tau)}{t-\tau}d\tau. \tag{2.2}$$

The Fourier transform of the analytic signal is single-sided, with a unique phase function. Since the instantaneous frequency is the time derivative of phase function, it is also unique, and is expressed by:

$$f(t) = \frac{1}{2\pi}\frac{d}{dt}\varphi(t). \tag{2.3}$$

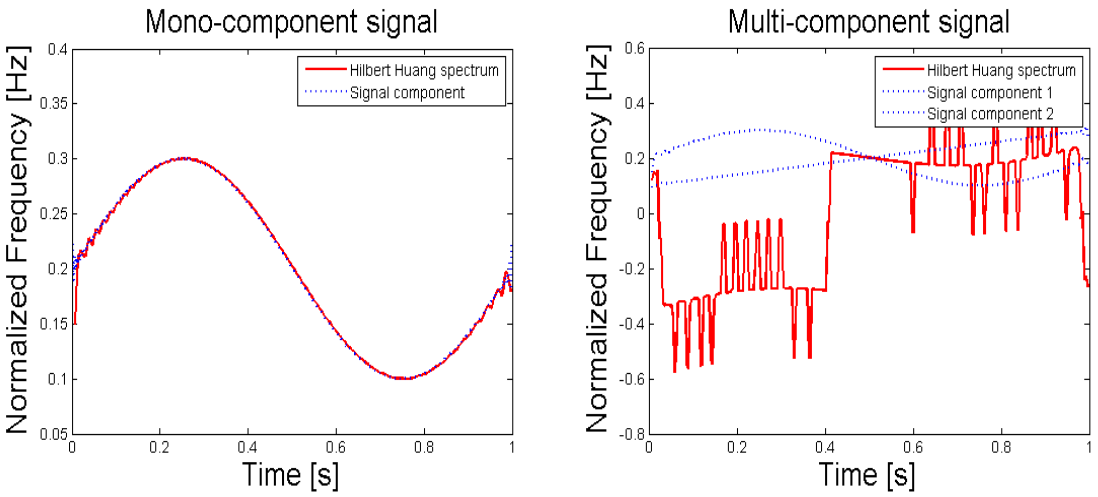
In the case of a discrete signal $s(n)$ $(n = 1, 2, ..., N)$, the analytic signal, the Hilbert transform and instantaneous frequency are defined as follows:

$$\begin{aligned} z(n) &= s(n) + jH\{s(n)\} = a(n)\exp(j\varphi(n)), \\ H\{s(n)\} &= \frac{1}{\pi}\sum_{b=-\infty}^{\infty}\frac{s(n)}{n-b}, \\ f(n) &= \frac{1}{2\pi}\frac{1}{2\Delta t}[\varphi(n+1) - \varphi(n-1)], \end{aligned} \tag{2.4}$$

where Δt is the sample interval. It is a pretty simple way to get the time-frequency analysis of the signal. However, the instantaneous frequency only gives one frequency value a time, and so, this method is only suitable for a mono-component signal, not for a multicomponent signal. The method is simulated with a single sinusoidal frequency modulated (FM) signal and a combination of a linear chirp and a sinusoidal FM. The sampling frequency is 256Hz. The approach performs well only in the former case as shown in Fig. 2.1.

2.2.2 Hilbert-Huang transform

In order to distinguish frequency distributions of a multiple-component signal, Huang et al. [15] introduced the concept of empirical mode decomposition (EMD) to separate a multi-component signal into many single component signals, which are called intrinsic mode functions (IMFs). The Hilbert transform is then applied for each IMF to obtain the TF analysis. Given the signal $s(t)$, the EMD algorithm is summarized as follows:



(a) (b)

Figure 2.1: Hilbert Huang spectrum for signal composed of (a) One component, and (b)Two components.

1. Identify all extrema (minima and maxima) of the signal $s(t)$.
2. Deduce an upper and a lower envelope by interpolation (for example linear or cubic splines).
   - Subtract the mean envelope from the signal.
   - Iterate until number of extrema=number of zeros ± 1.
3. Subtract the so-obtained IMF from the signal.
4. Iterate on the residual.

This method is model-free and fully data driven. As it is based on sifting, it is vulnerable to noise, and it requires oversampling for getting accurate interpolation. Moreover, the EMD method produces oscillatory or poorly-defined Hilbert-spectra, often with notable mode mixing. Importantly, this approach lacks a general mathematical theory [16]. Later, Olhede and Walden introduced a wavelet packet-based decomposition as a replacement for the EMD in preprocessing the multi-component

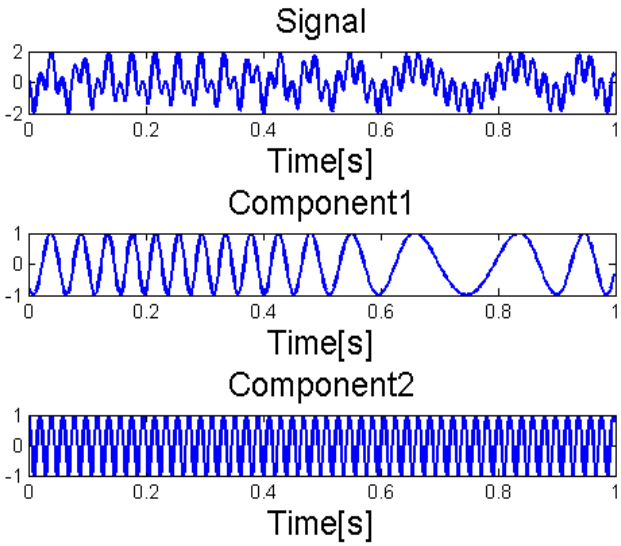


Figure 2.2: Multi-component signal.

signals. For a multi-component signal in Fig. 2.2, the signal decomposition is shown in Fig. 2.3, and the time-frequency analysis is displayed in Fig. 2.4. The second example illustrates how EMD works in the case where noise is present. A sinusoid signal which is contaminated by Gaussian noise with $SNR = 5$dB is used here. It is expected to get only one IMF, but we obtain more than that. The obtained IMFs are cluttered with noise, ans so the signal cannot be seen by the Hilbert transform. The simulation results are shown in Fig. 2.5 and Fig. 2.6.

2.3 Quadratic Time-Frequency (TF) Analysis

2.3.1 Short-time Fourier transform and spectrogram

We all know that the Fourier transform only offers all the frequency content inside the considered duration, and does not reveal when the frequency actually appears. In order to obtain both temporal and spectral information, the straightforward solution is to break up the signal into small time segments and perform the Fourier transform each time segment to ascertain the frequencies that existed in that segment. The totality of such spectra indicates how the spectrum is varying with time [5]. However,

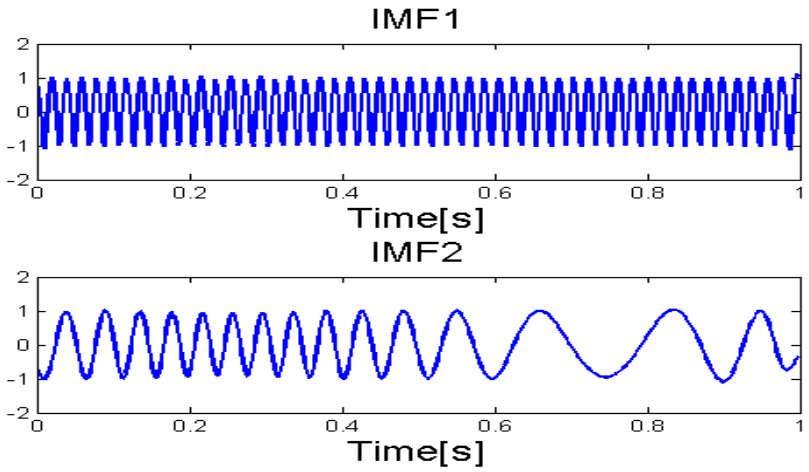


Figure 2.3: EMD decomposition of the multi-component signal in Fig. 2.2.

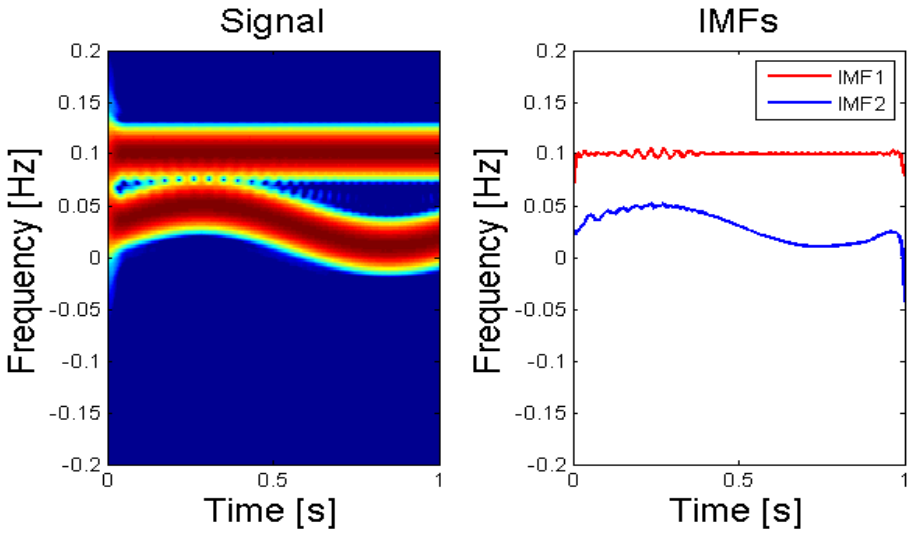


Figure 2.4: Time-frequency representation and IMFs for the signal in Fig. 2.2.

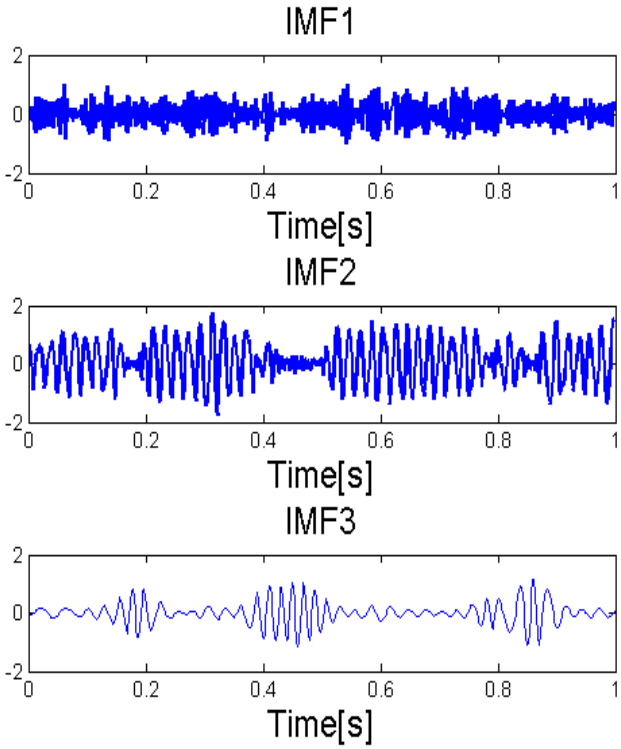


Figure 2.5: EMD decomposition of a sinusoid signal in the presence of noise.

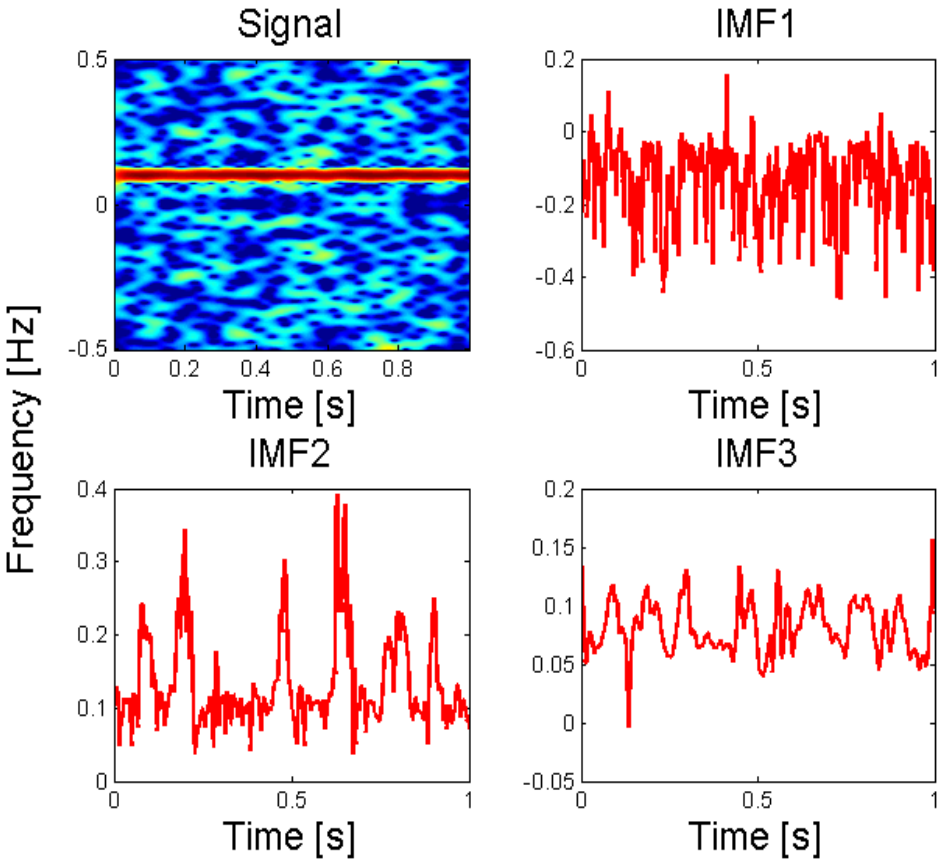


Figure 2.6: The TFD and the corresponding IMFs in the TF domain of the noisy sinusoidal signal.

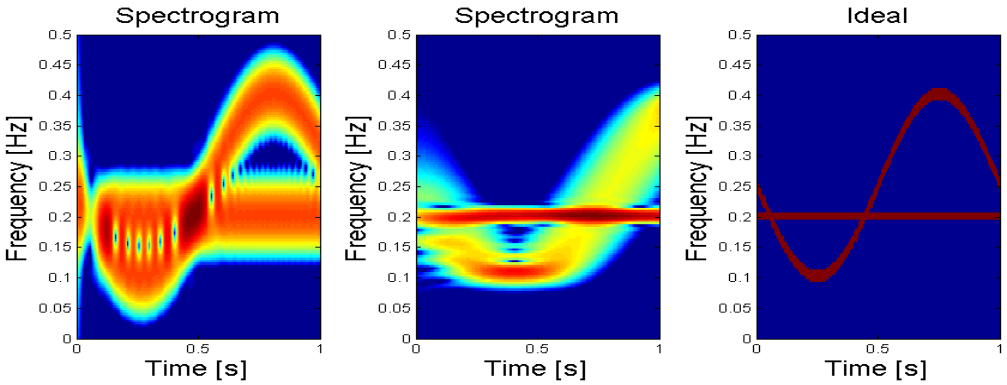


Figure 2.7: Time-frequency representation: (a) Spectrogram with small window length $N_w = 20$; (b) Spectrogram with large window length $N_w = 100$; (c) Ideal time-frequency distribution.

the method confronts the trade-off between time and frequency resolution. We cannot achieve finer and finer time localization because short signal duration means large bandwidth, and the spectra of such short duration signals have very little to do with the properties of the original signal. The short-time Fourier transform of signal $s(t)$ is expressed as:

$$S_t(\omega) = \frac{1}{\sqrt{2\pi}} \int s(\tau) h(\tau - t) e^{-j\omega\tau} d\tau, \tag{2.5}$$

where $h(t)$ is a window function, centered at t. The energy desity spectrum at time t is as follows:

$$P_{SP}(t, \omega) = |S_t(\omega)|^2 = \left| \frac{1}{2\pi} \int s(\tau) h(\tau - t) e^{-j\omega\tau} d\tau \right|^2. \tag{2.6}$$

For different times, different sprectra are obtained. When all spectra assemble together, we get the TF distribution, or spectrogram. The drawback of the spectrogram is illustrated with a signal composed of a sinusoidal FM and a chirp, and sampling frequency $F_s = 256$ Hz. The results are given in Fig. 2.7.

2.3.2 Wigner-Ville distribution

The Wigner distribution is the prototype of a distribution which is qualitatively differently from the spectrogram. The discovery of its strength and shortcomings

has become a major issue in the development of the field. The Wigner distribution in terms of the signal $s(t)$ or its Fourier transform, $S(\omega)$, is [5],

$$\begin{aligned} W(t,\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} s^*(t-\frac{1}{2}\tau)s(t+\frac{1}{2}\tau)e^{-j\tau\omega}d\tau \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} S^*(\omega+\frac{1}{2}\theta)S(\omega-\frac{1}{2}\theta)e^{-jt\phi}d\theta, \end{aligned} \tag{2.7}$$

where $\omega = 2\pi f$ is angular frequency, τ is time lag, θ is frequency shift and * denotes the complex conjugate operation. The Wigner distribution is bilinear in term of the signal because the signal $s(t)$ enters twice in its calculation. (2.7) can be written as the Fourier transform over the lag variable of the signal bilinear product,

$$\begin{aligned} W(t,\omega) &= \mathbb{F}_{\tau\to\omega}\left\{ s^*(t-\frac{1}{2}\tau)s(t+\frac{1}{2}\tau)\right\} \\ &= \frac{1}{2\pi}\left[2S^*(2\omega)\exp(j\omega t)\right] \otimes_\omega \left[2S(2\omega)\exp(j\omega t)\right] \\ &= \frac{2}{\pi}\left[S^*(2\omega)\exp(j\omega t)\right] \otimes_\omega \left[S(2\omega)\exp(j\omega t)\right], \end{aligned} \tag{2.8}$$

where $\otimes_\omega$ and $\mathbb{F}$ denotes convolution in the frequency plane and the Fourier transform, respectively. Equation (2.8) is based on the following Fourier transform properties [17],

$$\begin{aligned} &\mathbb{F}\left[s(a\tau)\right] = \frac{1}{a}S(\omega/a) \\ &\mathbb{F}\left[s(\tau-\tau_0)\right] = S(\omega)\exp(-j\omega\tau_0) \\ &\mathbb{F}\left[s^*(\tau)\right] = S^*(-\omega) \\ &\mathbb{F}\left[s(-\tau)\right] = S^*(\omega). \end{aligned} \tag{2.9}$$

The use of $\pm\tau/2$ in (2.7, 2.8) ensures that the frequency content is correctly positioned. Consider a monochromatic signal to see how the signal is distributed in the Wigner-Ville domain,

$$\begin{aligned} s(t) &= \cos(\omega_0 t) \\ S(\omega) &= \pi\delta(\omega \pm \omega_0). \end{aligned} \tag{2.10}$$

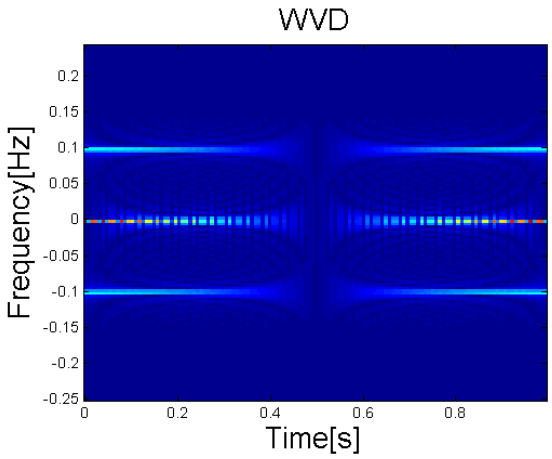


Figure 2.8: WVD of a monochromatic signal

Since $s(t) \otimes K\delta(t-t_0) = Ks(t-t_0)$, (2.8) is expressed as:

$$\begin{aligned} W(t,\omega) &= 2\pi\left[\delta(2\omega+\omega_0)+\delta(2\omega-\omega_0)\right]\exp(j\omega t)\otimes_\omega \\ &\quad \left[\delta(2\omega+\omega_0)+\delta(2\omega-\omega_0)\right]\exp(j\omega t) \\ &= 2\pi\left[\exp(-j\frac{\omega_0}{2}t)+\exp(j\frac{\omega_0}{2}t)\right]\otimes_\omega\left[\exp(-j\frac{\omega_0}{2}t)+\exp(j\frac{\omega_0}{2}t)\right] \\ &= 2\pi\left[\exp(-j\omega_0 t)\delta(\omega-\omega_0)+\exp(j\omega_0 t)\delta(\omega+\omega_0)+2\delta(\omega)\right] \end{aligned} \tag{2.11}$$

The last expression shows energy contained at $\omega = 0$, which does not correspond to the actual signal. They are called cross-terms due to the bilinear product. For illustration, supposing $\omega_0 = 0.1F_s$ (F_s is sampling frequency), the WVD is portrayed in Fig. 2.8. It can be observed that besides the two expected signal components, there is energy located in the centre at $\omega = 0$. The cross-term is a major drawback of the WVD. It appears when the input signal is multi-component, cluttering the time-frequency signal representation. This could, in turn, lead to a misinterpretation of local signal power concentrations and a misreading of the signal time-frequency signature, including the instantaneous frequency. To have a better understanding about the cross-terms, we consider a signal composing of two components,

$$s(t) = s_1(t) + s_2(t). \tag{2.12}$$

Substituting this into the definition in (2.7), we have:

$$\begin{aligned} W(t,\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(s_1^*(t-\frac{1}{2}\tau)+s_2^*(t-\frac{1}{2}\tau)\right)\left(s_1(t+\frac{1}{2}\tau)+s_2(t+\frac{1}{2}\tau)\right)e^{-j\tau\omega}d\tau \\ &= W_{11}(t,\omega)+W_{22}(t,\omega)+W_{12}(t,\omega)+W_{21}(t,\omega), \end{aligned} \tag{2.13}$$

where

$$W_{12}(t,\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} s_1^*(t-\frac{1}{2}\tau)s_2(t+\frac{1}{2}\tau)e^{-j\tau\omega}d\tau. \tag{2.14}$$

This is called the cross Wigner distribution. As $W_{12}(t,\omega) = W_{2,1}(t,\omega)$, hence

$$W(t,\omega) = W_{11}(t,\omega)+W_{22}(t,\omega)+2\Re(W_{12}(t,\omega)). \tag{2.15}$$

It can be seen that the Wigner distribution of the sum of two signals is not the sum of the Wigner distribution of each signal but has an additional term $2\Re(W_{12}(t,\omega))$. This term is called a cross-term or interference, which gives the artifact in the time-frequency domain.

In the case of a discrete signal $s(n)$ $(n = 1, 2, ..., N)$, the Wigner distribution is defined as follows:

$$W(k,n) = \sum_{b=-N}^{N-1} s(n+\frac{1}{2}b)s^*(n-\frac{1}{2}b)\exp(-j2\pi kb), \tag{2.16}$$

where n and k denote discrete time and frequency and b is the discrete time-lag ($b = -N/2, -N/2+1, ..., N/2-1$ if N is even). However, it is impossible to calculate $W(k,n)$ if using (2.16) because $s(n+\frac{1}{2}b)$ (b odd) does not exist. Thus, the equation is altered as follows:

$$W(k,n) = \sum_{b=-N/2}^{N/2-1} s(n+b)s^*(n-b)\exp(-j2\pi kb). \tag{2.17}$$

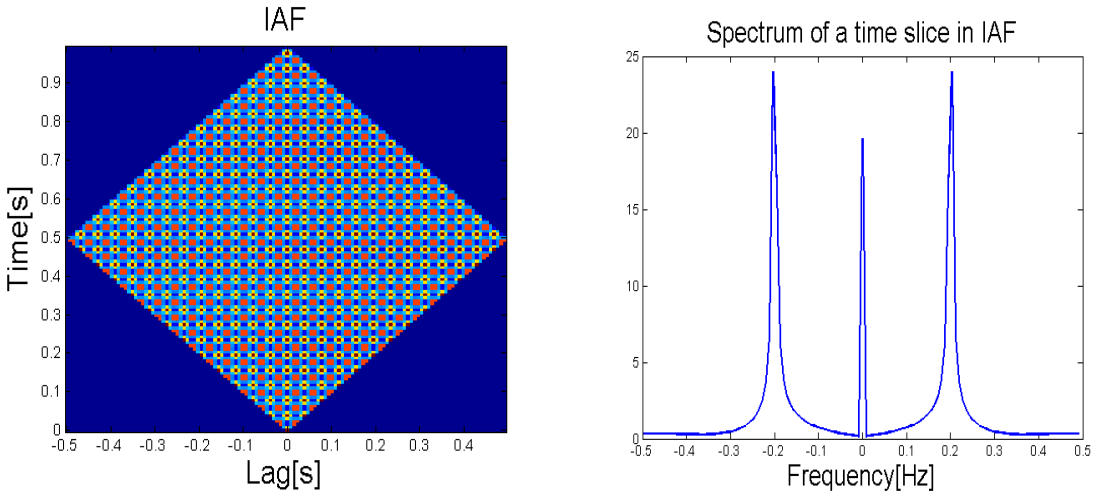
The bilinear product of the signal, $s(n+b)s^*(n-b)$, is called the instantaneous autocorrelation function (IAF). It is a joint time and lag domain. If N (N even) is length of signal $s(n)$, then the values of lag b are in the range of $[-N/2, N/2-1]$. For each value of b, n needs to satisfy:

$$0 \le n+b \le N$$
$$0 \le n-b \le N$$

or $|b| \leq n \leq N - |b|$. According to (2.17), the Wigner distribution is thus simply the Fourier transform of the IAF over the lag variable. The Fourier transform of the IAF over the time variable is called the ambiguity function (AF), which we will consider in the next section. The IAF of the cosine signal with frequency ω_0 is displayed in Fig. 2.9(a), and is expressed as,

$$\begin{aligned} IAF &= \cos[\omega_0(n+b)]\cos[\omega_0(n-b)] \\ &= \frac{1}{2}\left[\exp(j\omega_0(n+b)) + \exp(-j\omega_0(n+b))\right]\frac{1}{2}\left[\exp(j\omega_0(n-b)) + \exp(-j\omega_0(n-b))\right] \\ &= \frac{1}{4}\left[\exp(j2\omega_0 b) + \exp(-j2\omega_0 b) + \exp(j2\omega_0 n) + \exp(-j2\omega_0 n)\right]. \end{aligned}$$

It can be seen that the larger the absolute value of lag, then the value range of time



(a) (b)

Figure 2.9: (a) The IAF of a cosine; (b) Spectrum of a time slice in IAF domain.

is smaller, and with a certain time slice, the IAF contains two sinusoids at $\pm 2\omega_0$ and a constant as illustrated in Fig. 2.9(b).

2.3.3 The Cohen's class

The spectrogram and Wigner-Ville Distribution are members of the general class of time-frequency representations known as Cohen's class. In the most general form, a

time-frequency representation in Cohen's class, $D(n,k)$ with kernel function $C(p,b)$ can be represented by [5]:

$$D(n,k) = \sum_{p=-N/2}^{N/2-1} \sum_{b=-N/2}^{N/2-1} AF_s(p,b)C(p,b)e^{-j2\pi pn}e^{-j2\pi kb}, \tag{2.18}$$

where n and k are discrete time and frequency values and p and b are the discrete Doppler shift and time lag. $AF(p,b)$ is the ambiguity function (AF) of the signal, expressed as follows:

$$AF_s(p,b) = \sum_{n=0}^{N-1} s(n+b)s^*(n-b)e^{-j2\pi np/N}. \tag{2.19}$$

It is obvious that the ambiguity function is the Fourier transform over the time variable of the signal bilinear product, or the IAF, and the Cohen TF representation is simply the two-dimensional Fourier transform of the multiplication of the signal AF and the kernel in the ambiguity domain over Doppler and lag variables. We can also say that the AF is the inner product between the analyzed signal and the time-frequency shifted signal $(T_{p,b}s)(n) = s(n-b)e^{-j2\pi pn/N}$ [2]. It is expressed as,

$$A(p,b) = \langle s, T_{p,b}s \rangle$$

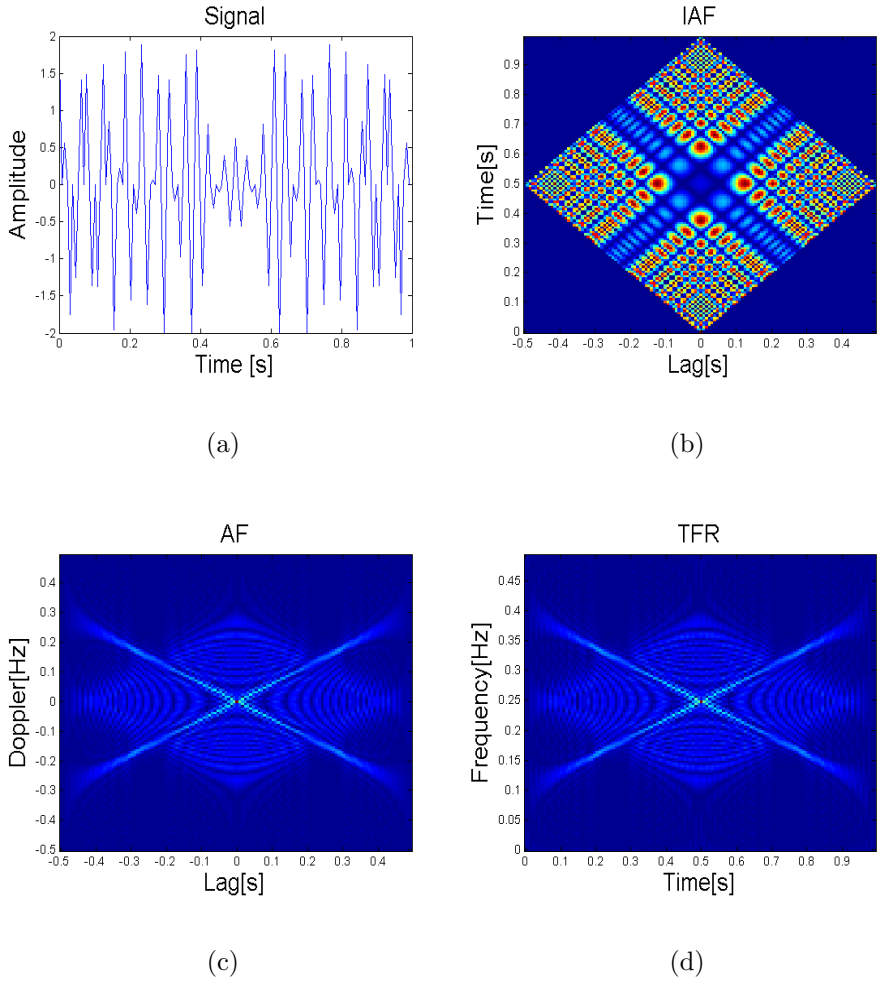
Therefore, the AF can be viewed as a TF correlation function, and have properties such as Hermitian symmetry, and the fact that its modulus is maximum at the origin. In the case of multiple signals, the total AF consists of auto-components neighbouring the origin of the plane and cross terms mostly locate some distance from the origin which directly depends on the TF separation between the signal components. To mitigate this interference and preserve the auto components, kernel functions are deployed. They are usually in the form of a two-dimensional low-pass filter as the auto-terms locate near the origin. Some well-known kernels includes Choi-Williams [18], Margenau-Hill [19] and Born-Jordan [19]. They are expressed in Table 2.1. To illustrate the four domains (time, ambiguity, instantaneous auto-correlation, TF) and Cohen's class TFR, we use a signal composed of two crossing linear chirps written as,

$$\begin{aligned} s(n) = &\exp\left\{j2\pi\left[(0.1F_s)\frac{n}{F_s} + (0.3F_s)\frac{n^2}{2F_s^2}\right]\right\} \\ &\exp\left\{j2\pi\left[(0.4F_s)\frac{n}{F_s} - (0.3F_s)\frac{n^2}{2F_s^2}\right]\right\}, \end{aligned} \tag{2.20}$$

Table 2.1: Examples of kernels

Distribution	**Kernel $\phi(p, b)$**
Choi-Williams	$\exp(-\frac{p^2b^2}{2\sigma^2})$
Margenau-Hill	$\cos(pb/2)$
Born-Jordan	$\mathrm{sinc}(pb/2)$

where $F_s = 128$ is sampling frequency. The four domains are shown in Fig. 2.10.

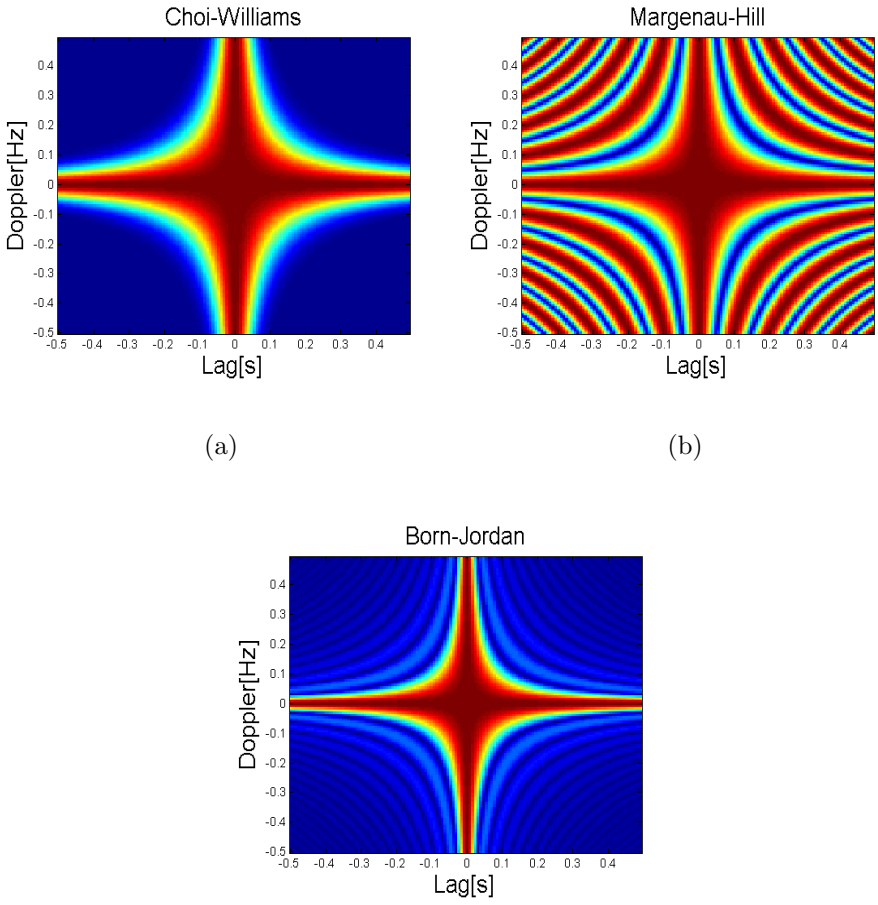


(a) (b)

(c) (d)

Figure 2.10: The signal in (2.20) in (a) The time domain; (b) The ambiguity domain; (c) The instantaneous auto-correlation domain; (d) The time-frequency domain.

The cross-terms are clearly present in the ambiguity domain due to the bilinear product of the multicomponent signal. This results in interference in the TF plane. The three aforementioned kernels are shown in Fig. 2.11. The Choi-Williams kernel uses $\sigma = 10$. The ambiguity domain after applying the kernel, and the resulting TF



(a) (b)

(c)

Figure 2.11: (a) Choi-Williams kernel; (b) Margenau-Hill kernel; (c) Born-Jordan kernel.

distribution are shown in Fig. 2.12.

For this signal, the Choi-Williams kernel is ineffective as only a small amount of the cross-terms are suppressed at the expense of omitting some parts of auto-terms.

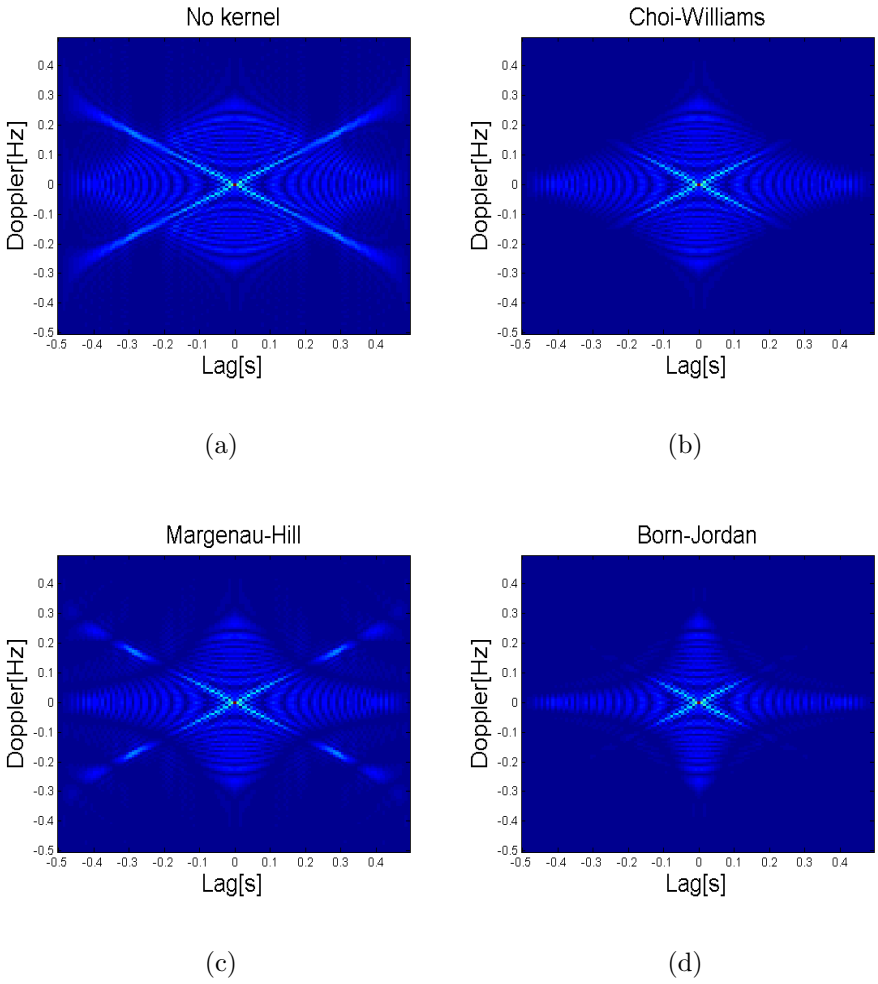


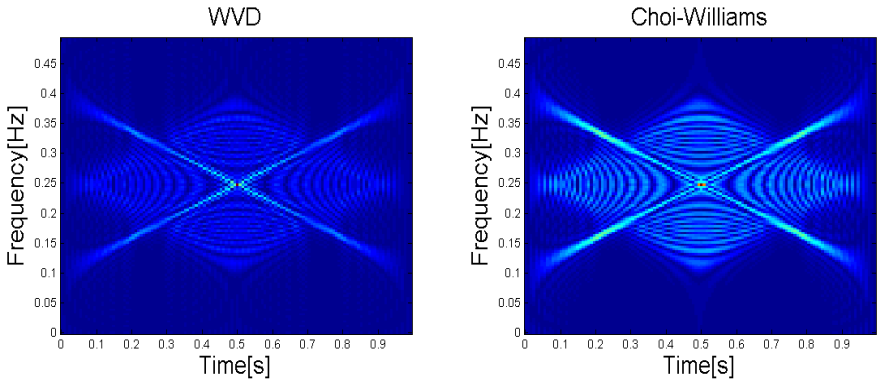
Figure 2.12: Ambiguity function with: (a) No kernel ; (b) Choi-Williams kernel; (c) Margenau-Hill kernel; (d) Born-Jordan kernel.

The other kernels are better, but also cannot eliminate the interference. The TF distributions are shown in Fig. 2.13.

The Wigner distribution is a member in the Cohen's class. The relation between the WVD and the Cohen's class can be written as:

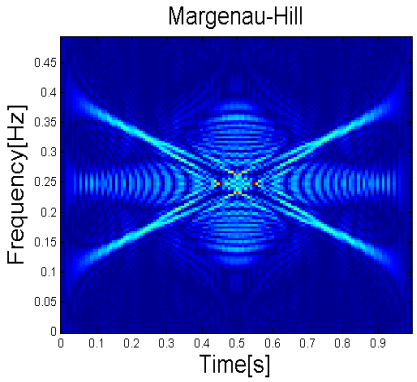
$$D(n,k) = \frac{1}{4\pi^2} \sum_u \sum_v \bar{C}(u-n, v-k) W(u,v), \tag{2.21}$$

where $\bar{C}(n,k)$ is the two-dimensional Fourier transform of the kernel $C(p,b)$. Thus,

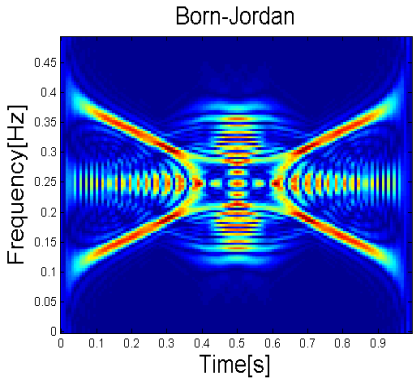


(a)

(b)



(c)



(d)

Figure 2.13: TF distribution when applying: (a) No kernel (WVD); (b) Choi-Williams kernel; (c) Margenau-Hill kernel; (d) Born-Jordan kernel.

Cohen's TFD is WVD if $\bar{C}(n,k) = \delta(n)\delta(k)$ or $C(p,b) = 1$. It can be seen there is no filtering effect in the WVD. The spectrogram also belongs to the Cohen's class. It is obvious if we rewrite (2.6),

$$\begin{aligned} P_{SP}(n,k) &= \left| \frac{1}{2\pi} \sum_{b} s(b)h(b-n)e^{-j2\pi kb} \right|^2 \\ &= \sum_{u} \sum_{v} W_h(u-n, v-k)W(n,k). \end{aligned} \tag{2.22}$$

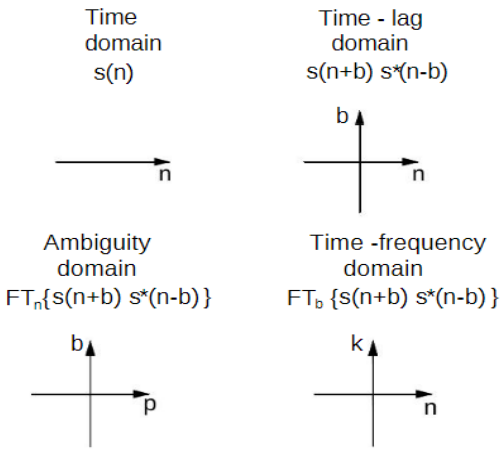


Figure 2.14: The four domains.

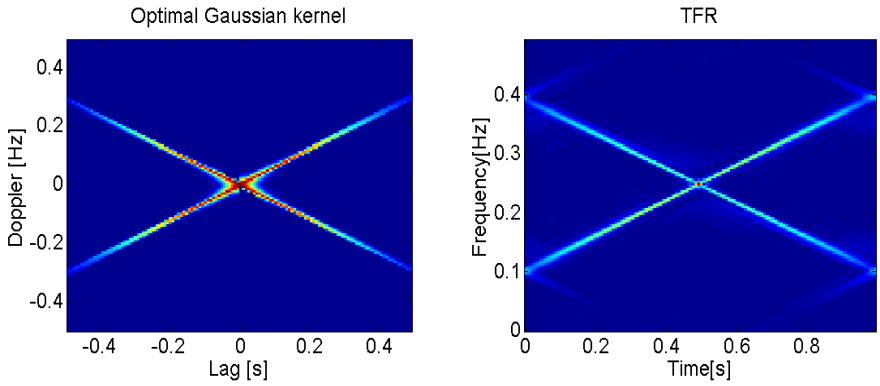
So, the spectrogram is also Cohen's TFD with kernel $\mathbf{C}(n,k) = W_h(n,k)$. The above conversion is based on the unitary property, which is written as:

$$\left| \int s(t)x^*(t)dt \right| = \int \int W_s(t,\omega)W_x^*(t,\omega)dtd\omega.$$

To summarize, the Cohen's class is built up from the relationship among four domains (time, time-frequency, time-lag, lag-Doppler). The relation is shown in Fig. 2.14. In general, the Cohen's class employs the signal-independent kernel, which cannot eliminate all interference and also simultaneously keep the desired terms. So, later researchers have suggested many improved methods, which are called adaptive kernels or signal-dependent kernels. The signal-dependent radially Gaussian kernel is the most well-known method in this category [20]. The filtered the ambiguity domain and the resulted TF signature are shown in Fig. 2.15. This algorithm almost removes all the interference, obtaining the best performance compared to the other mentioned kernels.

2.4 Conclusion

This chapter has provided the basic of TFDs. All these conventional TFD approaches have their own strengths and weaknesses. The Hilbert transform is the



(a) (b)

Figure 2.15: Signal-dependent radially Gaussian kernel: (a)Filtered ambiguity domain; (b)TF representation.

most simple method to detect the signal's instantaneous frequency. But it can only achieve the correct estimation if the signal has only one component. The Hilbert-Huang transform can detect the time-frequency signature of a multicomponent signal, but it is vulnerable to noise. Quadratic TFDs including the short-time Fourier transform (or spectrogram) and the Wigner-Ville distribution can be applied to a wide range of signals, and give pretty good results. However, the former method suffers from the trade-off between the time and the frequency resolutions. The latter witnesses cross-terms, which clutter the ambiguity and TF domain and hide the desired signal structures. By applying some kernels, which belong to Cohen's class or optimal masks, the TF signature estimations get improved. However, these methods are not designed to deal with missing samples, which can frequently occur due to noise, fading, or intentionally under-sampling. Thus, other methods which are robust with incomplete data are now needed.

Chapter 3

Sparse Time-Frequency Distribution Fundamentals

3.1 Introduction

As we know, TFD is the most efficient way to process a non-stationary signal as it reveals the signal's signatures in such a vivid manner. We have discussed conventional TFD methods such as the spectrogram, the WVD and the Cohen's class approaches. Each method has its own weaknesses, and all of them give noisy TFRs in the presence of missing data. The fact that most non-stationary signals are sparse in the TF domain enables compressive sensing (CS) to be applied in TFD, which we call sparse TFD. The CS allows reconstruction of the entire signal from its small randomly chosen set of measurements. Therefore, with CS techniques, we can obtain reliable TFRs even in the case of missing data. In this chapter, we will study closely the motivation of using CS, especially using CS in TFDs, the fundamentals of CS techniques and some recent TFD approaches using CS.

3.2 Motivation for Compressive Sensing

Now as we are entering the era of big data, the resolution or typical number of samples in a certain signal (image, video, etc.) get larger. Denote N as the necessary samples. The continuing growth of N places a burden on every stage of the data processing pipeline from acquisition, storage to analysis. In order to control and reduce

the cost for collecting and processing high-dimensional data, it is essential to exploit models that could encapsulate prior information regarding the signal's interest [1]. Based on the fact that N-sample signals actually have far fewer then N degree of freedom, called $E(E \ll N)$, there have been many dimensionality reduction techniques suggested. They include data compression, feature extraction and parameter estimation. Nevertheless, these approaches require N samples of signals to be collected before we can identify and exploit its intrinsic low-dimensional structure. So, even if these signals have E degrees of freedom, we cannot exploit this information, transform coefficients for example, until N samples are obtained. It means that the sensing processing for many signals may be unnecessarily wasteful. And after all, sensing high-bandwidth and high-resolution signals can require expensive hardware and consume valuable power, etc. Therefore, it raises a question of whether it is possible to incorporate the dimensionality reduction in the sensing process itself. It means that we deliberately measure fewer samples with the expectation that the missing samples can later be reconstructed from the recorded ones. This is also the central idea that underlies compressive sensing (CS), taking a small number M of linear measurements of a signal (slightly more than E, but far fewer than N), and from these measurements reconstruct the complete set of all N samples that a conventional sensor would have recorded.

Usually, non-stationary signals are locally sparse in TF domain. It means that most of the transform coefficients are small and only a few (say, E) are large. Therefore, we can apply CS to obtain the TFD especially in the case of incomplete data.

3.3 Compressive Sensing Overview

3.3.1 Sparsity, compressibility and norms [1]

A real- or complex-valued length N vector $\mathbf{x}$ is said to be E sparse if it contains only E non-zero entries. Or we can say its l_0 norm is E, i.e., $\|\mathbf{x}\|_0 = E$. Often we deal with approximations to sparse signals, called compressible signals. We refer to the set of positions of the non-zero entries of $\mathbf{x}$ as the *support* of $\mathbf{x}$ and denote this by $\text{supp}(\mathbf{x})$. For any $\mathbf{x}$, $|\text{supp}(\mathbf{x})| = \|\mathbf{x}\|_0$.

For any vector $\mathbf{x} \in \mathbb{R}^N$ or $\mathbb{C}^N$, we let $\mathbf{x}_E$ denote the nearest E-sparse vector to $\mathbf{x}$. This can be obtained simply by keeping the E entries of $\mathbf{x}$ with the largest magnitude and setting all remaining entries to 0. If the distance from $\mathbf{x}$ to $\mathbf{x}_E$ is small (but not necessary zero), $\mathbf{x}$ is said to be compressible. The CS technique is used to find the sparsest solution that is close to the signal. Because the l_0 norm does not meet the formal mathematical definition of a norm, we express the CS problem in l_1 and l_2 norm instead. The l_1 norm measures the absolute sum of entries of $\mathbf{x}$:

$$\|\mathbf{x}\|_1 = \sum_{n=1}^{N} |\mathbf{x}_n|, \tag{3.1}$$

while the l_2 norm measures the sum of the squared magnitudes of the entries of $\mathbf{x}$:

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{n=1}^{N} |\mathbf{x}_n|^2}. \tag{3.2}$$

The l_1 norm has a special connection to sparsity: it tends to be small for sparse signals. For two vector with the same l_2 norm, the one with fewer non-zero entries will generally have the smaller l_1 norm. Therefore, we can find the sparsest solution by searching for the solution with smallest l_1 norm.

3.3.2 Compressive sensing problem in a nutshell [1]

Let f denote a desired collection of N samples of a signal. For convenience, we arrange these samples into a vector $\mathbf{f}$, $\mathbf{f} \in \mathbb{R}^N$ or $\mathbb{C}^N$; in cases where these are samples of an image or other multidimensional signal, the pixel values can be stacked into the vector $\mathbf{f}$ using any ordering rule. Sometimes, the signal vector $\mathbf{f}$ may itself be sparse or compressible. It means it may contain just a few significant entries, or most of their entries are zero or nearly zero. For example, an astronomical image could have only a few pixels illuminated by stars [21]. However, the sparse structure of a signal is commonly revealed only when that signal is transformed into an appropriate domain. For instance, if a signal f is composed of a few harmonic tones, then the vector containing the discrete Fourier transform (DFT) coefficients of f or the discrete cosine transform (DCT) coefficients of f will be sparse. Denote

$\mathbf{\Psi}$ as a NxN real- or complex-valued basis matrix and denote $\mathbf{x}$, $\mathbf{x} \in \mathbb{R}^N$ or $\mathbb{C}^N$ as the sparse coefficient vector, then $\mathbf{f}$ is expressed as:

$$\mathbf{f} = \mathbf{\Psi x} \tag{3.3}$$

In case that $\mathbf{f}$ is already sparse, $\mathbf{\Psi}$ is identity matrix.

A typical CS problem looks like this: we do not attempt to record the N entries in $\mathbf{f}$ directly. Instead we record a smaller number $M (M \ll N)$ of linear measurements of $\mathbf{f}$; suppose these are arranged into an Mx1 vector we call $\mathbf{y}$. Notice that linear measurements are the result of linear operations on the signal: filtering, modulation, sampling, etc. Since these measurements are linear, we can represent the measurement vector $\mathbf{y}$ in the presence of noise as:

$$\mathbf{y} = \mathbf{\Phi f} + \mathbf{v} = \mathbf{\Phi \Psi x} + \mathbf{v} = \mathbf{Ax} + \mathbf{v}, \tag{3.4}$$

where $\mathbf{\Phi}$ is an MxN matrix we refer to as the measurement matrix and $\mathbf{v} \in \mathbb{R}^N$ or $\mathbb{C}^N$ is a vector of measurement noise. These vectors and matrices can be real- or complex-valued. Because we are interested in scenarios where the number of measurements M is smaller then the number of samples N, the vector $\mathbf{y}$ is often said to contain compressive measurements of $\mathbf{f}$, and the problem in (3.4) is underdetermined, which has an infinite number of candidate solutions. CS is proposed to obtain the closest estimation to $\mathbf{f}$.

Typically in CS, $\mathbf{\Phi}$ is designed with some element of randomness. In some cases, it can be appropriate to collect just a random set, say, 10% of the pixel in an image. In this case, $\mathbf{\Phi}$ is a binary matrix containing a single randomly positioned 1 in each row. In other cases, it can be appropriate to collect a random set of frequency domain measurements of a signal, for example by recording 10% of the Fourier transform of $\mathbf{f}$. In this case, $\mathbf{\Phi}$ contains a random set of M rows of the NxN discrete Fourier transform (DFT) matrix. In other cases, every measurement in $\mathbf{y}$ might be a random linear combination of all of the entries in $\mathbf{f}$. In this case, $\mathbf{\Phi}$ might be populated with independent and identically distributed (i.i.d) Gaussian or Rademacher (± 1) random variables.

Most CS recovery algorithms can be interpreted as solving for a sparse vector $\hat{\mathbf{x}}$ that satisfies $\mathbf{y} \approx \mathbf{A}\hat{\mathbf{x}}$ as closely as possible. Once the sparse coefficient vector

has been estimated, one can synthesize a signal estimate via multiplication by $\mathbf{\Psi}$, $\hat{\mathbf{f}} = \mathbf{\Psi}\hat{\mathbf{x}}$.

The solution is often expressed as:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|\mathbf{x}\|_1, \qquad \text{s.t.} \qquad \|\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}\|_2 \leq \epsilon, \tag{3.5}$$

where ϵ is a small predefined parameter. The techniques for recovering $\mathbf{f}$ from $\mathbf{y}$ and $\mathbf{\Phi}$ can be interpreted as searching among the candidate solutions to the equation $\mathbf{y} = \mathbf{\Phi f}$ for the one that best matches the low-dimensional model. In the case of a sparse signal model, which is by far the most commonly used model in CS, one might look for the sparsest candidate solution in some known basis. There are a variety of algorithms available for searching for this sparsest candidate solution; some involve convex optimization, while the others involve iterative greedy methods. Under certain assumptions on the random measurement protocol, an E-sparse signal $\mathbf{f}$ can be recovered from a number of measurements M which is proportional to $E\log(N/E)$. Thus the number of measurements can be significantly smaller than N and only greater than the information level E by a logarithmic factor; this logarithmic factor is the price one pays for not knowing the locations of sparse coefficient in advance. Remarkably, in the absence of noise and assuming $\mathbf{f}$ is exactly sparse, the recovery is exact. In the presence of noise or assuming $\mathbf{f}$ is nearly sparse, the recovery is provably robust.

3.3.3 Conditions for reliable recovery

Matrices that satisfy a condition known as the restricted isometry property (RIP) can be proved to allow the recovery of sparse signals via efficient algorithms. The matrix $\mathbf{A} = \mathbf{\Phi\Psi}$ is said to satisfy the RIP of order E if there exists a constant $\delta_E \in (0, 1)$ such that:

$$(1 - \delta_E)\|\mathbf{x}\|_2^2 \leq \|\mathbf{Ax}\|_2^2 \leq (1 + \delta_E)\|\mathbf{x}\|_2^2, \tag{3.6}$$

holds for all coefficient vectors $\mathbf{x}$ with $\|\mathbf{x}\|_0 \leq E$. The parameter δ_E is known as the isometry constant of order E. The RIP is essentially a requirement that in the matrix $\mathbf{A}$ any sub-matrix containing E columns will be approximate isometric (its E columns will be approximately orthonormal). However, it has been shown

that checking the RIP holds for a given matrix with a specified isometry constant is NP-hard in general [1]. Fortunately, this property can be guaranteed to hold with very high probability under suitable conditions. For example, let $\mathbf{\Psi}$ be an arbitrary fixed $N \,\mathrm{x}\, N$ dictionary in $\mathbb{R}^N$ or $\mathbb{C}^N$, and $\mathbf{\Phi}$ be a $M \,\mathrm{x}\, N$ matrix populated with i.i.d. sub-Gaussian entries having mean zero and variance $1/M$. If

$$M \geq C_1 \left(E \log \frac{N}{E} + \log \frac{1}{\rho} \right), \tag{3.7}$$

then with probability at least $1 - \rho$, A will satisfy the RIP of order E with isometry constant δ_E. Typically, $C_1 = 1/\delta_E$.

3.3.4 CS algorithms and orthogonal matching pursuit OMP

As discussed above, there are various algorithms for searching the sparsest candidate solutions: some relate to convex optimization, others involve the iterative greedy methods. Here, we introduce a popular greedy method, orthogonal matching pursuit (OMP) which is usually employed in sparse TF reconstruction due to its simplicity.

We now address the problem about how to recover a signal $\mathbf{f}$ from a vector $\mathbf{y}$, $\mathbf{y} = \mathbf{\Phi f} + \mathbf{v}$. Assumed that $\mathbf{f}$ is E-sparse or compressible in some orthonormal basis $\mathbf{\Psi}$. Thus, we can write:

$$\mathbf{y} = \mathbf{\Phi f} + \mathbf{v} = \mathbf{\Phi \Psi x} + \mathbf{v} = \mathbf{Ax} + \mathbf{v}, \tag{3.8}$$

where
$\mathbf{y}$ is vector of length M,
$\mathbf{A} = \mathbf{\Phi \Psi}$ is a MxN matrix,
$\mathbf{x}$ is a E-sparse or compressible vector of length N, and
$\mathbf{v}$ is a noise vector of length M.

Orthogonal matching pursuit (OMP) is a greedy algorithm for CS recovery [22, 23, 24, 25]. The intuition behind OMP is roughly as follows. Since $\mathbf{x}$ has only E non-zero components, the data vector $\mathbf{y}$ is a linear combination of E columns from the measurement matrix $\mathbf{A}$. In the language of sparse approximation, we say that $\mathbf{y}$ has an E-term representation over the dictionary $\mathbf{A}$. To identify the signal $\mathbf{x}$, we need to determine which column of $\mathbf{A}$ participates in data vector $\mathbf{y}$. The idea behind this algorithm is to pick columns in a greedy fashion. At each iterations,

we choose the column of $\mathbf{A}$ which is most strongly correlated to the remaining part of $\mathbf{y}$. Then we subtract off its contribution to $\mathbf{y}$ and iterate on the residual. After E iterations, we will identify the correct set of columns, which also means we will determine the vector $\mathbf{x}$. Also, OMP can also stop when the residual is less than a threshold. The threshold is often the noise level, which can be determined by SNR as following:

$$A_{\text{Noise}} = \frac{A_{\text{Signal}}}{(\frac{SNR}{20})^{10}}, \tag{3.9}$$

where A_{Noise} and A_{Signal} are the norm 2 of the noise and the signal. The algorithm is summarized as follows:

INPUT:

- An $[M \,\text{x}\, N]$ measurement matrix $\mathbf{A}$
- An $[M \,\text{x}\, 1]$ data vector $\mathbf{y}$
- Sparsity level E of unknown signal $\mathbf{x}$ $[N \,\text{x}\, 1]$

OUTPUT:

- An estimate $\hat{\mathbf{x}}$
- A vector $\mathbf{\Lambda}$ containing the index of E non-zero elements of $\mathbf{x}$

PROCEDURE:

1. Call $\mathbf{A}_i$, $\mathbf{\Lambda}_i$ the matrix and the index set of the chosen columns from measurement matrix $\mathbf{A}$ after i^{th} iteration. Denote $a_j (j = 1, 2, ..., N)$ as a column of matrix $\mathbf{A}$. Initialize the residual $\mathbf{r}_0 = \mathbf{y}$, $\mathbf{A}_0 = \emptyset$, $\mathbf{\Lambda}_0 = \emptyset$, and the iteration counter $i = 1$ $(i \leq E)$.
2. Find the index λ_i by $\lambda_i = \arg\max_{j=1,2,...,N} | < \mathbf{r}_{i-1}, a_j > |$.
3. $\mathbf{\Lambda}_i = \mathbf{\Lambda}_{i-1} \cup \lambda_i$, $\mathbf{A}_i = [\mathbf{A}_{i-1}, a_{\lambda_i}]$.
4. Solve the least squares problem to obtain a new signal estimate, $\mathbf{z}_i = \arg\min_z \|\mathbf{A}_i \mathbf{z} - \mathbf{y}\|_2$.
5. Calculate new approximation of the data $\mathbf{y}_i$, and new residual $\mathbf{r}_i$: $\mathbf{y}_i = \mathbf{A}_i \mathbf{z}_i$, $\mathbf{r}_i = \mathbf{y} - \mathbf{y}_i$.

6. Increment i, and return to step 2 if $i < E$ (or when $\|r\|_2$ is less than a predefined threshold).
7. Column vector $\mathbf{x}$ has non-zero elements at index listed in $\mathbf{\Lambda}$, and their non-zero values are shown in vector $\mathbf{z}$.

Our example used to illustrate the OMP method is from [26]. The basic $\mathbf{\Psi}$ is the discrete cosine transform (DCT). The signal generated by the "A" key on a touch-tone telephone is the sum of two sinusoids with different integer frequencies,

$$f(t) = \sin(2\pi\, 697\, t) + \sin(2\pi\, 1633\, t).$$

If we sample this tone for 1/8 of a second at sampling rate $F_s = 40000$ Hz, the result is a column vector $\mathbf{f}$ of length $N = 5000$. Call $\mathbf{x}$ is the coefficient vector obtained by taking the inverse DCT transform of $\mathbf{f}$. The signal can be expressed as:

$$\mathbf{f} = \mathbf{\Psi x}, \tag{3.10}$$

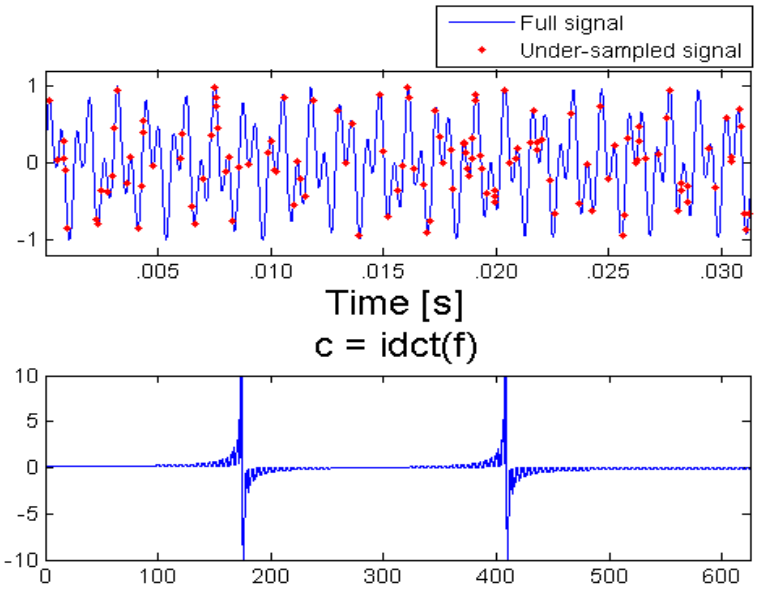
where $\mathbf{\Psi}$ is an NxN DCT matrix. The signal $\mathbf{f}$ and coefficient vector $\mathbf{x}$ are plotted in Fig. 3.1(a). We randomly pick $M = \lfloor N/10 \rfloor$ samples from $\mathbf{f}$ to get the measurement vector $\mathbf{y}$. We construct a matrix $\mathbf{A}$ by extracting M rows from the DCT matrix $\mathbf{\Psi}$ [NxN] in the same rule with which we get $\mathbf{y}$. We can write:

$$\mathbf{y} = \mathbf{\Phi \Psi x} = \mathbf{Ax}, \tag{3.11}$$

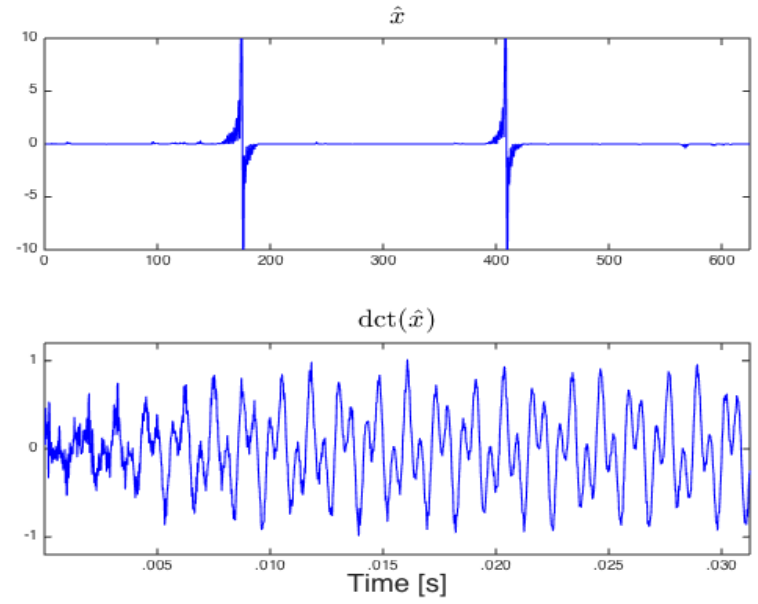
where $\mathbf{\Phi}$ [MxN] is a binary matrix containing a single randomly positioned 1 in each row. The simulation result is shown in figure 3.1.

3.4 Motivation for Applying CS in TFD

A large number of signals that appear in real applications (array signal processing, indoor and synthetic aperture radar imaging, communications, remote sensing, and biomedical and multimedia applications) are sparse in the TF domain. In general, a signal that is E-sparse in a specific domain can be completely characterized by M measurements ($M > E$) and CS techniques, although the total number of samples required by the Shanon-Nyquist theorem is far above M. Therefore, most non-stationary signals can be recovered by a small number of samples. The questions



(a)



(b)

Figure 3.1: (a) Signal and inverse DCT coefficients obtained from the full signal $\mathbf{f}$; (b) Recovered $\hat{\mathbf{x}}$, and $\hat{\mathbf{f}}$ by OMP.

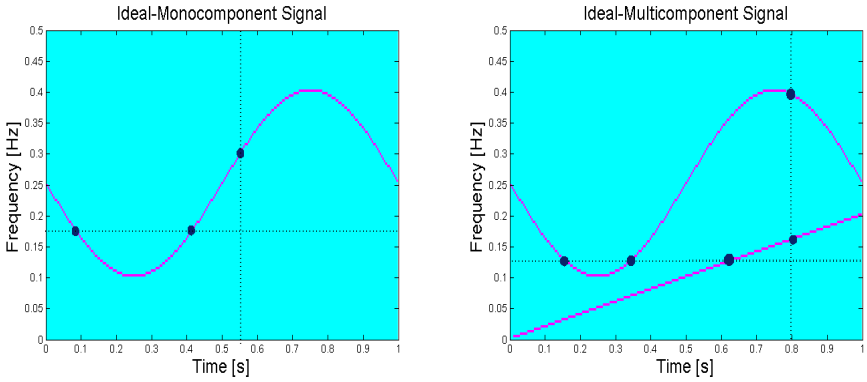
are why we only have a few samples and what happens if we apply conventional TFD like the spectrogram and the WVD, not using CS techniques in the case of incomplete data. As discussed above, missing entries may appear because we deliberately undersample to reduce the burden on the data processing pipeline from acquisition, storage and analysis. They also occur due to the removal of "unwanted components" such as clutter and noise. In both scenarios, the missing samples produce noise-like effect, which significantly degrades signal representation in the TF domain if we use the traditional TFD methods. Thus, a sparsity-aware TFD approach is of significance in TF analysis for non-stationary signals.

3.4.1 Sparsity property of non-stationary signals

An ideal TF representation of a single frequency modulated (FM) signal in Fig. 3.2(a) shows [27]:

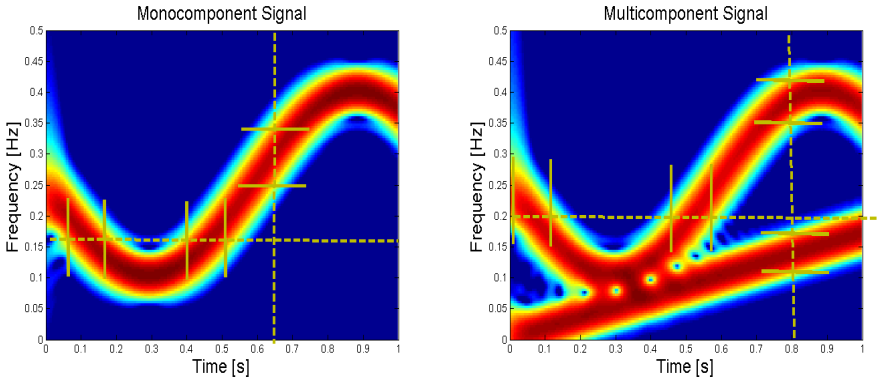
- Sparsity in the two-dimensional (2D) TF domain by the virtue of perfect power localization of FM signals, making it appear as a wavy line in a plane populated by zero values. As such, for a single FM component, the joint-variable TF representation is N-sparse.
- Local frequency sparsity, or the number of non-zero entries on a time-slice for a single FM component, is one.
- Local time sparsity, or the number of non-zero entries on a frequency-slice for a single FM component, is slightly more than one.

The above three cases are depicted in Fig. 3.2(a) using a sinusoid FM signal as an example. For multi-component signals, where each component is defined by a frequency law, there would be more non-zeros entries, which results in a reduction in sparsity when compared with mono-component signals. This is also illustrated in Fig. 3.2(b). With non-ideal representation, the time and frequency points are replaced by short windows and narrow-band filters as shown in Fig. 3.3. However, in general, non-stationary representations in the TF domain are populated by zero values. Thus, we can say that non-stationary signals are sparse locally and globally in the TF domain.



(a) (b)

Figure 3.2: Ideal TFDs for (a) A mono-component signal; (b) A multi-component signal.



(a) (b)

Figure 3.3: Spectrogram for the given (a) Mono-component signal; (b) Multi-component signal in Fig. 3.2.

3.4.2 Missing data effects

We know that missing samples generally cause artifacts in both the TF domain and the ambiguity domain. So, why do these noise-like effects appear? As discussed

above, TFD can be calculated as the Fourier transform of the instantaneous autocorrelation function (IAF) over the lag variable, and the AF is the Fourier transform of the IAF over the time variable. Also, the TFD can be obtained by taking the two-dimensional Fourier transform of the AF. So, we can say the Fourier transform of incomplete data might be the reason. So, in this part, we will discuss two issues:

- Missing samples in time lend themselves to missing entries in the instantaneous autocorrelation domain.
- Analysis of missing samples in the Fourier transform domain.

IAF with missing samples in time [28]

Missing samples in time generate noise-like artifacts in the ambiguity and the time-frequency domains, but only cause missing entries in autocorrelation domain. This part discusses how these missing points are distributed in the instantaneous autocorrelation domain and the maximum number of missing points presented in a time slice or at a specific time lag.

Consider the above discrete signal $s(n)$, $n = 1, 2, ..., N$. The IAF of $s(n)$ is as follows:

$$IAF(n, b) = s(n + b)s^*(n - b), \tag{3.12}$$

where b is lag variable, $-N/2 \ll b \ll N/2 - 1$. The IAF is affected by the window effect due to zero-padding. The length of each time slice in IAF depends on lag b and is express as:

$$W(b) = N - 2|b|. \tag{3.13}$$

With higher values of $|b|$, number of entries gets smaller. And this figure gets smaller if we have missing samples. Denote $m(n)$, $miss(n)$ and N_{miss} $(0 \leq N_{miss} < N)$ as observation data, missing data and number of missing samples, respectively. So, $m(n)$ can be expressed by the product of $s(n)$ and an observation mask, $M(n)$:

$$m(n) = s(n)M(n), \tag{3.14}$$

where

$$M(n) = \begin{cases} 1, & \text{if } n \in \mathbb{S}, \\ 0, & \text{if } n \notin \mathbb{S}, \end{cases} \tag{3.15}$$

where $\mathbb{S} \subset \{1, ..., N\}$ is the set of observed time instants and its cardinality is $|\mathbb{S}| = N - N_{miss}$. Similarly, $miss(n)$ is written as the product of the full signal $s(n)$ and the missing data mask $Miss(n)$:

$$miss(n) = s(n)Miss(n), \tag{3.16}$$

where

$$Miss(n) = S(n) - M(n) \qquad \text{with} \qquad S(n) = 1, \ \forall n \in \{1, ..., N\}. \tag{3.17}$$

To facilitate the analysis, we express the missing data mask as:

$$Miss(n) = \sum_{n_i \notin \mathbb{S}} \delta(n - n_i). \tag{3.18}$$

To examine the effect of missing samples, we look at the difference in the IAF between missing data case compared with complete data case,

$$\begin{aligned} &IAF_D(n,b) = IAF_{SS}(n,b) - IAF_{MM}(n,b) \\ &= IAF_{SS}(n,b) - (IAF_{SS}(n,b) + IAF_{MissMiss}(n,b) - IAF_{SMiss}(n,b) - IAF_{MissS}(n,b)) \\ &= IAF_{SMiss}(n,b) + IAF_{MissS}(n,b) - IAF_{MissMiss}(n,b). \end{aligned} \tag{3.19}$$

From the definitions, we obtain:

$$\begin{aligned} IAF_{MissMiss}(n,b) &= \sum_{n_i \notin \mathbb{S}} \delta(n - n_i + b) \sum_{n_j \notin \mathbb{S}} \delta(n - n_j - b) \\ &= \sum_{n_i \notin \mathbb{S}} \delta(n - n_i, b) + \sum_{\substack{n_i, n_j \notin \mathbb{S}, \\ n_i - n_j > 0, \text{even}}} \delta\left(n - \frac{n_i + n_j}{2}, b \pm \frac{n_i - n_j}{2}\right) \\ &= \sum_{n_i \notin \mathbb{S}} \delta(n - n_i, b) + \sum_{n_i \notin \mathbb{S}} \delta\left(b - (n_i - n), b - (n - n_i)\right). \end{aligned} \tag{3.20}$$

The first term in the righ-hand side includes the entries in the time axis ($b = 0$), and the second term represents entries off the time axis due to missing samples. These entries are actually the intersections of lines $b = n_i - n$ and $b = n - n_i$ with $n_i \notin \mathbb{S}$. Similarly, the cross-terms $IAF_{MissS}(n,b)$ and $IAF_{SMiss}(n,b)$ are obtained

as follows:

$$\begin{aligned} IAF_{MissS} &= \sum_{n_i \notin \mathbb{S}} \sum_{j=1}^{N} \delta(n - n_j, b - (n_i - n)) \\ IAF_{SMiss} &= \sum_{n_i \notin \mathbb{S}} \sum_{j=1}^{N} \delta(n - n_j, b - (n - n_i)). \end{aligned} \tag{3.21}$$

We can interpret that the IAF_{MissS} and IAF_{SMiss} are straight lines across all values of $n \in [1, N]$, where b satisfies $b = n_i - n$ and $b = n - n_i$. We can see the non-zero entries of $IAF_{MissMiss}(n, b)$ are located at the intersection of the two IAF cross-terms.

Because off time-axis entries of $IAF_{MissMiss}$ are intersections of two IAF cross-terms, the maximum number of missing entries in a time slice at specific b is expressed as:

$$\tilde{N}_{miss_{IAF}} = 2|b| + 2N_{miss}. \tag{3.22}$$

For illustration, four IAF masks are shown in Fig. 3.4 with $N_{miss} = 5$.

Analysis of missing samples in the Fourier transform domain [29]

As we said above, the ambiguity and the TF domain of incomplete signals are filled with noise-like effects, which can severely influence the true TF signature. The underlying reasons are:

- While the Fourier transform (FT) of the full sinusoidal signal is sparse, the FT of a randomly chosen set of samples is not sparse. In other words, missing samples in the time domain will cause a noisy spectral representation.
- Missing samples in time lend themselves to missing entries in IAF.
- The TFD and the AF are obtained by taking the Fourier transform of IAF over lag and time variables.

In short, we can say that the FT of incomplete data leads to the artifacts witnessed in the TF and the ambiguity domains. By analyzing the statistical properties of the Fourier coefficients, we can characterize the variance of the noise-like effect and also

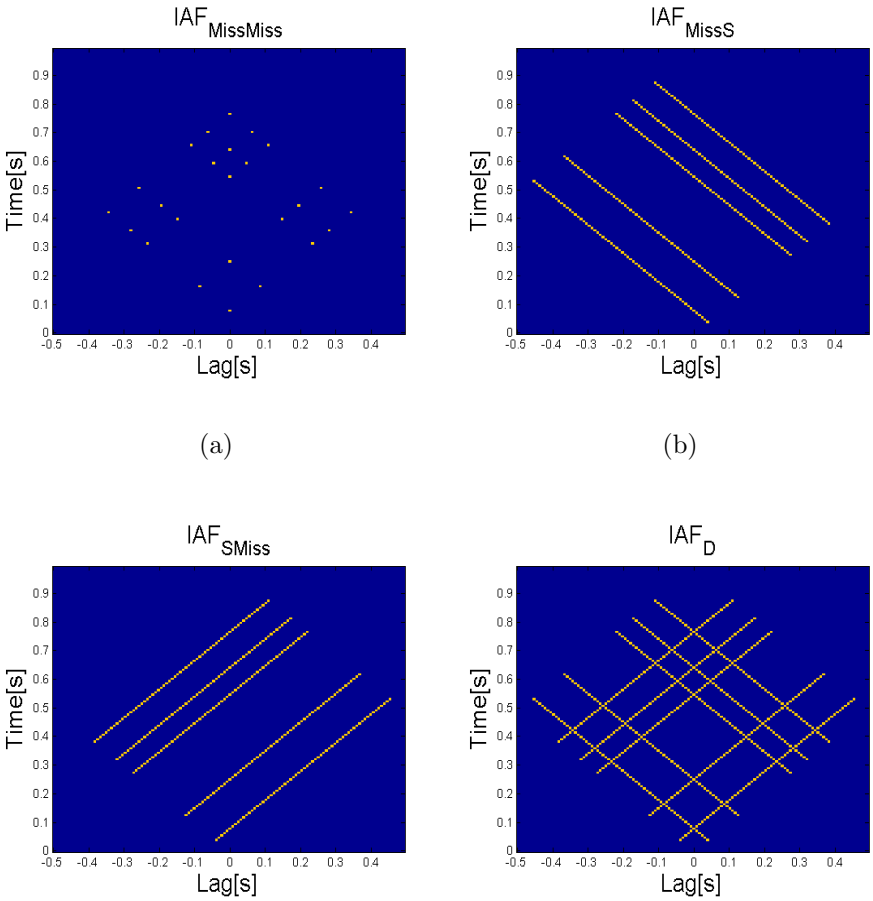


Figure 3.4: Mask IAF with 5 missing samples in time domain (a) $IAF_{MissMiss}$; (b) IAF_{SMiss}; (c) IAF_{MissS}; (d) IAF_D.

its relation to the number of missing data values. Observe a signal consisting of E sinusoid components in the form:

$$s(n) = \sum_{i=1}^{E} a_i e^{j2\pi k_i n/N}. \tag{3.23}$$

where a_i and k_i denote amplitudes and frequencies of the i^{th} signal components, respectively. The DFT of such signal can be written as:

$$S(k) = N \sum_{n=1}^{N} \sum_{i=1}^{E} a_i e^{-j2\pi(k-k_i)n/N}. \tag{3.24}$$

In a compressive sensing scenario we are dealing with just a small subset of samples from $s(n)$ taken at the random positions defined by the following set:

$$\boldsymbol{N}_M = \{n_1, n_2, ..., N_M\} \subset \mathbf{N} = \{1, 2, ..., N\}. \tag{3.25}$$

Therefore, $\boldsymbol{N}_M$ represents the positions of measurements. Following the definition of the DFT in (3.24), let us observe the product of samples and Fourier basis functions of the full and incomplete signal:

$$\mathbf{s}^{\dagger} = \left\{ s^{\dagger}(n) = \sum_{i=1}^{E} a_i e^{-j2\pi(k-k_i)n/N}, \qquad n = 1, 2, ..., N \right\} \tag{3.26}$$

$$\mathbf{y} = \left\{ y(n_m) = \sum_{i=1}^{E} a_i e^{-j2\pi(k-k_i)n_m/N}, \qquad n_m \in \mathbf{N}_M \right\} \tag{3.27}$$

Let start with the simplest case that signal is composed of one component $k_i = k_1$, and assume $a_1 = 1$. The set of samples given by (3.26) and (3.27) become:

$$\begin{aligned} s^{\dagger}(n) &= e^{-j2\pi(k-k_1)n/N},\ n \in \mathbf{N} \\ y(n_m) &= e^{-j2\pi(k-k_1)n_m/N},\ n_m \in \mathbf{N}_M. \end{aligned} \tag{3.28}$$

Considering $s^{\dagger}(n)$, we have:

$$\sum_{n=1}^{N} e^{-j2\pi(k-k_1)n/N} = \frac{1 - (e^{-j2\pi(k-k_1)/N})^N}{1 - e^{-j2\pi(k-k_1)/N}} = \begin{cases} N, & \text{if } k = k_1, \\ 0, & \text{if } k \neq k_1. \end{cases} \tag{3.29}$$

The incomplete signal is different. When $k = k_1$, $y(n_1) + y(n_2) + ... + y(n_M) = M$ and when $k \neq k_1$, $y(n_1) + y(n_2) + ... + y(n_M) \neq 0$. The variance of noise-like effect appearing in the FT of incomplete data can be calculated as follows:

$$\begin{aligned} \sigma^2(\mathbf{Y}_{k \neq k_1}) &= E\left\{ [y(1) + ... + y(M)][y(1) + ... + y(M)]^* \right\} \\ &= ME\left\{ y(n)y^*(n) \right\} + M(M-1)E\left\{ y(n)y^*(m) \right\}_{n \neq m} \\ &= M + M(M-1)\frac{-1}{N-1} \\ &= M\frac{N-M}{N-1}. \end{aligned} \tag{3.30}$$

The detailed derivation of variance can be found in [29]. Thus, in the case of a sparse E-component signal, the variance of noise-like effect that appears in the spectral domain as a consequence of missing data is:

$$\sigma^2(\mathbf{Y}_{k\neq k_i}) = M\frac{N-M}{N-1}\sum_{i=1}^{E} a_i^2. \tag{3.31}$$

The variance of noise-like effect produced in the spectral domain depends on the number of missing samples $(N-M)$. It can be seen that if $M=N$, the variance of the noise-like effect becomes zero. However, for $M \ll N$, we have $\frac{N-M}{N-1} \to 1$, or $\sigma^2(\mathbf{Y}_{k\neq k_i}) \approx M\sum_{i=1}^{E} a_i^2$. We can say for low value of M, the noise-like effect level exceeds the values of some (or all) signal components. Thus, when too many samples are missing, the TFD cannot be reconstructed accurately. For illustration, we use a linear chirp signal, and the effects of missing data us on the TF domain and the AF domain are shown in Fig. 3.5 and Fig. 3.6.

3.5 Literature Review of Sparse-Aware TFDs

In general, we have two approaches in applying CS in the TFDs. They are parametric and non-parametric methods. The former is used when the signal format is known before hand. It gives fast and accurate results. In the case that the signals are unknown, a non-parametric method is deployed. In this part, we introduce both approaches. The first three techniques are non-parametric and the last one is parametric.

3.5.1 Sparse kernel design [2]

The general form of Cohen's class has been expressed in (2.18) and rewritten as follows:

$$\begin{aligned} D(n,k) &= \sum_{p=-N/2}^{N/2-1}\sum_{b=-N/2}^{N/2-1} AF_s(p,b)C(p,b)e^{-j2\pi pn}e^{-j2\pi kb} \\ &= \mathbb{F}_{2D}\left\{AF_s(p,b)C(p,b)\right\}. \end{aligned} \tag{3.32}$$

Or we can say that the filtered AF is the result of the inverse two-dimensional FT of $D(n,k)$:

$$AF_s(p,b)C(p,b) = \mathbb{F}_{2D}^{-1}\{D(n,k)\}. \tag{3.33}$$

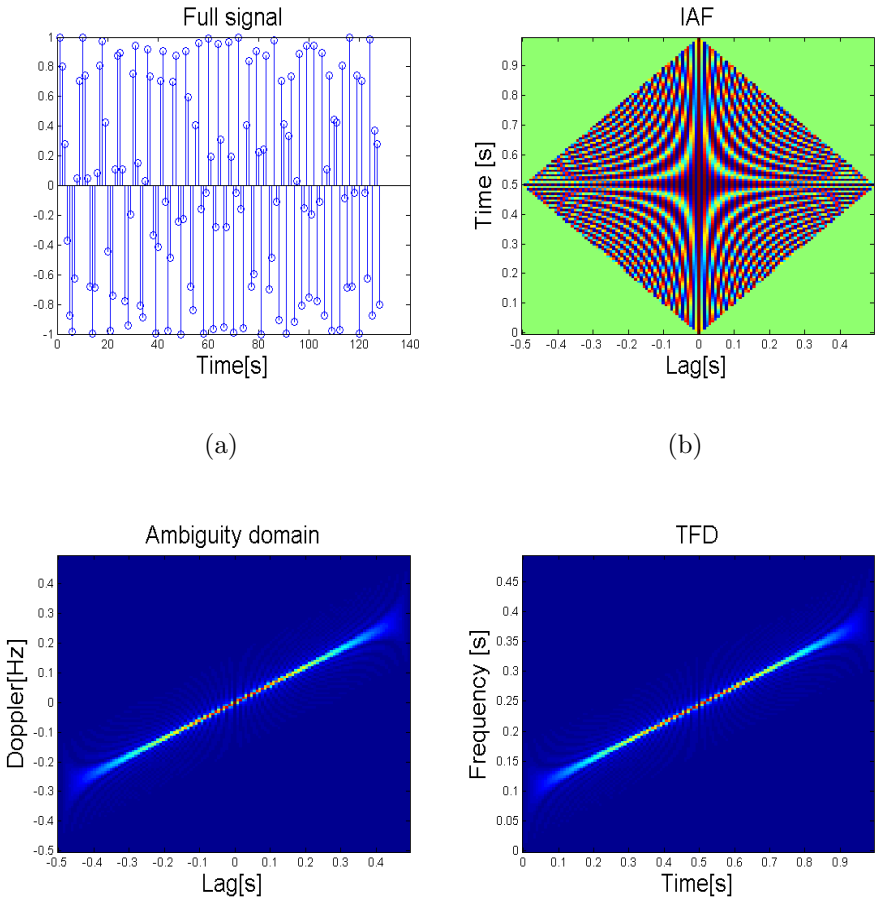
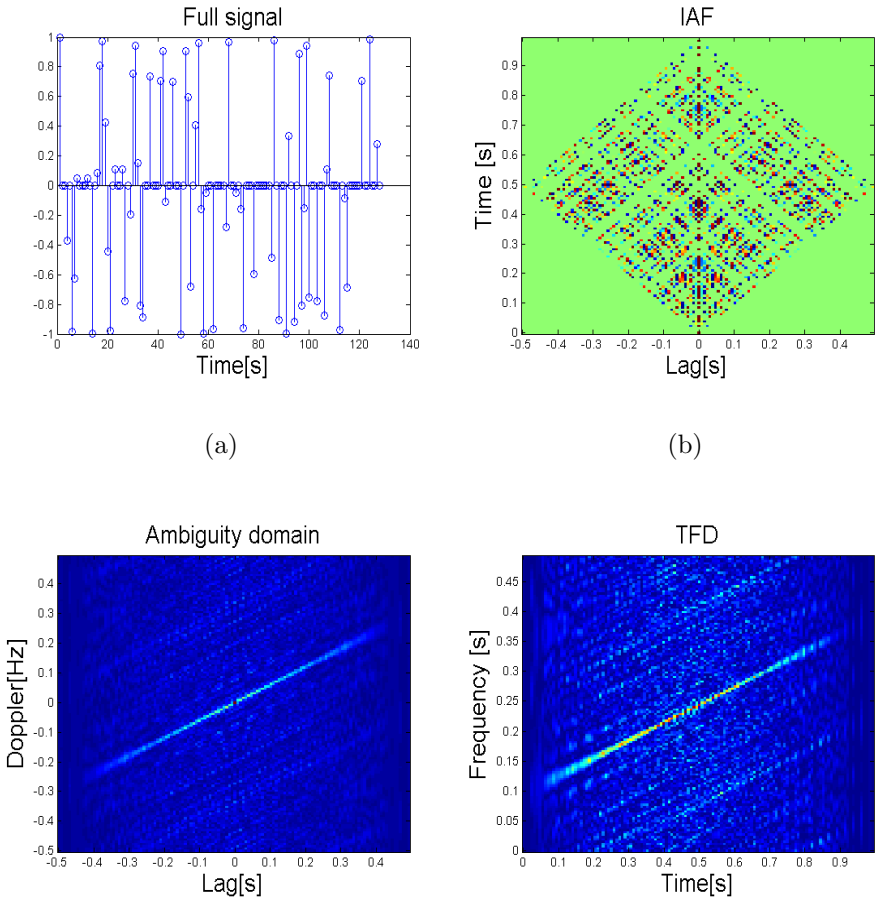


Figure 3.5: A full-data linear chirp in (a) Time domain; (b) IAF; (c) AF; (d) TFD.



(a) (b)

(c) (d)

Figure 3.6: A 50% incomplete linear chirp signal in (a) Time domain; (b) IAF; (c) AF; (d) TFD.

As most signals appearing in a real application, their TF representations contain a small number of non-zero values and thus TFD can be accurately recovered with incomplete measurements by compressive sensing. As we know, in the ambiguity domain, auto-terms mostly locate near the center, cross-terms often stay far away from the origin. By only taking measurements around the origin of the ambiguity plane not the total plane, the cross-terms can be attenuated. And compressive sensing helps obtain an improved TFD with few observations of the ambiguity domain near the origin. The l_1 problem is formed as follows. Observation vector $\mathbf{y}$ is obtained by a small number of measurements taken from specific ambiguity region $\mathbf{\Omega}$ around the origin of the ambiguity domain $AF \in \mathbb{C}(N\text{x}N)$. (3.33) can be written in vector form as:

$$\mathbf{y} = \mathbb{F}_{2D}^{-1}(\Omega)TF, \tag{3.34}$$

where $TF \in \mathbb{C}(N^2\text{x}1)$. The system defined by (3.34) is under-determined, and can have infinite solutions. According to the theory, the TFD with the smallest number of non-zero coefficients can be obtained as a solution of l_0-norm minimization. However, in practice we may use the near optimal solutions based on the l_1 norm minimization:

$$\hat{TF} = \min \|TF\|_1 \qquad \text{subject to} \qquad \mathbf{y} = \mathbb{F}_{2D}^{-1}(\Omega)TF \tag{3.35}$$

OMP or Lasso is used to solve l_1 problem. An example is presented to illustrate the above theory. Two crossing chirps are used for illustration. The result is displayed in Fig. 3.7.

3.5.2 TF estimation using a sinusoidal dictionary[3]

Consider an arbitrary continuous-time non-stationary signal $s_c(t)$, which consists of E components:

$$s_c(t) = \sum_{i=1}^{E} A_i(t) \exp(j\omega_i(t)t + v_c(t)), \quad 0 \leq t \leq T \tag{3.36}$$

where $A_i(t)$ and $\omega_i(t)$ are the time-varying positive amplitude and phase of the i^{th} component, and $v_c(t)$ is an additive white noise, and T is total observation time.

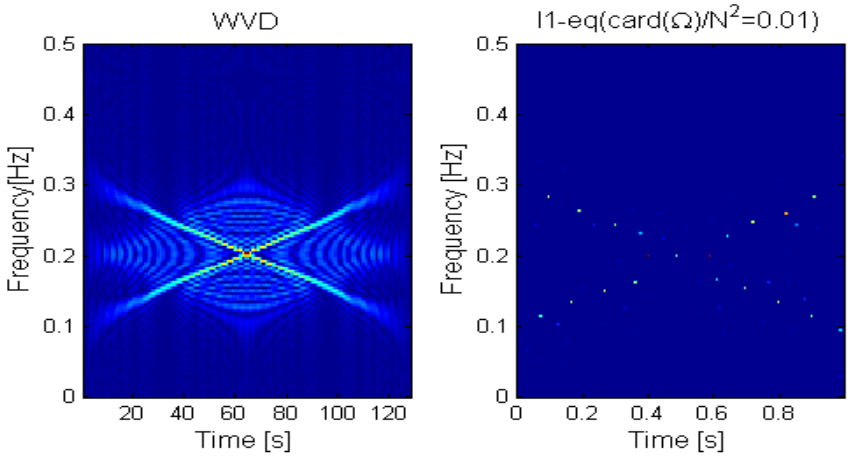


Figure 3.7: TFD obtained by the WVD and the sparse kernel method.

To avoid aliasing, the continuous-time signal is first passed through a low-pass filter to remove out-of-band noise and then sampled with the rate $F_s \geq 2F_{\max}$. The discrete-time signal is:

$$\begin{aligned} s(n) &= \sum_{i=1}^{E} A_i(nT_s)\exp(j\omega nT_s) + v(nT_s) \\ n &= 0, 1, ..., N-1, \end{aligned} \tag{3.37}$$

where $T_s = 1/F_s$ is the sampling period, $s(n)$ and $v(n)$ are the discrete versions of $s_c(t)$ and $v_c(t)$, and $N = \lfloor T/T_s \rfloor$.

The proposed approach builds on estimation of the local frequency contents, which are referred to the middle point of the sliding window, as expressed (3.38) and in Fig. 3.8.

$$s_m(n) \approx \sum_{i=1}^{E} C_{i,m}\exp(j2\pi k_{i,m}nT_s) + v_m(n) \qquad n = 0, 1, ..., N_w - 1, \tag{3.38}$$

where m is window index, $C_{i,m}$ is the complex amplitude of i^{th} component in m^{th} window and N_w is window length.

In vector form, the signal over the m^{th} window can be expressed as:

$$\mathbf{S}_m = \mathbf{\Psi}\mathbf{X}_m + \mathbf{V}_m, \tag{3.39}$$

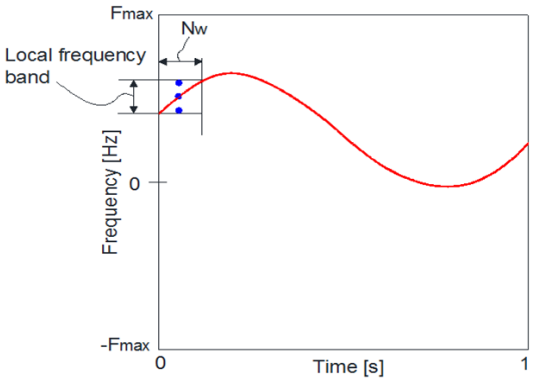


Figure 3.8: TF estimation based on the sinusoid dictionary principle

where $\mathbf{S}_m = [s_m(0), ..., s_m(N_w - 1)]^T$, $\mathbf{V}_m = [v_m(0), ..., v_m(N_w - 1)]^T$, $\mathbf{X}_m$ is a E-sparse amplitude vector of length F_s, and the dictionary matrix $\mathbf{\Psi}$, is defined as:

$$\begin{aligned} \mathbf{\Psi} &= [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, ..., \boldsymbol{\psi}_{F_s}] \\ \boldsymbol{\psi}_i &= \exp(j2\pi k_i n T_s) \\ k_i &= 0, 1, ..., F_s - 1 \\ n &= 0, 1, ..., N_w - 1. \end{aligned} \tag{3.40}$$

Since $E \ll N_w \ll F_s$, $\mathbf{X}_m$ in (3.39) can be solved as a compressive sensing problem, expressed as:

$$\hat{\mathbf{X}}_m = \arg\min \|\mathbf{X_m}\|_1 \; s.t. \; \|\mathbf{S}_m - \mathbf{\Psi}\mathbf{X}_m\|_2^2 \leq \epsilon, \tag{3.41}$$

where ϵ is noise level. The problem in (3.41) is solved by convex optimization or iterative greedy search for example OMP (Orthogonal matching pursuit). For illustration, we use a multi-component signal expressed as follows:

$$\begin{aligned} s(n) = &\exp\left\{ j(0.15F_s)\cos(2\pi\frac{n}{F_s} + \pi) + j2\pi(0.25F_s)\frac{n}{F_s} \right\} \\ &+ \exp\left\{ j2\pi \left[(0.1F_s)\frac{n}{F_s} + (0.3F_s)\frac{n^2}{2F_s^2} \right] \right\}. \end{aligned} \tag{3.42}$$

The signal is randomly shortened by 50%. The incomplete signal, and its TFDs obtained by spectrogram, WVD, and the sinusoidal dictionary method are depicted in Fig. 3.9. The sinusoidal dictionary method provides enhanced TFD compared with the spectrogram and the WVD, in which cross-terms and artifacts due to missing data are removed.

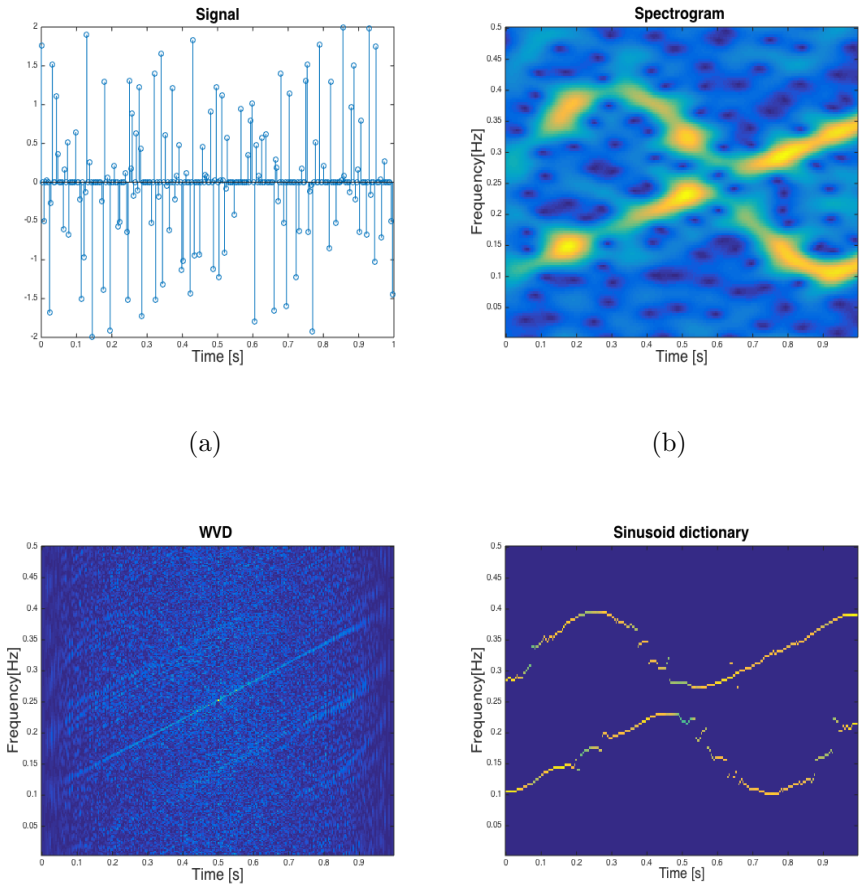


Figure 3.9: A multi-component signal with 50% data missing: (a) Time domain signal; (b) The spectrogram; (c) The WVD; (d) The TFD obtained by CS method with the sinusoidal dictionary.

3.5.3 Sparse reconstruction using multiple measurement vector [4]

This method performs sparse reconstruction of TFD from random time-domain samples using windows, reminiscent of the multiple window spectrogram. The reconstruction corresponding to different windows can be averaged to provide an improved TFD over a single window-based reconstruction. The overlapping nature of the different window reconstructions amounts to a common sparse support property and inspires the use of multiple measurement vector (MMV) techniques within the CS paradigm. This enables achieving enhanced signal localization in the TF domain over a single measurement vector (SMV). The MMV problem is solved using the complex multitask Bayesian compressive sensing method, or Block OMP.

The general form of Cohen's class of a signal $s(n)$ can be expressed as:

$$D(n,k) = \sum_{p}\sum_{b}\sum_{u} C(p,b)s(u+b)s^*(u-b)e^{j2\pi(pu-kb-pn)}, \tag{3.43}$$

where $C(p,b)$ is a kernel in the ambiguity domain, and b is time lag variable. Call $\bar{C}(n,b)$ the Fourier transform of $C(p,b)$ with respect to p, or the kernel in instantaneous autocorrelation domain. It is expressed as:

$$\bar{C}(n,b) = \sum_{p} C(p,b)e^{-j2\pi pn} \tag{3.44}$$

Called $\bar{C}^{45}$ is the 45-degree-rotated kernel from $\bar{C}(n,b)$ for all values n, b. Eigen decomposition $\bar{C}^{45}$, we have:

$$\bar{C}^{45} = \sum_{r=1}^{\text{rank}_{\bar{C}^{45}}} \lambda_r e_r e_r^H. \tag{3.45}$$

The TFD $D(n,k)$ can be re-written as:

$$D(n,k) = \sum_{r=1}^{\text{rank}_{\bar{C}^{45}}} \lambda_r \left| \sum_{i=0}^{N-1} s(n+i)e_r^*(i)e^{-j2\pi ki} \right|^2, \tag{3.46}$$

where $e_r^*(i)$ is the i^{th} element of e_r^*. As such, the TFD is obtained by the weighted sum of the spectrogram. For each signal segment, sparse reconstruction and the

sinusoidal matrix are used to estimate the frequency component at the instant corresponding to the middle point of the segment, which is similar to TF estimation using the sinusoidal dictionary mentioned above. However, here multiple windows are used, and the final estimation is the weighted sum of all resulted calculated TF. The result is proved to be better than using a single window. To illustrate, $C(p, b)$ is chosen as Choi-William kernel,

$$C(p, b) = e^{-p^2 b^2/\sigma}$$

With $\sigma = 100$, the time and frequency smoothing Hanning windows' length are 15 and 123, the eigen-decomposition of the rotated kernel is shown in Fig. 3.10, which reveals the eigenvalues decay very fast and those after the 6^{th} term are negligible. The six eigenvectors corresponded to the six largest eigenvalues are shown in Fig.

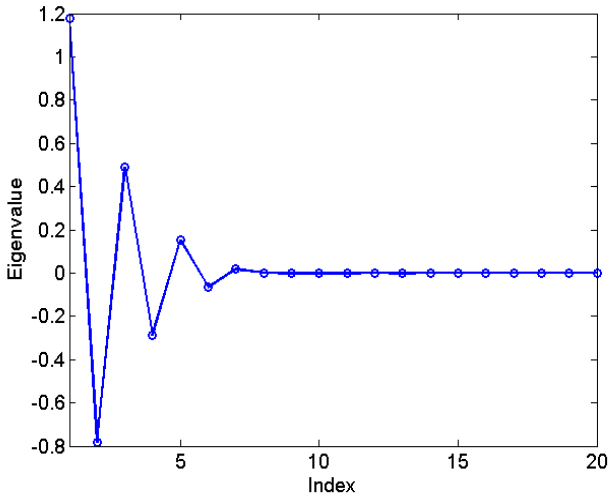


Figure 3.10: The first 6 dominant eigenvalues of the rotated $C(p, b)$.

3.11. The simulated input signal is expressed below:

$$x(n) = \exp\left[j2\pi\left((0.25F_s)\frac{n^2}{2} + (0.2F_s)n\right)\right] + \exp\left[j2\pi(0.1F_s)t\right]$$
$$n = 0, 1, ..., F_s - 1; \quad F_s = 128.$$

TFDs of signal with 75% data missing are shown in Fig. 3.12 and Fig. 3.13. The result obtained by sparse reconstruction with MMV is superior than one from sparse reconstruction with a single window.

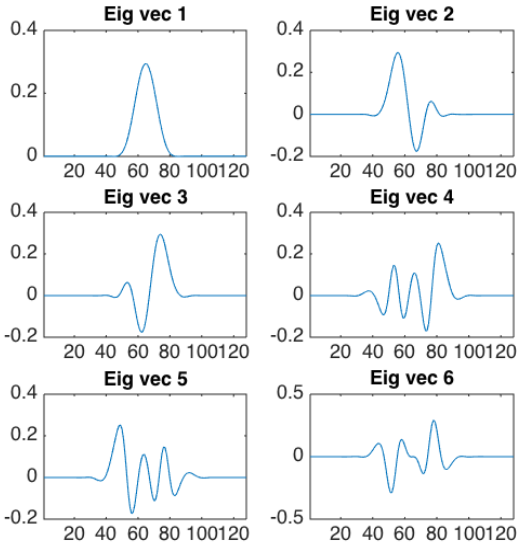


Figure 3.11: The 6 dominant eigenvectors of the rotated $C(p, b)$.

3.5.4 Parametric sparse recovery

The parametric approach for sparse-aware TF analysis was first proposed in [30], and then in [31] with pruned orthogonal matching pursuit (POMP). This method focuses on separation of multiple sinusoid FM micro-Doppler components and the estimation of their parameters, such as the Doppler repetition period, the Doppler amplitude and the initial phase. This method is used in the case that all radar reflections are from the same rotation, vibrating, or oscillating targets. The radar echo from a coning target is given by,

$$y(n) = \sum_{i=1}^{E} a_i \exp\{j\frac{4\pi}{\lambda} d_i \sin(\omega n + \theta_i)\} \quad n = n_1, n_2, ..., n_M \tag{3.47}$$

where ω refers to the rotation angular velocity of the target, a_i is the complex reflectivity of the i-th scatter, λ is the radar wavelength, d_i is dependent on the spatial position of the i-th scatter, θ_i is the initial phase which also depends on the spatial position.

To discretise the MD signal, d_i and θ_i are uniformly divided into I, and thus J discrete values, i.e., $d_i \in \{d_1, d_2, ..., d_I\}$and $\theta_i \in \{\theta_1, \theta_2, ..., \theta_j, ..., \theta_J\}$, and thus we

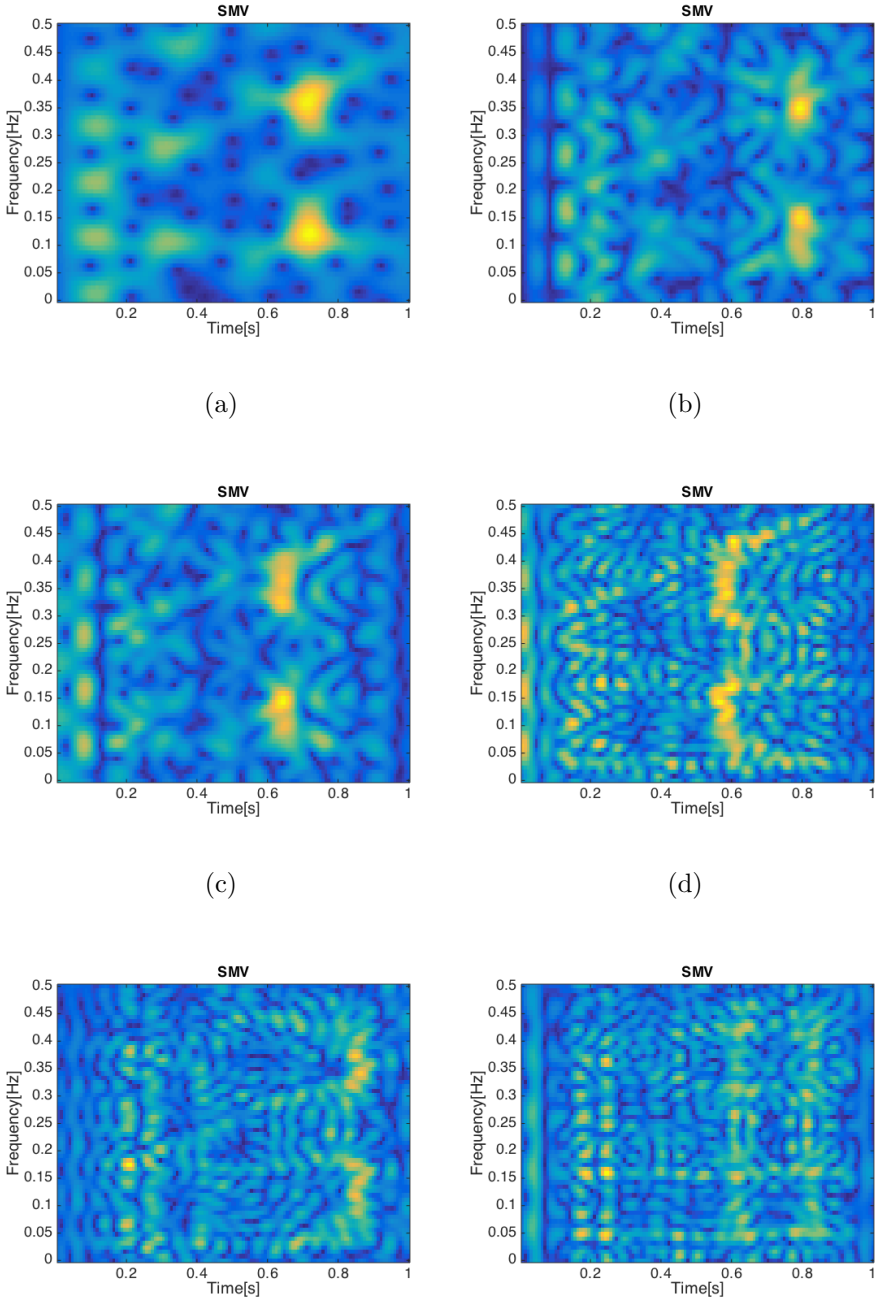


Figure 3.12: TFDs with a single window being: (a) Eigenvector 1; (b) Eigenvector 2; (c) Eigenvector 3; (d) Eigenvector 4; (e) Eigenvector 5; (f) Eigenvector 6.

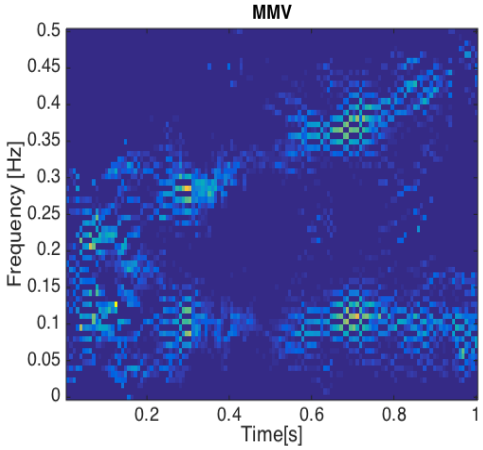


Figure 3.13: TFD obtained by MMV.

obtain matrix $[I \,\mathrm{x}\, J]$. The received signal in (3.47) is rewritten as:

$$\mathbf{Y} = \mathbf{A}(\omega)\mathbf{X}. \tag{3.48}$$

where $\mathbf{Y} = [y(n_1), y(n_2), ..., y(n_M)]^T$, $\mathbf{A}(\omega) \in C^{M\,\mathrm{x}(IJ)}$, and its element is:

$$\mathbf{A}(\omega)_{m,i+(j-1)I} = \exp\{j\frac{4\pi}{\lambda}d_i \sin(\omega n_m + \theta_j)\}. \tag{3.49}$$

$\mathbf{X} \in \mathbb{C}^{(IJ)\,\mathrm{x}\,1}$ is a E-sparse vector, and its non-zero element $\mathbf{X}_{i+(j-1)I} = a_i$ if and only if $d_i = d_i, \theta_j = \theta_i$ $(E < M < IJ)$. Thus, the MD separation becomes sparse recovery problem. For each value of ω, (3.48) is solved, using OMP:

$$\{\omega, \mathbf{X}\} = \arg\min \|\mathbf{X}\|_0 \quad \text{subject to} \quad \mathbf{Y} = \mathbf{A}(\omega)\mathbf{X}. \tag{3.50}$$

The correct value of ω is estimated by choosing most sparse $\mathbf{X}$. The estimation is based on (3.51). The value of ω with smallest $S_{\mathbf{X}}(\omega)$ is chosen, where:

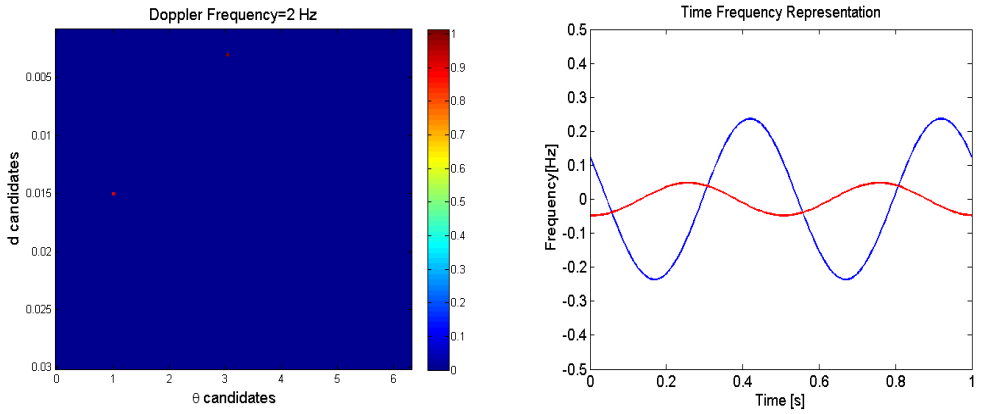
$$S_{\mathbf{X}}(\omega) = -\sum_{n=1}^{IJ} \frac{|\mathbf{X}_{\omega(n)}|^2}{\sum_{n=1}^{IJ}|\mathbf{X}_{\omega(n)}|^2} \log \frac{|\mathbf{X}_{\omega(n)}|^2}{\sum_{n=1}^{IJ}|\mathbf{X}_{\omega(n)}|^2}. \tag{3.51}$$

The simulation result with parameters mentioned in [30] is shown in Figure 3.14.

In [31], prune orthogonal matching pursuit (POMP) is proposed to replace OMP in [30] in order to avoid unnecessary computation with wrong candidate values of ω, thus reduce the computation load of the approach proposed in [30]. The result obtained is similar (Fig. 3.14) with less time for computation.

Table 3.1: The signal's parameters for the simulation of the parametric sparse recovery.

Received signal	$a_i \exp j\frac{4\pi}{\lambda} d_i \sin(2\pi f_{mD} n + \theta_i) \quad n = n_1, n_2, ..., n_M$
λ	0.008
f_{mD}	2
Number of samples per second F_s	200
Number of samples M	200
d_i	0.015 and 0.003
ϕ_k	1 [rad] and 3 [rad]



(a) (b)

Figure 3.14: (a) Estimated parameters; (b) Sinusoid TF signature of the two signals using the estimated parameters.

3.6 Conclusion

The fact that most of non-stationary signals are sparse in the TF domain has invited CS to play an important part in TFDs. CS techniques, or sparse TFDs, enable reconstruction of the entire signal from a small randomly chosen set of measurements. Thus, they can provide reliable TF estimation even in the case of missing and ran-

domly sampled data. This chapter has presented the key concepts of CS as well as the recovery condition and a popular approach to solve the CS problem, the OMP. Missing data effects are also included that underline the significance of CS in recovering the signal TF signature. Some recent sparse approaches are represented and are illustrated by simulations with synthetic signals.

Chapter 4

Sparse Reconstruction of Time-Frequency Signature using The Chirp Dictionary

4.1 Introduction

4.1.1 Motivation

Non-stationary signals arise in a broad class of active sensing modalities, including sonar, radar, and ultrasound. They are the preferred type of smart jamming and also characterize many passive sensing problems, such as speech and electromyographic recordings [32, 33, 34, 35]. In particular, non-stationarity underlines Doppler and micro Doppler signals which represent radar returns from moving targets [36, 37, 38, 39]. Time-frequency signal representations (TFSRs) reveal the signal local structure which changes with time. Therefore, they enable separations of non-stationary signals that are mixed in both time and frequency domains, where traditional approaches fail to capture or distinguish between individual signal components. TFSRs are commonly obtained using linear basis signal decomposition [40], [41], and quadratic TF distributions (QTFDs), generally referred to as Cohen's class [5], [42]. The latter have their roots in the nonparametric WVD.

QTFDs are defined by two-dimensional (2D) kernels which convolve the WVD for interference reduction. The reduced interference distribution (RID) kernels act on

preserving the true signal power terms, referred to as auto-terms, and eliminating, or at least considerably attenuating, the undesired cross-terms. Cross-terms represent false power concentrations and are generated from the data bilinear lag products underlying QTFDs. It has been analytically shown that missing and randomly sampled non-stationary signals give rise to artifacts in both the TF domain and the ambiguity domain [1, 28, 43]. These artifacts clutter the signal components and hide pertinent signal structure, including the instantaneous frequencies. Efforts and attempts to use traditional RID kernels to reduce the type of clutter induced by missing samples along with mitigations of signal cross-terms have proved both unsuccessful and ineffective.

In compressive sensing (CS), a sparse representation of a signal is projected onto a much lower dimensional measurement space. This leads, in general, to decreasing the data acquisition requirements from a time, logistic and hardware complexity perspectives. It is then possible to record a small number of linear measurements of a signal and then reconstruct the complete set of all samples. The required number of observations is slightly more than the signal sparsity level but much less than the signal dimension. Although applied in many applications, little consideration has been given to CS and sparse reconstructions of non-stationary signals. Owing to their instantaneous narrow band characteristics, the signatures of a large class of non-stationary signals occupy small regions in the TF domain. This property casts these signals as sparse in the joint-variable representations and has recently invited sparse signal reconstruction and compressive sensing techniques to play an important and fundamental role in TF signal analysis and processing, especially when we have incomplete data [2, 3, 44].

4.1.2 Related work

For most single and multicomponent FM signals, local reconstruction of TF signatures from few random observations is deemed to outperform global signal reconstructions, which deals with a much broader signal bandwidth, i.e., lower sparsity. One of the most straightforward sparse reconstructions of local signal frequency characteristics is achieved by applying a sliding window, reminiscent of the STFT [3], [27]. Using a partial Fourier basis, one can proceed to apply greedy algorithms

or convex optimization techniques to find the sparsest frequency content that describe the observations within the time window. This approach involves a sinusoidal dictionary that relates the windowed compressed observations to their local sparse frequencies. The method, however, suffers from the trade-off between necessary measurements for accurate recovery and sparsity when considering the window size, and also the picket-fence effect when there is a non-integer period in the analyzed data segments. Thus, another measurement basis is required to obtain more stable and reliable results. In this chapter, we introduce a sparse TF estimation based on a novel chirp dictionary, which provides better performance when compared with the Fourier basis.

4.1.3 Contribution

In many situations, the non-stationary signal frequency law is more properly approximated by piece-wise second-order polynomials rather than fixed frequency sinusoids. In this case, a chirp dictionary, in lieu of a sinusoidal dictionary, is better suited for sparse reconstruction problems when dealing with FM signals. Further, compared to reconstruction techniques using parameterized atoms [31], which also directly operate on the data, the proposed chirp dictionary does not assume any specific signal structure and, as such, is able to maintain its desirable performance for a wide class of non-stationary signals. In this chapter, we introduce a sparse TFD method using the chirp dictionary. The chirp dictionary is built in two ways. The first includes all possible chirps which can appear in any signal segment. The second is also composed of all chirps, but they are formed from sinusoids which are rotated in all eligible angles by the fractional Fourier transform. Although the dictionary construction procedure is different, the two ways actually lead to the same results. The purpose of presenting the second method is to introduce an alternative way to build the chirp dictionary. Also, its theory is also the foundation for other applications in the following chapters. Several simulations with synthetic and real signal are presented to prove the method's efficiency over the sinusoidal dictionary approach and the discrete chirp Fourier transform (DCFT).

4.1.4 Chapter outline

Section 4.2 discusses the first chirp dictionary approach. The range of chirp rate α and initial frequency β, the chirp dictionary as well as the instantaneous frequency estimation by solving the sparse problem are presented in this section. The second way to form a chirp dictionary is introduced in section 4.3. The fractional Fourier transform (FRFT), the key technique, is represented in detail. The value range of the FRFT angle and the frequency of the harmonic signal in the fractional domain will be discussed here. Section 4.4 compares the chirp dictionary and the sinusoidal dictionary approaches to underline the advantages obtained by the chirp atom approach. Section 4.5 focuses on the RIP associated with the chirp dictionary and provides the lower bound on the number of observations for exact recovery. Section 4.6 includes simulation results. The conclusions are given in section 4.7.

4.2 Chirp Dictionary

4.2.1 Signal modelling

Consider an arbitrary continuous-time non-stationary signal $s_c(t)$, which consists of E components:

$$s_c(t) = \sum_{e=1}^{E} A_e(t) \exp\left(j\phi_e(t)\right) + v_c(t), \;\; 0 \leq t < T, \tag{4.1}$$

where $A_e(t)$ and $\phi_e(t)$ are the time-varying positive amplitude and phase of the e^{th} component, $v_c(t)$ is an additive white noise and T is the total observation interval. It is assumed that the phase time-variations are much faster than those of the amplitudes. The continuous-time instantaneous frequency (IF) of the e^{th} component is defined as:

$$F_e(t) = \frac{1}{2\pi} \frac{d\phi_e(t)}{dt}. \tag{4.2}$$

We assume that it is known a priori that the absolute IFs do not exceed $F_{\max}$ i.e. $|F_e(t)| \leq F_{\max}$. We also assume that the IFs do not vary abruptly but rather vary smoothly over time, which is a reasonable assumption in many applications including radar.

To avoid aliasing, the continuous-time signal is first passed through a low-pass filter to remove out-of-band energy, and then sampled with a rate $F_s \geq 2F_{\max}$. The discrete-time signal is:

$$\begin{aligned} s(n) &= \sum_{e=1}^{E} A_e(nT_s)\exp(j\phi_e nT_s) + v(n), \\ n &= 0, 2, ..., N-1, \end{aligned} \tag{4.3}$$

where $T_s = 1/F_s$ is the sampling period, and $s(n)$ and $v(n)$ are the discrete-time versions of $s_c(t)$ and $v_c(t)$, and $N = \lfloor T/T_s \rfloor$.

4.2.2 Non-stationary signal approximation with chirps

The proposed approach builds on the local approximation of each signal component as a chirp. That is, by dividing the observation time interval into (possibly overlapping) time windows of a chosen duration, T_w, the discrete-time signal over each window is approximated by:

$$\begin{aligned} s_m(n) \approx \sum_{e=1}^{E} C_{e,m} \exp\left\{ j2\pi \left[\alpha_{e,m}\frac{n^2}{2F_s^2} + \beta_{e,m}\frac{n}{F_s} \right] \right\} \\ + v_m(n), \qquad 0 \leq n < N_w - 1. \end{aligned} \tag{4.4}$$

where m is the window index, $C_{e,m}$, $\alpha_{e,m}$ and $\beta_{e,m}$ are the complex amplitudes, the chirp rate, and the initial frequency of the e^{th} component/chirp over the m^{th} window, $s_m(n) = s(m\varsigma + n)$ and $v_m(n) = v(m\varsigma + n)$, with ς being the shift between two consecutive windows in terms of number of samples, and $N_w = \lfloor T_w/T_s \rfloor$.

Since $|F_e(n)| \leq F_{\max}$, the initial frequency $|\beta| \leq F_{\max}$, and frequency change in a period of T_w cannot exceed $F_{\max}$, thus the chirp rate α has a range value:

$$\alpha \in \left[-F_{\max}F_s/N_w,\ F_{\max}F_s/N_w \right]. \tag{4.5}$$

The parameter space of interest is (see Fig. 4.1):

$$\begin{aligned} \mathbf{\Omega} = \{ &(\alpha, \beta) \text{ such that} \\ &|\alpha| \leq F_{\max}F_s/N_w,\ |\beta| \leq F_{\max} \text{ and } |\alpha T_w + \beta| \leq F_{\max} \}. \end{aligned} \tag{4.6}$$

The discrete dictionary, to be used in CS, is designed by uniformly sampling the

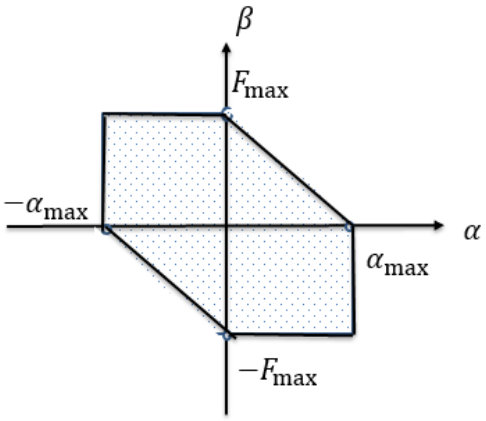


Figure 4.1: 2D space $\mathbf{\Omega}$ for α, β.

2D parameter space $\mathbf{\Omega}$. Let I denote the total number of chirp rate values in the discrete dictionary. For the i^{th} chirp rate value in the dictionary, which we denote as $\tilde{\alpha}_i$, let $\tilde{\beta}_{i,j}$ denote the corresponding possible values for the initial frequency, where $j = 1, ..., J_i$. Since the shape of the parameter space $\mathbf{\Omega}$ is not rectangular, the J_i's values are not all equal.

By performing a sparse component analysis within each window, we can track the time-variations of the chirp parameters of each component (i.e. chirp rate, initial frequency and complex amplitude), thus estimating arbitrary IFs.

In vector form, the signal over the m^{th} window in (4.4) can be expressed as:

$$\mathbf{S}_m = \mathbf{\Psi}\mathbf{X}_m + \mathbf{V}_m \tag{4.7}$$

where $\mathbf{S}_m = [s_m(0), ..., s_m(N_w - 1)]^T$, $\mathbf{V}_m = [v_m(0), ...v_m(N_w - 1)]^T$, $\mathbf{X}_m$ is a E-sparse amplitude vector of length $\sum_{i=1}^{I} J_i$, and the dictionary matrix, $\mathbf{\Psi}$, is defined as:

$$\begin{aligned}
&\mathbf{\Psi} = [\mathbf{\Psi}_1, \mathbf{\Psi}_2, ..., \mathbf{\Psi}_I] \\
&\mathbf{\Psi}_i = [\boldsymbol{\psi}_{i,1}, \boldsymbol{\psi}_{i,2}, ..., \boldsymbol{\psi}_{i,J_i}] \\
&\boldsymbol{\psi}_{i,j}|_n = \exp\left(j2\pi(\tilde{\alpha}_i \frac{n^2}{2F_s^2} + \tilde{\beta}_{i,j}\frac{n}{F_s})\right) \\
&i = 1, 2, ..., I;\ j = 1, 2, ..., J_i;\ n = 0, 1, ..., N_w - 1.
\end{aligned} \tag{4.8}$$

Since $E < N_w \ll \sum_{i=1}^{I} J_i$, solving for $\mathbf{X}_m$ in equation (4.7) becomes a sparse recovery

(or CS) problem, which can be solved by:

$$\hat{\mathbf{X}}_m = \arg\min \|\mathbf{X}_m\|_1 \quad s.t. \quad \|\mathbf{S}_m - \mathbf{\Psi}\mathbf{X}_m\|_2^2 \leq \epsilon. \tag{4.9}$$

where ϵ is the noise level. The solution for (4.9) can be obtained by a greedy algorithm such as the Orthogonal Matching Pursuit (OMP) or the linear programming [24, 45]. The proposed method is basically using the chirp dictionary, and selecting the atoms which best match the local structure of the signal, which is similar to the matching pursuit algorithm. However, sparse reconstruction considers the sparsity level of the signal, as well as minimum observations required for exact recovery.

In addition to employing different dictionaries, the process of obtaining the final signal TF signatures is also different for sinusoidal and chirp atoms. In the case of sinusoidal atoms, or dictionary, the sparse reconstruction algorithm, whether it is OMP or convex optimization, returns the local frequency content, which is referred to the center point of the sliding window, similar to the generation of the spectrograms. On the other hand, for the case of chirp dictionary, the chirp parameters returned by the sparse reconstructions describe the segment of the data captured by the window and, as such, represent the local signal behaviour over the entire window extent, and not only the center point. Since overlapping windows generate overlapping chirps, some averaging process is in order and must be performed to render unique answers at each time sample. In essence, for every TF point (t, f), we sum all the magnitudes of the reconstructed chirps provided by all corresponding sliding windows which include the time sample, t. In so doing, any chirp anomaly will be de-emphasized, whereas accurate frequency representations of the underlying signal will be strengthened. All TF points having a summed magnitude smaller than a certain threshold are ignored and so will not be considered further.

4.3 FRFT Based Chirp Dictionary Approach

4.3.1 Background

FRFT

Time and frequency represent two fundamental physical variables of signal analysis and processing. The Fourier transform (FT), which provides a mapping between

time and frequency domains of a signal, has been used extensively in signal processing applications. The FT of a continuous time domain signal $s(t)$ is defined as:

$$(\mathbb{F}s)(f) = S(f) = \int s(t)e^{-j2\pi ft}dt, \tag{4.10}$$

where $\mathbb{F}$ denotes Fourier transform, t and f are continuous time and frequency variables. The FT can be considered as a reformation of the time domain signal $s(t)$ with respect to the frequency variable f. Thus FT helps to reveal the frequency content of the signal $s(t)$. As time and frequency form the orthogonal coordinates of the TF plane, the FT of a time-domain signal can be viewed as a $\pi/2$-radian counter-clock rotation of the signal plane.

Following this interpretation, the fractional Fourier transform (FRFT) was developed as a generalization of the FT through an angle parameter ϕ [46, 47, 48]. For each value of angle ϕ, the corresponding FRFT rotates the time domain counter-clockwise by an angle of ϕ. Thus, for $\phi = 0$, the FRFT is the identity transform, which is the time representation of the signal. For $\phi = \pi/2$, it becomes FT. For other value of ϕ, the FRFT provides a representation of the signal with respect to a fractional variable, say x, of a fractional domain, between the time and frequency domains. Denote (x, y) the axes of the new reference plane, then the FRFT is illustrated in Fig. 4.2, from which we can see the (x, y) axes are the (t, f) axes rotated counter-clockwise by an angle ϕ. The FRFT of a time domain signal $s(t)$ is defined as:

$$(\mathbb{F}^{\phi}s)(x) = S^{\phi}(x) = \begin{cases} \sqrt{1 - j\cot(\phi)}e^{j\pi x^2\cot(\phi)} \int s(t)e^{j\pi t^2\cot(\phi)}e^{-j2\pi tx\csc(\phi)}dt, & \phi \neq l\pi \\ s(x), & \phi = 2l\pi \\ s(-x), & \phi = (2l+1)\pi, \end{cases} \tag{4.11}$$

where $\mathbb{F}^{\phi}$ is the FRFT operator associated with angle ϕ, $S^{\phi}(x)$ denotes the fractional Fourier transformed signal, l is an integer, t is time and x is the fractional variable. Based on (4.11), the FRFT can also be interpreted as a signal expansion onto a linear FM (chirp) function having a chirp rate of $\cot(\phi)$.

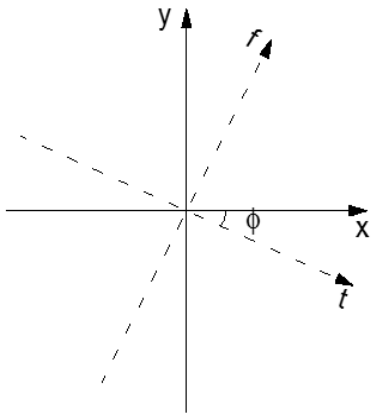


Figure 4.2: Counter-clockwise rotation of the TF plane (t, f), forming a new reference plane (x, y).

FRFT of a windowed signal

According to the discrete FRFT calculation method in Appendix A, all the coordinates appearing in the definition of the FRFT, Wigner distribution, etc., are all dimensionless quantities. Assume that the signal is approximately confined to the time interval $[0, \Delta t]$ and its frequency representation is confined to the interval $[-\Delta f/2, \Delta f/2]$. To obtain dimensionless quantities, we introduce a scale parameter $sl = \sqrt{\Delta t/\Delta f}$. Denote a the order of the FRFT, $\phi = a\frac{\pi}{2}$. So, if $a = 0$, the FRFT is equivalent to the identity transform. If $a = 1$, the FRFT is the Fourier transform. The scaled coordinates for time and frequency are x_0 and x_1, respectively, which are expressed below:

$$\begin{aligned} x_0 &= t/sl \\ x_1 &= fsl. \end{aligned} \tag{4.12}$$

Thus, the length of both intervals for x_0 and x_1 are confined to Δx which is equal to $\sqrt{\Delta t \Delta f}$. In the newly defined coordinates, the signal can be represented in both domains with number of samples $\mu = \Delta x^2$ and samples spaced $1/\Delta x$. Denote μ as the time-bandwidth product, $\mu = \Delta f \Delta t$. We have two examples to clarify the scaled coordinates. For a signal of length 1 second or $\Delta f = F_s$ Hz and $\Delta t = 1$ second, the time-bandwidth product is $\mu = F_s$ and the interval length for x_0 ad x_1

is $\Delta x = \sqrt{F_s}$. The signal in the time and frequency domains will have F_s samples with sampling space of $(1/\sqrt{F_s})$. For a windowed signal of length N_w/F_s seconds $(N_w \leq F_s)$, $\Delta f = F_s$ but $\Delta t = N_w/F_s \leq 1$. The time-bandwidth product is $\mu = N_w$, and the interval length for x_0 and x_1 is $\Delta x = \sqrt{N_w}$. The signal will have N_w samples in both time and frequency domain with sampling space of $(1/\sqrt{N_w})$. Let (x_a, y_{a+1}), or (x, y) in general, be the new coordinate of (x_0, x_1) after performing the FRFT with $\phi = a\frac{\pi}{2}$. The new dimensionless coordinates for the fractional domain as well as the Wigner distribution are plotted in Fig. 4.3.

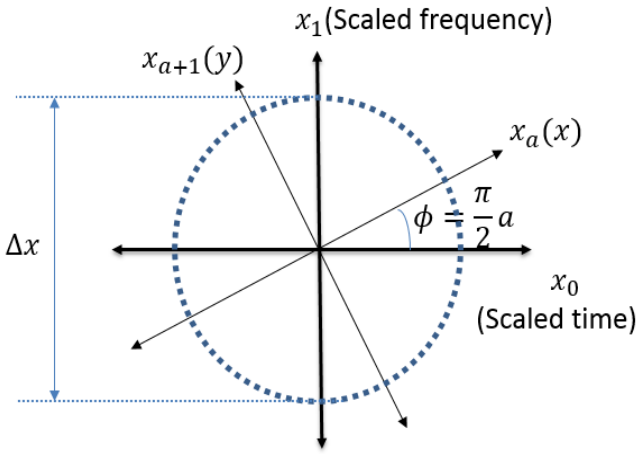
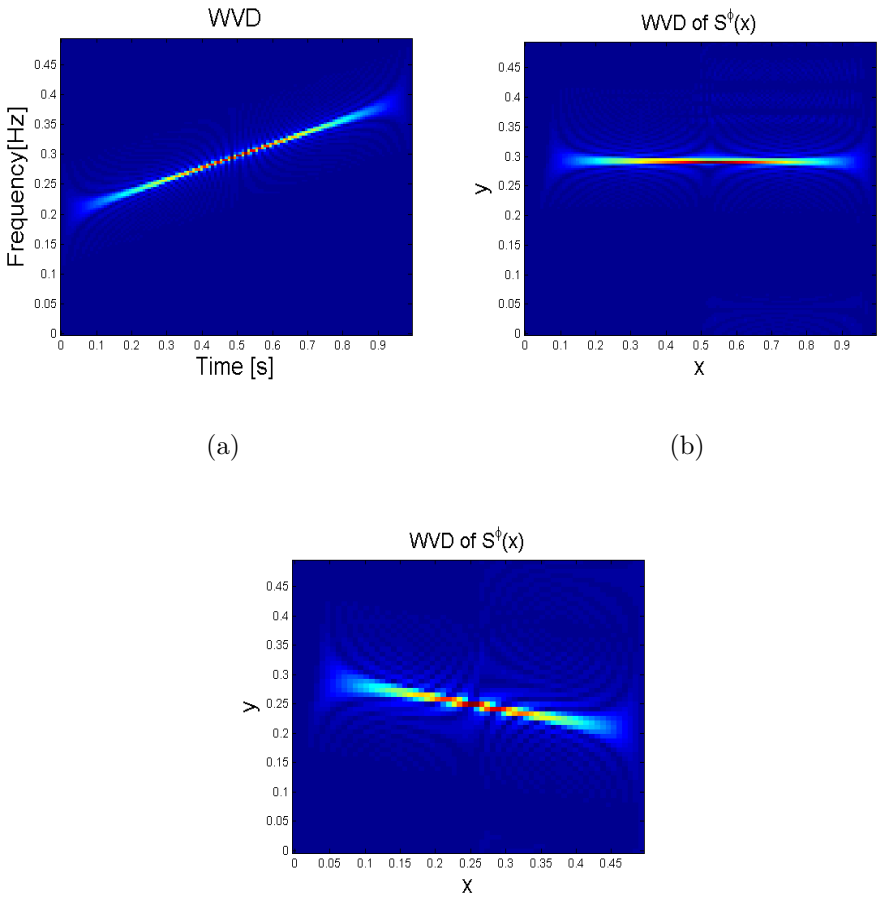


Figure 4.3: Dimensionless coordinates for the fractional domain.

To calculate the FRFT for a windowed signal, a similar method (described in Appendix A) is used. However, we have to notice that the FRFT angle ϕ because this angle changes with window length, N_w. A chirp with a chirp rate of $\alpha = 0.2F_s$ and an initial frequency of $\beta = 0.2F_s$ is used for illustration. We apply the FRFT with an angle $\phi = \tan^{-1}(\alpha/F_s)$ onto the signal in two cases: $N_w = F_s$ and $N_w = F_s/2$. Our purpose is to rotate the chirp signal into a sinusoid. The result is shown in Fig. 4.4. Fig. 4.4 shows that with the same FRFT angle, the result is not the same for different window lengths. Fig. 4.5 gives an explanation for this phenomenon. In Fig. 4.5, it is assumed that $\xi = F_s/N_w$, thus the scale becomes $sl = \sqrt{\frac{\Delta t}{\Delta f}} = \sqrt{\frac{1}{\xi F_s}}$. A windowed signal of any length is confined to $\Delta x = \sqrt{N_w}$. According to Fig. 4.5,



(a) (b)

(c)

Figure 4.4: (a) WVD of the signal ($\alpha = 0.2F_s, \beta = 0.2F_s$, $N = F_s$); (b) WVD of the FRFT of the signal in (a) with $\phi = \tan^{-1}(\alpha/F_s)$; (c) WVD of the FRFT of the windowed signal of (a), the window length is of $N_w = F_s/2$, and $\phi = \tan^{-1}(\alpha/F_s)$.

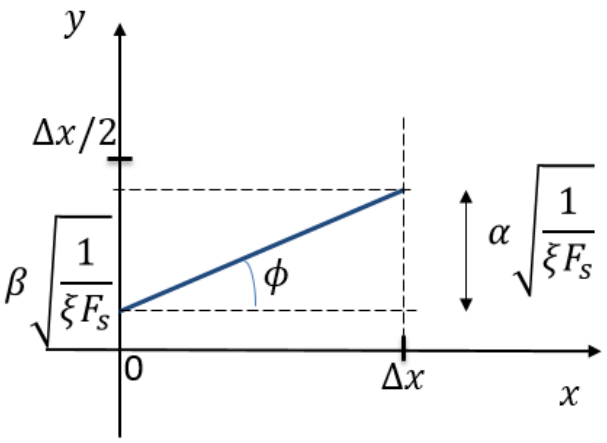


Figure 4.5: Geometric explanation for changing the FRFT angle with respect to window length.

the FRFT angle is obtained by:

$$\tan\phi = \frac{\alpha\sqrt{\frac{1}{\xi F_s}}}{\Delta x} = \frac{\alpha\sqrt{\frac{1}{\xi F_s}}}{\sqrt{N_w}} = \frac{\alpha}{F_s}. \tag{4.13}$$

Now, (4.13) shows that the FRFT calculation method considers all windowed signals with different window lengths $N_w (N_w \leq F_s)$ similar to the signal of length one second. However, the amount of frequency change in this interval is α/ξ, not α. The accurate the FRFT angle thus will be:

$$\tan\phi = \frac{\frac{\alpha}{\xi}\sqrt{\frac{1}{\xi F_s}}}{\sqrt{N_w}} = \frac{\alpha}{\xi F_s} = \frac{N_w\alpha}{F_s^2}. \tag{4.14}$$

The relation between the FRFT angle and the window length can also be explained in time-frequency coordinates as in Fig. 4.6. According to Fig. 4.6, the FRFT angle ϕ is obtained by:

$$\begin{aligned}\tan\phi &= \frac{\alpha/\delta f}{1/\delta t} \\ &= \frac{\alpha N_w}{F_s^2},\end{aligned} \tag{4.15}$$

where $\delta f = F_s/N_w$ is the frequency resolution and $\delta t = 1/F_s$ is the time resolution.

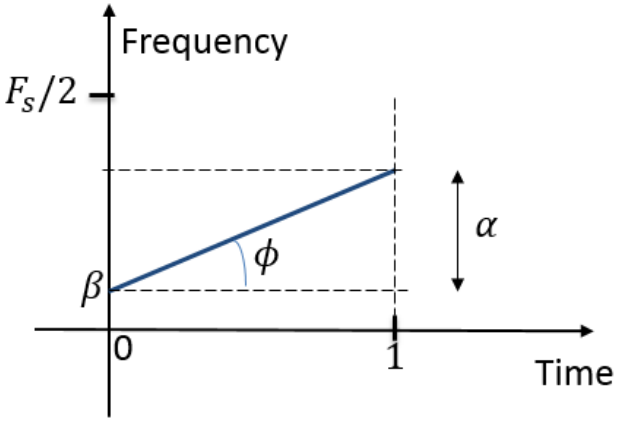


Figure 4.6: Geometric explanation for changing the FRFT angle with respect to window length.

Relation between a chirp and a sinusoid

From Fig. 4.7, any chirp in time domain is a sinusoid in the fractional domain. In particular, a chirp with a chirp rate of α and a sinusoid in the fractional domain are related by the FRFT with an angle $\phi = -\tan^{-1}\frac{N_w\alpha}{F_s^2}$. In this part, we will find the sinusoidal frequency in the fractional domain in relation to the chirp rate α and initial frequency β of the chirp in the time domain.

According to [49], the discrete FRFT rotates the TF plane (x_0, x_1) around a point C, defined by the intersection of the zero-frequency axis with half of the total duration of the time domain signal. With the new dimensionless coordinates, the relation is plotted in Fig. 4.7. Call d the sinusoidal frequency in the (x, y) domain, and then d is expressed as follows:

$$\begin{aligned} d &= \cos(\phi)\left(\beta\sqrt{\frac{1}{\xi F_s}} + \frac{\Delta x}{2}\tan\phi\right) \\ &= \cos(\phi)\left(\beta\frac{\sqrt{N_w}}{F_s} + \frac{\sqrt{N_w}}{2}\tan\phi\right) \\ &= \sqrt{N_w}\cos(\phi)\left(\frac{\beta}{F_s} + \frac{\tan\phi}{2}\right). \end{aligned} \tag{4.16}$$

So, in short, any sinusoid with frequency of $d = \sqrt{N_w}\cos(\phi)\left(\frac{\beta}{F_s} + \frac{\tan\phi}{2}\right)$ in any

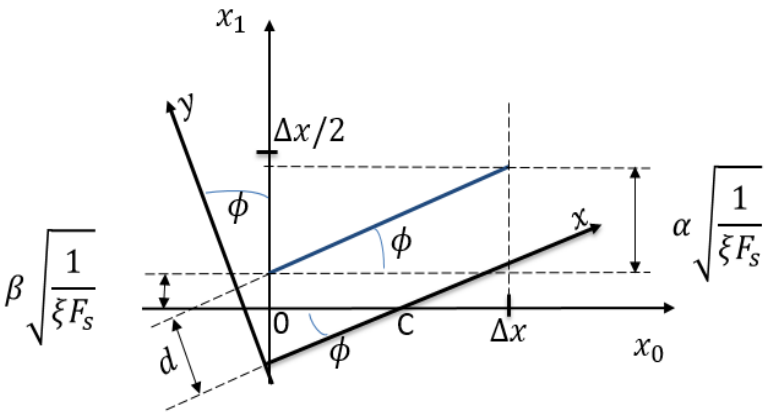


Figure 4.7: Geometric schematic for calculating the sinusoidal frequency in the fractional domain.

fractional domain (x, y) could become a chirp with chirp rate α and initial frequency β by a FRFT with angle $\phi = -\arctan \frac{\alpha N_w}{F_s^2}$.

Simplified calculation technique for the FRFT

The discrete FRFT calculation is presented in Appendix A. Although fast computation of the FRFT of a signal takes not much longer when compared to the Fourier transform. Still, it involves three steps: x2 Shannon interpolation, convolution and then decimation, which makes it more complex than the Fourier transform. In this part, we introduce a simple method for the discrete FRFT calculation.

According to [50], the windowed Fourier transform of the FRFT of a signal corresponds to the short-time Fourier transform of the signal itself, with the window being the fractional Fourier transform of the initial one.

$$\begin{aligned}(\mathbb{F}^{\phi_i} s_m)(u) &= \mathrm{DFT}^{-1}\left\{\mathrm{DFT}[\mathbb{F}^{\phi_i} s_m(n)]\right\} \\ &= \mathrm{DFT}^{-1}\left\{\mathrm{DFT}\left[s_m(n)(\mathbb{F}^{\phi} h_m(n))\right]\right\} \\ &= s_m(n)(\mathbb{F}^{\phi} h_m(n)),\end{aligned} \tag{4.17}$$

where $s_m(n)$ is the m^{th} windowed signal, $h_m(n)$ is the window for the signal $s_m(n)$ and u is the discrete variable for x in the fractional plane. So if we ascertain the window type, we can build a dictionary of FRFT of the window with different values

of FRFT angle. The FRFT of any signal is thus simply obtained by projecting the signal onto the dictionary.

4.3.2 Problem formulation

Similarly, in part 4.2.1, the m^{th} signal segment of $s(n)$ is written as:

$$s_m(n) \approx \sum_{e=1}^{E} A_{e,m} \exp\left\{ j2\pi \left[\alpha_{e,m} \frac{n^2}{2F_s^2} + \beta_{e,m} \frac{n}{F_s} \right] \right\} + v(n), \tag{4.18}$$

where $0 \leq n \leq N_w - 1$, $A_{e,m}$, $\alpha_{e,m}$ and $\beta_{e,m}$ are respectively the complex amplitude, the chirp rate and the initial frequency of the e^{th} chirp over the m^{th} window.

The windowed signal in (4.18) can be interpreted as a sum of harmonic signals rotated by certain angles. Thus, (4.18) can be rewritten as,

$$s_m(n) \approx \sum_{e=1}^{E} A_{e,m} \mathbb{F}^{\phi_{e,m}} \left[\exp(j2\pi d_{e,m} \frac{n}{F_s}) \right] + v(n), \tag{4.19}$$

where $d_{e,m}$ is the frequency value of a sinusoid in the fractional plane which corresponds to the e^{th} chirp in the m^{th} window (see Fig. 4.7). From part (4.3.1), we have:

$$d_{e,m} = \sqrt{N_w} \cos(\phi_{e,m}) \left(\frac{\beta_{e,m}}{F_s} + \frac{\tan\phi}{2} \right) \tag{4.20}$$

and

$$\phi_{e,m} = \arctan \frac{\alpha_{e,m} N_w}{F_s^2}. \tag{4.21}$$

From (4.6) and (4.21), the FRFT angle ϕ in general has to be in the range:

$$\begin{aligned} &\tan^{-1}\left(-\frac{\alpha_{\max} N_w}{F_s^2} \right) \leq \phi \leq \tan^{-1}\left(\frac{\alpha_{\max} N_w}{F_s^2} \right) \\ &\leftrightarrow \tan^{-1}\left(-\frac{1}{2} \right) \leq \phi \leq \tan^{-1}\left(\frac{1}{2} \right). \end{aligned} \tag{4.22}$$

From (4.6), we can say the parameter space of interest is:

$$\boldsymbol{\Omega} = (\phi, d) \text{ such that } \begin{cases} |\phi| \leq \tan^{-1}\left(\frac{1}{2}\right), \\ d = \sqrt{N_w}\cos(\phi)\left(\frac{\beta}{F_s} + \frac{\tan\phi}{2}\right), \\ -F_{\max} \leq \beta \leq F_{\max} - |\alpha| T_w & \alpha \geq 0, \\ -F_{\max} + |\alpha| T_w \leq \beta \leq F_{\max} & \alpha < 0, \\ \alpha = \frac{F_s^2}{N_w} \tan\phi. \end{cases} \tag{4.23}$$

Because $|\alpha|T_w = F_s \tan\phi$, (4.23) is equivalent with:

$$\mathbf{\Omega} = (\phi, d) \text{ such that } \begin{cases} |\phi| \leq \tan^{-1}\left(\frac{1}{2}\right), \\ d = \sqrt{N_w}\cos(\phi)\left(\frac{\beta}{F_s} + \frac{\tan\phi}{2}\right), \\ -F_{\max} \leq \beta \leq F_{\max} - F_s|\tan\phi|, & \phi \geq 0 \\ -F_{\max} + F_s|\tan\phi| \leq \beta \leq F_{\max}, & \phi < 0. \end{cases} \tag{4.24}$$

The discrete dictionary, to be used in CS, is designed by uniformly sampling the 2D parameter space $\mathbf{\Omega}$. Let I and J denote the total number of FRFT angle values and the total number of sinusoidal frequencies in the discrete dictionary. For the i^{th} FRFT angle in the dictionary, which we denote as $\tilde{\phi}_i$, let $\tilde{d_{i,j}}$ denote the corresponding possible values for the sinusoid frequency, where $j = 1, ..., Ji$. The 2D space of d and ϕ is illustrated in Fig. 4.8. We can see that the range of fractional frequency d is largest when $\phi = 0$, and it is smallest when $\phi = \pm 1/2$. Although the region of interest of d and ϕ is different from that of β and α (see Fig. 4.1), the dictionary content is similar because any chirp in TF domain can be expressed by a sinusoid in fractional domain with a certain rotation angle. By performing a

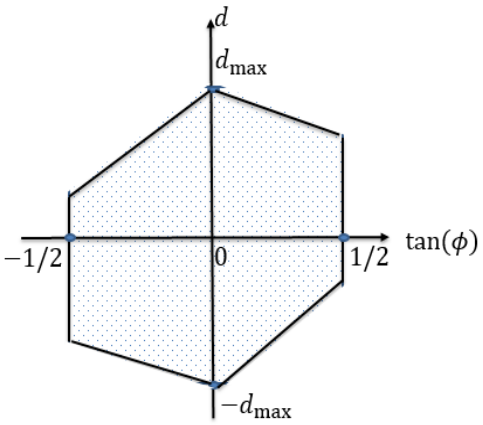


Figure 4.8: 2D space $\mathbf{\Omega}$ for d and ϕ.

sparse component analysis within each window, we can track the time variations of the chirp parameters of each component, thus estimating arbitrary IFs.

In vector form, the signal over the m^{th} window in (4.19) can be expressed as:

$$\mathbf{S}_m = \mathbf{\Psi}\mathbf{X}_m + \mathbf{V}_m, \tag{4.25}$$

where $\mathbf{S}_m = [s_m(0), ..., s_m(N_w - 1)]^T$, $\mathbf{V}_m = [v_m(0), ..., v_m(N_w - 1)]^T$. The FRFT dictionary $\mathbf{\Psi}$ is expressed as follows:

$$\begin{aligned} &\mathbf{\Psi} = [\mathbf{\Psi}_1, \mathbf{\Psi}_2, ..., \mathbf{\Psi}_I] \\ &\mathbf{\Psi}_i = [\boldsymbol{\psi}_{i,1}, \boldsymbol{\psi}_{i,2}, ..., \boldsymbol{\psi}_{i,J}] \\ &\boldsymbol{\psi}_{i,j}|_n = \mathbb{F}^{\tilde{\phi}_i}(\exp(j2\pi\tilde{d}_j \frac{n}{F_s})) \\ &i = 1, ..., I, j = 1, ..., J, n = 1, ..., Nw. \end{aligned} \tag{4.26}$$

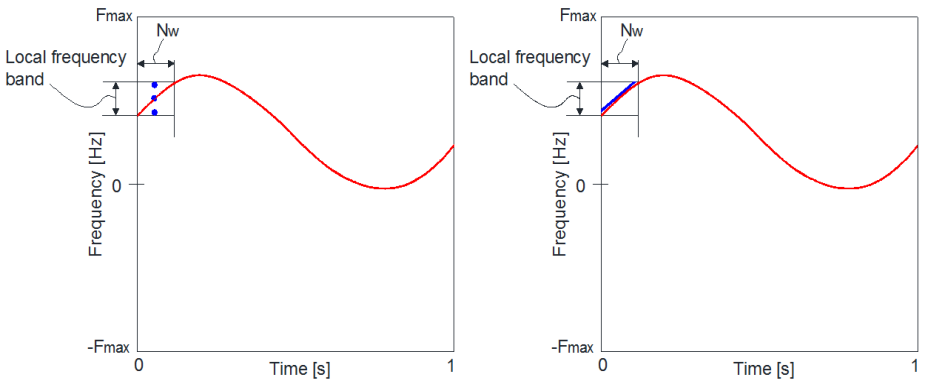
Since $E < N_w \ll \sum_{i=1}^{I} J$, then solving for $\mathbf{X}_m$ in (4.25) becomes a sparse recovery (or CS) problem, which can be solved by:

$$\hat{\mathbf{X}}_m = \arg\min \|\mathbf{X}_m\|_1 \quad s.t. \quad \|\mathbf{S}_m - \mathbf{\Psi}\mathbf{X}_m\|_2^2 \leq \epsilon \tag{4.27}$$

where ϵ is the noise level. The solution for (4.27) can be obtained by greedy algorithms such as the OMP or the linear programming [24, 45].

4.4 Chirp Dictionary and Sinusoid Dictionary Comparison

The sinusoid dictionary has been examined in the sparse signal reconstruction [3, 27]. Its drawback lies in the adverse window length and sparsity interlocking, namely, the longer the window, the lower sparsity of the local TF signature due to inclusion of a larger signal bandwidth. In the case of missing samples, longer windows are required to obtain a sufficient number of observations for stable recovery. This trade-off between sparsity and the required number of observations renders the sinusoidal dictionary ineffective for non-stationary signal reconstruction. The chirp dictionary does not suffer from this trade-off, or at least is less sensitive to it. When using the chirp dictionary, the sparsity of a signal segment depends only on the number of chirps contained in that window. It is illustrated in Fig. 4.9. In Fig. 4.9(a), the sparsity is defined by local frequency sparsity (local frequency sparsity = $\frac{\text{Local frequency band}}{\text{Frequency grid}}$). Thus, the longer window can result in the larger local bandwidth and reduction in sparsity. However, in Fig. 4.9(b) the sparsity is 1. Therefore, the chirp dictionary method has an advantage over the sinusoidal one when long window is necessary.



(a) (b)

Figure 4.9: Sparsity of a windowed signal when using: (a) Sinusoidal dictionary; (b) Chirp dictionary.

The chirp dictionary approach, however, may give an inaccurate TF signal reconstruction due to violation of the chirp piece-wise approximation of the signal TF signature when applying long windows. This problem is mitigated by applying an averaging over consecutive windows. The reconstructed values at each TF point corresponding to an overlapping window are added and a threshold is applied to remove small values. An accurate IF estimate would, therefore, benefit from persistent high values for a given TF point across neighbouring windows. In other words, accumulation of values at (n_i, f_i) strengthens the f_i estimate, whereas non-accumulative values are deemphasized.

Another advantage of the chirp dictionary approach over the sinusoid counterpart is that it is less sensitive to the picket fence effect. This phenomenon is illustrated and explained in Fig. 4.10 and Fig. 4.11. When the sinusoid dictionary is used and the signal frequency is 102.4 Hz, 102.5 Hz, 102.7 Hz, OMP selects either the atom of 102 Hz, or 103 Hz (see Fig. 4.10). Due to the low correlation between the signal and the chosen atom, large residual values are evident, and further iterations are executed, resulting in frequency content at false locations. The chirp dictionary addresses this inaccuracy (see Fig. 4.11). With the inclusion of the factor α, atoms with better matching with the signal are selected, which leads to a smaller residue,

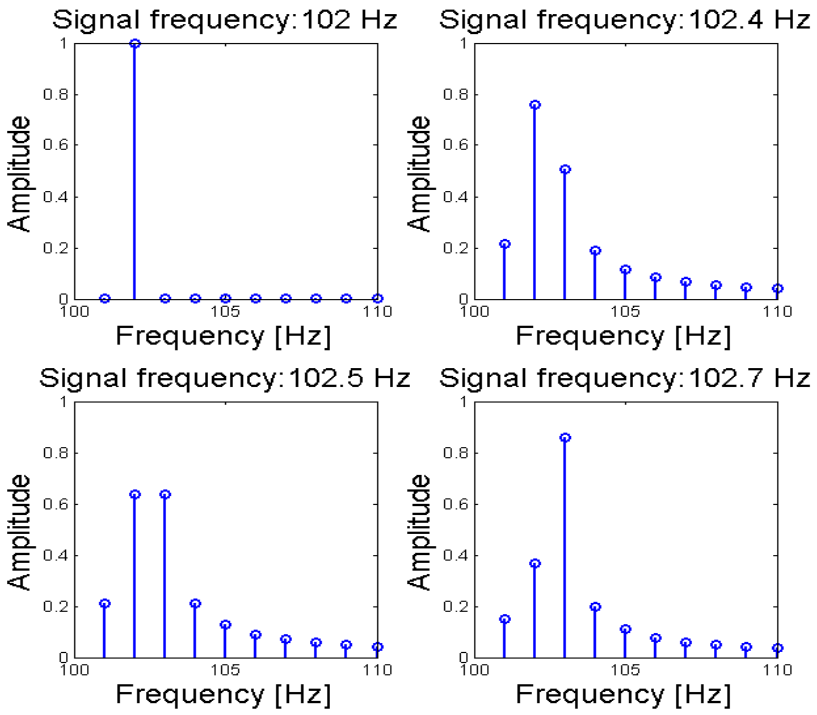


Figure 4.10: Expanded DFT.

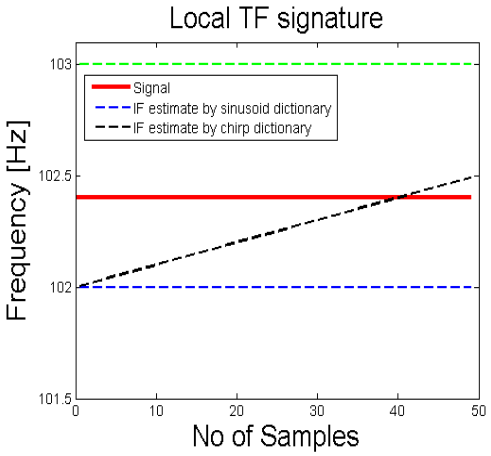


Figure 4.11: Local TF signature.

forcing the OMP to a halt. If further iterations continue, the result would be insignificant, close to noise floor.

The picket fence effect can be reduced by increasing the resolution or number of DFT points. However, an integer number of periods in the data segment being analyzed is never guaranteed. Moreover, as the radar return is typically a non-stationary signal with changing frequency over time, the picket fence effect is inevitable. Therefore, the chirp dictionary approach would outperform its sinusoidal counterpart most of the time.

4.5 Restricted Isometric Property (RIP) Analysis of The Chirp Dictionary

In this section, we examine the RIP associated with the chirp dictionary used in the previous section. Similar to the work in [51], we consider the bounds on the eigenvalues of the outer product of the dictionary matrix. We show that these bounds compete with those of Gaussian random dictionaries and as such lead to the same conditions on sparsity and compressed observations. The analysis follows closely that of [51] but differs in the final results due to differences in dictionary structure. Let $Q = \sum_{i=1}^{I} J_i$ and $\mathbb{Q} = \{1, ..., Q\}$. The structure of matrix $\mathbf{\Psi}$ is described in (4.8).

A matrix $\mathbf{\Psi}$ is said to satisfy the RIP of order E if there exists a $\delta_E \in (0,1)$ such that:

$$(1-\delta_E)\|\mathbf{x_\Gamma}\|_2^2 \leq \|\mathbf{\Psi}_\Gamma \mathbf{x}_\Gamma\|_2^2 \leq (1+\delta_E)\|\mathbf{x_\Gamma}\|_2^2, \tag{4.28}$$

where $\mathbf{x}_\Gamma$ is the vector obtained by retaining only the entries in $\mathbf{x}$ corresponding to the columns' indices in Γ. Because $\mathbf{\Psi}_\Gamma^H \mathbf{\Psi}_\Gamma$ is a Hermitian matrix, (4.28) is equivalently given by:

$$\begin{aligned} & 1-\delta_E \leq \frac{\|\mathbf{\Psi}_\Gamma \mathbf{x}_\Gamma\|_2^2}{\|\mathbf{x_\Gamma}\|_2^2} \leq 1+\delta_E \\ \iff & 1-\delta_E \leq \frac{\left\langle \mathbf{\Psi}_\Gamma^H \mathbf{\Psi}_\Gamma \mathbf{x}_\Gamma, \mathbf{x}_\Gamma \right\rangle}{\|\mathbf{x_\Gamma}\|_2^2} \leq 1+\delta_E \\ \iff & 1-\delta_E \leq \frac{\langle \lambda \mathbf{x}_\Gamma, \mathbf{x}_\Gamma \rangle}{\|\mathbf{x_\Gamma}\|_2^2} \leq 1+\delta_E \\ \iff & 1-\delta_E \leq \lambda_{min} \leq \lambda_{max} \leq 1+\delta_E, \end{aligned} \tag{4.29}$$

where $\lambda_{\max}$, λ_{min} are the maximum and minimum eigenvalues of $\mathbf{\Psi}_\Gamma^H \mathbf{\Psi}_\Gamma$. In order for the chirp dictionary to satisfy the RIP, the Grammian matrix $\mathbf{\Psi}_\Gamma^H \mathbf{\Psi}_\Gamma$ has all of its eigenvalues in $(0,2)$. Since the chirp dictionary $\mathbf{\Psi}$ is deterministic, the above requires checking all $\binom{Q}{E}$ possible Γ, which can be a computationally formidable problem. According to [52], a Gaussian random matrix $\mathbf{G} \in \mathbb{C}^{N_w \mathrm{x} Q}$ with entries of zero mean and variance $1/N_w$ can satisfy $\delta_E < 1$ with a number of measurements $O(E\log(Q/E))$. Therefore, we will compare the bounds on the eigenvalues of $\mathbf{\Psi}_\Gamma^H \mathbf{\Psi}_\Gamma$ with those of $\mathbf{G}_\Gamma^H \mathbf{G}_\Gamma$. In the simulations, $N_w = 50$, $Q = 16512$ and 100000 random realizations of subset Γ are used to estimate the eigenvalue statistics. The simulations are repeated for different cardinalities of subset Γ. Fig. 4.12 shows the bounds (sample mean + 3 STD for the maximum eigenvalue and sample mean -3 STD for the minimum eigenvalue) for both the chirp and the Gaussian dictionaries.

Fig. 4.12 shows that the bounds for the two dictionaries are very close to each other. This implies that $\mathbf{\Psi}$ can also satisfy condition $\delta_E < 1$ with high probability if the minimum number of observations is $O(E\log(Q/E))$.

4.6 Simulation

This section evaluates the performance of sparse reconstruction of the signal TF signature using a chirp dictionary, especially in the case of missing samples. We

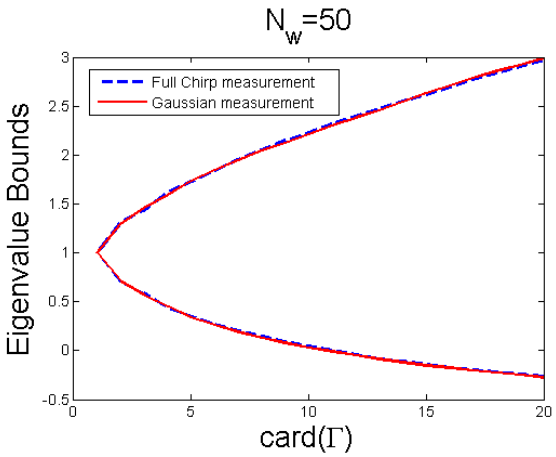


Figure 4.12: Eigenvalue bounds of $\mathbf{\Psi}_\Gamma^H \mathbf{\Psi}_\Gamma$ and $\mathbf{G}_\Gamma^H \mathbf{G}_\Gamma$.

compare the proposed approaches with sinusoid atoms, and the discrete chirp Fourier transform (DCFT). The DCFT is formed by projecting the signal onto the chirp dictionary. But it does not calculate the rest for the next projection like OMP. Thus, IF estimates from the first iteration of OMP are the same as those provided by the location of highest value in the DCFT. However, sparsity is not known a-priori, and it is usually larger than one; thus many iterations are needed in the implementation, which represents an advantage of the chirp dictionary over the DCFT. The two chirp dictionary approaches are proved to give a similar performance. To access the accuracy of sparse reconstruction algorithms, concentration level (ζ) is deployed. This is the ratio of the sum of the pixel magnitudes in the TF domain along the ground-truth (i.e., the actual IF) and the rest of the TF values. The higher ζ, the more accurate is the result. This is illustrated in Fig. 4.13.

In the following examples, the signals are sampled at the Nyquist rate with a sampling frequency $F_s = 256$Hz and the total signal length is $N = 256$. The data is then randomly under-sampled to create the incomplete data to be processed. The input signal is corrupted by white Gaussian noise, and the signal-to-noise ratio is $SNR = 30$ dB. A rectangular window is applied. The resulting image is normalized and transfered to the energy version for display. As we have discussed in section 4.5, the chirp dictionary can satisfy the RIP condition if the minimum number of observation is $E \log(Q/E)$. In this simulation part, we assume $E = 5$, but the

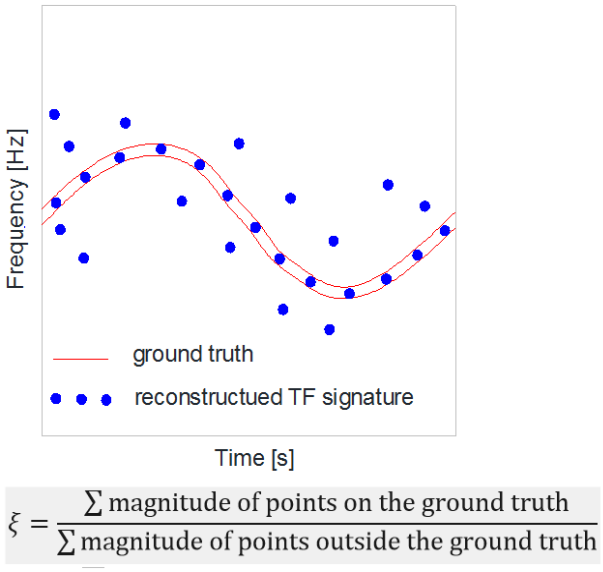


Figure 4.13: Illustration of the concentration level ζ.

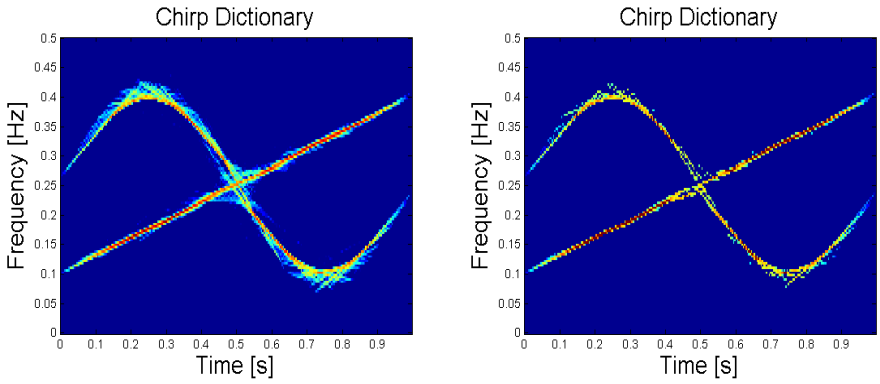
maximum actual signal components in the simulation is 3. Thus we need at least 30 measurements per window to obtain reliable results. Depending on the signal type, different window length is chosen. The average window length is 70, and we often cut 50% of data, which gives us around 35 samples per window. If we cut more samples, inaccurate estimations can happen.

4.6.1 Effect of averaging in TFRs obtained by the chirp dictionary approach

The first example illustrates the TF signature obtained by chirp atoms in the two cases with and without using the averaging method. The input signal is expressed as:

$$\begin{aligned} s(n) = \exp\left\{ j\left[(0.15F_s)\cos(2\pi\frac{n}{F_s} + \pi) + 2\pi(0.25F_s)\frac{n}{F_s} \right] \right\} \\ + \exp\left\{ j2\pi\left[(0.1F_s)\frac{n}{F_s} + (0.3F_s)\frac{n^2}{2F_s^2} \right] \right\} + v(n), \end{aligned} \tag{4.30}$$

where $n = 0, 1, ..., N-1$. We randomly discard 50% of the data. The window length is set to $N_w = 50$. The simulation results are shown in Fig. 4.14. The combination



(a) (b)

Figure 4.14: Local reconstruction of a two-component signal in (4.30) obtained by: (a) Chirp atoms without averaging; (b) Chirp atoms, using averaging.

of the chirp dictionary and the averaging method provides more accurate local signal frequency structure with $\zeta = 31.6678$, compared with $\zeta = 9.0554$ without averaging. So, from now on, we apply averaging on TFRs obtained by the chirp atom method.

4.6.2 Comparisons between the chirp and the sinusoid dictionary approaches

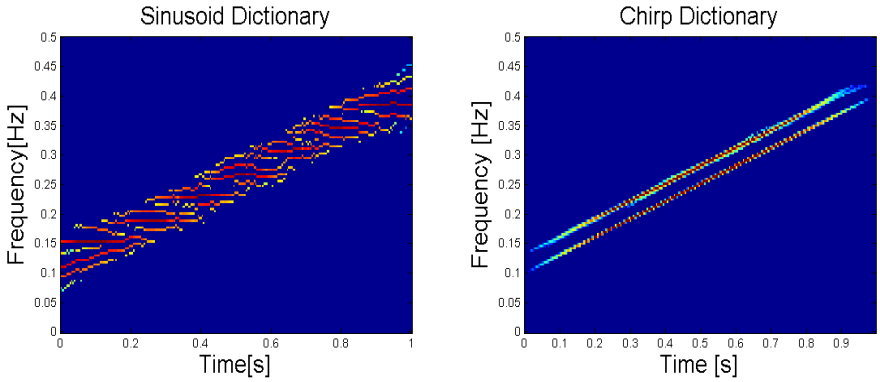
Relationship between the sparsity and the window length

In this example, the signal consists of two closely-parallel chirps. Its discrete-time version is expressed as:

$$\begin{aligned} s(n) &= \exp\left\{j2\pi[(0.1F_s)\frac{n}{N} + (0.3F_s)\frac{n^2}{2N^2}]\right\} \\ &+ \exp\left\{j2\pi[(0.13F_s)\frac{n}{N} + (0.3F_s)\frac{n^2}{2N^2}]\right\} + v(n) \end{aligned} \tag{4.31}$$

where $n = 0, 1, ..., N-1$. To capture enough data to resolve the two chirps, the window size is set to a large value, $N_w = 90$. The sparsity level is assumed to be $E = 5$. The result in Fig. 4.15(a) shows the failure of local reconstruction using the sinusoidal dictionary due to lack of sparsity in frequency with $\zeta = 3.5$. In contrast,

when the chirp dictionary is used, the sparsity remains constant, irrespective of the window size N_w. The two chirps are clearly resolved as evident from Fig. 4.15(b). The concentration level is $\zeta = \infty$.



(a) (b)

Figure 4.15: Local reconstruction of a two-component signal in (4.31) when 50 % of data is missing using (a) Chirp atoms; (b) Sinusoid atoms.

Picket fence effect

In the next example, we show that the chirp approach outperforms its sinusoidal dictionary counterpart even for sinusoid signals. For off-bin-center sine waves and $\Delta f = 1$Hz, the chirp approach obtains more accurate IF estimates, as it can alleviate the picket fence effect. The simulated signal is described in Table. 4.1. The window length is $N_w = 50$, and a rectangular window is deployed for both dictionary approaches. 50% of the signal is randomly shorted. When a full data is available, no noise, and the signals' periods are both integer, both methods provide perfect frequency localization with $\zeta = \infty$, as illustrated in Fig. 4.16(a,b). The chirp dictionary shows clear advantage over its sinusoidal counterpart with added noise and data missing. The corresponding performance measurements are $\zeta = 180$ and $\zeta = 41$, respectively, and the results are shown in Fig. 4.16(c,d). With a non-integer period signal, the sinusoid dictionary method significantly suffers, showing a noisy

Table 4.1: Signal $s(n)$

Signal $s(n)$	f_1	f_2
$\exp\left\{j2\pi\left[f_1\frac{n}{F_s}+f_2\frac{n}{F_s}\right]\right\}$	100	25
$\exp\left\{j2\pi\left[f_1\frac{n}{F_s}+f_2\frac{n}{F_s}\right]\right\}$	102.4	25.6

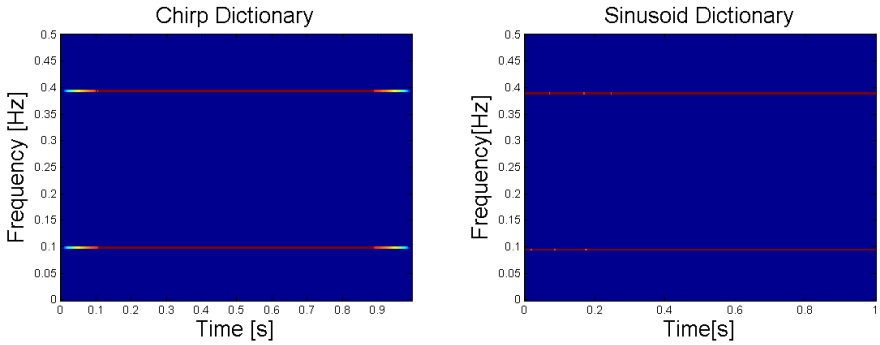
time-frequency signature as depicted in Fig. 4.16(f) in contrast with the result of the chirp dictionary shown in Fig. 4.16(e). The values of ζ of chirp and sinusoid dictionary approaches are $\zeta = 143$ and $\zeta = 5$, respectively. Higher resolution or smaller frequency grid helps mitigate the picket fence effect. With $\Delta f = 0.5$ Hz, the same signal expressed in Table. 4.1, $SNR = 30$dB, and 50% data missing, the concentration level is $\zeta = 28$ compared with $\zeta = 5$ when $\Delta f = 1$ Hz. The result is shown in Fig. 4.17. As the signal frequency value is unlikely to fall exactly in a frequency bin, the chirp dictionary provides more reliable IF estimations.

4.6.3 Comparisons between the chirp dictionary approach and the DCFT

This simulation compares the accuracy of a local TF signature when the chirp dictionary and DCFT are used. Two input signals $s_1(n)$ and $s_2(n)$ are employed,

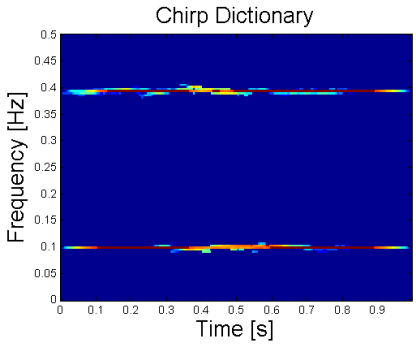
$$\begin{aligned} s_1(n) &= \exp\left\{j2\pi\left[(0.4F_s)\frac{n}{N}-(0.3F_s)\frac{n^2}{2N^2}\right]\right\}+v(n) \\ s_2(n) &= \exp\left\{j\left[(0.15F_s)\cos(2\pi\frac{n}{N})+2\pi(0.25F_s)\frac{n}{N}\right]\right\} \\ &+\exp\left\{j\left[(0.15F_s)\cos(2\pi\frac{n}{N}+\pi)+2\pi(0.25F_s)\frac{n}{N}\right]\right\} \\ &+\exp\left\{j2\pi\left[(0.25F_s)\frac{n}{N}\right]\right\}+v(n), \end{aligned} \tag{4.32}$$

where $n = 0, 1, ..., N-1$. The window size is $N_w = 50$ and a rectangular window and averaging method are utilized in both methods. The results are shown in Fig. 4.18. When the input signal is a mono-component $s_1(n)$ and assuming that the signal

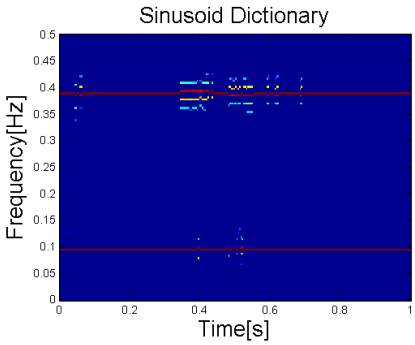


(a)

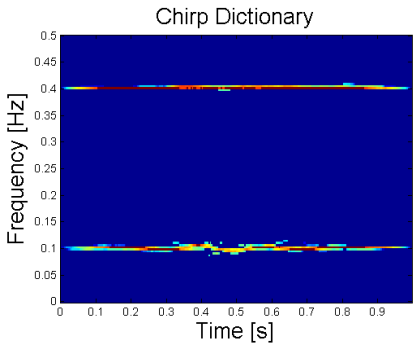
(b)



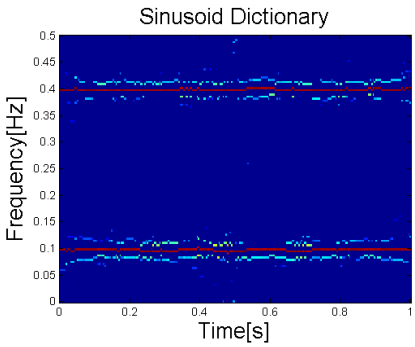
(c)



(d)



(e)



(f)

Figure 4.16: Sparse reconstruction of bin-center sine waves (Table 4.1) using chirp and sinusoid dictionary approaches, $\Delta f = 1$Hz: (a,b) Full data of integer period signal; (c,d) Integer period signal with 50 % samples missing; (e,f) Non-integer period signal with 50 % samples missing.

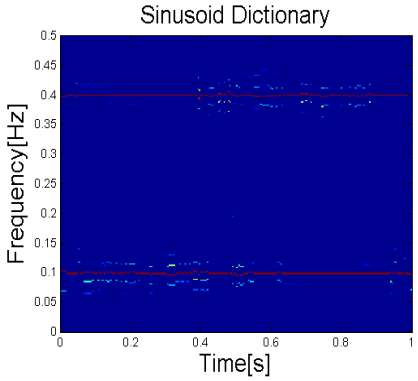


Figure 4.17: Sparse reconstruction of off-bin-center waves using sinusoid dictionary, and $\Delta f = 0.5$ Hz.

sparsity is known, then the concentration levels of the two approach become $\zeta = \infty$. The chirp approach provides better TF signature in the case of a multi-component input $s_2(n)$ with $\zeta = 62$, compared with $\zeta = 1.5$ resulting from the DCFT.

4.6.4 Comparisons among the two chirp dictionary approaches, the sinusoidal dictionary and the DCFT

The first example displays the performance of the four methods when 50% data are missing randomly. The window length is set to $N_w = 50$. The signal is expressed as:

$$\begin{aligned} s(n) = \exp\left\{j2\pi[(0.1F_s)\frac{n}{F_s} + (0.3F_s)\frac{n^2}{2F_s^2}]\right\} + \\ \exp\left\{j2\pi[(0.1F_s)\frac{n}{F_s}]\right\} + v(n), \end{aligned} \tag{4.33}$$

with $n = 0, 1, ..., N-1$. The result is shown in Fig. 4.19.

It can be seen that the two chirp approaches give pretty similar TF estimations, with the concentration level being around 35. They successfully resolve the TF signature under noise and incomplete data. Fig. 4.19(a) shows the failure of local reconstruction using the sinusoid dictionary due to lack of sparsity in frequency and the picket fence effect. The concentration level is $\zeta = 1.73$. The DCFT gives a little bit better result with $\zeta = 2.87$. Notice that when DCFT is applied, in some

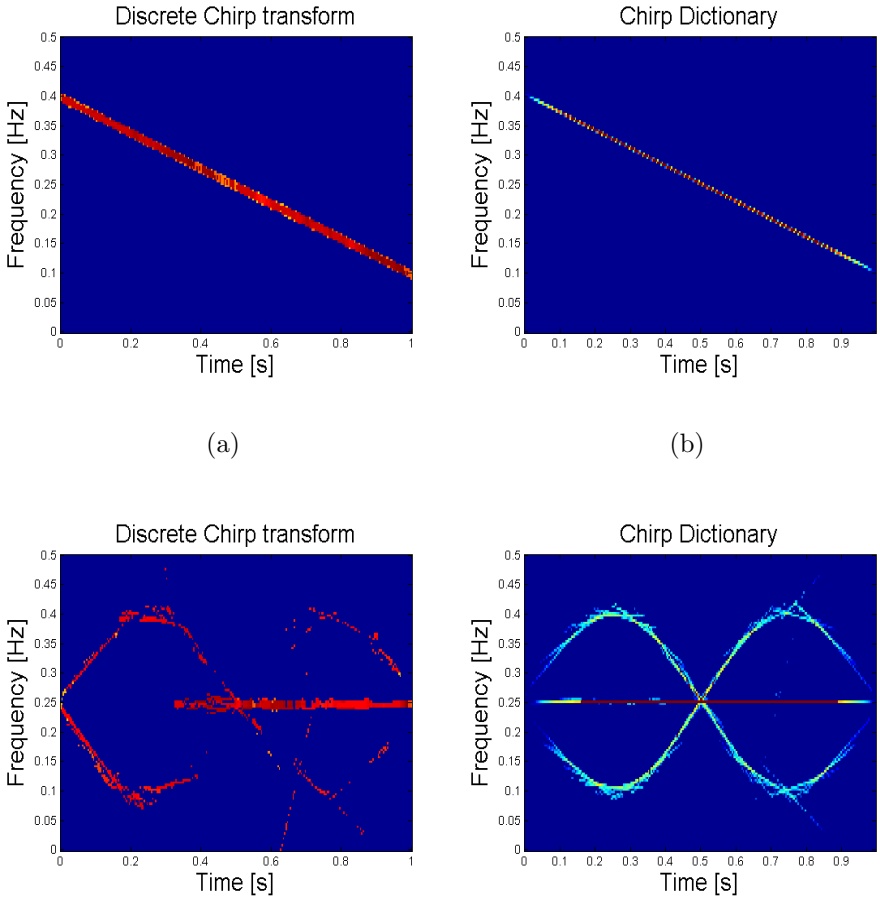
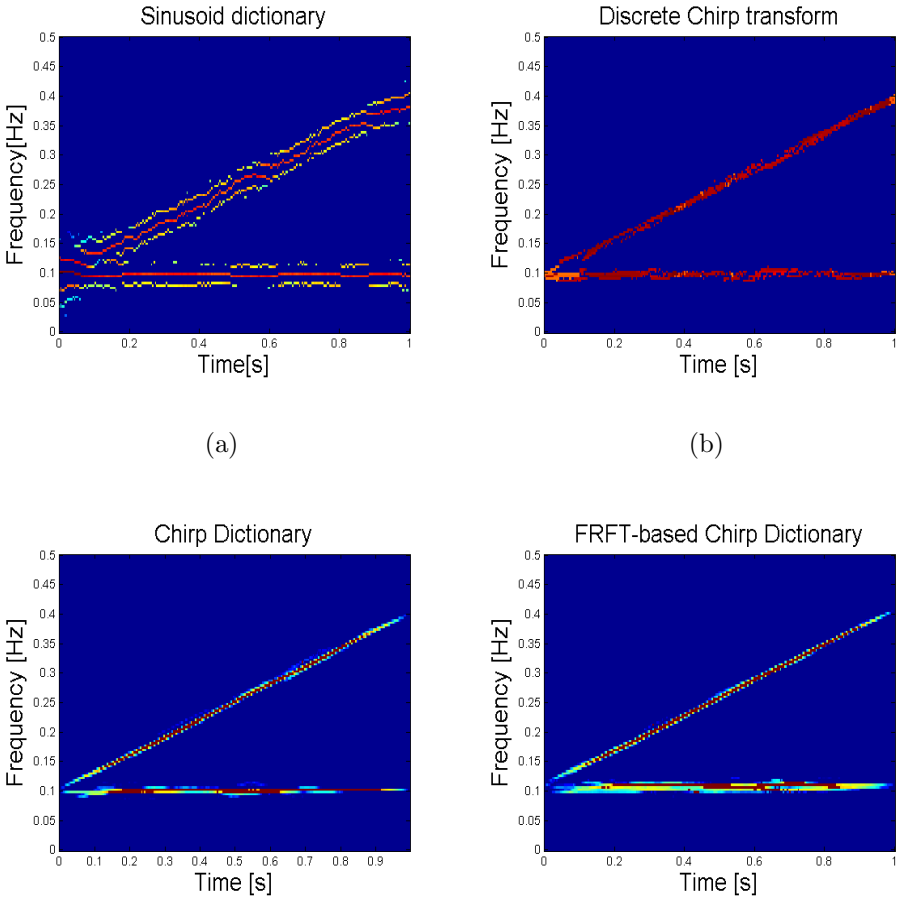


Figure 4.18: TFRs of $s_1(n)$ and $s_2(n)$ in (4.32) with: (a)(c) DCFT; (b)(d) Chirp atoms.



(a) (b)

(c) (d)

Figure 4.19: TFR of signal in (4.33) with 50% data missing, $N_w = 50$ obtained by: (a) Sinusoidal dictionary; (b) DCFT; (c) Chirp dictionary; (d) FRFT based chirp dictionary.

windowed segments, only one component is detected. It is because the signal is only projected onto the dictionary one time, and only the component with the highest coefficient is chosen.

The third example displays a signal composed of three components, expressed as follows:

$$\begin{aligned} s(n) &= \exp\left\{ j(0.15F_s)\cos(2\pi\frac{n}{F_s}+\pi) + j2\pi(0.25F_s)\frac{n}{F_s} \right\} \\ &+ \exp\left\{ j(0.15F_s)\cos(2\pi\frac{n}{F_s}) + j2\pi(0.25F_s)\frac{n}{F_s} \right\} \\ &+ \exp\left\{ j2\pi[(0.25F_s)\frac{n}{F_s}] \right\} + v(n), \end{aligned} \tag{4.34}$$

where $n = 0, 1, ..., N-1$. Similarly, 50% data is randomly removed. The window length is $N_w = 50$. The results are shown in Fig. 4.20. Similar to other examples, the sinusoidal dictionary method fails to give an accurate TF estimation with $\zeta = 1.658$. This example also shows the advantage of the chirp approach over the DCFT when there are multiple components. They obtain a ζ of approximately 19, compared with $\zeta = 1.619$ for the DCFT.

The last simulation uses the data from human gait radar returns. The data has 20000 samples, which is first sampled at a Nyquist rate $F_s = 1000$Hz, and then randomly thinned by discarding 50% of data samples. The window length is set to $N_w = 128$. A rectangular window is applied for the discrete chirp transform and the chirp dictionary, whereas a Hanning window is used for the approach of sinusoid atoms. The result in Fig. 4.21 shows that the chirp dictionary obtains the best results representing the torso and limb's micro-Doppler.

4.7 Conclusion

The accurate piece-wise chirp approximations to the TF signature of many Doppler and micro-Doppler signals motivate the use of the chirp dictionary for sparse reconstruction of the signal's local frequency structure under full and incomplete data. This chapter has presented two approaches to build the chirp dictionary. The first one simply includes all chirps with possible values of chirp rate and initial frequency.

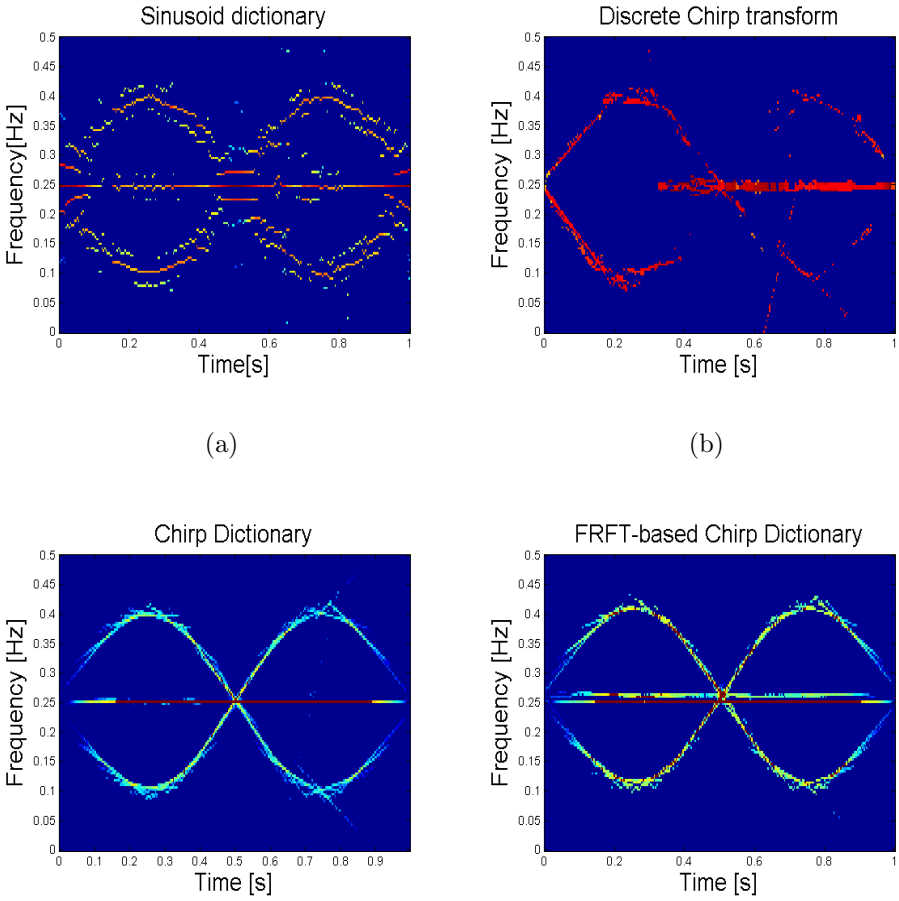


Figure 4.20: TFR of signal in (4.34) with 50% data missing, $N_w = 50$ obtained by: (a) Sinusoidal dictionary; (b) DCFT; (c) Chirp dictionary; (d) FRFT based chirp dictionary.

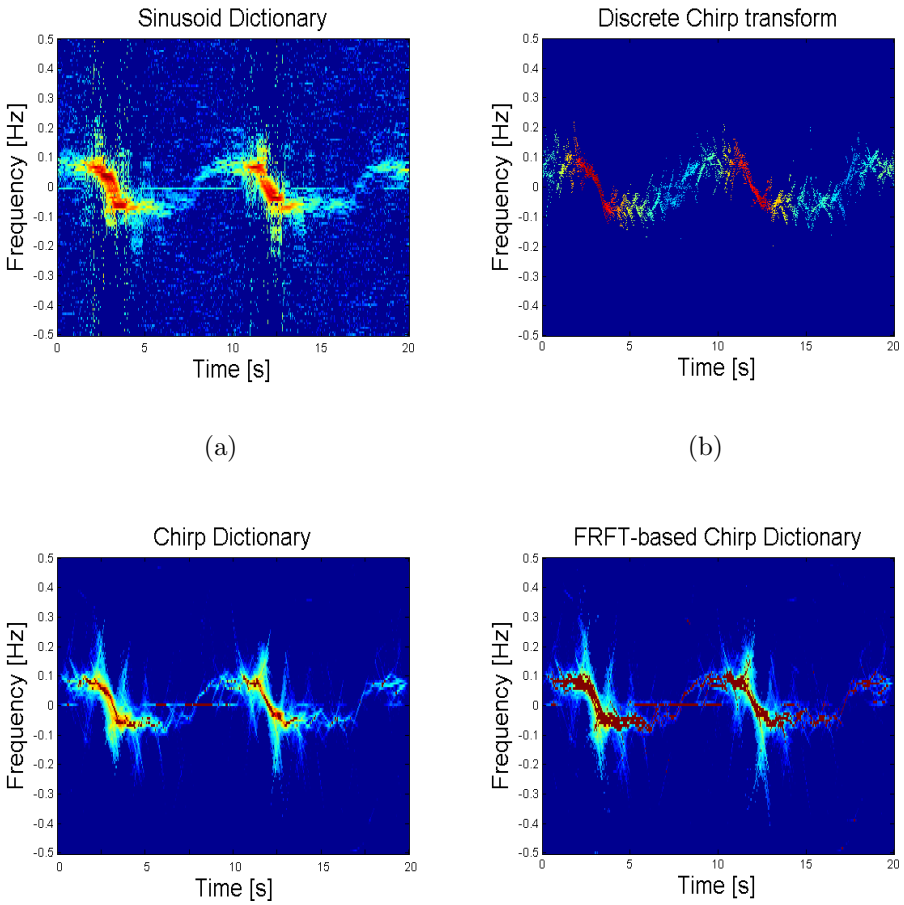


Figure 4.21: Local reconstruction of a real signal returned from a human with one arm when 50% data is missing: (a) Sinusoid dictionary; (b) Discrete Chirp transform; (c) Chirp dictionary; (d) FRFT based Chirp dictionary.

The second one is also composed of all possible chirps which are formed from sinusoids in the fractional domain with all eligible values of FRFT angles. The chirp dictionary formed by each way are shown similar performances. We have examined the chirp and sinusoidal dictionary performance over synthetic and real data in the case of incomplete observations. Although both methods are based on the reconstruction of data within short overlapping time intervals defined by a sliding window, the chirp dictionary outperforms the sinusoidal dictionary. The chirp dictionary approach can relax the drawbacks of its sinusoid counterpart, which are the picket fence effect, and converse requests on the number of measurements for exact recovery, and sparsity. Therefore, it provides more accurate approximations to the TF signature of non-stationary signals. In comparison with the discrete chirp Fourier transform, the chirp dictionary method also provides superior performance when the input is a multiple-component signal.

Chapter 5

Simplified Chirp Dictionary

5.1 Introduction

5.1.1 Motivation

Non-stationary signals are typically deployed to characterize speech, biomedical signals, sonar and radar returns (to name but a few) [32, 34, 35]. These signals' frequency can be constant, or linear/non-linear functions of time. In radar applications, they are commonly referred to as Doppler and micro-Doppler signals [36, 53, 54, 55]. Being able to analyze these Doppler and micro-Doppler frequencies correctly is extremely critical in the radar field [33, 37, 56, 57, 58]. For example, we can measure the velocity and direction of a bulk motion or the vibration of targets' structures by examining the frequency shifts or frequency modulations on the reflected signals.

There have been numerous methods of TFD analysis proposed. The short-time Fourier transform is the simplest linear TF signal representation [5, 47]. The TF estimation is obtained by computing the Fourier transform over a sliding window in time. The square modulus of the short-time Fourier transform is the spectrogram. The major drawback of the spectrogram is that its efficiency depends on the employed window size and shape. Improvement in resolution can be achieved by using WVD, which is obtained by calculating the Fourier transform of the bilinear product of the signal. WVD is ideal for a chirp because its bilinear product is a sinusoid. However, in the case of multiple-component signals, WVD suffers from cross-terms appearing between different components. In order to suppress the cross-terms in

the WVD, a class of reduced interference distribution, which belongs to Cohen's class, has been used. The TFD can also be obtained by the use of matching pursuit algorithm [40, 41]. This technique decomposes any signal into a linear expansion of waveforms that are selected from a redundant dictionary of functions. These waveforms are chosen in order to best match the signal structures. These methods, however, are not designed for compressed data. Thus, they deliver unreliable TFDs in the case of missing data.

The exponential growth of data demands new ways of collecting, representing and analyzing samples. In recent years, compressive sensing (CS), which helps deal with few observations, has attracted widespread interest. Incomplete samples, or random sampling in the field of radar can frequently happen due to range ambiguity, discarding noisy measurements, hardware simplification, sampling frequency limitations, or co-existence of various wireless services with active or passive sensing models[1, 59, 60]. Thus, a TFD which is robust with missing data is of significance.

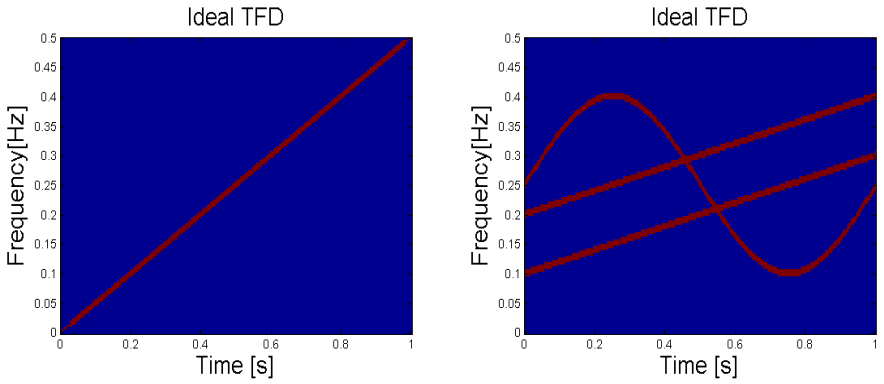
Consider the following linear model,

$$\mathbf{s} = \mathbf{\Psi}\mathbf{x}, \tag{5.1}$$

where $\mathbf{\Psi}$ maps the Nyquistly sampled data given by vector $\mathbf{x}$ into compressed measurements expressed by vector $\mathbf{s}$. Since (5.1) is an under-determined problem, additional information is necessary for finding a solution. The extra condition to obtain an accurate answer for (5.1) is that $\mathbf{x}$ is sparse. Fortunately, non-stationary signals are sparse in TF domain as seen in Fig. 5.1. Thus, CS has been used for obtaining good estimation of a signal's instantaneous frequency in the case of incomplete data [2, 3, 27, 28, 44].

5.1.2 Related work

Several approaches have been proposed to achieve sparse TFRs [2, 27, 28, 30, 31, 44, 61, 62, 63, 64, 65, 66, 67]. Generally, we can divide them into two categories: parametric and non-parametric methods. The former [30, 31] requires knowledge of the signal's structure. It only works well when there is a good match between the assumed and the actual signal characteristics. Non-parametric methods can be applied to any signal's structures. The work in [2] is an example. The compressed



(a) (b)

Figure 5.1: Ideal TFDs for: (a) Single-component signal; (b) Multiple-component signal.

measurements are defined in the ambiguity domain through the application of an appropriate mask. This approach does not consider the case of random or missing samples in time. Due to incomplete data, the signal ambiguity domain is contaminated by noise-like artifacts. Thus, any mask in this domain becomes unreliable.

To obtain reliable TFRs in the case of incomple data, a straightforward method is sparse reconstruction from windowed data in the time domain, deploying a partial Fourier basis [3]. This is similar to the spectrogram but compressive techniques are used and thus better results are attained. However, this method suffers from the trade-off between necessary measurements for accurate recovery and sparsity when considering the window size, and also the picket fence effect when there is a non-integer period in the analyzed data segments. Thus, another measurement basis is required to obtain more stable and reliable results.

In many situations, the frequency law of non-stationary signal segments can be represented as a weighted sum of piece-wise linear chirps, in which most of the parameter coefficients are zero. In this respect, they are sparse in the joint time-frequency domain. Thus, the segments' time-frequency signature can be recovered by sparse reconstruction with the measurement dictionary being chirp atoms [13, 68]. Greedy algorithms or convex optimization techniques are used to obtain the

sparsest chirp combination that best describes the windowed signals. Compared with the sinusoidal dictionary method, a better performance is obtained under both full and limited data because signals are more sparse in this representation, and longer windows can be employed without sensible reduction in sparsity. It is also not susceptible to the picket fence effect.

The chirp approach, nevertheless, deploys a very large dimension measurement dictionary. Since there are two parameters to be estimated (i.e. the chirp rate and the initial frequency), the dictionary dimension can be equal to the square of the dimension when using the sinusoid atom. This very large atom set leads to a much heavier computation burden and a longer calculation time. Therefore, in order to obtain good TF estimation at low computational complexity, chirp dictionary simplification methods are needed.

5.1.3 Chapter contribution

In this chapter, we introduce two approaches which reduce the chirp dictionary dimension, thus lower the calculation load. In the first approach, we estimate the chirp rate through the DTFT of the bilinear product at a certain time lag. The initial frequency is solved in the time domain, with a lower dimensional dictionary than the computationally complex of the full chirp atom. This approach allows us to obtain a good estimation of the TF features of incomplete and full-data signals at lower computational complexity. In the second approach, the fractional Fourier transform (FRFT) is used to obtain an initial frequency for each chirp-rate. This leads to a much simpler chirp atom set. As the FRFT can be executed with similar computational complexity as the Fourier transform, the advantage of this approach is that we get a good performance but with an improved computational efficiency.

5.1.4 Chapter outline

The paper is organized as follows. Section 5.2 discusses the computational requirement of the full chirp dictionary approach. The first simplified chirp atom method is then introduced in section 5.3. This part discusses CS and calculating the IAF for the whole signal and for the windowed signal. The algorithm which gives estimation of chirp rate in IAF and builds the compact chirp dictionary is introduced after

that. Next, we focus on the RIP associated with the simplified chirp dictionary and provide a lower bound on the number of observations for exact recovery. Simulations with synthetic and real signals will end this section. Section 5.4 talks about the second method for simplifying the chirp dictionary. It starts with the concept of the FRFT, which is followed by a description of the simplified chirp dictionary using the FRFT. The RIP associated with the simplified chirp atom using the FRFT is also considered. The section ends with a simulation involving multiple signals. The conclusions are given in section 5.5.

5.2 Calculation Load in The Full Chirp Dictionary Approach

In the full chirp dictionary approach, a discrete signal segment of length N_w is approximated as a sum of E ($E \geq 1$) chirps. This means that we have to estimate the chirp rates (α) and the initial frequencies (β) of the E chirps in each data segment. This task is carried out by CS techniques with a full chirp atom $\mathbf{\Psi}_F$. The parameter space of interest is [68], [13]:

$$\begin{aligned} \mathbf{\Omega} = \{&(\alpha, \beta) \quad \text{such that} \\ &|\alpha| \leq F_{\max} F_s / N_w \quad \text{and} \quad |\alpha N_w / F_s + \beta| \leq F_{\max}\}, \end{aligned} \tag{5.2}$$

where $F_{\max}$ is the maximum frequency of the signal and $F_s = 2F_{\max}$ is the sampling frequency.

Now (5.2) can be rewritten as:

$$\begin{aligned} \mathbf{\Omega} = \{&(\alpha, \beta) \quad \text{such that} \\ &|\alpha| \leq F_{\max} F_s / N_w \\ &-F_{\max} \leq \beta \leq F_{\max} - |\alpha| T_w \qquad \text{if} \qquad \alpha \geq 0 \\ &-F_{\max} + |\alpha| T_w \leq \beta \leq F_{\max} \qquad \text{if} \qquad \alpha < 0 \\ &\}. \end{aligned} \tag{5.3}$$

From (5.3), we can see that range of the values for initial frequency β changes with the value of chirp rate α. The matrix $\mathbf{\Psi}_F$ is designed by uniformly sampling the 2D parameter space $\mathbf{\Omega}$. Let I denote the total number of chirp rate values in the

discrete dictionary $\mathbf{\Psi}_F$. For the i^{th} chirp rate value in the dictionary, which we denote as $\tilde{\alpha}_i$, let $\tilde{\beta}_{i,j}$ denote the corresponding possible values for initial frequency, where $j = 1, 2, ..., J_i$. Note that the "˜" refers to "dictionary values". The full chirp dictionary $\mathbf{\Psi}_F$ is defined as [68], [13]:

$$\begin{aligned}
&\mathbf{\Psi}_F = [\mathbf{\Psi}_1, \mathbf{\Psi}_2, ..., \mathbf{\Psi}_I] \\
&\mathbf{\Psi}_i = [\boldsymbol{\psi}_{i,1}, \boldsymbol{\psi}_{i,2}, ..., \boldsymbol{\psi}_{i,J_i}] \\
&\boldsymbol{\psi}_{i,j}|_n = \exp\left(j2\pi(\tilde{\alpha}_i \frac{n^2}{2F_s^2} + \tilde{\beta}_{i,j}\frac{n}{F_s})\right) \\
&i = 1, ..., I,\ j = 1, ..., J_i,\ n = 0, 1, ..., N_w - 1.
\end{aligned} \tag{5.4}$$

From (5.4), it can be seen that for each value of α_i, there are J_i columns in the dictionary corresponding to J_i values of initial frequency.

Assume that the number of chirp rate values is $I = F_s + 1$ (including '0' value) and initial frequency resolution is $\Delta\tilde{\beta} = 1$. Now (5.3) shows the $\tilde{\alpha}$ ranges inside $[-F_{\max}F_s/N_w : F_{\max}F_s/N_w]$, or $[-F_{\max}/T_w : F_{\max}/T_w]$ ($N_w = T_sF_s$). We uniformly divide this range into $I = F_s + 1$ discrete values of $\tilde{\alpha}$. The chirp rate resolution becomes $\Delta\tilde{\alpha} = 1/T_w$. The discrete parameter space used in the chirp dictionary becomes:

$$\begin{aligned}
\mathbf{\Omega} = \{&(\tilde{\alpha}, \tilde{\beta}) \quad \text{such that} \\
&-\frac{F_{\max}}{T_w} \le \tilde{\alpha} \le -\frac{F_{\max}}{T_w} + F_s\Delta\tilde{\alpha} \\
&-F_{\max} \le \tilde{\beta} \le \lfloor F_{\max} - |\alpha|T_w \rfloor \qquad \text{if} \qquad \alpha \ge 0 \\
&\lfloor -F_{\max} + |\alpha|T_w \rfloor \le \tilde{\beta} \le F_{\max} \qquad \text{if} \qquad \alpha < 0 \\
&\}.
\end{aligned} \tag{5.5}$$

As $\Delta\tilde{\beta} = 1$, for each value of $\tilde{\alpha}$, there are $\lfloor F_s - |\tilde{\alpha}|T_w + 1 \rfloor$ values of $\tilde{\beta}$. The dimension

of the chirp dictionary Q is expressed as:

$$\begin{aligned}
Q &= \lfloor \sum_{\tilde{\alpha}=-\frac{F_{\max}}{T_w}}^{-\frac{F_{\max}}{T_w}+F_s\Delta\tilde{\alpha}} (F_s - T_w|\tilde{\alpha}| + 1)\rfloor \\
&= \lfloor F_s(F_s+1) + (F_s+1) - T_w \sum_{a=0}^{F_s} \left| -\frac{F_{\max}}{T_w} + a\Delta\tilde{\alpha} \right| \rfloor \\
&= \lfloor F_s(F_s+1) + (F_s+1) - T_w \sum_{a=1}^{F_{\max}} |a\Delta\tilde{\alpha}| \rfloor \\
&= \lfloor F_s(F_s+1) + (F_s+1) - 2T_w \sum_{a=1}^{(F_s/2)} a\Delta\tilde{\alpha} \rfloor \\
&= \lfloor F_s(F_s+1) + (F_s+1) - 2\frac{F_{\max}(F_{\max}+1)}{2} \rfloor \\
&= \lfloor (F_s+1)^2 - \frac{F_s}{2}(\frac{F_s}{2}+1) \rfloor \\
&= \lfloor \frac{3}{4}(F_s+1)^2 + \frac{1}{4} \rfloor \\
&= \lfloor \frac{3}{4}I^2 + \frac{1}{4} \rfloor,
\end{aligned} \tag{5.6}$$

where $\lfloor . \rfloor$ is the "floor operator". As the dictionary dimension is $\lfloor \frac{3}{4}I^2 + \frac{1}{4} \rfloor$, in each segment, $E\lfloor \frac{3}{4}I^2 + \frac{1}{4} \rfloor$ projections are implemented. Furthermore, as sliding windows are used, the chirp method faces a large computational burden, $O\left(E(\frac{3}{4}I^2 + \frac{1}{4})\right)$. With the same manner of operation, the sinusoidal dictionary method only requires $O(EI)$ projections (E times of DFT) in each windowed signal, which gives it a big advantage over its chirp counterpart when large amounts of data are concerned.

5.3 Simplify The Full Chirp Dictionary by Estimating The Chirp-Rate in The IAF Domain

5.3.1 Background

This approach reduces the dictionary dimension by estimating the chirp-rate of a chirp component in the signal segment first. The chirp-rate is approximated in IAF domain. Then, the dictionary is built with only the estimated value of the chirp-rate

and the corresponding values of initial frequency. Compressive sensing techniques are used to find the chirp components in the windowed signal. Therefore, this part will give an background about compressive sensing, some chirp-rate estimation methods and the IAF of a windowed signal.

Compressive sensing and sparse reconstruction

For a brief review of the principle underlying CS, consider a signal vector $\mathbf{x} \in \mathbb{C}^Q$, which is E sparse ($E \ll Q$). Only $M (E < M \ll Q)$ linear measurements of $\mathbf{x}$ are recorded, which means:

$$\mathbf{y} = \mathbf{\Psi}\mathbf{x} + \boldsymbol{\epsilon}, \tag{5.7}$$

where $\mathbf{y} \in \mathbb{C}^M$ is measurement vector, $\mathbf{\Psi} \in \mathbb{C}^{M \mathrm{x} Q}$ is a known dictionary matrix , $\boldsymbol{\epsilon}$ is Gaussian noise vector.

Since $\mathbf{\Psi}$ is a 'fat' matrix, solving $\mathbf{x}$ given $\mathbf{y}$ is ill posed with an infinite number of solutions. However, because $\mathbf{x}$ is sparse, it can be uniquely recovered with a high probability by l_1 convex optimization or iterative greedy algorithms such as OMP provided that the matrix $\mathbf{\Psi}$ satisfies certain restricted isometry properties and the dimension M of measurement vector $\mathbf{y}$ is at least of the order of $E \log(Q/E)$ [69]. The l_1 optimization is as follows [45].

$$\begin{aligned} &\min_{\mathbf{x}} \quad \|\mathbf{x}\|_1 \\ &\text{subject to } \|\mathbf{y} - \mathbf{\Psi}\mathbf{x}\|_2^2 \leq \epsilon_0, \end{aligned} \tag{5.8}$$

where $\|\|_1$ denotes l_1 norm, $\|\|_2$ denotes Euclidean l_2 norm, ϵ_0 is a regularizer.

Chirp-rate estimation methods

One of the accurate chirp-rate estimation is based on the fractional Fourier transform [70, 71, 72]. In the fractional domain, there is an optimum transform angle associated with the chirp rate, which concentrates the energy and keeps the signal compact. At this angle, the signal is transformed to an impulse with a maximum magnitude. Therefore, searching for the maximum magnitude in the FrFT domains allows us to estimate the chirp rate. An improved technique to obtain fast and accurate in noisy environment is proposed in [73], which utilizes the fractional Fourier transform and the golden section search algorithm (GSS). Using the GSS accelerates the search

process by computing the fractional Fourier transform at only certain limited angles. Error tolerance of the algorithm determines accuracy of the result.

Another method to estimate the local chirp-rate is based on STFT of a signal segment using a differentiable analysis window [74]. It differentiates the STFT over time, and rewrites the results in two parts. The chirp-rate then can be expressed by these two parts. It is called one partial derivative estimation. To get more accurate result, it also proposes two partial derivative estimator.

In this chapter, we introduce a method which estimates the chirp-rate of a chirp component in a windowed signal in IAF domain.

IAF of a windowed signal

Consider a discrete signal $s(n)$, $n = 0, 1, ..., N-1$. Let $N = \lfloor T/T_s \rfloor$, with T and T_s being total observation time and sampling period, respectively. The IAF of $s(n)$ is defined as:

$$R_{ss}(b, n) = s(n + b)s^*(n - b), \tag{5.9}$$

where b is the time lag, $b = \lfloor -N/2 \rfloor ... \lfloor N/2 - 1 \rfloor$. As $0 \leq n \pm b \leq N - 1$, for a certain value of time lag, the range of value for time n is:

$$|b| \leq n \leq N - 1 - |b|. \tag{5.10}$$

With a signal consisting of two parallel chirps and sampling frequency $F_s = 128$, its IAF is shown in Fig. 5.2.

IAF of a windowed signal of length N_w

The IAF of a windowed signal can be obtained by calculating (5.9) with $N = N_w$. If we need the IAF of every sliding window, this is quite time-consuming. The much simpler way to obtain the IAF of all windowed signals is through extraction from the IAF of the full data. Call $s_m(n)$ the m^{th} windowed signal of $s(n)$. The range of n is $n_1 \leq n|_{s_m} \leq n_1 + N_w - 1$, where n_1 and N_w are the starting point of the signal segment and the window length, respectively. Parameters n_1, N_w and b need to satisfy the following conditions:

$$\begin{cases} 0 \leq n_1 \leq N - 1 - N_w \\ 1 \leq N_w \leq N \\ \lfloor -N_w/2 \rfloor \leq b \leq \lfloor N_w/2 - 1 \rfloor. \end{cases} \tag{5.11}$$

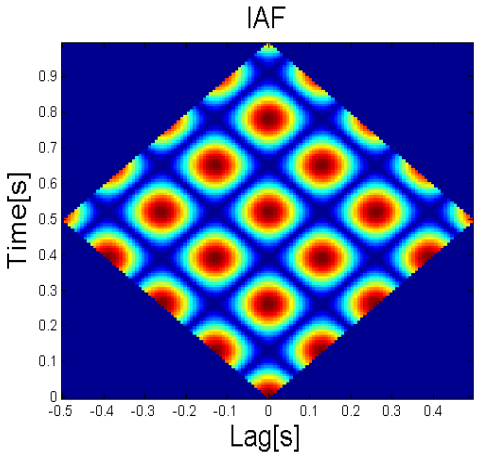


Figure 5.2: IAF of two parallel chirps.

As $n_1 \leq n \pm b \leq n_1 + N_w - 1$, a possible value for time n is:

$$|b| + n_1 \leq n \leq n_1 + N_w - 1 - |b|. \tag{5.12}$$

(5.11) and (5.12) show that the range of time lag b and time n when calculating the IAF of a segment reside in those when calculating the IAF of the whole signal. Thus, IAF $R_{s_m s_m}(b, n)$ of a windowed signal $s_m(n)$ is simply a part of the IAF of the whole signal $R_{ss}(b, n)$. This is illustrated in Fig. 5.3 with the signal used in Fig. 5.2.

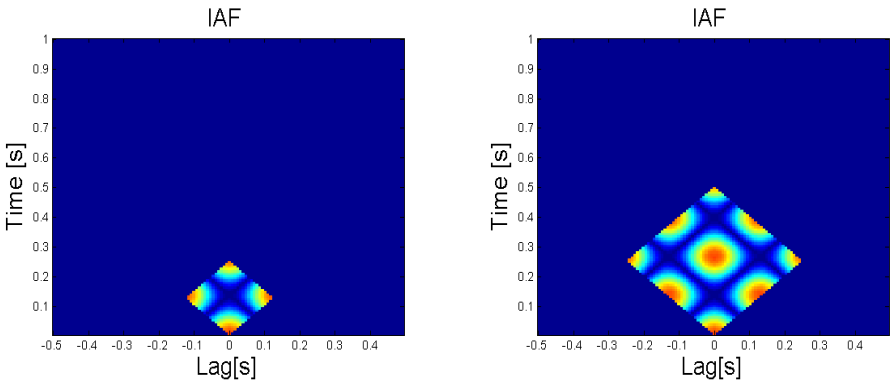
5.3.2 Simplified chirp dictionary approach

Chirp rate estimation

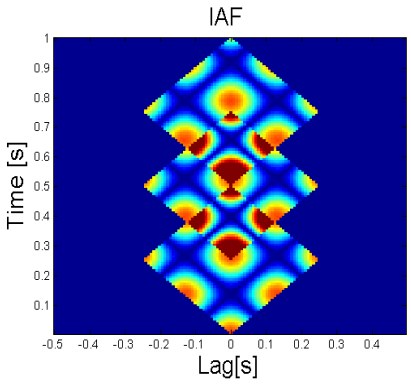
Consider an arbitrary continuous signal $s_c(t)$, which consists of $E \geq 1$ components:

$$s_c(t) = \sum_{e=1}^{E} A_e(t) \exp(j\phi_e(t)) + v_c(t), \tag{5.13}$$

where $0 \leq t \leq T$, $A_e(t)$ and $\phi_e(t)$ are the time-varying amplitude and phase of the e^{th} component, and $v_c(t)$ is white Gaussian noise. Sampling the continuous signal at the Nyquist rate $F_s = 2F_{\max}$, where $F_{\max}$ is the maximum frequency of the signal



(a) (b)



(c)

Figure 5.3: IAF of: (a) A windowed signal $N_w = 32, n_1 = 0$; (b) A windowed signal $N_w = 64, n_1 = 0$; (c) Three windowed signals $(N_w = 64, n_1 = 0)$, $(N_w = 64, n_1 = 32)$, $(N_w = 64, n_1 = 64)$ of the signal used in Fig. 5.2.

$s_c(t)$, we have:

$$s(n) = \sum_{e=1}^{E} A_e(nT_s)\exp(j\phi_e(nT_s)) + v(nT_s), \tag{5.14}$$

where $n = 0, 1, ..., \lfloor T/T_s \rfloor$, and $T_s = 1/F_s$.

Similar to the full chirp atom approach, in this method, observations inside a short-time window are also approximated by the sum of piece-wise chirps. So breaking $s(n)$ into N_w-length blocks $\{s_m(n)\}_{n=0}^{N_w-1}$, the m^{th} block is calculated as:

$$s_m(n - \varsigma(m-1)) = s(n)h(n - \varsigma(m-1)), \tag{5.15}$$

where $n = \varsigma(m-1), \varsigma(m-1)+1, ..., \varsigma(m-1)+N_w-1$, $1 \leq \varsigma \leq N_w$ is the shift between two consecutive windows, $m = 1, 2, ...$ is the window index, and $h(n)$ is a rectangular window which is non-zero only for $0 \leq n \leq N_w - 1$. Then the chirp-approximated m^{th} signal segment of $s(n)$ is written as:

$$\begin{aligned} s_m(n) &\approx \sum_{e=1}^{E} C_{e,m} \exp\left\{ j2\pi \left[\alpha_{e,m}\frac{n^2}{2F_s^2} + \beta_{e,m}\frac{n}{F_s} \right] \right\} \\ &+ v_m(n) = \sum_{e=1}^{E} s_{e,m} + v_m(n), \end{aligned} \tag{5.16}$$

where $0 \leq n \leq N_w - 1$, $C_{e,m}$, $\alpha_{e,m}$ and $\beta_{e,m}$ are respectively the complex amplitude, the chirp rate, and the initial frequency of the e^{th} chirp over the m^{th} window. Now $s_{e,m}$ is the chirp with parameters specified by $C_{e,m}$, $\alpha_{e,m}$ and $\beta_{e,m}$. The instantaneous autocorrelation function (IAF) of $s_m(n)$ is expressed as:

$$\begin{aligned} R_{s_m s_m}(b, n) &= s_m(n+b)s_m^*(n-b) \\ &= \sum_{e=1}^{E} s_{e,m}(n+b)s_{e,m}^*(n-b) \\ &+ \sum_{\substack{i,j=1 \\ i \neq j}}^{E} s_{i,m}(n+b)s_{j,m}^*(n-b) \\ &= \sum_{e=1}^{E} AT_{e,m}(b,m) + \sum_{g=1}^{E(E-1)} CT_{g,m}(b,m), \end{aligned} \tag{5.17}$$

where b is time lag, $AT_{e,m}$ and $CT_{g,m}$ contain auto-terms and cross-terms, respectively, and are expressed as:

$$\begin{aligned} AT_{e,m}(b,n) &= \exp\left(j2\pi\frac{2\alpha_{e,m}b}{F_s}\frac{n}{F_s}\right)\exp\left(j2\pi\frac{2\beta_{e,m}b}{F_s}\right) \\ CT_{g,m}(b,n) &= \exp\left(j2\pi(\alpha_{i,m}-\alpha_{j,m})\frac{n^2}{2F_s^2}\right) \\ &\exp\left(j2\pi\frac{(\alpha_{i,m}+\alpha_{j,m})b}{F_s}\frac{n}{F_s}\right)\exp\left(j2\pi(\beta_{i,m}-\beta_{j,m})\frac{n}{F_s}\right) \\ &\exp\left\{j2\pi\left[\frac{(\alpha_{i,m}-\alpha_{j,m})b^2}{2F_s^2}+\frac{(\beta_{i,m}+\beta_{j,m})b}{2F_s^2}\right]\right\}, \end{aligned} \tag{5.18}$$

where $i,j\in[1,E]$, $i\neq j$, $g\in[1,E(E-1)]$. Let $AT_{e,m}(k)|_{b=b_1}$, and $CT_{e,m}(k)_{b=b_1}$ be the DTFT of $AT_{e,m}(b,n)$ and $CT_{g,m}(b,n)$ at $b=b_1$, then:

$$\begin{aligned} |AT_{e,m}(k)|_{b=b_1}| &= \delta\left(k-2\alpha_{e,m}\frac{b_1}{F_s}\right) \\ |CT_{g,m}(k)|_{b=b_1}| &= W(k)*\delta\left(k-\frac{b_1(\alpha_{i,m}+\alpha_{j,m})}{F_s}\right)* \\ &\delta\left(k-(\beta_{i,m}-\beta_{j,m})\right), \end{aligned} \tag{5.19}$$

where $W(k)$ is the DTFT of $\exp\left(j2\pi(\alpha_{i,m}-\alpha_{j,m})\frac{n^2}{(2F_s^2)}\right)$. Now (5.19) shows that the spectral representation of the auto-terms are delta functions whose locations are determined by the chirp rates. So if $k_{AT_{e,m}}$ corresponds to the frequency location of the spectrum of the auto-terms then the chirp rates are approximated by:

$$\hat{\alpha}_{e,m} = \frac{k_{AT_{e,m}}F_s}{2b_1}. \tag{5.20}$$

In addition, (5.19) shows that any cross-term is also a chirp located between two auto-term components. Notice that we should choose a small value of $b_1(b_1\neq 0)$ to get more data in the IAF domain.

Simplified chirp dictionary algorithm

In the vector form, the signal over the m^{th} window in (5.16) can be expressed as:

$$\mathbf{S}_m = \mathbf{\Psi}_c\mathbf{X}_m + \mathbf{V}_m, \tag{5.21}$$

where $\mathbf{S}_m = [s_m(0), ..., s_m(N_w - 1)]^T$, $\mathbf{V}_m = [v_m(0), ..., v_m(N_w - 1)]^T$ and $\mathbf{X}_m$ has E non-zero components. The compact dictionary matrix, $\mathbf{\Psi}_c$, is designed for each signal component inside the windowed data. The chirp rate value in $\mathbf{\Psi}_c$ is estimated by the algorithm in part 5.3.2, and is denoted as $\hat{\alpha}$. Let $\tilde{\beta}_j$ denote the corresponding possible values for initial frequency, where $j = 1, 2, ..., J$. The compact chirp dictionary $\mathbf{\Psi}_c$ is defined as:

$$\begin{aligned} &\mathbf{\Psi}_c = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, ..., \boldsymbol{\psi}_J] \\ &\boldsymbol{\psi}_j|_n = \exp\left(j2\pi(\hat{\alpha}\frac{n^2}{2F_s^2} + \tilde{\beta}_j\frac{n}{F_s})\right) \\ &j = 1, ..., J;\ n = 0, ..., N_w - 1. \end{aligned} \tag{5.22}$$

As $\hat{\alpha}$ has only one value, the compact chirp dictionary only has J columns, where from $\mathbf{\Omega}$ in (5.2), $J = \lfloor F_s - |\hat{\alpha}|T_w + 1 \rfloor$. Assume that $J = F_s + 1$ ($|\hat{\alpha}| = 0$), then the dimension of the compact chirp atom now becomes:

$$d_{\mathbf{\Psi}_c} = J = F_s + 1. \tag{5.23}$$

As the preprocessing takes E times DFT or EF_s projection, we can say the number of calculations when using the simplified chirp dictionary is about $O(E(2F_s + 1))$. Since $E < N_w \ll J$, then $\mathbf{X}_m$ is highly sparse and solving for $\mathbf{X}_m$ in equation (5.21) becomes a sparse recovery (or CS) problem. The algorithm of the simplified chirp dictionary used in this paper is based on Orthogonal Matching Pursuit and has following steps:

INPUT:

- Signal $s(n)$ of length N.
- Signal vector $\mathbf{S} = [s(0), ..., s(N - 1)]^T$.
- Windowed signal vector $\mathbf{S}_m = \mathbf{S}((m-1)\varsigma + n)$, $0 \leq n \leq N_w - 1$. Initialize $m = 1$.
- Lag value $b = b_1$ (Choose b_1 close to 0)

OUTPUT:

- Matrix of selected chirp $\mathbf{\Phi}$.

PROCEDURE:

1. Initialize the residual $\mathbf{r_0} = \mathbf{S}_m$, matrix of selected chirps $\mathbf{\Phi}_i = \emptyset$, and the iteration counter $i = 1$.
2. Calculate IAF at $b = b_1$ ($R_{r_i r_i}(b, n)|_{b=b_1}$).
3. Calculate DTFT of $R_{r_i r_i}$. Estimate the chirp rate by (5.20) and build the compact chirp dictionary $\mathbf{\Psi}_c$ by (5.22).
4. Find the index λ_i, $\lambda_i = \arg\max_{j=1,\ldots,J} | < \boldsymbol{r}_{i-1}, \boldsymbol{\psi}_j > |$.
5. Store the selected chirp $\boldsymbol{\psi}_{\lambda_i}$, $\mathbf{\Phi}_i = [\mathbf{\Phi}_{i-1}\, \boldsymbol{\psi}_{\lambda_i}]$.
6. Solve a least square problem to find the residue after subtracting the chirp

$$\mathbf{x}_i = \arg\min_{\mathbf{x}} \|\mathbf{\Phi}_i \mathbf{x} - \mathbf{S}_m\|_2$$
$$\mathbf{r}_i = \mathbf{S}_m - \mathbf{\Phi}_i \mathbf{x}_i.$$

7. Increment i, and return to step 2 if $i < E$ or $\|\mathbf{r}_i\|_2 > 0.05\|\mathbf{S}_m\|_2$. The magnitude of the selected chirps is stored in $\mathbf{x}_i$. If $i = E$ or$\|\mathbf{r}_i\|_2 \leq 0.05\|\mathbf{S}_m\|_2$, move to the next windowed signal, increment m, and return to step 1.

There is another way to estimate the initial frequencies after the chirp rates are verified. According to (5.18), the initial frequencies can be approximated by the magnitude of the auto-terms' frequencies. However, this magnitude can be easily affected by noise, and thus its results are unreliable. The drawback of the simplified chirp dictionary method is that it does not perform well if too much data (over 50% of observations) is absent. This is because the missing samples in the IAF at any time lag can be double the number of missing samples for $s(n)$ in the time domain.

5.3.3 Restricted isometry properties (RIP) analysis of the chirp dictionary

In this section, we examine the RIP associated with chirp and simplified the chirp dictionary. Let $\mathbf{\Psi} \in C^{N_w \mathrm{x} Q}$ be the chirp dictionary in general. Similar to the work in [51], we consider the boundaries of the eigenvalues of Grammian matrix. We show

that these bounds compete with those of the Gaussian random dictionary, and as such lead to the same conditions on sparsity and compressed observations.

Consider the under-determined problem mentioned in (5.21) with infinite solutions. Denote the set of column indices by $\mathbb{Q} = \{1, ..., Q\}$. It has been shown in [45] that if $\mathbf{\Psi}$ satisfies a certain RIP, then $\mathbf{x}$ can be recovered uniquely using an l_1 minimization provided that the number of measurements is $O(E \log(Q/E))$. Given any set Γ of column indices ($\Gamma \subset \mathbb{Q}, \text{card}(\Gamma) \leq E$), we denote $\mathbf{\Psi}_\Gamma$ as the matrix composed of these columns. A matrix $\mathbf{\Psi}$ is said to satisfy the RIP of order E if there exists a $\delta_E \in (0, 1)$ such that:

$$(1 - \delta_E)\|\mathbf{x}_\mathbf{\Gamma}\|_2^2 \leq \|\mathbf{\Psi}_\Gamma \mathbf{x}_\Gamma\|_2^2 \leq (1 + \delta_E)\|\mathbf{x}_\mathbf{\Gamma}\|_2^2, \tag{5.24}$$

where $\mathbf{x}_\Gamma$ is the vector obtained by retaining only the entries in $\mathbf{x}$ corresponding to the columns' indices in Γ. Because $\mathbf{\Psi}_\Gamma^H \mathbf{\Psi}_\Gamma$ is a Hermitian matrix, (5.24) is rewritten as:

$$1 - \delta_E \leq \lambda_{min} \leq \lambda_{max} \leq 1 + \delta_E \tag{5.25}$$

In order for the chirp dictionary to satisfy RIP, the Grammian matrix $\mathbf{\Psi}_\Gamma^H \mathbf{\Psi}_\Gamma$ must have all of its eigenvalues in $(0, 2)$. Since the chirp dictionary $\mathbf{\Psi}$ is deterministic, the above condition requires a check of all $\binom{Q}{E}$ possible set Γ, which is a formidable problem. According to [52], a Gaussian random matrix $\mathbf{G} \in C^{N_w \text{x} Q}$ whose column's entries have zero mean, and variance $1/N_w$ can satisfy $\delta_E < 1$ with a minimum number of measurements equal to $O(E \log(Q/E))$. Therefore, we will compare the bounds on the eigenvalue of $\mathbf{\Psi}_\Gamma^H \mathbf{\Psi}_\Gamma$ with $\mathbf{G}_\Gamma^H \mathbf{G}_\Gamma$.

In the simulation, we consider the simplified chirp dictionary with $N_w = 50, Q = 129$, and random Gaussian dictionaries with corresponding sizes. For each value of $E (0 \leq E \leq 20)$, 100000 random realisations of subset Γ are used to estimate the eigenvalue statistic. Figure 5.4 shows the bounds (sample mean + 3 STD for maximum eigenvalues and sample mean - 3 STD for minimum eigenvalues).

Figure 5.4 shows that the bounds for the simplified chirp dictionary is closer to 1. This implies that $\mathbf{\Psi}$ can satisfy condition $\delta_K < 1$ with high probability if the minimum observation is $O(E \log(Q/E))$.

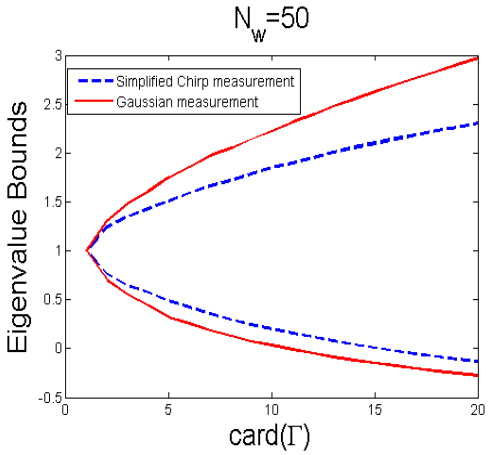


Figure 5.4: Eigenvalue bounds of $\mathbf{\Psi}_\Gamma^H \mathbf{\Psi}_\Gamma$ and $\mathbf{G}_\Gamma^H \mathbf{G}_\Gamma$.

5.3.4 Simulation results

For illustration purposes, we use 3 examples with four different TFD methods including the WVD, the sinusoidal dictionary, the full chirp dictionary and the simplified chirp dictionary. The WVD represents the quadratic TFD, which is vulnerable to missing samples and cross-terms and so it is unable to deliver an accurate TF estimation under compressed data. The WVD is simulated in order to compare it with the CS-related methods. As discussed above, the dictionary Ψ satisfies the RIP condition if the minimum observation is $E \log(Q/E)$. If $F_s = 256$Hz, and we have $E = 3$ components in each windowed signal, then we need at least 30 measurements for a reliable recovery. If the average window length is around 70, then 50% of data can obtain good TF estimation by CS technique. However, the chirp rate is estimated in the IAF domain and the missing samples in the IAF domain can double the missing data in the time domain (see 3.4.2). If 50% data in the time domain is removed, 75% of data is missing in a time of IAF, which gives unreliable estimation of the chirp rate. Thus in this simulation part, only 40% of data is removed. Notice that the RIP condition discussed in 5.3.3 is obtained when there is no noise. If noise is present, we need more data to recover well the TF signature. Here we cut 40% of the data (42 samples per window if the window length is 70), and add some noise with $SNR = 20$dB. If the signal contains more noise (smaller SNR), the estimation would not be reliable.

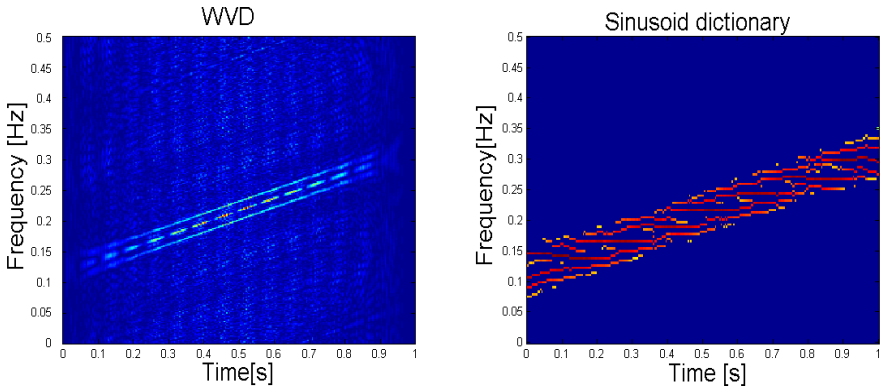
The signals in three examples are firstly sampled at the Nyquist rate, and then some samples are randomly removed. The sampling frequency is $F_s = 256Hz$, the total signal length is $N = 256$, and 60% of the data is used to obtain the time frequency signature of the signal. When computing the TF representation, a rectangular window is used for the chirp dictionary, and WVD, whereas the sinusoidal approach utilizes the Hanning window. A parameter of concentration level ζ is used to assess the accuracy of the resulting TF representations. So ζ is the ratio of the sum of pixel magnitudes along the actual instantaneous frequency, with respect to the rest of the TF values. So, the higher ζ, the better is the TF estimation.

In the first example, the signal consists of two closely-alligned chirps and is expressed as:

$$\begin{aligned} s(n) &= \exp\left\{j2\pi[(0.1F_s)\frac{n}{F_s} + (0.3F_s)\frac{n^2}{2F_s^2}]\right\} \\ &+ \exp\left\{j2\pi[(0.13F_s)\frac{n}{F_s} + (0.33F_s)\frac{n^2}{2F_s^2}]\right\} + v(n), \end{aligned} \tag{5.26}$$

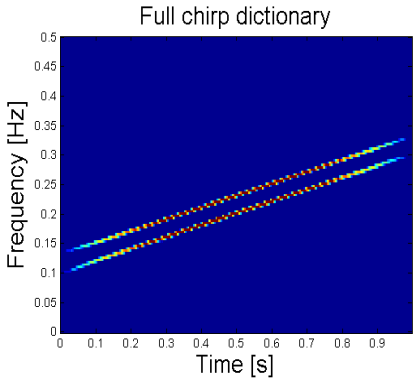
where $n = 0, 1, ..., N-1$. To capture enough data to resolve the two close parallel chirps, the window size is set to a large value, $N_w = 100$. The WVD suffers cross-terms and artifacts caused by the missing samples. The concentration level is very low with $\zeta = 0.3024$. These issues are mitigated when CS related methods are used. However, Fig. 5.11(b) shows the failure of local reconstruction of the sinusoidal method due to lack of sparsity when employing a long window. The concentration level of the resulted TFR is $\zeta = 2.19$. The sparsity, on the other hand, when chirp methods are in use, only depends on the number of piece-wise chirps inside the considered segment. Thus the two chirp dictionary methods are less sensitive to this issue, and the signal is clearly resolved as shown in Fig. 5.5(c) and Fig. 5.5(d). Thus, the TFRs obtained by the two methods have similar concentration level $\zeta = \infty$.

In the second example, we use a three-component FM signal, which is expressed

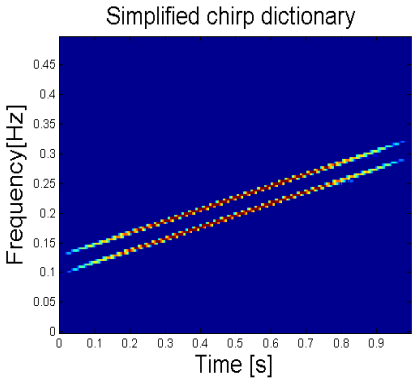


(a)

(b)



(c)



(d)

Figure 5.5: TF (frequency normalized) signature for $s(n)$ in (5.26) with 40% data missing: (a) WVD; (b) Sinusoidal dictionary; (c) Chirp dictionary; (d) Simplified chirp dictionary.

as:

$$
\begin{aligned}
s(n) = &\exp\left\{j(0.1F_s)\cos(2\pi\frac{n}{F_s}+\pi) + j2\pi(0.2F_s)\frac{n}{F_s}\right\} \\
&+ \exp\left\{j(0.1F_s)\cos(2\pi\frac{n}{F_s}+\pi) + j2\pi(0.3F_s)\frac{n}{F_s}\right\} \\
&+ \exp\left\{j2\pi[(0.1F_s)\frac{n}{F_s} + (0.3F_s)\frac{n^2}{2F_s^2}]\right\} + v(n),
\end{aligned}
\tag{5.27}
$$

where $n = 0, 1, ..., N-1$. The window length is set to $N_w = 70$. The results are given in Fig. 5.6. It is evident in Fig. 5.6(a) that cross-terms and noise-like artifacts clutter the signal component and hide the pertinent signal structure when the WVD is employed. The concentration level is $\zeta = 0.2$. The sinusoidal dictionary approach reveals inaccuracy in the TF signature estimation since besides insufficient sparsity, it is vulnerable to the picket fence effect [13], resulting in frequency content at false locations. Its concentration level is $\zeta = 1.52$. The chirp dictionary approach can address this failure and the instantaneous frequency laws are resolved as seen in Fig. 5.6(c) and Fig. 5.6(d). The simplified chirp dictionary has some inaccurate approximation due to limited samples in the instantaneous autocorrelation domain, but the result is acceptable. The concentration level of the TFR is $\zeta = 3$, which is lower than the one obtained by the full chirp dictionary $\zeta = 20$. Compared with its sinusoidal counterpart, it gives a better performance but with a similar calculation effort. In the third example, we use data from human gait radar returns obtained at the Radar Imaging Lab of the Center for Advanced Communications at Villanova University, USA. The data is first uniformly sampled at the Nyquist rate with $F_s = 1000$ Hz, and then thinned by randomly removing 40% of samples. The sparsity level is assumed to be $E = 30$. The window length is $N_w = 128$, and we only use 128 frequency components to display the TF signature in order to zoom in on the instantaneous frequencies, and so partly mitigate drawbacks of the sinusoidal dictionary method. The results in Fig. 5.7 show that the simplified chirp dictionary approach can describe Micro-Doppler TF presentations of the torso and limbs under compressed observations.

5.4 Simplify The Full Chirp Dictionary using The Fractional Fourier Transform (FRFT)

5.4.1 Chirp rate and initial frequency estimation of chirps using FRFT

Consider a discrete single chirp of length $T = 1$ (second) expressed as:

$$s(n) = \exp\left[j2\pi\left(\alpha\frac{n^2}{2F_s^2} + \beta\frac{n}{F_s}\right)\right], \tag{5.28}$$

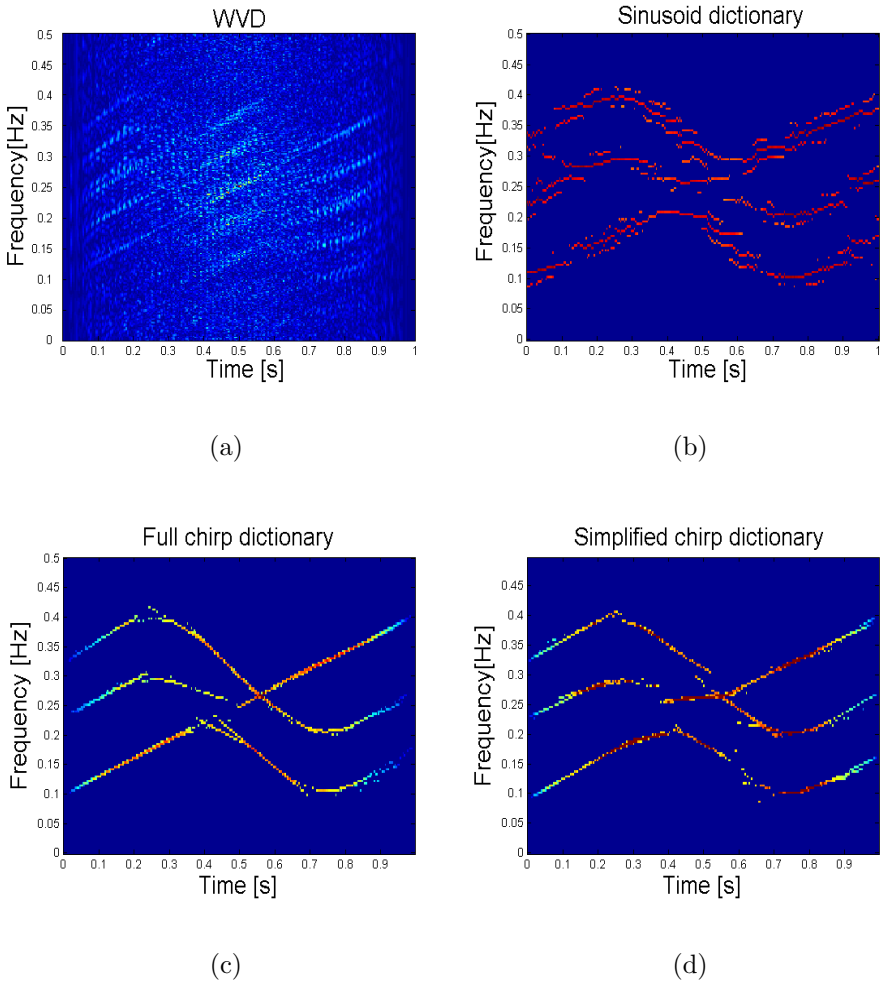


Figure 5.6: TF signature (frequency normalized) for $s(n)$ in (5.27) when 40 % samples are missing: (a) WVD; (b) Sinusoidal dictionary; (c) Chirp dictionary; (d) Simplified chirp dictionary.

where F_s is the sampling frequency, α and β are values of the chirp rate and the initial frequency, $n = 0, 1, ..., \lfloor T/T_s \rfloor$, and $T_s = 1/F_s$. We aim to use the FRFT for the α and β estimations. The principle is that we tune the FRFT angle ϕ. When $\phi = \phi_{\text{opt}}$, the fractional axis x is matched to the chirp rate of the signal (see Fig. 5.8), or the chirp becomes a sinusoid in the new plane the (x, y). Denote d as the sinusoidal frequency in (x, y) domain. The magnitude response (i.e. the

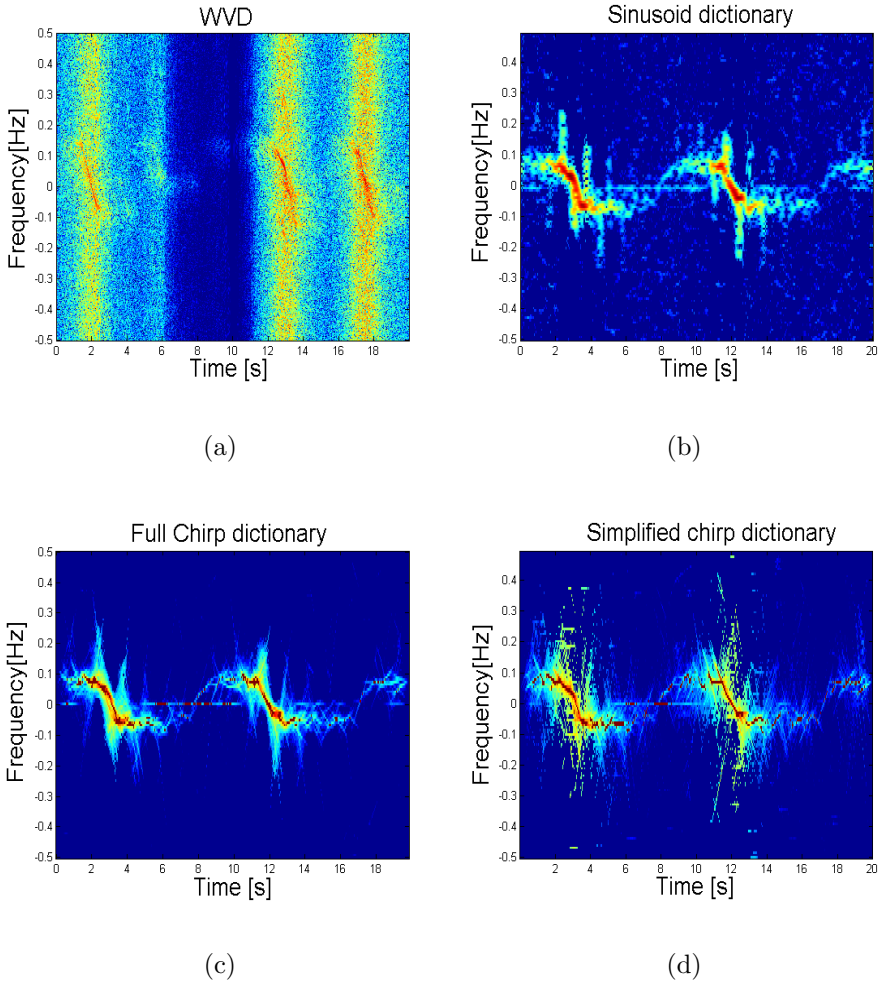


Figure 5.7: TF (frequency normalized) signature of human gait radar return with 40% data missing: (a) WVD; (b)Sinusoidal dictionary; (b) Chirp dictionary; (c) Simplified chirp dictionary.

absolute value of the FT) of $(\mathbb{F}^{\phi_{\text{opt}}}s)(x)$ reaches its maximum at d. Thus, we can get the estimation of d (see Fig. 5.9 (a)). The relationship between d and the initial frequency β has been discussed in the subsection 4.3.1. To remind, the relationship

is portrayed in Fig. 5.8 and expressed as:

$$\begin{aligned} d &= \sqrt{N_w}\cos(\phi)\left(\frac{\beta}{F_s}+\frac{\tan\phi}{2}\right) \\ \leftrightarrow \beta &= F_s\left(\frac{d}{\sqrt{N_w}\cos(\phi)}-\frac{\tan\phi}{2}\right), \end{aligned} \tag{5.29}$$

where N_w is the number of samples in a window. So, we can say that the FRFT can

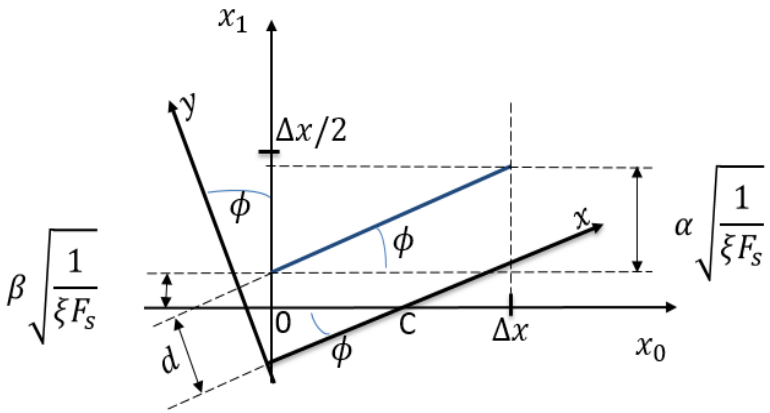
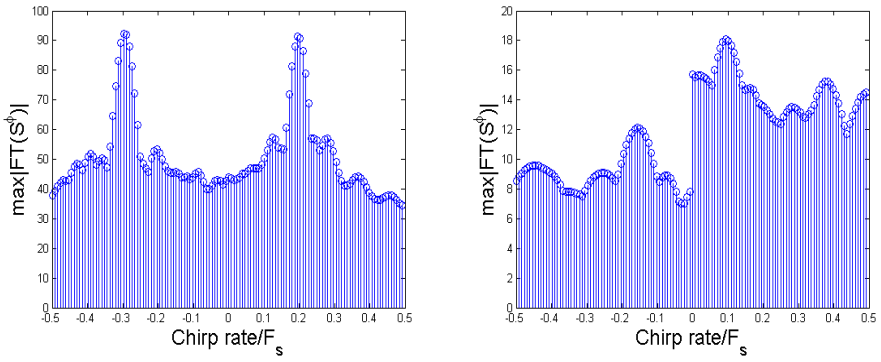


Figure 5.8: Geometric schematic for calculating sinusoidal frequency in the fractional domain.

help us to estimate the chirp rate and the initial frequency of a chirp, or a signal composed of chirps. As a windowed non-stationary signal can be approximated as a sum of different chirps [13, 68], the FRFT can be used to estimate the instantaneous frequency characteristic of the non-stationary signal segment. However, this method only works in the case that we have full data and the number of signal components is a known a-priori. When we only have limited observations, the method is unable to deliver accurate results because the magnitude responses do not always obtain a maximum when $\phi = \phi_{\text{opt}}$. Fig. 5.9(b) illustrates this situation for a signal composed of two chirps whose chirp rate values are $-0.3F_s$, and $0.2F_s$, $F_s = 128$. The maximum values of the magnitude response do not show the optimum FRFT angles. Thus, it is incapable of estimating the two chirp rate values.

So here we propose using the FRFT to build the chirp dictionary for sparse reconstruction instead. For each value of chirp rate, the FRFT gives a corresponding



(a) (b)

Figure 5.9: The maximum values of the magnitude response versus chirp rate values or the FRFT angle values: (a) Full length signal N=128; (b) Windowed signal of length $N_w = 64$, and randomly missing 50% of the data.

initial chirp value. Thus, the chirp dictionary dimension gets smaller, just the same as that of the sinusoidal dictionary. Moreover, the discrete FRFT algorithm proposed in [49] has a computational load of $O(N \log N)$ for a discrete-time signal of length N, which is at a similar complexity with the conventional FT. Therefore, the proposed method is more computationally efficient than the full chirp atom approach.

Simplified calculation technique for the DFT of the FRFT of a signal

The discrete FRFT calculation is presented in Appendix A. Although fast computation of the FRFT of a signal takes not much longer when compared to the Fourier transform. Still, it involves three steps: x2 Shannon interpolation, convolution and then decimation, which makes it more complex than the Fourier transform. In this part, we introduce a simple method for the DFT of the discrete FRFT calculation.

According to [50], the windowed Fourier transform of the FRFT of a signal corresponds to the short-time Fourier transform of the signal itself, with the window being the fractional Fourier transform of the initial one.

$$\mathrm{DFT}(\mathbb{F}^{\phi_i} s) = \mathrm{DFT}\left[s(n)(\mathbb{F}^{\phi} h(n))\right]. \tag{5.30}$$

where $h(n)$ is the window for the signal $s(n)$. Now, (5.30) can also be expressed in matrix format as follows (with $k_1, ..., k_N$ being values of discrete frequency k in the frequency plane):

$$\begin{bmatrix} \mathrm{DFT}(\mathbb{F}^{\phi_i} s)(v_1) \\ \mathrm{DFT}(\mathbb{F}^{\phi_i} s)(v_2) \\ \vdots \\ \mathrm{DFT}(\mathbb{F}^{\phi_i} s)(v_N) \end{bmatrix} =$$
$$\begin{bmatrix} \exp(-j2\pi k_1 \frac{n_1}{N})(\mathbb{F}^{\phi} h(n_1))(u_1) & \dots & \exp(-j2\pi k_1 \frac{n_N}{N})(\mathbb{F}^{\phi} h(n_N))(u_N) \\ \exp(-j2\pi k_2 \frac{n_1}{N})(\mathbb{F}^{\phi} h(n_1))(u_1) & \dots & \exp(-j2\pi k_2 \frac{n_N}{N})(\mathbb{F}^{\phi} h(n_N))(u_N) \\ \vdots & & \\ \exp(-j2\pi k_N \frac{n_1}{N})(\mathbb{F}^{\phi} h(n_1))(u_1) & \dots & \exp(-j2\pi k_N \frac{n_N}{N})(\mathbb{F}^{\phi} h(n_N))(u_N) \end{bmatrix} \begin{bmatrix} s(n_1) \\ s(n_2) \\ \vdots \\ s(n_N) \end{bmatrix}$$
$$= \mathbf{\Phi}^T \mathbf{s}, \tag{5.31}$$

where $\mathbf{\Phi}$ is called the FRFT matrix, and $\mathbf{s}$ is the signal vector, u, v are the discrete fractional variables which correspond to the continuous ones x, y. $v_1, ..., v_N$ are values of v and $u_1, ..., u_N$ are values of u. With a certain window and a value of the FRFT angle (or chirp rate), we can get the estimation of d and the chirp rate β by projecting the signal onto the matrix $\mathbf{\Phi}$ and finding the frequency corresponding to the maximum value.

5.4.2 Sparse reconstruction of non-stationary time frequency signature based on the FRFT

Consider an arbitrary continuous-time, non-stationary signal $s_c(t)$, which consists of E components:

$$s_c(t) = \sum_{e=1}^{E} A_e(t) \exp\left(j\varphi_e(t)\right) + v_c(t), \quad 0 \le t < T, \tag{5.32}$$

where $A_e(t)$ and $\varphi_e(t)$ are the time-varying positive amplitude and phase of the e^{th} component, $v_c(t)$ is an additive white noise, and T is the total observation interval. The continuous-time instantaneous frequency (IF) of the e^{th} component is defined as:

$$F_e(t) = \frac{1}{2\pi} \frac{d\varphi_e(t)}{dt}. \tag{5.33}$$

We assume that it is known a-priori that the absolute IFs do not exceed $F_{\max}$ i.e. $|F_e(t)| \leq F_{\max}$, where $F_{\max}$ is the maximum frequency of the signal $s_c(t)$. Sampling $s_c(t)$ at the Nyquist rate $F_s(F_s = 2F_{\max})$, then we have:

$$s(n) = \sum_{e=1}^{E} A_e(nT_s) \exp(j\varphi_e nT_s) + v(n), \tag{5.34}$$

where $n = 0, 1, ..., \lfloor T/T_s \rfloor$, and $T_s = 1/F_s$.

Similar to the chirp dictionary method, this approach also approximates the windowed signal by the sum of piece-wise chirps. The m^{th} signal segment of length N_w is obtained by:

$$s_m(n - \varsigma(m-1)) = s(n)h(n - \varsigma(m-1)), \tag{5.35}$$

where $n = \varsigma(m-1), \varsigma(m-1)+1, ..., \varsigma(m-1)+N_w-1$, $\varsigma(1 \leq \varsigma \leq N_w)$ is the shift between two consecutive windows, m is the window index, and $h(n)$ is a rectangular window which is non-zero only for $0 \leq n \leq N_w - 1$.

Then the chirp-approximated m^{th} signal segment of $s(n)$ is written as:

$$\begin{aligned} s_m(n) \approx & \sum_{e=1}^{E} A_{e,m} \exp\left\{ j2\pi \left[\alpha_{e,m} \frac{n^2}{2F_s^2} + \beta_{e,m} \frac{n}{F_s} \right] \right\} \\ & + v_m(n), \end{aligned} \tag{5.36}$$

where $0 \leq n \leq N_w - 1$, $A_{e,m}$, $\alpha_{e,m}$, $\beta_{e,m}$ are respectively the complex amplitude, the chirp rate, and the initial frequency of the e^{th} chirp over the m^{th} window.

Since $|F_e(n)| \leq F_{\max}$, the chirp rate α and the initial frequency β have to satisfy:

$$\begin{cases} |\beta| \leq F_{\max}, \\ |\alpha| \leq F_{\max} F_s / N_w. \end{cases} \tag{5.37}$$

In vector form, the signal over the m^{th} window can be expressed as:

$$\mathbf{S}_m = \mathbf{\Psi} \mathbf{X}_m + \mathbf{V}_m, \tag{5.38}$$

where $\mathbf{S}_m = [s_m(0), ..., s_m(N_w - 1)]^T$, $\mathbf{V}_m = [v_m(0), ..., v_m(N_w - 1)]^T$. The dictionary matrix, $\mathbf{\Psi}$, is designed by uniformly sampling the chirp rate space. Let I denote the total number of chirp rate values, $\tilde{\alpha}_i$ is i^{th} chirp rate value in the dictionary, and $\tilde{\beta}_i$ is corresponding initial frequency value for each $\tilde{\alpha}_i$ in the dictionary. The chirp atom $\mathbf{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, ..., \boldsymbol{\psi}_I]$ is obtained by:

PROCEDURE: For each value of $\tilde{\alpha}_i$

1. Calculate the FRFT angle $\tilde{\phi}_i = \arctan \frac{\tilde{\alpha}_i N_w}{F_s^2}$ (see (4.15)).
2. Calculate the magnitude response $|\mathrm{FT}(S^{\tilde{\phi}_i}(x))|$.
3. Find value of y at which $|\mathrm{FT}(S^{\tilde{\phi}_i}(x))|$ obtains a maximum, or d in short.
4. Calculate the corresponding value of the chirp rate (see (5.29)).
5. $\boldsymbol{\psi}_i = \exp[j2\pi(\tilde{\alpha}_i \frac{n^2}{2F_s^2} + \tilde{\beta}_i \frac{n}{F_s})]$.

Since $E < N_w \ll I$, $\mathbf{X}_m$ is highly sparse and solving for $\mathbf{X}_m$ in equation (5.38) becomes a sparse recovery (or CS) problem, which can be solved by:

$$\hat{\mathbf{X}}_m = \arg\min \|\mathbf{X}_m\|_1 \quad s.t. \quad \|\mathbf{S}_m - \boldsymbol{\Psi}\mathbf{X}_m\|_2^2, \leq \epsilon \tag{5.39}$$

where ϵ is the noise level. The solution for (5.39) can be obtained by a greedy algorithm such as Orthogonal Matching Pursuit (OMP) or linear programming. With the simplified calculation technique for the DFT of the FRFT of the signal, we only calculate the FRFT of the window (Gaussian window for example) with all values of the FRFT angle at one time. If we have $(F_s + 1)$ values of the FRFT angle, the calculation load approximates to performing the DFT $(F_s + 1)$ times. The simplified chirp dictionary is formed by projecting the signal onto the FRFT dictionary $\boldsymbol{\Phi}$, with the calculation load being (F_s+1) times that of the DFT, or $(F_s+1)^2$ projections. After we have the simplified chirp dictionary, the computational complexity is the same as the sinusoidal dictionary approach. As we have to form a new dictionary for each window signal, the computational load is $O(EI+I^2)$, which is much less than that of the full chirp dictionary, $(O(E(\frac{3}{4}I^2+\frac{1}{4})))$, especially when we have many components in the windowed signal.

5.4.3 Restricted isometry properties (RIP) analysis of the simplified chirp dictionary

Similar to the section 5.3.3, we examine the RIP associated with the simplified chirp dictionary using the FRFT. In Fig. 5.10, we compare the boundaries of the eigenvalues of the inner product of the simplified chirp matrix with those of the Gaussian random dictionary. As these bounds compete with those of the Gaussian

matrix and the Gaussian matrix satisfies the RIP with minimum $O(E\log(d/E))$ number of measurements, the simplified chirp dictionary matrix using the FRFT also satisfies the RIP with the same condition.

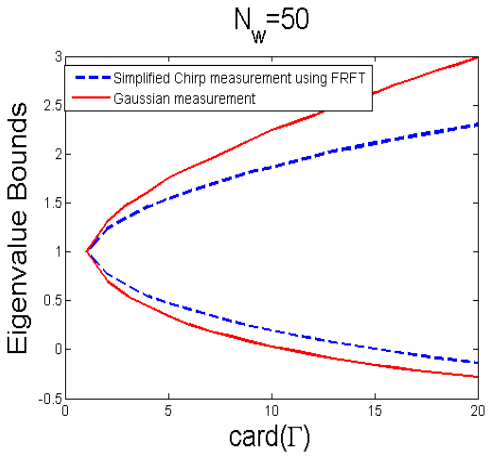


Figure 5.10: Eigenvalue bounds of $\mathbf{\Psi}_\Gamma^H\mathbf{\Psi}_\Gamma$ and $\mathbf{G}_\Gamma^H\mathbf{G}_\Gamma$ where $\mathbf{\Psi}$ is the simplified chirp dictionary using the FRFT, and $\mathbf{G}$ is the Gaussian random matrix.

5.4.4 Simulation

This section evaluates the performance of the FRFT based chirp dictionary in sparse reconstruction of non-stationary signals. We compare the proposed method with the full chirp dictionary, the sinusoidal dictionary and the FRFT approaches. In the FRFT approach, in each window, we only calculate the magnitude response of the FRFT for angles $\tilde{\phi}_i$, and choose the values of $\tilde{\phi}_i$ or $\tilde{\alpha}_i$ which have the largest magnitude responses. In the following examples, signals are sampled at the Nyquist rate and then 50% of samples are randomly removed. The sampling frequency is $F_s = 256$, the total signal length is $N = 256$. The observations are corrupted by white Gaussian noise, and the signal to noise ratio is set to $SNR = 20$dB. A rectangular window of length $N_w = 64$ is applied. The resulting images are normalized and transferred to energy versions for display. Similar to chapter 4, the concentration level ζ is used to assess the accuracy of the resulting TF representations. And the higher ζ, the better is the TF estimation. We assume that signals have $E = 5$ components.

In the first sample, the signal consists of two crossing chirps, which is expressed as:

$$s(n) = \exp\left\{j2\pi[(0.1F_s)\frac{n}{F_s} + (0.3F_s)\frac{n^2}{2F_s^2}]\right\} + \exp\left\{j2\pi[(0.4F_s)\frac{n}{F_s} - 0.3F_s\frac{n^2}{2F_s^2}]\right\} + v(n), \tag{5.40}$$

with $n = 0, 1, ..., N-1$. The results are shown in Fig. 5.11. The chirp dictionary and the FRFT based chirp dictionary provide perfect frequency localization with $\zeta = \infty$. The FRFT is unable to recover the TF representation of the whole signal because of missing data and noise, although one signal component is accurately displayed with $\zeta = 3000$. The sinusoidal method reveals inaccuracies in the TF signature estimation with $\zeta = 3$ since besides insufficient sparsity, it is also vulnerable to the picket fence effect [13, 68], resulting in frequency content at false locations.

Similar results are obtained in the next example where we use a signal composed of three components expressed as:

$$\begin{aligned} s(n) = &\exp\left\{j(0.1F_s)\cos(2\pi\frac{n}{F_s} + \pi) + j2\pi(0.2F_s)\frac{n}{F_s}\right\} \\ &+ \exp\left\{j(0.1F_s)\cos(2\pi\frac{n}{F_s} + \pi) + j2\pi(0.3F_s)\frac{n}{F_s}\right\} \\ &+ \exp\left\{j2\pi[(0.1F_s)\frac{n}{F_s} + (0.3F_s)\frac{n^2}{2F_s^2}]\right\} + v(n), \end{aligned} \tag{5.41}$$

with $n = 0, 1, ..., N-1$. The TF signature approximations of the four methods are displayed in Fig. 5.12. The FRFT based chirp dictionary and the normal chirp dictionary have pretty similar concentration levels of $\zeta = 20$, whereas the sinusoidal dictionary and the FRFT have lower concentration levels with $\zeta = 3$ and $\zeta = 8$, respectively.

The last example uses data from human gait radar returns obtained at the Radar Imaging Lab of the Center for Advanced Communication at Villnova University, USA. The data is first uniformly sampled at the Nyquist rate with $F_s = 1000$Hz, and then thinned by randomly removing 50% of the samples. The sparsity level is assumed to be $E = 30$. The window length is $N_w = 128$. The TF signature estimation is obtained by the FRFT based chirp dictionary, full-chirp and sinusoidal dictionary. The results in Fig. 5.13 shows that the FRFT based chirp dictionary can

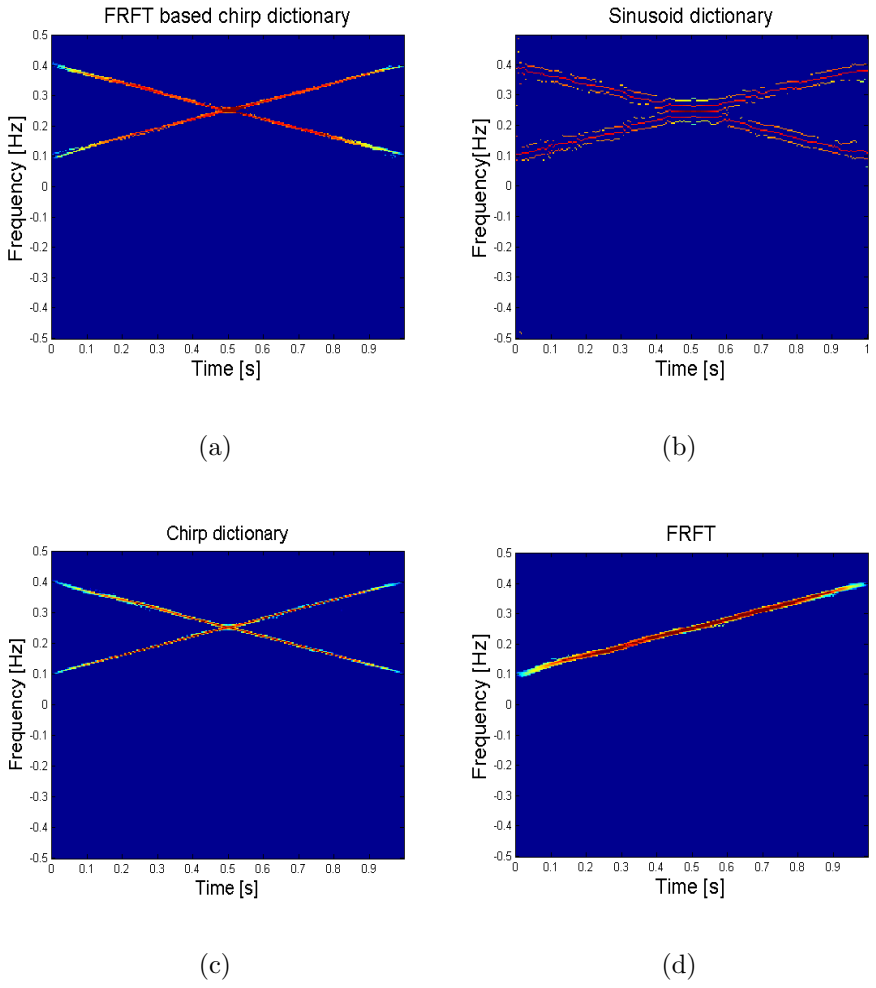
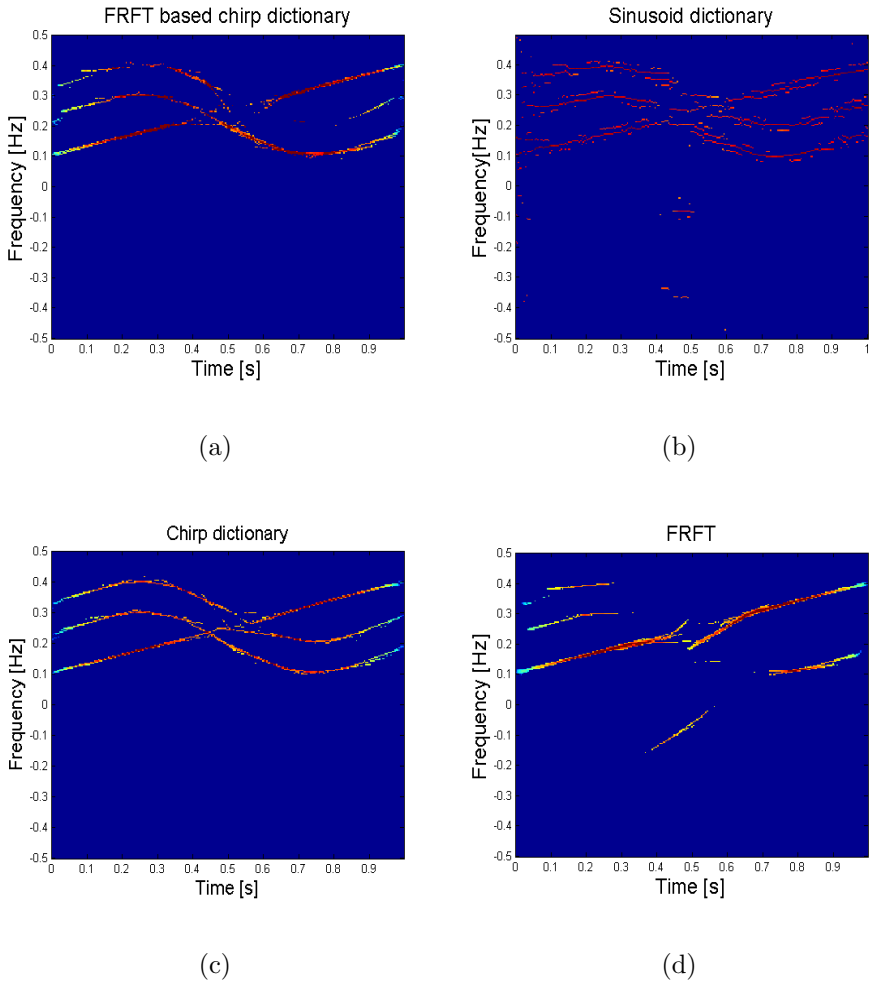


Figure 5.11: TF (frequency normalized) signature for $s(n)$ in (5.40) with 50% data missing: (a) FRFT based chirp dictionary; (b) Sinusoidal dictionary; (c) Chirp dictionary; (d) FRFT.

be used to describe the micro-Doppler TF presentations of torso and limbs under compressed observations.



(a) (b)

(c) (d)

Figure 5.12: TF (frequency normalized) signature for $s(n)$ in (5.41) with 50% data missing: (a) FRFT based chirp dictionary; (b) Sinusoidal dictionary; (c) Chirp dictionary; (d) FRFT.

5.5 Conclusion

The accurate piece-wise chirp approximations to the time-frequency signature of many Doppler and micro-Doppler signals motivate the use of a chirp dictionary for sparse reconstruction of the signals' local frequency structure under full and incomplete data. Compared with the sinusoidal dictionary method, the chirp dictionary approach can relax the converse request on the number of necessary observations

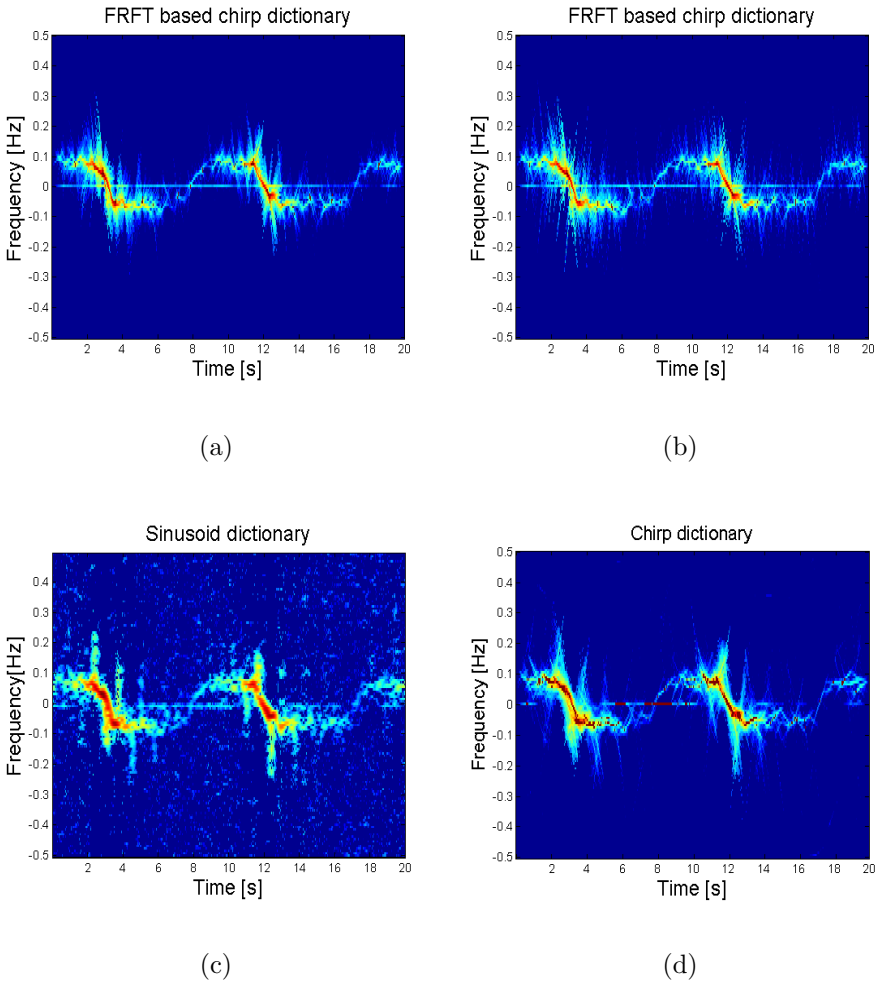


Figure 5.13: TF (frequency normalized) signature for $s(n)$ in (5.41):(a) FRFT based chirp dictionary with full data; (b) FRFT based chirp dictionary with 50% data missing; (c) Sinusoidal dictionary with 50% data missing; (d) Chirp dictionary with 50% data missing.

for exact solution and sparsity, and it can address the picket fence effect. Thus, the chirp atom gives (in general) a better performance. The simplified chirp dictionary is proposed in order to reduce the calculation burden. In summary, the dictionary is built for each window signal. In the first approach, the chirp rate of a chirp component inside the segment is estimated before building the dictionary. The second

method deploys the FRFT to find a corresponding initial value for each chirp rate value. Thanks to these two methods, the dictionary dimension gets as small as the sinusoidal atom set with better TF estimations.

Chapter 6

Reduced Interference Chirp-based Time-Frequency Distribution for Limited Data

6.1 Introduction

6.1.1 Motivation

Joint time-frequency distributions (TFDs) is a powerful tool to analyze non-stationary signals like radar returns from static or moving object, ECG, etc., [9, 36, 39, 75, 76, 77]. These non-stationary signals arise in many different applications,thus no single time-frequency (TF) estimation approach can be ideal in all cases.

The short-time Fourier transform (STFT) is the most widely used and simplest method for studying non-stationary signals [5, 78]. The major limitation is the trade-off between time and frequency resolution. The Wigner-Ville distribution (WVD) achieves better resolution, but due to the bilinear product, the WVD experiences cross-terms between different components as well as those between the same component. In order to eliminate (or at least attenuate) the cross-terms in the WVD, a class of reduced interference distributions (RIDs) has been defined which belongs to the general Cohen's class. As auto-terms generally are located near the origin in ambiguity domain, the RID utilizes kernels, which are two-dimensional low-pass filters, to reduce the undesired cross-terms, while still preserving the true signal

power terms. Traditionally, fixed kernels are used, for example Choi-Williams or Born Jordan kernels [18]. However, these fixed masks do not always work effectively because there are signals with their cross-terms residing near the origin, for example, a signal composed of multiple crossing chirps. In this case, these Cohen RID kernels let part of the cross-terms through, thus hiding the signals' TF signatures.

As the distribution of cross- and auto-components in the ambiguity domain actually depends upon the analyzed data, signal-dependent kernels have been proposed to guarantee a good performance for a large class of signals. The radially-Gaussian kernel (RGK) is an outstanding example in this category. It actually solves an optimization problem based on two performance criteria in order to achieve an optimal kernel [20]. The first requirement constrains the kernel to be a low-pass filter, and the second one limits the area of the kernel to a kernel volume so that the cross-components can be suppressed. It is very important to choose a proper value for the kernel volume because it controls the trade-off between cross-term suppression and auto-term smearing. The principle of this method is that we keep the magnitude of the kernel in the ambiguity domain large wherever that of the ambiguity function of the signal is large, regardless of whether the peaks correspond to auto- or cross-components. There is also a modified version of this approach, called the adaptive optimal-kernel (AOK), which allows on-line implementation [79]. This method employs a time-localized short-time ambiguity function (STAF), and the radially Gaussian kernel optimization procedure to obtain the optimal kernel. The TF slice at the middle of the duration is obtained by implementing the two-dimensional Fourier transform of the STAF-kernel product. In general, these signal-dependent kernels algorithms construct a mask that automatically matches with the signal's auto-terms, and so they largely remove the cross-terms and perform well with many types of signals. However, these methods are vulnerable with incomplete signals. Missing samples introduce noise-like artifacts in the whole ambiguity domain. They distort the signals' auto-terms, and thus misguide the RGK or AOK to capture the wrong regions in the ambiguity domain. Therefore, the optimization process becomes unreliable. Thus, we need other TF estimation methods that can deliver accurate results even in the case of missing data.

6.1.2 Related work

There have been a lot of TFDs proposed which can give good results in the case of incomplete data. Most of them are based on compressive sensing (CS) techniques. In recent year, CS has attracted considerable attention in many areas of the discipline by suggesting that it is possible to record a small number of linear measurements of a signal and then reconstruct the complete set of all samples if the signal itself is sparse. The required number of observations is slightly more than the signal sparsity level but less than the signal dimension. As non-stationary signals generally exhibit small occupancy in the TF domain, with small number of samples either in the time domain or in the ambiguity or in the instantaneous autocorrelation domains, TF signatures can be recovered with high accuracy with compressive sensing techniques. The first TFD approach is to perform sparse reconstruction from windowed data in the time domain. The measurement dictionary can be sinusoids or chirps [3, 13, 14, 27]. The second approach, within the Cohen's class, deploys a low-pass filter to mitigate the cross-components, but with sparsity in the TF domain, it yields robustness to missing data [2, 28, 64, 67, 80, 81].

In the latter category, there are some methods which modify the signal-dependent kernel AOK to get the sparsest TFRs. In [67], the author uses multiple sensors instead of one sensor. The auto-terms TFDs are real and also positive for meaningful TF points where the signal energy is concentrated. On the other hand, the values and signs of the cross-terms between different signal components depend on the relative phase between the contributing signals. This phase changes across the antenna array due to respective propagation delays. As such, the simple averaging of TFDs over different antennas does not favour the cross-terms and enhances the auto-terms. Meanwhile, the missing data samples yield artifacts that are randomly spread over the entire ambiguity domain. When different sampling patterns are adopted in each sensor, then, averaging the AF over all sensors effectively reduces the contributions of the artifacts induced from missing samples. Thus, the optimization of kernel parameters is proposed using the AF averaging over all sensors instead of the AF obtained in each individual sensor. The TFD can be computed as the two-dimensional Fourier transform of the kernelled and averaged AF. Alternatively, we can also obtain the TFD through a sparse reconstruction from the same kernelled and averaged AF. This method, however, requires more sensors, and thus need more

computational load. In [64, 81, 82], the authors introduce a modified AOK, which is robust to cross-terms and missing samples. Beside constructing a kernel which best matches with the signal's auto-terms, the method considers sparsity in the TF domain as an optimization object. Although it performs well, the computational complexity is high.

6.1.3 Contribution

In this chapter, we introduce novel fixed and signal-dependent kernels in the ambiguity domain, which can efficiently remove cross-term interference and partially combat missing sample effects without using compressive sensing techniques. These kernels are applied on windowed signals to facilitate online implementation, or processing long signals. According to [13, 14], any non-stationary signal segment can be approximated by a sum of chirps. Additionally, chirps' auto-terms always reside in only half of the ambiguity domain which do not cover the Doppler axis. By removing the areas where auto-terms do not lie in, part of interference and artifacts are mitigated. Moreover, the analysis of artifacts distributions shows that the artifacts always appear along the Doppler axis. By removing the region along the Doppler axis, our chirp-based kernels give satisfactory TFRs in the case of incomplete data.

6.1.4 Chapter outline

So this paper is organized as follows. Section 6.2 presents a review of reduced interference distributions including the fixed and adaptive kernels. Section 6.3 includes discussions about missing sample effects on the ambiguity domain, and the unsuitability of using traditional RIDs for incomplete data. Section 6.4 introduces the fixed and adaptive chirp-based kernels. These proposed kernels are based on the chirp property in the ambiguity domain, which is also expressed here. Section 6.5.1 gives some details about fast implementation, which helps save the computational load. Section 6.6 shows simulation results. Finally, conclusions are given in section 6.7.

6.2 Conventional Reduced Interference Kernels

One of the major drawbacks of the WVD is the cross-terms located between signal components, which can obscure desired signal TF information. Thus, RIDs have been defined in order to suppress the cross-terms as well as preserve the auto-terms [5, 42]. For a complex-valued signal sampled with period T, i.e., $s(n) = s(nT)$, the RID $D(n,k)$ is obtained by the two-dimensional Fourier transform of the product of the AF $A(n,k)$ and the kernel function $C(p,b)$ as follows:

$$D(n,k) = \sum_{p=-N/2}^{N/2-1} \sum_{b=-N/2}^{N/2-1} C(p,b)A(p,b)e^{j(-bk-pn)2\pi/N}, \tag{6.1}$$

where

$$A(p,b) = \sum_{n=0}^{N-1} s(n+b)s^*(n-b)e^{-j2\pi pn/N}, \tag{6.2}$$

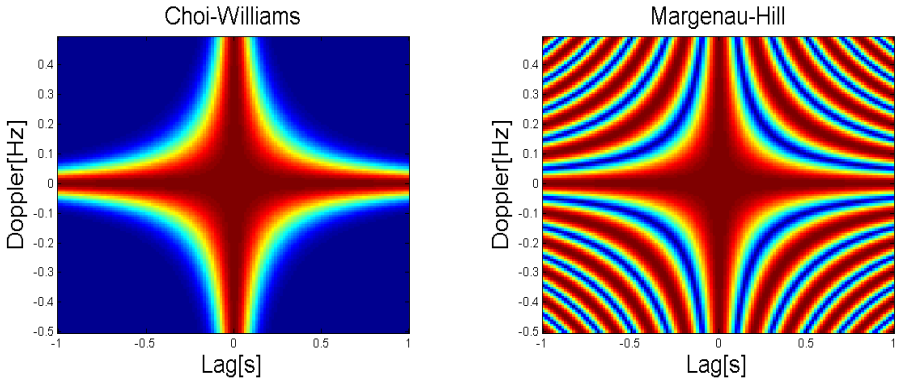
where n,k are discrete time and frequency variables, $n = 0,1,2,...,N-1$, $k = 0,1,2,...,N-1$ and p and b denote the frequency shift (Doppler frequency) and time lag. The data bilinear product in (6.2) is the instantaneous autocorrelation function (IAF). As such, the AF is the Fourier transform of the IAF over the time variable.

$$R_{ss}(n,b) = s(n+b)s^*(n-b). \tag{6.3}$$

In the ambiguity domain, most of the desired auto-terms are located at and around the origin, whereas the cross-terms reside at distant positions. The kernel function acts as a low-pass filter in the ambiguity domain and places different weights on the ambiguity samples. As the result, the auto-terms are retained while the cross-terms are substantially removed.

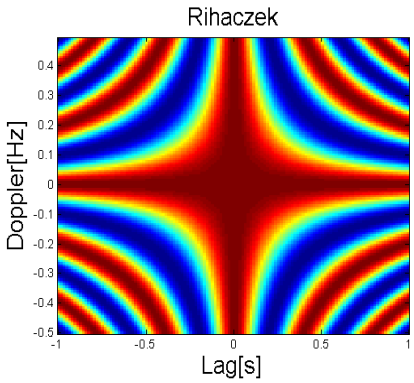
Existing TF kernels can be divided into two forms: signal-independent and signal-dependent. The former includes Choi-Williams kernel [18], Margenau-Hill kernel [19], Rihaczek kernel [83] and Born-Jordan kernel [84], etc. The kernel functions in the ambiguity domain of the aforementioned kernels are written in Table. 6.1 and plotted in Fig. 6.1.

To maintain most of the desirable properties of the WVD, the kernel is required to satisfy the marginal properties and realness. The former requires the kernel to be unity along lag and Doppler axis ($p = 0$ and $b = 0$). The latter is acquired

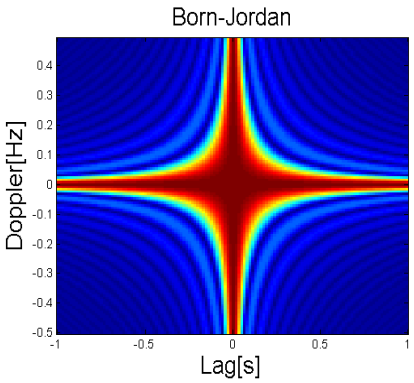


(a)

(b)



(c)



(d)

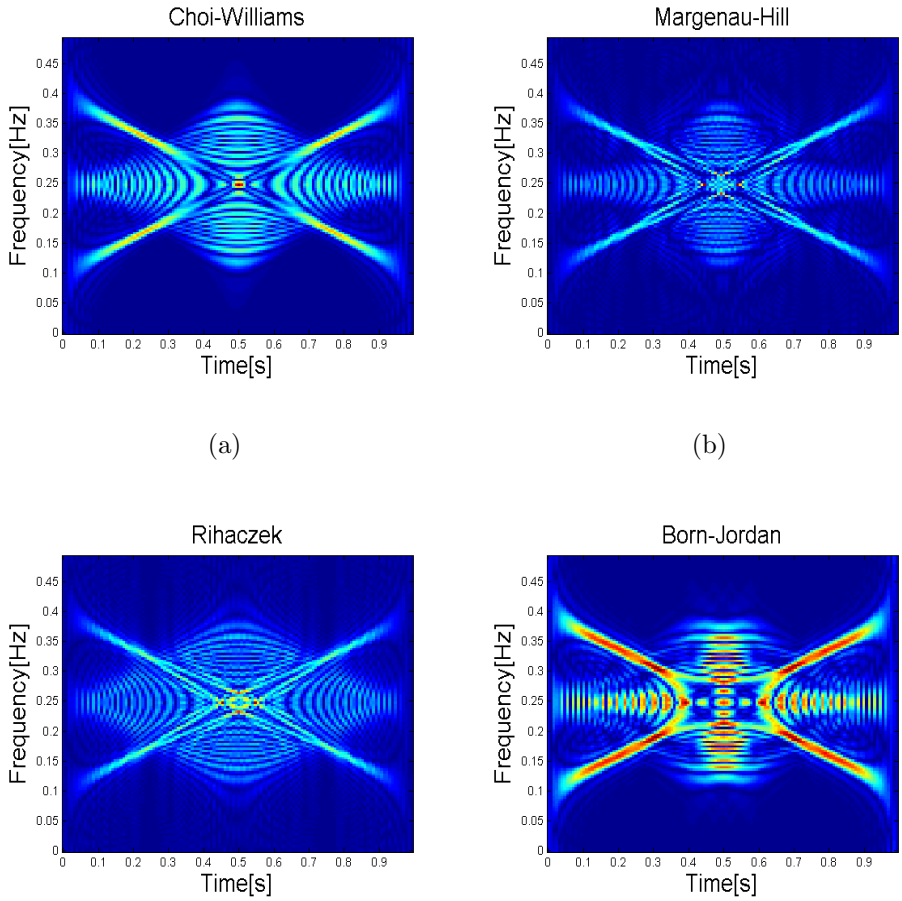
Figure 6.1: (a) Choi-Williams kernel; (b) Margenau-Hill kernel; (c) Rihaczek kernel; (d) Born-Jordan kernel.

Table 6.1: Some signal-dependent distributions and their kernels

Distribution	**Kernel $C(p, b)$**
Choi-Williams	$\exp(-p^2b^2/\sigma)$
Margenau-Hill	$\cos(pb/2)$
Rihaczek	$\exp(jpb/2)$
Born-Jordan	$\mathrm{sinc}(\frac{1}{2}pb)$

if $C(p, b) = C^*(-p, -b)$. The major drawback of the signal-independent signal is that they have the fixed shapes in the ambiguity domain. The fixed shapes result in inflexibility for a large class of signals. For example, the cross-terms of a signal composed of two crossing chirps reside near the origin, while there are some auto-terms locating far away from the center. With its low-pass property, the signal-dependent kernel does not perform efficiently for this signal. To illustrate, the four kernels in Table. 6.1 are used, then the resulting TFRs of the crossing-chirp signal are contaminated shown in Fig. 6.2.

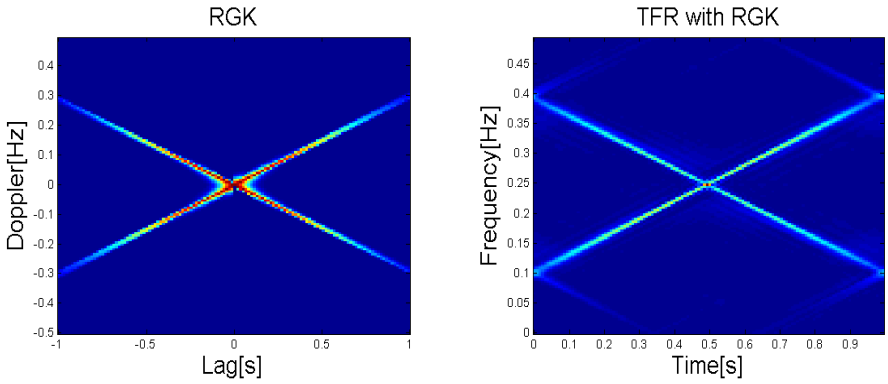
Signal-dependent or adaptive kernels tune their representation to each signal to offer good performance for a large class of signals [20]. A natural approach to tune is via optimization; this requires an objective function or performance measure to express what "we want to see" and a class of representations to choose from. There are two broad categories: adaptive TFRs based on the linear short time Fourier transform, and those based on the quadratic WD. In an adaptive quadratic TFR, we adapt the form of the kernel to match the shape of the signal's components in the ambiguity domain. In particular, we want the kernel to be close to 1 in the vicinity of the signal's auto-component, and close to 0 in the vicinity of the cross-components. The most well-known example is the radially-Gaussian kernel (RGK). Its objective is to maximize the energy in the ambiguity domain in order to suppress the cross-terms and to pass the auto-terms with as little distortion as possible. The



(a) (b)

(c) (d)

Figure 6.2: RIDs of a signal composed of two crossing chirps obtained by: (a) Choi-Williams distribution; (b) Margenau-Hill distribution; (c) Rihaczek distribution; (d) Born-Jordan distribution.



(a) (b)

Figure 6.3: (a) The RGK with $vo = 3$; (b) The resulting TFR for a signal composed of two crossing chirps.

RGK is adapted to a signal by solving the following optimization problem:

$$\begin{aligned} \max_{\sigma(q),\forall r,\forall q} \quad & \sum_{r}\sum_{q} r\,|A_{ss}(r,q)C(r,q)|^2 \\ \text{subject to} \quad & C(r,q) = e^{-\frac{r^2}{2\sigma^2(q)}} \\ & \sum_{q=0}^{N-1} \sigma^2(q) \leq vo, \end{aligned} \tag{6.4}$$

where σ is the standard deviation of the Gaussian kernel. The first condition in (6.4) forces the kernel to have low-pass filter characteristics, which lends itself to cross-term reduction. The second condition limits the area of the kernel to the kernel volume "vo". Cross-term suppression and auto-term preservation trade-off depends on the choice of "vo". The optimization problem is performed in polar coordinate in terms of the radius r and the aspect angle q. RGK has proved superior to data-independent kernels. A simulation is carried out with the above crossing-chirp signal. The results in Fig. 6.3 show that the RGK kernel has covered most of the auto-terms and eliminates most of the cross-terms.

While RGK generally performs well, it is block-oriented technique that calculates only one kernel for an entire signal. For analyzing signals with characteristics that change over time, and for real-time, online operation, or for very long signals,

adaptive signal-dependent kernels are required. To track the local signal characteristics over time, the kernel optimization procedure needs to consider the local signal characteristics. With STAF, we can apply RGK into a windowed signal centered at n_0 to produce the AOK for the signal segment. The frequency slice at time n_0 is obtained by the two-dimensional Fourier transform of the product of the obtained AOK and the STAF. Since the algorithm alters the kernel at each time to achieve optimal local performance, better tracking of signal changes results.

The traditional signal-independent and signal-dependent kernels work properly when we have full data. In the case of random or missing samples, these distributions give undesirable TFRs. The signal-independent kernels with fixed shapes in the ambiguity domain can efficiently suppress the cross-terms which often locate far from the origin, but not the noise-like artifacts caused by missing samples which spread over the entire ambiguity domain. The dependent-signal kernels also perform inappropriately in the presence of missing data because their artifacts can fool the AOK and RGK into capturing the wrong areas of the AF.

There have been many methods proposed to confront missing sample effects. They mostly use compressive sensing techniques to obtain TFRs which are robust to missing data [2, 3, 4, 13, 13, 27, 28, 82, 85, 86, 87, 87, 88, 89].

In this chapter, we propose a novel method which works efficiently in case of compressed data without using CS techniques. Based on the windowed non-stationary signal properties in the ambiguity domain and the artifacts' distribution, our method designs a fixed kernel and an adaptive kernel for each short-time signal in the ambiguity domain.

6.3 The Effect of Missing Samples on The Ambiguity Domain

In this section, we analyze the effect of missing samples on the ambiguity function [81]. Signal with missing samples can be represented as a modulated version of the original signal $x(n)$,

$$s(n) = x(n)\varphi_p(n), \tag{6.5}$$

where $\varphi_p(n)$ is the sampling pattern in time and can be represented as the sum of impulses at random positions n_p, i.e., $\varphi_p(n) = \sum_{n_p} \delta(n - n_p)$. These impulses determine the positions of available samples. The corresponding pattern of missing samples $\varphi_{miss}(n)$ at positions n_{miss} is given by,

$$\varphi_{miss}(n) = \sum_{n_{miss}} \delta(n - n_{miss}) = 1 - \varphi_p(n). \tag{6.6}$$

Thus, the signal can be represented by missing samples as follows:

$$s(n) = x(n)(1 - \varphi_{miss}(n)) = x(n)\left(1 - \sum_{n_{miss}} \delta(n - n_{miss})\right). \tag{6.7}$$

The ambiguity function of $s(n)$ becomes:

$$\begin{aligned}
&A_s(p,b)\\
&= \sum_{n=-N/2}^{N/2-1} x(n+b)(1-\varphi_{miss}(n+b))x^*(n-b)(1-\varphi_{miss}(n-b))e^{-j2\pi np/N}\\
&= \sum_{n=-N/2}^{N/2-1} R_{xx}(n,b)(1-\varphi_{miss}(n+b))(1-\varphi_{miss}(n-b))e^{-j2\pi np/N}\\
&= \sum_{n=-N/2}^{N/2-1} R_{xx}(n,b)e^{-j2\pi np/N} - \sum_{n=-N/2}^{N/2-1} R_{xx}(n,b)\varphi_{miss}(n-b)e^{-j2\pi np/N}\\
&- \sum_{n=-N/2}^{N/2-1} R_{xx}(n,b)\varphi_{miss}(n+b)e^{-j2\pi np/N}\\
&+ \sum_{n=-N/2}^{N/2-1} R_{xx}(n,b)\varphi_{miss}(n+b)\varphi_{miss}(n-b)e^{-j2\pi np/N}\\
&= A_x(p,b) + V_1(p,b) + V_2(p,b) + V_3(p,b) = A_x(p,b) + \Delta(p,b).
\end{aligned} \tag{6.8}$$

$R_{xx}(n,b)$ and $A_x(p,b)$ denote the IAF and AF of the original signal $x(n)$. Thus, the ambiguity function of the compressed observations can be represented as the sum of the AF of the full data $A_x(p,b)$ and the three artifacts caused by the missing samples.

From (6.8),we can define the ideal kernel, $C_s(p,b)$, for incomplete data $s(n)$ as follows:

$$A_s(p,b)C_s(p,b) = A_x(p,b)C_x(p,b), \tag{6.9}$$

where $C_x(p,b)$ is a desirable RID kernel, which performs well on the full signal, appropriately suppressing the cross-terms and preserve the auto-terms. From (6.8) and (6.9), we obtain:

$$\begin{aligned} C_s(p,b) &= A_x(p,b)C_x(p,b)/A_s(p,b) \\ &= \frac{A_x(p,b)C_x(p,b}{A_x(p,b)+\Delta(p,b)} \\ &= \frac{C_x(p,b)}{1+\frac{\Delta(p,b)}{A_x(p,b)}}. \end{aligned} \tag{6.10}$$

(6.10) implies that the ideal kernel is a signal-dependent kernel, and it is complex which, therefore, violates the realness property of the traditionally used kernels.

By further developing each of the three artifacts' terms, we obtain the following expressions:

$$\begin{aligned} V_1(p,b) &= -\sum_{n=-N/2}^{N/2-1} x(n+b)x^*(n-b)\sum_{n_{miss}}\delta(n-b-n_{miss})e^{-j2\pi np/N} \\ &= -\sum_{n_{miss}} x(2b+n_{miss})x^*(n_{miss})e^{-j2\pi(b+n_{miss})p/N}. \end{aligned} \tag{6.11}$$

$$\begin{aligned} V_2(p,b) &= -\sum_{n=-N/2}^{N/2-1} x(n+b)x^*(n-b)\sum_{n_{miss}}\delta(n+b-n_{miss})e^{-j2\pi np/N} \\ &= -\sum_{n_{miss}} x(n_{miss})x^*(n_{miss}-2b)e^{-j2\pi(n_{miss}-b)p/N}. \end{aligned} \tag{6.12}$$

$$\begin{aligned} V_3(p,b) &= \sum_{n=-N/2}^{N/2-1} x(n+b)x^*(n-b)\sum_{n_{miss}}\delta(n+b-n_{miss})\sum_{n_{miss}}\delta(n-b-n_{miss})e^{-j2\pi np/N} \\ &= \sum_{n_{miss}}\delta(2b)x(n_{miss})x^*(n_{miss})e^{-j2\pi n_{miss}p/N} \\ &+ \sum_{b=-N/2}^{N/2}\sum_{n_{miss}}\delta(n-n_{miss}+b)x(n+b)\sum_{n_l\neq n_{miss}}\delta(n-n_l-b)x^*(n-b). \end{aligned} \tag{6.13}$$

We can see that the first two terms (V_1 and V_2) can exist for $\forall b$. A component of the third term is always located at $b=0$, or along the Doppler frequency axis.

This discourages the use of traditional RID kernels, which capture all values along $b = 0$ due to marginal properties. With the same signal used as in Fig. 6.2 but with 50% of samples removed, the TFRs are severely contaminated as shown in Fig. 6.4. The signal-dependent kernel AOK or RGK is "mislead" by the noise-like effects, especially the artifacts along the Doppler frequency axis. This leads to inaccurate TFRs. Fig. 6.5 illustrates these artifacts with the same signal used in Fig. 6.3 but with 50% of samples absent.

We can conclude that the conventional signal-independent and signal-dependent kernels cannot offer reliable signal TF signature in the case of missing data.

6.4 RID Chirp-Based Kernel Design

6.4.1 Properties of chirps in the ambiguity domain

Consider a certain chirp with a chirp-rate α and initial frequency β as follows:

$$s(n) = \exp\left[j2\pi(\alpha\frac{n^2}{2F_s^2} + \beta\frac{n}{F_s})\right], \tag{6.14}$$

where F_s is the sampling frequency, n is the discrete time, $n = 0, 1, ..., \lfloor T/T_s \rfloor$, T is the total observation time, $T_s = 1/F_s$ is sampling period. Let N be the length of the signal, $N = \lfloor T/T_s \rfloor$.

The corresponding IAF is expressed as:

$$\begin{aligned} R_{ss}(n,b) &= s(n+\frac{b}{2})s^*(n-\frac{b}{2}) \\ &= \exp\left[j2\pi\left(\alpha\frac{n^2+b^2/4+nb}{2F_s^2} + \beta\frac{n+b/2}{F_s}\right)\right] \\ &\quad \exp\left[-j2\pi\left(\alpha\frac{n^2+b^2/4-nb}{2F_s^2} + \beta\frac{n-b/2}{F_s}\right)\right] \\ &= \exp\left[j2\pi\left(\alpha\frac{nb}{F_s^2} + \beta\frac{b}{F_s}\right)\right]. \end{aligned} \tag{6.15}$$

The WVD of $s(n)$ is given as:

$$\begin{aligned} D(n,\omega) &= \sum_b R_{ss}(n,b)e^{-jb\omega} \\ &= \delta\left[\frac{\omega}{2\pi} - (\beta + \alpha\frac{n}{F_s})\right]. \end{aligned} \tag{6.16}$$

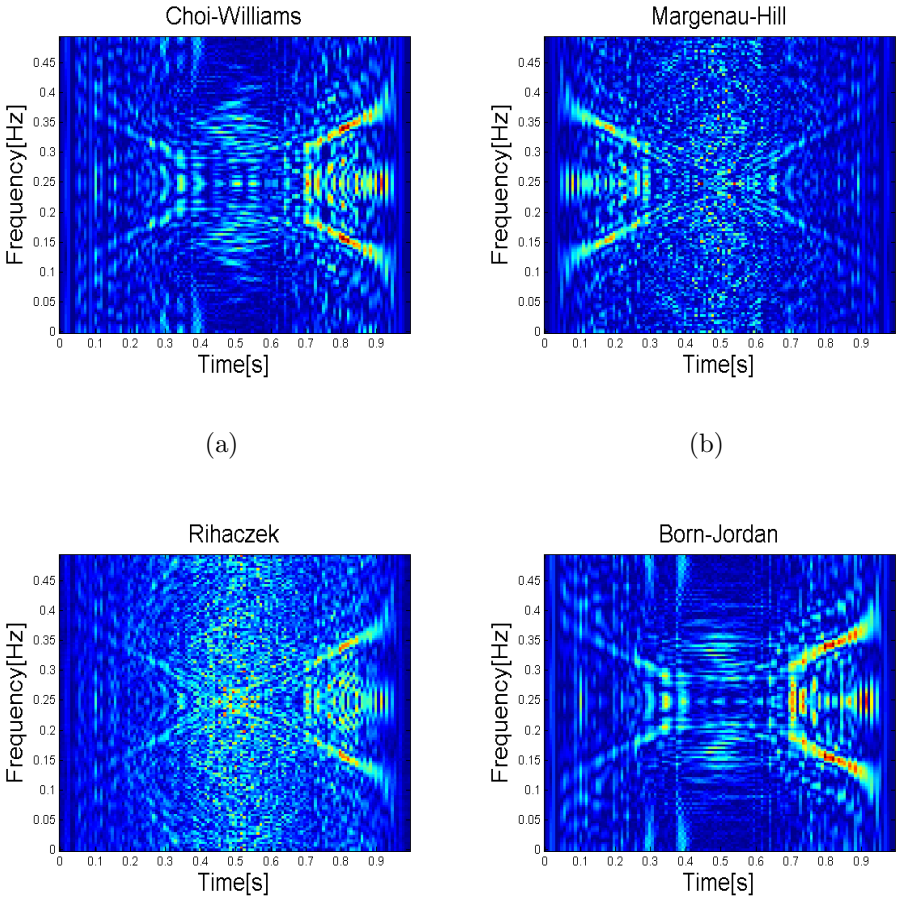
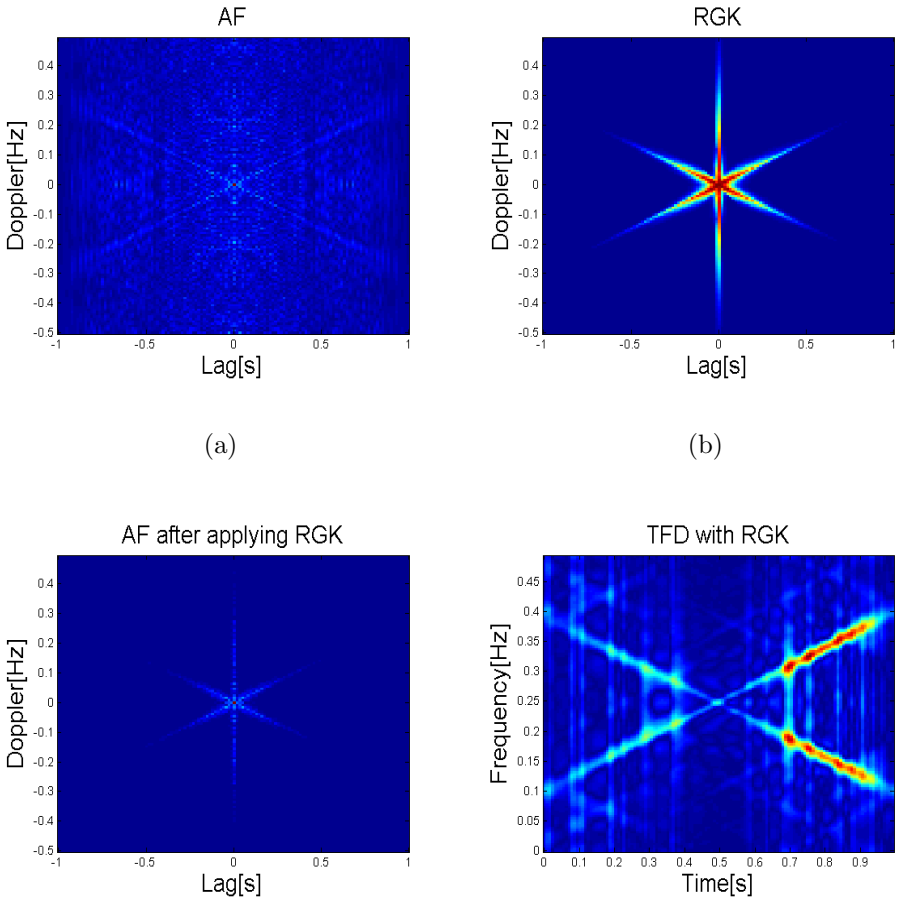


Figure 6.4: RIDs of an incomplete signal composed of two crossing chirps obtained by: (a) Choi-Williams distribution; (b) Margenau-Hill distribution; (c) Rihaczek distribution; (d) Born-Jordan distribution.



(a) (b)

(c) (d)

Figure 6.5: (a) AF; (b) RGK; (c) AF after being filtered by RGK; (d) TFR with RGK for a signal composed of two crossing chirps when 50% of the data is missing.

Thus, the instantaneous frequency of the chirp signal $s(n)$ is:

$$F(n) = \alpha \frac{n}{F_s} + \beta. \tag{6.17}$$

Assume that the signal is sampled at the Nyquist rate, i.e. the sampling frequency is double the maximum frequency of the signal, $F_s = 2F_{\max}$. Because $F_{\max}$ is the maximum frequency of the signal, so $F(n) \leq F_{\max}$. Also, the maximum frequency change in (N/F_s) second is $F_{\max}$. As the chirp-rate is the frequency change of a chirp in one second, the maximum chirp-rate is as follows:

$$|\alpha_{\max}| = F_{\max} \frac{F_s}{N}. \tag{6.18}$$

So if the signal is of one second length, $N = F_s$, the maximum chirp rate is $F_{\max}$.

The chirp signal $s(n)$ is expressed in ambiguity domain as follows:

$$\begin{aligned} A(\omega', b) &= \sum_n R_{ss}(n, b) e^{-jn\omega'} \\ &= \exp\left(j2\pi\beta \frac{b}{F_s}\right) \delta\left(\frac{\omega'}{2\pi} - \alpha \frac{b}{F_s}\right), \end{aligned} \tag{6.19}$$

where ω' is the Doppler angular frequency. (6.19) shows that the AF of all chirps has a linear support that passes through the origin of the ambiguity plane. The chirp auto-term lies at a certain angle to the horizontal line which is determined by the chirp-rate. Furthermore, since the chirp-rate is inside $[-F_{\max}\frac{F_s}{N}, F_{\max}\frac{F_s}{N}]$, the angle slope between the chirp and the horizontal in the ambiguity domain is also restricted. The chirp signal $s(n)$ in the ambiguity domain is plotted in Fig. 6.6. Based on Fig. 6.6, the slope between the chirp line and the horizontal line in the ambiguity domain is as follows:

$$\begin{aligned} \phi &= \arctan \frac{\alpha/\delta f}{1/\delta b} \\ &= \arctan \frac{2\alpha N}{F_s^2}, \end{aligned} \tag{6.20}$$

where $\delta_f = F_s/N$ is the frequency resolution and $\delta_b = 2/F_s$ is the lag resolution. Denote ϕ as the slope between the chirp and horizontal line in ambiguity domain. From (6.18) and (6.20), if we consider the positive lag plane, ϕ is restricted to:

$$-\pi/4 \leq \phi \leq \pi/4. \tag{6.21}$$

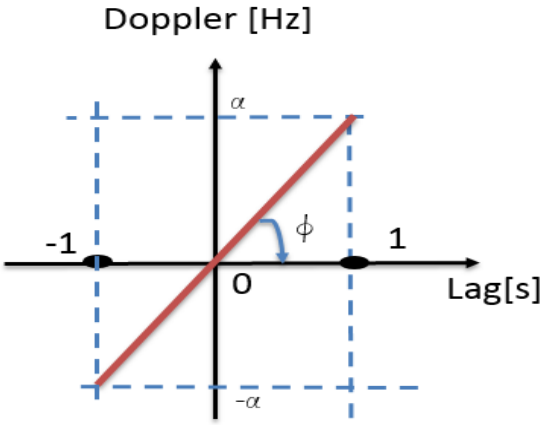


Figure 6.6: A chirp signal in ambiguity domain.

Similarly, for the negative lag plane, ϕ :

$$3\pi/4 \leq \phi \leq 5\pi/4. \tag{6.22}$$

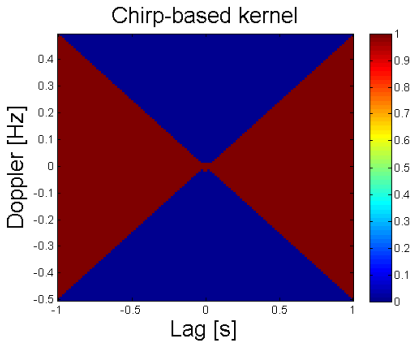
We can conclude that the auto-term of any chirp has either $|\phi| \leq \pi/4$ or $3\pi/4 \leq \phi \leq 5\pi/4$.

6.4.2 Kernel design for chirp signals

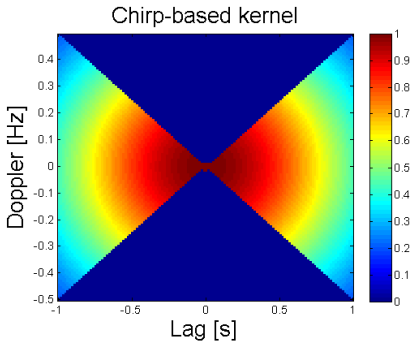
When the signal is composed of many chirps, and has missing samples, its AF is filled with cross-terms and noise-like artifacts spreading all over the ambiguity plane. However, as discussed in section 6.4.1, the auto-terms always locate inside $|\phi| \leq \pi/4$ and $3\pi/4 \leq \phi \leq 5\pi/4$, and so the kernel can filter out the rest, which corresponds to half of the ambiguity domain. Notice that the regions $|\phi| \leq \pi/4$ and $3\pi/4 \leq \phi \leq 5\pi/4$ do not include the Doppler axis, where the artifacts always lie along. Therefore, it substantially mitigates the adverse effects of the missing samples.

Based on this fact, we design a chirp-based kernel, which is basically the Gaussian kernel modified such that all components outside the regions $|\phi| \leq \pi/4$ and $3\pi/4 \leq \phi \leq 5\pi/4$ are zeros. A two-dimensional, radially-Gaussian kernel with a spread parameter σ [20] is expressed as:

$$C(p, b) = e^{-\frac{p^2+b^2}{2\sigma^2}}, \tag{6.23}$$



(a)



(b)

Figure 6.7: The proposed kernel in the AF domain: (a) $\sigma = \infty$; (b) $\sigma = 50$.

where p and b are Doppler and lag variables. The kernel is easily expressed in polar coordinates by using $r^2 = p^2 + b^2$ as the radius variable:

$$C(r, \phi) = e^{-\frac{r^2}{2\sigma^2}}. \tag{6.24}$$

The modified kernel is expressed as follows:

$$C(r, \phi) = \begin{cases} e^{-\frac{r^2}{2\sigma^2}}, & |\phi| \leq \pi/4 \quad \& \quad 3\pi/4 \leq \phi \leq 5\pi/4 \\ 0, & \text{otherwise.} \end{cases} \tag{6.25}$$

The proposed kernel is illustrated in Fig. 6.7. After applying a kernel $C(p, b)$, the resulting TFD is obtained through the 2D Fourier transform:

$$TFD(n, k) = \sum_{p} \sum_{b} A(p, b) C(p, b) e^{-j2\pi np/N} e^{-j2\pi bk/N}. \tag{6.26}$$

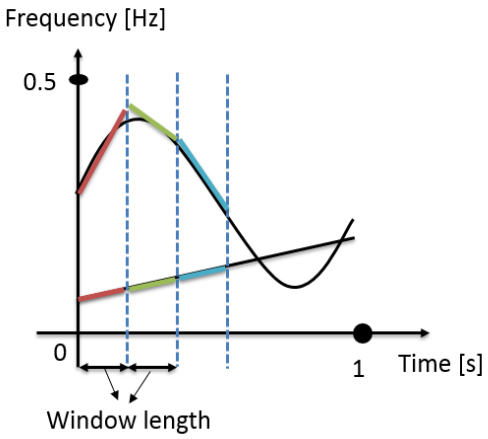


Figure 6.8: Non-stationary windowed signal approximated by chirps.

6.4.3 Windowed chirp-based kernel

The chirp-based kernel is basically applied when the input signals are chirps. Since other types of non-stationary signals have auto-terms locating anywhere in the ambiguity plane, this kernel results in inaccurate TF approximation. Nevertheless, according to [13, 14], the frequency law of any non-stationary windowed signal can be approximated as a sum of chirps as illustrated in Fig. 6.8. Therefore, for each non-stationary segments, the chirp-based kernel can be used. The algorithm proceeds as follows. The chirp-based kernel is first computed with the predefined window length N_w. At each time n, we compute the STAF centered at time n, $AF(n;p,b)$. $AF(n;p,b)$ is given by:

$$\begin{aligned} &AF(n;p,b) \\ &= \sum_u s^*(u-b/2)w^*(u-n-b/2)s(u+b/2)w(u-n+b/2)e^{j2\pi up/N_w} \\ &= \sum_u IAF(n;u,b)w^*(u-n-b/2)w(u-n+b/2)e^{j2\pi up/N_w}, \end{aligned} \tag{6.27}$$

where $w(u)$ is a symmetrical window function which is 0 when $|u| > N_w/2$ and $IAF(n;u,b)$ is the IAF of the windowed signal centered at n, which is illustrated in Fig. 6.9. The current-time slice of the TFR is computed as one slice (at time n only) of the two-dimensional Fourier transform of the STAF-kernel product, expressed as

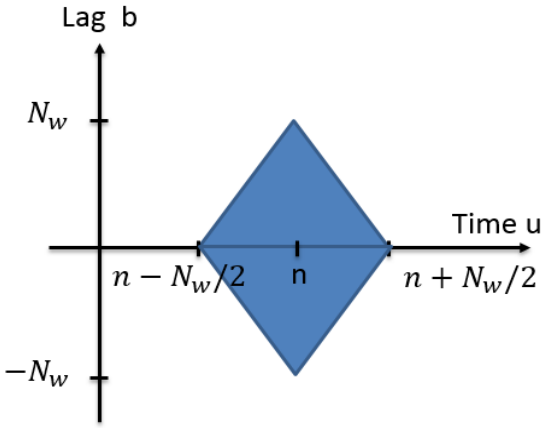


Figure 6.9: Illustration of the IAF of the windowed signal of length N_w.

follow:

$$TFR(n,k) = \sum_{p}\sum_{b} A(n;p,b)C(n;p,b)e^{-j2\pi np/N_w}e^{-j2\pi bk/N_w}. \tag{6.28}$$

Similar to the AOK, this method can be used to extract signals' TFRs for real-time, online implementation or for very long signals.

6.4.4 Chirp-based adaptive optimal kernel

This approach combines the chirp-based kernel and the AOK to design an optimal kernel for each windowed signal. The objective of the designed kernel is: 1) mitigating cross-components, 2) preserving auto-components and 3) combating with adverse effects of incomplete data in the time-frequency domain.

According to [20, 79], AOK solves the optimization problem in polar coordinates. It samples the angle ϕ and the radius r into Q, and G discrete values, respectively. The discrete kernel in polar coordinates is expressed as follows:

$$\begin{aligned} &C(g,q) = e^{-\frac{(g\Delta r)^2}{2\sigma(q\Delta\phi)^2}} \\ &g = 0,...,G-1,\ q = 0,...,Q-1, \end{aligned} \tag{6.29}$$

where g and q are the radius and angle indices, Δr and $\Delta\phi$ are the radius and angle step sizes. The discrete kernel is parameterized by a positive spread parameter,

$\sigma(q\Delta\phi)$, in such a way that it is large at a certain angle if the magnitude of the AF is large. Initially, the normal Gaussian kernel is used with equal spread parameter σ for every angle. Then, the spread parameter is updated by the gradient ascent methods to match with signals in the ambiguity plane. The AOK is implemented on windowed signals in order to suit online and real-time applications. A major drawback of the AOK in the case of missing data is that it capture wrong regions, especially along the Doppler frequency lag, where noise-like artifacts are always present.

As discussed above, for any short-time non-stationary signal, its auto-terms in ambiguity domain reside significantly inside the regions $|\phi| \leq \pi/4$ and $3\pi/4 \leq \phi \leq 5\pi/4$. Thus, the AOK should only carry out the optimization in these regions in order to lower the probability of choosing an incorrect area. The optimization problem in [20, 79] is amended as follows:

$$\begin{aligned} \max_{\sigma(q\Delta\phi)} \quad & \sum_{g=0}^{G-1}\sum_{q=0}^{Q-1}(g\Delta_r)\left|A_{ss}(n;g,q)C(n;g,q)\right|^2 \\ \text{subject to} \quad & C(n;r,q) = e^{-\frac{(g\Delta r)^2}{2\sigma^2(q\Delta\phi)}} \\ & \sum_{q=0}^{N-1}\sigma^2(q) \leq vo \\ & |q\Delta\phi| \leq \pi/4 \;\&\; 3\pi/4 \leq q\Delta\phi \leq 5\pi/4 \end{aligned} \tag{6.30}$$

where vo is the kernel volume. A set of reasonable upper and lower bounds for vo is given by:

$$1 \leq vo \leq 5. \tag{6.31}$$

Similarly, the current-time slice of the TFR is computed as one slice (at time n only) of the two-dimensional Fourier transform of the STAF-kernel product, expressed as follows:

$$TFR(n,k) = \sum_{p}\sum_{b} A(n;p,b)C(n;p,b)e^{-j2\pi np/N_w}e^{-j2\pi bk/N_w}. \tag{6.32}$$

6.5 Fast Implementation

Although straightforward conceptually, direct implementation of the chirp-based kernel would be quite expensive computationally. Fortunately, there are a number of tricks which can be applied to reduce the computation load required for this algorithm.

6.5.1 STAF computation

Fast implementation of the STAF depends on an asymmetrical form of the AF, for three reasons:

1. It supports a causal implementation suitable for on-line computation.
2. It allows computationally efficient recursive implementation.
3. It supports full-rate sampling in the correlation lag variable, which provides an alias-free distribution for signals sampled at or above the Nyquist rate.

For a signal $s(n)$ of length N, the positive asymmetric ambiguity function is given by [90]:

$$\chi(p,b) = \sum_{n=0}^{N-1} s(n)s(n-b)e^{j2\pi pn}. \tag{6.33}$$

With each lag $b \in [-N, N]$, range value of time is:

$$\begin{cases} b \le n \le N & b \ge 0 \\ 0 \le n \le N+b & b < 0. \end{cases} \tag{6.34}$$

We can say that the symmetric AF $A_s(p,b)$ can be easily obtained from the asymmetric AF by multiplying it with a factor $e^{j2\pi pb/2}$. The two-dimensional Fourier transform of the asymmetric AF is called the Rihaczek distribution. Assume that we have a signal composed of a sinusoid and a chirp, the symmetric AF, asymmetric AF and their two-dimensional Fourier transform are shown in Fig. 6.10.

Recursion allows fast implementation of the STAF. In order to facilitate recursive computation, the symmetric AF in (6.27) is amended to obtain asymmetric STAF

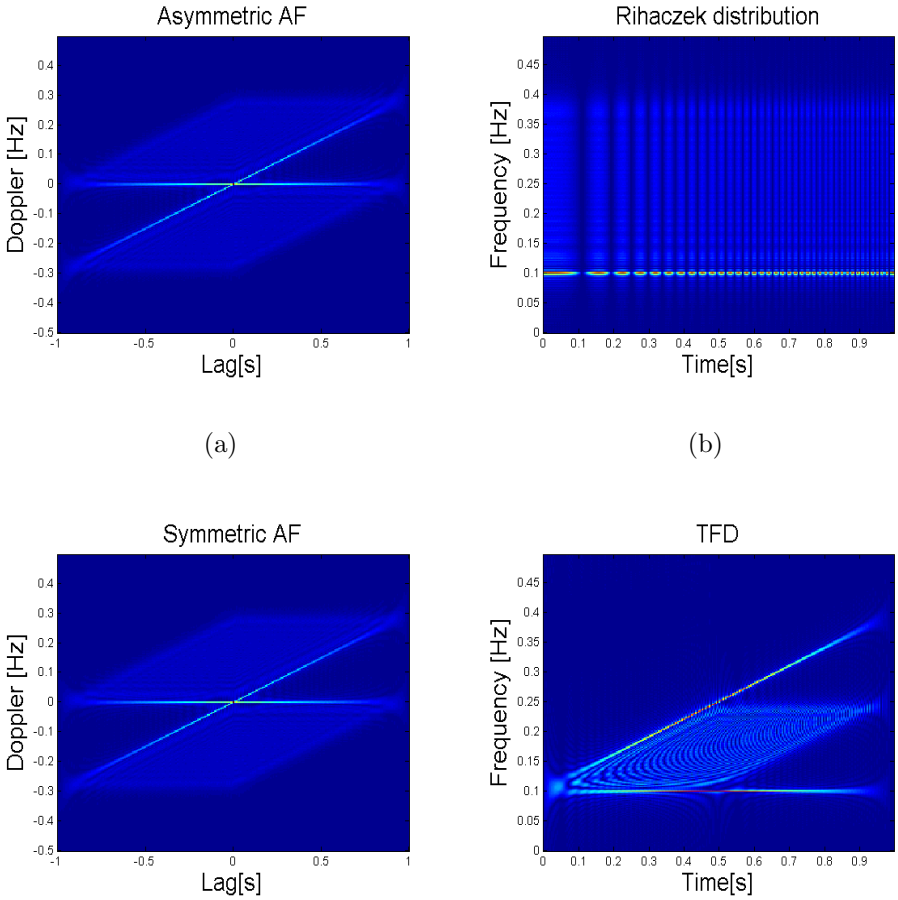


Figure 6.10: (a) Asymmetric AF; (b) Rihaczek distribution (2D Fourier transform of the asymmetric AF); (c) Symmetric AF; (d) TFD.

by replacing $u = n + N_w/2 - b/2 - v$. The asymmetric STAF is expressed as:

$$\chi(n;p,b) = \sum_{v} s^*(n+N_w-v-b)w^*(-v+N_w-b)s(n+N_w-v)w(-v+N_w)e^{-j2\pi vp/N_w}. \tag{6.35}$$

The range of v needs to satisfy:

$$\begin{aligned} & \begin{cases} n - N_w/2 + 1 \le u - b/2 \le n - N_w/2 \\ n - N_w/2 + 1 \le u + b/2 \le n - N_w/2 \end{cases} \\ \Longleftrightarrow\ & \begin{cases} n - N_w/2 + 1 \le n + N_w/2 - b - v \le n - N_w/2 \\ n - N_w/2 + 1 \le n + N_w/2 - v \le n - N_w/2 \end{cases} \\ \Longleftrightarrow\ & 0 \le v \le N_w - b - 1. \end{aligned} \tag{6.36}$$

For simplicity, we consider the special case of a rectangular window, and replacing $n_s = n + N_w$ then the asymmetrical STAF in (6.35) becomes:

$$\chi(n;p,b) = \sum_{v=0}^{N_w-b-1} s^*(n_s - v - b)s(n_s - v)e^{-j2\pi vp/N_w}. \tag{6.37}$$

According to [57, 79, 91, 92, 93], the asymmetric STAF $\chi(n;p,b)$ can be obtained by recursion as follows:

$$\begin{aligned} \chi(n_s;p,b) &= \gamma e^{-j2\pi p}\chi(n_s-1;p,b) + s(n_s)s^*(n_s-b) - \gamma^{N_w-b}e^{-2\pi p(N_w-b)} \\ & \quad s(n_s-(N_w-k))s^*(n_s-b-(N_w-b)) \\ &= \gamma e^{-j2\pi p}\chi(n_s-1;p,b) + R(n_s,b) - \gamma^{N_w-b}e^{-2\pi p(N_w-b)}R(n_s-(N_w-b),b), \end{aligned} \tag{6.38}$$

where γ is a positive real number just less than one. $R_d(n_s,b)$ is the asymmetric instantaneous correlation function:

$$R_d(n_s,b) = s(n_s)s^*(n_s-b). \tag{6.39}$$

Normally, we insert "$N_w - 1$" zeros in front of the signal. In this way, the initial value for $\chi(n_s-1;p,b)$ is a zero matrix of dimension $[N_w \mathrm{x} N_w]$ and the initial values for n and n_s are $N_w/2$ and N_w. The first asymmetric instantaneous correlation function is calculated with first one value of the signal. The relationship between the asymmetric and the symmetric STAF is as follows:

$$A(n;p,b) = \chi(n;p,b)e^{j2\pi p(n+N_w-b/2)/N_w}. \tag{6.40}$$

We calculate the asymmetrical STAF from (6.38) and then use (6.40) to set the symmetrical STAF. Once the discrete instantaneous asymmetric correlation function is computed, it can then be reused extensively, thus making the cost per STAF in (6.38) only two complex multiplications and two complex additions. Furthermore, only STAF samples corresponding to positive lags need to be computed or stored, because negative lag values can be obtained from the positive lags using the symmetry relation:

$$A(n;p,b) = A^*(n;-p,-b). \tag{6.41}$$

An additional benefit of this scheme is that half-sample lag values are avoided, thus preventing the need for interpolation or oversampling to obtain unaliased TFRs.

6.5.2 TFR time-slice computation

The current-time slice of the TFR is computed as one slice (at time n only) of the two-dimensional Fourier transform of the symmetrical STAF. Therefore, it is not necessary to compute the frequency content at every time instant inside the window. An efficient way to obtain the frequency slice at time n only is as follows:

$$\begin{aligned} D(n,k) &= \sum_b \sum_p A(n;p,b) C(n;p,b) e^{-j2\pi bk/N_w} e^{-j2\pi pn/N_w} \\ &= \sum_b \sum_p \chi(n;p,b) e^{j2\pi p(n+N_w-b/2)/N_w} C(n;p,b) e^{-j2\pi bk/N_w} e^{-j2\pi pn/N_w} \\ &= \sum_b \left(\sum_p \chi(n;p,b) e^{j2\pi p(N_w-b/2)/N_w} C(n;p,b) \right) e^{-j2\pi bk/N_w}. \end{aligned} \tag{6.42}$$

By using the asymmetric AF, the Fourier transform over the Doppler variable is replaced by a simple inner sum and thus the computational load is smaller. After summing the product over the Doppler variable, we get a vector changing with lag variable. Its one-dimensional Fourier transform gives only one frequency slice.

6.6 Simulation Results

This section evaluates the performance of the proposed RIDs, windowed chirp-based kernel and chirp-based adaptive optimal kernel, with various types of non-stationary

signals. The signals are sampled at the Nyquist rate, and then randomly shortened to create the incomplete data to be processed. The two representatives of signal-independent and signal-dependent kernels, Choi-Williams and the AOK, are simulated with the same signals to get a visual comparison with our methods. Notice that all methods will be applied on sliding windowed signals. The resulting images are normalized and transferred to the energy version to display. A parameter of concentration level ζ is used to access the accuracy of the resulting TFR. ζ is the ratio of the sum of pixel magnitude along the actual instantaneous frequency of the signals with respect to the rest of the TF values. So, the higher ζ, the more accurate the TF approximation. It is shown that the TFR based on chirp-based kernel design provides improved TF estimations when compared to the conventional RIDs. In all plots, the frequency axis is normalized with respect to the sampling frequency F_s.

1) Example 1: The first example considers a signal composed of a chirp and a sinusoid, which is expressed below:

$$\begin{aligned} s(n) = \exp\left\{ j2\pi[(0.1F_s)\frac{n}{F_s} + (0.3F_s)\frac{n^2}{2F_s^2}] \right\} + \\ \exp\left\{ j2\pi[(0.1F_s)\frac{n}{F_s}] \right\} + v(n), \end{aligned} \tag{6.43}$$

with the sampling frequency $F_s = 256$Hz. The signal's length is one second, or $N = F_s$, and $n = 0, ..., N-1$. The signal is corrupted by white Gaussian noise $v(n)$ with the signal-to-noise ratio (SNR) set to 30dB. A rectangular window of length $N_w = 64$ is used. We randomly remove 50% of the signal samples to have the compressed observations.

Fig. 6.11 shows the time-frequency signature of the full signal obtained by the proposed approach as well as other methods for comparison. Fig. 6.11(a) shows the windowed Choi-Williams distribution using full data. This method calculates the STAF first, then builds the Choi-Williams kernel with a predefined window length and obtains the TFRs by the two-dimensional FFT. It can be seen that with the fixed Choi-Williams kernel, cross-terms still show their strong existence in the TF domain. The concentration level $\zeta = 2.84$. The windowed chirp-based kernel gives an improved TFR with a higher $\zeta = 5.91$. It is reasonable since the chirp-based kernel only considers half of the ambiguity plane where the auto-terms reside. Under the full data condition, the two adaptive optimal kernels get good TF estimations.

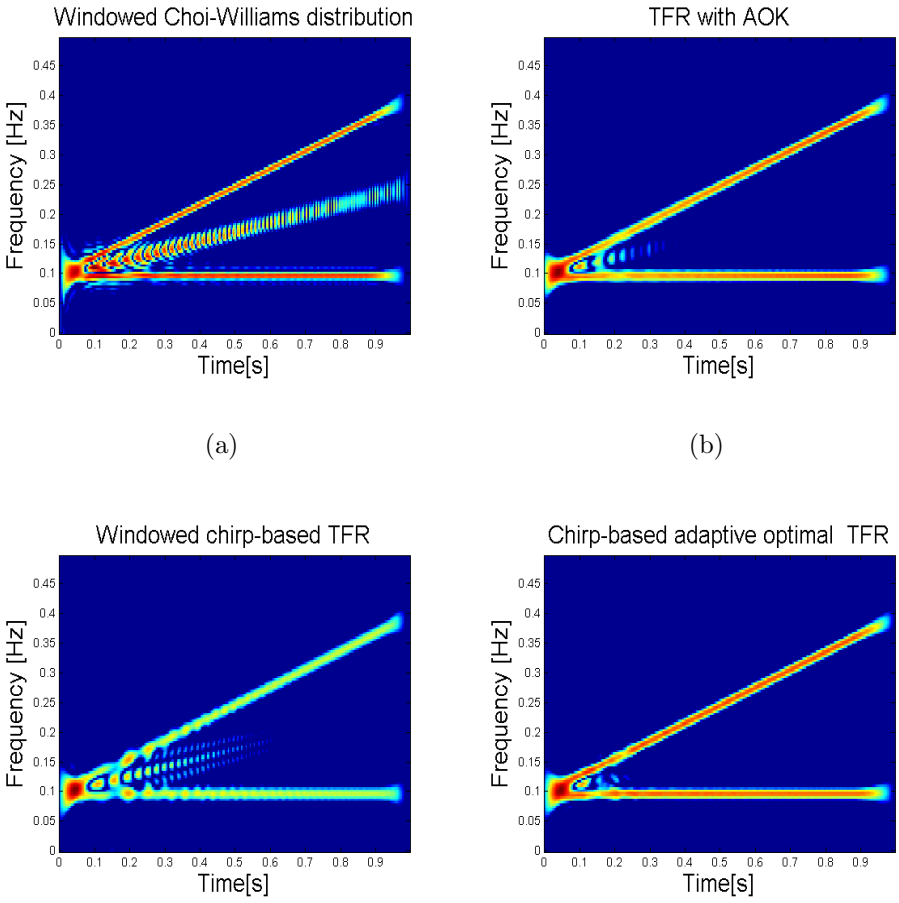


Figure 6.11: (a)Windowed Choi-Williams distribution; (b) TFR obtained using the AOK; (c) Windowed chirp-based TFR; (d) Chirp-based adaptive optimal TFR of the full signal in (6.43).

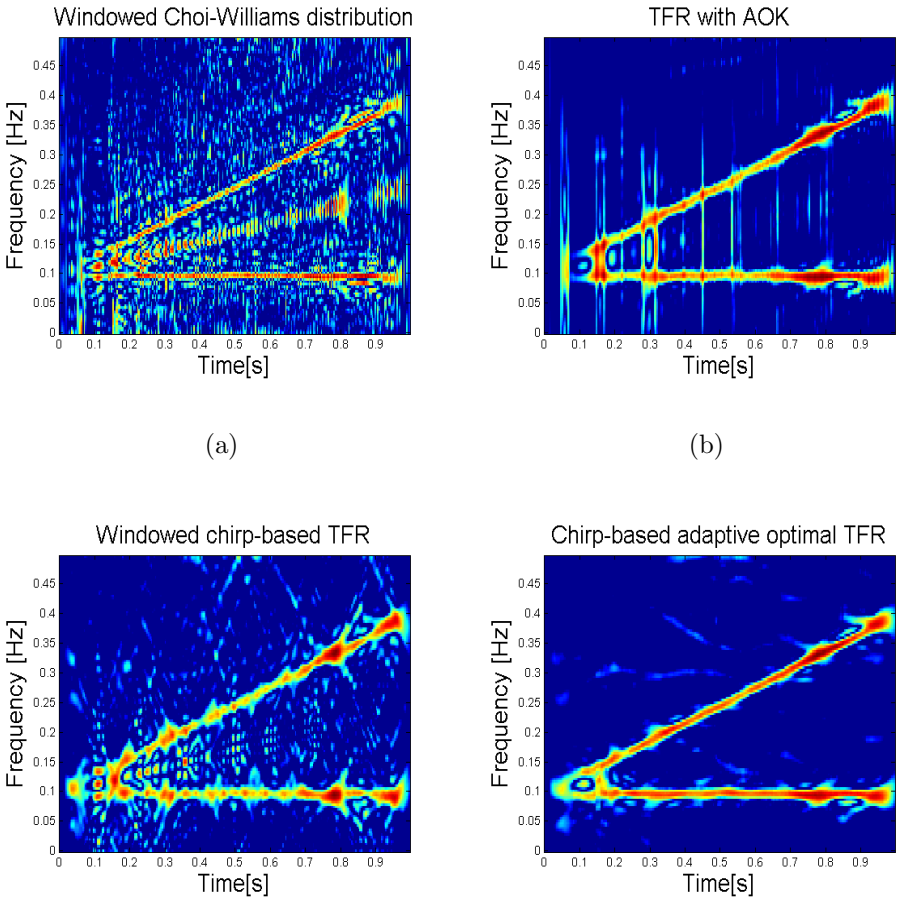


Figure 6.12: (a)Windowed Choi-Williams distribution; (b) TFR obtained using the AOK; (c) Windowed chirp-based TFR; (d) Chirp-based adaptive optimal TFR of the signal (6.43) when 50% data is missing.

The conventional AOK has $\zeta = 20.31$ while the chirp-based adaptive optimal TFR has $\zeta = 24.1$.

To get a comparison with the full data case, Fig. 6.12 shows the signal's instantaneous frequency obtained by the same techniques when 50% of data is missing. Missing samples introduce noise-like artifact which clutters both the time and the frequency domains, as evident in Fig. 6.12(a). The traditional signal-independent Choi-Williams kernel allows through not only part of the cross-terms but also artifacts, especially ones along the Doppler axis. Thus, the TF signature is corrupted with noise-like effect and cross-terms. The concentration levels in the case of missing data are obviously lower than those in the case of full data. The concentration level of the TFR obtained by the fixed Choi-Williams kernel is only $\zeta = 0.73$. The fixed windowed chirp-based kernel gives a better performance as seen in Fig. 6.12(c) because it suppresses more cross-terms and artifacts. In particular, it removes all the artifact along the Doppler axis. The concentration level is $\zeta = 1.74$. Fig. 6.12(b) depicts many vertical lines in the TFR obtained by the AOK. These lines are impulses caused by components captured along the Doppler axis. The dependent-kernel AOK performs better than the two independent-kernels with $\zeta = 3.39$. By removing the artifacts along the Doppler axis, the chirp-based adaptive optimal kernel achieves the most reliable result among the four methods with $\zeta = 6$.

2) Example 2: The second example observes a multi-component signal as follows:

$$\begin{aligned} s(n) = &\exp\left\{j(0.15F_s)\cos(2\pi\frac{n}{F_s} + \pi) + j2\pi(0.25F_s)\frac{n}{F_s}\right\} \\ &+ \exp\left\{j(0.15F_s)\cos(2\pi\frac{n}{F_s}) + j2\pi(0.25F_s)\frac{n}{F_s}\right\} + v(n). \end{aligned} \tag{6.44}$$

Similarly, we set $SNR = 30$dB, $F_s = 256$ Hz, $N = 256$, $n = 0, ..., N-1$. A rectangular window of length $N_w = 64$ is used. The signal is also randomly shortened by 50%. The signal TF signature is also obtained in both the full data case and the missing data case by four similar methods: the windowed Choi-Williams kernel,the chirp-based kernel, the AOK, the chirp-based adaptive optimal kernel. The simulation results in Fig. 6.13 and Fig. 6.14 show that both fixed and adaptive kernels based on the chirp give a better performance than the traditional ones.

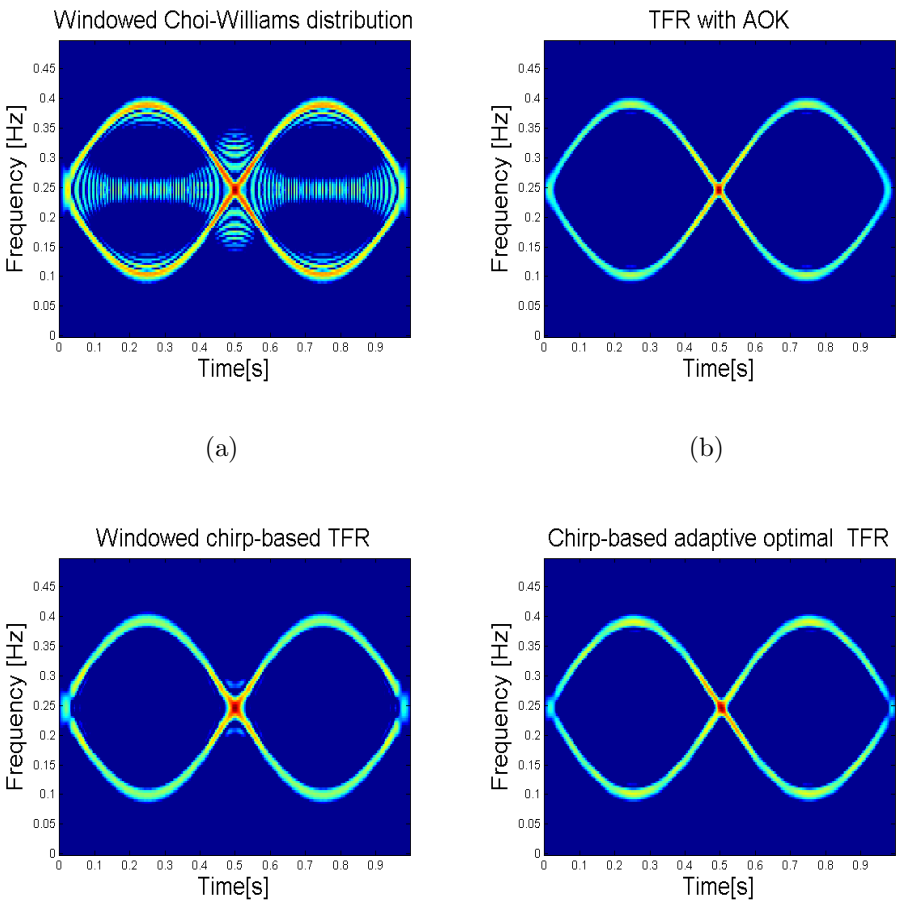


Figure 6.13: (a) Windowed Choi-Williams distribution; (b) TFR obtained using AOK; (c) Windowed chirp-based TFR; (d) Chirp-based adaptive optimal TFR of the full signal in (6.44).

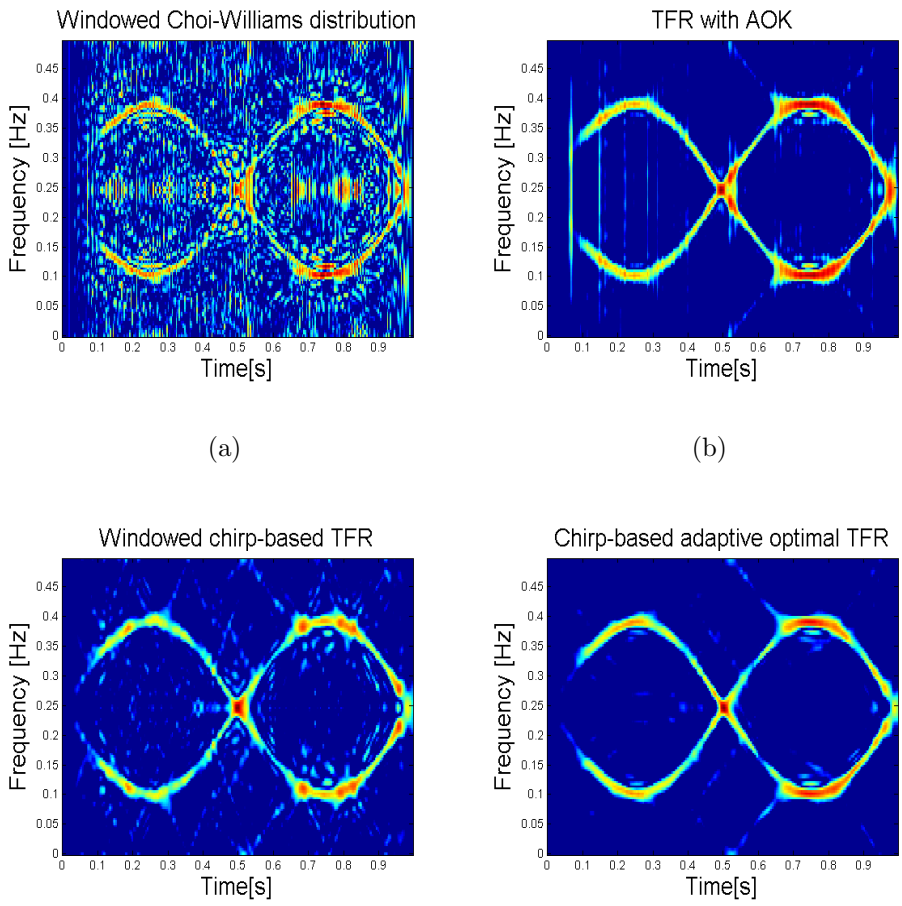


Figure 6.14: (a) Windowed Choi-Williams distribution; (b) TFR obtained using the AOK; (c) Windowed chirp-based TFR; (d) Chirp-based adaptive optimal TFR of the signal (6.44) when 50% data is missing.

It is evident that the windowed Choi-Williams distribution is severely influenced by the cross-terms and the noise-like effect caused by the missing samples, which can be seen in Fig. 6.13(a) and Fig. 6.14(a). The concentration level when we have all the data is $\zeta = 1.85$, which is larger than $\zeta = 0.52$ obtained when 50% of the data is absent. The fixed chirp-based kernel gives superior results compared with the fixed Choi-William kernel with $\zeta = 11.11$ with full data and $\zeta = 2.48$ with incomplete data. Similarly with example 1, the AOK works well under the full data condition with $\zeta = 12.65$. In the case of incomplete data, the TFR shows some vertical lines as the result of wrong areas captured in the ambiguity domain especially along the Doppler axis. The concentration level gets lower at $\zeta = 3.5$. The TF estimations get improved when chirp-based optimal kernel is used with $\zeta = 13$ for full data, and $\zeta = 5.8$ for incomplete data. It is because the optimization is performed only in regions where the auto-terms reside, and it excludes the regions covering the Doppler axis.

3) Example 3: In the third example, we observe a signal composed of three components as follows:

$$\begin{aligned} s(n) = &\exp\left\{j(0.15F_s)\cos(2\pi\frac{n}{F_s}+\pi) + j2\pi(0.25F_s)\frac{n}{F_s}\right\} \\ &+ \exp\left\{j2\pi[(0.1F_s)\frac{n}{F_s} + (0.2F_s)\frac{n^2}{2F_s^2}]\right\} \\ &+ \exp\left\{j2\pi[(0.2F_s)\frac{n}{F_s} + (0.2F_s)\frac{n^2}{2F_s^2}]\right\} + v(n). \end{aligned} \tag{6.45}$$

Similar parameter settings are used. Simulations are also carried out with both full data and missing samples. The windowed Choi-Williams kernel and the AOK are also deployed to get a comparison between the two proposed methods, which are based on the chirp property in the ambiguity domain. The results are plotted in Fig. 6.15 and Fig. 6.16. The TFRs obtained by the four methods (the windowed Choi-Williams kernel, the AOK, the windowed chirp-based TFR and the chirp-based adaptive optimal kernel) get worse when we use incomplete data because of the noise-like effect. The windowed Choi-William kernel gives the worst performance as it lets much of cross-terms and missing sample artifacts go through. The concentration levels in the full data case and the incomplete data case are 0.5 and 1.2, respectively. The fixed windowed chirp-based kernel offers more accurate TF estimations with

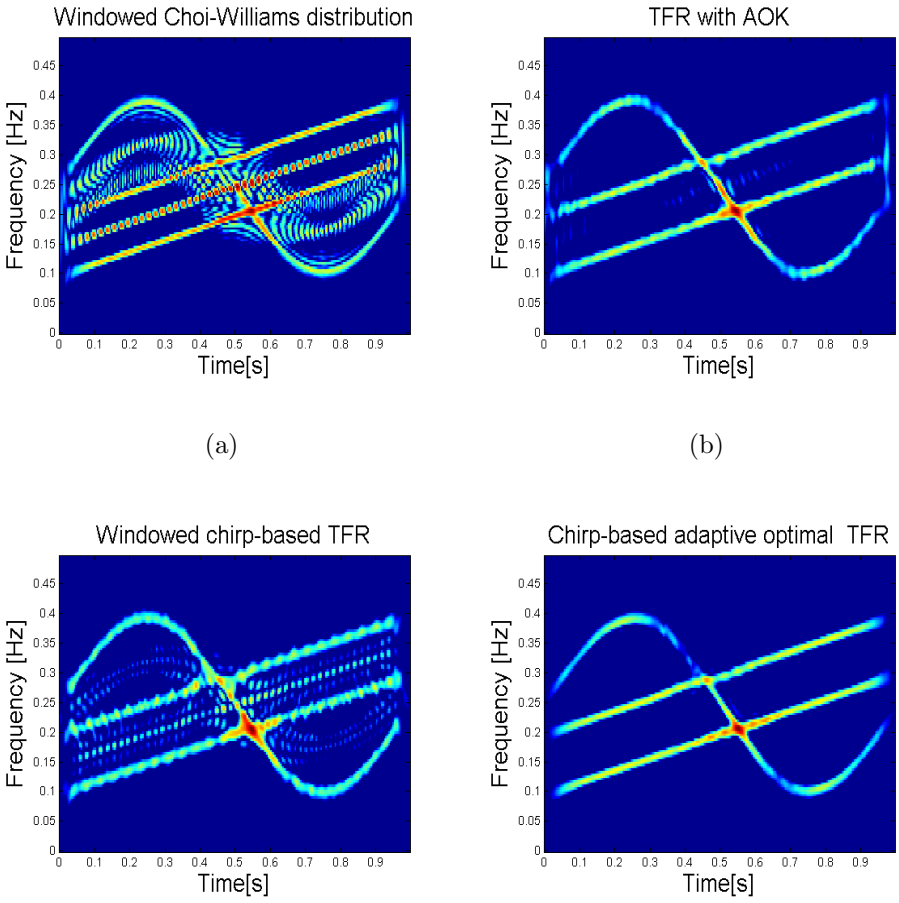


Figure 6.15: (a) Windowed Choi-Williams distribution; (b) TFR obtained using the AOK; (c) Windowed chirp-based TFR; (d) Chirp-based adaptive optimal TFR of the full signal in (6.45).

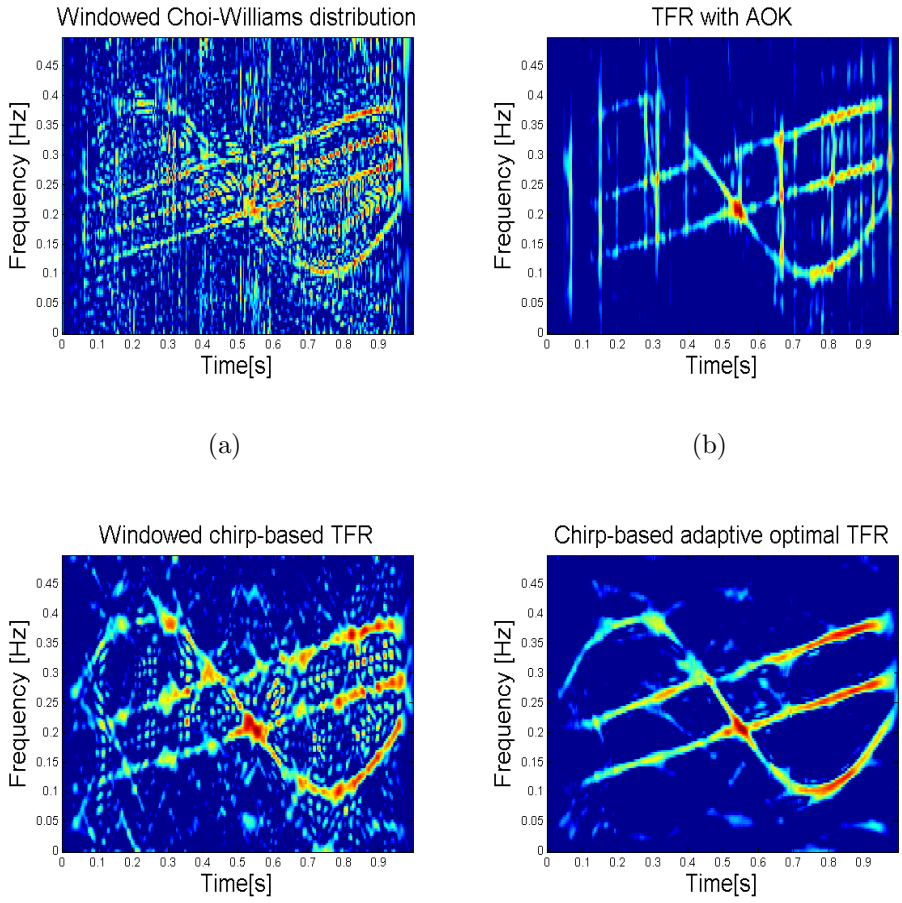


Figure 6.16: (a) Windowed Choi-Williams distribution; (b) TFR obtained using the AOK; (c) Windowed chirp-based TFR; (d) Chirp-based adaptive optimal TFR of the signal (6.45) when 50% data is missing.

$\zeta = 1.4$ and $\zeta = 3.7$. The signal-dependent kernels outperform the fixed kernels. The AOK achieves concentration levels in the missing and full data conditions of $\zeta = 1.64$ and $\zeta = 15.6$, respectively. The chirp-based adaptive optimal kernel give the best TFRs with the highest concentration level, $\zeta = 5.2$ in the case of limited data and $\zeta = 41.24$ in case of full data.

6.7 Conclusion

This chapter has introduced novel methods of designing signal-independent and signal-dependent kernels in the ambiguity domain. Similar to the AOK, they operate on windowed signals. The frequency slice at the middle point of the window is obtained by a two-dimensional Fourier transform of the STAF and the kernel product. The proposed methods also give superior results when compared with the traditional kernels both in the case of complete and incomplete data. It is because the kernels remove half of the ambiguity plane where the signals' auto-term do not reside. Especially, the removed region includes the Doppler axis, where the noise-like artifacts always appear. Therefore, these artifacts are largely mitigated. The proposed signal-independent kernel builds a mask with the predefined window length, and this mask is used in the whole TF estimation procedure. The chirp-based adaptive kernel, on the another hand, calculates a mask for each signal segment. This method can not only track the local change of the signal but also combat the adverse effects of compressed data.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This thesis has provided an overview of conventional as well as modern TFDs like the short time Fourier transform, the Wigner-Ville distribution, the reduced interference Cohen's class and the sparse TF reconstruction from full and windowed signals in both the time and ambiguity domains. However, each method has its own drawbacks, which leads to undesired results. So here we have introduced two new TFDs which can bring about more reliable estimations of the TF signature of non-stationary signals.

The first approach performs sparse reconstruction from windowed data in the time domain with a novel chirp dictionary. As in many situations, the non-stationary signal frequency law is more properly approximated by piece-wise second-order polynomials than fixed frequency sinusoids. Thus, the chirp dictionary, instead of the sinusoidal dictionary, is better suited for sparse reconstruction problems dealing with these signals. The chirp dictionary is built in two ways. The first way includes all possible chirps which can appear in any signal segment. The second way is also composed of all chirps, but they are formed from sinusoids which are rotated through all eligible angles by the fractional Fourier transform. Although the dictionary construction procedure is different, the two ways actually lead to the same results.

The chirp dictionary approach has been proven to provide more reliable TF estimations compared with the sinusoidal approach. The chirp method nevertheless,

deploys a very large dimension measurement dictionary. Since there are two parameters to be estimated (i.e. the chirp rate and the initial frequency), the dictionary dimension can be equal to the square of the dimension when using the sinusoid atom. This very large atom set leads to a much heavier computational burden and a longer calculation time. Therefore, in order to obtain good TF estimation at low computational complexity, chirp dictionary simplification methods are needed. Therefore, we introduce two approaches which reduce the chirp dictionary dimension, thus lowering the calculation load. In the first one, we estimate the chirp rate through the DTFT of the bilinear product at a certain time lag. The initial frequency is solved in the time domain, with a lower dimensional dictionary than the computationally complex full chirp atom. In the second approach, the fractional Fourier transform (FRFT) is used to obtain an initial frequency for each chirp-rate. This leads to a much simpler chirp atom set.

The second method in this thesis introduces a novel fixed and signal-dependent kernel in the ambiguity domain, which can efficiently remove cross-term interference and partially combat missing sample effects without using compressive sensing techniques. These kernels are applied on windowed signals to facilitate online implementation, or processing long signals, since any non-stationary signal segment can be approximated by a sum of chirps. Additionally, the chirps' auto-terms always reside in only half of the ambiguity domain which does not cover the Doppler axis. By removing the areas where the auto-terms do not lie, part of interference and artifacts are mitigated. Moreover, the analysis of the distribution of the artifacts shows that artifacts always appear along the Doppler axis. Filtering out the region along the Doppler axis results in obviously improved TFRs.

7.2 Future Work

During the past decade, the technology associated with "seeing" through walls has witnessed a growing interest. The objectives of sensing through walls and inside an enclosed structure range from determining building layouts, discerning the nature of activities inside the buildings and imaging building interiors to detect, identify, classify and track the whereabouts of humans and moving objects. These attributes are highly desirable for a range of organizations, including police, fire and rescue

personnel, first responders and defence forces [1, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103]. To achieve these objectives, electromagnetic waves are considered very effective due to their ability to penetrate man-made building materials and image targets behind opaque structures. Through-the-wall radar imaging is a multifaceted technology. It requires a blending of several disciplines in the field of electrical engineering, especially those that involve signal, array and image processing as well as radars, antennas and electromagnetic waves. Suggestions for future work focus on two important topics:

(a) Target range estimation and location of static and moving targets using compressive sensing techniques and time-frequency analysis [53].

Radar imaging achieves its range and cross-range resolutions of targets through the exploitation of, respectively, the bandwidth of the signals and the aperture of an array of sensors used in the operation. The aperture of the array can be achieved using a collocated physical array or a synthetic array. Depending on the application requirement and the system feasibility, an antenna array can be designed in one or two dimensions, or as a set of sparsely distributed sensors. The data collection at different sensors may be processed coherently or non-coherently.

For localization and tracking of a small number of targets in a sparse scene, it becomes possible to exploit compressive sensing in data collection and processing. Such techniques include the reduction of the number of time samples, thinned arrays, and thinned frequency steps.

Targets in motion demonstrate Doppler signatures, which, through an appropriate exploitation of time-frequency analysis can be useful for signal enhancement and target discrimination. Targets with different Doppler signatures may be separated in the time-frequency domain, and their localization becomes feasible, even with only a small number of frequency and array antennas.

The target range estimation deals with the high amount of multi-path in indoor environments. A variety of multi-path exploitation approaches, both under conventional and sparse reconstruction frameworks, have been recently proposed in the literature. However, these methods require prior knowledge

of the interior layout of the building to eliminate ghost targets (accumulation of unwanted energy at incorrect target location) and provide enhanced quality. In practice, this information may not be available. Thus, significant more research in this field needs to be carried out.

(b) Characterization and classification of human motion through radar micro-Doppler signatures.

Recent research and developments for in home radar monitoring have shown real promise of the technology in detecting normal and abnormal gross-motor activities of humans inside their residences and at private homes. It has been shown that typical interior walls do not significant alter the radar time-frequency signature of a fall or normal walking, and the radar signal return is only slightly weakened by wall penetration. Thus, radar can be used as a remote fall-monitoring [104]. As the focus shifts to radar system integration, operations and installation, attention is being paid to:

- The required number of radar units for a given residence.
- The radar unit is designed to be elderly or patient specific in the sense that the detection and classification algorithms of a fall and other gross-motor human activities do not use features extracted from training data associated with a large population. Rather, these algorithm are self-tuned to the person being monitoring.
- Classification methods, which can be based on physical features like extreme frequency dynamic range [105] or PCA-based classification [106].

Appendix A

A.1 Digital Computation of The Fractional Fourier Transform (FRFT)

A.1.1 Compactness in the time domain, frequency domain and Wigner space

A function will be referred to as compact if its support is so [49]. The support of a function is the subset of the real axis in which the function is not equal to zero. In other words, a function is compact if and only if its nonzero values are confined in a finite interval. We all know that a function in time and its Fourier transform cannot be both compact except the case that they are identically zero. However, in practice, we can work with a finite time interval and a finite bandwidth. Thus, the difference between the mathematical idealization and the real world is usually not a problem when we work with signals of large time-bandwidth product.

The time-bandwidth product here is defined as the product of the temporal extent of the signal and its doubled-sided bandwidth. Let μ denote the time-bandwidth product. For example, a signal has the sampling frequency F_s and the number of samples is $N = F_s$. It means that the signal is confined to the time interval of $[0, 1]$, and frequency interval of $[-F_s/2, F_s/2]$. The time-bandwidth product of this signal is $\mu = (2\text{x}F_s/2)\text{x}1 = F_s$.

To calculate the discrete FRFT, we will scale frequency, time and other fractional domains to the one variable. We will consider the coordinate of the new domain x. Assume that the signal in the time domain is approximately confined to the

interval $[-\Delta t/2, \Delta t/2]$ and its frequency representation is confined to the interval $[-\Delta f/2, \Delta f/2]$. The time-bandwidth product is $\mu = \Delta t \Delta f$. In the new domain, we use the scale coordinates $x = t/sl$, $y = fsl$, where sl is scaling parameter. With these new coordinates, the time and frequency domain representation will be confined to intervals of length $\Delta t/sl$ and $\Delta f sl$. Choose $sl = \sqrt{\Delta t/\Delta f}$, both intervals are now equal to $\sqrt{\Delta t \Delta f} = \sqrt{\mu}$, which we denote by Δx ($\Delta x = \sqrt{\mu}$). In other words, the new domain x is confined in the interval $[-\Delta x/2, \Delta x/2]$, and thus the sample space is $1/\Delta x = 1/\sqrt{N}$. Its orthogonal domain y is also confined to $[-\Delta x/2, \Delta x/2]$.

From now on, we will assume that this dimensional normalization has been performed and that the coordinates appearing in the definition of the FRFT, the Wigner distribution, etc., are all dimensionless quantities.

If a representation of the signal in a certain domain, x for instance, is confined to a certain interval around the origin, then the Wigner distribution will be confined to an infinite strip perpendicular to x defined by that interval. Thus, assuming that the representation of the signal in all domains is confined to an interval of length Δx around the origin, this is equivalent to assuming that the Wigner distribution is confined within a circle of diameter Δx. With this, we mean that a sufficient large percentage of the energy of the signal is contained within that circle.

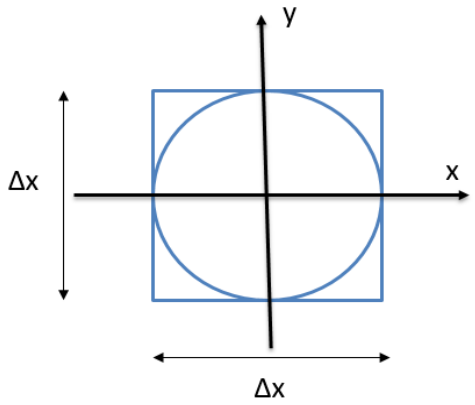


Figure A.1: Circular support of a signal in TF space. The square bounding the circle is also shown.

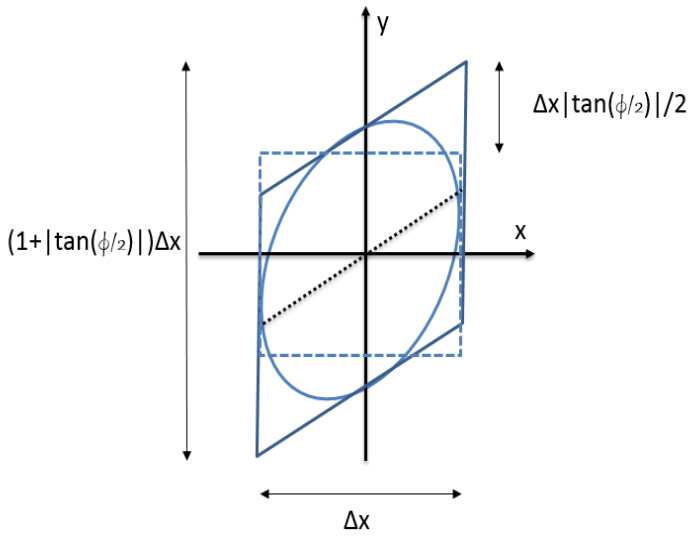


Figure A.2: The effect of multiplication of the time domain representation of the signal by a chirp function with chirp rate of $\tan(\phi/2)$.

A.1.2 Effect of chirp multiplication and convolution on compact signals

A signal is referred to as compact if its Wigner distribution is compact, that is, if it is confined to a circle of some finite diameter Δx (see Fig. A.1).

Now, let us assume that the time domain representation of our signal $f(x)$ is multiplied by the chirp function $\exp[j2\pi \tan(\phi/2)\frac{x^2}{2}]$, where $\pi/2 \leq \phi \leq \pi/2$. Mathematically, this corresponds to convolving the original Wigner distribution with the Wigner distribution of the chirp function $\delta[y - \tan(\phi/2)x]$. Its effect is to shear the Wigner distribution in the y direction, as shown in Fig. A.2.

We see that the support of the Wigner distribution remains compact. Whatever percentage of the signal was confined to the bandwidth Δx, now it is confined to the bandwidth $\Delta x(1 + |\tan(\phi/2)|)$. If $\phi = \pi/2$, the new bandwidth will be double the old one. So, multiplying a function $f(x)$ with the chirp function $\exp[j2\pi \tan(\phi/2)\frac{x^2}{2}]$ results in a convolution of their Fourier transforms, resulting in an overall double-sided bandwidth. Thus, we have to sample the modulated signal at interval of $\frac{1}{2\Delta x}$.

If the samples of $f(x)$ are spaced at $\frac{1}{\Delta x}$, we can interpolate these and then multiply by the samples of the chirp function to obtain the desired number of samples.

A similar argument holds for convolution of a function with a chirp. But this time, shearing of the Wigner distribution in the x-direction is involved.

A.1.3 Methods of computing the continuous Fractional Fourier Transform

Let $\{\mathbb{F}f\}(x)$ denote the Fourier transform of $f(x')$. The a^{th} order fractional Fourier transform $\{\mathbb{F}^a f\}(x)$ of the function $f(x')$ may be defined for $0 \leq |a| \leq 2$ as,

$$\begin{aligned} \{\mathbb{F}^a f\}(x) &= \int_{-\infty}^{\infty} B_a(x, x') f(x') dx' \\ B_a(x, x') &= A_\phi \exp\left[j\pi(x^2 \cot\phi - 2xx' \csc\phi + x'^2 \cot\phi)\right] \\ A_\phi &= \frac{\exp(-j\pi \text{sgn}(\sin\phi)/4 + j\phi/2)}{|\sin\phi|^{1/2}}, \end{aligned} \tag{A.1}$$

where

$$\phi = \frac{a\pi}{2}. \tag{A.2}$$

This definition is easily extended outside the interval $[-2, 2]$ by remembering that $\mathbb{F}^{4i}$ is the identity operator for any integer i and that the FRFT is additive in index, that is, $\mathbb{F}^{a_1}\mathbb{F}^{a_2} = \mathbb{F}^{a_1+a_2}$.

(A.1) can rarely be evaluated analytically; therefore, numerical integration is called for. Numerical integration of quadratic exponentials, which often appear in diffraction theory, requires a very large number of samples if conventional methods are to be employed, due to the rapid oscillations of the kernel. The problem is particularly pronounced when a is close to 0 or ± 2. If we assume both the function and its Fourier transform to be confined to a finite interval, then we can solve the problem as follows [107, 108, 109]:

- If $a \in [0.5, 1.5]$ or $a \in [2.5, 3.5]$, we evaluate the integral directly.
- If $a \in [-0.5, 0.5]$ or $a \in [1.5, 2.5]$, we use the property $\mathbb{F}^a = \mathbb{F}^1\mathbb{F}^{a-1}$. Notice that $\mathbb{F}^{a-1}$ can be evaluated directly.

Another method of evaluating (A.1) would be to use the spectral decomposition of the kernel [110, 110, 111]. This is equivalent to first expanding the function $f(x)$ as $\sum_{n=0}^{\infty} c_n \psi_n(x)$, multiplying the expansion coefficients c_n, respectively, with $e^{-ian\pi/2}$, and summing the components.

Although both ways of evaluating the fractional Fourier transform may be expected to give accurate results, we do not consider them further since they take $O(\mu^2)$, where μ is time-bandwidth product.

A.1.4 Digital computation of the fractional Fourier transform

The FRFT is a member of a more general class of transformations that are sometimes called linear canonical transformations or quadratic-phase transforms [112]. Members of this class of transformations can be broken down into a succession of simpler operations, such as chirp multiplication, chirp convolution, scaling and ordinary transformation.

The defining equation for the fractional Fourier transform in (A.1) can be put in the form:

$$\{\mathbb{F}^a f\}(x) = A_\phi e^{j\pi\gamma x^2} \int_{-\infty}^{\infty} e^{-i2\pi\kappa x x'} [e^{i\pi\gamma x'^2} f(x')] dx', \tag{A.3}$$

where $\gamma = \cot\phi$ and $\kappa = \csc\phi$, x' and x are the variables of the function before and after performing FRFT. We are again assuming that the Wigner distribution of $f(x')$ is zero outside a circle of diameter Δx centered around the origin. Under this assumption, and by limiting the order a to the interval $0.5 \leq |a| \leq 1.5$, the amount of vertical shear in the Wigner space resulting from the chirp modulation is bounded by $\Delta x/2$ (one side). Then, the modulated function $e^{i\pi\gamma x'^2} f(x')$ is band-limited to Δx (one side) in the frequency domain. It means that we have to sample $f(x')$ at a sampling frequency $2\Delta x$ to get rid of the alias. Let n_f be the discrete variable for x' variable. Thus, $e^{i\pi\gamma x'^2} f(x')$ can be represented by Shannon's interpolation formula,

$$e^{i\pi\gamma x'^2} f(x') = \sum_{n_f=-\mu}^{\mu} e^{i\pi\gamma(\frac{n_f}{2\Delta x})} f(\frac{n_f}{2\Delta x}) \operatorname{sinc}\left(2\Delta x(x' - \frac{n_f}{2\Delta x})\right), \tag{A.4}$$

where μ is the time-bandwidth product, $\mu = (\Delta x)^2$. The summation goes from $-\mu$ to μ since $f(x')$ is assumed to be zero outside $[-\Delta x/2, \Delta x/2]$. From (A.3) and (A.4), we have:

$$\begin{aligned}\{\mathbb{F}^a f\}(x) = A_\phi e^{j\pi\gamma x^2} \sum_{n_f=-\mu}^{\mu} e^{i\pi\gamma(\frac{n_f}{2\Delta x})} f(\frac{n_f}{2\Delta x}) \\ \int_{-\infty}^{\infty} e^{-i2\pi\kappa x x'} \text{sinc}\left(2\Delta x(x' - \frac{n_f}{2\Delta x})\right) dx' \\ = \frac{A_\phi}{2\Delta x} e^{j\pi\gamma x^2} \sum_{n_f=-\mu}^{\mu} e^{i\pi\gamma(\frac{n_f}{2\Delta x})} f(\frac{n_f}{2\Delta x}) e^{-i2\pi\kappa x(n_f/2\Delta x)}.\end{aligned} \tag{A.5}$$

Then, the samples of the transformed function are obtained as:

$$\{\mathbb{F}^a f\}(\frac{u_f}{2\Delta x}) = \frac{A_\phi}{2\Delta x} \sum_{n_f=-\mu}^{\mu} \exp\left(j\pi\gamma(\frac{u_f}{2\Delta x})^2 - i2\pi\kappa\frac{u_f n_f}{(2\Delta x)^2} + i\pi\gamma(\frac{n_f}{2\Delta x})\right) f(\frac{n_f}{2\Delta x}). \tag{A.6}$$

Direct computation of this form would require $O(\mu^2)$ multiplications. An $O(\mu \log \mu)$ can be obtained as follows. We put (A.6) into the following form after some algebraic manipulations:

$$\{\mathbb{F}^a f\}(\frac{u_f}{2\Delta x}) = \frac{A_\phi}{2\Delta x} e^{i\pi(\alpha-\beta)(\frac{u_f}{2\Delta x})^2} \sum_{n_f=-\mu}^{\mu} e^{i\pi\kappa(\frac{u_f - n_f}{2\Delta x})^2} e^{i\pi(\alpha-\beta)(\frac{n_f}{2\Delta x})^2} f(\frac{n_f}{2\Delta x}). \tag{A.7}$$

It can be seen that the calculation includes a chirp multiplication followed by a chirp convolution followed by another chirp multiplication.

By assuming appropriate x2 interpolation and decimation, the procedure starts with μ samples spaced at $1/\Delta x$, which uniquely characterizes the function $f(x')$ and returns the same for $\{\mathbb{F}^a f\}(x)$. Let $\boldsymbol{f}$, $\boldsymbol{f}_a$ denote column vectors with μ samples of $f(x')$ and $\{\mathbb{F}^a f\}(x)$. The overall procedure can be represented as:

$$\begin{aligned}\boldsymbol{f}_a &= \mathbb{F}^a \boldsymbol{f}, \\ \mathbb{F}^a &= \mathbf{D}\mathbf{K}_a\mathbf{J},\end{aligned} \tag{A.8}$$

where $\mathbf{D}$ and $\mathbf{J}$ are matrices representing the decimation and interpolation operator. $\mathbf{K}_a$ is expressed as:

$$K_a(u_f, n_f) = \frac{A_\phi}{2\Delta x} e^{j\pi\left\{\alpha(u_f/2\Delta x)^2 - 2\kappa[n_f u_f/(2\Delta x)^2] + \alpha(n_f/2\Delta x)^2\right\}}, \tag{A.9}$$

for $|n_f| \leq \mu$ and $|u_f| \leq \mu$. The convolution can be computed in $O(\mu \log \mu)$ by using the FFT. The output samples can be obtained by a final chirp multiplication. Hence, the overall complexity is $O(\mu \log \mu)$.

We has assumed $0.5 \leq |a| \leq 1.5$ in deriving this algorithm. Using the index additivity property of the FRFT, we can extend this range to all values of a easily. For instance, for the range $0 \leq a \leq 0.5$, we use the formula:

$$\mathbb{F}^a = \mathbb{F}^{a-1}\mathbb{F}^1. \tag{A.10}$$

A.2 Relationship between FRFT and WVD

TF representations are signal transformations that describe how the spectral content of a time-varying signal or a non-stationary signal varies with time. One of the most popular example is the Wigner Ville distribution (WVD) which is defined as:

$$WD_s(t,f) = \int s(t+\tau/2)s^*(t-\tau/2)e^{-j2\pi f\tau}d\tau. \tag{A.11}$$

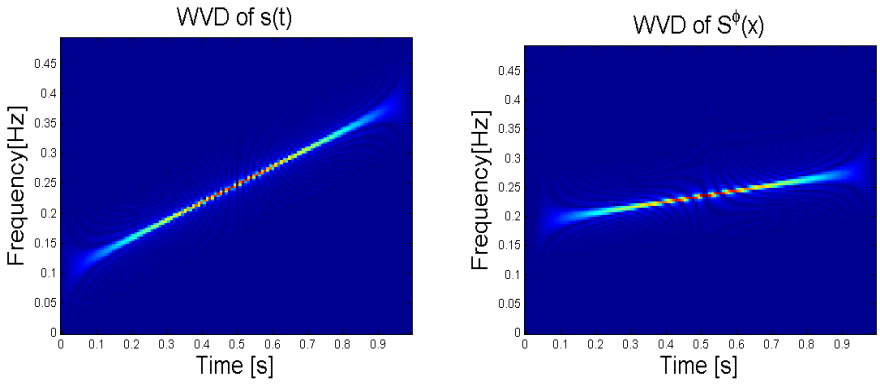
The relationship between the FRFT and the WVD is expressed as:

$$WD_{S^\phi}(x,y) = WD_s(t\cos(\phi) - f\sin(\phi), t\sin(\phi) + f\cos(\phi)), \tag{A.12}$$

where $WD_{S^\phi}(x,y)$ denotes the WD of the FRFT signal $S^\phi(x)$, and $WD_s(t,f)$ is the WD of the original time domain signal $s(t)$. According to (A.12), $WD_{S^\phi}(x,y)$ is simply a ϕ rotation of $WD_s(t,f)$. To illustrate, a chirp $s(t)$ with the initial frequency $\beta = 0.1F_s$ and the chirp rate $\alpha = 0.3F_s$ is employed. We plot the WVD of the time-domain signal and the WVD of the FRFT signal $S^\phi(x)$, with $\phi = \pi/16$. The figure is shown in Fig. A.3. This can be explained mathematically in [113]. The time domain signal can be obtained from the 2D WVD by the following two ways:

$$\begin{aligned} &\int WD_s(t,f)df = |s(t)|^2 \\ &\int WD_s(t,f)\exp(j2\pi ft)df = (\mathbb{F}^{-1}WD_s)(t) = s(t)s(0)^*, \end{aligned} \tag{A.13}$$

where $s(0)^* = \int WD_s(0,f)df$ is a constant phase factor. From (A.13), the relationship between the FRFT and the WVD can be explained in the following manner:



(a) (b)

Figure A.3: Relationship between FRFT and WVD

$$
\begin{aligned}
&s(t) \rightarrow WD_s(t,f) (\text{Wigner-Ville distribution}) \\
&WD_s(t,f) \rightarrow WD_s(t\cos(\phi) - f\sin(\phi), t\sin(\phi) + f\cos(\phi)) = WD_{S^\phi}(x,y) \\
&(\text{Rotate } WD_s(t,f) \text{ by an angle } \phi \text{ clockwise}) \\
&WD_{S^\phi}(x,y) \rightarrow (\mathbb{F}^{-1} WD_{S^\phi})(x) = S^\phi(x). \\
&(\text{Apply (A.13) to return the original form of the signal})
\end{aligned} \tag{A.14}
$$

In short, the WVD of the FRFT with angle ϕ of a signal $s(t)$ is simply the WVD of the signal itself rotated by an angle ϕ.

References

[1] M. Amin, *Compressive sensing for urban radar*. CRC Press, 2014.

[2] P. Flandrin and P. Borgnat, "Time-frequency energy distributions meet compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 2974–2982, 2010.

[3] B. Jokanović, M. Amin, and S. Stanković, "Instantaneous frequency and time-frequency signature estimation using compressive sensing," in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2013, pp. 871 418–871 418.

[4] M. G. Amin, Y. D. Zhang, and B. Jokanovic, "Time-frequency signature reconstruction from random observations using multiple measurement vectors," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 345–349.

[5] L. Cohen, *Time-frequency analysis*. Prentice hall, 1995, vol. 1, no. 995,299.

[6] A. Eden, "The search for Christian Doppler," in *The Search for Christian Doppler*. Springer, 1992, pp. 1–4.

[7] T. P. Gill, *The Doppler Effect: The Theory and Applications*, 1965.

[8] M. Saeed, M. Villarroel, A. T. Reisner, G. Clifford, L.-W. Lehman, G. Moody, T. Heldt, T. H. Kyaw, B. Moody, and R. G. Mark, "Multiparameter intelligent monitoring in intensive care ii (mimic-ii): A public-access intensive care unit database," *Critical care medicine*, vol. 39, no. 5, p. 952, 2011.

[9] I. Christov, G. Gómez-Herrero, V. Krasteva, I. Jekova, A. Gotchev, and K. Egiazarian, "Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification," *Medical engineering & physics*, vol. 28, no. 9, pp. 876–887, 2006.

[10] Y. D. Shirman, *Computer simulation of aerial target radar scattering, recognition, detection, and tracking*. Artech House, 2002.

[11] D. Gabor, "Theory of communication. Part 1: The analysis of information," *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, vol. 93, no. 26, pp. 429–441, 1946.

[12] S. Qian and D. Chen, *Joint time-frequency analysis: Methods and applications*. Prentice-Hall, Inc., 1996.

[13] Y. T. Nguyen, M. G. Amin, M. Ghogho, and D. McLernon, "Time-frequency signature sparse reconstruction using chirp dictionary," in *SPIE Sensing Technology+ Applications*. International Society for Optics and Photonics, 2015, pp. 94 840H–94 840H.

[14] ——, "Local sparse reconstructions of Doppler frequency using chirp atoms."

[15] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971. The Royal Society, 1998, pp. 903–995.

[16] S. Olhede and A. Walden, "The Hilbert spectrum via wavelet projections," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 460, no. 2044. The Royal Society, 2004, pp. 955–975.

[17] W. C. Van Etten, *Introduction to Random signals and Noise*. John Wiley & Sons, 2006.

[18] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 6, pp. 862–871, 1989.

[19] B. Boashash, *Time-frequency signal analysis*. Prentice Hall, 1991.

[20] R. G. Baraniuk and D. L. Jones, "A radially-Gaussian, signal-dependent time-frequency representation," in *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*. IEEE, 1991, pp. 3181–3184.

[21] J. Bobin, J.-L. Starck, and R. Ottensamer, "Compressed sensing in astronomy," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 718–726, 2008.

[22] G. M. Davis, S. G. Mallat, and Z. Zhang, "Adaptive time-frequency decompositions," *Optical engineering*, vol. 33, no. 7, pp. 2183–2191, 1994.

[23] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE, 1993, pp. 40–44.

[24] J. Tropp, A. C. Gilbert *et al.*, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.

[25] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Transactions on Information theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[26] C. Moler, "Magic reconstruction: Compressed sensing," *Mathworks News & Notes*, 2010.

[27] M. Amin, B. Jokanovic, and T. Dogaru, "Reconstruction of locally frequency sparse nonstationary signals from random samples," in *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*. IEEE, 2014, pp. 1771–1775.

[28] Y. D. Zhang, M. G. Amin, and B. Himed, "Reduced interference time-frequency representations and sparse reconstruction of undersampled data," in *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*. IEEE, 2013, pp. 1–5.

[29] S. Stanković, I. Orović, and E. Sejdić, *Multimedia Signals and Systems: Basic and Advanced Algorithms for Signal Processing*. Springer, 2015.

[30] G. Li, R. Zhang, W. Rao, and X. Wang, "Separation of multiple micro-Doppler components via parametric sparse recovery," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*. IEEE, 2013, pp. 2978–2981.

[31] G. Li and P. K. Varshney, "Micro-Doppler parameter estimation via parametric sparse representation and pruned orthogonal matching pursuit," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 7, no. 12, pp. 4937–4948, 2014.

[32] M. G. Amin and W. Sun, "A novel interference suppression scheme for global navigation satellite systems using antenna array," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 5, pp. 999–1012, 2005.

[33] C. Clemente, A. Balleri, K. Woodbridge, and J. J. Soraghan, "Developments in target micro-Doppler signatures analysis: radar imaging, ultrasound and through-the-wall radar," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, pp. 1–18, 2013.

[34] W. C. Knight, R. G. Pridham, and S. M. Kay, "Digital signal processing for sonar," *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1451–1506, 1981.

[35] T. Misaridis and J. A. Jensen, "Use of modulated excitation signals in medical ultrasound. Part i: Basic concepts and expected benefits," *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on*, vol. 52, no. 2, pp. 177–191, 2005.

[36] V. C. Chen, "Analysis of radar micro-Doppler with time-frequency transform," in *Statistical Signal and Array Processing, 2000. Proceedings of the Tenth IEEE Workshop on.* IEEE, 2000, pp. 463–466.

[37] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using a support vector machine," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, no. 5, pp. 1328–1337, 2009.

[38] C.-P. Lai and R. M. Narayanan, "Through-wall imaging and characterization of human activity using ultrawideband (UWB) random noise radar," in *Proc. of SPIE Vol*, vol. 5778, 2005, p. 187.

[39] P. Setlur, M. Amin, and T. Thayaparan, "Micro-Doppler signal estimation for vibrating and rotating targets," in *Signal Processing and Its Applications, 2005. Proceedings of the Eighth International Symposium on*, vol. 2. IEEE, 2005, pp. 639–642.

[40] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3397–3415, 1993.

[41] S. Ghofrani, M. G. Amin, and Y. D. Zhang, "High-resolution direction finding of non-stationary signals using matching pursuit," *Signal Processing*, vol. 93, no. 12, pp. 3466–3478, 2013.

[42] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Transactions on Signal Processing*, vol. 40, no. 2, pp. 402–412, 1992.

[43] C. Luo and J. H. McClellan, "Discrete random sampling theory," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on.* IEEE, 2013, pp. 5430–5434.

[44] L. Stankovic, I. Orovic, S. Stankovic, and M. Amin, "Compressive sensing based separation of nonstationary and stationary signals overlapping in time-frequency," *Signal Processing, IEEE Transactions on*, vol. 61, no. 18, pp. 4562–4572, 2013.

[45] E. J. Candes and T. Tao, "Decoding by linear programming," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203–4215, 2005.

[46] L. B. Almeida, "The fractional Fourier transform and time-frequency representations," *IEEE Transactions on signal processing*, vol. 42, no. 11, pp. 3084–3091, 1994.

[47] B. Boashash, *Time-frequency signal analysis and processing: A comprehensive reference*. Academic Press, 2015.

[48] H. M. Ozaktas, D. Mendlovic, L. Onural, and B. Barshan, "Convolution, filtering, and multiplexing in fractional Fourier domains and their relation to chirp and wavelet transforms," *JOSA A*, vol. 11, no. 2, pp. 547–559, 1994.

[49] H. M. Ozaktas, O. Arikan, M. A. Kutay, and G. Bozdagi, "Digital computation of the fractional fourier transform," *IEEE Transactions on signal processing*, vol. 44, no. 9, pp. 2141–2150, 1996.

[50] L. Stanković, T. Alieva, and M. J. Bastiaans, "Time-frequency signal analysis based on the windowed fractional Fourier transform," *Signal Processing*, vol. 83, no. 11, pp. 2459–2468, 2003.

[51] L. Applebaum, S. D. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," *Applied and Computational Harmonic Analysis*, vol. 26, no. 2, pp. 283–290, 2009.

[52] R. Calderbank, S. Howard, and S. Jafarpour, "Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 358–374, 2010.

[53] M. G. Amin, *Through-the-wall radar imaging*. CRC press, 2011.

[54] V. Chen, *The micro-Doppler effect in radar*. Artech House, 2011.

[55] V. C. Chen and H. Ling, *Time-frequency transforms for radar imaging and signal analysis*. Artech House, 2001.

[56] F. Ahmad, M. G. Amin, and P. Setlur, "Through-the-wall target localization using dual-frequency cw radars," in *Defense and Security Symposium*. International Society for Optics and Photonics, 2006, pp. 62 010H–62 010H.

[57] M. G. Amin, "Time-frequency spectrum analysis and estimation for nonstationary random-processes," in *Time-Frequency Signal Analysis*, 1992, pp. 208–232.

[58] G. Matz and F. Hlawatsch, *Time-varying power spectra of nonstationary random processes*. na, 2003.

[59] C. Law, G. Jones, D. Backer, W. Barott, G. Bower, C. Gutierrez-Kraybill, P. Williams, and D. Werthimer, "Millisecond imaging of radio transients with the pocket correlator," *The Astrophysical Journal*, vol. 742, no. 1, p. 12, 2011.

[60] J. M. Peha, "Wireless communications and coexistence for smart environments," *Personal Communications, IEEE*, vol. 7, no. 5, pp. 66–68, 2000.

[61] L. F. Chaparro, E. Sejdic, A. Can, O. A. Alkishriwo, S. Senay, and A. Akan, "Asynchronous representation and processing of nonstationary signals: A time-frequency framework," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 42–52, 2013.

[62] J. F. Gemmeke, H. Van Hamme, B. Cranen, and L. Boves, "Compressive sensing for missing data imputation in noise robust speech recognition," *IEEE Journal of selected topics in Signal Processing*, vol. 4, no. 2, pp. 272–287, 2010.

[63] B. Jokanovic, M. G. Amin, and Y. D. Zhang, "Reducing noise in the time-frequency representation using sparsity promoting kernel design," in *SPIE*, 2014.

[64] B. Jokanovic, M. G. Amin, Y. D. Zhang, and F. Ahmad, "Time-frequency kernel design for sparse joint-variable signal representations," in *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*. IEEE, 2014, pp. 2100–2104.

[65] G. E. Pfander and H. Rauhut, "Sparsity in time-frequency representations," *Journal of Fourier Analysis and Applications*, vol. 16, no. 2, pp. 233–260, 2010.

[66] L. Stankovic, S. Stankovic, and M. Amin, "Missing samples analysis in signals for applications to L-estimation and compressive sensing," *Signal Processing*, vol. 94, pp. 401–408, 2014.

[67] Y. D. Zhang, L. Guo, Q. Wu, and M. G. Amin, "Multi-sensor kernel design for time-frequency analysis of sparsely sampled nonstationary signals," in *Radar Conference (RadarCon), 2015 IEEE.* IEEE, 2015, pp. 0896–0900.

[68] Y. T. Nguyen, M. G. Amin, M. Ghogho, and D. McLernon, "Local sparse reconstructions of doppler frequency using chirp atoms," in *Radar Conference (RadarCon), 2015 IEEE.* IEEE, 2015, pp. 1280–1284.

[69] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[70] C. Capus, Y. Rzhanov, and L. Linnett, "The analysis of multiple linear chirp signals," 2000.

[71] C. Capus and K. Brown, "Fractional Fourier transform of the Gaussian and fractional domain signal support," *IEE Proceedings-Vision, Image and Signal Processing*, vol. 150, no. 2, pp. 99–106, 2003.

[72] L. Qi, R. Tao, S. Zhou, and Y. Wang, "Detection and parameter estimation of multicomponent LFM signal based on the fractional Fourier transform," *Science in China series F: information sciences*, vol. 47, no. 2, p. 184, 2004.

[73] A. Serbes and O. Aldimashki, "A fast and accurate chirp rate estimation algorithm based on the fractional Fourier transform," in *Signal Processing Conference (EUSIPCO), 2017 25th European.* IEEE, 2017, pp. 1105–1109.

[74] D. Fourer, F. Auger, K. Czarnecki, S. Meignen, and P. Flandrin, "Chirp rate and instantaneous frequency estimation: application to recursive vertical synchrosqueezing," *IEEE Signal Processing Letters*, vol. 24, no. 11, pp. 1724–1728, 2017.

[75] F. Miwakeichi, E. Martınez-Montes, P. A. Valdés-Sosa, N. Nishiyama, H. Mizuhara, and Y. Yamaguchi, "Decomposing ECG data into space–time–frequency components using parallel factor analysis," *NeuroImage*, vol. 22, no. 3, pp. 1035–1045, 2004.

[76] M. Stridh, L. Sörnmo, C. J. Meurling, and S. B. Olsson, "Sequential characterization of atrial tachyarrhythmias based on ECG time-frequency analysis," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 1, pp. 100–114, 2004.

[77] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *Signal Processing, IEEE transactions on*, vol. 52, no. 7, pp. 1830–1847, 2004.

[78] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *Information Theory, IEEE Transactions on*, vol. 36, no. 5, pp. 961–1005, 1990.

[79] D. L. Jones and R. G. Baraniuk, "An adaptive optimal-kernel time-frequency representation," *IEEE Transactions on Signal Processing*, vol. 43, no. 10, pp. 2361–2371, 1995.

[80] M. G. Amin, *Through-the-wall radar imaging*. CRC press, 2016.

[81] B. Jokanovic and M. Amin, "Reduced interference sparse time-frequency distributions for compressed observations," *IEEE transactions on signal processing*, vol. 63, no. 24, pp. 6698–6709, 2015.

[82] M. G. Amin, B. Jokanovic, Y. D. Zhang, and F. Ahmad, "A sparsity-perspective to quadratic time-frequency distributions," *Digital Signal Processing*, 2015.

[83] A. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Transactions on Information Theory*, vol. 14, no. 3, pp. 369–374, 1968.

[84] R. Abeysekera and B. Boashash, "Time-frequency domain features of ECG signals: Their application in p wave detection using the cross Wigner-Ville

distribution," in *Acoustics, Speech, and Signal Processing, 1989. ICASSP-89., 1989 International Conference on.* IEEE, 1989, pp. 1524–1527.

[85] M. G. Amin, "Spectral decomposition of time-frequency distribution kernels," *IEEE Transactions on Signal Processing*, vol. 42, no. 5, pp. 1156–1165, 1994.

[86] M. Bayram and R. G. Baraniuk, "Multiple window time-frequency and time-scale analysis," in *Proc. SPIE Int. Soc. Opt. Eng*, 1996.

[87] G. Fraser and B. Boashash, "Multiple window spectrogram and time-frequency distributions," in *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, vol. 4. IEEE, 1994, pp. IV–293.

[88] B. Jokanovic, M. Amin, and T. Dogaru, "Time-frequency signal representations using interpolations in joint-variable domains," *Geoscience and Remote Sensing Letters, IEEE*, vol. 12, no. 1, pp. 204–208, 2015.

[89] J. W. Pitton, "Time-frequency spectrum estimation: An adaptive multitaper method," in *Time-Frequency and Time-Scale Analysis, 1998. Proceedings of the IEEE-SP International Symposium on.* IEEE, 1998, pp. 665–668.

[90] N. Levanon, "Radar principles," *New York, Wiley-Interscience, 1988, 320 p.*, 1988.

[91] M. Amin, "Computationally lag-invariant recursive spectrum estimators," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 12, pp. 1713–1724, 1987.

[92] M. G. Amin, "Recursion in Wigner distribution," in *32nd Annual Technical Symposium.* International Society for Optics and Photonics, 1988, pp. 221–231.

[93] M. Unser, "Recursion in short-time signal analysis," *Signal Processing*, vol. 5, no. 3, pp. 229–240, 1983.

[94] F. Ahmad, R. M. Narayanan, and D. Schreurs, "Application of radar to remote patient monitoring and elder-care," *IET Radar, Sonar & Navigation*, vol. 9, no. 2, pp. 115–115, 2015.

[95] F. Ahmad, A. Cetin, K. Ho, and J. Nelson, "Special section on signal processing for assisted living," *IEEE Sig. Process. Mag.*, vol. 33, no. 2, pp. 25–94, 2016.

[96] A. Gadde, M. G. Amin, Y. D. Zhang, and F. Ahmad, "Fall detection and classifications based on time-scale radar signal characteristics," in *Proc. SPIE*, vol. 9077, 2014, p. 907712.

[97] G. Gennarelli, I. Catapano, and F. Soldovieri, "Rf/microwave imaging of sparse targets in urban areas," *IEEE Antennas and Wireless Propagation Letters*, vol. 12, pp. 643–646, 2013.

[98] S. Kidera, T. Sakamoto, and T. Sato, "Extended imaging algorithm based on aperture synthesis with double-scattered waves for UWB radars," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 12, pp. 5128–5139, 2011.

[99] M. Leigsnering, F. Ahmad, M. G. Amin, and A. M. Zoubir, "Compressive sensing based specular multipath exploitation for through-the-wall radar imaging," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 6004–6008.

[100] ——, "Compressive sensing-based multipath exploitation for stationary and moving indoor target localization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1469–1483, 2015.

[101] L. Liu, M. Popescu, M. Skubic, M. Rantz, T. Yardibi, and P. Cuddihy, "Automatic fall detection based on Doppler radar motion signature," in *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*. IEEE, 2011, pp. 222–225.

[102] P. Setlur, M. Amin, and F. Ahmad, "Multipath model and exploitation in through-the-wall and urban radar sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 4021–4034, 2011.

[103] B. Y. Su, K. Ho, M. J. Rantz, and M. Skubic, "Doppler radar fall activity detection using the wavelet transform," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 3, pp. 865–875, 2015.

[104] M. Amin, F. Ahmad, and B. Jokanovic, "Personalized fall detection and classification through walls and in heavy indoor clutter," in *SPIE Defense+ Security*. International Society for Optics and Photonics, 2015, pp. 94 610H–94 610H.

[105] P. P. Markopoulos and F. Ahmad, "Indoor human motion classification by l1-norm subspaces of micro-Doppler signatures," in *Radar Conference (RadarConf), 2017 IEEE*. IEEE, 2017, pp. 1807–1810.

[106] B. Jokanovic, M. Amin, F. Ahmad, and B. Boashash, "Radar fall detection using principal component analysis," in *SPIE Defense+ Security*. International Society for Optics and Photonics, 2016, pp. 982 919–982 919.

[107] J. C. Wood and D. T. Barry, "Linear signal synthesis using the Radon-Wigner transform," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2105–2111, 1994.

[108] ——, "Radon transformation of time-frequency distributions for analysis of multicomponent signals," *IEEE Transactions on signal processing*, vol. 42, no. 11, pp. 3166–3177, 1994.

[109] ——, "Tomographic time-frequency analysis and its application toward time-varying filtering and adaptive kernel design for multicomponent linear-FM signals," *IEEE transactions on signal processing*, vol. 42, no. 8, pp. 2094–2104, 1994.

[110] H. M. Ozaktas and D. Mendlovic, "Fourier transforms of fractional order and their optical interpretation," *Optics Communications*, vol. 101, no. 3-4, pp. 163–169, 1993.

[111] ——, "Fractional Fourier transforms and their optical implementation. ii," *JOSA A*, vol. 10, no. 12, pp. 2522–2531, 1993.

[112] ——, "Fractional Fourier optics," *JOSA A*, vol. 12, no. 4, pp. 743–751, 1995.

[113] A. W. Lohmann and B. H. Soffer, "Relationships between the Radon-Wigner and fractional Fourier transforms," *JOSA A*, vol. 11, no. 6, pp. 1798–1801, 1994.