2011

Compressed sensing based cyclic feature spectrum sensing for cognitive radios

Yohannes Z. Tafesse
Michigan Technological University

COMPRESSED SENSING BASED CYCLIC FEATURE SPECTRUM SENSING FOR COGNITIVE RADIOS

by

Yohannes Z. Tafesse

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

(Electrical Engineering)

MICHIGAN TECHNOLOGICAL UNIVERSITY

2011

This thesis, “Compressed Sensing Based Cyclic Feature Spectrum Sensing for Cognitive Radios,” is hereby approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE IN ELECTRICAL ENGINEERING.

Department of Electrical and Computer Engineering

Signatures:

Thesis Advisor _____

Dr. Zhi (Gerry) Tian

Department Chair _____

Dr. Daniel R. Fuhrmann

Date _____

To my

parents

with love

Table of Contents

Page

List of Figures vii

List of Tables ix

Acknowledgement x

Abstract xi

Chapter

1 Introduction 1

1.1 Cognitive Radio 1

1.2 Spectrum Sensing in Cognitive Radio 3

1.2.1 Matched Filter Spectrum Detection 4

1.2.2 Energy Detector 5

1.2.3 Cyclostationary Feature Detection 5

1.3 Motivation 6

1.4 Organization of the thesis 9

2 Theoretical Background 10

2.1 Second-order cyclostationarity 10

2.2 Compressed Sensing 15

3 Compressive Cyclic Spectrum Sensing 19

3.1 Signal Model and Problem Statement 19

3.2 Recovery of Sparse Cyclic Spectrum 22

3.2.1 Vector Relationship 23

3.2.2 Problem Formulation 27

3.2.3 Implementation 28

3.3 Recovery of Power Spectrum for Stationary Signals 29

3.4 Comparison of Spectrum, PSD and Cyclic Spectrum Estimation Techniques 31

4 Cyclic-based Spectrum Occupancy Estimation 37

4.1 Multi-cycle GLRT Based Band-by-Band Approach 37

4.2 A Fast Algorithm for BPSK Signals 39

5 Simulations 43

5.1 Simulation Setup 43

5.2 Robustness to Rate Reduction 44

5.3 Robustness to Noise Uncertainty 47

5.4 Comparison of Spectrum Occupancy Estimation Techniques 49

5.5 The Effect of The Number of Blocks Used 50

6 Conclusion and Future Work 51

6.1 Conclusion 51

6.2 Future Work 52

References 54

Appendix

A Mapping Matrices 58

B Noise Covariance Matrix Estimation 60

List of Figures

1.1 Measured Spectrum Occupancy in Chicago and New York, [2] 2

2.1 Surface plot of the SCD magnitude for an AM signal 13

2.2 Surface plot of the SCF magnitude of WGN 14

2.3 Surface plot of the SCF magnitude of a BPSK signal 15

3.1 Cyclic spectrum of digital samples 20

3.2 Original and folded cyclic spectrums of a signal 24

3.3 Comparison of Spectrum, PSD and Cyclic Spectrum Estimation Techniques 36

4.1 Non-zero region of support for the cyclic spectrum of BPSK. 40

5.1 Normalized MSE of reconstruction versus compression ratio M/N. 45

5.2 Probability of detection at a constant false alarm ratio of 0.01. 46

5.3 ROC of the cyclic feature approach (“Cyclo”), PSD approach (“PSD”) and energy detection (“ED”). 48

5.4 ROC of the two spectrum occupancy estimation approaches. 49

5.5 ROC of proposed sensing method for different L. $M = 16$, $N = 32$. 50

List of Tables

3.1 35

Acknowledgement

I would like to express my deep-felt gratitude to my advisor, Dr. Zhi Tian of Electrical and Computer engineering department, for her guidance, support, encouragement, and most of all her continued motivation. Her door was always open and she always made time to address my questions and doubts.

I would like to thank my advisory committee members Dr. Daniel Fuhrmann and Dr. Richard Donovan for their feedback and support. I am very thankful to Dr. Donovan for giving me access to his computing machine which made my work much easier and saved me a lot of time. I would also like to thank Mr. Robert Emperley of Advanced Research Support, Michigan Tech University for his assistance in setting up the access to the computing machine and addressing all the problems that came up promptly. My thanks and appreciation also goes to all faculty and staff of electrical and computer engineering department of Michigan Tech for all they have done for me throughout the time of my study.

Most of all, I would like to thank my parents for their never stopping love and support which cannot be described in words and for putting my education ahead of all their plans. I am also very grateful to all my family members and relatives who provided their support at one time or another. I especially thank Belayhune Asrese for all his support in helping me pursue graduate school. Last but not least I want to thank my friends who made my stay here memorable.

Abstract

Spectrum sensing is currently one of the most challenging design problems in cognitive radio. A robust spectrum sensing technique is important in allowing implementation of a practical dynamic spectrum access in noisy and interference uncertain environments. In addition, it is desired to minimize the sensing time, while meeting the stringent cognitive radio application requirements. To cope with this challenge, cyclic spectrum sensing techniques have been proposed. However, such techniques require very high sampling rates in the wideband regime and thus are costly in hardware implementation and power consumption. In this thesis the concept of compressed sensing is applied to circumvent this problem by utilizing the sparsity of the two-dimensional cyclic spectrum. Compressive sampling is used to reduce the sampling rate and a recovery method is developed for reconstructing the sparse cyclic spectrum from the compressed samples. The reconstruction solution used, exploits the sparsity structure in the two-dimensional cyclic spectrum domain which is different from conventional compressed sensing techniques for vector-form sparse signals. The entire wideband cyclic spectrum is reconstructed from sub-Nyquist-rate samples for simultaneous detection of multiple signal sources. After the cyclic spectrum recovery two methods are proposed to make spectral occupancy decisions from the recovered cyclic spectrum: a band-by-band multi-cycle detector which works for all modulation schemes, and a fast and simple thresholding method that works for Binary Phase Shift Keying (BPSK) signals only. In addition a method for recovering the power spectrum of stationary signals is developed as a special case. Simulation results demonstrate that the proposed spectrum sensing algorithms can significantly reduce sampling rate without sacrificing performance. The robustness of the algorithms to the noise uncertainty of the wireless channel is also shown.

Chapter 1

Introduction

1.1 Cognitive Radio

The electromagnetic radio frequency (RF) spectrum is the most valuable resource of wireless communication systems. Driven by the ever increasing spectrum demand from different applications, the radio frequency spectrum has come to be a scarce resource. A closer investigation of this issue reveals that, even though there appears to be spectrum shortage, the radio frequency band is in fact underutilized. According to the Federal Communication Commission (FCC), currently almost all RF spectrum is allocated, but most of it is underutilized or not used at all. The spectrum usage survey conducted by the FCC reveals that utilization of the RF spectrum ranges from 15% to 85% depending on population density [1]. To reinforce this fact Figure 1.1 shows the results of spectrum occupancy measurements conducted in Chicago and New York cities [2]. According to the statistics, Chicago and New York in fact have two of the highest spectrum occupancies of the RF band ranging from 30MHz to 3GHz. In other regions this spectrum is even less utilized [3].

To alleviate this spectrum underutilization problem, cognitive radio (CR) proposed in [6], [7], has emerged as a promising solution. Cognitive radio is a notion that is based on dynamic spectrum sharing. In CR framework, a cognitive radio, usually referred to as the secondary user, searches for and uses unoccupied spectrum bands that are already licensed

to a primary user (PU). These unused frequency bands are known as spectrum holes. A definition of a spectrum hole as given in [4] is: "a band of frequencies assigned to a primary user, but at a particular time and specific geographic location, is not being utilized by that user". One promising application of the concept of CR is IEEE 802.22 standard for Wireless Regional Area Network (WRAN). The aim of this standard is developing specifications that allow CRs to utilize unused spectrum allocated to the Television Broadcast Service [8].

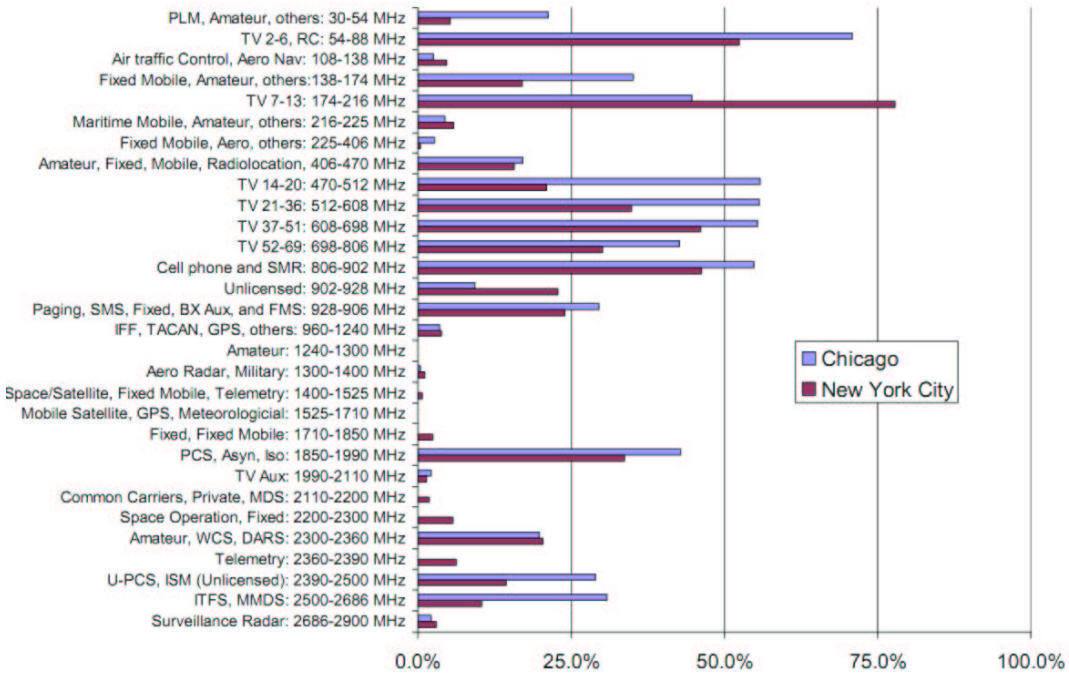


Figure 1.1: Measured Spectrum Occupancy in Chicago and New York, [2]

A cognitive radio is an intelligent wireless communication system that is capable of sensing its surrounding RF environment and adapting its transmission parameters including transmit power and frequency accordingly [4]. There are three fundamental tasks that need to be performed by a CR [4]. The first is radio-scene analysis, which involves spectrum sensing to detect spectrum holes. The second task is channel identification which includes

estimation of channel-state information (CSI) and channel capacity prediction. The third task is transmit-power control and dynamic spectrum management. Accordingly spectrum sensing is the foremost task that needs to be performed by a CR. To achieve good spectrum utilization and support the required quality of service, CRs need to sense a wide spectrum and utilize unoccupied bands while avoiding interference to any PU. To accomplish these tasks, efficient spectrum sensing is a key enabler [5]. The following section provides a brief introduction of spectrum sensing and reviews the main spectrum sensing approaches proposed in the literature.

1.2 Spectrum Sensing in Cognitive Radio

The RF environment of a CR can be classified in to three categories based on the power detected at a particular band [4].

1. Black spaces: These are spectrum bands occupied by a high power signal some of the time.
2. Gray spaces: These are bands that are partially occupied by low-power interferers.
3. White spaces: These are spectrum bands that are unoccupied and are free except for the presence of ambient noise.

Obviously white spaces are the best candidates to be used by unlicensed secondary users. Gray spectral bands also present a good opportunity for CRs as long as the interference from the CR to PUs stays below a certain threshold. The black spaces are to be avoided whenever they are being used by the licensed primary user. Simply put spectrum sensing in CR refers to the task of sensing the RF environment and categorizing RF bands in to one of

these groups. In general, spectrum sensing boils down to deciding the presence or absence of a PU signal in a given frequency band. The basic hypothesis model for such a detection problem is

$$x(t) = \begin{cases} w(t), & H_0 \\ hs(t) + w(t), & H_1 \end{cases} \tag{1.1}$$

where $w(t)$ is additive noise, $s(t)$ is the signal transmitted by the primary user, h is the channel gain and $x(t)$ is the signal received by the CR. H_1 is the alternative hypothesis that stands for the presence of a PU signal whereas H_0, the null hypothesis, states there is no licensed PU signal in the monitored band. In the literature there are three main approaches to this spectrum detection problem; matched filter detection, energy detection and cyclostationary feature detection. A brief review of each follows.

1.2.1 Matched Filter Spectrum Detection

When the signal of the primary user is known to the CR, the matched filter is the best detector for the presence of the PU signal. The matched filter is the optimal linear filter that maximizes the Signal to Noise ratio (SNR) at the receiver of the CR in the presence of additive Gaussian noise [5]. In time domain, this is equivalent to convolving the received signal with the time-reversed version of the PU transmitted signal. The biggest advantage of the matched filter is the short sensing time it requires to achieve a certain required probability of detection. However, in a dynamic radio environment this method is highly impractical since it requires a perfect knowledge of the PU signal. In most practical situations a CR may not even know the existing PUs let alone their signal waveforms.

1.2.2 Energy Detector

Unlike the matched filter, the energy detector does not require prior knowledge of PU signals and is easy to implement. Energy detector measures the energy of the received signal by squaring and integrating the output of a band pass filter of bandwidth W in a time interval T [5]. Then the output of the integrator, Y_E, is compared with a threshold η_E to decide the presence of a PU signal. Assuming a non-fading environment the probability of detection P_d and false alarm P_f are given by [9]:

$$P_d = PY_E > \eta_E|H_1 = Q_m(\sqrt{2\gamma}, \sqrt{\eta_E}) \tag{1.2}$$

$$P_f = PY_E > \eta_E|H_0 = \frac{\mathbf{\Gamma}(m, \eta_E/2)}{\mathbf{\Gamma}(m)} \tag{1.3}$$

Where γ is the SNR, $Q_m(.)$ is the generalized Marcum Q-function, and $\mathbf{\Gamma}(.)$ and $\mathbf{\Gamma}(.,.)$ are complete and incomplete gamma functions.

Owing to its easy implementation the energy detector is the most widely adopted spectrum sensing scheme in CR. However, energy detector suffers significantly from noise and interference uncertainty and does not provide any means of differentiating between different PU signals [11], [12].

1.2.3 Cyclostationary Feature Detection

Cyclostationary feature detection utilizes the cyclostationary property of communication signals that arises due to the periodicities of sine wave carriers, pulse trains, spreading codes, etc. The cyclic features of communication signals are exploited using spectral cor-

relation analysis as a method for spectrum sensing. The major advantage of this scheme is its ability to differentiate between noise and modulated signals. This is possible because unlike communication signals, noise is a stationary process with no correlation and does not give rise to cyclic features. As a result, cyclic feature detection is robust to noise and interference uncertainty problems. On the other hand this method is computationally complex and requires longer sensing time in comparison with the energy detector [14], [17]. Cyclic feature detection is the spectrum sensing approach used in this thesis and will be discussed in more detail in later chapters.

1.3 Motivation

Spectrum sensing is the foremost task that needs to be performed by cognitive radio networks to utilize spectrum holes efficiently. The traditional spectrum sensing techniques discussed in the previous chapter have major limitations for wideband spectrum sensing applications. The matched filter is optimal for signal detection under Gaussian noise, but requires prior knowledge of all the possible PU signals in the wide band, which is infeasible in practice. An energy detector, although popular because of its easy implementation, is highly susceptible to noise uncertainty and has unacceptable performance under low signal to noise ratio (SNR) conditions. Cyclic feature detectors can suppress the stationary noise and thus are quite robust to noise uncertainty [13]–[15]. They can also differentiate various modulation types and hence potentially separate PUs from CRs and categorize signals based on the interference level they can tolerate. However, they typically require sampling at a higher-than-Nyquist rate in order to induce cyclostationarity, and generally take long sensing time to acquire the cyclic statistics. Due to the high sampling rate requirements, all these techniques are typically carried out on a channel-by-channel basis over small bandwidth, in combination with a frequency shifter that scans one channel at a time over the

wide band, or using a bank of filters followed by narrowband processors , each of which sense only a narrow bandwidth [10]. However, for fast and reliable adaptation to the highly dynamic wideband fading channels, it is desired to perform wideband spectrum sensing that simultaneously detects all signal sources over a wide band of interest. In this thesis wideband radio architecture is adopted in order to detect multiple sources over the entire band of interest.

The focus of this thesis is development of wideband spectrum sensing techniques that take advantage of the capabilities of cyclic feature detectors, while resolving the high sampling rate requirements. The main goal is to resolve the high sampling rate issue of cyclic feature detectors, by exploiting the two-dimensional sparsity of a signal's cyclic spectrum. To achieve this, the unique sparsity property of the two-dimensional (2D) cyclic spectrum of communication signals is exploited to alleviate the sampling rate requirements. The main contribution of this work is the reconstruction of the two dimensional cyclic spectrum from reduced number of samples.

According to the theory of compressed sensing a signal which is sparse in a certain domain can be recovered using computationally feasible algorithms from a reduced number of samples [18], [19]. Since the whole concept of CR is based on the fact that the frequency spectrum of interest is sparse (underutilized), it is an ideal problem to which compressed sensing (CS) is a natural solution. Accordingly, compressive sampling techniques have been developed for wideband sensing in CR networks [27]–[30], which effectively reconstruct the sparse frequency spectrum from sub-Nyquist-rate samples collected from random linear samplers [18], [19], [22], [24]–[26]. Afterwards, spectrum occupancy is estimated based on the recovered frequency or power spectrum, which are still prone to estimation errors caused by noise variations. In this work, apart from the sparsity of the frequency spectrum, we take advantage of an additional dimension of sparsity which appears due to the modulation-dependent structure of the cyclic spectrum. This work exploits the signal

sparsity in both dimensions to extract second-order statistics, which departs from existing compressed sensing techniques for recovering one-dimensional sparse signals. Accordingly the 2D cyclic spectrum recovery problem is reformulated so that it can be feasibly solved via convex ℓ_1-norm minimization. The proposed compressive cyclic spectrum estimator can simultaneously capture all communication signals over a very wide band, and meanwhile remove non-cyclic noise and interference. Power spectral density (PSD) estimation for stationary signals from reduced samples is also presented as a special case of cyclic spectrum estimation [32]. This formulation process also reveals that we can use compressive sampling to recover the PSD from reduced samples even when the PSD is not sparse. This is due to the utilization of all the time-varying cross-correlation terms of compressed samples. Motivated by this phenomenon, we investigate the minimum number of measurements needed for adequate performance of the proposed cyclic spectrum and PSD recovery algorithms according to the principles of compressed sensing. We also compare these values with the number of measurements that are needed for the frequency response recovery algorithm proposed in [27].

This thesis also contributes to estimating the spectrum occupancy of wideband signals from the cyclic spectrum when multiple signal sources are present. On a single channel, binary signal detection based on the cyclic spectrum has been extensively studied. Subject to unknown noise levels, a family of cyclic feature detectors have been developed and compared based on various design principles [13], [14]. The maximum-SNR single cycle detector has been described in [14]. However, it is little studied and remains a challenge to use the cyclic spectrum for simultaneous estimation of the spectrum occupancy states of multiple signals spreading on a wide band. In this work a band-by-band detection approach is proposed, where a multi-cycle generalized likelihood ratio test (GLRT) is used on each channel to detect PU presence [34]. To obtain the noise statistics used in the GLRT detector, a blind estimator is derived to estimate the noise variance from the available compressive samples. Further, when the spectrum is occupied by (multiple) BPSK signals, a fast and effective

method is developed to detect the presence of BPSK signals and estimate their bandwidth using cyclic features that contain useful information about their carrier frequencies and transmission rates. The next section presents the organization of the rest of this thesis.

1.4 Organization of the thesis

The rest of this thesis is organized as follows. Chapter 2 gives the theoretical background of cyclic spectral analysis and compressed sensing, which are used to develop the desired spectrum sensing techniques. The first part of chapter 3 introduces the problem statement and describes the signal model used in this thesis. The proposed cyclic spectrum estimation and power spectral density estimation techniques are also presented in this chapter. The last section of chapter 3 presents comparisons of the minimum number of measurements needed among the proposed cyclic spectrum recovery algorithm, PSD recovery algorithm and frequency response recovery algorithm found in [27]. Chapter 4 presents two cyclic-based spectrum occupancy estimation techniques. Chapter 5 presents the simulation setups and simulation results used to demonstrate the effectiveness of the proposed spectrum sensing methods. The last chapter, chapter 6, gives concluding remarks and potential future work.

Chapter 2

Theoretical Background

2.1 Second-order cyclostationarity

In this section a high level introduction of second-order cyclostationarity and definition of related terminologies are presented. Traditionally communication signals are assumed to be stationary processes; specifically wide sense stationarity (WSS) is assumed most of the time. A random process is said to be WSS if its first moment (mean) is constant and its second moment is a function of the time difference only. Even if these two conditions do not hold true for the entire time duration of the signal, small segments of the signal can have a constant mean and time invariant second moment. The principle of cyclostationarity on the other hand states that most manmade signals are better modeled as cyclostationary random processes whose statistical parameters vary periodically with time [13]. Although the message contained in a communication signal is unknown and is appropriately modeled as a stationary random process, the periodicity arises from the carrier sine wave, the symbol period, repeating spreading codes etc. If there is more than one periodicity with periods that are not integer multiples of each other, the process is called almost cyclostationary. Cyclostationary modeling is very advantageous for signal interception and related low power signal detection problems. Most of the background material presented in this section is referred from [13] and [14].

A signal is said to be cyclostationary (in the wide sense) if and only if a certain nonlinear transformation of the signal generates a spectral line [13]. Here the focus is on second order cyclostationarity where a quadratic transformation of the signal yields spectral lines in the power spectral density (PSD). For a real time signal $x(t)$ a linear combination of the delay products is a generalization of time invariant quadratic transformation. A real signal $x(t)$ is second order cyclostationary if and only if the PSD of the delay product signal $x(t-\tau_1)x(t-\tau_2)$ contains spectral lines at some nonzero frequency α. The fundamental parameter of second order periodicity is the cyclic autocorrelation function (CAF), which is given by:

$$R_x^{\alpha}(\tau) \triangleq \hat{E}\left\{x(t)x^*(t-\tau)e^{-j2\pi\alpha(t-\tau)}\right\} \tag{2.1}$$

Where $\hat{E}\{.\}$ denotes the time averaging operation given by

$$\hat{E}\{.\} \triangleq \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}(.)dt$$

α is called cyclic frequency and is a discrete quantity unlike spectral frequency. It shows the spectral correlation between the frequency shifted versions of $x(t)$ and is drawn on a separate axis. For $\alpha = 0$, the CAF reduces to the conventional autocorrelation function

$$R_x^0 = \hat{E}\{x(t)x^*(t-\tau)\}. \tag{2.2}$$

A useful interpretation of the CAF is to see, $R_x^{\alpha}(\tau)$ as a conventional crosscorrelation function of frequency translates of $x(t)$. That is, if $u(t)$ and $v(t)$ are frequency translates of $x(t)$ as:

$$u(t) = x(t)e^{-j\pi\alpha t}$$
$$v(t) = x(t)e^{-j\pi\alpha t} \tag{2.3}$$

Then $R_x^\alpha(\tau)$ is

$$R_u^\alpha v(\tau) \triangleq \hat{E}\{u(t)v^*(t-\tau)\} = R_x^\alpha(\tau) \tag{2.4}$$

The PSDs of $u(t)$ and $v(t)$ are given by

$$S_u(f) = S_x(f + \frac{\alpha}{2})$$
$$S_v(f) = S_x(f - \frac{\alpha}{2}) \tag{2.5}$$

From this it follows that $x(t)$ exhibits second-order periodicity if and only if there is a correlation between the frequency translates of $x(t)$, i.e. (2.4) is not identically zero as a function of τ for some $\alpha \neq 0$. The Fourier transform of the CAF is called the Cyclic Spectral Correlation Density function (SCD) and is denoted by $S_x^\alpha(f)$.

$$S_x^\alpha(f) = \int_{-\infty}^{\infty} R_x^\alpha(\tau)e^{-j2\pi f\tau}d\tau \tag{2.6}$$

For $\alpha \neq 0$, $S_x^\alpha(f)$ is the density of the spectral correlation between the spectral components of $x(t)$ at the frequencies $f - \frac{\alpha}{2}$ and $f + \frac{\alpha}{2}$. For $\alpha = 0$ the cyclic spectral correlation density function reduces to the conventional power spectral density function. As an example, the derivation of the cyclic spectrum of an Amplitude modulated (AM) signal is shown from [13]. Let $s(t)$ be a noise free AM signal given by:

$$s(t) = a(t)cos(2\pi f_c t + \phi_0) \tag{2.7}$$

From the definition of CAF it can be shown that

$$R_s^\alpha(\tau) = \begin{cases} \frac{1}{2}R_a^0(\tau)cos(2\pi f_c\tau), & \alpha = 0 \\ \frac{1}{4}R_a^0(\tau)e^{j2\phi_0}, & \alpha = \pm 2f_c \\ 0, & otherwise \end{cases} \tag{2.8}$$

Fourier transforming (2.8) results in the cyclic spectral density function:

$$S_s^{\alpha}(f) = \begin{cases} \frac{1}{4}[S_a^0(f - f_c) + S_a^0(f + f_c)], & \alpha = 0 \\ \frac{1}{4}S_a^0(f)e^{j2\phi_0}, & \alpha = \pm 2f_c \\ 0, & otherwise \end{cases} \tag{2.9}$$

The magnitude plot for the SCF of such an AM signal is shown in figure 2.1.

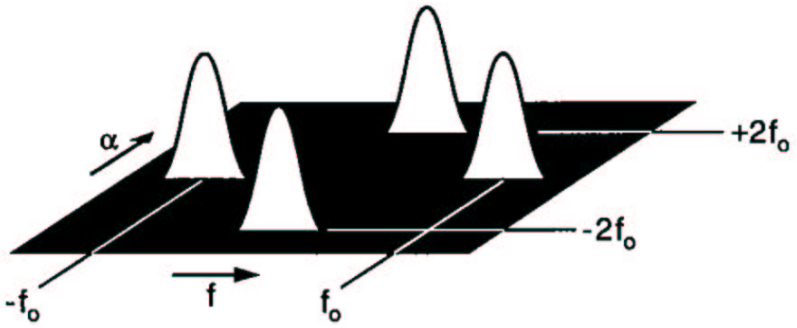


Figure 2.1: Surface plot of the SCD magnitude for an AM signal

As part of the introduction of second-order cyclostationarity we briefly mention some of the major advantages of cyclic spectral analysis. The first and the one most relevant to this thesis is, its discriminatory capabilities. Different signals have different cyclic features which can be used to separate one from the other. Hence signals with overlapping PSDs can be separated in the cyclic frequency domain. The second advantage, which is related to the first one, is suppression of background Gaussian and non-Gaussian noise with symmetric PDFs. Consider a signal (non-Gaussian) received along with additive Gaussian noise; according to the principles of cyclostationary processing, the background noise will not have any features at non-zero cycle frequencies. Therefore there will be no background noise effect, except process noise, at the cyclic frequencies where the cyclic features of the signal are expected to appear. Hence cyclostationary signal processing provides detec-

tion and parameter estimation advantages over the conventional stationary processing. As an example of this advantage of cyclic spectrum, figure 2.1 shows the Spectral Coherence function SCF of white Gaussian noise and figure 2.1 shows the SCF of a BPSK signal. As can be seen from the respective figures the SCF of white Gaussian noise is zero for $\alpha \neq 0$ while BPSK has non zero spectral values at non-zero cyclic frequency values.

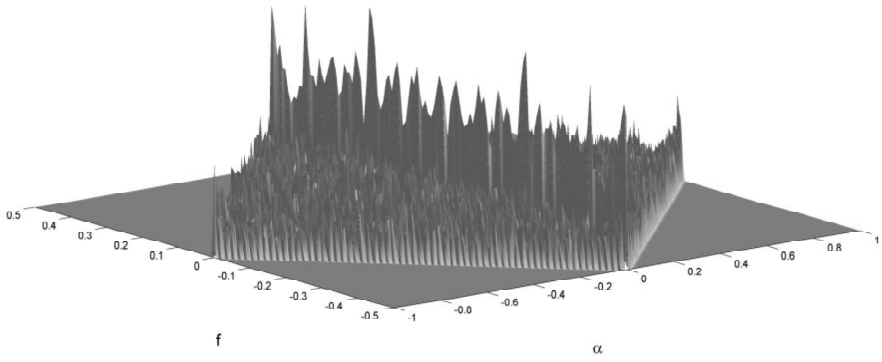


Figure 2.2: Surface plot of the SCF magnitude of WGN

In general the performance of cyclic feature detectors depends on the strength of the features they are trying to detect. Among these features the most exploited for signal detection are related to the carrier frequency and symbol period of signals. Symbol rate related features can be exploited after RF-to-baseband conversion. Whereas Features related to the carrier frequency are exploited in the RF domain before conversion to baseband. An example of such a feature that is widely exploited is the spectral correlation associated with a carrier frequency f_c of a communication signal, at $\alpha = \pm 2f_c + K/T_0$. Here T_0 stands for symbol period and K is an integer. This feature appears in modulation schemes like BPSK, that do not have quadrature components .

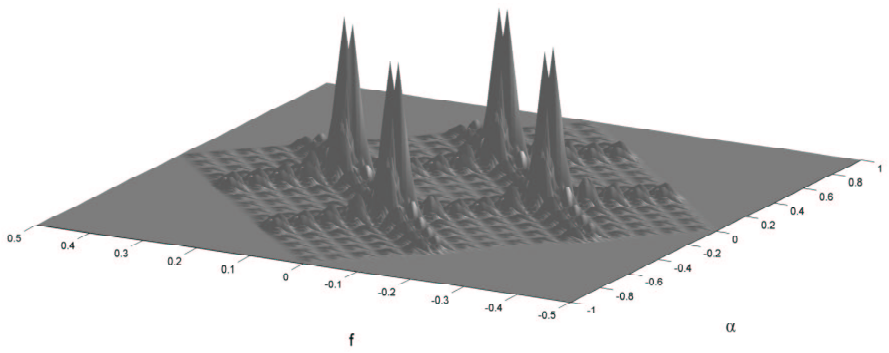


Figure 2.3: Surface plot of the SCF magnitude of a BPSK signal

2.2 Compressed Sensing

The well known Shannon's sampling theorem for signal acquisition states that the sampling rate must be at least twice the maximum frequency present for correct representation (reconstruction of the signal). This sampling rate is known as the Nyquist sampling rate. It is easy to see how this principle poses a challenge for a cognitive radio which needs to perform spectrum sensing over a wide frequency bandwidth of potentially up to several gigahertzes. The traditional radio architecture uses a method of sweeping a single narrowband filter over the entire wideband of concern or uses multiple narrowband filters in parallel. These narrowband solutions are slow, require a lot of additional circuitry and are not responsive to practical environments. Another method is to use a wideband RF front-end followed by digital signal processing (DSP) blocks; but this method requires a very high speed Analog to digital converter (ADC) which may be very costly if not infeasible. This is where compressed sensing (CS) becomes a very useful solution. According to the theory of CS, a signal representing the whole wideband of interest can be recovered from

far fewer samples than required by the traditional Nyquist sampling theorem, given the signal of interest meets certain sparsity requirements. In order to show how CS is used for wideband spectrum sensing, this section gives a brief review of the theory of compressed sensing.

The concept of compressed sensing is one that takes the concept of data compression one step further. The high compression rates achieved by various data compression techniques have demonstrated that, most of the data we acquire can be thrown away without a significant loss. CS is a method by which just the important information about the signal of interest is acquired to begin with instead of collecting a lot of data and then compressing it. To make this work there are two main criteria that need to be met, sparsity and incoherence.

Sparsity addresses the fact that the information rate of a signal is lower than its bandwidth. This means, representation of the signal in some domain has few non-zero elements. According to CS most natural signals fulfill this criterion. An example is a digital image which when acquired using Nyquist rate sampling has many samples, but remains convincingly unchanged after throwing away many of the samples using different compression techniques [20]. Of course cognitive radio is another example which works on the principle that at any given time there are fewer number of frequency bands occupied in the RF spectrum, making the signal sparse in the frequency domain.

Let's explain the concept of sparsity mathematically. Let $\mathbf{x}$ be a real-valued, finite length, one-dimensional, discrete-time signal of size $N \times 1$ in $\mathcal{R}^N$. This vector $\mathbf{x}$ can be represented in terms of an $N \times N$ orthonormal basis matrix $\mathbf{\Psi} = [\psi_1 \psi_2 ... \psi_N]$, which is formed by stacking the vectors $\{\psi_i\}$ as columns.

$$\mathbf{x} = \sum_{i=1}^{N} \beta_i \psi_i \tag{2.10}$$

where $\{\beta_i\}_{i=1}^{N}$ are weighting coefficients with $\beta_i = \langle \mathbf{x}, \psi_i \rangle = \psi_i^T \mathbf{x}$. Here $.^T$ denotes transpose operation. $\mathbf{x}$ and β are representations of the same signal in different domains. we say that $\mathbf{x}$ has a sparse representation when it is a linear combination of only K basis vectors, with $K \ll N$, i.e, only K of β_i are non-zero. In CS a linear measurement process is used to acquire $M < N$ measurements as:

$$\mathbf{y} = \mathbf{\Phi x} \tag{2.11}$$

Where $\mathbf{y}$ is the compressed $M \times 1$ measurement and $\mathbf{\Phi}$ is a measurement matrix. Here, the measurement matrix $\mathbf{\Phi}$ does not depend on the signal $\mathbf{x}$ in any way, making (2.11) a non-adaptive measurement process.

The incoherence criterion is the property which must be satisfied between the measurement matrix $\mathbf{\Phi}$ and the signal's representation matrix $\mathbf{\Psi}$. It addresses the idea that objects with sparse representation in $\mathbf{\Psi}$ must be spread in the domain in which they are acquired [21]. In other words we want the largest correlation between any two elements of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ to be small, or for $\mathbf{\Phi}$ and $\mathbf{\Psi}$ to have maximum incoherence between each other. Equation (2.11) can be re-written as

$$\mathbf{y} = \mathbf{\Phi x} = \mathbf{\Phi \Psi} \beta = \mathbf{\Theta} \beta \tag{2.12}$$

where $\mathbf{\Theta} = \mathbf{\Phi \Psi}$ is an $M \times N$ matrix. The goal is to design a measurement matrix $\mathbf{\Phi}$ and a reconstruction algorithm that allow us to require only $M \approx K$ measurements of a K sparse compressible signal in order to capture all the needed information content of the signal. To achieve these goals, two tasks need to be performed. First we need to design a measurement matrix $\mathbf{\Phi}$ that preserves the needed information content in $\mathbf{x}$ while acquiring a reduced vector $\mathbf{y}$. In order to realize this for a given sparsifying basis $\mathbf{\Psi}$, we need to

construct a measurement matrix $\mathbf{\Phi}$ such that $\mathbf{\Theta} = \mathbf{\Phi\Psi}$ satisfies the so-called *restricted isometry property* (RIP). Without going in to too much detail, in CS we obtain a matrix that meets these conditions by choosing $\mathbf{\Phi}$ as a random matrix. An example of such a matrix can be constructed by drawing its elements from independent and identically distributed (iid) Gaussian random variables [20]. For such a Gaussian $\mathbf{\Phi}$, the matrix $\mathbf{\Theta} = \mathbf{\Phi\Psi}$ is iid for any sparsifying basis matrix $\mathbf{\Psi}$, making Gaussian measurements $\mathbf{\Phi}$ universal in the sense that $\mathbf{\Theta}$ has the RIP with high probability for any $\mathbf{\Psi}$.

The second task is development of a reconstruction algorithm to recover $\mathbf{x}$ from $\mathbf{y}$. Since $M < N$ in (2.12), this is an ill posed problem. This is because there are infinitly many $\beta\prime$ that satisfy $\mathbf{\Theta}\beta\prime = \mathbf{y}$. However the assumption of the sparsity of the signal in the basis $\mathbf{\Psi}$ makes the recovery possible. Under the sparsity assumption the goal is to find the signal with l_0-sparsest coefficients β that agree with the observed measurement $\mathbf{y}$ [22]. l_p stands for the p-th norm of a vector, say $\mathbf{x}$, defined as

$$||\mathbf{x}||_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}. \tag{2.13}$$

Unfortunately solving this l_0 optimization problem is NP-hard but for $M \geq cKlog(N/K)$ iid Gaussian measurements we need only solve the l_1-sparsest coefficients β that agree with the observed measurement $\mathbf{y}$ [20].

$$\begin{aligned} \hat{\beta} =& \arg\, min||\beta||_1 \\ & \text{s.t. } \mathbf{y} = \mathbf{\Phi\Psi}\beta \end{aligned} \tag{2.14}$$

This is a convex optimization problem, also known as *Basis Pursuit* [22]. Another algorithm known as Orthogonal Matching Pursuit (OPM) can also be applied to the recovery problem [23].

Chapter 3

Compressive Cyclic Spectrum Sensing

3.1 Signal Model and Problem Statement

Consider a CR monitoring a wide band of interest in the frequency range $[-f_{\max}, f_{\max}]$, where $f_{\max}$ is very large. There are I active PU signals emitting over this wide band, where the i-th signal is denoted by $x_i(t)$, $i \in \{1, 2, \ldots, I\}$. We have no information regarding the number of signals present. The waveform, bandwidth and carrier frequency of each signal is not known either. The CR is assumed to be equipped with an (ideal) wideband antenna that passes all signal components within the frequency range of interest, $[-f_{\max}, f_{\max}]$. Hence, the received signal is given by

$$x(t) = \sum_{i=1}^{I} x_i(t) + w(t) \tag{3.1}$$

where $w(t)$ is the additive ambient noise. The cyclic spectrum $S(\alpha, f)$ of $x(t)$, where f is the frequency and α is the cyclic frequency is nonzero only for $|f| + \frac{|\alpha|}{2} \leq f_{\max}$ [14].

In a traditional analog-to-digital conversion (ADC) process, $x(t)$ is sampled uniformly every $T_s = \frac{1}{f_s}$ seconds, where f_s is the sampling rate in Hz. The cyclic spectrum of the digital samples becomes (Eq. (62) in [14]):

$$\tilde{S}(\alpha, f) = \frac{1}{T_s} \sum_{m,n=-\infty}^{\infty} S(\alpha + \frac{m}{T_s}, f - \frac{m}{2T_s} - \frac{n}{T_s}). \tag{3.2}$$

As shown in Fig. 3.1, replicas of the original cyclic spectrum show up due to spectrum folding. Accordingly to avoid any aliasing in the cyclic spectrum, the minimum sampling rate should be $2f_{\max}$ [13]:

$$f_s = \frac{1}{T_s} \geq 2f_{\max}. \tag{3.3}$$

Because $f_{\max}$ is very large in our CR sensing task, the required sampling rate f_s has to be very high, causing high energy consumption and high hardware costs of ADC. On the other hand, we observe that the cyclic spectrum of the signal is in fact sparse, which motivates us to utilize the compressive sampling (CS) approach to reduce the sampling rate.

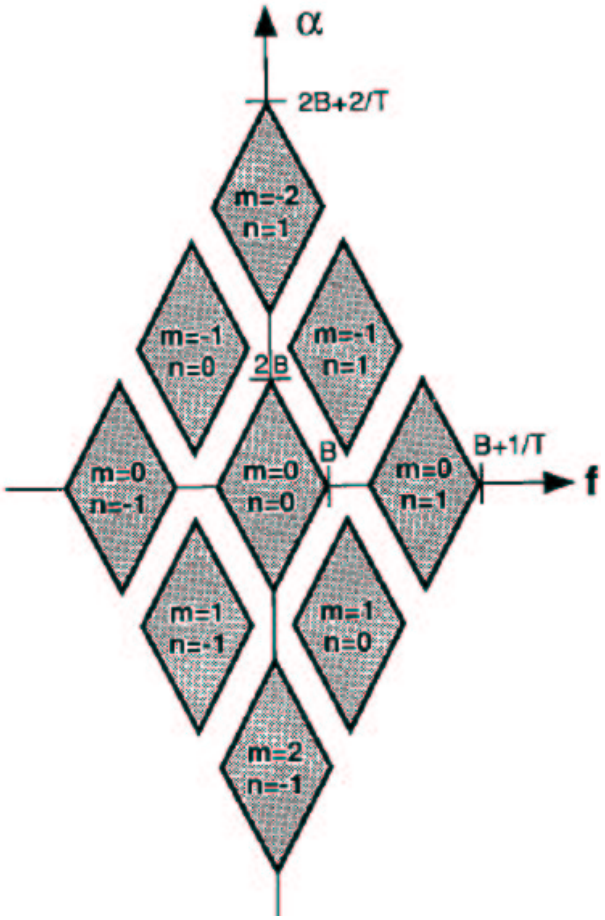


Figure 3.1: Cyclic spectrum of digital samples

In CS, $x(t)$ is not directly sampled, but is passed through a set of properly designed analog filters $\{a_m(t)\}_m$ before sampling at a reduced rate $f_{s,c} = \frac{1}{T_{s,c}}$, yielding $z[m] = a_m(t) * x(t)|_{t=mT_{s,c}}$ ($*$ denotes convolution). Advances in CS assert that $x(t)$ can be recovered error free with high probability even when $f_{s,c}$ is lower than the Nyquist rate, as long as $\{a_m(t)\}_m$ satisfy certain restricted isometric property [18], [19]. Essentially these sampling filters need to be broadband and mutually incoherent, and the number of samples needed (hence the sampling rate) is determined by the sparsity order of the signal of interest. Examples include the parallel compressive samplers in [24]–[26] and the serial analog-to-information converter (AIC) in [22].

For a sensing time-block of T_B seconds, the traditional Nyquist-rate ADC would yield N samples as $\mathbf{x}_t = [x[0], ..., x[N-1]]^T$, for $N = T_B/T_s$ and $T_s \leq 1/(2f_{\max})$. Whereas a compressive sampler would yield M samples, which we collect into $\mathbf{z}_t = [z[0], ..., z[M-1]]^T$ for $M = T_B/T_{s,c} \leq N$. Apparently the actual sampling rate $f_{s,c} = \frac{1}{T_{s,c}}$ is a fraction of f_s, with $f_{s,c} = (M/N)f_s$. Also, $\mathbf{z}_t$ is related to $\mathbf{x}_t$ through a linear sampling matrix $\mathbf{A} \in \mathcal{R}^{M \times N}$, whose rows are given by the digital representations of the sampling filters $\{a_m(t)\}_m$ respectively. In discrete-time domain, the compressive sampling process is described by

$$\mathbf{z}_t = \mathbf{A}\mathbf{x}_t. \tag{3.4}$$

Note that we directly collect samples $\mathbf{z}_t$ from $x(t)$ at a sub-Nyquist rate $f_{s,c} = (M/N)f_s$, without having to generate Nyquist-rate samples $\mathbf{x}_t$ first. The first challenge after collecting the compressive samples is to reconstruct the 2D cyclic spectrum $S(\alpha, f)$ for $|f| + \frac{\alpha}{2} \leq f_{\max}$, directly from the compressive measurements $\mathbf{z}_t$. Then, based on the recovered cyclic spectrum, we extract useful features to estimate the spectrum occupancy of the wide band of interest.

3.2 Recovery of Sparse Cyclic Spectrum

Consider a real-valued communication signal $x(t)$. The digital sequence $\mathbf{x}_t$ is zero-mean cyclostationary, which means that its time-varying covariance $r_x(n,\nu) = \mathrm{E}\{x(nT_s)x(nT_s + \nu T_s)\} = \mathrm{E}\{\mathbf{x}_t(n)\mathbf{x}_t(n+\nu)\}$ is periodic in n with some integer P as the cyclic period ($\mathrm{E}\{\cdot\}$ denotes expectation), that is, $r_x(n,\nu) = r_x(n+kP,\nu)$, $\forall n,k,\nu$ being integers. The cyclic period reveals the hidden periodicity in the signal, as discussed in chapter 2. When multiple signal components are present in $x(t)$, P is the least common multiple of the cyclic periods of all components, and N is set to be an integer multiple of P.

The cyclic covariance function of $x(t)$ is given by the Fourier series of $r_x(n,\nu)$ with respect to time n. Since $\mathbf{x}_t$ has a finite length N, its cyclic covariance can be estimated as

$$\tilde{r}_x^{(c)}(a,\nu) = \left\{ \frac{1}{N} \sum_{n=0}^{N-1-\nu} r_x(n,\nu) e^{-j\frac{2\pi}{N}an} \right\} e^{-j\frac{\pi}{N}a\nu} \tag{3.5}$$

where $a \in [0, N-1]$ indicates the cyclic frequency $\alpha = \frac{1}{NT_s}a$, and the adjustment term $e^{-j\frac{\pi}{N}a\nu}$ makes the sequence symmetric with respect to a. This is a biased estimate, but its estimation variance is smaller than that of an unbiased one [17]. It is assumed that all the cyclic periods of all the signal components of $x(t)$ belong to a set of allowable cyclic periods $\{\frac{NTs}{a}\}_{a=0}^{N-1}$ corresponding to cyclic frequencies $\{\frac{a}{NTs}\}_{a=0}^{N-1}$. That is for any component $x_i(t)$ it shall satisfy, $r_{x_i}(n,\nu) = r_{x_i}(n+T_i,\nu)$, where $T_i = \frac{NT_s}{a}$ for some integer a.

The cyclic spectrum is the Fourier transform of the cyclic covariance with respect to time-delay ν, given by

$$s_x^{(c)}(a,b) = \textstyle\sum_{\nu=0}^{N-1} \tilde{r}_x^{(c)}(a,\nu) e^{-j\frac{2\pi}{N}b\nu} \tag{3.6}$$

where $b \in [0, N-1]$ indicates the frequency $f = \frac{1}{NT_s}b$. The problem of cyclic spectrum recovery amounts to estimating $\{s_x^{(c)}(a,b)\}_{a,b=0}^{N-1}$ from compressive samples $\mathbf{z}_t$.

To facilitate analysis, we represent those 2D quantities $\{r_x(n,\nu)\}_{n,\nu=0}^{N-1}$, $\{\tilde{r}_x^{(c)}(a,\nu)\}_{a,\nu=0}^{N-1}$ and

$\{s_x^{(c)}(a,b)\}_{a,b=0}^{N-1}$ as matrices $\mathbf{R}_x$, $\tilde{\mathbf{R}}_x^{(c)}$ and $\mathbf{S}_x^{(c)}$ respectively, each of size $N \times N$. The nonzero support region of the 2D cyclic spectrum $\mathbf{S}_x^{(c)}$ is shown in Fig. 3.2(b), which corresponds to the folded cyclic spectrum within the square region $\alpha, f \in [0, f_s]$. Once $\mathbf{S}_x^{(c)}$ is acquired, it can be wrapped into the original cyclic spectrum $S(\alpha, f)$ in the diamond-shaped region of Fig. 3.2(a), by properly mapping the three subregions marked by I, II and III. Further, for real signals, since the cyclic spectrum is conjugate symmetric about the origin, all the useful information is contained in the sub region I of the cyclic spectrum plots in Fig. 3.2. As a result only the shaded region in Fig. 3.2(b) needs to be recovered in order to reconstruct the original complex valued cyclic spectrum of real valued signals. Let $\mathbf{S}_{xT}^{(c)}$ denote the cyclic spectrum matrix for this region, which is of size $N \times (\frac{N}{2}+1)$ for N even, and $N \times (\frac{N+1}{2})$ for N odd. In the remainder of this thesis N is assumed to be even. It is straight forward to show that $\mathbf{S}_x^{(c)}$ is related to $\mathbf{S}_{xT}^{(c)}$ as:

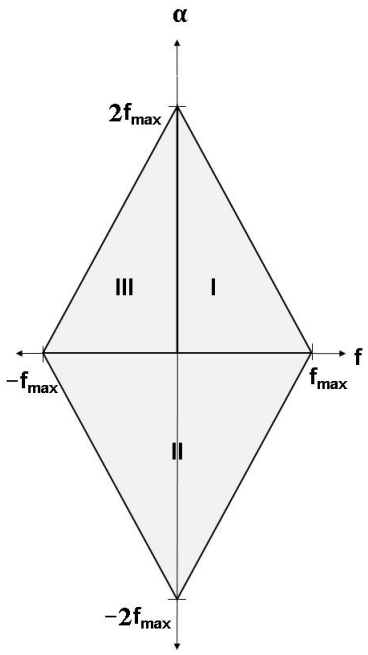
$$\mathbf{S}_x^{(c)} = \mathbf{S}_{xT}^{(c)}\mathbf{D} \tag{3.7}$$

where $\mathbf{D} = [\mathbf{I}_{(\frac{\mathbf{N}}{\mathbf{2}}+\mathbf{1})}\ \mathbf{J}]$. Here $\mathbf{I}_{(\frac{\mathbf{N}}{\mathbf{2}}+\mathbf{1})}$ is an identity matrix of size $\frac{N}{2}+1$ and $\mathbf{J}$ is formed by removing the first and last columns of an $\frac{N}{2}+1$ reverse identity matrix.

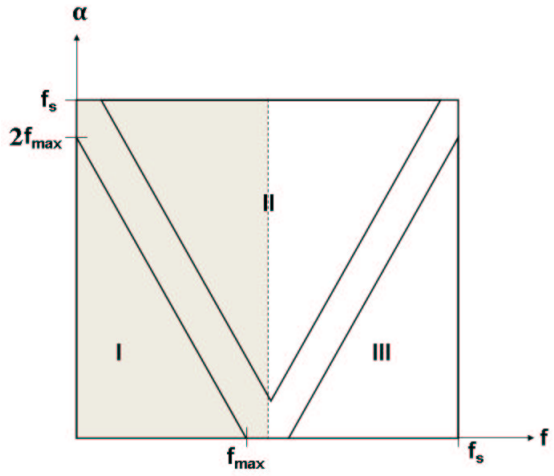
3.2.1 Vector Relationship

The time-varying covariance matrix $\mathbf{R}_x = E\{\mathbf{x}_t\mathbf{x}_t^T\}$ has the form:

$$\mathbf{R}_x = \begin{bmatrix} r_x(0,0) & r_x(0,1) & r_x(0,2) & \cdots & r_x(0,N-1) \\ r_x(0,1) & r_x(1,0) & r_x(1,1) & \cdots & r_x(1,N-2) \\ r_x(0,2) & r_x(1,1) & r_x(2,0) & \cdots & r_x(2,N-3) \\ \vdots & & & \ddots & \vdots \\ r_x(0,N-1) & \cdots & \cdots & \cdots & r_x(N-1,0) \end{bmatrix} \tag{3.8}$$



(a) Original cyclic spectrum $S(\alpha, f)$ of the analog signal $x(t)$.



(b) The folded cyclic spectrum $\{s_x^{(c)}(a,b)\}_{a,b=0}^{N-1}$ of digital samples $\mathbf{x}_t$, shown within $0 \leq \alpha, f \leq f_s$.

Figure 3.2: Original and folded cyclic spectrums of a signal

which is symmetric for real-valued signals. Due to the symmetry, there are $\frac{N(N+1)}{2}$ degrees of freedom in $\{r_x(n,\nu)\}_{n,\nu}$ for $n+\nu<N$, which can be arranged into the vector:

$$\mathbf{r}_x = [r_x(0,0), r_x(1,0), \cdots, r_x(N-1,0), r_x(0,1), r_x(1,1), \\ \cdots, r_x(N-2,1), \cdots\cdots, r_x(0,N-1)]^T.$$

To relate $\mathbf{r}_x$ with $\mathbf{R}_x$, we inspect $\text{vec}\{\mathbf{R}_x\} \in \mathcal{R}^{N^2}$, where the operator $\text{vec}\{\cdot\}$ stacks all the columns into a vector. Evidently, $\text{vec}\{\mathbf{R}_x\}$ contains the same $\frac{N(N+1)}{2}$ elements of $\mathbf{r}_x$, and hence can be related to $\mathbf{r}_x$ using mapping matrices $\mathbf{P}_N \in \{0,1\}^{N^2\times\frac{N(N+1)}{2}}$ and $\mathbf{Q}_N \in \{0,\frac{1}{2},1\}^{\frac{N(N+1)}{2}\times N^2}$ (shown in Appendix A) that map the entries of $\mathbf{r}_x$ with corresponding ones in $\text{vec}\{\mathbf{R}_x\}$:

$$\mathbf{r}_x = \mathbf{Q}_N \text{vec}\{\mathbf{R}_x\} \quad \text{and} \quad \text{vec}\{\mathbf{R}_x\} = \mathbf{P}_N \mathbf{r}_x. \tag{3.9}$$

Alternatively, to facilitate the representation of (3.5), we form another covariance-related matrix $\mathbf{R}$:

$$\mathbf{R} = \begin{bmatrix} r_x(0,0) & r_x(0,1) & r_x(0,2) & \cdots & r_x(0,N-1) \\ r_x(1,0) & r_x(1,1) & r_x(1,2) & \cdots & 0 \\ r_x(2,0) & r_x(2,1) & r_x(2,2) & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ r_x(N-1,0) & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{3.10}$$

which contains the same $\frac{N(N+1)}{2}$ elements of $\mathbf{r}_x$ besides zeros. Thus, $\text{vec}\{\mathbf{R}\} \in \mathcal{R}^{N^2}$ can be mapped to $\mathbf{r}_x$ as:

$$\text{vec}\{\mathbf{R}\} = \mathbf{B}\mathbf{r}_x \tag{3.11}$$

where $\mathbf{B} \in \{0,1\}^{N^2\times\frac{N(N+1)}{2}}$ is the binary-valued mapping matrix that maps the entries of $\mathbf{r}_x$ with corresponding ones in $\text{vec}\{\mathbf{R}\}$. Specifically, it is constructed by removing $\frac{N(N-1)}{2}$ columns of the $N^2 \times N^2$ identity matrix, for any column index $q = \nu N + n$ satisfying $n+\nu \geq N$, $\forall n,\nu \in [0, N-1]$, or equivalently satisfying $\lfloor q/N \rfloor + (q - \lfloor q/N \rfloor N) \geq N$.

Following (3.5) and (3.6), the cyclic covariance matrix $\tilde{\mathbf{R}}_x^{(c)}$ and the cyclic spectrum matrix $\mathbf{S}_x^{(c)}$ are related to $\mathbf{R}$ through Fourier series transformation, as follows:

$$\tilde{\mathbf{R}}_x^{(c)} = \sum_{\nu=0}^{N-1} \mathbf{G}_\nu \mathbf{R} \mathbf{D}_\nu \tag{3.12}$$

$$\mathbf{S}_x^{(c)} = \tilde{\mathbf{R}}_x^{(c)} \mathbf{F} \tag{3.13}$$

where $\mathbf{F} = \left[e^{-j\frac{2\pi}{N}\nu(b)}\right]_{(\nu,b)}$ is the N-point DFT matrix, $\mathbf{G}_\nu = \left[\frac{1}{N} e^{-j\frac{2\pi}{N}a(n+\frac{\nu}{2})}\right]_{(a,n)} \in C^{N\times N}$, and $\mathbf{D}_\nu$ is an $N \times N$ matrix with only its (ν, ν)-th diagonal element being 1 and all other elements being 0.

The cyclic spectrum of interest can be fully represented in a vector form as $\mathbf{s}_x^{(c)} = \text{vec}\left\{\mathbf{S}_x^{(c)}\right\}$. From (3.7), $\mathbf{s}_x^{(c)}$ can also be expressed as:

$$\mathbf{s}_x^{(c)} = (\mathbf{D}^T \otimes \mathbf{I_N}) \text{vec}\left\{\mathbf{S}_{xT}^{(c)}\right\} = \mathbf{K}\mathbf{s}_{xT}^{(c)} \tag{3.14}$$

where $\mathbf{K} = (\mathbf{D}^T \otimes \mathbf{I_N})$ and $\mathbf{s}_{xT}^{(c)} = \text{vec}\left\{\mathbf{S}_{xT}^{(c)}\right\}$ is of size $N(\frac{N}{2} + 1)$. $\otimes$ represents Kronecker product. Similarly we can vectorize the cyclic covariance matrix and express it in terms of $\mathbf{s}_{xT}^{(c)}$ using (3.13) and (3.14) as follows:

$$\text{vec}\left\{\tilde{\mathbf{R}}_x^{(c)}\right\} = \text{vec}\left\{\mathbf{S}_x^{(c)} \mathbf{F}^{-1}\right\} = (\mathbf{F}^{-T} \otimes \mathbf{I_N})\mathbf{K}\mathbf{s}_{xT}^{(c)} = \mathbf{W}\mathbf{s}_{xT}^{(c)} \tag{3.15}$$

where $\mathbf{W} = (\mathbf{F}^{-T} \otimes \mathbf{I_N})\mathbf{K}$ is of size $N^2 \times N(\frac{N}{2} + 1)$, $\mathbf{F}^{-1}$ is the inverse DFT matrix and $\mathbf{F}^{-T}$ is its transpose. Meanwhile, we can also express $\text{vec}\left\{\tilde{\mathbf{R}}_x^{(c)}\right\}$ in terms of $\mathbf{r}_x$ from (3.11) and (3.12), as follows:

$$\text{vec}\left\{\tilde{\mathbf{R}}_x^{(c)}\right\} = \sum_{\nu=0}^{N-1} \mathbf{G}_\nu \mathbf{R} \mathbf{D}_\nu = \sum_{\nu=0}^{N-1} (\mathbf{D}_\nu^T \otimes \mathbf{G}_\nu)\mathbf{B}\mathbf{r}_x = \mathbf{H}\mathbf{r}_x \tag{3.16}$$

where $\mathbf{H} = \sum_{\nu=0}^{N-1} (\mathbf{D}_\nu^T \otimes \mathbf{G}_\nu)\mathbf{B}$ is of size $N^2 \times N(N+1)/2$. From (3.15) and (3.16), it holds that

$$\mathbf{r}_x = \mathbf{H}^\dagger \mathbf{W} \mathbf{s}_{xT}^{(c)} \tag{3.17}$$

where $\mathbf{H}^\dagger$ is the pseudo inverse of $\mathbf{H}$. $\mathbf{H}$ is left invertible since its rank is $\frac{N(N+1)}{2}$, which is also the rank of $\mathbf{B}$.

Next, we aim to relate the unknown $\mathbf{s}_{xT}^{(c)}$ with the data vector $\mathbf{z}_t$ by deriving the relationship between $\mathbf{r}_x$ and the time-varying data covariance matrix $\mathbf{R}_z = \mathrm{E}\{\mathbf{z}_t\mathbf{z}_t^T\} \in \mathcal{R}^{M\times M}$. Similar to $\mathbf{R}_x$ in (3.8), $\mathbf{R}_z$ is symmetric with $\frac{M(M+1)}{2}$ degrees of freedom, which can be organized into the vector:

$$\mathbf{r}_z = [r_z(0,0), r_z(1,0), \cdots, r_z(M-1,0), r_x(0,1), r_z(1,1), \\ \cdots, r_z(M-2,1), \cdots\cdots, r_z(0,M-1)]^T.$$

Similar to (3.9), it holds that $\mathbf{r}_z = \mathbf{Q}_M \mathrm{vec}\{\mathbf{R}_z\}$, where $\mathbf{Q}_M$ of size $\frac{M(M+1)}{2} \times M^2$ is defined similarly as $\mathbf{Q}_N$ in (3.9) and given in Appendix A. Meanwhile, it holds from (3.4) that $\mathbf{R}_z = \mathbf{A}\mathbf{R}_x\mathbf{A}^T$. Thus, it can be derived from (3.9) that

$$\mathbf{r}_z = \mathbf{Q}_M \mathrm{vec}\{\mathbf{A}\mathbf{R}_x\mathbf{A}^T\} = \mathbf{Q}_M(\mathbf{A}\otimes\mathbf{A})\mathrm{vec}\{\mathbf{R}_x\} = \mathbf{\Phi}\mathbf{r}_x \tag{3.18}$$

where $\mathbf{\Phi} = \mathbf{Q}_M(\mathbf{A}\otimes\mathbf{A})\mathbf{P}_N$ is of size $\frac{M(M+1)}{2} \times \frac{N(N+1)}{2}$.

Now we can relate the measurements $\mathbf{r}_z$ and the vector form cyclic spectrum $\mathbf{s}_x^{(c)}$ using (3.17) and (3.18) as:

$$\mathbf{r}_z = \mathbf{\Phi}\mathbf{H}^\dagger\mathbf{W}\mathbf{s}_{xT}^{(c)} = \mathbf{\Psi}\mathbf{s}_{xT}^{(c)} \tag{3.19}$$

where $\mathbf{\Psi} = \mathbf{\Phi}\mathbf{H}^\dagger\mathbf{W}$ is of size $\frac{M(M+1)}{2} \times N(\frac{N}{2}+1)$. Utilizing the conjugate symmetric property, only the shaded region in Figure 3.2 is recovered, which reduces the number of complex-valued unknowns in (3.19) to $N(\frac{N}{2}+1)$.

3.2.2 Problem Formulation

Given the available data covariance vector $\mathbf{r}_z$, the cyclic spectrum recovery problem boils down to solving $\mathbf{s}_x^{(c)}$ from (3.19), which is an under-determined linear system. To accurately solve this problem, we make use of the prior knowledge about wideband communication signals. We observe that the 2D cyclic spectrum is highly sparse, which can be incorporated

into the spectrum estimator by penalizing the ℓ_1-norm of the vector $\mathbf{s}_x^{(c)}$ to induce a sparse solution. Hence, we formulate an unconstrained ℓ_1-norm regularized least squares (LR-LS) problem for reconstructing $\mathbf{s}_x^{(c)}$, as follows:

$$\min_{\mathbf{s}_x^{(c)}} \left\| \mathbf{r}_z - \mathbf{\Psi} \mathbf{s}_{xT}^{(c)} \right\|_2^2 + \lambda \left\| \mathbf{s}_{xT}^{(c)} \right\|_1 \tag{3.20}$$

In (3.20), $\lambda > 0$ is a weighting scalar that balances between the sparsity of the solution induced by the ℓ_1-norm term and the data reconstruction error reflected by the ℓ_2-norm LS term.

The optimization formulation in (3.20) is attractive since it addresses the sparsity of the cyclic spectrum and is a convex problem, which can be solved by convex solvers such as *cvx* [37].

3.2.3 Implementation

In practical implementations with finite data samples, the data vector $\mathbf{r}_z$ in (3.20) is replaced by its finite-sample estimate $\hat{\mathbf{r}}_z$. Specifically, the steps for sparse cyclic spectrum recovery are:

1. Observe the received analog signal $x(t)$ for a total time of T, which is chosen according to the desired sensing time and the time needed to produce reliable statistics;

2. Divide the total time into L blocks, each of time-block length $T_B = T/L$. Within each block l, the same random sampler $\mathbf{A}$ (e.g., the AIC [22]) is used to obtain a compressive sample vector $\mathbf{z}_t(l) \in \mathcal{R}^M$ directly from $x(t)$ at a reduced rate of $f_{s,c} = M/T_B = (M/N)f_s$, where $f_s = 1/T_s \geq 2f_{\max}$ and $N = T_B/T_s$, with a compression ratio $M/N < 1$. Estimate the covariance matrix $\mathbf{R}_z$ of the compressive samples by

the finite-sample average across all blocks as $\hat{\mathbf{R}}_z = \frac{1}{L}\sum_{l=0}^{L-1} \mathbf{z}_t(l)\mathbf{z}_t^T(l)$, and form $\hat{\mathbf{r}}_z = \mathbf{Q}_M \text{vec}\{\hat{\mathbf{R}}_z\}$.

3. Use $\hat{\mathbf{r}}_z$ to replace $\mathbf{r}_z$, and solve (3.20) to recover $\hat{\mathbf{s}}_{xT}^{(c)}$.

The choice of L reflects the design tradeoff, via $L = Tf_s/N$ for given T and f_s. As L increases, the estimation variance of $\hat{\mathbf{r}}_z$ is reduced and the finite-sample effect is alleviated, but the frequency resolution of the recovered $\mathbf{s}_{xT}^{(c)}$ also suffers, because the resolution increases linearly in N.

3.3 Recovery of Power Spectrum for Stationary Signals

Besides cyclic spectrum estimation, the developed formulations can be used to efficiently estimate the power spectrum (PSD) in the one-dimensional frequency domain, which is a special case of the 2D cyclic spectrum when the signal $x(t)$ is (treated as) stationary. Power spectrum estimation is useful in many well-known spectral estimators, such as energy detectors.

When $x(t)$ is stationary, the covariance function $r_x(n, \nu)$ becomes time invariant, which means that

$$r_x(n, \nu) = \bar{r}_x(\nu), \quad \forall n. \tag{3.21}$$

Let $\bar{s}_x(b)$ denote the discrete power spectrum with respect to frequency b. It is given by the Fourier series of $\{\bar{r}_x(\nu)\}_{\nu=0}^{N-1}$, in the form

$$\bar{s}_x(b) = \sum_{\nu=0}^{N-1} \bar{r}_x(\nu) e^{-j\frac{2\pi}{N}b\nu}, \quad b \in [0, N-1]. \tag{3.22}$$

Defining $\bar{\mathbf{s}}_x = [\bar{s}_x(0), \ldots, \bar{s}_x(N-1)]^T$ and $\bar{\mathbf{r}}_x = [\bar{r}_x(0), \ldots, \bar{r}_x(N-1)]^T$ results in

$$\bar{\mathbf{s}}_x = \mathbf{F}\bar{\mathbf{r}}_x \tag{3.23}$$

where $\mathbf{F}$ is the N-point DFT matrix.

Because of the stationarity property in (3.21), the time-varying covariance matrix in (3.8) reduced to a stationary one as follows:

$$\mathbf{R}_x = \begin{bmatrix} \bar{r}_x(0) & \bar{r}_x(1) & \bar{r}_x(2) & \cdots & \bar{r}_x(N-1) \\ \bar{r}_x(1) & \bar{r}_x(0) & \bar{r}_x(1) & \cdots & \bar{r}_x(N-2) \\ \bar{r}_x(2) & \bar{r}_x(1) & \bar{r}_x(0) & \cdots & \bar{r}_x(N-3) \\ \vdots & & & \ddots & \vdots \\ \bar{r}_x(N-1) & \cdots & \cdots & \cdots & \bar{r}_x(0) \end{bmatrix}. \tag{3.24}$$

Mapping between (3.8) and (3.24) reveals that

$$\begin{aligned} [\mathrm{vec}\{\mathbf{R}_x\}]_{(n-\nu)N+n} &= [\mathbf{R}_x]_{(n,n-\nu)} = \bar{r}_x(\nu) \\ [\mathrm{vec}\{\mathbf{R}_x\}]_{nN+n-\nu} &= [\mathbf{R}_x]_{(n-\nu,n)} = \bar{r}_x(\nu) \\ \nu \in [0, N-1],\ n &\in [\nu, N-1]. \end{aligned} \tag{3.25}$$

Apparently, $\mathrm{vec}\{\mathbf{R}_x\}$ can be represented linearly by $\bar{\mathbf{r}}_x$ through a binary-valued mapping matrix $\bar{\mathbf{P}}_N \in \{0,1\}^{N^2 \times N}$ as follows:

$$\mathrm{vec}\{\mathbf{R}_x\} = \bar{\mathbf{P}}_N \bar{\mathbf{r}}_x \tag{3.26}$$

where $\bar{\mathbf{P}}_N$ can be deduced from (3.25) as: for any column $\nu \in [1, N-1]$, there are $2(N-\nu)$ elements of 1's at row indices $((n-\nu)N+n)$ and $(nN+n-\nu)$, $n = \nu, \ldots, N-1$, while all the rest elements are zero; for the column $\nu = 0$, there are N elements of 1's at row indices $(nN+n), \forall n \in [0, N-1]$, while the rest elements on this column are zero.

Although $x(t)$ is stationary, $\mathbf{z}_t$ is not, because each element $z[m]$ is generated by filtering $x(t)$ with an individual filter $a_m(t)$ indicated by the corresponding row of $\mathbf{A}$, and $\{a_m(t)\}_m$ are

mutually incoherent as explained for (3.4). Therefore, there are still $M(M+1)/2$ degree of freedom in $\mathbf{R}_z$, represented by $\mathbf{r}_z$. Using $\mathbf{R}_z = \mathbf{A}\mathbf{R}_x\mathbf{A}^T$ deduced from (3.4), and replacing the expression of vec$\{\mathbf{R}_x\}$ in (3.9) by the stationary version in (3.26), (3.18) becomes

$$\mathbf{r}_z = \mathbf{Q}_M(\mathbf{A} \otimes \mathbf{A})\text{vec}\{\mathbf{R}_x\} = \bar{\mathbf{\Phi}}\bar{\mathbf{r}}_x \tag{3.27}$$

where $\bar{\mathbf{\Phi}} = \mathbf{Q}_M(\mathbf{A} \otimes \mathbf{A})\bar{\mathbf{P}}_N$ is of size $\frac{M(M+1)}{2} \times N$.

Putting together (3.23) and (3.27), we reach

$$\mathbf{r}_z = \bar{\mathbf{\Phi}}\mathbf{F}^{-1}\bar{\mathbf{s}}_x. \tag{3.28}$$

In (3.28), there are $M(M+1)/2$ linear equations and N unknowns. As long as $M(M+1)/2 > N$, this is an over-determined system that can be solved using simple LS. This is a revealing observation, suggesting that sub-Nyquist-rate sampling with $M/N < 1$ can be employed to permit a unique LS solution to the PSD vector $\bar{\mathbf{s}}_x$, even when $\bar{\mathbf{s}}_x$ is non-sparse. This is due to the use of all cross-correlations of the compressive samples in recovering the second-order statistics, which is also exploited in [31], [32]. With (3.28), the PSD can be estimated as

$$\hat{\bar{\mathbf{s}}}_x = \arg \min_{\bar{\mathbf{s}}_x} \|\mathbf{r}_z - \bar{\mathbf{\Phi}}\mathbf{F}^{-1}\bar{\mathbf{s}}_x\|_2^2 + \lambda\|\bar{\mathbf{s}}_x\|_1. \tag{3.29}$$

By setting $\lambda = 0$, (3.29) reduces to a sparsity-agnostic LS estimator that can be used when the PSD is non-sparse and $M(M+1)/2 > N$; on the other hand, adopting $\lambda > 0$ results in a sparsity-aware LR-LS estimator useful when the PSD is sparse, as in CR sensing applications.

3.4 Comparison of Spectrum, PSD and Cyclic Spectrum Estimation Techniques

In this chapter we have presented cyclic spectrum and as a special case PSD estimation techniques using compressive sensing. The PSD reconstruction problem has revealed that

$M < N$ compressive samples can be used to recover not only a parse but also non-sparse PSD. In this section we compare the minimum number of measurements needed to solve the cyclic spectrum and PSD recovery problems. In the literature compressive sensing has also been applied for estimating the frequency response of sparse signals [27], [29], [30]. In this section we also compare the frequency spectrum and edge spectrum (derivative of the PSD) recovery formulations found in [27] with the PSD and cyclic spectrum recovery formulations proposed in this paper in terms of the number of known and unknown variables found in the inverse problems of the respective optimization problem formulations. These comparisons reveal the needed reduced number of measurements M for each method.

Consider our wideband signal $\mathbf{x}_t$ which is sparse in the frequency domain and has a total bandwidth of B_T. Let B_O stand for the actual occupied bandwidth of the signal. Then the sparsity ratio of the signal in the frequency domain is given by $K_p = B_O/B_T$. K_p represents the one dimensional frequency domain sparsity ratio of the signal. The cyclic spectrum of $\mathbf{x}_t$ on the other hand adds an additional dimension of sparsity, the cyclic frequency. Thus the cyclic spectrum domain sparsity of $\mathbf{x}_t$ is different from K_p and is dependent on the modulation schemes of the signal components found in $\mathbf{x}_t$. Lets denote the cyclic domain sparsity ration of $\mathbf{x}_t$ by K_c. K_c and K_p can be related through a constant α_c for a give signal as $K_c = \alpha_c K_P$. Since the cyclic domain is sparser than the frequency domain, $0 < \alpha_c < 1$ and $K_c < K_p$. For example a frequency band of 300MHz containing two BPSK signals of symbol period $T_0 = 0.04\mu$s has a spectral occupancy of 33.3% and cyclic domain occupancy of 3.125%.

Let,

$$\tilde{\mathbf{y}} = \tilde{\mathbf{A}}\tilde{\mathbf{x}} \tag{3.30}$$

represent a general inverse problem presented in a compressed sensing recovery formulation. Where $\tilde{\mathbf{y}}$ denotes a vector of known measurements of size $\tilde{M} \times 1$, $\tilde{\mathbf{x}}$ denotes the vector of unknowns of size $\tilde{N} \times 1$ that need to be estimated and $\tilde{\mathbf{A}}$ is a measurement matrix of size

$\tilde{M} \times \tilde{N}$. Note that $\tilde{M} < \tilde{N}$. According to the principles of compressed sensing we need

$$\tilde{M} \geq cKlog(\tilde{N}/K) \tag{3.31}$$

knowns to solve this under determined inverse problem using minimum l_1 norm reconstruction. Here c is a small constant which depends on the measurement matrix $\tilde{\mathbf{A}}$ and K represents the sparsity of the signal. In our cyclic spectrum recovery formulation the inverse problem presented in (3.20) is

$$\mathbf{r}_z = \mathbf{\Psi}\mathbf{s}_{xT}^{(c)} \tag{3.32}$$

where $\mathbf{r}_z$ is known and is of size $\frac{M(M+1)}{2} \times 1$ and $\mathbf{s}_{xT}^{(c)}$ is the unknown cyclic spectrum of size $\frac{N(N+1)}{2} \times 1$. Substituting $\tilde{M} = \frac{M(M+1)}{2}$, $\tilde{N} = \frac{N(N+1)}{2}$ and $K = K_c = \alpha_c K_p$ in (3.31), we find that

$$M \geq \frac{\sqrt{1 + 8c_1\alpha_c K_p log(\frac{N(N+1)}{2K_c})} - 1}{2} \tag{3.33}$$

time domain measurements are required.

Similarly, in the PSD recovery formulation presented in (3.28), the inverse problem is

$$\mathbf{r}_z = \bar{\mathbf{\Phi}}\mathbf{F}^{-1}\bar{\mathbf{s}}_x = \bar{\mathbf{\Theta}}\bar{\mathbf{s}}_x \tag{3.34}$$

where $\bar{\mathbf{s}}_x$ represents the unknown PSD of interest which of size $N \times 1$. Again substituting the number of knowns, $\tilde{M}$, by $\frac{M(M+1)}{2}$, using $\tilde{N} = N$ and $K = K_p$, the number of measurements needed for the PSD recovery is

$$M \geq \frac{\sqrt{1 + 8c_2 K_p log(\frac{N}{K_p})} - 1}{2}. \tag{3.35}$$

In [27], compressive sensing is used to recover the frequency response as well as edge spectrum of a signal. In the frequency spectrum recovery, first compressed random sampling is used to generate reduced $M \times 1$ measurements $\mathbf{z}_t$ from the discrete sequence $\mathbf{x}_t$ of size $N \times 1$ (taken at the Nyquist sampling rate), using a sampling matrix $\mathbf{V}_z$. $\mathbf{z}_t$ are then used

for the reconstruction of the frequency response of the signal $\mathbf{x}_f$. The spectrum recovery formulation in this approach is given as

$$\hat{\mathbf{x}}_f = \arg \min_{\mathbf{x}_f} \|\mathbf{x}_f\|_1, \quad s.t. \ (\mathbf{V}_z^T \mathbf{F}_N^{-1})\mathbf{x}_f = \mathbf{z}_t \tag{3.36}$$

where $\mathbf{x}_f$ is the frequency response of $\mathbf{x}_t$ and $\mathbf{F}_N$ is N point discrete Fourier transform matrix. The linear transformation equality $(\mathbf{V}_z^T \mathbf{F}_N^{-}1)\mathbf{x}_f = \mathbf{z}_t$ in equation (3.36) is an inverse problem with M knowns and N unknowns. Hence, to solve this recovery problem $M \geq c_3 K log(N/K_p)$ measurements are required.

In the edge spectrum recovery approach proposed in [27], the $N \times 1$ discrete vector $\mathbf{h}_s$, representing the edge spectrum is recovered directly from the reduced measurements $\mathbf{x}_t$. First the discrete wavelet transform, $\mathbf{y}_s$, is found as

$$\mathbf{y}_s = \mathbf{F}_N \text{diag}\{\Phi_s\}\mathbf{x}_t. \tag{3.37}$$

where Φ_s is the discrete representation of the inverse transform of a wavelet function. The edge spectrum, $\mathbf{h}_s$, which is the derivative of the wavelet of $\mathbf{x}_f$ is related to $\mathbf{y}_s$ as $\mathbf{h}_s = \Gamma \mathbf{y}_s$, where Γ is a differentiation matrix of size $N \times N$ given by

$$\Gamma = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -1 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \\ 0 & \cdots & -1 & 1 \end{bmatrix} \tag{3.38}$$

Accordingly $\mathbf{x}_t$ and $\mathbf{h}_s$ are related as:

$$\mathbf{x}_t = (\mathbf{F}_N \text{diag}\Phi_s)^{-1}\mathbf{y}_s = (\mathbf{F}_N \text{diag}\Phi_s)^{-1} \cdot \Gamma^{-1} \cdot \mathbf{h}_s. \tag{3.39}$$

Then noting the sparsity of $\mathbf{h}_s$, defining $\mathbf{G} = (\mathbf{F}_N \text{diag}\Phi_s)^{-1} \cdot \Gamma^{-1}$ and using $\mathbf{z}_t = \mathbf{V}_z^T \mathbf{x}_t$, the following formulation is given

$$\hat{\mathbf{h}}_s = \arg \min_{\mathbf{h}_f} \|\mathbf{h}_f\|_1, \quad s.t. \ \mathbf{x}_z^T = (\mathbf{V}_z^T)\mathbf{h}_s. \tag{3.40}$$

Table 3.1

Frequency spectrum, Cyclic spectrum and PSD recovery comparison summary

Recovered Quantity	Sparsity	# Knowns	# Unknowns	Minimum # Measurements
Cyclic Spectrum	$K_c = \alpha_c K_p$	$\frac{M(M+1)}{2}$	$\frac{N(N+1)}{2}$	$M \geq \frac{\sqrt{1+8c_1 K_c log(\frac{N(N+1)}{2K_c})}-1}{2}$
PSD	K_p	$\frac{M(M+1)}{2}$	N	$M \geq \frac{\sqrt{1+8c_2 K_p log(\frac{N}{2K_p})}-1}{2}$
Frequency Spectrum	K_p	M	N	$M \geq c_3 K_p log(N/K_p)$

Afterwards, $\mathbf{h}_s$ is used to detect and estimate frequency band locations. The inverse problem in the optimization formulation of (3.40) also has M knowns and N unknowns.

Table 3.1 gives a summary of the comparisons of the frequency spectrum, cyclic spectrum and PSD recovery formulations presented in this section. It should be noted that the constants c_1, c_2 and c_3 are different from each other since they depend on the respective measurement matrix $\tilde{\mathbf{A}}$ even when the sampling matrix $\mathbf{A}$ (in $\mathbf{z}_t = \mathbf{A}\mathbf{x}_t$) is the same for each of them. An example of the comparison presented here is shown in figure 3.4. The figure shows the minimum number of required measurements for the three estimation techniques corresponding to different values of N. For this example, $c_1 = c_2 = c_3 = 1$ is used, frequency sparsity ratio, $K_p = 0.5$ and $\alpha_c = 0.1$. As can be seen from the plot for a reasonably large enough N, the minimum number of required measurements for the cyclic spectrum estimation is the least.

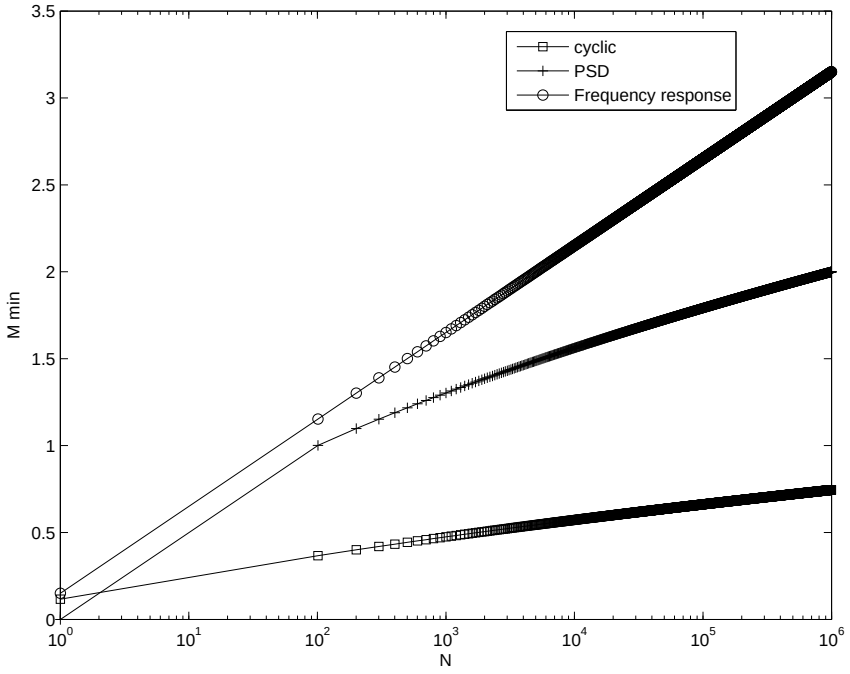


Figure 3.3: Comparison of Spectrum, PSD and Cyclic Spectrum Estimation Techniques

Chapter 4

Cyclic-based Spectrum Occupancy Estimation

After recovering the sparse 2D cyclic spectrum $S(\alpha, f)$ in its vectorized form $\hat{\mathbf{s}}_x^{(c)}$, we now want to simultaneously estimate the spectrum occupancy of all frequency sub-bands within the monitored wide band. We develop two spectrum detection algorithms: one adopts a band-by-band multi-cycle generalized likelihood ratio test (GLRT) framework that works for all types of modulation and waveform patterns, and the other is a simple and fast approach tailored for BPSK signals.

4.1 Multi-cycle GLRT Based Band-by-Band Approach

The goal of spectrum occupancy estimation is to decide whether a specific frequency location $f^{(n)}$ is occupied or not. We set $f^{(n)} = \frac{n}{N} f_s$, $\forall n \in [0, \frac{N}{2}]$, according to the frequency resolution of the discrete cyclic spectrum $s_x^{(c)}(a, b)$ in (3.6). The sensing task amounts to a band-by-band inspection of the spectrum occupancy over the entire frequency range $|f| \leq f_s/2$, with $f_s/2 \geq f_{\max}$. It is important to note that such a band-by-band inspection is a computational approach for processing the data collected simultaneously from a wideband antenna, which is fundamentally different from a narrowband approach in which narrow-

band antennas scan the wide spectrum one by one using frequency shifters and narrowband processors.

We now focus on the occupancy decision on a single band $f^{(n)}$. An active signal on this band would occupy a region $\mathcal{I}^{(n)}$ of the 2D cyclic spectrum map defined by the cyclic-frequency and frequency pairs (α, f) [14]:

$$f + \frac{\alpha}{2} = f^{(n)} \text{ and } |f| + \frac{|\alpha|}{2} \leq f_{\max}, \quad \forall (\alpha, f) \in \mathcal{I}^{(n)}. \tag{4.1}$$

In discrete-time domain, this region is represented by discrete points $(\alpha_i, f_i) \in \mathcal{I}^{(n)}$, which correspond to a set of integer-valued indices (a_i, b_i), that is, $S(\alpha_i, f_i) = s_x^{(c)}(a_i, b_i)$, $\forall i \in \mathcal{I}_d^{(n)}$. Because $\alpha_i = \frac{1}{NT_s} a_i$ and $f_i = \frac{1}{NT_s} b_i$ by definition, the index set $\mathcal{I}_d^{(n)}$ can be deduced from (4.1) as:

$$b_i + \frac{a_i}{2} = n, \text{ and } |b_i| + \frac{|a_i|}{2} \leq \frac{f_{\max} N}{f_s} \leq \frac{N}{2}, \ \forall i \in \mathcal{I}_d^{(n)}. \tag{4.2}$$

We stack the estimated $\{\hat{s}_x^{(c)}(a_i, b_i)\}_{i \in \mathcal{I}_d^{(n)}}$ into a vector $\hat{\mathbf{c}}^{(n)}$ of size $|\mathcal{I}_d^{(n)}|$, which is formed by selected entries of the vectorized cyclic spectrum $\hat{\mathbf{s}}_x^{(c)}$ from (3.20):

$$\hat{\mathbf{c}}^{(n)}[i] = \hat{s}_x^{(c)}(a_i, b_i), \quad i \in \mathcal{I}_d^{(n)}. \tag{4.3}$$

To test for the presence of a PU signal at the frequency $f^{(n)}$, the following binary hypothesis test is formulated:

$$\begin{aligned} H_1 &: \hat{\mathbf{c}}^{(n)} = \mathbf{c}^{(n)} + \boldsymbol{\epsilon}, \\ H_0 &: \hat{\mathbf{c}}^{(n)} = \boldsymbol{\epsilon}, \end{aligned} \tag{4.4}$$

where $\mathbf{c}^{(n)}$ is the non-random true vector of cyclic spectrum values, and $\boldsymbol{\epsilon}$ is asymptotically Gaussian distributed, i.e., $\lim_{LN \to \infty} \sqrt{LN}\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}^{(n)})$, where $\boldsymbol{\Sigma}^{(n)}$ is the asymptotic covariance matrix. Because $\boldsymbol{\Sigma}^{(n)}$ is not readily available, we derive a blind estimator for $\hat{\boldsymbol{\Sigma}}^{(n)}$ using the available reduced-rate measurements $\{\mathbf{z}_t(l)\}_{l=0}^{L}$, shown in Appendix B. Replacing $\boldsymbol{\Sigma}^{(n)}$ by $\hat{\boldsymbol{\Sigma}}^{(n)}$ results in a data-adaptive GLRT detector.

Treating $\mathbf{c}^{(n)}$ as an unknown nuisance parameter, the adaptive GLRT can be derived for (4.4), which yields the following test statistic:

$$\mathcal{T}^{(n)} = (\hat{\mathbf{c}}^{(n)})^T (\hat{\boldsymbol{\Sigma}}^{(n)})^{-1} \hat{\mathbf{c}}^{(n)}. \tag{4.5}$$

Next, $\mathcal{T}^{(n)}$ is compared with a threshold $\eta^{(n)}$ to make decisions on the occupancy at frequency $f^{(n)}$, where $\eta^{(n)}$ is chosen to meet a certain constant false alarm rate (CFAR). This is a multi-cycle GRLT, because $\hat{\mathbf{c}}^{(n)}$ contains multiple cyclic frequencies α_i for $i \in \mathcal{I}_d^{(n)}$. Putting together all the occupied bands, we are able to draw the spectrum occupancy map over the entire wide band. Overall, the proposed spectrum occupancy estimation algorithm is implemented by the following steps:

1. Set $n = 0$. Form the cyclic spectrum matrix $\hat{\mathbf{S}}_x^c$ from $\text{vec}\{\hat{\mathbf{S}}_x^{(c)}\} = \mathbf{K}\hat{\mathbf{s}}_{xT}^c$;
2. Let $f^{(n)} = \frac{n}{N} f_s$. Find the cyclic spectrum values of interest from (4.1), and calculate $\hat{\mathbf{c}}^{(n)}$ from (4.3);
3. Calculate the test statistic in (4.5) using $\hat{\boldsymbol{\Sigma}}^{(n)}$ calculated from (B.1) and (B.3). If it is larger than a pre-determined threshold $\eta^{(n)}$, then a PU is declared present at frequency $f^{(n)}$, otherwise PU absence is declared.
4. If $n < \frac{N}{2}$, increase n by 1 and go to step (2).

4.2 A Fast Algorithm for BPSK Signals

For a BPSK signal with carrier frequency f_c and symbol rate $\frac{1}{T_0}$, the major cyclic feature is a lobe at $\alpha = 2f_c$, ranging from $f = -\frac{1}{T_0}$ to $f = \frac{1}{T_0}$ (Fig. 14 (a) in [14]). The non-zero region of support for such signal is shown in Fig. 4.2.

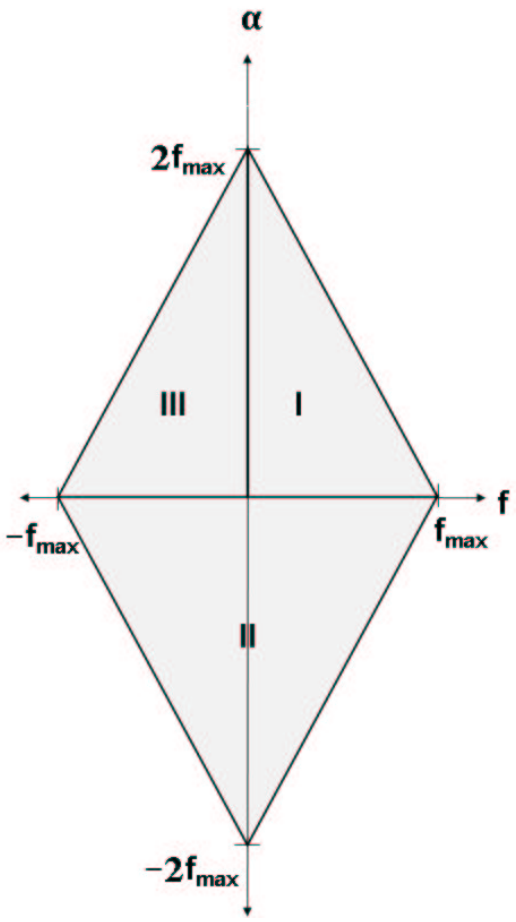


Figure 4.1: Non-zero region of support for the cyclic spectrum of BPSK.

Our idea is to identify the modulation-dependent parameters T_0 and f_c by finding the lobe locations on the estimated 2D cyclic spectrum. First, we can simply search over the cyclic frequency α along the axis $f = 0$. If $|\hat{S}(\alpha, 0)| > \eta$ (η is a preset threshold), then we claim that there is a BPSK signal with estimated carrier frequency $\hat{f}_c = \alpha/2$.

Next, we let $\mathcal{A} = \{\alpha | \, |S(\alpha, 0)| > \eta\}$. For each $\tilde{\alpha} \in \mathcal{A}$, we search along the line $\alpha = \tilde{\alpha}$ to find the double-sided width of the lobe, denoted by $2w$, such that all the points in the lobe have absolute values equal to or greater than η. The bandwidth of the corresponding BPSK signal is also $2w$, and hence we can claim that the frequency band $[\tilde{\alpha}/2 - w, \tilde{\alpha}/2 + w]$ is occupied, and the estimated symbol period is $\hat{T}_0 = 1/w$.

By now, we have rapidly identified not only the carrier frequency f_c but also the bandwidth $1/T_0$ of a BPSK signal. The procedure is applicable to the entire wide band to identify all BPSK signals. Combining all the occupied frequency bands, we are able to draw the spectrum occupancy map.

The details of this approach are:

1. Choose a threshold η for the recovered cyclic spectrum.
2. Search over α along the axis $f = 0$. If $|S(\alpha, 0)| > \eta$, we claim that there is a BPSK signal with carrier frequency $\alpha/2$. Let $\mathcal{A} = \{\alpha | \, |S(\alpha, 0)| > \eta\}$.
3. For each $\tilde{\alpha} \in \mathcal{A}$, search along the line $\alpha = \tilde{\alpha}$ to find the double-sided width of the lobe, denoted by $2w$, such that all the points in the lobe have absolute values equal to or greater than η. The bandwidth of the corresponding BPSK signal will then be $2w$, and we claim that the frequency band $[\tilde{\alpha}/2 - w, \tilde{\alpha}/2 + w]$ is occupied by a PU.
4. Combining all the occupied frequency bands, we draw the spectrum occupancy map.

The fast approach can be extended to other modulation types using their respective cyclic spectrum features. For instance, an SQPSK signal [14] with carrier frequency f_c and symbol rate $\frac{1}{T}$, two peaks with similar heights will appear at $\alpha = 2f_c - \frac{1}{T}$ and $\alpha = 2f_c + \frac{1}{T}$, and this feature can be used for fast detection.

Chapter 5

Simulations

5.1 Simulation Setup

In this chapter Monte Carlo simulation setups and results are presented that testify the effectiveness of the proposed algorithms under sampling rate reduction and noise uncertainty. In the simulations the frequency band monitored has $f_{\max} = 500$MHz, in which two signal sources PU_1 and PU_2 appear at the center frequencies 187.5MHz and 375MHz respectively. Two separate simulation setups are considered:

a) Both PU_1 and PU_2 are BPSK signals.

b) PU_1 is BPSK and PU_2 is QPSK.

The symbol period for each source is fixed at $T_0 = 0.1$ μs. The reference (Nyquist) sampling rate is $f_s = 1$GHz, and the non-compressed discrete-time representation $\mathbf{x}_t$ is of length $N = 32$ in each time block. Although both setups a) and b) have the same nonzero spectral occupancy of 8%, their nonzero cyclic spectral occupancies are not the same, due to the different cyclic features of BPSK and QPSK signals. In fact, Setup a) has a 0.75% cyclic spectral occupancy, and Setup b) has a cyclic spectral occupancy of 0.625%. These occupancy ratios also demonstrate that the 2D cyclic spectrum is much sparser than the 1D

power spectrum.

5.2 Robustness to Rate Reduction

A key parameter of interest is the compression ratio M/N for sampling rate reduction. Its impact on the performance of the proposed cyclic spectrum recovery algorithm is tested on Setup b) with a mixture of BPSK and QPSK signals, in a noise-free case. The performance metric is the normalized mean-square error (MSE) of the reconstructed cyclic spectrum, $\hat{\mathbf{s}}_x^{(M/N)}$, for a given compression ratio M/N compared with the uncompressed one, $\hat{\mathbf{s}}_x^{(c)}$. That is,

$$\text{MSE} = \text{E}\left\{ \| \frac{\hat{\mathbf{s}}_x^{(M/N)} - \mathbf{s}_x^{(c)}\|_2^2}{\|\mathbf{s}_x^{(c)}\|_2^2} \right\}. \tag{5.1}$$

Similarly the MSE performance for the proposed PSD reconstruction is also evaluated as:

$$\text{MSE} = \text{E}\left\{ \| \frac{\hat{\bar{\mathbf{s}}}_x^{(M/N)} - \bar{\mathbf{s}}_x\|_2^2}{\|\bar{\mathbf{s}}_x\|_2^2} \right\} \tag{5.2}$$

where $\hat{\bar{\mathbf{s}}}_x^{(M/N)}$ represents the power spectrum reconstructed from compressive samples with compression ratio of M/N and $\bar{\mathbf{s}}_x$ stands for the uncompressed one. Figure 5.2 depicts the MSE versus the compression ratio curves, for $L = 40, 200$ and 400 data blocks of the cyclic spectrum reconstruction. The MSE performance of the PSD reconstruction algorithm is also shown in the same figure for $L = 200$. The figure shows that the MSE curves start to flatten out when $M/N \geq 0.4$ for all cases, which is the rate-reduction region that the PSD and cyclic spectrum recovery algorithm are robust. It can be seen that there is some MSE improvement achieved by using $L = 200$ instead of $L = 40$ in the cyclic spectrum reconstruction. However, the estimation performance for $L = 400$ is very close to the one for $L = 200$. This suggests that the additional averaging does not offer significant

improvement. Since $L = 200$ corresponds to a much shorter sensing time than $L = 400$, it is a more reasonable choice whenever the performance loss can be tolerated.

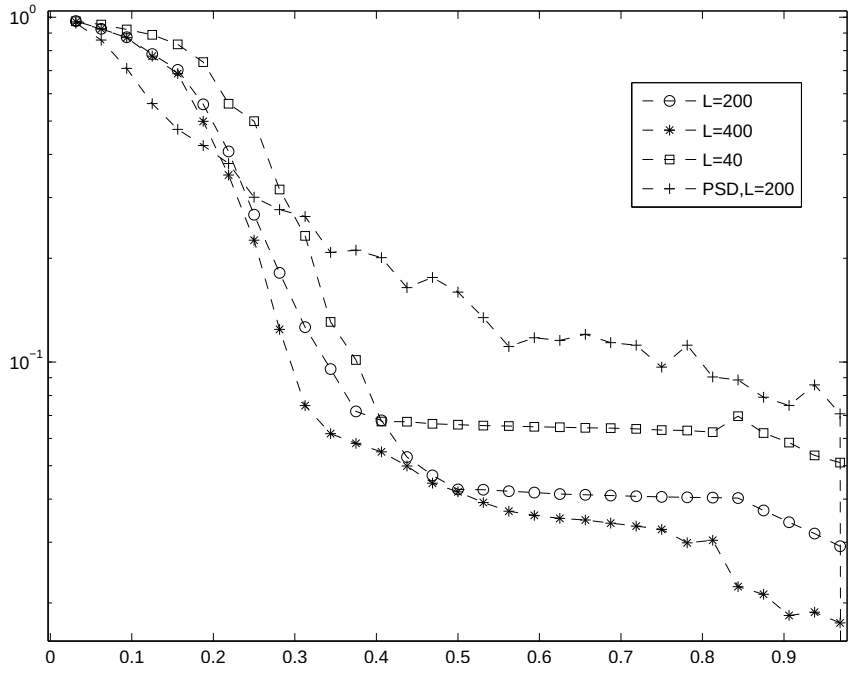


Figure 5.1: Normalized MSE of reconstruction versus compression ratio M/N.

After cyclic spectrum recovery, the GLRT-based approach is utilized to estimate the frequency occupancy. Figure 5.2 depicts the probability of detection P_d versus M/N, computed using $L = 200$ for CFAR $= 0.01$. It shows that the compression ratio can go as low as 0.25 while keeping $P_d \geq 0.9$. As a result, when the frequency occupancy is 8%, the sampling rate can be reduced by three-quarter, while still maintaining good estimation performance. When the frequency occupancy is sparser, the effective sampling rate can be reduced even further.

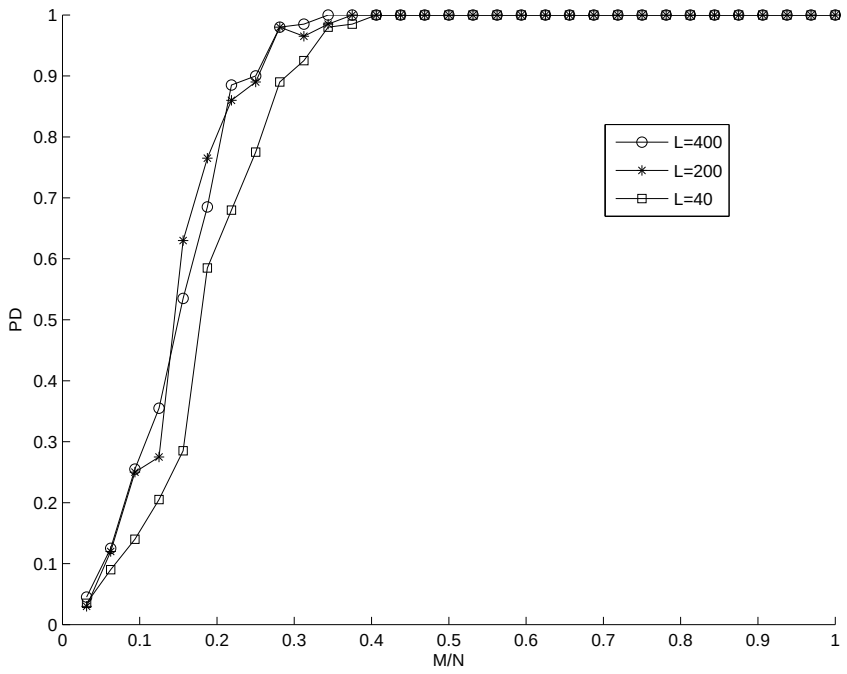


Figure 5.2: Probability of detection at a constant false alarm ratio of 0.01.

5.3 Robustness to Noise Uncertainty

To testify the robustness of spectrum sensing under noise uncertainty, we compare the receiver operating characteristic (ROC) of the proposed compressive cyclic feature detector to that of an energy detector (ED). A traditional ED can be combined with existing compressive sampling techniques through a two-step procedure. First, within each data block l, the frequency response $\mathbf{x}_f(l) = \mathbf{F}\mathbf{x}_t(l)$ is reconstructed from compressive samples $\mathbf{z}_t(l) = \mathbf{A}\mathbf{x}_t(l)$ using the LR-LS formulation as

$$\hat{\mathbf{x}}_f(l) = \arg \min_{\mathbf{x}_f(l)} \left\| \mathbf{z}_t(l) - \mathbf{A}\mathbf{F}^{-1}\mathbf{x}_f(l) \right\|_2^2 + \lambda \left\| \mathbf{x}_f(l) \right\|_1 . \tag{5.3}$$

Then, the estimated $\hat{\mathbf{x}}_f(l)$ over all L blocks are used to compute the power spectrum $\bar{\mathbf{s}}_x$ as

$$\hat{\bar{\mathbf{s}}}_x = \frac{1}{L} \sum_{l=1}^{L} \text{diag} \left\{ \hat{\mathbf{x}}_f(l) \hat{\mathbf{x}}_f^H(l) \right\}$$

where diag$\{\cdot\}$ returns the diagonal of a matrix and ^{H} denotes conjugate transpose. In energy detection, each element of $\hat{\bar{\mathbf{s}}}_x$ serves as the energy-related test statistic for the signal component at the corresponding frequency band, which is passed through a decision threshold to detect the frequency occupancy band by band.

In spectrum detection, the noise power level is varying and never known exactly, due to thermal noise change, amplifier gain change, calibration error, and fluctuating interference levels in CR networks. This lack of accurate knowledge is known as noise uncertainty [12]. In the simulation, noise uncertainty is modeled using the method of robust statistics [35]. In this method, the upper limit of the noise PSD level that corresponds to the lower limit of the SNR is used to calculate the probability of false alarm, whereas the lower limit of the noise PSD corresponding to the upper limit of the SNR is used to calculate the probability of detection. Such robust statistics assess the detection performance under worst-case scenarios. We use a nominal SNR of -3dB with noise uncertainty levels of $0, 1, 2$ and 3 dB. A total of $L = 200$ data blocks are used for noise averaging, and the number

of compressive measurements in a block is $M = 16$, which corresponds to a compression ratio of 0.5. Fig. 5.3 depicts the ROCs of the proposed compressive cyclic feature detector and PSD detector, compared with ED. The figure shows that the proposed cyclic feature detector is quite insensitive to noise uncertainty, because the varying yet non-cyclic noise does not appear at $\alpha \neq 0$. The ROC performance considerably outperforms that of energy detection, while the latter degrades noticeably as the noise uncertainty level increases.

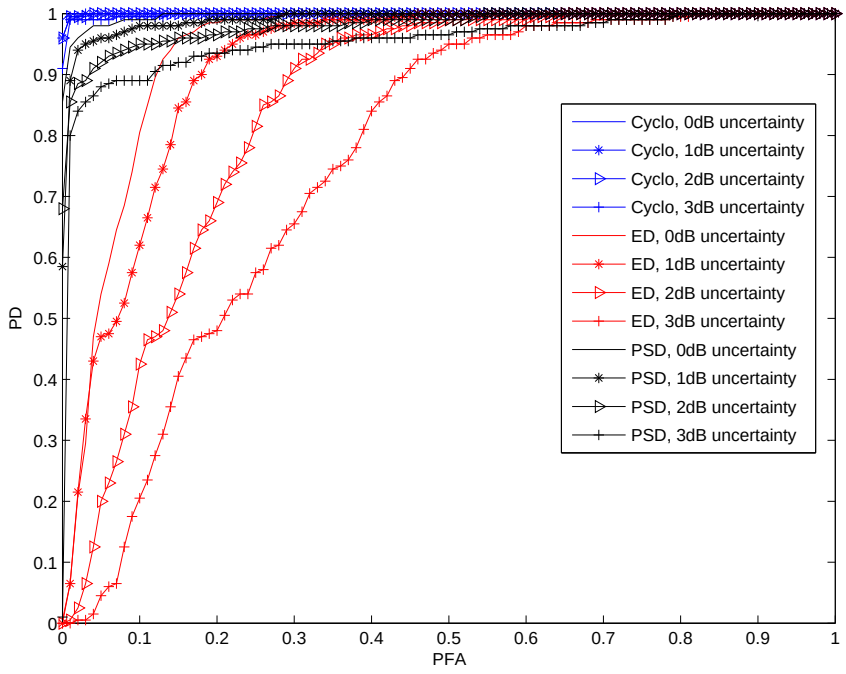


Figure 5.3: ROC of the cyclic feature approach ("Cyclo"), PSD approach ("PSD") and energy detection ("ED").

5.4 Comparison of Spectrum Occupancy Estimation Techniques

Here the GLRT based and the fast cyclic-based spectrum occupancy estimation techniques are compared, using Setup a) where both PUs are BPSK signals. The test parameters are: the SNR is -5dB without uncertainty, $M = 16$ (i.e., $M/N = 50\%$) and $L = 200$. Figure 5.4 shows the ROC curves for both algorithms. From the figure we learn that the GLRT approach provides a slightly better performance. Considering that the GLRT-based approach applies to general modulation types, it is preferred whenever enough computational resources are available.

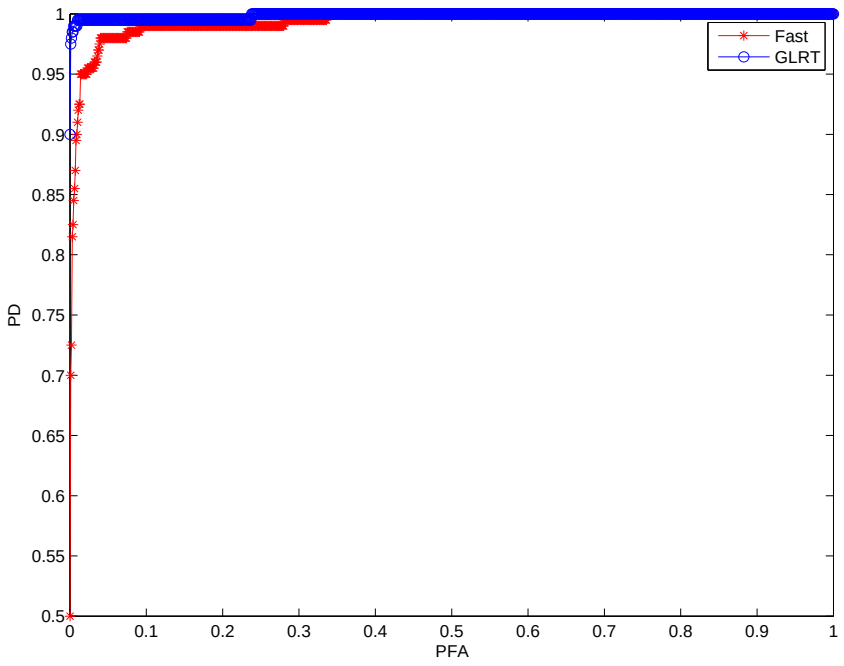


Figure 5.4: ROC of the two spectrum occupancy estimation approaches.

5.5 The Effect of The Number of Blocks Used

The last parameter we investigate is the total sensing time. When the block size N and the sampling rate f_s are fixed, the total sensing time depends on the number of blocks L. Increasing L means more sensing time and less time for transmission, but it could also lead to improved sensing performance. Figure 5.5 shows the impact of L on sensing performance, for SNR = −6dB without uncertainty, and 50% rate compression with $M = 16$ using setup (b). Compared are tested values of $L = 40$, $L = 200$ and $L = 400$. The figure shows that $L = 200$ has better performance than $L = 40$, but increasing L further to 400 does not bring significant additional performance improvement. The same conclusion was drawn from MSE performance comparison presented in section 5.2.

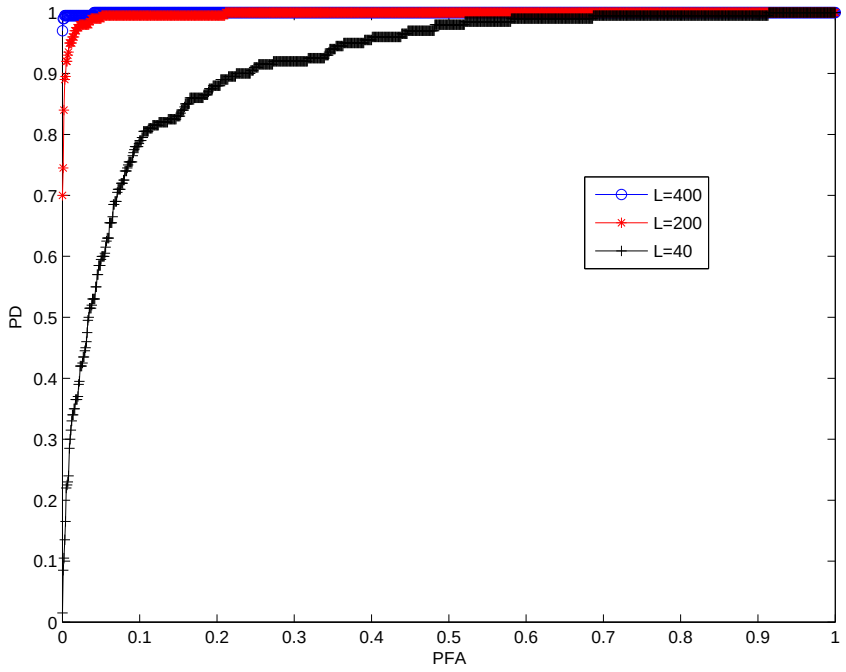


Figure 5.5: ROC of proposed sensing method for different L. $M = 16$, $N = 32$.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

This thesis has presented a wideband spectrum sensing technique that alleviates the high sampling cost of cognitive radios utilizing the concepts of Cyclostationary processing and compressive sampling. A method of recovering the sparse 2-D cyclic spectrum from a reduced number of time-domain samples has been developed using the concept of compressed sensing. The cyclic spectrum is vectorized and reformulated to take linear relationship with the covariance function of the compressive samples, which enables an effective recovery of the 2D cyclic spectrum via convex ℓ_1-norm minimization. Accordingly, a practical algorithm is developed that reduces computational complexity and accounts for process noise effect. As a special case of the compressive cyclic spectrum estimator, a power spectrum estimator for stationary signals is also developed. With the recovered cyclic spectrum, two techniques have been developed to estimate the spectrum occupancy of a wide band with an unknown number of active sources: a band-by-band multi-cycle GLRT detector, and a fast thresholding technique for BPSK signals. The proposed spectrum occupancy estimation techniques demonstrate robustness to sampling rate reduction and noise uncertainty.

6.2 Future Work

This thesis has addressed the basic development of wideband spectrum sensing model for cognitive radios. Some further improvements and possible future work directions are given below.

1. The spectrum sensing model developed in this thesis considers an additive Gaussian noise operating environment, whereas in practice one has to consider other factors including multipath effect and shadowing. So in the future, algorithm modifications can be made to account for these issues.

2. The developed PSD estimation technique has revealed that we can compress signals even when they are not sparse. Motivated by this future work can be conducted to evaluate the developed spectrum sensing models in non-sparse RF environments.

3. It is known that cognitive radios have to abide by stringent spectrum sensing and access requirements in order to minimize interference with primary users. For example the first CR standard under development, 802.22, requires that CRs operating in the TV white spectrum be able to sense primary user transmissions as low as -114dBm. In addition CRs should have very short sensing times in order to make efficient use of the spectrum and provide significant data rates. These requirements suggest that spectrum sensing algorithms should be adaptive to the RF environment. Accordingly, the developed spectrum sensing algorithms should be evaluated with respect to operational requirements including operating characteristics in low SNR environments. Implementation complexity of the algorithms should also be evaluated. Based on the results of the evaluation modifications can be made to make the model more efficient and practical. Such modifications may include the integration of modulation classification with the spectrum sensing. Additional parameters or features that can be identified from the cyclic spectrum should also be investigated in order to make the

CR "smarter". In the developed algorithms the block size N, the number of blocks L and the compression ratio M/N determine the sensing time, frequency resolution and the test statistics estimation errors. Additional tests should be performed to find out the detailed tradeoffs of these factors in various environments. The proposed algorithms can also be made self adaptive in choosing these values online during spectrum sensing based on the specific parameters of the RF environment at hand including the spectrum occupancy ratio.

4. The proposed spectrum sensing algorithms could also be enhanced to perform spread spectrum signal detection in low SNR conditions. It has been found that DSSS signals give rise to unique cyclic spectrum features that are not found in cyclic spectrums of narrowband signals. These features can be exploited to perform DSSS signal detection in low SNR environments. The detection enhancement techniques will allow identification of the spreading gain in DSSS signals, without having to perform dispreading using the spreading codes. The proposed wideband monitoring scheme, the unique features found in DSSS signals and the ability of cyclic statistics to remove Gaussian noise create a very ideal setup for identifying DSSS signals hidden in noise.

5. After the discussed high level implementation issues are evaluated more detailed implementation analysis needs to be done. This would require understanding of the specific purpose and intended operating environment of the CRs. Even though cognitive radios are required to be adaptive and dynamic in their operating characteristics their design requires knowledge of expected operating conditions in order to make optimum resource allocation.

References

[1] Federal Communications Commission, "Spectrum Policy task force," ET Docket No. 02-155, Nov. 2002

[2] T. Erpek, M. Lofquist, K. Patton, "Spectrum Occupancy Measurements Loring Commerce Center Limestone, Maine," Shared Spectrum Company, Sep. 2007

[3] M. McHenry, D. McCloskey, "Spectrum Occupancy Measurements Chicago, Illinois," Shared Spectrum Company, Nov. 2005

[4] S. Haykin, "Cognitive radio: brain-empowered wireless communications", *IEEE journal on selected areas in communications*, vol. 23, pp. 201-220, February 2005.

[5] I. F. Akyildiz, Won-Yeol Lee, M. C. Vuran, Shantidev Mohanty "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey", *Computer Networks*, vol. 50, no. 13, pp. 2127-2159, Sep. 2006.

[6] J. Mitola, G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Personal Communications*, vol. 6, no.4, pp. 13-18, Aug. 1999

[7] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," Doctor of Technology, Royal Ins. Technol. (KTH), Stockholm, Sweden, 2000.

[8] IEEE P802.22/D0.5, "Draft Standard for Wireless Regional Area Networks Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Policies and procedures for operation in the TV Bands," Mar. 2008.

[9] F. Digham, M. Alouini, M. Simon, "On the energy detection of unknown signals over fading channels," *Proc. of IEEE ICC 2005*, vol. 5, May 2003, pp. 3575-3579

[10] D. Cabric, S. M. Mishra, R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios", *Asilomar Conference on Signals, Systems and Computers*, vol. 1 pp. 772- 776, 2004.

[11] A. Sonnenschein and P. Fishman, "Radiometric detection of spread-spectrum signals in noise of uncertain power", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28 (3) pp. 654-660, 1992.

[12] R. Tandra and A. Sahai, "Fundamental limits on detection in low SNR under noise uncertainty", *IEEE Conf. on Wireless Networks, Communications and Mobile Computing*, vol. 1 pp. 464- 469, 2005.

[13] W. A. Gardner, "Signal interception: a unifying theoretical framework for feature detection", *Communications, IEEE Transactions on*, vol. 36 (8) pp. 897-906, 1988.

[14] W. Gardner, "Exploitation of spectral redundancy in cyclostationary signals", *IEEE Signal Processing Magazine*, vol. 8 (2), pp. 14-36, April 1991.

[15] F. Gini and G. B. Giannakis, "Frequency offset and symbol timing recovery in flat-fading channels: a cyclostationary approach", *IEEE Transactions on Communications*, vol. 46 (3), pp. 400-411, March 1998.

[16] C.L. Nikias, A. P. Petropulu, *Higher-Order Spectral Analysis*, Prentice Hall Signal Processing series, 1993.

[17] A. V. Dandawate and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Transactions on Signal Processing*, vol. 42 (9), pp. 2355-2369, 1994.

[18] E. J. Candes, J. Romberg, T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information", *IEEE Transactions on Information Theory*, vol. 52 (2) pp. 489- 509, 2006.

[19] D. L. Donoho, "Compressed sensing", *IEEE Transactions on Information Theory*, vol. 52 (4) pp. 1289 - 1306, 2006.

[20] R. G. Baraniuk, “Compressive Sensing [Lecture Notes],” *Signal Processing Magazine, IEEE*, vol. 24, pp. 118-121, 2007.

[21] Candes, E.J. Wakin, “An Introduction To Compressive Sampling,” *Signal Processing Magazine, IEEE* , vol.25, no.2, pp.21-30, March 2008.

[22] S. Kirolos et al, “Analog-to-Information Conversion via Random Demodulation”, *IEEE Workshop on Design, Applications, Integration and Software*, pp. 71-74, 2006.

[23] J. Tropp and A. C. Gilbert, “Signal recovery from partial information via orthogonal matching pursuit,” Apr. 2005, preprint

[24] Y. Bresler, “Spectrum-blind sampling and compressive sensing for continuous-index signals,” *Proc. of Information Theory and Applications Workshop* (ITA 2008), pp. 547-554, January 2008.

[25] Z. Yu, S. Hoyos, and B. M. Sadler, “Mixed-signal parallel compressed sensing and reception for cognitive radio”, *IEEE ICASSP Conf.* , pp. 3861 - 3864, 2008.

[26] X. Chen, Z. Yu, S. Hoyos, B. M. Sadler, and J. Silva-Martinez, “A sub-Nyquist rate sampling receiver exploiting compressive sensing,” *IEEE Trans. on Circuits and Systems I*, 2011 (to appear).

[27] Z. Tian, and G.B. Giannakis, “Compressed sensing for wideband cognitive radios,Ó *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Vol.4, Honolulu, pp.IV/1357– IV/1360, April 2007.

[28] Z. Tian, “Compressed Wideband Sensing in Cooperative Cognitive Radio Networks”, *IEEE GLOBECOM Conference*, Dec. 2008.

[29] Y. L. Polo, Y. Wang, A. Pandharipande and G. Leus, “Compressive wideband spectrum sensing,” *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*,Taipei, pp.2337– 2340, April 2009.

[30] J.-A. Bazerque, and G. B. Giannakis, “Distributed Spectrum Sensing for Cognitive Radio Networks by Exploiting Sparsity,” *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1847–1862, March 2010.

[31] G. Leus and D.D. Arianandа, “Power Spectrum Blind Sampling,” submitted to *IEEE Signal Processing Letters*, 2011.

[32] D. D. Arianandа, G. Leus and Z. Tian, “Multi-coset Sampling for Power Spectrum Blind Sensing,” *IEEE Intl. Conf. on Digital Signal Processing*, July 2011.

[33] K. Kim et al. “Cyclostationary Approaches to Signal Detection and Classification in Cognitive Radio”, *IEEE DySPAN Conf.*, pp. 212-215, 2007.

[34] J. Lunden, V. Koivunen, A. Huttunen, and H. V. Poor. “Collaborative Cyclostationary Spectrum Sensing for Cognitive Radio Systems”, *IEEE Transactions on Signal Processing*, vol. 57 (11), pp. 4182 - 4195, 2009.

[35] Steve Shellhammer, Rahul Tandra. “Performance of the Power Detector with Noise Uncertainty”, *IEEE 802.22-06/0075r0*, July 2006.

[36] J. Leech, “On the representation of 1, 2, ... , n by differences,Ó *Journal of the London Mathematical Society*,Vol. 31, pp.160– 169, April 1956.

[37] M. Grant, S. Boyd, “CVX: Matlab Software for Disciplined Convex Programming,” http://cvxr.com/cvx/

[38] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, Wiley-Interscience, 2001.

[39] H. L. Van Trees, *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*, Wiley-Interscience, 2002.

Appendix A

Mapping Matrices

The appendix shows the mapping matrices $\mathbf{P}_N$ and $\mathbf{Q}_N$ used in (3.9). The results apply for any value of N.

Let us consider the (n, ν)-th element of $\mathbf{R}$ in (3.10), denoted as $[\mathbf{R}]_{(n,\nu)}$, and is nonzero $\forall n \in [0, N-1], \nu \in [0, N-1-n]$. Its value $r_x(n, \nu)$ appears at two symmetric locations inside $\mathbf{R}_x$ of (3.8) for $\nu > 0$, and appears at one location on the diagonal of $\mathbf{R}_x$ when $\nu = 0$. Specifically,

$$\begin{aligned} &[\mathbf{R}]_{(n,\nu)} = [\mathbf{R}_x]_{(n,n+\nu)} = [\mathbf{R}_x]_{(n+\nu,n)} = r_x(n, \nu), \quad \nu > 0; \\ &[\mathbf{R}]_{(n,\nu)} = [\mathbf{R}_x]_{(n,n)} = r_x(n, 0), \quad \nu = 0. \end{aligned} \tag{A.1}$$

Since $\mathbf{r}_x$ stacks all the vectors of $\mathbf{R}$ after removing those zero entries at the lower-right triangle, the (n, ν)-th element of $\mathbf{R}$ shows up as the p-th element of $\mathbf{r}_x$ with $p(n, \nu) = \sum_{l=0}^{\nu-1}(N-l) + n$. That is, $[\mathbf{R}]_{(n,\nu)} = [\mathbf{r}_x]_{p(n,\nu)}$, where

$$p(n, \nu) = \nu N - \frac{\nu(\nu-1)}{2} + n. \tag{A.2}$$

Meanwhile, $[\mathbf{R}_x]_{(n,n+\nu)}$ and $[\mathbf{R}_x]_{(n+\nu,n)}$ in (A.1) appear at the q_1-th and q_2-th rows of $\text{vec}\{\mathbf{R}_x\}$ respectively, where

$$\begin{aligned} q_1(n, \nu) &= (n+\nu)N + n, \\ q_2(n, \nu) &= nN + n + \nu. \end{aligned} \tag{A.3}$$

Linking the above relationships among $\mathbf{R}, \mathbf{R}_x, \mathbf{r}_x$ and $\text{vec}\{\mathbf{R}_x\}$, we now summarize the expressions for $\mathbf{P}_N$ and $\mathbf{Q}_N$ used in (3.9). For $\text{vec}\{\mathbf{R}_x\} = \mathbf{P}_N\mathbf{r}_x$ to hold, $\mathbf{P}_N$ has only two nonzero entries of value one at the $q_1(n,\nu)$-th and $q_2(n,\nu)$-th rows while the rest rows are zero, for a given column indexed by $p(n,\nu)$. When $\nu = 0$, the two nonzero row locations collide into one location at $q_1(n,0) = q_2(n,0) = nN + n$, on the column $p(n,0) = n$. In mathematical form, $\mathbf{P}_N \in \{0,1\}^{N^2 \times \frac{N(N+1)}{2}}$ can be expressed as

$$\begin{cases} [\mathbf{P}_N]_{(q_1(n,\nu),p(n,\nu))} = [\mathbf{P}_N]_{(q_2(n,\nu),p(n,\nu))} = 1 \\ [\mathbf{P}_N]_{(q,p(n,\nu))} = 0, \quad \forall q \in [0, N^2-1]\backslash\{q_1(n,\nu), q_2(n,\nu)\} \\ \qquad\qquad\qquad\qquad n \in [0, N-1],\ \nu \in [0, N-1-n]. \end{cases} \tag{A.4}$$

To construct $\mathbf{r}_x = \mathbf{Q}_N\text{vec}\{\mathbf{R}_x\}$, we note from (A.1) that given (n,ν), the $p(n,\nu)$-th entry of $\mathbf{r}_x$ is equal to both the $q_1(n,\nu)$-th and $q_2(n,\nu)$-th entries of $\text{vec}\{\mathbf{R}_x\}$, and hence can be computed by the averaged sum of these two entries. When $\nu = 0$, these two entries degenerate to one entry at $q_1(n,0) = q_2(n,0) = nN + n$. Letting $\delta_{i,j}$ denote the Kronecker delta, $\mathbf{Q}_N \in \left\{0, \frac{1}{2}, 1\right\}^{\frac{N(N+1)}{2} \times N^2}$ is given by

$$\begin{cases} [\mathbf{Q}_N]_{(p(n,\nu),q_1(n,\nu))} = [\mathbf{Q}_N]_{(p(n,\nu),q_2(n,\nu))} = \frac{1}{2} + \frac{1}{2}\delta_{\nu,0} \\ [\mathbf{Q}_N]_{(p(n,\nu),q)} = 0, \quad \forall q \in [0, N^2-1]\backslash\{q_1(n,\nu), q_2(n,\nu)\} \\ \qquad\qquad\qquad\qquad n \in [0, N-1],\ \nu \in [1, N-1-n]. \end{cases} \tag{A.5}$$

Appendix B

Noise Covariance Matrix Estimation

This appendix derives a non-data-aided estimator for the noise covariance matrix $\mathbf{\Sigma}^{(n)}$ used in the cyclic feature detector in (4.4).

First, we evaluate the error covariance $\mathbf{\Sigma}_z = \mathrm{E}\{(\hat{\mathbf{r}}_z - \mathbf{r}_z)(\hat{\mathbf{r}}_z - \mathbf{r}_z)^H\}$ of the finite-sample estimate $\hat{\mathbf{r}}_z$ obtained from $\{\mathbf{z}_t\}_{l=0}^{L-1}$. Define $\hat{\mathbf{R}}_z(l) = \mathbf{z}_t(l)\mathbf{z}_t^T(l)$ and form $\hat{\mathbf{r}}_z(l) = \mathbf{Q}_M \mathrm{vec}\{\hat{\mathbf{R}}_z(l)\}$ for each block l, $l = 0, 1, \ldots, L-1$. Then, the true covariance vector $\mathbf{r}_z$ of the compressive samples can be estimated as $\hat{\mathbf{r}}_z = \frac{1}{L}\sum_{l=0}^{L-1} \hat{\mathbf{r}}_z(l)$. Each summand can be expressed as $\hat{\mathbf{r}}_z(l) = \hat{\mathbf{r}}_z + \epsilon_z(l)$, where $\epsilon_z(l)$ is the estimation error from the l-th block. An unbiased estimate of $\mathbf{\Sigma}_z$ can be obtained from $\{\epsilon_z(l) = \mathbf{r}_z(l) - \hat{\mathbf{r}}_z\}$ as follows [38]:

$$\hat{\mathbf{\Sigma}}_z = \frac{1}{L-1}\sum_{l=0}^{L-1}(\mathbf{r}_z(l) - \hat{\mathbf{r}}_z)(\mathbf{r}_z(l) - \hat{\mathbf{r}}_z)^H. \tag{B.1}$$

Next, we estimate the error covariance $\mathbf{\Sigma}_s = \mathrm{E}\left\{(\hat{\mathbf{s}}_x^{(c)} - \mathbf{s}_x^{(c)})(\hat{\mathbf{s}}_x^{(c)} - \mathbf{s}_x^{(c)})^H\right\}$ where the estimate $\hat{\mathbf{s}}_x^{(c)}$ results from $\hat{\mathbf{r}}_z$ via (3.19). The inverse problem for (3.19) is solved by (3.20), but it is not convenient for performance analysis. Alternatively, we relax the ℓ_1-norm term in (3.20) by its ℓ_2-norm, which yields $\hat{\mathbf{s}}_x^{(c)} = (\mathbf{\Psi}^H\mathbf{\Psi} + \lambda\mathbf{I})^{-1}\mathbf{\Psi}^H\hat{\mathbf{r}}_z$. Defining $\mathbf{T} = (\mathbf{\Psi}^H\mathbf{\Psi} + \lambda\mathbf{I})^{-1}\mathbf{\Psi}^H$, the estimation error covariance of $\hat{\mathbf{s}}_x^{(c)}$ can be obtained from that of $\hat{\mathbf{r}}_z$ as

$$\hat{\mathbf{\Sigma}}_s = \mathbf{T}\hat{\mathbf{\Sigma}}_z\mathbf{T}^H. \tag{B.2}$$

Finally, we note from (4.3) that the cyclic spectrum vector of interest $\hat{\mathbf{c}}^{(n)}$ is given by se-

lecting $|\mathcal{I}_d^{(n)}|$ rows from $\hat{\mathbf{s}}_x^{(c)}$, for row-indices $\forall i \in \mathcal{I}_d^{(n)}$. This row selection operation can be expressed as $\hat{\mathbf{c}}^{(n)} = \mathbf{J}_n \hat{\mathbf{s}}_x^{(c)}$, where the binary-valued selection matrix $\mathbf{J}_n \in \{0,1\}^{|\mathcal{I}_d^{(n)}| \times N^2}$ is obtained from the $N^2 \times N^2$ identity matrix by retaining its rows with indices $\forall i \in \mathcal{I}_d^{(n)}$ only. With this linear relationship, the error covariance $\mathbf{\Sigma}^{(n)}$ of $\hat{\mathbf{c}}^{(n)}$ can be estimated as

$$\hat{\mathbf{\Sigma}}^{(n)} = \mathbf{J}_n \hat{\mathbf{\Sigma}}_s \mathbf{J}_n^T = \mathbf{J}_n \mathbf{T} \hat{\mathbf{\Sigma}}_z \mathbf{T}^H \mathbf{J}_n^T. \tag{B.3}$$

With (B.1) and (B.3), we have found the error covariance estimate $\hat{\mathbf{\Sigma}}^{(n)}$ that can be computed from the available compressive sample vectors $\{\mathbf{z}_t(l)\}_{l=0}^{L-1}$.