**Circuits and Systems**
Mekelweg 4,
2628 CD Delft
The Netherlands
http://ens.ewi.tudelft.nl/


# M.Sc. Thesis

# Wideband Spectrum Sensing Techniques for Wireless Sensors

**Sundeep Prabhakar Chepuri**

**Abstract**

The limited availability of radio frequency spectrum demands for more efficient ways to utilize it in future wireless networks. Spectrum sharing radios are an interesting solution to the spectral scarcity problem, where the available resources are adaptively used across time and frequency without affecting other user's transmissions. In this context, sensing the spectrum for its occupancy is needed to increase the awareness among technologies that share the same spectrum.

In a typical wireless sensor network, each node senses and transmits data constrained by a very low power budget. At the same time, they should be capable of finding a free frequency channel with minimal latency. A solution to this problem is to make radios capable of sensing multiple frequency bands, in the order of a few hundred MHz, all at once. The technical challenge lies in the design of low-complexity wideband spectrum sensing techniques that increase context awareness at the wireless node.

In this thesis, we address this problem with two approaches. The first approach is based on Compressed Sampling (CS) theory, where a new perspective is taken, different to conventional methods that estimate the spectrum and perform detection on the reconstructed spectrum. Instead a direct detection is performed on the sub-Nyquist rate sampled wideband signal. In the second part of this thesis, an alternative approach to reduce the power at an architectural level is proposed, by avoiding the Nyquist rate wideband Analog-to-Digital Converter (ADC) and pushing the conventional digital processing to the analog domain.

# Wideband Spectrum Sensing Techniques for Wireless Sensors

Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science

in

Electrical Engineering

by

Sundeep Prabhakar Chepuri
born in Bangalore, India

This work was performed in:

Wireless Group
Holst Centre
High Tech Campus 31
5656 AE Eindhoven, The Netherlands.

**Delft University of Technology**

Delft University of Technology
Department of
Microelectronics & Computer Engineering

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled **"Wideband Spectrum Sensing Techniques for Wireless Sensors"** by **Sundeep Prabhakar Chepuri** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: 28-July-2011

Chairman:

Dr. ir. Geert Leus, CAS, TU Delft

Advisor:

Dr. ir. Ruben de Francisco, Holst Centre/imec

Committee Member:

Dr. ir. Homayoun Nikookar, IRCTR, TU Delft

# Acknowledgments

I would like to thank all of them who contributed directly or indirectly to this thesis. Only with their support, this work could be accomplished.

First of all, I would like to thank my supervisor, Dr. Ruben de Francisco, for giving me the opportunity to do this work. Thank you for your enthusiasm, dedication, time, positive attitude, valuable comments and suggestions. Thanks Ruben, for being a very pleasant supervisor. I have learnt a lot from you.

I am extremely grateful to my professor, Dr. Geert Leus, for his time, support, encouragement, positive attitude, his invaluable guidance, and constructive criticism. Geert gives instantaneous response for any of your queries in spite of his busy schedule. I really learned a lot from you. Thanks Geert!

Also, I would like to thank my committee member Dr. Homayoun Nikookar for his time and interest to review this work.

I would like to thank all my colleagues and friends at the Holst Centre/imec for providing a friendly environment to work. Thank you all for both the technical and the non-technical discussions.

Finally and most importantly, I want to thank my family and friends for their support, their encouragement, their interest and their presence.

Sundeep Prabhakar Chepuri
Delft, The Netherlands
28-July-2011

# Contents

# List of Figures

# List of Tables

# Notations

| | |
|---|---|
| $x$ | Single element |
| $\mathbf{x}$ | Vector of elements |
| $\mathbb{E}(\mathbf{x})$ | Expectation operator |
| $\mathbf{X}$ | Matrix |
| $\bar{\mathbf{x}}$ | Conjugate transpose |
| $\mathbf{x}^T$ | Transpose |
| $\mathbf{x}^H$ | Hermitian |
| $\mathbf{X}^{-1}$ | Inverse of a matrix $\mathbf{X}$ |
| $\mathbf{X}^\dagger$ | Moore-Penrose pseudo-inverse of a matrix $\mathbf{X}$ |
| $tr(\mathbf{X})$ | Trace of a matrix $\mathbf{X}$ |
| $\lvert\mathbf{x}\rvert$ | Cardinality of the vector $\mathbf{x}$ |
| $\lVert\mathbf{x}\rVert_p$ | $p$-norm of the vector $\mathbf{x}$ |
| $\mathbf{I}_N$ | $N \times N$ identity Matrix |
| $\odot$ | Hadamard or element-wise product of two matrices or vectors |
| $\imath$ | $\sqrt{-1}$ |

# Abbreviations

| | |
|---|---|
| **ADC** | Analog-to-Digital Converter |
| **AIC** | Analog-to-Information Converter |
| **AWGN** | Additive White Gaussian Noise |
| **BPF** | Band-Pass Filter |
| **CD** | Compressed Detector |
| **CDMA** | Code Division Multiple Access |
| **CR** | Cognitive radio |
| **CS** | Compressed Sampling |
| **DAC** | Digital-to-Analog Converter |
| **DFT** | Discrete Fourier Transform |
| **DSA** | Dynamic Spectrum Access |
| **DSM** | Dynamic Spectrum Management |
| **FC** | Fusion Center |
| **FFT** | Fast Fourier Transform |
| **GLRT** | Generalized Likelihood Ratio Test |
| **i.i.d.** | Independently and Identically Distributed |
| **ISM** | Industrial, Scientific and Medical |
| **LLR** | Log-Likelihood Ratio |
| **LML** | Local Maximum Log-Likelihood Ratio |
| **LNA** | Low-Noise Amplifier |
| **LTE** | Long Term Evolution |
| **MBAN** | Medical Body Area Network |
| **MHT** | Multiple Hypothesis Testing |
| **ML** | Maximum Likelihood |
| **MSE** | Mean Squared Error |
| **MU** | Multi-User |
| **MWC** | Modulated Wideband Converter |
| **OFDM** | Orthogonal Frequency Division Multiplexing |
| **OMP** | Orthogonal Matching Pursuit |
| **PDF** | Probability Density Function |
| **PLL** | Phase-Lock Loop |
| **PSD** | Power Spectral Density |
| **PU** | Primary User |
| **RF** | Radio-Frequency |
| **RV** | Random Variable |
| **RIP** | Restricted Isometry Property |
| **SNR** | Signal-to-Noise Ratio |
| **w.h.p.** | With High Probability |
| **WS** | Wideband Sensing |
| **WiMAX** | Worldwide Interoperability for Microwave Access |

# 1 Introduction

Innovations in wireless communications and the increasing need for wireless devices for various applications are creating a tremendous load on the capacity of wireless networks, and demands for sustainable solutions to this continuously deteriorating problem. Almost all the usable portions of the Radio-Frequency (RF) spectrum are allocated to licensed users or Primary Users (PUs) [1], while most of it is either unused or significantly under-utilized. The average percentage spectrum occupancy can be seen in Fig. 1.1 (details regarding the measurements can be found in [2]).

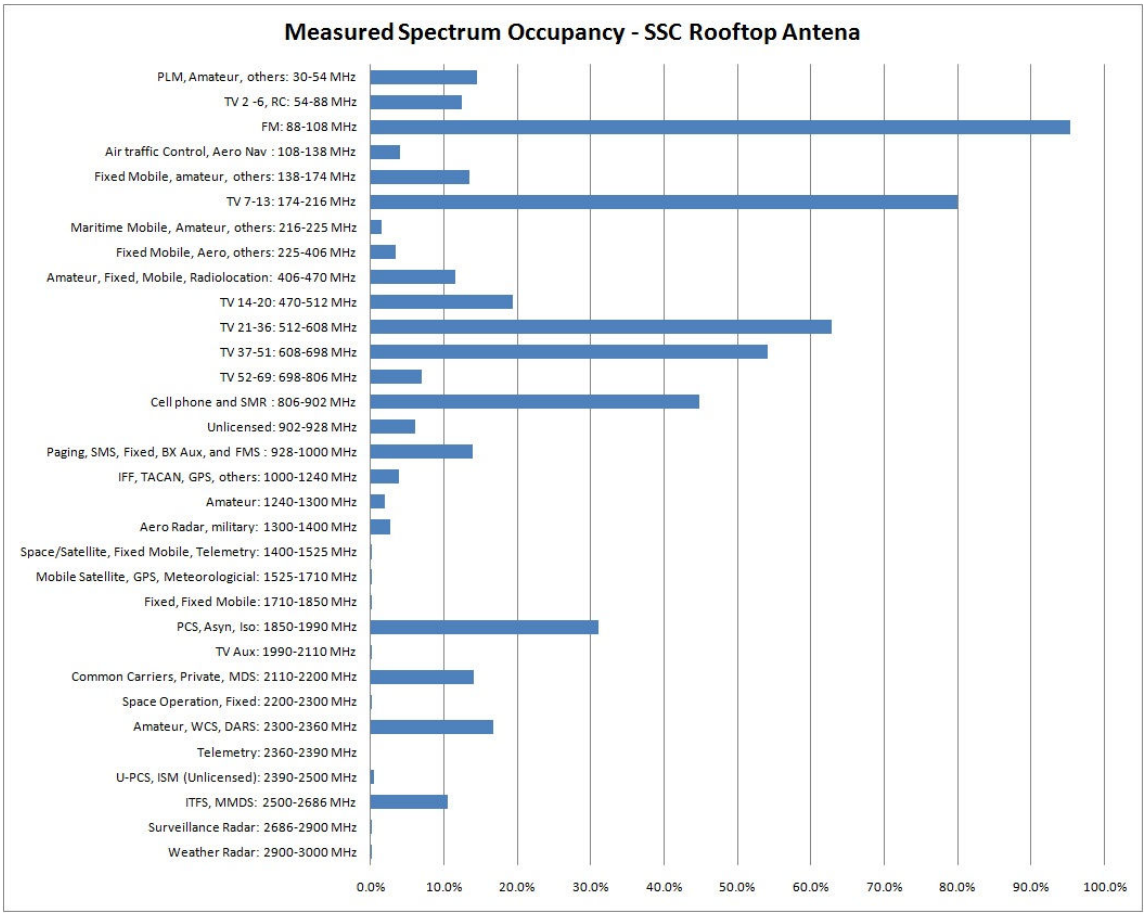



Figure 1.1: Average percentage spectrum occupancy [2].[1]

At any given instance of time there is always an unused frequency band, often termed a *spectrum hole* (which can either be non-overlapping frequency bands or a portion of

[1]Reprinted with permission.

the spectrum that is not used for a certain duration). This can be further categorized into black, gray or white spaces, corresponding to the Power Spectral Density (PSD) levels being high, medium or low [3]. In this thesis, we focus only on the black and white spaces in the spectrum. The frequency and duration of the spectrum holes can be

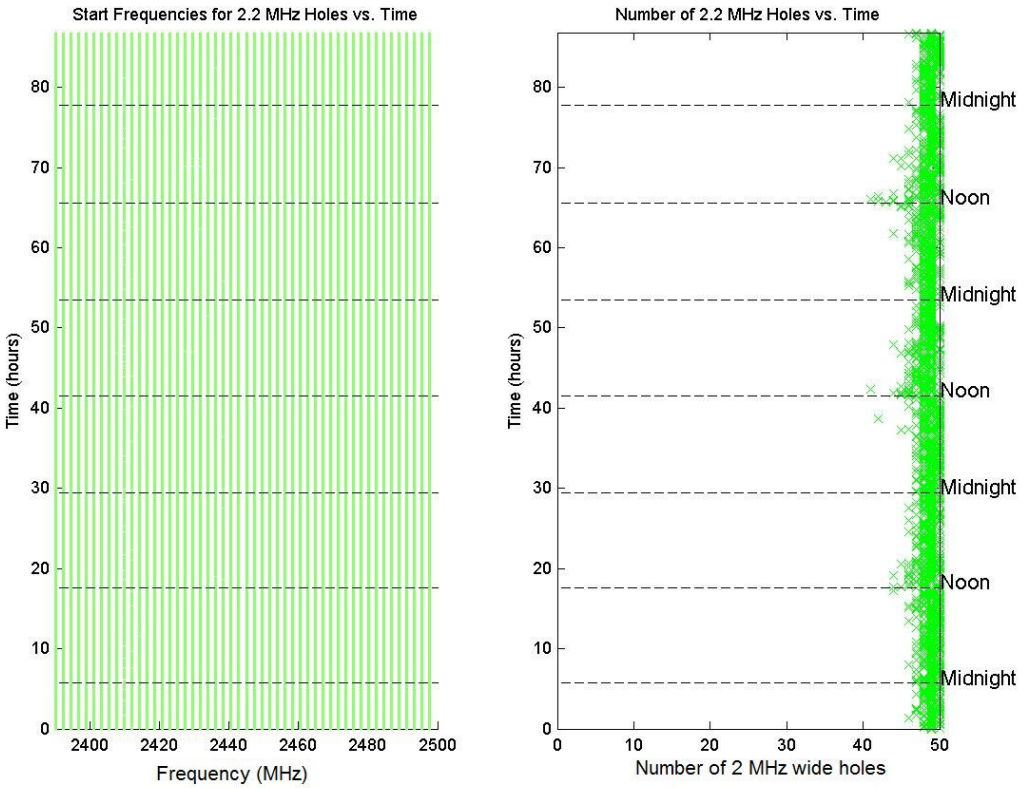



Figure 1.2: Spectrum hole analysis for 2390-2500 MHz band [2].[1]

seen in Fig. 1.2 for the 2390-2500 MHz band (plot on the left-side). The white spaces between the green vertical lines show the non-overlapping frequency bands. The number of 2 MHz wide spectrum holes can be seen in the plot (on the right-side) which varies between 45 to 50. Here we consider 2 MHz wide spectrum holes to support smallest bandwidths supported by LTE and WiMAX (i.e., 1.4 MHz and 1.75 MHz, respectively), and to support usage of IEEE 802.15.4/Zigbee in the 2360-2400 MHz lower ISM band for Medical Body Area Network (MBAN) applications [4]. For bandwidths greater than 2 MHz, groups of these continuous 2 MHz channels can be considered.

Spectrum sharing radios or Cognitive Radios (CRs) are proposed as a solution to this spectrum scarcity problem with an aim to utilize the spectrum more efficiently [5], [6]. The secondary radios should be aware of the environment and capable of sensing it. The secondary radios should learn from these tasks to utilize the spectrum adaptively and efficiently without affecting the primary radio links. Secondary radio links should

[1]Reprinted with permission.

be capable of dynamically utilizing the spectrum hole across time, frequency and space, without causing any harmful interference to PUs, and vacate the frequency band on sensing the PU activity. This reduces the harmful interference to the PUs and also improves the performance (e.g. reduces collisions) of the secondary radio link in case of licensed free bands (e.g. the 2400 MHz ISM band). Such radios enable Dynamic Spectrum Access (DSA), Dynamic Spectrum Management (DSM), co-existence with other wireless technologies and allow for secondary spectrum usage. This requires the development of wireless spectral detection and estimation techniques to sense and identify the spectrum holes.

In the following section, the motivation behind this thesis is discussed.

## 1.1 Motivation: wideband spectrum sensing

Spectrum sharing radios should be flexible to adaptively operate over a wide range of frequencies. The frequency support of a multiband signal lies within several continuous intervals spread over a wide spectrum, each consisting of a small number of narrowband transmissions as depicted in Fig. 1.3.

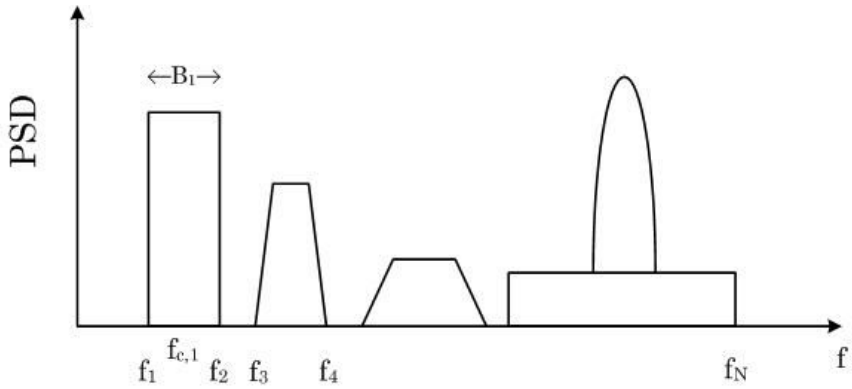



Figure 1.3: Multiband signal at the receiver.

Spectrum sensing can be classified at a very high-level into two categories, namely, narrowband sensing and wideband sensing. In narrowband sensing, the entire bandwidth is modeled as a train of consecutive sub-bands (narrowband channels) [7] and sensing can be done on these sub-bands. The detection techniques over these individual narrowband channels have been extensively studied in literature under two categories; energy detection and feature detection [8]. To detect free channels in a given wide band of interest, spectrum sensing is performed over individual narrowband channels either sequentially or at random [9] until a free channel is found. If we consider the random free channel search proposed in [9], the average number of trials to find the first free channel is given by

$$N_{trials} = \frac{N(N+1)\binom{N-1}{N_{free}-1}}{N_{free}(N_{free}+1)\binom{N}{N-N_{free}}} \tag{1.1}$$

where, $N$ indicates the total number of channels and $N_{free}$ indicates the number of free

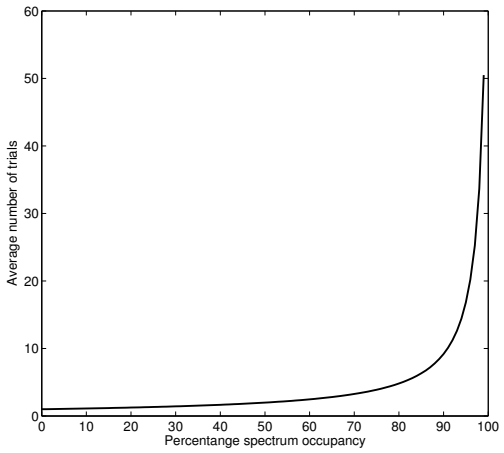
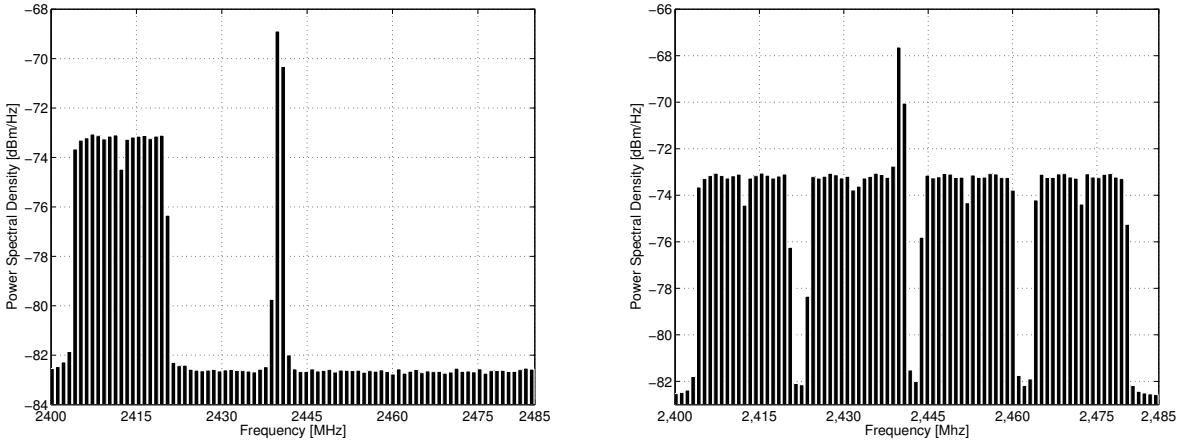



Figure 1.4: Average number of channel searches to find a free channel for different occupancies.

channels.

In Fig. 1.5, we show two scenarios in the 2400-2486 MHz ISM band, where each bar indicates a 1 MHz wide narrowband channel. We illustrate a low spectrum occupancy in Fig. 1.5a, where we have one IEEE 802.11g/WiFi node and one IEEE 802.15.4/Zigbee node. An example of a more crowded spectrum is shown in Fig. 1.5b where there are four IEEE 802.11g/WiFi nodes and one IEEE 802.15.4/Zigbee node. Depending on the occupancy of the channel, the latency in finding a free channel increases (with latency being more when the occupancy is higher), which is intuitive from this illustrations and can also be seen in Fig. 1.4 which is obtained using (1.1).




(a) Scenario 1: Low spectrum occupancy.

(b) Scenario 2: High spectrum occupancy.

Figure 1.5: Spectrum occupancy example in the 2400 MHz band.

The disadvantage of narrowband sensing method is the latency in finding a free band, since the local oscillator needs to be locked at a new frequency for every channel

search. In addition, it has an inherent power consumption for every channel search as the full receiver chain has to be powered each time a channel is sensed.

In wideband sensing, the entire band of interest is processed at once to find a free channel, with either a single Nyquist rate Analog-to-Digital Converter (ADC) or a bank of sub-Nyquist rate ADCs, both followed by digital processing. These typically consume a lot of power and radios with limited power budget cannot afford it. Compressive Sampling or Compressed Sensing (CS) [10], is a recently emerging approach for wideband sensing [11], which samples the signal at the information rate rather than at the Nyquist rate. CS requires knowledge of the sparsity level (ratio of the number of busy channels to the total number of channels). Usually, detection with CS is preceded by a coarse or a fine spectrum estimation. Estimating the spectrum using CS generally requires $\ell_1$-norm optimization and is usually carried out using high-complexity recursive algorithms (e.g., the interior point linear program solver of [12]). Alternatively, this convex optimization problem can be solved using greedy/suboptimal algorithms (e.g., Orthogonal Matching Pursuit (OMP) [13]).

In this thesis, we focus on designing techniques for wideband spectrum sensing to tackle these two key problems mentioned above. In summary, the research addresses the following questions :

1. Can we do a direct detection on the samples obtained from CS, without a fine or a coarse spectrum reconstruction?

2. Can we substantially reduce the power consumption for wideband spectrum sensing, by processing in the analog domain yet achieve a good detection performance?

Next, we discuss the challenges associated with wideband spectrum sensing.

## 1.2 Challenges in wideband spectrum sensing

Spectrum sensing is a key feature to enable the concept of spectrum sharing radios. The idea of such radios are to find out a frequency channel free of PU's signal transmissions, reliably and quickly, in order to utilize the spectrum opportunistically. Together with the reliable and quick detection of free frequency band, the secondary radios should also be aware of the re-transmissions from the PU. This means, either the PU traffic activity (typically difficult to classify) should be known to secondary radio a priori, or it should perform the spectrum sensing more frequently. To perform this on the wireless sensor nodes with a limited power budget, spectrum sensing technique has to be less complex. Spectrum sensing with constraints on reliable detection, low-latency, sensing interval and low-complexity together makes spectrum sensing not only an interesting problem, but also a challenging one. In order to find a free channel quickly, the secondary radios should be able to process the entire band of interest all at once, which needs a paradigm shift from conventional narrowband sensing engines to wideband architectures. Next, we analyze the challenges associated with spectrum sensing (wideband, in particular) under three categories :

1. **Latency and complexity**: As discussed in Section 1.1, in order to minimize the latency, the radios should adopt wideband architectures to search over multiple

frequency channels all at once. It is also necessary for the secondary radios to be aware of the PU retransmission. Hence, sensing has to be repeated at certain intervals, which also demands for low-complexity techniques, which in turn will result in power saving. Realizing low-complexity wideband sensing techniques that can be afforded by sensor nodes is a challenging task, which is addressed in this work.

2. **Reliable detection**: Even though spectrum sharing radios allow secondary spectrum usage and co-existence with other technologies, protection of the PU from the harmful interference and minimizing degradation of the PU's performance due to this secondary radio link, always has the top priority. The interference to the PU due to the secondary radio link is often measured in terms of miss-detection probability (to detect a channel as free, when the channel is actually busy). The receiver that performs sensing could be affected due to multipath, fading and shadowing in the channel, or the PU could be hidden to the sensing receiver [6]. These effects limit the detection performance and interfere with the PU. In addition to this, the receiver sensitivity plays a key role for a reliable detection. This becomes important especially while detecting nodes with lower transmit power. Receiver sensitivity decreases with an increase in the receiver bandwidth, as the receiver noise increases with the bandwidth ($N_0 = -174 + 10 \log B + NF$, where $N_0$ is the receiver noise power in dB, $NF$ is the Noise Figure and $B$ is the bandwidth in Hz). Achieving good receiver sensitivity with wideband architectures, is relatively difficult.

3. **Wideband RF front-end**: Designing a low-complexity wideband RF front-end is a challenging task and different approaches have been proposed in the literature. Multiple narrowband Band-Pass Filters (BFPs) could be employed to realize a filterbank, followed by a decision device to perform wideband sensing [14], but this architecture would require a large number of bulky components and the filter bandwidth of the BPFs (usually determined by the bank of capacitors) is preset. An alternative approach is to use a wideband Nyquist rate ADC, followed by digital processing. In order to achieve better sensitivity, the ADCs should have a higher dynamic range, which means a larger number of bits. Thus, wideband sensing requires high-rate and high resolution ADCs, which typically consume a lot of power. In case of sparse signals, the sampling rate can be relaxed and the acquisition can be done at a sub-Nyquist rate (significantly lower than the Nyquist rate). Later optimization algorithms can be used to recovery the signal without forgoing perfect reconstruction in the noiseless case. This is often referred to as a CS problem. However, current techniques demand signal recovery before detection.

## 1.3 Thesis outline and contributions

In this section, the major contributions of this thesis are highlighted. We focus on designing low-complexity techniques for wideband spectrum occupancy detection, in order to find a spectral hole with a very low latency. For this purpose, two techniques

are proposed. In the first technique, we demonstrate how to solve the signal detection problem given incoherent measurements (often the measurements are much smaller than the number of Nyquist rate samples) without reconstructing the signals under the CS framework. To do this, we propose Multiple Hypothesis Testing (MHT) under a Neyman-Pearson-like criterion to solve the compressed detection problem. In the second part of this thesis, a low-power wideband spectrum sensing architecture based on analog processing is proposed as an alternative approach for multiband occupancy detection.

Next, we describe the content of the thesis chapter by chapter.

**Chapter 2: Background:** In this chapter, we provide a brief survey of the existing approaches to wideband spectrum sensing, along with their computational complexity order. Next, we introduce the CS framework, the techniques proposed in literature to acquire sub-Nyquist rate samples, and the conventional approach involving detection on the compressive estimate of the signal. At the end of this chapter, we discuss the MHT problem and the necessary background required to understand the MHT problem under the Neyman-Pearson criterion.

**Chapter 3: Wideband sensing through multiple hypothesis testing:** The current literature on CS has focused almost extensively on problems in sparse signal reconstruction and estimation. In this chapter, we focus on direct signal detection using sub-Nyquist rate samples, with application to multiband occupancy detection. We formulate a MHT problem under a Neyman-Pearson-like criterion to solve the detection problem. To understand the MHT more clearly, we first develop the detector for a signal sampled at the Nyquist rate (complete frequency information is available). Next, this is extended to reduced dimensionality signals acquired at sub-Nyquist rate (complete frequency information in not available).

**Chapter 4: Low-power wideband spectrum sensing architecture:** As an alternative approach to multiband spectral occupancy detection under the CS framework, we propose an architecture at the system level to reduce the power consumption. Here the conventional digital processing is done in the analog domain to reduce the power consumption. The performance analysis of the proposed system is done in terms of power saving, spectrum estimation and detection.

**Chapter 5: Conclusions:** This chapter summarizes the major results of this work and gives suggestions for future research.

# 2 Background

This thesis focuses on the physical node level paradigms for wideband spectrum sensing, where the aim is to find an available frequency channel free from signal transmission within a wide spectral range. This has gained a lot of interest recently due to its importance in the field of spectrum sharing radio networks. In such wireless networks, the radio should be capable of sensing bandwidths of the order of a few hundred MHz with low-latency and low-complexity. This chapter gives a brief literature survey of the existing approaches to wideband sensing. The necessary background about the Compressive Sampling or Compressed Sensing (CS) framework along with its mathematical formulation is provided. The proposed Multiple Hypothesis Testing (MHT) problem in the Neyman-Pearson approach is introduced as a background to solve the compressed detection problem that is addressed in this thesis. The following section provides a brief overview of existing techniques to wideband spectrum sensing.

## 2.1 Existing approaches to wideband spectrum sensing

The current research on wideband spectrum sensing can be categorized based on two signal acquisition techniques; Nyquist rate and sub-Nyquist rate. In Nyquist rate signal acquisition techniques, the signal is sampled at the Nyquist rate according to the Whittaker-Kotelnikov-Shannon-Nyquist theorem. According to this theorem, a signal with a frequency support between $-f_i$ and $f_i$ can be perfectly recovered from its samples if the sampling rate is more than or equal to the Nyquist rate, i.e, $2f_i$. In the sub-Nyquist rate approaches for spectrum sensing, properties like sparsity of the spectrum, or the edge spectrum, are used to relax the sampling rate requirements. These are often casted into a CS problem. Sequential narrowband sensing, is an alternative approach by sampling individual narrowband channels, but has more latency and inherent power consumption as discussed in Section 1.1.

In [15], the entire wide bandwidth of interest is modeled as a union of subbands. The task of spectrum sensing is treated as a spectral edge detection problem with wavelet-based techniques for detecting irregular edges in the signal Power Spectral Density (PSD). These edges were used to characterize the number of subbands, their locations and the intensity of the spectral usage to enable opportunistic sharing. In [9], a gradient based search technique with an adaptive step size was proposed to identify the white spaces in the spectrum. This technique proposes a sequential empty channel search in the frequency domain, and reduces the complexity by computing only the required frequency domain coefficient instead of all the coefficients. Sampling signals of very large bandwidth at Nyquist rate requires power-hungry high-rate Analog-to-Digital Converters (ADCs).

A hot favorite tool in signal processing, CS, is used to reduce the requirements on

the ADCs and decrease the latency associated with spectral sensing. The sampling rate is often reduced using techniques like multicoset sampling or modulated wideband converters (discussed in more detail in the next section), and the signals can be reconstructed perfectly without sacrificing much information. The average minimum sampling rate required to perfectly reconstruct the multiband signal is equal to the Nyquist rate multiplied by the frequency occupancy given by the Landau lower bound (as in [16]). This means that an upper-bound on the sparse support of the spectrum should be known before hand. The reconstruction can be achieved either with one of the famous sparse recovery techniques (e.g., basis pursuit [17]), or with traditional spectral estimation techniques like multiple signal classification (MUSIC), or the minimum variance distortionless response (MVDR) method [18]. In [19], a sub-Nyquist rate sampling technique was proposed to reconstruct the power spectrum of the signal, without any constraints on the power spectrum. Once the compressive estimate of the signal is obtained which could be either coarse or fine, the occupancy of the bands can be detected with a threshold [20], [11].

There are other approaches for multiband spectrum sensing under the co-operative sensing model, or distributed detection framework as in [21]. However, we restrict ourselves to sensing with a single node in this thesis. In the following subsection, we discuss the complexity of some of the sensing algorithms.

### 2.1.1 Computational complexity

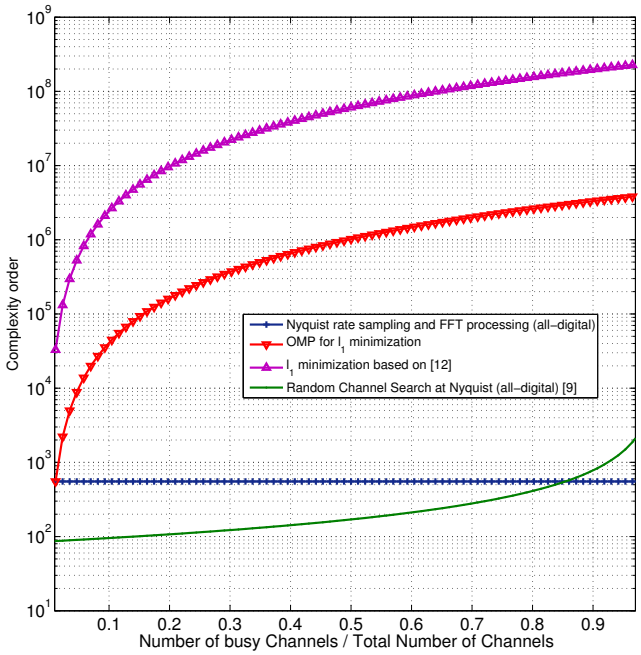



Figure 2.1: Complexity order of existing algorithms.

To understand the computational complexity of wideband sensing we consider the following techniques. First, the traditional way of Nyquist rate sampling using a high-

rate ADC, followed by digital FFT processing to reconstruct the spectrum is considered. To compute all the frequency coefficients efficiently would require $O(N \log N)$ computations, for $N$ channels. Next, we consider a random channel search, still using high-rate Nyquist rate sampling as in [9]. This technique needs $O(N_{trial}N)$ computations, where $N_{trials}$ is the average number of trials required to find an empty channel computed using equation (1.1). In general, sparse signal reconstruction under the CS framework would require high-complexity recursive algorithms to solve the $\ell_1$-norm minimization problem. Here, the interior point linear program solver [12] is considered. The optimization problem can also be solved in a sub-optimal way using greedy algorithms (e.g., Orthogonal Matching Pursuit (OMP)).

The computational complexity order of these sensing techniques for different spectral occupancy levels is shown in Fig. 2.1. Even though Nyquist rate sampling appears to be the best compromise for wideband sensing, performing all-digital processing consumes a lot of power. Hence, there is a need for wideband sensing algorithms with lower complexity or lower power consumption, especially in scenarios where the spectrum occupancy is higher. This is addressed in Chapter 4 of this thesis.

In the following section, the CS framework along with its mathematical formulation is discussed.

## 2.2 Compressive sensing

Compressive sensing is a method in which signals are acquired through a set of a few non-adaptive[1] linear measurements and reconstructed efficiently from this incomplete set of measurements [10]. Usually in signal processing, the entire signal is first acquired and then compressed later on, either to store or to transmit. This often needs an enormous effort to acquire the entire signal which could be avoided, as the insignificant information is thrown away during the compression. The key motivation behind CS is to combine both the signal acquisition and the compression process by directly sensing the essential part of the signal using fewer linear measurements. From linear algebra it is known that it is not possible to reconstruct an arbitrary signal from an incomplete set of linear measurements. To enable reconstruction under CS, the signals are constrained to be *sparse* in some basis, i.e., those signals that only have a few non-zero coordinates in some basis (the sparsifying basis).

Sparse signals lie in a lower dimensional subspace, which may be represented by a few linear measurements. However, it is difficult to determine which lower dimension subspace the signal lies in. In other words, we may know the signal has a few non-zero coordinates, but the locations of these coordinates are not known. This makes the sparse recovery techniques more complex. To aid the sparse recovery, CS techniques rely on the *incoherence* between the measurement basis and the sparsifying basis. The following subsection presents the mathematical formulation of the CS problem.

[1]Non-adaptive means the measurement process does not depend on the signal being measured.

### 2.2.1 Problem formulation

The CS framework relies on the sparsity of the signal. Hence, first we need to quantify the sparsity of a vector.

**Definition 2.1.** Sparsity [22]. *An $N$-dimensional vector, $\mathbf{s} \in \mathbb{C}^N$ is said to be $K$-sparse if it has $K$ or fewer non-zero coordinates, i.e.,*

$$\|\mathbf{s}\|_0 := \lim_{p \to 0} \quad (|s_1|^p + |s_2|^p \cdots + |s_N|^p)^{\frac{1}{p}} \leq K \ll N, \quad \mathbf{s} \in \mathbb{C}^N \tag{2.1}$$

*where, $\|\cdot\|_0$ denotes the $\ell_0$-norm which just counts the number of non-zero components in the vector and $p$ is a constant that is traditionally used to parameterize the standard $\ell_p$-norm.*

In practice, the signals often encountered are not exactly sparse, but are compressible (close to being sparse).

**Definition 2.2.** Compressibility [23]. *A vector is called compressible if its entries obey a power decay law*

$$|s|_{(i)} \leq R_r i^{-r} \tag{2.2}$$

*where, $|s|_{(i)}$ is the ith largest value of $\mathbf{s}$, i.e., $(|s|_{(1)} \geq |s|_{(2)} \geq \cdots \geq |s|_{(N)})$, $r > 1$, and $R_r$ is a positive constant which depends only on $r$.*

This means only a few entries of a compressible vector are large while most of them are small. It should be noted that sparse signals are compressible.

Let us consider a discrete-time signal $\mathbf{x} \in \mathbb{C}^N$, which we can expand in terms of an $N \times 1$ orthonormal basis (e.g. wavelet, Fourier) vectors $\boldsymbol{\psi}_{xi}$ with $i = 1 \ldots N$ as

$$\mathbf{x} = \sum_{i=1}^{N} s_i \boldsymbol{\psi}_{xi} \tag{2.3}$$

where, $s_i$ with $i = 1 \ldots N$ are the entries of the coefficient sequence of $\mathbf{x}$. Alternatively, stacking $\boldsymbol{\psi}_{xi}$ for $i = 1 \ldots N$ as columns results in an $N \times N$ *sparsifying basis* matrix $\boldsymbol{\Psi}_x = \begin{bmatrix} \boldsymbol{\psi}_{x1} & \boldsymbol{\psi}_{x2} \cdots \boldsymbol{\psi}_{xN} \end{bmatrix}$. Then (2.3) can be rewritten in matrix-vector form as

$$\mathbf{x} = \boldsymbol{\Psi}_x \mathbf{s} \tag{2.4}$$

In case of the Fourier basis, $\boldsymbol{\Psi}_x = \mathbf{F}^H$, where $\mathbf{F}$ is be the Discrete Fourier Transform (DFT) matrix, i.e. $\mathbf{F} = \exp(-\imath \frac{2\pi kn}{N})$, with $k, n = 0, \ldots, N-1$ and $\imath = \sqrt{-1}$. Hence, the vector $\mathbf{s} = \mathbf{x}_f$ will be the frequency response of $\mathbf{x}$. In this thesis, the Fourier basis is used as a sparsifying basis with the notation

$$\mathbf{x} = \boldsymbol{\Psi}_x \mathbf{x}_f \tag{2.5}$$

Sensing, i.e. of the time domain signal $\mathbf{x}$ is done by collecting measurements by correlating $\mathbf{x}$ with some sensing vectors $\boldsymbol{\phi}_i$ (waveforms in case of the continuous-time domain), i.e.,

$$\tilde{x}_i = \langle \mathbf{x}, \boldsymbol{\phi}_i \rangle, \qquad i = 1, 2, \ldots, M. \tag{2.6}$$

Depending on the sensing vectors the entries $\tilde{x}_i$ of the vector $\tilde{\mathbf{x}} \in \mathbb{C}^N$ will be the Fourier coefficients (if the sensing vectors are sinusoids), or a subsampled vector (for Dirac delta sensing vectors). Stacking the sensing vectors $\boldsymbol{\phi}_i$ as columns results in a *sensing matrix*[1], $\boldsymbol{\Phi} \in \mathbb{R}^{M \times N}$. In that case, (2.6) can be rewritten in matrix-vector form as

$$\tilde{\mathbf{x}} = \boldsymbol{\Phi}\mathbf{x} = \boldsymbol{\Phi}\boldsymbol{\Psi}_x\mathbf{x}_f = \tilde{\boldsymbol{\Psi}}_x\mathbf{x}_f \tag{2.7}$$

where $\tilde{\boldsymbol{\Psi}}_x \in \mathbb{C}^{M \times N}$ is the new basis of the signal $\tilde{\mathbf{x}}$. In the case of CS the number of available measurements $M$ is much smaller than the dimension $N$ of the signal $\mathbf{x}$, resulting in an under-determined system. However, the sparsity constraint helps recovering the signal exactly with high probability (w.h.p.).

The sparse vector $\mathbf{s}$ can be recovered w.h.p. (for some $\epsilon > 0$) by solving the optimization problem

$$\min_{\mathbf{s}} \quad \|\mathbf{s}\|_0 \quad \text{s.t.} \quad \|\boldsymbol{\Phi}\boldsymbol{\Psi}_x\mathbf{s} - \tilde{\mathbf{x}}\|_2^2 < \epsilon \tag{2.8}$$

Even though the non-convex $\ell_0$-optimization problem in (2.8) works perfectly in theory, it is not numerically feasible and is NP-Hard (non-deterministic polynomial-time hard) in general [24]. An extensive research on CS has resulted in numerous alternatives to this problem.

Assuming the coefficient vector $\mathbf{x}_f$ (or $\mathbf{s}$) is $K$-sparse, then $\mathbf{x}_f$ and, hence $\mathbf{x} = \boldsymbol{\Psi}_x\mathbf{x}_f$, can be recovered from $\tilde{\mathbf{x}}$, if the matrices $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}_x$, and $\tilde{\boldsymbol{\Psi}}_x$ satisfy certain properties. For exact signal recovery, it is clear that the measurement process should not damage the information content in the signal, and should just reduce the dimensionality of the signal ($\mathbf{x} : \mathbb{C}^N \mapsto \mathbb{C}^M$). In addition to the sparsity constraint, CS relies on *incoherence* between the sensing matrix and the sparsifying basis.

**Definition 2.3.** Mutual coherence [22]**.** *The mutual coherence between two orthonormal bases (e.g., the sensing matrix $\boldsymbol{\Phi}$ and the sparsifying basis $\boldsymbol{\Psi}_x$) of $\mathbb{R}^N$ is*

$$\begin{aligned} \mu(\boldsymbol{\Phi}, \boldsymbol{\Psi}_x) &= \sup\{|\langle \boldsymbol{\phi}_i, \boldsymbol{\psi}_{xi} \rangle| : \boldsymbol{\phi}_i \in \boldsymbol{\Phi}, \boldsymbol{\psi}_{xi} \in \boldsymbol{\Psi}_x\} \\ &\in [1, \sqrt{N}] \end{aligned} \tag{2.9}$$

The mutual coherence between two bases is quantified using the parameter $\mu$. The two bases are said to be mutually incoherent if they have small values of $\mu$ (closer to 1), which further guarantees the possibility of signal recovery [25]. One more very useful tool called the *Restricted Isometry Property* (RIP) is used for determining the sufficient condition that guarantees sparse signal recovery w.h.p., and also to study the general robustness of CS [26], [22].

**Definition 2.4.** Restricted isometry property [22]**.** *For each integer $K = 1, 2, \ldots,$ the isometry constant $\delta_K$ of a matrix $\tilde{\boldsymbol{\Psi}}_x$ is defined as the smallest number such that*

$$(1 - \delta_K)\|\mathbf{s}\|_2^2 \leq \|\tilde{\boldsymbol{\Psi}}_x\mathbf{s}\|_2^2 \leq (1 + \delta_K)\|\mathbf{s}\|_2^2 \tag{2.10}$$

*holds for all $K$-sparse vectors $\mathbf{s}$.*

[1]Although similar results hold for the sensing matrix $\boldsymbol{\Phi} \in \mathbb{C}^{M \times N}$, for simplicity we consider measurements taken over $\mathbb{R}$-space and also assume that the measurement process is noise free.

It can be loosely said that a matrix $\tilde{\mathbf{x}}$ obeys the RIP of order $K$ if $\delta_K$ is not too close to 1. When RIP holds, the Euclidean length of $K$-sparse signals is approximately preserved by $\tilde{\mathbf{\Psi}}_x$. This means that the $K$-sparse vectors cannot be in the nullspace of $\tilde{\mathbf{\Psi}}_x$. Some of the matrices that satisfy the RIP (of order 2), i.e. matrices with column vectors taken from arbitrary subsets being nearly orthogonal [10], [22], [26] are :

1. *Gaussian*: Matrix whose elements are drawn independently and identically distributed (i.i.d.) from a random Gaussian distribution of zero mean and variance $\frac{1}{M}$.
2. *Bernoulli*: Matrix whose elements are drawn i.i.d. from a symmetric Bernoulli distribution ($Pr(\phi_{i,j} = \pm\frac{1}{\sqrt{M}}) = \frac{1}{2}$).
3. *Random selection*: Matrix constructed by selecting $M$ columns uniformly at random from $\mathbf{I}_N$ and multiplied with a normalization factor of $\sqrt{\frac{N}{M}}$.
4. *Toeplitz Gaussian*: A Toeplitz structured matrix with elements drawn from the same distribution as the Gaussian matrix above.

If the above conditions are met, the $\ell_0$-norm in (2.8) can be relaxed and instead can be reformulated as

$$\min_{\mathbf{s}} \quad \|\mathbf{s}\|_p \quad \text{s.t.} \quad \|\mathbf{\Phi}\mathbf{\Psi}_x\mathbf{s} - \tilde{\mathbf{x}}\|_2^2 < \epsilon \tag{2.11}$$

Ridge regression [27] (i.e. Tikhonov regularization) uses $p = 2$, while basis pursuit [17] and LASSO [28] use $p = 1$, to solve this optimization problem. The other approaches include greedy algorithms such as Orthogonal Matching pursuit [13], Stagewise Orthogonal Matching pursuit [29], or Iterative re-weighted algorithms [30]. Most of these algorithms calculate the support of the signal iteratively, and work for a specific number of measurements

$$M = cK\log\left(\frac{N}{K}\right) \tag{2.12}$$

where $c$ is the over-measuring factor ($c > 0$, varies between 2 and 20 [10] depending on the recovery algorithm).

To practically realize the CS and mitigate the high sampling rate problems sub-Nyquist rate sampling architectures have been proposed. These are more often referred to as Analog-to-Information Convertors (AICs) by the CS research community. The sampling from these architectures could be modeled by a sensing matrix $\mathbf{\Phi}$. In [31] and [32] an AIC architecture based on random sampling, for signals sparse in the Fourier basis was proposed. This consists of a bank of parallel low-rate Analog-to-Digital Converters (ADCs), which were enabled at random. This can be viewed as a *random selection* compression matrix. An AIC based on random filtering was proposed in [33]. In this architecture the signal is convolved with a random-tap filter (of certain length), and then the filtered signal is downsampled by a factor $\lfloor\frac{N}{M}\rfloor$. The filter taps were drawn either from a Gaussian or Bernoulli distribution. In [34], [35], and [36] AICs were realized by pseudo-random demodulation. Here, the signals were initially spread with a high-rate pseudo-random sequence. The spread signal is low-pass filtered, and sampled at a relatively lower rate. In [16], sub-Nyquist sampling based on multicoset

sampling was proposed for the case of multiband signals (similar to the signal shown in Fig. 1.3). Similarly, a non-uniform sampling for multiband analog signals called Modulated Wideband Converter (MWC) was proposed in [37]. This architecture consists of multiple channels, each consisting of different demodulating mixing functions followed by low-pass filtering and low-rate uniform sampling. Unlike [16], the architecture in [37] does not impose any limitation on the knowledge of the band locations for recovery of multiband signals.

The following section presents the hypothesis testing problem that is used for compressed detection.

## 2.3 Multiple hypothesis testing

In case of occupancy detection, we are interested in detecting the presence or absence (event) of a primary user (PU) signal in a particular band. Often in detection theory this is formulated as a hypothesis testing problem to decide which event is more likely to occur. The absence of the signal (noise-only case) is referred to as the *null hypothesis* denoted by $\mathcal{H}_0$ and the presence of the signal (signal-plus-noise case) as the *alternative hypothesis* denoted by $\mathcal{H}_1$. This problem is known as the *binary* hypothesis test, since we must choose between two hypotheses.

A multiband spectrum will have all possible combinations of the frequency bands being free (indicated by "0") and the frequency bands being occupied (indicated by "1"). The detection of spectrum occupancy for multiband signals all at once, would require solving multiple alternative hypotheses. To illustrate such a scenario we consider the toy example described in Table 2.1. For two channels, there are 4 possible states of them being free and/or busy as indicated by hypotheses $\mathcal{H}_0$, $\mathcal{H}_1$, $\mathcal{H}_2$, and $\mathcal{H}_3$. Similarly a channel length of $N$ would result in one null hypothesis $\mathcal{H}_0$, and $m = 2^N - 1$, alternative hypotheses indicated by $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_m$. Such a hypothesis testing problem is referred to as a Multiple Hypothesis Testing (MHT) problem. Solving the MHT problem would give the occupancy of all the $N$ channels at once, and would avoid solving a binary hypothesis problem on every channel.

Table 2.1: A toy example to illustrate MHT for multiband sensing

| **Hypotheses** | **Channel 1** | **Channel 2** |
|---|---|---|
| $\mathcal{H}_0$ | 0 | 0 |
| $\mathcal{H}_1$ | 0 | 1 |
| $\mathcal{H}_2$ | 1 | 0 |
| $\mathcal{H}_3$ | 1 | 1 |

The MHT problems arises in various applications including wireless communications, radar, pattern recognition, classification, etc. The optimal Bayesian detector for MHT is a well-known result in classical detection theory [38], [39]. In [40], MHT was tackled sequentially in the context of multiple resolution element radar. The problem of classifying non-stationary Gaussian signals using MHT with a constraint on the probability of incorrect classification of one class was studied in [41].

In the Bayesian framework, one typically assigns probabilities to various hypotheses. This means, there is a prior belief in the likelihood of the hypotheses. This is a reasonable assumption for e.g. digital communications where the transmission of a "0" is equally likely as the transmission of a "1". However, this is not a valid assumption in case of impulsive event detection which happens with radars or occupancy detection. Hence people often resort to the Neyman-Pearson criterion, where there is no assumption on the prior probabilities of a certain hypothesis.

**Theorem 2.1** (Neyman-Pearson approach to the binary hypothesis problem)**.**

$$\begin{aligned} \max \quad & Pr(\mathcal{H}_1; \mathcal{H}_1) \\ s.t. \quad & Pr(\mathcal{H}_1; \mathcal{H}_0) = \alpha. \end{aligned} \tag{2.13}$$

*where the notation $Pr(\mathcal{H}_i; \mathcal{H}_j)$ indicates the probability of deciding hypothesis $\mathcal{H}_i$ when hypothesis $\mathcal{H}_j$ is true.*

The classical Neyman-Pearson approach to binary hypothesis testing suggests to maximize $Pr(\mathcal{H}_1; \mathcal{H}_1)$ with an upper-bound constraint on $Pr(\mathcal{H}_1; \mathcal{H}_0)$. The $Pr(\mathcal{H}_1; \mathcal{H}_0)$ can be constrained by choosing an appropriate threshold for the decision [39]. We now define the Likelihood Ratio (LR), which is required to understand the MHT.

**Definition 2.5.** Likelihood ratio [39]**.** *Let $\mathbf{y} \in \mathbb{C}^N$ be an observed vector of i.i.d. random variables from a certain distribution. The* likelihood ratio *is then given by*

$$\Lambda = \frac{Pr(\mathbf{y}|\mathcal{H}_1)}{Pr(\mathbf{y}|\mathcal{H}_0)} \tag{2.14}$$

The function $\Lambda$ indicates for each value of $\mathbf{y}$ the likelihood of $\mathcal{H}_1$ versus the likelihood of $\mathcal{H}_0$. The logarithm of the function $\Lambda$ is referred to as the Log-Likelihood Ratio (LLR).

Although MHT can be formulated in the Neyman-Pearson sense [42] it is seldom used in practice. The proposed extension of the classical Neyman-Pearson approach for binary hypothesis testing to MHT is discussed in Chapter 3.

As discussed in Section 2.1, the classical approaches for detection based on CS, recover the signal first by solving the optimization problem given in equation (2.11). Once the signal is estimated (either coarse or fine) detection is performed. For example this could be a simple threshold to just detect the presence or the absence of a signal. We try to avoid this two-step approach with redundant estimation of the signal, and do a direct detection on the reduced dimensionality signal. We formulate this as a MHT problem. This could be seen as an $N$-dimensional signal classifier with respect to a reference signal (noise). This problem becomes more interesting when only $M$-dimensional observations ($M < N$) are available.

The conclusions of this chapter are presented in the following section.

## 2.4 Conclusions

In this chapter, we provided a brief literature survey of existing approaches to wideband spectrum sensing, which could be categorized based on the signal acquisition technique.

Typically, high-rate Nyquist rate ADCs consume a lot of power, and the requirements on the ADCs could be relaxed, by acquiring the signal at the information rate rather than the Nyquist rate. This is often formulated as a CS problem. The necessary background on CS has been provided. However, CS constrains the signals to be sparse and demands the knowledge of the sparse support for recovery. The algorithms for sparse recovery are in general complex and the classical approach includes estimation of the signal even for detection. With the motivation to perform direct detection on compressed samples, the necessary background required for the MHT problem under the Neyman-Pearson criterion has been discussed in this chapter.

# Wideband sensing through multiple hypothesis testing

# 3

## 3.1 Introduction

Efficient spectrum sharing can be achieved in wireless sensor networks through wideband spectrum sensing, by identifying available channels within a large frequency range. In order to sense large bandwidths, in the order of a few hundred MHz, high-rate Analog-to-Digital Converters (ADC) or complex receiver front-ends are required. Alternatively, the wideband channel can be seen as a train of narrowband channels, where detection can be performed on the individual channels. Depending on the channel occupancy, this may increase the latency in finding a free channel and also consume more power, as the full receiver chain has to be powered each time a channel is sensed.

Currently, there is a great interest in reducing the sampling rate for sparse signals and relax the requirements on the ADCs. These are often casted as a Compressive Sensing (CS) problem, where the data is acquired at a rate significantly lower than the Nyquist rate. The signal can be recovered with one of the many available sparse recovery algorithms with a little or no loss of information. However, an important aspect is that in order to solve a detection problem the signal or its statistical measures need not be reconstructed. Instead, the detection can be performed directly on the reduced set of data samples. Such detection problems appear in various fields such as event detection in radar, Multi-User (MU) detection in communications, imaging, and spectrum sensing for Cognitive Radios (CRs).

Here, we consider a multiband occupancy detection problem. In a multiband spectrum, each band could be either busy or free. The multiband occupancy is combinatorial in the number of channels $N$. We formulate this detection problem as a Multiple Hypothesis Testing (MHT) problem, with each hypothesis describing one possible combination of all the channels being busy and/or free. This is explained with an example in Table. 2.1 of Chapter 2. Solving the MHT problem gives the occupancy of $N$ channels all at once.

In this chapter, we develop a detector for two cases:

- When the signal is acquired using Nyquist rate sampling, i.e., $M = N$, where $M$ is the number of available measurements.
- When the signal is acquired using sub-Nyquist rate sampling, i.e., $M < N$ (reduced dimension).

The detector for $M = N$ is of linear complexity, but still would require Nyquist rate sampling. The detector for the case $M < N$ is of more practical interest, resulting in a Compressed Detector (CD).

An optimal algorithm for $M = N$ case is proposed. For the $M < N$ case, the optimal detector is of complexity order $O(2^N)$. Hence, we propose a sub-optimal greedy

algorithm. The performance of these detectors is analyzed through simulations in terms of the detection probability, the false alarm probability and the compression rate, which is defined as the ratio of the number of the available measurements to total number of Nyquist rate measurements, i.e., $\frac{M}{N}$.

In the following section, we present the detection model considered for the MHT problem.

## 3.2 Detection model

In this chapter, we propose a detector based on the MHT problem for multiband sensing. Before presenting the MHT using the Neyman-Pearson approach, we define the probability of detection, $P_d$, and the probability of false alarm, $P_{fa}$, which are used to analyze the detection performance for wideband sensing.

**Definition 3.1.** Probability of detection and probability of false alarm**.** *Consider a wideband spectrum of $B$ Hz segmented into $N$ channels, such that each channel has a bandwidth $\frac{B}{N}$ Hz. The channels are indexed from $1, \ldots, N$. These channels can be either busy or free, depending on whether there is or there is no Primary User (PU) signal transmission, respectively. The indices of such $N_{busy}$ busy channels are collected in a vector*

$$\mathbf{b} = \begin{bmatrix} b_1 & b_2 \cdots b_{N_{busy}} \end{bmatrix}^T \tag{3.1}$$

*with $|\mathbf{b}| = N_{busy}$. The complement of the set $\mathbf{b}$ is denoted by $\mathbf{b}^c$,*

$$\mathbf{b}^c = \begin{bmatrix} b_1^c & b_2^c \cdots b_{N_{free}}^c \end{bmatrix}^T \tag{3.2}$$

*with $|\mathbf{b}^c| = N - N_{busy} = N_{free}$. The detected busy channel set, $\hat{\mathbf{b}}$, is determined by solving the hypothesis testing problem. The probability of detection is then defined as*

$$P_d = \frac{1}{N_{busy}} \sum_{j=1}^{N_{busy}} Pr(b_j \in \hat{\mathbf{b}} | b_j \in \mathbf{b}) \tag{3.3}$$

*and the probability of false alarm as*

$$P_{fa} = \frac{1}{N_{free}} \sum_{j=1}^{N_{free}} Pr(b_j^c \in \hat{\mathbf{b}} | b_j^c \in \mathbf{b}^c) \tag{3.4}$$

For example consider the following channel occupancy $\begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}^T$, with $N = 6$. Here the busy channel set will be $\mathbf{b} = \begin{bmatrix} 3 & 4 \end{bmatrix}^T$ and its complement set is $\mathbf{b}^c = \begin{bmatrix} 1 & 2 & 5 & 6 \end{bmatrix}^T$, with $N_{busy} = 2$ and $N_{free} = 4$.

We next present the proposed extension of the classical Neyman-Pearson approach for binary hypothesis testing to MHT.

**Proposition 3.1.** MHT under the Neyman-Pearson-like criterion**.** *Let $\mathbf{y} \in \mathbb{C}^N$ be an observed vector of independently and identically distributed (i.i.d.) random variables*

*from a certain distribution. Let the null hypothesis be denoted by, $\mathcal{H}_0$, and the alternate hypotheses by, $\mathcal{H}_i$, with $i = 1, \ldots, 2^N - 1$. Then the MHT detector in the Neyman-Pearson sense is the most likely hypothesis with respect to the null hypothesis, that would optimize the probability of detection, $P_d$, of (3.3) with a constraint on the probability of false alarm, $P_{fa}$ (as defined in (3.4)). However, this detector is complicated to derive. The simplified MHT detector under the Neyman-Pearson-like criterion is*

$$i^* = \arg\max_i \quad \Lambda(i) = \ln\left(\frac{Pr(\mathbf{y}|\mathcal{H}_i)}{Pr(\mathbf{y}|\mathcal{H}_0)}\right) \quad \textit{with} \quad i \in \{1, \ldots, m = 2^N - 1\}. \tag{3.5}$$

$$\Lambda(i^*) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{i^*}}{\gtrless}} \gamma_{th} \tag{3.6}$$

In case of multiple alternative hypotheses, it is not required to consider $\mathcal{H}_0$ as a hypothesis to be explicitly tested, but, rather used as a reference or a dummy hypothesis [43]. Hence testing $\mathcal{H}_1$ against $\mathcal{H}_2$ can be accomplished by comparing the Likelihood Ratios (LRs)

$$\frac{Pr(\mathbf{y}|\mathcal{H}_1)}{Pr(\mathbf{y}|\mathcal{H}_0)} \lessgtr \frac{Pr(\mathbf{y}|\mathcal{H}_2)}{Pr(\mathbf{y}|\mathcal{H}_0)} \tag{3.7}$$

This can be generalized to $i$-hypotheses [44]. The optimization problem (3.5) will result in the most likely hypothesis with respect to the null hypothesis. We choose a threshold, $\gamma_{th}$, based on simulations, to maintain a certain $P_{fa}$ of (3.4), and achieve the desired $P_d$, defined in (3.3).

In the next section, we present the signal model considered to solve the detection problem.

## 3.3 Signal model

Let the time domain signal representing $N$ frequency channels be denoted by the $N \times 1$ vector $\mathbf{x} \in \mathbb{C}^N$ and the noise by the $N \times 1$ vector $\mathbf{v} \in \mathbb{C}^N$. The signal $\mathbf{x}$ can be written in terms of its frequency response $\mathbf{x}_f \in \mathbb{C}^N$ as $\mathbf{x} = \mathbf{F}^H \mathbf{x}_f$, where $\mathbf{F} \in \mathbb{C}^{N \times N}$ is the normalized Discrete Fourier Transform (DFT) matrix. Similarly, the noise $\mathbf{v}$ can be written in terms of its frequency response $\mathbf{v}_f \in \mathbb{C}^N$ as $\mathbf{v} = \mathbf{F}^H \mathbf{v}_f$. The Fourier basis is denoted using the matrix $\mathbf{\Psi}_y = \mathbf{F}^H$. The passband signal at the receiver can be written as

$$\mathbf{y} = \mathbf{\Psi}_y(\mathbf{x}_f + \mathbf{v}_f) \tag{3.8}$$

We acquire the received signal through a linear measurement process modeled by the *sensing matrix*, $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$, where $M$ is the number of available measurements. The acquired signal is denoted by the $M \times 1$ vector

$$\begin{aligned} \tilde{\mathbf{y}} = \mathbf{\Phi}\mathbf{y} &= \mathbf{\Phi}\mathbf{\Psi}_y(\mathbf{x}_f + \mathbf{v}_f) \\ &= \tilde{\mathbf{\Psi}}_y(\mathbf{x}_f + \mathbf{v}_f) \end{aligned} \tag{3.9}$$

where, $\tilde{\mathbf{\Psi}}_y = \mathbf{\Phi}\mathbf{\Psi}_y$ is the new basis of the signal $\tilde{\mathbf{y}}$, often referred to as the *holographic basis*.

The occupancy of these $N$ channels are detected all at once using the MHT approach. It should be noted that these $N$ frequency channels could have all possible combinations of each channel being busy and/or free. This results in one *null hypothesis* denoted by $\mathcal{H}_0$, with all channels being signal-free (noise-only). All the remaining combinations of each channel being busy (signal and noise) or free will result in $2^N - 1$ *alternative hypotheses* denoted by $\mathcal{H}_i$, with $i = 1, \ldots, m = 2^N - 1$. Let the combination of $N$ frequency bands being free (indicated by "0") and the frequency bands being occupied (indicated by "1") for the $i$th hypothesis be denoted by the $N \times 1$ vector $\mathbf{c}_{x|\mathcal{H}_i}$. Such that

$$\begin{aligned}
i = 0 : \mathbf{c}_{x|\mathcal{H}_0} &= \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \end{bmatrix}^T \\
i = 1 : \mathbf{c}_{x|\mathcal{H}_1} &= \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix}^T \\
i = 2 : \mathbf{c}_{x|\mathcal{H}_2} &= \begin{bmatrix} 0 & 0 & \cdots & 1 & 0 \end{bmatrix}^T \\
&\vdots \\
i = 2^N - 1 : \mathbf{c}_{x|\mathcal{H}_{2^N-1}} &= \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \end{bmatrix}^T
\end{aligned} \tag{3.10}$$

The number of non-zero entries of the vector $\mathbf{c}_{x|\mathcal{H}_i}$ is denoted by its $\ell_0$-norm, i.e., $\|\mathbf{c}_{x|\mathcal{H}_i}\|_0$. The variance of any active channel is modeled as $\sigma_x^2$, and the variance of the noise in each channel as $\sigma_v^2$. It is assumed that the channels are uncorrelated. The covariance matrices $\mathbf{\Sigma}_{x|\mathcal{H}_i} \in \mathbb{R}^{N \times N}$ and $\mathbf{\Sigma}_v \in \mathbb{R}^{N \times N}$ are defined as

$$\mathbf{\Sigma}_{x|\mathcal{H}_i} = \mathbb{E}[(\mathbf{x}_f \mathbf{x}_f^H)|\mathcal{H}_i] = \sigma_x^2 \text{diag}(\mathbf{c}_{x|\mathcal{H}_i}) \tag{3.11}$$

$$\mathbf{\Sigma}_v = \mathbb{E}[\mathbf{v}_f \mathbf{v}_f^H] = \sigma_v^2 \mathbf{I}_N \tag{3.12}$$

We model both the signal and noise as i.i.d. Gaussian random variables. Hence, for the $i$th hypothesis, the signal can be written as $\mathbf{x}^{(i)} \sim \mathcal{CN}(0, \mathbf{\Psi}_y \mathbf{\Sigma}_{x|\mathcal{H}_i} \mathbf{\Psi}_y^H)$ and the noise as $\mathbf{v} \sim \mathcal{CN}(0, \sigma_v^2 \mathbf{I}_N)$. We define the matrix $\mathbf{C}_{x|\mathcal{H}_i} = \text{diag}(\mathbf{c}_{x|\mathcal{H}_i})$.

In the following section, we develop a multiband occupancy detector based on the MHT problem.

## 3.4 Optimization problem

The solution to the MHT problem will decide on one of the following hypotheses

$$\begin{aligned}
\mathcal{H}_0 : &\quad \tilde{\mathbf{y}} = \tilde{\mathbf{\Psi}}_y \mathbf{v}_f \\
\mathcal{H}_i : &\quad \tilde{\mathbf{y}}^{(i)} = \tilde{\mathbf{\Psi}}_y (\mathbf{x}_f^{(i)} + \mathbf{v}_f) \quad \text{for} \quad i = 1, \ldots, m = 2^N - 1.
\end{aligned} \tag{3.13}$$

We develop a detector under the Neyman-Pearson-like criterion by solving the optimization problems (3.5) and (3.6). In the following subsections, this is done for two cases

1. When the number of available measurements $M = N$. This means the complete frequency information is available.

2. Compressed detection problem: When the number of available measurements $M < N$, for e.g., compressed samples obtained using sub-Nyquist rate sampling. The frequency information is incomplete in such cases.

### 3.4.1 Multiple hypothesis testing with complete frequency information ($M = N$)

We consider $N$ samples obtained using Nyquist rate sampling, with the number of available measurements $M = N$. Taking an $N$-point DFT, we obtain a frequency domain signal of length $N$, each representing one channel. In other words, the complete frequency information about the signal is available. To realize such a measurement process, we consider the sensing matrix $\mathbf{\Phi} = \mathbf{I}_N$. Therefore, $\tilde{\mathbf{\Psi}}_y = \mathbf{\Psi}_y$ and $\tilde{\mathbf{y}} = \mathbf{y}$. We can rewrite (3.13) as

$$\begin{aligned} \mathcal{H}_0 : &\quad \mathbf{y} = \mathbf{\Psi}_y \mathbf{v}_f \\ \mathcal{H}_i : &\quad \mathbf{y}^{(i)} = \mathbf{\Psi}_y(\mathbf{x}_f^{(i)} + \mathbf{v}_f) \quad \text{with} \quad i = 1, \ldots, m = 2^N - 1. \end{aligned} \tag{3.14}$$

Let the objective function $\Lambda$ of (3.5) for the Nyquist rate sampling case be denoted by $\Lambda_N$, which is given by

$$\Lambda_N(i) = \ln\left(\frac{Pr(\mathbf{y}|\mathcal{H}_i)}{Pr(\mathbf{y}|\mathcal{H}_0)}\right) \quad , \quad i \in \{1, \ldots, m\}. \tag{3.15}$$

From (3.11), (3.12) and (3.13) we can write

$$Pr(\mathbf{y}|\mathcal{H}_0) = \frac{|\mathbf{\Sigma}_v|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} \exp\left(-\frac{1}{2}\mathbf{y}^H \mathbf{\Psi}_y \mathbf{\Sigma}_v^{-1} \mathbf{\Psi}_y^H \mathbf{y}\right) \tag{3.16}$$

$$Pr(\mathbf{y}|\mathcal{H}_i) = \frac{|\mathbf{\Sigma}_{x|\mathcal{H}_i} + \mathbf{\Sigma}_v|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} \exp\left(-\frac{1}{2}\mathbf{y}^H \mathbf{\Psi}_y (\mathbf{\Sigma}_{x|\mathcal{H}_i} + \mathbf{\Sigma}_v)^{-1} \mathbf{\Psi}_y^H \mathbf{y}\right) \tag{3.17}$$

Substituting (3.16) and (3.17) in (3.15), and scaling appropriately will result in

$$\begin{aligned} \Lambda_N(i) &= \ln\left(\frac{|\mathbf{\Sigma}_v|}{|\mathbf{\Sigma}_{x|\mathcal{H}_i} + \mathbf{\Sigma}_v|}\right) + \mathbf{y}^H \mathbf{\Psi}_y (\mathbf{\Sigma}_v^{-1} - (\mathbf{\Sigma}_{x|\mathcal{H}_i} + \mathbf{\Sigma}_v)^{-1}) \mathbf{\Psi}_y^H \mathbf{y} \quad , \quad i \in \{1, \ldots, m\}. \\ &= \ln\left(\frac{|\mathbf{\Sigma}_v|}{|\mathbf{\Sigma}_{x|\mathcal{H}_i} + \mathbf{\Sigma}_v|}\right) + \mathbf{y}_f^H (\mathbf{\Sigma}_v^{-1} - (\mathbf{\Sigma}_{x|\mathcal{H}_i} + \mathbf{\Sigma}_v)^{-1}) \mathbf{y}_f \end{aligned} \tag{3.18}$$

where, $\mathbf{y}_f = \mathbf{\Psi}_y^H \mathbf{y}$ is the frequency response of $\mathbf{y}$.

Using the matrix inversion lemma

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}^{-1}(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}$$

we can then write

$$\mathbf{\Sigma}_v^{-1} - (\mathbf{\Sigma}_{x|\mathcal{H}_i} + \mathbf{\Sigma}_v)^{-1} = \mathbf{\Sigma}_v^{-1} \mathbf{\Sigma}_{x|\mathcal{H}_i} (\mathbf{\Sigma}_{x|\mathcal{H}_i} + \mathbf{\Sigma}_v)^{-1} \tag{3.19}$$

From (3.19) we can write

$$\Lambda_N(i) = \ln\left(\frac{|\mathbf{\Sigma}_v|}{|\mathbf{\Sigma}_{x|\mathcal{H}_i} + \mathbf{\Sigma}_v|}\right) + \mathbf{y}_f^H \mathbf{\Sigma}_v^{-1} \mathbf{\Sigma}_{x|\mathcal{H}_i} (\mathbf{\Sigma}_{x|\mathcal{H}_i} + \mathbf{\Sigma}_v)^{-1} \mathbf{y}_f \quad , \quad i \in \{1, \ldots, m\}. \tag{3.20}$$

To simplify the determinant and the inverse, we define $\gamma = \frac{\sigma_x^2}{\sigma_v^2}$. The objective function (3.20) can be written as

$$\Lambda_N(i) = \ln\left(\frac{1}{(1+\gamma)^{\|\mathbf{c}_{x|\mathcal{H}_i}\|_0}}\right) + \frac{\gamma}{\sigma_v^2(1+\gamma)} \mathbf{y}_f^H \mathbf{C}_{x|\mathcal{H}_i} \mathbf{y}_f \quad , \quad i \in \{1, \ldots, m\}. \tag{3.21}$$

We define the matrix $\mathbf{Y} = \text{diag}(\mathbf{y}_f \odot \mathbf{y}_f)$. Here the $N \times 1$ vector $\mathbf{a} \odot \mathbf{b}$ denotes the element-wise product of vectors $\mathbf{a}$ and $\mathbf{b}$. In terms of the matrix $\mathbf{Y}$, (3.21) can be written as

$$\Lambda_N(i) = \ln\left(\frac{1}{(1+\gamma)^{\|\mathbf{c}_{x|\mathcal{H}_i}\|_0}}\right) + \frac{\gamma}{\sigma_v^2(1+\gamma)} \mathbf{c}_{x|\mathcal{H}_i}^T \mathbf{Y} \mathbf{c}_{x|\mathcal{H}_i} \quad , \quad i \in \{1, \ldots, m\}. \tag{3.22}$$

We can then write the optimization problem (3.5) as

$$i^* = \arg \max_i \quad (\|\mathbf{Y}^{\frac{1}{2}} \mathbf{c}_{x|\mathcal{H}_i}\|_2^2 - \lambda \|\mathbf{c}_{x|\mathcal{H}_i}\|_0) \tag{3.23}$$

where, $\lambda = \frac{\sigma_v^2(1+\gamma)\ln(1+\gamma)}{\gamma}$. The optimization problem (3.23) is highly non-linear and non-convex, and naive strategies involves sifting through all the $2^N$ possibilities. This gets impractical for large values of $N$.

Next, we use some of the properties of the vector $\mathbf{c}_{x|\mathcal{H}_i}$, to reduce the complexity of this problem. For $\mathbf{c}_{x|\mathcal{H}_i} \in \{0,1\}^N$, we know that

$$\|\mathbf{c}_{x|\mathcal{H}_i}\|_0 = \|\mathbf{c}_{x|\mathcal{H}_i}\|_1 = \|\mathbf{c}_{x|\mathcal{H}_i}\|_2^2 \tag{3.24}$$

Using (3.24), (3.23) can be written as

$$i^* = \arg \max_i \quad (\mathbf{c}_{x|\mathcal{H}_i}^T (\mathbf{Y} - \lambda \mathbf{I}_N) \mathbf{c}_{x|\mathcal{H}_i}) \tag{3.25}$$

Since $(\mathbf{Y} - \lambda \mathbf{I}_N)$ is a diagonal matrix, the optimization problem (3.25) can be written as

$$i^* = \arg \max_i \quad (\mathbf{c}_{x|\mathcal{H}_i}^T \mathbf{\Theta} \mathbf{c}_{x|\mathcal{H}_i}) = \sum_{j=1}^{N} [\mathbf{\Theta}]_{jj} [\mathbf{c}_{x|\mathcal{H}_i}]_j^2 \tag{3.26}$$

where, $\mathbf{\Theta} = (\mathbf{Y} - \lambda \mathbf{I}_N)$. It should be noted that $[\mathbf{c}_{x|\mathcal{H}_i}]_j^2 = [\mathbf{c}_{x|\mathcal{H}_i}]_j$ as the elements of the vector $\mathbf{c}_{x|\mathcal{H}_i}$ will have only the binary values, i.e. either 0 or 1. This will result in

$$i^* = \arg \max_i \quad \sum_{j=1}^{N} [\mathbf{\Theta}]_{jj} [\mathbf{c}_{x|\mathcal{H}_i}]_j \tag{3.27}$$

Thus the optimization problem (3.27) is linear in $[\mathbf{c}_{x|\mathcal{H}_i}]_j$, with a computational complexity of $O(N)$. The optimal solution to the problem is given by

$$[\mathbf{c}_{x|\mathcal{H}_{i^*}}]_j = \begin{cases} 1 & \text{if } [\mathbf{\Theta}]_{jj} \geq 0, \\ 0 & \text{if } [\mathbf{\Theta}]_{jj} < 0. \end{cases} \tag{3.28}$$

Alternatively,

$$[\mathbf{c}_{x|\mathcal{H}_{i^*}}]_j = \begin{cases} 1 & \text{if } [\mathbf{Y}]_{jj} \geq \lambda, \\ 0 & \text{if } [\mathbf{Y}]_{jj} < \lambda. \end{cases} \tag{3.29}$$

After the vector $\mathbf{c}_{x|\mathcal{H}_{i^*}}$ is obtained using (3.28) or (3.29), the corresponding $i^*$ or simply $\Lambda_N(i^*)$ can be determined using

$$\Lambda_N(i^*) = \ln\left(\frac{1}{(1+\gamma)^{\|\mathbf{c}_{x|\mathcal{H}_{i^*}}\|_0}}\right) + \frac{\gamma}{\sigma_v^2(1+\gamma)}\mathbf{c}_{x|\mathcal{H}_{i^*}}^T \mathbf{Y} \mathbf{c}_{x|\mathcal{H}_{i^*}} \tag{3.30}$$

And finally, we decide on a hypothesis $\mathcal{H}_{i^*}$ based on

$$\Lambda_N(i^*) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_{i^*}}{\gtrless}} \gamma_{th,N} \tag{3.31}$$

where $\gamma_{th,N}$ ($N$ in the subscript denotes the threshold for Nyquist rate sampling) is a certain threshold used to achieve a certain $P_{fa}$. We term this algorithm as MHT based wideband sensing for the Nyquist rate (MHT-WS:N). The entries along the diagonal of the matrix $\mathbf{Y}$, $[\mathbf{Y}]_{jj} = (\text{abs}([\mathbf{y}_f]_j))^2$, where $\text{abs}(\cdot)$ denotes the absolute value. The detection is based on the energy of the received signal per frequency bin, $(\text{abs}([\mathbf{y}_f]_j))^2$. Hence, the optimal detector, for the $M = N$ case is the energy detector. The detector developed is summarized in Table. 3.1.

Table 3.1: MHT-WS:N detector for $M = N$

*Objective:* Decide on a hypothesis $\mathcal{H}_i$, with $i \in \{0, \ldots, m = 2^N - 1\}$
*Given:* Received signal $\mathbf{y}$ of length $N$, power of an active channel $\sigma_x^2$, noise power per channel $\sigma_v^2$, $\gamma = \frac{\sigma_x^2}{\sigma_v^2}$, $\gamma_{th,N}$ corresponding to a certain $P_{fa}$, and the number of channels $N$.

- *Compute:*
  - $\mathbf{y}_f = \mathbf{\Psi}_y^H \mathbf{y}$.
  - $\lambda = \frac{\sigma_v^2(1+\gamma)\ln(1+\gamma)}{\gamma}$
- *Update:* for $j = 1, \ldots, N$

  $$[\mathbf{c}_{x|\mathcal{H}_{i^*}}]_j = \begin{cases} 1 & \text{if } (\text{abs}([\mathbf{y}_f]_j))^2 \geq \lambda, \\ 0 & \text{if } (\text{abs}([\mathbf{y}_f]_j))^2 < \lambda. \end{cases}$$

- *Compute:*
  $\Lambda_N(i^*)$ as specified in (3.30)
- if $\Lambda_N(i^*) \geq \gamma_{th,N}$ then choose $\mathcal{H}_{i^*}$ otherwise choose $\mathcal{H}_0$

In the following subsection, we discuss the performance of this detector through simulations.

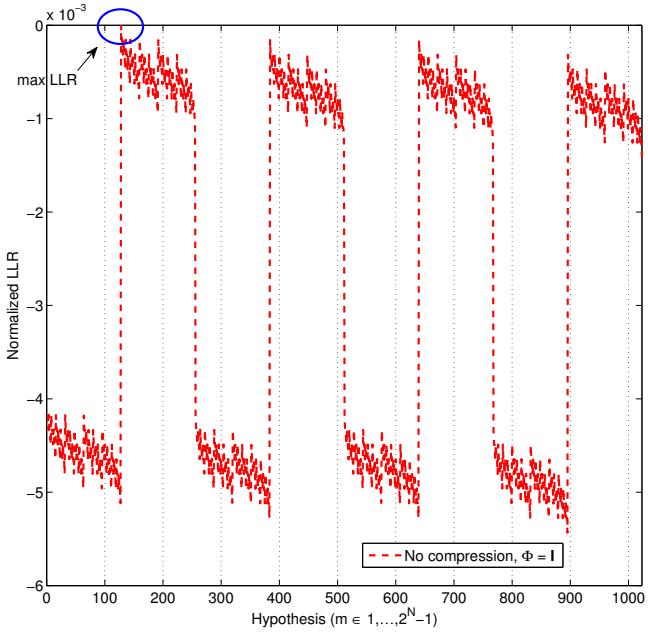



Figure 3.1: Normalized LLR for $M = N = 10$, $K = 1$, SNR = 10 dB, and a static channel occupancy of $[0\,0\,1\,0\,0\,0\,0\,0\,0\,0]$, corresponding to $i = 128$.

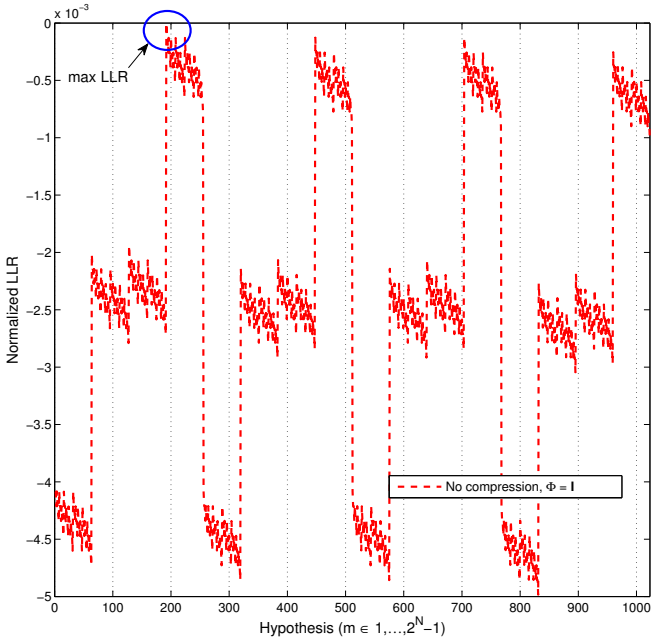



Figure 3.2: Normalized LLR for $M = N = 10$, $K = 2$, SNR = 10 dB, and a static channel occupancy of $[0\,0\,1\,1\,0\,0\,0\,0\,0\,0]$, corresponding to $i = 192$.

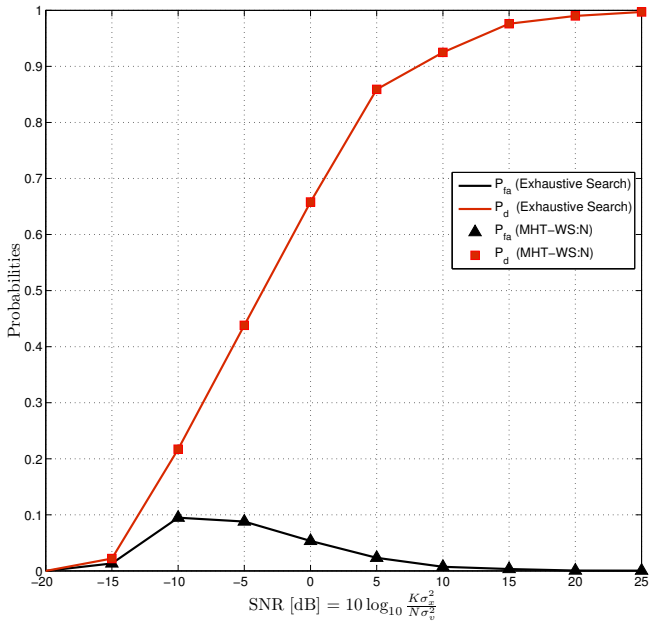



Figure 3.3: Performance of MHT based wideband sensing for M=N=10, and K=1.

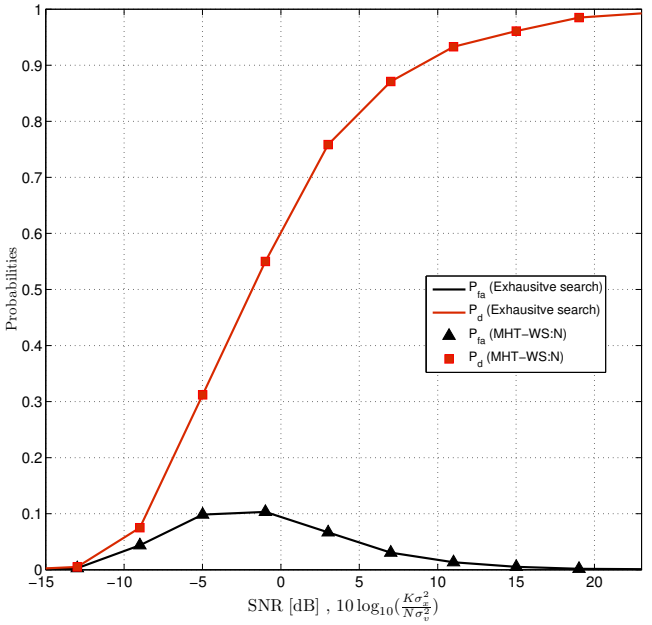



Figure 3.4: Performance of MHT based wideband sensing for M=N=10, and K=2.

#### 3.4.1.1 Simulation results

The proposed MHT-WS:N detector is tested for an example with the following parameters. The number of channels $N = 10$, the number of available measurements $M = 10$, and for two cases of the number of active channels (or $K$-sparse): i)$N_{busy} = K = 1$ and ii) $N_{busy} = K = 2$. The simulation results are averaged over 1000 trials. In each trial, the vector $\mathbf{x}_f$ is randomly generated with i.i.d. Gaussian distributed entries of zero mean and variance according to a certain static channel occupancy ($\sigma_x^2$ for busy channels and zero for free channels). The vector $\mathbf{v}_f$ is generated with i.i.d. Gaussian distributed entries of zero mean and variance $\sigma_v^2$ in each trial. The time domain vector is then obtained using the DFT matrix $\mathbf{\Psi_y}$. The variances are set according to the Signal-to-Noise Ratio (SNR), which is given by $10 \log_{10}\left(\frac{K\sigma_x^2}{N\sigma_v^2}\right)$ dB.

Fig. 3.1 and Fig. 3.2 show the normalized Log-Likelihood Ratios (LLRs) for the busy channel set (specified in (3.1)) $\mathbf{b} = [3]$ and $\mathbf{b} = [3 \quad 4]^T$, respectively. An SNR of 10 dB is considered here. In Fig. 3.1, the maximum LLR is obtained for the hypothesis index $i = 128$ corresponding to the channel occupancy vector considered. Similarly, the maximum LLR is obtained for index $i = 192$ in Fig. 3.2, where the two channels are active. By looking at the simulation results, it can be noticed that, the hypotheses which have small Hamming distance between the channel occupancy vectors, results in close values for the LLRs. We make use of this property, to develop a sub-optimal detector when $M < N$, discussed in the next subsection.

The performance of the detector is measured in terms of the probability of detection, $P_d$, and the probability of false alarm, $P_{fa}$, for different SNRs. The $P_d$ and the $P_{fa}$ are defined in (3.3) and (3.4), respectively. In Fig. 3.3, a static channel occupancy with $\mathbf{b} = [3]$ is considered, and the threshold $\gamma_{th,N}$ is chosen, based on simulations, to keep the values of $P_{fa}$ below 10%. In Fig. 3.4, we consider two channels to be active with $\mathbf{b} = [3 \quad 4]^T$. The simulations based on both exhaustive search to solve (3.23) and the MHT-WS:N detector as described in Table. 3.1 are shown in Fig. 3.3 and Fig. 3.4. It can be seen that the curves of the exhaustive search based detector and the MHT-WS:N detector overlap.

It should be noted that the detector is based on the energy of the received signal per frequency channel. Therefore, as any energy detector it suffers from the SNR wall and noise uncertainty issues, with a poor performance in the low-SNR regime.

In the following subsection, we extend the MHT based wideband sensing detector for the case where the number of available measurements is significantly smaller than the number of Nyquist rate samples.

### 3.4.2 Multiple hypothesis testing with incomplete frequency information ($M < N$)

In Section 2.1, we provided an overview of techniques for multiband occupancy detection, some of which use the classical Compressive Sensing (CS) framework discussed in Section 2.2. In these techniques, either the entire signal or its statistical measures are reconstructed first. Then, in the second stage, detection is performed on this compressive estimate. This reduces the complexity by avoiding Nyquist rate sampling, but

still requires high complexity sparse recovery algorithms. Here, we avoid this two stage estimation-detection approach, and instead perform a direct detection on the compressed samples obtained using sub-Nyquist rate sampling. The sampling rate could be reduced by using the architectures reported in the literature, such as multi-coset sampling, or modulated wideband converters, as discussed in Section 2.2. These are often referred to as Analog-to-Information Converters (AICs). This linear measurement process is modeled using the sensing matrix, $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$, with $M < N$. For the sub-Nyquist rate sampling, $\mathbf{\Phi}$ is a fat matrix, resulting in an under-determined system.

In this case, the hypothesis testing problem is the same as in (3.13), given by

$$\begin{aligned} \mathcal{H}_0: &\quad \tilde{\mathbf{y}} = \tilde{\mathbf{\Psi}}_y \mathbf{v}_f \\ \mathcal{H}_i: &\quad \tilde{\mathbf{y}}^{(i)} = \tilde{\mathbf{\Psi}}_y (\mathbf{x}_f^{(i)} + \mathbf{v}_f) \quad \text{with} \quad i = 1, \ldots, 2^N - 1. \end{aligned} \tag{3.32}$$

where, the $M \times 1$ vector $\tilde{\mathbf{y}}$ is the observation of reduced dimensionality.

The acquired signal can be written in terms of the compressed signal denoted by $\tilde{\mathbf{x}}^{(i)} = \tilde{\mathbf{\Psi}}_y \mathbf{x}_f^{(i)}$, and the compressed noise[1] $\tilde{\mathbf{v}} = \tilde{\mathbf{\Psi}}_y \mathbf{v}_f$. Since, the measurement process is linear, the compressed vectors will be Gaussian random variables. The covariance matrix of the signal $\tilde{\mathbf{x}}$ for the $i$th hypothesis is given by

$$\tilde{\mathbf{\Sigma}}_{x|\mathcal{H}_i} = \mathbb{E}[(\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H)|\mathcal{H}_i] = \tilde{\mathbf{\Psi}}_y \mathbf{\Sigma}_{x|\mathcal{H}_i} \tilde{\mathbf{\Psi}}_y^H \quad \text{with} \quad \tilde{\mathbf{\Sigma}}_{x|\mathcal{H}_i} \in \mathbb{C}^{M \times M} \tag{3.33}$$

and the noise $\tilde{\mathbf{v}}$ is given by

$$\begin{aligned} \tilde{\mathbf{\Sigma}}_v = \mathbb{E}[\tilde{\mathbf{v}}\tilde{\mathbf{v}}^H] &= \tilde{\mathbf{\Psi}}_y \mathbf{\Sigma}_v \tilde{\mathbf{\Psi}}_y^H \quad \text{with} \quad \tilde{\mathbf{\Sigma}}_v \in \mathbb{R}^{M \times M} \\ &= \sigma_v^2 \mathbf{\Phi}\mathbf{\Phi}^H \end{aligned} \tag{3.34}$$

Hence, the covariance matrix of the acquired signal is given by

$$\tilde{\mathbf{\Sigma}}_{y|\mathcal{H}_i} = \mathbb{E}[(\tilde{\mathbf{y}}\tilde{\mathbf{y}}^H)|\mathcal{H}_i] = \tilde{\mathbf{\Psi}}_y \mathbf{\Sigma}_{x|\mathcal{H}_i} \tilde{\mathbf{\Psi}}_y^H + \sigma_v^2 \mathbf{\Phi}\mathbf{\Phi}^H \quad \text{with} \quad \tilde{\mathbf{\Sigma}}_{y|\mathcal{H}_i} \in \mathbb{C}^{M \times M} \tag{3.35}$$

The objective function $\Lambda$ in (3.5) for the compressed case is denoted by $\Lambda_C$, and is given by

$$\Lambda_C(i) = \ln\left(\frac{Pr(\tilde{\mathbf{y}}|\mathcal{H}_i)}{Pr(\tilde{\mathbf{y}}|\mathcal{H}_0)}\right) \quad , \quad i \in \{1, \ldots, m\}. \tag{3.36}$$

The vector to be recovered $\mathbf{c}_{x|\mathcal{H}_i}$ has length $N$. In other words, the number of hypotheses to be tested is $2^N$, similar to the case $M = N$. The only difference is that, we have fewer number of observations.

From (3.33), (3.34) and (3.35) we can write

$$Pr(\tilde{\mathbf{y}}|\mathcal{H}_0) = \frac{|\tilde{\mathbf{\Sigma}}_v|^{-\frac{1}{2}}}{(2\pi)^{(\frac{N}{2})}} \exp\left(-\frac{1}{2}\tilde{\mathbf{y}}^H \tilde{\mathbf{\Sigma}}_v^{-1} \tilde{\mathbf{y}}\right) \tag{3.37}$$

$$Pr(\tilde{\mathbf{y}}|\mathcal{H}_i) = \frac{|\tilde{\mathbf{\Sigma}}_{y|\mathcal{H}_i}|^{-\frac{1}{2}}}{(2\pi)^{(\frac{N}{2})}} \exp\left(-\frac{1}{2}\tilde{\mathbf{y}}^H (\tilde{\mathbf{\Psi}}_y \mathbf{\Sigma}_{x|\mathcal{H}_i} \tilde{\mathbf{\Psi}}_y^H + \tilde{\mathbf{\Sigma}}_v)^{-1} \tilde{\mathbf{y}}\right) \tag{3.38}$$

[1]It should be noted that the noise is not sparse in either basis. The word compressed noise is used to maintain the notational consistency.

Substituting (3.37) and (3.38) in (3.36), and scaling appropriately we get

$$\Lambda_C(i) = \ln\left(\frac{|\tilde{\mathbf{\Sigma}}_{y|\mathcal{H}_i}|}{|\tilde{\mathbf{\Sigma}}_v|}\right) + \tilde{\mathbf{y}}^H\left(\tilde{\mathbf{\Sigma}}_v^{-1} - (\tilde{\mathbf{\Psi}}_y\mathbf{\Sigma}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}_y^H + \tilde{\mathbf{\Sigma}}_v)^{-1}\right)\tilde{\mathbf{y}} \quad , \quad i \in \{1,\ldots,m\}. \tag{3.39}$$

Using the matrix inversion lemma

$$(\tilde{\mathbf{\Psi}}_y\mathbf{\Sigma}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}_y^H + \tilde{\mathbf{\Sigma}}_v)^{-1} = \tilde{\mathbf{\Sigma}}_v^{-1} - \tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}}_y(\mathbf{\Sigma}_{x|\mathcal{H}_i}^{-1} + \tilde{\mathbf{\Psi}}_y^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}}_y)^{-1}\tilde{\mathbf{\Psi}}_y^H\tilde{\mathbf{\Sigma}}_v^{-1} \tag{3.40}$$

Further factorizing (3.40) we have

$$\begin{aligned}&(\tilde{\mathbf{\Psi}}_y\mathbf{\Sigma}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}_y^H + \tilde{\mathbf{\Sigma}}_v)^{-1}\\ &= \tilde{\mathbf{\Sigma}}_v^{-1} - \tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}}_y\mathbf{\Sigma}_{x|\mathcal{H}_i}^{\frac{1}{2}}(\mathbf{I}_N + \mathbf{\Sigma}_{x|\mathcal{H}_i}^{\frac{1}{2}}\tilde{\mathbf{\Psi}}_y^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}}_y\mathbf{\Sigma}_{x|\mathcal{H}_i}^{\frac{1}{2}})^{-1}\mathbf{\Sigma}_{x|\mathcal{H}_i}^{\frac{1}{2}}\tilde{\mathbf{\Psi}}_y^H\tilde{\mathbf{\Sigma}}_v^{-1}\end{aligned} \tag{3.41}$$

Using (3.41) and simplifying the determinant, we can rewrite the objective function (3.39) as

$$\Lambda_C(i) = \tilde{\mathbf{y}}^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}}_y\mathbf{\Sigma}_{x|\mathcal{H}_i}^{\frac{1}{2}}(\mathbf{I}_N + \mathbf{\Sigma}_{x|\mathcal{H}_i}^{\frac{1}{2}}\tilde{\mathbf{\Psi}}_y^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}}_y\mathbf{\Sigma}_{x|\mathcal{H}_i}^{\frac{1}{2}})^{-1}\mathbf{\Sigma}_{x|\mathcal{H}_i}^{\frac{1}{2}}\tilde{\mathbf{\Psi}}_y^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{y}} - \|\mathbf{c}_{x|\mathcal{H}_i}\|_0\ln(1+\gamma), \quad i \in \{1,\ldots,m\}. \tag{3.42}$$

where $\gamma = \frac{\sigma_x^2}{\sigma_v^2}$ as defined before.

Similar to the previous case of $M = N$, for $M < N$ also, the exhaustive search would require $O(2^N)$ computations, and is practically not feasible for larger $N$. We term this algorithm based on exhaustive search as MHT-WS:CD (MHT based Compressed Detection (CD) for Wideband Sensing (WS)). Alternatively, (3.42) can be written as

$$\begin{aligned}\Lambda_C(i) = \sigma_x^2\tilde{\mathbf{y}}^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}}_y\mathbf{C}_{x|\mathcal{H}_i}(\mathbf{I}_N + \sigma_x^2\mathbf{C}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}_y^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{\Psi}}_y\mathbf{C}_{x|\mathcal{H}_i})^{-1}\mathbf{C}_{x|\mathcal{H}_i}\tilde{\mathbf{\Psi}}_y^H\tilde{\mathbf{\Sigma}}_v^{-1}\tilde{\mathbf{y}}\\ -\|\mathbf{c}_{x|\mathcal{H}_i}\|_0\ln(1+\gamma), \quad i \in \{1,\ldots,m\}.\end{aligned} \tag{3.43}$$

It is mathematically intricate to factor out the matrix $\mathbf{C}_{x|\mathcal{H}_i}$ from the matrix inverse in (3.43). This makes the optimization an involved non-convex non-linear integer programming problem of high complexity. Therefore, we propose a sub-optimal greedy algorithm to solve this optimization problem, based on certain heuristics.

Before presenting the proposed greedy algorithm, we provide some definitions.

**Definition 3.2.** Neighborhood. *For $\forall\, \mathbf{c}_{x|\mathcal{H}_i} \in \{0,1\}^N$, the neighborhood of $\mathbf{c}_{x|\mathcal{H}_i}$ with a size $S$ is defined as the set*

$$N_S(\mathbf{c}_{x|\mathcal{H}_i}) = \{\mathbf{c}_{x|\mathcal{H}_j} \in \{0,1\}^N \mid \|\mathbf{c}_{x|\mathcal{H}_i} - \mathbf{c}_{x|\mathcal{H}_j}\|_1 = S\} \tag{3.44}$$

*where, $\|\mathbf{c}_{x|\mathcal{H}_i} - \mathbf{c}_{x|\mathcal{H}_j}\|_1$ denotes the Hamming distance between $\mathbf{c}_{x|\mathcal{H}_i}$ and $\mathbf{c}_{x|\mathcal{H}_j}$.*

This means, the vectors $\mathbf{c}_{x|\mathcal{H}_j}$ and $\mathbf{c}_{x|\mathcal{H}_i}$ differs by $S$ bits $\forall\, \mathbf{c}_{x|\mathcal{H}_i} \in N_S(\mathbf{c}_{x|\mathcal{H}_j})$. For any $\mathbf{c}_{x|\mathcal{H}_i}$ the total number of vectors in $N_S(\mathbf{c}_{x|\mathcal{H}_i})$ will be $|N_S(\mathbf{c}_{x|\mathcal{H}_i})| = \binom{N}{S}$, with $\sum_{r=0}^{N}|N_r(\mathbf{c}_{x|\mathcal{H}_i})| = \sum_{r=0}^{N}\binom{N}{r} = 2^N$.

**Definition 3.3.** Local Maximum LLR (LML) point. *Consider* $\mathbf{c}_{x|\mathcal{H}_i} \in \{0,1\}^N$ *with a neighborhood* $N_S(\mathbf{c}_{x|\mathcal{H}_i})$*, and the LLR* $\Lambda_C(i)$ *as specified in (3.43). If* $\Lambda_C(i^*) \geq \Lambda_C(j)$ $\forall \mathbf{c}_{x|\mathcal{H}_j} \in N_S(\mathbf{c}_{x|\mathcal{H}_i})$*, then* $\mathbf{c}_{x|\mathcal{H}_{i^*}}$ *results in the local maximum LLR point,* $\Lambda^*_{C|S}(i)$ *within the neighborhood size* $S$*.*

**Property 3.1.** *Consider* $\mathbf{c}_{x|\mathcal{H}_i} \in \{0,1\}^N$ *with a neighborhood* $N_S(\mathbf{c}_{x|\mathcal{H}_i})$ *and LML point* $\Lambda^*_{C|S}(i)$*. Then* $\Lambda^*_{C|1}(i) \leq \Lambda^*_{C|2}(i) \cdots \leq \Lambda^*_{C|N}(i)$*.*

As the neighborhood size increases, the distance between the LML points increases. For $\mathbf{c}_{x|\mathcal{H}_i} \in \{0,1\}^N$ with a neighborhood $N_S(\mathbf{c}_{x|\mathcal{H}_i})$, the LLR values between $\mathbf{c}_{x|\mathcal{H}_i}$ and $N_S(\mathbf{c}_{x|\mathcal{H}_i})$ are closer to each other for smaller $S$. As $S \to N$, the difference between the LLR values increases. This property is used in the proposed sub-optimal greedy algorithm. Due to the compression, the LLRs are distorted as can be seen in Fig. 3.5a and Fig. 3.5b (these plots are described in more detail in the next subsection). To increase the performance of the algorithm, we choose higher neighborhood size (of 3 in this case, to achieve the desired performance). Next, we present the algorithm based on the property discussed above. Since this algorithm is based on the LML and used to perform CD, we call it LML-CD.

A similar approach is proposed in the literature for Maximum Likelihood (ML) MU detection for Code Division Multiple Access (CDMA) communications systems [45] and [46]. This approach is extended to LLRs for the considered MHT problem.

Table 3.2: LML-CD algorithm for MHT problem for WS ($M < N$)

*Objective:* Decide on a hypothesis $\mathcal{H}_i$, with $i \in \{1, \ldots, 2^N - 1\}$

*Initialization:* Start with the initial vector $\mathbf{c}_{x|\mathcal{H}_i} \in \{0,1\}^N$ uniformly at random.

- *Generate $N$ vectors:*
  - $\mathbf{c}_{x|\mathcal{H}_u} \in \{0,1\}^N$, for $u = 1, \ldots, N$, choosing each vector uniformly at random from $N_u(\mathbf{c}_{x|\mathcal{H}_i})$.
- *Observation space:*
  - $U_1 = \{\mathbf{c}_{x|\mathcal{H}_i} \cup \mathbf{c}_{x|\mathcal{H}_u} \mid u = 1, \ldots, N\}$, with $|U_1| = N + 1$
- *Compute:* $\Lambda_C$ as specified in (3.43)
  - $i^* = \arg \max_{i \in \{1,\ldots,N+1\}} \Lambda_C(i), \quad \text{with} \quad \mathbf{c}_{x|\mathcal{H}_i} \in U_1$ (3.45)
- *Update observation space:*
  - $U_2 = \{N_{\{S=1\}}(\mathbf{c}_{x|\mathcal{H}_{i^*}}) \cup N_{\{S=2\}}(\mathbf{c}_{x|\mathcal{H}_{i^*}}) \cup N_{\{S=3\}}(\mathbf{c}_{x|\mathcal{H}_{i^*}})\}$, with $|U_2| = \sum_{r=1}^{3} \binom{N}{r}$
- *Compute:* $\Lambda_C$ as specified in (3.43)
  - $i^* = \arg \max_{i \in \{1,\ldots,N\}} \Lambda_C(i), \quad \text{with} \quad \mathbf{c}_{x|\mathcal{H}_i} \in U_2$ (3.46)
- if $\Lambda_C(i^*) \geq \gamma_{th,C}$ then choose $\mathcal{H}_{i^*}$ otherwise choose $\mathcal{H}_0$

The algorithm LML-CD is initialized with a vector $\mathbf{c}_{x|\mathcal{H}_i}$, at random out of $2^N$ possible vectors. Next, we choose $N$ vectors uniformly at random, such that, the $N$

vectors have all possible Hamming distances from 1 to $N$, with the initial vector. An observation space with these $N$ vectors and the initial vector is formed. The maximum LLR, and hence the hypothesis $\mathcal{H}_{i^*}$, is determined within this observation space. Next, the observation space is updated with all the possible vectors that are at a Hamming distance of 1, 2, and 3 from $\mathbf{c}_{x|\mathcal{H}_{i^*}}$, i.e. neighborhood $N_S(\mathbf{c}_{x|\mathcal{H}_{i^*}})$, with a size $S \in \{1, 2, 3\}$. The LML point is computed with the new $i^*$. If the LLR value $\Lambda_C(i^*)$ exceeds the pre-determined threshold $\gamma_{th,C}$ ($C$ in the subscript denotes the threshold for the CS case) then we decide on hypothesis $\mathcal{H}_{i^*}$ otherwise we choose $\mathcal{H}_0$. The threshold is determined to achieve a certain $P_{fa}$. The algorithm LML-CD is summarized in Table 3.2.

The algorithm LML-CD requires $\frac{1}{6}(N^3 + 11N + 6)$ computations, with complexity order of $O(N^3)$. In other words, if the neighborhood size $S$ is increased from one , two, etc., up to $N$, the computational complexity is linear, quadratic, etc., up to exponential in the number of channels (the performance also increases with the complexity). Note that, these algorithms can also be used for the $M = N$ case. Since, the optimal algorithm for $M = N$ case is of linear complexity, we do not consider these sub-optimal algorithms in that case. The complexity of these algorithms is summarized in Table 3.3.

Table 3.3: Complexity of MHT based wideband sensing algorithms

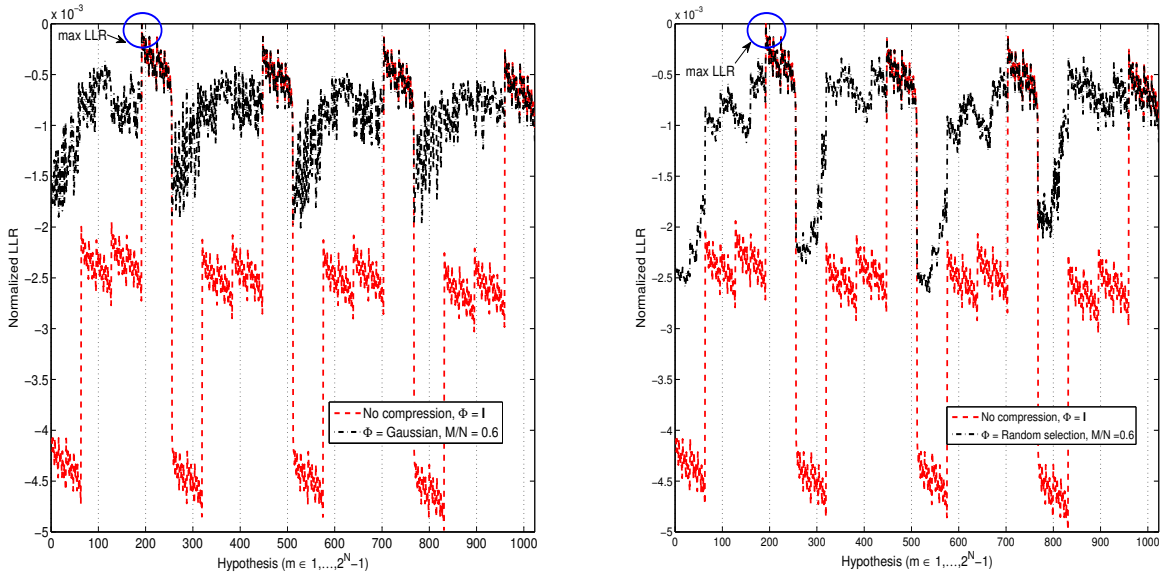
| Algorithms | Complexity order |
|---|---|
| MHT-WS:N | $O(N)$ |
| MHT-WS:CD | $O(2^N)$ |
| LML-CD | $O(N^3)$ |

In the following section, we analyze the performance of the proposed MHT detector for wideband sensing, with $M < N$ through simulations.

#### 3.4.2.1 Simulation results

The proposed MHT-WS:CD detector based on an exhaustive search and the proposed sub-optimal LML-CD based on a greedy search are tested in this section. For the simulations, the following parameters are considered: the number of channels $N = 10$, and two fixed sensing matrices, $\mathbf{\Phi}$, i) *Gaussian* and ii) *Random selection* (these matrices are generated as described in Chapter 2). The *Gaussian* and *random selection* matrices are used in the simulations, as these are the standard favorable sensing matrices (that satisfy the Restricted Isometry Property (RIP), discussed in section 2.2) used in the CS framework for recovery with $\ell_1$-norm optimization. However, for the direct detection, the properties that the sensing matrices should satisfy have to be analyzed and this is subject of future work. The simulations are averaged over 1000 trials. In every trial, the vector $\mathbf{x}_f$ is randomly generated with i.i.d. Gaussian distributed entries of zero mean and variance according to a certain static channel occupancy ($\sigma_x^2$ for busy channels and zero for free channels). The vector $\mathbf{v}_f$ is generated with i.i.d. Gaussian distributed entries of zero mean and variance $\sigma_v^2$ in each trial. The time domain vectors $\mathbf{x}$ and $\mathbf{v}$ are then obtained using the DFT matrix $\mathbf{\Psi}_y$. The variances are set according to the

required SNR, given by $10\log_{10}\left(\frac{K\sigma_x^2}{N\sigma_v^2}\right)$ dB. A static channel occupancy is considered, with a busy channel set $\mathbf{b} = [3 \quad 4]^T$, number of free channels $N_{free} = 8$ and busy channels $N_{busy} = 2$. The threshold $\gamma_{th,C}$ is chosen so as to keep the $P_{fa}$ below 10%.




(a) *Gaussian* sensing matrix.

(b) *Random selection* sensing matrix.

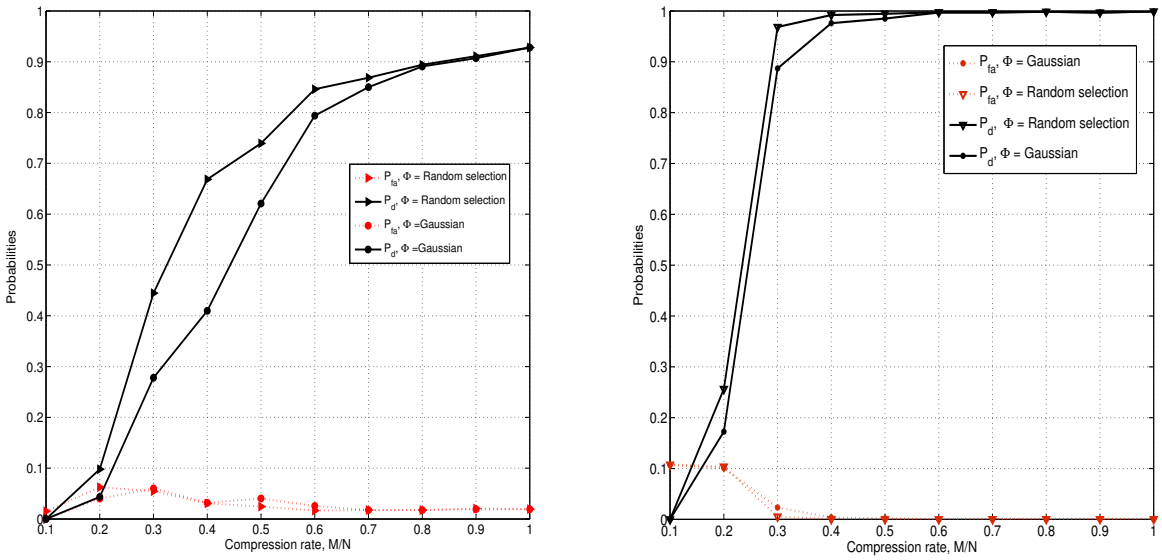
Figure 3.5: Normalized LLR for $\frac{M}{N} = 0.6$ with $K = 2$, SNR = 10 dB, and a static channel occupancy of $[0\,0\,1\,1\,0\,0\,0\,0\,0\,0]$, corresponding to $i = 192$.

Fig. 3.5a and Fig. 3.5b show the normalized LLRs for *Gaussian* and *Random selection* compression matrices, respectively. The Normalized LLRs without compression are also shown in both figures. It can be seen that the compression distorts the LLR values, with the most likely and less likely hypothesis getting much closer compared to the case when $M = N$. It should also be noted that, the LLRs are very sensitive to noise. Here a compression rate of 0.6 is considered, where the compression rate is defined as $\frac{M}{N}$. The maximum LLR is obtained for the hypothesis index $i = 192$, which corresponds to the channel occupancy considered in this example.

Next, the performance of the MHT-WS:CD for different compression rates for an SNR of 10 dB and 30 dB, is shown in Fig. 3.6a and Fig. 3.6b, respectively. For detection with high probability (w.h.p.), we can say that an over-measuring factor $c$ ($M := cK$) of 2 to 3 would be required at high SNRs. And, at a relatively low SNR of 10 dB, an over-measuring factor of 4 is needed to achieve a $P_d \approx 0.9$.

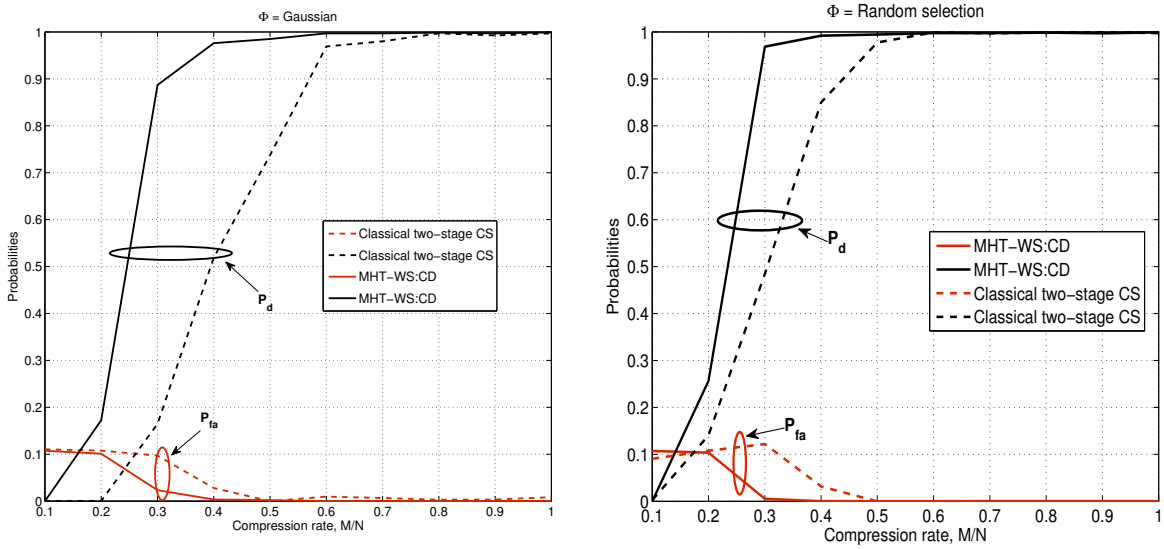
The proposed MHT based wideband sensing for $M < N$ is compared with the conventional approach for CS based spectrum sensing of [20]. In this conventional approach, the signal is first acquired using sub-Nyquist rate sampling. The acquired signal is recovered by solving the optimization problem

$$\min_{\mathbf{y}_f} \quad \|\mathbf{y}_f\|_1 \quad \text{s.t.} \quad \mathbf{\Phi\Psi}_y\mathbf{y}_f = \tilde{\mathbf{y}} \tag{3.47}$$




(a) SNR = 10 dB. (b) SNR = 30 dB.

Figure 3.6: Performance of MHT based wideband sensing (exhaustive search) for $(M < N)$, with $N = 10$, $K = 2$.




(a) *Gaussian* matrix. (b) *Random selection* matrix.

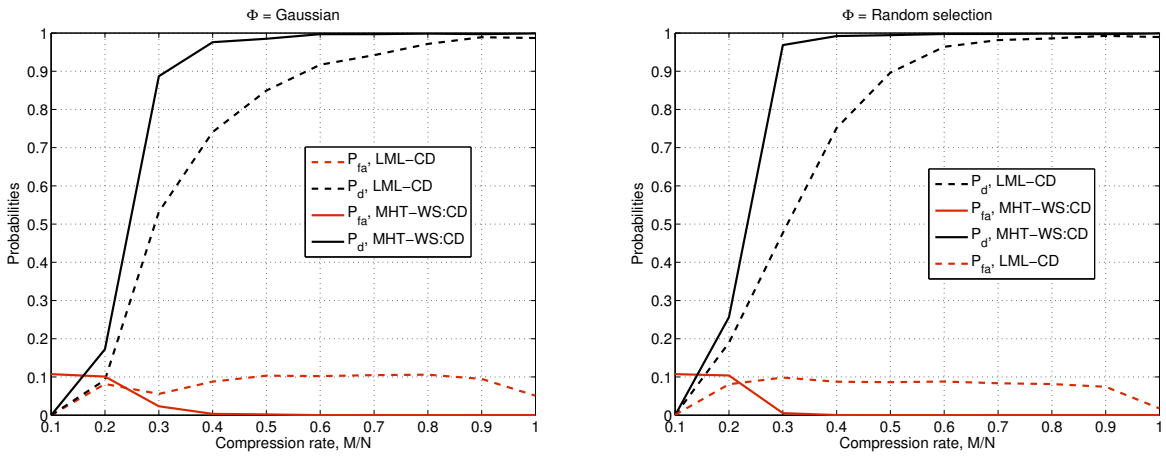
Figure 3.7: Comparison of conventional estimation-detection two stage approach of CS with MHT based wideband sensing for $(M < N)$, with $N = 10$, $K = 2$, and SNR = 30 dB.

Then the detection is performed on the frequency response estimate (power or amplitude). We make use of the Regularized Orthogonal Matching Pursuit (ROMP) [47] to

solve the convex program in (3.47), and then perform energy detection to determine the occupancy. We make use of ROMP, as it has the speed of the greedy iterative methods (matching pursuits) and robustness of $\ell_1$-minimization. The complexity order of ROMP is $O(NMK)$[2]. Additionally, in the second stage, threshold based detection has a complexity order of $O(N)$.

The comparison is shown in Fig. 3.7a and Fig. 3.7b for *Gaussian* and *random selection* compression matrices, respectively. The detector MHT-WS:CD based on exhaustive search performs better than the classical two-stage approach. The $P_d$ of 0.9 is achieved for a compression rate of $\approx 0.3$ with the MHT-WS:CD. Using the classical two-stage approach, the $P_d$ of 0.9 is achieved for a compression rate of $\approx 0.6$ and $\approx 0.5$ with *Gaussian* and *random selection* compression matrices, respectively.

Using a small size of $N$ in the simulation reveals an important aspect of CS. In the CS literature, *Gaussian* or any other random matrices are suggested as a favorable choice for signal recovery with $\ell_1$-norm optimization, but this choice holds mostly for $N \to \infty$ [10]. Structured matrices like *random selection* matrices perform better in cases of smaller $N$ [48], which could be used more often in digital communications, as can be seen in Fig. 3.7b. The performance of the proposed MHT-WS:CD is better than the classical two-stage approach.




(a) *Gaussian* matrix.

(b) *Random selection* matrix.

Figure 3.8: Performance of the sub-optimal LML-CD algorithm for $(M < N)$, with $N = 10$, $K = 2$, and SNR $= 30$ dB.

Fig. 3.8a and Fig. 3.8b show the performance of the proposed sub-optimal algorithms for *Gaussian* and *random selection* compression matrices, respectively. The $P_d \approx 0.9$ is obtained for a compression rate of $\approx 0.5$ and $\approx 0.6$ for *random selection* and *Gaussian* sensing matrices, respectively. The performance of the cubic complexity LML-CD algorithm is similar to that of the classical two-stage approach.

[2]When the exact sparsity level is known.

## 3.5 Conclusions

In this chapter, we considered wideband spectral sensing in the multiple hypothesis sense. Each hypothesis corresponds to one of the possible combinations of all channels being busy and/or free. To solve this detection problem, the extension of the classical Neyman-Pearson approach for binary hypothesis testing to MHT has been proposed. The motivation to formulate multiband occupancy detection as an MHT problem was to reduce the latency involved in solving multiple binary hypothesis test on each frequency channel. An optimal detector was developed for observations with complete frequency information. The energy detector is the optimal detector for Nyquist rate sampling, and is of linear complexity. However, it would still require a high-rate ADC to acquire the samples. To further reduce the complexity and capitalize on the sparsity of the spectrum, the sampling rate is reduced as in the CS framework with observations having incomplete frequency information. Here, we avoid the conventional CS based sensing, which usually involves reconstruction of the compressed signal before detection. Instead a direct detection based on the compressed samples is performed, resulting in a CD. The optimal detector MHT-WS:CD was developed and requires $O(2^N)$ computations. The optimal detector MHT-WS:CD performs better than the classical two-stage approach, but is impractical for larger $N$. Hence, we proposed a sub-optimal greedy algorithm, LML-CD, based on the properties of the local maximum LLRs. The LML-CD is of complexity order $O(N^3)$, and has a performance comparable to that of the conventional two-stage approach (of complexity order $O(KMN)$). The exact knowledge of the sparsity level is not required for direct detection.

# Low-power architecture for wideband spectrum sensing

# 4

## 4.1 Introduction

The frequency spectrum is a scarce resource and has to be utilized efficiently to foster innovations in wireless communications. Cognitive radio technology enables this by supporting secondary spectrum usage, coexistence, and Dynamic Spectrum Access (DSA) by sensing the spectrum for the occupancy and adaptively using the free frequency band for a certain duration without affecting the performance of the licensed Primary User (PU) [6]. In case of low-power radios, spectrum sensing is usually an overhead for the radio to enable co-existence or secondary spectrum usage and should consume minimal power.

Wideband spectrum sensing poses serious challenges for low-power cognitive radios which cannot afford to use high-rate Analog-to-Digital Converters (ADCs) to sample the signals at Nyquist rate and process them digitally thereby spending a lot of power. On the other hand, typical narrowband sensing requires the full receiver chain to be powered each time a channel is sensed. In addition, it has an associated latency depending on the spectrum occupancy, since the local oscillator needs to be locked at a new frequency for every channel search. As discussed in Chapter 2, the minimum average sampling rate required for the perfect reconstruction of a multiband signal is same as the spectral occupancy [16]. For scenarios where the spectrum is not so sparse as in Fig. 1.5b, the sampling rate cannot be reduced significantly than the Nyquist rate. In this chapter, an alternative approach to wideband spectrum sensing to reduce the complexity (power consumption) at the architecture level is proposed for such scenarios.

A closer look at the developments in ADC [49] indicate that the state-of-the-art ADCs are ultra low-power and are not major contributors to the power consumption of the radio. Digital processing and Phase-Lock Loops (PLLs) consume a major portion of the power [50]. However, it should be noted that high-rate ADCs output many samples that have to be processed.

The idea of analog FFT processors was initially proposed for low-power Orthogonal Frequency Division Multiplexing (OFDM) receivers operating at giga-samples per second [51], [52], to reduce the total information processed by the ADCs, and make them power efficient. This motivates the proposed architecture, pushing the conventional ADC and digital processing to the analog domain, at the same time saving a considerable amount of power. In addition, a periodogram can be reconstructed in the analog domain using envelope detectors, which in turn provides averaging for each energy estimate and reduces the effect of fading.

Here we propose an analog/mixed signal topology that replaces the conventional Nyquist ADCs and digital Fast Fourier Transform (FFT) core with a bank of Sam-

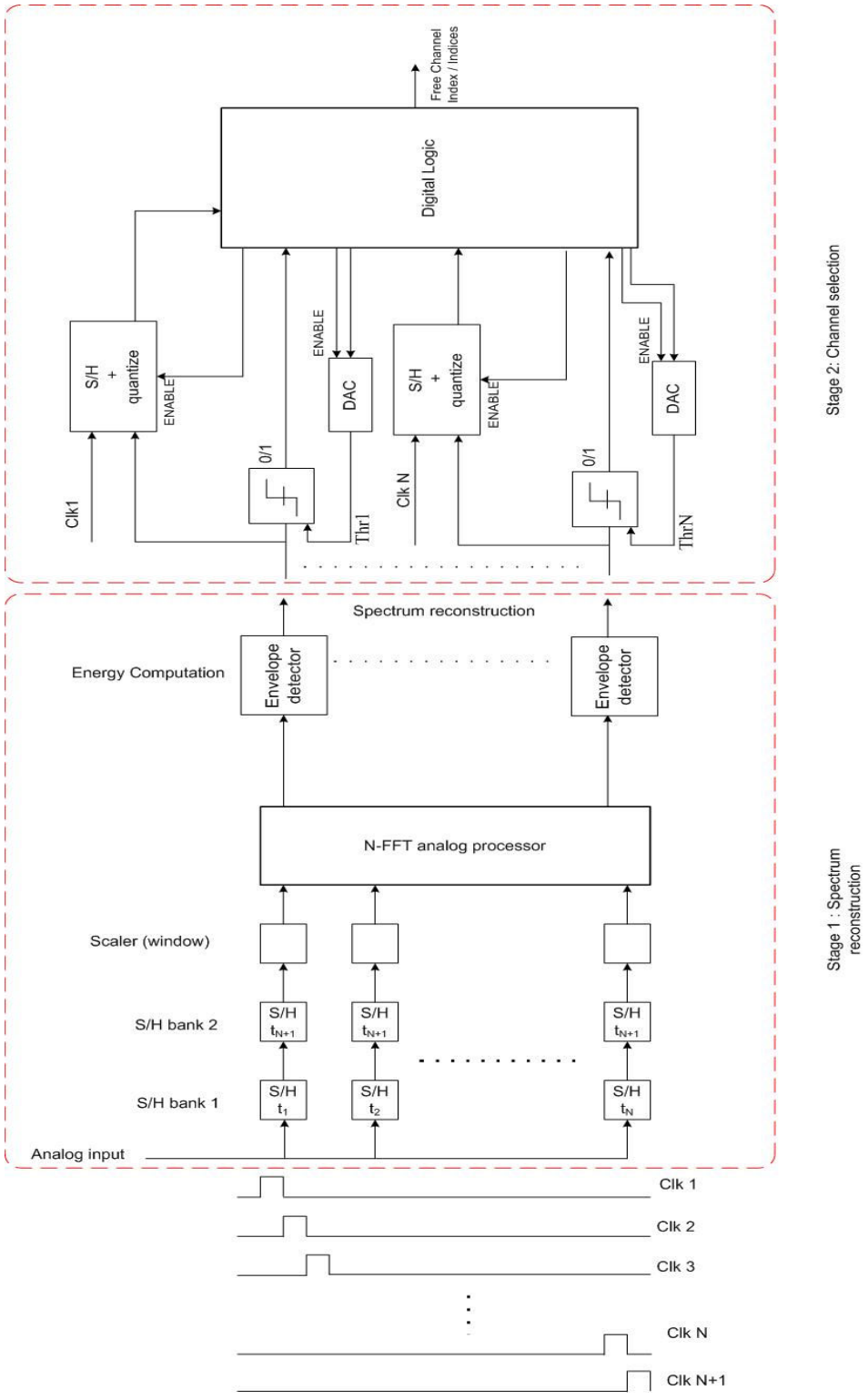



Figure 4.1: Architecture for low-power wideband spectrum reconstruction and channel selection using an analog FFT.

ple and Hold (S/H) circuits, each operating at sub-Nyquist rate, and an all-analog FFT processor. The proposed low-power cross-layered system architecture is used for wideband coarse spectrum reconstruction and free channel selection.

Typically in wideband spectrum sensing, to detect weak and strong signals (for e.g., separated by 50 dB) requires a large dynamic range for the ADCs, to accommodate strong signals while still providing sufficient quantization performance for the weak signals (low-power cognitive radio itself that performs sensing for instance). The power consumption of an ADC increases linearly with sampling frequency and exponentially with the resolution [53]. This makes the front-end circuitry more complex (or high-power) for sensing signals with large dynamic range. The dynamic range issues in cognitive radios are addressed in [54], where a single high resolution ADC is substituted with low-resolution low-rate ADCs to cancel the strong signal (or interference) in the first stage, and another low-resolution ADC to acquire the weak signals. Processing in the analog domain, using S/H circuits offers a low-complexity solution to mitigate the ADC resolution issues associated with the multiband occupancy detection.

In this work, meaningful scenario of 86 MHz in the 2.4 GHz ISM band is considered for analysis of the proposed architecture. The proposed architecture is based on a bank of S/H circuits, analog FFT processing, envelope detectors for spectrum reconstruction, and analog decision thresholds for channel selection.

## 4.2 Proposed system architecture

The block diagram of the proposed architecture is illustrated in Fig. 4.1. The architecture comprises of two stages, stage 1 for course spectrum reconstruction and stage 2 for channel selection. The analog input baseband signal after the Low-Noise Amplifier (LNA) is discretized using a bank of S/H circuits indicated by S/H bank 1 in the block diagram. The S/H bank 1, consists of $N$ S/H circuits respectively operating at clk1, clk2, ..., clk$N$. Each S/H circuit should operate at a rate, $\frac{2B}{N}$ to monitor a spectrum of $B$ Hz. The second bank of S/H indicated by S/H bank 2, also consists of $N$ S/H circuits and re-samples the data from the S/H bank 1, at clk$(N+1)$. The two banks of S/Hs are used to achieve a serial to parallel conversion, which is required for the $N$-FFT analog processor. The output of the S/H bank 2 is scaled to realize a time domain windowing (e.g., Hamming) to reduce spectral leakage. The $N$-FFT analog processor is an analog implementation of an $N$-point FFT. More details regarding the implementation of an analog FFT processor can be found in [51], [52], and [55]. The outputs of the $N$-FFT analog processor are an evolution of $N$ discrete frequency bins in time. Each of these $N$ branches, are fed to $N$ envelope detectors to compute the energy in each frequency bin. An envelope detector can be modeled as a squaring function followed by a low-pass filter. The $N$ branches viewed together would result in a coarse spectrum reconstruction based on a modified periodogram [56]. It should be noted that the low-power envelope detector after the analog FFT output helps in reducing the multi-path and fading effects [57] associated with spectrum sensing.

To find an empty channel in the spectrum, the output of the branches of the envelope detectors are compared to a threshold computed on the basis of a noise power estimate. The thresholds are set to achieve a maximum probability of detection $P_d$ (to detect a

channel as busy, when the channel is indeed busy) subject to a probability of false alarm $P_{fa}$ (decision that a channel is occupied, when the channel is in fact free) constraint. The definitions of $P_{fa}$ and $P_d$ for wideband sensing are given in Chapter 3. The noise power estimation and threshold updates are controlled via digital logic and could be performed during the initialization stage. When the threshold has to be updated, the output of the envelope detector is quantized whenever a channel is detected as free. This quantized output is used in the digital logic to compute a new threshold. The threshold can be different or can be the same for all the frequency bins. The blocks used to set the threshold are i) S/H + quantize and ii) Digital-to-Analog Converter (DAC), which can be reused in a time interleaved manner for all the $N$ branches to conserve power. The output of the digital logic at stage 2 will be the indices of the free channels.

## 4.3 Detection

### 4.3.1 System model

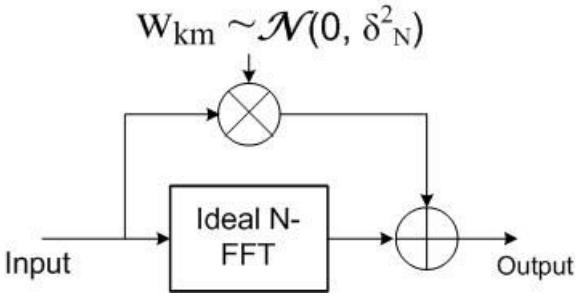



Figure 4.2: Mismatch model for CMOS based analog FFT [51].

The signal at the $k^{th}$ branch (indicating the $k^{th}$ frequency bin) before the envelope detector during time $m$ is denoted by $y_{km}$. These are collected for $M$ time instances in the $M \times 1$ vector $\mathbf{y}_k$ given by $\mathbf{y}_k = [y_{k1} \quad y_{k2} \ldots y_{km}]^{\mathrm{T}}$, where $(\cdot)^T$ denotes the transpose. At time $m$ and frequency $k$, the signal is denoted by $x_{km} \sim \mathcal{N}(0, \sigma_x^2)$ and the noise by $v_{km} \sim \mathcal{N}(0, \sigma_v^2)$. The device mismatch for the $N$-FFT analog processor is modeled as Additive White Gaussian Noise (AWGN) with a mismatch variance $\delta_N^2$ and is denoted at time $m$ and frequency $k$ by the Random variable (RV) $w_{km}$. For the 128-point FFT processor $\delta_N^2$ is 52.3 [51]. An illustration of the model is provided in Fig. 4.2. The mismatch model accounts for the noise propagation from stage to stage in the analog circuitry which is very sensitive to the FFT size, and the device impairments in realizing the FFT. The mismatch model at the $k^{th}$ branch during time $m$ will result in $w_{km}v_{mk}$ for the noise and in $w_{km}x_{mk}$ for the signal. These are collected for $M$ time instances in the $M \times 1$ vectors $\mathbf{q}$ and $\mathbf{r}$ given by

$$\mathbf{q} = \begin{bmatrix} w_{k1}v_{1k} & w_{k2}v_{2k} \cdots w_{km}v_{nk} \end{bmatrix}^{\mathrm{T}}$$

$$\mathbf{r} = \begin{bmatrix} w_{k1}x_{1k} & w_{k2}x_{2k} \cdots w_{km}x_{nk} \end{bmatrix}^{\mathrm{T}}$$

Let $\mathbf{f}_k$ be the $1 \times N$ vector indicating the $k^{th}$ row of a Discrete Fourier Transform (DFT) matrix, $\mathbf{F} = \exp(-\frac{2\pi ikn}{N})$, with $k, n = 0, \ldots, N-1$. Let $[\mathbf{X}]_{mk} = x_{mk}$ and $[\mathbf{V}]_{mk} = v_{mk}$, with $m = 1, \ldots, M$ and $k = 1, \ldots, N$, be the $M \times N$ signal and noise matrices, respectively, indicating $M$ (number of time snapshots) vectors of $N$ values discretized by the S/H bank.

The spectrum sensing engine decides on the occupancy of the channel by solving the binary hypothesis denoted by $\mathcal{H}_0$ indicating the channel is free and hypothesis $\mathcal{H}_1$ which indicates that the channel is occupied. The hypothesis testing problem is given by

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{y}_k = \mathbf{V}\mathbf{f}_k^T + \mathbf{q} \\ \mathcal{H}_1 &: \mathbf{y}_k = \mathbf{X}\mathbf{f}_k^T + \mathbf{V}\mathbf{f}_k^T + \mathbf{r} + \mathbf{q} \end{aligned} \tag{4.1}$$

This system model can also be viewed as a detection problem with multiplicative Gaussian noise.

### 4.3.2 Probability of false alarm and threshold

Here, the Neyman-Pearson criterion is considered, where we set a constraint on the probability of false alarm $P_{fa}$ and determine the detection threshold $\gamma_{th}$. The corresponding probability of detection $P_d$ for different Signal-to-Noise Ratios (SNRs) is shown through simulations. To determine the threshold for a certain $P_{fa}$, we next derive the distribution of the signal at the $k^{th}$ branch after the envelope detector (indicating its energy) under the $\mathcal{H}_0$ hypothesis.

The probability of false alarm can be written as

$$P_{fa} = Pr(E \geq \gamma_{th}|\mathcal{H}_0) \quad = \int_{\gamma_{th}}^{\infty} f_Z(z)dz \tag{4.2}$$

where, $E = \sum_{i=1}^{\eta} (y_{ki})^2$ is the energy at the output of the envelope detector and $\eta$ denotes the averaging achieved with the low-pass filter. $E$ can be modeled as a process defined by the RV, $z = \sum_{i=1}^{\eta}(q_i + v_i)^2 \approx \sum_{i=1}^{\eta}(u_i + v_i)^2$, where $q_i$, $u_i$ and $v_i$ are general RVs with a certain distribution.

The entries of $\mathbf{q}$ have a normal product distribution [58], i.e., $q_i \sim \mathcal{N}(0, \delta_N^2) \cdot \mathcal{N}(0, \frac{\sigma_v^2}{N})$. To simplify the derivation, we approximate the normal product distribution with a sum of two Gaussian functions denoted by the Probability Density Function (PDF) $f_U$. The simulations in Fig. 4.3 show that the sum of two Gaussian functions is a good approximation for a normal product distribution. The RVs $v_1, v_2, \ldots, v_\eta$ and $u_1, u_2, \ldots, u_\eta$ are independently and identically distributed (i.i.d.) with PDFs $f_V$ and $f_U$ respectively. The normal product distribution after approximation with the sum of two Gaussian functions is given by the PDF

$$f_U(u) = \frac{1}{\sqrt{\pi}(a_1\sigma_1 + a_2\sigma_2)}(a_1 \exp -(\frac{u}{\sigma_1})^2 + a_2 \exp -(\frac{u}{\sigma_2})^2) \tag{4.3}$$

and the PDF of the noise is given by

$$f_V(v) = \frac{1}{\sqrt{\pi}(\sigma_3)} \exp -(\frac{v}{\sigma_3})^2 \tag{4.4}$$

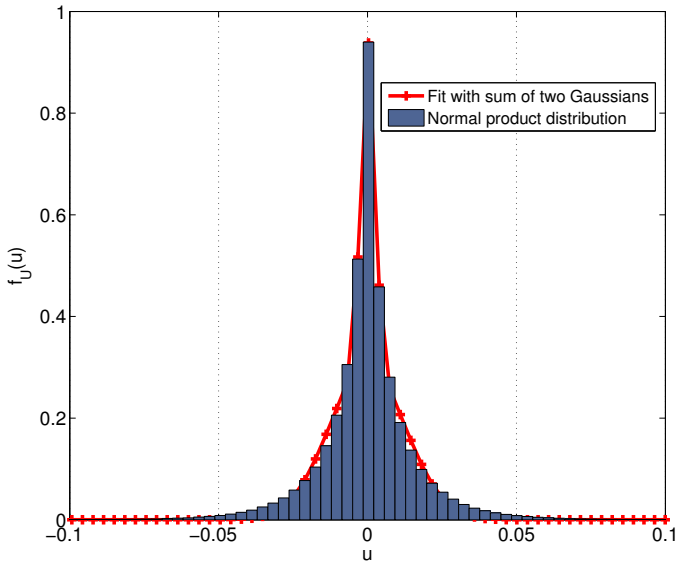



Figure 4.3: Normal product distribution vs. the best fit with a sum of two Gaussian functions.

where $a_1$ and $a_2$ are the weights, and $\sigma_1$ and $\sigma_2$ are the standard deviations of the Gaussian functions in (4.3); and $\sigma_3 = \sqrt{2}\sigma_v$ is just a scaled standard deviation of the noise.

Next, we derive the PDF $f_Z(z)$. To do this, we first derive the PDF $f_Z(z)$ for $\eta = 1$. Using the convolution property and square law [59], for $\eta = 1$, the PDF $f_Z(z)$ is given by

$$f_Z(z|\eta=1) = \frac{1}{\sqrt{\pi z(a_1\sigma_1 + a_2\sigma_2)}} \left( \frac{a_1\sigma_1}{\sqrt{\sigma_1^2+\sigma_3^2}} \exp(\frac{-z}{\sigma_1^2+\sigma_3^2}) + \frac{a_2\sigma_2}{\sqrt{\sigma_2^2+\sigma_3^2}} \exp(\frac{-z}{\sigma_2^2+\sigma_3^2}) \right) \tag{4.5}$$

The characteristic function of the RV, $z$ for $\eta = 1$ will be

$$\begin{aligned} \Omega_Z(\omega|\eta=1) &= \int_{-\infty}^{\infty} \exp(iz\omega) f_Z(z) dz \qquad \text{with} \quad \omega \in \mathbb{R}, \\ &= \frac{1}{\sqrt{a_1\sigma_1 + a_2\sigma_2}} \left( \frac{a_1\sigma_1}{\sqrt{1 - i(\sigma_1^2+\sigma_3^2)\omega}} + \frac{a_2\sigma_2}{\sqrt{1 - i(\sigma_2^2+\sigma_3^2)\omega}} \right) \end{aligned} \tag{4.6}$$

Hence, the characteristic function of the RV $z$ for a generic $\eta$ will be, $\Omega_Z(\omega) = \prod_{i=1}^{\eta} \Omega_Z(\omega|\eta = i) = (\Omega_Z(\omega|\eta = 1))^{\eta}$.

Using (4.6), the PDF $f_Z(z)$ can be written as

$$f_Z(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-iz\omega) \Omega_Z(\omega) d\omega \tag{4.7}$$

Substituting (4.6) in (4.7), we have

$$f_Z(z) = \frac{1}{2\pi(\varphi_1+\varphi_2)^{\eta}} \int_{-\infty}^{\infty} \exp(-iz\omega) \left( \frac{\varphi_1}{\sqrt{1-i\alpha\omega}} + \frac{\varphi_2}{\sqrt{1-i\beta\omega}} \right)^{\eta} d\omega \tag{4.8}$$

where, $a_1\sigma_1 = \varphi_1$ , $a_2\sigma_2 = \varphi_2$, $\sigma_1^2 + \sigma_3^2 = \alpha$, $\sigma_2^2 + \sigma_3^2 = \beta$.

Using binomial theorem, we can re-write (4.8) as

$$f_Z(z) = \frac{1}{2\pi(\varphi_1+\varphi_2)^\eta} \sum_{k=0}^{\eta} \binom{\eta}{k} \varphi_1^{\eta-k}\varphi_2^k \int_{-\infty}^{\infty} \exp(-iz\omega)(1-i\alpha\omega)^{\frac{\eta-k}{2}}(1-i\beta\omega)^{\frac{-k}{2}} d\omega \tag{4.9}$$

From [60], and further simplification, the PDF $f_Z(z)$ is given by

$$f_Z(z) = \frac{\exp(\frac{-z}{\alpha})z^{(\frac{\eta}{2}-1)}}{\Gamma(\frac{\eta}{2})(\varphi_1+\varphi_2)^\eta} \sum_{k=0}^{\eta} \binom{\eta}{k} \left(\frac{\varphi_1}{\sqrt{\alpha}}\right)^{\eta-k} \left(\frac{\varphi_2}{\sqrt{\beta}}\right)^k \; {}_1F_1\left(\frac{k}{2}; \frac{\eta}{2}; \frac{(\beta-\alpha)z}{\alpha\beta}\right) \tag{4.10}$$

where, $\Gamma(\theta) = \int_0^\infty \tau^{\theta-1}\exp(-\tau)d\tau$ is the Gamma function, and ${}_1F_1(a;b;z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)} \int_0^1 \exp^{zt} t^{(a-1)}(1-t)^{(b-a-1)}dt$ is the confluent hypergeometric function of the first kind [61].

The integral (4.2) can be solved numerically using a Chebyshev polynomial expansion [62], using standard software packages (e.g., by Mathematica). The threshold for a certain $P_{fa}$ can be obtained by solving the lower tail probability using Newton's method [63].

Fig. 4.4 shows the theoretical value of $P_{fa}$ obtained from (4.2) for different SNRs and the actual $P_{fa}$ for a fixed threshold. Deviations below 5 dB between the theoretical and actual values are obtained, which are due to the approximation of the normal product distribution with the sum of two Gaussian functions.

## 4.4 Performance evaluation and analysis

### 4.4.1 Simulations

A network in the 2400 MHz ISM band with 86 frequency bins of 1 MHz resolution centered at $f_c = 2400, 2401, \ldots, 2485$ MHz is simulated. The simulated scenario has four WiFi (IEEE 802.11g) nodes centered at 2412, 2432, 2452, 2472 MHz respectively and one Zigbee (IEEE 802.15.4) node centered at 2440 MHz. An illustration is provided in Fig. 4.5, which is constructed with a high resolution FFT (FFT length = 183430). The signals considered are present at all times with fixed transmit power.

The proposed system architecture is simulated such that the resolution of the samples mimic the analog signal, and the all-analog FFT is realized using the mismatch model as in Fig. 4.2. The spectrum reconstruction with the conventional (all-digital) method using a Nyquist rate ADC and a 128-point digital FFT, as well as the spectrum reconstruction obtained at the stage 1 of the proposed architecture are shown in Fig. 4.6. It can be seen that using the proposed architecture a coarse spectrum estimate can be obtained, with a reduction in the dynamic range compared to the conventional method. However, the worse spectrum reconstruction with the proposed architecture compared to the conventional approaches can be traded for a considerable power saving. Fig. 4.7 shows the spectrum reconstruction performance in terms of the Mean

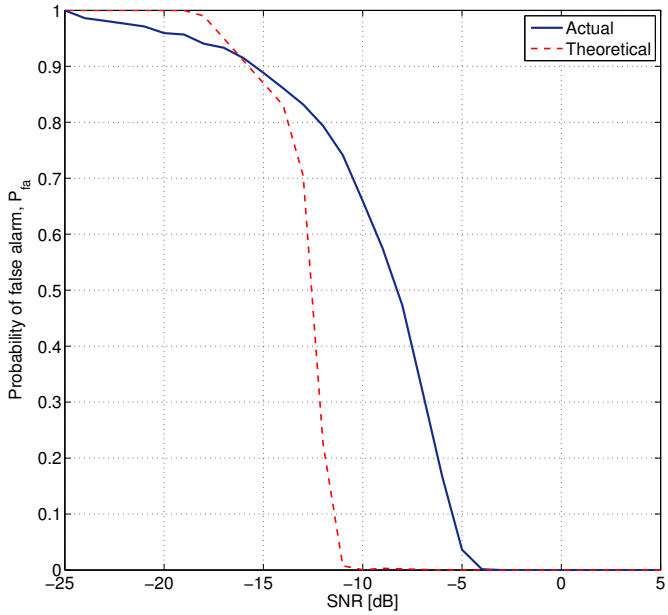



Figure 4.4: Actual and theoretical values of the probability of false alarm $P_{fa}$ for a fixed threshold and different SNRs.

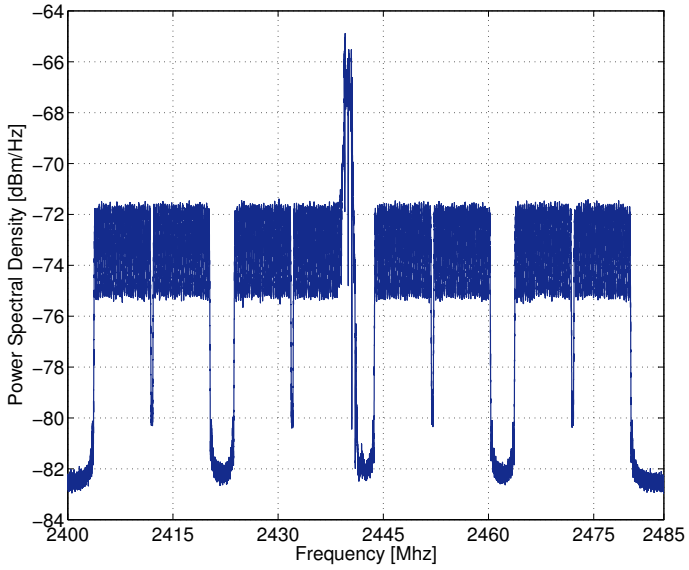



Figure 4.5: Spectrum with four IEEE 802.11g/WiFi nodes and 1 IEEE 802.15.4/Zigbee node.

Squared Error (MSE) between the high resolution FFT and the digital 128-point FFT for the conventional method and between the high resolution FFT and the 128-point analog FFT for the proposed method. The analog processing results in a deteriora-

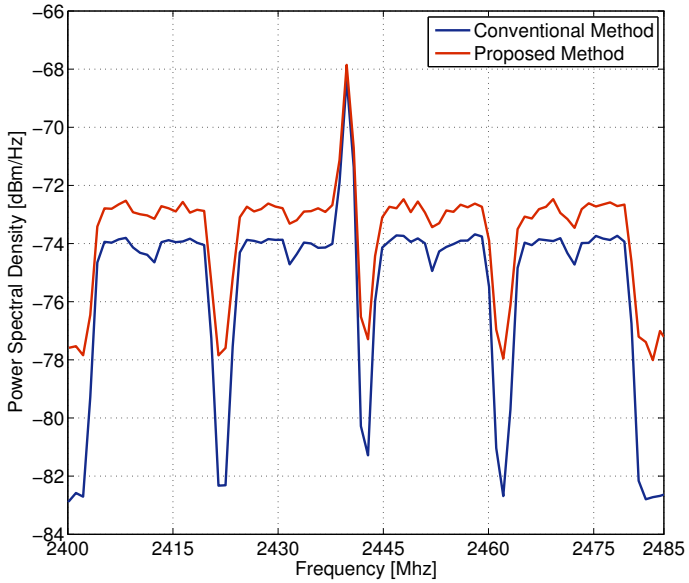



Figure 4.6: Spectrum reconstruction using both the conventional and proposed method with a 128-point FFT.

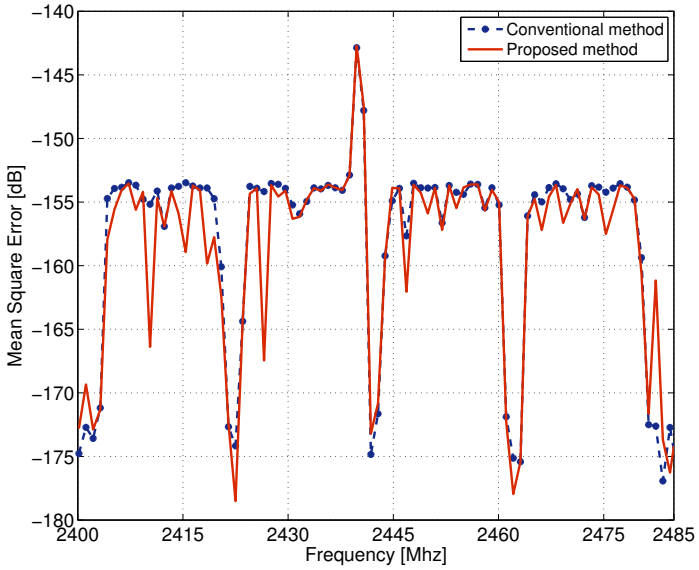



Figure 4.7: Mean squared error of the spectrum reconstruction.

tion of the signal strength, and introduces noise between the stages. An actual static channel occupancy of around 80% with a frequency resolution of 1 MHz shown in Fig. 4.8, is used to evaluate the detection performance. The detection performance in terms of $P_{fa}$ and $P_d$ for different SNRs is shown in Fig. 4.9. The threshold is set so as to

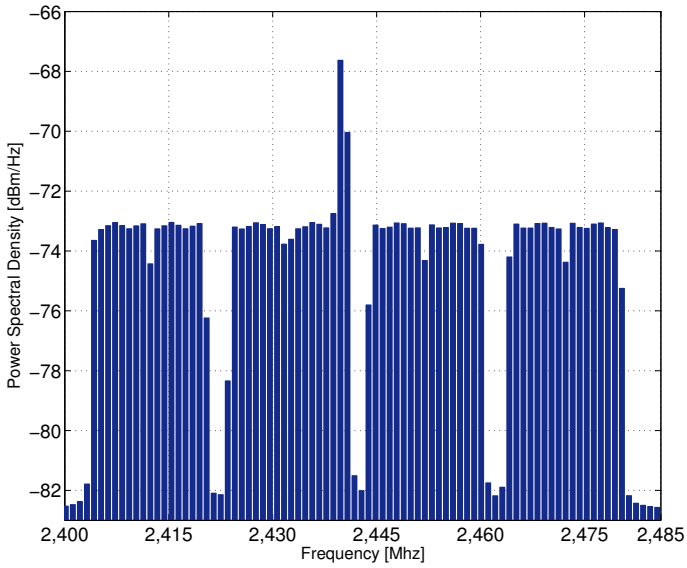



Figure 4.8: Smoothed periodogram indicating 86 frequency bins, with 1 MHz frequency resolution.

maintain a $P_{fa}$ below 5%. With the proposed architecture, a detection performance comparable to that of the conventional method can be achieved, with losses below 1 dB in the observed SNR range.

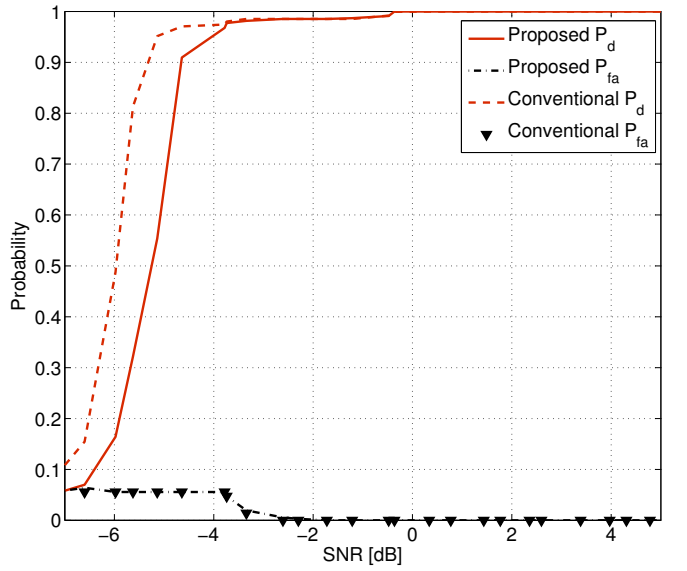



Figure 4.9: Detection performance.

### 4.4.2 Power consumption comparison: conventional vs. proposed

In this section, we compare the power consumption of the conventional all-digital approach, which involves sampling at Nyquist rate using a high-rate 8-bit ADC, with the application of an FFT using an 8-bit 128-point digital FFT processor. An 8-bit ADC based on 90nm CMOS technology is considered. The power consumption of the ADC is approximately $27\mu W/10$ Megasamples/second, i.e., 2.7pJ/conversion step [49]. The 8-bit state-of-the-art energy aware digital FFT processor proposed for low-power sensor nodes [64] consumes around 33nJ/FFT (a scaling factor of $(0.66)^3$ [65] for scaling the energy from 180nm to 90nm CMOS technology is used).

The proposed method uses S/H circuits which on average can be estimated to consume 10% of the ADC's power [66]. The analog 128-point FFT processor requires 512 differential inputs and 7 stages for a radix-2 implementation of the butterfly structure, and the number of multipliers from the 128-FFT stage to 8-FFT stage would be about 576 [55]. With a bias current of 100nA and $V_{dd} = 1.2V$ (for 90nm CMOS technology), the power consumption would be of the order of $(512\times7+576)(0.12\mu W) = 0.5mW$, and considering the FFT processor speed of 1 MHz, the energy/FFT will be 500 pJ/FFT. A summary of the comparison of the estimated power between the proposed and conventional method is given in Table 4.1.

Table 4.1: Estimated energy and power consumption for 128-point FFT systems

| Estimated | **Conventional** | **Proposed** |
|---|---|---|
| pJ/conversion | (ADC) **2.7** | (S/H) **2× 0.27 = 0.54** |
| pJ/FFT | (Digital) **33000** | (Analog) **≈500** |
| Power (mW) | (ADC + Digital FFT) **33.7** | (S/H + Analog FFT) **≈0.64** |

In the proposed architecture, to realize the periodogram estimate and channel selection, additional power is consumed. Envelope detectors can consume a power below $1.5\mu$W for state-of-the-art designs [67]. On the other hand, for the conventional method, these steps are done digitally using different algorithms with different complexity order, as discussed in section 2.1.1 of Chapter 2.

In the proposed architecture we choose energy detector, since it simple and of low-complexity. As any energy detector, this system also suffers from noise uncertainty and SNR wall issues. To enhance the performance in the low-SNR regimes, a two-stage sensing can be performed as suggested in the literature [68]. Typically, the second stage performs feature detection (e.g., cyclostationarity, pilot-tone detection) on the detected free narrowband channels, to improve the performance in low SNRs.

## 4.5 Conclusions

In this chapter, we have proposed an architecture for low-power wideband spectrum sensing. With the proposed architecture a large portion of the spectrum (e.g., 128 MHz) can be sensed at once to obtain a coarse estimate of the spectrum and/or search for an empty channel. The major high-power consuming processes are pushed to the analog domain which include the high-rate ADC and the digital FFT operations, which are

replaced by a bank of S/H circuits and an analog FFT processor, respectively. The simulation results show that even though analog processing leads to lower performance in spectrum reconstruction with respect to the conventional approaches, a good detection performance can still be achieved with a considerable reduction in power consumption. A closed-form expression for the probability of false alarm is also provided, as well as the threshold needed to achieve a target false alarm rate.

# 5 Conclusions

## 5.1 Conclusions

With the large portions of the usable spectrum being under-utilized and the ever increasing innovations of wireless devices, there is a need to efficiently utilize the scarce frequency spectrum. Cognitive Radios (CRs) or spectrum sharing radios have been identified as one of the solutions to the spectrum scarcity problem. These radios adaptively utilize the vacant frequency bands without causing interference to licensed Primary Users (PUs), and vacate these bands on sensing the PU activity. To realize this efficiently with low-latency, these radios should be capable of sensing a wide spectral range, in the order of a few hundred MHz. This requires wideband Radio-Frequency (RF) front-ends with high-rate Analog-to-Digital Converters (ADCs), followed by digital processing. This typically consumes high power. In case of sparse spectrum, the sampling rates can be reduced significantly depending on its sparse support and the requirements on the ADCs can be relaxed. These are often formulated as Compressive Sensing (CS) problems, where the conventional detection is done on the compressive estimate of the signal. Unfortunately, recovery of such compressed samples are computationally expensive.

In this thesis, we addressed two parallel techniques for wideband spectral sensing:

1. Algorithmic approach: To determine the occupancy of a sparse spectrum, we perform a direct detection on the compressive samples, termed as Compressed Detection (CD), and avoid the classical estimation-detection two-stage approach.

2. Architectural approach: When the spectrum is not necessarily sparse, the sampling rate cannot be relaxed much. In such cases (irrespective of the spectrum sparsity), a low-power architecture to reduce the complexity (power consumption) has been proposed.

In Chapter 3, we formulated the multiband occupancy detection as a Multiple Hypothesis Testing (MHT) problem under the Neyman-Pearson-like criterion. We developed the detector for two types of the acquired signals (observations). In the first case, the signals were acquired at the Nyquist rate. A linear complexity optimal detector MHT-WS:N (MHT based wideband sensing for Nyquist rate samples) has been developed for Nyquist rate observations. The optimal detector is the energy detector for $M = N$ case with double threshold for the Neyman-Pearson-like approach. Next, the detector was developed for the reduced dimensionality observations, i.e., the signals acquired at sub-Nyquist rate, as in the CS framework. The optimal detector MHT-WS:CD (MHT based wideband sensing: compressed detector) has a complexity of the order of $O(2^N)$ and outperforms the conventional CS two-stage approach. However,

the complexity gets impractical with larger $N$. Hence, we proposed a sub-optimal algorithm (LML-CD) to perform CD based on observed properties of the Local Maximum Log-likelihood (LML) ratio test points. The performance of the LML-CD algorithm is comparable to that of the classical two-stage approach. The complexity order of LML-CD algorithm is $O(N^3)$, and does not require the exact knowledge of the sparsity level.

In Chapter 4, an alternative low-power technique for wideband spectral sensing at the architecture level has been proposed. For scenarios where the spectrum is not so sparse, we cannot capitalize on the sub-Nyquist rate sampling approaches, but there is still need for low-power and low-complexity technique for wideband occupancy detection. For this purpose, we proposed an analog/mixed signal topology that replaces the conventional Nyquist rate ADCs and digital Fast Fourier Transform (FFT) core with a bank of Sample and Hold (S/H) circuits, each operating at sub-Nyquist rate, and an all-analog FFT processor. The results show that even though sub-optimal analog processing leads to lower performance in spectral reconstruction with respect to the conventional approaches, good detection performances can be achieved along with a substantial reduction in the power consumption.

In the next section, we conclude this thesis with some suggestions for future research.

## 5.2 Suggestions for future research

- **Generalized Likelihood Ratio Test (GLRT) approach for MHT-WS detector**:
  In Chapter 3, while deriving the detector for wideband sensing based on multiple hypothesis, we assumed that the active channel powers $\sigma_x^2$ were known. Alternatively, the unknown active channel powers can be replaced by their maximum likelihood estimates $\hat{\sigma}_x^2$ as in the GLRT approach, and then develop a detector.

- **Distributed detection**: In this thesis, we considered a non-cooperative setup, where only a single node was involved in performing the detection. To improve the detection under unfavorable channel conditions, and capitalize on the spatial diversity gain, a collaborative setup can be considered. In such distributed sensing schemes, a number of radios cooperate to improve the sensing performance. Each of these radios perform sensing locally and send either a hard or a soft value to a Fusion Center (FC) and the FC computes a global spectrum occupancy based on these local values. It is interesting to look at the performance gains when the radio solves the MHT problem locally, and a global decision is taken at the FC. Alternatively, a joint detection can be performed on (compressive) measurements obtained from the cooperating radios.

- **Properties of the sensing matrices for compressed detection**: It is well-known from the CS theory that for perfect reconstruction of signals sampled at sub-Nyquist rate using $\ell_1$-norm optimization, the sensing matrices should satisfy certain properties such as restricted isometry and incoherence. In order to perform a direct detection on the compressive measurements, it is important to study the properties that the sensing matrices have to satisfy. And also, it is interesting to

analyze the sensitivity to the measurement noise, and techniques to increase the robustness to the measurement noise.

- **Optimal and sub-optimal reduced complexity algorithms and performance guarantees for the sub-optimal algorithm**: In this thesis, we proposed sub-optimal algorithms based on certain heuristics. The performance of these sub-optimal algorithms was studied through simulations. More investigation and mathematical analysis is required on the convergence and stability of these sub-optimal algorithms. It is important to further look into the optimal algorithm with more effective techniques and/or approximations to solve (3.43). It is important to also look into other sub-optimal greedy algorithms with low-complexity that will approach the performance of the proposed optimal wideband sensing algorithm, MHT-WS:CD.

- **Architecture extension to CS framework**: In Chapter 4, the proposed architecture acquires the signal at the Nyquist rate, by periodically and uniformly enabling the sample and hold (S/H) circuits. The emerging interests in CS, motivates to extend and explore this low-power architecture for compressed detection. One intuitive way is to enable the S/H circuits in a pseudo-random manner. This results in a low-power realization of the structured sensing matrix (or an Analog-to-Information converter).

# Publications

1. (Pre-study)
   S. P. Chepuri, R. de Francisco, and G. Leus. Performance evaluation of an IEEE 802.15.4 cognitive radio link in the 2360-2400 MHz band. In Proc. of the IEEE Wireless Communications and Networking Conference (WCNC 2011), Cancun, Mexico, March 2011.

2. S. P. Chepuri, R. de Francisco, and G. Leus. Low-power architecture for wideband spectrum sensing. To be submitted to ICC 2012.

3. S. P. Chepuri, R. de Francisco, and G. Leus. Wideband sensing through multiple hypothesis testing. To be submitted to ICASSP 2012.

4. S. P. Chepuri, R. de Francisco, and G. Leus. Multiband occupancy detection in wireless sensors. Journal in preparation.

# Bibliography


[1] *FCC: Spectrum policy task force report.* ET. Docket No. 02-155, Nov. 2002.

[2] "General survey of radio frequency bands (30 MHz to 3 GHz): Vienna, virginia, september 1-5, 2009," Sep. 2010. [Online]. Available: http://www.sharedspectrum.com/papers/spectrum-reports/

[3] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201 – 220, feb. 2005.

[4] S. P. Chepuri, R. de Francisco, and G. Leus, "Performance evaluation of an IEEE 802.15.4 cognitive radio link in the 2360-2400 MHz band," in *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*, march 2011, pp. 2155 –2160.

[5] I. Mitola, J., "Cognitive radio for flexible mobile multimedia communications," in *Mobile Multimedia Communications, 1999. (MoMuC '99) 1999 IEEE International Workshop on*, 1999, pp. 3 –10.

[6] D. Cabric, I. O'Donnell, M.-W. Chen, and R. Brodersen, "Spectrum sharing radios," *Circuits and Systems Magazine, IEEE*, vol. 6, no. 2, pp. 30 –45, 2006.

[7] K. C. Chen and R. Prasad, *Cognitive Radio Networks.* John Wiley and Sons, Chichester, West Sussex., 2009.

[8] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 1, pp. 116 –130, quarter 2009.

[9] R. de Francisco, "Sequential search of available channels in cognitive radios," in *IEEE GLOBECOM Workshops (GC Wkshps), 2010*, dec. 2010, pp. 1162 –1166.

[10] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289 –1306, 2006.

[11] Y. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wide-band spectrum sensing," in *Proc. of ICASSP 2009*, 2009, pp. 2337 –2340.

[12] I. E.Nesterov, A. Nemirovskii, and Y. Nesterov, *Interior-Point Polynomial Algorithms in Convex Programming.* SIAM, 1994.

[13] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655 –4666, dec. 2007.

[14] A. Sahai and D. Cabric, "Spectrum sensing: fundamental limits and practical challenges," *A tutorial presented at IEEE DySpan conference*, Nov. 2005.

[15] Z. Tian and G. B. Giannakis, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *Cognitive Radio Oriented Wireless Networks and Communications, 2006. 1st International Conference on*, june 2006, pp. 1 –5.

[16] R. Venkataramani and Y. Bresler, "Optimal sub-Nyquist nonuniform sampling and reconstruction for multiband signals," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2301 –2313, oct 2001.

[17] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.

[18] Y. Wang, A. Pandharipande, and G. Leus, "Compressive sampling based mvdr spectrum sensing," in *Cognitive Information Processing (CIP), 2010 2nd International Workshop on*, june 2010, pp. 333 –337.

[19] G. Leus and D. Ariananda, "Power spectrum blind sampling," *IEEE Signal Process. Lett.*, vol. 18, no. 8, pp. 443 –446, aug. 2011.

[20] Z. Tian and G. Giannakis, "Compressed sensing for wideband cognitive radios," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 4, april 2007, pp. IV–1357 –IV–1360.

[21] S. M. Mishra, R. Tandra, and A. Sahai, "The case for multiband sensing," in *45th Allerton Conference Communication, Control, and Computing, Monticello, IL.*, sep 2007.

[22] E. Candes and M. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21 –30, march 2008.

[23] E. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406 –5425, dec. 2006.

[24] S. Muthukrishnan, *Data Streams: Algorithms and Applications*, 2005.

[25] D. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845 –2862, nov 2001.

[26] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489 – 509, feb. 2006.

[27] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. pp. 55–67, 1970.

[28] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. pp. 267–288, 1996.

[29] D. L. Donoho, Y. Tsaig, I. Drori, and J. L. Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit (StOMP)," 2007.

[30] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 31 2008-april 4 2008, pp. 3869 –3872.

[31] T. Ragheb, S. Kirolos, J. Laska, A. Gilbert, M. Strauss, R. Baraniuk, and Y. Massoud, "Implementation models for analog-to-information conversion via random sampling," in *Circuits and Systems, 2007. MWSCAS 2007. 50th Midwest Symposium on*, aug. 2007, pp. 325 –328.

[32] J. Laska, S. Kirolos, Y. Massoud, R. Baraniuk, A. Gilbert, M. Iwen, and M. Strauss, "Random sampling for analog-to-information conversion of wideband signals," in *Design, Applications, Integration and Software, 2006 IEEE Dallas/-CAS Workshop on*, oct. 2006, pp. 119 –122.

[33] J. Tropp, M. Wakin, M. Duarte, D. Baron, and R. Baraniuk, "Random filters for compressive sampling and reconstruction," in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 3, may 2006, p. III.

[34] S. Kirolos, J. Laska, M. Wakin, M. Duarte, D. Baron, T. Ragheb, Y. Massoud, and R. Baraniuk, "Analog-to-information conversion via random demodulation," in *Design, Applications, Integration and Software, 2006 IEEE Dallas/CAS Workshop on*, oct. 2006, pp. 71 –74.

[35] T. Ragheb, J. Laska, H. Nejati, S. Kirolos, R. Baraniuk, and Y. Massoud, "A prototype hardware for random demodulation based compressive analog-to-digital conversion," in *Circuits and Systems, 2008. MWSCAS 2008. 51st Midwest Symposium on*, aug. 2008, pp. 37 –40.

[36] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, "Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals," *Information Theory, IEEE Transactions on*, vol. 56, no. 1, pp. 520 –544, jan. 2010.

[37] M. Mishali and Y. Eldar, "From theory to practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals," *IEEE Trans. Signal Process.*, vol. 4, no. 2, pp. 375 –391, april 2010.

[38] H. Trees, *Detection, Estimation, and Modulation Theory: Part 1.* John Wiley & Sons, NY, 2001.

[39] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. II: Detection Theory.* Prentice-Hall, Englewood Cliffs, NJ, 2004.

[40] M. Marcus and P. Swerling, "Sequential detection in radar with multiple resolution elements," *IEEE Trans. Inf. Theory*, vol. 8, no. 3, pp. 237 –245, april 1962.

[41] M. Green and A. Zoubir, "Multiple hypothesis testing for time-varying nonlinear system identification," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, vol. 1, 2000, pp. 560 –563 vol.1.

[42] E. Lehmann, *Testing Statistical Hypotheses*. Wiley, 1959.

[43] H. Trees, *Detection, Estimation, and Modulation Theory, Part III, Radar-Sonar Signal Processing and Gaussian Signals in Noise*. John Wiley & Sons, NY, 2003.

[44] S. Kay, A. Nuttall, and P. Baggenstoss, "Multidimensional probability density function approximations for detection, classification, and model order selection," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2240 –2252, oct 2001.

[45] Y. Sun, "Local maximum likelihood multiuser detection for cdma communications," in *Information Technology: Coding and Computing, 2001. Proceedings. International Conference on*, apr 2001, pp. 307 –311.

[46] ——, "A family of likelihood ascent search multiuser detectors: an upper bound of bit error rate and a lower bound of asymptotic multiuser efficiency," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1743 –1752, june 2009.

[47] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," vol. 4, no. 2, pp. 310 –316, april 2010.

[48] S. Gishkori, G. Leus, and H. Delic and, "Energy detectors for sparse signals," in *Signal Processing Advances in Wireless Communications (SPAWC), 2010 IEEE Eleventh International Workshop on*, june 2010, pp. 1 –5.

[49] P. Harpe, C. Zhou, X. Wang, G. Dolmans, and H. de Groot, "A 12fJ/conversion-step 8bit 10MS/s asynchronous SAR ADC for low energy radios," in *Proc. of ESSCIRC*, 2010, pp. 214 –217.

[50] M. Vidojkovic and et al., "A 2.4GHz 90nm CMOS ultra-low power OOK SoC transceiver for wireless body area network (WBAN) applications," in *Proc. of ISSCC*, Feb. 2011.

[51] N. Sadeghi, V. Gaudet, and C. Schlegel, "Analog DFT processors for OFDM receivers: Circuit mismatch and system performance analysis," *IEEE Trans. Circuits Syst. I*, vol. 56, no. 9, pp. 2123 –2131, 2009.

[52] M. Lehne and S. Raman, "An analog mixed signal fourier transform pre-processor for OFDM receivers," in *Proc. of WAMICON*, 2006, pp. 1 –5.

[53] R. Walden, "Analog-to-digital converter survey and analysis," *Selected Areas in Communications, IEEE Journal on*, vol. 17, no. 4, pp. 539 –550, apr 1999.

[54] J. Yang, R. Brodersen, and D. Tse, "Addressing the dynamic range problem in cognitive radios," in *Communications, 2007. ICC '07. IEEE International Conference on*, june 2007, pp. 5183 –5188.

[55] N. Sadeghi, *Analog FFT interface for ultra low power analog receiver architectures.* M.Sc. thesis, Univ. Alberta, Edmonton, Canada, 2007.

[56] M. H. Hayes, *Statistical digital signal processing and modeling.* John Wiley and Sons, New York, 1996.

[57] D. Cabric, S. Mishra, and R. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proc. of ASILOMAR*, vol. 1, 2004, pp. 772 – 776 Vol.1.

[58] E. W. Weisstein, *Normal Product Distribution.* From MathWorld–A Wolfram Web Resource.

[59] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, New York, 2002.

[60] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products.* Academic, New York, 1994.

[61] M. Abramowitz. and I. A. E. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* New York: Dover, 1972.

[62] "Computation of oscillating integrals," *Journal of Computational and Applied Mathematics*, vol. 1, no. 3, pp. 153 – 164, 1975.

[63] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge University Press, New York, NY, 2004.

[64] B. Calhoun, D. Daly, N. Verma, D. Finchelstein, D. Wentzloff, A. Wang, S.-H. Cho, and A. Chandrakasan, "Design considerations for ultra-low energy wireless microsensor nodes," *IEEE Trans. Comput.*, vol. 54, no. 6, pp. 727 – 740, Jun. 2005.

[65] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits - A Design Perspective.* Prentice-Hall International, 1996.

[66] P. Harpe, *Concepts for Smart AD and DA Converters.* PhD thesis, 2010.

[67] M. Durante and S. Mahlknecht, "An ultra low power wakeup receiver for wireless sensor nodes," in *Proc. of SENSORCOMM*, 2009, pp. 167 –170.

[68] S. Maleki, A. Pandharipande, and G. Leus, "Two-stage spectrum sensing for cognitive radios," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, march 2010, pp. 2946 –2949.