Adaptive Spectrum Sensing for Cognitive Radio Networks

Arash Vakili

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

October 2011

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Engineering.

2011/10/24

Abstract

Spectrum sensing is an important functionality of cognitive radio as a means to detect the presence or absence of the primary user (PU) in a certain spectrum band. Energy detection is a widely used spectrum sensing technique based on the assumption that the PU is either present or absent during the whole sensing period. However, this assumption is not realistic in a dynamic environment where the PU could appear or disappear at any time. The performance of the conventional energy detector (ED) actually deteriorates in the scenario where the PU activity status changes during the sensing period. Therefore, it is crucial to design a detector which can adapt to such an environment and reliably detect a change in the PU activity. Several *sequential* change detection techniques already exist in the literature; however, change detection in a *fixed* sensing duration has not been given enough attention.

In this dissertation, three adaptive EDs are proposed to improve the detection performance in dynamic environments, where there is a *single* change in the PU activity during a *fixed* sensing period. In particular, we address the change detection problem using an exponential weighting approach and two theoretical approaches based on the *composite hypothesis testing*. In the first case, an intuitive idea of exponential weighting of the received energies is applied to design an adaptive ED that aims to satisfy the Neyman-Pearson (NP) criterion. The performance analysis and simulation results prove that the proposed adaptive ED outperforms the conventional ED and also the only existing adaptive ED in the literature that deals with the aforementioned issue. In the second case, two theoretical approaches based on the composite hypothesis testing are used to design two additional adaptive EDs that improve the change detection during the sensing period. The first approach, known as the generalized likelihood ratio test (GLRT), uses the maximum likelihood estimation (MLE) of the unknown change location in a likelihood ratio test. In this case, an iterative method is proposed to reduce the computational complexity of the MLE process. The second approach, referred to as composite-Bayesian, assumes that the unknown change location is a discrete random variable whose probability mass function (PMF) is available. The PU channel access pattern is modelled as a two-state Markov chain to obtain the PMF of the change location and the probability of occurrence of the two hypotheses. The resultant adaptive ED based on the GLRT approach aims to satisfy the NP criterion while the adaptive ED based on the composite-Bayesian approach aims to

minimize the probability of error. It is demonstrated through simulations that these two proposed adaptive EDs have superior performance over the conventional ED. Furthermore, the GLRT-based adaptive ED outperforms the first proposed adaptive ED based on the exponential weighting approach.

Sommaire

La détection de spectre est une fonctionnalité importante de la radio cognitive car elle permet de vérifier la présence ou l'absence d'un utilisateur principal (PU) sur une bande de spectre donnée. La détection de l'énergie est une méthode fréquemment utilisée pour y parvenir. Cette dernière s'appuie sur l'hypothèse que le PU est présent ou absent pour la totalité de la période de mesure. Cependant, cette hypothèse n'est pas réaliste pour un environnement dynamique dans lequel le PU peut apparaître ou disparaître à n'importe quel instant. En effet, les performances d'un détecteur d'énergie conventionnel (ED) se détériorent lorsque l'état du PU varie au cours de la période durant laquelle les mesures sont effectuées. C'est donc pour cette raison qu'il est nécessaire de concevoir un détecteur qui s'adapte bien à ce genre d'environnement et qui permet de détecter de manière fiable tout changement dans l'activité du PU. Plusieurs techniques de détection de changements séquentiels existent dans la littérature mais la détection de changement pour une durée fixe n'a pas été explorée suffisamment en détails.

Dans le cadre de ce mémoire, trois EDs adaptatifs sont proposés dans le but d'améliorer les performances dans un environnement dynamique au sein duquel il y a un seul changement au niveau de l'activité du PU et ce durant une période de mesure de durée fixe. Pour tenter de résoudre cette problématique, une approche à pondération exponentielle et deux approches théoriques en lien avec le test d'hypothèse composée sont proposées. Dans le premier cas, une approche intuitive exploitant la pondération exponentielle de l'énergie mesurée est utilisée afin de concevoir un ED adaptatif qui satisfait le critère de Neyman-Pearson (NP). L'analyse des performances et des résultats de simulation prouvent que cette stratégie offre de meilleures performances par rapport aux ED conventionnels. Il s'agit également du seul ED adaptatif présent dans la littérature qui tente de résoudre la problématique précédemment mentionnée. Dans le second cas, deux approches théoriques fondées sur le test d'hypothèse composée sont utilisées afin de concevoir deux nouveaux EDs adaptatifs qui améliorent la détection de changements durant la période de mesure. La première approche s'appuie sur le test généralisé de vraisemblance (GLRT) et utilise une estimation de la vraisemblance maximale (MLE) de la position inconnue du changement. Dans ce cas, une méthode itérative est proposée pour réduire la complexité de calcul du processus de MLE. La deuxième approche, dite composée bayésienne, prend pour acquis que la position inconnue du changement est une variable aléatoire discrète dont la loi

de probabilité (PMF) est connue. Pour cette dernière approche, les accès au canal sont modélisés par un modèle de Markov à deux états afin d'obtenir la PMF de la position du changement et la probabilité d'occurrence des deux hypothèses. Le ED adaptatif utilisant le GLRT tente de satisfaire le critère de NP tandis que le ED adaptatif utilisant l'approche de la composée bayésienne tente de minimiser la probabilité d'une erreur. Il est démontré à l'aide de simulations que ces deux EDs adaptatifs offrent des performances supérieures à celles du ED conventionnel. En outre, le ED adaptatif utilisant le GLRT surpasse le ED adaptive utilisant l'approche pondération exponentielle.

Acknowledgments

Throughout my graduate study, I have been blessed with support and friendship of many whose insights and advices have tremendously enlightened my way of thinking. First, I would like to express my deepest gratitude to my supervisor, Dr. Benoît Champagne, for his consistent support, guidance, and encouragement. I have been extremely fortunate to have a supervisor who made the arduous journey through graduate school a rewarding and continuous learning experience that would last with me forever. Without his vast knowledge, great experience, and remarkable insights, this dissertation would have never been successful.

I have also been blessed to receive unconditional guidance and support from my colleagues at the Telecommunications and Signal Processing laboratories whose names are provided in the alphabetic order as follows: Reza Abdolee, Dr. Neda Ehtiati, Khalid Hossein, Siavash Rahimi, Fang Shang, Yunhao Tian, Siamak Yousefi, Chao Zho. Very special thanks are reserved for Dr. Ayman Asra for his remarkable comments and my friend, Francois Grondin, for providing the French translation of the abstract.

I would like to acknowledge that my Master research was supported by Le Fonds Québécois de la Recherche sur la Nature et les Technologies and I also acknowledge the support from InterDigital Canada Ltée., the Natural Sciences and Engineering Research Council of Canada, and the Government of Québec under the PROMPT program.

Finally, I extend my most gratefulness to my family who supported me emotionally and financially throughout my studies. To my mother and father, Farah and Ghodrat, who tolerated countless days and nights being away from me throughout my entire studies at McGill University. To my brother, Rouzbeh, for his friendship, encouragement, and advices.

To my beloved aunt, Parivash Aghabozorgi...

Contents

List of Figures

List of Abbreviations

AWGN	Additive White Gaussian Noise
CAF	Cyclic Autocorrelation Function
CR	Cognitive Radio
CSCG	Circularly Symmetric Complex Gaussian
DFT	Discrete Fourier Transform
ED	Energy Detector
FFC	Federal Communication Commission
FFT	Fast Fourier Transform
GLRT	Generalized Likelihood Ratio Test
i.i.d	Independent and Identically Distributed
ISM	Industrial Science and Medical
LR	Likelihood Ratio
MLE	Maximum Likelihood Estimation
NP	Neyman-Pearson
PMF	Probability Mass Function
PDF	Probability Density Function
PU	Primary User
SCF	Spectral Correlation Function
SDR	Software-Defined Radio
SU	Secondary User
SSC	Spectrum Shared Company
SNR	Signal-to-Noise Ratio
TV	Television
TVBD	TV Band Device

UMP	Uniformly Most Powerful
WLAN	Wireless Local Area Network
WSS	Wide-Sense Stationary

List of Notations

x^*	Complex conjugate of the number x
Re$[x]$	Real part of the complex number x
Im$[x]$	Imaginary part of the complex number x
$\cup$	Union (set theory)
$\cap$	Intersection (set theory)
$\subseteq$	Subset (set theory)
$\emptyset$	Empty set
$\mathbb{R}$	Set of real numbers
$\mathbb{C}$	Set of all complex numbers
$\mathbb{Z}^+$	Set of all positive integers:{1, 2, 3,...}
ln(.)	Natural logarithm (to base e)
$\equiv$	Equivalent to (notational)
$\mathcal{N}(\mu, \sigma^2)$	Real-valued Gaussian distribution with mean μ and variance σ^2
$\mathcal{CN}(\mu, \sigma^2)$	CSCG distribution with mean μ and variance σ^2
$E[.]$	Expectation operator
$Var[.]$	Variance operator
$Q(.)$	Complementary cumulative distribution function of standard normal
$\chi^2(v)$	Central chi-square variate with v degrees of freedom

Chapter 1

Introduction

The radio spectrum, which expands from 3 KHz to 300 GHz, is typically regulated by governmental agencies in most countries, such as the Federal Communication Commission (FCC) and the National Telecommunications and Information Administration in United States or the Industry Canada in this country. Based on the static spectrum allocation policy, regulatory bodies divide the radio spectrum into licensed and unlicensed bands. An exclusively assigned frequency band to a licensee for a particular use in a geographical region, is referred to as the *licensed band*. The licensed bands are generally reserved, sold, or leased to a specific user. Conversely, the unlicensed bands are available to all users free of charges as long as they adhere to the published requirements by the regulatory body.

In the past two decades, there has been a vast development in personal wireless devices. Some of these devices such as mobile phones, television (TV) receivers, and global positioning systems use the licensed frequency bands, while others such as wireless local area network (WLAN) devices and cordless phones use the unlicensed bands, most notably the industrial science and medical (ISM) bands. Up to now, most of the radio spectrum has already been licensed and the unlicensed ISM bands have been filling up very rapidly due to the popular Wi-Fi devices. Therefore, the wireless spectrum resources have become scarce in recent years as more wireless devices/applications have emerged in the market.

Nonetheless, measurement campaigns have shown that significant portion of the allocated spectrum are not often being used efficiently [1–4]. Such a recent spectrum measurement in United States was done by Shared Spectrum Company (SSC) in the prime radio frequency band of 30 MHz to 3 GHz over a three-and-half day period in Vienna, VA, in the

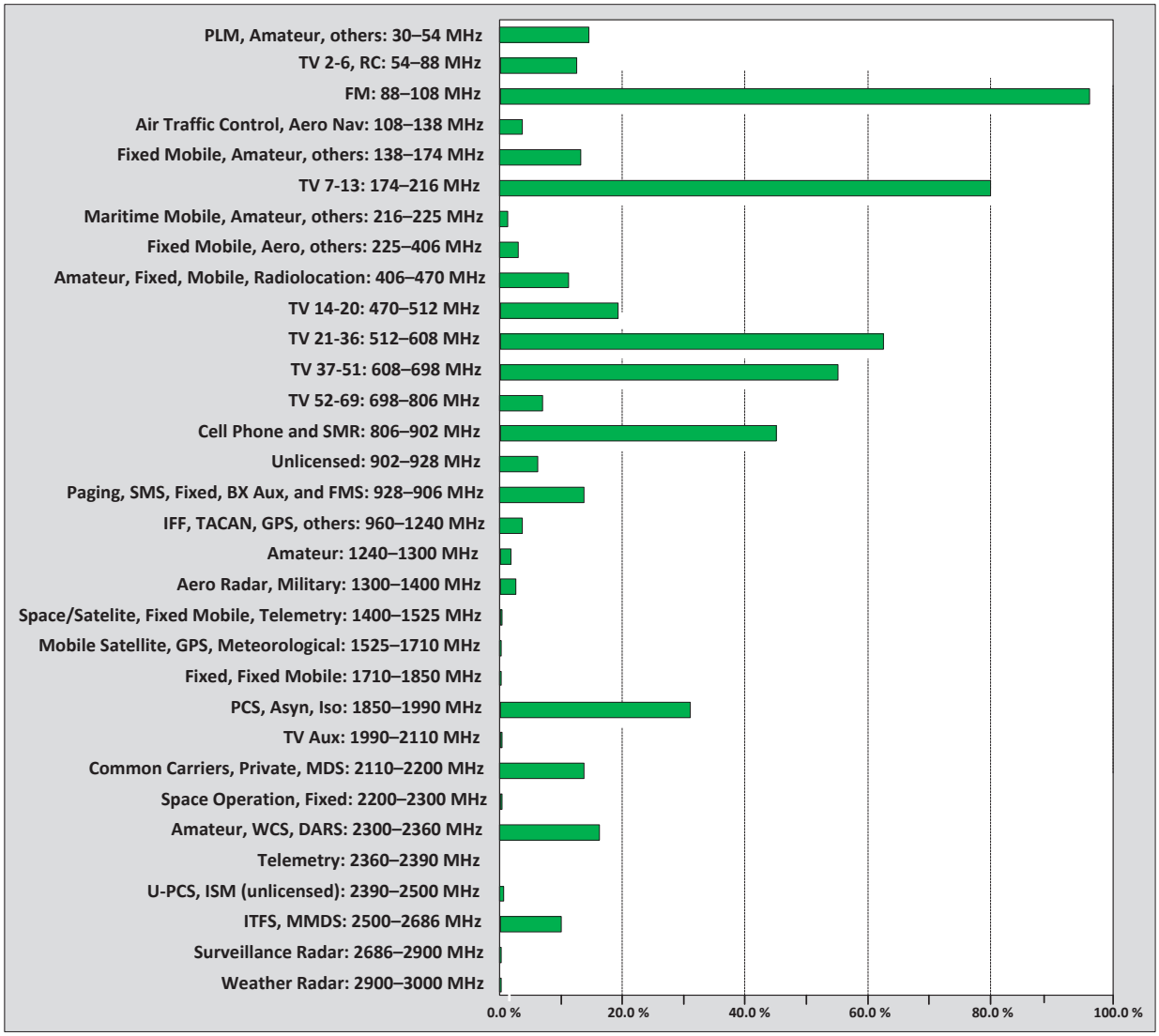


Fig. 1.1 Summary of spectrum occupancy measurements by SSC in 2009 [3].

fall of 2009 [3]. The result of this measurement, summarized as the percentage of the time that the measured power in a given frequency band exceeds a specified threshold, is shown in Fig. 1.1, where it is observed that all the prime radio spectrum is underutilized. Even though, the degree of underutilization in general varies greatly depending on the location, frequency band, and time; this and similar result tend to prove that the static spectrum allocation policy is highly inefficient.

The opportunistic use of the vacant frequency bands has been proposed as a promising solution to the spectrum underutilization and scarcity [5–9]. In the past few years, several important developments have been made in the spectrum policy and regulations to accelerate opportunistic spectrum usage. The most recent ones are the publication of the National Broadband Plan in March 2010 [10], the final report of FCC for unlicensed use of TV white spaces in September 2010 [11], and the ongoing proceeding on secondary use of the $2360 - 2400$ MHz bands for medical body area networks [12]. Cognitive radio is the key technology behind opportunistic spectrum usage which supports a variety of emerging wireless applications, ranging from public safety, smart grid, and broadband cellular, to medical applications. Several standardization efforts based on cognitive radio technology are under progress, including IEEE 802.22 [13], IEEE 802.11af, ECMA 392 [14], IEEE SCC41, and ETSI RRS [15].

1.1 The Cognitive Radio Paradigm

Over the years, the notion of radio has evolved from a pure hardware based design to a combination of hardware and software. Mitola coined the term "software radio", now commonly known as "software-defined radio" (SDR), to indicate the transition from a predominately analog radio to a multi-band, multi-mode, and multi-carrier radio whose functionalities are mostly realized in software [16]. SDR basically consists of a reconfigurable device such as field-programmable gate array or digital signal processor, an analog-to-digital converter, and a radio frequency front-end with software-controlled tuner. In 2000, Mitola introduced cognitive radio (CR) as an integrated agent architecture for SDR when addressing the broad issue of wireless personal digital assistants in his Ph.D dissertation [17], where he defined CR as:

> "*The point in which wireless personal digital assistants (PDAs) and the related networks are sufficiently computationally intelligent about radio resources and*

related computer-to-computer communications to detect user communications needs as a function of use context, and to provide radio resources and wireless services most appropriate to those needs".

Ever since, CR has gained considerable attention as the enabling technology to solve the spectrum scarcity and underutilization problems by providing the capabilities to use or share the spectrum in an opportunistic manner. However, a clear description of the CR network architecture is essential to understand its concept.

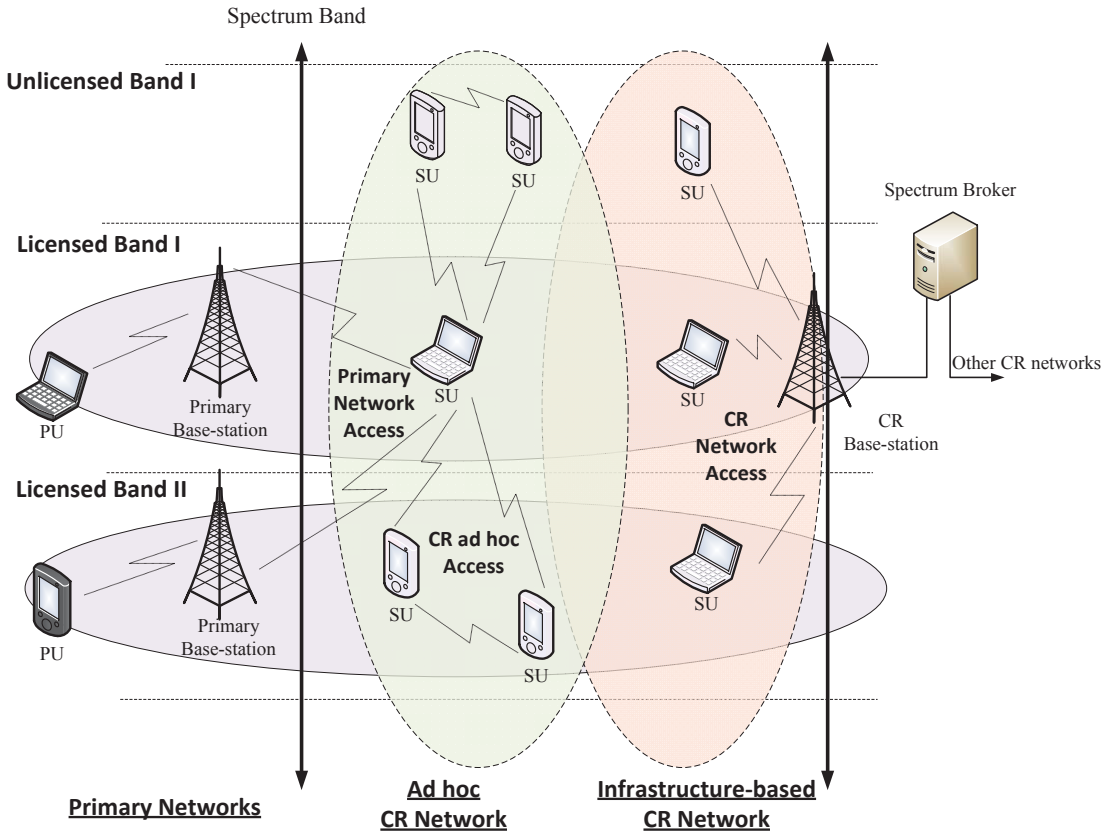


Fig. 1.2 CR network architecture [18].

CR network architecture can be classified into primary network and CR network (or secondary network) as shown in Fig. 1.2 [18]. The primary network is composed of several primary users (PUs) who have the exclusive right to operate in a certain spectrum band under coordination of the primary base-station and whose transmissions should not be interfered by any other unlicensed users. The primary base-station is a fixed infrastructure

network component which does not have the capability to share the licensed spectrum; however, it can be configured to have additional protocol for the primary network access of CR users. The CR network is only allowed to access the licensed or unlicensed spectrum bands in an opportunistic manner and its components are CR users, CR base-station, and spectrum broker. CR users, also referred to as secondary users (SUs), have additional functionalities to use or share the spectrum opportunistically. The CR base-station is a fixed infrastructure network component which has the same functionalities as CR users. The spectrum broker is the central entity responsible for sharing a common spectrum band among several CR networks.

In a CR network, the spectrum can be accessed in three different ways. In the first approach referred to as *CR network access*, a SU can access the licensed and unlicensed bands through a single hop connection to the CR base-station. The second approach, known as *CR ad hoc access*, allows several SUs to communicate with each other on both licensed and unlicensed bands in a multi-hop manner. Finally, in a *primary network access*, a SU can access a licensed band through a primary base-station that is configured for secondary usage. Based on the aforementioned access protocols, CR networks can be further classified into infrastructure-based and ad hoc networks. In an infrastructure-based network, the CR base-station is responsible to manage the operation of the SUs and provide access to other networks. More precisely, the CR base-station receives observations of the individual SUs in the network and makes the final decision on how to avoid interference to the PU in the licensed band or the other SUs in the unlicensed band. The SUs in turn configure their communication parameters according to this decision. In an ad hoc CR network, several CR users can communicate with each other in a multi-hop manner or can access the primary base-station through the primary network access point. The ad hoc network does not have any infrastructure component; therefore, each SU is responsible to determine its action based on its own observation. However, cooperative scheme could be deployed in this case, where the SUs exchange their local observations among each other to increase the reliability of their decisions [19–21].

CR networks can be designed to operate in the unlicensed bands in order to improve the efficiency in this portion of the spectrum; however, the most important challenge is to design a CR network to share the licensed spectrum with the PU. In this thesis, we consider CR network designs that exploit the temporarily-unused licensed spectrum (see Fig. 1.3), which is also referred to as spectrum hole or white space [23], without causing

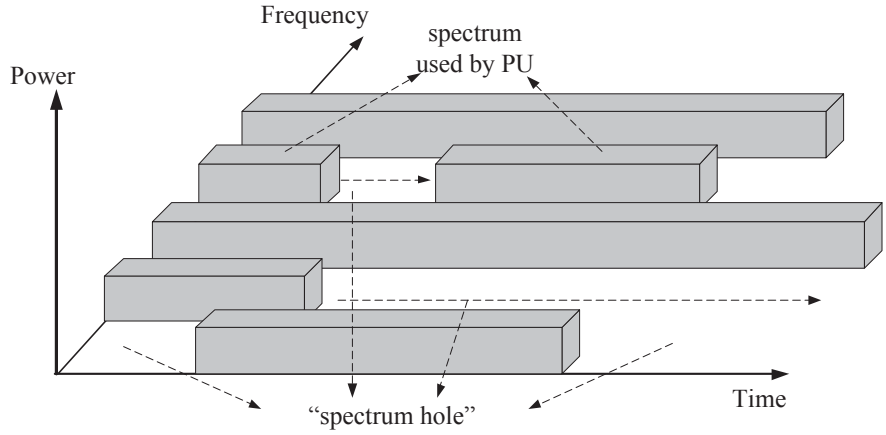


Fig. 1.3 Spectrum hole concept [22].

harmful interference to the PU. The main functions of the CR network to fulfil this task are described as follows:

1. *Spectrum Sensing* is the fundamental task of monitoring the licensed spectrum to reliably detect the spectrum holes and the PU emergence for the purpose of opportunistic transmission and interference avoidance, respectively.

2. *Spectrum Management* is the process of selecting the best spectrum band among all the vacant bands, which meets the CR user communication requirements and quality of service requirements, and maximizes the spectrum efficiency.

3. *Spectrum Mobility* allows CR nodes to vacate the channel when the PU emerges and continue transmission on another vacant channel. It also provides the opportunity for CR to switch to a better channel when the current one does not meet the user's requirements, and maintains a seamless communication requirement during the transition.

4. *Spectrum Sharing* aims to fairly distribute the vacant spectrum among a large number of CR nodes within the network.

1.2 Spectrum Sensing

Spectrum sensing is the essential function of CR network as a means to detect the presence or absence of the PUs in a certain spectrum band at a given time. Although FCC has eliminated spectrum sensing for TV band devices (TVBDs) that use geo-location and database access, this technology still offers a significant promise for spectrum efficiency and access both in TV broadcast bands and other spectrum bands. Indeed, FCC still retains the provision in their rules that permits the authorization and operation of personal/portable TVBDs that rely only on spectrum sensing to determine a list of available channels [11]. On the same note, the other regulatory bodies around the world, such as the Office of Communications in the United Kingdom and the Electronic Communications Committee of the European Conference of Postal and Telecommunications Administrations in Europe, also suggested sensing as the possible solution for incumbent protection in TV broadcast bands [24, 25]. Hence, spectrum sensing remains an important task for the establishment of opportunistic spectrum usage.

Spectrum sensing techniques are typically based on two approaches, namely block-based detection and sequential detection. In the block-based detection, the CR takes a block of samples, computes a statistic and compares it with a threshold to decide on the occupancy state of the spectrum. The ultimate goal in the *non-Bayesian* formulation is to maximize the probability of PU signal detection subject to a constraint on the probability of false alarm, that is the probability of falsely declaring the presence of the PU; whereas the *Bayesian* formulation considers the ensemble minimization of miss detection and false alarm probabilities. Various block-based detection schemes have been proposed in the literature, according to the SU's knowledge on the transmitted signal and the required receiver complexity. Matched filter detection [26, 27], cyclostationary feature detection [28–30] and energy detection [31–33] are the representative ones.

The matched filter is the optimal detector in stationary Gaussian noise when the transmitted signal is known to the SU [34]. However, its implementation complexity and power consumption is extremely high, because matched filter requires separate receivers for all types of signals and their corresponding receiver algorithms have to be executed [35, pp. 17-18]. The cyclostationary feature detector exploits the inherent periodicity of the transmitted signal to distinguish the PU signal with a particular modulation type from the stationary noise [36]. It requires a longer processing time and a more complex receiver al-

gorithm as compared to the other block-based techniques. Energy detection is widely used in practice because it requires no *a priori* knowledge about the transmitted signal and it is simple to implement. A novel wideband spectrum sensing technique based on energy detection is introduced in [37], which jointly detects the PU signal over multiple frequency bands rather than one band at a time. This technique is further improved in [38] where the correlation between the sub-band occupancies is taken into account. Energy detection is also used in [39] for wideband spectrum sensing where the size of the observation interval is dynamically adjusted to maximize the overall throughput of the system. All the aforementioned block-based detectors assume that the PU activity status remains constant during the whole observation block; therefore, the detection accuracy is improved by increasing the number of observations. However, these detectors can fail in the low signal-to-noise ratio (SNR) region when there is modeling uncertainties, such as noise variance uncertainty, no matter how long the observation interval is. This phenomenon is referred to as SNR wall [40].

Sequential detectors take an alternative view to the problem, in which the number of observations is not determined in advance. The observations are made sequentially and the decision on the occupancy state of the channel is based on the current and all the previously made observations. In this case, the detector stops taking new observations as soon as it can declare the presence or absence of the PU signal with reasonable certainty. A merit of the sequential detection approach is that it requires, on the average, a smaller number of observations than equally reliable block-based detectors. Sequential detectors are generally realized either based on the classical sequential testing method developed by Walds [41], or the sequential change-point detection method [42, 43]. The objective in the classical sequential testing method is to distinguish between two hypotheses from a sequence of statistically *homogeneous* observations; that is, the observations only belong to one of the two hypotheses during the entire period of observation. Several detectors based on this method are proposed in the literature [44–48]. In [44], the classical sequential testing method is applied to cyclostationary feature detection to reduce the detection time. [47] studies the performance of a sequential energy detection technique and shows its efficiency gain over the fixed sample size energy detector. In [46], a joint PHY-MAC layer spectrum sensing technique exploiting the classical sequential testing method is proposed for fast detection of the PU emergence in cognitive radio networks. Contrarily to the classical sequential testing method, the change-point detection method assumes that the observations are not homo-

geneous and raises an alarm when an inhomogeneity occurs. This method is well suited in the context of CR networks, because the change in the PU signal activity status results into a change of the received signal distribution. The quickest detection technique, based on change-point detection method, minimizes the detection delay for a given probability of false alarm [49–52]. In [49,50] , the quickest detection technique is applied in a CR network to detect the appearance of the PU signal in a single spectrum band. [51] presents a fresh variation to the change-point detection method by considering the quickest detection of an idle period in *multiple* spectrum bands.

1.3 Thesis Objective and Contributions

The conventional blocked-base energy detector (ED) is widely used for spectrum sensing in CR applications since it only requires the energy of the received signal and it is easy to implement. This detector is based on the assumption that the PU activity status remains constant during the entire sensing period. This assumption could be valid for analog television broadcasting where a PU occupies and leaves the spectrum for a long time. However, it may not be realistic in digital television broadcasting or other practical cases where a burst mode transmission is used. It has been shown in [53] that the performance of the conventional ED deteriorates in a dynamic environment where the status of the PU changes during the observation interval. This limitation of the conventional ED has brought its credibility under question. The popularity of the ED in CR applications and its shortcoming in dynamic environments have urged the need for adaptive EDs that could reliably detect a change within the observation interval.

The authors in [53] have proposed an adaptive ED to improve the detection performance in the scenario where the PU appears during the sensing time. In their approach, a side detector is deployed to estimate the PU appearance time in order to discard the accumulated noise energy when the PU is absent and to calculate the test threshold that maintains a desired probability of false alarm. However, there is still room to improve this adaptive ED since the estimation error associated with the side detector is not taken into account. Furthermore, the equally important situation where the PU disappears within the sensing period is not considered.

The goal of this dissertation is to incorporate adaptive mechanisms into the existing ED in order to improve its change detection performance in the following scenarios:

- Scenario 1: PU appears during the sensing period
- Scenario 2: PU disappears during the sensing period.

In particular, we aim to achieve this goal by dynamically adjusting the test statistic and/or test threshold and avoiding expensive computations as much as possible.

Note that the quickest change detection techniques, which are based on the *sequential* detection method, deal with the change detection problem from a different perspective, where the number of observations are not fixed in advanced and the goal is to minimize the detection delay (the required number of observations to declare a change) subject to a certain probability of false alarm [49, 51, 54]. However, in this work, we are more concerned about reliable detection of a single change within a *fixed* observation interval.

The main contributions of this dissertation are listed below:

1. In Chapter 3, the performance deterioration of the conventional ED, under the condition that the PU activity status changes during the sensing period, is first proven through theoretical analysis in the two scenarios of interest.

2. In Chapter 4, the problem of a single change detection (i.e., PU appearance or disappearance) within a fixed observation interval is first handled by incorporating the exponential weighting approach into the energy detection technique. More specially, the received signal energies are weighted by an exponential window whose coefficients are dynamically adapted based on the estimated change location. The test threshold of the system is accordingly adjusted, based on the Neyman-Pearson (NP) criterion, to either maximize the probability of the PU detection subject to a constraint on the probability of false alarm in the appearing scenario or to maximize the probability of the spectrum hole detection subject to a constraint on probability of miss detection in the disappearing scenario. The analytical evaluation and simulation results prove the superiority of the proposed adaptive ED over the conventional ED and the only available adaptive ED in the literature [53] that addresses the problem of the PU appearance within the sensing period.

3. In Chapter 5, the problem of change detection in a fixed observation interval is formulated as a composite hypothesis testing problem. Two common approaches to the composite hypothesis testing problem are used to design two additional adaptive EDs.

The first approach is known as the generalized likelihood ratio test in the literature, which uses the maximum likelihood estimation (MLE) of the unknown change location in the likelihood ratio test. In addition, an iterative method is proposed to reduce the computational complexity of the MLE process. The second approach, referred to as composite-Bayesian approach in this dissertation, considers the change location as a discrete random variable with a known probability mass function (PMF). In the latter approach, the PU channel access pattern is modelled as a two-state Markov chain to obtain the PMF of the change location. The use of these two approaches results into two adaptive EDs, where the first one aims to satisfy the NP criterion while the second one aims to minimize the probability of error. The simulation results prove that the two proposed adaptive EDs outperform the conventional ED in both appearing and disappearing scenario.

The first two contributions have resulted into the following publication [55].

1.4 Thesis Outline

This thesis is organized as follows. Chapter 2 covers the background on block-based signal detection theory and reviews the most common blocked-based spectrum sensing techniques used in CR applications. The system model and the practical limitations of the conventional ED in dynamic environments are discussed in Chapter 3. The first adaptive ED based on an intuitive approach is proposed in Chapter 4. Chapter 5 covers the other two proposed adaptive EDs based on the composite hypothesis testing and finally, Chapter 6 concludes this thesis with some discussion and potential future research directions. Related derivations are provided in the Appendices. In this thesis, we use $\mathcal{CN}(\mu, \sigma^2)$ to denote the distribution of a circularly symmetric complex Gaussian random variable with mean μ and variance σ^2, and $\chi^2(v)$ to denote a central chi-square random variable with v degrees of freedom.

Chapter 2

Background

As discussed in the previous chapter, spectrum sensing techniques are based on two frameworks, namely block-based detection and sequential detection. This chapter covers the basic background on blocked-based signal detection theory and reviews the most common blocked-based spectrum sensing techniques used in CR applications. Furthermore, the optimality of the ED under certain conditions is proven.

2.1 Signal Detection in Noise

The detection of a signal in noise follows directly from the theory of binary hypothesis testing in which one needs to distinguish between the signal corrupted by noise and the noise signal alone, based on a vector of observations. The mathematical representations of the two hypotheses for N statistically homogeneous observations, $\mathbf{r} = [r(1), ..., r(N)] \in Z$, are given as follows:

$$\begin{aligned} \mathcal{H}_0 &: r(n) = w(n), \qquad && n = 1, ..., N \\ \mathcal{H}_1 &: r(n) = x(n) + w(n), \qquad && n = 1, ..., N \end{aligned} \tag{2.1}$$

where hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ denote the absence and presence of the signal, respectively. In (2.1), n represents the observation index, $x(n)$ is either a deterministic or a random signal to be detected, $w(n)$ is the additive noise which is often modelled as a white Gaussian random variable, and the domain Z of vector $\mathbf{r}$ is called the observation space. In this section, we assume $Z \subseteq \mathbb{R}^N$, although it is also possible to have $Z \subseteq \mathbb{C}^N$. Thus, the

vector $\mathbf{r}$ is a collection of random variables with a certain probability density function (PDF) under each hypothesis. A given observation vector can be thought of as a point in an N-dimensional space which is generated in accord with one of two possible conditional PDF $f(\mathbf{r}|\mathcal{H}_0)$ and $f(\mathbf{r}|\mathcal{H}_1)$ [1] . The detection problem can be categorized into two groups, namely *simple* and *composite* hypothesis testing problems. In the simple hypothesis testing problem, $f(\mathbf{r}|\mathcal{H}_0)$ and $f(\mathbf{r}|\mathcal{H}_1)$ are completely known. However, a more general problem arises when these conditional PDFs have some unknown parameters, which is referred to as the composite hypothesis testing problem.

In general, the objective is to use the conditional PDFs to develop a suitable decision rule, that is: to partition the total observation space Z into sets Z_0 and Z_1 (see Fig. 2.1), such that $\mathcal{H}_i$ is selected when $\mathbf{r} \in Z_i$ for $i = 0, 1$. There are several decision rules for any detection problem; however, we would like to employ decision rules that are optimal in some sense. The optimality definition varies from one approach to the other; hence the details of the most common approaches for the binary hypotheses testing are given in the subsequent parts of this section.

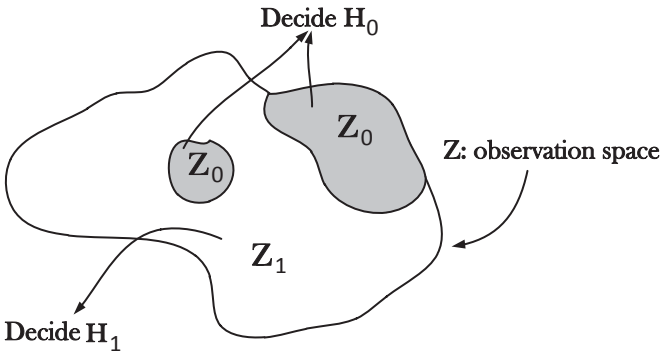


Fig. 2.1 Decision regions.

2.1.1 Simple Hypothesis Testing

The primary approaches to the simple hypothesis testing problem are the *Bayesian* approach, based on minimization of the Bayes risk, and the *Neyman-Pearson (NP)* approach,

[1] We have adopted the notations used in [56], where no explicit distinction is made between the random variables and their realizations. However, the readers should note the difference between the two from the context.

based on the NP criterion [56, Ch. 3]. The choice of an appropriate approach is dictated by the amount of information available to the detector.

Bayesian Approach

This approach is based on two assumptions. The first is that the occurrence of the two hypotheses is governed by *a priori* probabilities, which are denoted by $\pi_0 = Pr(\mathcal{H}_0)$ and $\pi_1 = Pr(\mathcal{H}_1)$, respectively. The second assumption is that there is a cost associated with our decisions. In particular, we use positive number C_{ij} for $i, j \in \{0, 1\}$, where C_{ij} denotes the cost incurred by deciding $\mathcal{H}_i$ when $\mathcal{H}_j$ is true; it is generally assumed that $C_{10} > C_{00}$ and $C_{01} > C_{11}$. The decision rule in this case aims to minimize the overall average cost or the Bayes risk, which is given as

$$\mathcal{R} = C_{00}\pi_0 Pr(\mathcal{H}_0|\mathcal{H}_0) + C_{10}\pi_0 Pr(\mathcal{H}_1|\mathcal{H}_0) + C_{01}\pi_1 Pr(\mathcal{H}_0|\mathcal{H}_1) + C_{11}\pi_1 Pr(\mathcal{H}_1|\mathcal{H}_1), \quad (2.2)$$

where $Pr(\mathcal{H}_i|\mathcal{H}_j)$ is the probability of deciding $\mathcal{H}_i$ given that $\mathcal{H}_j$ is true. As shown in Fig. 2.1, the observation space needs to be partitioned into two complementary sets, Z_0 and Z_1; therefore, we can rewrite the expression for the Bayes risk in terms of the unknown decision regions and the conditional probabilities of the observations as follows:

$$\begin{aligned}\mathcal{R} = C_{00}\pi_0 \int_{Z_0} f(\mathbf{r}|\mathcal{H}_0)d\mathbf{r} + C_{10}\pi_0 \int_{Z_1} f(\mathbf{r}|\mathcal{H}_0)d\mathbf{r} \\ + C_{01}\pi_1 \int_{Z_0} f(\mathbf{r}|\mathcal{H}_1)d\mathbf{r} + C_{11}\pi_1 \int_{Z_1} f(\mathbf{r}|\mathcal{H}_1)d\mathbf{r}. \end{aligned} \quad (2.3)$$

The integrals in (2.3) are N-dimensional integrals since the observation vector $\mathbf{r} \in Z \subseteq \mathbb{R}^N$. Using the fact that $Z = Z_0 \cup Z_1$ and $Z_0 \cap Z_1 = \emptyset$, (2.3) is rewritten as follows:

$$\begin{aligned}\mathcal{R} = C_{00}\pi_0 \int_{Z-Z_1} f(\mathbf{r}|\mathcal{H}_0)d\mathbf{r} + C_{10}\pi_0 \int_{Z_1} f(\mathbf{r}|\mathcal{H}_0)d\mathbf{r} \\ + C_{01}\pi_1 \int_{Z-Z_1} f(\mathbf{r}|\mathcal{H}_1)d\mathbf{r} + C_{11}\pi_1 \int_{Z_1} f(\mathbf{r}|\mathcal{H}_1)d\mathbf{r}. \end{aligned} \quad (2.4)$$

Observing that

$$\int_Z f(\mathbf{r}|\mathcal{H}_0)d\mathbf{r} = \int_Z f(\mathbf{r}|\mathcal{H}_1)d\mathbf{r} = 1, \quad (2.5)$$

(2.4) reduces to

$$\mathcal{R} = C_{00}\pi_0 + C_{01}\pi_1 + \int_{Z_1} [(C_{10} - C_{00})\pi_0 f(\mathbf{r}|\mathcal{H}_0) + (C_{11} - C_{01})\pi_1 f(\mathbf{r}|\mathcal{H}_1)]d\mathbf{r}. \tag{2.6}$$

It is clear that the Bayes risk consists of a fixed cost and a cost that depends on the choice of the set Z_1. Therefore, to minimize the Bayes risk, all values of $\mathbf{r}$ which contribute a negative amount to the integral should be included in Z_1. In other words, $\mathcal{H}_1$ is selected if

$$(C_{10} - C_{00})\pi_0 f(\mathbf{r}|\mathcal{H}_0) < (C_{01} - C_{11})\pi_1 f(\mathbf{r}|\mathcal{H}_1). \tag{2.7}$$

Finally, the decision rule is given by

$$L(\mathbf{r}) = \frac{f(\mathbf{r}|\mathcal{H}_1)}{f(\mathbf{r}|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{(C_{10} - C_{00})\pi_0}{(C_{01} - C_{11})\pi_1} = \eta, \tag{2.8}$$

where $L(\mathbf{r})$ is the *likelihood ratio* (LR) and η is the LR test threshold. Note that if $C_{00} = C_{11} = 0$ and $C_{01} = C_{10} = 1$, the Bayes risk in (2.2) corresponds to the average probability of error; in this case, (2.8) is the minimum-probability-of-error decision scheme.

Neyman-Pearson Approach

In the Bayesian approach, the optimality of the decision rule was defined in terms of minimization of the overall expected cost or the Bayes risk. However, in many practical situations, it is difficult to define a realistic cost structure for decisions or *a priori* probabilities for the hypotheses. In such cases, an alternative approach based on the NP criterion can be used, as explained below.

In a binary hypothesis testing, an error occurs if either $\mathcal{H}_1$ or $\mathcal{H}_0$ is rejected falsely. The probability of occurrence of these two errors can be written in terms of the two complementary sets of the observation space, Z_0 and Z_1, as

$$P_f = \int_{Z_1} f(\mathbf{r}|\mathcal{H}_0)d\mathbf{r}, \tag{2.9}$$

$$P_m = \int_{Z_0} f(\mathbf{r}|\mathcal{H}_1)d\mathbf{r}, \tag{2.10}$$

where P_f is the probability of false alarm (i.e., the probability of deciding on $\mathcal{H}_1$ given that

$\mathcal{H}_0$ is true) and P_m is the probability of miss detection (i.e., the probability of accepting $\mathcal{H}_0$ while $\mathcal{H}_1$ is true). There is an inherent trade-off between the two probabilities of error, because one can always be made arbitrary small at expense of the other. Hence, the NP criterion places a constraint on the probability of false alarm, $P_f = \alpha$, and minimizes P_m (or maximizes the probability of signal detection, $P_d = 1 - P_m$). The method of Lagrange multiplier [57, pp. 745-746] can be used to solve this constrained optimization problem as follows:

$$\begin{aligned} F = P_d + \lambda(P_f - \alpha) &= \int_{Z_1} f(\mathbf{r}|\mathcal{H}_1)d\mathbf{r} + \lambda\left(\int_{Z_1} f(\mathbf{r}|\mathcal{H}_0)d\mathbf{r} - \alpha\right) \\ &= \int_{Z_1} (f(\mathbf{r}|\mathcal{H}_1) + \lambda f(\mathbf{r}|\mathcal{H}_0))d\mathbf{r} - \lambda\alpha. \end{aligned} \tag{2.11}$$

In order to maximize F, the values of $\mathbf{r}$ that contribute a positive amount to the above integral should be included in Z_1, which is given mathematically by

$$Z_1 = \{\mathbf{r} \in Z | f(\mathbf{r}|\mathcal{H}_1) + \lambda f(\mathbf{r}|\mathcal{H}_0) > 0\}. \tag{2.12}$$

Note that the $>$ sign can be replaced with $\geq$, because the probability of occurrence of the event $f(\mathbf{r}|\mathcal{H}_1) + \lambda f(\mathbf{r}|\mathcal{H}_0) = 0$ is zero if we assume the PDFs are continuous. Thus, $\mathcal{H}_1$ is selected if

$$L(\mathbf{r}) = \frac{f(\mathbf{r}|\mathcal{H}_1)}{f(\mathbf{r}|\mathcal{H}_0)} \geq -\lambda, \tag{2.13}$$

where the Lagrangian multiplier satisfies $\lambda \leq 0$, since $L(\mathbf{r})$ is always non-negative. Finally, the decision rule based on the NP criterion is given for $\eta = -\lambda$ as

$$L(\mathbf{r}) = \frac{f(\mathbf{r}|\mathcal{H}_1)}{f(\mathbf{r}|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{2.14}$$

where $\eta > 0$ is the LR test threshold that is obtained from the constraint $P_f = \alpha$ as follows. Let $f(L|\mathcal{H}_0)$ denote the PDF of $L(\mathbf{r})$ under $\mathcal{H}_0$, then η is found by solving the following equation

$$P_f = \int_{\eta}^{\infty} f(L|\mathcal{H}_0)dL = \alpha. \tag{2.15}$$

It is apparent from (2.14) that decreasing η results into an increase in the region Z_1 where $\mathcal{H}_1$ is selected. This is equivalent to an increase in P_d. Thus, η in (2.15) is decreased in

order to maximize P_d until the constraint $P_f = \alpha$ is reached. The value of η given by (2.15) will be non-negative because $f(L|\mathcal{H}_0)$ will be zero for negative value of η.

Two approaches to the binary hypothesis testing problem were introduced in this subsection. The optimum decision rule in both approaches consist of processing the received observation vector $\mathbf{r}$ to obtain the LR $L(\mathbf{r})$ and then comparing it to a specific threshold to make the final decision. The choice of the appropriate threshold ultimately defines the difference between the two approaches. In the Bayesian approach, the threshold is dependent on *a priori* information regarding the decision cost and the probabilities of occurrence of the two hypotheses; while, the threshold in the NP approach is selected considering a constraint on P_f . Here, the conditional PDFs of the observation vector, $f(\mathbf{r}|\mathcal{H}_0)$ and $f(\mathbf{r}|\mathcal{H}_1)$, are assumed to be completely known; however, a more general problem in which these PDFs have some unknown parameters is considered in the following subsection.

2.1.2 Composite Hypothesis Testing

In the simple hypothesis testing problems, it is assumed that a complete description of the PDF of the observations is available under the two hypotheses. However, this is not a realistic assumption in practical signal detection problems, because the statistical model of the received signal often includes some unknown parameters such as the signal amplitude, phase, frequency, or the additive noise variance. The problems of this type are referred to as composite hypothesis testing problems.

The first step to solve a composite hypothesis testing problem is to express the conditional PDFs in terms of the unknown parameters and construct the LR test. In some rare cases, the LR test can be manipulated such that the resulting test is independent of the unknown parameters. This type of test is called a *uniformly most powerful* (UMP) test and it achieves the best performance over any other tests. However, in many practical problems, a UMP test does not exist and one must resort to suboptimal tests. In such cases, there are two common approaches that are based on joint estimation and detection techniques. The first approach employs the Bayesian parameter estimation philosophy described in [58, Ch. 10], where the unknown parameter is considered as a random variable with a known *a priori* PDF. This is referred to as the composite-Bayesian approach in this dissertation to distinguish it from the Bayesian approach used in the simple hypothesis testing problem. The second approach employs the well-known joint estimation and detec-

tion method called generalized likelihood ratio test (GLRT). In this approach, the MLE of the unknown parameters is used to construct the LR test. GLRT is widely used in practice due to its ease of implementation and less restrictive assumptions [59–61].

To explain the details of these two approaches, we consider a binary hypothesis testing problem in which the PDF of the observation vector $\mathbf{r}$ depends on a set of unknown parameters $\mathcal{U}_0$ and $\mathcal{U}_1$ under each hypothesis as follows:

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{r} \sim f(\mathbf{r}|\mathbf{u}_0, \mathcal{H}_0), \quad \mathbf{u}_0 \in \mathcal{U}_0, \\ \mathcal{H}_1 &: \mathbf{r} \sim f(\mathbf{r}|\mathbf{u}_1, \mathcal{H}_1), \quad \mathbf{u}_1 \in \mathcal{U}_1. \end{aligned} \tag{2.16}$$

Composite-Bayesian Approach

In this approach, *a priori* PDFs, $f_0(\mathbf{u}_0)$ and $f_1(\mathbf{u}_1)$, are assigned to the unknown parameters $\mathbf{u}_0$ and $\mathbf{u}_1$, respectively. Therefore, the dependency of the PDF of the observation vector on the unknown parameters can be removed by obtaining the marginal PDF as follows:

$$\begin{aligned} f(\mathbf{r}|\mathcal{H}_0) &= \int f(\mathbf{r}|\mathbf{u}_0, \mathcal{H}_0) f_0(\mathbf{u}_0) d\mathbf{u}_0, \\ f(\mathbf{r}|\mathcal{H}_1) &= \int f(\mathbf{r}|\mathbf{u}_1, \mathcal{H}_1) f_1(\mathbf{u}_1) d\mathbf{u}_1, \end{aligned} \tag{2.17}$$

where the dimension of each integral is equal to the dimension of the unknown parameter vector under each hypothesis. So far, we have assumed that the unknown parameters are continuous random variables; however, they can also be discrete. In such cases, the joint PMFs, $p_0(\mathbf{u}_0)$ and $p_1(\mathbf{u}_1)$, are used to obtain the marginal PDF of the observation vector:

$$\begin{aligned} f(\mathbf{r}|\mathcal{H}_0) &= \sum f(\mathbf{r}|\mathbf{u}_0, \mathcal{H}_0) p_0(\mathbf{u}_0), \\ f(\mathbf{r}|\mathcal{H}_1) &= \sum f(\mathbf{r}|\mathbf{u}_1, \mathcal{H}_1) p_1(\mathbf{u}_1). \end{aligned} \tag{2.18}$$

Finally, the LR test is constructed to choose between the two hypotheses:

$$L(\mathbf{r}) = \frac{f(\mathbf{r}|\mathcal{H}_1)}{f(\mathbf{r}|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta. \tag{2.19}$$

At this point, $f(\mathbf{r}|\mathcal{H}_1)$ and $f(\mathbf{r}|\mathcal{H}_0)$ are completely defined; therefore, the approaches introduced in Section 2.1.1 are applicable. The composite-Bayesian approach requires multidi-

mensional integration or summation which is not usually possible in closed form. Furthermore, assigning appropriate *a priori* probabilities to the unknown parameters is difficult in practical situations. Hence, it is usually a less attractive approach for composite hypothesis testing problems.

Generalized Likelihood Ratio Test

The GLRT replaces the unknown parameters in (2.16) with their MLEs and decides between the two hypotheses as follows:.

$$L_G(\mathbf{r}) = \frac{f(\mathbf{r}|\hat{\mathbf{u}}_1, \mathcal{H}_1)}{f(\mathbf{r}|\hat{\mathbf{u}}_0, \mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{2.20}$$

where

$$\hat{\mathbf{u}}_j = \arg\max_{\mathbf{u}_j \in \mathcal{U}_j} f(\mathbf{r}|\mathbf{u}_j, \mathcal{H}_j), \quad j = 0, 1. \tag{2.21}$$

The simple hypothesis testing approaches introduced in Section 2.1.1 can now be used to select the appropriate threshold. The GLRT appears to perform well in general, even though there is no optimality associated with it. The GLRT is shown to be invariant under any transformation which leaves the detection problem invariant [62]. Furthermore, it is asymptotically (as $N \to \infty$) optimal in the sense that it outperforms all the other tests that are invariant [63]. This is due to the fact that under mild conditions, the MLE converges almost surely to the true parameter value as the number of observations N becomes large [64, p. 113].

Two approaches to composite hypothesis testing problems were introduced in this section. The composite-Bayesian approach requires the statistical knowledge of the unknown parameters to construct the LR test, while the GLRT uses the MLE of the unknown parameters in the LR test. Although the GLRT is more attractive due to its less restrictive requirements, the MLE of the unknown parameters might not always be easy to evaluate. Therefore, the selection of the appropriate approach is problem oriented.

2.2 PU Signal Detection in CR Applications

In the previous section, detection of a signal in additive white Gaussian noise (AWGN) was discussed from a theoretical point of view, which is essential to understand the subsequent

parts of this dissertation. However, in this section, we focus more on the PU signal detection techniques used in CR applications. First, we start by discussing the CR structure and its signal reception and transmission process. Then, we review the most common blocked-based spectrum sensing techniques.

2.2.1 CR Structure

As mention in Section 1.1, CR is a SDR with additional functionalities to sense and share the spectrum. The basic components of the CR, based on the *direct conversion SDR* [65, p. 137], are shown in Fig. 2.2. In this structure, the received RF signals are first passed through a low-noise amplifier (LNA) and directly down converted to baseband by a mixer. The baseband signals are then low-pass filtered, and gain controlled before they are digitally sampled by an analog-to-digital converter (ADC). The automatic gain controller (AGC) is an adaptive device responsible for keeping the signals level in the dynamic range of the ADC. The analog low-pass filter (LPF) is an anti-aliasing filter which passes a broad frequency range and the desired band within that range is selected by a narrow-band digital filter, as shown conceptually in Fig. 2.3. This is advantageous for CRs operating in multiple frequency bands since they can tune the digital filter in software to cover several bands instead of using a set of dedicated narrow-band analog filters. The digital signal processor (DSP) runs spectrum sensing algorithms on the digital filter's output to decide on the occupancy state of the spectrum. The higher layer functions such as spectrum management and mobility are usually realized by embedded processor programs; whereas, the micro controllers handles multi-media and user interface applications [66, p. 43]. Once a spectrum hole is detected, the CR switches into the transmission mode to opportunistically use the available spectrum band as follows. The digital signals are filtered, passed through a digital-to-analog converter (DAC) which is followed by a reconstructive low-pass filter (anti-imaging filter), up-converted into RF band, amplified by a high power amplifier (HPA), and transmitted through the antenna. Note that simultaneous transmission and reception is only possible if the CR is equipped with more than one antenna.

2.2.2 Block-Based Spectrum Sensing Techniques

Spectrum sensing has a critical role in operation of the CR since it is responsible for reliable detection of the spectrum hole and the PU emergence. A wealth of the literature on

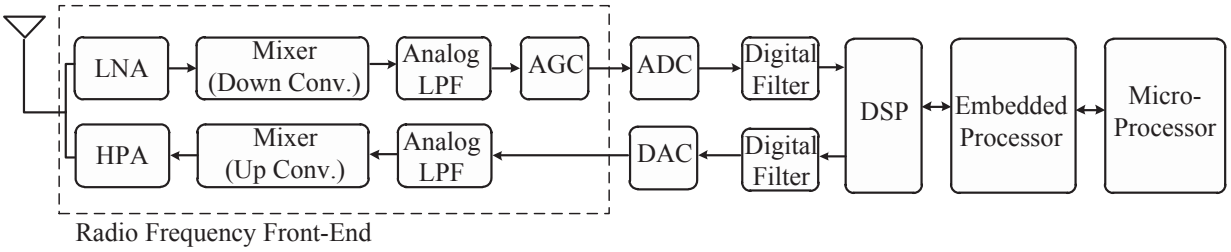


Fig. 2.2 CR transceiver based on the direct conversion SDR.

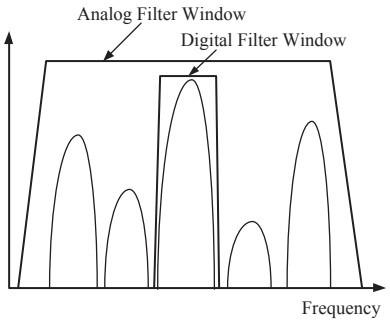


Fig. 2.3 Selection of the desired frequency band by a narrow-band digital filter within a broadband analog filter [65, Fig. 6.5].

spectrum sensing focuses on primary transmitter detection based on a *fixed* number of local measurements of the SU [27–33,36–39], which is referred to as *block-based* spectrum sensing in this dissertation to distinguish it from the *sequential* approach. The three common block-based spectrum sensing techniques are: matched filter detection, cyclostationary feature detection, and energy detection. The selection of the appropriate technique is dictated by *a priori* knowledge available to the SU and the resulting receiver complexity and accuracy. Here, we provide an overview on these techniques and prove the optimality of the ED under certain conditions based on the detection theory introduced in Section 2.1.1. We base our discussion on the simplest CR network, where a single CR is trying to share the licensed spectrum with a single PU. The CR uses a block-based spectrum sensing technique to decide between the two hypotheses given in (2.1), where $r(n)$ represents the complex baseband received sample at the output of the digital filter in Fig. 2.2. We consider the case of an ideal flat fading channel; accordingly, in (2.1), $x(n)$ is equal to $hs(n)$, where $s(n)$ is the PU signal transmitted through an ideal channel between the PU and the SU with a complex gain of h. For simplicity, we assume that h is completely known to the SU.

If the channel is not ideal (i.e., frequency selective), h and $s(n)$ are convolved instead of multiplied [66, p. 184].

Matched Filtering based sensing

If the CR has *a priori* knowledge of the PU signal $s(n)$, then the optimal detector is the well-known matched filter [67, 68]. A matched filter is defined by the convolution of the received signal $r(n)$ with the time-reversed version of the known signal $x(n)$, which is equivalent to correlation of $x(n)$ and $r(n)$. Therefore, the binary decision is given by:

$$\sum_{n=1}^{N} r(n)x^*(n) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \tag{2.22}$$

where γ is the user defined threshold and $*$ denotes the complex conjugate. In wireless communication theory, the matched filter is well understood to be the optimal linear filter to maximize the SNR in the presence of AWGN [69, pp. 238-243]. However, the theory introduced in Section 2.1.1 could also be applied on the received observation vector $\mathbf{r}$ to prove that the binary test in (2.22) is optimal in the NP or the Bayesian sense. The matched filter requires a fewer number of observations, in comparison with other block-based techniques, to achieve a certain performance criterion that could be defined in terms of the probabilities of false alarm and miss detection [70]. However, the required number of observations grow as the received SNR decreases [40]. Furthermore, the computational complexity can become very large if different PU signals are to be detected because dedicated matched-filters are required for all possible types of the PU signals. In addition, the matched filter requires the perfect knowledge of the PU signal, and any inaccuracies in this knowledge could result into performance deterioration. Therefore, the use of this technique is severely limited in CR applications since the complete knowledge of the PU signals is rarely available to the CR [71, p. 119].

Cyclostationarity-based sensing

The cyclostationary feature detector exploits the *inherent* periodicity of the modulated PU signal such as sine wave carriers, pulse trains, hoping sequences, or cyclic prefixes to differentiate the PU signal from the wide-sense stationary (WSS) noise [28–30]. There are also some cases, where periodicity are intentionally *induced* on the PU signal to assist

the spectrum sensing [72]. If the PU signal is present, the received signal samples exhibit correlation between wide spread spectral components due to spectral redundancy caused by the PU signal periodicity [73, p. 362]; whereas, such a cyclostationary feature is missing in the noise only signal. Thus, the cyclostationary feature detector aims to detect the spectral correlation in the received signal to decide on the presence or absence of the PU signal. The details of this procedure is summarized as follows. First, the cyclic autocorrelation function (CAF) of the received signal $r(n)$, denoted by $R_r^\alpha(\tau)$, is obtained as

$$R_r^\alpha(\tau) = E[r(n+\tau)r^*(n-\tau)e^{j2\pi\alpha n}], \quad \tau \in \mathbb{Z} \tag{2.23}$$

where $E[.]$ is the expectation operation and α is called the *cyclic frequency*. In the next step, the discrete Fourier transform (DFT) of the CAF is computed to obtain the spectral correlation function (SCF), $S(\alpha, f)$, as

$$S(\alpha, f) = \sum_{\tau=-\infty}^{\infty} R_r^\alpha(\tau)e^{-j2\pi f\tau}, \tag{2.24}$$

which is a two-dimensional function in terms of frequency f and cyclic frequency α. Under $\mathcal{H}_1$, the SCF, also known as cyclic spectral density, has peaks when α equals a multiple of the fundamental frequency of the transmitted signal $s(n)$, i.e., $\alpha = i/T_0$ where T_0 is the fundamental period of $s(n)$ and $i \in \mathbb{Z}^+$. Note that the fundamental frequency is either assumed to be known [74] or additional strategies are required to extract this information [75]. Under $\mathcal{H}_0$, the SCF has no peak since the WSS noise is not a cyclostationary signal. Finally, the GLRT [76], or a peak detector [77] can be used to choose between the two hypotheses. This technique can robustly detect a PU signal in low SNR regions as opposed to the other blocked-based detectors. Furthermore, it can distinguish among different types of the PU signals. However, its high computational complexity limits its widespread usage.

Energy Detector based sensing

Energy detection is the most commonly used blocked-based spectrum sensing technique because it is easy to implement and requires no *a priori* knowledge about the PU signal [37, 38, 55]. It can be implemented both in time domain or frequency domain. In this section, we cover the time domain formulation and we defer the alternative formulation

until the next chapter. The binary test in this case is given by summing the energy of N received samples:

$$\sum_{n=1}^{N} |r(n)|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \tag{2.25}$$

where γ is the pre-determined threshold parameter. The drawback associated with the ED is that it can only detect the PU signal if the energy of the received signal is over a threshold. Therefore, the selection of the appropriate test threshold is problematic as it is susceptible to the received SNR. Furthermore, the ED cannot distinguish the PU signal from the other signal sources in the network, which can result into false alarms and missing potential transmission opportunities. However, due to its low implementation cost, simplicity, and less restrictive requirements (no *a priori* knowledge is required about the PU signal), the ED is most commonly used in practice. In addition, as proven in the following section, the ED is the optimal decision rule under certain conditions.

2.3 Optimality of the ED

In the previous section, we discussed three common block-based spectrum sensing techniques. If the PU signal is completely known to the CR, the matched filter is the optimal sensing technique. However, if the PU signal is unknown, the ED is mostly preferred. The unknown PU signal could be either deterministic or modelled as a random process depending on the application. Here, we will prove that the ED is optimal if the unknown PU signal $s(n)$ (or equivalently $x(n)$, since the channel gain h is assumed to be known) is modelled as zero-mean, white Gaussian process.

Let $x(n)$ and $w(n)$ be two independent, zero mean, white circularly symmetric complex Gaussian (CSCG) random processes with variances σ_x^2 and σ_w^2, respectively. Based on this model, the statistically homogeneous received signal samples $r(n)$ in (2.1) are considered to be independent and identically distributed (i.i.d) as $\mathcal{CN}(0, \sigma_w^2)$ under $\mathcal{H}_0$ and $\mathcal{CN}(0, \sigma_x^2+\sigma_w^2)$ under $\mathcal{H}_1$. In a simple hypothesis testing problem, the optimal decision rules by the NP or the Bayesian approach consist of constructing the LR and comparing it against a threshold as follows:

$$L(\mathbf{r}) = \frac{f(\mathbf{r}|\mathcal{H}_1)}{f(\mathbf{r}|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\lessgtr}} \eta. \tag{2.26}$$

Since the received signal samples are i.i.d, the conditional probabilities of the observation

vector can be obtained as the product of the marginal density function of the individual samples. As the result, we have

$$L(\mathbf{r}) = \frac{\frac{1}{[\pi(\sigma_x^2+\sigma_w^2)]^N}\exp\left[-\frac{1}{(\sigma_x^2+\sigma_w^2)}\sum_{n=1}^{N}|r(n)|^2\right]}{\frac{1}{(\pi\sigma_w^2)^N}\exp\left[-\frac{1}{(\sigma_w^2)}\sum_{n=1}^{N}|r(n)|^2\right]}. \tag{2.27}$$

The log-likelihood ratio test becomes

$$l(\mathbf{r}) = N\ln(\frac{\sigma_w^2}{\sigma_x^2+\sigma_w^2}) + \frac{\sigma_x^2}{\sigma_w^2(\sigma_x^2+\sigma_w^2)}\sum_{n=1}^{N}|r(n)|^2. \tag{2.28}$$

Finally, the optimal decision rule can be manipulated into the form of energy detection as follows:

$$T(\mathbf{r}) = \sum_{n=1}^{N}|r(n)|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{\sigma_w^2(\sigma_x^2+\sigma_w^2)}{\sigma_x^2}\left[\ln(\eta) - N\ln(\frac{\sigma_w^2}{\sigma_x^2+\sigma_w^2})\right] = \gamma, \tag{2.29}$$

where γ is the ED's test threshold, selected to minimize the probability of error (i.e., P_m+P_f) or to maximize P_d for a given P_f. Hence, energy detection is the optimal detection technique in the simple hypothesis testing problem where the signal $x(n)$ is modelled as a zero mean Gaussian process. This justifies the use of ED in practical applications where such signal model is appropriate.

The performance of the ED can be analyzed in terms of the probabilities of false alarm and miss detection as defined in (2.9) and (2.10), once the distribution the test statistic $T(\mathbf{r})$ is known. It is observed in (2.29) that $T(\mathbf{r})$ is the summation of squared magnitude of i.i.d CSCG random variables; therefore, it has a generalized chi-squared distribution under the two hypothesis:

$$T(\mathbf{r}) \sim \begin{cases} \frac{\sigma_w^2}{2}\chi^2(2N), & \mathcal{H}_0 \\ \frac{(\sigma_w^2+\sigma_x^2)}{2}\chi^2(2N), & \mathcal{H}_1. \end{cases} \tag{2.30}$$

According to central limit theorem [78, p. 30], for a large N, $T(\mathbf{r})$ is approximately normally distributed with mean

$$E[T(\mathbf{r})] = \begin{cases} N\sigma_w^2, & \mathcal{H}_0 \\ N(\sigma_w^2+\sigma_x^2), & \mathcal{H}_1, \end{cases} \tag{2.31}$$

and variance

$$Var[T(\mathbf{r})] = \begin{cases} N\sigma_w^4, & \mathcal{H}_0 \\ N(\sigma_w^2 + \sigma_x^2)^2, & \mathcal{H}_1. \end{cases} \tag{2.32}$$

Using the decision rule in (2.29) and the definition of P_f and P_d, the latter can be approximately expressed as

$$P_f = Pr(T(\mathbf{r}) > \gamma | \mathcal{H}_0) = Q(\frac{\gamma - N\sigma_w^2}{\sigma_w^2\sqrt{N}}), \tag{2.33}$$

and

$$P_d = Pr(T(\mathbf{r}) > \gamma | \mathcal{H}_1) = Q(\frac{\gamma - N(\sigma_x^2 + \sigma_w^2)}{(\sigma_x^2 + \sigma_w^2)\sqrt{N}}), \tag{2.34}$$

where $Q(.)$ denotes the complementary cumulative distribution function of a standard normal distribution, defined in [79, p. 219] as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-y^2/2} dy. \tag{2.35}$$

In the context of CR networks, it is desirable to have P_f as low as possible to maximize the spectral efficiency, since P_f measures the percentage of the vacant spectrum that is declared busy by the detector. On the other hand, $P_m = 1 - P_d$ should be maintained at a very low level in order to avoid interference to the PU. The choice of γ is critical as it leads to a tradeoff between P_m and P_f. In particular, a smaller γ results into a smaller P_m, but a larger P_f. The test threshold that minimizes the probability of error (i.e., $P_e = P_m + P_f$) is obtained by the Bayesian approach as

$$\gamma = \frac{\sigma_w^2(\sigma_x^2 + \sigma_w^2)}{\sigma_x^2} \left[\ln(\frac{\pi_0}{\pi_1}) - N \ln(\frac{\sigma_w^2}{\sigma_x^2 + \sigma_w^2}) \right]. \tag{2.36}$$

In the case that the *a priori* probabilities π_0 and π_1 are not known, the NP criterion is used to find a test threshold that maximizes P_d for a fixed P_f as follows:

$$\gamma = Q^{-1}(P_f)\sigma_w^2\sqrt{N} + N\sigma_w^2. \tag{2.37}$$

2.4 Summary

In this chapter, we first covered the basic background for signal detection in AWGN for simple and composite hypothesis testing problems and we reviewed the most common blocked-based spectrum sensing techniques used in CR applications. Furthermore, we proved that the ED is the optimal decision rule for a zero-mean Gaussian signal in AWGN. This proof was based on the assumption that the received signal samples are statistically homogeneous, that is they belong to one of the two hypotheses during the entire period of observations. However, the optimality of the ED does not hold if the observations are non-homogeneous (heterogeneous). This issue is further investigated in the subsequent chapter.

Chapter 3

System Model and Problem Formulation

In this chapter, we first present the system model which incorporates energy detection as the core of its detection technique and then describe the practical limitations of the conventional ED in dynamic environments.

3.1 System Model

We consider an ad hoc CR network where the SUs can communicate with each other over a licensed spectrum band as shown in Fig. 3.1. The decentralized ad hoc CR network does not have an infrastructure component to control its operation; therefore, each SU is responsible to determine its action solely based on its own observations. The SUs are equipped with a single antenna transceiver which is used for both purpose of transmission and sensing. It is assumed that the SUs are aware of the bandwidth and the operating frequency range of the PU. A block-based detection scheme is considered where each SU tunes its digital filter into the operating frequency band of the PU and collects a fixed number of samples to decide on the occupancy state of the spectrum. Ideally, a SU should continue sensing the spectrum during its transmission in order to detect the PU emergence. However, the SUs are not able to perform the transmission and sensing operations simultaneously with their single antenna transceivers. Therefore, the SUs operate in a periodic fashion with a period of T as shown in Fig. 3.2. In this structure, each SU senses the spectrum for a fixed duration D and decides to either transmit or wait for the remaining duration $T - D$

until the next sensing cycle. $T - D$ defines the maximum time that a SU disregards the PU activity in the network and could potentially cause harmful interference to the PU transmission. It is further assumed that the transmission and sensing schedules of all SUs are synchronized and exits a medium access control protocol to share the spectrum holes among the SUs.

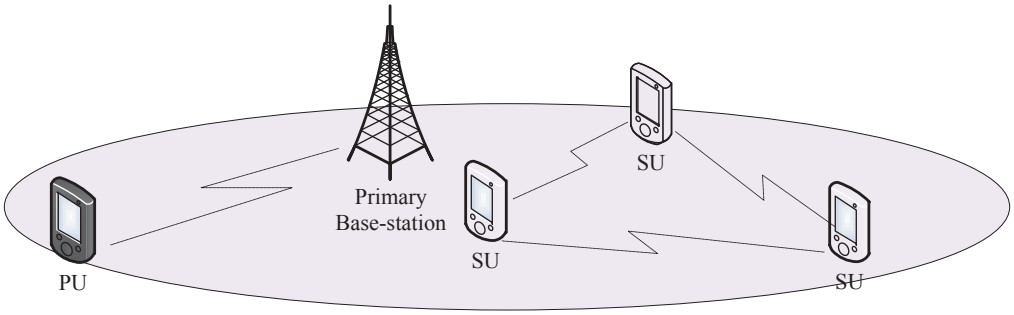


Fig. 3.1 Ad hoc CR network over the licensed band.

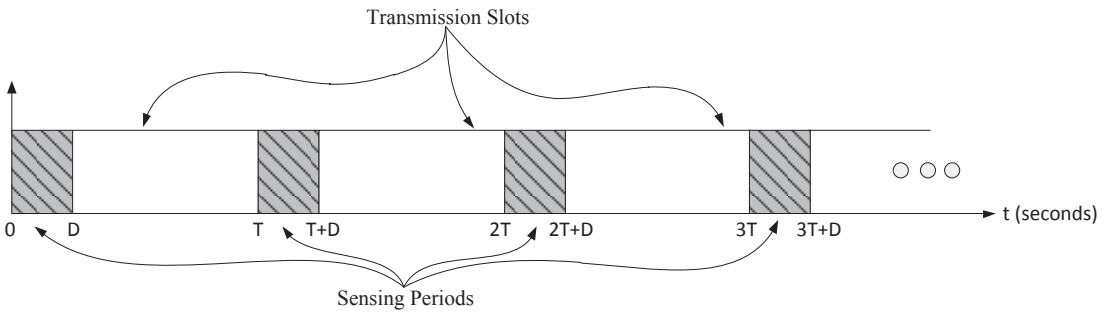


Fig. 3.2 Periodic sensing scheme [80].

Once the overall system model is known, the next step is to model the received signal by individual SUs and present the details of the appropriate block-based detection technique.

3.1.1 Received Signal

We let $r(n)$ denote the complex baseband received signal after down-conversion and sampling at the rate F_s. In the presence of the PU, this signal is represented as

$$r(n) = x(n) + w(n), \quad x(n) = \sum_{l=0}^{L-1} h(l)s(n-l), \tag{3.1}$$

where n denotes the sampling time index, $s(n)$ is the PU signal, $w(n)$ is the additive receiver noise, and $x(n)$ is the received PU signal at the output of the multipath fading channel between the PU and the SU as shown in Fig. 3.3. This wireless channel is modelled as a linear time-invariant system with finite impulse response $h(n)$ of length L, where L represents the number of resolvable paths. It is further assumed that $x(n)$ and $w(n)$ are two independent white CSCG random processes with zero means and variances σ_x^2 and σ_w^2, respectively. It is assumed here that these variances are known from *a priori* estimation.

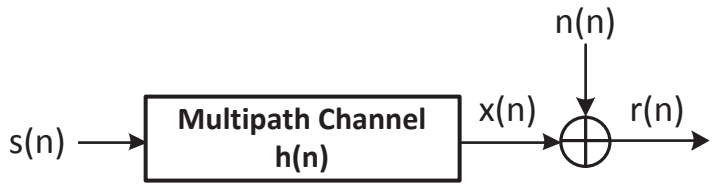


Fig. 3.3 The received PU signal at the output of the multipath channel and corrupted with noise.

3.1.2 PU Signal Detection

The received PU signal is unknown to the SU; therefore, according to the discussion in Section 2.2.2, the ED is the appropriate block-based detection technique. Furthermore, since we are considering the detection of zero mean white Gaussian signal in AWGN, the ED is optimum if the received signal samples are *homogeneous*. In this work, a frequency domain energy detection structure is considered as illustrated in Fig. 3.4. The received signal samples are divided into consecutive frames of K samples, and inputted to a K-point DFT to obtain the narrow-band frequency components. The latter are represented analytically by:

$$R_m(k) = \frac{1}{\sqrt{K}} \sum_{n=0}^{K-1} r(mK+n) e^{-j2\pi nk/K}, \tag{3.2}$$

where $k = 0, ..., K-1$ is the frequency index (or frequency bin), $m = 1, ..., M$ is the frame index, and M is the total number of frames being processed by the SU. Here, $M = \lfloor \frac{DF_s}{K} \rfloor$, where $\lfloor . \rfloor$ is the floor operator. Similarly, let $X_m(k)$ and $W_m(k)$ denote the kth DFT coefficient of the mth frame of the channel output and the additive noise, i.e. $x(n)$ and

$w(n)$ respectively. Then, it follows from (3.1) that

$$R_m(k) = X_m(k) + W_m(k). \tag{3.3}$$

Since $x(n)$ and $w(n)$ are assumed to be independent processes, the DFT coefficients $\{X_m(k)\}$ and $\{W_m(k)\}$ can be modelled as two independent random processes. The samples of the individual process are also independent across frequency and frame indices and follow a zero-mean CSCG distribution with variances σ_x^2 and σ_w^2, respectively. Note that the normalization factor in (3.2) is necessary to preserve the variances of $x(n)$ and $w(n)$ after discrete Fourier transformation.

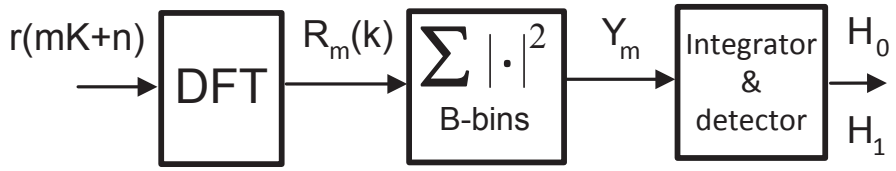


Fig. 3.4 Energy detection procedure (in practice, the DFT block is implemented by means of a fast Fourier transform (FFT)).

The energy of each frame is obtained by summing the squared magnitude of B frequency coefficients based on Parseval's theorem [81, p. 114], which for the mth frame takes the form

$$Y_m = \sum_{k \in \mathcal{B}} |R_m(k)|^2, \tag{3.4}$$

where $\mathcal{B}$ is a set of B frequency bins corresponding to the bandwidth of the PU. Finally, the frame energies Y_m, for $m = 1, ..., M$, are passed to a detection module where they are integrated and used to make a decision about the presence of the PU signal.

3.2 Problem Formulation

Conventional ED assumes that the PU activity status is constant during the entire sensing period of M frames, and a choice is made between two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, which represent the absence and presence of the PU signal, respectively:

$$\begin{aligned} \mathcal{H}_0 &: R_m(k) = W_m(k), \\ \mathcal{H}_1 &: R_m(k) = X_m(k) + W_m(k), \end{aligned} \tag{3.5}$$

where the given condition holds for all $m = 1, ..., M$ and $k \in \mathcal{B}$. Thus, under each hypothesis, the observed frequency samples $R_m(k)$ are i.i.d CSCG random variables, with $R_m(k) \sim \mathcal{CN}(0, \sigma_w^2)$ under $\mathcal{H}_0$ and $R_m(k) \sim \mathcal{CN}(0, \sigma_x^2 + \sigma_w^2)$ under $\mathcal{H}_1$.

For the above homogeneous observations, the optimal binary detector computes a test statistic, T_M, by summing the measured energies of M frames, and comparing it against a threshold γ to choose between $\mathcal{H}_0$ and $\mathcal{H}_1$:

$$T_M = \sum_{m=1}^{M} Y_m \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma. \tag{3.6}$$

Based on the received signal model, Y_m in (3.4) is the summation of squared magnitudes of i.i.d CSCG random variables and therefore, has a chi-square distribution under each hypothesis, with mean and variance given as follows [1]:

$$\mu_{Y|0} = B\sigma_w^2, \quad \mu_{Y|1} = B(\sigma_w^2 + \sigma_x^2), \tag{3.7a}$$

$$\sigma_{Y|0}^2 = B\sigma_w^4, \quad \sigma_{Y|1}^2 = B(\sigma_w^2 + \sigma_x^2)^2, \tag{3.7b}$$

where $|0$ and $|1$ denote conditioning on the hypothesis where the PU signal is absent and present during a *single* frame, respectively. From there, the distribution of T_M and the performance of the detector in (3.6) can be obtained.

Unfortunately, the above assumption that the PU activity status remains unchanged during the sensing period is not realistic in a dynamic environment where the PU could appear or disappear at any time. Two scenarios of particular interest in this work are illustrated in Fig. 3.5 for the case that the SU allocates M frames for sensing. In both scenarios, the PU activity status is constant until it is changed on the Jth frame and it is assumed that the PU keeps that status for a period longer than sensing time. Part a) corresponds to the appearance of the PU, while part b) corresponds to its disappearance.

The performance of the classical ED in (3.6) is now examined under the two aforementioned scenarios. The distribution of T_M should first be obtained for this purpose. Assuming that M or B are relatively large, the distribution of T_M can be approximated by the central limit theorem as Gaussian under both scenarios, but with different first and second moments. The details of the analysis are provided below for each scenario

[1] The distribution of Y_m and its mean and variance under each frame-hypothesis are provided in Appendix A

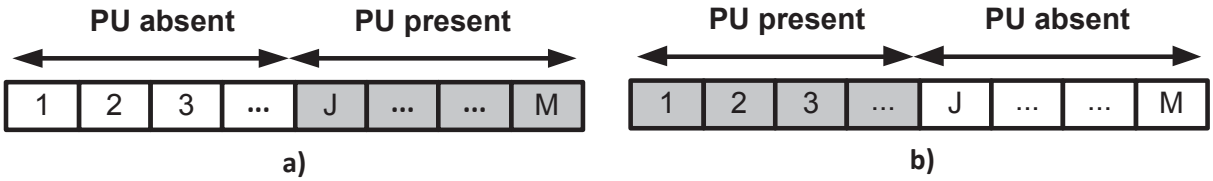


Fig. 3.5 Change of the PU activity status during the sensing period: a) PU appearance b) PU disappearance.

separately.

3.2.1 Appearing Scenario

Let $\mathcal{H}_0$ and $\mathcal{H}_1 \equiv \mathcal{H}_1(J)$ represent the absence of the PU signal (during the entire sensing period) and its appearance at the Jth frame, respectively. The mean and variance of T_M under these two hypotheses are obtained as follows [1]:

$$\begin{aligned}
\mu_{T|\mathcal{H}_0} &= M\mu_{Y|0}, \\
\mu_{T|\mathcal{H}_1} &= (J-1)\mu_{Y|0} + (M-J+1)\mu_{Y|1}, \qquad (3.8a)\\
\sigma^2_{T|\mathcal{H}_0} &= M\sigma^2_{Y|0}, \\
\sigma^2_{T|\mathcal{H}_1} &= (J-1)\sigma^2_{Y|0} + (M-J+1)\sigma^2_{Y|1}. \qquad (3.8b)
\end{aligned}$$

The probability of PU signal detection P_d, the probability of false alarm P_f, and the probability of the PU signal miss detection P_m are then obtained as

$$P_d = P(T_M > \gamma|\mathcal{H}_1) = Q(\frac{\gamma - \mu_{T|\mathcal{H}_1}}{\sigma_{T|\mathcal{H}_1}}), \qquad (3.9)$$

$$P_f = P(T_M > \gamma|\mathcal{H}_0) = Q(\frac{\gamma - \mu_{T|\mathcal{H}_0}}{\sigma_{T|\mathcal{H}_0}}), \qquad (3.10)$$

$$P_m = 1 - P_d, \qquad (3.11)$$

where the test threshold γ can be selected either based on the NP approach to maximize P_d for a given P_f, or based on the Bayesian approach to minimize the probability of error

[1] Derivation of the mean and the variance of T_M in the appearing scenario is provided in Appendix B.1

$P_e = P_f + P_m$. In the former approach, γ is given by

$$\gamma = Q^{-1}(P_f)\sigma_{T|\mathcal{H}_0} + \mu_{T|\mathcal{H}_0}, \tag{3.12}$$

while the latter approach requires *a priori* probabilities of occurrence of the two hypotheses (i.e., π_0 and π_1) to construct γ as

$$\gamma = \frac{\sigma_w^2(\sigma_x^2 + \sigma_w^2)}{\sigma_x^2}\left[\ln(\frac{\pi_0}{\pi_1}) - MB\ln(\frac{\sigma_w^2}{\sigma_x^2 + \sigma_w^2})\right]. \tag{3.13}$$

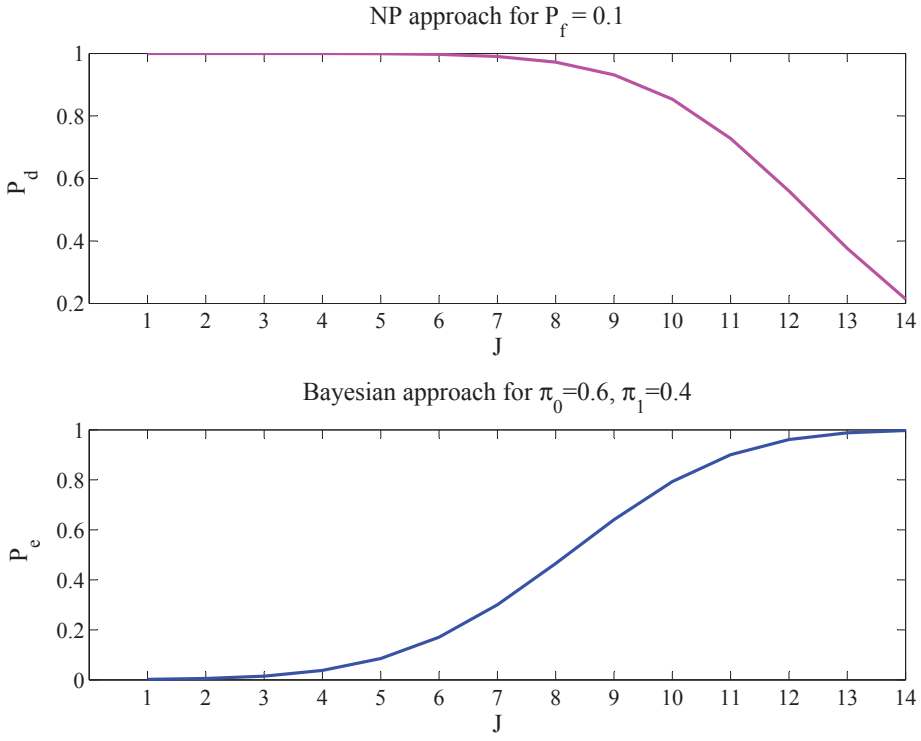


Fig. 3.6 Performance degradation of the conventional ED under the appearing scenario for both NP and Bayesian approach.

As explained in Section 2.3, the energy detection techniques with the above test thresholds are the optimal decision rules for homogeneous observations. However, the optimality does not hold for heterogeneous observations, which is the case in this scenario. More specially, the test statistic T_M under H_1 is corrupted by the accumulation of noise power

in the first $J-1$ frames, which causes the deterioration of P_d (or an increase in P_m). This fact has been confirmed in a specific environment where $M = 14$, $B = 128$, SNR $= -8dB$, $\sigma_w^2 = 1$, and J is varied from 1 to 14. In the NP approach, P_d in (3.9) is evaluated for a fixed $P_f = 0.1$. For the Bayesian approach, $\pi_0 = 0.6$ and $\pi_1 = 0.4$, and the probability of error P_e is computed based on P_f and P_m given in (3.10) and (3.11), respectively. The results of these analyses are shown in Fig. 3.6, where it is observed that P_d and P_e are degraded significantly as J increases, specially for $J > M/2$. It should be noted that the deterioration of P_e is caused mainly by the increase in P_m as J increases.

3.2.2 Disappearing Scenario

In the disappearing scenario, let $\mathcal{H}_1$ and $\mathcal{H}_0 \equiv \mathcal{H}_0(J)$ represent the presence of the PU signal (during the entire sensing period) and its disappearance at the Jth frame, respectively. The mean and variance of T_M under these two hypotheses are given as [2]

$$\begin{aligned} \mu_{T|\mathcal{H}_0} &= (J-1)\mu_{Y|1} + (M-J+1)\mu_{Y|0}, \\ \mu_{T|\mathcal{H}_1} &= M\mu_{Y|1}, \end{aligned} \tag{3.14a}$$

$$\begin{aligned} \sigma^2_{T|\mathcal{H}_0} &= (J-1)\sigma^2_{Y|1} + (M-J+1)\sigma^2_{Y|0}, \\ \sigma^2_{T|\mathcal{H}_1} &= M\sigma^2_{Y|1}. \end{aligned} \tag{3.14b}$$

In this case, the probability of spectrum hole detection P_h is of particular interest as opposed to P_d in the appearing scenario. Therefore, the NP criterion introduced in Section 2.1.1, is now modified to maximize P_h for a fixed P_d. For this approach, P_h can be obtained as

$$P_h = P(T_M < \gamma|\mathcal{H}_0) = 1 - Q(\frac{\gamma - \mu_{T|\mathcal{H}_0}}{\sigma_{T|\mathcal{H}_0}}), \tag{3.15}$$

where the test threshold γ corresponds to a desired P_d as follows:

$$\gamma = Q^{-1}(P_d)\sigma_{T|\mathcal{H}_1} + \mu_{T|\mathcal{H}_1}. \tag{3.16}$$

By contrast, the formulation of the Bayesian approach is not affected in this scenario, as it aims to minimize $P_e = P_f + P_m$, which is equivalent to maximization of $P_h + P_d$. Thus,

[2] Derivation of the mean and the variance of T_M in the disappearing scenario is provided in Appendix B.2

its test threshold is still equal to (3.13); moreover, P_f and P_m are obtained by replacing (3.14a)-(3.14b) into (3.10)-(3.11). However, the optimality of these two approaches do not hold in this scenario either, due to the heterogeneity of the observations. Particularly, the test statistic under $\mathcal{H}_0$ is corrupted by the accumulation of the PU signal energies in the first $J-1$ frames. Therefore, P_h is degraded (or P_f is increased), if $J > 1$. To verify this fact, the same conditions as in the appearing scenario are used, in which $M = 14$, $B = 128$, SNR $= -8$ dB, $\sigma_w^2 = 1$, and J is varied from 1 to 14. In the NP approach, P_h in (3.15) is evaluated for $P_d = 0.98$. In the Bayesian approach, P_e is evaluated for π_0 and π_1 set to 0.4 and 0.6, respectively. The results are illustrated in Fig. 3.7, where it is observed that P_h and P_e are both degraded as J increases. Note that, the deterioration of P_e is due to the increase in P_f as J increases.

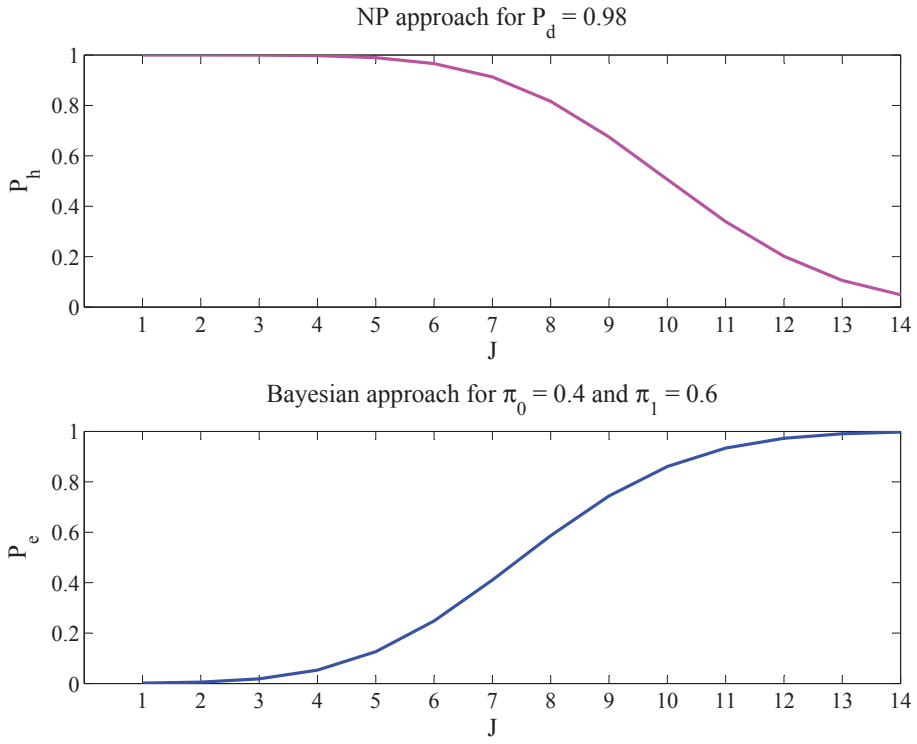


Fig. 3.7 Performance degradation of the conventional ED under the disappearing scenario for both NP and Bayesian approach.

3.2.3 Existing Solution in the Literature

The authors in [53] have proposed an adaptive ED to improve the detection performance in the appearing scenario. The adaptive ED uses a *side detector* to continuously monitor the spectrum and detect the PU emergence on a frame-by-frame basis. The side detector is also an ED which uses the measured energy of each frame, Y_m, in a binary hypothesis testing as follows:

$$Y_m \underset{0}{\overset{1}{\gtrless}} \gamma_s, \tag{3.17}$$

where the frame-hypothesis 0 and 1 denote the absence and presence of the PU during a *single* frame, respectively. In (3.17), γ_s is the side detector's test threshold which is given for a desired probability of false alarm within a *single* frame, P_{fs}, as

$$\gamma_s = Q^{-1}(P_{fs})\sigma_{Y|0} + \mu_{Y|0}, \tag{3.18}$$

where $\mu_{Y|0}$ and $\sigma^2_{Y|0}$ are the mean and variance of Y_m under the frame-hypothesis 0 that are provided in (3.7a)-(3.7b). In particular, the side detector outputs $\hat{J}$ as the index of the first frame whose energy is over the test threshold γ_s; otherwise, $\hat{J}$ is set to 1. The side detector's output, $\hat{J}$, is used to construct a binary test to decide on the occupancy state of the spectrum during M sensing frames as follows:

$$T_{M|\hat{J}} = \sum_{m=\hat{J}}^{M} Y_m \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \tag{3.19}$$

where γ is the test threshold given for a fixed P_f as

$$\gamma = Q^{-1}(P_f)\sigma_{T|\mathcal{H}_0} + \mu_{T|\mathcal{H}_0}. \tag{3.20}$$

Here, $\mu_{T|\mathcal{H}_0}$ and $\sigma^2_{T|\mathcal{H}_0}$ are the mean and variance of the test statistic $T_{M|\hat{J}}$ in (3.19) under $\mathcal{H}_0$, which are given as

$$\mu_{T|\mathcal{H}_0} = (M - \hat{J} + 1)\mu_{Y|0}, \tag{3.21}$$

$$\sigma^2_{T|\mathcal{H}_0} = (M - \hat{J} + 1)\sigma^2_{Y|0}. \tag{3.22}$$

The simulations in [53] prove that this technique outperforms the conventional ED in the appearing scenario.

The idea of this adaptive ED is to use the side detector to obtain a rough estimate of the unknown change location J, and to update the test statistic and the test threshold accordingly. However, this technique does not take into account the estimation error associated with the side detector because it blindly discards all the measured energies before $\hat{J}$. In the next chapter, we propose an adaptive ED which makes a better use of all the measured energies by weighting them exponentially.

3.3 Summary

In this chapter, we first introduced the system model where the energy detection technique was used for the protection of the PU against interference and the search of spectrum holes. Furthermore, we showed through theoretical analysis that the performance of the conventional ED deteriorates in dynamic environments where there is a single change during the sensing period. The subsequent chapters of this thesis cover the design of adaptive energy detection techniques that can overcome this practical limitation of the conventional ED.

Chapter 4

Adaptive Energy Detector Based on Exponential Weighting †

As discussed in the previous chapter, the performance of the conventional ED is degraded when there is a change in the PU activity status during the observation interval. The intuitive idea of exponential weighting of the observations has been previously used as a possible solution to the change detection problem arising in digital signal processing applications, which is commonly known as geometric moving average algorithm [82, 83]. Thus, in this chapter, this concept is incorporated into the conventional ED to improve its change detection performance. Furthermore, two adaptive mechanisms are included to make the design more compatible to the dynamic environments.

4.1 Proposed Adaptive ED

The structure of the proposed adaptive ED, based on exponential weighting, is shown in block diagram form in Fig. 4.1. We assume that the occurrence probability of the two hypotheses, i.e., $\pi_0 = Pr(\mathcal{H}_0)$ and $\pi_1 = Pr(\mathcal{H}_1)$, are not available; therefore, the NP formulation of the problem is considered throughout this chapter.

In this design, the side detector compares individual frame energies Y_m, as defined in (3.4), to a pre-set threshold γ_s, to obtain a rough estimate of the true (but unknown) frame index J where the PU activity status changes. This estimate, $\hat{J}$, is applied to the input of

†This chapter has been accepted for publication in part in the Proc. of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'11) [55].

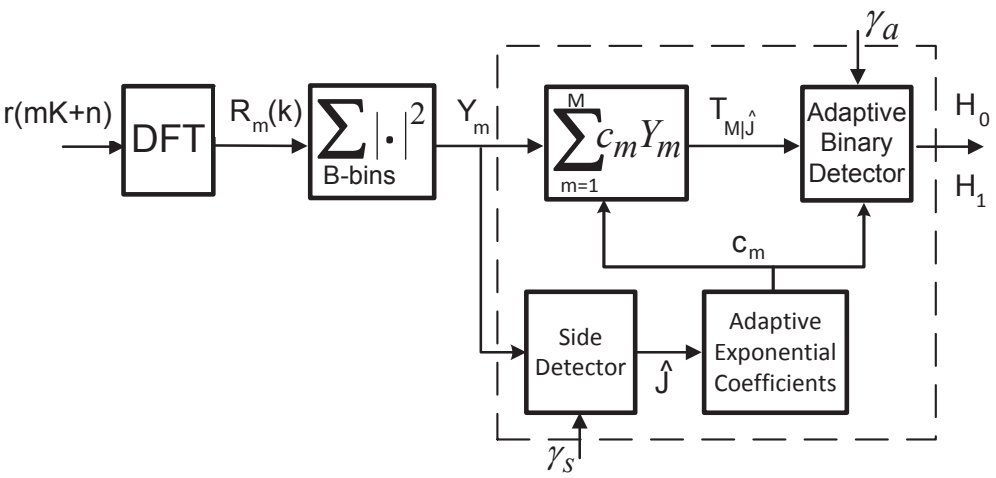


Fig. 4.1 Proposed adaptive ED procedure based on the exponential weighting approach.

an adaptive control mechanism which adjusts a set of exponential weighting coefficients, c_m for $m \in \{1, \ldots, M\}$, as follows:

$$c_m = \begin{cases} a^{(M-m)/(M-\hat{J})}, & \hat{J} < M \\ b^{M-m}, & \hat{J} = M, \end{cases} \tag{4.1}$$

where a and b are user-defined parameters limited to the range $(0, 1)$. In the upper branch of the diagram, the collected frame energies are weighted by the exponential coefficients, and added together to obtain a conditional test statistic as follows:

$$T_{M|\hat{J}} = \sum_{m=1}^{M} c_m Y_m. \tag{4.2}$$

The weighting coefficients c_m in (4.2) have the effect of reducing the accumulated energy in the first $J - 1$ frames, which causes the deterioration of the detection probability in dynamic environments. These coefficients are in turns used by the adaptive binary detector mechanism to calculate the test threshold γ_a that maintains a desired performance, as given by P_f in the appearing scenario and P_d in the disappearing scenario. The conditional test

statistics is then used to make the final decision via a binary test:

$$T_{M|\hat{J}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma_a. \tag{4.3}$$

The functionality of the proposed adaptive ED is further detailed below for the two scenarios of interest.

4.1.1 Appearing Scenario

In the appearing scenario, the side detector attempts to estimate the frame index, J, where the PU signal appears by applying ED on a frame-by-frame basis, using the threshold

$$\gamma_s = Q^{-1}(P_{fs})\sigma_{Y|0} + \mu_{Y|0}, \tag{4.4}$$

where $P_{fs} = P(Y_m > \gamma_s|0)$ is the desired probability of false alarm (for single frame processing) and $\mu_{Y|0}$ and $\sigma_{Y|0}$ are defined in (3.7a)-(3.7b). The rough estimate $\hat{J}$ is obtained as the first value of $m \in \{1, ..., M\}$, starting with $m = 1$, such that $Y_m > \gamma_s$; otherwise, if no such m can be observed, we set $\hat{J} = 1$. The estimate $\hat{J}$ is used to compute the exponential weights c_m as in (4.1).

The distribution of $T_{M|\hat{J}}$ under $\mathcal{H}_0$ is required to obtain the test threshold of the adaptive detector block. As explained in Section 3.2, Y_m has a chi-square distribution under $\mathcal{H}_0$ and $\mathcal{H}_1$ for $1 \leq m \leq M$; however, upon conditioning on $\hat{J}$, $Y_{\hat{J}}$ is no longer chi-square distributed under $\mathcal{H}_0$. To comprehend this, recall that the side detector declares the PU appearance at the $\hat{J}$th sensing frame, if $Y_{\hat{J}} > \gamma_s$. Thus, under such knowledge, $Y_{\hat{J}}$ is a variate that follows the portion of a centralized chi-square distribution that takes values only greater than γ_s. Therefore, $Y_{\hat{J}}$ is considered as a bias term under $\mathcal{H}_0$ and results into an increase in P_f of the overall system which is further investigated in Section 4.3. Nevertheless, for simplicity, we assume that $Y_{\hat{J}}$ is also chi-square distributed under $\mathcal{H}_0$. Therefore, the conditional test statistic in (4.2) is just a weighted sum of independent chi-squared random variables under both hypotheses. Because the exact distribution of such a weighted sum is difficult to obtain in general, various approximations have been proposed [84]. One relatively simple and widely used approximation is to model this sum as a scaled chi-square random variable [85]. Specifically, based on this approximation, the distribution of

$T_{M|\hat{J}}$ in (4.2) is given by $\zeta_M \frac{\sigma_R^2}{2}\chi^2(2MB)$, where $\zeta_M = \frac{1}{M}\sum_{m=1}^{M} c_m$, and σ_R is the variance of $R_m(k)$ under the given hypothesis. The conditional mean and variance of $T_{M|\hat{J}}$ under $\mathcal{H}_0$ are then given by [1]

$$\mu_{T|\mathcal{H}_0} = \zeta_M M \mu_{Y|0}, \tag{4.5a}$$

$$\sigma^2_{T|\mathcal{H}_0} = \zeta^2_M M \sigma^2_{Y|0}. \tag{4.5b}$$

The adaptive detector block is informed of the choice of $\{c_m\}$ and applies them into the above equations to update the test threshold, γ_a, to make the final decision. Assuming that M or B are relatively large, then $T_{M|\hat{J}}$ can be approximated as a Gaussian random variable [85]. Therefore, for a given P_f, the threshold γ_a is obtained as

$$\gamma_a = Q^{-1}(P_f)\sigma_{T|\mathcal{H}_0} + \mu_{T|\mathcal{H}_0}. \tag{4.6}$$

If $T_{M|\hat{J}}$ exceeds γ_a, then hypotheses $\mathcal{H}_1$ is selected; otherwise, $\mathcal{H}_0$ is declared as the current state of the channel.

The idea of deploying the side detector to estimate J for the appearing scenario, has already been used in [53]. However, in that work, the output of the side detector, $\hat{J}$, is used to eliminate the first $\hat{J}-1$ frames in order to improve the PU signal detection performance. In the proposed model, the side detector output, $\hat{J}$, is used instead to adjust the exponential weighting coefficients, c_m, in such a way that the accumulated energy in the first $\hat{J}-1$ frames are weighted less in comparison with the energy of the frames in the rest of the sensing period. This makes our proposed technique more robust to the estimation errors of the side detector, as it will be shown in section 4.3.

4.1.2 Disappearing Scenario

In the disappearing scenario, the threshold used by the side detector is

$$\gamma_s = Q^{-1}(P_{ds})\sigma_{Y|1} + \mu_{Y|1}, \tag{4.7}$$

[1] The variance of $T_{M|\hat{J}}$ can be directly obtained from (4.2) using the fact that the random variables $Y_m, m = 1, ..., M$ are independent. The resulting variance $\sum_{m=1}^{M} c_m^2 \sigma^2_{Y|0}$ differs from the variance obtained by the approximation given in [85]. However, we have observed through simulations that the variance in (4.5b) results into better performance.

where $P_{ds} = Pr(Y_m > \gamma_s|1)$ is the desired probability of PU signal detection for a single frame. The rough estimate $\hat{J}$ is obtained as the first value of m such that $Y_m < \gamma_s$; otherwise, we set $\hat{J} = 1$.

In this scenario, $Y_{\hat{J}}$ under $\mathcal{H}_1$ is no longer chi-square distributed due to similar reasons as explained in Section 4.1.1. Here, $Y_{\hat{J}}$ is the portion of a centralized chi-square variate that takes values only less than γ_s; thus, it represents bias term under $\mathcal{H}_1$ and causes a decrease in P_d. For simplicity, we assume that $Y_{\hat{J}}$ is a chi-square variate and approximate the distribution of $T_{M|\hat{J}}$ as a Gaussian using the approximation given in [85]. Thus, the mean and variance of $T_{M|\hat{J}}$ under $\mathcal{H}_1$ are given by

$$\mu_{T|\mathcal{H}_1} = \zeta_M M \mu_{Y|1}, \tag{4.8a}$$

$$\sigma^2_{T|\mathcal{H}_1} = \zeta^2_M M \sigma^2_{Y|1}. \tag{4.8b}$$

Using the above results, the adaptive binary detector adjusts the test threshold, γ_a, for a desired P_d as follows:

$$\gamma_a = Q^{-1}(P_d)\sigma_{T|\mathcal{H}_1} + \mu_{T|\mathcal{H}_1}, \tag{4.9}$$

and chooses between the two hypothesis using (4.3).

We finally note that the thresholds of the two adaptive mechanisms are different depending on the scenario of interest. A SU should employ the appearing scenario's thresholds for the case that it is already transmitting data and it is sensing the spectrum for the PU appearance. On the other hand, a SU should use the disappearing scenario's thresholds for the case that it is sensing for an unoccupied spectrum band.

4.2 Performance Analysis

The analytical performance of the proposed adaptive ED is investigated under the two scenarios of interest. This means that the parameter J is considered to be known and the probability of detection is calculated analytically for each scenario.

4.2.1 Appearing Scenario

As it was explained in Section 4.1.1, the operation of the adaptive binary detector is based on the calculated weighting coefficients c_m. These coefficients are dependent on $\hat{J}$; therefore,

using the statistics of $T_{M|\hat{J}}$, the probability of the PU signal detection *conditioned* on $\hat{J}$ is obtained as follows:

$$P_{d|\hat{J}} = Pr(T_{M|\hat{J}} > \gamma_a|\mathcal{H}_1) = Q(\frac{\gamma_a - \mu_{T|\mathcal{H}_1}}{\sigma_{T|\mathcal{H}_1}}), \tag{4.10}$$

where γ_a is the threshold of the adaptive binary detector block in (4.6). Here, $\mu_{T|\mathcal{H}_1}$ and $\sigma^2_{T|\mathcal{H}_1}$ represent the mean and the variance of $T_{M|\hat{J}}$ under $\mathcal{H}_1$ and are given by

$$\mu_{T|\mathcal{H}_1} = \zeta_0 A_0 \mu_{Y|0} + \zeta_1 A_1 \mu_{Y|1}, \tag{4.11a}$$
$$\sigma^2_{T|\mathcal{H}_1} = \zeta_0^2 A_0 \sigma^2_{Y|0} + \zeta_1^2 A_1 \sigma^2_{Y|1}, \tag{4.11b}$$

where $A_0 = J - 1$, $A_1 = M - J + 1$, $\zeta_0 = \frac{1}{A_0}\sum_{m=1}^{J-1} c_m$, and $\zeta_1 = \frac{1}{A_1}\sum_{m=J}^{M} c_m$. Note that ζ_0 and ζ_1 implicitly depend on $\hat{J}$ through the choice of $\{c_m\}$.

The probability of PU signal detection is obtained by averaging (4.10) as

$$P_d = \sum_{\hat{J}=1}^{M} P_{d|\hat{J}}\, p(\hat{J}|\mathcal{H}_1), \tag{4.12}$$

where $p(\hat{J}|\mathcal{H}_1)$ is the PMF of $\hat{J}$ under $\mathcal{H}_1$, obtained as follows. The side detector applies energy detection on every frame until the first frame energy that exceeds γ_s; therefore, this process is modelled as a Bernoulli distribution with probabilities of "success" (i.e., $Y_m > \gamma_s$) under the two frame-hypotheses as

$$p_0 = Pr(Y_m > \gamma_s|0) = Q(\frac{\gamma_s - \mu_{Y|0}}{\sigma_{Y|0}}), \tag{4.13}$$
$$p_1 = Pr(Y_m > \gamma_s|1) = Q(\frac{\gamma_s - \mu_{Y|1}}{\sigma_{Y|1}}). \tag{4.14}$$

Here, γ_s is the side detector's threshold given in (4.4). Consequently, the probability of a "failure" ($Y_m < \gamma_s$) under frame-hypothesis 0 and 1 are defined as $q_0 = 1 - p_0$ and $q_1 = 1 - p_1$, respectively. Therefore, $\hat{J}$ has a "generalized" geometric distribution with

PMF under $\mathcal{H}_1$ given as

$$p(\hat{J}|\mathcal{H}_1) = \begin{cases} p_0 + q_0^{J-1} q_1^{M-J+1}, & \hat{J} = 1 \\ q_0^{\hat{J}-1} p_0, & 1 < \hat{J} < J \\ q_0^{J-1} q_1^{\hat{J}-J} p_1, & J \leq \hat{J} \leq M. \end{cases} \tag{4.15}$$

Finally, the performance of the proposed adaptive ED under the appearing scenario can be evaluated analytically based on (4.12). The evaluation results will be presented in Section 4.3.

4.2.2 Disappearing Scenario

The same procedure as in Section 4.2.1 is followed here with the exception that now the probability of hole detection, P_h, is under consideration, which is given by

$$P_h = \sum_{\hat{J}=1}^{M} P_{h|\hat{J}} p(\hat{J}|\mathcal{H}_0). \tag{4.16}$$

The conditional probability of hole detection, $P_{h|\hat{J}}$, is obtained as follows:

$$P_{h|\hat{J}} = Pr(T_{M|\hat{J}} < \gamma_a|\mathcal{H}_0) = 1 - Q(\frac{\gamma_a - \mu_{T|\mathcal{H}_0}}{\sigma_{T|\mathcal{H}_0}}), \tag{4.17}$$

where γ_a is given in (4.9). $\mu_{T|\mathcal{H}_0}$ and $\sigma_{T|\mathcal{H}_0}$ are the mean and the variance of $T_{M|\hat{J}}$ under $\mathcal{H}_0$, which are given by

$$\mu_{T|\mathcal{H}_0} = \zeta_0 A_0 \mu_{Y|1} + \zeta_1 A_1 \mu_{Y|0}, \tag{4.18a}$$

$$\sigma^2_{T|\mathcal{H}_0} = \zeta_0^2 A_0 \sigma^2_{Y|1} + \zeta_1^2 A_1 \sigma^2_{Y|0}, \tag{4.18b}$$

where A_0, A_1, ζ_0, and ζ_1 are defined in Section 4.2.1. The PMF of $\hat{J}$ under $\mathcal{H}_0$, $p(\hat{J}|\mathcal{H}_0)$, is obtained using a similar procedure as in the appearing scenario. The probability that the side detector declares the PU signal disappearance is given under the two frame-hypotheses as follows:

$$p_0 = Pr(Y_m < \gamma_s|0) = 1 - Q(\frac{\gamma_s - \mu_{Y|0}}{\sigma_{Y|0}}), \tag{4.19}$$

$$p_1 = Pr(Y_m < \gamma_s|1) = 1 - Q(\frac{\gamma_s - \mu_{Y|1}}{\sigma_{Y|1}}), \tag{4.20}$$

where γ_s is the threshold of the side detector in (4.7). Therefore, the PMF of $\hat{J}$ under $\mathcal{H}_0$ is obtained as

$$p(\hat{J}|\mathcal{H}_0) = \begin{cases} p_1 + q_1^{J-1} q_0^{M-J+1}, & \hat{J} = 1 \\ q_1^{\hat{J}-1} p_1, & 1 < \hat{J} < J \\ q_1^{J-1} q_0^{\hat{J}-J} p_0, & J \leq \hat{J} \leq M. \end{cases} \tag{4.21}$$

Finally, (4.16) is used to evaluate the performance of the proposed adaptive ED analytically under the disappearing scenario.

4.3 Simulation Results

In this section, the performance of the proposed adaptive ED in the two scenarios of interest is compared with that of the conventional ED whose test threshold is provided in (3.12) and (3.16) for appearing and disappearing scenarios, respectively. A PU signal with a bandwidth of 200 kHz is assumed to be sampled at a frequency of 1600 kHz, and a 1024-point FFT is used to obtain the frequency representation of the signal. The number of frequency bins corresponding to the PU bandwidth is 128, i.e., $B = 128$. The received signal SNR is equal to -8dB, with desired $P_f = P_{fs} = 0.1$ and $P_d = P_{ds} = 0.98$ in the appearing and disappearing scenarios, respectively. The parameters in (4.1) are set to $a = 0.2$ and $b = 0.5$, based on preliminary experiments. In all the simulations, J is set to 9 and M is varied from 9 to 14. The simulation parameters are summarized in Table 4.1, where BW_{PU} represents the bandwidth of the PU signal and SNR $= 10\log(\frac{\sigma_x^2}{\sigma_w^2})$.

For both the conventional ED and the proposed adaptive ED, random data are generated based on the model introduced in Section 3.1 and the simulations are run for 1000 independent trials to obtain an estimation of P_d and P_h for the appearing and disappearing scenarios, respectively. The results of these experiments are illustrated with the solid lines for the two scenarios in Fig. 4.2 and 4.3. These figures also include the theoretical results

(dashed lines) which are based on the analytical formulae obtained in Sections 3.2 and 4.2. It is observed that the proposed adaptive ED achieves a better performance compared to the conventional ED in both scenarios. Furthermore, the experimental results closely follow the analytical results; however, there are some discrepancies between the two results for the proposed adaptive ED in the disappearing scenario, which is caused by the approximation made regarding the distribution of $T_{M|\hat{J}}$. The proposed adaptive ED is able to reduce the corrupting energies in the first $J-1$ frames and dynamically adapt the test threshold to improve its change detection capabilities. However, a trade-off is associated with the proposed technique. As explained in Section 4.1.1 and Section 4.1.2, the test statistic contains a bias term, which results into an increase in P_f and a decrease in P_d under the appearing and disappearing scenarios, respectively.

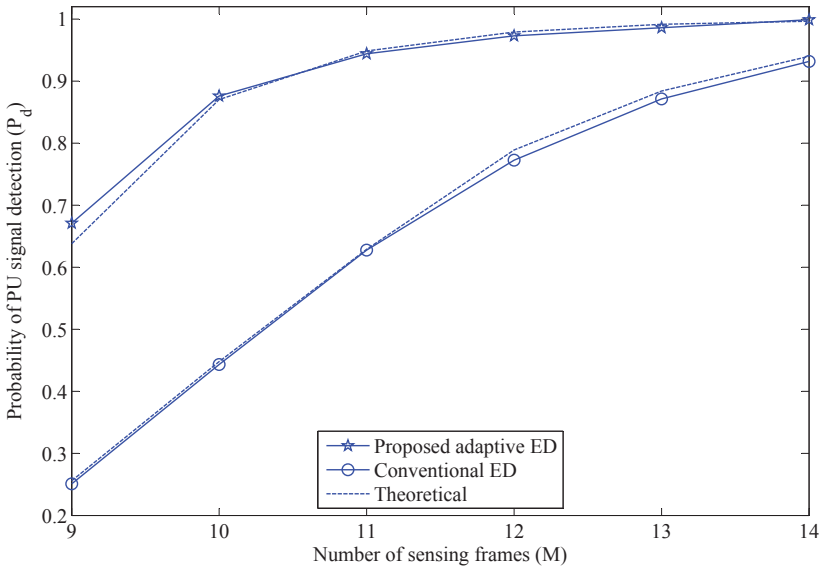


Fig. 4.2 Probability of PU signal detection for the proposed adaptive ED and conventional ED in the appearing scenario.

Table 4.1 Simulation parameters for appearing and disappearing scenarios

Appearing & Disappearing Scenario									Appearing Scenario	Disappearing Scenario
BW_{PU}	F_s	K	B	SNR	σ_w^2	J	a	b	P_f,P_{fs}	P_d,P_{ds}
200 KHz	1600 KHz	1024	128	-8 dB	1	9	0.2	0.5	0.1	0.98

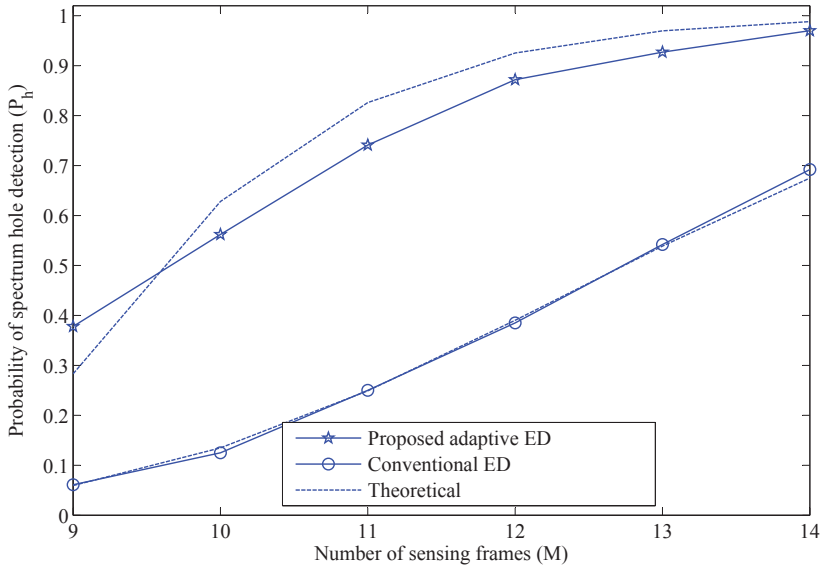


Fig. 4.3 Probability of spectrum hole detection for the proposed adaptive ED and conventional ED in the disappearing scenario.

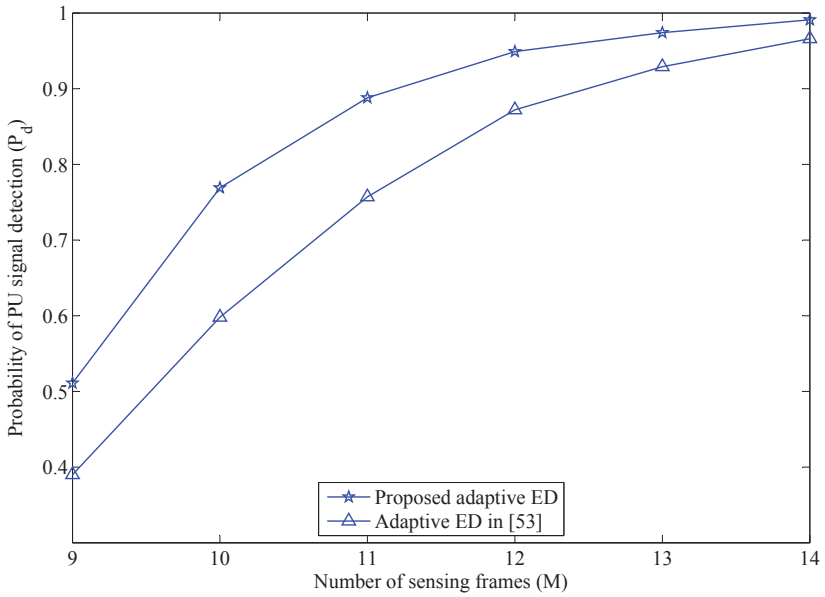


Fig. 4.4 Performance comparison of the proposed adaptive ED and the alternative adaptive ED in the appearing scenario.

As introduced in Section 3.2.3, the adaptive ED in [53] improves the detection probability in the appearing scenario; therefore, it is reasonable to compare its performance to that of the adaptive ED proposed in this chapter. These two adaptive EDs achieve a better performance, compared to the conventional ED, in exchange of an increase in P_f. Hence, the test thresholds obtained by both adaptive detectors need to be adjusted through simulations in order to maintain a desired P_f and thus a fair comparison. The result of such comparison is shown in Fig. 4.4 and it is observed that the proposed adaptive ED outperforms the alternative approach. This result was expected as the proposed adaptive ED considers the uncertainty in the side detector estimation, $\hat{J}$, and makes a better use of the measured energies by exponentially weighting them as opposed to the other technique which blindly disregards the frame energies before $\hat{J}$.

4.4 Summary

The performance of the conventional ED is degraded in a dynamic environment where the PU activity status changes during the sensing period. An adaptive ED was proposed to improve the probability of detection in such environments. The proposed technique applies an exponential weighting window over the measured energies of M frames and adjusts the weighting coefficients based on the estimated location of the frame where the PU activity status changes. Analytical performance evaluation and simulation results have proven the superiority of this technique over the conventional ED and the adaptive ED in [53]. In the next chapter, a more theoretical approach is employed to improve the detection performance furthermore.

Chapter 5

Adaptive Energy Detectors Based on the Composite Hypothesis Testing

In the previous chapter, we have incorporated the exponential weighting approach into the energy detection technique in order to improve its change detection performance. In this chapter, we aim to achieve the same goal from a more theoretical point of view. In the two dynamic environments of interest, the PU's activity status changes suddenly at the Jth frame during a fixed observation interval. Therefore, the PDF of the observation vector depends on the unknown parameter J. According to the theory exposed in Section 2.1.2, this type of problems can be formulated as a composite hypothesis testing. Hence, the goal of this chapter is to design detectors, based on the composite hypothesis testing, to overcome the practical limitations of the conventional ED in the two scenarios of interest.

Recall from Section 2.1.2 that the first step in the composite hypothesis testing problem is to look for the UMP test, that is, to express the PDFs of the observations under each hypothesis in terms of the unknown parameters, construct the LR test, and manipulate it such that the resultant test is independent of the unknown parameters. Here, we first prove that the UMP test does not exist for this particular composite hypothesis testing problem. Then, we deploy the two alternative suboptimal approaches which were introduced in Section 2.1.2, namely the GLRT and the composite-Bayesian approach. Both approaches have already been used in the literature to overcome the uncertainty in noise power, PU signal power, or/and channel gain [86–89]; however, to the best of the author's knowledge, they have never been considered in the change detection problems of interest. The use of

these two approaches results into two adaptive EDs which are able to improve the detection performance in the two dynamic environments introduced in Section 3.2.

5.1 UMP Test

In this section, we investigate the existence of the UMP test in both appearing and disappearing scenarios. We continue working with the frequency representation of the observations; hence, we let the vector $\mathbf{R}$ represent the set of DFT coefficients of the M sensing frames in the frequency range of interest (i.e., $\mathbf{R} = \{R_m(k); 1 \leq m \leq M, k \in \mathcal{B}\}$). Based on the received signal model, the enteries of $\mathbf{R}$ are i.i.d CSCG variates with zero mean and variances $\sigma^2_{R|0} = \sigma^2_w$ when the PU is absent and $\sigma^2_{R|1} = \sigma^2_x + \sigma^2_w$ in its presence. Therefore, the PDF of $\mathbf{R}$ can be obtained as the product of the marginal PDFs under the both hypotheses in the two scenarios of interest. Finally, the LR can be constructed to investigate the existence of the UMP test as explained in the followings.

5.1.1 Appearing Scenario

In this scenario, the PU appears at the Jth sensing frame; therefore, it is convenient to introduce the following hypotheses about the frequency representation of the observations, $R_m(k)$, for $1 \leq m \leq M$ and $k \in \mathcal{B}$ as follows:

$$\mathcal{H}_0 : \quad R_m(k) = W_m(k), \tag{5.1}$$

$$\mathcal{H}_1 : \quad R_m(k) = X_m(k)u(m - J) + W_m(k), \tag{5.2}$$

where J is the unknown frame index in the range of $[1, M]$ and $u(m)$ is the unit step function defined as

$$u(m) = \begin{cases} 0, & m < 0 \\ 1, & m \geq 0. \end{cases} \tag{5.3}$$

The PDF of $\mathbf{R}$ under the two hypotheses are then given by [1]

$$f(\mathbf{R}|\mathcal{H}_0) = (\frac{1}{\pi\sigma^2_{R|0}})^{BM} \exp\left[\frac{-\sum_{m=1}^{M} Y_m}{\sigma^2_{R|0}}\right], \tag{5.4}$$

$$f(\mathbf{R}|J,\mathcal{H}_1) = (\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}})^{B(J-1)} \exp\left[\frac{-\sum_{m=1}^{J-1} Y_m}{\sigma^2_{R|0}}\right] (\frac{1}{\pi\sigma^2_{R|1}})^{BM} \exp\left[\frac{-\sum_{m=J}^{M} Y_m}{\sigma^2_{R|1}}\right], \tag{5.5}$$

where we recall the definition of Y_m, i.e. : $Y_m = \sum_{k\in\mathcal{B}} |R_m(k)|^2$. Finally, the LR test is constructed and simplified as

$$L(\mathbf{R}) = \frac{f(\mathbf{R}|J,\mathcal{H}_1)}{f(\mathbf{R}|\mathcal{H}_0)} = \left(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}}\right)^{B(J-1-M)} \exp\left[\frac{\sigma^2_{R|1}-\sigma^2_{R|0}}{\sigma^2_{R|1}\sigma^2_{R|0}} \sum_{m=J}^{M} Y_m\right] \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{5.6}$$

which can be further manipulated into an energy detector of the form:

$$\sum_{m=J}^{M} Y_m \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{\sigma^2_{R|1}\sigma^2_{R|0}}{\sigma^2_{R|1}-\sigma^2_{R|0}} \left[\ln(\eta) - B(J-1-M)\ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}})\right]. \tag{5.7}$$

Clearly, both the test statistic and the test threshold are dependent on J; hence, the UMP test does not exist in this scenario.

5.1.2 Disappearing Scenario

The two hypotheses in this scenario are given for $1 \le m \le M$ and $k \in \mathcal{B}$ by

$$\mathcal{H}_0 : \quad R_m(k) = X_m(k)(1-u(m-J)) + W_m(k), \tag{5.8}$$

$$\mathcal{H}_1 : \quad R_m(k) = X_m(k) + W_m(k), \tag{5.9}$$

[1] The Derivation of the PDFs of $\mathbf{R}$ under the two hypotheses in the appearing scenario are provided in Appendix C.1

where J is any integer in the range $[1, M]$. The PDF of $\mathbf{R}$ under the above hypotheses are given by [2]

$$f(\mathbf{R}|J, \mathcal{H}_0) = (\frac{\sigma^2_{R|0}}{\sigma^2_{R|1}})^{B(J-1)} \exp\left[\frac{-\sum_{m=1}^{J-1} Y_m}{\sigma^2_{R|1}}\right] (\frac{1}{\pi\sigma^2_{R|0}})^{BM} \exp\left[\frac{-\sum_{m=J}^{M} Y_m}{\sigma^2_{R|0}}\right], \quad (5.10)$$

$$f(\mathbf{R}|\mathcal{H}_1) = (\frac{1}{\pi\sigma^2_{R|1}})^{BM} \exp\left[\frac{-\sum_{m=1}^{M} Y_m}{\sigma^2_{R|1}}\right]. \quad (5.11)$$

Once the PDFs are expressed as above, the LR test can be constructed and simplified into:

$$L(\mathbf{R}) = \frac{f(\mathbf{R}|\mathcal{H}_1)}{f(\mathbf{R}|J, \mathcal{H}_0)} = (\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}})^{B(J-1-M)} \exp\left[\frac{\sigma^2_{R|1} - \sigma^2_{R|0}}{\sigma^2_{R|1}\sigma^2_{R|0}} \sum_{m=J}^{M} Y_m\right] \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \quad (5.12)$$

which is equivalent to (5.6). Similar to the appearing scenario, the LR can be further manipulated into an ED that is dependent on J; therefore, the UMP test does not exist in this scenario either.

We have ensured that the UMP test does not exist in this particular problem; therefore, we can continue solving the change detection problem using the suboptimal approaches. More specially, the GLRT and the composite-Bayesian approaches are used to design two adaptive EDs that can achieve a significant performance improvement in the two dynamic environments in comparison with the conventional ED.

5.2 Generalized Likelihood Ratio Test Approach

In this approach, the unknown parameter J in (5.5) and (5.10) is replaced with its corresponding MLE, which is denoted by $\hat{J}$. Therefore, the LR in (5.6) and (5.12) can now be constructed in terms of $\hat{J}$ instead of the unknown change location J. The use of the MLE converts the composite hypothesis testing problem back into a simple binary hypothesis testing problem, where the previously mentioned approaches in Section 2.1.1 can be deployed. In this section, we assume that the detector has no *a priori* knowledge about the probabilities of occurrence of the two hypotheses; thus, the NP formulation of the problem is considered through out this section. The GLRT approach can be used to modify the

[2]The Derivation of the PDFs of $\mathbf{R}$ under the two hypotheses in the disappearing scenario are provided in Appendix C.2

previously proposed adaptive ED in order to further improve its performance in the two dynamic environments. The procedure of modified adaptive ED is shown in Fig. 5.1, where the side detector is now replaced by a MLE block. The output of the MLE block $\hat{J}$ is used as side information to construct the conditional test statistic:

$$T_{M|\hat{J}} = \sum_{m=\hat{J}}^{M} Y_m, \tag{5.13}$$

and it is also applied to the input of the adaptive binary detector block to calculate the test threshold γ_a, which maintains a desired performance as given by P_f in the appearing scenario and P_d in the disappearing scenario. Finally, a binary test is used to choose between the two hypotheses:

$$T_{M|\hat{J}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma_a, \tag{5.14}$$

where γ_a is the adaptive test threshold. Below, the functionality of the proposed adaptive ED is further studied under the two scenarios of interest.

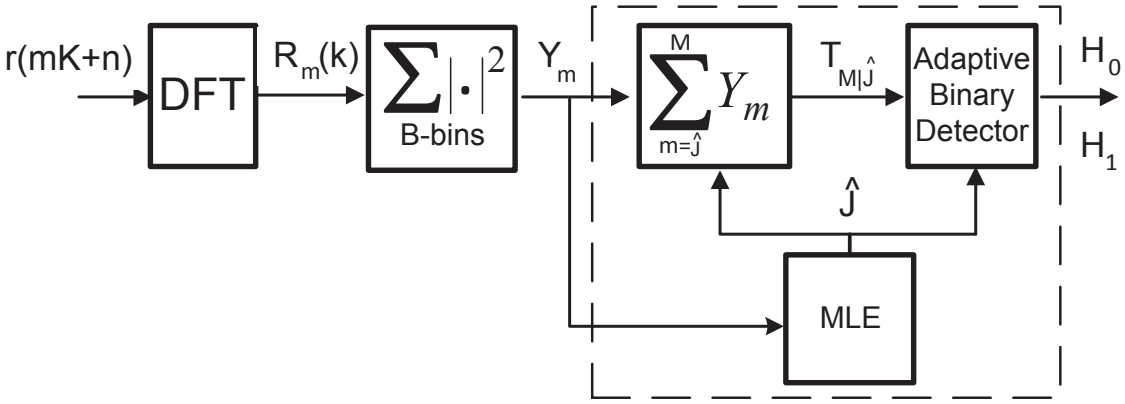


Fig. 5.1 Proposed GLRT-based adaptive ED procedure.

5.2.1 Appearing Scenario

Based on the two hypotheses in (5.1) and (5.2), the GLRT of the vector $\mathbf{R}$ is given by

$$L_G(\mathbf{R}) = \frac{f(\mathbf{R}|\hat{J},\mathcal{H}_1)}{f(\mathbf{R}|\mathcal{H}_0)} = \left(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}}\right)^{B(\hat{J}-1-M)} \exp\left[\frac{\sigma^2_{R|1}-\sigma^2_{R|0}}{\sigma^2_{R|1}\sigma^2_{R|0}} \sum_{m=\hat{J}}^{M} Y_m\right] \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{5.15}$$

where $\hat{J}$ is the MLE of the unknown change location J under $\mathcal{H}_1$:

$$\hat{J} = \underset{1\leq J\leq M}{\operatorname{argmax}}\ f(\mathbf{R}|J,\mathcal{H}_1) \equiv \underset{1\leq J\leq M}{\operatorname{argmax}}\ \ln(f(\mathbf{R}|J,\mathcal{H}_1)). \tag{5.16}$$

The objective function in the above optimization problem is either $f(\mathbf{R}|J,\mathcal{H}_1)$ which is provided in (5.5), or equivalently, its logarithm

$$\ln(f(\mathbf{R}|J,\mathcal{H}_1)) = B(J-1)\ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}}) - \frac{1}{\sigma^2_{R|0}}\sum_{m=1}^{J-1} Y_m + BM\ln(\frac{1}{\pi\sigma^2_{R|1}}) - \frac{1}{\sigma^2_{R|1}}\sum_{m=J}^{M} Y_m, \tag{5.17}$$

since the logarithm is a monotonic function. The terms $-B\ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}})$ and $BM\ln(\frac{1}{\pi\sigma^2_{R|1}})$ in (5.17) are independent of J; thus, we can ignore them in the objective function and write (5.16) as

$$\hat{J} = \underset{1\leq J\leq M}{\operatorname{argmax}}\ \left\{ BJ\ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}}) - \frac{1}{\sigma^2_{R|0}}\sum_{m=1}^{J-1} Y_m - \frac{1}{\sigma^2_{R|1}}\sum_{m=J}^{M} Y_m \right\}, \tag{5.18}$$

which can be further simplified into

$$\begin{aligned}
\hat{J} &= \underset{1\leq J\leq M}{\operatorname{argmax}}\ \left\{ BJ\ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}}) + \frac{1}{\sigma^2_{R|0}}(-\sum_{m=1}^{M} Y_m + \sum_{m=J}^{M} Y_m) - \frac{1}{\sigma^2_{R|1}}\sum_{m=J}^{M} Y_m \right\} \\
&= \underset{1\leq J\leq M}{\operatorname{argmax}}\ \left\{ BJ\ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}}) + \frac{\sigma^2_{R|1}-\sigma^2_{R|0}}{\sigma^2_{R|1}\sigma^2_{R|0}}\sum_{m=J}^{M} Y_m \right\}.
\end{aligned} \tag{5.19}$$

The computational complexity [3] of MLE process in (5.19) is of order M^2; therefore, an iterative method is proposed to reduce it to the order of M as follows. We let C_J denote the objective function in the above optimization problem:

$$C_J = BJ\ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}}) + \frac{\sigma^2_{R|1}-\sigma^2_{R|0}}{\sigma^2_{R|1}\sigma^2_{R|0}}\sum_{m=J}^{M} Y_m. \tag{5.20}$$

[3] A more detailed computational complexity analysis is provided in Appendix D.

Then the values of C_J for $J = M$ to $J = 1$ can be computed iteratively as follows:

$$C_M = BM\ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}}) + \frac{\sigma^2_{R|1} - \sigma^2_{R|0}}{\sigma^2_{R|1}\sigma^2_{R|0}}Y_M, \qquad J = M \tag{5.21}$$

$$C_J = C_{J+1} - B\ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}}) + \frac{\sigma^2_{R|1} - \sigma^2_{R|0}}{\sigma^2_{R|1}\sigma^2_{R|0}}Y_J, \quad 1 \leq J < M. \tag{5.22}$$

Going back to the Fig. 5.1, the MLE block iteratively computes C_J for $1 \leq J \leq M$ and declares $\hat{J}$ as the index corresponding to the maximum. Thus, the computation complexity of this process is only of order M.

Taking the natural logarithm of the GLRT in (5.15), the latter can be manipulated into an ED of the form given in (5.14). The distribution of the test statistic $T_{M|\hat{J}}$ under $\mathcal{H}_0$ is required to obtain the adaptive test threshold γ_a. As explained in Appendix A, the frame energies, $\{Y_m; 1 \leq m \leq M\}$, are i.i.d chi-square random variables with $2B$ degrees of freedom; therefore, $T_{M|\hat{J}}$ in (5.13) is also chi-square distributed with $2B(M-\hat{J}+1)$ degrees of freedom with the following mean and variance under $\mathcal{H}_0$:

$$\mu_{T|\mathcal{H}_0} = (M - \hat{J} + 1)\mu_{Y|0}, \qquad \sigma^2_{T|\mathcal{H}_0} = (M - \hat{J} + 1)\sigma^2_{Y|0}, \tag{5.23}$$

where $\mu_{Y|0}$ and $\sigma^2_{Y|0}$ are defined in (3.7a)-(3.7b). Based on the central limit theorem, $T_{M|\hat{J}}$ is approximately normally distributed for large values of B; therefore, γ_a is obtained for a desired P_f as follows:

$$\gamma_a = Q^{-1}(P_f)\sigma_{T|\mathcal{H}_0} + \mu_{T|\mathcal{H}_0}. \tag{5.24}$$

The MLE block uses the proposed iterative method to estimate the unknown change location J and its output, $\hat{J}$, is used to update the test statistic $T_{M|\hat{J}}$ in (5.13) and the test threshold γ_a in (5.24).

5.2.2 Disappearing scenario

The two hypotheses in this scenario are provided in (5.8) and (5.9). In contrast with the appearing scenario, the PDF of $\mathbf{R}$ under $\mathcal{H}_0$ is dependent on the unknown change location

J; therefore, the GLRT is given by

$$L_G(\mathbf{R}) = \frac{f(\mathbf{R}|\mathcal{H}_1)}{f(\mathbf{R}|\hat{J},\mathcal{H}_0)} = (\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}})^{B(\hat{J}-1-M)} \exp\left[\frac{\sigma^2_{R|1}-\sigma^2_{R|0}}{\sigma^2_{R|1}\sigma^2_{R|0}} \sum_{m=\hat{J}}^{M} Y_m\right] \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{5.25}$$

where $\hat{J}$ is obtained as

$$\hat{J} = \underset{1\le J\le M}{\operatorname{argmax}}\ f(\mathbf{R}|J,\mathcal{H}_0) \equiv \underset{1\le J\le M}{\operatorname{argmax}}\ \ln(f(\mathbf{R}|J,\mathcal{H}_0)). \tag{5.26}$$

Here, the objective function $f(\mathbf{R}|J,\mathcal{H}_0)$ is given in (5.10), where its logarithm is

$$\ln(f(\mathbf{R}|J,\mathcal{H}_0)) = B(J-1)\ln(\frac{\sigma^2_{R|0}}{\sigma^2_{R|1}}) - \frac{1}{\sigma^2_{R|1}}\sum_{m=1}^{J-1} Y_m + BM(\frac{1}{\pi\sigma^2_{R|0}}) - \frac{1}{\sigma^2_{R|0}}\sum_{m=J}^{M} Y_m. \tag{5.27}$$

Similar to the appearing scenario, the terms in (5.27) that are independent of J are ignored; thus (5.26) can be simplified to

$$\hat{J} = \underset{1\le J\le M}{\operatorname{argmax}}\ BJ\ln(\frac{\sigma^2_{R|0}}{\sigma^2_{R|1}}) - \frac{1}{\sigma^2_{R|1}}\sum_{m=1}^{J-1} Y_m - \frac{1}{\sigma^2_{R|0}}\sum_{m=J}^{M} Y_m, \tag{5.28}$$

and rewritten as

$$\begin{aligned}\hat{J} &= \underset{1\le J\le M}{\operatorname{argmax}}\ BJ\ln(\frac{\sigma^2_{R|0}}{\sigma^2_{R|1}}) + \frac{1}{\sigma^2_{R|1}}(-\sum_{m=1}^{M} Y_m + \sum_{m=J}^{M} Y_m) - \frac{1}{\sigma^2_{R|0}}\sum_{m=J}^{M} Y_m \\ &= \underset{1\le J\le M}{\operatorname{argmax}}\ BJ\ln(\frac{\sigma^2_{R|0}}{\sigma^2_{R|1}}) + \frac{\sigma^2_{R|0}-\sigma^2_{R|1}}{\sigma^2_{R|1}\sigma^2_{R|0}}\sum_{m=J}^{M} Y_m.\end{aligned} \tag{5.29}$$

Notice that the objective function in (5.29) is of the same magnitude but of different sign than the objective function in (5.19); thus, the iterative method proposed for the appearing scenario can be applied here by negating (5.21) and (5.22) in order to reduce the computational complexity to the order of M. Similar to the appearing scenario, the GLRT in (5.25) can be manipulated and transformed into an adaptive ED given in (5.14), where the test statistic $T_{M|\hat{J}}$ is approximately normally distributed for large values of B.

Therefore, the mean and variance of $T_{M|\hat{J}}$ under $\mathcal{H}_1$ are given by

$$\mu_{T|\mathcal{H}_1} = (M - \hat{J} + 1)\mu_{Y|1}, \qquad \sigma^2_{T|\mathcal{H}_1} = (M - \hat{J} + 1)\sigma^2_{Y|1}, \tag{5.30}$$

where $\mu_{Y|1}$ and $\sigma^2_{Y|1}$ are defined in (3.7a)-(3.7b). Using the above results, the adaptive test threshold γ_a in (5.14) is computed for a desired P_d as

$$\gamma_a = Q^{-1}(P_d)\sigma_{T|\mathcal{H}_1} + \mu_{T|\mathcal{H}_1}. \tag{5.31}$$

The proposed adaptive ED solves the change detection problem using the GLRT approach and the simulation results in Section 5.4.1 prove its superiority over the conventional ED and also the proposed adaptive ED in Chapter 4. In the next section, the composite-Bayesian approach introduced in Section 2.1.2 is used to solve the change detection problem.

5.3 Composite-Bayesian Approach

In the previous section, the change location J was assumed to be an unknown parameter, where its MLE is used to construct the LR test. In this section, the parameter J is modelled as a discrete random variable with a known PMF; therefore, we can use the composite-Bayesian approach to solve the composite hypothesis testing problem. In order to obtain the PMF of J, the PU channel access pattern is modelled as a Markov chain.

The Markov chain has been widely used in the literature to model the PU's activity status at a given *instant* [4] [46, 51, 80, 90], and the authors in [91] have validated its use based on real-time measurements. Therefore, in this section, a two-state Markov chain is used to model the PU's activity at a given *sensing frame* during the whole observation interval. More specially, we let $\{S_m; 1 \leq m \leq M\}$ be the discrete-time Markov process with $S_m \in \{1, 0\}$, representing channel states during the sensing period of M frames, where $S_m = 0$ and $S_m = 1$ correspond to the idle state (PU inactive) and the busy state (PU active) of the mth frame, respectively. Based on this model, the channel is at a certain state at a given *instant* (i.e., sensing frame), with the state changing according to specific transition probabilities between instants. As depicted in Fig. 5.2, the transition matrix for

[4] In the literature, the unit of instant can be a sampling point, a received symbol, a packet, or a sensing frame. In this thesis, the time evolutions of the PU's activity is described in the units of a sensing frame, which corresponds to K consecutive signal samples.

this model is given by

$$P = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}, \tag{5.32}$$

where p_{ij} is the homogeneous transition probability from state i to state j. The two scenarios of interest are special cases of the above model where the probability that a PU changes its activity status at a given sensing frame is geometrically distributed. In other words, the random variable J is geometrically distributed with the probability of "success" (probability that the PU changes its activity status) given by p_{01} and p_{10} in the appearing and disappearing scenario, respectively. We assume that these transition probabilities are known from *a priori* estimation of the PU's activity under each scenario; therefore, the PMF of J is completely known in both scenarios.

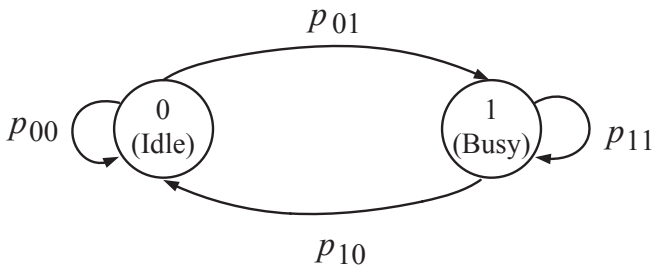


Fig. 5.2 The Markov channel state model.

The composite-Bayesian approach uses the PMF of J to obtain the marginal PDF of the vector $\mathbf{R}$ in (5.5) and (5.10), as explained in Section 2.1.2. Once the dependency on the variable J is removed, the two resultant conditional PDFs of the vector $\mathbf{R}$, $f(\mathbf{R}|\mathcal{H}_0)$ and $f(\mathbf{R}|\mathcal{H}_1)$, under the two scenarios are completely known. Therefore, the *composite* hypothesis testing problem is transformed into the *simple* hypothesis testing one, where the approaches introduced in Section 2.1.1 are applicable. In this case, the probability of occurrence of each hypothesis under the two scenarios can be obtained according to the available knowledge about the channel occupancy model. Hence, the Bayesian formulation of the problem is carried out throughout this section.

The use of this approach results into an adaptive ED as it is shown in Fig. 5.3. Here,

the energies of all M sensing frames are used to construct the test statistic

$$T_m = \sum_{m=1}^{M} Y_m. \tag{5.33}$$

The adaptive binary detector also receives the frame energies Y_m and dynamically adjusts the test threshold γ_a to choose between the two hypotheses as follows:

$$T_M \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma_a. \tag{5.34}$$

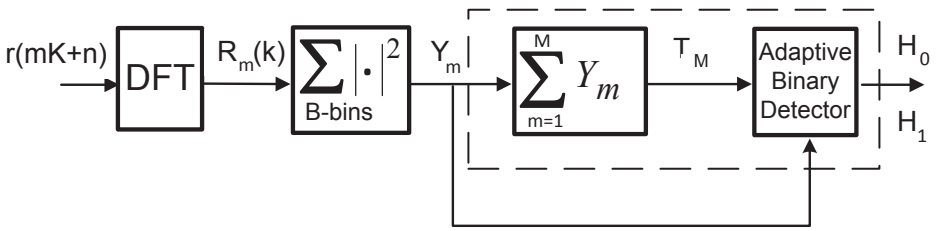


Fig. 5.3 Proposed adaptive ED procedure based on the composite-Bayesian approach.

The detail functionality of the proposed adaptive ED is further studied under the appearing and disappearing scenarios.

5.3.1 Appearing Scenario

The two hypotheses are provided in (5.1) and (5.2), where it is assumed that the PU is inactive before the sensing starts and suddenly becomes active at the beginning of the Jth sensing frame. Furthermore, the PU is assumed to remain active for a period longer than sensing time. Therefore, according to the channel occupancy model, the transition matrix in this case becomes

$$P = \begin{pmatrix} p_{00} & p_{01} \\ 0 & 1 \end{pmatrix}. \tag{5.35}$$

Based on the above model, the change location J is a random variable with geometric distribution parametrized by p_{01}. Therefore, the PMF of J is given by

$$p(J) = (1 - p_{01})^{J-1} p_{01} = p_{00}^{J-1} p_{01}, \tag{5.36}$$

which is used according to the composite-Bayesian approach to obtained the marginal PDF of $f(\mathbf{R}|J, \mathcal{H}_1)$ as follows:

$$f(\mathbf{R}|\mathcal{H}_1) = \sum_{J=1}^{M} f(\mathbf{R}|J, \mathcal{H}_1) p(J). \tag{5.37}$$

At this point, both $f(\mathbf{R}|\mathcal{H}_1)$ and $f(\mathbf{R}|\mathcal{H}_0)$ are completely known; therefore, the LR test can then be obtained as

$$L_B(\mathbf{R}) = \frac{f(\mathbf{R}|\mathcal{H}_1)}{f(\mathbf{R}|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{5.38}$$

where the test threshold η is selected based on the Bayesian approach (see Section 2.1.1) to minimize the probability of error. In particular, the probabilities of occurrence of the two hypotheses $Pr(\mathcal{H}_0) = \pi_0$ and $Pr(\mathcal{H}_1) = \pi_1$ are used to obtain η as $\frac{\pi_0}{\pi_1}$. Based on the definition of the two hypotheses, it should be clear that $\mathcal{H}_1$ occurs if any of the following mutually exclusive events takes place, where either $J = 1$, $J = 2$, ..., or $J = M$. Therefore, π_1 can be obtained as the summation of the probabilities of these M mutually exclusive events:

$$\pi_1 = \sum_{J=1}^{M} p(J), \tag{5.39}$$

where $p(J)$ is the PMF of J given in (5.36). Consequently, π_0 is equal to $1 - \pi_1$.

Finally, by taking the natural logarithm of the LR in (5.38) and further manipulation we obtain the adaptive ED given in (5.34), where γ_a is [5]

$$\gamma_a = \frac{\sigma_{R|0}^2 \sigma_{R|1}^2}{\sigma_{R|1}^2 - \sigma_{R|0}^2} \left[\ln(\frac{\pi_0}{\pi_1 p_{01}}) + BM \ln(\frac{\sigma_{R|1}^2}{\sigma_{R|0}^2}) - \ln\left(1 + \sum_{J=2}^{M} p_{00}^{J-1} (\frac{\sigma_{R|1}^2}{\sigma_{R|0}^2})^{B(J-1)} \psi(J)\right)\right], \tag{5.40}$$

for

$$\psi(J) = \exp\left[\frac{-\sigma_{R|1}^2 + \sigma_{R|0}^2}{\sigma_{R|0}^2 \sigma_{R|1}^2} \sum_{m=1}^{J-1} Y_m\right]. \tag{5.41}$$

[5] The Derivation of γ_a in the appearing scenario is provided in Appendix E.1

The adaptive binary detector received the measured energies of M frames and adjust its test threshold γ_a according to (5.40) in order to minimize the probability of error.

5.3.2 Disappearing Scenario

The two hypotheses are provided in (5.8) and (5.9). In contrast to the appearing scenario, we assume that the PU is active before the sensing starts and a sudden change in the PU's activity occurs at the beginning of the Jth frame. The PU is further assumed to maintain its inactive status for a time longer than the sensing period. Thus, the transition matrix in this case is given by

$$P = \begin{pmatrix} 1 & 0 \\ p_{10} & p_{11} \end{pmatrix}, \tag{5.42}$$

where p_{10} denotes the probability of the event that the PU becomes inactive at the beginning of the current sensing frame given that it was active in the previous one. Based on the above model, the random variable J is geometrically distributed with the probability of "success" (probability of the PU becoming inactive) given by p_{10}. Thus, the PMF of J is obtained for $1 \leq J \leq M$ as

$$p(J) = (1 - p_{10})^{J-1} p_{10} = p_{11}^{J-1} p_{10}. \tag{5.43}$$

In this scenario, the PDF of $\mathbf{R}$ under $\mathcal{H}_0$ is dependent on the unknown parameter J. Therefore, the composite-Bayesian approach is used to remove the dependency of $f(\mathbf{R}|J, \mathcal{H}_0)$ on J, and construct the LR test as follows:

$$L_B(\mathbf{R}) = \frac{f(\mathbf{R}|\mathcal{H}_1)}{\sum_{J=1}^{M} f(\mathbf{R}|J, \mathcal{H}_0) p(J)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{5.44}$$

where the test threshold η is equal to $\frac{\pi_0}{\pi_1}$ in order to minimize the probability of error. Similar to the appearing scenario, the probability of occurrence of $\mathcal{H}_0$, π_0, is calculated using $p(J)$ as follows:

$$\pi_0 = \sum_{J=1}^{M} p(J), \tag{5.45}$$

while, the probability of occurrence of $\mathcal{H}_1$ is equal to $\pi_1 = 1 - \pi_0$. Finally, applying the logarithm function on both sides of the LR test in (5.44) along with further manipulations result into the adaptive ED given in (5.34), where γ_a is given by [6]

$$\gamma_a = \frac{\sigma^2_{R|0}\sigma^2_{R|1}}{\sigma^2_{R|1} - \sigma^2_{R|0}} \left[\ln(\frac{\pi_0 p_{10}}{\pi_1}) + BM \ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}}) + \ln\left(1 + \sum_{J=2}^{M} p_{11}^{J-1} (\frac{\sigma^2_{R|0}}{\sigma^2_{R|1}})^{B(J-1)} \psi(J)\right)\right], \tag{5.46}$$

where

$$\psi(J) = \exp\left[\frac{-\sigma^2_{R|0} + \sigma^2_{R|1}}{\sigma^2_{R|0}\sigma^2_{R|1}} \sum_{m=1}^{J-1} Y_m\right]. \tag{5.47}$$

The proposed adaptive ED solves the change detection problem using the composite-Bayesian approach and it uses the *a priori* probabilities of occurrence of the two hypotheses to minimize the probability of error. The performance of the proposed adaptive ED is compared with that of the conventional ED whose test threshold is provided in (3.13) and the result of this comparison is provided in the following section.

5.4 Simulation Results

In this section, the performance of the two adaptive EDs proposed in Section 5.2 and 5.3 are evaluated through simulations and compared with the performance of the conventional ED in the two scenarios of interest. The simulation details are provided for the two adaptive EDs in the followings.

5.4.1 GLRT-Based Adaptive ED

The adaptive test threshold in this approach is selected based on the NP criterion in order to either maximize P_d subject to a constraint on the P_f in the appearing scenario or to maximize P_h subject to a constraint on P_d in the disappearing scenario. Therefore, its performance should be compared with that of the conventional ED whose test threshold is also selected based on the NP criterion as provided in (3.12) and (3.16) for appearing and disappearing scenario, respectively. The simulation parameters are chosen to be identical to the ones specified in Table 4.1, where $P_f = 0.1$ and $P_d = 0.98$ are the constraints used in the NP criterion. However, the parameters a, b, P_{fs}, and P_{ds} in Table 4.1 are only used for

[6] The derivation of γ_a in the disappearing is provided in Appendix E.2

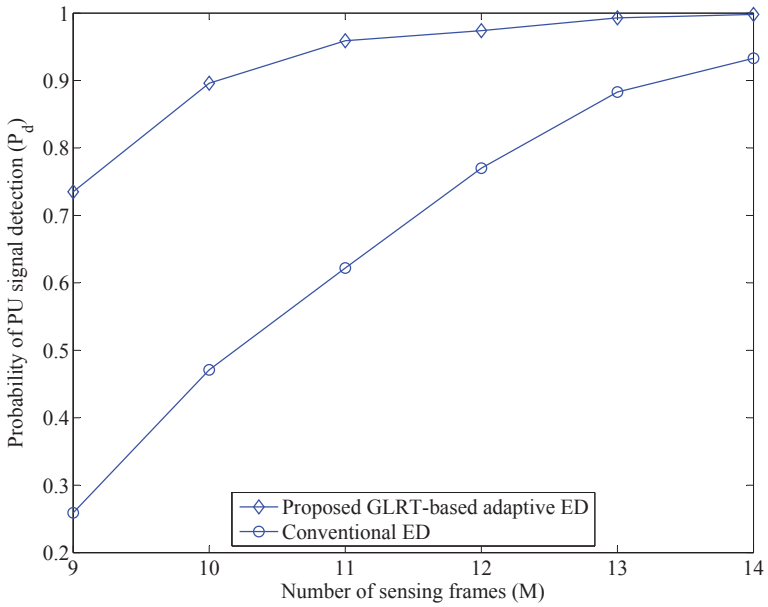


Fig. 5.4 Performance comparison of the GLRT-based adaptive ED and the conventional ED in the appearing scenario.

the adaptive ED proposed in Chapter 4 and they are not required by the GLRT approach. Similar to what has been done in Section 4.3, J is set to 9 and M is varied from 9 to 14 in all the simulations. For each value of M, random data are generated based on the model introduced in Section 3.1 and the simulations are run for 1000 independent trials to obtain an estimate of P_d and P_h for appearing and disappearing scenarios, respectively. The results of these experiments are shown in Fig. 5.4 and Fig. 5.5 for the two scenarios of interest and it is observed that the proposed adaptive ED outperforms the conventional ED in both cases. As explained in Section 3.2, the corrupting energies in the first $J-1$ frames caused the performance deterioration of the conventional ED in the dynamic environments; therefore, the proposed adaptive ED uses the MLE of the unknown change location J to discard those energies and to adjust its test threshold, which is shown to improve the detection performance significantly.

Note that the proposed adaptive ED in Chapter 4 is also formulated based on the NP criterion; therefore, we can compare its performance with that of the proposed adaptive ED based on the GLRT approach. However, as explained in Chapter 4, there is a trade-off associated with the former adaptive ED, which causes an increase in P_f in the appearing

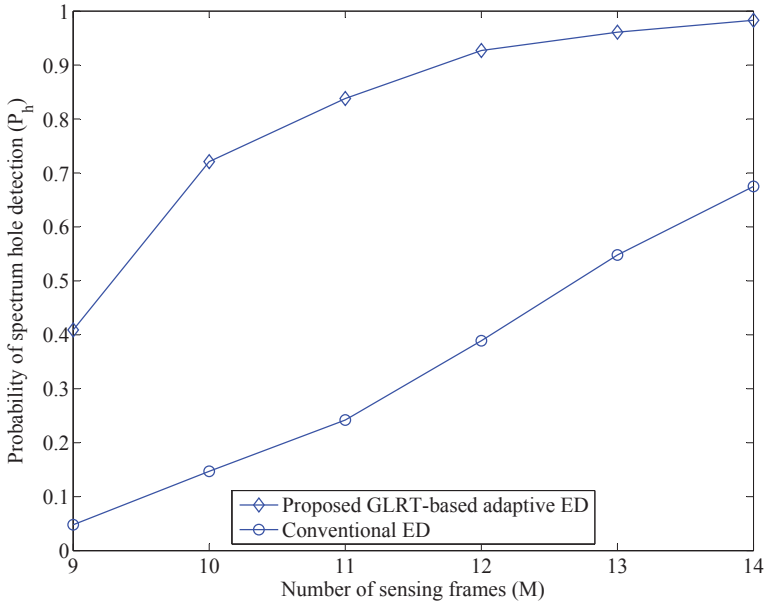


Fig. 5.5 Performance comparison of the GLRT-based adaptive ED and the conventional ED in the disappearing scenario.

scenario and a decrease in P_d in the disappearing scenario. Therefore, to have a fair comparison, its test threshold γ_a is adjusted through an exhaustive search to maintain a desired P_f and P_d in the appearing and disappearing scenario, respectively. The simulation results are illustrated in Fig. 5.6 and Fig. 5.7 and they prove that the GLRT-based adaptive ED has superior performance over the adaptive ED based on the exponential weighting approach. The GLRT-based adaptive ED uses the MLE of the unknown change location J to eliminate the corrupting energies in the first $J-1$ frames and the simulation results prove that this is a more effective approach than applying exponential weighting window over the measured energies.

5.4.2 Composite-Bayesian-Based Adaptive ED

In this approach, the adaptive test thresholds is selected based on the Bayesian approach to minimize the probability of error P_e; therefore, its performance is compared with that of the conventional ED whose test threshold is also chosen based on the Bayesian approach as provided in (3.13). Here, we also use the same simulation parameters as provided in Table 4.1; however, parameters J, a, b, P_{fs}, P_f, P_{ds}, and P_d are no longer required for

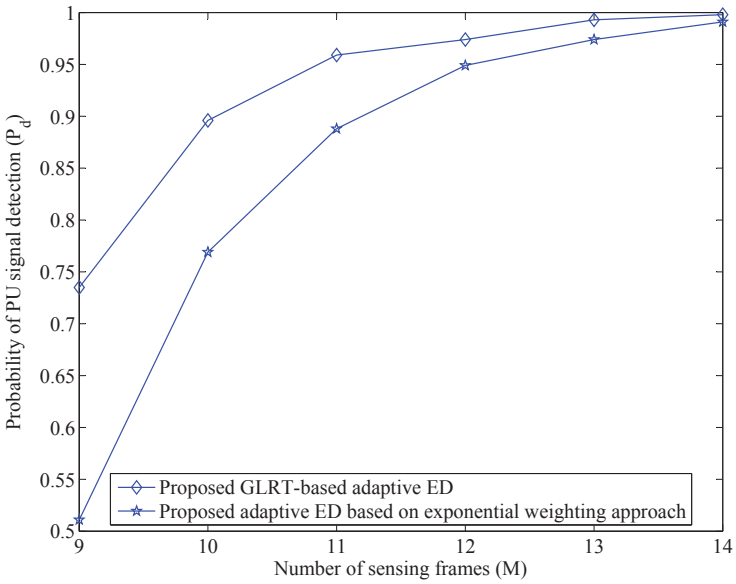


Fig. 5.6 Performance comparison of the GLRT-based adaptive ED and the adaptive ED based on the exponential weighting approach in the appearing scenario after γ_a is adjusted to maintain $P_f = 0.1$.

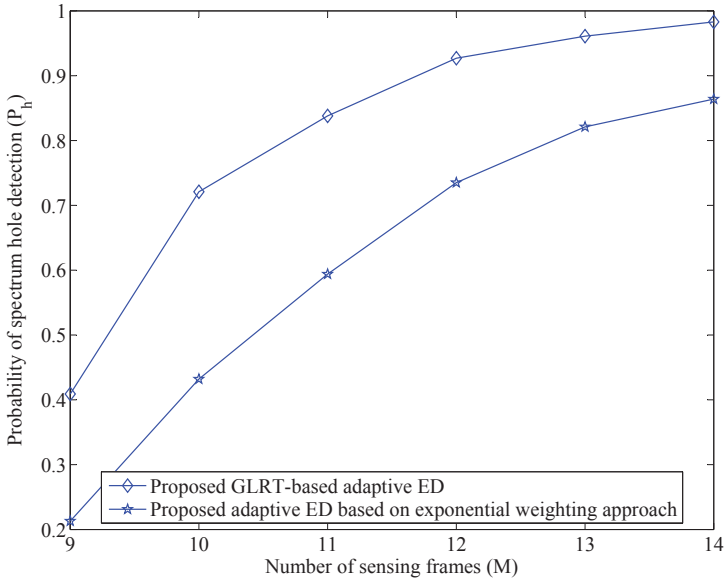


Fig. 5.7 Performance comparison of the GLRT-based adaptive ED and the adaptive ED based on the exponential weighting approach in the disappearing scenario after γ_a is adjusted to maintain $P_d = 0.98$.

the composite-Bayesian approach. In addition, we let $p_{01} = 0.125$ and $p_{10} = 0.125$ define the transition matrix P in (5.35) and (5.42); therefore, the parameter J is modelled as a geometric random variable parametrized by p_{01} and p_{10} in the appearing and disappearing scenario, respectively. In all simulations, M is varied from 9 to 16. For each value of M, random data are generated according to the system model introduced in Section 3.1, while the activity status of the PU is changed from one frame to another according to the specified transition matrix P under each scenario. The simulations are run for 1000 independent trials under each value of M to obtain an estimate of P_m and P_f which are finally added together to estimate P_e. The results are shown in Fig. 5.8 and Fig. 5.9 for appearing and disappearing scenarios respectively and it is observed that the proposed adaptive ED outperforms the conventional ED. The performance improvement achieved by the proposed adaptive ED was expected since the composite-Bayesian approach takes advantage of the *a priori* knowledge about the PMF of the change location J to construct the LR test.

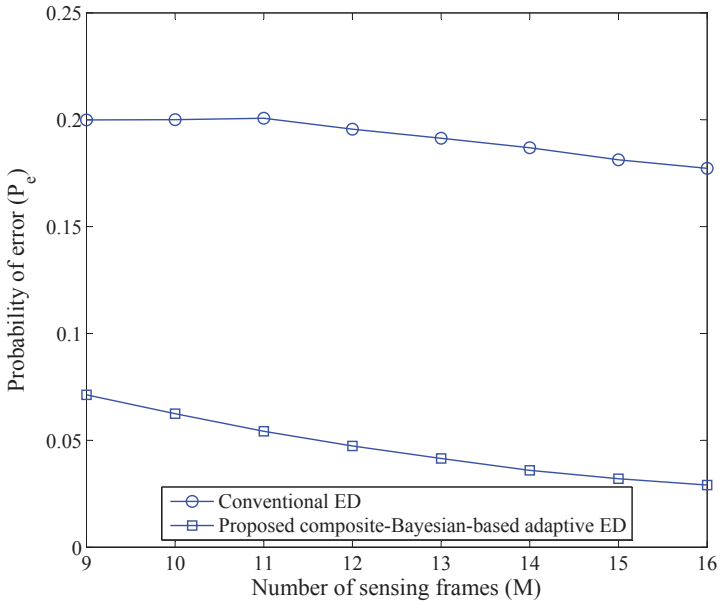


Fig. 5.8 Performance comparison of the composite-Bayesian-based adaptive ED and the conventional ED in the appearing scenario.

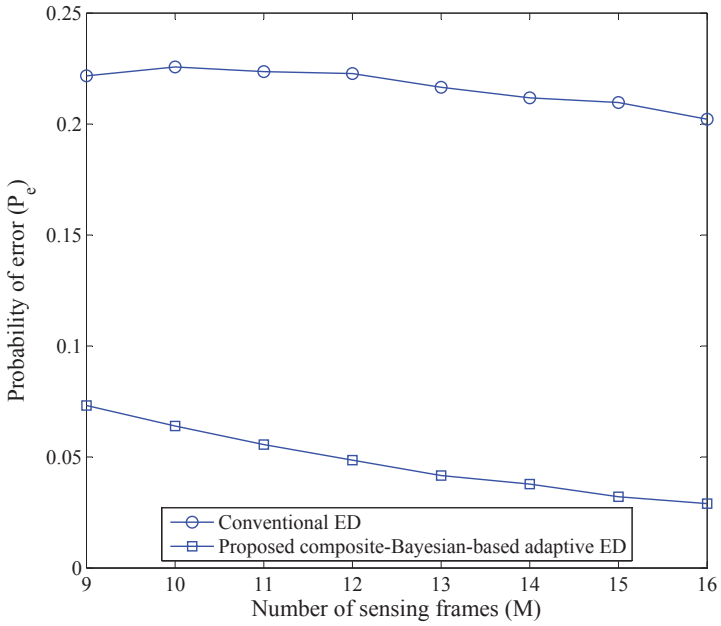


Fig. 5.9 Performance comparison of the composite-Bayesian-based adaptive ED and the conventional ED in the disappearing scenario.

5.4.3 Performance Comparison of the Two Proposed Techniques

Finally, the performance of the two proposed adaptive EDs are compared in the environment where the change location J is modelled as a geometric random variable. The simulation parameters provided in Table 4.1 are used here; however, the parameters J, a, b, P_{fs}, and P_{ds} are no longer required. Note that $P_f = 0.1$ and $P_d = 0.98$ are the desired probabilities of false alarm and signal detection respectively, which are used by the GLRT-based adaptive ED. In addition, we let $p_{01} = 0.125$ and $p_{10} = 0.125$ define the transition matrix P in (5.35) and (5.42). It is assumed that the composite-Bayesian-based adaptive ED has *a priori* knowledge of the transition matrix P under each scenario. Similar to the simulations in Section 5.4.2, random observation frames are generated for a given value of M in the range $[9, 16]$ such that the activity status of the PU changes from one frame to another according to the specified transition matrix under each scenario. The simulations are run for 1000 independent trials to obtain an estimate of P_m and P_f which are finally added together to estimate P_e. The simulation results are illustrated in Fig. 5.10 and Fig. 5.11 for appearing and disappearing scenarios respectively and it is observed that the composite-Bayesian-based adaptive ED outperforms the GLRT-based adaptive ED in

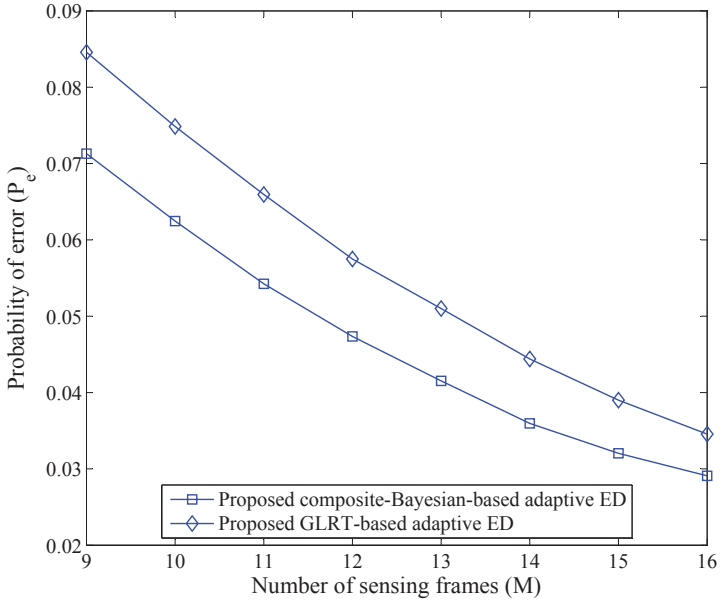


Fig. 5.10 Performance comparison of the composite-Bayesian-based adaptive ED and the GLRT-based adaptive ED in the appearing scenario.

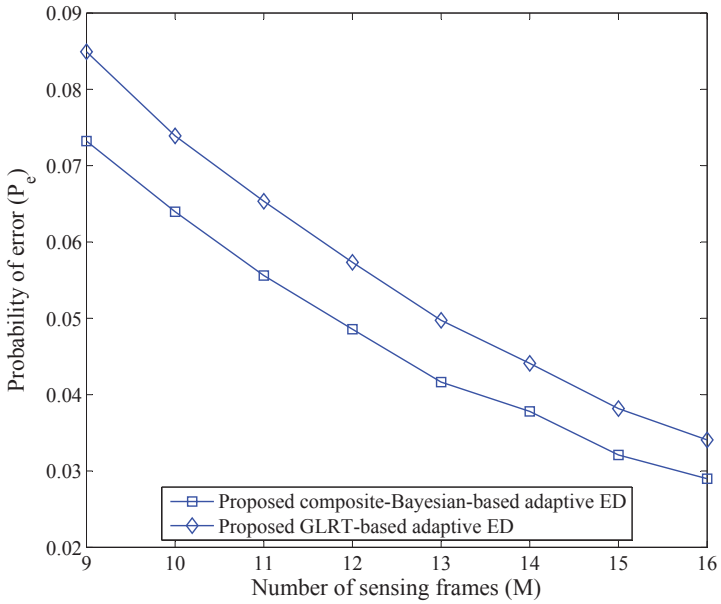


Fig. 5.11 Performance comparison of the composite-Bayesian-based adaptive ED and the GLRT-based adaptive ED in the disappearing scenario.

both scenarios. The superiority of the composite-Bayesian-based adaptive ED is expected due to the fact that this technique makes use of *a priori* knowledge of the transition matrix P to construct the test threshold γ_a as given in (5.40) and (5.46) for appearing and disappearing scenarios, respectively. As observed in the simulation results, the performance improvement may be significant depending on the application. However, the GLRT-based adaptive ED can still be used as an alternative technique to the composite-Bayesian-based adaptive ED if it is desired to save computations by avoiding *a priori* estimation of the transition matrix components.

5.5 Summary

In this chapter, the change detection problem within a fixed observation interval was formulated as the composite hypothesis testing and the two approaches introduced in Section 2.1.2 were employed to solve the given problem. The first approach, known as the GLRT, uses the MLE of the unknown change location to construct the LR test. The use of this approach resulted into an adaptive ED of the form (5.14), where the test threshold is selected to satisfy the NP criterion. The second approach, referred to as the composite-Bayesian, considers the change location J as a discrete random variable with a known PMF to construct the LR test. To obtain the PMF of J in the latter case, we modelled the PU channel access pattern as a two-state Markov chain whose transition probabilities are assumed to be known from *a priori* estimation. Applying the composite-Bayesian approach resulted into an adaptive ED of the form (5.34), where the test threshold is computed based on the Bayesian approach to minimize the probability of error. For both proposed adaptive EDs, the derived thresholds are different in appearing and disappearing scenarios. Therefore, a SU should employ the appearing scenario's thresholds for the case that it is already transmitting data and it is sensing the spectrum for the PU appearance. On the other hand, a SU should use the disappearing scenario's thresholds for the case that it is sensing for an unoccupied spectrum band. The simulation results have proved that the two proposed adaptive EDs outperform the conventional ED in both appearing and disappearing scenarios. Furthermore, the GLRT-based adaptive ED outperforms the previously proposed adaptive ED in Chapter 4. As the final remark, the composite-Bayesian based adaptive ED should be deployed to minimize the probability of error if the SU has *a priori* knowledge of the PU's activity within the sensing period. However, in the lack of such knowledge, the

GLRT-based adaptive ED should be deployed to achieve a better detection performance in comparison with the conventional ED.

Chapter 6

Conclusion

This dissertation addressed the practical limitations of the conventional ED in dynamic environments. In particular, we proved the performance deterioration of the conventional ED in two specific scenarios, where there was a single change in the PU's activity. The goal was to design adaptive EDs to improve the detection performance in such environments. In this chapter, we summarize the thesis, provide concluding points, and suggest some research directions that may be considered for future work in this area.

6.1 Thesis Summary

This thesis began with a background on signal detection in AWGN, which was further extended to spectrum sensing techniques for CR applications. The most common block-based spectrum sensing techniques were briefly covered, namely matched filter detection, cyclostationary feature detection, and energy detection. We mentioned that the matched filter technique requires *a priori* knowledge of the PU signal while the cyclostationary feature detection exploits the inherent periodicity of the modulated PU signal with a high computational complexity. The aforementioned drawbacks limit the use of these two techniques for CR applications. The energy detection is the most attractive spectrum sensing technique due to its ease of implementation, less restrictive assumptions (requires no *a priori* knowledge of the PU signal), and low computational complexity. We proceeded to prove the optimality of the conventional ED under the assumption that the PU received signals are modelled as a white Gaussian process corrupted by AWGN. The proof was based on the assumption that the received signal samples are homogeneous, that is, they belong to

one of the two hypotheses through out the sensing period. This assumption is not realistic in dynamic environments where the PU signal could appear and disappear at any time. We proved that the performance of the conventional ED actually deteriorates in dynamic environments where there is a single change in the PU's activity during the sensing period (i.e., the received samples are heterogeneous). More specially, we divided the analyses into two separate scenarios: (1) The PU appearance scenario; and (2) The PU disappearance scenario. The goal of the thesis was set to design adaptive EDs that could improve the detection performance in both of the aforementioned scenarios.

The work in [53] was introduced as the only existing adaptive ED in the literature that improves the detection performance in the appearing scenario. However, there was no optimality associated with that work, which provided the possible opportunity for further performance improvement. Furthermore, it did not considered the equally important scenario where the PU disappears during the sensing period. Therefore, we proposed three adaptive EDs to improve the detection performance in both appearing and disappearing scenarios. The first adaptive ED was based on an intuitive approach, where the measured energies of all sensing frames were exponentially weighted. In this case, the most recent measured energies are emphasized more as compared to the past ones. The exponential weighting approach has already been used to detect an abrupt change in digital signal processing applications [82, 83]; hence, it was applied as the first approach to the change detection problem in this thesis. In addition, the exponential weighting coefficients are dynamically adjusted according to the estimated change location in order to make the detector more compatible to the two dynamic scenarios of interest. Furthermore, the test threshold of the proposed detector is selected to satisfy the NP criterion. The analytical evaluation and simulation results proved the superiority of the proposed adaptive ED over the conventional ED in both scenarios. In the appearing scenario, the proposed adaptive ED also outperforms the work in [53].

In the next part of this study, the problem of change detection within the observation interval was formulated as the composite hypothesis testing. It was proven that the UMP test does not exist for this particular problem; therefore, two suboptimal approaches were considered. The first approach, known as the GLRT, uses the MLE of the unknown change location to construct the LR test. In this case, the test threshold is selected to satisfy the NP criterion similar to what has been done in the first proposed adaptive ED. In addition, an iterative method was proposed to reduce the computational complexity of the

MLE process. The performance of the resultant adaptive ED was compared with that of the conventional ED and also the first proposed adaptive ED. The simulation results proved that the GLRT-based adaptive ED has superior performance over the two alternative detectors in both scenarios. Therefore, we can conclude that the GLRT-based adaptive ED is the most suitable detector for these types of problems as it attains the best performance with a relatively low computational complexity. The second approach, referred to as the composite-Bayesian, considers the change location as a discrete random variable with a known PMF. In this case, the PU channel access pattern was modelled as a two-state Markov chain in order to obtain the PMF of the change location in both scenarios. The transition matrix of the Markov chain was assumed to be completely defined from *a priori* estimation of the PU's activity. The composite-Bayesian approach uses the PMF of the change location to construct the LR test. In this case, the *a priori* information regarding the PU channel access pattern was also used to obtain the probability of occurrence of the two hypotheses; therefore, the test threshold is selected based on the Bayesian approach to minimize the probability of error. The simulation results proved that the resultant adaptive ED outperforms the conventional ED in both scenarios. Finally, we can conclude that the composite-Bayesian-based adaptive ED should be deployed in the applications where the SU has *a priori* knowledge about the PMF of the change location within the observation interval.

6.2 Future Research Directions

We have proposed three adaptive EDs to improve the detection performance of the energy detection technique in dynamic environments; however, there is still opportunity to further extend this work for future research, which are outlined as follows:

- In this thesis, we considered the problem of change detection within a fixed observation interval from the individual SU's point of view. This idea can be incorporated into cooperative spectrum sensing as well in order to overcome the hidden terminal and channel fading problems.

- The reliable detection of a *single* change was considered in this thesis which leaves the detection of a sequence of changes an open area of research.

Appendix A

The Distribution of Y_m, its Mean and Variance

Here, we first prove that Y_m has a generalized chi-square distribution, that is $Y_m \sim \frac{\sigma_R^2}{2}\chi^2(2B)$, where σ_R^2 is the variance of $R_m(k)$. Then, we derive its mean and variance under each frame-hypothesis as provided in (3.7a)-(3.7b). The following theorems are used for this purpose:

Theorem 1 (Chi-square distribution [92, p. 373]) *Let U be a random variable defined by:*

$$U = \sum_{i=1}^{k} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2, \tag{A.1}$$

where the X_i are real, normally and independently distributed with means μ_i and variances σ_i^2. Then U has a chi-square distribution with k degrees of freedom, that is $U \sim \chi^2(k)$.

Theorem 2 (Sum of chi-square variates [93, p. 293]) *If $U_1, ..., U_k$ are independent and chi-square distributed random variables, $U_i \sim \chi^2(v_i)$, then $Y = U_1 + ... + U_k$ has chi-square distribution with v degrees of freedom, where $v = v_1 + ... + v_k$.*

Theorem 3 (First and second moment of chi-square variate [94, p. 571]) *The mean and variance of a chi-square random variable with v degrees of freedom, $\chi^2(v)$, are:*

$$E[\chi^2(v)] = v, \tag{A.2}$$

and

$$Var[\chi^2(v)] = 2v. \tag{A.3}$$

Y_m is the energy of the mth sensing frame which is given by

$$Y_m = \sum_{k\in\mathcal{B}} |R_m(k)|^2, \tag{A.4}$$

where $\{R_m(k)\}$ is a random process whose samples are independent across frequency and frame indices and follow a zero mean CSCG distribution with variances $\sigma^2_{R|1} = \sigma^2_x + \sigma^2_w$ and $\sigma^2_{R|0} = \sigma^2_w$ under the frame-hypothesis that the PU signal is present and absent, respectively. In turn, the squared magnitude of $R_m(k)$ is given by

$$|R_m(k)|^2 = (\mathrm{Re}[R_m(k)])^2 + (\mathrm{Im}[R_m(k)])^2, \tag{A.5}$$

where the real and imaginary part of $R_m(k)$ are jointly Gaussian, uncorrelated (which also implies independence in this case), and have the same variance equal to half of the variance of $R_m(k)$ (i.e., $\sigma^2_R/2$). Scaling Y_m by $\dfrac{2}{\sigma^2_R}$ results into the following:

$$\begin{aligned} Y'_m = \frac{2}{\sigma^2_R} Y_m &= \frac{2}{\sigma^2_R}\sum_{k\in\mathcal{B}} (\mathrm{Re}[R_m(k)])^2 + \frac{2}{\sigma^2_R}\sum_{k\in\mathcal{B}} (\mathrm{Im}[R_m(k)])^2 \\ &= \sum_{k\in\mathcal{B}} Z_{m,r}(k)^2 + \sum_{k\in\mathcal{B}} Z_{m,i}(k)^2, \end{aligned} \tag{A.6}$$

where $Z_{m,r}(k) = \dfrac{2(\mathrm{Re}[R_m(k)])^2}{\sigma^2_R}$ and $Z_{m,i}(k) = \dfrac{2(\mathrm{Im}[R_m(k)])^2}{\sigma^2_R}$ are independent standard normal random variables. Since the cardinality of the set $\mathcal{B}$ is B, it follows that Y'_m consists of two summations of B i.i.d standard normal random variables. Based on Theorem 1, each of these summations is a chi-square random variable with B degrees of freedom. These two random variables are independent of each other; therefore, it can be concluded based on Theorem 2 that Y'_m has a chi-square distribution with $2B$ degrees of freedom. According to (A.6), Y'_m is the scaled version of Y_m; thus, Y_m is also chi-square distributed as

$$Y_m \sim \frac{\sigma^2_R}{2}\chi^2(2B). \tag{A.7}$$

We can finally obtained the mean and variance of Y_m under each frame-hypothesis using Theorem 3 as follows:

$$E[Y_m|0] = \mu_{Y|0} = \frac{\sigma^2_{R|0}}{2}2B = \sigma^2_w B, \tag{A.8}$$

$$Var[Y_m|0] = \sigma^2_{Y|0} = (\frac{\sigma^2_{R|0}}{2})^2 4B = \sigma^4_w B, \tag{A.9}$$

$$E[Y_m|1] = \mu_{Y|1} = \frac{\sigma^2_{R|1}}{2}2B = (\sigma^2_x + \sigma^2_w)B, \tag{A.10}$$

$$Var[Y_m|1] = \sigma^2_{Y|1} = (\frac{\sigma^2_{R|1}}{2})^2 4B = (\sigma^2_x + \sigma^2_w)^2 B. \tag{A.11}$$

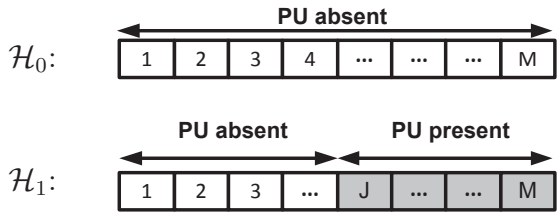
Appendix B

Derivation of the Mean and the Variance of T_M

In this appendix, the mean and the variance of the conventional ED's test statistic, T_M, is derived under the appearing and the disappearing scenarios.

B.1 Appearing Scenario

In this scenario, the PU appears on the Jth frame during the sensing period. Thus, the two hypotheses are represented graphically for M sensing frames as follows:



The conventional ED sums up the energy of M frames to construct the test statistic, i.e, $T_M = \sum_{m=1}^{M} Y_m$ in (3.6). The frame energies $\{Y_m : 1 \leq m \leq M\}$ are i.i.d chi-square random

variables. Therefore, the mean and variance of T_M is calculated under each hypothesis as

$$\mu_{T|\mathcal{H}_0} = E[T_M|\mathcal{H}_0] = E[\sum_{m=1}^{M} Y_m|\mathcal{H}_0] = \sum_{m=1}^{M} E[Y_m|0], \tag{B.1}$$

$$\sigma^2_{T|\mathcal{H}_0} = Var[T_M|\mathcal{H}_0] = Var[\sum_{m=1}^{M} Y_m|\mathcal{H}_0] = \sum_{m=1}^{M} Var[Y_m|0], \tag{B.2}$$

$$\mu_{T|\mathcal{H}_1} = E[T_M|\mathcal{H}_1] = E[\sum_{m=1}^{M} Y_m|\mathcal{H}_1] = \sum_{m=1}^{J-1} E[Y_m|0] + \sum_{m=J}^{M} E[Y_m|1], \tag{B.3}$$

$$\sigma^2_{T|\mathcal{H}_1} = Var[T_M|\mathcal{H}_1] = Var[\sum_{m=1}^{M} Y_m|\mathcal{H}_1] = \sum_{m=1}^{J-1} Var[Y_m|0] + \sum_{m=J}^{M} Var[Y_m|1], \tag{B.4}$$

where (B.2) and (B.4) follow from the independence of the random variables Y_m. Finally, defining $\mu_{Y|i} = E[Y_m|i]$ and $\sigma^2_{Y|i} = Var[Y_m|i]$ for $i = 0, 1$, the above equations result into

$$\mu_{T|\mathcal{H}_0} = M\mu_{Y|0}, \tag{B.5}$$

$$\sigma^2_{T|\mathcal{H}_0} = M\sigma^2_{Y|0}, \tag{B.6}$$

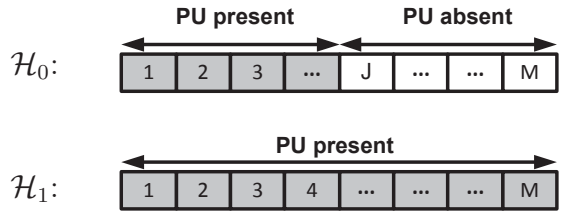
$$\mu_{T|\mathcal{H}_1} = (J-1)\mu_{Y|0} + (M-J+1)\mu_{Y|1}, \tag{B.7}$$

$$\sigma^2_{T|\mathcal{H}_1} = (J-1)\sigma^2_{Y|0} + (M-J+1)\sigma^2_{Y|1}, \tag{B.8}$$

where $\mu_{Y|0}$, $\sigma^2_{Y|0}$, $\mu_{Y|1}$, and $\sigma^2_{Y|1}$ are given in (A.8)-(A.11).

B.2 Disappearing Scenario

In this scenario, the two hypotheses are represented graphically as



Similar to the previous scenario, the mean and variance of T_M is obtained as follows:

$$\mu_{T|\mathcal{H}_0} = E[T_M|\mathcal{H}_0] = E[\sum_{m=1}^{M} Y_m|\mathcal{H}_0] = \sum_{m=1}^{J-1} E[Y_m|1] + \sum_{m=J}^{M} E[Y_m|0], \tag{B.9}$$

$$\sigma^2_{T|\mathcal{H}_0} = Var[T_M|\mathcal{H}_0] = Var[\sum_{m=1}^{M} Y_m|\mathcal{H}_0] = \sum_{m=1}^{J-1} Var[Y_m|1] + \sum_{m=J}^{M} Var[Y_m|0], \tag{B.10}$$

$$\mu_{T|\mathcal{H}_1} = E[T_M|\mathcal{H}_1] = E[\sum_{m=1}^{M} Y_m|\mathcal{H}_1] = \sum_{m=1}^{M} E[Y_m|1], \tag{B.11}$$

$$\sigma^2_{T|\mathcal{H}_1} = Var[T_M|\mathcal{H}_1] = Var[\sum_{m=1}^{M} Y_m|\mathcal{H}_1] = \sum_{m=1}^{M} Var[Y_m|1], \tag{B.12}$$

which result into

$$\mu_{T|\mathcal{H}_0} = (J-1)\mu_{Y|1} + (M-J+1)\mu_{Y|0}, \tag{B.13}$$

$$\sigma^2_{T|\mathcal{H}_0} = (J-1)\sigma^2_{Y|1} + (M-J+1)\sigma^2_{Y|0}, \tag{B.14}$$

$$\mu_{T|\mathcal{H}_1} = M\mu_{Y|1}, \tag{B.15}$$

$$\sigma^2_{T|\mathcal{H}_1} = M\sigma^2_{Y|1}. \tag{B.16}$$

Appendix C

Derivation of the PDF of R

In this appendix, the PDF of $\mathbf{R}$ is derived under the two hypotheses in the two scenarios of interest. Recall that $\mathbf{R} = \{R_m(k); 1 \leq m \leq M, k \in \mathcal{B}\}$, where the enteries are i.i.d CSCG random variables with zero mean and variance σ_R^2. More specially, $\sigma_{R|1}^2 = \sigma_w^2 + \sigma_x^2$ and $\sigma_{R|1}^2 = \sigma_w^2$ are the variances of $R_m(k)$ when the PU is present and absent, respectively. Due to the independence of the variates, the PDF of $\mathbf{R}$ is obtained as the product of the marginal PDFs of $R_m(k)$. Since $R_m(k)$ is CSCG distributed, its conditional PDF is given in a generic form by

$$f(R_m(k)|i) = \frac{1}{\pi\sigma_{R|i}^2} \exp\left[\frac{-|R_m(k)|^2}{\sigma_{R|i}^2}\right], \ i \in \{0,1\} \tag{C.1}$$

The next step is to use (C.1) to obtain PDF of $\mathbf{R}$ in both appearing and disappearing scenarios in terms of Y_m defined in (3.4).

C.1 Appearing Scenario

Based on (5.1)-(5.2), the PDFs of $\mathbf{R}$ under the two hypotheses are derived as follows:

$$\begin{aligned} f(\mathbf{R}|\mathcal{H}_0) &= \prod_{m=1}^{M}\prod_{k\in\mathcal{B}} f(R_m(k)|0) = \prod_{m=1}^{M}\prod_{k\in\mathcal{B}} \frac{1}{\pi\sigma_{R|0}^2} \exp\left[\frac{-|R_m(k)|^2}{\sigma_{R|0}^2}\right] \\ &= \prod_{m=1}^{M} (\frac{1}{\pi\sigma_{R|0}^2})^B \exp\left[\frac{-Y_m}{\sigma_{R|0}^2}\right] = (\frac{1}{\pi\sigma_{R|0}^2})^{BM} \exp\left[\frac{-\sum_{m=1}^{M} Y_m}{\sigma_{R|0}^2}\right], \end{aligned} \tag{C.2}$$

$$
\begin{aligned}
f(\mathbf{R}|J,\mathcal{H}_1) &= \prod_{m=1}^{J-1}\prod_{k\in\mathcal{B}} f(R_m(k)|0) \prod_{m=J}^{M}\prod_{k\in\mathcal{B}} f(R_m(k)|1) \\
&= \prod_{m=1}^{J-1}\prod_{k\in\mathcal{B}} \frac{1}{\pi\sigma_{R|0}^2}\exp\left[\frac{-|R_m(k)|^2}{\sigma_{R|0}^2}\right] \prod_{m=J}^{M}\prod_{k\in\mathcal{B}} \frac{1}{\pi\sigma_{R|1}^2}\exp\left[\frac{-|R_m(k)|^2}{\sigma_{R|1}^2}\right] \\
&= \prod_{m=1}^{J-1}(\frac{1}{\pi\sigma_{R|0}^2})^B \exp\left[\frac{-Y_m}{\sigma_{R|0}^2}\right] \prod_{m=J}^{M}(\frac{1}{\pi\sigma_{R|1}^2})^B \exp\left[\frac{-Y_m}{\sigma_{R|1}^2}\right] \\
&= (\frac{\sigma_{R|1}^2}{\sigma_{R|0}^2})^{B(J-1)} \exp\left[\frac{-\sum_{m=1}^{J-1} Y_m}{\sigma_{R|0}^2}\right] (\frac{1}{\pi\sigma_{R|1}^2})^{BM} \exp\left[\frac{-\sum_{m=J}^{M} Y_m}{\sigma_{R|1}^2}\right]. \quad \text{(C.3)}
\end{aligned}
$$

C.2 Disappearing Scenario

In this scenario, the two hypotheses are given in (5.8)-(5.9), where $\mathcal{H}_0$ is now conditioned on J. Similar to the appearing scenario, the PDF of $\mathbf{R}$ is derived as

$$
\begin{aligned}
f(\mathbf{R}|J,\mathcal{H}_0) &= \prod_{m=1}^{J-1}\prod_{k\in\mathcal{B}} f(R_m(k)|1) \prod_{m=J}^{M}\prod_{k\in\mathcal{B}} f(R_m(k)|0) \\
&= \prod_{m=1}^{J-1}\prod_{k\in\mathcal{B}} \frac{1}{\pi\sigma_{R|1}^2}\exp\left[\frac{-|R_m(k)|^2}{\sigma_{R|1}^2}\right] \prod_{m=J}^{M}\prod_{k\in\mathcal{B}} \frac{1}{\pi\sigma_{R|0}^2}\exp\left[\frac{-|R_m(k)|^2}{\sigma_{R|0}^2}\right] \\
&= \prod_{m=1}^{J-1}(\frac{1}{\pi\sigma_{R|1}^2})^B \exp\left[\frac{-Y_m}{\sigma_{R|1}^2}\right] \prod_{m=J}^{M}(\frac{1}{\pi\sigma_{R|0}^2})^B \exp\left[\frac{-Y_m}{\sigma_{R|0}^2}\right] \\
&= (\frac{\sigma_{R|0}^2}{\sigma_{R|1}^2})^{B(J-1)} \exp\left[\frac{-\sum_{m=1}^{J-1} Y_m}{\sigma_{R|1}^2}\right] (\frac{1}{\pi\sigma_{R|0}^2})^{BM} \exp\left[\frac{-\sum_{m=J}^{M} Y_m}{\sigma_{R|0}^2}\right], \quad \text{(C.4)} \\
f(\mathbf{R}|\mathcal{H}_1) &= \prod_{m=1}^{M}\prod_{k\in\mathcal{B}} f(R_m(k)|1) \\
&= \prod_{m=1}^{M}\prod_{k\in\mathcal{B}} \frac{1}{\pi\sigma_{R|1}^2}\exp\left[\frac{-|R_m(k)|^2}{\sigma_{R|1}^2}\right] \\
&= \prod_{m=1}^{M}(\frac{1}{\pi\sigma_{R|1}^2})^B \exp\left[\frac{-Y_m}{\sigma_{R|1}^2}\right] \\
&= (\frac{1}{\pi\sigma_{R|1}^2})^{BM} \exp\left[\frac{-\sum_{m=1}^{M} Y_m}{\sigma_{R|1}^2}\right]. \quad \text{(C.5)}
\end{aligned}
$$

Appendix D

Computational Complexity Analysis of the MLE Process

We analyse the computational complexity of the MLE process, which is represented with the associated number of real arithmetic operations as

$$\hat{J} = \underset{1\leq J\leq M}{\operatorname{argmax}} \quad \underbrace{\underbrace{BJ\ln(\frac{\sigma^2_{R|1}}{\sigma^2_{R|0}})}_{\text{1 mult.+1 log.}} + \underbrace{\frac{\sigma^2_{R|1}-\sigma^2_{R|0}}{\sigma^2_{R|1}\sigma^2_{R|0}}\sum_{m=J}^{M} Y_m}_{\text{1 mult.+}(M-J+1)\text{ summ.}}}_{\text{1 summ.}}, \tag{D.1}$$

where *mult.*, *summ.*, and *log.* stand for multiplication, summation, and logarithm operation, respectively. Thus, $5 + M - J$ arithmetic operations are required to obtain the objective function for a given J. Finally, the total computational complexity of the MLE process is obtained by summing the number of operations over $1 \leq J \leq M$ as follows:

$$\sum_{J=1}^{M}(4 + M - J) = M(4 + M) - \sum_{J=1}^{M} J = 4M + M^2 - \frac{M^2 + M}{2} = \frac{7M}{2} + \frac{M^2}{2}, \tag{D.2}$$

using the fact that

$$\sum_{i=1}^{n} i = \frac{n^2 + n}{2}, \tag{D.3}$$

for all natural number n. Based on (D.2), the computational complexity of MLE process is $\frac{7M}{2} + \frac{M^2}{2}$, which is dominated by $\frac{M^2}{2}$. Therefore, M^2 represents the order of the computation complexity.

Appendix E

Derivation of the Adaptive Test Threshold Based on the Composite-Bayesian Approach

In this appendix, we prove that the use of the composite-Bayesian approach in the two scenarios of interest results into an adaptive ED of the form (5.34), where the test threshold is given by (5.40) and (5.46) for the appearing and disappearing scenario, respectively.

E.1 Appearing Scenario

We first start by taking the natural logarithm on both sides of (5.38):

$$\ln(L_B(\mathbf{R})) = \ln\left(\sum_{J=1}^{M} f(\mathbf{R}|J,\mathcal{H}_1)p(J)\right) - \ln(f(\mathbf{R}|\mathcal{H}_0)) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \ln(\eta), \tag{E.1}$$

where $f(\mathbf{R}|\mathcal{H}_0)$, $f(\mathbf{R}|J,\mathcal{H}_1)$, and $p(J)$ are provided in (5.4), (5.5), and (5.36), respectively. Thus, the first term on the left hand side of the inequality in (E.1) is computed as

$$\begin{aligned}\ln\left(\sum_{J=1}^{M} f(\mathbf{R}|J,\mathcal{H}_1)p(J)\right) &= -BM\ln(\pi\sigma_{R|1}^2)+ \\ &\ln\left(\sum_{J=1}^{M}\left[(\frac{\sigma_{R|1}^2}{\sigma_{R|0}^2})^{B(J-1)}\exp\left[\frac{-\sum_{m=1}^{J-1}Y_m}{\sigma_{R|0}^2}\right]\exp\left[\frac{-\sum_{m=J}^{M}Y_m}{\sigma_{R|1}^2}\right]p_{00}^{(J-1)}p_{01}\right]\right).\end{aligned} \tag{E.2}$$

We let a_J be equal to the expression inside the summation on the right hand side of the equality in (E.2), and use the following logarithmic identity:

$$\ln(\sum_{J=1}^{M} a_J) = \ln(a_1) + \ln(1 + \sum_{J=2}^{M} \frac{a_J}{a_1}), \tag{E.3}$$

to rewrite (E.2) as

$$\ln\left(\sum_{J=1}^{M} f(\mathbf{R}|J,\mathcal{H}_1)p(J)\right) = -BM\ln(\pi\sigma_{R|1}^2) - \sum_{m=1}^{M}\frac{Y_m}{\sigma_{\mathbf{R}|1}^2} + \ln(p_{01}) + \ln\left(1 + \sum_{J=2}^{M}\left[(\frac{\sigma_{R|1}^2}{\sigma_{R|0}^2})^{B(J-1)}\exp\left[\frac{\sigma_{R|0}^2 - \sigma_{R|1}^2}{\sigma_{R|0}^2\sigma_{R|1}^2}\sum_{m=1}^{J-1} Y_m\right] p_{00}^{J-1}\right]\right). \tag{E.4}$$

The second term on the left hand side of the inequality in (E.1) is given as

$$\ln(f(\mathbf{R}|\mathcal{H}_0)) = -MB\ln(\pi\sigma_{R|0}^2) - \sum_{m=1}^{M}\frac{Y_m}{\sigma_{R|0}^2}. \tag{E.5}$$

Finally, substituting (E.5) and (E.4) back into (E.1) and further manipulation result into the adaptive ED of the form (5.34) where the test threshold is given by (5.40).

E.2 Disappearing Scenario

Similar to the appearing scenario, we start by taking the logarithm of both side in (5.44):

$$\ln(L_B(\mathbf{R})) = \ln(f(\mathbf{R}|\mathcal{H}_1)) - \ln\left(\sum_{J=1}^{M} f(\mathbf{R}|J,\mathcal{H}_0)p(J)\right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \ln(\eta), \tag{E.6}$$

where

$$\ln(f(\mathbf{R}|\mathcal{H}_1)) = -BM\ln(\pi\sigma_{R|1}^2) - \sum_{m=1}^{M}\frac{Y_m}{\sigma_{R|1}^2}, \tag{E.7}$$

$$\ln\left(\sum_{J=1}^{M} f(\mathbf{R}|J,\mathcal{H}_0)p(J)\right) = -BM\ln(\pi\sigma_{R|0}^2)+ \\ \ln\left(\sum_{J=1}^{M}\left[(\frac{\sigma_{R|0}^2}{\sigma_{R|1}^2})^{B(J-1)}\exp\left[\frac{-\sum_{m=1}^{J-1}Y_m}{\sigma_{R|1}^2}\right]\exp\left[\frac{-\sum_{m=J}^{M}Y_m}{\sigma_{R|0}^2}\right]p_{11}^{(J-1)}p_{10}\right]\right). \tag{E.8}$$

We use the same logarithmic identity provided in (E.3), to rewrite (E.8) as

$$\ln\left(\sum_{J=1}^{M} f(\mathbf{R}|J,\mathcal{H}_0)p(J)\right) = -BM\ln(\pi\sigma_{R|0}^2) - \sum_{m=1}^{M}\frac{Y_m}{\sigma_{R|0}^2} + \ln(p_{10})+ \\ \ln\left(1+\sum_{J=2}^{M}\left[(\frac{\sigma_{R|0}^2}{\sigma_{R|1}^2})^{B(J-1)}\exp\left[\frac{-\sigma_{R|0}^2+\sigma_{R|1}^2}{\sigma_{R|0}^2\sigma_{R|1}^2}\sum_{m=1}^{J-1}Y_m\right]p_{11}^{J-1}\right]\right). \tag{E.9}$$

Substituting (E.9) and (E.7) back into (E.6) plus further manipulation result into an adaptive ED of the form (5.34) where γ_a is given by (5.46).

References

[1] K. Patil, R. Prasad, and K. Skouby, "A survey of worldwide spectrum occupancy measurement campaigns for cognitive radio," in *Proc. Int. Conf. Devices and Commun.*, Feb. 2011, pp. 1–5.

[2] R. de Francisco and A. Pandharipande, "Spectrum occupancy in the 2.36 to 2.4 GHz band: Measurements and analysis," in *Proc. European Wireless Conf.*, Apr. 2010, pp. 231–237.

[3] Shared Spectrum Company, "General survey of radio frequency bands: 30 MHz to 3 GHz," Shared Spectrum Co., Rep., Aug. 2009.

[4] M. A. McHenry, "NSF spectrum occupancy measurements project summary," Shared Spectrum Co., Rep., Dec. 2005.

[5] S. Tang and B. Mark, "Modeling and analysis of opportunistic spectrum sharing with unreliable spectrum sensing," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 1934–1943, Apr. 2009.

[6] A. T. Hoang, Y.-C. Liang, D. Wong, Y. Zeng, and R. Zhang, "Opportunistic spectrum access for energy-constrained cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 1206–1211, Mar. 2009.

[7] J. Sun, H. Rashvand, and H.-B. Zhu, "Opportunistic spectrum access framework for video over ad hoc wireless networks," *IET Commun.*, vol. 4, no. 11, pp. 1269–1276, July 2010.

[8] X. Zhou, G. Li, D. Li, D. Wang, and A. Soong, "Probabilistic resource allocation for opportunistic spectrum access," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2870–2879, Sep. 2010.

[9] B. Hamdaoui, "Adaptive spectrum assessment for opportunistic access in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 922–930, Feb. 2009.

[10] Connecting America, "The national broadband plan," FCC, Rep., Mar. 2010.

[11] Federal Communication Commission, "Unlicensed operation in the TV broadcast bands: second memorandum opinion and order," FCC, Rep. ET Docket No. 04-186, Sep. 2010.

[12] ——, "Amendment of the commission's rules to provide spectrum for the operation of medical body area networks," FCC, Rep. ET Docket No. 08-59, June 2009.

[13] C. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. Shellhammer, and W. Caldwell, "IEEE 802.22: The first cognitive radio wireless regional area network standard," *IEEE Commun. Mag.*, vol. 47, pp. 130–138, Jan. 2009.

[14] *MAC and PHY for Operation in TV White Space*, Std. ECMA 392, Dec. 2009.

[15] M. Mueck, A. Piipponen, K. Kallioja, G. Dimitrakopoulos, K. Tsagkaris, P. Demestichas, F. Casadevall, J. Perez-Romero, O. Sallent, G. Baldini, S. Filin, H. Harada, M. Debbah, T. Haustein, J. Gebert, B. Deschamps, P. Bender, M. Street, S. Kandeepan, J. Lota, and A. Hayar, "ETSI reconfigurable radio systems: Status and future directions on software defined radio and cognitive radio standards," *IEEE Commun. Mag.*, vol. 48, pp. 78 –86, Sep. 2010.

[16] J. Mitola, "Software radios-survey, critical evaluation and future directions," presented at the Nat. Telesyst. Conf., May 1992.

[17] ——, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Royal Institute of Technology, Sweden, May 2000.

[18] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks J.*, vol. 50, pp. 2127–2159, May 2006.

[19] S.-J. Kim, E. Dall'Anese, and G. Giannakis, "Cooperative spectrum sensing for cognitive radios using Kriged Kalman filtering," *IEEE J. Sel. Topics Signal Process.*, vol. 5, pp. 24–36, Feb. 2011.

[20] S. Atapattu, C. Tellambura, and H. Jiang, "Energy detection based cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 1232–1241, Jan. 2011.

[21] Q. Chen, M. Motani, W.-C. Wong, and A. Nallanathan, "Cooperative spectrum sensing strategies for cognitive radio mesh networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, pp. 56–67, Feb. 2011.

[22] I. F. Akyildiz, W. Y. Lee, and K. R. Chowdhury, "CRAHNs: Cognitive radio ad hoc networks," *Ad Hoc Networks*, vol. 7, pp. 810–836, Jan. 2009.

[23] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas in Commun.*, vol. 23, pp. 201–220, Feb. 2005.

[24] Office of Communications, "Digital dividend: cognitive access," Ofcom, Rep., Feb. 2009.

[25] Electronic Communications Committee, "Technical and operational requirements for the possible operation of cognitive radio systems in the white space of the frequency band 470-790 MHz," CEPT, Rep. 159, Jan. 2011.

[26] A. Bouzegzi, P. Ciblat, and P. Jallon, "Matched filter based algorithm for blind recognition of OFDM systems," in *Proc. IEEE 68th Veh. Technol. Conf.*, Sep. 2008, pp. 1–5.

[27] N. Kamil and X. Yuan, "Detection proposal schemes for spectrum sensing in cognitive radio," *Wireless Sensor Network*, vol. 2, pp. 365–72, May 2010.

[28] P. Sutton, K. Nolan, and L. Doyle, "Cyclostationary signatures in practical cognitive radio applications," *IEEE J. Sel. Areas in Commun.*, vol. 26, pp. 13–24, Jan. 2008.

[29] B. Deepa, A. Iyer, and C. Murthy, "Cyclostationary-based architectures for spectrum sensing in IEEE 802.22 WRAN," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2010, pp. 1–5.

[30] K.-L. Du and W. H. Mow, "Affordable cyclostationarity-based spectrum sensing for cognitive radio with smart antennas," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 1877–1886, May 2010.

[31] J. Shen, S. Liu, Y. Wang, G. Xie, H. Rashvand, and Y. Liu, "Robust energy detection in cognitive radio," *IET Commun.*, vol. 3, no. 6, pp. 1016–1023, June 2009.

[32] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Trans. Commun.*, vol. 55, pp. 21–24, Jan. 2007.

[33] K. Kim, Y. Xin, and S. Rangarajan, "Energy detection based spectrum sensing for cognitive radio: An experimental study," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2010, pp. 1–5.

[34] L. Hu, V. Iversen, and L. Dittmann, "Survey of PHY and LINK layer functions of cognitive radio networks for opportunistic spectrum sharing," *Commun. in Comput. and Inform. Sci.*, vol. 26, pp. 10–24, Jan. 2009.

[35] R. K. J. Liu and B. Wang, *Cognitive Radio Networking and Security.* Cambridge, UK: Cambridge University Press, 2011.

[36] W.-Y. Lee and I. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 3845–3857, Oct. 2008.

[37] Z. Quan, S. Cui, A. Sayed, and V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, pp. 1128–1140, Mar. 2009.

[38] K. Hossain and B. Champagne, "Wideband spectrum sensing for cognitive radios with correlated subband occupancy," *IEEE Signal Process. Lett.*, vol. 18, pp. 35–38, Jan. 2011.

[39] P. Paysarvi-Hoseini and N. Beaulieu, "Optimal wideband spectrum sensing framework for cognitive radio systems," *IEEE Trans. Signal Process.*, vol. 59, pp. 1170–1182, Mar. 2011.

[40] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Topics Signal Process.*, vol. 2, pp. 4–17, Feb. 2008.

[41] A. Walds, *Sequential Analysis*. NY, USA: Wiley, 1947.

[42] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Applications*. NJ, USA: Prentice-Hall, 1993.

[43] H. V. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge, UK: Cambridge University Press, 2009.

[44] K. W. Choi, W. S. Jeon, and D. G. Jeong, "Sequential detection of cyclostationary signal for cognitive radio systems," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 4480–4485, Sep. 2009.

[45] Y. Xin, H. Zhang, and S. Rangarajan, "SSCT: A simple sequential spectrum sensing scheme for cognitive radio," in *Proc. IEEE Global Telecomm. Conf.*, Dec. 2009, pp. 1–6.

[46] G. Feng, W. Chen, and Z. Cao, "A joint PHY-MAC spectrum sensing algorithm exploiting sequential detection," *IEEE Signal Process. Lett.*, vol. 17, pp. 703–706, Aug. 2010.

[47] N. Kundargi and A. Tewfik, "Doubly sequential energy detection for distributed dynamic spectrum access," in *Proc. IEEE Int. Conf. Comm.*, May 2010, pp. 1–5.

[48] ——, "A performance study of novel sequential energy detection methods for spectrum sensing," in *Proc. IEEE Int. Conf. Acoust. Speech and Signal Process.*, Mar. 2010, pp. 3090–3093.

[49] L. Lai, Y. Fan, and V. Poor, "Quickest detection in cognitive radio: A sequential change detection framework," in *Proc. IEEE Global Telecomm. Conf.*, Dec. 2008, pp. 1–5.

[50] H. Li, C. Li, and H. Dai, "Quickest spectrum sensing in cognitive radio," in *Proc. 42nd Annu. Conf. on Inform. Sci. and Syst.*, Mar. 2008, pp. 203–208.

[51] Q. Zhao and J. Ye, "Quickest detection in multiple ON-OFF processes," *IEEE Trans. Signal Process.*, vol. 58, pp. 5994–6006, Dec. 2010.

[52] Z. Chen, Z. Hu, and R. Qiu, "Quickest spectrum detection using hidden markov model for cognitive radio," in *Proc. IEEE Military Comm. Conf.*, Oct. 2009, pp. 1–7.

[53] T. S. Shehata and M. El-Tanany, "A novel adaptive structure of the energy detector applied to cognitive radio networks," in *Proc. 11th Canadian Workshop Inform. Theory*, June 2009, pp. 95–98.

[54] V. Veeravalli, "Decentralized quickest change detection," *IEEE Trans. Inf. Theory*, vol. 47, pp. 1657–1665, May 2001.

[55] A. Vakili and B. Champagne, "An adaptive energy detection technique applied to cognitive radio networks," in *Proc. 22nd IEEE Int. Symp. Personal, Indoor, and Mobile Radio Commun.*, To be published. Sept 2011.

[56] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory.* NJ, USA: Prentice Hall, 1998, vol. 2.

[57] M. C. Fong, D. D. Kee, and N. P. Kaloni, *Advanced Mathematics for Engineering and Science.* Singapore: World Scientific Publishing Co., 2003.

[58] S. M. Kay, *Fundamentals of Statistical Signal Processing: EstimationTheory.* NJ, USA: Prentice Hall, 1993, vol. 1.

[59] J. Font-Segura and X. Wang, "GLRT-based spectrum sensing for cognitive radio with prior information," *IEEE Trans. Comm.*, vol. 58, pp. 2137–2146, July 2010.

[60] R. Zhang, T. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Trans. Commun.*, vol. 58, pp. 84–88, Jan. 2010.

[61] T. J. Lim, R. Zhang, Y. C. Liang, and Y. Zeng, "GLRT-based spectrum sensing for cognitive radio," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2008, pp. 1–5.

[62] S. Kay and J. Gabriel, "An invariance property of the generalized likelihood ratio test," *IEEE Signal Process. Lett.*, vol. 10, pp. 352–355, Dec. 2003.

[63] E. Lehmann, *Testing Statistical Hypotheses*. NY, USA: Willey, 1959.

[64] B. Levy, *Principles of Signal Detection and Parameter Estimation*. NY, USA: Springer, 2008.

[65] N. Morinaga, R. Kohno, and S. Sampei, *Wireless Communication Technologies: New Multimedia Systems*. NY, USA: Kluwer, 2002.

[66] K.-C. Chen and R. Prasad, *Cognitive Radio Networks*. West Sussex, UK: Wiley, 2009.

[67] A. H. Gholamipour, A. Gorcin, H. Celebi, B. U. Toreyin, M. A. R. Saghir, F. Kurdahi, and A. Eltawil, "Reconfigurable filter implementation of a matched-filter based spectrum sensor for cognitive radio systems," in *Proc. IEEE Int. Symp. Circuits and Syst.*, May 2011, pp. 2457–2460.

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

[69] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. Boston, USA: McGraw Hill, 2008.

[70] R. Tandra and A. Sahai, "Fundamental limits on detection in low SNR under noise uncertainty," in *Proc. Int. Conf. Wireless Netw. Commun. and Mobile Computing*, vol. 1, June 2005, pp. 464–469.

[71] E. Hossain and V. Bhargava, *Cognitive Wireless Communication Networks*. NY, USA: Springer, 2007.

[72] H. Harada, H. Fujii, S. Miura, and T. Ohya, "Cyclostationarity-inducing transmission for OFDM-based spectrum sharing systems," in *Proc. Int. Symp. Commun. and Inf. Technol.*, Oct. 2010, pp. 956 –961.

[73] F. H. P. Fitzek and M. D. Katz, *Cognitive Wireless Networks: Concepts, Methodologies and Visions Inspiring the Age of Enlightenment of Wireless Communication*. Dordrecht, Netherlands: Springer, 2007.

[74] J. Chen, A. Gibson, and J. Zafar, "Cyclostationary spectrum detection in cognitive radios," in *Proc. IET Seminar Cognitive Radio and Software Defined Radios: Technologies and Techniques*, Sept. 2008, pp. 1–5.

[75] Q. Yuan, P. Tao, W. Wenbo, and Q. Rongrong, "Cyclostationarity-based spectrum sensing for wideband cognitive radio," in *Proc. Int. Cof. Commun. and Mobile Computing*, vol. 1, Jan. 2009, pp. 107–111.

[76] J. Lunden, V. Koivunen, A. Huttunen, and V. Poor, "Spectrum sensing in cognitive radios based on multiple cyclic frequencies," in *Proc. Int. Conf. Cognitive Radio Oriented Wireless Netw. and Commun.*, Aug. 2007, pp. 37–43.

[77] L. P. Goh, Z. Lei, and F. Chin, "Feature detector for DVB-T signal in multipath fading channel," in *Proc. Int. Conf. Cognitive Radio Oriented Wireless Netw. and Commun.*, Aug. 2007, pp. 234–240.

[78] O. Johnson, *Information Theory and the Central Limit Theorem.* London, UK: Imperial College Press, 2004.

[79] A. J. Gubner, *Probability and Random Processes for Electrical and Computer Engineers.* Cambridge, UK: Cambridge University Press, 2006.

[80] N. Moayeri and H. Guo, "How often and how long should a cognitive radio sense the spectrum?" in *Proc. IEEE Symp. New Frontiers in Dynamic Spectrum*, Apr. 2010, pp. 1–10.

[81] D. Sundararajan, *A Practical Approach to Signals and Systems.* Singapore: Wiley, 2008.

[82] S. Stankovic and, N. Ilic and, M. Stankovic and, and K. Johansson, "Distributed change detection based on a randomized consensus algorithm," in *Proc. 5th European Conf. Circuits and Syst. for Commun.*, Nov. 2010, pp. 51–54.

[83] W. Hines, "A simple monitor of a system with sudden parameter changes," *IEEE Trans. Inf. Theory*, vol. 22, pp. 210–216, Mar. 1976.

[84] H. Solomon and M. A. Stephens, "Distribution of a sum of weighted chi-square variables," *J. of the American Statistical Association*, vol. 72, no. 360, pp. 881–885, Dec. 1977.

[85] K.-H. Yuan and P. Bentler, "Two simple approximation to the distribution of quadratic forms," *British J. of Math. and Statistical Psychology*, vol. 63, no. 2, pp. 273–291, May. 2010.

[86] M. Kamalian and A. Tadaion, "Invariant detection of OFDM signals with unknown parameters for cognitive radio applications," in *Proc. IEEE Int. Conf. Signal Process.*, Oct. 2010, pp. 1507–1511.

[87] N. Han, H. Li, and J. Fang, "GLRT based cooperative spectrum sensing with location information," in *Proc. Asilomar Conf. on Signals, Syst. and Comput.*, Nov. 2010, pp. 496–500.

[88] G. Sanjeev, K. Chaythanya, and C. Murthy, "Bayesian decentralized spectrum sensing in cognitive radio networks," in *Proc. IEEE Int. Conf. Signal Process. and Commun.*, July 2010, pp. 1–5.

[89] E. Axell and E. Larsson, "A bayesian approach to spectrum sensing, denoising and anomaly detection," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process.*, Apr. 2009, pp. 2333–2336.

[90] Q. Zhao, L. Tong, and A. Swami, "Decentralized cognitive MAC for dynamic spectrum access," in *Proc. IEEE Int. Symp. New Frontiers in Dynamic Spectrum Access Netw.*, Nov. 2005, pp. 224–232.

[91] C. Ghosh, C. Cordeiro, D. Agrawal, and M. Rao, "Markov chain existence and hidden markov models in spectrum sensing," in *Proc. IEEE Int. Conf. Pervasive Computing and Commun.*, Mar. 2009, pp. 1–6.

[92] G. B. Dantzig and M. N. Thapa, *Linear Programming 2: Theory and Extensions.* NY, USA: Springer, 2003.

[93] R. Bartoszynski and M. Niewiadomska-Bugaj, *Probability and Statistical Inference.* NJ, USA: Wiley, 2008.

[94] H. Sahai and M. I. Ageel, *Analysis of Variance: Fixed, Random, and Mixed Models.* MA, USA: Birkhauser Boston, 2000.