Spectrum Sensing in Cognitive Radio Networks

by

Sepideh Zarrin

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Electrical and Computer Engineering
University of Toronto

Copyright © 2011 by Sepideh Zarrin
Abstract

Spectrum Sensing in Cognitive Radio Networks

Sepideh Zarrin

Doctor of Philosophy

Graduate Department of Electrical and Computer Engineering

University of Toronto

2011

This thesis investigates different aspects of spectrum sensing in cognitive radio (CR) technology. First a probabilistic inference approach is presented which models the decision fusion in cooperative sensing as a probabilistic inference problem on a factor graph. This approach allows for modeling and accommodating the uncertainties and correlations in the cooperative sensing system.

A constraint in the cognitive radios is the lack of knowledge about the primary signal and channel gain statistics at the secondary users. Therefore, a practical composite hypothesis approach is proposed which does not require any prior knowledge or estimates of these unknown parameters.

Detection delay is an important performance measure in spectrum sensing. Quickest detection aiming to minimize detection delay has been studied in other contexts, and we apply it here to spectrum sensing. To combat the destructive channel conditions such as fading, various cooperative schemes based on the cumulative sum (CUSUM) algorithm are considered in this thesis. Furthermore, cooperative quickest sensing with imperfectly known parameters is investigated and a new solution is derived, which does not require any parameter estimation or iterative algorithm.

In cognitive radios, there is a fundamental trade-off between the achievable throughput by the CRs and the level of protection for the primary user. In this thesis, this trade-off is formulated for the quickest sensing-based CRs. By throughput analysis, it
is shown that for the same protection level to the primary user, the quickest sensing approach results in significantly higher average throughput compared to that of the conventional block sensing approach.
Acknowledgements

First and foremost, I would like to express my sincerest gratitude to my supervisor Professor Teng Joon Lim for his invaluable and insightful guidance throughout the course of my PhD study at the University of Toronto. I am always thankful for his support and all I have learned from his great professional vision and scientific insight. This would not be possible without him.

I wish to thank my thesis committee members, Professor Adve, Professor Valaee, Professor Gulak and Professor Blostein for their invaluable comments. I also thank Diane B. Silva for her the administrative support and kindness through the course my PhD study.

I would not have been here today if it were not for the love and support of my family. I want to express my deepest thanks to my parents for their unconditional love, devotion and support. And my most special thanks to my husband, Saeed, for his love, support, patience, and encouragement through my academic experience and above all, for always being by my side through this journey of my life.
Contents

List of Symbols ix

1 Introduction and Background 1
   1.1 Introduction ......................................................... 1
   1.2 Spectrum Sensing .................................................... 2
      1.2.1 Matched Filtering .............................................. 3
      1.2.2 Cyclostationary Feature Detection ............................ 3
      1.2.3 Likelihood Ratio Test (LRT) ................................. 4
      1.2.4 Energy Detection .............................................. 4
   1.3 Cooperative Spectrum Sensing ...................................... 6
   1.4 Spectrum Sensing with Unknown Parameters ......................... 7
   1.5 Quickest Sensing ................................................... 8
      1.5.1 Cooperative Quickest Sensing ................................. 9
      1.5.2 Quickest Sensing with Unknown Parameters .................. 10
   1.6 Throughput of Cognitive Radio Networks ............................ 11
   1.7 Example Application: Femtocells .................................. 12
   1.8 Thesis Organization and Contributions ............................. 14

2 Probabilistic Inference for Cooperative Sensing 17
   2.1 Introduction ......................................................... 17
   2.2 Model and Notation .................................................. 19
2.2.1 System Model ............................................. 19
2.2.2 Notation .................................................. 21

2.3 Relating Probabilistic Inference to Cooperative Spectrum Sensing .... 21
  2.3.1 Neyman-Pearson and Bayesian Methods ......................... 21
  2.3.2 Probabilistic Inference On a Factor Graph ..................... 23
  2.3.3 Belief Propagation for Cooperative Spectrum Sensing ....... 25

2.4 NP-based LRT at the FC ...................................... 27
  2.4.1 Hard Local Decisions .................................. 28
  2.4.2 Soft Local Decisions .................................. 30
  2.4.3 Quantized Soft Local Decisions ............................ 32

2.5 Simulation Results .......................................... 33

2.6 Chapter Summary ........................................... 38

3 Cooperative Sensing with Unknown Parameters .......................... 40
  3.1 Introduction .............................................. 40
  3.2 Model and Formulation ..................................... 42
    3.2.1 System Model ....................................... 42
    3.2.2 NP-Based LRT at the Fusion Center ....................... 43
  3.3 The Proposed Composite Hypothesis Testing Approach ............. 43
  3.4 The Proposed Test for Cooperative Spectrum Sensing ............ 47
    3.4.1 Unknown Primary Signal and Channel Statistics ............ 47
    3.4.2 Unknown Primary Signal Statistics and Known Channel Statistics 51
    3.4.3 Threshold Setting at the Fusion Center .................... 53
  3.5 Simulation Results ........................................ 56
  3.6 Chapter Summary .......................................... 62

4 Cooperative Quickest Sensing with Unknown Parameters .............. 64
  4.1 System Model ............................................. 67
4.2 Cooperative CUSUM with Full Knowledge of Parameters 68
  4.2.1 Global CUSUM 69
  4.2.2 Global CUSUM with Quantized Local Decision 69
  4.2.3 Hard Fusion of Local CUSUM 71
4.3 Cooperative CUSUM with Unknown Parameters 72
  4.3.1 The GLRT for Quickest Detection 72
  4.3.2 The Rao Test for Quickest Detection Schemes 72
  4.3.3 The Proposed Linear Test for Quickest Detection Schemes 73
4.4 Performance Analysis and Threshold Setting 75
  4.4.1 Threshold Setting for LRT-based CUSUM algorithms 77
  4.4.2 Threshold Setting for Linear-based CUSUM Algorithms 78
4.5 Simulation Results 79
4.6 Chapter Summary 84

5 Throughput of Cognitive Radio Networks 85
  5.1 System Model and Problem Formulation 87
    5.1.1 System Description 87
    5.1.2 Problem Formulation 88
    5.1.3 CUSUM algorithm 89
  5.2 Derivation of $R(T_0)$, $P_c$ and $I_R$ 91
    5.2.1 Quickest Spectrum Sensing 91
    5.2.2 Block-Based Sensing 96
  5.3 Solving the Optimization Problem 99
    5.3.1 Quickest Sensing 99
    5.3.2 Block-Based Sensing 101
  5.4 Simulation Results 102
  5.5 Chapter Summary 107
## List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{H}_0$</td>
<td>Null hypothesis</td>
</tr>
<tr>
<td>$\mathcal{H}_1$</td>
<td>Alternative hypothesis</td>
</tr>
<tr>
<td>$z$</td>
<td>Primary signal</td>
</tr>
<tr>
<td>$x_k$</td>
<td>Observed signal at the $k$-th secondary user</td>
</tr>
<tr>
<td>$h_k$</td>
<td>Channel gain at the $k$-th secondary user</td>
</tr>
<tr>
<td>$n_k$</td>
<td>Additive noise at the $k$-th secondary user</td>
</tr>
<tr>
<td>$u_k$</td>
<td>Observation summary signal at the $k$-th secondary user</td>
</tr>
<tr>
<td>$y_k$</td>
<td>Signal received at the fusion center from the $k$-th secondary user</td>
</tr>
<tr>
<td>$w_k$</td>
<td>Channel noise from the $k$-th secondary user to the fusion center</td>
</tr>
<tr>
<td>$\alpha_k$</td>
<td>Crossover probability of the $k$-th binary symmetric channel</td>
</tr>
<tr>
<td>$\tau_k$</td>
<td>Local threshold at the $k$-th secondary user</td>
</tr>
<tr>
<td>$K$</td>
<td>The number of secondary users</td>
</tr>
<tr>
<td>$N$</td>
<td>The number of sensing samples</td>
</tr>
<tr>
<td>$A$</td>
<td>The maximum transmit voltage</td>
</tr>
<tr>
<td>$M$</td>
<td>Modulation order</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Vector of unknown parameters</td>
</tr>
<tr>
<td>$T_0$</td>
<td>The cognitive radio frame duration</td>
</tr>
<tr>
<td>$T$</td>
<td>Sensing time</td>
</tr>
<tr>
<td>$S$</td>
<td>Collision time</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Summary comparison of spectrum sensing schemes. .................. 5

4.1 Average detection delay in samples for various cooperative quickest sensing schemes. ................................................. 82
List of Figures

1.1 Spectrum utilization measurements [1]. ........................................... 2
1.2 Femtocell network architecture [2]. ............................................... 13

2.1 Block diagram of complete system with $K$ secondary users and one primary
user. The final decision made by the fusion center is denoted by $\hat{\theta}$. ........ 20
2.2 Partial factor graph depicting the joint distribution (2.9). ................. 24
2.3 Complementary ROC ($P_m$ vs. $P_f$) under AWGN SU-FC channels with
$\sigma_n^2 = 1$, $N = 200$ sensing samples, $\sigma_h^2 = 2$ and $\sigma_n^2 = 8$. .............. 33
2.4 Complementary ROC ($P_m$ vs. $P_f$) under binary symmetric SU-FC chan-
nels with $\alpha = 0.1$, $N = 500$ sensing samples, $K = 6$ secondary users,
$\sigma_h^2 = 2$ and $\sigma_n^2 = 8$. ............................................................... 34
2.5 Complementary ROC ($P_m$ vs. $P_f$) under binary symmetric SU-FC chan-
nels with $N = 500$ sensing samples, $K = 6$ secondary users, $\sigma_{h_1}^{2K} = \sigma_{n_1}^{2K}$,
$\sigma_{n_1}^{2K} = \{1.28, 0.32, 4.5, 0.5, 2, 0.08\}$ and $\alpha_1^K = \{0.2, 0.3, 0.1, 0.05, 0.2, 0.4\}$. .. 35
2.6 Complementary ROC ($P_m$ vs. $P_f$) under soft local decisions, with $N = 100$
sensing samples, $K = 6$ secondary users, $\sigma_n^2 = 1$, $\sigma_h^2 = 2$ and $\sigma_n^2 = 4$. ... 36
2.7 Complementary ROC ($P_m$ vs. $P_f$) under quantized local decisions, with
$N = 500$ sensing samples, $K = 6$ secondary users, $\sigma_{h_1}^{2K} = \sigma_{n_1}^{2K}$, $\sigma_{n_1}^{2K} = \{1.28, 0.32, 4.5, 0.5, 2, 0.08\}$, and $\alpha_1^K = \{0.1, 0.3, 0.2, 0.05, 0.1, 0.2\}$. .... 37
3.1 Block diagram of cooperative spectrum sensing system with $K$ secondary users. ................................................................. 42
3.2 Complementary ROC ($P_m$ vs. $P_f$) curves for hard local decisions with unknown primary statistics and $K = 3$. .................. 57
3.3 Complementary ROC ($P_m$ vs. $P_f$) curves for hard local decisions with unknown primary statistics and $K = 3$. .................. 58
3.4 Complementary ROC ($P_m$ vs. $P_f$) curves for hard local decisions with unknown primary statistics and $K = 6$. .............. 59
3.5 Complementary ROC ($P_m$ vs. $P_f$) curves for soft local decisions and $K = 6$. ................................................................. 60
3.6 Complementary ROC ($P_m$ vs. $P_f$) curves for 2-bit quantized local decisions with unknown primary statistics. .................. 61
3.7 Complementary ROC ($P_m$ vs. $P_f$) curves for 2-bit quantized local decisions with unknown primary and channel gain statistics. ........ 62
3.8 Probability distribution function of the linear test statistic under the null hypothesis................................................................. 63
4.1 Probability of false alarm, $P_f$, versus average detection delay $T_d$. ...... 79
4.2 Probability of false alarm, $P_f$, versus average detection delay $T_d$. ...... 80
4.3 Probability of false alarm, $P_f$, versus threshold $q$. ....................... 83
5.1 A CR frame, showing the sensing interval in black and the data transmission phase in a lighter shade. The primary user is inactive at the start of the frame (defined as event $H_0$ in the text), and comes on at time $T + X$ from the start of the frame, leading to collision over a duration $S = T_0 - T - X$. ................................................................. 88
5.2 Average normalized throughput, $R_Q / C_0$, and probability of collision, $P_{cQ}$, versus frame duration $T_0$ for quickest sensing. ............ 103
5.3 Average normalized throughput, $R_Q/C_0$, and primary interference time ratio (PITR), $I_{R_Q}$, versus frame duration $T_0$ for quickest sensing . . . . . . 104

5.4 Average normalized CR throughput, versus probability of collision . . . 105

5.5 Average normalized CR throughput, versus probability of collision . . . 106
Chapter 1

Introduction and Background

1.1 Introduction

Studies have shown that most of the licensed radio-wave spectral bands are under-utilized in time and space domain [1,3,4], resulting in unused “white spaces” in the time-frequency grid at any particular location. The spectrum utilization is mainly around certain parts of the spectrum whereas a considerable amount of the spectrum is unutilized as depicted in Figure 1.1. As can be observed, spectrum utilization is more intense and competitive at frequencies below 3 GHz whereas the spectrum is under-utilized in the 3-6 GHz bands [1]. The Federal Communications Commission (FCC) has also reported the temporal and geographic variations in spectrum utilization to range from 15% to 85% [3]. On the other hand, fixed spectrum allocation policies do not allow for reusing of the rarely used spectrum allocated to licensed users by unlicensed users. This problem coupled with the rapidly increasing demand for wireless services and radio spectrum has led to spectrum scarcity for wireless applications.

This has necessitated a new communication standard that allows unlicensed (secondary) users to utilize the vacant bands which are allocated to licensed (primary) users. However, this opportunistic access should be in a manner that does not interrupt any
primary process in the band. Therefore, the secondary users must be aware of the activity of the primary user in the target band. They should spot the spectrum holes and the idle state of the primary users in order to exploit the free bands and also promptly vacate the band as soon as the primary user becomes active. Cognitive radio, encompasses this awareness by dynamically interacting with the environment and altering the operating parameters with the mission of exploiting the unused spectrum without interfering with the primary users [5, 6]. Showing support for the cognitive radio idea, the FCC allowed for usage of the unused television spectrum by unlicensed users wherever the spectrum is free [6,7]. IEEE has also supported the cognitive radio paradigm by developing the IEEE 802.22 standard for wireless regional area network (WRAN) which works in unused TV channels [8].

1.2 Spectrum Sensing

In order to maintain the primary users’ right to interference-free operation, the secondary users need to regularly sense the allocated band and reliably detect the presence of the
primary users’ signals with little delay. In the IEEE 802.22 standard, for example, the secondary users need to detect the TV and wireless microphone signals and upon their detection, they are required to vacate the channel within two seconds [8]. For TV primary signals, a probability of detection of 90% and a probability of false alarm of 10% should be maintained [8]. Therefore, spectrum sensing plays a crucial role in the cognitive radio technology to prevent damaging interference to the primary users and to reliably and quickly spot the white spaces in the spectrum and utilize the opportunity.

Various spectrum sensing methods are used in literature depending on how much information about the primary signal is available to the secondary users, as discussed in the following.

1.2.1 Matched Filtering

Matched filtering-based methods are optimal for stationary Gaussian noise scenarios as they maximize the received SNR [9, 10]. For this optimal performance, they require perfect knowledge of the channel responses from the primary user to the secondary user and the structure and waveforms of the primary signal (including modulation type, frame format and pulse shape) as well as accurate synchronization at the secondary user [9–12]. In cognitive radios, however, such knowledge is not readily available to secondary users and implementation cost and complexity of this detector is high especially as the number of primary bands increases. Therefore, this method is not practical and applicable to cognitive radio technology.

1.2.2 Cyclostationary Feature Detection

Another detection method that can be applied for spectrum sensing is the cyclostationary feature detector. Cyclostationary feature detectors can distinguish between modulated signals and noise [9–14]. This detector exploits the fact that the primary modulated signals are cyclostationary with spectral correlation due to the built-in redundancy of
signal periodicity (e.g., sine wave carriers, pulse trains, and cyclic prefixes), while the noise is a wide-sense stationary signal with no correlation [13, 14]. This task can be performed by analyzing a spectral correlation function. Therefore, cyclostationary feature detectors are robust to the uncertainty in noise power [9–14]. This is at the price of excessive computational complexity and long observation times. Moreover, it requires the knowledge of the cyclic frequencies of the primary users, which may not be available to the secondary users.

1.2.3 Likelihood Ratio Test (LRT)

Spectrum sensing is a binary hypothesis testing problem, with the null and alternative hypotheses

\[ H_0 : \text{Primary user not active} \]
\[ H_1 : \text{Primary user active} \]  \hspace{1cm} (1.1)

Based on the Neyman-Pearson (NP) theorem [15], the test statistic that maximizes the probability of detection for a given probability of false alarm is the likelihood ratio test (LRT) defined as

\[ L(X) = \frac{p(X|H_1)}{p(X|H_0)} \]  \hspace{1cm} (1.2)

where \( X \) denotes the received signal vector and \( p(·) \) denotes the probability density function (PDF).

The LRT, which is proven to be NP optimal [15], requires the exact distributions of primary signal and noise and channel gains which makes it practically intractable.

1.2.4 Energy Detection

For a Gaussian noise model, when the noise power is known to the secondary user energy detection can be applied to detect the existence of the primary signal. This simple scheme accumulates the energy of the received signal during the sensing interval and declares the band to be occupied if the energy surpasses a certain threshold. This threshold is set
### Spectrum Sensing Scheme

<table>
<thead>
<tr>
<th>Spectrum Sensing Scheme</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matched Filtering</td>
<td>Optimum performance</td>
<td>Requires full primary signal knowledge, high power consumption and implementation complexity</td>
</tr>
<tr>
<td>Cyclo Feature Detection</td>
<td>Robust to interference and noise uncertainty</td>
<td>High computational complexity, vulnerable to sampling clock offsets and model uncertainties, long observation time</td>
</tr>
<tr>
<td>LRT detection</td>
<td>NP-optimal performance</td>
<td>Requires exact distributions and parameters</td>
</tr>
<tr>
<td>Energy Detection</td>
<td>Low complexity, no primary knowledge required</td>
<td>Vulnerable to noise uncertainty</td>
</tr>
</tbody>
</table>

Table 1.1: Summary comparison of spectrum sensing schemes.
based on the desired probability of false alarm [15,16]. Energy detection, unlike the other schemes, does not require any information about the primary signal and channel gains and is robust to unknown fading channel. Compared to other methods it has simpler implementation and hence is less expensive. Therefore, in literature energy detection is mainly adopted for spectrum sensing [9–12,17–20].

Table 1.1 presents a brief comparison of the above spectrum sensing schemes.

1.3 Cooperative Spectrum Sensing

In practice, the reliability of spectrum sensing in detecting weak primary signals is hard to maintain due to the destructive channel conditions such as multi-path fading or shadowing between the primary user and the secondary users. As different users in distant locations undergo independent fading channels, allowing cooperation among them can improve the overall detection reliability by utilizing the inherent diversity [9,10,17–19,21–24]. In this manner, the individual secondary users require less detection sensitivity whereas the overall detection reliability is improved. This means having more accurate and agile detection, fewer false alarms and more protection to the primary user [12,20,23–25]. The problem of hidden terminals can also be alleviated through cooperation among fairly distant secondary users [12,20,23,25]. These advantages come at the price of added communication overhead due to cooperation. The FCC has also acknowledged the need for cooperative spectrum sensing for TV bands [26,27].

In a widely studied form of cooperative spectrum sensing, the secondary users provide locally-sensed information on the primary user’s activity status to a decision-making fusion center (FC) which can be an access point or base station or another secondary user [9,10,17–19,21–24]. The fusion center analyzes the information and determines the activity status of the primary user.

The cooperative sensing schemes can be mainly categorized based on the type of fusion
method used at the fusion center. Hard-decision combining schemes such as AND/OR and $M$-out-of-$K$ rules are considered in [17, 18] where their performances are compared. It is shown that when one node has much higher SNR compared to other nodes then cooperative spectrum sensing performs worse than the single node spectrum sensing. Soft decision combining rules such as linear combination of local energies [19, 22] and LRT approach [17, 21] are also considered in the literature. In LRT-based schemes, the local decision statistics are either assumed to be available [17], or estimated [21] at the fusion center. The summation of the secondary users’ received signal energies has been considered as a soft decision based cooperative sensing method in [19]. A cooperative sensing scheme based on linear combination of the local test statistics was proposed in [22] where the combining weights were optimized to improve the detection performance.

In the above studies, the SU-FC channels were assumed to be error- or noise-free. There was also no mention of how to make use of knowledge about the modulation format used at the primary transmitter, or correlations in time or space of the channels linking the primary transmitter to the secondary users. For a more rigorous study, we propose a probabilistic inference approach for cooperative spectrum sensing in Chapter 2.

### 1.4 Spectrum Sensing with Unknown Parameters

An important constraint in cognitive radios is the lack of knowledge about the primary signal and channel statistics at the secondary users and the fusion center. Therefore the LRT-based methods which perform better than other approaches cannot be applied in cognitive radios due to unknown parameters and uncertainties in likelihood functions. These unknown parameters in cognitive radios, which are mainly the statistics of the primary signal and channel gains, appear in the likelihood function under $\mathcal{H}_1$. We denote them here by $\beta$. In detection theory, hypothesis testing in the presence of uncertain
parameters is known as composite hypothesis testing [15]. The main approach to tackle this problem is the generalized likelihood ratio test (GLRT) [15]. This test first finds the maximum likelihood estimate (MLE) of the unknown parameters under $H_1$, and then forms the GLRT statistic:

$$L_G(X) = \frac{p(X|\beta = \hat{\beta}, H_1)}{p(X|H_0)} \quad (1.3)$$

where $\hat{\beta} = \arg\max_\beta p(X|\beta, H_1)$ and $\beta = 0$ under $H_0$.

An asymptotically (as the number of observations increase) equivalent test for GLRT is the Wald test with statistic

$$T_W(X) = \hat{\beta}^T I(\hat{\beta})^{-1} \hat{\beta} \quad (1.4)$$

where $I(\cdot)$ is the Fisher information matrix (see [28] pg. 40). The Wald test requires the derivation of the MLEs. However, the analytical derivation of the MLEs is complicated and even intractable for a general cooperative noisy channel model.

Another asymptotic equivalent test for GLRT is the Rao test which does not require the MLE of the unknown parameters and is simpler computationally [15]. The Rao test statistic is given by

$$T_R(X) = \frac{\partial \ln p(X|\beta)}{\partial \beta}^T |_{\beta=0} \frac{\partial \ln p(X|\beta)}{\partial \beta} |_{\beta=0}^{-1} \frac{\partial \ln p(X|\beta)}{\partial \beta} |_{\beta=0} \quad (1.5)$$

The Rao test, in fact, is not practical as the Fisher information matrix is not easily attainable or invertible in many system models. Dealing with multiple unknown parameters in cooperative spectrum sensing, even further adds to the complexity and infeasibility of calculating this inverse matrix. In Chapter 3, we propose a practical composite hypothesis testing approach which estimates the unknown parameters, and further simplify it so that it does not even require these estimates.

## 1.5 Quickest Sensing

The cooperative spectrum sensing approaches that we discussed so far had a classic detection framework, in which the goal is to minimize the miss-detection probability
subject to constraints on the false alarm probability. Their detection approach is block-based, where the secondary users take a block of samples, and decide on the activity state of the primary user based on these samples. However, other than the probability of detection, detection delay is also crucial to cognitive radios in order to act quickly upon vacancy or occupancy of the primary band. The vacancy of the primary band should be sensed quickly so that the secondary transmission starts right away and the opportunity is utilized efficiently. On the other hand, upon the activity of the primary user the secondary user should be able to sense and act quickly and vacate the band promptly in order not to disrupt the primary transmission. Therefore a detection framework is required which allows for analysis and minimization of the detection delay for a certain level of false alarm [29].

In essence, the spectrum sensing is detection of change in spectrum activity (emergence of the primary radio and/or vacancy of the band). Therefore, we can apply the theory of sequential change detection, which performs a statistical test to detect the change of distribution in observations as quickly as possible, in order to attain an agile and robust spectrum sensing [30]. The adaptation of this approach to spectrum sensing is known as quickest sensing in literature [30–32]. In the sequential change detection literature, the well-known Page’s cumulative sum (CUSUM) algorithm [33] has been shown to be optimal in the sense of minimizing the detection delay while maintaining an acceptable level of false alarm [34–36]. In [30–32], the CUSUM algorithm has been applied to quickest spectrum sensing in cognitive radios.

1.5.1 Cooperative Quickest Sensing

Due to fading and path loss effects in cognitive radio networks, single user sensing is proven to be unreliable and the detection performance improves with the number of available sensors that make independent observations [32,37,38]. The study of distributed quickest detection has been based on two different formulations in the detection literature.
One is based on a Bayesian formulation in which the change point (time at which the change occurs) is assumed to have a known prior distribution [29, 39, 40]. In this case, the joint optimization of local users and the fusion center policies over time becomes hard or intractable [40]. However, a tractable test design is developed for a geometric a priori distribution for the change time in [29, 39]. The other formulation is the minimax formulation, proposed by Lorden [34], for which Pages CUSUM procedure is the optimal scheme. The asymptotically optimal cooperative CUSUM-based approaches for various local memory and transmission rate scenarios have been studied by Mei [37]. We apply and compare three CUSUM-based cooperative spectrum sensing schemes in cognitive radios with noisy channels: global CUSUM with soft local decisions, global CUSUM with quantized local decisions and hard fusion of local CUSUM to cognitive radios as described in Chapter 4.

1.5.2 Quickest Sensing with Unknown Parameters

In practice, when the primary user starts transmission there are unknown parameters in the distribution of the observed signals, e.g. the variance of the observed signals, due to unknown primary signal statistics and channel gains between the primary and secondary users. Therefore, the CUSUM-based approaches that are based on perfectly known distributions cannot be directly applied to spectrum sensing in cognitive radios. Therefore, quickest detection schemes that can accommodate unknown parameters should be developed.

The GLRT was adopted by Lorden [34] for CUSUM with unknown parameters. However, due to the non-recursive expression of the GLRT-based CUSUM and the need to store all the observations and re-estimate the unknown parameters in all time slots, this algorithm turns out to be infeasible and impractical. The parallel CUSUM test [41] requires a complicated recursive algorithm and does not achieve the optimal performance. The successive refinement algorithm [31] which performs estimation while detecting em-
beds incremental estimation in the parallel CUSUM test for the unknown parameter and is suitable only for scenarios with unknown parameters that can be estimated incrementally. Otherwise, it needs to store all samples like the GLRT algorithm. In Chapter 4 we introduce our linear method which does not require any complicated algorithm or parameter estimation and accommodates multiple unknown parameters.

1.6 Throughput of Cognitive Radio Networks

Associated with spectrum sensing are two parameters: probability of detection and probability of false alarm. The higher the detection probability, the better the primary users can be protected. However, from the secondary users perspective, the lower the false alarm probability, the more chances the channel can be reused when it is available, thus the higher the achievable throughput for the secondary users. Thus there exists a fundamental tradeoff between sensing capability and achievable throughput for the secondary network.

In block sensing schemes\(^1\), the sensing time can be chosen to achieve the desired throughput for the CRs while satisfying the regularity constraints for primary user’s protection from interference. In [42], the optimum sensing time has been formulated under both presence and absence of the primary signal. Upon finding an available channel, the optimization problem is to find the optimum in-band sensing time such that the maximum average throughput is achieved. The sensing-throughput trade-off is also considered in [43] where the CR’s average throughput is maximized for a certain level of certainty in detecting the primary signal.

In [43], the block sensing duration was designed to maximize CR throughput with a given frame duration, while protecting the primary user by keeping the probability of detection higher than a certain threshold. Similar problem of optimizing the frame

\(^{1}\text{Where the sensing phase is fixed in duration.}\)
duration for a given sensing time has been considered in [44, 45]. In [44], the objective of the optimization was to maximize the CR’s throughput while keeping the probability of collision with the primary user below a certain threshold, assuming that primary user activity in successive frames are independent of each other. In [45], the throughput of the CR was maximized while ensuring that the primary interference time ratio (PITR) is less than a certain threshold. Unlike [44], it assumed dependent primary signals in successive frames by modeling the primary user activity as a continuous time Markov process. Both [44] and [45] assume perfect error-free block sensing which is not realistic and they neglect the effect of interference from the primary user on the CR’s throughput.

While in all of the above throughput analysis only block sensing approach is considered, in Chapter 5, we investigate the sensing-throughput tradeoff for quickest sensing-based CR networks. We analyze the average CR throughput for both quickest and block sensing approaches, and unlike the previous work, we consider a drop in the CR throughput due to interference from the primary user.

1.7 Example Application: Femtocells

A practical application and supporting infrastructure for cognitive radio networks is the femtocell network [46]. Femtocells are home base stations which improve indoor coverage and service quality in cellular networks [2, 46]. These low-power access points are installed at homes or small public places and by improving the spectral efficiency and indoor coverage they reduce the overall and operational costs [46]. Figure 1.2 depicts an example femtocell network. The femtocells backhauls data to the cellular operator network or the Internet backbone through a broadband connection such as DSL, TV cable, fiber, power line or Ethernet [46, 47]. The existing wired backhaul infrastructure makes femtocells an ideal candidate for cognitive radio implementation by allowing cooperation among users. Spectrum sensing can be performed by individual femtocells and the sensing results can
be shared by other users. By utilizing the existing infrastructure of femtocell networks for CRs, the issues in deployment, hardware, security, and network management of CRs can be tackled easily and cost efficiently.

In [46], the conventional femtocell concept was extended to enable cognitive operation. The femtocell base station (FBS), provides backhauling service via a wired broadband IP connection to the users in its coverage region. The FBS performs local sensing to find the unused spectrum by the primary users in its region and the neighboring femtocells and to control interference to the underlaying macrocell and nearby femtocells. The only change required to the current handsets is a software update to support the cognitive mode and provide application level tools and protocol modifications.

In order to improve sensing reliability and decrease interference to nearby femtocells and underlying macrocell, cooperative sensing can be performed in femtocell networks. As mentioned before, the existing wired backhaul in femtocells is a favorable infrastructure for cooperative spectrum sensing among femtocells. Cooperative sensing among nearby FBSs improves the sensing reliability and increases the sensing domain. Moreover, this could alleviate interference among femtocells and between femtocells and the
underlying macrocell which results in higher system capacity.

1.8 Thesis Organization and Contributions

In Chapter 2 of this thesis, the probabilistic approach for cooperative spectrum sensing is presented. The main contributions of this chapter are outlined as follows.

- A probabilistic inference approach for cooperative spectrum sensing is proposed. The cooperative sensing system is probabilistically modeled on a representative factor graph, and the decision fusion problem is approached as a probabilistic inference problem on a factor graph that can be tackled by message passing algorithms like belief propagation.

- The proposed approach allows for modeling and accommodating the uncertainties and correlations in the cooperative sensing system.

- Unlike most studies in this field, nonideal transmission channels between secondary users and fusion center as well as the presence of fading in links between primary and secondary users are considered.

- The likelihood ratio test statistics are derived for one-bit, multi-bit or infinite precision of local decisions.

In Chapter 3, the spectrum sensing problem with unknown parameters is investigated. The main contributions of this chapter are as follows.

- A novel composite hypothesis testing approach for spectrum sensing with unknown parameters is proposed which does not require any prior knowledge about the primary signal and channel statistics.

- A simple spectrum sensing scheme is proposed for cognitive radios with unknown parameters, which does not require any parameter estimation or iterative algorithm.
The proposed test has been applied to cooperative spectrum sensing with hard, soft and quantized local decisions and the corresponding test statistics are derived under the unknown primary signal and channel statistics scenario.

The locally most powerful (LMP) detector has been adopted for decision making at the fusion center for weak constant-modulus primary signals and its corresponding test statistic has been derived.

An analytical threshold setting scheme is presented for the proposed test as well as the LMP.

Performance comparison between the proposed test and LRT is provided. The results illustrate the close performance of the proposed test to the optimal LRT.

In Chapter 4, cooperative quickest spectrum sensing is investigated. The main contributions of this chapter are as follows.

Several cooperative quickest sensing schemes based on the CUSUM algorithm are presented for cognitive radios and their test statistics under noisy and noiseless channel scenarios are derived.

A linear test for cooperative spectrum sensing with unknown parameters is presented, which does not require any prior knowledge or estimation of the primary signal or channel statistics. This test is very simple and does not entail any iterative or complicated algorithms.

A threshold setting method for CUSUM-based algorithms and the proposed test is presented for a target probability of false alarm.

Chapter 5 presents the throughput analysis of cognitive radio networks. The main contributions of this chapter are outlined as follows.

The throughput-sensing tradeoff is formulated and analyzed for quickest sensing.
• Analytical expressions for the average throughput, probability of collision and PITR (primary interfered time ratio) for both quickest and block-based sensing approaches are derived for the realistic assumption that collision with PUs causes a drop in throughput for the CR.

• The optimal frame durations of quickest and block-based sensing CRs to achieve the maximum CR throughput are found, while protecting the primary user by keeping the collision probability or PITR below a certain threshold.

• The maximum achievable throughput with block and quickest sensing schemes are compared and shown that for the same protection level to the primary network, the quickest sensing approach results in significantly higher average throughput.

Finally, the thesis is concluded in Chapter 6 by presenting some final remarks and future research directions.
Chapter 2

Probabilistic Inference Approach for Cooperative Sensing

2.1 Introduction

Cooperative spectrum sensing has been introduced to address the problem of detection reliability of weak primary signals under fading and hidden terminal problem conditions [17–19, 21, 22]. In a widely studied form of cooperative spectrum sensing, the secondary users provide locally-sensed information on the primary user’s activity status to a decision-making fusion center (FC) which can be an access point or base station. Intuitively, the resulting diversity improves detection performance by deciding based on multiple observations of the same signal. The problem of hidden terminals can also be alleviated by cooperation among distant secondary users from each other. Cooperative sensing is helpful when the cooperating users are close enough together that the same primary activity level applies to them all, but still far enough apart that the channels are independent.

The cooperative spectrum sensing schemes studied in the literature are mostly based on either hard-decision combining schemes such as logical AND/OR and $M$-out-of-$K$
rules [17,18], or soft decision combining rules such as linear combination of local energies [22] or LRT approaches [17,21]. In these studies, the SU-FC channels were assumed to be error and noise free, which is unrealistic especially for the infinite-precision energy-combining methods, in which the FC is assumed to know the energy measured at the secondary users precisely.

In this chapter, we propose a probabilistic inference approach for cooperative spectrum sensing – we probabilistically model the cooperative sensing system with a factor graph, and approach the decision fusion or detection problem as a probabilistic inference problem on a factor graph that can be tackled by belief propagation (BP) [48]. The likelihood ratio test at the FC, based on Neyman-Pearson (NP) or Bayesian criteria [15], is then straightforward to compute. This approach can accommodate the uncertainties and correlations in cooperative sensing and provide fresh insight into statistical cooperative spectrum sensing. For instance, with the proposed method we are able to consider Rayleigh fading PU-SU channels as well as noisy SU-FC channels for added realism compared to the existing studies in the CR literature.

Using the proposed approach, we derive the optimal NP-based LRT at the FC for hard, soft and quantized local decisions. For the hard decision scenario, we consider two different models for the SU-PU channels: binary symmetric channels (BSCs) and additive white Gaussian noise (AWGN) channels. In the BSC scenario, we show that under the same channel conditions and local thresholds for all the secondary users, the NP-based LRT is equivalent to the $M$-out-of-$K$ rule where the licensed band is declared to be occupied if at least $M$ out of the $K$ secondary users have detected the primary signal. Under the AWGN channel scenario, we show that the LRT performs better than the $M$-out-of-$K$, AND and OR fusion rules. For the soft local decisions scenario where the secondary users send their local test statistics to the FC, we consider AWGN SU-FC channels and derive the LRT at the FC. We show that the LRT outperforms the adding, maximal ratio combining and selection combining methods.
Finally, we discuss the transmission of multi-bit quantized soft local decisions from each secondary user to the FC. Given that the received energy is the sufficient test statistic for Gaussian channels [49], we consider energy quantization at the secondary users and derive the LRT at the FC using the BP approach. We show that by sending only two bits of information on the signal energy from all the secondary users, the performance of the NP-based LRT improves significantly compared to the case of hard (one-bit) local decisions. The results of this chapter have been published partly in [50].

The remainder of this chapter is organized as follows. The system model and notations are described in Section 2.2. The relation between detection and probabilistic inference and the proposed graphical model approach are presented in Section 2.3. The derivation of NP-based LRT at the FC for hard, soft and quantized local decisions is provided in Section 2.4. The simulation results and performance comparisons are provided in Section 2.5. Finally, Section 2.6 concludes the chapter.

2.2 Model and Notation

2.2.1 System Model

We consider a centrally coordinated cognitive radio network with $K$ secondary users. We denote the baseband-equivalent signal transmitted by the primary user over a sensing interval of $N$ samples by $\tilde{z}$, where the underscore denotes a vector of $N$ samples in time. This signal is propagated to the $k$-th secondary user over a frequency non-selective channel that is time invariant over $N$ sampling intervals\footnote{Here we prefer to present a simple channel model to illustrate the technique proposed. Nevertheless, time-varying channels can also be included in the model.}. The observed complex baseband-equivalent signal at the $k$-th secondary user over $N$ sampling intervals is

\[
\tilde{x}_k = h_k \tilde{z} + n_k
\]
where $n_k$ denotes the zero-mean additive white Gaussian noise (AWGN) i.e., $n_k \sim \mathcal{CN}(0, \sigma^2_{n_k})$ and $h_k$ represents the circularly symmetric complex Gaussian (CSCG) channel gain i.e. $h_k \sim \mathcal{CN}(0, \sigma^2_{h_k})$. Moreover, $\tilde{z}$, $h_k$ and $n_k$ are assumed to be independent, which is reasonable from a practical perspective.

The observed signal at the $k$-th secondary user is mapped onto an observation summary signal $u_k$ over $N$ sampling intervals, i.e.,

$$u_k = \gamma_k(z_k)$$  \hspace{1cm} (2.2)

In this chapter, we assume three different scenarios for the local mappings at the secondary users; hard-decision cooperation, where $u_k$ is the output of energy detection at the $k$-th secondary user, soft-decision cooperation where $u_k$ is the energy of $x_k$, and quantized-decision cooperation where $u_k$ is the quantized energy of $x_k$.

![Figure 2.1: Block diagram of complete system with $K$ secondary users and one primary user. The final decision made by the fusion center is denoted by $\hat{\theta}$.](image)

As shown in Figure 2.1, $u_k$ is subsequently transmitted to the fusion center where the final spectrum occupancy decision is made. We assume the channel between each secondary user and the FC to be noisy and characterized by $p(y_k|u_k)$, where $y_k$ denotes the
signal received at the FC from \( k \)-th secondary user. This characterization is general and different models can be considered for the SU-FC channels. In this chapter, we consider two different models for the SU-FC channels: binary symmetric channels with crossover probabilities \( \alpha_k \)s, as well as AWGN channels with additive noise \( w_k \sim \mathcal{CN}(0, \sigma^2_{w_k}) \). In the latter case, \( p(y_k|u_k) = p_w(y_k - u_k) = \frac{1}{\pi \sigma^2_{w_k}} \exp\left(-\frac{|y_k - u_k|^2}{\sigma^2_{w_k}}\right) \). The evidence available to the FC to make the global decision is the set of SU-FC channel outputs \( y^K_1 = \{y_1, \ldots, y_K\} \).

### 2.2.2 Notation

We define the following notations which are used throughout this chapter. \( v^K_1 \) denotes the set \( \{v_1, \ldots, v_K\} \) which can also be collected into the vector form \([v_1, \ldots, v_K]\). The underscored variable \( v \) denotes vector of \( N \) samples in time. There is no distinction made between random variables and their values, e.g. \( p(z) \) is the probability density function of the random variable \( z \), except in the factor graph in Figure 2.2 and discussions referring to Figure 2.2, where capital letters are used to denote random variables by convention.

### 2.3 Relating Probabilistic Inference to Cooperative Spectrum Sensing

#### 2.3.1 Neyman-Pearson and Bayesian Methods

Spectrum sensing is a binary hypothesis testing problem, with the null and alternative hypotheses

\[
\mathcal{H}_0 : \text{Primary user not active : } \underline{z} = 0, \\
\mathcal{H}_1 : \text{Primary user active : } \underline{z} \neq 0.
\]

There are two well-known strategies for solving a binary hypothesis testing problem; Baysian method and Neyman-Pearson (NP) method [15]. The Bayesian method is based on the minimization of a Bayesian risk function which, under equal costs of a false
alarm and missed detection, reduces to the minimization of the probability of error. By introducing \( \theta \in \{0, 1\} \) as indices to the null and alternative hypotheses, the probability of error can be represented by \( P(\hat{\theta} \neq \theta) \), where \( \hat{\theta} \) is the decision on \( \theta \). The NP method is based on the maximization of the probability of detection \( P(\hat{\theta} = 1|\theta = 1) \), while maintaining a given probability of false alarm \( P(\hat{\theta} = 1|\theta = 0) \).

The optimal decision that minimizes the probability of error is the maximum a posteriori (MAP) decision [15]

\[
\hat{\theta} = \arg \max_{\theta \in \{0, 1\}} P(H_\theta | y^K_1).
\] (2.4)

If no prior information on \( H_\theta \) is available, the MAP decision is equivalent to the maximum likelihood (ML) decision which is equivalent to choosing \( H_1 \) if and only if

\[
\frac{p(y^K_1 | z_0 \neq 0)}{p(y^K_1 | z_0 = 0)} > \frac{1}{2}.
\] (2.5)

The optimal NP method similarly employs a likelihood ratio test (LRT) [15] by choosing \( H_1 \) if and only if

\[
L(y^K_1) = \frac{p(y^K_1 | z_0 \neq 0)}{p(y^K_1 | z_0 = 0)} > \lambda,
\] (2.6)

where the threshold \( \lambda \) is found by solving (numerically or analytically)

\[
P_f = \int_{y^K_1 : L(y^K_1) > \lambda} p(y^K_1 | z_0 = 0) dy^K_1
\] (2.7)

for a given desired probability of false alarm \( P_f \).

In both (2.5) and (2.6), we need to evaluate the likelihood function \( p(y^K_1 | z) \) for \( z = 0 \) and otherwise. From Bayes’ rule, we have

\[
p(y^K_1 | z) \propto \frac{p(z | y^K_1)}{p(z)}
\] (2.8)

or in other words, the likelihood can be found from the a posteriori probability (APP) distribution of \( z \). The detection problem can therefore be decomposed into a Bayesian inference problem, namely finding \( P(z = 0 | y^K_1) \) and \( P(z \neq 0 | y^K_1) \).
2.3.2 Probabilistic Inference On a Factor Graph

As discussed above, the cooperative spectrum sensing problem can be reduced to finding the APP for null and alternative hypotheses. The APP $p(z|y^K_1)$ can be obtained by marginalizing over all other random variables in the model. For the model presented in Section 2.2, the joint distribution is

$$p(z, u^K_1, x^K_1, h^K_1|y^K_1),$$

and the required APP is

$$p(z|y^K_1) = \int p(z, u^K_1, x^K_1, h^K_1|y^K_1)du^K_1 dx^K_1 dh^K_1.$$  \hspace{1cm} (2.10)

For factorizable joint probability distributions, with each factor being a function of only a small subset of the total set of variables, marginalization may be efficiently performed using the so-called belief propagation\(^2\) with the aid of a factor graph [48]. A factor graph is a bipartite graph that represents the factorization of a function. A factor graph consists of variable nodes, factor nodes and edges. For a certain factorization of a function, the edges connect the factor nodes to the variable nodes that they are a function of [51]. BP obtains exact results for loop-free or tree graphs, and when the graph is “locally” tree-like but not actually a tree, very accurate results arise in iterative BP. In the following, we will demonstrate that the factor graph for (2.9) is, in fact, a tree (if channel coefficients $h_j$ and $h_k$ are independent for $j \neq k$). Thus, BP can be used to obtain a cooperative sensing algorithm that is based on firm probabilistic foundations.

From the conditional probability chain rule, we have

$$p(z, u^K_1, x^K_1, h^K_1|y^K_1) = p(u^K_1|y^K_1)p(x^K_1|u^K_1, y^K_1)p(h^K_1|x^K_1, u^K_1, y^K_1)$$

\(^2\)also known as message passing, and the sum-product algorithm
Examining each term on the right-hand side in turn,

\[ p(u^K_k | y^K_1) \propto \prod_k p(y_k | u_k) \]  
(2.11)  
\[ p(x^K_1 | u^K_1, y^K_1) = \prod_k I(u_k = \gamma_k(x_k)) \]  
(2.12)  
\[ p(h^K_1, z | x^K_1, u^K_1, y^K_1) = p(h^K_1, z | x^K_1) \]  
\[ \propto p(z) \prod_k p(x_k | h_k, z) p(h_k) \]  
(2.13)  

where \( I(A) = 1 \) if event \( A \) is true, and \( I(A) = 0 \) otherwise. The first equation arises because we assume no prior on \( u_k \) and also independent SU-FC channels; the second comes from the fact that the secondary users operate independently and the mapping from \( z^K_1 \) to \( u^K_1 \) consists of \( K \) independent mappings; and the last equation is deduced from the observation that \((z, h_k) \rightarrow x_k \rightarrow u_k \rightarrow y_k \) is a Markov chain and the realistic assumption of independent additive noises and fading channels at the secondary users.

Therefore, the representative factor graph for (2.9) is as depicted in Figure 2.2, where the variables associated with only two secondary users are drawn for neatness. If the secondary users are very close to each other, it may be possible for adjacent PU-SU
channels, $H_k$ and $H_{k-1}$, to be correlated. This would be depicted in the factor graph with a connection between $H_k$ and $H_{k-1}$, represented by the factor node $p(h_k|h_{k-1})$. Alternatively, it may be more realistic to assume that all $H_k$s are independent, which would simplify the graph and the resulting message passing algorithm. The factor graph can also be easily modified to model correlations in time, for instance expressed in a Markov chain for each $H_k$.

Finally, if $u_k$ is dependent only on a function of $x_k$, e.g. its energy $\|x_k\|^2$, the variable node $x_k$ in the factor graph may be replaced by that function, say $t_k$. In other words, $I(u_k = \gamma_k(x_k)) \rightarrow I(u_k = \Gamma_k(t_k))$ and $p(x_k|h_{k-1}, z) \rightarrow p(t_k|h_{k-1}, z)$, where $t_k = \|x_k\|^2$, and $\Gamma_k(\|x_k\|^2) = \gamma_k(x_k)$. The joint distribution of interest is now

$$p(z, u^K_1, t^K_1, h^K_1|y^K_1) \propto p(z) \prod_{k=1}^{K} p(y_k|u_k) I(u_k = \Gamma_k(t_k)) p(t_k|h_k, z)p(h_k).$$

(2.14)

In this chapter, we assume that energy detection is used by each secondary user, and make use of the above transformation in the next section.

### 2.3.3 Belief Propagation for Cooperative Spectrum Sensing

Having identified the factor graph characterizing the cooperative sensing problem, we proceed by deriving the BP calculations required at each node. For independent PU-SU channels, our graphical model in Figure 2.2 is a cycle-free graph and hence, the sum-product algorithm performs exact marginalization. In this algorithm, the messages are initialized from the leaf nodes and each node, when all but one message from its neighbors have been received, computes the message for the remaining neighbor [48].

As a first step, all the $p(y_k|u_k)$ factor nodes send to their attached $U_k$ variable nodes the function (or message)

$$\mu_{p(y_k|u_k)\rightarrow U_k} = p(y_k|u_k)$$

(2.15)

which is subsequently forwarded to the $I(u_k = \Gamma_k(t_k))$ node by the $U_k$ node. This message
is multiplied by $I(u_k = \Gamma_k(t_k))$ at the $I(u_k = \Gamma_k(t_k))$ factor node, and summed over\(^3\) $u_k$ to produce the outgoing message to $T_k$. For the hard local decision scenario, this message is derived as

$$
\mu_{I(u_k=\Gamma_k(t_k)) \rightarrow T_k} = p(y_k|0)I(t_k < \tau_k) + p(y_k|1)I(t_k > \tau_k)
$$

where $p(y_k|1)$ and $p(y_k|0)$, respectively, represent $p(y_k|u_k = 1)$ and $p(y_k|u_k = 0)$ and $\tau_k$ is the local energy detection threshold of the $k$-th secondary user. Here we are assuming that $u_k \in \{0, 1\}$ but extension to other cases is straightforward in principle, as shown, for example, in later sections. This message is then passed to the $p(t_k|h_k, \bar{z})$ node by the $T_k$ node. The prior density $p(h_k)$ is also forwarded from the corresponding factor node to the $H_k$ variable node and subsequently to the $p(t_k|h_k, \bar{z})$ node. At this node, the outgoing message to $Z$ is computed as a product of the incoming messages from $T_k$ and $H_k$ with $p(t_k|h_k, \bar{z})$, integrated over $H_k$ and $T_k$ as follows:

$$
\mu_{p(t_k|h_k, \bar{z}) \rightarrow Z} = \int_{t_k} \int_{h_k} [p(y_k|0)I(t_k < \tau_k) + p(y_k|1)I(t_k > \tau_k)] p(t_k|h_k, \bar{z})p(h_k)dh_kdt_k. \tag{2.16}
$$

In computing the above message, we notice that by first marginalizing over $h_k$ we have $p(t_k|\bar{z}) = \int_{h_k} p(t_k|\bar{z}, h_k)p(h_k)dh_k$. Therefore, the message (2.16) reduces to

$$
\mu_{p(t_k|h_k, \bar{z}) \rightarrow Z} = p(y_k|0) \int_{t_k < \tau_k} p(t_k|\bar{z})dt_k + p(y_k|1) \int_{t_k > \tau_k} p(t_k|\bar{z})dt_k. \tag{2.17}
$$

For a general primary signal, by the central limit theorem (CLT), the local energy $t_k = \|x_k\|^2$ is asymptotically Gaussian distributed under either $\mathcal{H}_0$ or $\mathcal{H}_1$. For i.i.d. received signals at the $k$th secondary user over $N$ sampling intervals, we have

$$
t_k \sim \mathcal{N} \left( NE[|x_k|^2], N\text{var}[|x_k|^2] \right), \tag{2.18}
$$

where $\text{var}[|x_k|^2]$ is the variance of random variable $|x_k|^2$. For zero-mean $h_k$ and $n_k$, we

\(^3\)For discrete $u_k$. If $u_k$ is assumed to be continuous, then the summation is replaced by an integral.
have
\[
E[|x_k|^2] = \begin{cases} 
\sigma_{nk}^2, & \text{under } \mathcal{H}_0 \\
\sigma_{hk}^2 E[|z|^2] + \sigma_{nk}^2, & \text{under } \mathcal{H}_1 
\end{cases},
\]
(2.19)

\[
\text{var}[|x_k|^2] = \begin{cases} 
\sigma_{nk}^4, & \text{under } \mathcal{H}_0 \\
\sigma_{hk}^4 \left(2E[|z|^4] - (E[|z|^4])^2\right) + 2\sigma_{hk}^2 \sigma_{nk}^2 E[|z|^2] + \sigma_{nk}^4, & \text{under } \mathcal{H}_1.
\end{cases}
\]
(2.20)

Therefore, the message (2.17) sent to the $Z$ node is derived as
\[
\mu_{p(t_k|h_k,z)\to Z} = \left[1 - Q\left(\frac{\tau_k - NE[|x_k|^2]}{\sqrt{N \text{var}[|x_k|^2]}}\right)\right] p(y_k|0) + Q\left(\frac{\tau_k - NE[|x_k|^2]}{\sqrt{N \text{var}[|x_k|^2]}}\right) p(y_k|1),
\]
where $Q(\cdot)$ is the $Q$-function defined as $Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$.

Finally, the marginal distribution of $\mathbf{z}$ conditioned on the observations is computed as the product of all the messages received at the $Z$ node:
\[
p(\mathbf{z}|y^K) = p(\mathbf{z}) \prod_{k=1}^{K} \left[p(y_k|0) + [p(y_k|1) - p(y_k|0)]Q\left(\frac{\tau_k - NE[|x_k|^2]}{\sqrt{N \text{var}[|x_k|^2]}}\right)\right].
\]
(2.21)

However, both (2.5) and (2.6) require the likelihood function and hence by (2.8), what is needed in the implementation of the cooperative detector is the product of all messages arriving at the $Z$ node, excluding the $p(\mathbf{z})$ message. Therefore, the required likelihood function is
\[
p(y_1^K|z) = \prod_{k=1}^{K} \mu_{p(t_k|h_k,z)\to Z}
\]
(2.22)

with $\mu_{p(t_k|h_k,z)\to Z}$ given by (2.21) for hard local decisions.

### 2.4 NP-based LRT at the FC

After receiving the observations from all the secondary users, the NP test can be applied at the FC to decide whether the primary user is active or not. A Bayesian-risk approach can alternatively be adopted. However, in this chapter we study only the NP criterion which uses the LRT given in (2.6). The threshold $\lambda$ can be found by solving (2.7), or through finding an empirical distribution of the test statistic from observations in an interval in which the primary user is known to be inactive.
2.4.1 Hard Local Decisions

In this section, we consider the scenario that the secondary users make hard (binary) decisions, i.e. \( u_k \in \{0, 1\} \) and transmit these to the FC. We assume energy detection for mapping at the \( k \)th secondary user with local threshold of \( \tau_k \). The local energy threshold is chosen to maintain a certain local probability of false alarm \( p_{f_k} \) and is determined as follows:

\[
\tau_k = \sigma_{n_k}^2 \sqrt{N} Q^{-1}(p_{f_k}) + N\sigma_{n_k}^2, \tag{2.23}
\]

where \( Q^{-1}(\cdot) \) is the inverse \( Q \)-function.

Using the likelihood function derived in previous section in (2.21) under \( z = 0 \) and \( z \neq 0 \), the NP-based LRT in (2.6) is derived as

\[
L(y_1^K) = \prod_{k=1}^{K} \frac{p(y_k|0) + [p(y_k|1) - p(y_k|0)] Q\left(\frac{\tau_k - m_{t_k,H_1}}{\sigma_{t_k,H_1}}\right)}{p(y_k|0) + [p(y_k|1) - p(y_k|0)] Q\left(\frac{\tau_k - N\sigma_{n_k}^2}{\sqrt{N}\sigma_{n_k}^2}\right)}, \tag{2.24}
\]

where from (2.19) and (2.20)

\[
m_{t_k,H_1} = N(E[|z|^2] \sigma_{h_k}^2 + \sigma_{n_k}^2), \tag{2.25}
\]

\[
\sigma_{t_k,H_1} = \sqrt{N((2E[|z|^4] - E[|z|^2]^2)\sigma_{h_k}^4 + 2E[|z|^2] \sigma_{n_k}^2 \sigma_{h_k}^2 + \sigma_{n_k}^4).} \tag{2.26}
\]

We can observe that the LRT depends on \( E[|z|^2] \) and \( E[|z|^4] \) which in turn depends on the modulation type and maximum amplitude of the primary signal. For example, for \( M \)-ary PSK (phase shift keying) modulated primary signals, which are constant-modulus, \( E[|z|^2] = A^2 \) and \( E[|z|^4] = (E[|z|^2])^2 = A^4 \), where \( A \) is the maximum transmit voltage. For \( M \)-ary ASK (amplitude shift keying) primary signals, which are not constant-modulus, we derive \( E[|z|^2] \) and \( E[|z|^4] \) as

\[
E[|z|^2] = \frac{A^2 M + 1}{3(M - 1)}, \tag{2.27}
\]

\[
E[|z|^4] = \frac{A^4 3M^3 + 3M^2 - 7M - 7}{15(M - 1)^3}. \tag{2.28}
\]
Assuming the binary symmetric SU-FC channels with crossover probabilities $p(y_k = 1|u_k = 0) = \alpha_k$, the LRT in (2.24) becomes

$$L(y^K_f) = \prod_{k=1}^{K} \alpha_k^{y_k} (1-\alpha_k)^{1-y_k} \left( \frac{1-p_{d_k}}{1-p_{f_k}} \right) + \alpha_k^{1-y_k} (1-\alpha_k)^{y_k} \left( \frac{1-p_{f_k}}{1-p_{d_k}} \right)$$  \hspace{1cm} (2.29)

where $p_{d_k} = Q\left( \frac{\tau_k - m_{t_k} \alpha_k}{\sqrt{N \sigma_n^2}} \right)$ and $p_{f_k} = Q\left( \frac{\tau_k + N \sigma_n^2}{\sqrt{N \sigma_n^2}} \right)$ are, respectively, the local probabilities of detection and false alarm at the $k$-th secondary user. Assuming that all the secondary users employ the same decision thresholds and experience the same channel conditions (i.e., $\sigma_h^2 = \sigma_h^2$, $\sigma_n^2 = \sigma_n^2$, and $\alpha_k = \alpha$ for $k = 1, \ldots, K$), the local probabilities of detection and false alarm will be the same for all the secondary users (i.e., $p_{d_k} = p_d$ and $p_{f_k} = p_f$ for $k = 1, \ldots, K$). In the following, we show that under these conditions, the NP-based LRT is identical to the $M$-out-of-$K$ rule.

We first note that the terms within the product operation in (2.29) can take only two forms, corresponding to $y_k = 0$ and $y_k = 1$. If $M$ of the $y_k$s are 1, we can then write

$$\log L(y^K_f) = M \log \frac{p_D(1-p_F)}{p_F(1-p_D)} + K \log \frac{1-p_D}{1-p_F}$$  \hspace{1cm} (2.30)

where

$$p_D = p(y_k = 1|H_1) = \alpha + p_d - 2\alpha p_d,$$  \hspace{1cm} (2.31)

$$p_F = p(y_k = 1|H_0) = \alpha + p_f - 2\alpha p_f.$$  \hspace{1cm} (2.32)

where $p_D$ and $p_F$ are not dependent on $k$ because of the i.i.d. assumption for SU-FC channels. We note that these are probabilities of detection and false alarm for the individual SU-FC channel outputs at the fusion center, with the bit error probabilities between secondary user and FC accounted for.

For the FC decision threshold $\lambda$ satisfying

$$K \log \frac{1-p_D}{1-p_F} + (M-1) \log \frac{p_D(1-p_F)}{p_F(1-p_D)} < \log \lambda < K \log \frac{1-p_D}{1-p_F} + M \log \frac{p_D(1-p_F)}{p_F(1-p_D)},$$  \hspace{1cm} (2.33)

Binary symmetric channel is an appropriate model for most wireless digital communication channels.
Chapter 2. Probabilistic Inference for Cooperative Sensing

the FC decides $\mathcal{H}_1$ if $M$ or more $y_k$s are 1 and decides $\mathcal{H}_0$ if $M - 1$ or fewer $y_k$s are 1. When (2.33) is satisfied, the NP test becomes equivalent to the $M$-out-of-$K$ rule, with an overall probability of false alarm of

$$P_f = \sum_{k=M}^{K} \binom{K}{k} P_F^k (1 - P_F)^{K-k}. \quad (2.34)$$

We have therefore established that the $M$-out-of-$K$ rule is optimal in the NP sense when:

1. All $K$ channels between primary user and secondary users are statistically identical
2. All $K$ channels from secondary user to FC are identical and independent binary symmetric channels
3. All the $K$ secondary users use the same local thresholds and send hard local decisions to the FC.

Different SU-FC channel models can be considered using the framework discussed in this section, by appropriately modifying $p(y_k|0)$ and $p(y_k|1)$ in (2.24). For instance, in an AWGN SU-FC channel with noise variable $w_k \sim \mathcal{N}(0, \sigma_{w_k}^2)$, $p(y_k|u_k) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp \left( -\frac{|y_k - u_k|^2}{2\sigma_n^2} \right)$ can be applied in (2.24).

2.4.2 Soft Local Decisions

In many of the previous studies on cooperative sensing, the local test statistics (soft decisions) are assumed to be transmitted with infinite precision to the FC over error-free control channels [17, 21, 22]. In [19, 22], for instance, the local test statistics (signal energies) were transmitted over error-free control channels to the fusion center, and a global decision was made based on a linear combination of the local test statistics. In this section, we investigate the factor graph approach in deriving the LRT at the FC for soft local decisions. For a more realistic model compared to others [17–19, 22], we assume the SU-FC channels to be noisy AWGN channels.
Chapter 2. Probabilistic Inference for Cooperative Sensing

The test statistic received at the FC from the $k$-th secondary user is $y_k = t_k + w_k$, where $w_k$ is assumed to be zero-mean Gaussian noise with variance $\sigma^2_{w_k}$. In order to use BP to derive the likelihood function, we first notice that in the soft local decision scenario $u_k = t_k$, i.e. $p(u_k|t_k) = 1$, and therefore the variable nodes $u_k$ and $t_k$ can be merged in the factor graph. The $p(y_k|t_k)$ factor node sends

$$p(y_k|t_k) = \frac{1}{\sqrt{2\pi\sigma_{w_k}}} \exp\left(-\frac{(y_k - t_k)^2}{2\sigma^2_{w_k}}\right)$$

to the $T_k$ node which subsequently forwards it to the $p(t_k|h_k,z)$ node. The outgoing message from the $p(t_k|h_k,z)$ node to the $Z$ node is computed as

$$\mu_{p(t_k|h_k,z)} \rightarrow Z = \int_{t_k} p(y_k|t_k)p(t_k|h_k,z)dh_kdt_k = \int_{t_k} p(y_k|t_k)p(t_k|z)dt_k = p(y_k|z). \quad (2.35)$$

Given the Gaussian distribution of $t_k$ as in (2.18), $y_k$ is also Gaussian distributed under $H_0$ or $H_1$:

$$y_k \sim \mathcal{N}\left(NE[|x_k|^2], N\text{var}[|x_k|^2] + \sigma^2_{w_k}\right), \quad (2.36)$$

where $E[|x_k|^2]$ and $\text{var}[|x_k|^2]$ under $H_0$ and $H_1$, are given in (2.19) and (2.20), respectively. Therefore, the message (2.35) reduces to

$$\mu_{p(t_k|h_k,z)} \rightarrow Z = \frac{\exp\left(-\frac{(y_k - NE[|x_k|^2])^2}{2(N\text{var}[|x_k|^2] + \sigma^2_{w_k})}\right)}{\sqrt{2\pi(N\text{var}[|x_k|^2] + \sigma^2_{w_k})}}. \quad (2.37)$$

By the arguments from Section 2.3.3, the likelihood function for the soft local decision scenario is attained by inserting (2.37) into (2.22). Finally, the NP test statistic is calculated as follows:

$$L(y^K_1) = \prod_{k=1}^K e^{\frac{(y_k - m_{t_k,H_1})^2}{2(\sigma^2_{t_k,H_1} + \sigma^2_{w_k})}} \cdot e^{\frac{(y_k - N\sigma^2_{w_k})^2}{2(N\sigma^2_{w_k} + \sigma^2_{w_k})}} \cdot \sqrt{\frac{\sigma^2_{t_k,H_1} + \sigma^2_{w_k}}{N\sigma^2_{w_k} + \sigma^2_{w_k}}}, \quad (2.38)$$

where $m_{t_k,H_1}$ and $\sigma^2_{t_k,H_1}$ are given in (2.25) and (2.26) respectively. The LRT decides $H_1$ if $L(y^K_1) > \lambda$, where the threshold $\lambda$ can be set using empirical distribution of the log-likelihood ratio under the null hypothesis.
2.4.3 Quantized Soft Local Decisions

The assumption of sending uncompressed soft local test statistics to the FC is not realistic due to bandwidth constraints in wireless communication systems. In fact, allocating wideband channels to transmit uncompressed soft information to FC contradicts the essential advantage of CRs which is to increase spectrum efficiency. In this section, we assume that the secondary users locally quantize their test statistics (received signal energies) and send them to the FC. Using BP, we derive the likelihood functions and subsequently the NP-based LRT test statistics at the FC.

Let us assume that the kth secondary user quantizes its received signal energy into \( M_k \) bits with quantization thresholds of \( a_{0k}, \ldots, a_{(M_k-1)k} \) and subsequently sends the quantization output \( u_k \in \{0, \ldots, M_k-1\} \) to the FC. In this case, the message from node \( I(u_k = \Gamma_k(t_k)) \) to node \( t_k \) in the message-passing approach is derived as

\[
\mu_{I(u_k = \Gamma_k(t_k)) \rightarrow T_k} = \sum_{u_k} p(y_k|u_k) I(u_k = \Gamma_k(t_k))
\]

where \( q_i = i - 1, a_{0k} = 0 \) and \( a_{M_kk} = \infty \). Consequently, the outgoing message to the \( Z \) node from \( p(t_k|h_k, z) \) is computed as

\[
\mu_{p(t_k|h_k, z) \rightarrow Z} = \sum_{i=1}^{M_k} p(y_k|q_i) \int_{a_{(i-1)k}}^{a_{ik}} p(t_k|z) dt_k.
\]

Using the fact that \( t_k \) has a Gaussian distribution as in (2.18), and also the fact that the likelihood function is obtained by

\[
p(y^K_1|z) = \prod_{k=1}^{K} \mu_{p(t_k|h_k, z) \rightarrow Z},
\]

the LRT statistic at the FC is derived as

\[
L(y^K_1) = \frac{\sum_{i=1}^{M_k} p(y_k|q_i) \left[ Q \left( \frac{a_{(i-1)k} - m_{t_k, H_1}}{\sigma_{t_k, H_1}} \right) - Q \left( \frac{a_{ik} - m_{t_k, H_1}}{\sigma_{t_k, H_1}} \right) \right] }{\sum_{i=1}^{M_k} p(y_k|q_i) \left[ Q \left( \frac{a_{(i-1)k} - N\sigma^2_{\alpha_k}}{\sqrt{N}\sigma^2_{\alpha_k}} \right) - Q \left( \frac{a_{ik} - N\sigma^2_{\alpha_k}}{\sqrt{N}\sigma^2_{\alpha_k}} \right) \right]}
\]

where \( m_{t_k, H_1} \) and \( \sigma_{t_k, H_1} \) are respectively given in (2.25) and (2.26).
Figure 2.3: Complementary ROC ($P_m$ vs. $P_f$) under AWGN SU-FC channels with $\sigma_n^2 = 1$, $N = 200$ sensing samples, $\sigma_h^2 = 2$ and $\sigma_n^2 = 8$.

### 2.5 Simulation Results

In this section, the performance of the NP-based LRT cooperative sensing is evaluated numerically for hard, soft and quantized local decisions and compared with several existing schemes. The primary signal and local binary decisions are mainly assumed to be BPSK modulated with unit power. The complementary receiver operating characteristic (ROC) curves (plots of probability of missed detection, $P_m = 1 - P_d$, vs. probability of false alarm, $P_f$) are used to characterize the performance of the cooperative spectrum sensing schemes.

Figure 2.3 shows the complementary ROC curves at different LRT threshold levels, $\lambda$, under the AWGN SU-FC channel scenario. The cases of $K = 6$ and $K = 12$ secondary
users and the performance of the OR and $M$-out-of-$K$ rules for $M = K/3, K/2, 2K/3$ are provided for comparison. It can be observed that the LRT outperforms the OR and $M$-out-of-$K$ rules in AWGN SU-FC channels scenario. For a certain choice of local thresholds, the NP-based LRT can achieve a range of performance in the complementary ROC using different global threshold values, while $M$-out-of-$K$ methods can only attain a fixed performance as their fusion rule is fixed and does not depend on any global threshold. By changing the local thresholds, the $M$-out-of-$K$ methods can achieve a range of performance.

Figure 2.4 provides plots of complementary ROC curves at different LRT threshold levels, $\lambda$, under the BSC scenario for SU-FC links and local probabilities of false alarm.
Figure 2.5: Complementary ROC ($P_m$ vs. $P_f$) under binary symmetric SU-FC channels with $N = 500$ sensing samples, $K = 6$ secondary users, $\sigma^2_{h_1} = \sigma^2_{n_1}$, $\sigma^2_{n_1} = \{1.28, 0.32, 4.5, 0.5, 2, 0.08\}$ and $\alpha^K = \{0.2, 0.3, 0.1, 0.05, 0.2, 0.4\}$.

of $p_f = 0.05$ and $p_f = 0.2$. We observe that, as shown in Section 2.4.1, under identical channel statistics and binary symmetric SU-FC channels, the NP-based LRT performs identically to the $M$-out-of-$K$ rule for certain choices of $\lambda$. We also observe that by increasing $M$ from 1 to $K$ the performance of $M$-out-of-$K$ rule moves from performance of OR rule to AND rule. As Figure 2.5 shows, for non-identical channel statistics and thresholds, the LRT fusion outperforms the $M$-out-of-$K$ rule.

Figure 2.6 shows the complementary ROC curves for the soft local decision scenario. The adding (ADD) [19], maximal ratio combining (MRC) and selection combining (SC) curves are provided for comparison. We observe that the soft-decision LRT scheme out-
Figure 2.6: Complementary ROC ($P_m$ vs. $P_f$) under soft local decisions, with $N = 100$ sensing samples, $K = 6$ secondary users, $\sigma_n^2 = 1$, $\sigma_h^2 = 2$ and $\sigma_n^2 = 4$. 
Figure 2.7: Complementary ROC ($P_m$ vs. $P_f$) under quantized local decisions, with $N = 500$ sensing samples, $K = 6$ secondary users, $\sigma_{h_1}^{2K} = \sigma_{n_1}^{2K} = \sigma_{n_1}^{2K} = \{1.28, 0.32, 4.5, 0.5, 2, 0.08\}$, and $\alpha_1^K = \{0.1, 0.3, 0.2, 0.05, 0.1, 0.2\}$.
performs the aforementioned schemes by having a lower probability of missed detection for the same probability of false alarm.

Figure 2.7 depicts the complementary ROC curves for a two-bit quantized local decision scenario. The quantization thresholds are set to be $a_{1k} = \tau_k/2$, $a_{2k} = \tau_k$ and $a_{3k} = 2\tau_k$, where $\tau_k$ is determined as (2.23). The primary signal is assumed to be 4-ASK modulated with maximum transmit voltage of $A = 1$. The $M$-out-of-$K$ and hard (one-bit) local decision LRT are provided for comparison. We observe that the two-bit quantized LRT scheme significantly outperforms the one-bit quantized LRT and $M$-out-of-$K$ schemes.

2.6 Chapter Summary

In this chapter, we presented a probabilistic inference approach for cooperative spectrum sensing by modeling the decision fusion in cooperative sensing as a probabilistic inference problem on a factor graph. This approach allows for modeling and accommodating the uncertainties and correlations in the cooperative sensing system and applying the well-known algorithms such as belief propagation for inferring the target probability distribution. Using belief propagation approach, we derived the LRT at the fusion center for hard, soft and quantized local decision scenarios. Considering nonideal noisy channels, we showed that the LRT cooperative sensing scheme outperforms conventional schemes such as AND, OR and $M$-out-of-$K$ rules and MRC and SC schemes. Unlike the hard decision based methods that can only have a certain performance for fixed local thresholds, the LRT method allows for achieving a wide range of performance by changing the global threshold. We also showed that the quantized local decisions can considerably improve the detection performance of the cognitive radios compared to the hard (one-bit) decision scenario.

As shown in this chapter, the optimal LRT requires the knowledge of the primary
signal and channel gains statistics which may not be available to the secondary users. In the next chapter, we present a new scheme overcome this problem.
Chapter 3

Cooperative Spectrum Sensing with Unknown Parameters

3.1 Introduction

As we showed in the previous chapter, the optimal LRT decision rule depends on the statistics of the primary signal and channel gains. However, in cognitive radios such information may not be readily available to the secondary users and the fusion center. Therefore, in this chapter we investigate the problem of spectrum sensing with unknown parameters. We consider the problem of spectrum sensing at a central node in a cooperative-sensing cognitive radio network, acting analogously to the fusion center in a distributed sensor network. The channels between the $K$ sensing nodes or secondary users and the FC are treated as noisy. The primary network, when active, is assumed to transmit a zero-mean signal unknown to the cognitive network, which is received over a Rayleigh flat-fading channel by each secondary user.

When the second- and fourth-order moments of the primary signal and the variance of channel gains are known, we will show that it is possible to derive the NP test at
the FC based on various ways\(^1\) of making local decisions from the observations at the sensing nodes. These results are derived using general concepts introduced in the context of sensor networks [40, 49, 52–55], specialized to the scenario we just described. However, in practice, it is unlikely that the SUs or the FC would have knowledge of the moments of the transmitted primary signal or the channel between the primary transmitter and each secondary user. In fact, this is an important constraint in cognitive radio networks.

Thus we go on to address the above problem, using first the notion of the generalized likelihood ratio test (GLRT) and its asymptotically (as the number of observations increase without bound) equivalent form, and then further simplifying it heuristically to the sum of log-likelihood derivatives. The resulting test statistic, which we call the linear test statistic, does not require maximum likelihood estimates (MLEs) of the unknown parameters, which are hard to derive analytically in many realistic scenarios, while showing a comparable performance to the known-parameter optimal NP test. Compared to other well-known asymptotically-equivalent tests to GLRT, such as the Wald and Rao tests [56], this scheme is much simpler and does not require the calculation of the Fisher information matrix which is not easily attainable or invertible in many practical scenarios.

The proposed simplified linear test is applied to the cases of (i) unknown primary signal and PU-SU channel statistics, and (ii) unknown primary signal statistics but known PU-SU channel statistics. In case (ii), for the practical case of weak constant-modulus primary signals, we additionally derive the LMP test [56]. Furthermore, for threshold setting and performance analysis, we derive the distributions of the linear test and LMP statistics under the null hypothesis. In this chapter, we consider soft (infinite-precision) and quantized local decision scenarios and derive the LRT and linear test statistics for each case. Our simulation results demonstrate the close performance of the linear test to the known-parameter LRT, without requiring the knowledge or estimation of the primary statistics. This chapter has been published in part in [57] and [58].

\(^1\)Binary, \(M\)-level and no quantization of the received signal energy.
Figure 3.1: Block diagram of cooperative spectrum sensing system with $K$ secondary users.

The remainder of this chapter is organized as follows. The system model and the NP-based LRT at the fusion center are presented in Section 3.2. The proposed linear test for cooperative spectrum sensing is presented in Section 3.3. Section 3.4 comprises the derivation of the linear test statistics for different types of local decisions under the unknown channel and primary signal statistics as well as the linear test and LMP statistics for known channel statistics and unknown primary signal moments along with the analytical threshold setting for linear test and LMP. Section 3.5 provides the simulation results and, finally, Section 3.6 concludes the chapter.

3.2 Model and Formulation

3.2.1 System Model

We consider a centrally coordinated cognitive radio network with $K$ secondary users. We assume the same system model and parameters as in Chapter 2. Figure 3.1 depicts the block diagram of the assumed system. The detailed system model is provided in Section 2.2.
3.2.2 NP-Based LRT at the Fusion Center

Spectrum sensing is a binary hypothesis testing problem. The optimal decision rule at the fusion center in the sense of maximizing the probability of detection for a given probability of false alarm is the NP-based LRT. This method decides $\mathcal{H}_1$ if and only if

$$L(y^K_1) = \frac{p(y^K_1|z \neq 0)}{p(y^K_1|z = 0)} > \lambda,$$

where the threshold $\lambda$ is found by solving (numerically or analytically)

$$P_f = \int_{y^K_1:L(y^K_1)\geq \lambda} p(y^K_1|z = 0) dy^K_1$$

for a given desired probability of false alarm $P_f$. We can apply the graphical model approach for cooperative sensing presented in Chapter 2 to obtain the likelihood functions $p(y^K_1|z \neq 0)$ and $p(y^K_1|z = 0)$, for different system model assumptions. We showed in (2.21) that the likelihood functions depend on $E[|x^2_k]$ and $\text{var}[|x^2_k]$ which were derived in (2.19) and (2.20). It can be observed from (2.19) and (2.20) that $E[|x^2_k]$ and $\text{var}[|x^2_k]$ depend on the second and forth order statistics of the primary signal and channel gain variances. Therefore, LRT-based methods are only applicable when $E[|x^2_k]$ and $\text{var}[|x^2_k]$ are known, in particular, when the second and forth order statistics of the primary signal and channel gain variances are known.

3.3 The Proposed Composite Hypothesis Testing Approach

As discussed in Section 3.2.2 the optimal NP-based method relies on knowledge of the primary signal and channel statistics. In practice, however, such knowledge may not be available to the cognitive radios and thus, the LRT-based schemes are not directly applicable. Therefore, in cognitive radios, detection schemes that are robust to unknown parameters should be employed. In a general scenario, we assume that no knowledge
Chapter 3. Cooperative Sensing with Unknown Parameters

of the primary signal statistics and channel gains is available at the fusion center. We denote $\beta_{1,k} = \sigma_{h_k}^2 E[|z|^2]$ and $\beta_{2,k} = \sigma_{h_k}^4 E[|z|^4]$, for $k = 1, \ldots, K$, as the unknown parameters in the likelihood ratio in (3.1). The vector of unknown parameters is then $\underline{\beta} = (\beta_{1,1}, \ldots, \beta_{1,K}, \beta_{2,1}, \ldots, \beta_{2,K})$.

A well-known composite hypothesis test which can replace the NP-based LR T in unknown parameter scenarios is the GLR T. Unlike the NP-based LR T, the GLR T is not optimal but has been found to work well in many scenarios. This test first finds the MLE of the unknown parameters under $H_1$, and then forms the GLR T statistic:

$$L_G(y^K_1) = \frac{p(y^K_1|\underline{\beta} = \hat{\beta}, H_1)}{p(y^K_1|\underline{\beta} = 0, H_0)}$$ (3.3)

where $\hat{\beta} = \arg\max_{\underline{\beta}} p(y^K_1|\underline{\beta}, H_1)$.

For finding $\hat{\beta}$, we can maximize $p(y^K_1|\underline{\beta})$ over the entire parameter space rather than merely over $H_1$ (see [56], pg. 232), since the probability of the MLE yielding the particular value 0 is zero. Therefore $\hat{\beta}$ is the unrestricted MLE of $\underline{\beta}$. The unrestricted MLE of $\underline{\beta}$ attains the Cramer-Rao lower bound, assuming the asymptotic pdf of the MLE is attained (see [56], pg. 232), thus it satisfies [28, 56]

$$\frac{\partial \ln p(y^K_1|\underline{\beta})}{\partial \underline{\beta}} = I(\underline{\beta})(\hat{\beta} - \underline{\beta})$$ (3.4)

The $i$th element of $\frac{\partial \ln p(y^K_1|\underline{\beta})}{\partial \underline{\beta}_i}$ is given by [56],

$$\frac{\partial \ln p(y^K_1|\underline{\beta})}{\partial \beta_i} = \sum_{j=1}^K [I(\underline{\beta})]_{ij} (\hat{\beta}_j - \beta_j)$$ (3.5)

By the first-order Taylor expansion

$$[I(\hat{\beta})]_{ij} = [I(\hat{\beta})]_{ij} + \frac{\partial [I(\underline{\beta})]_{ij}}{\partial \underline{\beta}} |_{\underline{\beta} = \hat{\beta}} (\beta_j - \hat{\beta}_j)$$ (3.6)

we have

$$\frac{\partial \ln p(y^K_1|\underline{\beta})}{\partial \beta_i} = \sum_{j=1}^K [I(\hat{\beta})]_{ij} (\hat{\beta}_j - \beta_j) + \sum_{j=1}^K \frac{\partial [I(\underline{\beta})]_{ij}}{\partial \beta_\star} |_{\beta_\star = \hat{\beta}} (\beta_j - \hat{\beta}_j)$$ (3.7)
By consistency of the unrestricted MLE \([28, 56, 59]\), as \(K \to \infty\), \(\hat{\beta} \to \beta\), and thus the last term which is of second order may be neglected \([56]\). Thus we have

\[
\frac{\partial \ln p(y^K_1|\beta)}{\partial \beta} = I(\hat{\beta})(\hat{\beta} - \beta) \tag{3.8}
\]

Integrating both sides of (3.8) with respect to \(\beta\) results in

\[
\ln p(y^K_1|\beta) = (\hat{\beta} - \beta)^T I(\hat{\beta})(\hat{\beta} - \beta) + C. \tag{3.9}
\]

The constant \(C\) is bound to be \(\ln p(y^K_1|\hat{\beta})\). Therefore the asymptotic likelihood function is

\[
p(y^K_1|\beta) = p(y^K_1|\hat{\beta}) \exp \left( -\frac{1}{2} (\hat{\beta} - \beta)^T I(\hat{\beta})(\hat{\beta} - \beta) \right) \tag{3.10}
\]

By replacing \(p(y^K_1|\hat{\beta})\) from (3.10) into (3.3), the GLRT statistic reduces to

\[
L_G(y^K_1) = \frac{p(y^K_1|\beta = \hat{\beta})}{p(y^K_1|\beta = 0)} = \frac{1}{\exp \left( -\frac{1}{2} \hat{\beta}^T I(\hat{\beta})\hat{\beta} \right)} \tag{3.11}
\]

Therefore, the asymptotic equivalent for GLRT can be represented as

\[
2 \ln L_G(y^K_1) = \hat{\beta}^T I(\hat{\beta})\hat{\beta} \tag{3.12}
\]

By substituting (3.8) into the asymptotic form of the GLRT in (3.12), the proposed asymptotically equivalent test to GLRT is derived as

\[
T_L(y^K_1) = \hat{\beta}^T \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta} \bigg|_{\beta = 0}. \tag{3.13}
\]

We call the proposed test as the linear test as its test statistic consists of a linear combination of \(\frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_k} \bigg|_{\beta = 0}\) terms. For the set of unknown parameters in our model, the linear test statistic can be written as

\[
T_L(y^K_1) = \sum_{k=1}^{K} \hat{\beta}_{1,k} \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{1,k}} \bigg|_{\beta = 0} + \sum_{k=1}^{K} \hat{\beta}_{2,k} \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{2,k}} \bigg|_{\beta = 0} \tag{3.14}
\]
In scenarios where it is infeasible or computationally demanding to solve
\[
\frac{\partial \ln p(y_1^K|\beta)}{\partial \beta} = 0 \tag{3.15}
\]
uniquely for $\beta$ under $\mathcal{H}_1$, we alternatively propose to substitute the $\hat{\beta}$ vector with a vector of ones, $1$,
\[
T_L(y_1^K) = 1^T \frac{\partial \ln p(y_1^K|\beta)}{\partial \beta} \bigg|_{\beta=0}.
\tag{3.16}
\]
This choice is intuitive as we assume no prior knowledge about the known parameters and when the MLEs are not available, they can be treated equally and the effect of all the MLEs in (3.14) can be equalized. This simplified test can also be treated as a generalized LMP with multiple unknown parameters. In our system model, the unknown parameters are non-negative which removes the necessity of the quadratic form in the test statistic and allows for a linear test statistic as in (3.16). Our numerical results support this suggested scheme and show a close performance for the proposed linear scheme to that of the optimal NP-based LRT with known parameters. This simplified linear test statistic for the assumed model can be rewritten as
\[
T_{SL}(y_1^K) = \sum_{k=1}^{K} \frac{\partial \ln p(y_1^K|\beta)}{\partial \beta_{1,k}} \bigg|_{\beta=0} + \sum_{k=1}^{K} \frac{\partial \ln p(y_1^K|\beta)}{\partial \beta_{2,k}} \bigg|_{\beta=0}.
\tag{3.17}
\]
This method can be used for any detection problem with unknown or uncertain parameters. Unlike the Rao and Wald tests (see Appendix C), the simplified linear test does not have a quadratic form and does not require the calculation of the Fisher information matrix which is a $2K \times 2K$ matrix for our system. Therefore it is easier and faster to implement compared to these schemes. In its simplified form (3.17), the linear test does not even require the MLEs of the unknown parameters which results in a much simpler and faster test compared to the Wald and GLRT tests. In fact, it is not feasible to analytically find the unrestricted MLEs of unknown parameters in most realistic scenarios [56], such as the soft and quantized local decision scenarios in this chapter. This linear test is also much simpler than the Rao test which has a quadratic form and requires the
inverse of the Fisher information matrix. The Rao test, in fact, is not practical as the Fisher information matrix is not easily attainable or invertible in many system models. Dealing with multiple unknown parameters (in this case $2K$ parameters), exacerbates the complexity and infeasibility of calculating this inverse matrix.

### 3.4 The Proposed Test for Cooperative Spectrum Sensing

#### 3.4.1 Unknown Primary Signal and Channel Statistics

In this section, we derive the NP-based LRT at the fusion center and apply the proposed linear test to cooperative fusion of quantized and soft local decisions with unknown primary signal moments and channel statistics.

**Quantized Soft Local Decisions**

In this section, we assume that the secondary users transmit locally quantized test statistics (received signal energies) and then derive the LRT and linear test statistics at the fusion center. In this chapter, we do not deal with the optimization of the local thresholds and refer the readers to [40,49,52,54] for threshold optimization algorithms in distributed sensor networks. We assume that the $k$th user quantizes its received signal energy with $M_k$ bits and preset local thresholds of $a_{1k}, \ldots, a_{(M_k-1)k}$ and sends $u_k \in \{0, \ldots, M_k - 1\}$ to the fusion center. In this case, the local decision sent to the fusion center based on the quantized local energy is determined as follows:

\[
 u_k = q_i, \text{ if } a_{(i-1)k} < t_k < a_{ik} \quad \text{for } i = 1, \ldots, M_k \tag{3.18}
\]
where \( q_i = i - 1 \), \( a_{0k} = 0 \) and \( a_{M_k,k} = \infty \). The likelihood function in this case is derived in Appendix B as

\[
p(y^K_1 | z) = \prod_{k=1}^K \left( \sum_{i=1}^{M_k} p(y_k | q_i) \left[ Q \left( \frac{a_{(i-1)k} - m_{tk}}{\sigma_{tk}} \right) - Q \left( \frac{a_{ik} - m_{tk}}{\sigma_{tk}} \right) \right] \right) \quad (3.19)
\]

where \( Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \) and \( m_{tk} \) and \( \sigma_{tk}^2 \) are respectively the mean and variance of the local energy \( t_k \) given as

\[
m_{tk, \mathcal{H}_1} = N(\beta_{1,k} + \sigma_{n_k}^2), \quad \sigma_{tk, \mathcal{H}_1} = \sqrt{N((2\beta_{2,k} - \beta_{1,k}^2) + 2\beta_{1,k}\sigma_{n_k}^2 + \sigma_{n_k}^4)}. \quad (3.20)
\]

Consequently, the LRT statistic at the fusion center is derived as,

\[
L(y^K_1) = \prod_{k=1}^K \frac{\sum_{i=1}^{M_k} p(y_k | q_i) \left[ Q \left( \frac{a_{(i-1)k} - m_{tk, \mathcal{H}_1}}{\sigma_{tk, \mathcal{H}_1}} \right) - Q \left( \frac{a_{ik} - m_{tk, \mathcal{H}_1}}{\sigma_{tk, \mathcal{H}_1}} \right) \right]}{\sum_{i=1}^{M_k} p(y_k | q_i) \left[ Q \left( \frac{a_{(i-1)k} - N\sigma_{n_k}^2}{\sqrt{N}\sigma_{n_k}^2} \right) - Q \left( \frac{a_{ik} - N\sigma_{n_k}^2}{\sqrt{N}\sigma_{n_k}^2} \right) \right]}. \quad (3.22)
\]

We observe that the NP-based LRT depends on \( \beta_{1,k} = \sigma_{h_k}^2 E[|z|^2] \) and \( \beta_{2,k} = \sigma_{h_k}^4 E[|z|^4] \). For \( M \)-ary PSK (phase shift keying) modulated primary signals, \( E[|z|^2] = A^2 \) and \( E[|z|^4] = A^4 \), where \( A \) is the maximum transmit voltage. For \( M \)-ary ASK (amplitude shift keying), we have

\[
E[|z|^2] = \frac{A^2 M + 1}{3M - 1}, \quad (3.23)
\]

\[
E[|z|^4] = \frac{A^4 3M^3 + 3M^2 - 7M - 7}{15(M - 1)^3}. \quad (3.24)
\]

Therefore, in order to apply the LRT, knowledge of the modulation type and maximum transmit voltage of the primary signal as well as the channel gain variance is required at the fusion center.

For the realistic scenario that the primary signal moments and channel statistics are not available at the cognitive radios, we propose to apply the linear test rather than the LRT in (3.22). For this reason, we derive the partial derivatives of the logarithm of the LRT in (3.22) with respect to \( \beta_{1,k} \)’s and \( \beta_{2,k} \)’s and evaluate them at \( \beta = 0 \) as follows:
The linear test statistic at the fusion center, in this case, is obtained from (3.29) as

\[
\frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{1,k}} \bigg|_{\beta = 0} = \frac{\sum_{i=1}^{M_k} p(y_k|q_i) A_{ik}}{\sqrt{2\pi N \sigma_{n_k}^2}}
\]

(3.25)

\[
\frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{2,k}} \bigg|_{\beta = 0} = \frac{\sum_{i=1}^{M_k} p(y_k|q_i) B_{ik}}{\sqrt{2\pi N \sigma_{n_k}^2}}
\]

(3.26)

where

\[
A_{ik} = a_{(i-1)k} e^{-\frac{(a_{(i-1)k} - N \sigma_{n_k}^2)^2}{2N \sigma_{n_k}^2}} - a_{ik} e^{-\frac{(a_{ik} - N \sigma_{n_k}^2)^2}{2N \sigma_{n_k}^2}},
\]

(3.27)

\[
B_{ik} = \left[ a_{(i-1)k} - N \sigma_{n_k}^2 \right] e^{-\frac{(a_{(i-1)k} - N \sigma_{n_k}^2)^2}{2N \sigma_{n_k}^2}} - \left[ a_{ik} - N \sigma_{n_k}^2 \right] e^{-\frac{(a_{ik} - N \sigma_{n_k}^2)^2}{2N \sigma_{n_k}^2}}.
\]

(3.28)

Derivation of the MLEs of \( \beta_{1,k} \) and \( \beta_{2,k} \) is not feasible analytically in this scenario. Therefore, we apply the simplified linear test and derive its test statistic as

\[
T_{SL}(y^K_1) = \sum_{k=1}^{K} \sum_{i=1}^{M_k} p(y_k|q_i) \left[ Q \left( \frac{a_{(i-1)k} - N \sigma_{n_k}^2}{\sqrt{N \sigma_{n_k}^2}} \right) - Q \left( \frac{a_{ik} - N \sigma_{n_k}^2}{\sqrt{N \sigma_{n_k}^2}} \right) \right]
\]

(3.29)

For the spacial case of binary (hard) local decisions, the LRT statistic in (3.22) reduces to

\[
L(y^K_1) = \prod_{k=1}^{K} \frac{p(y_k|0) + [p(y_k|1) - p(y_k|0)] Q \left( \frac{\tau_k - m_k}{\sigma_{k,n_1}} \right)}{p(y_k|0) + [p(y_k|1) - p(y_k|0)] Q \left( \frac{\tau_k - N \sigma_{n_k}^2}{\sqrt{N \sigma_{n_k}^2}} \right)},
\]

(3.30)

where \( \tau_k \) is the local threshold at the kth user. Based on the distribution of the local energy in (2.18), the local threshold \( \tau_k \) that result in local probability of false alarm of \( p_{f_k} \) is determined as

\[
\tau_k = \sigma_{n_k}^2 \sqrt{N} Q^{-1}(p_{f_k}) + N \sigma_{n_k}^2.
\]

(3.31)

The linear test statistic at the fusion center, in this case, is obtained from (3.29) as

\[
T_{SL}(y^K_1) = \sum_{k=1}^{K} \frac{p(y_k|1) - p(y_k|0)}{p(y_k|0) + [p(y_k|1) - p(y_k|0)] Q \left( \frac{\tau_k - N \sigma_{n_k}^2}{\sqrt{N \sigma_{n_k}^2}} \right)} \left[ \frac{\tau_k - m_k \sigma_{k,n_1}}{\sigma_{k,n_1}} \right]
\]

(3.32)
Soft Local Decisions

In many previous studies in cooperative sensing [17, 19, 21, 22] the local test statistics (signal energies) were assumed to be transmitted with infinite precision to the fusion center over error-free control channels. In this section, we derive the LRT and then the linear test at the fusion center for such soft local decisions with noisy AWGN channels from the users to the fusion center.

The signal received at the fusion center from the \( k \)-th sensor is \( y_k = t_k + w_k \), where \( w_k \) is assumed to be zero-mean Gaussian noise with variance \( \sigma_{w_k}^2 \), so that

\[
p(y_k | t_k) = \frac{1}{\sqrt{2\pi}\sigma_{w_k}} \exp\left( -\frac{(y_k - t_k)^2}{2\sigma_{w_k}^2} \right). \tag{3.33}
\]

Given the Gaussian distribution of \( t_k \) as in (2.18), \( y_k \) is also Gaussian distributed under \( H_0 \) or \( H_1 \):

\[
y_k \sim \mathcal{N}(NE[|x_k|^2], N\text{var}[|x_k|^2] + \sigma_{w_k}^2), \tag{3.34}
\]

where \( E[|x_k|^2] \) and \( \text{var}[|x_k|^2] \) under \( H_0 \) and \( H_1 \), are given in (2.19) and (2.20) respectively.

Therefore, by independence of the channels, the LRT statistic is derived as follows:

\[
L(y^K_1) = \prod_{k=1}^{K} \frac{p(y_k | H_1)}{p(y_k | H_0)} = \prod_{k=1}^{K} e^{-\frac{(y_k - \sqrt{N}(\beta_1 + \beta_2^2))}{2(N(2\beta_2^2 + 2\beta_1\sigma_{w_k}^2 + \sigma_{x_k}^2) + \sigma_{w_k}^2)} + \frac{(y_k - \sqrt{N}\sigma_{w_k}^2)}{2(N\sigma_{n_k}^2 + \sigma_{w_k}^2)}} \sqrt{\frac{N(2\beta_2^2 + 2\beta_1\sigma_{w_k}^2 + \sigma_{x_k}^2) + \sigma_{w_k}^2}{N\sigma_{n_k}^2 + \sigma_{w_k}^2}}. \tag{3.35}
\]

The LRT decides \( H_1 \) if \( L(y^K_1) > \lambda \), where the threshold \( \lambda \) can be set using an empirical distribution of the log-likelihood ratio under the null hypothesis.

We next derive the unknown-parameter linear test for soft local decisions by evaluating the derivatives of the log-likelihood function at \( \beta = 0 \)

\[
\frac{\partial \ln p(y^K_1 | \beta)}{\partial \beta_{1,k}} \bigg|_{\beta=0} = \frac{N(y_k - \sqrt{N}\sigma_{w_k}^2)(\sigma_{n_k}^2 y_k + \sigma_{w_k}^2)}{(N\sigma_{n_k}^2 + \sigma_{w_k}^2)^2} - \frac{1}{\sigma_{n_k}^4} \tag{3.36}
\]

\[
\frac{\partial \ln p(y^K_1 | \beta)}{\partial \beta_{2,k}} \bigg|_{\beta=0} = \frac{N(y_k - \sqrt{N}\sigma_{w_k}^2)^2}{(N\sigma_{n_k}^2 + \sigma_{w_k}^2)^2} - \frac{1}{\sigma_{n_k}^4} \tag{3.37}
\]
and substituting (3.36) and (3.37) into (3.17). Solving \( \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{1,k}} = 0 \) and \( \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{2,k}} = 0 \) for \( \beta_{1,k} \) and \( \beta_{2,k} \) results in a third-order equation with respect to \( \beta_{1,k} \) which must be solved numerically. Thus, we alternatively use the simplified linear test for this scenario and derive its statistic as

\[
T_{SL}(y^K_1) = \sum_{k=1}^{K} \frac{N(y_k - N\sigma^2_{n_k})}{(N\sigma^4_{n_k} + \sigma^2_{w_k})^2} [y_k(\sigma^2_{n_k} + 1) + \sigma^2_{w_k} - N\sigma^2_{n_k}] - \sum_{k=1}^{K} \frac{\sigma^2_{n_k} + 1}{\sigma^4_{n_k}}. \tag{3.38}
\]

Note that \( T_{SL}(y^K_1) \) depends on \( \sigma^2_{w_k} \) and \( \sigma^2_{n_k} \), the AWGN variances in the SU-FC and PU-SU channels respectively, but not on the primary transmitted signal statistics or the PU-SU fading channel parameter.

### 3.4.2 Unknown Primary Signal Statistics and Known Channel Statistics

**Linear Test**

It may be practical to assume that the second-order statistics of the channel gain from the primary user to secondary users are available at the fusion center. Therefore, the set of unknown parameters in our model will reduce to \( \beta_1 = E[|z|^2] \) and \( \beta_2 = E[|z|^4] \).

In this case, unlike in the previous section where the \( \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{1,k}} \)’s and \( \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{2,k}} \)’s only consisted of terms that depend on \( k \), the \( \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_1} \) and \( \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_2} \) consist of summations of nonlinear terms in \( \beta_1 \) and \( \beta_2 \) over \( k \). Therefore the analytical derivation of the MLEs is infeasible for our general cooperative noisy channel model. Hence, we propose to apply the simplified linear test with statistic

\[
T_{SL}(y^K_1) = \left. \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_1} \right|_{\beta=0} + \left. \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_2} \right|_{\beta=0}. \tag{3.39}
\]

From now on, we use the term linear test instead of the simplified linear test for conciseness. We can apply the linear test to all local decision scenarios discussed in Section 3.4.1 when the channel gain statistics are available but the primary signal statistics are not.
For the binary local decision scenario, for instance, the derivatives of the log-likelihood ratio with respect to $\beta_1$ and $\beta_2$ are derived and evaluated at $\beta = 0$ as follows:

\[
\frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_1} \bigg|_{\beta = 0} = K \sum_{k=1}^{K} \left[ p(y_k|1) - p(y_k|0) \right] e^{-\frac{(\tau_k - N \sigma_n^2)^2}{2N \sigma_n^4}} \frac{\sigma_n^2 \tau_k}{\sqrt{2\pi N \sigma_n^4}}, \tag{3.40}
\]

\[
\frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_2} \bigg|_{\beta = 0} = K \sum_{k=1}^{K} \left[ p(y_k|1) - p(y_k|0) \right] e^{-\frac{(\tau_k - N \sigma_n^2)^2}{2N \sigma_n^4}} \frac{2\sigma_n^4 \tau_k}{\sqrt{2\pi N \sigma_n^4}}, \tag{3.41}
\]

The linear test statistic in this case is derived by inserting (3.40) and (3.41) into (3.39).

We can similarly derive the linear test statistics of the soft and quantized local decisions for known channel parameters and unknown primary signal statistics, but do not show them here for conciseness.

**LMP Test**

When the channel gains are known and the only unknown parameters are the primary signal statistics, we can use the LMP test under certain assumptions. The LMP test is a one-sided scalar-parameter test for small departures of the parameter from zero [56].

The LMP test statistic is derived as

\[
T_{LMP}(y^K_1) = \frac{\partial \ln p(y^K_1|\beta_1, \mathcal{H}_1)}{\partial \beta_1} \bigg|_{\beta_1 = 0} \tag{3.42}
\]

if $\beta_1$ is the only unknown parameter. We use this scalar-parameter test for the case of constant-modulus primary signal with unknown small power of $\beta_1$. In this case, $\beta_2 = \beta_2^2$ and the only unknown parameter is $\beta_1$. This test can be applied to hard, soft and quantized local decision scenarios for constant-modulus primary signals. Here we derive the LMP test statistic for the binary local decision case as an example. For constant-modulus (e.g. $M$-PSK modulated) primary signals, the LRT statistic in (3.30) reduces
Chapter 3. Cooperative Sensing with Unknown Parameters

(3.43)

\[ L(y^K_1) = \prod_{k=1}^{K} \frac{p(y_k|0) + [p(y_k|1) - p(y_k|0)] Q \left( \frac{\tau_k - N(\beta_1 \sigma^2_{hk} + \sigma^2_{nk})}{\sqrt{N(\beta_1 \sigma^2_{hk} + \sigma^2_{nk})}} \right)}{p(y_k|0) + [p(y_k|1) - p(y_k|0)] Q \left( \frac{\tau_k - N\sigma^2_{hk}}{\sqrt{N\sigma^2_{hk}}} \right)}. \]

Therefore, the LMP test statistic is derived as

(3.44)

\[ T_{LMP}(y^K_1) = \sum_{k=1}^{K} \frac{[p(y_k|1) - p(y_k|0)] e^{-\frac{(\tau_k - N\sigma^2_{hk})^2}{2N\sigma^2_{hk}}} \frac{\sigma^2_{hk}}{\sqrt{2\pi N \sigma^2_{hk}}} \frac{\tau_k}{\sqrt{2\pi N \sigma^2_{hk}}}}{p(y_k|0) + [p(y_k|1) - p(y_k|0)] Q \left( \frac{\tau_k - N\sigma^2_{hk}}{\sqrt{N\sigma^2_{hk}}} \right)}. \]

We observe that the LMP test statistic for weak constant-modulus primary signals has a similar expression to the first part of the linear test statistic in (3.40) for general primary signals. For the case of quantized local decisions, we derive the LMP test statistic as

(3.45)

\[ T_{LMP}(y^K_1) = \sum_{k=1}^{K} \frac{\sum_{i=1}^{M_k} p(y_k|q_i) A_{ik} \frac{\sigma^2_{hk}}{\sqrt{2\pi N \sigma^2_{hk}}}}{\sum_{i=1}^{M_k} p(y_k|q_i) \left[ Q \left( \frac{a_{(i-1)k} - \mu_{tk}}{\sigma_{tk}} \right) - Q \left( \frac{a_{ik} - \mu_{tk}}{\sigma_{tk}} \right) \right]}, \]

where \( A_{ik} \) is given in (3.27).

By and large, the proposed linear test can be applied for all ranges of signal power, whereas the LMP is restricted to small signal powers. Moreover, the linear test accommodates multiple unknown parameters, while the LMP is confined to scalar parameter tests.

3.4.3 Threshold Setting at the Fusion Center

Linear Test

In order to set the threshold at the fusion center when using the linear test, we need the distribution of the test statistic \( T_{SL}(y^K_1) \) under \( \mathcal{H}_0 \). Under \( \mathcal{H}_0 \) for which \( \beta = 0 \), by regularity conditions [28, pg. 67] we have

(3.46)

\[ E \left[ \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta} \right]_{\beta=0} = 0 \]
and also
\[
E \left[ \left( \frac{\partial \ln p(y_1^K|\beta)}{\partial \beta} \bigg|_{\beta=0} \right)^2 \right] = I(0) \quad (3.47)
\]
which is the Fisher information matrix at $\beta = 0$. Moreover, the linear test statistic can be written as
\[
T_{SL}(y^K_1) = \sum_{k=1}^{K} \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{1,k}} \bigg|_{\beta=0} + \sum_{k=1}^{K} \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{2,k}} \bigg|_{\beta=0}. \quad (3.48)
\]
Therefore, by CLT we have
\[
T_{SL}(y^K_1) \sim N(0, \sum_{k=1}^{K} \{ I_{\beta_{1,k}}(0) + I_{\beta_{2,k}}(0) + 2 I_{\beta_{1,2,k}}(0) \})
\]
under $H_0$ \quad (3.49)

where
\[
I_{\beta_{i,k}}(0) = E \left[ \left( \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{i,k}} \bigg|_{\beta=0} \right)^2 \right] \quad \text{for } i = 1, 2, \quad (3.50)
\]
and
\[
I_{\beta_{1,2,k}}(0) = E \left[ \left( \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{1,k}} \bigg|_{\beta=0} \right) \left( \frac{\partial \ln p(y^K_1|\beta)}{\partial \beta_{2,k}} \bigg|_{\beta=0} \right) \right]. \quad (3.51)
\]
For binary local decisions and binary symmetric channels between the users and fusion center, with crossover probability of $\alpha_k$, we derive $I_{\beta_{1,k}}(0)$, $I_{\beta_{2,k}}(0)$ and $I_{\beta_{1,2,k}}(0)$ as follows:
\[
I_{\beta_{1,k}}(0) = (1 - 2\alpha_k)^2 e^{-\frac{(\tau_k - N\sigma^2_n) \sqrt{2 \pi N \sigma^2_n}}{\alpha_k}} \left( \frac{\tau_k}{\sqrt{2 \pi N \sigma^2_n}} \right)^2 \frac{x_k(1 - x_k)}{x_k(1 - x_k)}, \quad (3.52)
\]
\[
I_{\beta_{2,k}}(0) = (1 - 2\alpha_k)^2 e^{-\frac{(\tau_k - N\sigma^2_n) \sqrt{2 \pi N \sigma^2_n}}{\alpha_k}} \left( \frac{2(\tau_k - N\sigma^2_n) \sqrt{2 \pi N \sigma^2_n}}{\alpha_k} \right)^2 \frac{x_k(1 - x_k)}{x_k(1 - x_k)}, \quad (3.53)
\]
\[
I_{\beta_{1,2,k}}(0) = (1 - 2\alpha_k)^2 e^{-\frac{(\tau_k - N\sigma^2_n) \sqrt{2 \pi N \sigma^2_n}}{\alpha_k}} \left( \frac{\tau_k}{\sqrt{2 \pi N \sigma^2_n}} \right)^2 \frac{x_k(1 - x_k)}{x_k(1 - x_k)}, \quad (3.54)
\]
where
\[
x_k = \alpha_k + (1 - 2\alpha_k)Q \left( \frac{\tau_k - N\sigma^2_n}{\sqrt{N \sigma^2_n}} \right). \quad (3.55)
\]
For the quantized local decision scenario, these information numbers can be derived as

\[
I_{\beta_1,k}(0) = \sum_{y_k=0}^{M-1} \frac{\left( \sum_{i=1}^{M_k} p(y_k|q_i) A_{ik} \frac{1}{\sqrt{2\pi N \sigma^2_{\eta_k}}} \right)^2}{\sum_{y_k=0}^{M-1} p(y_k|q_i) \left[ Q \left( \frac{a_i - N\sigma^2_{\eta_k}}{\sqrt{N\sigma^2_{\eta_k}}} \right) - Q \left( \frac{a_i - N\sigma^2_{\eta_k}}{\sqrt{N\sigma^2_{\eta_k}}} \right) \right]}, \tag{3.56}
\]

\[
I_{\beta_2,k}(0) = \sum_{y_k=0}^{M-1} \frac{\left( \sum_{i=1}^{M_k} p(y_k|q_i) B_{ik} \frac{2}{\sqrt{2\pi N \sigma^2_{\eta_k}}} \right)^2}{\sum_{y_k=0}^{M-1} p(y_k|q_i) \left[ Q \left( \frac{a_i - N\sigma^2_{\eta_k}}{\sqrt{N\sigma^2_{\eta_k}}} \right) - Q \left( \frac{a_i - N\sigma^2_{\eta_k}}{\sqrt{N\sigma^2_{\eta_k}}} \right) \right]}, \tag{3.57}
\]

\[
I_{\beta_1,2,k}(0) = \sum_{y_k=0}^{M-1} \frac{\left( \sum_{i=1}^{M_k} p(y_k|q_i) A_{ik} \frac{1}{\sqrt{2\pi N \sigma^2_{\eta_k}}} \right) \left( \sum_{i=1}^{M_k} p(y_k|q_i) B_{ik} \frac{2}{\sqrt{2\pi N \sigma^2_{\eta_k}}} \right)}{\sum_{y_k=0}^{M-1} p(y_k|q_i) \left[ Q \left( \frac{a_i - N\sigma^2_{\eta_k}}{\sqrt{N\sigma^2_{\eta_k}}} \right) - Q \left( \frac{a_i - N\sigma^2_{\eta_k}}{\sqrt{N\sigma^2_{\eta_k}}} \right) \right]}, \tag{3.58}
\]

where \(A_{ik}\) and \(B_{ik}\) are given in (3.27) and (3.28). Thus, using the derived distribution of the linear test statistic under \(H_0\), the probability of false alarm for global threshold of \(\lambda\) is given by

\[
P_f = P\{T_{SL}(y^K_1) > \lambda|H_0\} = Q\left( \frac{\lambda}{\sqrt{\sum_{k=1}^{K} \{I_{\beta_1,k}(0) + I_{\beta_2,k}(0) + 2I_{\beta_1,2,k}(0)\}}} \right). \tag{3.59}
\]

In this way, we can analytically set the global threshold at the fusion center to maintain the target probability of false alarm, \(P_f\), as

\[
\lambda = Q^{-1}(P_f) \sqrt{\sum_{k=1}^{K} \{I_{\beta_1,k}(0) + I_{\beta_2,k}(0) + 2I_{\beta_1,2,k}(0)\}}. \tag{3.61}
\]

**LMP Test**

The LMP test statistic in (3.42) can be rewritten as

\[
T_{LMP}(y^K_1) = \sum_{k=1}^{K} \frac{\partial \ln p(y_k|\beta_1, H_1)}{\partial \beta_1} |_{\beta_1=0}. \tag{3.62}
\]

Therefore, by CLT and regularity condition

\[
E \left[ \frac{\partial \ln p(y_k|\beta_1)}{\partial \beta_1} |_{\beta_1=0} \right] = 0, \tag{3.63}
\]
we have

\[ T_{\text{LMP}}(y^K_1) \sim \mathcal{N}(0, \sigma^2_{T_{\text{LMP}}}) , \text{under } \mathcal{H}_0 \]  

(3.64)

where

\[
\sigma^2_{T_{\text{LMP}}} = \sum_{k=1}^{K} E \left[ \left( \frac{\partial \ln p(y_k | \beta)}{\partial \beta_1} |_{\beta_1=0} \right)^2 \right].
\]  

(3.65)

Thus, the global threshold at the fusion center to maintain a probability of false alarm of \( P_f \) is analytically determined as

\[
\lambda = Q^{-1}(P_f) \sqrt{\sum_{k=1}^{K} E \left[ \left( \frac{\partial \ln p(y_k | \beta)}{\partial \beta_1} |_{\beta_1=0} \right)^2 \right]}. \]  

(3.66)

In the quantized local decision scenario, for example, the variance of the LMP test statistic under the null hypothesis is derived as

\[
\sigma^2_{T_{\text{LMP}}} = \sum_{k=1}^{K} \sum_{y_k=0}^{M-1} \sum_{i=1}^{M_k} p(y_k | q_i) A_{ik} \frac{\sigma^2_{hk}}{2\pi N \sigma^2_{nk}} \left[ Q \left( \frac{a_{(i-1)k} - N \sigma^2_{nk}}{\sqrt{N \sigma^2_{nk}}} \right) - Q \left( \frac{a_{ik} - N \sigma^2_{nk}}{\sqrt{N \sigma^2_{nk}}} \right) \right]. \]  

(3.67)

### 3.5 Simulation Results

In this section, we present our simulation results to evaluate the performance of the linear test in cooperative spectrum sensing. We consider the cases of hard, soft and quantized local decisions and use the performance of NP-based LRT cooperative sensing as an upperbound for comparison. The primary signal is assumed to be 4-ASK modulated (multi-level) from the set \(-A, -A/3, A/3, A\) where \(A\) is the maximum primary signal voltage. This modulation is used in order to examine the non-constant-modulus scenario for the primary signal where unlike the constant-modulus scenario, there are more than one unknown primary parameters. The local decisions at the secondary users are BPSK modulated and sent to the fusion center through independent binary symmetric channels for hard and quantized local decisions.

Fig. 3.2 depicts the complementary ROC curves of the NP-based LRT and the linear test with unknown primary statistics for hard local decisions. The points of these curves
Figure 3.2: Complementary ROC ($P_m$ vs. $P_f$) curves for hard local decisions with unknown primary statistics and $K = 3$.

are achieved by changing the global threshold at the fusion center. We consider $K = 3$ secondary users in the network which independently sense the primary user’s spectrum. The Rayleigh fading channel coefficients are assumed to be independently drawn from $\mathcal{CN}(0, \sigma^2_{h_k})$, where the variances $\sigma^2_{h_k}$ s for the three PU-SU channels are assumed to be 0.32, 4.5 and 0.08 respectively. The local energy detection thresholds are chosen to achieve a target local probability of false alarm of $p_f = 0.1$, where the local noise variances are set to be $\sigma^2_n = \sigma^2_{h_k}$. The number of sensing samples are assumed to be $N = 100$. The crossover probabilities for the six SU-FC BSC channels are 0.05, 0.1 and 0.05 respectively. We observe that the linear test has performance close to that of the optimal NP-based LRT fusion for the cases of both $A^2 = 2$ and $A^2 = 4$. In the first case $\beta_1 = 1.11$ and
Figure 3.3: Complementary ROC ($P_m$ vs. $P_f$) curves for hard local decisions with unknown primary statistics and $K = 3$.

$\beta_2 = 2.02$, whereas the difference between the two parameters is even higher in the second case where $\beta_1 = 2.22$ and $\beta_2 = 8.09$.

Fig. 3.3 depicts the complementary ROC curves of the NP-based LRT and the linear test with unknown primary statistics for hard local decisions and different values of $N$. We consider $K = 3$ secondary users and the same channel and noise parameters as the previous figure. We observe that the simplified linear test performs closely to the optimal NP-based LRT fusion for small as well as large numbers of sensing samples.

Fig. 3.4 depicts the complementary ROC curves of the NP-based LRT and the linear test with unknown primary statistics for hard local decisions. We consider $K = 6$ secondary users in the network which independently sense the primary user’s spectrum with
channel gain variances $\sigma_{h_k}^2$ assumed to be 1.28, 0.32, 4.5, 0.5, 2 and 0.08 respectively. The local noise variances are set to be $\sigma_n^2 = \sigma_{h_k}^2$. The crossover probabilities for the six SU-FC BSC channels are 0.1, 0.15, 0.1, 0.2, 0.15 and 0.1 respectively. We observe that the simplified linear test performs closely to the optimal NP-based LRT fusion for $K = 6$ secondary users.

Fig. 3.5 depicts the complementary ROC curves of the linear test with unknown primary statistics and NP-based LRT for soft local decisions. The local noise variances of $\sigma_n^2 = 1$ and $\sigma_n^2 = 2$ have been considered. The AWGN channels are generated from a Gaussian distribution of $\mathcal{N}(0, \sigma_w^2)$ where the variances for the six SU-FC channels are assumed to be 0.64, 1, 2.25, 0.81, 1 and 1.21 respectively. We observe that the linear test
performs closely to the optimal NP-based LRT fusion for the soft local decision scenario.

Fig. 3.6 depicts the complementary ROC curves of NP-based LRT and the linear test with unknown primary statistics for 2-bit quantized local decisions. The quantization thresholds are set to be $a_{1k} = \tau_k/2$, $a_{2k} = \tau_k$ and $a_{3k} = 2\tau_k$, where $\tau_k$ is determined as (3.31) for $p_f = 0.05$, and $A$ is assumed to be 0.8. The crossover probabilities of $\alpha = 0.05$ and $\alpha = 0.1$ are considered for the BSC SU-FC channels. We observe that the linear test performs closely to the optimal NP-based LRT.

Fig. 3.7 depicts the complementary ROC curves of NP-based LRT and linear test with unknown channel gains and primary statistics for 2-bit quantized local decisions. The BSC cross-over probabilities of 0.1 and 0.05, local noise and channel gain variance of 1 and maximum transmit voltage of 1 have been assumed. We observe that the linear test
performs comparably to the optimal NP-based LRT with known primary and channel parameters. As the channel gain variance is not known to the FC, the performance gap between the LRT and the linear test is more considerable compared to the previous figure. Our simulation results show that with knowledge of the channel gain variance at the FC, the performance of the linear test becomes closer to that of the LRT.

Fig. 3.8 shows the analytical and numerical probability distribution of the linear test statistic under the null hypothesis. We observe that the analytical pdf derived in (3.49) is very close to the numerical pdf which represents the exact pdf.
Figure 3.7: Complementary ROC ($P_m$ vs. $P_f$) curves for 2-bit quantized local decisions with unknown primary and channel gain statistics.

### 3.6 Chapter Summary

In this chapter, we presented a composite hypothesis testing approach for cooperative spectrum sensing in cognitive radios. For non-ideal transmission channels, we derived the LRT statistic based on the Neyman-Pearson criterion at the fusion center. However, due to lack of knowledge about the primary signal and channel statistics at the secondary networks, the optimal LRT-based scheme which depends on these parameters, cannot be applied in most cognitive radio networks. For this reason, we proposed linear composite hypothesis testing methods, which do not require any prior knowledge about the primary signal and channel statistics. In its simplified form, this test is more practical and simpler compared to the GLRT which requires ML estimates of the unknown parameters. We
Figure 3.8: Probability distribution function of the linear test statistic under the null hypothesis.

derived the linear test statistic for decision making at the fusion center under hard, soft and quantized local decision scenarios. Furthermore, we applied the LMP detector at the fusion center for weak constant-modulus primary signals and derived its corresponding test statistic. Deriving the pdf of the linear test and LMP statistics under the signal-absent hypothesis, we analytically determined the linear test and LMP thresholds at the fusion center. We showed that applying the linear test for cooperative sensing results in very close performance to that of the known-parameter NP-based LRT, without requiring prior knowledge of the primary statistics at the fusion center.
Chapter 4

Cooperative Quickest Spectrum Sensing with Unknown Parameters

A large and growing body of work on spectrum sensing is focused on block detection methods, in which the detector makes its decision based on a window or block of observations, using classical statistical detection theory (see e.g. [9, 10, 17–19, 21–24]). However, in practice, changes in primary-network activity within the frequency band of interest should be sensed as quickly as possible, in order to maximally utilize any vacated spectrum, as well as to minimize interference to the primary network after it resumes transmission. Therefore, detection delay is an important criterion in spectrum sensing, especially in bursty primary networks.

Sequential change-point detection (quickest detection) provides a theoretical framework for minimizing detection delay, and is therefore well-matched to spectrum sensing in cognitive radios [30]. The well-known Page’s cumulative sum (CUSUM) algorithm [33] has been shown to be optimal in the sense of minimizing the detection delay while maintaining an acceptable level of false alarm [34]. The CUSUM algorithm was applied to spectrum sensing, in the context of a single sensor, in [30] and [31]. However, at low SNRs, spectrum sensing based on a single user’s observations may not be reliable, and a
diversity of sensors at various locations is known to give better results through cooperative or collaborative sensing.

In this chapter, we first present cooperative quickest spectrum sensing schemes based on the CUSUM algorithm. The study of distributed quickest detection has been based on two different formulations in the detection literature. One is based on a Bayesian formulation in which the change point (i.e. time at which the change occurs) is assumed to have a known prior distribution [60]. In this case, the joint optimization of local users and the fusion center policies over time becomes intractable [37]. The other formulation is the minimax formulation, proposed by Lorden [34], for which Page’s CUSUM procedure is the optimal scheme. The asymptotically optimal decentralized CUSUM approaches for limited and full local memory scenarios have been studied by Mei [37]. In this chapter, we apply and compare three CUSUM-based cooperative spectrum sensing schemes in cognitive radios mainly based on Mei’s studies [37]: global CUSUM with soft local decisions, global CUSUM with quantized local decisions and hard fusion of local CUSUM. We derive the local and global test statistics in each scheme assuming noisy channels between the secondary users and the fusion center.

Next, we address the issue of limited knowledge about the primary signal at the secondary users. In particular, when the primary user starts transmission, there are unknown parameters in the distribution of the observed signals, e.g. the variance of the observed signals, due to unknown primary signal statistics and channel gains between the primary and secondary users. Therefore, the CUSUM-based approaches that are based on perfectly known distributions cannot be directly applied to spectrum sensing in cognitive radios. The generalized likelihood ratio test (GLRT) was adopted by Lorden [15] for CUSUM with unknown parameters. However, due to the non-recursive expression of the GLRT-based CUSUM and the need to store all the observations and re-estimate the unknown parameters in all time slots, this algorithm turns out to be infeasible and impractical.
As an alternative, we propose a linear test for quickest spectrum sensing which does not require maximum likelihood (ML) estimates of the unknown parameters and has the same asymptotic performance as the GLRT. We derive the test statistics of the linear test version of the CUSUM-based cooperative sensing schemes and show that the CUSUM-based schemes using the linear test perform comparably to the optimal CUSUM-based schemes, without requiring knowledge of the primary signal or channel gains. Unlike the successive refinement algorithm in [31] and the adaptive parameter tracking approach in [61], our scheme does not require any iterative algorithm or any parameter estimation and, thus, it is much faster and less complex to implement. It has also a much simpler implementation and much better performance compared to the nonparametric algorithm adopted in [30].

For performance evaluation and threshold setting, we apply the matrix approach in [62] which is based on finding the cumulative distribution function (CDF) of the stopping time through an underlying state-update process. We provide the threshold-setting scheme for both known and unknown-parameter CUSUM-based algorithms. This chapter has been partly published in [63].

The remainder of this chapter is organized as follows. The system model is described in Section 4.1. The CUSUM-based cooperative quickest sensing schemes with known parameters are presented in Section 4.2. The cooperative quickest detection schemes with unknown parameters are presented in Section 4.3 where the linear test is applied to the CUSUM-based schemes. The threshold setting scheme is described in Section 4.4. The simulation results are provided in Section 4.5. Finally, Section 4.6 concludes the chapter.
4.1 System Model

We assume $L$ secondary users are monitoring the frequency band allocated to the primary user. Let $x_{l,i}$ denote the observation at the $l$th secondary user at time $i$. When the primary user is not active $x_{l,i} = n_{l,i}$, where $n_{l,i}$ is zero-mean white Gaussian noise with variance $\sigma_n^2$. When the primary user is using the band $x_{l,i} = z_{l,i} + n_{l,i}$, where $z_{l,i}$ is the faded received primary signal at the $l$th secondary user. We assume time-invariant channel gain $h_l$ between the primary user and the $l$th secondary user, and $z_{l,i} = h_l s_i$ to be zero-mean Gaussian with variance $\sigma_z^2$.

In this chapter, we focus on the detection of the emergence of the primary signal in the band. The detection of vacancy in the band can be tackled similarly as we discuss in the next chapter. We assume that the primary user is initially inactive, and at an unknown time $\tau$ it becomes active. Thus there is a change in the distribution of the observations from $\mathcal{N}(0, \sigma_n^2)$ to $\mathcal{N}(0, \sigma_z^2 + \sigma_n^2)$ at unknown time $\tau$. The sequential change detection modeling of this problem is to detect this change as quickly as possible while maintaining a certain false alarm probability. Defining the detection time $T$ as the time at which the change is detected, if $T > \tau$ then the detection delay is $\delta = T - \tau$. If $T < \tau$, a false alarm has occurred with the average run length (ARL) of false alarm being $T_f = E[T|T < \tau]$. The detection of the primary user vacating the band can be approached similarly. The non-Bayesian approach to this problem based on Lorden’s formulation [34] minimizes the average detection delay,

$$T_d = \sup_{\tau \geq 1} E_{\tau} [T - \tau | T \geq \tau]$$

while maintaining the ARL of false alarms larger than a certain threshold, $T_f \geq \lambda$.

The algorithm that finds the minimum $T_d$ in this problem, is Page’s CUSUM algorithm (a.k.a. the Page test) [33]. The stopping time in this algorithm for a single user detection problem is determined as follows:

$$T(q) = \inf \{ n : C_n \geq q \}$$
with cumulative statistic

\[ C_n = \max_{1 \leq k \leq n} \sum_{i=k}^{n} L_i(x_i) \]  

(4.3)

where

\[ L_i(x_i) = \log \frac{f_1(x_i)}{f_0(x_i)} \]  

(4.4)

and \( f_0(x_i) \) and \( f_1(x_i) \) are the distributions of \( x_i \) before and after the change, respectively.

The cumulative statistic \( C_n \) can be recursively calculated for \( n \geq 1 \) as follows:

\[ C_n = \max(C_{n-1}, 0) + L_n(x_n) \]  

(4.5)

where \( C_0 = 0 \). In other words, this algorithm finds the first \( n \) for which \( C_n \geq q > 0 \).

The only analytical results on quickest detection performance bounds and their relations to the threshold \( q \) deal with asymptotic scenarios where \( q \to \infty \) [29, 37], which are not practical. The analysis in [37] may be used to asymptotically bound the performance of the schemes in the next section.

### 4.2 Cooperative CUSUM with Full Knowledge of Parameters

Due to fading and shadowing effects in cognitive radio networks, detection performance improves with the number of available sensors that make independent observations. In this section, we introduce cooperative quickest sensing schemes for cognitive radios, given full knowledge of all parameters. We apply CUSUM-based approaches mainly based on [37] to various local memory and transmission rate scenarios, and derive the necessary test statistics. In all these methods, the local decisions are sent to a decision making fusion center, which makes the global final decision. In the next section, we will tackle the problem of unknown parameters.
4.2.1 Global CUSUM

Assuming noiseless error-free channels between the secondary users and the fusion center, the local observations can be completely provided to the fusion center. Thus, the global CUSUM can be applied at the fusion center based on the local observations. In this case, based on [37], the CUSUM test statistic at the fusion center can be written as

\[ C_n = \max_{1 \leq k \leq n} \sum_{i=k}^{n} \sum_{l=1}^{L} L_i^l(x_{l,i}). \]  

(4.6)

where \( L_i^l(x_{l,i}) \) represents the log-likelihood function at time \( i \) at the \( l \)th user. In our assumed model, \( f_{l,0}^1(x_{l,i}) = \frac{1}{\sqrt{2\pi\sigma_{n_l}}} e^{-\frac{x_{l,i}^2}{2\sigma_{n_l}^2}} \) and \( f_{l,1}^1(x_{l,i}) = \frac{1}{\sqrt{2\pi(\sigma_{n_l}^2 + \sigma_{z_l}^2)}} e^{-\frac{x_{l,i}^2}{2(\sigma_{n_l}^2 + \sigma_{z_l}^2)}} \), thus the global CUSUM statistic is derived as

\[ C_n = \max_{1 \leq k \leq n} \sum_{i=k}^{n} \sum_{l=1}^{L} \left( \frac{1}{2} \log \left( \frac{\sigma_{n_l}^2}{\sigma_{n_l}^2 + \sigma_{z_l}^2} \right) + \frac{x_{l,i}^2 \sigma_{n_l}^2}{2\sigma_{n_l}^2 (\sigma_{n_l}^2 + \sigma_{z_l}^2)} \right). \]  

(4.7)

The fusion center declares that the primary user is active as soon as \( C_n \) exceeds threshold \( q \). In practice, due to communication constraints in cognitive radios, the exact local observations cannot be completely provided to the fusion center. Therefore, we consider the quantized local observation scenario in the next section.

4.2.2 Global CUSUM with Quantized Local Decision

The case of unlimited bandwidth between the fusion center and the secondary users have been considered in the previous section. However, bandwidth constraints in wireless communications implies quantization of the local observations before they are sent to the fusion center. For memoryless local users, it has been shown that the monotone likelihood ratio quantizer (MLRQ) is the optimal local mapping [37]. In this mapping, the \( l \)th user local decision at time \( n \) is \( U_{l,n} = d \) if and only if

\[ a_{l,d} \leq \frac{f_1(x_{l,n})}{f_0(x_{l,n})} < a_{l,d+1} \]  

(4.8)
where \( 0 = a_{l,0} \ldots \leq a_{l,D_l-1} \leq a_{l,D_l} = \infty \) are the \( l \)th user quantization thresholds. These local decisions are sent to the fusion center over noisy channels characterized by \( p(y_{l,k} | U_{l,k}) \). The fusion center applies Page’s CUSUM scheme with stopping time

\[
T(q) = \inf \{ n : C_n \geq q \} \tag{4.9}
\]

with \( C_0 = 0 \) and

\[
C_n = \max(C_{n-1}, 0) + \sum_{l=1}^{L} L_l^i(y_{l,n}), \tag{4.10}
\]

where \( L_l^i(y_{l,n}) \) is the log-likelihood ratio of \( y_{l,n} \) at the fusion center. Using the graphical approach in Chapter 2, we derive this log-likelihood ratio as follows:

\[
L_l^i(y_{l,k}) = \log \frac{\sum_{d=1}^{D_l} p(y_{l,k} | d) P[a_{l,d-1} \leq lr(x_{l,k}) \leq a_{l,d} | \mathcal{H}_1]}{\sum_{d=1}^{D_l} p(y_{l,k} | d) P[a_{l,d-1} \leq lr(x_{l,k}) \leq a_{l,d} | \mathcal{H}_0]} \tag{4.11}
\]

where \( lr(x_{l,k}) = \frac{f_r(x_{l,k})}{f_0(x_{l,k})} \) and \( \mathcal{H}_0 \) and \( \mathcal{H}_1 \) represent \( x_{l,n} \sim \mathcal{N}(0, \sigma_{n_l}^2) \) and \( x_{l,n} \sim \mathcal{N}(0, \sigma_{n_l}^2 + \sigma_{z_l}^2) \) respectively. In order to calculate \( P[a_{l,d-1} \leq lr(x_{l,k}) \leq a_{l,d} | \mathcal{H}_\theta] \), for \( \theta = 1, 2 \), we need to translate the MLRQ thresholds into thresholds for the observations \( x_{l,k} \). For Gaussian distributed local observations, the likelihood ratio is derived as

\[
lr(x_{l,k}) = \frac{\sigma_{n_l}}{\sqrt{\sigma_{n_l}^2 + \sigma_{z_l}^2}} \exp \left( \frac{\sigma_{z_l}^2 x_{l,k}^2}{(\sigma_{n_l}^2 + \sigma_{z_l}^2) \sigma_{n_l}^2} \right). \tag{4.12}\]

Thus, \( lr(x_{l,k}) \geq a_{l,k} \) is equivalent to

\[
x_{l,k}^2 \geq \log \left( \frac{a_{l,d} \sqrt{\sigma_{n_l}^2 + \sigma_{z_l}^2}}{\sigma_{n_l}} \right) \left( \frac{\sigma_{n_l}^2 + \sigma_{z_l}^2}{\sigma_{z_l}^2} \right) \sigma_{n_l}^2 \sigma_{z_l}^2 \tag{4.13}
\]

where the right hand side of the inequality (4.13) is the equivalent threshold for \( x_{l,k}^2 \), which we call \( a'_{l,d} \). Given the Gaussian distribution of \( x_{l,k} \), \( x_{l,k}^2 \) will be exponentially distributed with parameters \( 2\sigma_{n_l}^2 \) and \( 2 \sqrt{\sigma_{n_l}^2 + \sigma_{z_l}^2} \), under \( \mathcal{H}_0 \) and \( \mathcal{H}_1 \) respectively. Therefore,

\[
P[a'_{l,d-1} \leq x_{l,k}^2 \leq a'_{l,d} | \mathcal{H}_0] = e^{-\frac{a'_{l,d-1}}{\sigma_{n_l}}} - e^{-\frac{a'_{l,d}}{\sigma_{n_l}}},
\]

\[
P[a'_{l,d-1} \leq x_{l,k}^2 \leq a'_{l,d} | \mathcal{H}_1] = e^{-\frac{a'_{l,d-1}}{2 \sqrt{\sigma_{n_l}^2 + \sigma_{z_l}^2}}} - e^{-\frac{a'_{l,d}}{2 \sqrt{\sigma_{n_l}^2 + \sigma_{z_l}^2}}}.\]
Consequently, the log-likelihood ratio is derived as

\[
L_l^k(y_{l,k}) = \log \frac{\sum_{d=0}^{D_l-1} p(y_{l,k} | d) \left[ e^{-\frac{a_l'}{\sigma_n l}} - e^{-\frac{a_l'}{\sigma_n l}} \right]}{\sum_{d=0}^{D_l-1} p(y_{l,k} | d) \left[ e^{-\frac{a_l'}{2\sigma_n l}} - e^{-\frac{a_l'}{2\sigma_n l}} \right]}.
\]

(4.14)

For the case of binary local decisions with thresholds \(a_l'\), the likelihood ratio becomes

\[
L_l^k(y_{l,k}) = \log \frac{p(y_{l,k} | 0) + [p(y_{l,k} | 1) - p(y_{l,k} | 0)] e^{-\frac{a_l'}{\sigma_n l}}}{p(y_{l,k} | 0) + [p(y_{l,k} | 1) - p(y_{l,k} | 0)] e^{-\frac{a_l'}{2\sigma_n l}}}.
\]

(4.15)

The algorithm stops and the presence of the primary user is detected right after \(C_n \geq q\).

### 4.2.3 Hard Fusion of Local CUSUM

In this section we assume that each secondary user has sufficient memory to individually perform the CUSUM algorithm and send its local decision to the fusion center. The CUSUM test statistic at the \(l\)th user is derived as

\[
C_n^l = \max_{1 \leq k \leq n} \sum_{i=k}^{n} L_l^i(x_{l,i}),
\]

(4.16)

where the log-likelihood ratio \(L_l^i(x_{l,i})\) is derived as

\[
L_l^i(x_{l,i}) = \frac{1}{2} \log \left( \frac{\sigma_n^2}{\sigma_n^2 + \sigma_z^2} \right) + \frac{x_{l,i}^2}{2\sigma_n^2 (\sigma_n^2 + \sigma_z^2)}.
\]

(4.17)

After each update the \(l\)th user sends its local decision

\[
U_{l,n} = \begin{cases} 
1 & C_n^l \geq q_l \\
0 & \text{otherwise},
\end{cases}
\]

(4.18)

to the fusion center. The fusion center can use any of the well-known hard-decision combining schemes, such as OR, AND, \(M\)-out-of-\(L\) rules, to make the final decision. The stopping time in the \(M\)-out-of-\(L\) rule scenario can be written as

\[
T(q) = \inf \{ n : C_n^l \geq q_l \text{ for at least } M \text{ out of } L \text{ users} \}
\]

(4.19)

where \(q = [q_1, \ldots, q_L]\) captures all \(L\) users’ local thresholds.
Chapter 4. Cooperative Quickest Sensing with Unknown Parameters

4.3 Cooperative CUSUM with Unknown Parameters

All the CUSUM-based approaches are based on perfect knowledge of the likelihood functions. In cognitive radios, however, the parameters of the distribution under $\mathcal{H}_1$ are not known due to unknown primary signal statistics and channel gains. Therefore, detection schemes that accommodate unknown parameters should be employed.

4.3.1 The GLRT for Quickest Detection

Lorden [34] proposed to apply the GLRT in CUSUM algorithms with uncertain parameters. Let us denote the unknown parameters by $\beta$. The GLRT replaces the unknown parameter with its ML estimate. Thus the basic CUSUM statistic in (4.3) can be replaced with

$$C_n = \max_{1 \leq k \leq n} \max_{\beta} \sum_{i=k}^{n} \log \frac{f_1(x_i|\beta)}{f_0(x_i)}. \quad (4.20)$$

Equation (4.20) implies that we can no longer have a recursive expression for the CUSUM as in (4.5), since the GLRT requires a re-calculation of $\beta$ for each $k$, using all the observations up to the current time $n$. Thus, in practice, the GLRT approach is hard or even infeasible to be implemented in CUSUM-based systems. Therefore, for quickest detection schemes, we propose to use the asymptotically equivalent tests for GLRT that do not require the ML estimate of the unknown parameters and are practical to implement.

4.3.2 The Rao Test for Quickest Detection Schemes

The Rao test is asymptotically equivalent to the GLRT, and does not require the ML estimate of the unknown parameters thereby making it simpler computationally [15]. In classical detection, the Rao test statistic for observation set $x = \{x_1, \ldots, x_n\}$ is given by

$$R(x) = \frac{\partial \log f_1(x|\beta)}{\partial \beta} \bigg|_{\beta=0}^T \left( \frac{\partial \log f_1(x|\beta)}{\partial \beta} \bigg|_{\beta=0} \right)^{-1} \frac{\partial \log f_1(x|\beta)}{\partial \beta} \bigg|_{\beta=0} \quad (4.21)$$
where $I(0)$ is the Fisher information matrix (see [28] pg. 40) evaluated at $\beta = 0$. However, as the quadratic form implies the Rao test cannot result in a recursive expression for the CUSUM as in (4.5) and thus, is not applicable to CUSUM-based approaches. The calculation of the inverse Fisher information matrix, on the other hand, is computationally demanding in multi-parameter scenarios. Therefore, in the next section we propose a linear test which results in a recursive expression for the test statistic and is simple to implement.

### 4.3.3 The Proposed Linear Test for Quickest Detection Schemes

The standard Rao test is a two-sided test, which necessitates the quadratic form in (4.20). In our model, however, the unknown parameters are $\sigma_2 z_l$ which are nonnegative. As a result, the parameter test is one-sided which eliminates the necessity of the quadratic form. Therefore, we propose a modified test with a linear statistic of

$$\Lambda(x) = \sum_{l=1}^{L} \frac{\partial \log f_1(x|\beta_l)}{\partial \beta_l} \bigg|_{\beta_l=0}.$$  \hspace{1cm} (4.22)

As (4.22) shows, the linear test replaces each log-likelihood ratio term depending on $\beta_l$, with its derivative at $\beta_l = 0$ which distributes over the summation in CUSUM test statistic unlike the ML estimation in the GLRT approach in (4.20). Therefore, unlike GLRT-based approaches it does not require storing the observations and re-estimating the parameters at every sampling interval. The linear test is easily applicable to CUSUM-based algorithms, unlike the Rao test which has a quadratic form and does not allow for a recursive expression for the test statistic.

In this section, we apply the linear test (4.22) to the cooperative CUSUM-based schemes discussed in Section 4.2, when the received signal variances are not known upon primary user activity. All the proposed linear-based schemes in this section, do not require any iterative algorithm and any parameter estimation. As a result, they are faster and less complicated compared to the successive refinement algorithm in [31] and the
nonparametric algorithm adopted in [30] which are applied for non-cooperative quickest detection.

**Global CUSUM with linear test**

Although the primary signal and channel statistics may not be known at the fusion center, we can still apply the global CUSUM algorithm by adopting the linear-based approach as follows. By denoting \( \beta_l = \sigma^2 z_l \) and replacing the \( l \)th log-likelihood ratio term in (4.7) with its derivative with respect to \( \beta_l \) evaluated at \( \beta_l = 0 \), we derive the recursive linear-based global CUSUM statistic as follows:

\[
C^\lambda_n = \max(C^\lambda_{n-1}, 0) + \sum_{l=1}^{L} \frac{1}{2\sigma^2_n} \left( \frac{x_{l,n}^2}{\sigma^2_n} - 1 \right)
\]  

(4.23)

where \( C^\lambda_0 = 0 \) and the superscript \( \lambda \) represents the linear test approach. The algorithm will stop as soon as it finds the first \( n \geq 1 \) that satisfies \( C^\lambda_n \geq q^\lambda \).

**Global CUSUM using quantized local decisions with linear test**

The linear-based algorithm for global CUSUM with local MLRQ can be developed as follows. We replace the \( l \)th log-likelihood ratio term in (4.10) with its derivative evaluated at \( \beta_l = 0 \) and derive the linear-based global CUSUM recursively as

\[
C^\lambda_n = \max(C^\lambda_{n-1}, 0) + \sum_{l=1}^{L} \Lambda^l_n(y_{l,n})
\]  

(4.24)

where

\[
\Lambda^l_n(y_{l,n}) = \frac{\sum_{d=0}^{D_l-1} p(y_{l,n}|d) \left[ a^l_{d} e^{\frac{a^l_{d}}{2\sigma^2_n}} - a^l_{d+1} e^{\frac{a^l_{d+1}}{2\sigma^2_n}} \right]}{\sum_{d=0}^{D_l-1} p(y_{l,n}|d) \left[ e^{\frac{a^l_{d}}{2\sigma^2_n}} - e^{\frac{a^l_{d+1}}{2\sigma^2_n}} \right]}
\]  

(4.25)

and \( C^\lambda_0 = 0 \). The primary user is declared to be present right after \( C^\lambda_n \) goes beyond threshold \( q^\lambda \).
Hard fusion of local CUSUM with linear test

When the secondary users have sufficient memory but no knowledge of the primary signal and channel statistics, they can individually perform the linear-based CUSUM algorithm. The test statistic at the $l$th user can be recursively derived by taking derivative from (4.16) with respect to $\beta_l = \sigma_z^2$ and evaluating it at $\beta_l = 0$:

$$C_{l,n}^\lambda = \max(C_{l,n-1}^\lambda, 0) + \frac{1}{2\sigma_n^2} \left( \frac{x_{l,n}^2}{\sigma_n^2} - 1 \right)$$

where $C_{l,0}^\lambda = 0$. The $l$th user decision sent to the fusion center at time $n$ is determined as

$$U_{l,n}^\lambda = \begin{cases} 1 & C_{l,n}^\lambda \geq q_l^\lambda \\ 0 & \text{otherwise} \end{cases}$$

The global stopping time depends on the combining rule at the fusion center. For OR combining rule, for example, the algorithm stops as soon as $C_{l,n}^\lambda$ exceeds $q_l^\lambda$ for at least one of the users.

### 4.4 Performance Analysis and Threshold Setting

The performance of the standard Page test for detection of a one-time permanent change has been studied in terms of average run length, mainly in asymptotic form [34]. In cognitive radios, the probability of false alarm is a useful performance metric. The probability of false alarm can be defined as $P[T(q) < \tau]$, i.e., the probability that the fusion falsely decides that the primary user has become active while it has not. Several methods have been proposed in [62] for performance analysis of the Page test in terms of probability of false alarm. These methods determine the CDF of the stopping time.

In this section, we use the matrix approach in [62] to find the probability of false alarm. In this approach the Page’s test update $(L_n(x_n))$ is uniformly quantized with a step size of $\Delta$. Let the quantization levels be $l_i = i\Delta$, $i = 0, \pm1, \pm2, \ldots$. Let $p_{i}^n$ denote the probabilities of the Page’s test update, $L_n(x_n)$, taking value $l_i$ at time $n$, under $\mathcal{H}_0$. 
We define a state vector $\mathbf{u}_n$ as the vector of probabilities of $L_n(x_n)$ lying in each quantization level at time $n$, i.e.

$$\mathbf{u}_n = [p^n_0, p^n_1, \ldots, p^n_{\gamma-1}]'$$

(4.28)

where $\gamma$ is the integer satisfying $l_{\gamma-1} < q < l_\gamma$ for the test threshold $q$ (i.e., $\gamma$ is the threshold index), the superscript $'$ represents the matrix transpose operation, and

$$p^n_i = P \{ i\Delta - \frac{\Delta}{2} \leq L_n(x_n) < i\Delta + \frac{\Delta}{2} | H_0 \}.$$  

(4.29)

As described in [62], the one step update can be written as

$$\mathbf{u}_{n+1} = \mathbf{D}_n \mathbf{u}_n$$

(4.30)

where $\mathbf{1}$ is a vector of ones and the square matrix $\mathbf{D}_n$ is

$$\mathbf{D}_n = \begin{bmatrix}
\sum_{i=-\infty}^{0} p^n_i & \sum_{i=-\infty}^{-1} p^n_i & \cdots & \sum_{i=-\infty}^{-(\gamma-1)} p^n_i \\
p^n_1 & p^n_0 & \cdots & p^n_{\gamma+2} \\
p^n_2 & p^n_1 & \cdots & p^n_{\gamma+3} \\
\vdots & \vdots & \ddots & \vdots \\
p^n_{\gamma-1} & p^n_{\gamma-2} & \cdots & p^n_0
\end{bmatrix}.$$  

(4.31)

The probability mass function of the quantized Page’s test statistic is contained in $\mathbf{u}_n$ [64], which is given by

$$\mathbf{u}_n = \mathbf{D}_n \mathbf{D}_{n-1} \cdots \mathbf{D}_0 \mathbf{u}_0$$

(4.32)

where $\mathbf{u}_0$ is a vector composed of the initial probabilities of observing $(l_0, l_1, \ldots, l_{\gamma-1})$.

The CDF of the stopping time $T$ which is the probability that the algorithm has stopped prior to or during the current time sample can be written as [64]

$$F_T(t) = 1 - P \{ \text{Test update takes on levels } (l_0, l_1, \ldots, l_{\gamma-1}) \text{ at time } t \}$$

$$= 1 - \mathbf{1}' \mathbf{D}_t \mathbf{D}_{t-1} \cdots \mathbf{D}_1 \mathbf{u}_0$$

(4.33)

under $H_0$. With the realistic assumption that the test starts with the zero state, $\mathbf{u}_0$ has a one as the first element and zeros elsewhere, i.e. $\mathbf{u}_0 = \mathbf{e}_1$. When the sensing nonlinearity
does not change with time (as in our case), all the probability transition matrices will be the same and equal to $D$ and thus the CDF of the stopping time $T$, under $H_0$, reduces to

$$F_T(t) = 1 - \mathbf{1}'D'\mathbf{e}_1$$  \hspace{1cm} (4.34)

On the other hand, the probability of one or more false alarms in a signal of length $n_d$ is

$$P_f = 1 - Pr\{\text{No false alarm in } n_d \text{ samples}\}$$
$$= 1 - Pr\{T > n_d\}$$
$$= F_T(n_d)$$  \hspace{1cm} (4.35)

Therefore, the false alarm performance of the Page test is quantified as

$$P_f = 1 - \mathbf{1}'D^{n_d}\mathbf{e}_1$$  \hspace{1cm} (4.36)

The probability of false alarm is a function of the Page test threshold through $\gamma$.

### 4.4.1 Threshold Setting for LRT-based CUSUM algorithms

In the known-parameter CUSUM algorithm, the optimal detector nonlinearity is the log-likelihood ratio (LLR) [64]. In order to apply (4.36) for threshold setting, we need to derive $p_i = Pr\{L(x) = i\Delta|H_0\}$, where $L(x)$ is the LLR at $x$. In our assumed model, for a single-user CUSUM algorithm $L(x)$ is

$$L(x) = \frac{1}{2} \log\left(\frac{\sigma_n^2}{\sigma_n^2 + \sigma_z^2}\right) + \frac{x^2\sigma_z^2}{2\sigma_n^2(\sigma_z^2 + \sigma_n^2)}.$$  \hspace{1cm} (4.37)

As $x$ has Gaussian distribution, $x^2$ will have a chi-square distribution with one degree of freedom and thus, the CDF of the LLR under $H_0$ is derived as

$$F_L(l) = \gamma \left(0.5, (l - \mu)\frac{\sigma_z^2 + \sigma_n^2}{\sigma_z^2}\right) \frac{1}{\Gamma(0.5)}$$  \hspace{1cm} (4.38)
where \( \mu = \frac{1}{2} \log \left( \frac{\sigma_z^2}{\sigma_n^2 + \sigma_z^2} \right) \), \( \gamma(k, z) \) is the lower incomplete Gamma function and \( \Gamma(k) \) is the Gamma function with \( \Gamma \left( \frac{1}{2} \right) = \sqrt{\pi} \). Hence \( p_i \) is obtained as follows:

\[
p_i = \frac{1}{\Gamma(0.5)} \left[ \gamma \left( 0.5, (i \Delta - \frac{\Delta}{2} - \mu) \frac{\sigma_z^2 + \sigma_n^2}{\sigma_z^2} \right) - \gamma \left( 0.5, (i \Delta - \frac{\Delta}{2} - \mu) \frac{\sigma_z^2 + \sigma_n^2}{\sigma_z^2} \right) \right] \tag{4.39}
\]

In global CUSUM algorithm, the update statistic is

\[
L(x_1, \ldots, x_L) = \sum_{i=1}^{L} \left( \frac{1}{2} \log \left( \frac{\sigma_n^2}{\sigma_n^2 + \sigma_z^2} \right) + \frac{x_i^2 \sigma_z^2}{2 \sigma_n^2 (\sigma_z^2 + \sigma_n^2)} \right) \tag{4.40}
\]

We again observe that \( L(x) \) has a chi-square distribution with \( L \) degrees of freedom and thus \( p_i \) is derived as

\[
p_i = \frac{1}{\Gamma \left( \frac{L}{2} \right)} \left[ \gamma \left( \frac{L}{2}, (i \Delta + \frac{\Delta}{2} - L\mu) \frac{\sigma_z^2 + \sigma_n^2}{\sigma_z^2} \right) - \gamma \left( \frac{L}{2}, (i \Delta - \frac{\Delta}{2} - L\mu) \frac{\sigma_z^2 + \sigma_n^2}{\sigma_z^2} \right) \right] \tag{4.41}
\]

Having found the \( p_i \)s, we can form the matrix \( D \) by (4.31) and obtain the probability of false alarm using (4.36).

### 4.4.2 Threshold Setting for Linear-based CUSUM Algorithms

For global CUSUM algorithm with unknown parameters where the linear test is applied, as proposed in previous section, the update statistic is

\[
\Lambda(x_1, \ldots, x_L) = \sum_{i=1}^{L} \frac{1}{2 \sigma_n^2} \left( \frac{x_i^2}{\sigma_n^2} - 1 \right) \tag{4.42}
\]

The update statistic has chi-square distribution and its CDF under \( \mathcal{H}_0 \) is

\[
F_{\Lambda}(\lambda) = \frac{\gamma \left( \frac{L}{2}, (\lambda \sigma_n^2 + \frac{L}{2}) \right)}{\Gamma \left( \frac{L}{2} \right)} \tag{4.43}
\]

Therefore \( p_i \) can be obtained from

\[
p_i = \frac{1}{\Gamma \left( \frac{L}{2} \right)} \left[ \gamma \left( \frac{L}{2}, ((i \Delta + \frac{\Delta}{2}) \sigma_n^2 + \frac{L}{2}) \right) - \gamma \left( \frac{L}{2}, ((i \Delta - \frac{\Delta}{2}) \sigma_n^2 + \frac{L}{2}) \right) \right] \tag{4.44}
\]

Consequently, we can form the matrix \( D \) by (4.31) and determine the probability of false alarm using (4.36).
4.5 Simulation Results

In this section, we present our simulation results to evaluate the performance of the cooperative quickest detection schemes presented in this chapter. We consider both cases of known and unknown parameters and show the performance of the proposed linear-based schemes. We assume that \( L = 6 \) secondary users independently sense the primary user’s spectrum. In order for the sensing samples to be independent for the CUSUM algorithm, the received signal should be sampled at a rate below the primary’s symbol rate. If the received signal is sampled with sampling rate no lower than the symbol rate, the primary signal will be over-sampled which makes the resulting signal samples to be correlated. Here, we assume that the received signal is sampled below the symbol rate and the signal samples are independent. The received signals at the secondary users are
Chapter 4. Cooperative Quickest Sensing with Unknown Parameters

Figure 4.2: Probability of false alarm, $P_f$, versus average detection delay $T_d$.

assumed to have a zero-mean Gaussian distribution with variance 1 and after change
time $\tau$, when the primary user becomes active, have a zero-mean Gaussian distribution
with variance 3. The change time $\tau$ is assumed to be geometrically distributed with
parameter $\rho = 0.0463$.

Fig. 4.1 shows the probability of false alarm versus the average detection delay for
the CUSUM-based cooperative quickest sensing schemes discussed above for the known-
parameter scenario. The local quantization thresholds are set as follows: for the one-bit
global CUSUM the local threshold is $a' = 1/3$, for the 2-bit global CUSUM $a'_1 = 1/2$,
$a'_2 = 1$ and $a'_3 = 3/2$, and local thresholds of the 3-bit Global CUSUM are $a'_1 = 1/3$,$a'_2 = 1/2$, $a'_3 = 2/3$, $a'_4 = 5/6$, $a'_5 = 7/6$, $a'_6 = 5/3$ and $a'_7 = 13/6$. As shown in [60],
constant local quantization thresholds would work as the optimal policy. From Fig. 4.1,
we observe that the global CUSUM scheme performs the best, as we expected, due to the assumption of availability of perfect local observations at the fusion center. The global CUSUM with one-bit local decisions performs the worst, as the global CUSUM is applied based on only one bit of information about the local observations. We can also observe that by having only three bits of information about the local observations, the global CUSUM performs better than the hard fusion of local CUSUM, where each user requires sufficient memory to perform CUSUM.

Fig. 4.2 depicts probability of false alarm versus average detection delay curves for cooperative quickest sensing schemes based on linear test under the unknown-parameter scenario. We observe that the linear-based cooperative quickest sensing schemes perform comparably to the their known-parameter counterparts without requiring the knowledge or estimation of $\sigma^2_{\overline{z}_l}$. We can also observe that the linear-based 1-bit global CUSUM performs the closest to its known-parameter counterpart compared to the linear-based hard fusion and global CUSUM schemes. This can be intuitively observed from difference in its test statistic (4.25) from the test statistics of linear-based hard fusion and global CUSUM in (4.23) and (4.26) respectively.

Table 4.1 shows the average detection delay for different cooperative quickest sensing schemes maintaining certain levels of average run length to false alarm. Once again, we observe that the global CUSUM with available observations at the fusion center outperforms the other schemes and the global CUSUM with one-bit local decision results in the highest detection delay among the discussed cooperative quickest sensing schemes. Moreover, we observe that the linear-based schemes perform closely to their known-parameter counterparts while not requiring any parameter estimation.

Fig. 4.3 depicts the probability of false alarm versus the Page test threshold for the global CUSUM scheme. The probability of false alarm has been calculated according to (4.36). Using these plots, the threshold can be chosen according to the target probability of false alarm for a certain signal length.
### Table 4.1: Average detection delay in samples for various cooperative quickest sensing schemes.

<table>
<thead>
<tr>
<th>Cooperative quickest sensing scheme</th>
<th>$T_f = 20$</th>
<th>$T_f = 25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global CUSUM</td>
<td>1.94</td>
<td>2.59</td>
</tr>
<tr>
<td>Global CUSUM with linear test</td>
<td>2.23</td>
<td>2.71</td>
</tr>
<tr>
<td>3-bit global CUSUM</td>
<td>2.70</td>
<td>3.66</td>
</tr>
<tr>
<td>3-bit global CUSUM with linear test</td>
<td>3.15</td>
<td>3.94</td>
</tr>
<tr>
<td>OR fusion of local CUSUM</td>
<td>3.12</td>
<td>4.68</td>
</tr>
<tr>
<td>OR fusion of local CUSUM with linear test</td>
<td>3.67</td>
<td>4.82</td>
</tr>
<tr>
<td>2-bit global CUSUM</td>
<td>4.35</td>
<td>5.23</td>
</tr>
<tr>
<td>2-bit global CUSUM with linear test</td>
<td>4.60</td>
<td>5.67</td>
</tr>
<tr>
<td>1-bit global CUSUM</td>
<td>5.42</td>
<td>7.02</td>
</tr>
<tr>
<td>1-bit global CUSUM with linear test</td>
<td>5.75</td>
<td>7.63</td>
</tr>
</tbody>
</table>
Figure 4.3: Probability of false alarm, $P_f$, versus threshold $q$. 
4.6 Chapter Summary

In this chapter, we presented a linear test approach for cooperative quickest sensing in cognitive radios. We presented several cooperative quickest sensing schemes based on the CUSUM algorithm and derived their test statistics under noisy and noiseless channel scenarios. However, due to unknown parameters in the distribution of the observations, such as primary signal variance, the CUSUM-based approaches are not directly applicable to cognitive radios. For this reason, we proposed a linear test for cooperative spectrum sensing, which does not require any prior knowledge or estimation of the primary signal or channel statistics. Unlike other proposed schemes, the linear-based schemes do not lead to any iterative algorithm or any parameter estimation and, thus, are much simpler and faster to implement. Furthermore, we presented a threshold setting method for CUSUM-based algorithms for a target probability of false alarm. We showed that linear-based cooperative quickest sensing schemes perform comparably to their known-parameter counterparts, while not relying on prior knowledge or estimation of the uncertain parameters.
Chapter 5

Throughput of Cognitive Radio Networks

In order to investigate the efficiency of quickest sensing approach and feasibility of deploying CR networks based on this approach in achieving satisfactory secondary rates while protecting the primary user, in this chapter, we analyze the achievable throughput by quickest-based CRs and investigate the sensing-throughput tradeoff when quickest sensing is applied. To the best of our knowledge, this analysis for quickest sensing has not been performed in the existing literature. By analyzing the same problem for block sensing-based CRs, we compare the achievable throughput and efficiency of these two approaches and illustrate the superiority and practicality of the quickest sensing approach.

As CR nodes cannot perform spectrum sensing and data transmission simultaneously in the same band, a sense and then transmit approach must be applied [43–45]. Therefore the basic frame structure used by CR networks consists of a sensing interval and a transmission interval. Such a frame structure leads to a fundamental tradeoff [43] between the throughput achievable by the CRs and the quality of sensing. The latter relates directly to the level of protection for the primary user (PU). This tradeoff is due to the following:
• Higher-quality sensing requires longer sensing times;

• For a fixed frame duration, a longer sensing time results in a shorter transmission time for the CR and hence a lower achievable throughput.

In [43], the block sensing duration was designed to maximize CR throughput with a given frame duration, while protecting the primary user by keeping the probability of detection higher than a certain threshold. The similar problem of optimizing the frame duration for a given sensing time has been considered in [44, 45]. In [44], the objective of the optimization was to maximize the CRs throughput while keeping the probability of collision with the primary user below a certain threshold, assuming that primary user activity in successive frames are independent of each other. In [45], the throughput of the CR was maximized while ensuring that the primary interference time ratio (PITR) is less than a certain threshold. Unlike [44], it assumed dependent primary user signals in successive frames by modeling the primary user activity as a continuous time Markov process. Both [44] and [45] assume perfect error-free block sensing which is not realistic and they neglect the effect of interference from the primary user on the CRs throughput.

While the existing research on throughput-sensing trade-off is focused on block sensing, it is clear at least at an intuitive level that the class of quickest (sequential) sensing schemes [29–31, 37, 60] can perform better than block-based sensing, given that the objective here is to minimize the delay between the occurrence of the event of interest and its detection. A comparison between block and sequential sensing CR networks is therefore investigated in this chapter. We are interested particularly in bursty primary applications, where the question of feasibility of providing cognitive radios with a certain throughput and primary users with a certain protection level has not been studied rigorously.

In this chapter, we derive the average throughput for both sequential and block sensing, given the probability distribution of the primary network’s active and inactive periods, and assuming the CR network’s throughput is lower in the event of collision with the
primary user. The sequential sensing method is the Page’s cumulative sum (CUSUM) algorithm [33,34] while the block sensing method is the energy detector. We consider two optimization problems: one with the probability of collision as the primary user protection metric and the other with the PITR as the constraint. Unlike [44,45] we realistically assume block sensing schemes with detection error and take into account the degradation of CR throughput in the event of collision between the CR and primary user signals. We solve the two optimization problems for quickest sensing and block sensing-based CRs. Our simulation results reveal that quickest sensing results in higher average throughput compared to block sensing, while requiring smaller frame duration and sensing time, when the interference to the primary user is constrained to be small (as it should be in practice). This shows quantitatively the superiority of quickest sensing compared to block sensing in CR networks. This chapter is partially published in [65].

The remainder of this chapter is organized as follows. The system model and problem formulation are described in Section 5.1. The derivation of the throughput, probability of collision and PITR is presented in Section 5.2 for both quickest and block-based sensing schemes. The solutions to the optimization problems are provided in Section 5.3. The simulation results are presented in Section 5.4. Finally, Section 5.5 concludes the chapter.

5.1 System Model and Problem Formulation

5.1.1 System Description

We assume the primary user has a continuous-time Markov process traffic model with the two states “on” and “off”. The on and off times are exponentially distributed with parameters $\beta$ and $\alpha$ respectively. This model has been assumed in previous studies [44,45, 66,67]. Sensing and transmission cannot be performed by the CR network simultaneously in the same band. Therefore, the CR network applies a sense-then-transmit strategy
Figure 5.1: A CR frame, showing the sensing interval in black and the data transmission phase in a lighter shade. The primary user is inactive at the start of the frame (defined as event $H_0$ in the text), and comes on at time $T + X$ from the start of the frame, leading to collision over a duration $S = T_0 - T - X$.

[43–45], where each frame consists of sensing and transmission phases, as depicted by the dark and light areas in Figure 5.1. The sensing duration is $T$, which is random in quickest sensing; the total frame duration is $T_0$. If the primary user is inactive and the detector correctly senses that state then only the CR uses the spectrum over the duration $X$ after sensing, and the primary user comes on for a period $S$ while the CR is still transmitting. $X$ and $S$ are random variables whether in block or sequential sensing since the primary network has random on and off times. Clearly, $T_0 = T + X + S$.

### 5.1.2 Problem Formulation

Using the above notation, we can explain the fundamental problem as follows: to maximize throughput for the CR without constraining interference to the primary user, we should maximize the fraction of time the CR is transmitting, i.e. $(T_0 - T)/T_0$, but to reduce interference to the primary user, we should increase the sensing time $T$ so as to improve sensing accuracy and reduce the probability of the CR and primary user transmissions colliding. For a given $T_0$, the former requirement calls for small sensing time $T$, but the latter calls for large $T$ – there is thus an optimal $T$ that balances both require-
ments. Alternatively, if we consider $T$ (or its distribution in the case of quickest sensing) to be fixed, then increasing $T_0$ leads to higher CR throughput but larger interference to the primary user since $S$ increases, so an optimal $T_0$ exists. In this chapter, we tackle the latter problem.

Specifically, the optimal frame duration may be found by solving this optimization problem:

$$T_{0 \text{opt}} = \arg \max_{T_0} R(T_0)$$

s.t. $P_c \leq P_c^{\text{max}}$ \hspace{1cm} (5.1)

where $R(T_0)$ is the average throughput of the CR network as a function of $T_0$, and $P_c^{\text{max}}$ is the maximum probability of collision that the primary user can tolerate. This tradeoff can also be formulated with the primary interference time ratio (PITR) as the constraint, i.e.

$$T_{0 \text{opt}} = \arg \max_{T_0} R(T_0)$$

s.t. $I_R \leq I_R^{\text{max}}$ \hspace{1cm} (5.2)

where $I_R^{\text{max}}$ is the maximum PITR, defined as the ratio of the average interference time experienced by primary user to the average transmission time of the primary user. The PITR has been used in [44, 45] as a measure of the effect of collision on the primary user. The rate $R(T_0)$ is a function not just of $T_0$, but also $\alpha$, $\beta$, and the performance of the sensor in terms of probabilities of false alarm and detection, as we show in the next section.

### 5.1.3 CUSUM algorithm

We apply the well-known Page’s CUSUM algorithm [33] as the quickest sensing engine. Assuming a zero-mean Gaussian primary user signal with variance $\sigma^2$ and AWGN noise
with variance $\sigma_n^2$, the log likelihood ratio (LLR) based on observation $z[m]$ is derived as

$$l(z[m]) = \frac{\sigma_s^2 z^2[m]}{2(\sigma_s^2 + \sigma_n^2)} + \frac{1}{2} \ln \left( \frac{\sigma_n^2}{\sigma_s^2 + \sigma_n^2} \right)$$

This LLR is accumulated in the CUSUM algorithm to form the test statistic, which is compared against a set threshold, as explained below, to decide if the primary user is active or inactive.

As the CUSUM algorithms for detecting a rising edge (primary user active) and a falling edge (primary user silent) are different and we are interested in both events, we need to run both algorithms in parallel. The rising edge detector compares the test statistic against threshold $h_0$, whereas the falling edge detector applies $h_1$ as its CUSUM threshold. The first of the two algorithms to return a positive test result will yield the decision on primary user activity. The algorithms are as follows:

**The rising edge detector** [64]

1. Set $g_0 = 0$, $m = 1$

2. If $g_{m-1} < h_0$
   - Set $g_m = \max\{0, g_{m-1} + l(z[m])\}$
   - If $g_m = 0$, set $t_1 = m$
   - If $g_m < h_0$, set $m = m + 1$ and goto 2)

3. If $g_m \geq h_0$,
   - A rising edge of the signal has been detected.
   - $t_1 = m$

**The falling edge detector** [64]

1. Set $g_0 = h_0 + h_1$, $m = 1$
2. If $g_{m-1} \geq h_0$
   - Set $g_m = \min\{h_0 + h_1, g_{m-1} + l(z[m])\}$
   - If $g_m = h_0 + h_1$, set $t_2 = m$
   - If $g_m > h_0$, set $m = m + 1$ and goto 2)

3. If $g_m \leq h_0$,
   - A falling edge of the signal has been detected.
   - $t_2 = m$

In the pseudo-codes above, $t_1$ and $t_2$ are respectively the times at which a positive result is returned by the rising and falling edge detectors. If $t_1 \leq t_2$ the algorithm declares that the primary user is active and otherwise declares the availability of the band. A key observation from simulations is that the probability that $t_2 > t_1$ when the primary user is idle is negligible, and so is the probability that $t_1 > t_2$ when the primary user signal is on. Therefore, we assume that our proposed parallel CUSUM rising and falling edge detection algorithm always correctly detects the activity of the primary user.

### 5.2 Derivation of $R(T_0)$, $P_c$ and $I_R$

In this section, we derive expressions for the three quantities in (5.1) and (5.2) for both quickest sensing and block sensing, so that the optimization problems can be rigorously formulated and then solved.

#### 5.2.1 Quickest Spectrum Sensing

**Preliminaries**

The primary user is either active or inactive at the start of a CR frame. In both of these scenarios we assume that the primary user’s activity state remains unchanged during the
sensing interval. This implies that the expected on and off durations of the primary user are much larger than the expected sensing time. This assumption was made in related previous studies [43–45].

In quickest sensing, the CR is assumed to detect the activity of the primary user correctly with probability one, as mentioned in the previous section. Therefore, when the primary user is active at the beginning of the CR frame, quickest sensing detects this activity after sensing time $T$, and refrains from transmission in that frame. In that case, the throughput of the CR is zero. When the primary user is idle, on the other hand, the CR starts transmission after sensing time $T$. However, the primary user may resume transmission during the CRs transmission phase, which results in collision. This collision degrades the throughput of the CR. Unlike most studies where the interference from the primary user to the CR has been neglected [44,45], we take into account the rate degradation to the CR due to collision. We denote the throughput of the CR network with no collision as $C_0$ and the throughput with collision as $C_1$. For instance, in a point-to-point transmission in the CR network with SNR of $\text{SNR}_c$ and interference SNR from the primary user to the CR of $\text{SNR}_p$, and Gaussian primary user and CR signals, we have $C_0 = \log_2(1 + \text{SNR}_c)$ and $C_1 = \log_2(1 + \frac{\text{SNR}_p}{1+\text{SNR}_p})$. It is clear that $C_0 > C_1$.

Referring to Figure 5.1, let $V = X + T$. Also, we denote by $\mathcal{H}_0$ the event that the primary network is not active at the start of the frame, and by $\mathcal{H}_1$ its complement. The achievable rate is then

\[ R(T_0) = P(\mathcal{H}_0)R_0(T_0) + P(\mathcal{H}_1)R_1(T_0) \]  \hspace{1cm} (5.4)

where $R_i(T_0)$ is the rate conditioned on $\mathcal{H}_i$, and $P(\mathcal{H}_i)$ is the probability of $\mathcal{H}_i$. Since the probability of missed detection (detection error) is negligible in the proposed quickest sensing scheme, $R_1(T_0) = 0$. Therefore we have that

\[ R(T_0) = P(\mathcal{H}_0)R_0(T_0) \]  \hspace{1cm} (5.5)

When the primary user is inactive at the start of the frame, there are three scenarios
to consider:

1. $T \leq T_0, V \leq T_0$: The CR transmits at the rate $C_0$ for time $X = V - T$, and at the rate $C_1$ for the remaining time $T_0 - V$;

2. $T \leq T_0, V > T_0$: The primary user does not return until after the end of the frame, so the CR transmits at the rate $C_0$ for the duration $T_0 - T$;

3. $T > T_0$: The CR takes so long to detect the availability of the band that the transmission opportunity is lost for that frame. The rate is 0.

**Exact Expressions**

With the above breakdown, we have

$$R_Q(T_0) = \left[ P(V \leq T_0) \left\{ E \left[ \frac{X}{T_0} \bigg| V \leq T_0 \right] C_0 + E \left[ \frac{T_0 - V}{T_0} \bigg| V \leq T_0 \right] C_1 \right\} \right] + P(V > T_0, T \leq T_0) \left\{ E \left[ \frac{T_0 - T}{T_0} \bigg| V > T_0, T \leq T_0 \right] C_0 \right\} P(H_0) \quad (5.6)$$

Thus the achievable rate depends on the distributions of the random variables $X, T$ and $V$, which we will now derive.

As mentioned before, the idle time is exponentially distributed with parameter $\alpha$, thus the probability density function (pdf) of $X$ is

$$f_X(x) = \alpha e^{-\alpha x}, \ x \geq 0. \quad (5.7)$$

(Recall that the exponential distribution has the memory-less property $P[T > t_1 + t_2 | T > t_2] = P[T > t_1]$, and so even though the start of the transmission phase in the CR frame does not coincide with the instant the channel becomes idle, the same exponential distribution applies.) The sensing time $T$ in the CUSUM algorithm is a random variable whose exact distribution is not known. For simplicity, we model $T$ as an exponential random variable with parameter $\lambda$. Our simulation results (in Section 5.4) verify this assumption.
Given that the primary user idle time is exponential with parameter $\alpha$ and its active time is exponential with parameter $\beta$, the prior probabilities of primary user activity at the start of each CR frame are

\[
P(H_0) = \frac{\beta}{\alpha + \beta} \quad \text{and} \quad P(H_1) = \frac{\alpha}{\alpha + \beta}. \tag{5.8}
\]

In order to derive an expression for the average throughput in (5.5), first we need to determine the joint pdf of $X$ and $T$. As the sensing time of the CR and the idle time of the primary user are independent exponential random variables we have

\[
f_{X,T}(x, t) = \begin{cases} 
\alpha e^{-\alpha x} \lambda e^{-\lambda t}, & x \geq 0, t \geq 0 \\
0 & \text{otherwise}
\end{cases}. \tag{5.10}
\]

Through straightforward integration of $f_{X,T}(x, t)$ over the appropriate regions in the $x$-$t$ plane, we derive the probability terms in (5.6) as

\[
P(V \leq T_0) = 1 + \frac{\alpha}{\lambda - \alpha} e^{-\lambda T_0} - \frac{\lambda}{\lambda - \alpha} e^{-\alpha T_0} \tag{5.11}
\]

and

\[
P(V > T_0, T \leq T_0) = \frac{\lambda}{\lambda - \alpha} \left( e^{-\alpha T_0} - e^{-\lambda T_0} \right). \tag{5.12}
\]

To evaluate the expected value terms in (5.6), we first find the conditional distributions of $X$ and $T$ conditioned on $\{V \leq T_0\}$ and $\{V > T_0, T \leq T_0\}$, as outlined in Appendix C. Then we obtain

\[
E\left[ \frac{X}{T_0} \middle| V \leq T_0 \right] = 1 - \frac{1 - e^{-\alpha T_0}}{T_0 \alpha} - \frac{\alpha e^{-\lambda T_0}}{\alpha - \lambda} + \frac{\alpha e^{-\lambda T_0} - \alpha e^{-\alpha T_0}}{T_0 (\alpha - \lambda)^2} \tag{5.13}
\]

\[
E\left[ \frac{T}{T_0} \middle| V \leq T_0 \right] = 1 - \frac{1 - e^{-\lambda T_0}}{T_0 \lambda} - \frac{\lambda e^{-\alpha T_0}}{\lambda - \alpha} + \frac{\lambda e^{-\alpha T_0} - \lambda e^{-\lambda T_0}}{T_0 (\lambda - \alpha)^2} \tag{5.14}
\]
\[ E\left[ \frac{T_0 - T}{T_0} \middle| V > T_0, T \leq T_0 \right] = \frac{1}{\lambda - \alpha} \left( \frac{e^{-\lambda T_0} - e^{-\alpha T_0}}{e^{-\alpha T_0} - e^{-\lambda T_0}} \right) + e^{-\alpha T_0}. \] (5.15)

Consequently, the average CR throughput is derived as

\[ R_Q(T_0) = \left[ (\alpha e^{-\lambda T_0} - \lambda e^{-\alpha T_0}) \left( \frac{C_1}{\lambda - \alpha} \right) + \frac{1}{T_0} \left( \frac{C_0}{\alpha} - C_1 \frac{\lambda + \alpha}{\lambda \alpha} \right) \right. \]
\[ + \frac{e^{-\lambda T_0}}{T_0} \left( \frac{C_0 - C_1}{\lambda - \alpha} + C_1 \frac{(\lambda + \alpha)}{\lambda \alpha} - \frac{C_0}{\alpha} \right) + \frac{e^{-\alpha T_0}}{T_0} \left( \frac{C_1 - C_0}{\lambda - \alpha} \right) + C_1 \right] P(H_0). \] (5.16)

Even if the primary user is idle during the sensing period, it may become active during the transmission period and collide with the CR. The interference to the primary user can be evaluated by the probability of collision which is defined as

\[ P_{cq} = P(H_0)P(V \leq T_0) \] (5.17)
or by the PITR which is the ratio of the average collision time to the average activity time of primary user, defined as

\[ I_{RQ} = \frac{P(H_0)E[S]}{\beta^{-1}}. \] (5.18)

These definitions are based on the assumptions that the CR network has full traffic and once the primary user becomes active it remains busy until the end of the frame. The latter assumption is realistic when the average primary user active time is larger than the frame duration. Using (5.8) and (5.11) in (5.17), we will obtain the probability of collision, so the only remaining problem is finding \( I_{RQ} \), which requires \( E[S] \), derived as follows:

\[ E[S] = P(H_0)E[T_0 - V|V \leq T_0]P(V \leq T_0) \]
\[ = P(H_0)[T_0P(V \leq T_0) - E[V|V \leq T_0]P(V \leq T_0)] \] (5.19)

To find \( E[V|V \leq T_0] = E[X + T|V \leq T_0] \), we can apply the derived expressions in (5.13) and (5.14). Furthermore, by substituting (5.11) into (5.20), \( I_{RQ} \) is derived as

\[ I_{RQ} = P(H_0) \left( T_0 \beta + \frac{\beta}{\lambda - \alpha} \left( \frac{\lambda}{\alpha}(1 - e^{-\alpha T_0}) - \frac{\alpha}{\lambda}(1 - e^{-\lambda T_0}) \right) \right). \] (5.21)
5.2.2 Block-Based Sensing

Preliminaries

Block sensing schemes have a non-negligible probability of making decision errors. This fact, together with the non-random nature of the sensing interval $T$, changes the analysis compared to the quickest sensing case of the last section.

In block-based sensing schemes, based on the activity state of the primary user at the beginning of the frame, two scenarios may occur. The first scenario is that the primary user is idle when the CR frame starts and it can potentially become active during the transmission slot of the primary user. In this case, unlike in the quickest sensing framework, the primary user can be falsely detected to be active (false alarm) and the CR would miss the transmission opportunity. The second scenario happens when the primary user is active during the sensing slot and can potentially become idle in the transmission slot. In this case, the activity of the primary user could be falsely detected as being idle (missed detection) by the block-based sensing scheme. In this case, the CR would start transmission after the sensing slot and would cause collision with the primary user. This collision lasts as long as the primary user is active during the frame.

In order to analyze the CR throughput under $H_1$, we denote the active time of the primary user by $Y$, which based on the exponential model for the active duration of the primary user, has pdf

$$f_Y(y) = \beta e^{-\beta y}, \ y \geq 0. \quad (5.22)$$

Under $H_1$, this variable replaces random variable $X$ in Figure 5.1. We define random variable $W = Y + T$ under $H_1$. Also, we define the probabilities of false alarm and detection respectively as $P_f = P(\hat{\theta} = 1|H_0)$ and $P_d = P(\hat{\theta} = 1|H_1)$, where $\hat{\theta}$ denotes the sensing result of the block-based sensing scheme.

For $N$-sample energy detection as the block sensing scheme, decision making is as
follows:

\[
\begin{align*}
\mathbf{u} &= \begin{cases} 
1, & \sum_{m=1}^{N} z[m]^2 \geq \epsilon \\
0, & \sum_{m=1}^{N} z[m]^2 < \epsilon,
\end{cases}
\end{align*}
\]  

(5.23)

where \( \epsilon \) is the energy threshold. The energy threshold is chosen to maintain a certain local probability of false alarm \( P_f \). The probabilities of false alarm and detection for energy detection are determined as [16]

\[
\begin{align*}
P_f &= Q \left( \frac{\epsilon - T f_s \sigma_n^2}{\sqrt{2 T f_s \sigma_n^2}} \right) \\
P_d &= Q \left( \frac{\epsilon - T f_s (\sigma_s^2 + \sigma_n^2)}{\sqrt{2 T f_s (\sigma_s^2 + \sigma_n^2)}} \right)
\end{align*}
\]

(5.24)

(5.25)

where \( Q(\cdot) \) is the inverse \( Q \)-function, \( \epsilon \) is the threshold, \( T \) is the sensing time, \( f_s \) is the sampling frequency and \( \sigma_n^2 \) and \( \sigma_s^2 \) are respectively the noise and signal power.

**Exact Expressions**

When the primary user is inactive at the start of the frame, the CR detects this opportunity with probability \( 1 - P_f \). In this case, there are three scenarios to be considered which are identical to the ones described for quickest sensing and will not be repeated here. On the other hand, under \( \mathcal{H}_1 \), block-based sensing could falsely declare that the primary user is inactive with probability \( 1 - P_d \) and the CR starts transmission while the primary user signal is on. In this case, there are two scenarios to consider:

1. \( W \leq T_0 \): The CR transmits at the rate \( C_1 \) for time \( Y = W - T \), and at the rate \( C_0 \) for the remaining time \( T_0 - W \);

2. \( W > T_0 \): The primary user remains active until after the end of the frame, so the CR transmits at the rate \( C_1 \) for the duration \( T_0 - T \);
With the above description, we derive the average throughput for CR with block-based sensing as follows

\[
R_B(T_0) = P(H_0)(1 - P_f) \left\{ P_0^c \left( C_0 E\left[ \frac{X}{T_0} \right] X < T_0 - T \right) + C_1 E\left[ \frac{T_0 - V}{T_0} \right] X + T < T_0 \right) + C_0(1 - P_0^c) \left( \frac{T_0 - T}{T_0} \right) \right\} + P(H_1)(1 - P_d) \left\{ P_1^c \left( C_1 E\left[ \frac{Y}{T_0} \right] Y < T_0 - T \right) + C_0 E\left[ \frac{T_0 - W}{T_0} \right] Y < T_0 - T \right) + C_1(1 - P_1^c) \left( \frac{T_0 - T}{T_0} \right) \right\}
\]

(5.26)

where \( P_0^c = P(X < T_0 - T) \) and \( P_1^c = P(Y < T_0 - T) \). To calculate the average throughput achievable by a block-based sensing CR network, we derive the expected values and probabilities in (5.26) as follows:

\[
E\left[ \frac{X}{T_0} \right] V < T_0 = 1 - \frac{T}{T_0} \left[ 1 - \frac{1 - e^{-\alpha(T_0 - T)}}{\alpha T_0} \right]
\]

(5.27)

\[
E\left[ \frac{Y}{T_0} \right] W < T_0 = 1 - \frac{T}{T_0} \left[ 1 - \frac{1 - e^{-\beta(T_0 - T)}}{\beta T_0} \right]
\]

(5.28)

\[
P_0^c = 1 - e^{-\alpha(T_0 - T)}
\]

(5.29)

\[
P_1^c = 1 - e^{-\beta(T_0 - T)}
\]

(5.30)

Consequently, the average throughput of a CR network with block-based sensing is derived as

\[
R_B(T_0) = P(H_0)(1 - P_f) \left[ C_1 + \frac{1}{T_0} \left( \frac{C_0 - C_1}{\alpha} - TC_1 \right) \right] + \frac{e^{-\alpha(T_0 - T)}}{T_0} \left( \frac{C_1 - C_0}{\alpha} \right) + P(H_1)(1 - P_d) \times \left[ C_0 + \frac{1}{T_0} \left( \frac{C_1 - C_0}{\alpha} - TC_0 \right) + \frac{e^{-\beta(T_0 - T)}}{T_0} \left( \frac{C_0 - C_1}{\beta} \right) \right]
\]

(5.31)
In finding the probability of collision, we should note that in block sensing CRs, there is a non-zero probability of collision under both \( H_0 \) and \( H_1 \) as described before and thus \( P_c \) can be written as

\[
P_{cb} = P(H_0)(1 - P_f)P_c^0 + P(H_1)(1 - P_d) \tag{5.32}
\]

\[
= P(H_0)(1 - P_f)(1 - e^{-\alpha(T_0 - T)}) + P(H_1)(1 - P_d) \tag{5.33}
\]

The PITR is derived as

\[
I_RB = \frac{P(H_0)E[S|H_0] + P(H_1)E[S|H_1]}{\beta^{-1}}
\]

\[
= \beta (P(H_0)E[T_0 - T - X] + P(H_1)E[T_0 - T - Y]).
\]

\[
= \beta (T_0 - T) + \frac{\beta}{\alpha} P(H_0) (1 - e^{-\alpha(T_0 - T)}) + P(H_1) (1 - e^{-\beta(T_0 - T)}). \tag{5.34}
\]

### 5.3 Solving the Optimization Problem

#### 5.3.1 Quickest Sensing

Using the expressions derived for the average CR throughput and the probability of collision respectively in (5.16) and (5.11), the optimization problem (5.1) can be expressed as

\[
T_{0_2}^{\text{opt}} = \arg \max_{T_0} R_Q(T_0)
\]

s.t. \( P(H_0) \left( 1 + \frac{\alpha}{\lambda - \alpha} e^{-\lambda T_0} - \frac{\lambda}{\lambda - \alpha} e^{-\alpha T_0} \right) \leq P_{cQ}^{\text{max}} \tag{5.35} \)

In order to check the existence of an optimal solution we first examine the first derivative of \( R_Q(T_0) \):

\[
\frac{dR_Q(T_0)}{dT_0} = \left[ (e^{-\alpha T_0} - e^{-\lambda T_0}) \left( \frac{\alpha \lambda C_1}{\lambda - \alpha} - \frac{1}{T_0^2} \left( \frac{C_0}{\alpha} - C_1 \frac{\lambda + \alpha}{\lambda \alpha} \right) \right) - \frac{(\lambda T_0 + 1)e^{-\lambda T_0}}{T_0} \left( \frac{C_0 - C_1}{\lambda - \alpha} + \frac{C_1(\lambda + \alpha)}{\lambda \alpha} - \frac{C_0}{\alpha} \right) \right. \\
\left. - \frac{(\alpha T_0 + 1)e^{-\alpha T_0}}{T_0} \left( \frac{C_1 - C_0}{\lambda - \alpha} \right) + C_1 \right] P(H_0) \tag{5.36}
\]
As we show in the next section, whether any $T_0$ drives the derivative above to zero depends on $\lambda$ and $C_1/C_0$. Intuitively, as the frame duration increases, the average rate eventually converges to $P(H_0)C_1$ since the probability of the primary user returning to the channel approaches 1 and the collision interval $S$ approaches $T_0 - T \approx T_0$. Conversely as $T_0$ shrinks, the rate goes to zero since both terms in $R_Q(T_0)$ disappear.

Considering the constraint in (5.35), the probability of collision $P_{cq}$ is a monotonically increasing function of the frame duration as its first derivative,

$$\frac{dP_{cq}}{dT_0} = P(H_0) \frac{\alpha \lambda}{\lambda - \alpha}(e^{-\alpha T_0} - e^{-\lambda T_0}),$$

(5.37)

is a positive function of $T_0$. Hence, we must have

$$T_0 \leq P_{cq}^{-1}(P_{cq}^{\max})$$

(5.38)

where $P_{cq}^{-1}(\cdot)$ is the inverse function of $P_{cq}(\cdot)$ in (5.17). Therefore, the probability of collision imposes an upper bound on the frame length. If $R_Q(T_0)$ does not have a maximum in the range (5.38), the optimal $T_0$ is this upper bound; otherwise the optimal $T_0$ is the stationary point of $R_Q(T_0)$ found by setting (5.36) to zero.

Considering the optimization problem in (5.2) and using the derived expression for the throughput and PITR in (5.16) and (5.21) we get

$$T_{0q}^{\text{opt}} = \arg \max_T R_Q(T_0)$$

s.t. $P(H_0) \left[ T_0 \beta + \frac{\beta}{\lambda - \alpha} \left( \frac{\lambda}{\alpha}(1 - e^{-\alpha T_0}) - \frac{\alpha}{\lambda}(1 - e^{-\lambda T_0}) \right) \right] \leq I_{RQ}^{\max}.$

(5.39)

By taking the derivative of $I_{RQ}$,

$$\frac{dI_{RQ}}{dT_0} = P(H_0) \left[ \beta + \frac{\beta}{\lambda - \alpha}(\lambda e^{-\alpha T_0} - \alpha e^{-\lambda T_0}) \right],$$

(5.40)

we observe that the PITR monotonically increases with $T_0$ as its derivative is a positive function of $T$. Therefore, for a certain threshold $I_{RQ}^{\max}$ the optimal frame duration is bounded by

$$T_{0q}^{\text{opt}} \leq I_{RQ}^{-1}(I_{RQ}^{\max})$$

(5.41)
where $I^{-1}_{R_0}(\cdot)$ is the inverse function of $I_{R_Q}(\cdot)$ in (5.21).

Again, the optimal $T_0$ is the right-hand side of (5.41) if there is no stationary point of $R_Q(T_0)$ in the range (5.41), otherwise it will be the value that drives (5.36) to zero.

### 5.3.2 Block-Based Sensing

The optimization problem (5.1) for block-based sensing schemes can be written as

$$T_{0\text{B}}^{\text{opt}} = \arg \max_{T_0} R_B(T_0)$$

s.t. $P(\mathcal{H}_0)(1-P_f)(1-e^{-\alpha(T_0-T)})+P(\mathcal{H}_1)(1-P_d) \leq P_{\text{cB}}^{\text{max}}$ \hspace{1cm} (5.42)

where $R_B(T_0)$ is derived in (5.31). By examining the first derivative of $R_B(T_0)$ as below

$$\frac{dR_B(T_0)}{dT_0} = P(\mathcal{H}_0)(1-P_f) \left[ -\frac{1}{T_0^2} \left( \frac{C_0 - C_1}{\alpha} - TC_1 \right) - \frac{(\alpha T_0 + 1)e^{-\alpha(T_0-T)}}{T_0^2} \left( \frac{C_1 - C_0}{\alpha} \right) \right] + P(\mathcal{H}_1)(1-P_d)$$

$$\times \left[ -\frac{1}{T_0^2} \left( \frac{C_1 - C_0}{\alpha} - TC_0 \right) - \frac{(\beta T_0 + 1)e^{-\beta(T_0-T)}}{T_0^2} \left( \frac{C_0 - C_1}{\beta} \right) \right].$$

(5.43)

We observe that depending on the values of $C_1/C_0$ and $T$ the rate function (5.31) potentially has a local maximum. We also investigate the behavior of the probability of collision by deriving its first derivative with respect to the frame duration,

$$\frac{dP_{\text{cB}}}{dT_0} = P(\mathcal{H}_0)(1-P_f)\alpha e^{-\alpha(T_0-T)}.$$

(5.44)

This derivative is a positive function of $T_0$, and hence, the probability of collision monotonically increases with the frame duration. Therefore, in this optimization problem, the optimal frame length is upper-bounded as

$$T_{0\text{B}}^{\text{opt}} \leq P_{\text{cB}}^{-1}(P_{\text{cB}}^{\text{max}})$$

(5.45)

where $P_{\text{cB}}^{-1}(\cdot)$ is the inverse function of $P_{\text{cB}}(\cdot)$ in (5.33). As we show in the next section, based on the value of $C_1/C_0$, the average rate function of the block-based sensing CR
either has a local maximum or is monotonically increasing with the frame length and it eventually converges to \( P(H_0)(1 - P_l)C_1 + P(H_1)(1 - P_d)C_0 \) as \( T_0 \) grows large. In the latter case, the optimal frame length is the bound in (5.45), whereas in the former case the optimal frame length is determined in a similar fashion as the quickest sensing scenario in Section 5.3.1.

Considering the optimization problem (5.2) for block sensing-based CRs results in similar solutions to the ones for quickest sensing, as again the first derivative of the PITR,

\[
\frac{dI_{RB}}{dT} = \beta + \beta P(H_0)e^{-\alpha(T_0 - T)} + \beta P(H_1)e^{-\beta(T_0 - T)}
\]

(5.46)

is a nonnegative function of the frame length. Thus we observe that the PITR monotonically increases with \( T_0 \) and for a certain threshold \( I_{RB}^{\max} \) the optimal frame duration must satisfy

\[
T_{0_{RB}}^{\text{opt}} \leq I_{RB}^{-1}(I_{RB}^{\max})
\]

(5.47)

where \( I_{RB}^{-1}(\cdot) \) is the inverse function of \( I_{RB}(\cdot) \) in (5.34). If the maximum is achieved at a lower frame length than the upper bound in (5.47), then the optimal frame length is the first zero of (5.43). Otherwise, the optimal frame length is

\[
T_{0_{RB}}^{\text{opt}} = I_{RB}^{-1}(I_{RB}^{\max}).
\]

(5.48)

### 5.4 Simulation Results

Consider a single CR network and a single primary user. We choose the activity parameters of the primary user according to VoIP traffic [44, 45] where the mean idle and active durations of the primary user are respectively \( \alpha^{-1} = 650 \text{ ms} \) and \( \beta^{-1} = 352 \text{ ms} \). The sampling rate is \( f_s = 1000 \text{ Hz} \) and the CR transmits one packet per slot. The primary user signal variance \( \sigma_s^2 \) and AWGN noise variance \( \sigma_n^2 \) are assumed to be 1.

Fig. 5.2 shows the normalized average throughput \( R_Q/C_0 \), and probability of collision \( P_{cQ} \), of the quickest sensing-based CR network versus the frame duration. The average
sensing time is $\lambda^{-1} = 7.2$ ms. We observe that the simulated average throughput is very close to the theoretical one derived in (5.16). This verifies the suitability of the exponential assumption for the distribution of the quickest sensing time in the average throughput analysis. We also observe that depending on $C_1/C_0$ the average rate may have a local maximum. The probability of collision, on the other hand, is a monotonically increasing function of frame duration. We observe that for acceptable levels of collision probability (e.g., below 0.1) the frame length is small and falls in the monotonically increasing part of the average rate function. Therefore the optimal frame length is directly determined by the constraint on the probability of collision as in (5.38).

Fig. 5.3 shows the normalized average throughput $R_Q/C_0$, and primary interference time ratio (PITR), $I_{R_Q}$, of the quickest sensing-based CR network versus the frame duration. The average sensing time is $\lambda^{-1} = 7.2$ ms. We observe that the PITR is a monotonically increasing function of frame duration. Depending on the acceptable levels
Figure 5.3: Average normalized throughput, $R_Q/C_0$, and primary interference time ratio (PITR), $I_{R_Q}$, versus frame duration $T_0$ for quickest sensing.

of PITR the optimal frame length may fall in the monotonically increasing part of the average rate function or to the right of its peak point. Therefore in the latter case, the optimal $T_0$ in the optimization problem (5.39) is the stationary point of $R_Q(T_0)$ found by setting (5.36) to zero, whereas in the former case the optimal $T_0$ is the upper bound imposed by the PITR constraint in (5.41).

Fig. 5.4 shows the average throughput of quickest sensing-based CR network as well as block sensing-based CR network versus probability of collision. The points of these curves are achieved by changing the frame duration $T_0$. The block sensing scheme considered is energy detection. For block sensing CR we consider two cases. In case 1, the sensing time of the block sensing is set to be the same as the average sensing time of the quickest sensing, $T = \lambda^{-1} = 7.2$ ms and the probabilities of false alarm and detection are respectively $P_f = 0.15$ and $P_d = 0.8$. In case 2 the probabilities of false alarm and detection are set to be respectively $P_f = 0.05$ and $P_d = 0.9$ which requires a longer sensing
Figure 5.4: Average normalized CR throughput, versus probability of collision.

time of $T = 17.7$ ms. We observe that the probability of collision in block sensing has a lower bound and cannot be reduced further due to its fixed sensing time. For example in case 1, where the block sensing time is the same as the average quickest sensing time, the probability of collision for block sensing cannot be lower than 0.07 whereas the probability of collision in quickest sensing can be lowered as much as desired. For the probability of collision of 0.05, the average throughput for the quickest sensing-based CR is 0.56 whereas the average throughput of the block sensing-based CR is 0.29 in case 2, where the block sensing time is more than two times of the average quickest sensing time. For the probability of collision of 0.08, the average throughput for the quickest sensing-based CR is 0.59 (for $C_1/C_0 = 0.8$) whereas the average throughput of the block sensing-based CR in case 1 (with the same sensing time as the quickest sensing) is 0.37. We observe that for the same probability of collision, the quickest sensing-based CR achieves a much higher average throughput compared to the block sensing-based CR even when the block sensing time is much larger than the average quickest sensing time.
Fig. 5.5 depicts the average throughput of quickest sensing-based CR network as well as block sensing-based CR network versus probability of collision for the case of $C_1/C_0 = 0.1$. We can observe that the average throughput in the quickest sensing scenario goes as low as 0.068 as the probability of collision converges to 0.64. This totally complies with the analytical statement in Section 5.3.1 that the average normalized throughput in the quickest sensing scenario converges to $P(H_0)C_1/C_0$ as the block length increases and the probability of collision converges to $P(H_0)$. For the block sensing scenario, the average normalized throughput converges to $P(H_0)(1 - P_f)C_1/C_0 + P(H_1)(1 - P_d)$ and the probability of collision converges to $P(H_0)(1 - P_f) + P(H_1)(1 - P_d)$ as the block length grows. This is verified by this figure where in the case 2 scenario, for example, the average normalized throughput converges to 0.09 as the probability of collision goes to 0.65. Once again, we observe that for acceptable values of the probability of collision (e.g., below 0.2), the quickest sensing achieves much higher average throughput compared to the block sensing scheme. For instance, for the probability of collision of 0.04, the quickest
sensing achieves average normalized throughput of 0.53 whereas the average normalized throughput of the block sensing-based CR in case 2 is 0.015.

5.5 Chapter Summary

In this chapter, we studied the optimization of the frame duration of quickest sensing-based and block sensing-based cognitive radios (CRs), to achieve the maximum CR throughput while protecting the primary user by keeping the collision probability or PITR (primary interfered time ratio) below a certain threshold. Analytical expressions for the average throughput, probability of collision and PITR for both quickest and block-based sensing approaches were found. These expressions were used to find the optimal frame duration for a certain quality of service for the primary user and the CR with both sensing approaches. We compared the maximum average throughput in block-based and quickest sensing schemes and showed that the quickest sensing approach results in a considerably higher CR throughput while requiring shorter frame length and sensing time at realistically low primary interference levels. These results show the superiority of quickest sensing approach compared to the conventional block sensing approach in cognitive radio networks.
Chapter 6

Conclusion

6.1 Summary of Conclusions

The increasing demand for wireless services and the static spectrum allocation policies have led to spectrum scarcity in wireless communications. On the other hand, studies show that most of the radio frequency spectrum is under-utilized [3]. This is mainly due to the spectrum allocation policies that traditionally have fixed assignment of frequency resources to licensed or primary users. These policies do not allow unlicensed or secondary users to utilize rarely used allocated frequencies, even when no interference with the assigned service would occur.

Recently, cognitive radios (CRs) have been introduced as a potential solution to this problem [5,6]. In a widely studied form of CR technology sometimes known as opportunistic or dynamic spectrum access, the secondary users, are allowed to opportunistically utilize a frequency band allocated to primary users, when it is not being used. The CR network has to be able to detect the presence of the primary network’s signals, so as to avoid interfering with it. Therefore, spectrum sensing plays a crucial role in the successful deployment of CRs.

In Chapter 2, we presented a probabilistic inference approach for cooperative spec-
Chapter 6. Conclusion

This approach allows for modeling and accommodating the uncertainties and correlations in the cooperative sensing system and applying belief propagation for inferring the likelihood functions. We derived the NP-based LRT at the FC for hard, soft and quantized local decisions. Unlike most existing studies, we consider the presence of fading and noise in the transmission channels.

Due to lack of primary signal and channel knowledge in cognitive radios, most existing results in the distributed detection literature which assume known transmit signals, cannot be directly applied to the cooperative sensing problem. Therefore, in Chapter 3 we proposed a linear composite hypothesis testing approach which does not require any prior knowledge or estimates of the unknown parameters and performs very closely to the known-parameter optimal LRT.

We also studied quickest change detection for spectrum sensing in cognitive radios in Chapter 4. Presenting several cooperative quickest sensing approaches in cognitive radios, we derived the optimal CUSUM test statistics in each scenario. To tackle the problem of unknown primary signal and channel statistics in cognitive radios, we proposed linear test-based CUSUM algorithms, which unlike the existing procedures, do not require any parameter estimation or iterative algorithms.

We formulated and studied the sensing-throughput trade-off of the quickest sensing approach in Chapter 5. We maximized the throughput of a cognitive radio network with respect to the frame length when quickest sensing is used. The amount of interference to the primary network, measured by the probability of collision with primary users or the primary interference time ratio (PITR), was constrained. The corresponding problems when CRs use block sensing were also solved, assuming that collision with primary users causes a drop in throughput for the CR. We then compared the maximum achievable throughput with block and quickest sensing schemes and showed that for the same protection level to the primary network, the quickest sensing approach results in significantly higher average throughput. Our results revealed that the quickest sensing
approach is potentially more suitable for cognitive radio technology than the conventional block sensing approach.

6.2 Future Work

This thesis presents the probabilistic inference and graphical modeling approach for cooperative spectrum sensing and highlights only some of the potential for graphical modeling and message passing useful tools. Ever since many papers inspired by our work, have tried to apply probabilistic inference and message passing algorithms into different aspects of cooperative spectrum sensing [68–76]. Further investigations into the effects of a mismatch between assumed and true distributions, correlations over time and space of fading channels and activity pattern of the primary user can be a direction of future work. Another interesting direction for future research is application of probabilistic inference approach to distributed (decentralized) spectrum sensing. For instance, the spatial interactions among the secondary users can be represented on a factor graph where individual beliefs of the primary user activity can be reached through an iterative message passing algorithm among the secondary users. Preliminary research in this direction has been presented in [70].

The unknown parameter scheme proposed in Chapter 3, can be applied to many other unknown-parameter sensing scenarios that may arise in cognitive radio networks. In particular, a future research direction can be towards applying the proposed approach to distributed spectrum sensing problems with unknown parameters.

Another possible future work is applying the proposed sensing approaches in a femtocell network framework. The quickest sensing approach seems appropriate for femtocells as the primary users are mainly cellular networks. Developing suitable cooperative quickest sensing schemes for cognitive femtocells can be an interesting future work.

The throughput analysis presented in Chapter 5 was the first step towards throughput-
sensing tradeoff analysis for CR networks with quickest sensing. This topic can be further investigated and extended to cooperative quickest spectrum sensing as a future research direction. This extension should accommodate the spectrum sharing algorithms and multiple-access protocols developed for multi-user cognitive radio networks.
Bibliography


Appendix A

Asymptotically equivalent tests for GLRT

A.0.1 The Wald test

An asymptotically equivalent test for GLRT is the Wald test with statistic

\[ T_W(y^K) = \beta^T \mathbf{I}(\hat{\beta}) \hat{\beta}, \quad (A.1) \]

The analytical derivation of the MLE is complicated and even infeasible for our general cooperative noisy channel model. Hence, for cooperative sensing, we propose to use the asymptotically equivalent tests for GLRT that do not require the MLE of the parameters and are easier to compute.

A.0.2 The Rao test

The Rao test is an asymptotically equivalent test for GLRT that does not require the MLE of the unknown parameters and is simpler computationally [15]. The Rao test statistic is given by

\[ T_R(y^K) = \frac{\partial \ln p(y^K | \beta)}{\partial \beta} \bigg|_{\beta = 0}^T \Gamma^{-1}(0) \frac{\partial \ln p(y^K | \beta)}{\partial \beta} \bigg|_{\beta = 0} \quad (A.2) \]
where $I(0)$ is the Fisher information matrix (see [28] pg. 40) evaluated at $\beta = 0$. The usual Rao test is a two-sided test, which necessitates the quadratic form in (A.2). In our problem, however, $\beta_{1,k}$’s and $\beta_{2,k}$’s are nonnegative which makes the parameter test one-sided and eliminates the necessity of the quadratic form.
Appendix B

Derivation of the Likelihood Function

Using the message-passing approach in [50], the message from node $I(u_k = \gamma_k(t_k))$ to node $t_k$ is derived as

$$
\mu_{I(u_k = \gamma_k(t_k)) \rightarrow T_k} = \sum_{u_k} p(y_k | u_k) I(u_k = \gamma_k(t_k)) = \sum_{i=1}^{M_k} p(y_k | q_i) I(a_{(i-1)k} < t_k < a_{ik})
$$

where $I(A) = 1$ if event A is true, and $I(A) = 0$ otherwise. Consequently, the outgoing message to the $Z$ node from $p(t_k | h_k, z)$ is computed as (see [50])

$$
\mu_{p(t_k | h_k, z) \rightarrow Z} = \sum_{i=1}^{M_k} p(y_k | q_i) \int_{a_{(i-1)k}}^{a_{ik}} p(t_k | z) dt_k.
$$

By multiplying all the messages arrived at node $Z$ and using the fact that $t_k$ has Gaussian distribution as in (2.18), the likelihood function is obtained by

$$
p(y^K_1 | z) = \prod_{k=1}^{K} \mu_{p(t_k | h_k, z) \rightarrow Z},
$$

$$
= \prod_{k=1}^{K} \left( \sum_{i=1}^{M_k} p(y_k | q_i) \left[ Q \left( \frac{a_{(i-1)k} - m_{tk}}{\sigma_{tk}} \right) \right. \right. \left. \left. - Q \left( \frac{a_{ik} - m_{tk}}{\sigma_{tk}} \right) \right] \right)
$$

(B.3)
where

\[
m_{t_k} = \begin{cases} 
N\sigma^2_{n_k}, & \text{under } \mathcal{H}_0 \\
 m_{t;\mathcal{H}_1}, & \text{under } \mathcal{H}_1 
\end{cases}
\]  \hspace{1cm} (B.5)

\[
\sigma^2_{t_k} = \begin{cases} 
N\sigma^4_{n_k}, & \text{under } \mathcal{H}_0 \\
 \sigma^2_{t_k;\mathcal{H}_1}, & \text{under } \mathcal{H}_1 
\end{cases}
\]  \hspace{1cm} (B.6)

with \( m_{t;\mathcal{H}_1} \) and \( \sigma_{t;\mathcal{H}_1} \) respectively given in (2.25) and (2.26)
Appendix C

Conditional distributions

In order to find the conditional expected values in (5.6), we can use the fact that for nonnegative random variables

\[ E[X] = \int_0^{\infty} (1 - F(x'))dx', \quad (C.1) \]

where \( F_X(x) \) is the cumulative distribution function (CDF) of random variable \( X \). Thus, we need to find the conditional CDFs of the random variables whose expected values are required.

C.1 Deriving the expected value in (5.13)

The conditional CDF required for \( E[X \mid X + T \leq T_0, T \leq T_0] \) is

\[ F_{X \mid X+T \leq T_0, T \leq T_0}(x) = \begin{cases} \frac{P(X \leq x, X+T \leq T_0, T \leq T_0)}{P(X \leq T_0, X+T \leq T_0, T \leq T_0)} & x \leq T_0 \\ 1 & x > T_0 \end{cases} \quad (C.2) \]

where \( P(X \leq x, Z \leq T_0, T \leq T_0) \) is derived as

\[
P(X \leq x, Z \leq T_0, T \leq T_0) = \int_0^x \int_0^{T_0-x'} f_{X,T}(x', t')dx'dt'
= 1 - e^{-\lambda x} - \frac{\lambda e^{-\alpha T_0}}{\lambda - \alpha} \left(1 - e^{-(\lambda - \alpha)x}\right) \quad (C.3)
\]
By substituting $t$ with $T_0$ we attain $P(X \leq T_0, X + T \leq T_0, T \leq T_0)$ in the denominator of (C.2). Taking the integral in (C.1), results in the expected value in (5.13).

C.2 Deriving the expected value in (5.14)

We derive the required conditional CDF as follows

$$F_{T|X+T\leq T_0, T\leq T_0}(t) = \begin{cases} P(T \leq t, X+T \leq T_0, T \leq T_0) & t \leq T_0 \\ P(T \leq t_0, X+T \leq T_0) & t > T_0 \end{cases} \quad (C.4)$$

where

$$P(T \leq t, Z \leq T_0, T \leq T_0) = \int_{0}^{t} \int_{0}^{T_0-t'} f_{X,T}(x', t') dx' dt'$$

$$= 1 - e^{-at} - \frac{\alpha e^{-\lambda T_0}}{\lambda - \alpha} \left( e^{(\lambda-\alpha)T_0} - 1 \right) \quad (C.5)$$

The denominator of (C.4) can be attained by evaluating (C.5) at $t = T_0$. Consequently, the expected value in (5.14) is derived by taking the integral in (C.1).

C.3 Deriving the expected value in (5.15)

The required conditional CDF for calculating $E[T|X + T > T_0, T \leq T_0]$ is

$$F_{T|X+T>T_0, T\leq T_0}(t) = \begin{cases} P(T \leq t, X+T > T_0, T \leq T_0) & t \leq T_0 \\ P(T \leq T_0, X+T > T_0) & t > T_0 \end{cases} \quad (C.6)$$

We can derive $P(T \leq t, Z > T_0, T \leq T_0)$ as follows

$$P(T \leq t, Z > T_0, T \leq T_0) = \int_{0}^{T_0} \int_{T_0-x'}^{t} f_{X,T}(x', t') dx' dt' + \int_{T_0-t}^{T_0} \int_{T_0-x'}^{t} f_{X,T}(x', t') dx' dt'$$

$$= \frac{e^{-aT_0}}{\lambda - \alpha} \left( 1 - e^{-(\lambda-\alpha)T_0} - e^{-(\lambda+a)T_0} (e^{at} - 1) + e^{-aT_0} (1 - e^{-\lambda T_0}) \right) \quad (C.7)$$

By evaluating (C.7) at $t = T_0$, we can calculate $P(T \leq T_0, X + T > T_0)$. Finally, (5.15) is achieved by calculating the integral in (C.1).