Quickest Spectrum Sensing with Multiple Antennas: Performance Analysis in Various Fading Channels

Effariza binti Hanafi

A thesis presented for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering
at the
University of Canterbury,
Christchurch, New Zealand.

30th May 2014

To my parents, Hanafi and Zuraini and

beloved husband, Nik Mohd Faiz

ABSTRACT

Traditional wireless networks are regulated by a fixed spectrum assignment policy. This results in situations where most of the allocated radio spectrum is not utilized. In order to address this spectrum underutilization, cognitive radio (CR) has emerged as a promising solution. Spectrum sensing is an essential component in CR networks to discover spectrum opportunities. The most common spectrum sensing techniques are energy detection, matched filtering or cyclostationary feature detection, which aim to maximize the probability of detection subject to a certain false alarm rate. Besides probability of detection, detection delay is also a crucial criterion in spectrum sensing. In an interweave CR network, quick detection of the absence of primary user (PU), which is the owner of the licensed spectrum, allows good utilization of unused spectrum, while quick detection of PU transmission is important to avoid any harmful interference.

This thesis consider quickest spectrum sensing, where the aim is to detect the PU with minimal detection delay subject to a certain false alarm rate. In the earlier chapters of this thesis, a single antenna cognitive user (CU) is considered and we study quickest spectrum sensing performance in Gaussian channel and classical fading channel models, including Rayleigh, Rician, Nakagami-m and a long-tailed channel. We prove that the power of the complex received signal is a sufficient statistic and derive the probability density function (pdf) of the received signal amplitude for all of the fading cases. The novel derivation of the pdfs of the amplitude of the received signal for the Rayleigh, Rician and Nakagami-m channels uses an approach which avoids numerical integration. We also consider the event of a mis-matched channel, where the cumulative sum (CUSUM) detector is designed for a specific channel, but a different channel is experienced. This scenario could occur in CR network as the channel may not be known and hence the CUSUM detector may be experiencing a different channel. Simulations results illustrate that the average detection delay depends greatly on the channel but very little on the nature of the detector.

Hence, the simplest time-invariant detector can be employed with minimal performance loss.

Theoretical expressions for the distribution of detection delay for the time-invariant CUSUM detector, with single antenna CU are developed. These are useful for a more detailed analysis of the quickest spectrum sensing performance. We present several techniques to approximate the distribution of detection delay, including deriving a novel closed-form expression for the detection delay distribution when the received signal experiences a Gaussian channel. We also derive novel approximations for the distribution of detection delay for the general case due to the absence of a general framework. Most of the techniques are general and can be applied to any independent and identically distributed (i.i.d) channel. Results show that different signal-to-noise ratio (SNR) and detection delay conditions require different methods in order to achieve good approximations of the detection delay distributions. The remarkably simple Brownian motion approach gives the best approximation for longer detection delays. In addition, results show that the type of fading channel has very little impact on long detection delays.

In later chapters of this thesis, we employ multiple receive antennas at the CU. In particular, we study the performance of multi-antenna quickest spectrum sensing when the received signal experiences Gaussian, independent and correlated Rayleigh and Rician channels. The pdfs of the received signals required to form the CUSUM detector are derived for each of the scenarios. The extension into multiple antennas allows us to gain some insight into the reduction in detection delay that multiple antennas can provide. Results show that the sensing performance increases with an increasing Rician K-factor. In addition, channel correlation has little impact on the sensing performance at high SNR, whereas at low SNR, increasing correlation between channels improves the quickest spectrum sensing performance. We also consider mis-matched channel conditions and show that the quickest spectrum sensing performance at a particular correlation coefficient or Rician K-factor depends heavily on the true channel irrespective of the number of antennas at the CU and is relatively insensitive to the channel used to design the CUSUM detector. Hence, a simple multi-antenna time-invariant detector can be employed. Based on the results obtained in the earlier chapters, we derive theoretical expressions for the detection delay distribution when multiple receive antennas are employed at the CU. In particular, the approximation of the detection delay distribution is based on the Brownian motion approach.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisors, Assoc. Prof. Philippa Martin, Prof. Peter Smith and Dr. Alan Coulson for their constant guidance, patience, and motivation through my doctoral work.

My appreciation also goes to the Ministry of Higher Education Malaysia and University of Malaya for their scholarship and financial supports to pursue my PhD. Also, thanks to Assoc. Prof. Philippa Martin and University of Canterbury for providing funding for me to attend the AusCTW 2013 and 2014, in Adelaide and Sydney, Australia, respectively. I am grateful to my supervisors who gave me the great opportunity to visit Hong Kong University of Science & Technology (HKUST). Then, I would also like to thank Assoc. Prof. Matthew McKay and Prof. Ross D. Murch from HKUST for hosting me during my visit to Hong Kong, and for exchanging useful information for my thesis.

My research experience in graduate school was greatly enriched by the intellectual environment in the Communications Lab. I would also like to thank several current and former members of the Communications Lab for valuable advice on research.

I would also like to thank my friends in Christchurch, who have been a constant source of encouragement and companionship. Last but not least, I would like to give my special thanks to my beloved family and husband, who were miles away yet had been giving continuous moral support and encouragement through every step in this endeavor.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AF	amount of fading.
ARQ	automatic repeat request.
AWGN	additive white Gaussian noise.
BM	Brownian motion.
cdf	cumulative distribution function.
CDMA	code division multiple access.
CSCG	circularly symmetric complex Gaussian.
CSI	channel state information.
CU	cognitive user.
CUSUM	cumulative sum.
CV	coefficient of variation.
DAS	distributed antenna system.
DPC	dirty paper coding.
DSA	dynamic spectrum access.
EGC	equal gain combining.
FAR	false alarm rate.
FCC	Federal Communications Commission.
FEC	forward error correction.
GLRT	generalized likelihood ratio test.
i.i.d	independent and identically distributed.
ISI	intersymbol interference.
LO	local oscillator.
LOS	line-of-sight.
LRT	likelihood ratio test.
MI	multipath interference.

MISO	multiple-input single-output.
MIMO	multiple-input multiple-output.
MRC	maximal ratio combining.
OFDM	orthogonal frequency division multiplexing.
OSA	opportunistic spectrum access.
pdf	probability density function.
PU	primary user.
QoS	quality of service.
RAU	remote antenna unit.
RF	radio frequency.
RW	random walk.
SC	scattered component.
SIC	successive interference cancellation.
SIMO	single-input multiple-output.
SINR	signal-to-interference-plus-noise ratio.
SISO	single-input single-output.
SNR	signal-to-noise ratio.
SPRT	sequential probability ratio test.
TDMA	time division multiple access.
TWDP	two-wave with diffuse power.
ULA	uniform linear array.
UWB	ultra-wide band.
WSS	wide-sense stationary.

LIST OF SYMBOLS

$\delta(1)$	Dirac delta function.
$\det(\mathbf{A})$	Determinant of matrix $\mathbf{A}$.
$\exp(.)$	Exponential function.
γ	Threshold.
$\Gamma(.)$	Gamma function.
$\Gamma(.,.)$	Incomplete Gamma function.
inf	Infimum.
λ	Parameter of an exponential distribution.
ln	Natural logarithm.
$\mathcal{CN}(\mu, \sigma^2)$	CSCG random variable with mean μ and variance σ^2.
$\mathcal{G}(\alpha, \beta)$	Gamma distributed random variable with shape parameter α and scale parameter β.
max	Maximum.
$\overline{T}_d$	Worst-case detection delay.
$\overline{T}_f$	Mean number of samples to false alarm.
$\sim$	Is distributed as.
τ	Sample number at which an abrupt change actually occurs.
$\mathbf{I}_n, \mathbf{I}$	$n \times n$ Identity matrix.
$\mathrm{B}(a, b)$	Beta function with parameters a and b
$\mathrm{diag}[a_1, ..., a_N]$	Square diagonal matrix with $a_1, ..., a_N$ on the main diagonal.
$\mathrm{F}(\alpha, \beta)$	F distributed random variable with parameters α and β.
$\mathrm{I}_\nu(.)$	Modified Bessel function of the first kind of order ν.
$\mathrm{J}_0(.)$	Zeroth order Bessel function of the first kind.

$\mathrm{K}_\nu(.)$	Modified Bessel function of the second kind of order ν.
Var(.)	Variance.
C_n	CUSUM statistic.
$D(. \mid\mid .)$	Kullback-Leibler divergence.
$E[.]$	Statistical expectation.
$f(.)$	The probability density function.
f_D	Maximum Doppler shift in Hertz.
j	$\sqrt{-1}$.
K	Rician K-factor.
M	Number of antennas.
m	Nakagami-m fading parameter.
R	Number of rectangles in the Riemann sum
T	Sample number at which an abrupt change is detected.
T_c	Channel coherence time.
x^+	$\max(x, 0)$
X^*	Complex conjugate of X.
$X^\dagger$	Conjugate transpose of X.
X^T	Transpose of X.
(.)!	Factorial.
ess sup	Essential supremum.

Chapter 1

INTRODUCTION

The explosive growth in the use of wireless devices and applications over recent years illustrates the huge and growing demand for high data rates. One of the most important resources required for wireless communications is spectrum. Currently, wireless networks are characterized by a fixed spectrum allocation policy, in which national regulatory bodies, such as the Federal Communications Commission (FCC) in the United States, exclusively allocate spectrum bands to specific licensed users on a long-term basis for large geographical regions and no violation from unlicensed users is allowed [6, 7]. Although the fixed spectrum allocation policy ensures licensed users do not interfere with each other, it cannot accommodate the dramatic increase in spectrum access required by high data rate devices and mobile services [8, 9]. Therefore, the limited wireless spectrum has traditionally been viewed as a scarce resource in high demand [10]. In addition to the limited available spectrum, the current policy suffers from inefficient spectrum utilization. Several studies and reports have shown that most of the allocated spectrum is not utilized and is sitting idle (i.e there are spectrum holes) [1, 7, 11]. Further, the FCC shows that the spectrum utilization for licensed frequency bands ranges from 15% to 85% depending on the temporal and geographic situation [1, 2, 12]. In addition, spectrum measurements are also conducted in Europe [13, 14], in particular France, Czech Republic and Spain. In general, results show that spectrum occupancy is moderate below 1 GHz and very low above 1 GHz.

The limited spectrum availability and the inefficient usage of the spectrum necessitate a new communication paradigm which would improve the efficiency of the utilization of the licensed spectrum while keeping up with growing demand in the wireless communication industry [6, 15]. Dynamic spectrum access (DSA) or opportunistic spectrum access (OSA) is proposed by the FCC as an alternative policy to solve these spectrum shortage and inefficiency problems [1,3,7,9]. In this thesis, the term DSA will be used to refer to this new communications paradigm. DSA is based on the concept that a portion of the licensed spectrum can be accessed opportunistically by a given radio when the spectrum is unused and hence this leads to an improved spectrum utilization [6, 16]. The key enabling technology for implementing efficient DSA is cognitive radio [6,7,9].

1.1 OVERVIEW OF THE COGNITIVE RADIO CONCEPT

Cognitive radio, originally proposed in [17], is a promising technology to avoid the under utilization of the wireless spectrum by dynamic access of available spectral opportunities [3]. In cognitive radio terminology, primary user is defined as the licensed or authorized owners of a given frequency band, which has a higher priority of using the spectrum band [8, 18]. On the other hand, cognitive user is defined as the unlicensed or secondary user that is allowed to opportunistically use the spectrum band when the primary user is absent and has a lower priority on the usage of the spectrum [8, 18]. According to the FCC, the term cognitive radio can be defined as follows [19]:

> "*A radio or system that senses its operational electromagnetic environment and can dynamically and autonomously adjust its radio operating parameters to modify system operation, such as maximize throughput, mitigate interference, facilitate interoperability, access secondary markets.*"

Based on this definition, the two main characteristics of a cognitive radio that distinguish it from conventional radio devices are cognitive capability and reconfigurability [1,6,7,20–23]. The cognitive capability of a cognitive radio refers to the cognitive radio's ability to sense and capture the information from the surrounding radio environment [1, 6, 7, 20]. This allows the cognitive user to be aware of the transmitted waveform, radio frequency (RF) spectrum, geographical information and subsequently analyze the gathered information to identify any unused spectrum at a specific time and location that could be exploited [6, 7]. With this capability, the cognitive radio can then determine the appropriate operating parameters, the best transmission strategy to

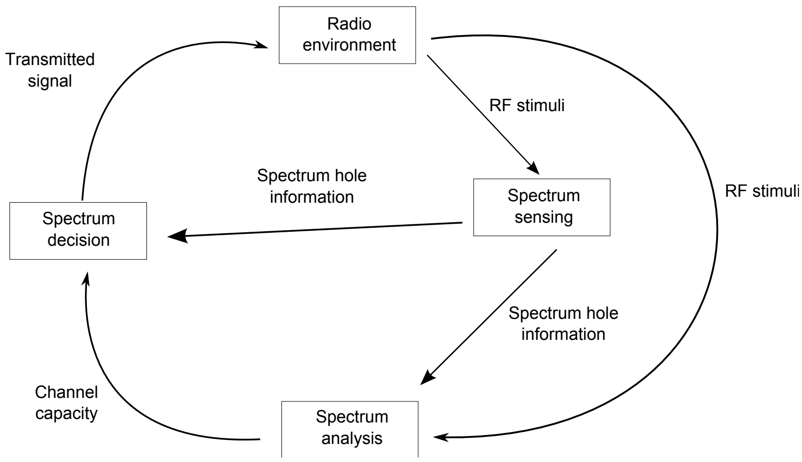

Figure 1.1 Cognitive cycle [1].

employ and the best available spectrum [3,6]. The cognitive radio tasks as well as its interaction with the radio environment are illustrated in Figure 1.1. As can be seen from Figure 1.1, the cognitive cycle consists of three major components which are spectrum sensing, spectrum analysis and spectrum decision. A brief overview of each of these components is given as follows [1,6]:

1. **Spectrum sensing**: In spectrum sensing, a cognitive radio observes the frequency band and gathers necessary information regarding its surrounding radio environment. Based on the information captured, the cognitive radio is then able to detect any spectrum holes.

2. **Spectrum analysis**: Once spectrum holes are detected using spectrum sensing, each of the spectrum bands is characterized based on the local observation of the cognitive radio as well as the statistical information of the primary user network. The spectrum holes characteristics are then analyzed and estimated.

3. **Spectrum decision**: Based on the spectrum analysis, the cognitive radio determines the operating parameters such as the data rate, the transmission mode and the bandwidth available for the transmission. The most appropriate spectrum band is selected based on the spectrum band characterization and the user requirements.

As mentioned earlier, the second key feature of a cognitive radio that distinguishes it from a traditional radio is reconfigurability. Reconfigurability refers to the ability of a cognitive radio to intelligently adapt to the radio environment by adjusting its operating parameters according

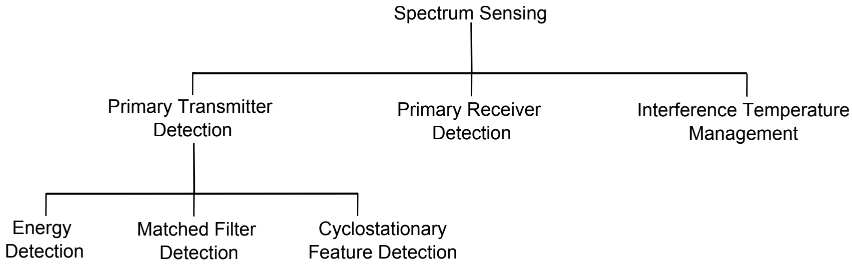

Figure 1.2 Classification of spectrum sensing.

to the sensed environment variations in order to achieve the optimal performance [1, 7, 21, 23]. For example, since the radio environment keeps changing due to the primary user starting and completing transmissions, there are changes in the frequency of the spectrum holes [23]. Therefore, a cognitive user equipped with cognitive radio capabilities must be able to re-tune its transmitting frequencies when the frequencies of the available spectrum band change. This is important in order to exploit the unused part of the spectrum [23]. This thesis will focus on one of the key cognitive capabilities of a cognitive radio which is spectrum sensing.

1.2 SPECTRUM SENSING FOR COGNITIVE RADIO

A cognitive radio is designed to measure, learn, sense, be aware of the changes in its surrounding and adapt itself to the radio's operating environment, which makes spectrum sensing an important component for the establishment of cognitive radio networks [2, 3]. Spectrum sensing enables a cognitive radio user to determine the spectrum availability in order to improve the spectrum's utilization without causing any harmful interference to the primary user [2, 3]. This capability is required in two scenarios. The first scenario is when the cognitive users detect that a certain frequency band is not being used by the primary user [3]. In this case, the primary user has stopped transmission and there exists a spectrum opportunity. The second scenario is when the cognitive users monitor the frequency band during its transmission to detect the existence of the primary user in order to vacate the channel without causing any significant interference to the primary user [3]. Generally, spectrum sensing techniques can be classified as primary transmitter detection, primary receiver detection and interference temperature management [2]. This is illustrated in Figure 1.2.

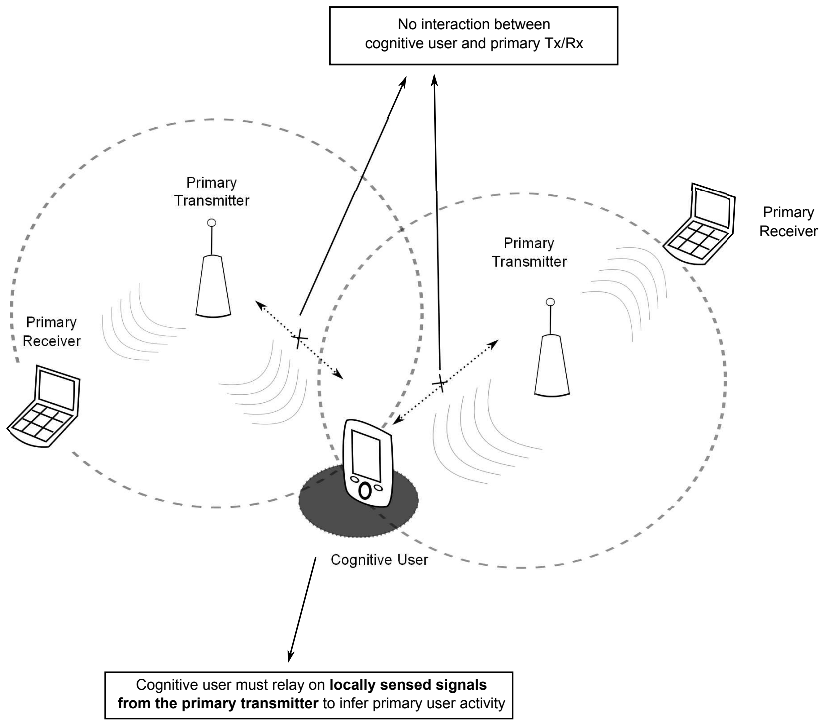

Figure 1.3 Primary transmitter detection [2, 3].

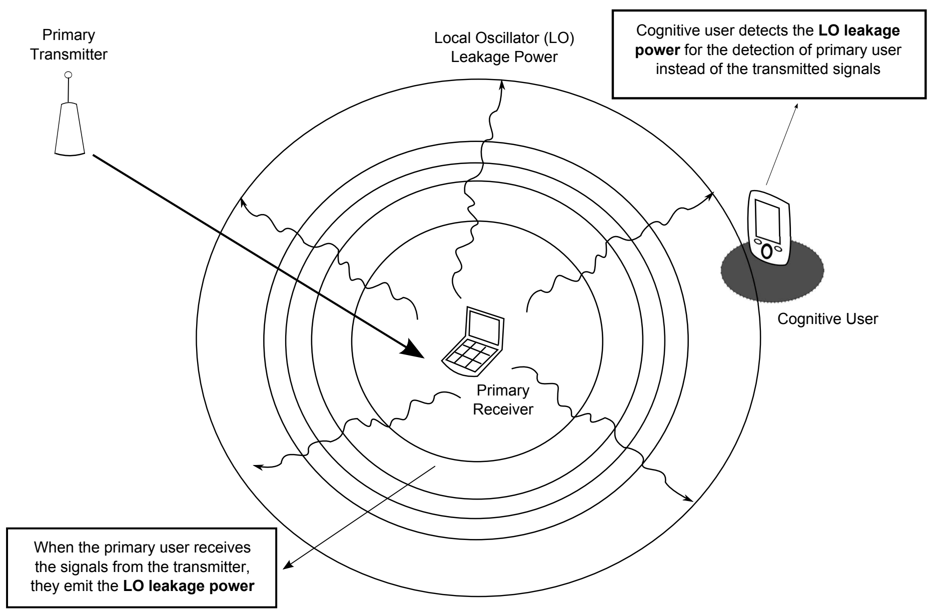

Figure 1.4 Primary receiver detection [2, 3].

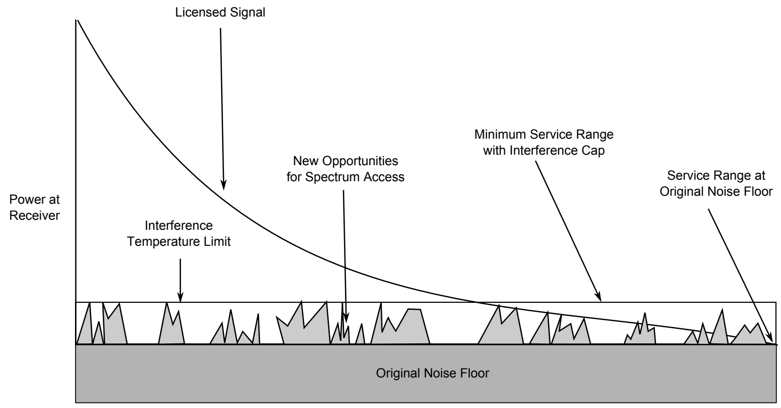

Figure 1.5 Interference temperature management [2].

Primary transmitter detection is based on detecting the primary user transmission using local observations at the cognitive user [2]. It is usually assumed that the cognitive user has no real-time interactions with the primary user transmitter and receiver [2, 3]. Thus, as shown in Figure 1.3, the cognitive user has no exact information on the current transmissions within the primary user network. In order to determine the spectrum availability, the cognitive user needs to detect the signal from the primary user transmitter based on its local observation through the frequency band [1–3].

Another way of detecting any unused spectrum is by employing the primary receiver detection techniques, which is shown in Figure 1.4. This technique aims at detecting the primary user that is receiving data within the communication range of the cognitive radio user [1–3]. The primary user's receiver in Figure 1.4 contains an inevitable reverse leakage and hence, some of the local oscillator (LO) power couples back through the input port and radiates out of the antenna [24]. This is illustrated in Figure 1.4, where the primary user usually emits LO leakage power from its RF front-end when it receives the signal from the primary transmitter [2, 3]. Therefore, in detecting the presence of the primary user by employing the primary receiver detection approach, the cognitive user detects the LO leakage power, instead of the signal from the primary transmitter. However, in practice, direct measurement of the channel between the primary transmitter and receiver is difficult to obtain [1].

Interference temperature is an interference assessment metric proposed by the FCC which aims to measure the interference experienced by the primary user [16]. Based on the interference temperature concept, the FCC established an interference temperature limit, which is shown in

Figure 1.5. The value of the interference temperature limit is set depending on the maximum amount of interference the primary user can tolerate in its frequency band [7, 16]. Thus, based on Figure 1.5, the cognitive user using the spectrum band must guarantee that its transmission in addition to the existing noise and interference shall not exceed the interference temperature limit at the primary receiver [2]. However, this model is difficult to employ since it requires the cognitive user to accurately measure the interference temperature [2, 3]. This is challenging because the cognitive user is not usually able to distinguish between the actual signals from the primary user and noise or interference [2, 3].

Based on the difficulties that lie in employing primary receiver detection and interference temperature management, most of the literature on spectrum sensing focuses on primary transmitter detection to identify the presence or absence of the primary user signal transmission [1, 3, 7]. In [2,3,7,8,15,25,26], detailed surveys of various spectrum sensing techniques for primary transmitter detection are presented. As shown in Figure 1.2, the most common spectrum sensing techniques for the primary transmitter detection are energy detection, matched filter detection and cyclostationary feature detector based sensing [1, 7, 27–29].

1.2.1 Energy detection

Energy detection is the most widely used sensing technique and it is the simplest form of sensing because of its simplicity and low computational and implementation complexities [7, 8, 16, 18, 23, 30–32]. It does not require any information on the primary user signals and hence, it is considered to be a blind detection technique [7,8,16,18,23,32]. In energy detection, the presence or absence of the primary user is detected based on the energy in the signal received by the cognitive user [2, 3, 26]. In particular, the detection statistic of the energy detector, which is defined as the average or total energy of a certain number of observed samples, is compared with a predetermined threshold in order to determine whether the primary user exists or is absent [1, 2, 7, 8, 16, 18, 23, 26, 30, 33]. The performance of the energy detector is evaluated in terms of the probability of detection and the probability of false alarm [1, 7, 8, 23]. The probability of detection is defined as the probability that the energy detector correctly decides on the primary user's existence in the spectrum while the probability of false alarm denotes the probability that the energy detector decides that the primary user is present while it is actually absent [8]. The goal of energy detection is to maximize the probability of detection subject to a constrained/predefined probability of false alarm [7, 23].

Although energy detector is easy to implement, there are some drawbacks to using it. Firstly, the threshold used in energy detection to detect the primary user transmission depends on the noise floor, which can change considerably over time [7, 8, 18, 26, 29, 31]. Thus, it is difficult to set the threshold level correctly [8, 18, 26, 31]. Another shortcoming of energy detection is that its performance is susceptible to uncertainty of the noise power which is due to the dependency on the signal-to-noise ratio (SNR) of the received signal [1–3, 16, 30, 33]. In addition, the energy detector cannot distinguish between a primary user signal and other types of signals (e.g. signals from cognitive users sharing the same channel with the primary user, noise and interference) since the decision on the primary user's existence is based on the received signal energy [2, 3, 7, 8, 16, 18, 26, 29–31, 33]. Therefore, it is prone to a high false alarm probability as the energy detector is triggered by signal sources other than the primary user [1–3, 7, 23, 33]. Moreover, energy detection has a poor performance in low SNR regimes, where the noise power is very high [7, 8, 18, 23].

1.2.2 Matched filtering

Matched filtering is the optimal detection approach when the information of the primary user transmission is known as it maximizes the received signal SNR [1–3, 7, 8, 23, 25, 26, 29, 31, 33–35]. The performance of matched filtering is optimal in an additive white Gaussian noise (AWGN) channel [1, 33]. Existence of the primary user is detected by comparing the matched filter output with a threshold [3, 26]. The objective of matched filtering is to maximize the probability of detection [32].

However, matched filtering suffers from a number of shortcomings. It has a high implementation complexity because it requires receivers for all types of signal [3, 7, 8, 26, 29, 31]. Moreover, the power consumption of the matched filter is too high as detection of the primary user transmission requires various receiver algorithms to be executed [7, 8]. Another disadvantage of using matched filtering is that it requires perfect knowledge of the primary user signalling characteristics such as operating frequency, bandwidth, modulation type and order, pulse shaping and the packet format [1, 2, 7, 8, 26, 29]. However, if the matched filter uses inaccurate information on the primary user signal, this will result in poor sensing performance [1, 2, 7, 33, 34]. Thus, due to these drawbacks, the matched filtering technique is not practical in the context of spectrum sensing [31].

1.2.3 Cyclostationary feature detection

Another commonly used sensing technique is cyclostationary feature detection which exploits the cyclostationary features of the received signals to detect the presence of the primary user in a given spectrum [8, 16, 18, 33, 35]. Modulated signals exhibit cyclostationary properties due to their statistical periodicity in mean and autocorrelation [1, 2, 8, 18, 29, 30]. In order to detect the existence of the primary user, the spectral correlation of the received signal is averaged over some interval and the result of this process is then compared with a test statistic to determine the spectrum occupancy [3].

There are some advantages in using a cyclostationary feature detector such as its ability to distinguish the primary user signal from the noise [1, 2, 7, 8, 18, 23, 26, 33]. This is due to the fact that the modulated signals are generally coupled with sine wave carriers, pulse trains, repeating spreading, hoping sequences or cyclic prefixes, which result in embedded signal periodicity while noise is a wide-sense stationary (WSS) signal without correlation [1, 2, 7, 8, 18, 26, 29, 32, 33]. In addition, the cyclostationary feature detector is also able to distinguish between different types of transmission and primary users since different primary user transmissions exhibit different cyclostationary features [2, 3, 7, 8, 30]. Moreover, the cyclostationary feature detector is robust to uncertainty in the value of the noise power and hence this leads to a better sensing performance in low SNR regimes [1–3, 7, 23, 26, 29, 30, 33]. However, cyclostationary feature detectors have a high implementation and computational complexity [1, 3, 16, 23, 26, 33]. Another drawback of cyclostationary feature detection is that it requires a significantly longer observation period for adequate detection performance [1–3, 23, 26, 29, 33].

The aforementioned spectrum sensing techniques are summarized and compared in Table 1.1.

Table 1.1 Summary of most common spectrum sensing techniques

Type	Advantages	Disadvantages
Energy detection	• Low computational and implementation complexities • Does not require information on primary user signals	• Difficult to set threshold level correctly • Performance susceptible to noise power uncertainty • Cannot distinguish a primary user from other signal sources • Prone to high false alarm probability • Poor performance in low SNR
Matched filtering	• Optimal detection approach when primary user's information is known • Optimal performance in AWGN channel	• High implementation complexity • Power consumption is too high • Requires perfect knowledge of primary user signalling characteristics
Cyclostationary feature detection	• Able to distinguish primary user signal from noise • Able to distinguish different types of transmission and primary users • Robust to noise uncertainty and better performance in low SNR	• High implementation and computational complexity • Requires longer observation period for adequate detection performance

1.3 QUICKEST SPECTRUM SENSING

The spectrum sensing approaches that we discussed above are based on a classical detection framework, which aim to maximize the probability of detection subject to a certain false alarm rate [36, 37]. These approaches use a fixed length of data sequence or in other words, a fixed number of required samples (i.e sensing time window) to make a decision on the presence or absence of the primary user [30, 37, 38]. Based on this block-based detection feature, these sensing techniques can be classified as block detection schemes, in which the secondary user observes a block of samples, computes a test statistic from the block of observations and finally makes a decision by comparing the test statistic with a threshold [36, 37, 39]. The difficulty of the block detection scheme lies in determining the size of the block, where a small block size may lead to an inaccurate and unreliable sensing decision [39]. Although sensing accuracy is increased with a larger block size, it may yield a longer detection delay [39, 40].

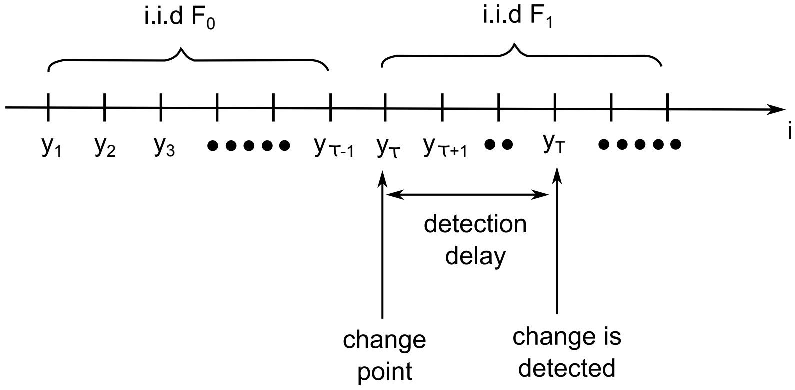

Figure 1.6 Illustration of quickest detection problem.

Besides the probability of detection, detection delay is also a crucial criterion in spectrum sensing [36,37]. When the primary user stops transmission, the cognitive user must be able to detect the absence of the primary user as soon as possible. This enables the cognitive user to fully utilize the unused spectrum for its transmission. In contrast, when the primary user starts using the spectrum band again, the cognitive user needs to detect the existence of the primary user as quickly as possible in order to vacate the channel without causing any significant interference to the transmission of the primary user. Therefore, a detection framework which facilitates minimal detection delay is of significant interest.

In quickest detection problems, samples are observed sequentially $\{Y[i] : i = 1, 2, ..\}$ [30]. In the standard quickest detection framework, the observations are initially independent and identically distributed (i.i.d) according to distribution F_0 and at some unknown sample number, τ, the observation's distribution changes abruptly to i.i.d F_1 such that $Y[i] \sim F_0$ for $i \leq \tau - 1$ and $Y[i] \sim F_1$ for $i \geq \tau$. This is illustrated in Figure 1.6. The distribution will change if the primary user becomes operational (i.e primary user starts transmitting). As can be seen from Figure 1.6, quickest detection refers to a detection framework which detects the abrupt change as soon as possible after the change occurs, where an algorithm will raise an alarm to declare that a change is detected [7,30,36,41–46]. The objective is to detect the occurrence of the change with minimal detection delay, such that the delay between the point at which the change actually occurs and the point at which the algorithm detects such a change is minimized, subject to a certain false alarm rate [30, 36, 43–46]. Hence, the quickest detection framework suits the cognitive radio scenario.

In the context of spectrum sensing, primary user activity changes at some unknown point which leads to a change in the distribution of the cognitive user's received signal [36, 37]. Thus, the quickest detection theory which detects abrupt changes in the observation distribution can be applied to spectrum sensing in order to detect the change in spectrum occupancy based on sequential observation [30,37,39,41–44]. The adoption of quickest detection theory into spectrum sensing for cognitive radio systems is known as quickest spectrum sensing in the literature [37, 39, 41, 42]. The goal of quickest spectrum sensing is to detect the existence or absence of a primary user based on the occurrence of a change in the cognitive user's received signal, using the fewest number of samples (i.e. with minimal detection delay) conditioned on the false alarm constraint [36, 37, 39, 41, 42]. By using this approach, agile and robust spectrum sensing is achieved [7, 41]. Hence, it will be the focus of this thesis.

1.4 PROBLEM STATEMENT AND FOCUS

There have been a number of studies on quickest spectrum sensing considering different scenarios [36,39,41,45–48]. In [36], various scenarios are considered depending on the prior information the cognitive user has about the primary user. A successive refinement test is proposed in [41], where it is assumed that the primary user signal is a sinusoid signal. In [39], cyclostationary features of the primary user signal are exploited and incorporated into the quickest spectrum sensing. Spectrum sensing based on the quickest detection framework over multiple frequency channels is considered in [47], where the author studies the procedure of the cognitive user vacating channels when the primary user starts transmitting as well as aiming to find the best channel(s) to sense. Another quickest detection problem in multiple frequency channel is considered in [45,46,48], where a Bayesian formulation of quickest change detection in multiple on-off processes is proposed within a decision-theoretic framework. In [49], a throughput-sensing tradeoff for quickest spectrum sensing is studied where the cognitive radio throughput is maximized with respect to frame length. However, no studies on quickest spectrum sensing have appeared which consider a range of fading channels. Furthermore, no studies have also given a theoretical expression for the distribution of detection delay, which is beneficial in analyzing the quickest spectrum sensing performance. In addition, all the studies mentioned above only consider single antenna cognitive users.

Therefore, this thesis contains an extensive investigation on the performance of quickest spectrum sensing for both single and multiple antenna cognitive users over a time-invariant channel and several fading channels, including the classical fading channels such as Rayleigh, Rician, Nakagami-m as well as a long-tailed channel, which models severe fading. Apart from independent channels, temporally and spatially correlated channels are also being considered in this thesis. Based on the result of these studies, novel theoretical expressions for the distribution of detection delay are developed for both single and multiple antenna cases which are useful in providing a more detailed analysis of the quickest spectrum sensing performance.

In this thesis, comparison with other existing algorithms, which are based on the classical detection schemes such as energy detection, matched filtering, cyclostationary feature detection and generalized likelihood ratio test (GLRT) is challenging. As mentioned in Section 1.3, the classical detection framework has a different goal than quickest spectrum sensing. In particular, the classical detection framework aims at maximizing the probability of detection subject to false alarm constraints. On the other hand, the quickest spectrum sensing algorithm aims to detect abrupt changes in the distribution of the received signals using the fewest received signals (i.e. minimizing the detection delay) while maintaining a certain false alarm rate. In addition, the classical detection schemes are based on a fixed block of samples, whereas the length of samples in quickest spectrum sensing varies depending on the information received from the cognitive user observations [30]. Thus, with very different objectives and requirements, it is hard to compare quickest spectrum sensing with the classical detection framework.

1.5 THESIS OUTLINE AND CONTRIBUTIONS

The organization and contributions of this thesis are as follows:

Chapter 2

This chapter presents relevant background material required for the subsequent chapters.

Chapter 3

This chapter investigates the performance of quickest spectrum sensing with single antenna cognitive users when the received signal experiences a time-invariant channel or various fading conditions, including Rayleigh, Rician, Nakagami-m and F channels. The received signal at the cognitive user is considered to be the product of two complex variables (the primary user signal and channel) with additive noise. We prove that the power of the complex received signal is a sufficient statistic and hence the log likelihood ratio can be computed on the basis of the received signal amplitude. This result allows us to derive a general form for the log likelihood ratio and the quickest spectrum sensing technique for various channel models. We derive the probability density function (pdf) of the received signal amplitude for the Rayleigh, Rician, Nakagami-m and long-tailed F channel scenarios. We use an approach which avoids numerical integration to derive the novel pdfs of the amplitude of the received signal for the most commonly used fading channel models, including Rayleigh, Rician and Nakagami-m. This approach is particularly useful in the Rician case where simple quadrature methods are unstable.

We also study the quickest spectrum sensing performance in the event of a mis-matched channel condition, where the cumulative sum (CUSUM) detector is designed for a specific channel, but experiences a different channel. This study allows us to gain some insights into the effects of the channel on the sensing performance as well as the robustness of the detector. Furthermore, we also consider the case of a temporally correlated Rayleigh channel. This study is useful for gaining further insights into the impact of various fade rates on the quickest spectrum sensing performance.

Chapter 4

This chapter presents several techniques to approximate the distribution of detection delay for a time-invariant CUSUM detector with a single antenna cognitive user when the received signal is transmitted over a Gaussian channel as well as over a mis-matched, Rayleigh channel. In particular, we derive a novel approximate closed-form expression for the distribution of detection delay for the Gaussian case. Furthermore, we derive novel approximations for the detection delay distribution for the general case due to the absence of a general framework. We also apply a simple random walk and Brownian motion theory with drift in deriving approximate expressions for the distribution of detection delay for both Gaussian and Rayleigh cases. Approximate methods are necessary since there is no exact solution available. Most of the methods discussed in the chapter are general and can be applied to any independent and identically distributed (i.i.d) channel. The validity of the approximate expressions for the detection delay distribution is verified using simulations.

The approximation methods that we presented are based on the modified detection delay statistic. Hence, we also analyze the accuracy of the modified detection delay statistic. This is valuable in understanding the limitations of the approximations used. Furthermore, we also approximate the probability of missed detection and provide an analysis of long detection delays to gain further insight into the factors that contribute to an increased probability of long detection delays.

Chapter 5

This chapter extends the study of quickest spectrum sensing into multiple receive antennas at the cognitive user when the received signal experiences Gaussian, Rayleigh and Rician fading channels. Apart from i.i.d channels, we also consider the case of an insufficient separation between multiple antennas on a cognitive user by looking at the effect of spatial-correlation for Rayleigh channels. We prove that the sum of the complex received signal powers at each antenna for the Gaussian and independent Rayleigh scenarios are sufficient statistics. Hence, the log likelihood ratio for both cases can be evaluated based on the complex received signal vector or the sum of the received signal powers. We derive the pdfs of the received signal for both Gaussian and independent Rayleigh fading scenarios using the sum of the received signal powers. The derivation of the pdf for the independent Rayleigh case uses an approach which avoids numerical integration over an infinite region. This approach gives a negligible error and performs better than numerical integration.

We also derive the joint pdfs of the received signal for the correlated Rayleigh and independent Rician cases. We also study the performance of multi-antenna quickest spectrum sensing in the event of mis-matched channel conditions. Furthermore, we analytically compute the upper bound and asymptotic worst-case detection delay for both Gaussian and independent Rayleigh cases. We also numerically evaluate the quickest spectrum sensing performance for the Rayleigh (independent and correlated) and Rician (independent and correlated) cases. The extension to multiple antennas allows us to gain further insights into the reduction in detection delay that multiple antennas can provide. In addition, this study allows us to explore the effects of the Rician K-factor or channel correlation on the performance of multi-antenna quickest spectrum sensing.

Chapter 6

Based on the results obtained from Chapters 4 and 5, in this chapter, we derive approximate expressions for the detection delay distribution for multi-antenna time-invariant CUSUM detector when the received signal experiences Rayleigh (independent and correlated) and Rician (independent and correlated) channels. In particular, we derive the expected value and variance of the received signal, which is transmitted over a correlated Rician channel, where these derivations are helpful in deriving the approximate expression for the distribution of detection delay The received signal considered here is a product of two complex Gaussian variables (the primary user signal and Rician channel) with additive noise. For all cases considered, the derivation of the approximate expression for the detection delay distribution are based on the theory of Brownian motion with drift. The validity of the approximate expressions for the distribution of detection delay for each case considered is verified via simulations. Furthermore, we also investigate the effects of multiple antennas at the cognitive user, channel correlation and line-of-sight (LOS) strength on the probability of detection delay, particularly for long detection delays or at low signal-to-noise ratios (SNRs).

Chapter 7

In this chapter, we summarize our novel contributions of the thesis and outline several possible area for future works.

1.6 PUBLICATIONS

The following are the list of journal articles and conference proceedings produced during the PhD study as well as an invited talk at an international conference.

Journal articles

1. **E. Hanafi**, P. A. Martin, P. J. Smith and A. J. Coulson, “Extension of quickest spectrum sensing to multiple antennas and Rayleigh channels,” *IEEE Commun. Lett.*, vol. 17, no. 4, pp. 625-628, Apr. 2013.

2. **E. Hanafi**, P. A. Martin, P. J. Smith and A. J. Coulson, “Quickest spectrum sensing with multiple antennas in Rician and correlated Rayleigh channels,” *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1455-1458, Aug. 2014.

3. **E. Hanafi**, P. A. Martin, P. J. Smith and A. J. Coulson, “On the distribution of detection delay for quickest spectrum sensing,” in preparation for *IEEE Trans. Commun.*

4. **E. Hanafi**, P. A. Martin, P. J. Smith and A. J. Coulson, “On the distribution of detection delay for multi-antenna quickest spectrum sensing,” in preparation for *IEEE Commun. Lett.*

Conference papers

1. **E. Hanafi**, P. A. Martin, P. J. Smith and A. J. Coulson, “Performance of quickest spectrum sensing over various fading channels,” *Proc. AusCTW*, pp. 69-74, Jan. 29 - Feb. 1, 2013. **(Awarded Best Student Poster Presentation)**

Invited talks

1. **E. Hanafi**, “On the distribution of detection delay for quickest spectrum sensing,” Invited talk at the Australian Communications Theory Workshop (AusCTW) 2014, Feb 2014, Macquarie University, Sydney, Australia

Chapter 2

BACKGROUND

This chapter provides the relevant background information required for the subsequent chapters. Section 2.1 presents the statistical wireless channel models and a characterization of the channels along with statistical measures to evaluate the severity of fading channels. The modelling of the wireless channel is important in cognitive radio systems as it affects the performance of the spectrum sensing techniques. In Section 2.2, a brief overview of multiple antenna wireless systems is given. The concept of diversity, different methods of achieving diversity and the main diversity combining techniques are also discussed. A classification of cognitive radio networks is presented in Section 2.3. In Section 2.4, a brief outline of quickest detection theory is given followed by a description of the cumulative sum (CUSUM) algorithm. Finally, Section 2.5 concludes this chapter with a short summary.

2.1 WIRELESS CHANNEL

The various paths between a transmitter and a receiver in a wireless communications environment characterize the wireless channel and these different paths (called multi-path) result in multiple versions of the transmitted signal being received at the receiver [4]. Figure 2.1 illustrates some different possible paths for the received signal. As shown in Figure 2.1, the three major radio wave propagation mechanisms in a wireless channel are reflection, diffraction and scattering [4]. These three propagation mechanisms act simultaneously on the transmitted signal traveling through the wireless channel and hence, the resulting received signal is a combination of reflected, diffracted and scattered signals from various obstacles along the propagation path [50]. The combined effect of these three propagation mechanisms leads to the received signal behaving like

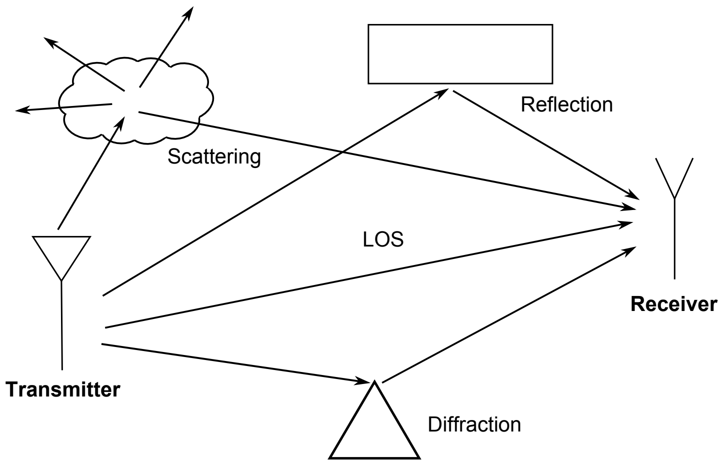

Figure 2.1 An example of different paths in a wireless channel [4].

a complex random process [50].

The communication path between transmitter and receiver is time and frequency variant [51,52]. The variations of the wireless channel can be classified into large-scale and small-scale fading [51]. Large-scale fading is caused by path loss (or propagation loss) and shadowing, where the variations of the received signal power due to these two phenomena occur over relatively large distances [51–55]. Path loss is due to the signal power attenuation in the propagation path when the distance between the transmitter and receiver is large [52]. Shadowing occurs when the receiver is shadowed by the presence of large objects (e.g. buildings, trees, bridges and hills) between the transmitter and receiver [51–53].

On the other hand, small-scale fading occurs over short distances, of the order of the carrier wavelength, and is due to the constructive and destructive addition of multiple copies of the transmitted signal traveling along multiple propagation paths before arriving at the receiver [51–55]. In this thesis, small-scale fading is considered.

We now describe small-scale fading channel models in more detail. All fading channels considered in this thesis are frequency-flat. If the transmitted signal occupies a bandwidth smaller than the coherence bandwidth of the channel, then the fading channel is referred to as a frequency-flat fading channel, or a frequency nonselective fading channel [5, 50, 53, 54, 56, 57]. The coherence bandwidth is the frequency range over which the channel passes all spectral components with approximately equal gain and linear phase [54]. In a flat-fading channel, all spectral components of the transmitted signal experience the same fading distortion [52, 53, 57] and the spectral characteristics of the transmitted signal are preserved at the receiver [50,54]. However, due to the variations of the channel gain caused by multipath, the received signal strength fluctuates in time [50, 54]. The flat-fading channel is also known as a narrowband channel since the bandwidth of the transmitted signal is narrow compared to the coherence bandwidth of the channel [50,54,57].

Let s denotes the transmitted signal. In general, the signal received at the output of a channel, which is affected by the channel and the additive white Gaussian noise (AWGN), can be expressed in the form

$$y = hs + n, \tag{2.1}$$

where h is the channel coefficient and n is the complex Gaussian noise [58]. If $h = 1$, then the channel is known as Gaussian or AWGN channel, which is the simplest wireless communications channel model [56]. In this thesis, we refer to the Gaussian channel as a time-invariant channel and both of these terms will be used interchangeably. For the Gaussian channel model, the received signal is given by

$$y = s + n, \tag{2.2}$$

where $n \sim \mathcal{CN}(0, N_0)$ is AWGN [5, 59]. This is illustrated in Figure 2.2. Here, the channel is characterized by the power spectral density of the noise component, denoted by

$$S_N(f) = \frac{N_0}{2} \quad W/Hz, \quad -\infty \leq f \leq \infty, \tag{2.3}$$

where N_0 is the average noise power per unit bandwidth measured at the front end of the receiver [55, 58, 60, 61].

In the following subsections, we will discuss various wireless fading channel models, including Rayleigh, Rician, Nakagami-m and F channel models. In this thesis, we normalize the amplitude channel distributions so that the mean power is unity, $E\left[|h|^2\right] = 1$, where $E[.]$ denotes statistical expectation.

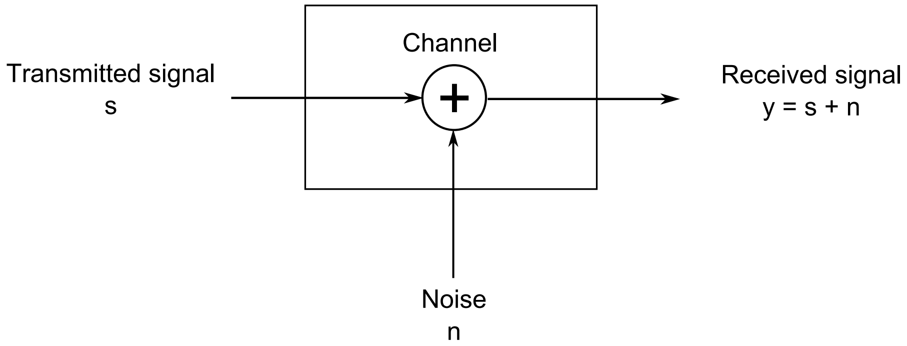

Figure 2.2 Model for the received signal passed through an AWGN channel [5].

2.1.1 Rayleigh fading

Multipath fading usually occurs when the transmitted signal arrives at the receiver via multiple paths due to the presence of reflecting objects and scatterers in the channel [54]. A combination of multiple delayed, reflected, scattered and diffracted versions of the transmitted signal are received [50, 54]. This results in constructive and destructive interference of the transmitted waves, which causes the signal to experience multipath fading [57].

The Rayleigh distribution is often used to model multipath fading when no direct line-of-sight (LOS) path exists between the transmitter and receiver, such as in heavily built-up urban environments [50, 54, 56, 56, 57]. With this model, the amplitude of fading channel, Y, is Rayleigh distributed with distribution, given by [53, 57]

$$f_Y(y) = 2ye^{-y^2}, \quad y \geq 0. \tag{2.4}$$

Here, $E\left[Y^2\right] = 1$, while the phase follows a uniform distribution defined over the interval $[0, 2\pi]$ [50]. As a result of (2.4), the channel fading power, $Z = Y^2$, is distributed according to an exponential distribution defined as [53]

$$f_Z(z) = e^{-z}, \quad z \geq 0. \tag{2.5}$$

2.1.2 Rician fading

The Rician fading channel is often used to model a multipath fading environment when there exists a LOS component along the propagation path together with other scattered components [50, 53, 54, 56]. In this case, the received signal is the superposition of the multipath components which arrive at different angles, and a stationary LOS component [50,53,54]. The channel fading amplitude for the Rician case, Y, follows the distribution given by [53, 57]

$$f_Y(y) = \frac{2y(K+1)}{\nu} e^{-K-\frac{(K+1)y^2}{\nu}} \mathrm{I}_0\left(2y\sqrt{\frac{K(K+1)}{\nu}}\right), \quad y \geq 0, \tag{2.6}$$

where K is the Rician K-factor, which measures the severity of the fading. The power distribution for the Rician fading model can be obtained by a change of variables, which yields

$$f_{Y^2}(z) = \frac{(K+1)}{\nu} e^{-\left(K+\frac{(K+1)z}{\nu}\right)} \mathrm{I}_0\left(2\sqrt{\frac{zK(K+1)}{\nu}}\right), \quad z \geq 0. \tag{2.7}$$

In this thesis, it is assumed that $\nu = E\left[Y^2\right] = 1$. Therefore, ν is normalized to unity. The Rician K-factor ranges from 0 to ∞. For $K = 0$, the Rician distribution reduces to the Rayleigh distribution whereas for $K = \infty$, the Rician distribution degenerates (collapsing to a fixed value), since the channel has no multipath and only consists of a LOS component [53, 56].

2.1.3 Nakagami-*m* fading

The Rayleigh and Rician models are both motivated by physical arguments [53] and as a result of their simplicity and physical motivation, they have become widely accepted. Nevertheless, several other simple channel models are also popular, although they do not have the same level of physical intuition. The Nakagami-m distribution is one such model as it has the ability to model a wider class of fading channel distributions and to fit a variety of empirical measurements [50,53]. For this model, the channel fading amplitude, Y, is distributed according to the Nakagami-m fading distribution, given by [53, 57]

$$f_Y(y) = \frac{2m^m y^{2m-1}}{\Gamma(m)\nu^m} e^{\frac{-my^2}{\nu}}, \quad m \geq 0.5, \quad y \geq 0, \tag{2.8}$$

where m is the Nakagami-m fading parameter, which determines the severity of the fading and $\Gamma(m)$ is the Gamma function defined as $\Gamma(m) = (m-1)!$ for an integer m. The distribution of the channel fading power can be obtained by a change of variables, which gives [53]

$$f_{Y^2}(z) = \left(\frac{m}{\nu}\right)^m \frac{z^{m-1}}{\Gamma(m)} e^{\frac{-mz}{\nu}}, \quad z \geq 0. \tag{2.9}$$

In this thesis, $\nu = E\left[Y^2\right]$ is assumed to be normalized to unity. The Nakagami-m fading parameter ranges from 1/2 to ∞. For $m = 1/2$, the Nakagami-m distribution reduces to the one-sided Gaussian distribution and for $m = 1$, it reduces to the Rayleigh distribution. The Nakagami-m distribution will be similar to a Rician distribution when $\frac{3}{2} \leq m \leq 3$. As $m \to +\infty$, the Nakagami-m distributed fading channel will converge to a nonfading AWGN channel. The Nakagami-m distribution can closely approximate the Rician distribution by a one-to-one mapping between the m parameter and the Rician K-factor parameter, where the mapping is given by [53]

$$m = \frac{(K+1)^2}{2K+1}, \qquad m \geq 1, K \geq 0. \tag{2.10}$$

2.1.4 F fading

Long-tailed distributions have been used in [62] as models for severe fading channels. In this thesis, we use the F-distribution, which is a long-tailed distribution, as a severe fading channel model in order to investigate the spectrum sensing performance under extreme fading conditions. The F channel, which is based on the F-distribution, is selected due to its simplicity as compared to other severe fading channel models such as the two-wave with diffuse power (TWDP) model of [63]. The probability density function of the F-distribution with ν_1 and ν_2 degrees of freedom is given by [64]

$$f_X(x) = \frac{(\nu_1/\nu_2)^{\nu_1/2}}{\mathrm{B}\left(\frac{1}{2}\nu_1, \frac{1}{2}\nu_2\right)} \frac{x^{(\nu_1/2)-1}}{\left(1+\nu_1\nu_2^{-1}x\right)^{(\nu_1+\nu_2)/2}}, \quad 0 < x, \tag{2.11}$$

where $\mathrm{B}(a,b)$ is the Beta function with parameters a and b. The mean and variance of the F-distribution are denoted respectively as [64]

$$E[X] = \frac{\nu_2}{\nu_2 - 2}, \quad \nu_2 > 2, \tag{2.12}$$

$$\mathrm{Var}(X) = \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}, \quad \nu_2 > 4, \tag{2.13}$$

where Var(.) denotes the variance. From [62], the F-distribution can be used to model the power of a severely fading channel. Throughout this thesis, all the classical fading channel models discussed in Sections 2.1.1 - 2.1.3 have channel powers normalized to unity and hence, for consistency, we set the power of the F channel to unity. One way to model a severely fading channel is with an F-distribution with $\nu_1 = 1$ and $\nu_2 = 3$. Here, the distribution is so long-tailed that the variance does not exist, hence providing an extreme case for spectrum sensing. The distribution of the channel fading power with 1 and 3 degrees of freedom can be written as

$$f_Z(z) = \frac{2}{\pi\sqrt{z}(1+z)^2}. \tag{2.14}$$

2.1.5 Fast and slow fading

The wireless channel can also be categorized as a fast or slow fading channel depending on how fast the transmitted signal changes as compared to the rate of change of the channel [54]. Thus, this classification is related to the channel coherence time, T_c, which is a statistical measure of the period of time over which the impulse response of the channel is invariant [54, 57]. The coherence time of the channel is also related to the channel Doppler spread, f_d, by [53, 54, 57]

$$T_c \simeq \frac{1}{f_d}. \tag{2.15}$$

A channel is referred to as a fast fading channel if the coherence time of the channel is smaller than the symbol period of the transmitted signal [50, 54, 57]. Therefore, the channel impulse response changes rapidly at approximately the symbol rate of the communication system [54].

On the other hand, the channel is classified as a slow fading channel if the symbol time duration of the transmitted signal is small as compared to the coherence time of the channel [50, 54, 57]. Therefore, in a slow fading channel, the channel stays unchanged for several symbols and a certain fade level will affect many consecutive symbols [50, 54, 57].

2.1.6 Coefficient of variation for fading channels

Coefficient of variation (CV) is a simple measure of the severity of the wireless fading channels [62, 65]. Generally, the CV is defined as the ratio of the standard deviation to the mean [66, 67]. The evaluation of fading severity using the CV can therefore be obtained from the moments of the channel fading distribution. Therefore, the CV can be defined in terms of the channel fading amplitude, Y as

$$\begin{aligned} \mathrm{CV} &= \frac{\sqrt{\mathrm{Var}(Y)}}{E[Y]} \\ &= \sqrt{\frac{1}{E\left[Y\right]^2} - 1}. \end{aligned} \tag{2.16}$$

Equation (2.16) follows since the fading channels are normalized to have unit power, $E\left[Y^2\right] = 1$.

Another measure that can be used to quantify the severity of the fading experienced by a particular channel model is the amount of fading (AF), introduced by Charash in [68], where AF is defined as [68, 69]

$$\begin{aligned} \mathrm{AF} &= \frac{\mathrm{Var}(Y)}{\left(E[Y]\right)^2} \\ &= \mathrm{CV}^2. \end{aligned} \tag{2.17}$$

Therefore, as can be seen from (2.17), AF is closely related to CV and hence, in this thesis, we will be using CV to measure the fading severity.

2.2 OVERVIEW OF THE MULTI-ANTENNA SYSTEM

In wireless communications, different antenna configurations are used in space-time systems. The most well-known and simplest wireless configuration is single-input single-output (SISO), which contains a single antenna at both transmitter and receiver [52, 70]. On the other hand, when multiple antennas are used at the receiver with a single transmitting antenna, the wireless system is known as single-input multiple-output (SIMO) [52, 70, 71]. In this thesis, we adopt both the SISO and SIMO systems into the cognitive radio network, which is illustrated in Figure

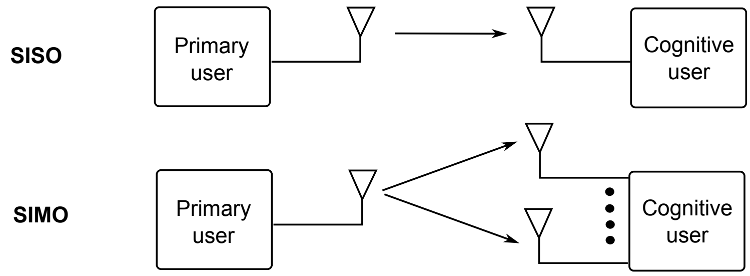

Figure 2.3 SISO and SIMO structures.

2.3, where single or multiple antennas may be employed at the cognitive user.

2.2.1 Diversity

Diversity can be achieved when the receiver receives multiple copies of the transmitted signal which travel through different wireless channels [52,53,71–73]. These multiple copies arrive at the receiver through different subchannels and experience different amplitudes and phases depending on the wireless channel [52, 74]. As a result of this, some copies of the transmitted signal may be deeply faded while others can be less attenuated [52]. Therefore, the probability that all copies of the transmitted signal simultaneously experience deep fades can be reduced significantly [53, 71, 72]. In addition, diversity can be used to mitigate the effect of multipath fading as well as improving both the instantaneous and average signal-to-noise ratio (SNR) at the receiver [53, 54, 72]. There are various ways of realizing diversity gain and the most common methods include temporal diversity, frequency diversity and spatial diversity [22, 53, 54, 70, 72, 73, 75].

In Chapters 5 and 6, the concept of spatial diversity will be used to investigate the performance of quickest spectrum sensing with multiple receive antennas at the cognitive user. However, for completeness, we discuss all the methods of achieving diversity in the following subsections.

2.2.1.1 Temporal diversity

Time diversity is one of the techniques that can be used to achieve multiple independently faded replicas of the transmitted signal. Time diversity involves transmission of the same transmit signal at different times, where the time difference is larger than the coherence time of the channel [53, 54, 72, 75]. This allows the multiple copies of the transmitted signal to undergo independent fading and as a result, the signals received at different time spacings are uncorrelated [22, 54, 72]. Although temporal diversity does not require high transmit power, it results in a lower data rate due to the repeated transmission of the same signal over time instead of transmission of new data [53]. In addition, one of the drawbacks of temporal diversity is that the transmission efficiency reduces based on the number of repeated signal transmissions [75]. Temporal diversity can be realized in various ways including repetition coding, which is the simplest technique, automatic repeat request (ARQ), interleaving and forward error correction (FEC) coding [22,53,72,73].

2.2.1.2 Frequency diversity

In frequency diversity, the same narrowband signal is transmitted simultaneously over multiple channels at different carrier frequencies, where the carrier frequencies are separated by more than the coherence bandwidth of the channel [22,53,54,72,75]. This ensures that the channels are uncorrelated and hence, the multiple versions of the transmitted signal will experience independent fading [54,70,72]. As a result, the probability that multiple copies of the signal are simultaneously in a deep fade is low [72]. However, in order to transmit the same signal over multiple frequency bands, the frequency diversity method requires additional transmit power [53]. Another shortcoming of frequency diversity is that the spectral efficiency of the system is greatly reduced due to repeated transmissions of the same signal at multiple frequencies [72, 75]. As a result, frequency diversity is rarely implemented in the traditional way [22, 72]. Instead, the information is spread over a wide frequency band and hence, different frequency components will carry small parts of the information [22, 72]. This spreading can be done by different techniques such as time division multiple access (TDMA), frequency hoping in conjunction with coding, code division multiple access (CDMA) and orthogonal frequency division multiplexing (OFDM) [22, 54, 72].

2.2.1.3 Spatial diversity

Spatial diversity, also known as antenna diversity, is the oldest and simplest diversity technique [54, 70, 72]. Spatial diversity can be achieved by transmitting the same signal to several antenna elements so that multiple copies of the transmitted signal are received at different antennas [70, 72]. In order to ensure that the fading amplitudes corresponding to each antenna are approximately independent, the antennas need to be spaced by at least half a wavelength [53,54,75]. Spatial diversity can be categorized into two types, namely transmit and receive diversity, depending on whether the diversity is realized at the transmitter or receiver [70]. In this thesis, as illustrated in Figure 2.3, multiple receive antennas are used to achieve spatial diversity and hence, this leads to receiver diversity. One of the advantages of receiver diversity is that multiple independent fading paths are realized without increasing the transmit signal power or bandwidth [53].

There are three main diversity combining techniques: selection combining, maximal ratio combining and equal gain combining [54, 70, 72, 76]. The combining techniques discussed in this section will focus on the combination of signals from different antennas at the receiver. It is assumed that the receiver is equipped with M sufficiently spaced antennas and hence, this leads to M independently fading signal paths.

2.2.1.3.1 Selection combining

Selection combining technique selects branch with the signal that has the highest instantaneous SNR among the M diversity branches for further processing [52,54,72,77]. Therefore, the output of the combiner is equal to the strongest incoming signal, assuming that the noise power at each branch is equally distributed [52, 53, 70, 77]. In practice, due to the difficulty in measuring the SNR, the branch with the highest total power in the received signal is selected [53, 54, 70].

Selection combining does not require any channel state information and hence, it can be used with either coherent or non-coherent signalling schemes [53,58]. However, since the receiver only exploits one of the M independent copies of the transmitted signal, the energy of the other $M-1$ replicas of the signal is wasted [72].

2.2.1.3.2 Maximal ratio combining

One of the techniques that employs a linear combination of the signals at the output of the M diversity branches is maximal ratio combining (MRC). Multiple copies of the same signal information are combined using the MRC method to achieve the maximum SNR at the combiner's output [52, 53, 58, 71, 72, 74]. In this method, the signal replicas at different diversity branches are weighted according to their individual SNRs and the resulting signals are then summed [53, 54, 70, 72]. The weights are chosen such that the output SNR is maximized, so that the weights are proportional to the branch SNRs [53, 72].

The MRC method produces an output SNR equal to the sum of the SNRs at the individual diversity branches [53, 54, 70–72]. Therefore, MRC has the advantage of achieving a reasonable SNR output even if the individual signals at each diversity branch are in a deep fade [54, 70]. However, MRC requires perfect knowledge of the channel state information at the receiver. This corresponds to obtaining the complex channel gains of each diversity branch, in order to estimate the phase of the received signal [52,58,70,71,77]. As a result, MRC can only be used for coherent detection [58, 78].

2.2.1.3.3 Equal gain combining

Equal gain combining (EGC) is the simplest linear combining technique [79]. In EGC, the signal replicas of the transmitted signals on different diversity branches are combined with equal weighting [53, 54, 70]. In particular, the gain or amplitude of the complex weights are all set to unity [54,80]. Since multiple signals are combined simultaneously from all the diversity branches, the combined signal at the output of the combiner can still have a reasonable SNR even if the the branch SNRs are too low for signal detection [54]. The advantage of EGC over MRC lies in its low implementation complexity as it does not require knowledge of the channel fading amplitudes for signal combining [53, 70, 80, 81]. Also, EGC performs better than selection combining and is often similar in performance to MRC [53, 54, 79]. In addition, as for selection combining, EGC does not require any channel state information (CSI) [58]. In this thesis, we employ EGC as it allows us to construct the correct CUSUM detector, which will be shown later in Chapter 5.

2.3 COGNITIVE RADIO NETWORK PARADIGMS

Cognitive radio networks can be categorized into three classes: underlay, overlay and interweave [30, 82–84]. The success of the cognitive radio paradigms in providing a communication framework for the interoperability between primary and cognitive user networks depends several factors. These include the regulatory constraints on spectrum usage, interference management mechanism as well as the nature of the available side information [30, 82, 84]. The side information available may include knowledge regarding the activity, channel, encoding strategies (i.e. codebook) and/or transmitted data sequences of the primary users (i.e. primary user messages) with which the cognitive user shares the spectrum. This information can be exploited by the cognitive user to improve the spectrum utilization [30, 82]. In this section, a description of each of the cognitive radio paradigms will be given.

2.3.1 Underlay paradigm

In an underlay paradigm, simultaneous transmissions of the primary and cognitive users is allowed as long as the interference generated by the cognitive user is kept well below some acceptable threshold [21,30,34,82–88]. In order to ensure that the amount of interference caused at the primary user is within tolerable limits, multiple antennas can be used for interference nulling (i.e. using beamforming techniques [16]), as long as certain quality of service (QoS) are met [82, 84]. However, the use of beamforming at the cognitive user requires knowledge of the primary user channel in order to perform beamsteering [16]. Alternatively, the signals transmitted by the cognitive user can be spread over a large bandwidth such that the interference generated is below the noise floor [30, 82, 84, 86]. These spread signals can then be despread at the cognitive receiver [30, 82]. This spreading technique is the basis of both spread spectrum and ultra-wide-band (UWB) communication systems [30, 34, 82, 84, 88].

The main advantage of this paradigm is that the cognitive user does not need to monitor the primary user activity [88]. However, the underlay paradigm is usually restricted to short-range communications due to the strict transmission power limitations [21, 30, 82, 88]. Another drawback of the underlay paradigm is that information concerning the fading channel gains between the cognitive transmitter and primary receiver is required by the cognitive user in order to perform the interference management task [84].

2.3.2 Overlay paradigm

The overlay paradigm is similar to the underlay paradigm in that it permits concurrent primary and cognitive user transmissions [21, 83, 86–88]. In this paradigm, the cognitive user is assumed to have sufficient knowledge of the channel gains, primary user codebooks and messages to either cancel or mitigate the interference at the cognitive and primary receivers [30, 82–88]. In particular, based on such information, the cognitive user can employ sophisticated adaptive signal processing techniques such as dirty paper coding (DPC) or successive interference cancellation (SIC) in order to remove the interference caused by the primary user's signal at the cognitive receiver [30,82,84,87,88]. In addition, knowledge of the primary user codebook and/or messages can be utilized by the cognitive radio to split its transmission power between its own transmission and relaying the primary user signal [30, 82–84, 86–88].

In order for the cognitive user to manage interference at the primary receiver, the choice of the transmission power split needs to be chosen carefully such that the primary user signal-to-interference-plus-noise ratio (SINR) remains the same regardless of the presence of the cognitive user transmission [84, 88]. Therefore, a sophisticated power control mechanism is needed to determine the allocation of the transmission power [88]. This paradigm also requires cooperation between the primary and cognitive users [21, 30].

2.3.3 Interweave paradigm

The interweave paradigm is the most popular paradigm adopted in cognitive radio research [88]. It is the original motivation for cognitive radio and is based on the idea of opportunistic communication [17], where there exists a transmission opportunity when the primary user is not transmitting [30, 82, 83, 86, 88]. In other words, there exists a portion of the spectrum which is unoccupied by the primary user, referred to as a spectrum hole or white space [16, 30, 88]. This spectrum hole is empty (i.e. free of any interfering signals) except for noise due to natural and/or artificial sources and hence, the cognitive user could utilize this vacant spectrum [16, 30, 88].

In this paradigm, the cognitive user monitors the primary user's frequency band, detects the occupancy of the spectrum (i.e. the existence or absence of the primary user) and opportunistically communicates over the spectrum holes with minimal interference to the primary user [30, 82, 84–87, 87]. Therefore, due to the opportunistic reuse of the spectrum holes, the utilization of the spectrum is enhanced [30, 82, 82, 83]. In the interweave paradigm, an accurate spectrum sensing mechanism is required to determine the spectrum holes so that the cognitive user does not cause any harmful interference to the primary user [86, 88].

Table 2.1 summarizes the differences among the underlay, overlay and interweave cognitive radio paradigms. In some of the literature, including [2, 7, 16, 34, 89], cognitive radio is classified into only underlay or overlay, where the overlay paradigm in this case refers to the interweave paradigm [86, 90]. In this thesis, a cognitive radio network model based on the interweave paradigm is presented.

Table 2.1 Comparison of underlay, overlay and interweave cognitive radio paradigms

Type	Simultaneous transmission	Network side information
Underlay	Cognitive user can transmits simultaneously with primary user as long as interference caused is below an acceptable limit.	Cognitive user knows information on fading channel gains between cognitive transmitter and primary receiver.
Overlay	Permits concurrent primary and cognitive user transmission; interference due to cognitive user transmission at primary receiver can be offset by using part of cognitive user's power to relay the primary user signal.	Cognitive user knows information on channel gains, primary user codebooks and messages.
Interweave	Cognitive user transmits simultaneously with a primary user only in the event of false spectral hole detection.	Cognitive user identifies spectrum holes in space, time and/or frequency when primary user is absent.

2.4 OVERVIEW OF QUICKEST DETECTION THEORY

Recall from Section 1.3 that the quickest detection problem is the problem of detecting an abrupt change in the observation distribution as quickly as possible. This can be illustrated by the following example. Let $Y[1], Y[2], \ldots$ be a sequence of independent random observations, where i is the sample number. Initially, the sequence of observations, $Y[1], Y[2]..., Y[\tau-1]$, are i.i.d following a common distribution, $F^{(0)}$. At an unknown sample number, τ, referred to as a change point, the subsequent observations $Y[\tau], Y[\tau+1], \ldots$ are i.i.d with another distribution, $F^{(1)}$. Let $f^{(0)}_{Y[i]}$ and $f^{(1)}_{Y[i]}$ denote the probability density functions corresponding to $F^{(0)}$ and $F^{(1)}$, respectively. Therefore, there is a change in the distribution of the observations at τ and the objective is to detect such changes with minimal detection delay subject to a certain false alarm constraint.

The two standard formulations in the quickest detection literature are Bayesian and non-Bayesian (or known as minimax formulations) [30,43,46,91,92]. The Bayesian quickest detection formulation was proposed by Shiryaev [93], where the change point is assumed to be a random variable with a known prior distribution, in particular a geometric distribution [30,43,45,46,91,92]. The objective of this formulation is to minimize the expected detection delay subject to an upper bound on the probability of false alarm [46,46,92]. On the other hand, the minimax formulation, which was proposed by Lorden [94], models the change point as an unknown deterministic quantity [30, 46, 91]. It was shown in [94] that the well-known Page's CUSUM algorithm [95] is asymptotically[1] optimal in the sense of minimax (i.e. it minimizes the worst-case detection delay while maintaining a certain level of false alarm) [30, 37, 43, 46].

It is worth noting that the quickest detection problem is different to but has a close relationship with the classical sequential detection problem, namely the sequential probability ratio test (SPRT), formulated by Wald [96]. Sequential detection is based on the theory of solving a hypothesis-testing problem, where the objective is to distinguish between two hypotheses from a statistically homogeneous sequence of i.i.d random observations as quickly as possible given a specified level of detection error [30, 36, 43, 43, 44]. Therefore, all the samples are drawn from an identical distribution [36, 43]. However, in the quickest detection problem, the sequence of

[1] Asymptotic here means that the mean number of samples between false alarms goes to infinity.

observations is not homogenous and the aim of the quickest detection is to detect the occurrence of an inhomogeneity in the random observations [36]. The connection between the quickest and sequential detection problem lies in the study of the CUSUM algorithm, where the algorithm can be regarded as a repeated one-sided SPRT according to the renewal property of the stopping rule [97].

As mentioned in Section 1.3, we apply quickest detection theory to spectrum sensing in a cognitive radio system to detect any changes in spectrum occupancy and this application is referred to as quickest spectrum sensing. In this thesis, we adopt the CUSUM algorithm as a detection strategy for quickest spectrum sensing.

2.4.1 CUSUM algorithm

This section describes an application of the CUSUM algorithm to quickest spectrum sensing, where we apply the general problem discussed in Section 2.4. Let the primary user be initially inactive and at an unknown sample number, τ, the primary user commences transmission. The cognitive user observes samples sequentially and employs the CUSUM algorithm to detect the primary user via a change in the distribution of the received signal. The CUSUM algorithm detects the abrupt change (i.e. the emergence of the primary user) at sample

$$T = \inf(n : C_n \geq \gamma), \tag{2.18}$$

where γ is a threshold and C_n is the CUSUM statistic defined as [36]

$$C_n = \max_{k \leq n} \sum_{i=k+1}^{n} l_{Y[i]}(y[i]). \tag{2.19}$$

$l_{Y[i]}(y[i])$ in (2.19) is the log likelihood ratio denoted by

$$l_{Y[i]}(y[i]) = \ln \left\{ \frac{f^{(1)}_{Y[i]}(y[i])}{f^{(0)}_{Y[i]}(y[i])} \right\}, \tag{2.20}$$

and $f^{(0)}_{Y[i]}(y[i])$ and $f^{(1)}_{Y[i]}(y[i])$ are the probability density functions of the received signal, $Y[i]$, when the primary user is absent and present, respectively. The CUSUM statistic, C_n can also be expressed in a recursive form given by [36, 41–43]

$$
\begin{aligned}
C_{n+1} &= \max_{k\leq n+1}\left\{\sum_{i=k+1}^{n+1} l_{Y[i]}(y[i])\right\} \\
&= \max\left\{\max_{k\leq n}\left\{\sum_{i=k+1}^{n+1} l_{Y[i]}(y[i])\right\}, 0\right\} \\
&= \max\left\{\max_{k\leq n}\left\{\sum_{i=k+1}^{n} l_{Y[i]}(y[i])\right\} + l_{Y[n+1]}(y[n+1]), 0\right\} \\
&= \{C_n + l_{Y[n+1]}(y[n+1])\}^{+},
\end{aligned}
\tag{2.21}
$$

where $x^{+} = \max(x, 0)$. Therefore, the CUSUM statistic can be computed recursively for $n \geq 0$ by setting $C_0 = 0$. The C_n statistic is compared to a threshold, γ, after each sample and the algorithm will raise an alarm when $C_n \geq \gamma$, which indicates the existence of the primary user.

If $T > \tau$, a detection delay, $\delta = T - \tau$ will occur. On the other hand, if $T < \tau$, a false alarm event will occur, where the mean number of samples between false alarms is denoted by [36, 37, 41, 43]

$$\overline{T}_f = E_{f^{(0)}_{Y[i]}}[T], \tag{2.22}$$

where $E_{f^{(0)}_{Y[i]}}[.]$ denotes the expectation operator when there is no change in the distribution of the observations. The false alarm rate (FAR) is defined as [98]

$$\text{FAR}(T) = \frac{1}{E_{f^{(0)}_{Y[i]}}[T]}. \tag{2.23}$$

Recall from Section 2.4 that the CUSUM algorithm is asymptotically minimax optimal with respect to Lorden's measure of the worst-case detection delay. Based on Lorden's formulation, the worst-case detection delay is denoted by [36, 37, 43, 98]

$$\overline{T}_d = \sup_{\tau\geq 1} \text{ess sup } E_{f^{(1)}_{Y[i]}}\left[\delta = T - \tau | T \geq \tau, Y[1], ..., Y[\tau]\right], \tag{2.24}$$

where $E_{f^{(1)}_{Y[i]}}$ denotes the expectation under the distribution of $F^{(1)}$ with the corresponding pdf

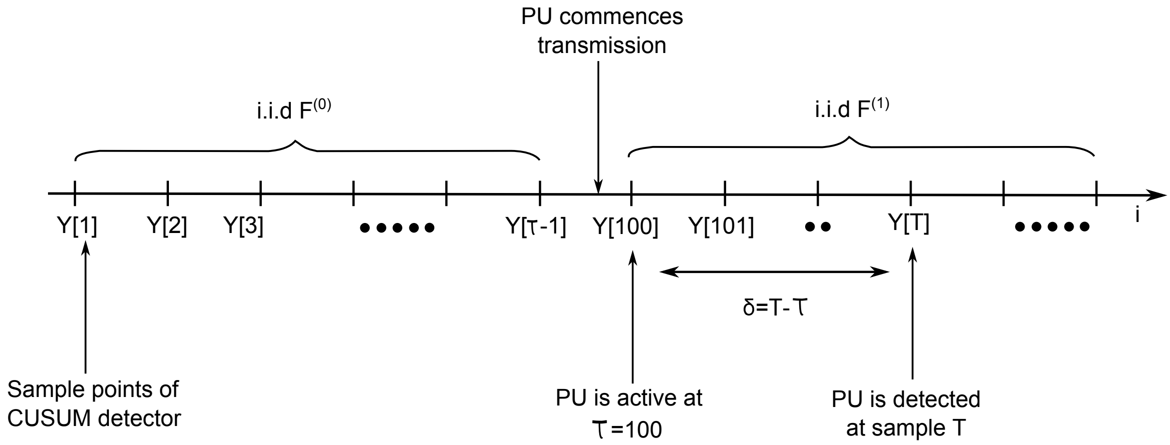

Figure 2.4 Illustration of the primary user (PU) detection based on the CUSUM algorithm, employed at the cognitive user.

of $f^{(1)}_{Y[i]}$. It is worth noting that ess sup is used in (2.24) so that $\overline{T}_d$ takes the worst-case value of the expected detection delay over all possible realizations of the observations of the received signals, Y, before the change occurs. The threshold, γ in (2.18) can be set based on the lower bound of $\overline{T}_f$, where the bound can be expressed as [36, 44].

$$\overline{T}_f \geq e^{\gamma}. \tag{2.25}$$

Alternatively, the threshold values can be set in an arbitrary way to give a desired range of average detection delay or false alarm rate. Based on the pre-determined threshold values, the average detection delay and false alarm rate can be measured.

Figure 2.4 illustrates the primary user detection based on the CUSUM algorithm, employed at the cognitive user. Here, it is assumed that the primary user is active at $\tau = 100$. Based on Figure 2.4, we can see that if the cognitive user detects the primary user in the first sample when the primary user is active (i.e at $T = 100$), then, $\delta = 0$. However, $\delta = 1$ occurs when the cognitive user detects the primary user using two samples (i.e at $T = 101$).

2.5 CHAPTER SUMMARY

This chapter has provided an overview of several concepts needed for the subsequent chapters in this thesis. Firstly, statistical models for wireless channels, a characterization of the wireless channel as well as statistical measures of fading severity were introduced. Secondly, a brief overview of multi-antenna wireless systems, a general concept of diversity along with various approaches of realizing diversity gain including several spatial diversity combining techniques were discussed. Furthermore, a classification of cognitive radio networks was given. Finally, an overview of quickest detection theory and the CUSUM algorithm were presented.

Chapter 3

PERFORMANCE OF QUICKEST SPECTRUM SENSING OVER VARIOUS FADING CHANNELS

3.1 INTRODUCTION

In this chapter, we employ a SISO system model, where both primary and cognitive users are equipped with a single receive antenna. Existing quickest spectrum sensing studies [36, 37, 42] usually assume that the received signal has a real, Gaussian distribution. Here, we assume that the received signal consists of a Gaussian complex signal in the presence of multiplicative fading and additive noise. For this general model, we prove that the power of the complex received signal is a sufficient statistic. This result enables us to derive a general form for the log likelihood ratio and the quickest spectrum sensing technique for a range of channel models.

Although there have been a number of studies evaluating the sensing performance of an energy detector over different fading channels [99, 100], no studies on quickest spectrum sensing have appeared which consider a range of fading channels. Thus, in this chapter, we investigate the performance of quickest spectrum sensing over the time-invariant channel and several fading channels, including Rayleigh, Rician, Nakagami-m and long-tailed channel models. As discussed in Section 2.1.4, we employ the F-distribution, which is a long-tailed distribution, as a severe fading channel model. We derive the probability density function (pdf) of the amplitude of the received signal for all the fading cases that we consider, where these pdfs are required to form the CUSUM detector. In deriving the pdfs for the most commonly used channel models, we employ a technique which avoids numerical integration. This approach is effective, particularly in the Rician case, where simple quadrature methods are found to be unstable. In addition, this technique provides much faster computation compared to numerical integration. These advantages will be discussed in detail in Section 3.4.

Mis-matched channel conditions are also considered, in which we evaluate the performance of different CUSUM detectors, where the each of the CUSUM detectors are designed for a specific channel, but the true channel is different. This study is useful for gaining further insights into the effects of the channel on sensing performance and the robustness of the detectors. In addition, we also consider the case when the fading channels are temporally correlated. This allows us to gain some insight into the quickest spectrum sensing performance in the presence of dependent observations at the cognitive user. Parts of this chapter have been published in [101].

The remainder of this chapter is organized as follows. In Section 3.2, the system model is described. Section 3.3 presents a proof that the power of the complex received signal is a sufficient statistic. Derivations of the pdfs of the received signal amplitude and the log likelihood ratios for the time-invariant as well as the fading channel scenarios that we considered are given in Section 3.4. Numerical results are provided in Section 3.5 for the performance of quickest spectrum sensing in the classical channel models, a severe fading model, mis-matched channels and temporally correlated channel conditions. Finally, Section 3.6 ends the chapter with some concluding remarks.

3.2 SYSTEM MODEL

An interweave cognitive radio network is considered, where a PU is initially inactive and a CU attempts to detect the presence of the PU's signal after it begins transmission. Once the PU is detected, the CU needs to vacate the channel. A small detection delay is crucial in this case so that the CU vacates the channel quickly. We focus on detection of the entrance of the PU to the licensed channel. The detection of the departure of a PU can be approached similarly.

Let $Y[i]$ denote the CU observation at sample number i. It is assumed that the CU observes samples sequentially. If the PU is not transmitting, then $Y[i] = N[i]$, where $N[i]$ is the noise. If the PU is active, $Y[i] = H[i] \times S[i] + N[i]$, where $H[i]$ is the channel coefficient and $S[i]$ is the PU's signal, which is assumed to be a narrowband complex Gaussian signal. We assume that $S[i]$ and $N[i]$ are independent circularly symmetric complex Gaussian (CSCG) random variables such that $S[i] \sim \mathcal{CN}(0,\sigma_S^2)$ and $N[i] \sim \mathcal{CN}(0,\sigma_N^2)$. The channel variance is defined as $E[|H[i]|^2] = \sigma_H^2$ so that the observed SNR at the CU is $\sigma_H^2\sigma_S^2/\sigma_N^2$. All channel models considered are assumed to be independent and identically distributed (i.i.d) between samples, i, which corresponds to fast fading channels. In Section 3.5.1, we consider temporally correlated fading channels with more realistic fade rates.

3.3 PROOF THAT $|Y[i]|^2$ IS A SUFFICIENT STATISTIC

Before deriving the pdfs of the received signal amplitude and the log likelihood ratio for each channel models, we first need to prove that the power of the complex received signal, $|Y[i]|^2$, is a sufficient statistic. Sufficient statistic is a function of the data such that the log likelihood ratio is only dependent on that function. In the presence of the PU signal, the conditional distribution of the received signal, $Y[i]$, given $H[i]$ is $\mathcal{CN}(0,|H[i]|^2\sigma_S^2 + \sigma_N^2)$. Hence, the pdf of the received signal can be written as

$$\begin{aligned} f_{Y[i]}^{(1)}(y[i]) &= \int f_{Y[i]|H[i]}^{(1)}(y[i]) f_{H[i]}(h[i])\, dh[i] \\ &= E\left[f_{Y[i]|H[i]}^{(1)}(y[i]) \right], \end{aligned} \tag{3.1}$$

where

$$f_{Y[i]|H[i]}^{(1)}(y[i]) = \frac{e^{\frac{-|y[i]|^2}{(|H[i]|^2\sigma_S^2+\sigma_N^2)}}}{\pi(|H[i]|^2\sigma_S^2 + \sigma_N^2)}. \tag{3.2}$$

Let $|H[i]|^2 = \sigma_H^2 Z[i]$, where $Z[i]$ is a normalized random variable that is unique for each channel model. The particular models for $Z[i]$ are provided in Section 3.4. Using the above notation along with $\sigma_T^2 = \sigma_H^2\sigma_S^2$, (3.2) can be written as

$$f_{Y[i]|H[i]}^{(1)}(y[i]) = \frac{e^{\frac{-|y[i]|^2}{(\sigma_T^2 Z[i]+\sigma_N^2)}}}{\pi(\sigma_T^2 Z[i] + \sigma_N^2)}. \tag{3.3}$$

Therefore, the expectation in (3.1) can be expressed as

$$f_{Y[i]}^{(1)}(y[i]) = \int_0^\infty \frac{e^{\frac{-|y[i]|^2}{(\sigma_T^2 z[i]+\sigma_N^2)}}}{\pi(\sigma_T^2 z[i] + \sigma_N^2)} f_{Z[i]}(z[i])\, dz[i], \tag{3.4}$$

and the log likelihood ratio used in the CUSUM algorithm (discussed in Section 2.4.1), follows as

$$l_{Y[i]}(y[i]) = \ln\left\{\sigma_N^2 e^{\frac{|y[i]|^2}{\sigma_N^2}} \int_0^\infty \frac{e^{\frac{-|y[i]|^2}{(\sigma_T^2 z[i]+\sigma_N^2)}}}{(\sigma_T^2 z[i]+\sigma_N^2)} f_{Z[i]}(z[i])\, dz[i]\right\}. \tag{3.5}$$

Since $l_{Y[i]}(y[i])$ in (3.5) is solely a function of $|y[i]|^2$, it follows that $|Y[i]|^2$ is a sufficient statistic. From [102], it follows that the log likelihood ratio is unchanged whether it is computed using the original complex signal, $Y[i]$, the amplitude of the observed signal, $|Y[i]|$ or the power of the observed signal, $|Y[i]|^2$. This holds true, irrespective of the channel model. Hence, the log likelihood ratio for various channel models can be computed based on the amplitude of the received signals, which will be shown in Section 3.4.

3.4 PDFS AND LOG LIKELIHOOD RATIOS FOR DIFFERENT CHANNEL MODELS

As shown in Section 3.3, the log likelihood ratio can be evaluated based on the statistics of $|Y[i]|$. Therefore, for the rest of the chapter, all the analysis will be based on $|Y[i]|$. In order to derive the pdf of the amplitude of the received signal when the PU is inactive, we first derive its cumulative distribution function (cdf). The received signal in the absence of the PU can be expressed as

$$Y[i] = (\sigma_N^2)^{1/2} J[i], \tag{3.6}$$

where $J[i] \sim \mathcal{CN}(0,1)$. Equation (3.6) can then be rewritten as

$$Y[i]^\dagger Y[i] = \sigma_N^2 J[i]^\dagger J[i]. \tag{3.7}$$

Let $V[i] = J[i]^\dagger J[i]$, where $V[i]$ is an exponential random variable such that $V[i] \sim \text{Exp}(1)$. With this notation, the cdf can be expressed as

$$\begin{aligned} P(|Y[i]| < k) &= P(|Y[i]|^2 < k^2) \\ &= P(\sigma_N^2 V[i] < k^2) \\ &= P(V[i] < \frac{k^2}{\sigma_N^2}) \\ &= 1 - e^{-\frac{k^2}{\sigma_N^2}}. \end{aligned} \tag{3.8}$$

The pdf of the amplitude of the received signal when the PU is absent can be obtained by taking the derivative of (3.8) to yield

$$f^{(0)}_{|Y[i]|}(|y[i]|) = \frac{2|y[i]|}{\sigma_N^2} e^{\frac{-|y[i]|^2}{\sigma_N^2}}. \tag{3.9}$$

When the PU starts transmitting, the pdf of the amplitude of the received signal, $f^{(1)}_{|Y[i]|}(|y[i]|)$, for a range of common channel models, can be expressed as

$$f^{(1)}_{|Y[i]|}(|y[i]|) = \int_0^\infty \frac{2|y[i]| e^{\frac{-|y[i]|^2}{(\sigma_T^2 z[i] + \sigma_N^2)}}}{(\sigma_T^2 z[i] + \sigma_N^2)} f_{Z[i]}(z[i])\, dz[i], \tag{3.10}$$

using the same development as in (3.1)-(3.4). Hence, the log likelihood ratio, $l_{Y[i]}(y[i])$, for various channels can be derived using (3.9) and (3.10), which yields the same results as (3.5). The only term in (3.10) which is not fixed is the pdf of $Z[i]$ which depends on the channel model used. $Z[i]$ is a standard exponential random variable with parameter $\lambda = 1$ for a Rayleigh channel [53]. For a Rician channel, $Z[i]$ is a non-central chi-square variable with 2 degrees of freedom. For a Nakagami-m channel, $Z[i]$ is a gamma distributed random variable, $Z[i] \sim \mathcal{G}(m, 1/m)$, in which m is the Nakagami-m fading parameter [72]. For an F channel, $Z[i]$ is an F distributed random variable with 1 and 3 degrees of freedom, $Z[i] \sim \text{F}(1, 3)$. Since $Z[i]$ is the normalized channel power, $E[Z[i]] = 1$. All the channel models have well-known pdfs for $Z[i]$, $f_{Z[i]}(z[i])$, which are tabulated in Table 3.1.

Table 3.1 Pdf of $Z[i]$ for all types of channel model

Channel model	Pdf for $Z[i]$, $f_{Z[i]}(z[i])$
Time-invariant	$\delta(1)$
Rayleigh	$e^{-z[i]}$
Rician	$(K+1)e^{-(K+(K+1)z[i])}\mathrm{I}_0(2\sqrt{z[i]K(K+1)})$
Nakagami-m	$\frac{m^m}{\Gamma(m)}z[i]^{m-1}e^{-mz[i]}$
$\mathrm{F}_{1,3}$	$\frac{2}{\pi\sqrt{z[i]}(1+z[i])^2}$

In Table 3.1, $\delta(1)$ is the Dirac delta function and $\Gamma(m)$ is the Gamma function defined as $\Gamma(m) = \int_0^\infty e^{-t}t^{m-1}\,dt$ for $\mathbb{R}(m) > 0$ [103]. $\Gamma(m)$ can also be written in a simplified form in the case of an integer m as $\Gamma(m) = (m-1)!$ [103]. Also shown in Table 3.1 is the scaled F distribution, $\mathrm{F}_{1,3}$, normalized to have unit mean. As discussed in Section 2.1.4, the F distribution considered in this thesis is by no means a standard channel model. It is long-tailed [64] and has $E[Z[i]] = 1$ and variance, $\mathrm{Var}(Z[i]) = \infty$. Previous work [62] considered various long-tailed distributions as models for severe fading channels. In this thesis, the F distribution is considered as an extreme case. The standard channel models all have pdfs for $Z[i]$ which decay exponentially fast. In contrast, the F distribution, $\mathrm{F}_{1,3}$, has a very slow decay rate and as a result, the variance does not exist.

In the following subsections, we derive the pdfs of the amplitude of the received signal in the presence of the PU and the log likelihood ratios for the various channel models that we consider. The derivation of the pdfs for the most commonly used channels, including Rayleigh, Rician and Nakagami-m channels, uses an approach which avoids numerical integration. The benefits of this approach will be discussed in Sections 3.4.2 and 3.4.3.

3.4.1 Time-invariant channel

In the case when the received signal is transmitted over a time-invariant channel, $X[i] = H \times S[i]$ is a circularly symmetric complex Gaussian variable with variance, σ_X^2, where for the time-invariant channel, $H = 1$. The pdf of the amplitude of the received signal when the PU is present is given by

$$f_G^{(1)}(|y[i]|) = \frac{2|y[i]|e^{\frac{-|y[i]|^2}{(\sigma_N^2+\sigma_X^2)}}}{(\sigma_N^2+\sigma_X^2)}, \tag{3.11}$$

where the subscript, G, in (3.11) denotes the time-invariant or Gaussian channel. It is worth noting that in later subsections, subscripts will be used to differentiate between various channel models considered. The log likelihood ratio can then be written using (3.11) and (3.9) to yield

$$\begin{aligned} l_G(|y[i]|) &= \ln\left\{2|y[i]|e^{\frac{-|y[i]|^2}{(\sigma_N^2+\sigma_X^2)}}\right\} - \ln\left\{\sigma_N^2+\sigma_X^2\right\} - \ln\left\{2|y[i]|e^{\frac{-|y[i]|^2}{\sigma_N^2}}\right\} + \ln\left\{\sigma_N^2\right\} \\ &= \frac{|y[i]|^2\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)} + \ln\left\{\frac{\sigma_N^2}{\sigma_N^2+\sigma_X^2}\right\}. \end{aligned} \tag{3.12}$$

3.4.2 Rayleigh channel

The pdf of the amplitude of the received signal transmitted over a Rayleigh fading channel can be derived using (3.10) and Table 3.1 which gives

$$f_{Ray}^{(1)}(|y[i]|) = \int_0^\infty \frac{2|y[i]|e^{\frac{-|y[i]|^2}{(\sigma_T^2 z[i]+\sigma_N^2)}}}{(\sigma_T^2 z[i]+\sigma_N^2)} e^{-z[i]}\, dz[i]. \tag{3.13}$$

Let $t = \sigma_T^2 z[i] + \sigma_N^2$, then (3.13) becomes

$$f_{Ray}^{(1)}(|y[i]|) = \frac{2|y[i]|e^{\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2} \int_{\sigma_N^2}^\infty \frac{e^{-\frac{|y[i]|^2}{t}-\frac{t}{\sigma_T^2}}}{t}\, dt. \tag{3.14}$$

The numerical integration over an infinite region can be avoided by rewriting (3.14) as the difference between the integral from 0 to ∞ and the definite integral with interval $[0, \sigma_N^2]$, which gives

$$f_{Ray}^{(1)}(|y[i]|) = \frac{2|y[i]|e^{\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2}\left[\int_0^\infty \frac{e^{-\frac{|y[i]|^2}{t}-\frac{t}{\sigma_T^2}}}{t}\, dt - \int_0^{\sigma_N^2} \frac{e^{-\frac{|y[i]|^2}{t}-\frac{t}{\sigma_T^2}}}{t}\, dt\right]. \tag{3.15}$$

With the aid of (3.471.9) in [103, p. 363] and the midpoint rule of the Riemann sum approximation [104], (3.15) can be expressed as

$$f_{Ray}^{(1)}(|y[i]|) = \frac{2|y[i]|e^{\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2}\left[2\mathrm{K}_0\left(\frac{|y[i]|}{\sigma_T/2}\right) - \frac{\sigma_N^2}{R}\sum_{r=1}^{R}\frac{e^{-\frac{|y[i]|^2}{s_r}-\frac{s_r}{\sigma_T^2}}}{s_r}\right], \tag{3.16}$$

in which

$$s_r = \left(r - \frac{1}{2}\right)\left(\frac{\sigma_N^2}{R}\right), \tag{3.17}$$

where R is the number of rectangles in the Riemann sum used to approximate the integral. Numerical tests show that with $R = 50$, evaluation of this expression is much faster than numerical integration and the error is negligible. The log likelihood ratio can therefore be written using (3.16) and (3.9) to give

$$l_{Ray}(|y[i]|) = \ln\left\{\frac{\sigma_N^2}{\sigma_T^2} e^{\frac{\sigma_N^2}{\sigma_T^2} + \frac{|y[i]|^2}{\sigma_N^2}} \left[2\mathrm{K}_0\left(\frac{|y[i]|}{\sigma_T/2}\right) - \frac{\sigma_N^2}{R}\sum_{r=1}^{R}\frac{e^{-\frac{|y[i]|^2}{s_r} - \frac{s_r}{\sigma_T^2}}}{s_r}\right]\right\}. \tag{3.18}$$

3.4.3 Rician channel

In the case when the received signal experiences a Rician fading channel, the pdf of the amplitude of the received signal can be written using (3.10) and Table 3.1 to yield

$$f_{Ric}^{(1)}(|y[i]|) = \int_0^\infty \frac{2|y[i]|e^{\frac{-|y[i]|^2}{(\sigma_T^2 z[i] + \sigma_N^2)}}}{(\sigma_T^2 z[i] + \sigma_N^2)}(K+1)e^{-(K+(K+1)z[i])}\mathrm{I}_0\left(2\sqrt{z[i]K\,(K+1)}\right)\,dz[i]. \tag{3.19}$$

Letting $t = \sigma_T^2 z[i] + \sigma_N^2$, (3.19) can be rewritten as

$$f_{Ric}^{(1)}(|y[i]|) = \frac{2|y[i]|(K+1)e^{-K+\frac{\sigma_N^2}{\sigma_T^2}(K+1)}}{\sigma_T^2} \times \mathcal{I}, \tag{3.20}$$

where

$$\mathcal{I} = \int_{\sigma_N^2}^\infty \frac{e^{\frac{-|y[i]|^2}{t} - \frac{t}{\sigma_T^2}(K+1)}}{t}\mathrm{I}_0\left(2\sqrt{\left(\frac{t-\sigma_N^2}{\sigma_T^2}\right)K(K+1)}\right)\,dt. \tag{3.21}$$

Using the Taylor series expansion given in (8.447.1) of [103, p. 909], the Bessel function in (3.21) can be expressed as

$$\mathrm{I}_0\left(2\sqrt{\left(\frac{t-\sigma_N^2}{\sigma_T^2}\right)K(K+1)}\right) = \sum_{a=0}^{\infty}\frac{(t-\sigma_N^2)^a K^a (K+1)^a}{(\sigma_T^2)^a (a!)^2}. \tag{3.22}$$

In (3.22), convergence occurs everywhere and is rapid due to the double factorial in the dominator. Substituting (3.22) into (3.21) yields

$$\mathcal{I} = \sum_{a=0}^{\infty} \frac{K^a (K+1)^a}{(\sigma_T^2)^a (a!)^2} \int_{\sigma_N^2}^{\infty} \frac{e^{\frac{-|y[i]|^2}{t} - \frac{t}{\sigma_T^2}(K+1)} (t-\sigma_N^2)^a}{t}\, dt. \tag{3.23}$$

Then, applying the binomial expansion, we can write (3.23) as

$$\mathcal{I} = \sum_{a=0}^{\infty} \frac{K^a (K+1)^a}{(\sigma_T^2)^a (a!)^2} \sum_{b=0}^{a} \binom{a}{b} (-\sigma_N^2)^b \int_{\sigma_N^2}^{\infty} \frac{e^{\frac{-|y[i]|^2}{t} - \frac{t}{\sigma_T^2}(K+1)}}{t^{-a+b+1}}\, dt. \tag{3.24}$$

The integral in (3.24) can be rewritten using the same approach as in Section 3.4.2 and hence, (3.24) becomes

$$\mathcal{I} = \sum_{a=0}^{\infty} \frac{K^a (K+1)^a}{(\sigma_T^2)^a (a!)^2} \sum_{b=0}^{a} \binom{a}{b} (-\sigma_N^2)^b \left[\int_{0}^{\infty} \frac{e^{\frac{-|y[i]|^2}{t} - \frac{t}{\sigma_T^2}(K+1)}}{t^{-a+b+1}}\, dt - \int_{0}^{\sigma_N^2} \frac{e^{\frac{-|y[i]|^2}{t} - \frac{t}{\sigma_T^2}(K+1)}}{t^{-a+b+1}}\, dt \right]. \tag{3.25}$$

Using (3.471.9) in [103, p. 363] and the midpoint rule of the Riemann sum approximation [104], (3.25) can be expressed as

$$\begin{aligned}\mathcal{I} = \sum_{a=0}^{\infty} \frac{K^a (K+1)^a}{(\sigma_T^2)^a (a!)^2} \sum_{b=0}^{a} \binom{a}{b} (-\sigma_N^2)^b \Bigg[2 \left(\frac{|y[i]|^2 \sigma_T^2}{(K+1)} \right)^{(a-b)/2} \mathrm{K}_{a-b} \left(\frac{|y[i]|\sqrt{K+1}}{\sigma_T/2} \right) - \\ \frac{\sigma_N^2}{R} \sum_{r=1}^{R} \frac{e^{\frac{-|y[i]|^2}{s_r} - \frac{s_r}{\sigma_T^2}(K+1)}}{(s_r)^{1+b-a}} \Bigg],\end{aligned} \tag{3.26}$$

where s_r is given in (3.17). The result for $\mathcal{I}$ in (3.26) can now be substituted into (3.20) and hence, the pdf of the amplitude of the received signal for the Rician case can be written as

$$\begin{aligned}f_{Ric}^{(1)}(|y[i]|) = \frac{2|y[i]|(K+1)e^{-K+\frac{\sigma_N^2}{\sigma_T^2}(K+1)}}{\sigma_T^2} \sum_{a=0}^{\infty} \frac{K^a (K+1)^a}{(\sigma_T^2)^a (a!)^2} \sum_{b=0}^{a} \binom{a}{b} (-\sigma_N^2)^b \times \\ \Bigg[2 \left(\frac{|y[i]|^2 \sigma_T^2}{(K+1)} \right)^{(a-b)/2} \mathrm{K}_{a-b} \left(\frac{|y[i]|\sqrt{K+1}}{\sigma_T/2} \right) - \frac{\sigma_N^2}{R} \sum_{r=1}^{R} \frac{e^{\frac{-|y[i]|^2}{s_r} - \frac{s_r}{\sigma_T^2}(K+1)}}{(s_r)^{1+b-a}} \Bigg].\end{aligned} \tag{3.27}$$

For $K = 0$, the pdf expression in (3.27) reduces to the pdf of the received signal on a Rayleigh channel in (3.16). In (3.27), the number of terms required to evaluate the summation for a specified numerical figure accuracy depends on the value of the Rician K-factor as well as the SNR. For example, in order to achieve six significant figure accuracy at SNR=10 dB, we require 10 summation terms for $K = 0$ dB and 19 terms for $K = 6$ dB. The log likelihood ratio for a Rician channel, $l_{Ric}(y[i])$ can be written using (3.27) and (3.9) to yield

$$l_{Ric}(|y[i]|) = \ln \left\{ \frac{\sigma_N^2 (K+1) e^{-K + \frac{\sigma_N^2}{\sigma_T^2}(K+1) + \frac{|y[i]|^2}{\sigma_N^2}}}{\sigma_T^2} \sum_{a=0}^{\infty} \frac{K^a (K+1)^a}{(\sigma_T^2)^a (a!)^2} \sum_{b=0}^{a} \binom{a}{b} (-\sigma_N^2)^b \times \right.$$
$$\left. \left[2 \left(\frac{|y[i]|^2 \sigma_T^2}{(K+1)} \right)^{(a-b)/2} \mathrm{K}_{a-b} \left(\frac{|y[i]| \sqrt{K+1}}{\sigma_T/2} \right) - \frac{\sigma_N^2}{R} \sum_{r=1}^{R} \frac{e^{\frac{-|y[i]|^2}{s_r} - \frac{s_r}{\sigma_T^2}(K+1)}}{(s_r)^{1+b-a}} \right] \right\}. \tag{3.28}$$

The approach in (3.19)-(3.28), which avoids numerical integration, is particularly useful in the Rician case where simple quadrature methods are problematic. The difficulty is in the integration of $\mathrm{I}_0(.)$ in (3.21) where the Bessel function grows as the argument tends to infinity. Without special numerical treatment, the integral is not well-behaved and returns unstable results. A solution is possible using a finite upper limit on the integral and an asymptotic formula for the Bessel function. However, such numerical issues are conveniently handled by the approach in (3.19)-(3.28).

3.4.4 Nakagami-m channel

If the received signal is transmitted over a Nakagami-m channel, the pdf of the amplitude of the received signal can be written using (3.10) and Table 3.1 in the form

$$f_{Nak}^{(1)}(|y[i]|) = \frac{m^m}{\Gamma(m)} \int_0^{\infty} \frac{2|y[i]| e^{\frac{-|y[i]|^2}{(\sigma_T^2 z[i] + \sigma_N^2)}}}{(\sigma_T^2 z[i] + \sigma_N^2)} z[i]^{m-1} e^{-mz[i]} \, dz[i]. \tag{3.29}$$

Let $t = \sigma_T^2 z[i] + \sigma_N^2$, then (3.29) can be written as

$$f_{Nak}^{(1)}(|y[i]|) = \frac{2|y[i]|m^m e^{\frac{m\sigma_N^2}{\sigma_T^2}}}{(\sigma_T^2)^m \Gamma(m)} \int_{\sigma_N^2}^{\infty} \frac{e^{\frac{-|y[i]|^2}{t} - \frac{mt}{\sigma_T^2}} (t-\sigma_N^2)^{m-1}}{t} \, dt. \tag{3.30}$$

Applying the binomial expansion, (3.30) can be expressed as

$$f_{Nak}^{(1)}(|y[i]|) = \frac{2|y[i]|m^m e^{\frac{m\sigma_N^2}{\sigma_T^2}}}{(\sigma_T^2)^m \Gamma(m)} \sum_{l=0}^{m-1} \binom{m-1}{l} (-\sigma_N^2)^l \int_{\sigma_N^2}^{\infty} e^{\frac{-|y[i]|^2}{t} - \frac{mt}{\sigma_T^2}} t^{m-2-l} \, dt. \tag{3.31}$$

The integral in (3.31) can be rewritten using the same approach taken in Section 3.4.2, which gives

$$\begin{aligned} f_{Nak}^{(1)}(|y[i]|) = {} & \frac{2|y[i]|m^m e^{\frac{m\sigma_N^2}{\sigma_T^2}}}{(\sigma_T^2)^m \Gamma(m)} \sum_{l=0}^{m-1} \binom{m-1}{l} (-\sigma_N^2)^l \times \\ & \left[\int_0^{\infty} e^{\frac{-|y[i]|^2}{t} - \frac{mt}{\sigma_T^2}} t^{m-2-l} \, dt - \int_0^{\sigma_N^2} e^{\frac{-|y[i]|^2}{t} - \frac{mt}{\sigma_T^2}} t^{m-2-l} \, dt \right]. \end{aligned} \tag{3.32}$$

With the aid of (3.471.9) in [103, p. 363] and the midpoint rule of the Riemann sum approximation [104], the pdf of the amplitude of the received signal, $f_{Nak}^{(1)}(|y[i]|)$, in (3.32) can then be expressed as

$$\begin{aligned} f_{Nak}^{(1)}(|y[i]|) = {} & \frac{2|y[i]|m^m e^{\frac{m\sigma_N^2}{\sigma_T^2}}}{(\sigma_T^2)^m \Gamma(m)} \sum_{l=0}^{m-1} \binom{m-1}{l} (-\sigma_N^2)^l \left[2 \left(\frac{|y[i]|^2}{m/\sigma_T^2} \right)^{(m-1-l)/2} \times \right. \\ & \left. \mathrm{K}_{m-1-l}\left(\frac{|y[i]|\sqrt{m}}{\sigma_T/2} \right) - \frac{\sigma_N^2}{R} \sum_{r=1}^{R} e^{\frac{-|y[i]|^2}{s_r} - \frac{ms_r}{\sigma_T^2}} (s_r)^{m-2-l} \right], \end{aligned} \tag{3.33}$$

where all terms in (3.33) have been defined previously. The log likelihood ratio, $l_{Nak}(y[i])$, can be expressed using (3.9) along with (3.33) which gives

$$\begin{aligned} l_{Nak}(|y[i]|) = \ln & \left\{ \frac{\sigma_N^2 m^m e^{\frac{m\sigma_N^2}{\sigma_T^2} + \frac{|y[i]|^2}{\sigma_N^2}}}{(\sigma_T^2)^m \Gamma(m)} \sum_{l=0}^{m-1} \binom{m-1}{l} (-\sigma_N^2)^l \left[2 \left(\frac{|y[i]|^2}{m/\sigma_T^2} \right)^{(m-1-l)/2} \times \right. \right. \\ & \left. \left. \mathrm{K}_{m-1-l}\left(\frac{|y[i]|\sqrt{m}}{\sigma_T/2} \right) - \frac{\sigma_N^2}{R} \sum_{r=1}^{R} e^{\frac{-|y[i]|^2}{s_r} - \frac{ms_r}{\sigma_T^2}} (s_r)^{m-2-l} \right] \right\}. \end{aligned} \tag{3.34}$$

3.4.5 F channel

The F channel represents an extreme case, where the received signal experiences severe fading. The amplitude of the received signal over the F channel can be derived using (3.10) and Table 3.1 to give

$$f_F^{(1)}(|y[i]|) = \int_0^\infty \frac{2|y[i]|e^{\frac{-|y[i]|^2}{(\sigma_T^2 z[i]+\sigma_N^2)}}}{(\sigma_T^2 z[i]+\sigma_N^2)} \frac{2}{\pi\sqrt{z[i]}(1+z[i])^2}\, dz[i]. \tag{3.35}$$

The log likelihood ratio can then be written using (3.9) and (3.35) as

$$l_F(|y[i]|) = \frac{2\sigma_N^2 e^{\frac{|y[i]|^2}{\sigma_N^2}}}{\pi} \int_0^\infty \frac{e^{\frac{-|y[i]|^2}{(\sigma_T^2 z[i]+\sigma_N^2)}}}{(\sigma_T^2 z[i]+\sigma_N^2)\sqrt{z[i]}(1+z[i])^2}\, dz[i]. \tag{3.36}$$

In order to detect the presence of the PU for a range of channel models, the log likelihood ratios, $l_G, l_{Ray}, l_{Ric}, l_{Nak}$ and l_F, can be substituted into the CUSUM algorithm in (2.21).

3.5 NUMERICAL RESULTS

In this section, the performance of quickest spectrum sensing in classical channel models as well as in a severe fading model is presented. All simulation results used 20000 trials to generate each point on the plot, with each trial consisting of 200 samples. It is worth noting that more trials are used in the thesis to improve simulation accuracy compared to [101]. It is assumed that the PU begins transmission at $\tau = 100$. This is illustrated in Figure 3.1, where we see that the PU could commence transmission at any time between sample 99 and sample 100. In terms of the CU sample times, this is equivalent to the PU being active at the largest sample number (i.e. $\tau = 100$). In this chapter, the thresholds are set to $\gamma = 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7$ by trial and error. These values were chosen to give a sensible range of average detection delay and false alarm rate. Based on these pre-determined threshold values, we measure the average detection delay and the false alarm rate.

Figures 3.2 and 3.3 illustrate the quickest spectrum sensing performance for SNR=5 dB and 10 dB, respectively, where the CUSUM detectors are designed specifically for the particular

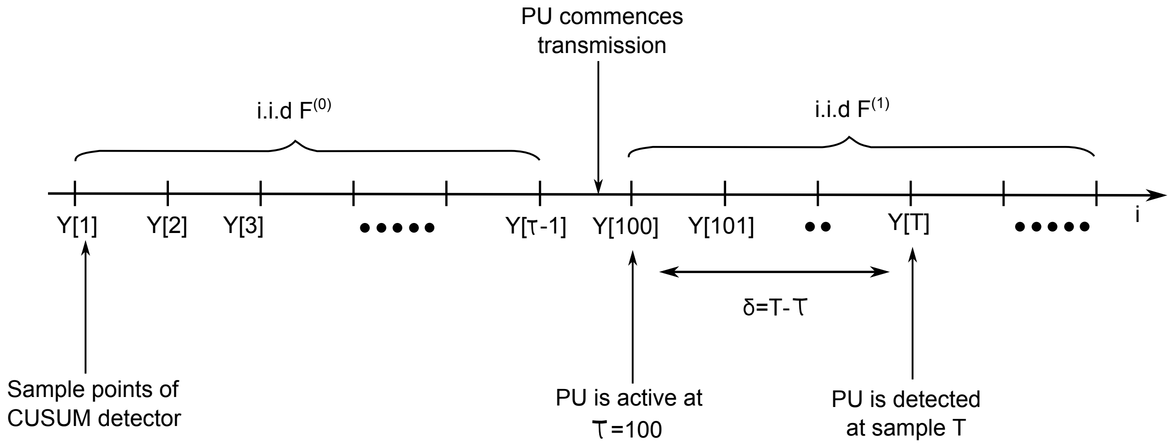

Figure 3.1 Illustration of the primary user (PU) detection based on the CUSUM algorithm, employed at the cognitive user.

channels. Based on both figures, since the time-invariant channel has no fading and the long-tailed F distribution can model severe fading, it is reasonable to suggest that detection delay increases with increasing severity of the fading. The CV, discussed in Section 2.1.6, is often used to measure the fading severity [62] and is defined as

$$\mathrm{CV} = \sqrt{\frac{1}{E\left[|H[i]|\right]^2} - 1}, \tag{3.37}$$

for channels normalized to have unit power, $E\left[|H[i]|^2\right] = 1$. Since the CV is only a function of the mean channel amplitude, standard results for the channel distributions can be used to obtain the value of the CV, which results in Table 3.2. Based on Table 3.2, the order of fading severity is given, from lowest to highest as, time-invariant, Nakagami-3, Rician (K=6 dB), Nakagami-2, Rician (K=0 dB), Rayleigh and F.

Comparing this ranking based on CV values with Figures 3.2 and 3.3, it can be seen that the ranking of the time-invariant, Nakagami-3, Rician (K=0 dB), Rayleigh, and F channels matches the simulated results for SNR=5 dB. However, for SNR=10 dB, the order of the time-invariant, Rayleigh, and the F channels matches the CV ranking. The slight variation in the ordering of the Nakagami-3, Rician (K=6 dB), Nakagami-2 and Rician (K=0 dB) channels between the CV and simulation results may be due to the fact that the expected values of the channel amplitude for

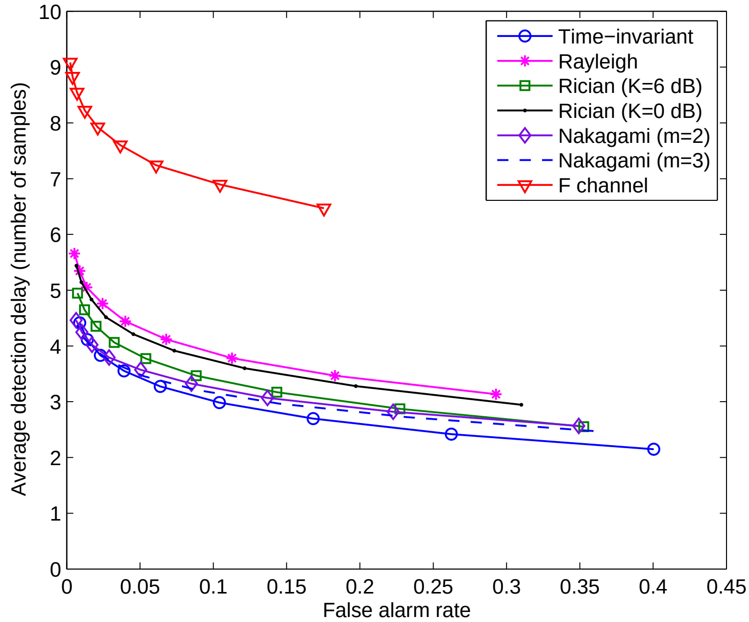

Figure 3.2 The effect of various channel conditions on the sensing performance at SNR=5 dB.

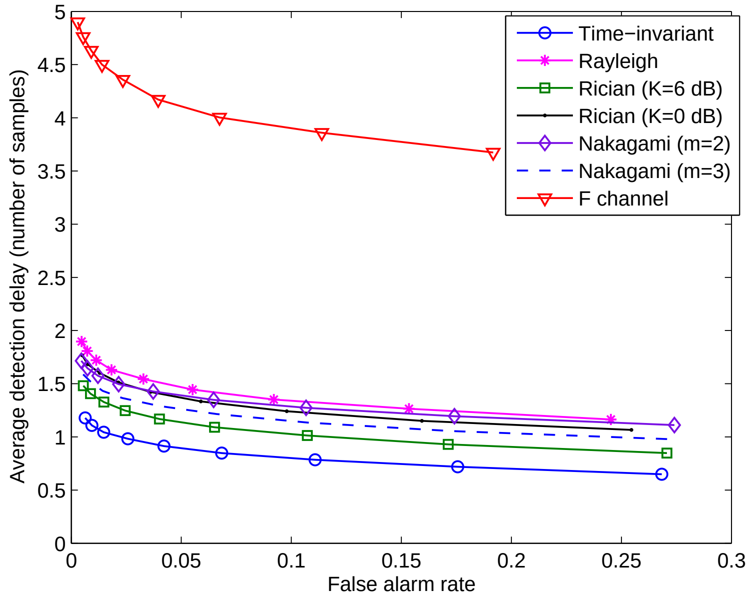

Figure 3.3 The effect of various channel conditions on the sensing performance at SNR=10 dB.

Table 3.2 CV values for the channel models

Channel model	$E[\|H[i]\|]$	CV
Time-invariant	1	0
Rayleigh	$\frac{\sqrt{\pi}}{2}$	$\sqrt{\frac{4}{\pi}-1}=0.52$
Rician(K)	$\sqrt{\frac{\pi}{4(K+1)}}e^{\frac{-K}{2}}\times$ $\left((1+K)\mathrm{I}_0\left(\frac{K}{2}\right)+K\mathrm{I}_1\left(\frac{K}{2}\right)\right)$	$\sqrt{\frac{\pi e^K}{4(K+1)}\left((1+K)\mathrm{I}_0\left(\frac{K}{2}\right)+K\mathrm{I}_1\left(\frac{K}{2}\right)\right)^2-1}$
Rician(K=0 dB)	0.91	0.46
Rician(K=6 dB)	0.95	0.32
Nakagami-m	$\frac{\Gamma(m+1/2)}{\Gamma(m)\sqrt{m}}$	$\sqrt{\frac{m\Gamma^2(m)}{\Gamma^2(m+1/2)}-1}$
Nakagami-2	0.94	0.36
Nakagami-3	0.96	0.29
$\mathrm{F}_{1,3}$	$\frac{2}{\pi}$	$\sqrt{\frac{\pi^2}{4}-1}=1.21$

these channels are very close (0.96, 0.95, 0.94, 0.91). Hence, such a simple ranking approach does not completely capture the full CUSUM behaviour. Nevertheless, we can see that the average detection delay tends to increase with the severity of the fading channels according to the CV metric.

In Figure 3.4, we present a comparison of the sensing performance when the CUSUM detector is designed for the time-invariant or Rayleigh channels, but a different channel is experienced. Based on Figure 3.4, we observe that in the event of such a channel mis-match, sensing performance depends highly on the channel, but the detector has very little impact. For example, the average detection delay when the received signal is transmitted over a Rayleigh channel is almost identical when the signal is detected using either a Rayleigh or a time-invariant detector. The same trends can be observed in Figure 3.5 for Rayleigh and Rician fading. Similar trends were observed in [105] in the event of a channel mis-match, but [105] senses the PU in a wideband spectrum using Bayesian sequential testing with dynamic update. A detailed description on the insensitivity of the designed detector to the performance of quickest spectrum sensing is given in Chapter 4.

Figure 3.6 shows the sensing performance of a Rayleigh detector under various mis-matched channels. It can be seen that for a Rayleigh detector, the ranking of the channels is the same as in Figure 3.3. We also observe that the duration of the detection delay depends on the fading channel conditions.

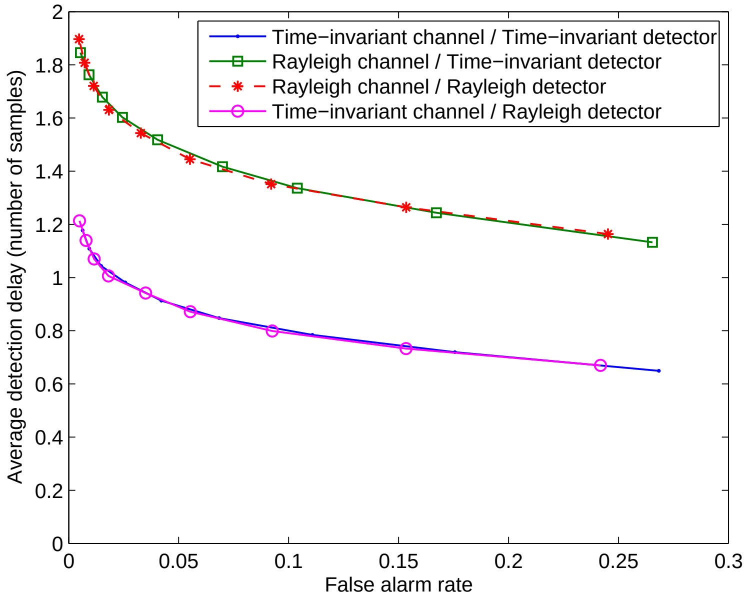

Figure 3.4 Performance comparison in the event of a mismatch between channel and detector (Rayleigh / Time-invariant cases) at SNR=10 dB.

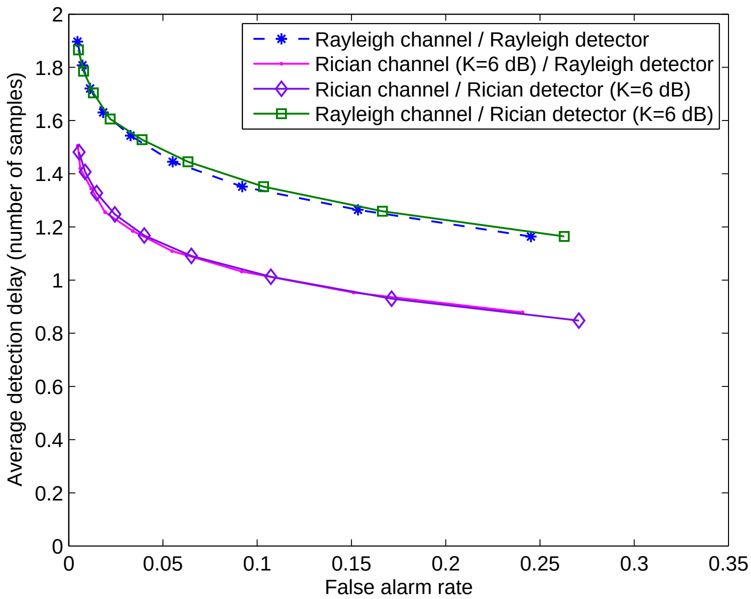

Figure 3.5 Performance comparison in the event of a mismatch between channel and detector (Rayleigh / Rician cases) at SNR=10 dB.

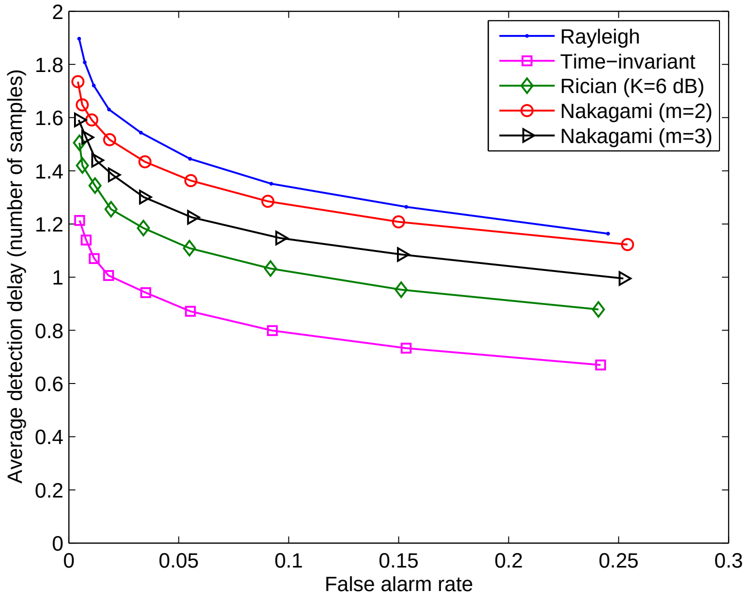

Figure 3.6 Performance of a Rayleigh detector for various channels at SNR=10 dB.

Recall from Section 1.4 that a comparison of the quickest spectrum sensing performance with energy detection in various fading channels is difficult as both approaches have different goals.

3.5.1 Temporally correlated channel

Since the classic quickest spectrum sensing approaches assume a sequence of independent observations at the CU, it is useful to investigate the sensing performance when this assumption is violated. Hence, in this section, we consider a PU signal observed over a temporally correlated channel. For simplicity, we consider Rayleigh fading channels in the correlated scenario. We assume the classic temporal behaviour of a Rayleigh channel with Jakes correlation structure defined as [106, 107]

$$E[H(t)H(t+\alpha)^*] = \mathrm{J}_0(2\pi f_D \alpha), \tag{3.38}$$

where $\mathrm{J}_0(.)$ is the zeroth order Bessel function of the first kind, f_D is the maximum Doppler shift in Hertz and α is the time separation.

Simulation results in the case when the received signal experiences a temporally correlated channel are illustrated in Figure 3.7. The signal is detected using either a time-invariant or

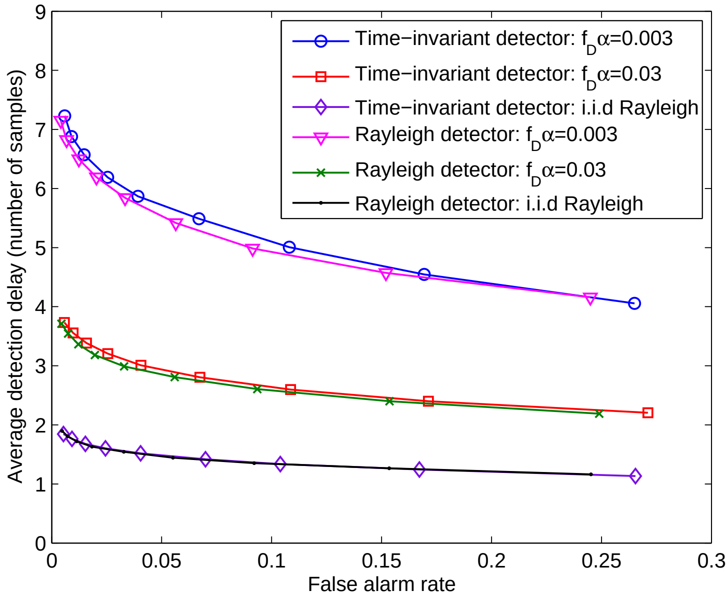

Figure 3.7 The effect of sensing performance on a temporally correlated Rayleigh channel with different fade rates, $f_D\alpha$ at SNR=10 dB.

Rayleigh detector. In this case, a time-varying Rayleigh channel is assumed and the fade rate is given by $f_D\alpha$. We consider the i.i.d Rayleigh channel (no correlation) as the extreme case. Based on Figure 3.7, we observe that for both detectors, the average detection delay increases with the level of temporal correlation. This is because when the correlated channels are in deep fades, the received signals are weak over many sample periods and hence detection takes longer for the CUSUM statistic, C_n to exceed the threshold, γ. This then creates a longer detection delay.

3.6 CHAPTER SUMMARY

In this chapter, we studied the performance of quickest spectrum sensing over a time-invariant, Rayleigh, Rician, Nakagami-m or F channel. The received signal at the CU is considered to be the product of two complex Gaussian variables (the PU signal and channel) with additive noise. We first proved that the power of the complex received signal is a sufficient statistic and hence the log likelihood can be computed on the basis of the signal amplitude. This allows us to derive a general log likelihood ratio for various channel conditions. We then derived the pdf of the received signal amplitude for the Rayleigh, Rician, Nakagami-m and F channel scenarios. A technique which avoids numerical integration was used in the derivation of the pdf of the received signal amplitude for the most commonly used fading channels, including Rayleigh, Rician and Nakagami-m channels.

We showed that the average detection delay increases with fading severity and the level of temporal correlation. Severe fading and highly temporally correlated channels cause high detection delays because the PU signal can be weak for many samples even if the average SNR is reasonable.

We also showed that in the event of a mis-matched channel, the average detection delay is sensitive to the true channel conditions, but not on the channel used to design the CUSUM detector. Thus, an appealing strategy is to employ the simplest time-invariant detector to detect the PU existence since in a cognitive radio network, the channel is usually unknown and the sensing performance in various channels is insensitive to the design of the detector. Since the time-invariant detector is robust and has general applicability, we will formulate novel theoretical expressions for the distribution of detection delay for a time-invariant detector in the next chapter.

Chapter 4

ON THE DISTRIBUTION OF DETECTION DELAY

4.1 INTRODUCTION

Several quickest spectrum sensing studies in the literature [37,39,42,101,108] evaluate the sensing performance in terms of average detection delay and false alarm rate. However, no studies have given a theoretical expression for the distribution of detection delay, which is useful for a more detailed analysis of sensing performance, including an assessment of the likelihood of long delays.

In Chapter 3, we studied the performance of quickest spectrum sensing over time-invariant as well as various fading channels, including Rayleigh, Rician, Nakagami-m and F channels. Results in Chapter 3 show that in the event of a mis-matched channel (where the CUSUM detector is designed for a specific channel, but the true channel is different), the quickest spectrum sensing performance depends heavily on the true channel, but very little on the channel assumed by the CUSUM detector. Hence, the simplest time-invariant detector can be employed with minimal performance loss.

Therefore, motivated by the result in Chapter 3, in this chapter, we derive mathematical expressions for the distribution of detection delay for the time-invariant CUSUM detector when the signal is transmitted over Gaussian and Rayleigh channels. Slowly varying channels can be modelled by a Gaussian channel model. Hence, the theoretical expression on the distribution of detection delay for the Gaussian scenario allows detailed analysis of the quickest spectrum sensing performance in an ideal or optimal condition. For each of the scenarios considered, we present several approximation methods based on a modified detection delay statistic. This is discussed in detail in Section 4.3. Approximate methods are necessary since there is no exact solution available. We focus on Gaussian and Rayleigh fading channels although the approximate methods can handle arbitrary channel models.

We also analyze the accuracy of the modified detection delay statistic. This is valuable in understanding the limitations of the approximations used. We also approximate the probability of missed detection and study the probability of long detection delays via analysis and simulation. This analysis of long detection delays is useful to gain some insight into the factors that contribute to an increased probability of long detection delays.

The rest of the chapter is organized as follows. Section 4.2 describes the system model and briefly introduces the modified detection delay statistic. Several techniques for approximating the distribution of detection delay are presented in Section 4.3. In Section 4.4, we analyze the performance of the modified detection delay statistic, present some numerical results to validate the approximate expressions, approximate the probability of missed detection and evaluate the probability of long detection delays. Section 4.5 concludes the chapter.

4.2 SYSTEM MODEL

We consider the PU signal to be a narrowband complex Gaussian signal. The PU is initially inactive, but subsequently commences transmission. Hence, the CU is attempting to detect the appearance of the PU. Let $Y[i]$ denote the received signal at the CU at sample number i. If the PU is absent, then $Y[i] = N[i]$, where $N[i]$ is the noise. If the PU is transmitting, then $Y[i] = H[i] \times S[i] + N[i]$, where $H[i]$ is the channel coefficient and $S[i]$ is the PU signal. We let $S[i]$ and $N[i]$ be independent circularly symmetric complex Gaussian random variables such that $S[i] \sim \mathcal{CN}(0, \sigma_S^2)$ and $N[i] \sim \mathcal{CN}(0, \sigma_N^2)$. We define $H[i] = H$ for the Gaussian case, where H is the time-invariant channel gain. For the Rayleigh case, $H[i]$ is the Rayleigh fading channel, where $H[i] \sim \mathcal{CN}(0, 1)$. For the Gaussian case, $X[i] = H \times S[i]$ is a circularly symmetric complex Gaussian variable with variance σ_X^2. For the Rayleigh case, $X[i] = H[i] \times S[i]$ has a more complex distribution (as seen in Section 3.4.2) with variance also denoted by σ_X^2.

Assume that the PU commences transmission at sample number τ and the CU uses the CUSUM algorithm (discussed in Section 2.4.1) to detect the PU via a change in the distribution of the received signal. Recall from Section 2.4.1, that the CUSUM algorithm detects the PU at sample T and the CUSUM statistic is given in (2.21) as

$$C_{n+1} = \{C_n + l(Y[n+1])\}^+, \tag{4.1}$$

where $x^+ = \max(x, 0)$ and $l(Y[i])$ is the log likelihood ratio, given in (2.20). If $T > \tau$, a detection delay of $T - \tau$ will occur. The probability that the detection delay is δ samples is denoted by $P(\delta)$ and this defines the distribution of detection delay. As denoted in (3.12) of Section 3.4.1, the log likelihood ratio, for a time-invariant CUSUM detector is given by

$$l(Y[i]) = l(|Y[i]|) = \frac{|Y[i]|^2\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)} + \ln\left\{\frac{\sigma_N^2}{\sigma_N^2+\sigma_X^2}\right\}, \tag{4.2}$$

since $|Y[i]|^2$ is a sufficient statistic. Hence, (4.2) is the time-invariant log likelihood ratio.

4.2.1 Modified detection delay statistic

The CUSUM statistic, C_n, is very difficult to handle analytically due to the max operation in (4.1). This creates a complex dependence between the C_n values and the log likelihood ratios, $l(Y[i])$, which makes exact analysis intractable in the general case. Hence, we approximate the CUSUM process by the modified detection delay statistic, $\mathcal{D}_n$, defined as

$$\begin{aligned}\mathcal{D}_n &= \sum_{i=1}^{n} l(Y[i]) \\ &= \sum_{i=1}^{n}\left(\frac{|Y[i]|^2\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)} + \ln\left\{\frac{\sigma_N^2}{\sigma_N^2+\sigma_X^2}\right\}\right).\end{aligned} \tag{4.3}$$

Note that the exact CUSUM statistic defined in (4.1) and (4.2) is used for all simulated results in this chapter. The modified process in (4.3) is simply used to provide analytical approximations. In (4.3), $\mathcal{D}_n$ is the sum of a fixed number of terms, n. Let $\widetilde{Z}_i = |Y[i]|^2$, $\bar{a} = \frac{\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)}$ and $b = -\ln\left\{\frac{\sigma_N^2}{\sigma_N^2+\sigma_X^2}\right\}$, then $\bar{a} > 0$, $b > 0$ and (4.3) can be rewritten as

$$\mathcal{D}_n = \sum_{i=1}^{n}(\bar{a}\widetilde{Z}_i - b), \tag{4.4}$$

which is identical to the CUSUM value when the $\mathcal{D}_n$ process remains positive. The approximation lies in the fact that $\mathcal{D}_n$ can go negative, whereas $C_n \geq 0$. However, in the presence of a PU, the log likelihood ratios have a positive mean and large negative values are unlikely. Hence, we approximate C_n by the $\mathcal{D}_n$ process which has i.i.d increments of the form $\bar{a}\widetilde{Z}_i - b$. At low signal-to-noise ratio (SNR), it is more likely the $\mathcal{D}_n$ process will become negative since the $\widetilde{Z}_i$ values tend to be smaller. In contrast, the CUSUM process in (4.1) is restricted to $C_n \geq 0$. Hence, the approximate analysis is likely to be least accurate in the low SNR scenario. This is considered analytically in Section 4.4.1 and via simulations in Section 4.4.2. These results demonstrate that the approximation is useful in all cases except for short delays at low SNR, where improvements are possible. However, this approach always provides useful approximations for the important case of long delays.

Let ω be the sample number at which the approximate CUSUM process, $\mathcal{D}_n$, detects the PU transmission such that $\omega = \inf(n : \mathcal{D}_n \geq \gamma)$, where γ is a threshold. Therefore,

$$\mathcal{D}_\omega = \sum_{i=1}^{\omega} (\bar{a}\widetilde{Z}_i - b). \tag{4.5}$$

$\mathcal{D}_\omega$ is then a random sum of random variables. We assume that the PU is inactive at $i = 0$ and becomes active at $i = 1$. This corresponds to $\tau = 1$ and since ω is the sample number at which the approximate CUSUM process, $\mathcal{D}_n$, detects the PU, the detection delay is denoted by $\omega - 1$. Therefore, we approximate the true detection delay probability, $P(\delta) = P(\text{CUSUM detection delay} = \delta)$, by the approximation $P(\omega = \delta + 1)$, where ω is defined based on the detection delay definition given in Section 4.2.1.

The process, $\mathcal{D}_n$, is a discrete time, continuous state space Markov process often referred to as a random walk [109, 110]. Hence, for $\mathcal{D}_n$, the detection delay problem is a classic first passage time or hitting problem, where ω is the time at which $\mathcal{D}_n$ exceeds a threshold, γ, for the first time [111]. For general problems of this kind, numerous steady state results are available, such as expected hitting times [110, 112] and hitting probabilities [110, 111, 113]. However, results on the distribution of hitting time are far more limited and are largely restricted to Brownian motion and simple special cases of random walks [110, 111, 114, 115]. Therefore, in this chapter, we derive a novel approximate closed-form expression for the detection delay distribution for the Gaussian case. In addition, we derive novel approximations for the distribution of detection delay for the general case due to the absence of a general framework. Known results based on simple random walks and Brownian motion theory are also employed in deriving approximate expressions for the detection delay distribution.

4.3 DISTRIBUTION OF DETECTION DELAY

In this section, we introduce some methods for constructing approximate expressions for the distribution of detection delay for both Gaussian and Rayleigh channels.

4.3.1 Gaussian channel

The Gaussian channel considered in this section corresponds to a constant PU to CU channel, $H = 1$, as discussed in Section 3.4.1. Normalizing $\widetilde{Z}_i$ to have unit mean gives

$$\begin{aligned} \mathcal{D}_n &= \sum_{i=1}^{n} \big(\bar{a}(\sigma_N^2 + \sigma_X^2) Z_i - b\big) \\ &= a \sum_{i=1}^{n} Z_i - nb, \end{aligned} \tag{4.6}$$

where $a = \sigma_X^2/\sigma_N^2$ and Z_i denotes the normalized version of $\widetilde{Z}_i$, in which Z_i is an exponential random variable with parameter $\lambda = 1$. Hence, (4.5) can be rewritten as

$$\mathcal{D}_\omega = a \sum_{i=1}^{\omega} Z_i - \omega b. \tag{4.7}$$

In all subsequent analysis of the approximate expression for the detection delay distribution for the Gaussian case, we will be using $\mathcal{D}_n$ in (4.6) and $\mathcal{D}_\omega$ in (4.7).

4.3.1.1 Closed-form expression

The approximate distribution of detection delay can be expressed as an integral of the joint pdf of the detection delay statistics. This gives

$$\begin{aligned}\widetilde{P}_c(\delta) &= P\left(\mathcal{D}_1 < \gamma, \mathcal{D}_2 < \gamma, ..., \mathcal{D}_{\delta+1} \geq \gamma\right) \\ &= P\big(aZ_1 - b < \gamma, a(Z_1+Z_2) - 2b < \gamma, ..., a(Z_1+Z_2+...+Z_{\delta+1}) - (\delta+1)b \geq \gamma\big) \\ &= P\left(Z_1 < \frac{\gamma+b}{a}, Z_1+Z_2 < \frac{\gamma+2b}{a}, ..., Z_1+Z_2+...+Z_{\delta+1} \geq \frac{\gamma+(\delta+1)b}{a}\right),\end{aligned} \tag{4.8}$$

where the subscript, c, in $\widetilde{P}_c(\delta)$ denotes the closed-form expression and tilde denotes an approximation of $P(\delta)$. Let $V_u = \frac{\gamma+ub}{a}$, then (4.8) yields

$$\begin{aligned}\widetilde{P}_c(\delta) &= P\left(Z_1 < V_1, Z_1+Z_2 < V_2, ..., Z_1+Z_2+...+Z_{\delta+1} \geq V_{\delta+1}\right) \\ &= \int_{z_1=0}^{V_1}\int_{z_2=0}^{V_2-z_1} ... \int_{z_\delta=0}^{V_\delta - z_1 - z_2 - ... - z_{\delta-1}} P\big(Z_{\delta+1} \geq V_{\delta+1} - z_1 - z_2 - ... - z_\delta\big) f(z_1) f(z_2) ... \\ &\quad f(z_\delta)\, dz_\delta ... \, dz_2 \, dz_1,\end{aligned} \tag{4.9}$$

where $f(z_i)$ is the pdf of Z_i. Substituting $f(z_i) = e^{-z_i}$ and $P(Z_{\delta+1} \geq V_{\delta+1} - z_1 - z_2 - ... - z_\delta) = e^{-(V_{\delta+1} - z_1 - z_2 - ... - z_\delta)}$ in (4.9) gives

$$\widetilde{P}_c(\delta) = e^{-V_{\delta+1}} \mathcal{I}_{\delta+1}, \tag{4.10}$$

where

$$\mathcal{I}_{\delta+1} = \int_{z_1=0}^{V_1}\int_{z_2=0}^{V_2-z_1} ... \int_{z_\delta=0}^{V_\delta - z_1 - z_2 - ... - z_{\delta-1}} 1 \, dz_\delta ... \, dz_2 \, dz_1. \tag{4.11}$$

The following theorem provides a closed-form expression for $\widetilde{P}_c(\delta)$.

Theorem 1. *Defining* $V = \frac{\gamma+b}{a}$, $\mathcal{I}_{\delta+1}$ *can be derived as*

$$\mathcal{I}_{\delta+1} = V \left(\frac{V_{\delta+1}^{\delta-1}}{\Gamma(\delta+1)}\right). \tag{4.12}$$

Substituting (4.12) into (4.10) gives

$$\widetilde{P}_c(\delta) = e^{-V_{\delta+1}V}\left(\frac{V_{\delta+1}^{\delta-1}}{\Gamma(\delta+1)}\right). \tag{4.13}$$

We will now prove (4.12) by induction. In order to prove (4.12), we let $r = \delta + 1$.

Proof. We know that for $r = 1$, $\mathcal{I}_1 = 1$. Substituting $r = 1$ into (4.12) also yields $\mathcal{I}_1 = 1$. Now, assume (4.12) is true for $r = k$, then for $r = k + 1$,

$$\mathcal{I}_{k+1} = \int_{z_1=0}^{V_1}\int_{z_2=0}^{V_2-z_1}\dots\int_{z_{k-1}=0}^{V_{k-1}-z_1-z_2-\dots-z_{k-2}}\int_{z_k=0}^{V_k-z_1-z_2-\dots-z_{k-1}} 1\,dz_k\,dz_{k-1}\dots dz_2\,dz_1. \tag{4.14}$$

In (4.14), the inner $k-1$ dimensional integral is given by

$$\mathcal{I} = \int_{z_2=0}^{V_1}\int_{z_3=0}^{V_2-z_2}\dots\int_{z_k=0}^{V_{k-1}-z_2-\dots-z_{k-1}} 1\,dz_k\dots dz_3\,dz_2, \tag{4.15}$$

using the relabelling: $V_1 = V_2 - z_1$, $V_2 = V_3 - z_1$,...,$V_{k-1} = V_k - z_1$. From the induction hypothesis and using $\varphi = b/a$, it can be seen that

$$\mathcal{I} = \frac{V(V + (k-1)\varphi)^{k-2}}{(k-1)!}, \tag{4.16}$$

which follows from equating $\mathcal{I}$ in (4.15) to $\mathcal{I}_{\delta+1}$ in (4.11). Then, inserting expression (4.16) into (4.14) and using the fact that $V = V_1 = V_2 - z_1$, (4.14) can be expressed as

$$\mathcal{I}_{k+1} = \frac{1}{(k-1)!}\int_{z_1=0}^{V_1}(V_2 - z_1)\times(V_2 - z_1 + (k-1)\varphi)^{k-2}\,dz_1. \tag{4.17}$$

Evaluating (4.17) using integration by parts gives

$$\begin{aligned}\mathcal{I}_{k+1} = \frac{1}{(k-1)!}\Bigg[&\frac{1}{k-1}\left(V_2(V_2 + (k-1)\varphi)^{k-1} - (V_2 - V_1)(V_2 - V_1 + (k-1)\varphi)^{k-1}\right) + \\ &\frac{1}{k(k-1)}\left((V_2 - V_1 + (k-1)\varphi)^k - (V_2 + (k-1)\varphi)^k\right)\Bigg].\end{aligned} \tag{4.18}$$

Substituting $V_1 = V$ and $V_2 = V + \varphi$ into (4.17) yields

$$\begin{aligned}\mathcal{I}_{k+1} &= \frac{1}{k(k-1)(k-1)!}\Big[k(V+\varphi)(V+k\varphi)^{k-1} - (V+k\varphi)^{k}\Big] \\ &= \frac{V(V+k\varphi)^{k-1}}{k!} \\ &= \frac{V(V_{k+1})^{k-1}}{\Gamma(k+1)}.\end{aligned} \tag{4.19}$$

This completes the proof. Hence, (4.12) holds for all $\delta + 1$. □

4.3.1.2 Approximation

Although a closed-form expression for the approximate delay distribution in (4.13) has been derived, it is only applicable to the Gaussian channel. Hence, in this section, as well as in Sections 4.3.1.3 and 4.3.1.4, we also consider approximate delay distributions which have a wider range of applicability. In deriving an approximate expression for $P(\delta)$, we consider $\mathcal{D}_\omega$ in (4.7) just before and just after a threshold crossing. Hence, we focus on any crossing of the threshold without requiring it to be the first crossing. This results in having an approximation which tends to be an overestimate. $P(\delta)$ can then be approximated by

$$\begin{aligned}\widetilde{P}_{a_G}(\delta) &= P(\mathcal{D}_\delta < \gamma, \mathcal{D}_{\delta+1} \geq \gamma) \\ &= P\left(a\sum_{i=1}^{\delta} Z_i - \delta b < \gamma, a\sum_{i=1}^{\delta+1} Z_i - (\delta+1)b < \gamma\right) \\ &= P\left(\sum_{i=1}^{\delta} Z_i < \frac{\gamma+\delta b}{a}, \sum_{i=1}^{\delta} Z_i + Z_{\delta+1} \geq \frac{\gamma+(\delta+1)b}{a}\right).\end{aligned} \tag{4.20}$$

The subscript, a_G, in $\widetilde{P}_{a_G}(\delta)$ denotes the approximation method for a Gaussian channel. Let $\mathcal{S}_\delta = \sum_{i=1}^{\delta} Z_i$, then (4.20) becomes

$$\begin{aligned}\widetilde{P}_{a_G}(\delta) &= P\left(\mathcal{S}_\delta < \frac{\gamma+\delta b}{a}, Z_{\delta+1} \geq \frac{\gamma+(\delta+1)b - a\mathcal{S}_\delta}{a}\right) \\ &= \int_0^{\frac{\gamma+\delta b}{a}} f_{\mathcal{S}_\delta}(z) P\left(Z_{\delta+1} \geq \frac{\gamma+(\delta+1)b}{a} - z\right) dz,\end{aligned} \tag{4.21}$$

where $f_{\mathcal{S}_\delta}(z)$ is the pdf of a chi-square distribution with 2δ degrees of freedom, defined as [116]

$$f_{\mathcal{S}_\delta}(z) = \frac{z^{\delta-1}e^{-z}}{\Gamma(\delta)}, \quad z \geq 0. \tag{4.22}$$

Using (4.22), the cumulative distribution function (cdf) of $Z_{\delta+1}$ and evaluating the resulting simple integral in (4.21), we obtain

$$\widetilde{P}_{a_G}(\delta) = \frac{e^{-\left(\frac{\gamma+(\delta+1)b}{a}\right)}\left(\frac{\gamma+\delta b}{a}\right)^{\delta}}{\Gamma(\delta+1)}. \tag{4.23}$$

4.3.1.3 Approximation based on a random walk approach

The classical random walk theory can also be applied in approximating $P(\delta)$. Recall that $\mathcal{D}_n$ in (4.6) has i.i.d continuous increments of the form $aZ_i - b$. Here, we model the CUSUM process $\{C_n\}$, via $\mathcal{D}_n$ in (4.6) by the highly simplified random walk with i.i.d jumps of fixed size, D, and fixed probabilities of positive or negative jumps. Hence, the sensing delay becomes equivalent to a first passage time for a simple random walk. To model $\{C_n\}$ effectively, we select the random walk parameters so that the jumps have the same mean and variance as the jumps in $\mathcal{D}_n$. From (4.6), the jumps in $\mathcal{D}_n$ are given by

$$J_{i_G} = aZ_i - b. \tag{4.24}$$

Thus, the expected value and variance of each J_{i_G} in (4.24) can be expressed respectively as

$$E[J_{i_G}] = a - b, \tag{4.25}$$

$$\mathrm{Var}(J_{i_G}) = a^2. \tag{4.26}$$

The equivalent random walk has jumps, J_{eq}, which are upward, $+D$, with probability q and downward, $-D$, with probability $p = 1 - q$. Thus, the expected value and variance of J_{eq} can be defined respectively as

$$E[J_{eq}] = D(2q-1), \tag{4.27}$$

$$\begin{aligned}
\mathrm{Var}(J_{eq}) &= E[J_{eq}^2] - E[J_{eq}]^2 \\
&= \left[D^2 q + (-D)^2(1-q)\right] - D^2(2q-1)^2 \\
&= D^2\left(1-(2q-1)^2\right).
\end{aligned} \tag{4.28}$$

Equating the means and variances in (4.25), (4.26), (4.27) and (4.28) yields

$$D = \frac{a-b}{2q-1}, \tag{4.29}$$

and

$$q = \frac{1}{2} + \frac{1}{2}\frac{1}{\sqrt{\eta+1}}, \tag{4.30}$$

where

$$\eta = \frac{a^2}{(a-b)^2}. \tag{4.31}$$

Normalizing the sensing process to multiples of D, we define the number of steps needed by the random walk to cross γ as $\mathcal{K} = \gamma/D$. Thus, the approximate expression for $P(\delta)$ can be written as [111, 114]

$$\widetilde{P}_{rw_G}(\delta) = \frac{\mathcal{K}\Gamma(\delta+2)}{(\delta+1)\Gamma(\frac{\delta-\mathcal{K}}{2}+\frac{3}{2})\Gamma(\frac{\delta+\mathcal{K}}{2}+\frac{3}{2})} p^{\frac{\delta+1-\mathcal{K}}{2}} q^{\frac{\delta+1+\mathcal{K}}{2}}, \quad \delta \geq \mathcal{K}-1. \tag{4.32}$$

The subscript, rw_G, in $\widetilde{P}_{rw_G}(\delta)$ denotes the random walk approximation for the Gaussian case.

4.3.1.4 Approximation based on a Brownian motion approach

We can also apply the theory of Brownian motion with drift in approximating $P(\delta)$. In this scenario, the CUSUM process, $\{C_n\}$, can be modelled by re-scaling $\mathcal{D}_n$ in (4.6) to construct a continuous Brownian motion process. This yields

$$Q(n) = \frac{\mathcal{D}_n}{\sqrt{\mathrm{Var}(J_{i_G})}}, \tag{4.33}$$

where $Q(n)$ is the sum of the re-scaled independent increments, J_{i_G}, and hence, in continuous time, $Q(t)$ can be approximated by a Brownian motion process. Thus, the detection delay is equivalent to the first passage time for Brownian motion with drift. The expected value and variance of $Q(n)$ are given respectively as

$$E\left[Q(n)\right] = \frac{nE[J_{i_G}]}{\sqrt{\mathrm{Var}(J_{i_G})}}, \tag{4.34}$$

$$\mathrm{Var}\left(Q(n)\right) = n. \tag{4.35}$$

The re-scaling (division by $\sqrt{\mathrm{Var}(J_{i_G})}$) is chosen such that $Q(t)$ has the same mean and variance as the Brownian motion process with drift μ, $\overline{W}_t$, defined as [115]

$$\overline{W}_t = W_t + \mu t, \tag{4.36}$$

where $0 \leq t < \infty$ and $W_t \sim \mathcal{N}(0, t)$. For this process, the expected value and variance of $\overline{W}_t$ can be written respectively as

$$E[\overline{W}_t] = \mu t, \tag{4.37}$$

$$\mathrm{Var}(\overline{W}_t) = t. \tag{4.38}$$

It can be observed from (4.34) and (4.37) that the drift, μ, can be expressed as

$$\mu = \frac{E[J_{i_G}]}{\sqrt{\mathrm{Var}(J_{i_G})}}. \tag{4.39}$$

As a result of the re-scaling process, the threshold is changed to

$$\psi = \frac{\gamma}{\sqrt{\mathrm{Var}(J_{i_G})}}. \tag{4.40}$$

We can then write the approximate expression for $P(\delta)$ as [115]

$$\widetilde{P}_{bm_G}(\delta) = \frac{|\psi|}{\sqrt{2\pi q^3}} e^{-\frac{(\psi-\mu q)^2}{2q}}\, dq, \tag{4.41}$$

where $q = \delta + \frac{1}{2}$ and we set $dq = 1$. The subscript, bm_G, in $\widetilde{P}_{bm_G}(\delta)$ denotes the Brownian motion approximation for the Gaussian channel.

4.3.2 Rayleigh channel

In this section, we consider a mis-matched channel, in which the received signal detected using a time-invariant detector is transmitted over a Rayleigh channel. For this Rayleigh scenario, $\mathcal{D}_n$ in (4.4) and $\mathcal{D}_\omega$ in (4.5) will be used in deriving the approximate expressions for $P(\delta)$.

4.3.2.1 Approximation

In this case, $\mathcal{S}_\delta$ in (4.21) becomes $\mathcal{S}_\delta = \sum_{i=1}^{\delta} \widetilde{Z}_i$. The distribution of $\widetilde{Z}_i$ is more complex than in Section 4.3.1 (as shown in Section 3.4.2) and the distribution of $\mathcal{S}_\delta$ is unknown. Hence, in order to use (4.21) as an approximation method, we approximate $\widetilde{Z}_i$ by a gamma variable and this leads to a gamma approximation for $\mathcal{S}_\delta$ also. The motivation for this approach lies in the fact that for virtually all mis-matched channels, $\widetilde{Z}_i$, will remain positive, unimodal and positively skewed, as is a gamma variable. Furthermore, as δ becomes large, both the gamma approximation and the $\mathcal{S}_\delta$ variable converge to a Gaussian limit. The gamma approximations are simply obtained by the method of moments. Hence, we can approximate $\widetilde{Z}_i$ by a gamma variable with mean, $\mu_{\widetilde{Z}_i}$ and variance, $\sigma^2_{\widetilde{Z}_i}$ defined respectively as

$$\mu_{\widetilde{Z}_i} = \sigma_X^2 + \sigma_N^2, \tag{4.42}$$

$$\begin{aligned} \sigma^2_{\widetilde{Z}_i} &= E[\widetilde{Z}_i^2] - E[\widetilde{Z}_i]^2 \\ &= (4\sigma_X^4 + 4\sigma_X^2\sigma_N^2 + 2\sigma_N^4) - (\sigma_X^2 + \sigma_N^2)^2 \\ &= 3\sigma_X^4 + 2\sigma_X^2\sigma_N^2 + \sigma_N^4. \end{aligned} \tag{4.43}$$

Equations (4.42) and (4.43) follow by using known results on the moments of $|H[i]|^2$, which has a standard exponential distribution. Similarly, the summation term in (4.4) can also be approximated by a gamma variable giving

$$\sum_{i=1}^{\delta+1} \widetilde{Z}_i \sim \mathcal{G}\left(\frac{(\delta+1)\mu^2_{\widetilde{Z}_i}}{\sigma^2_{\widetilde{Z}_i}}, \frac{\sigma^2_{\widetilde{Z}_i}}{\mu_{\widetilde{Z}_i}}\right), \tag{4.44}$$

where $\mathcal{G}(k,\theta)$ is a gamma distribution with shape parameter k and scale parameter θ. The derivation of the approximate expression for $P(\delta)$ in the Rayleigh case follows the approach taken in Section 4.3.1.2 using (4.5) along with (4.44). Hence, $P(\delta)$ can be approximated by

$$\widetilde{P}_{a_{Ray}}(\delta) = \int_0^{\frac{\gamma+\delta b}{\bar{a}}} f_{\mathcal{S}_\delta}(z) P\left(\widetilde{Z}_{\delta+1} \geq \frac{\gamma+(\delta+1)b}{\bar{a}} - z\right) dz, \tag{4.45}$$

where

$$\mathcal{S}_\delta \sim \mathcal{G}\left(\frac{\delta\mu^2_{\widetilde{Z}_i}}{\sigma^2_{\widetilde{Z}_i}}, \frac{\sigma^2_{\widetilde{Z}_i}}{\mu_{\widetilde{Z}_i}}\right), \tag{4.46}$$

and

$$\widetilde{Z}_{\delta+1} \sim \mathcal{G}\left(\frac{\mu^2_{\widetilde{Z}_i}}{\sigma^2_{\widetilde{Z}_i}}, \frac{\sigma^2_{\widetilde{Z}_i}}{\mu_{\widetilde{Z}_i}}\right). \tag{4.47}$$

Let $\alpha_1 = \frac{\delta\mu^2_{\widetilde{Z}_i}}{\sigma^2_{\widetilde{Z}_i}}$, $\alpha_2 = \frac{\mu^2_{\widetilde{Z}_i}}{\sigma^2_{\widetilde{Z}_i}}$ and $\beta = \frac{\sigma^2_{\widetilde{Z}_i}}{\mu_{\widetilde{Z}_i}}$, then $f_{\mathcal{S}_\delta}(z)$ in (4.45) is the pdf of a gamma distribution defined as

$$f_{\mathcal{S}_\delta}(z) = \frac{z^{\alpha_1-1}e^{-z/\beta}}{\Gamma(\alpha_1)\beta^{\alpha_1}}. \tag{4.48}$$

We first evaluate the second term in the integral expression in (4.45), which gives

$$\begin{aligned} P\left(\widetilde{Z}_{\delta+1} \geq \frac{\gamma+(\delta+1)b}{\bar{a}} - z\right) &= \int_{\frac{\gamma+(\delta+1)b}{\bar{a}}-z}^{\infty} f_{\widetilde{Z}_{\delta+1}}(u)\, du \\ &= \frac{1}{\Gamma(\alpha_2)\beta^{\alpha_2}} \int_{\frac{\gamma+(\delta+1)b}{\bar{a}}-z}^{\infty} u^{\alpha_2-1}e^{-u/\beta}\, du \\ &= \frac{1}{\Gamma(\alpha_2)}\Gamma\left(\alpha_2, \frac{\frac{\gamma+(\delta+1)b}{\bar{a}}-z}{\beta}\right), \end{aligned} \tag{4.49}$$

where $\Gamma(.,.)$ is the incomplete Gamma function. Inserting (4.48) and (4.49) into (4.45) yields

$$\widetilde{P}_{a_{Ray}}(\delta) = \frac{1}{\beta^{\alpha_1}\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^{\frac{\gamma+\delta b}{\bar{a}}} z^{\alpha_1-1}e^{-z/\beta}\Gamma\left(\alpha_2, \frac{\frac{\gamma+(\delta+1)b}{\bar{a}}-z}{\beta}\right) dz. \tag{4.50}$$

Thus, $\widetilde{P}_{a_{Ray}}(\delta)$ has a simple form which requires only a single numerical integration. To increase accuracy, a reasonable approach would be to approximate $\mathcal{S}_\delta$ by a generalized gamma random variable and fit the generalized gamma approximation by using the method of moments. However, in order to approximate $P(\delta)$, we need to re-estimate all three parameters of the generalized gamma for each δ. Since the method of moments requires an iterative solution and this must be recomputed for every δ, this approach is computationally expensive. Preliminary results indicate that the increased accuracy is not sufficient to justify the complexity.

4.3.2.2 Approximation based on a random walk approach

Following the approach taken in Section 4.3.1.3, we can also approximate $P(\delta)$ for the Rayleigh scenario using random walk theory. From (4.4), the jump in $\mathcal{D}_n$ can be defined as

$$J_{i_{Ray}} = \bar{a}\widetilde{Z}_i - b, \tag{4.51}$$

where the expected value and variance of $J_{i_{Ray}}$ can be expressed respectively as

$$E[J_{i_{Ray}}] = \bar{a}\mu_{\widetilde{Z}_i} - b, \tag{4.52}$$

$$\mathrm{Var}(J_{i_{Ray}}) = \bar{a}^2\sigma^2_{\widetilde{Z}_i}, \tag{4.53}$$

in which $\mu_{\widetilde{Z}_i}$ and $\sigma^2_{\widetilde{Z}_i}$ are given in (4.42) and (4.43). Equating the means and variances of $J_{i_{Ray}}$ in (4.52), (4.53) and J_{eq} in (4.27) and (4.28) gives

$$D_{Ray} = \frac{\bar{a}\mu_{\widetilde{Z}_i} - b}{2q-1}, \tag{4.54}$$

$$q_{Ray} = \frac{1}{2} + \frac{\sqrt{\vartheta - 4\bar{a}^2\sigma^2_{\widetilde{Z}_i}}}{2\sqrt{\vartheta}}, \tag{4.55}$$

in which

$$\vartheta = 4\bar{a}^2\sigma^2_{\widetilde{Z}_i} + 4\bar{a}^2\mu^2_{\widetilde{Z}_i} - 8\bar{a}\mu_{\widetilde{Z}_i}b + 4b^2. \tag{4.56}$$

It is worth noting that the subscripts in (4.54) and (4.56) are used to differentiate between the Gaussian and Rayleigh case. The number of steps needed to reach γ can be written as $W = \gamma/D_{Ray}$. Hence, we can express the approximate expression for the distribution of detection delay for the Rayleigh case as [111, 114]

$$\widetilde{P}_{rw_{Ray}}(\delta) = \frac{W\Gamma(\delta+2)}{(\delta+1)\Gamma(\frac{\delta-W}{2}+\frac{3}{2})\Gamma(\frac{\delta+W}{2}+\frac{3}{2})} p_{Ray}^{\frac{\delta+1-W}{2}} q_{Ray}^{\frac{\delta+1+W}{2}}, \quad \delta \geq W-1 \tag{4.57}$$

where $p_{Ray} = 1 - q_{Ray}$.

4.3.2.3 Approximation based on a Brownian motion approach

We can also follow the approach taken in Section 4.3.1.4 to approximate $P(\delta)$ in the Rayleigh scenario by employing the theory of Brownian motion with drift. In this case, we re-scale $\mathcal{D}_n$ in (4.4) to yield

$$Q_{Ray}(n) = \frac{\mathcal{D}_n}{\sqrt{\mathrm{Var}(J_{i_{Ray}})}}. \tag{4.58}$$

The expected value and variance of $Q_{Ray}(n)$ can be written respectively as

$$E\left[Q_{Ray}(n)\right] = \frac{nE[J_{i_{Ray}}]}{\sqrt{\mathrm{Var}(J_{i_{Ray}})}}, \tag{4.59}$$

$$\mathrm{Var}\left(Q_{Ray}(n)\right) = n. \tag{4.60}$$

The drift, μ_{Ray}, can be obtained by comparing the means in (4.59) and (4.37), which gives

$$\mu_{Ray} = \frac{E[J_{i_{Ray}}]}{\sqrt{\mathrm{Var}(J_{i_{Ray}})}}. \tag{4.61}$$

Due to the re-scaling of $\mathcal{D}_n$, the threshold now becomes

$$\psi_{Ray} = \frac{\gamma}{\sqrt{\mathrm{Var}(J_{i_{Ray}})}}. \tag{4.62}$$

Thus, the approximate expression for $P(\delta)$ can be expressed as [115]

$$\widetilde{P}_{bm_{Ray}}(\delta) = \frac{|\psi_{Ray}|}{\sqrt{2\pi q^3}} e^{-\frac{(\psi_{Ray}-\mu_{Ray} q)^2}{2q}}\, dq, \tag{4.63}$$

where $q = \delta + \frac{1}{2}$ and $dq = 1$.

4.4 NUMERICAL ANALYSIS AND RESULTS

In this section, we provide some further analysis of the modified detection delay statistic for the Gaussian case and present some numerical results to validate the approximate expressions for the distribution of detection delay for a time-invariant CUSUM detector when the received signals experience Gaussian and Rayleigh channels. We also theoretically approximate the probability of missed detection and provide a detailed analysis of the likelihood of long detection delays. In all simulations, the threshold is set to be $\gamma = 3$, following from Section 3.5 and each point on the plots is generated using 20000 trials, where each trial consists of 200 samples. For $i < 100$ samples, the CUSUM detector receives only noise. However, at $\tau = 100$, the PU starts transmitting and the received signal at the detector consists of the faded signal plus noise. Detection delay is measured from when the PU becomes active, at τ, until detection by the CUSUM detector.

4.4.1 Analysis of the $\mathcal{D}_n$ process: Gaussian case

The $\mathcal{D}_n$ process only differs from the true CUSUM process if it becomes negative before it crosses the threshold, γ. The probability of this event occurring for the first time at sample n can be defined as

$$\begin{aligned} &P(0 < \mathcal{D}_1 < \gamma, 0 < \mathcal{D}_2 < \gamma, ..., 0 < \mathcal{D}_{n-1} < \gamma, \mathcal{D}_n < 0) \\ &\leq P(\mathcal{D}_1 > 0, \mathcal{D}_2 > 0, ..., \mathcal{D}_{n-1} > 0, \mathcal{D}_n < 0) = P_{neg}(n). \end{aligned} \tag{4.64}$$

The upper bound, $P_{neg}(n)$, can be evaluated for small values of n using standard techniques. This gives

$$\begin{aligned} P_{neg}(1) &= P(\mathcal{D}_1 < 0) \\ &= P(Z_1 < \frac{b}{a}) \\ &= 1 - e^{-\frac{b}{a}}. \end{aligned} \tag{4.65}$$

$$\begin{aligned} P_{neg}(2) &= P(\mathcal{D}_1 > 0, \mathcal{D}_2 < 0) \\ &= P\left(Z_1 > \frac{b}{a}, Z_2 < \frac{2b}{a} - Z_1\right) \\ &= \int_{z_1=\frac{b}{a}}^{\frac{2b}{a}} e^{-z_1} P\left(Z_2 < \frac{2b}{a} - z_1\right) dz_1 \\ &= e^{-\frac{b}{a}} - \left(1 + \frac{b}{a}\right) e^{-\frac{2b}{a}}. \end{aligned} \tag{4.66}$$

$$\begin{aligned}
P_{neg}(3) &= P(\mathcal{D}_1 > 0, \mathcal{D}_2 > 0, \mathcal{D}_3 < 0) \\
&= P\left(Z_1 > \frac{b}{a}, Z_2 > \frac{2b}{a} - Z_1, Z_3 < \frac{3b}{a} - Z_1 - Z_2\right) \\
&= \int_{z_1=\frac{b}{a}}^{\frac{2b}{a}} e^{-z_1} P\left(Z_2 > \frac{2b}{a} - z_1, Z_3 < \frac{3b}{a} - z_1 - z_2\right) dz_1 + \int_{z_1=\frac{2b}{a}}^{\frac{3b}{a}} e^{-z_1} \times \\
&\quad P\left(Z_2 > 0, Z_3 < \frac{3b}{a} - z_1 - z_2\right) dz_1 \\
&= \left(1 + \frac{b}{a}\right) e^{-\frac{2b}{a}} - \left(1 + \frac{2b}{a} + \frac{3b^2}{2a^2}\right) e^{-\frac{3b}{a}}.
\end{aligned} \tag{4.67}$$

$$\begin{aligned}
P_{neg}(4) &= P(\mathcal{D}_1 > 0, \mathcal{D}_2 > 0, \mathcal{D}_3 > 0, \mathcal{D}_4 < 0) \\
&= P\left(Z_1 > \frac{b}{a}, Z_2 > \frac{2b}{a} - Z_1, Z_3 > \frac{3b}{a} - Z_1 - Z_2, Z_4 < \frac{4b}{a} - Z_1 - Z_2 - Z_3\right) \\
&= \int_{z_1=\frac{b}{a}}^{\frac{2b}{a}} e^{-z_1} P\left(Z_2 > \frac{2b}{a} - z_1, Z_3 > \frac{3b}{a} - z_1 - z_2, Z_4 < \frac{4b}{a} - z_1 - z_2 - z_3\right) dz_1 + \\
&\quad \int_{z_1=\frac{2b}{a}}^{\frac{3b}{a}} e^{-z_1} P\left(Z_2 > 0, Z_3 > \frac{3b}{a} - z_1 - z_2, Z_4 < \frac{4b}{a} - z_1 - z_2 - z_3\right) dz_1 + \\
&\quad \int_{z_1=\frac{3b}{a}}^{\frac{4b}{a}} e^{-z_1} P\left(Z_2 > 0, Z_3 > 0, Z_4 < \frac{4b}{a} - z_1 - z_2 - z_3\right) dz_1 \\
&= \left(1 + \frac{2b}{a} + \frac{3b^2}{2a^2}\right) e^{-\frac{3b}{a}} - \left(1 + \frac{3b}{a} + \frac{4b^2}{a^2} + \frac{8b^3}{3a^3}\right) e^{-\frac{4b}{a}}.
\end{aligned} \tag{4.68}$$

There is a clear pattern in the $P_{neg}(n)$ results in (4.65)-(4.68), but the precise form for the polynomial terms in b/a appears to be difficult to obtain. Nevertheless, the values of $P_{neg}(n)$ for $n = 1, 2, 3$ and 4 give us some insight into the limitations of the approximation. In Table 4.1, we compare the theoretical and simulation results for $P_{neg}(n)$ at SNR=5, 0 and -5 dB. We can see that the theoretical results are validated by the simulations. We also observe that the $\mathcal{D}_n$ process has the highest probability of going negative in the first sample compared to the other samples. Due to the small value of $P_{neg}(n)$ for $n \geq 2$, this indicates that the approximation approaches are most likely to be in error in the first sample (i.e. $\delta = 0$).

Although $P_{neg}(1)$ is substantial, the size of the negative jump tends to be small, so that the approximation remains reasonable as shown in Figures 4.2 and 4.3. It is possible to adjust for the probability of $\mathcal{D}_n$ becoming negative by using a re-scaled threshold. For example, the

Table 4.1 Results for $P_{neg}(n)$ at SNR=5, 0 and -5 dB

n	Theoretical			Simulation		
	5 dB	0 dB	-5 dB	5 dB	0 dB	-5 dB
1	0.3630	0.5	0.5806	0.3676	0.5031	0.5807
2	0.0482	0.0767	0.0907	0.0483	0.0761	0.0910
3	0.0183	0.0349	0.0432	0.0183	0.0347	0.0431
4	0.0088	0.0203	0.0264	0.0098	0.02	0.0255

threshold could be reduced by the mean negative excursion of the $\mathcal{D}_n$ process. Therefore, the re-scaled threshold can be written as

$$h = \gamma + \varsigma, \qquad \varsigma < 0, \tag{4.69}$$

where $\varsigma = P_{neg}(1) \times \mathbf{E}[aZ_i - b|aZ_i - b < 0]$, in which $\mathbf{E}[aZ_i - b|aZ_i - b < 0]$ is the mean negative excursion of the $\mathcal{D}_n$ process, defined as

$$\begin{aligned}
\mathbf{E}[aZ_i - b|aZ_i - b < 0] &= a\left[\int_0^{\frac{b}{a}} \frac{z_i f(z_i)}{P(Z_i < \frac{b}{a})}\, dz_i\right] - b \\
&= \frac{a}{1 - e^{-\frac{b}{a}}} \int_0^{\frac{b}{a}} z_i e^{-z_i}\, dz_i - b \\
&= \frac{a\left(1 - \left(1 + \frac{b}{a}\right) e^{-\frac{b}{a}}\right)}{1 - e^{-\frac{b}{a}}} - b.
\end{aligned} \tag{4.70}$$

Figure 4.1 compares the simulated results with the approximate expressions for $P(\delta)$ based on the original threshold γ and the re-scaled threshold, h, when the received signals experience a Gaussian channel at SNR=5 dB. As can be seen from Figure 4.1, results obtained using the re-scaled threshold do not provide much improvement over the standard approximation using γ. Hence, in Section 4.4.2, we present numerical results for the distribution of detection delay based on the threshold γ, and do not consider the re-scaling method any further.

4.4.2 Results

Figure 4.2 compares simulated results with the approximate expressions for $P(\delta)$ when the received signals experience a Gaussian channel at SNR=5, 0 and -5 dB. Based on Figure 4.2, we observe that the closed-form expression for $\widetilde{P}_c(\delta)$ is close to the simulated results at SNR=5

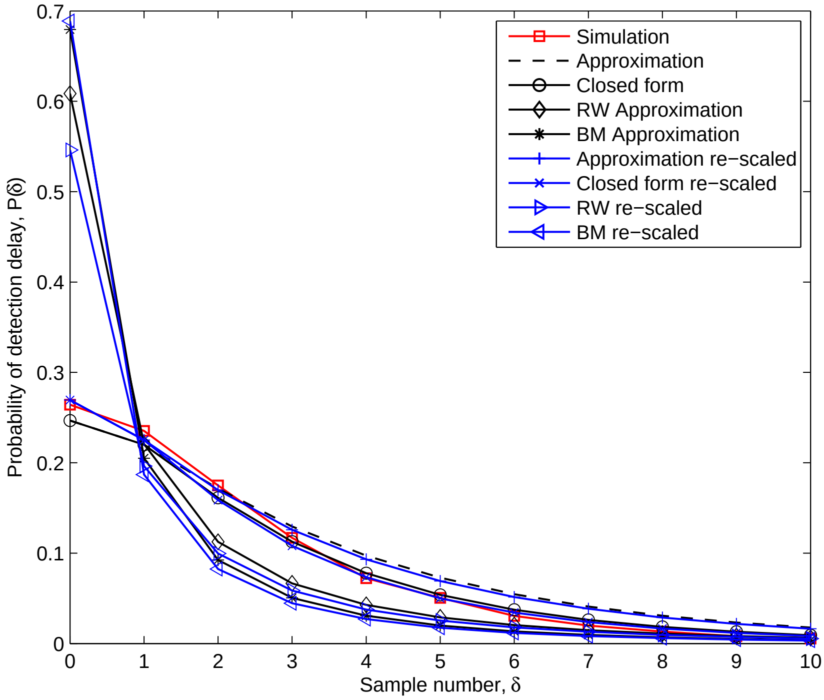

Figure 4.1 A comparison of the simulated results and the approximate distribution of detection delay based on threshold γ and the re-scaled threshold, h, for the Gaussian scenario at SNR=5 dB.

and 0 dB. However, at SNR=-5 dB, $\widetilde{P}_c(\delta)$ is only accurate for long delays. We can see that at all three SNRs considered, the approximate expression, $\widetilde{P}_{a_G}(\delta)$ behaves as an upper bound on the closed-form expression for $\widetilde{P}_c(\delta)$ and the approximation is reasonably good at SNR=5 and 0 dB. However, the estimation error is high at SNR=-5 dB. In Figure 4.2, we see that at all SNRs, the approximate expression based on the random walk (RW) approach, $\widetilde{P}_{rw_G}(\delta)$ provides a slightly better approximation for long detection delays as compared to $\widetilde{P}_{a_G}(\delta)$. At SNR=0 dB, the Brownian motion (BM) approximation, $\widetilde{P}_{bm_G}(\delta)$ provides a good estimate of $P(\delta)$ for long delays. At SNR=-5 dB, $\widetilde{P}_{bm_G}(\delta)$ gives the best approximation for $P(\delta)$ among all the other approximation methods.

Figure 4.3 shows a comparison of the simulated results and the approximate expressions for $P(\delta)$ when a mis-matched channel occurs, i.e the received signal is transmitted over a Rayleigh channel. We observe from Figure 4.3 that the approximate expression, $\widetilde{P}_{a_{Ray}}(\delta)$, provides a good estimate of $P(\delta)$ at SNR=5 dB compared to the random walk approximation, $\widetilde{P}_{rw_{Ray}}(\delta)$ and the Brownian motion approximation, $\widetilde{P}_{bm_{Ray}}(\delta)$. At SNR=0 dB, $\widetilde{P}_{a_{Ray}}(\delta)$ gives an accurate

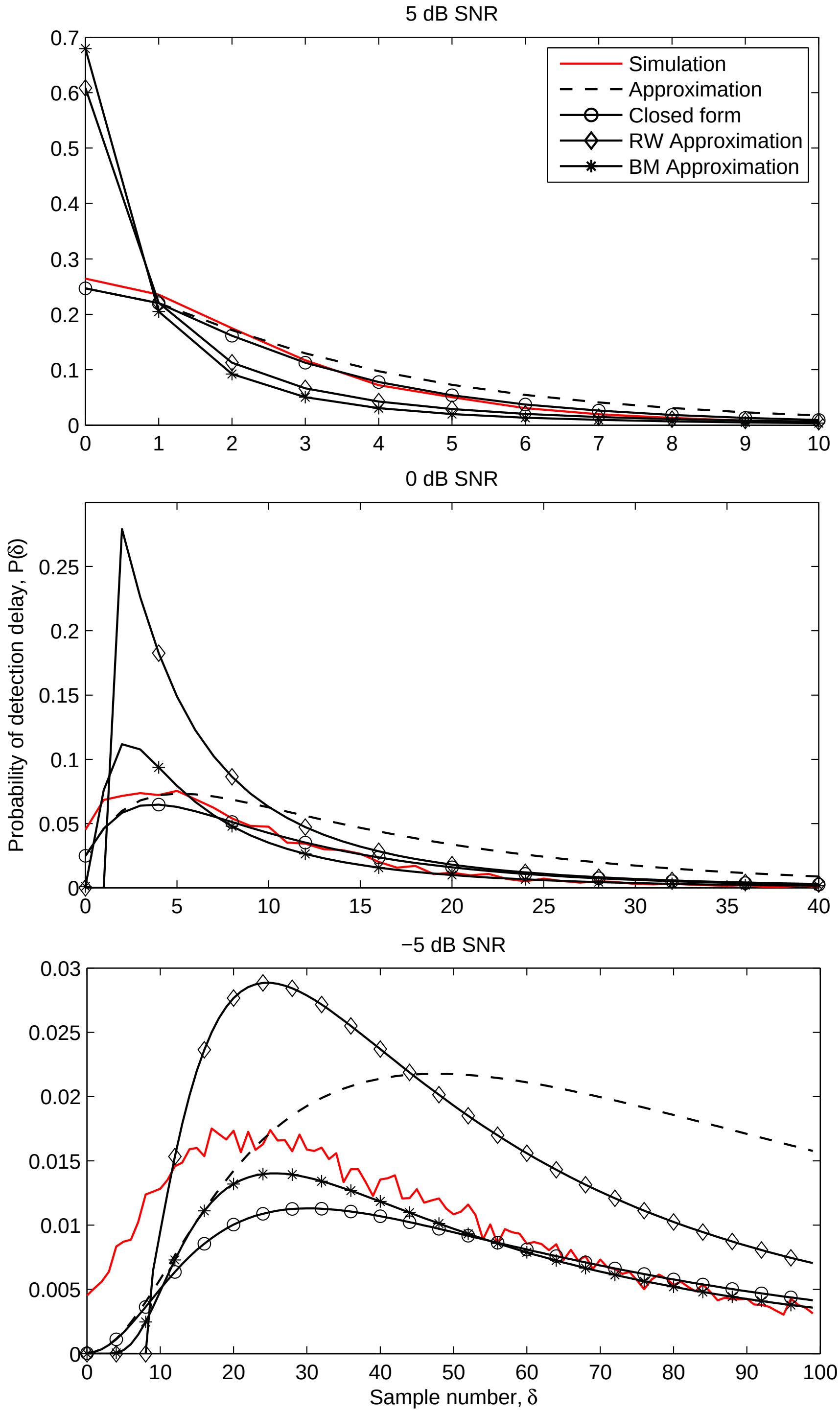

Figure 4.2 A comparison of the simulated results and the approximate distribution of detection delay for the Gaussian scenario at SNR=5,0 and -5 dB

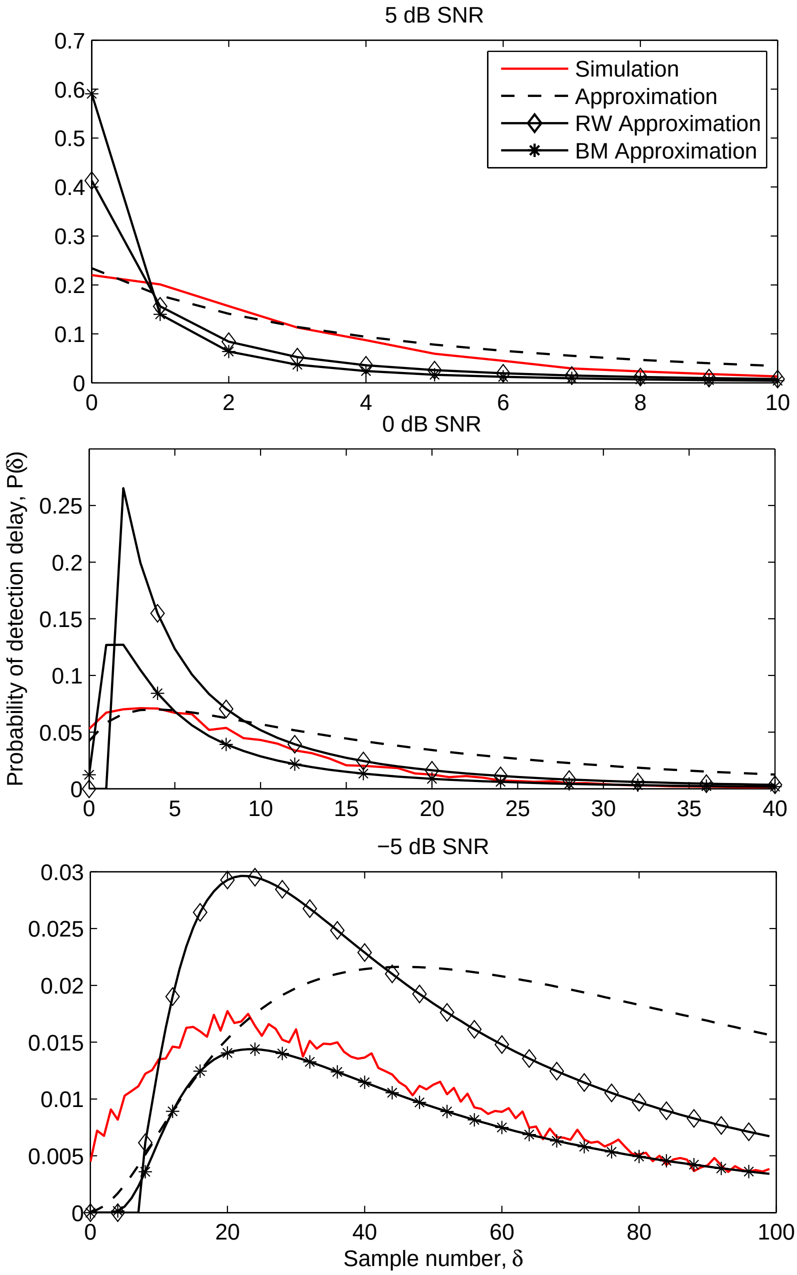

Figure 4.3 A comparison of the simulated results and the approximate distribution of detection delay for the Rayleigh scenario at SNR=5,0 and -5 dB. This is a mis-matched channel scenario.

approximation for short delays whereas $\widetilde{P}_{rw_{Ray}}(\delta)$ and $\widetilde{P}_{bm_{Ray}}(\delta)$ are close to the simulated results at long delays. However, at SNR=-5 dB, $\widetilde{P}_{bm_{Ray}}(\delta)$ provides a good approximation, especially at long delays, as compared to the other approximation methods.

In short, different SNR and detection delay conditions require different approaches in order to achieve good approximations of the distribution of detection delay. Furthermore, the remarkably simple Brownian motion approach provides the best approximation for longer delays.

4.4.3 Probability of missed detection

It is known that the CUSUM detector will continuously sense the frequency spectrum until it detects the PU's transmission. In the event that the CU is sensing the existence of the PU, the CU will declare that the PU is absent at sample r, if detection has not occurred after a certain period of time, even if the PU is active. In other words, the CU fails to detect the existence of the PU leading to a missed detection. We denote r as the number of samples taken during which no PU is detected by the CU and after which the CU decides to transmit. We can approximate the probability of missed detection based on $\widetilde{P}_{bm_G}(\delta)$, since the Brownian motion approach gives a good approximation of the distribution of detection delay for longer delays. We define the approximate probability of missed detection by

$$\begin{aligned}\widetilde{P}_{md}(r) &= \sum_{\delta \geq r} \widetilde{P}_{bm_G}(\delta) \\ &= \sum_{\delta \geq r} \frac{|\psi|}{\sqrt{2\pi q^3}} e^{-\frac{(\psi-\mu q)^2}{2q}}\, dq,\end{aligned} \tag{4.71}$$

where r is the sample number at which the CU declares the absence of the PU. ψ, μ, q and dq in (4.71) are defined in Section 4.3.1.4.

The results for $\widetilde{P}_{md}(r)$ at various SNR are shown in Table 4.2. Based on Table 4.2, it can be seen that there is a low probability of missed detection at SNR=5 dB as compared to the other SNRs, with the highest $\widetilde{P}_{md}(r)$ observed at SNR=-5 dB. This is because as the received signal gets weaker, there is a higher probability that the detection delay will be longer and this leads to a higher $\widetilde{P}_{md}(r)$. The remarkably simple Brownian motion approximation enables a very rapid quantification of these missed detection probabilities. It is worth noting that for small values of r, $\widetilde{P}_c(\delta)$ in (4.13) or $\widetilde{P}_a(\delta)$ in (4.23) could be used in deriving the approximate expression for the probability of missed detection as both of the approximate methods provide a good approximation of the detection delay distribution for small detection delays.

Table 4.2 Missed detection probabilities, $\widetilde{P}_{md}(r)$ at SNR=5, 0 and -5 dB

r	5 dB	0 dB	-5 dB
20	5.5062×10^{-4}	0.1101	0.6817
40	0	0.0197	0.4143
60	0	0.0033	0.2169
80	0	0	0.0867
100	0	0	0

4.4.4 Analysis of long detection delays

Here, we extend our investigation to general i.i.d channels and analyze the likelihood of long detection delays for the general case. In both Figures 4.2 and 4.3, the Brownian motion approach gives a reasonably good estimate of $P(\delta)$ for longer detection delays. Therefore, the evaluation of long detection delays will be based on the Brownian motion approach. In order to analyze the likelihood of long detection delays, we derive a general approximate expression for $P(\delta)$ based on the Brownian motion approach, where this general expression can be applied to any i.i.d channel. For this general case, each jump in $\mathcal{D}_n$ can be written as

$$J_i = aK_i - b, \tag{4.72}$$

where $K_i = \frac{|Y[i]|^2}{E[|Y[i]|^2]}$, $a = \sigma_X^2/\sigma_N^2$ and $b = -\ln\left\{\frac{\sigma_N^2}{\sigma_N^2+\sigma_X^2}\right\}$. The expected value and variance of J_i can be written respectively as

$$E[J_i] = a - b, \qquad \text{Var}(J_i) = a^2\sigma_{K_i}^2, \tag{4.73}$$

where

$$\begin{aligned}\sigma_{K_i}^2 &= E[K_i^2] - E[K_i]^2 \\ &= \left(\frac{E\left[|X|^4\right] - 2\sigma_X^4}{\sigma_X^4 + 2\sigma_X^2\sigma_N^2 + \sigma_N^4} + 2\right) - (1)^2 \\ &= \frac{E\left[|X|^4\right] - 2\sigma_X^4}{\sigma_X^4 + 2\sigma_X^2\sigma_N^2 + \sigma_N^4} + 1.\end{aligned} \tag{4.74}$$

Based on Section 4.3.1.4 and 4.3.2.3, we define a general approximate expression for $P(\delta)$, which can be expressed as

$$\widetilde{P}_{bm_{Gen}}(\delta) = \frac{|\psi_{Gen}|}{\sqrt{2\pi q^3}} e^{-\frac{(\psi_{Gen}-\mu_{Gen}q)^2}{2q}}\, dq, \tag{4.75}$$

where $q = \delta + \frac{1}{2}$ and $dq = 1$. The general drift, μ_{Gen} and new threshold, ψ_{Gen} can be written using (4.73) and (4.74) to yield

$$\begin{aligned} \mu_{Gen} &= \frac{E[J_i]}{\sqrt{\mathrm{Var}(J_i)}} \\ &= \frac{1 + \xi \ln\left\{\frac{\xi}{1+\xi}\right\}}{\sigma_{K_i}}, \end{aligned} \tag{4.76}$$

$$\begin{aligned} \psi_{Gen} &= \frac{\gamma}{\sqrt{\mathrm{Var}(J_i)}} \\ &= \frac{\gamma\xi}{\sigma_{K_i}}, \end{aligned} \tag{4.77}$$

in which $\xi = \sigma_N^2/\sigma_X^2$. Substituting (4.76) and (4.77) into (4.75), $\widetilde{P}_{bm_{Gen}}(\delta)$ can be re-written as

$$\widetilde{P}_{bm_{Gen}}(\delta) = \frac{|\frac{\gamma\xi}{\sigma_{K_i}}|}{\sqrt{2\pi q^3}} e^{-\frac{\left(\frac{\gamma\xi - Rq}{\sigma_{K_i}}\right)^2}{2q}}\, dq, \tag{4.78}$$

where

$$R = 1 + \xi \ln\left\{\frac{\xi}{1+\xi}\right\}. \tag{4.79}$$

Note that (4.78) is valid for an arbitrary channel. As $\delta \to \infty$, $q \to \infty$ and hence, (4.78) becomes

$$\widetilde{P}_{bm_{Gen}}(\delta) = \gamma\xi \times \frac{e^{-\frac{(R\sqrt{q})^2}{2\sigma_{K_i}^2}}}{\sqrt{2\pi\sigma_{K_i}^2}} \times \frac{dq}{q^{3/2}}. \tag{4.80}$$

The form of (4.80) allows us to consider the factors which lead to high probabilities of long delays. The first term in (4.80) is $\gamma\xi$ so that we have the simple conclusion that high thresholds (large γ) or low SNR (high ξ) lead to longer delays. The second term is in the form of a Gaussian density and for large q, this is increased by large values of σ_{K_i} and small values of R. For R, it is straight forward to show that low SNR (high ξ) leads to smaller R values, so again low SNR is a factor. Hence, the only factor that relates to the actual channel is σ_{K_i}. Using (4.74), in order for $\sigma_{K_i}^2$ to be large, we require

$$E[|X|^4] \gg 3\sigma_X^4. \tag{4.81}$$

Let $M = |X|^2$, then (4.81) can be re-written as

$$\frac{E[M^2] - E[M]^2}{E[M]^2} \gg 2. \tag{4.82}$$

The left-hand side of the inequality can be written in terms of the coefficient of variation (CV), which is discussed in Section 2.1.6. CV is usually used to measure the severity of the fading [62]. This then yields

$$\text{CV} \gg \sqrt{2}. \tag{4.83}$$

From Table 3.2 in Section 3.5, most of the classical channel models including Gaussian, Rayleigh, Rician and Nakagami-m channels have a small value of CV, with the Rayleigh having the highest CV value of 0.52. In contrast, the long-tailed F distribution, which models a severe fading channel, has a CV of 1.21. Consider another type of long-tailed distribution, the log-normal distribution, where the CV is defined as $\text{CV} = \sqrt{e^{\sigma^2} - 1}$ [117]. In order to satisfy (4.83), $\sigma^2 \gg \log_e 3$. Therefore, we observe that in order for the channel to create long delays, it needs an extremely high CV value that can result from severe fading channels (log-normal shadow fading or the F channel), but does not occur with the traditional channel models (Rayleigh, Rician and Nakagami-m fading channels). Hence, we conclude that the channel has very little impact on long detection delays unless it experiences unusually severe fading.

It is worth noting that the detection delay is mainly a function of the first two moments of the received signal (i.e. mean and variance). This is because the Brownian motion approach, which works well in approximating the distribution of detection delay for long detection delay is based on matching of two moments. In addition, our work on investigating the performance of quickest spectrum sensing with a single receive antenna over various fading channels in Chapter 3 shows that CV affects the sensing performance. Again, CV is a function of the first two moments and these first two moments are not wildly different for most channel models. Furthermore, in this thesis, we compare channel models with the same power and hence, one of the moments is exactly the same. Therefore, the performance of the quickest spectrum sensing is insensitive to the designed detector.

4.5 CHAPTER SUMMARY

In this chapter, we derived approximate expressions for the detection delay distribution for a time-invariant CUSUM detector when the received signal is transmitted over a Gaussian channel as well as over a mis-matched, Rayleigh channel. In particular, for the Gaussian case, we derived a novel approximate closed-form expression for the distribution of detection delay. In addition, we also derived novel approximate expressions for the detection delay distribution for the general case due to the absence of a general framework. We also applied the simple random walk and Brownian motion theory with drift to derive the approximate expressions for the detection delay distribution for both Gaussian and Rayleigh cases. Most of the approximate expressions that we formulated are general and can be applied to any i.i.d channel.

Numerical results illustrate that in order to achieve good approximations, different approximate methods are needed for different SNR and detection delay conditions. In particular, at high SNR, the closed-form and approximation approaches provide good approximations for the Gaussian and Rayleigh cases. At moderate SNR, both closed-form and approximation approaches provide good approximation for short delays, but the random walk and Brownian motion approaches give better approximations for long delays for both of the cases considered. In addition, at low SNR, the Brownian motion approach gives the best approximation among all the other methods. Furthermore, of all the approximations, the Brownian motion approach provides a remarkably simple and accurate approximation for longer delays. The analysis of long detection delays shows that an increased probability of a long detection delay can be obtained if the threshold value is large or the received signal is weak due to low SNR. However, long detection delays are normally insensitive to the type of fading channel. In the next chapter, we will investigate the quickest spectrum sensing performance when multiple antennas are employed at the receiver.

Chapter 5

EXTENSION OF QUICKEST SPECTRUM SENSING TO MULTIPLE ANTENNAS

5.1 INTRODUCTION

In Chapters 3 and 4, we studied and analyzed the performance of quickest spectrum sensing, where the CU is equipped with a single antenna. There have been a number of studies on spectrum sensing using an energy detector and a GLRT detector employing multiple receive antennas [118–122]. However, no studies using quickest spectrum sensing consider CUs with multiple antennas. Therefore, in this chapter, we investigate quickest spectrum sensing performance with multiple antenna CUs when the received signal is transmitted over Gaussian, Rayleigh and Rician channels.

Slowly varying channels are modeled by a time-invariant channel gain so that a Gaussian signal in noise gives an overall Gaussian received signal. Fast fading channels are modeled by a Rayleigh or a Rician channel so that the received signal is the product of two complex Gaussian variables (channel and signal) with additive noise. We prove that the sum of the complex received signal powers at each antenna for the independent Rayleigh scenario is a sufficient statistic. This result allows us to derive the pdf of the received signal experiencing independent Rayleigh channel based on the sum of the received signal powers. Hence, we can employ an EGC before applying standard CUSUM sensing techniques (as discussed in Section 2.4.1). The pdf of the received signal is required to form the optimal CUSUM detector. The derivation of the pdf for the independent Rayleigh scenario uses a technique which avoids numerical integration. The benefits of using this technique will be discussed in Section 5.5.2. Besides the independent Rayleigh scenario, we also employ an EGC for the Gaussian case as the sum of the complex received signal powers at each antenna in this case is also a sufficient statistic.

We also consider the case of insufficient separation between multiple antennas on a CU by looking at the effect of spatial-correlation for Rayleigh channels. Existing studies, including [123, 124], show that the sensing performance of an energy detector degrades with antenna correlation. However, correlation improves the sensing performance of a GLRT detector [125]. Therefore, in this chapter, we also study quickest spectrum sensing performance with multiple antenna CUs in correlated Rayleigh channels. We derive the joint pdf of the received signal for the correlated Rayleigh case and analyze the effect of correlation on the quickest spectrum sensing performance.

We also consider the case when the received signal experiences a LOS condition, resulting in a Rician channel. We derive the joint pdf of the received signal for the independent Rician scenario. For the Rayleigh (independent and correlated) and the Rician (independent) cases, we also study the quickest spectrum sensing performance in the event of a mis-matched channel, where the CUSUM detector is designed for a specific channel, but experiences a different channel. We analytically compute the upper bound and asymptotic worst-case detection delay for both independent Rayleigh and Gaussian cases. We also numerically evaluate the sensing performance for the Rayleigh (independent and correlated) and the Rician (independent) cases. The results in this study provide us with new insights into the minimum detection delay that can be obtained by adding more antennas at the CU in various types of channels. In addition, the results allow us to gain further insights into the effects of channel correlation or the Rician K-factor on multi-antenna quickest spectrum sensing.

The rest of the chapter is organized as follows. Section 5.2 describes the system model. Sections 5.4, 5.5, 5.6 and 5.7 study the Gaussian, independent Rayleigh, correlated Rayleigh and independent Rician channel scenarios, respectively, where derivation of the pdfs of the received signal and the log likelihood ratio are given for each of the scenarios. Analytical and numerical results are presented in Sections 5.8 and 5.9. Finally, Section 5.10 provides some concluding remarks.

5.2 SYSTEM MODEL

In this chapter, we model the signal transmitted by the PU as a narrowband complex Gaussian signal. An interweave cognitive radio network is considered where a CU, which is equipped with multiple antennas, monitors the channel allocated to the PU based on its own observations at each antenna. It is assumed that the PU is initially inactive and that the CU observes samples sequentially and attempts to detect the PU signal. Therefore, this study focuses on the detection of the entrance of the PU to the licensed channel. The detection of the departure of a PU can be approached similarly.

The CU observation at antenna m is denoted by $Y_m[i]$ for $m = 1, 2, ..., M$, where i is the sample number of the received signal and M is the number of antennas. If the PU is not active, $Y_m[i] = N_m[i]$, where $N_m[i] \sim \mathcal{CN}(0,\sigma_N^2)$ is an independent circularly symmetric complex white Gaussian noise. If the PU is transmitting, $Y_m[i] = H_m[i]\times S[i]+N_m[i]$, where $H_m[i]$ is the channel coefficient and the PU signal is $S[i] \sim \mathcal{CN}(0,\sigma_S^2)$, an independent circularly symmetric complex Gaussian random variable with variance σ_S^2. At an unknown sample, τ, the PU commences transmission resulting in a change in the distribution of the received signal. The PU signal is detected by the CUSUM algorithm (discussed in Section 2.4.1) at sample T.

5.3 CUSUM DETECTOR WITH VECTOR RECEIVED SIGNALS

We define $\mathbf{Y}[i] = [Y_1[i], ..., Y_M[i]]^T$. Recall from Section 2.4.1 that the log likelihood ratio required to construct the CUSUM statistic in (2.19) can be written as

$$l_{\mathbf{Y}[i]}(\mathbf{y}[i]) = \ln\left\{\frac{f^{(1)}_{\mathbf{Y}[i]}(y_1[i], y_2[i], ..., y_M[i])}{f^{(0)}_{\mathbf{Y}[i]}(y_1[i], y_2[i], ..., y_M[i])}\right\}. \tag{5.1}$$

Alternatively, if a single sufficient statistic, $z[i]$, exists then $z[i]$ is a scalar function of $\mathbf{Y}[i]$ and (5.1) can be replaced by a simpler ratio, which can be expressed as

$$l_{z[i]}(z[i]) = \ln\left\{\frac{f^{(1)}_{z[i]}(z[i])}{f^{(0)}_{z[i]}(z[i])}\right\}. \tag{5.2}$$

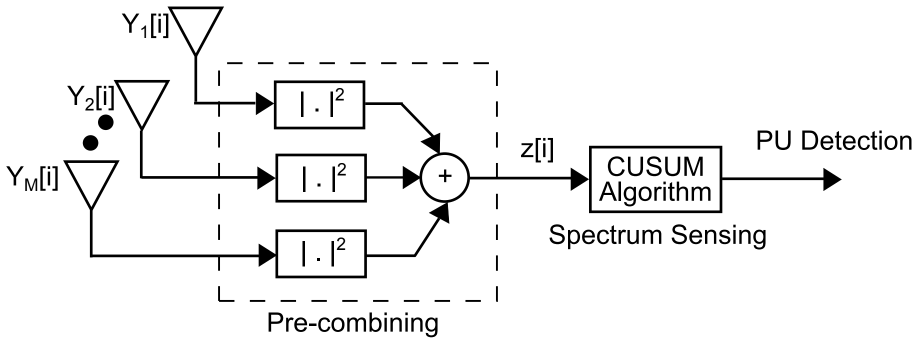

Figure 5.1 Block diagram of CUSUM detection with multiple receive antennas at the CU employing EGC for pre-combining.

If a single sufficient statistic exists, the log likelihood ratio can also be computed using this sufficient statistic, i.e. by using (5.2), which results in the same value as that computed based on the joint pdfs of the received signal, as in (5.1). In this chapter, a single sufficient statistic can be found in the Gaussian and independent Rayleigh cases (which will be discussed and shown in Sections 5.4 and 5.5, respectively), where the sufficient statistic can be expressed as

$$z[i] = \mathbf{Y}[i]^{\dagger}\mathbf{Y}[i] = \sum_{m=1}^{M} |Y_m[i]|^2 . \tag{5.3}$$

It can be observed from (5.3), that $z[i]$ can be obtained from a linear combining technique. As discussed in Section 2.2.1.3, there are various types of multi-antenna combining techniques including selection combining, EGC and MRC [54, 70, 72, 76]. The sufficient statistic in (5.3) can be produced by an EGC and so $z[i]$ is the output of an EGC as shown in Figure 5.1.

5.4 MULTI-ANTENNA SENSING WITH GAUSSIAN CHANNELS

In this section, $H_m[i] = H_m$, where H_m is a time-invariant channel gain. Let $X_m[i] = H_m \times S_m[i]$, where $X_m[i]$ is a circularly symmetric complex Gaussian variable with variance σ_X^2. Hence, the received signal at antenna m when the PU is transmitting can be rewritten as $Y_m[i] = X_m[i] + N_m[i]$. As shown from Figure 5.1, at each sample, i, EGC is applied giving

$$z[i] = \sum_{m=1}^{M} |Y_m[i]|^2 . \tag{5.4}$$

Then, $z[i]$ is processed by the CUSUM algorithm (discussed in Section 2.4.1) to determine the PU's existence.

5.4.1 Pdfs and log likelihood ratio

Whether or not the PU signal is present, $z[i]$ has a chi-square distribution with $2M$ degrees of freedom since $Y_m[i]$ is zero-mean Gaussian. Initially, the signal observed by the CU contains only noise because of the absence of the PU. At an unknown sample number, τ, the PU becomes active and the pdf of the combined signals switches instantaneously. The pdfs when the PU is absent and present are, respectively,

$$f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i]) = \frac{z[i]^{M-1}}{(\sigma_N^2)^M\Gamma(M)} e^{-\frac{z[i]}{\sigma_N^2}}, \tag{5.5}$$

$$f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i]) = \frac{z[i]^{M-1}}{(\sigma_N^2+\sigma_X^2)^M\Gamma(M)} e^{-\frac{z[i]}{\sigma_N^2+\sigma_X^2}}. \tag{5.6}$$

We can then derive the log likelihood ratio, $l_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])$ used in the CUSUM algorithm as

$$\begin{aligned} l_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i]) &= \ln\left\{\frac{f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])}{f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])}\right\} \\ &= \ln\left\{z[i]^{M-1}e^{-\frac{z[i]}{\sigma_N^2+\sigma_X^2}}\right\} - \ln\left\{(\sigma_N^2+\sigma_X^2)^M\Gamma(M)\right\} - \ln\left\{z[i]^{M-1}e^{-\frac{z[i]}{\sigma_N^2}}\right\} + \\ &\quad \ln\left\{(\sigma_N^2)^M\Gamma(M)\right\} \\ &= \frac{z[i]\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)} + M\ln\left\{\frac{\sigma_N^2}{\sigma_N^2+\sigma_X^2}\right\}. \end{aligned} \tag{5.7}$$

It can be observed from (5.7) that $l_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])$ is solely a function of $z[i]$ and hence, it follows that $z[i]$ is a sufficient statistic. It is also comforting to see that for $M = 1$, (5.7) would result in the same result as in (3.12), computed using the amplitude of the received signal, $|Y[i]|$.

5.5 MULTI-ANTENNA SENSING WITH INDEPENDENT RAYLEIGH CHANNELS

Here, the Rayleigh channel coefficient is $H_m[i] \sim \mathcal{CN}(0,\sigma_H^2)$. It is assumed that the antennas experience mutually uncorrelated channels and therefore, the Rayleigh fading channel is independent and identically distributed (i.i.d) between samples i and across antennas.

5.5.1 Proof that $\sum_{m=1}^{M} |Y_m[i]|^2$ is a sufficient statistic

The joint pdf of the received signal when the PU is absent is denoted by

$$\begin{aligned} f^{(0)}_{\mathbf{Y}[i]}(\mathbf{y}[i]) &= \frac{e^{-\mathbf{y}[i]^{\dagger}\mathbf{y}[i]/\sigma_N^2}}{\left(\pi\sigma_N^2\right)^M} \\ &= \frac{e^{-\sum_{m=1}^{M}|y_m[i]|^2/\sigma_N^2}}{\left(\pi\sigma_N^2\right)^M}. \end{aligned} \tag{5.8}$$

In the presence of the PU, the distribution of $Y_m[i]$ conditioned on $S[i]$, is $\mathcal{CN}(0, \sigma_H^2|S[i]|^2 + \sigma_N^2)$. Therefore, the conditional joint pdf of $Y_1[i], Y_2[i], ..., Y_M[i]$, given $S[i]$ can be written as

$$f^{(1)}_{\mathbf{Y}[i]|S[i]}(\mathbf{y}[i]) = \frac{e^{-\sum_{m=1}^{M}|y_m[i]|^2/\left(\sigma_H^2|S[i]|^2+\sigma_N^2\right)}}{\left[\pi\left(\sigma_H^2|S[i]|^2+\sigma_N^2\right)\right]^M}. \tag{5.9}$$

Let $|S[i]|^2 = \sigma_S^2 U[i]$, where $U[i]$ is a standard exponential random variable such that $U[i] \sim \mathrm{Exp}(1)$. Thus, (5.9) can be expressed as

$$f^{(1)}_{\mathbf{Y}[i]|S[i]}(\mathbf{y}[i]) = \frac{e^{-\sum_{m=1}^{M}|y_m[i]|^2/\left(\sigma_H^2\sigma_S^2U[i]+\sigma_N^2\right)}}{\left[\pi\left(\sigma_H^2\sigma_S^2U[i]+\sigma_N^2\right)\right]^M}. \tag{5.10}$$

The joint pdf of the received signal when the PU is present can be written using (5.10) as

$$\begin{aligned} f^{(1)}_{\mathbf{Y}[i]}(\mathbf{y}[i]) &= \int f^{(1)}_{\mathbf{Y}[i]|S[i]}(\mathbf{y}[i]) f_{U[i]}(u[i])\, du[i] \\ &= \int_0^{\infty} \frac{e^{-\sum_{m=1}^{M}|y_m[i]|^2/\left(\sigma_H^2\sigma_S^2u[i]+\sigma_N^2\right)}}{\left[\pi\left(\sigma_H^2\sigma_S^2u[i]+\sigma_N^2\right)\right]^M} e^{-u[i]}\, du[i]. \end{aligned} \tag{5.11}$$

Therefore, using (5.8) and (5.11), the log likelihood ratio can be written as

$$l_{\mathbf{Y}[i]}(\mathbf{y}[i]) = \ln\left\{ \left(\sigma_N^2\right)^M e^{\sum_{m=1}^{M} |y_m[i]|^2/\sigma_N^2} \int_0^\infty \frac{e^{-\sum_{m=1}^{M} |y_m[i]|^2/(\sigma_H^2\sigma_S^2 u[i]+\sigma_N^2)}}{(\sigma_H^2\sigma_S^2 u[i] + \sigma_N^2)^M} e^{-u[i]}\, du[i] \right\}. \quad (5.12)$$

Based on (5.12), $l_{\mathbf{Y}[i]}(\mathbf{y}[i])$ is solely a function of $\sum_{m=1}^{M} |y_m[i]|^2$ and hence, it follows that $\sum_{m=1}^{M} |Y_m[i]|^2$ is a sufficient statistic. From Theorem 8.2.4 in [126], we conclude that a likelihood ratio test (LRT) based on $\sum_{m=1}^{M} |Y_m[i]|^2$ is equivalent to the LRT based on the vector $Y_1[i], Y_2[i], ..., Y_M[i]$. Thus, the log likelihood ratio in (5.12) can also be computed using the pdf of $\sum_{m=1}^{M} |Y_m[i]|^2$, which is based on a chi-square density and this will be shown in Section 5.5.2.

5.5.2 Pdfs and log likelihood ratio

As shown in Section 5.5.1, the log likelihood ratio can also be evaluated based on the statistics of $\sum_{m=1}^{M} |Y_m[i]|^2$. Therefore, in this section, we will derive the pdf and subsequently the log likelihood ratio based on $\sum_{m=1}^{M} |Y_m[i]|^2$. This then allows us to apply CUSUM detection, as in Figure 5.1. As in Section 5.4, in the absence of the PU, the combined received signal, $z[i] = \sum_{m=1}^{M}|Y_m[i]|^2$, has a chi-square distribution with $2M$ degrees of freedom and the pdf can be written as in (5.5).

In order to derive the pdf of the combined signal, when the PU signal is present, we first derive its cdf. Conditioned on $S[i]$, $\mathbf{Y}[i] = [Y_1[i], ..., Y_M[i]]^T$ has a complex Gaussian distribution and can be written as

$$\mathbf{Y}[i] = \left(\sigma_H^2 |S[i]|^2 + \sigma_N^2\right)^{1/2} \mathbf{J}[i], \quad (5.13)$$

where $\mathbf{J}[i]$ is a $M \times 1$ vector with independent $\mathcal{CN}(0,1)$ entries. Using this representation,

$$\begin{aligned} z[i] &= \mathbf{Y}[i]^\dagger \mathbf{Y}[i] \\ &= \left(\sigma_H^2 |S[i]|^2 + \sigma_N^2\right) \mathbf{J}[i]^\dagger \mathbf{J}[i]. \end{aligned} \quad (5.14)$$

Defining $Q[i] = \mathbf{J}[i]^\dagger \mathbf{J}[i]$ and $|S[i]|^2 = \sigma_S^2 U[i]$, (5.14) can be rewritten as

$$z[i] = (\sigma_H^2 \sigma_S^2 U[i] + \sigma_N^2) Q[i], \quad (5.15)$$

where $U[i]$ is a standard exponential random variable, such that $U[i] \sim \text{Exp}(1)$ and $Q[i]$ is a standard chi-square variable with $2M$ degrees of freedom. The densities of $U[i]$ and $Q[i]$ are respectively $f_{U[i]}(u[i])$ and $f_{Q[i]}(q[i])$. We let $f(q[i], u[i])$ be the joint pdf of $Q[i]$ and $U[i]$. Thus, the cdf of the combined signal when the PU is present can be expressed as

$$\begin{aligned} P(z[i] < x) &= P\left(\left(\sigma_H^2\sigma_S^2 U[i] + \sigma_N^2\right) Q[i] < x\right) \\ &= P\left((Q[i], U[i]) \in \mathbb{D}\right), \end{aligned} \tag{5.16}$$

where $\mathbb{D} = \{Q[i], U[i] : z[i] < x\}$. Hence $P(z[i] < x)$ can be expressed as

$$\begin{aligned} P(z[i] < x) &= \iint_{\mathbb{D}} f(q[i], u[i])\, dq[i]\, du[i] \\ &= \iint_{\mathbb{D}} f(q[i]|u[i])\, dq[i] f_{U[i]}(u[i])\, du[i] \\ &= E\left[P\left(Q[i] < \frac{x}{\sigma_H^2\sigma_S^2 U[i] + \sigma_N^2}\right)|U[i]\right] \\ &= E\left[1 - \int_0^\infty e^{\frac{-x}{\sigma_H^2\sigma_S^2 U[i]+\sigma_N^2}} \sum_{k=0}^{M-1} \frac{1}{k!}\left(\frac{x}{\sigma_H^2\sigma_S^2 U[i] + \sigma_N^2}\right)^k |U[i]\right] \\ &= 1 - \int_0^\infty e^{\frac{-x}{\sigma_H^2\sigma_S^2 u[i]+\sigma_N^2}} \sum_{k=0}^{M-1} \frac{1}{k!}\left(\frac{x}{\sigma_H^2\sigma_S^2 u[i] + \sigma_N^2}\right)^k e^{-u[i]}\, du[i]. \end{aligned} \tag{5.17}$$

Let $t = \sigma_T^2 u[i] + \sigma_N^2$, where $\sigma_T^2 = \sigma_H^2\sigma_S^2$, then

$$P(z[i] < x) = 1 - \frac{e^{\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2} \sum_{k=0}^{M-1} \frac{x^k}{k!} \int_{\sigma_N^2}^\infty \frac{e^{\frac{-x}{t} - \frac{t}{\sigma_T^2}}}{t^k}\, dt. \tag{5.18}$$

The pdf of the combined signal when the PU signal is present is obtained by taking the derivative of (5.18) to yield

$$f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i]) = \frac{e^{\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2} \sum_{k=0}^{M-1} \frac{z[i]^k}{k!} \int_{\sigma_N^2}^\infty \left(\frac{e^{\frac{-z[i]}{t} - \frac{t}{\sigma_T^2}}}{t^{k+1}} - \frac{k}{z[i]} \times \frac{e^{\frac{-z[i]}{t} - \frac{t}{\sigma_T^2}}}{t^k}\right) dt. \tag{5.19}$$

Computation of (5.19) is assisted by avoiding numerical integration over an infinite region. Hence, we rewrite (5.19) as the difference of the integral from 0 to ∞ and the finite integral from

0 to σ_N^2, which yields

$$
\begin{aligned}
f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) &= \frac{e^{\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2} \sum_{k=0}^{M-1} \frac{z[i]^k}{k!} \left[\int_0^\infty \frac{e^{\frac{-z[i]}{t} - \frac{t}{\sigma_T^2}}}{t^{k+1}} dt - \int_0^{\sigma_N^2} \frac{e^{\frac{-z[i]}{t} - \frac{t}{\sigma_T^2}}}{t^{k+1}} dt \right. \\
&\quad \left. - \frac{k}{z[i]} \left(\int_0^\infty \frac{e^{\frac{-z[i]}{t} - \frac{t}{\sigma_T^2}}}{t^k} dt - \int_0^{\sigma_N^2} \frac{e^{\frac{-z[i]}{t} - \frac{t}{\sigma_T^2}}}{t^k} dt \right) \right] \\
&= \frac{e^{\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2} \sum_{k=0}^{M-1} \frac{z[i]^k}{k!} \left[\int_0^\infty \frac{e^{\frac{-z[i]}{t} - \frac{t}{\sigma_T^2}}}{t^{k+1}} dt - \frac{k}{z[i]} \int_0^\infty \frac{e^{\frac{-z[i]}{t} - \frac{t}{\sigma_T^2}}}{t^k} dt \right. \\
&\quad \left. + \frac{k}{z[i]} \int_0^{\sigma_N^2} \frac{e^{\frac{-z[i]}{t} - \frac{t}{\sigma_T^2}}}{t^k} dt - \int_0^{\sigma_N^2} \frac{e^{\frac{-z[i]}{t} - \frac{t}{\sigma_T^2}}}{t^{k+1}} dt \right].
\end{aligned}
\tag{5.20}
$$

The integrals from 0 to ∞ in (5.20) can be evaluated with the aid of (3.471.9) in [103, p. 363]. Since σ_N^2 is never large, a simple Riemann sum approximation with the mid-point rule [104] works well, where rectangles are used to approximate the area under the curve. With this approach, the pdf in (5.20) can be written to any degree of approximation as

$$
\begin{aligned}
f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) = \frac{e^{\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2} \sum_{k=0}^{M-1} \frac{z[i]^k}{k!} &\left[2 \left(z[i]\sigma_T^2 \right)^{\frac{-k}{2}} \left(\mathrm{K}_{-k} \left(\frac{\sqrt{z[i]}}{\frac{\sigma_T}{2}} \right) - \frac{k\sigma_T}{\sqrt{z[i]}} \mathrm{K}_{1-k} \left(\frac{\sqrt{z[i]}}{\frac{\sigma_T}{2}} \right) \right) + \right. \\
&\left. \frac{\sigma_N^2}{R} \sum_{r=1}^{R} \frac{e^{\frac{-z[i]}{s_r} - \frac{s_r}{\sigma_T^2}}}{(s_r)^k} \left(\frac{k}{z[i]} - \frac{1}{s_r} \right) \right],
\end{aligned}
\tag{5.21}
$$

where

$$
s_r = \left(r - \frac{1}{2} \right) \left(\frac{\sigma_N^2}{R} \right), \tag{5.22}
$$

in which R is the number of rectangles. Numerical tests show that this approach gives a negligible error and is much faster than numerical integration with $R = 50$.

The log likelihood ratio can now be calculated using the chi-square density in (5.5) along with

(5.21), which gives

$$l_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i]) = \ln\left\{ \frac{(\sigma_N^2)^M \Gamma(M) e^{\frac{z[i]}{\sigma_N^2}+\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2 z[i]^{M-1}} \sum_{k=0}^{M-1} \frac{z[i]^k}{k!} \left[2\left(z[i]\sigma_T^2\right)^{\frac{-k}{2}} \left(\mathrm{K}_{-k}\left(\frac{\sqrt{z[i]}}{\frac{\sigma_T}{2}}\right) - \frac{k\sigma_T}{\sqrt{z[i]}} \mathrm{K}_{1-k}\left(\frac{\sqrt{z[i]}}{\frac{\sigma_T}{2}}\right)\right) + \frac{\sigma_N^2}{R} \sum_{r=1}^{R} \frac{e^{\frac{-z[i]}{s_r}-\frac{s_r}{\sigma_T^2}}}{(s_r)^k} \left(\frac{k}{z[i]} - \frac{1}{s_r}\right)\right]\right\}. \tag{5.23}$$

Therefore, this log likelihood ratio, $l_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])$, can be substituted into (2.21) to detect the presence of the PU.

5.6 MULTI-ANTENNA SENSING WITH CORRELATED RAYLEIGH CHANNELS

Here, we assume that $H_m[i]$ is correlated between antennas, but independent between samples. The spatially correlated Rayleigh fading channel, $\mathbf{H}[i] = [H_1[i], H_2[i], ..., H_M[i]]^T$, can be modelled by

$$\mathbf{H}[i] = \mathbf{R}^{1/2}\mathbf{H}_w[i], \tag{5.24}$$

where $\mathbf{H}_w[i]$ is a spatially white $M \times 1$ vector with i.i.d $\mathcal{CN}(0,1)$ entries and $\mathbf{R}$ is the $M \times M$ antenna correlation matrix denoted by

$$\mathbf{R} = E\left[\mathbf{H}[i]\mathbf{H}[i]^\dagger\right]. \tag{5.25}$$

We assume that

$$\mathbf{R}_{jk} = \rho^{|j-k|}, \tag{5.26}$$

where $j, k = 1, 2, .., M$ and $0 \leq \rho \leq 1$. This is the well-known exponential correlation model [127] where ρ is the correlation coefficient between the channels on adjacent antennas.

5.6.1 Insufficient statistic

In Section 5.5.1, we prove that the sum of the complex received signal power at each antenna is a sufficient statistic. This is true for the case when the received signal is transmitted over independent Rayleigh channels. However, when there is an insufficient separation between multiple antennas, resulting in correlated Rayleigh channels, we will show in this section that a sufficient statistic cannot be isolated from the log likelihood ratio.

Initially, the PU is inactive and hence, the joint pdf of the received signal when the PU is absent has a complex Gaussian distribution which can be written as in (5.8). When the PU is active, the distribution of the observed signal at the CU, $\mathbf{Y}[i]$, conditioned on $S[i]$, has a complex Gaussian distribution given by $\mathcal{CN}(0, |S[i]|^2\mathbf{R} + \sigma_N^2\mathbf{I})$. Therefore, the joint pdf of the received signal can be written as

$$\begin{aligned} f^{(1)}_{\mathbf{Y}[i]}(\mathbf{y}[i]) &= \int f^{(1)}_{\mathbf{Y}[i]|S[i]}(\mathbf{y}[i]) f_{S[i]}(s[i])\, ds[i] \\ &= E\left[f^{(1)}_{\mathbf{Y}[i]|S[i]}(\mathbf{y}[i]) \right], \end{aligned} \tag{5.27}$$

where the expectation is taken over $S[i]$. The joint pdf of $\mathbf{Y}[i]$ conditioned on $S[i]$ which appears in (5.27) is given by

$$f^{(1)}_{\mathbf{Y}[i]|S[i]}(\mathbf{y}[i]) = \frac{e^{-\mathbf{y}[i]^\dagger(|S[i]|^2\mathbf{R}+\sigma_N^2\mathbf{I})^{-1}\mathbf{y}[i]}}{\pi^M \det\left(|S[i]|^2\mathbf{R} + \sigma_N^2\mathbf{I}\right)}, \tag{5.28}$$

where $\det(\mathbf{A})$ is the determinant of a matrix $\mathbf{A}$. Let $|S[i]|^2 = \sigma_S^2 U[i]$, where $U[i]$ is a standard exponential random variable. Therefore, (5.28) can be rewritten as

$$f^{(1)}_{\mathbf{Y}[i]|S[i]}(\mathbf{y}[i]) = \frac{e^{-\mathbf{y}[i]^\dagger(\sigma_S^2 U[i]\mathbf{R}+\sigma_N^2\mathbf{I})^{-1}\mathbf{y}[i]}}{\pi^M \det\left(\sigma_S^2 U[i]\mathbf{R} + \sigma_N^2\mathbf{I}\right)}. \tag{5.29}$$

Using (5.29), the joint pdf of the received signal in (5.27) can be expressed as

$$f^{(1)}_{\mathbf{Y}[i]}(\mathbf{y}[i]) = \frac{1}{\pi^M}\int_0^\infty \frac{e^{-\mathbf{y}[i]^\dagger(\sigma_S^2 u[i]\mathbf{R}+\sigma_N^2\mathbf{I})^{-1}\mathbf{y}[i]}}{\det\left(\sigma_S^2 u[i]\mathbf{R} + \sigma_N^2\mathbf{I}\right)} e^{-u[i]}\, du[i]. \tag{5.30}$$

The log likelihood ratio can be written using (5.8) and (5.30) to give

$$l_{\mathbf{Y}[i]}(\mathbf{y}[i]) = \ln\left\{ \left(\sigma_N^2\right)^M e^{\mathbf{y}[i]^\dagger\mathbf{y}[i]/\sigma_N^2} \int_0^\infty \frac{e^{-\mathbf{y}[i]^\dagger(\sigma_S^2 u[i]\mathbf{R}+\sigma_N^2\mathbf{I})^{-1}\mathbf{y}[i]}}{\det\left(\sigma_S^2 u[i]\mathbf{R} + \sigma_N^2\mathbf{I}\right)} e^{-u[i]}\, du[i] \right\}. \tag{5.31}$$

It can be observed from (5.31) that $\mathbf{Y}[i]$ cannot be isolated from $U[i]$ and so a single sufficient statistic cannot be found. As a result of this, the rest of the analysis in this section will be based on the pdfs and the log likelihood ratio of the vector $\mathbf{Y}[i]$. In Section 5.6.2, we will present an alternative method of deriving the joint pdf of the received signal in the presence of the PU.

This result is useful as it is more suitable for numerical integration than the original result using (5.30).

5.6.2 Pdfs and log likelihood ratio

In this subsection, we present an alternative form of the joint pdf of the received signal, $f^{(1)}_{\mathbf{Y}[i]}(\mathbf{y}[i])$. An eigendecomposition is performed on $\mathbf{R}$ in (5.29), which gives

$$\mathbf{R} = \boldsymbol{\phi}\boldsymbol{\Lambda}\boldsymbol{\phi}^{\dagger}, \tag{5.32}$$

where $\boldsymbol{\phi}$ is a unitary matrix, $\boldsymbol{\Lambda} = \mathrm{diag}[\lambda_1, \lambda_2, .., \lambda_M]$ and λ_m is the mth eigenvalue of $\mathbf{R}$. Thus, (5.29) becomes

$$f^{(1)}_{\mathbf{Y}[i]|S[i]}(\mathbf{y}[i]) = \frac{e^{-\mathbf{v}[i]^{\dagger}(\sigma_S^2 U[i]\boldsymbol{\Lambda}+\sigma_N^2\mathbf{I})^{-1}\mathbf{v}[i]}}{\pi^M \prod_{m=1}^{M}(\lambda_m \sigma_S^2 U[i] + \sigma_N^2)}, \tag{5.33}$$

where $\mathbf{v}[i] = \boldsymbol{\phi}^{\dagger}\mathbf{y}[i]$. We can also rewrite (5.33) as

$$f^{(1)}_{\mathbf{Y}[i]|S[i]}(\mathbf{y}[i]) = \frac{e^{-\sum_{m=1}^{M} v_m^* v_m/(\sigma_S^2 U[i]\lambda_m+\sigma_N^2)}}{\pi^M \prod_{m=1}^{M}(\lambda_m \sigma_S^2 U[i] + \sigma_N^2)}. \tag{5.34}$$

Using (5.27) and (5.34), the joint pdf of the received signal can be written as

$$f^{(1)}_{\mathbf{Y}[i]}(\mathbf{y}[i]) = \frac{1}{\pi^M}\int_0^{\infty} \frac{e^{-\sum_{m=1}^{M} v_m^* v_m/(\sigma_S^2 u[i]\lambda_m+\sigma_N^2)}}{\prod_{m=1}^{M}(\lambda_m \sigma_S^2 u[i] + \sigma_N^2)} e^{-u[i]}\, du[i]. \tag{5.35}$$

Note that no closed form solution for (5.35) exists, so a single numerical integration is necessary. Furthermore, series expansions lead to multiple infinite summations and approximations are complicated by the fact that the density is a multidimensional function of σ_N^2 and the elements of $\mathbf{R}$. Finally, note that the form given in (5.35) is more suitable for numerical integration than the original using (5.30).

The log likelihood ratio can be expressed using (5.8) and (5.35), which gives

$$l_{\mathbf{Y}[i]}(\mathbf{y}[i]) = \ln\left\{ \left(\sigma_N^2\right)^M e^{\mathbf{y}[i]^{\dagger}\mathbf{y}[i]/\sigma_N^2} \int_0^{\infty} \frac{e^{-\sum_{m=1}^{M} v_m^* v_m/(\sigma_S^2 u[i]\lambda_m+\sigma_N^2)}}{\prod_{m=1}^{M}(\lambda_m \sigma_S^2 u[i] + \sigma_N^2)} e^{-u[i]}\, du[i] \right\}. \tag{5.36}$$

This log likelihood ratio, $l_{\mathbf{Y}[i]}(\mathbf{y}[i])$, is then used in the CUSUM algorithm in (2.21) to detect the existence of the PU.

5.7 MULTI-ANTENNA SENSING WITH INDEPENDENT RICIAN CHANNELS

In this section, it is assumed that the antennas experience mutually uncorrelated channels so that $H_m[i]$ is i.i.d between samples and across all antennas. The Rician channel can be modelled by

$$H_m[i] = \mu_m[i] + \nu_m[i], \tag{5.37}$$

where

$$\mu_m[i] = \sqrt{\frac{K}{K+1}} H_m^{\mathrm{LOS}}[i], \tag{5.38}$$

and

$$\nu_m[i] = \sqrt{\frac{1}{K+1}} H_m^{\mathrm{SC}}[i]. \tag{5.39}$$

In (5.38) and (5.39), $H_m^{\mathrm{LOS}}[i]$ is the LOS component, satisfying $\left|H_m^{\mathrm{LOS}}[i]\right|^2 = 1$ and $H_m^{\mathrm{SC}}[i]$ is the scattered component (SC), such that $H_m^{\mathrm{SC}}[i] \sim \mathcal{CN}(0,1)$ is a zero-mean circularly symmetric complex Gaussian random variable.

5.7.1 Pdfs and log likelihood ratio

The joint pdf of the observed signal when the PU is absent can be written as in (5.8). Since $E\left[|H_m[i]|^2\right] = 1$, the distribution of $Y_m[i]$ in the presence of the PU, conditioned on $S[i]$, is $\mathcal{CN}\left(\mu_m[i]S[i], \left(1 - |\mu_m[i]|^2\right)|S[i]|^2 + \sigma_N^2\right)$. Thus, the conditional joint pdf of $Y_1[i], Y_2[i], ..., Y_M[i]$, given $S[i]$ can be expressed as

$$f^{(1)}_{\mathbf{Y}[i]|S[i]}(\mathbf{y}[i]) = \frac{1}{\left[\pi\left(\alpha|S[i]|^2 + \sigma_N^2\right)\right]^M} \times \exp\left(-\frac{\sum_{m=1}^{M} |y_m[i] - \mu_m[i]S[i]|^2}{\alpha\,|S[i]|^2 + \sigma_N^2}\right), \tag{5.40}$$

where $\alpha = 1 - |\mu_m[i]|^2 = \frac{1}{K+1}$ and exp(.) denotes the exponential function. We let $S[i] = |S[i]|e^{j\theta}$, where $\theta = \arg(S[i])$ and θ is uniformly distributed over $[0, 2\pi]$, $\theta \sim U(0, 2\pi)$. Using (5.40) and the polar expression for $S[i]$, the joint pdf of the received signal can be written as

$$f^{(1)}_{\mathbf{Y}[i]}(\mathbf{y}[i]) = \frac{1}{2\pi^{M+1}} \int_{|s[i]|=0}^{\infty} \int_{\theta=0}^{2\pi} \frac{1}{\left(\alpha|s[i]|^2+\sigma_N^2\right)^M} \times \exp\left(-\frac{\sum_{m=1}^{M} \left|y_m[i]-\mu_m[i]\,|s[i]|\,e^{j\theta}\right|^2}{\alpha\,|s[i]|^2+\sigma_N^2}\right)$$
$$\times f_{|S[i]|}\left(|s[i]|\right)\,d\theta\,d|s[i]|, \tag{5.41}$$

where $f_{|S[i]|}\left(|s[i]|\right)$ is the pdf of the amplitude of the PU signal. In order to derive $f_{|S[i]|}\left(|s[i]|\right)$, we first derive its cdf. The PU signal can be expressed as

$$S[i] = \left(\sigma_S^2\right)^{1/2} J[i], \tag{5.42}$$

where $J[i] \sim \mathcal{CN}(0,1)$. We can rewrite (5.42) as

$$S[i]^{\dagger} S[i] = \sigma_S^2 J[i]^{\dagger} J[i]. \tag{5.43}$$

Let $Q[i] = J[i]^{\dagger} J[i]$, where $Q[i]$ is an exponential random variable such that $Q[i] \sim \text{Exp}(1)$. Therefore, the cdf can be written as

$$\begin{aligned} P(|S[i]|<c) &= P\left(|S[i]|^2<c^2\right) \\ &= P\left(\sigma_S^2 Q[i]<c^2\right) \\ &= P\left(Q[i]<\frac{c^2}{\sigma_S^2}\right) \\ &= 1-e^{-\frac{c^2}{\sigma_S^2}}. \end{aligned} \tag{5.44}$$

The pdf of the amplitude of the PU signal can then be obtained by taking the derivative of (5.44) to give

$$f_{|S[i]|}(|s[i]|) = \frac{2|s[i]|}{\sigma_S^2} e^{-|s[i]|^2/\sigma_S^2}. \tag{5.45}$$

The inner integral in (5.41) is given by

$$\begin{aligned} A &= \int_{\theta=0}^{2\pi} \exp\left(-\frac{\sum_{m=1}^{M} \left|y_m[i] - \mu_m[i]\,|s[i]|\,e^{j\theta}\right|^2}{\alpha\,|s[i]|^2 + \sigma_N^2}\right) d\theta \\ &= \exp\left(-\frac{\sum_{m=1}^{M} \left(|y_m[i]|^2 + |\mu_m[i]|^2\,|s[i]|^2\right)}{\alpha\,|s[i]|^2 + \sigma_N^2}\right) \times \mathcal{I}, \end{aligned} \tag{5.46}$$

where

$$\mathcal{I} = \int_{\theta=0}^{2\pi} \exp\left(\frac{\sum_{m=1}^{M} y_m[i]^\dagger \mu_m[i]|s[i]|e^{j\theta}}{\alpha\,|s[i]|^2 + \sigma_N^2}\right) \times \exp\left(\frac{\sum_{m=1}^{M} y_m[i]\mu_m[i]^\dagger |s[i]|e^{-j\theta}}{\alpha\,|s[i]|^2 + \sigma_N^2}\right) d\theta. \tag{5.47}$$

With the aid of Euler's formula and (3.339) in [103, p. 336], $\mathcal{I}$ in (5.47) becomes

$$\mathcal{I} = 2\pi \mathrm{I}_0\left(\frac{2|s[i]|}{\left|\alpha\,|s[i]|^2 + \sigma_N^2\right|}\left|\sum_{m=1}^{M} y_m[i]^\dagger \mu_m[i]\right|\right). \tag{5.48}$$

Substituting (5.48) into (5.46) yields

$$A = \exp\left(-\frac{\sum_{m=1}^{M} \left(|y_m[i]|^2 + |\mu_m[i]|^2\,|s[i]|^2\right)}{\alpha\,|s[i]|^2 + \sigma_N^2}\right) \times 2\pi \mathrm{I}_0\left(\frac{2|s[i]|}{\left|\alpha\,|s[i]|^2 + \sigma_N^2\right|}\left|\sum_{m=1}^{M} y_m[i]^\dagger \mu_m[i]\right|\right). \tag{5.49}$$

The result for A in (5.49) can be substituted into (5.41) and hence, using (5.45), the joint pdf of the received signal can be written as

$$\begin{aligned} f_{\mathbf{Y}[i]}^{(1)}(\mathbf{y}[i]) = \frac{2}{\sigma_S^2 \pi^M} \int_{|s[i]|=0}^{\infty} \frac{|s[i]|e^{-|s[i]|^2/\sigma_S^2}}{\left(\alpha|s[i]|^2 + \sigma_N^2\right)^M} \exp\left(-\frac{\sum_{m=1}^{M} \left(|y_m[i]|^2 + |\mu_m[i]|^2\,|s[i]|^2\right)}{\alpha\,|s[i]|^2 + \sigma_N^2}\right) \times \\ \mathrm{I}_0\left(2|s[i]|\left|\frac{\sum_{m=1}^{M} y_m[i]^\dagger \mu_m[i]}{\alpha\,|s[i]|^2 + \sigma_N^2}\right|\right) d|s[i]|. \end{aligned} \tag{5.50}$$

As in (5.35), the joint density in (5.50) cannot be given in a closed form. Also, the integrands in (5.50) and (5.35) are non-oscillatory and decay rapidly to zero making them ideal for numerical integration.

The log likelihood ratio can then be written using (5.8) along with (5.50) to give

$$l_{\mathbf{Y}[i]}(\mathbf{y}[i]) = \ln\Bigg\{ \frac{2\left(\sigma_N^2\right)^M}{\sigma_S^2} e^{\sum_{m=1}^{M} |y_m[i]|^2/\sigma_N^2} \int_{|s[i]|=0}^{\infty} \frac{|s[i]|e^{-|s[i]|^2/\sigma_S^2}}{\left(\alpha|s[i]|^2+\sigma_N^2\right)^M} \times \\ \exp\left(- \frac{\sum_{m=1}^{M} \left(|y_m[i]|^2 + |\mu_m[i]|^2 \, |s[i]|^2 \right)}{\alpha \, |s[i]|^2 + \sigma_N^2} \right) \times \mathrm{I}_0 \left(2|s[i]| \left| \frac{\sum_{m=1}^{M} y_m[i]^\dagger \mu_m[i]}{\alpha \, |s[i]|^2 + \sigma_N^2} \right| \right) d|s[i]| \Bigg\}. \tag{5.51}$$

It is worth noting that the log likelihood ratio, $l_{\mathbf{Y}[i]}(\mathbf{y}[i])$, does not contain a simple sufficient statistic. Hence, $l_{\mathbf{Y}[i]}(\mathbf{y}[i])$ is based on the vector $\mathbf{Y}[i]$ and it can be substituted into the CUSUM algorithm in (2.21) to detect the PU transmission.

5.8 PERFORMANCE ANALYSIS

Recall from Section 2.4.1 that the threshold, γ in (2.21), can be set either theoretically, corresponding to an approximate false alarm rate value or in an arbitrary manner, where γ is chosen by trial and error to give a sensible range of average detection delay and false alarm rate. In this section and the following section, we will investigate both of the threshold setting techniques. In particular, in this section, we present a theoretical performance analysis of quickest spectrum sensing employing multiple receive antennas for the Gaussian and independent Rayleigh scenarios. In particular, we analytically compute the upper bound and asymptotic worst-case detection delay for both of the cases. The upper bound and asymptotic worst-case detection delay could also be derived for the correlated Rayleigh and independent Rician cases, but these derivations may be complicated to evaluate due to the existence of multiple integrals in the Kullback-Leibler divergence (contained in both of the derivations). These multiple integrals are caused by the pdfs of the received signals in the presence of the PU given in (5.35) and (5.50) and the log likelihood ratios given in (5.36) and (5.51) for both correlated Rayleigh and independent Rician cases, respectively, which involve numerical integration. Therefore, analyzing the theoretical sensing performance for both correlated Rayleigh and independent Rician cases can be considered for future work. However, the sensing performance for both of these cases will be evaluated numerically in Section 5.9.

Let T denote the sample number at which the change is detected and τ be the sample number when the change actually occurs. If $T > \tau$, then the detection delay is $\delta = T - \tau$. Recall from Section 2.4 that the minimax formulation proposed by Lorden [94] models the change-point as an unknown deterministic quantity. Lorden subsequently showed that the well-known Page's CUSUM algorithm [95] is asymptotically[2] optimal in minimizing the worst-case detection delay [37,43,44]. Based on Lorden's formulation [94], the worst-case detection delay for Gaussian and independent Rayleigh channels (described in Sections 5.4 and 5.5, respectively) are given by

$$\overline{T}_{d_G} = \sup_{\tau\geq 1} \text{ess sup } E_{f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}} \left[\delta = T-\tau | T \geq \tau, z[1], ..., z[\tau]\right], \tag{5.52}$$

$$\overline{T}_{d_{Ray}} = \sup_{\tau\geq 1} \text{ess sup } E_{f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}} \left[\delta = T-\tau | T \geq \tau, z[1], ..., z[\tau]\right], \tag{5.53}$$

where $E_{f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}}[.]$ denote the expectation operators when the change occurs at sample number τ and the pdf of $f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}$ is given in (5.6) for the Gaussian case and in (5.21) for the independent Rayleigh case[3]. The subscript, d_G, in (5.52) and d_{Ray}, in (5.53) denote the Gaussian and independent Rayleigh cases, respectively.

Alternatively, if $T < \tau$, a false alarm event will occur with the mean number of samples to false alarm defined as $\overline{T}_{f_G} = E_{f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}}[T]$ for the Gaussian scenario and $\overline{T}_{f_{Ray}} = E_{f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}}[T]$ for the independent Rayleigh scenario [36,37]. $E_{f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}}[.]$ is the expectation operators when the change never happens, where $f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}$ is given in (5.5) for both Gaussian and independent Rayleigh scenarios. The false alarm rates are then defined as $\text{FAR}_G(T) = 1/\overline{T}_{f_G}$ and $\text{FAR}_{Ray}(T) = 1/\overline{T}_{f_{Ray}}$ [98].

The assumption of independence between the received signals allows us to express the lower bound on the mean number of samples between false alarms as $\overline{T}_{f_G} \geq e^\gamma$, and $\overline{T}_{f_{Ray}} \geq e^\gamma$ [36, 44, 94]. We now proceed to derive the upper bound for the worst-case detection delay for the Gaussian channel, $\overline{T}_{d_G}$. Using Wald's equation, (Theorem 1 in [128]) and the fact that the combined signal, $z[i]$, follows a chi-square distribution, we can express the upper bound on $\overline{T}_{d_G}$

[2]Asymptotic here means that the mean number of samples between false alarms goes to infinity.

[3]Essential supremum (ess sup) is used in (5.52) and (5.53) so that $\overline{T}_{d_G}$ and $\overline{T}_{d_{Ray}}$ takes the worst-case value over all possible realizations of the z's before the change [43].

as

$$\overline{T}_{d_G} \leq \frac{\gamma + \varphi}{D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}\right)}, \tag{5.54}$$

where

$$\varphi = \frac{\int_{\upsilon}^{\infty} l^2_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])\, dz[i]}{D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}\right)}. \tag{5.55}$$

In (5.55), υ is the zero of the log likelihood ratio function, $l_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])$ in (5.7), $f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])$ is given in (5.6) and $D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}\right)$ is the Kullback-Leibler divergence of $f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}$ from $f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}$. Solving $l_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) = 0$ in (5.7) gives υ as

$$\upsilon = -\frac{\sigma_N^2(\sigma_N^2 + \sigma_X^2)}{\sigma_X^2} \times M \ln\left\{\frac{\sigma_N^2}{\sigma_N^2 + \sigma_X^2}\right\}. \tag{5.56}$$

$D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}\right)$ in (5.54) and (5.55) can be written as

$$\begin{aligned} D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}\right) &= \int_0^{\infty} f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) \ln\left\{\frac{f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])}{f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])}\right\} dz[i] \\ &= \int_0^{\infty} f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) \ln\left\{f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])\right\} dz[i] - \int_0^{\infty} f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) \ln\left\{f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])\right\} dz[i]. \end{aligned} \tag{5.57}$$

Substituting $f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])$ of (5.5) and $f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])$ of (5.6) into (5.57) yields

$$\begin{aligned} D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}\right) &= \Bigg(-M \ln\left\{\sigma_N^2 + \sigma_X^2\right\} - \ln\left\{\Gamma(M)\right\} - M + (M-1)\times \\ &\int_0^{\infty} \ln\{z[i]\} \frac{z[i]^{M-1}}{(\sigma_N^2 + \sigma_X^2)^M \Gamma(M)} e^{-\frac{z[i]}{\sigma_N^2 + \sigma_X^2}}\, dz[i]\Bigg) - \Bigg(-\ln\left\{(\sigma_N^2)^M \Gamma(M)\right\} - \frac{M(\sigma_N^2 + \sigma_X^2)}{\sigma_N^2} + \\ &(M-1) \times \int_0^{\infty} \ln\{z[i]\} \frac{z[i]^{M-1}}{(\sigma_N^2 + \sigma_X^2)^M \Gamma(M)} e^{-\frac{z[i]}{\sigma_N^2 + \sigma_X^2}}\, dz[i]\Bigg). \end{aligned} \tag{5.58}$$

In order to evaluate the integral in (5.58), we first take the derivative of the gamma function, which gives

$$\begin{aligned}\Gamma'(M) &= \frac{\partial}{\partial M}\Gamma(M)\\ &= \frac{\partial}{\partial M}\int_0^\infty \frac{z[i]^{M-1}}{(\sigma_N^2+\sigma_X^2)^M} e^{-\frac{z[i]}{\sigma_N^2+\sigma_X^2}}\, dz[i]\\ &= \int_0^\infty \ln\{z[i]\} e^{-\frac{z[i]}{\sigma_N^2+\sigma_X^2}} \frac{z[i]^{M-1}}{(\sigma_N^2+\sigma_X^2)^M}\, dz[i] - \ln\left\{\sigma_N^2+\sigma_X^2\right\}\Gamma(M).\end{aligned} \tag{5.59}$$

Dividing (5.59) with the gamma function, $\Gamma(M)$ gives

$$\frac{\Gamma'(M)}{\Gamma(M)} = \int_0^\infty \ln\{z[i]\} e^{-\frac{z[i]}{\sigma_N^2+\sigma_X^2}} \frac{z[i]^{M-1}}{(\sigma_N^2+\sigma_X^2)^M\Gamma(M)}\, dz[i] - \ln\left\{\sigma_N^2+\sigma_X^2\right\}. \tag{5.60}$$

Thus, (5.60) can be rewritten as

$$\int_0^\infty \ln\{z[i]\} \frac{z[i]^{M-1}}{(\sigma_N^2+\sigma_X^2)^M\Gamma(M)} e^{-\frac{z[i]}{\sigma_N^2+\sigma_X^2}}\, dz[i] = \frac{\Gamma'(M)}{\Gamma(M)} + \ln\left\{\sigma_N^2+\sigma_X^2\right\}. \tag{5.61}$$

The result of the integral in (5.61) can be substituted in (5.58) to give the Kullback-Leibler divergence, which can be expressed as

$$\begin{aligned}D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}\right) &= -M\ln\left\{\sigma_N^2+\sigma_X^2\right\} - \ln\left\{\Gamma(M)\right\} - M + (M-1)\times\\ &\quad \left[\frac{\Gamma'(M)}{\Gamma(M)} + \ln\left\{\sigma_N^2+\sigma_X^2\right\}\right] - \Bigg(-\ln\left\{(\sigma_N^2)^M\Gamma(M)\right\} - \frac{M(\sigma_N^2+\sigma_X^2)}{\sigma_N^2} +\\ &\quad (M-1)\times\left[\frac{\Gamma'(M)}{\Gamma(M)} + \ln\left\{\sigma_N^2+\sigma_X^2\right\}\right]\Bigg)\\ &= -M\ln\{\sigma_N^2+\sigma_X^2\} + \frac{M\sigma_X^2}{\sigma_N^2} + M\ln\{\sigma_N^2\}.\end{aligned} \tag{5.62}$$

Therefore, using (5.62), $l_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])$ of (5.7) and $f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])$ of (5.6), we can re-write the upper bound on $\overline{T}_{d_G}$ in (5.54) as

$$\overline{T}_{d_G} \le \frac{\gamma+\varphi}{-M\ln\{\sigma_N^2+\sigma_X^2\} + \frac{M\sigma_X^2}{\sigma_N^2} + M\ln\{\sigma_N^2\}}, \tag{5.63}$$

where

$$\varphi = \frac{\int_{\upsilon}^{\infty}\left(\frac{z[i]\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)} + M\ln\left\{\frac{\sigma_N^2}{\sigma_N^2+\sigma_X^2}\right\}\right)^2 \times \frac{z[i]^{M-1}}{(\sigma_N^2+\sigma_X^2)^M\Gamma(M)} e^{-\frac{z[i]}{\sigma_N^2+\sigma_X^2}}\,dz[i]}{-M\ln\{\sigma_N^2+\sigma_X^2\} + \frac{M\sigma_X^2}{\sigma_N^2} + M\ln\{\sigma_N^2\}}$$
$$= \frac{1}{(\sigma_N^2+\sigma_X^2)^M\Gamma(M)\times\left(-M\ln\{\sigma_N^2+\sigma_X^2\} + \frac{M\sigma_X^2}{\sigma_N^2} + M\ln\{\sigma_N^2\}\right)}\times \tag{5.64}$$
$$\int_{\upsilon}^{\infty}\left(\frac{z[i]\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)} + M\ln\left\{\frac{\sigma_N^2}{\sigma_N^2+\sigma_X^2}\right\}\right)^2 \times z[i]^{M-1} e^{-\frac{z[i]}{\sigma_N^2+\sigma_X^2}}\,dz[i],$$

in which υ is given by (5.56).

The upper bound for the worst-case detection delay for the independent Rayleigh channel, $\overline{T}_{d_{Ray}}$ can be written with the aid of Wald's equation and Theorem 1 in [128] along with the fact that the combined signal from each antenna, $z[i]$, is chi-square distributed, which yields

$$\overline{T}_{d_{Ray}} \leq \frac{\gamma+\Phi}{D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}\right)}, \tag{5.65}$$

where

$$\Phi = \frac{\int_{\eta}^{\infty} l^2_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i]) f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])\,dz[i]}{D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}\right)}. \tag{5.66}$$

In (5.66), η is the zero of $l_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])$ in (5.23), $f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])$ is given by (5.21) and the Kullback-Leibler divergence, $D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}\right)$, can be expressed using (5.5) and (5.21) as

$$D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]} \| f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}\right) = \frac{e^{\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2}\int_0^{\infty}\sum_{k=0}^{M-1}\frac{z[i]^k}{k!}\left[2\left(z[i]\sigma_T^2\right)^{\frac{-k}{2}}\left(\mathrm{K}_{-k}\left(\frac{\sqrt{z[i]}}{\frac{\sigma_T}{2}}\right) - \right.\right.$$
$$\left.\left.\frac{k\sigma_T}{\sqrt{z[i]}}\mathrm{K}_{1-k}\left(\frac{\sqrt{z[i]}}{\frac{\sigma_T}{2}}\right)\right) + \frac{\sigma_N^2}{R}\sum_{r=1}^{R}\frac{e^{\frac{-z[i]}{s_r}-\frac{s_r}{\sigma_T^2}}}{(s_r)^k}\left(\frac{k}{z[i]}-\frac{1}{s_r}\right)\right] \times \ln\left\{\frac{(\sigma_N^2)^M\Gamma(M)e^{\frac{z[i]}{\sigma_N^2}+\frac{\sigma_N^2}{\sigma_T^2}}}{\sigma_T^2 z[i]^{M-1}}\times\right.$$
$$\sum_{k=0}^{M-1}\frac{z[i]^k}{k!}\left[2\left(z[i]\sigma_T^2\right)^{\frac{-k}{2}}\left(\mathrm{K}_{-k}\left(\frac{\sqrt{z[i]}}{\frac{\sigma_T}{2}}\right) - \frac{k\sigma_T}{\sqrt{z[i]}}\mathrm{K}_{1-k}\left(\frac{\sqrt{z[i]}}{\frac{\sigma_T}{2}}\right)\right) + \right.$$
$$\left.\left.\frac{\sigma_N^2}{R}\sum_{r=1}^{R}\frac{e^{\frac{-z[i]}{s_r}-\frac{s_r}{\sigma_T^2}}}{(s_r)^k}\left(\frac{k}{z[i]}-\frac{1}{s_r}\right)\right]\right\}dz[i], \tag{5.67}$$

where s_r is given by (5.22) and R is the number of rectangles. It is worth noting that η in (5.66) is calculated numerically. The resulting Kullback-Leibler divergence in (5.67), $l_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])$ of (5.23) and $f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}(z[i])$ of (5.21) can then be substituted into (5.65) and (5.66) to give

$$\overline{T}_{d_{Ray}} \leq \frac{\sigma_T^2 e^{\frac{-\sigma_N^2}{\sigma_T^2}}(\gamma+\Phi)}{\int_0^\infty \mathcal{I}\left(\ln\left\{\left((\sigma_N^2)^M\Gamma(M)\exp\left(\frac{z[i]}{\sigma_N^2}+\frac{\sigma_N^2}{\sigma_T^2}\right)\right)/\left(\sigma_T^2 z[i]^{M-1}\right)\right\}+\ln\{\mathcal{I}\}\right)dz[i]}, \tag{5.68}$$

where

$$\begin{aligned}\mathcal{I} = \sum_{k=0}^{M-1}\frac{z[i]^k}{k!}\Bigg[&2\left(z[i]\sigma_T^2\right)^{\frac{-k}{2}}\left(\mathrm{K}_{-k}\left(\frac{\sqrt{z[i]}}{\frac{\sigma_T}{2}}\right)-\frac{k\sigma_T}{\sqrt{z[i]}}\mathrm{K}_{1-k}\left(\frac{\sqrt{z[i]}}{\frac{\sigma_T}{2}}\right)\right)+\\ &\frac{\sigma_N^2}{R}\sum_{r=1}^{R}\frac{e^{\frac{-z[i]}{s_r}-\frac{s_r}{\sigma_T^2}}}{(s_r)^k}\left(\frac{k}{z[i]}-\frac{1}{s_r}\right)\Bigg],\end{aligned} \tag{5.69}$$

and

$$\Phi = \frac{\int_\eta^\infty \mathcal{I}\times\left(\ln\left\{\left((\sigma_N^2)^M\Gamma(M)\exp\left(\frac{z[i]}{\sigma_N^2}+\frac{\sigma_N^2}{\sigma_T^2}\right)\right)/\left(\sigma_T^2 z[i]^{M-1}\right)\right\}+\ln\{\mathcal{I}\}\right)^2 dz[i]}{\int_0^\infty \mathcal{I}\left(\ln\left\{\left((\sigma_N^2)^M\Gamma(M)\exp\left(\frac{z[i]}{\sigma_N^2}+\frac{\sigma_N^2}{\sigma_T^2}\right)\right)/\left(\sigma_T^2 z[i]^{M-1}\right)\right\}+\ln\{\mathcal{I}\}\right)dz[i]}. \tag{5.70}$$

As $\overline{T}_{f_G}, \overline{T}_{f_{Ray}} \to \infty$, $\gamma \to \infty$. Therefore, it is desirable in practice to analyze the detection performance asymptotically as it will give a low false alarm rate [44]. With the aid of Theorem 1 in [94], the asymptotic worst-case detection delay can be approximated [43, 44]. Hence, the asymptotic detection delay, for Gaussian and independent Rayleigh channels can be written respectively using (5.62) and (5.67) as

$$\begin{aligned}\overline{T}_{d_G} &\sim \frac{\gamma}{D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}\|f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}\right)}\\ &\sim \frac{\gamma}{-M\ln\{\sigma_N^2+\sigma_X^2\}+\frac{M\sigma_X^2}{\sigma_N^2}+M\ln\{\sigma_N^2\}},\end{aligned} \tag{5.71}$$

$$\begin{aligned}\overline{T}_{d_{Ray}} &\sim \frac{\gamma}{D\left(f^{(1)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}\|f^{(0)}_{\mathbf{Y}[i]^\dagger\mathbf{Y}[i]}\right)}\\ &\sim \frac{\sigma_T^2 e^{\frac{-\sigma_N^2}{\sigma_T^2}}\gamma}{\int_0^\infty \mathcal{I}\left(\ln\left\{\left((\sigma_N^2)^M\Gamma(M)\exp\left(\frac{z[i]}{\sigma_N^2}+\frac{\sigma_N^2}{\sigma_T^2}\right)\right)/\left(\sigma_T^2 z[i]^{M-1}\right)\right\}+\ln\{\mathcal{I}\}\right)dz[i]},\end{aligned} \tag{5.72}$$

where $\mathcal{I}$ is given in (5.69).

Therefore, the upper bound for the worst-case detection delay given by (5.63) and (5.68) and the asymptotic detection delay given by (5.71) and (5.72) for both Gaussian and independent Rayleigh cases, respectively, can be used to evaluate the quickest spectrum sensing performance. These are shown in Figures 5.2 and 5.3, which will be discussed in detail in Section 5.9.

5.9 ANALYTICAL AND NUMERICAL RESULTS

In this section, we present some analytical and simulated results to evaluate quickest spectrum sensing performance in Gaussian and independent Rayleigh channels, employing multiple receive antennas at the CU. The analytical results follow from the theoretical analysis of the sensing performance discussed in Section 5.8. In addition, we also present some numerical results to evaluate the sensing performance in Rayleigh (independent and correlated) and Rician (independent) channels when the CU is equipped with multiple antennas. We assume that the PU begins transmission at $\tau = 100$. In all simulations, each point on the plot represents the average values observed over 20000 trials, where each trial consists of 2500 samples. The number of rectangles required for the Riemann sum in (5.21) is set to $R = 50$.

5.9.1 Analytical results

In order to analyze the theoretical sensing performance, the thresholds, γ, for both Gaussian and independent Rayleigh cases that we consider (i.e. for Figures 5.2, 5.3 and 5.4) are set using the respective lower bounds on the mean number of samples between false alarms, $\overline{T}_{f_G}$ and $\overline{T}_{f_{Ray}}$.

Figure 5.2 compares the simulated and analytical results for the case when the CU is equipped with single and multiple antennas (M=3) and the Gaussian signal is observed over a time-invariant channel at SNR = 4.77 dB[4]. We observe that with M=3 antennas at the CU, the detection delay, $\overline{T}_{d_G}$ reduces substantially compared to M=1. The reduction in detection delay is at least 6 samples which represents at least a 6-fold improvement.

[4]The SNR value of 4.77 dB is chosen to allow comparison with [36].

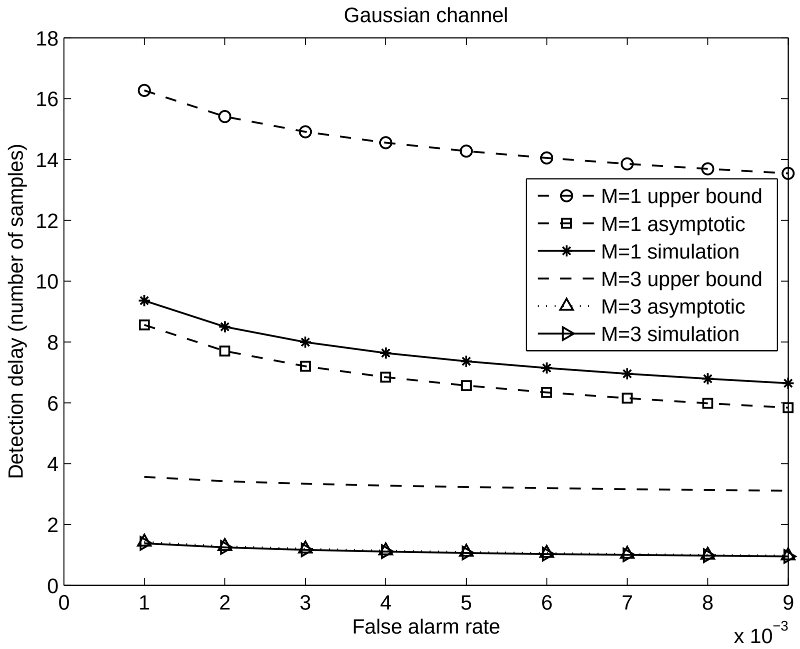

Figure 5.2 Comparison between the performance of the CU equipped with single (M=1) and multiple (M=3) antennas for a Gaussian channel with SNR=4.77 dB.

In Figure 5.3, we present a comparison between the simulated and analytical sensing performance in Gaussian and independent Rayleigh channels at SNR=10 dB when the CU is equipped with M=3 antennas. Comparing $\overline{T}_{d_G}$ with $\overline{T}_{d_{Ray}}$ in Figure 5.3, we can see that the sensing performance degrades in the Rayleigh fading channel as compared to a Gaussian channel due to the faded received signals. However, the impact of the propagation conditions is small compared to the effects of different numbers of antennas as shown from a comparison of Figures 5.2 and 5.3.

Figure 5.4 shows the performance improvements due to adding more antennas at the CU for Gaussian and independent Rayleigh scenarios. The improvements are due to the increased spatial diversity provided. In Figures 5.2 and 5.3, the simulation results are close to the asymptotic analysis and the upper bound is very loose. Hence, the asymptotic result provides a good indicator of sensing performance. Deviations of the asymptotic results from the simulations are due to the fact that the theory involves a large threshold value and an asymptotically small false alarm rate.

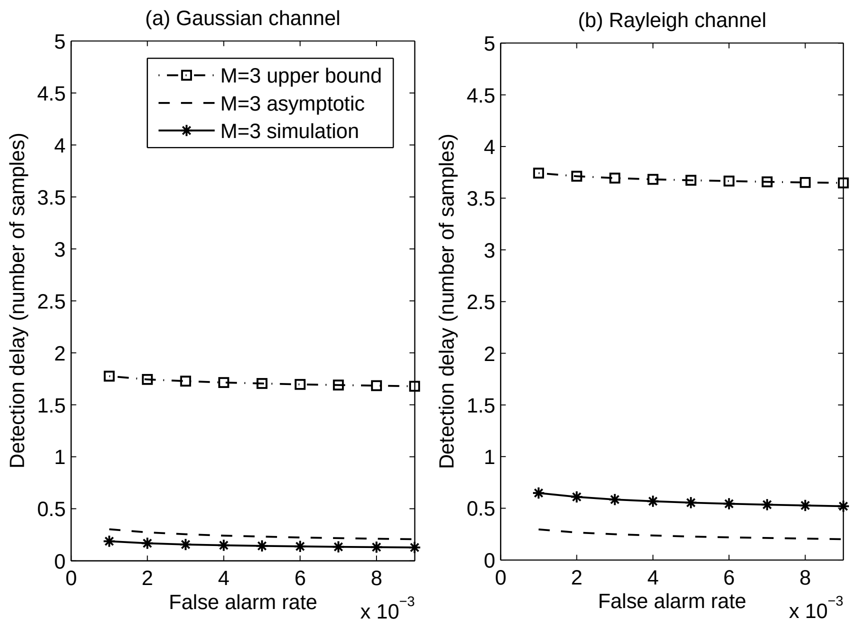

Figure 5.3 Simulation and analytical sensing performance in (a) Gaussian and (b) independent Rayleigh channels at SNR=10 dB.

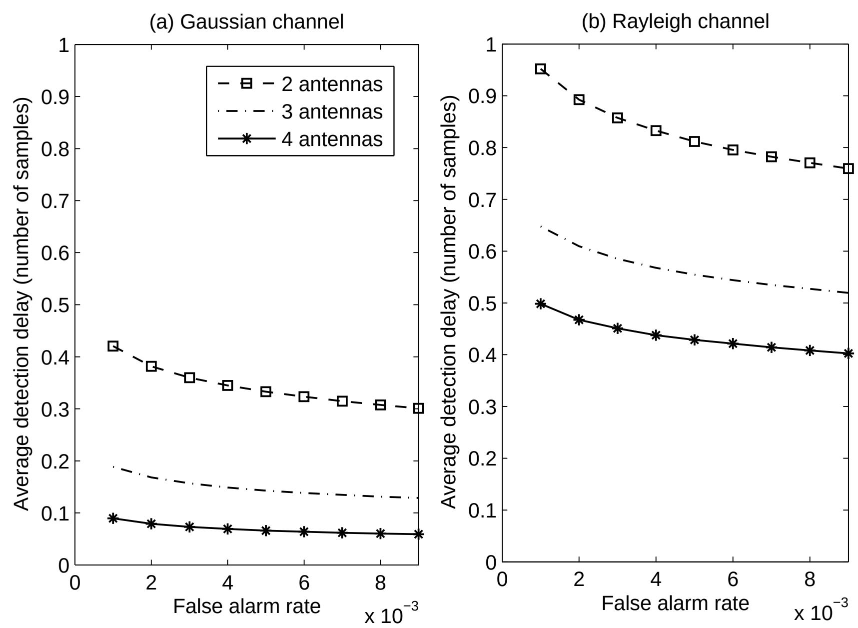

Figure 5.4 The effect of different number of antennas on the simulated performance of sensing in (a) Gaussian and (b) independent Rayleigh channels at SNR=10 dB.

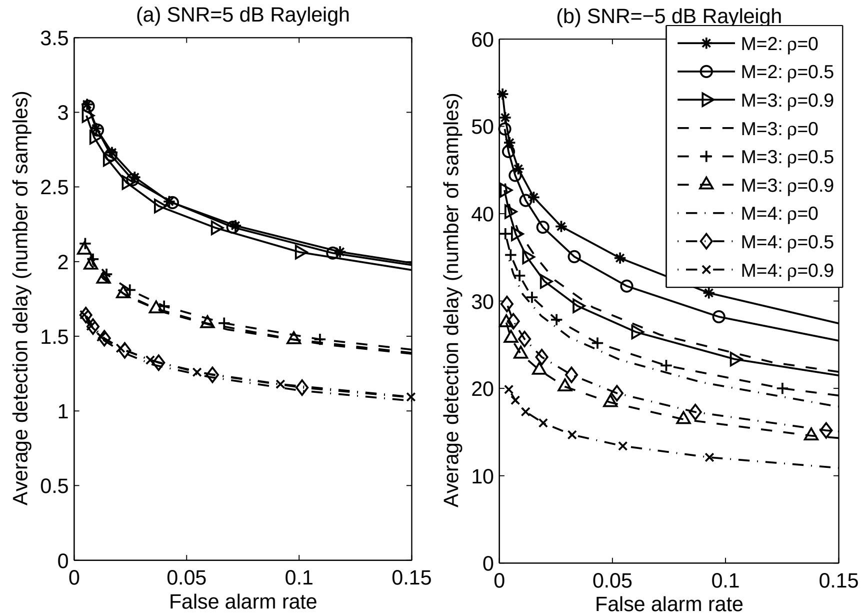

Figure 5.5 Comparison between the sensing performance in i.i.d Rayleigh ($\rho = 0$) and correlated Rayleigh channels ($\rho = 0.5$ and 0.9) with multiple antenna CUs at SNR=5 dB in (a) and SNR=-5 dB in (b).

5.9.2 Numerical results

Here, we consider CUSUM threshold values of $\gamma \in \{3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7\}$, following the threshold values in Chapter 3, for Figures 5.5, 5.7, 5.9 and 5.10, whereas for Figure 5.8, we consider $\gamma \in \{0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5\}$. The threshold is set in these ranges (chosen by trial and error) so that a reasonable range of false alarm rate and average detection delay values are achieved for the numerical results.

Figure 5.5 compares the sensing performance of a multiple antenna CU in i.i.d and correlated Rayleigh channels for various values of ρ at SNR=5 dB and -5 dB. Results show that performance improves as the number of antennas and SNR increases. In addition, results show that increased channel correlation has little impact at SNR=5 dB, but helps to reduce the average detection delay at SNR=-5 dB.

The effect of correlation on the sensing performance can be explained by looking at the moments of the total power of the signal. Consider the extreme cases of i.i.d ($\rho = 0$) and perfectly correlated ($\rho = 1$) Rayleigh channels. Here, the total power of the signal for both of the cases can be expressed as

$$P_{iid} = \sum_{m=1}^{M} |H_m[i]|^2 |S[i]|^2, \tag{5.73}$$

and

$$P_{corr} = \sum_{m=1}^{M} |H[i]|^2 |S[i]|^2, \tag{5.74}$$

where $E\left[|H_m[i]|^2\right] = 1$, $H[i] = H_1[i] = ... = H_M[i]$ and the subscripts in (5.73) and (5.74) denote the i.i.d and perfectly correlated cases respectively. Let $|S[i]|^2 = \sigma_S^2 U[i]$, where $U[i]$ is a standard exponential random variable. The mean and variance of P_{iid} and P_{corr} can be written respectively as

$$\begin{aligned} E[P_{iid}] &= E\left[\sum_{m=1}^{M} |H_m[i]|^2 |S[i]|^2\right] \\ &= M\sigma_S^2, \end{aligned} \tag{5.75}$$

$$\begin{aligned} \mathrm{Var}(P_{iid}) &= \sigma_S^4 \mathrm{Var}\left(U[i] \sum_{m=1}^{M} |H_m[i]|^2\right) \\ &= \sigma_S^4 \left(E\left[U[i]^2\right] E\left[\left(\sum_{m=1}^{M} |H_m[i]|^2\right)^2\right] - (E[U[i]])^2 \left(E\left[\sum_{m=1}^{M} |H_m[i]|^2\right]\right)^2\right) \\ &= \sigma_S^4 \left(2E\left[\sum_{m=1}^{M} |H_m[i]|^2 \sum_{n=1}^{M} |H_n[i]|^2\right] - M^2\right) \\ &= \sigma_S^4 \left(2\left[M \times E\left[|H_m[i]|^4\right] + M(M-1)E\left[|H_m|^2\right] E\left[|H_n|^2\right]\right] - M^2\right) \\ &= \sigma_S^4 (M^2 + 2M), \end{aligned} \tag{5.76}$$

$$\begin{aligned} E[P_{corr}] &= E\left[\sum_{m=1}^{M} |H[i]|^2 |S[i]|^2\right] \\ &= M\sigma_S^2, \end{aligned} \tag{5.77}$$

$$\begin{aligned} \mathrm{Var}(P_{corr}) &= M^2 \sigma_S^4 \mathrm{Var}\left(U[i] |H[i]|^2\right) \\ &= M^2 \sigma_S^4 \left(E\left[|H[i]|^4\right] E\left[U[i]^2\right] - \left(E\left[|H[i]|^2\right]\right)^2 (E(U[i]))^2\right) \\ &= 3M^2 \sigma_S^4. \end{aligned} \tag{5.78}$$

It is worth noting that (5.75)-(5.78) are derived using the fact that $|H_m[i]|^2$, $|H_n[i]|^2$ and $|S[i]|^2$ have exponential distributions. Based on these results, we observe that although P_{iid} and P_{corr}

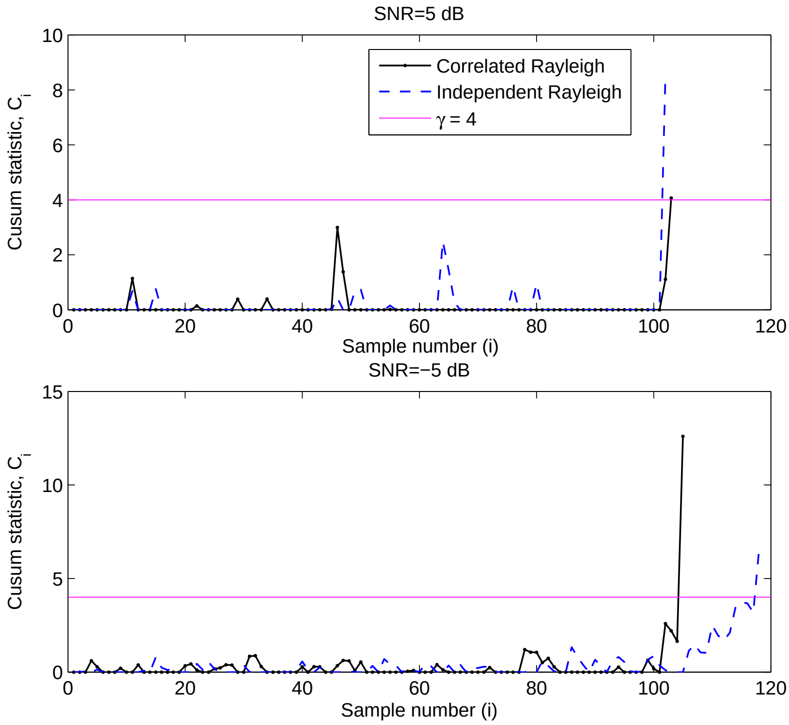

Figure 5.6 Simulated CUSUM statistic for i.i.d and correlated Rayleigh cases at SNR=5 dB and -5 dB when threshold, $\gamma = 4$.

have the same mean, P_{corr} has a higher variance than P_{iid}. When the SNR is low, the CUSUM detector has small average jumps and takes a long time to cross the threshold. In this situation, a larger variance and, therefore the occasional larger jump is beneficial, which accelerates the threshold crossing. This is illustrated by traces of the CUSUM process at SNR=-5 dB in Figure 5.6, which showed a few larger jumps dominating the threshold crossing. Hence, at low SNR, correlation leading to a higher variance is beneficial. In contrast, at high SNR, the CUSUM jumps are larger and a threshold crossing occurs rapidly even with average sized jumps. Hence, the occasional large jump has little effect and correlation causing increased variance has little effect. Again, this can also be seen by the traces of the CUSUM process at SNR=5 dB from Figure 5.6, which did not show an important role being played by unusually large jumps.

Table 5.1 shows the minimum number of antennas needed to detect the PU transmission at a given channel correlation, ρ, in less than a 2 sample average delay for a false alarm rate of 0.05 at various SNR. The key conclusion is that multiple antennas can be used to reduce detection delay for weaker PU signals. However, the need for antennas is much greater when the spatial correlations are small. Note that sensing in the low SNR region is assisted by

Table 5.1 Minimum number of antennas required for an average detection delay ≤ 2 samples with FAR = 0.05 at various SNR

ρ	Minimum number of antennas			
	10 dB	5 dB	0 dB	-5 dB
0	1	3	12	82
0.5	1	2	11	60
0.7	1	2	10	45
0.9	1	2	9	30

correlation, whereas data transmission traditionally prefers independent channels. Hence, there are competing demands on the antenna array, which may be better met using a reconfigurable antenna array.

Figure 5.7 compares performance of single, discussed in Chapter 3, and multiple antenna CUs in independent Rician channels with K=0 dB and 6 dB at SNRs of 5 dB and -5 dB. We can see that at high SNR, the performance gain achieved by employing multiple antennas gives a reduction of at least 2 samples in the average delay. Also, at both SNRs, increased values of K tend to improve sensing performance. As expected, a LOS component stabilizes the received signal and makes it easier to detect. The M=1, SNR=-5 dB scenario is an exception to this trend. Here, the received SNR is lowest (as there is no diversity gain for M=1) and the signal is noise-limited. As a result, the detection results are insensitive to the type of channel and the K=0 dB and K=6 dB results are indistinguishable. From Figures 5.5 and 5.7, we observe the benefits of both correlation and LOS. This motivates correlated Rician channels as an area for future research.

The same trends as in Figures 5.5 and 5.7 can be observed at SNR=-20 dB in Figure 5.8. However, in order for a CUSUM detector to have a reasonable false alarm rate, lower threshold values are required due to small increments in the CUSUM detector, resulting in long delays. Thus, the average detection delay in this case is substantially longer than Figures 5.5 and 5.7. This motivates cooperative quickest spectrum sensing with multiple receive antennas for future work, which could possibly help the sensing performance at SNR=-20 dB, where the received signal is very weak. This possible extension will be discussed in Section 7.2.

Figures 5.9 and 5.10 compare sensing performance for the correlated Rayleigh (ρ = 0.5) and independent Rician case (K=0 dB) at SNR=10 dB and 5 dB, respectively, for mis-matched

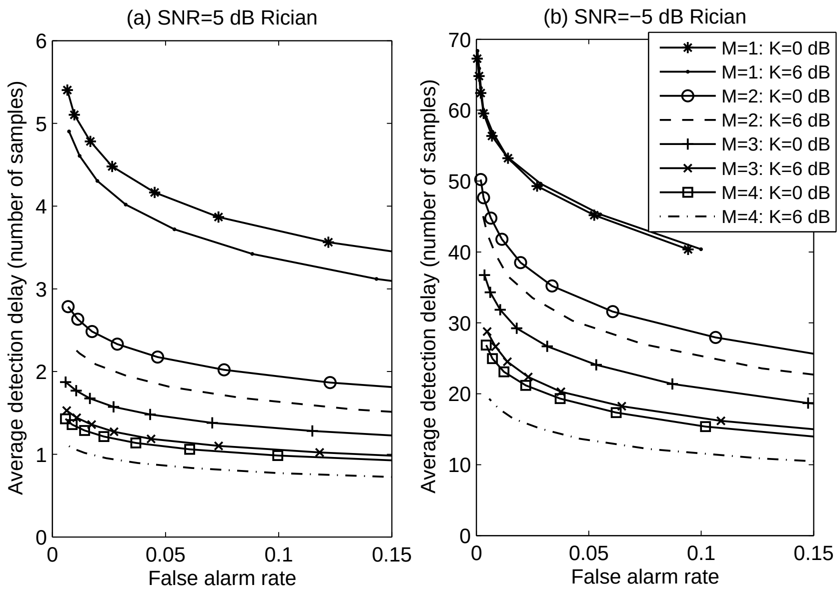

Figure 5.7 Simulated sensing performance for different numbers of antennas in independent Rician channels with K=0 dB, 6 dB at (a) SNR=5 dB (b) SNR=-5 dB.

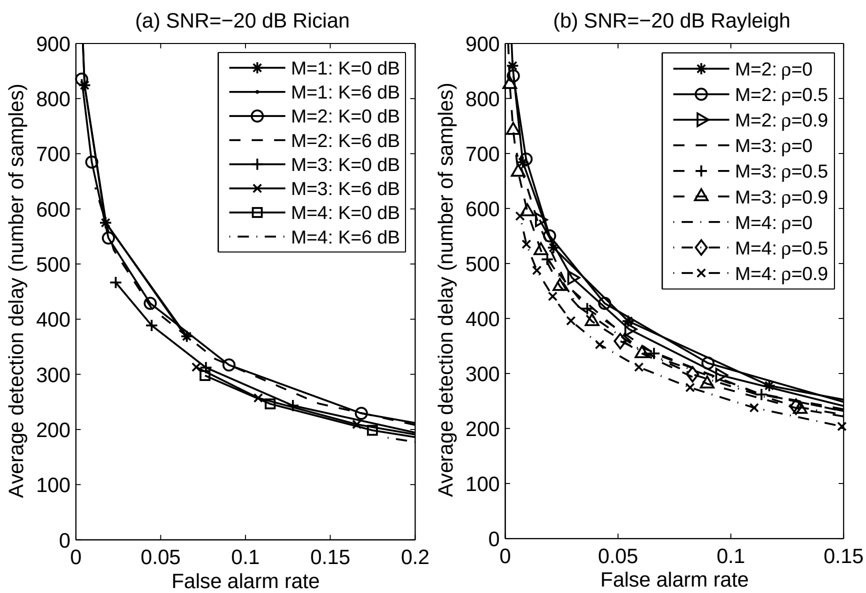

Figure 5.8 Simulated sensing performance in (a) Rician and (b) i.i.d Rayleigh ($\rho = 0$) and correlated Rayleigh channels ($\rho = 0.5$ and 0.9) with multiple antenna CUs at SNR=-20 dB.

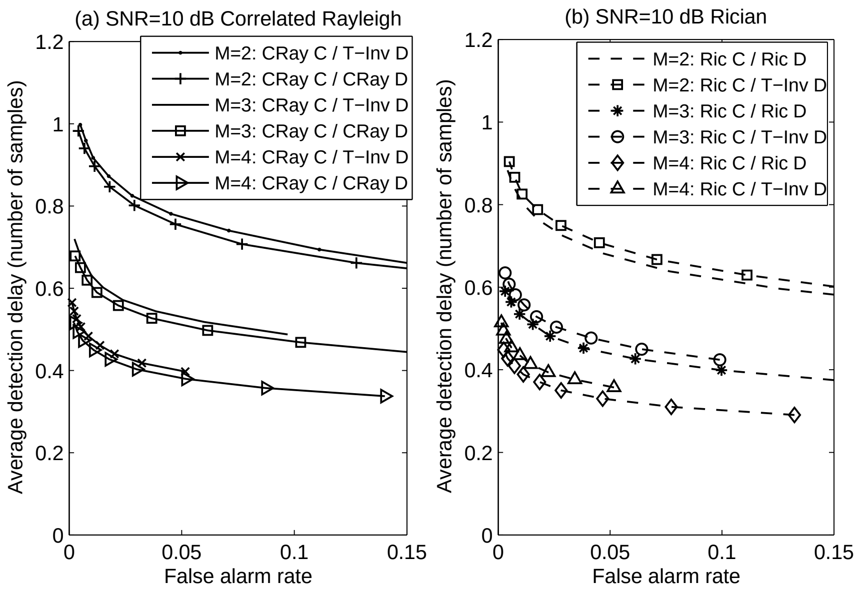

Figure 5.9 Performance comparison between correct and mis-matched detectors (D) in different channels (C) at SNR=10 dB (correlated Rayleigh (CRay) with $\rho = 0.5$ / independent Rician (Ric) with K=0 dB / Time-invariant (T-Inv) cases).

channels, in which the CUSUM detector is designed for one channel, but experiences another. In both Figures 5.9 and 5.10, we observe that for both correlated Rayleigh and Rician cases, the actual channel experienced by the PU signal has a greater impact on the sensing performance than the channel used to design the detector. For example, the average detection delay for the correlated Rayleigh detector is close to the time-invariant detector when the received signal is transmitted over correlated Rayleigh channels at both SNRs. Similar trends at both SNRs can be observed over independent Rician channels and as the number of antennas varies in both scenarios. Hence, the detection technique is robust to errors in channel identification.

5.10 CHAPTER SUMMARY

In this chapter, we have studied quickest spectrum sensing for CUs equipped with multiple receive antennas when the received signal is transmitted over Gaussian, independent and correlated Rayleigh and independent Rician fading channels. We first showed and proved that for the Gaussian and independent Rayleigh scenario, the sum of the complex received signal powers at each antenna is a sufficient statistic. This allowed us to use a log likelihood ratio based on the sum of the received signal powers. Hence, we derived the pdf of the power sum (the

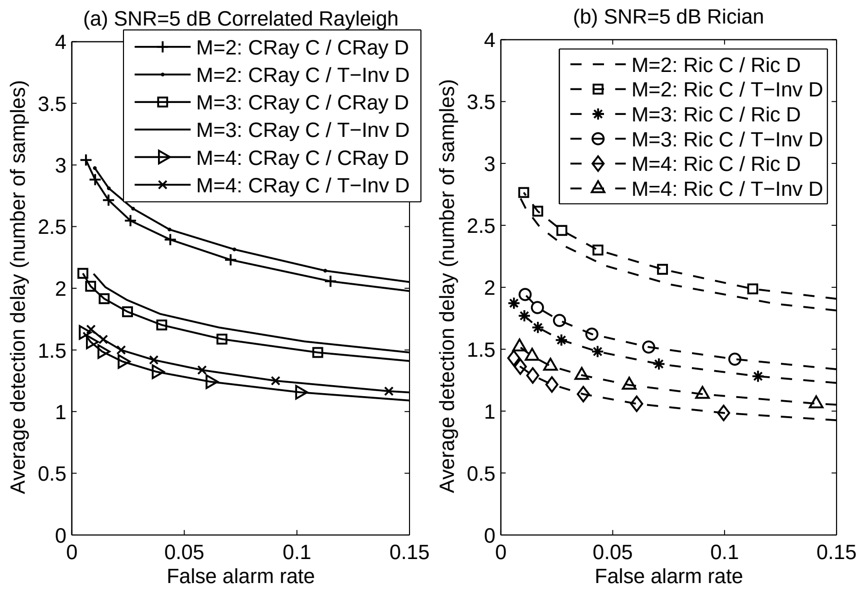

Figure 5.10 Performance comparison between correct and mis-matched detectors (D) in different channels (C) at SNR=5 dB (correlated Rayleigh (CRay) with $\rho = 0.5$ / independent Rician (Ric) with K=0 dB / Time-invariant (T-Inv) cases).

output of the EGC) in order to construct the log likelihood ratio. The derivation of the pdf for the independent Rayleigh scenario uses a technique which avoids numerical integration. We also derived the joint pdfs of the received signals for the correlated Rayleigh and independent Rician fading scenarios. In addition, we derived an analytical performance analysis, including the upper bound and asymptotic worst-case detection delay for both Gaussian and independent Rayleigh scenarios. The sensing performance for the Rayleigh (independent and correlated) and independent Rician cases were evaluated numerically.

The numerical analysis and simulation results illustrate the performance gains that can be achieved in all cases by employing multiple antennas at the CU due to the spatial diversity provided. Furthermore, channel correlation has little impact on the sensing performance at high SNR, whereas at low SNR, increasing correlation between Rayleigh channels improves sensing performance. In particular, the detection of weak signals is significantly assisted by spatial correlation at the CU array. Simulation results show that the sensing performance increases with an increasing Rician K-factor value. In the event of mis-matched channels, simulations demonstrated that at a particular correlation coefficient or Rician K-factor, the sensing performance is sensitive to the true channel condition irrespective of the number of CU antennas.

Chapter 6

DISTRIBUTION OF DETECTION DELAY FOR MULTI-ANTENNA QUICKEST SENSING

6.1 INTRODUCTION

In Chapter 5, we studied quickest spectrum sensing performance for cognitive users with multiple receive antennas when the received signal experiences Gaussian, Rayleigh (independent and spatially correlated) and independent Rician channels. The quickest spectrum sensing performance was evaluated theoretically and numerically in terms of detection delay and false alarm rate. No studies have appeared which give a theoretical expression for the distribution of detection delay for multi-antenna quickest spectrum sensing. Such an expression would be beneficial in further analyzing the sensing performance with multiple receive antennas at the CU, especially for longer delays or in the lower SNRs.

Results from Chapter 5 illustrated that in the event of a mis-matched channel, where the CUSUM detector is designed for a specific channel, but experiences a different channel, the quickest spectrum sensing performance at a particular correlation coefficient or Rician K-factor depends greatly on the true channel and is relatively insensitive to the CUSUM detector. Since the fading channel is usually unknown in the cognitive radio network, and quickest spectrum sensing performance is insensitive to the designed detector, it is reasonable to employ a simple multi-antenna time-invariant detector to detect the PU transmission. Recall from Chapter 4 that the Brownian motion approach provides the best approximation of the distribution of detection delay for longer delays.

Therefore, in this chapter, we derive theoretical expressions for the distribution of detection delay for a multiple antenna time-invariant CUSUM detector when the received signal experiences Rayleigh (independent and spatially correlated) and Rician (independent and spatially correlated) channels. In particular, we derive the expected value and variance of the received signal transmitted over a correlated Rician channel, which is useful in deriving the approximate expression for the distribution of detection delay. The derivation of the approximate expressions for the detection delay distribution for all of the cases are based on the Brownian motion approach. We verify the validity of the approximate expressions for the distribution of detection delay for each of the cases, employing multiple antennas at the CU, using simulations. In addition, we also study the effects of channel correlation, LOS strength and the employment of multiple receive antennas at the CU on the probability of detection delay, especially for long delays or at low SNR.

The rest of the chapter is organized as follows. Section 6.2 describes the system model including a brief description of the modified detection delay statistic. Section 6.3 presents the derivation of the theoretical expressions for the distribution of detection delay for a multi-antenna time-invariant CUSUM detector, when the received signal is transmitted over Rayleigh (independent and spatially correlated) and Rician (independent and spatially correlated) channels. Numerical results to validate the approximate expressions for each of the cases are given in Section 6.4. Finally, Section 6.5 ends the chapter with some concluding remarks.

6.2 SYSTEM MODEL

It is assumed that the PU is initially inactive and a CU equipped with multiple receive antennas attempts to detect the PU transmission. The PU signal is assumed to be a narrowband complex Gaussian signal. Let $Y_m[i]$ denote the received signal at antenna m, where $m = 1, 2, ..., M$ and i is the sample number of the received signal. If the PU is inactive, then

$$Y_m[i] = N_m[i], \tag{6.1}$$

where $N_m[i]$ is independent circularly symmetric complex white Gaussian noise such that $N_m[i] \sim \mathcal{CN}(0, \sigma_N^2)$. If the PU is active, then

$$Y_m[i] = H_m[i] \times S[i] + N_m[i], \tag{6.2}$$

where $H_m[i]$ is the channel coefficient and $S[i]$ is the PU signal such that $S[i] \sim \mathcal{CN}(0, \sigma_S^2)$ is an independent circularly symmetric complex Gaussian random variable with variance σ_S^2. Let $X_m[i] = H_m[i] \times S[i]$, where $X_m[i]$ is a circularly symmetric complex Gaussian variable with variance σ_X^2.

We first consider the case when there is insufficient separation between the antennas at the CU resulting in spatially correlated Rician fading channels. Hence, $H_m[i]$ is correlated between antennas, but independent between samples. The spatially correlated Rician fading channel, $\mathbf{H}[i] = [H_1[i], H_2[i], ..., H_m[i]]^T$, can be modelled by

$$\mathbf{H}[i] = \beta \mathbf{V}[i] + \alpha \mathbf{R}^{1/2} \mathbf{U}[i], \tag{6.3}$$

where

$$\beta = \sqrt{\frac{K}{K+1}}, \qquad \mathbf{V}[i] = \mathbf{H}^{\mathrm{LOS}}[i], \tag{6.4}$$

and

$$\alpha = \sqrt{\frac{1}{K+1}}, \qquad \mathbf{U}[i] = \mathbf{H}^{\mathrm{SC}}[i]. \tag{6.5}$$

$\mathbf{R}$ in (6.3) is the $M \times M$ antenna correlation matrix denoted by $E\left[\mathbf{H}[i]\mathbf{H}[i]^\dagger\right]$. It is assumed that $\mathbf{R}$ follows an exponential correlation model such that $\mathbf{R}_{qr} = \rho^{|q-r|}$, where $q, r = 1, 2, ..., M$ and $0 \leq \rho \leq 1$. Note that this model is most relevant to uniform linear arrays (ULAs). In (6.4), $\mathbf{H}^{\mathrm{LOS}}[i]$ is the LOS component given by [129]

$$\mathbf{H}^{\mathrm{LOS}}[i] = \left[1, e^{j\Delta}, ..., e^{j(M-1)\Delta}\right]^T, \tag{6.6}$$

where

$$\Delta = 2\pi d \cos\theta, \tag{6.7}$$

in which d is the antenna spacing in wavelengths and θ is the angle of the LOS component at the receiver. Here we are assuming a ULA is deployed. In (6.5), $\mathbf{H}^{\mathrm{SC}}[i]$ is the scattered component (SC) which is a spatially white $M \times 1$ vector such that $\mathbf{H}^{\mathrm{SC}}[i] \sim \mathcal{CN}(0, 1)$. When $\rho = 0$, the correlated Rician channel, $\mathbf{H}[i]$ in (6.3) reduces to an independent and identically distributed (i.i.d) Rician channel. When $K = 0$, $\mathbf{H}[i]$ in (6.3) reduces to a correlated Rayleigh channel, whereas, when $K = 0$ and $\rho = 0$, it reduces to i.i.d Rayleigh. Therefore, in this chapter, the derivation of the mathematical expression for the distribution of detection delay will be conducted for the correlated Rician case and based on the corresponding result, the theoretical expression for detection delay distribution for the other cases (i.e. i.i.d Rician, i.i.d Rayleigh and correlated Rayleigh) can be obtained by changing the parameters K or ρ.

Let $\mathbf{Y}[i] = [Y_1[i], Y_2[i], ..., Y_M[i]]^T$ and $z[i] = \mathbf{Y}[i]^\dagger \mathbf{Y}[i]$. It is assumed that the PU begins transmission at some sample τ and the PU signal is detected by the CU using the CUSUM algorithm. Recall from Section 2.4.1 that the PU signal is detected by the CUSUM algorithm at sample T and the recursive form of the CUSUM statistic is denoted in (2.21) as

$$C_{n+1} = \{C_n + l_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[n+1])\}^+, \tag{6.8}$$

where $x^+ = \max(x, 0)$ and $l_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])$ is the log likelihood ratio, which can be expressed as $l_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) = \ln \left\{ f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) / f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) \right\}$, where $f^{(1)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])$ and $f^{(0)}_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])$ are the probability density function (pdf) of the received signal, $\mathbf{Y}[i]$, when the PU is present and absent, respectively. Recall from Section 5.4 that the log likelihood ratio for a multi-antenna time-invariant CUSUM detector can be written as in (5.7), which gives

$$l_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) = \frac{z[i]\sigma_X^2}{\sigma_N^2(\sigma_N^2 + \sigma_X^2)} + M \ln \left\{ \frac{\sigma_N^2}{\sigma_N^2 + \sigma_X^2} \right\}. \tag{6.9}$$

If $T > \tau$, then a detection delay, $\delta = T - \tau$ will occur. Let $P(\delta)$ denote the probability that the detection delay is δ samples. Hence, this defines the distribution of detection delay for multi-antenna quickest spectrum sensing.

6.2.1 Modified detection delay statistic

As discussed earlier in Section 4.2.1, it is difficult to analytically handle the CUSUM statistic, C_n, due to the max operation in (6.8). This is because the max operation causes a complex dependence between C_n and the log likelihood ratio, $l_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i])$. Therefore, in general, no exact analysis is possible. Following the approach taken in Section 4.2.1, we could approximate the CUSUM process by the modified detection delay statistic, $\mathcal{D}_n$, denoted by

$$\begin{aligned}\mathcal{D}_n &= \sum_{i=1}^{n} l_{\mathbf{Y}[i]^\dagger \mathbf{Y}[i]}(z[i]) \\ &= \sum_{i=1}^{n} \left(\frac{z[i]\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)} + M \ln \left\{ \frac{\sigma_N^2}{\sigma_N^2+\sigma_X^2} \right\} \right),\end{aligned} \tag{6.10}$$

where D_n is a sum of a fixed number of terms. It is worth noting that the modified CUSUM process in (6.10) is only used to analytically approximate the CUSUM process whereas the exact CUSUM algorithm in (6.8) and (6.9) is used for all simulated results. Let $z_i = z[i]$, $a = \frac{\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)}$ and $b = -M \ln \left\{ \frac{\sigma_N^2}{\sigma_N^2+\sigma_X^2} \right\}$. Hence, $a > 0$, $b > 0$ and we can rewrite (6.10) as

$$\mathcal{D}_n = \sum_{i=1}^{n} (a z_i - b)\,. \tag{6.11}$$

As can be seen from (6.11), the approximation of the CUSUM process based on $\mathcal{D}_n$ lies in the fact that $\mathcal{D}_n$ can take negative values whereas C_n in (6.8) is restricted to $C_n \geq 0$. However, when $\mathcal{D}_n$ has positive values, it is the same as C_n. Recall that the validity of this approximation has been discussed and shown in Section 4.2.1 and 4.4.1.

6.3 DISTRIBUTION OF DETECTION DELAY BASED ON THE BROWNIAN MOTION APPROACH

As shown in Section 4.4, the Brownian approximation provides a remarkably simple and accurate approximation for the distribution of detection delay for longer delays. Therefore, following the approach taken in Section 4.3.1.4 and 4.3.2.3, we can also approximate $P(\delta)$ for multi-antenna quickest spectrum sensing in correlated Rician channels by employing the theory of Brownian motion with drift. From (6.11), the $\mathcal{D}_n$ process has i.i.d increments of the form $az_i - b$ and we denote each increment or jump as

$$J_i = az_i - b. \tag{6.12}$$

In order to find the expected value and variance of each J_i, we first need to find the expected value and variance of z_i. Let $\mathbf{Y}[i] = \mathbf{H}_i \times S_i + \mathbf{N}_i$ and thus, using this new notation, $\mathbf{H}[i]$ in (6.3) can be expressed as $\mathbf{H}_i = \beta \mathbf{V}_i + \alpha \mathbf{R}^{1/2}\mathbf{U}_i$. Since $z_i = z[i] = \mathbf{Y}[i]^\dagger \mathbf{Y}[i]$, we can rewrite z_i as

$$\begin{aligned} z_i = \beta^2 S_i^\dagger \mathbf{V}_i^\dagger \mathbf{V}_i S_i + \beta\alpha S_i^\dagger \mathbf{V}_i^\dagger \mathbf{R}^{1/2}\mathbf{U}_i S_i + \beta\alpha S_i^\dagger \mathbf{U}_i^\dagger \mathbf{R}^{1/2}\mathbf{V}_i S_i + \alpha^2 S_i^\dagger \mathbf{U}_i^\dagger \mathbf{R}\mathbf{U}_i S_i + \\ \beta S_i^\dagger \mathbf{V}_i^\dagger \mathbf{N}_i + \alpha S_i^\dagger \mathbf{U}_i^\dagger \mathbf{R}^{1/2}\mathbf{N}_i + \beta \mathbf{N}_i^\dagger \mathbf{V}_i S_i + \alpha \mathbf{N}_i^\dagger \mathbf{R}^{1/2}\mathbf{U}_i S_i + \mathbf{N}_i^\dagger \mathbf{N}_i. \end{aligned} \tag{6.13}$$

Taking the expectation of z_i in (6.13) yields

$$\begin{aligned} E[z_i] &= E\Big[\beta^2 S_i^\dagger \mathbf{V}_i^\dagger \mathbf{V}_i S_i + \beta\alpha S_i^\dagger \mathbf{V}_i^\dagger \mathbf{R}^{1/2}\mathbf{U}_i S_i + \beta\alpha S_i^\dagger \mathbf{U}_i^\dagger \mathbf{R}^{1/2}\mathbf{V}_i S_i + \alpha^2 S_i^\dagger \mathbf{U}_i^\dagger \mathbf{R}\mathbf{U}_i S_i + \\ &\qquad \beta S_i^\dagger \mathbf{V}_i^\dagger \mathbf{N}_i + \alpha S_i^\dagger \mathbf{U}_i^\dagger \mathbf{R}^{1/2}\mathbf{N}_i + \beta \mathbf{N}_i^\dagger \mathbf{V}_i S_i + \alpha \mathbf{N}_i^\dagger \mathbf{R}^{1/2}\mathbf{U}_i S_i + \mathbf{N}_i^\dagger \mathbf{N}_i\Big] \\ &= E\left[\beta^2 S_i^\dagger \mathbf{V}_i^\dagger \mathbf{V}_i S_i + \alpha^2 S_i^\dagger \mathbf{U}_i^\dagger \mathbf{R}\mathbf{U}_i S_i + \mathbf{N}_i^\dagger \mathbf{N}_i\right]. \end{aligned} \tag{6.14}$$

Using the fact that $E\left[|V_{i_m}|^2\right] = 1$ and using the notation $\mathbf{V}_i = [V_{i_1}, V_{i_2}, ..., V_{i_M}]^T$, the term $E\left[\mathbf{V}_i^\dagger \mathbf{V}_i\right]$ in (6.14) can be expressed as

$$\begin{aligned} E\left[\mathbf{V}_i^\dagger \mathbf{V}_i\right] &= E\left[\sum_{m=1}^{M} |V_{i_m}|^2\right] \\ &= M. \end{aligned} \tag{6.15}$$

In order to evaluate $E\left[\mathbf{U}_i^\dagger \mathbf{R}\mathbf{U}_i\right]$ in (6.14), we perform an eigendecomposition of $\mathbf{R}$, which gives

$$\mathbf{R} = \phi \Lambda \phi^\dagger, \tag{6.16}$$

where ϕ is a unitary matrix and $\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, ..., \lambda_M]$, where λ_m is the mth eigenvalue of $\mathbf{R}$.

Let $\overline{\mathbf{U}}_i^{\dagger} = \mathbf{U}_i^{\dagger}\phi$, where $\overline{\mathbf{U}}_i^{\dagger} \sim \mathcal{CN}(0,1)$. Therefore, using the notation $\overline{\mathbf{U}}_i = \left[\overline{U}_{i_1}, \overline{U}_{i_2}, ..., \overline{U}_{i_M}\right]^T$,

$$\begin{aligned} E\left[\mathbf{U}_i^{\dagger}\mathbf{R}\mathbf{U}_i\right] &= E\left[\mathbf{U}_i^{\dagger}\phi\Lambda\phi^{\dagger}\mathbf{U}_i\right] \\ &= E\left[\sum_{m=1}^{M} |\overline{U}_{i_m}|^2\lambda_m\right] \\ &= \sum_{m=1}^{M}\lambda_m \\ &= M. \end{aligned} \tag{6.17}$$

In (6.14), $E\left[\mathbf{N}_i^{\dagger}\mathbf{N}_i\right]$ is given by

$$E\left[\mathbf{N}_i^{\dagger}\mathbf{N}_i\right] = M\sigma_N^2. \tag{6.18}$$

Using (6.15), (6.17) and (6.18), the expected value of z_i in (6.14) can be written as

$$\begin{aligned} E[z_i] &= E\left[\beta^2 S_i^{\dagger}\mathbf{V}_i^{\dagger}\mathbf{V}_i S_i + \alpha^2 S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}\mathbf{U}_i S_i + \mathbf{N}_i^{\dagger}\mathbf{N}_i\right] \\ &= \beta^2 M\sigma_S^2 + \alpha^2 M\sigma_S^2 + M\sigma_N^2. \end{aligned} \tag{6.19}$$

The variance of z_i is denoted by

$$\mathrm{Var}(z_i) = E\left[z_i^2\right] - \left(E[z_i]\right)^2, \tag{6.20}$$

where $E[z_i]$ can be obtained from (6.19). Using (6.13), $E\left[z_i^2\right]$ in (6.20) can be expressed as

$$\begin{aligned} E\left[z_i^2\right] = E\Big[&\left(\beta^2 S_i^{\dagger}\mathbf{V}_i^{\dagger}\mathbf{V}_i S_i\right)^2 + \beta^2\alpha^2 S_i^{\dagger}\mathbf{V}_i^{\dagger}\mathbf{V}_i S_i S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}\mathbf{U}_i S_i + \beta^2 S_i^{\dagger}\mathbf{V}_i^{\dagger}\mathbf{V}_i S_i\mathbf{N}_i^{\dagger}\mathbf{N}_i \\ &\beta^2\alpha^2 S_i^{\dagger}\mathbf{V}_i^{\dagger}\mathbf{R}^{1/2}\mathbf{U}_i S_i S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}^{1/2}\mathbf{V}_i S_i + \beta^2\alpha^2 S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}^{1/2}\mathbf{V}_i S_i S_i^{\dagger}\mathbf{V}_i^{\dagger}\mathbf{R}^{1/2}\mathbf{U}_i S_i + \\ &\beta^2\alpha^2 S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}\mathbf{U}_i S_i S_i^{\dagger}\mathbf{V}_i^{\dagger}\mathbf{V}_i S_i + \alpha^4 S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}\mathbf{U}_i S_i S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}\mathbf{U}_i S_i + \alpha^2 S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}\mathbf{U}_i S_i\mathbf{N}_i^{\dagger}\mathbf{N}_i + \\ &\beta^2 S_i^{\dagger}\mathbf{V}_i^{\dagger}\mathbf{N}_i\mathbf{N}_i^{\dagger}\mathbf{V}_i S_i + \alpha^2 S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}^{1/2}\mathbf{N}_i\mathbf{N}_i^{\dagger}\mathbf{R}^{1/2}\mathbf{U}_i S_i + \beta^2\mathbf{N}_i^{\dagger}\mathbf{V}_i S_i S_i^{\dagger}\mathbf{V}_i^{\dagger}\mathbf{N}_i + \\ &\alpha^2\mathbf{N}_i^{\dagger}\mathbf{R}^{1/2}\mathbf{U}_i S_i S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}^{1/2}\mathbf{N}_i + \beta^2\mathbf{N}_i^{\dagger}\mathbf{N}_i S_i^{\dagger}\mathbf{V}_i^{\dagger}\mathbf{V}_i S_i + \alpha^2\mathbf{N}_i^{\dagger}\mathbf{N}_i S_i^{\dagger}\mathbf{U}_i^{\dagger}\mathbf{R}\mathbf{U}_i S_i + \left(\mathbf{N}_i^{\dagger}\mathbf{N}_i\right)^2\Big], \end{aligned} \tag{6.21}$$

where $E\left[|S_i|^4\right]$ can be obtained by using known results on the moments of an exponential distribution, which yields

$$E\left[|S_i|^4\right] = 2\sigma_S^4. \tag{6.22}$$

In (6.21), $E\left[\left(\mathbf{U}_i^\dagger \mathbf{R} \mathbf{U}_i\right)^2\right]$ can be evaluated by performing an eigendecomposition of $\mathbf{R}$ as in (6.16), which gives

$$E\left[\left(\mathbf{U}_i^\dagger \mathbf{R} \mathbf{U}_i\right)^2\right] = E\left[\left(\mathbf{U}_i^\dagger \phi \Lambda \phi^\dagger \mathbf{U}_i\right)^2\right]. \tag{6.23}$$

Let $\overline{\mathbf{U}}_i^\dagger = \mathbf{U}_i^\dagger \phi$, where $\overline{\mathbf{U}}_i^\dagger \sim \mathcal{CN}(0,1)$. Thus, (6.23) becomes

$$\begin{aligned} E\left[\left(\mathbf{U}_i^\dagger \mathbf{R} \mathbf{U}_i\right)^2\right] &= E\left[\sum_{m=1}^{M} |\overline{U}_{i_m}|^2 \lambda_m \sum_{n=1}^{M} |\overline{U}_{i_n}|^2 \lambda_n\right] \\ &= E\left[\sum_{m=1}^{M} |\overline{U}_{i_m}|^4 \lambda_m^2 + \sum_{m\neq n}^{M} |\overline{U}_{i_m}|^2 |\overline{U}_{i_n}|^2 \lambda_m \lambda_n\right] \\ &= 2\sum_{m=1}^{M} \lambda_m^2 + \sum_{m\neq n}^{M} \lambda_m \lambda_n \\ &= \sum_{m=1}^{M} \lambda_m^2 + \left(\sum_{m=1}^{M} \lambda_m\right)^2 \\ &= \sum_{m=1}^{M} \lambda_m^2 + M^2. \end{aligned} \tag{6.24}$$

$E\left[\mathbf{V}_i^\dagger \mathbf{R} \mathbf{V}_i\right]$ in (6.21) can be written as

$$E\left[\mathbf{V}_i^\dagger \mathbf{R} \mathbf{V}_i\right] = E\left[\sum_{q=1}^{M} |V_{i_q}|^2 + \sum_{q=1}^{M-1} \sum_{r\neq q}^{M} \rho^{|q-r|} \left(V_{i_q}^* V_{i_r} + V_{i_r}^* V_{i_q}\right)\right]. \tag{6.25}$$

Since $|V_{i_q}|^2 = 1$, (6.25) can be expressed with the aid of Euler's formula as

$$\begin{aligned} E\left[\mathbf{V}_i^\dagger \mathbf{R} \mathbf{V}_i\right] &= E\left[\sum_{q=1}^{M} 1 + \sum_{q=1}^{M-1} \sum_{r\neq q}^{M} \rho^{|q-r|} 2\cos\left((q-r)\Delta\right)\right] \\ &= E\left[\sum_{q=1}^{M} \sum_{r=1}^{M} \rho^{|q-r|} \cos\left((q-r)\Delta\right)\right]. \end{aligned} \tag{6.26}$$

Using (6.15), (6.17), (6.18), (6.22), (6.24), (6.26) and since $E\left[\mathbf{U}_i \mathbf{U}_i^\dagger\right] = \mathbf{I}$ and $E\left[\mathbf{N}_i^\dagger \mathbf{N}_i\right] = \sigma_N^2 \mathbf{I}$, (6.21) can be rewritten as

$$
\begin{aligned}
E\left[z_i^2\right] &= 2\beta^4 M^2\sigma_S^4 + 4\beta^2\alpha^2 M^2\sigma_S^4 + 2\beta^2 M^2\sigma_S^2\sigma_N^2 + 4\beta^2\alpha^2\sigma_S^4 \sum_{q=1}^{M}\sum_{r=1}^{M} \rho^{|q-r|} E\left[\cos\left((q-r)\Delta\right)\right] \\
&+ 2\alpha^4\sigma_S^4\left(M^2 + \sum_{m=1}^{M}\lambda_m^2\right) + 2\alpha^2 M^2\sigma_S^2\sigma_N^2 + 2\beta^2 M\sigma_S^2\sigma_N^2 + 2\alpha^2 M\sigma_S^2\sigma_N^2 + M(M+1)\sigma_N^4.
\end{aligned} \tag{6.27}
$$

Therefore, using (6.19) and (6.27), the variance of z_i in (6.20) becomes

$$
\begin{aligned}
\mathrm{Var}(z_i) &= E\left[z_i^2\right] - \left(E[z_i]\right)^2 \\
&= 2\beta^4 M^2\sigma_S^4 + 4\beta^2\alpha^2 M^2\sigma_S^4 + 2\beta^2 M^2\sigma_S^2\sigma_N^2 + 4\beta^2\alpha^2\sigma_S^4 \sum_{q=1}^{M}\sum_{r=1}^{M} \rho^{|q-r|} E\left[\cos\left((q-r)\Delta\right)\right] \\
&+ 2\alpha^4\sigma_S^4\left(M^2 + \sum_{m=1}^{M}\lambda_m^2\right) + 2\alpha^2 M^2\sigma_S^2\sigma_N^2 + 2\beta^2 M\sigma_S^2\sigma_N^2 + 2\alpha^2 M\sigma_S^2\sigma_N^2 + M(M+1)\sigma_N^4 \\
&- \left(\beta^4 M^2\sigma_S^4 + 2\beta^2\alpha^2 M^2\sigma_S^4 + 2\beta^2 M^2\sigma_S^2\sigma_N^2 + \alpha^4 M^2\sigma_S^4 + 2\alpha^2 M^2\sigma_S^2\sigma_N^2 + M^2\sigma_N^4\right) \\
&= \sigma_S^4 M^2 + 2M\sigma_S^2\sigma_N^2 + M\sigma_N^4 + 4\beta^2\alpha^2\sigma_S^4 \sum_{q=1}^{M}\sum_{r=1}^{M} \rho^{|q-r|} E\left[\cos\left((q-r)\Delta\right)\right] + 2\alpha^4\sigma_S^4 \sum_{m=1}^{M}\lambda_m^2.
\end{aligned} \tag{6.28}
$$

Having computed $E[z_i]$ and $\mathrm{Var}(z_i)$, the expected value and variance of each J_i can now be written respectively as

$$
\begin{aligned}
E[J_i] &= aE\left[z_i\right] - b \\
&= aM\left(\beta^2\sigma_S^2 + \alpha^2\sigma_S^2 + \sigma_N^2\right) - b \\
&= aM\left(\sigma_S^2 + \sigma_N^2\right) - b,
\end{aligned} \tag{6.29}
$$

$$
\begin{aligned}
\mathrm{Var}(J_i) &= a^2\mathrm{Var}(z_i) \\
&= a^2\left(\sigma_S^4 M^2 + 2M\sigma_S^2\sigma_N^2 + M\sigma_N^4 + 4\beta^2\alpha^2\sigma_S^4 \sum_{q=1}^{M}\sum_{r=1}^{M} \rho^{|q-r|} E\left[\cos\left((q-r)\Delta\right)\right] + \right. \\
&\left. 2\alpha^4\sigma_S^4 \sum_{m=1}^{M}\lambda_m^2\right).
\end{aligned} \tag{6.30}
$$

Substituting $\beta^2\alpha^2 = K/(K+1)^2$ and $\alpha^4 = 1/(K+1)^2$ into (6.30) yields

$$\mathrm{Var}(J_i) = a^2\left(\varsigma + \frac{\sigma_S^4}{(K+1)^2}\left[4K\sum_{q=1}^{M}\sum_{r=1}^{M}\rho^{|q-r|}E\left[\cos\left((q-r)\Delta\right)\right] + 2\sum_{m=1}^{M}\lambda_m^2\right]\right), \tag{6.31}$$

where

$$\varsigma = \sigma_S^4 M^2 + 2M\sigma_S^2\sigma_N^2 + M\sigma_N^4. \tag{6.32}$$

6.3.1 Application of Brownian motion with drift

In this subsection, we follow the approach taken in Sections 4.3.1.4 and 4.3.2.3 to approximate $P(\delta)$ for all the cases that we consider by applying the theory of Brownian motion with drift.

6.3.1.1 Correlated Rician channel

In this case, $\mathcal{D}_n$ in (6.11) can be re-scaled to give

$$G(n) = \frac{\mathcal{D}_n}{\sqrt{\mathrm{Var}(J_{i_{corrRice}})}}, \tag{6.33}$$

where $\mathrm{Var}(J_{i_{corrRice}})$ is the variance of the jump for the correlated Rician case which is given by (6.31). It is worth noting that the subscript is introduced to differentiate between the different scenarios that we consider. The expected value and variance of $G(n)$ can be denoted respectively by

$$E[G(n)] = \frac{nE[J_i]}{\sqrt{\mathrm{Var}(J_{i_{corrRice}})}}, \tag{6.34}$$

$$\mathrm{Var}(G(n)) = n. \tag{6.35}$$

By comparing the means in (4.37) and (6.34), the drift, μ can be written as

$$\begin{aligned}\mu &= \frac{E[J_i]}{\sqrt{\mathrm{Var}(J_{i_{corrRice}})}}\\ &= \frac{aM\left(\sigma_S^2+\sigma_N^2\right) - b}{a\sqrt{\varsigma + \frac{\sigma_S^4}{(K+1)^2}\left[4K\sum_{q=1}^{M}\sum_{r=1}^{M}\rho^{|q-r|}E\left[\cos\left((q-r)\Delta\right)\right] + 2\sum_{m=1}^{M}\lambda_m^2\right]}},\end{aligned} \tag{6.36}$$

where ς is given in (6.32). The re-scaling of $\mathcal{D}_n$ results into a re-scaled threshold, which is given by

$$\begin{aligned} \upsilon &= \frac{\gamma}{\sqrt{\mathrm{Var}(J_{i_{corrRice}})}} \\ &= \frac{\gamma}{a\sqrt{\varsigma + \frac{\sigma_S^4}{(K+1)^2}\left[4K\sum_{q=1}^{M}\sum_{r=1}^{M}\rho^{|q-r|}E\left[\cos\left((q-r)\Delta\right)\right] + 2\sum_{m=1}^{M}\lambda_m^2\right]}}. \end{aligned} \tag{6.37}$$

Therefore, the approximate expression for $P(\delta)$ for the correlated Rician case can be expressed as

$$\widetilde{P}_{bm_{corrRice}}(\delta) = \frac{|\upsilon|}{\sqrt{2\pi q^3}} e^{-\frac{(\upsilon-\mu q)^2}{2q}}\, dq, \tag{6.38}$$

where $q = \delta + \frac{1}{2}$ and we set $dq = 1$. It is worth noting that tilde denotes an approximation of $P(\delta)$ and the subscript, $bm_{corrRice}$ in (6.38) denotes the Brownian motion approximation for the correlated Rician case. As can be seen from (6.38), the only term in $\widetilde{P}_{bm_{corrRice}}(\delta)$ that depends on the parameters K and ρ is $\sqrt{\mathrm{Var}(J_{i_{corrRice}})}$, which is embedded in both μ and υ, given in (6.36) and (6.37), respectively. Therefore, in order to obtain the approximate expression for $P(\delta)$ for all the other cases, $\mathrm{Var}(J_{i_{corrRice}})$ in (6.31) needs to be changed accordingly.

6.3.1.2 Independent Rician channel

For the special case of no correlation between the antennas (i.e. $\rho = 0$), the channel reduces to an i.i.d Rician channel and the variance of each jump can be written using (6.31) as

$$\mathrm{Var}(J_{i_{indRice}}) = a^2\left(\varsigma + \frac{\sigma_S^4}{(K+1)^2}\left[4K\sum_{q=1}^{M}\sum_{r=1}^{M}1 + 2\sum_{m=1}^{M}\lambda_m^2\right]\right). \tag{6.39}$$

In (6.39), for $\rho = 0$, $\sum_{m=1}^{M}\lambda_m^2$ can be written as

$$\sum_{m=1}^{M}\lambda_m^2 = \mathrm{tr}(\mathbf{R}^2) = M, \tag{6.40}$$

where $\mathrm{tr}(\mathbf{A})$ is the trace of the matrix $\mathbf{A}$. Using (6.40), $\mathrm{Var}(J_{i_{indRice}})$ in (6.39) can be re-written as

$$\begin{aligned}\mathrm{Var}(J_{i_{indRice}}) &= a^2\left(\varsigma + \frac{\sigma_S^4}{(K+1)^2}\left[4K\sum_{q=1}^{M}\sum_{r=1}^{M}1 + 2M\right]\right) \\ &= a^2\left(\varsigma + \frac{2M\sigma_S^4}{(K+1)^2}\left[2K+1\right]\right),\end{aligned} \tag{6.41}$$

where ς is given in (6.32). Therefore, using (6.41), the drift, μ in (6.36) and the re-scaled threshold, υ in (6.37) can then be re-written respectively as

$$\begin{aligned}\mu_{indRice} &= \frac{E[J_i]}{\sqrt{\mathrm{Var}(J_{i_{indRice}})}} \\ &= \frac{aM\left(\sigma_S^2+\sigma_N^2\right)-b}{a\sqrt{\varsigma + \frac{2M\sigma_S^4}{(K+1)^2}\left[2K+1\right]}},\end{aligned} \tag{6.42}$$

$$\begin{aligned}\upsilon_{indRice} &= \frac{\gamma}{\sqrt{\mathrm{Var}(J_{i_{indRice}})}} \\ &= \frac{\gamma}{a\sqrt{\varsigma + \frac{2M\sigma_S^4}{(K+1)^2}\left[2K+1\right]}},\end{aligned} \tag{6.43}$$

It is worth noting that the subscripts in (6.42) and (6.43) are used to differentiate between the correlated and independent Rician cases. Hence, we can express the approximate expression for $P(\delta)$ for the i.i.d Rician case as

$$\widetilde{P}_{bm_{indRice}}(\delta) = \frac{\left|\upsilon_{indRice}\right|}{\sqrt{2\pi q^3}}e^{-\frac{(\upsilon_{indRice}-\mu_{indRice}q)^2}{2q}}\,dq, \tag{6.44}$$

where $q = \delta + \frac{1}{2}$, $dq = 1$, and $\mu_{indRice}$ and $\upsilon_{indRice}$ are given in (6.42) and (6.43) respectively.

6.3.1.3 Correlated Rayleigh channel

When there is no LOS component (i.e. K=0), leading to a correlated Rayleigh channel, the variance of each jump can be expressed using (6.31) to give

$$\mathrm{Var}(J_{i_{corrRay}}) = a^2\left(\varsigma + 2\sigma_S^4\sum_{m=1}^{M}\lambda_m^2\right), \tag{6.45}$$

where ς is given in (6.32). The drift, μ in (6.36) and the re-scaled threshold, υ in (6.37) can be re-written respectively using (6.45) as

$$\begin{aligned}\mu_{corrRay} &= \frac{E[J_i]}{\sqrt{\mathrm{Var}(J_{i_{corrRay}})}} \\ &= \frac{aM\left(\sigma_S^2+\sigma_N^2\right)-b}{a\sqrt{\varsigma+2\sigma_S^4\sum_{m=1}^{M}\lambda_m^2}},\end{aligned} \tag{6.46}$$

$$\begin{aligned}\upsilon_{corrRay} &= \frac{\gamma}{\sqrt{\mathrm{Var}(J_{i_{corrRay}})}} \\ &= \frac{\gamma}{a\sqrt{\varsigma+2\sigma_S^4\sum_{m=1}^{M}\lambda_m^2}}.\end{aligned} \tag{6.47}$$

Thus, the approximate expression for $P(\delta)$ for the correlated Rayleigh scenario can be expressed as

$$\widetilde{P}_{bm_{corrRay}}(\delta) = \frac{|\upsilon_{corrRay}|}{\sqrt{2\pi q^3}} e^{-\frac{(\upsilon_{corrRay}-\mu_{corrRay}q)^2}{2q}}\, dq, \tag{6.48}$$

where $q=\delta+\frac{1}{2}$, $dq=1$ and $\mu_{corrRay}$ and $\upsilon_{corrRay}$ are given in (6.46) and (6.47) respectively.

6.3.1.4 Independent Rayleigh channel

In the case when there is no correlation between the antennas and no LOS component (i.e. when $\rho=0$ and $K=0$), the correlated Rician channel reduces to an independent Rayleigh channel. Here, the variance of each jump can be written using (6.45) and (6.40) to yield

$$\mathrm{Var}(J_{i_{indRay}}) = a^2\left(\varsigma+2M\sigma_S^4\right), \tag{6.49}$$

where ς is given in (6.32). Using (6.49), the drift, μ in (6.36) and the re-scaled threshold, υ in (6.37) can be expressed respectively as

$$\begin{aligned}\mu_{indRay} &= \frac{E[J_i]}{\sqrt{\mathrm{Var}(J_{i_{indRay}})}} \\ &= \frac{aM\left(\sigma_S^2+\sigma_N^2\right)-b}{a\sqrt{\varsigma+2M\sigma_S^4}},\end{aligned} \tag{6.50}$$

$$\begin{aligned}\upsilon_{indRay} &= \frac{\gamma}{\sqrt{\mathrm{Var}(J_{i_{indRay}})}} \\ &= \frac{\gamma}{a\sqrt{\varsigma + 2M\sigma_S^4}},\end{aligned} \tag{6.51}$$

Therefore, the approximate expression for $P(\delta)$ for the independent Rayleigh scenario can be written as

$$\widetilde{P}_{bm_{indRay}}(\delta) = \frac{|\upsilon_{indRay}|}{\sqrt{2\pi q^3}} e^{-\frac{(\upsilon_{indRay} - \mu_{indRay} q)^2}{2q}} \, dq, \tag{6.52}$$

where $q = \delta + \frac{1}{2}$, $dq = 1$, and μ_{indRay} and υ_{indRay} are given in (6.50) and (6.51) respectively.

6.4 NUMERICAL RESULTS

In this section, we present some numerical results to validate the approximate expressions for the distribution of detection delay based on the Brownian motion approach for a time-invariant CUSUM detector when the received signal is transmitted over (independent and correlated) Rayleigh and (independent and correlated) Rician channels. The threshold, γ, is set to be $\gamma = 3$, following Chapter 4. Simulations use 20000 trials to generate each point, where each trial has 200 samples. We consider the case when the PU starts transmitting at $\tau = 100$. Therefore, the CUSUM detector receives only noise when $i < 100$ whereas the faded PU signal and noise will be received by the CUSUM detector when $i \geq 100$. Detection delay is measured from when the PU begins transmission, at τ, until the CU detects the PU signal using the CUSUM detector. It is worth noting that $\delta = 0$ means that the CU successfully detects the PU transmission in the first sample from when the PU begins transmitting. For example, since in this case the PU begins at $\tau = 100$, if the CU detects the PU at $T = 100$ (i.e. $i = 100$), this means that the CU successfully detects the PU existence in the first sample of the PU transmission. Hence, in this case, no detection delay occurs.

Figures 6.1 and 6.2 compare the simulated and the approximate distribution of detection delay based on the Brownian motion (BM) approach for all the cases at SNR=-5 dB when the CU is equipped with M=2 and 4 antennas, respectively. It is worth noting that a K-factor of 6 dB is selected for the Rician (independent and correlated) case as it is a typical K-factor value used [130]. It can be seen from both figures that for all of the cases, $\widetilde{P}_{bm_{indRay}}(\delta)$, $\widetilde{P}_{bm_{corrRay}}(\delta)$,

$\widetilde{P}_{bm_{indRice}}(\delta)$ and $\widetilde{P}_{bm_{corrRice}}(\delta)$ provide accurate approximations of $P(\delta)$, especially for moderate and longer detection delays. However, for short delays, a reasonably good approximation can still be observed for all of the cases in both figures. Comparing both Figures 6.1 and 6.2, it can be observed that as M increases, the detection delay distribution becomes narrower as a result of increasing variance for the jumps. The increased variance can be observed from (6.31) for the correlated Rician case, which gives

$$\mathrm{Var}(J_{i_{corrRice}}) = a^2 \left(\varsigma + \frac{\sigma_S^4}{(K+1)^2} \left[4K \sum_{q=1}^{M} \sum_{r=1}^{M} \rho^{|q-r|} E\left[\cos\left((q-r)\Delta\right)\right] + 2 \sum_{m=1}^{M} \lambda_m^2 \right] \right), \quad (6.53)$$

where $a = \frac{\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)}$ and $\varsigma = \sigma_S^4 M^2 + 2M\sigma_S^2\sigma_N^2 + M\sigma_N^4$ (which is given in (6.32)). It can be observed from (6.53) that $\mathrm{Var}(J_{i_{corrRice}})$ increases with M. From (6.38), we can see that $\widetilde{P}_{bm_{corrRice}}(\delta)$ is inversely proportional to $\sqrt{\mathrm{Var}(J_{i_{corrRice}})}$ and $\sqrt{q^3}$, where $q = \delta + \frac{1}{2}$. Therefore, as $\mathrm{Var}(J_{i_{corrRice}})$ and δ increases, $\widetilde{P}_{bm_{corrRice}}(\delta)$ reduces. Hence, this suggest that employing multiple antennas reduces the probability of longer detection delays due to the spatial diversity provided.

Based on Figures 6.1 and 6.2, we can see that the effects of the correlation coefficient, ρ, and the Rician K-factor are small at SNR=-5 dB. The size of these effects can be explained by looking at the variance of each jump, J_i, for each of the scenarios. As discussed in Section 6.3.1.1, the variance of each jump is the only term in the approximate expression for the detection delay distribution that depends on the parameters ρ and K. The variance of each jump can be written as

$$\mathrm{Var}(J_i) = a^2(\varsigma + \xi), \quad (6.54)$$

where we recall from Section 6.2.1 that $a = \frac{\sigma_X^2}{\sigma_N^2(\sigma_N^2+\sigma_X^2)}$ and ς is given in (6.32). In (6.54), both a and ς depend only on the SNR, whereas ξ varies for (independent and correlated) Rayleigh and (independent and correlated) Rician cases. Hence, in order to differentiate between these cases, we introduced subscripts for ξ and J_i to denote each case. In particular, we let $\xi_{indRay} = 2M\sigma_S^4$, $\xi_{corrRay} = 2\sigma_S^4 \sum_{m=1}^{M} \lambda_m^2$, $\xi_{indRice} = \frac{2M\sigma_S^4}{(K+1)^2}[2K+1]$ and $\xi_{corrRice} = \frac{\sigma_S^4}{(K+1)^2}\left[4K \sum_{q=1}^{M} \sum_{r=1}^{M} \rho^{|q-r|} E\left[\cos\left((q-r)\Delta\right)\right] + 2\sum_{m=1}^{M} \lambda_m^2\right]$, where the subscripts *indRay*, *corrRay*, *indRice* and *corrRice* denote the independent Rayleigh, correlated Rayleigh, indepen-

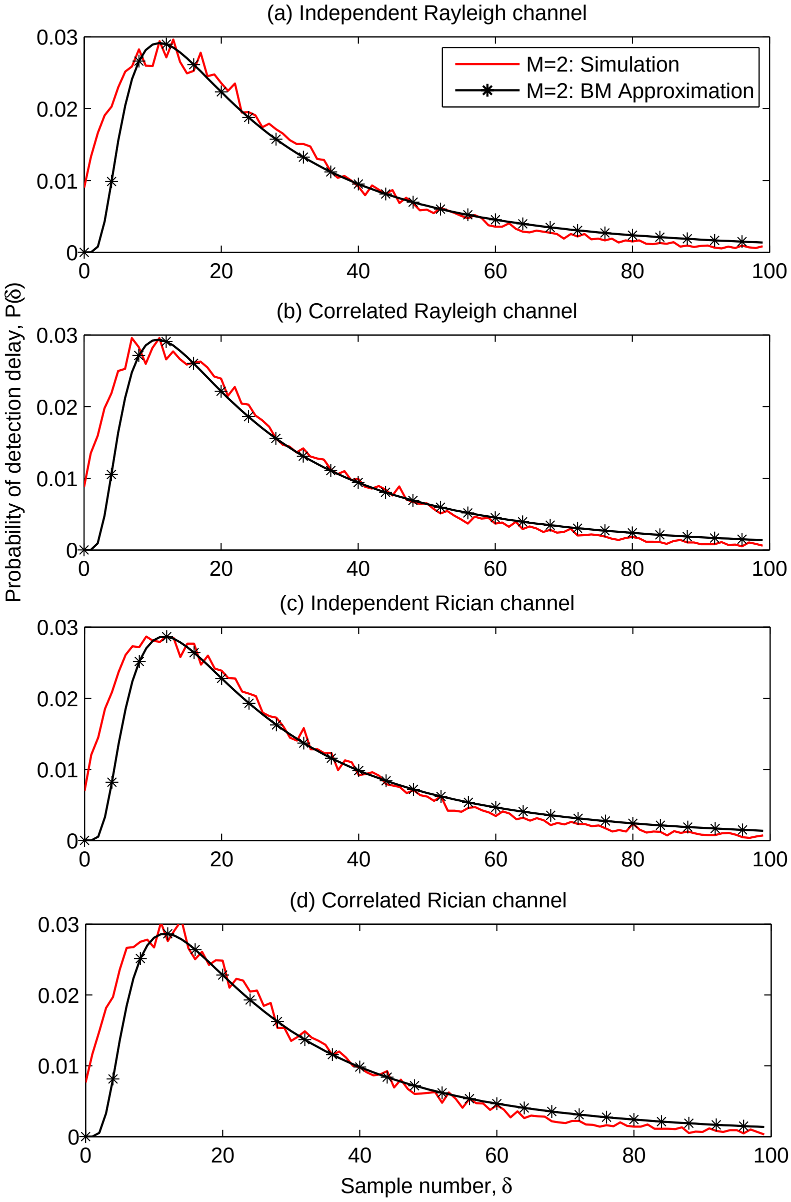

Figure 6.1 A comparison of the simulated results and the approximate distribution of detection delay for (a) independent Rayleigh ($\rho = 0$, $K = 0$), (b) correlated Rayleigh ($\rho = 0.5$, $K = 0$), (c) independent Rician ($\rho = 0$, K=6 dB) and (d) correlated Rician ($\rho = 0.5$, K=6 dB) scenarios with M=2 antennas at SNR=-5 dB.

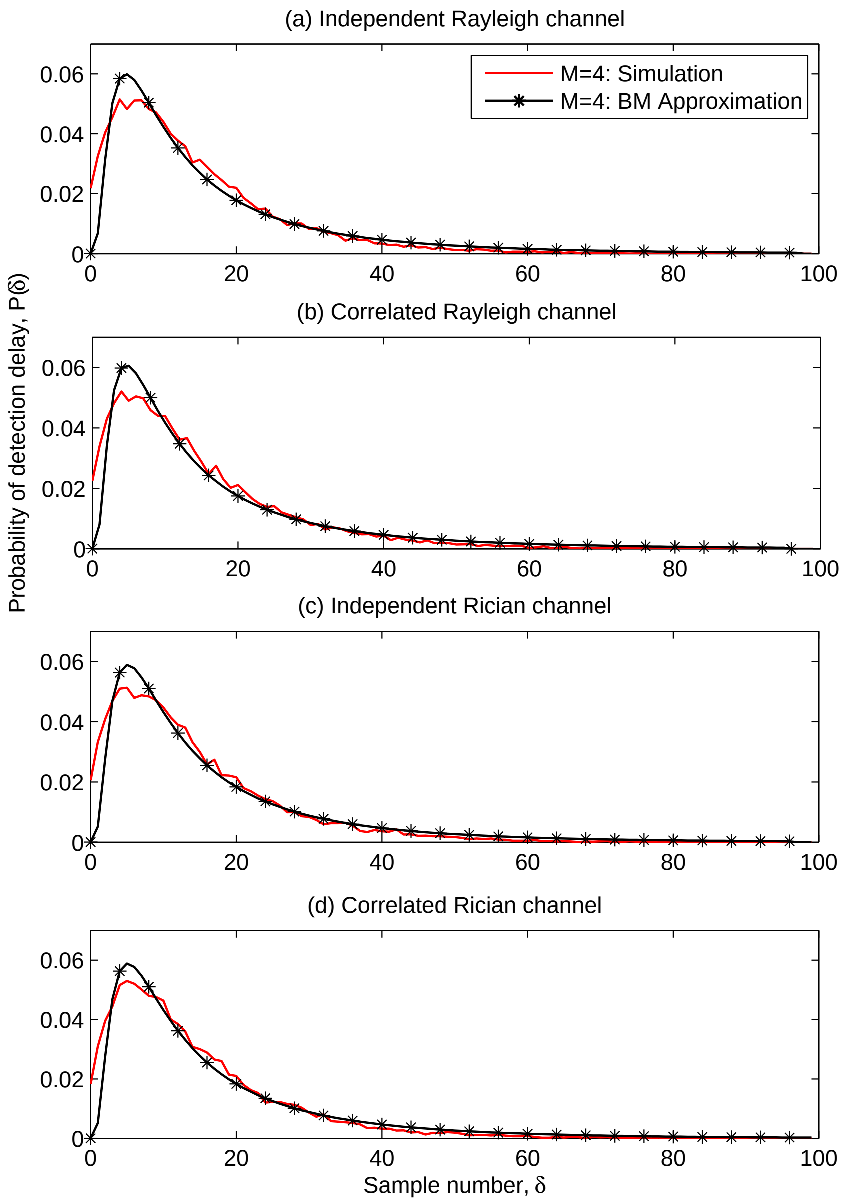

Figure 6.2 A comparison of the simulated results and the approximate distribution of detection delay for (a) independent Rayleigh ($\rho = 0$, $K = 0$), (b) correlated Rayleigh ($\rho = 0.5$, $K = 0$), (c) independent Rician ($\rho = 0$, K=6 dB) and (d) correlated Rician ($\rho = 0.5$, K=6 dB) scenarios with M=4 antennas at SNR=-5 dB.

Table 6.1 Variance of each jump, J_i for different scenarios at SNR=5, 0 and -5 dB with M=4 antennas

SNR	ρ	K	a^2	ς	ξ	$\mathrm{Var}(.) = a^2(\varsigma + \xi)$
5 dB	0	0	0.5772	189.2982	$\xi_{indRay} = 80$	$\mathrm{Var}(J_{i_{indRay}}) = 155.4389$
	0.5	0	0.5772	189.2982	$\xi_{corrRay} = 115.6250$	$\mathrm{Var}(J_{i_{corrRay}}) = 176.0017$
	0.9	0	0.5772	189.2982	$\xi_{corrRay} = 250.9456$	$\mathrm{Var}(J_{i_{corrRay}}) = 254.1087$
	0	0 dB	0.5772	189.2982	$\xi_{indRice} = 60$	$\mathrm{Var}(J_{i_{indRice}}) = 143.8949$
	0	6 dB	0.5772	189.2982	$\xi_{indRice} = 28.8972$	$\mathrm{Var}(J_{i_{indRice}}) = 125.9424$
	0.5	0 dB	0.5772	189.2982	$\xi_{corrRice} = 65.4434$	$\mathrm{Var}(J_{i_{corrRice}}) = 147.0369$
	0.9	0 dB	0.5772	189.2982	$\xi_{corrRice} = 96.3440$	$\mathrm{Var}(J_{i_{corrRice}}) = 164.8727$
	0.5	6 dB	0.5772	189.2982	$\xi_{corrRice} = 28.1600$	$\mathrm{Var}(J_{i_{corrRice}}) = 125.5169$
	0.9	6 dB	0.5772	189.2982	$\xi_{corrRice} = 31.8423$	$\mathrm{Var}(J_{i_{corrRice}}) = 127.6423$
0 dB	0	0	0.25	28	$\xi_{indRay} = 8$	$\mathrm{Var}(J_{i_{indRay}}) = 9$
	0.5	0	0.25	28	$\xi_{corrRay} = 11.5625$	$\mathrm{Var}(J_{i_{corrRay}}) = 9.8906$
	0.9	0	0.25	28	$\xi_{corrRay} = 25$	$\mathrm{Var}(J_{i_{corrRay}}) = 13.25$
	0	0 dB	0.25	28	$\xi_{indRice} = 6$	$\mathrm{Var}(J_{i_{indRice}}) = 8.5$
	0	6 dB	0.25	28	$\xi_{indRice} = 2.8897$	$\mathrm{Var}(J_{i_{indRice}}) = 7.7224$
	0.5	0 dB	0.25	28	$\xi_{corrRice} = 6.5440$	$\mathrm{Var}(J_{i_{corrRice}}) = 8.6360$
	0.9	0 dB	0.25	28	$\xi_{corrRice} = 9.6110$	$\mathrm{Var}(J_{i_{corrRice}}) = 9.4028$
	0.5	6 dB	0.25	28	$\xi_{corrRice} = 2.8176$	$\mathrm{Var}(J_{i_{corrRice}}) = 7.7044$
	0.9	6 dB	0.25	28	$\xi_{corrRice} = 3.1529$	$\mathrm{Var}(J_{i_{corrRice}}) = 7.7882$
-5 dB	0	0	0.0577	8.1298	$\xi_{indRay} = 0.8$	$\mathrm{Var}(J_{i_{indRay}}) = 0.5152$
	0.5	0	0.0577	8.1298	$\xi_{corrRay} = 1.1562$	$\mathrm{Var}(J_{i_{corrRay}}) = 0.5358$
	0.9	0	0.0577	8.1298	$\xi_{corrRay} = 2.5095$	$\mathrm{Var}(J_{i_{corrRay}}) = 0.6139$
	0	0 dB	0.0577	8.1298	$\xi_{indRice} = 0.6$	$\mathrm{Var}(J_{i_{indRice}}) = 0.5037$
	0	6 dB	0.0577	8.1298	$\xi_{indRice} = 0.2890$	$\mathrm{Var}(J_{i_{indRice}}) = 0.4858$
	0.5	0 dB	0.0577	8.1298	$\xi_{corrRice} = 0.6550$	$\mathrm{Var}(J_{i_{corrRice}}) = 0.5069$
	0.9	0 dB	0.0577	8.1298	$\xi_{corrRice} = 0.9645$	$\mathrm{Var}(J_{i_{corrRice}}) = 0.5247$
	0.5	6 dB	0.0577	8.1298	$\xi_{corrRice} = 0.2824$	$\mathrm{Var}(J_{i_{corrRice}}) = 0.4854$
	0.9	6 dB	0.0577	8.1298	$\xi_{corrRice} = 0.3173$	$\mathrm{Var}(J_{i_{corrRice}}) = 0.4874$

dent Rician and correlated Rician cases, respectively.

Based on Section 6.3.1, we can measure the variance of each jump for all of the scenarios at various SNR and this is illustrated in Table 6.1. The results in Table 6.1 give us further insights into the effect of increasing or decreasing SNR, channel correlation, K-factor and variance on the distribution of detection delay. Furthermore, Table 6.1 is also useful in providing some insights into the relationship between a, ς, ξ and the variance of each jump for different cases.

Based on Table 6.1, it can be observed that at SNR=-5 dB, the variance of each jump is very similar for all of the cases that we consider. This is because, the value of ξ for all of the cases are

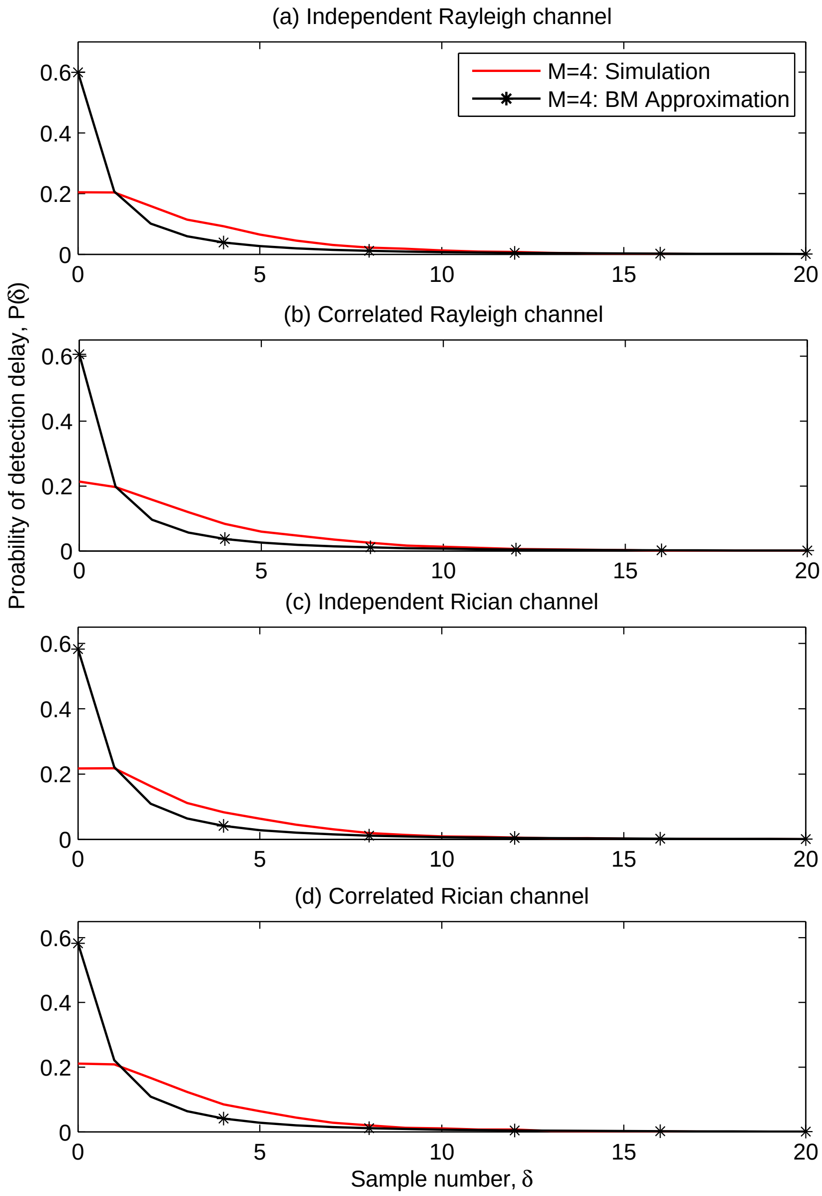

Figure 6.3 A comparison of the simulated results and the approximate distribution of detection delay for (a) independent Rayleigh ($\rho = 0$, $K = 0$), (b) correlated Rayleigh ($\rho = 0.5$, $K = 0$), (c) independent Rician ($\rho = 0$, K=6 dB) and (d) correlated Rician ($\rho = 0.5$, K=6 dB) scenarios with M=4 antennas at SNR=0 dB.

small and hence the constant term, ς dominates, where ς depends only on the SNR. Therefore, channel correlation and LOS have little impact on the approximation of the distribution of detection delay and this is shown in both Figures 6.1 and 6.2. Similar trends can also be observed for SNR=0 dB shown in Table 6.1 and Figure 6.3.

In Figure 6.3, we can see that at SNR=0 dB, $\widetilde{P}_{bm_{indRay}}(\delta)$, $\widetilde{P}_{bm_{corrRay}}(\delta)$, $\widetilde{P}_{bm_{indRice}}(\delta)$ and $\widetilde{P}_{bm_{corrRice}}(\delta)$ provide good approximations of $P(\delta)$ for longer delays. Recall from the analysis in

Section 4.4.1, that the approximation approach of using the modified detection delay statistic, $\mathcal{D}_n$, in modelling the CUSUM process, C_n is most likely to be in error in the first sample (i.e. $\delta = 0$). This error is more likely to occur at lower SNR, since the values of the received signal, z_i, tend to be smaller, resulting in a higher probability of the $\mathcal{D}_n$ process becoming negative whereas $C_n \geq 0$. It can be seen from Figure 6.3 at $\delta = 0$, that the approximate expression of the distribution of detection delay differs by almost three times the true detection delay distribution. Similar trends at $\delta = 0$ can also be observed in Figures 6.1 and 6.2 for SNR=-5 dB, but with a much smaller difference between the approximated and the true distribution of detection delay, which may be caused by the fact that the Brownian motion approximation is a better fit at SNR=-5 dB.

Figure 6.4 shows a comparison of the simulated results and the approximate distribution of detection delay for all of the cases at SNR=5 dB when M=4 antennas are employed at the CU. It can be observed from Figure 6.4 that $\widetilde{P}_{bm_{indRay}}(\delta)$, $\widetilde{P}_{bm_{corrRay}}(\delta)$, $\widetilde{P}_{bm_{indRice}}(\delta)$ and $\widetilde{P}_{bm_{corrRice}}(\delta)$ are only accurate for long delays. In order to achieve a good approximation of $P(\delta)$ for short delays at SNR=5 dB, a new approximate expression for the distribution of detection delay for CU with multiple antennas could be derived based on the gamma approximation, following the approach taken in Sections 4.3.1.2 and 4.3.2.1. This can be considered in future work.

Based on Table 6.1, we can see that at SNR=5 dB, the variance of the jumps differ from each other and although $\xi < \varsigma$, ξ is not negligible. Since the parameter ξ depends on the correlation coefficient and the Rician K-factor, whereas ς only depends on the SNR, the distribution of detection delay is dependent on the channel correlation and LOS strength. The highly correlated Rayleigh case, with $\rho = 0.9$, is particularly notable since here, $\xi > \varsigma$, and hence, the variance of the jump is the highest amongst all the other cases. Figure 6.5 shows the approximate distribution of detection delay for the independent and correlated Rayleigh cases, when $\rho = 0$ and 0.9, respectively. By comparing both of these cases in Figure 6.5, we can see the as the channel correlation increases, $\widetilde{P}_{bm_{corrRay}}(\delta)$ reduces for short delays. However, for longer delays, channel correlation has little impact on $\widetilde{P}_{bm_{corrRay}}(\delta)$. Therefore, based on this result, we conclude that high channel correlation reduces the probability of detection delay for short delays although this effect is not pronounced. Note that at such short delays, the approximations in Figures 6.4 and

6.5 are not accurate.

In addition, at SNR=5 dB, it can also be observed from Table 6.1 that as the LOS strength increases, the variance of the jump reduces. This effect can also be observed from Figure 6.5 by comparing $\widetilde{P}_{bm_{indRay}}(\delta)$ and $\widetilde{P}_{bm_{indRice}}(\delta)$ for the independent Rayleigh and independent Rician cases. We can see that when the value of the K-factor increases, $\widetilde{P}_{bm_{indRice}}(\delta) > \widetilde{P}_{bm_{indRay}}(\delta)$. However, LOS strength has little impact for long delays. Hence, despite the fact that ξ is a function of both ρ and K, the correlation and K-factor values have only a minor effect on detection delay distribution.

6.5 CHAPTER SUMMARY

In this chapter, we derived theoretical expressions to approximate the distribution of detection delay for a multi-antenna time-invariant CUSUM detector when the received signal experiences (independent and correlated) Rayleigh and (independent and correlated) Rician channels. This was based on Brownian motion theory with drift. The validity of the approximate expressions for the distribution of detection delay for each case considered was verified via simulations. Numerical results illustrate that the Brownian motion approach provides a good approximation for all channel models considered in the low SNR region, whereas at high SNR, the approximation is only accurate for long delays. In addition, results show that channel correlation and LOS strength have little impact on the distribution of detection delay at moderate and low SNR region. However, for short delays at high SNR, there are minor effects caused by varying ρ and K. However, for longer delays at high SNR, the channel correlation and LOS strength have little impact on the distribution of detection delay. Furthermore, the likelihood of long detection delays reduces as the number of receive antennas increases. In practical cognitive radio systems, the limiting factor is not the rapid detection of strong PU signals but the more challenging detection of weak PU signals. Hence, the most important case is when the received signal experiences low SNR, where there is a higher possibility of long detection delay. Hence, our method of approximating the distribution of detection delay is useful in analyzing the quickest spectrum sensing performance at such a low SNR.

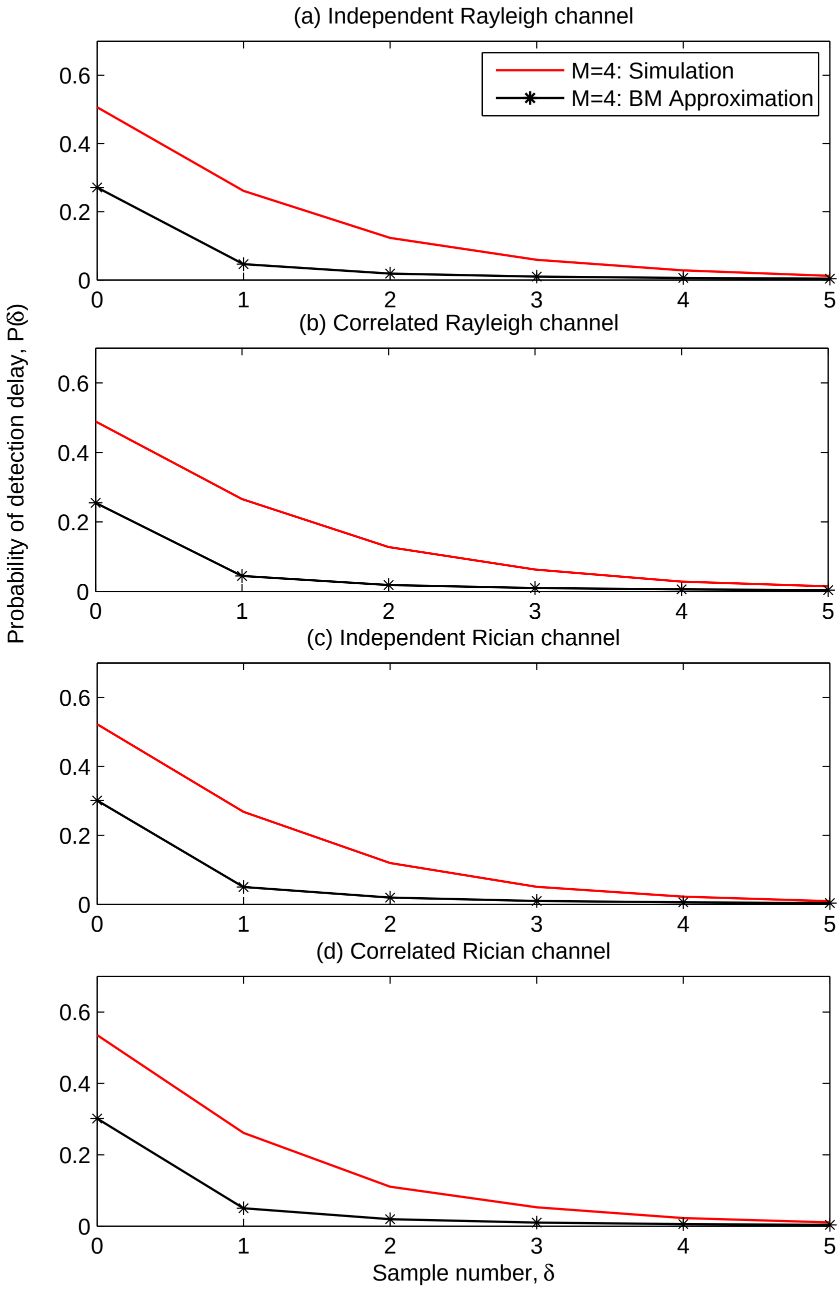

Figure 6.4 A comparison of the simulated results and the approximate distribution of detection delay for (a) independent Rayleigh ($\rho = 0$, $K = 0$), (b) correlated Rayleigh ($\rho = 0.5$, $K = 0$), (c) independent Rician ($\rho = 0$, K=6 dB) and (d) correlated Rician ($\rho = 0.5$, K=6 dB) scenarios with M=4 antennas at SNR=5 dB.

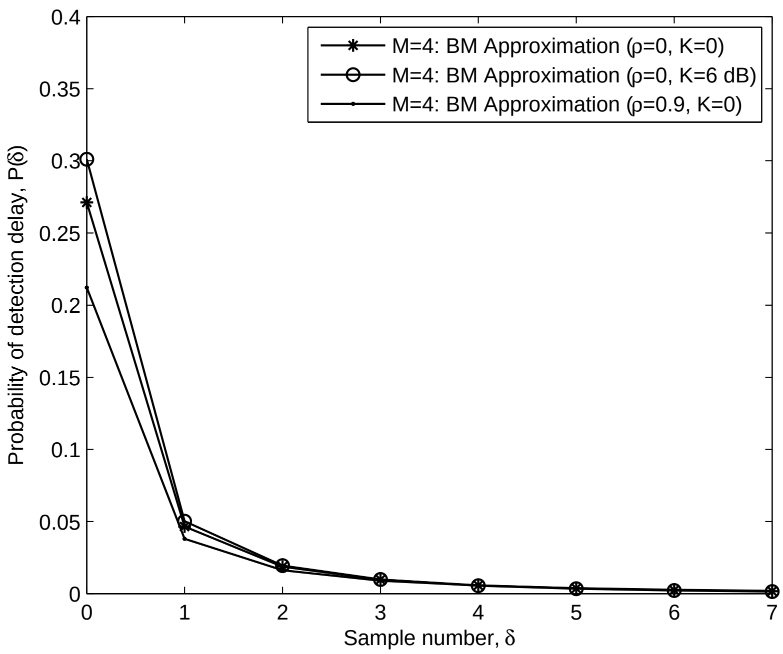

Figure 6.5 Approximate distribution of detection delay for independent Rayleigh ($\rho = 0$, $K = 0$), independent Rician ($\rho = 0$, K=6 dB) and correlated Rayleigh ($\rho = 0.9$, K=0) scenarios with M=4 antennas at SNR=5 dB.

Chapter 7

CONCLUSIONS AND FUTURE WORK

This chapter summarizes the novel contributions of this thesis and highlights several potential extensions and open problems for future work.

7.1 CONCLUSIONS

The demand for wireless spectrum is constantly increasing as new wireless communication services appear and existing services grow. These trends have led to the existing spectrum scarcity. Current spectrum allocation policies aggravate the spectrum scarcity in wireless communications since particular spectrum bands are usually dedicated to licensed (primary) users for specific services. This approach results in much of the allocated radio spectrum being underutilized or sitting idle (i.e. spectrum holes). Therefore, cognitive radio has emerged as a potential solution to improve the spectrum utilization efficiency, where unlicensed (cognitive) users attempt to access the spectrum in such a way that the primary users are unaffected.

This thesis primarily considers interweave cognitive radio systems where the CU is allowed to access the spectrum licensed to the PU only when the PU does not occupy the spectrum. Therefore, the CU needs to monitor and subsequently detect the occupancy of the spectrum in order to opportunistically communicate over the vacant spectrum without interfering with the PU. Hence, spectrum sensing plays a crucial role in the deployment of a cognitive radio system and many spectrum sensing techniques have been proposed in the literature. In spectrum sensing, when the PU is absent, the CU needs to detect the spectrum holes as quickly as possible to fully utilize the unused spectrum. When the PU starts transmitting, the CU needs to detect the existence of the PU in order to vacate the frequency band quickly to avoid any harmful interference to the PU. Thus, detection delay is a crucial criterion in spectrum sensing.

In this thesis, we consider quickest spectrum sensing, which is based on the theory of quickest detection, where the aim is to detect the PU with minimal detection delay (i.e. using the least number of samples) subject to a certain false alarm rate. The main focus of this thesis is the investigation of the quickest spectrum sensing performance for both single and multiple receive antennas at the CU over Gaussian and several fading channels, including the classical fading channels such as Rayleigh, Rician, Nakagami-m and a long tailed channel, which is based on the F-distribution. Apart from independent channels, temporally and spatially correlated channels are also being considered in this thesis. The results of this investigation has led us to develop novel theoretical expressions for the distribution of detection delay for both single and multiple antenna scenarios, which are beneficial in providing a more detailed analysis of the performance of quickest spectrum sensing.

In particular, in Chapter 3, we studied the performance of quickest spectrum sensing with single antenna CUs when the receive signal is transmitted over time-invariant as well as various fading channel conditions, including Rayleigh, Rician, Nakagami-m and a more severe fading channel, the F channel. The received signal at the CU is considered to be the product of two complex variables (the PU signal and channel) with additive noise. In Chapter 3, we proved that the power of the complex received signal is a sufficient statistic. Therefore, the log likelihood ratio could be computed based on the amplitude of the received signal. The pdfs of the amplitude of the received signal were derived for the Rayleigh, Rician, Nakagami-m and F channels. The novel derivation of these pdfs (excluding F channel) used a technique which avoids numerical integration.

Results in Chapter 3 show that the quickest spectrum sensing performance degrades with the severity of the fading channels as well as the level of temporal correlation. In addition, results also show that in the event of mis-matched channel condition (where the CUSUM detector is designed for a specific channel, but the true channel is different), the quickest spectrum sensing performance depends heavily on the true channel, but very little on the channel used to design the CUSUM detector. Since in a cognitive radio network, the channel is usually unknown and the sensing performance in various channels is insensitive to the designed detector, a time-invariant detector can be employed with minimal performance loss.

Motivated by these results, in Chapter 4, we derived theoretical expressions for the distribution of detection delay for a time-invariant CUSUM detector with single antenna CU when the received signal experiences Gaussian and Rayleigh channels. In each of the cases considered, several techniques to approximate the distribution of detection delay are presented, where these approximate methods are necessary since there is no exact solution possible. In particular, we derived a novel approximate closed-form expression for the detection delay distribution for the Gaussian case. Furthermore, we also derived novel approximations for the distribution of detection delay for the general case due to the absence of a general framework. For both Gaussian and Rayleigh cases, we applied simple random walks and Brownian motion theory with drift in deriving the approximate expressions for the distribution of detection delay. Most of the approximate expressions that we formulated are general and can be applied to any i.i.d channel.

Moreover, in Chapter 4, we analyzed the accuracy of the modified detection delay statistic and derived an approximate expression for the probability of missed detection. The probability of long detection delays is also investigated via analysis and simulation. Results show that in order to achieve a good approximation of the distribution of detection delay, different approximate approaches are required for different SNR and detection delay conditions. The Brownian motion approach provides the best approximation for longer delays. In addition, the analysis of long detection delays illustrate that the probability of a long detection delay is increased if the threshold value is large or the received signal is weak due to low SNR. However, the type of fading channel has very little impact on long detection delays. This is because in order for the channel to create long detection delays, it needs an extremely high CV value that can result from severe fading channels (e.g. log-normal shadow fading or the F channel), but does not occur with the traditional channel models, such as Rayleigh, Rician and Nakagami-m fading channels.

We employed multiple receive antennas at the CU in Chapter 5 and studied the performance of multi-antenna quickest spectrum sensing when the received signal is transmitted over Gaussian, Rayleigh and Rician channels. Apart from i.i.d channels, we also considered the case of an insufficient spatial separation between multiple antennas on a CU, resulting in a spatially correlated channel. In particular, we considered a correlated Rayleigh channel. We proved that the sum of the complex received signal powers at each antenna for the Gaussian and independent Rayleigh cases are sufficient statistics. Based on this result, the log likelihood ratio can be computed on the basis of the sum of the received signal powers. Therefore, we derived the pdfs of the received signals experiencing both Gaussian and independent Rayleigh channels based on the sum of the received signal powers, where the derivation of the pdfs for the independent Rayleigh case uses an approach which avoids numerical integration. Furthermore, we derived the joint pdf of the received signal for the correlated Rayleigh and independent Rician scenarios.

In Chapter 5, we also derived an analytical performance analysis, where the upper bound and asymptotic worst-case detection delay are derived for both Gaussian and Rayleigh cases. The sensing performance for the correlated Rayleigh and independent Rician cases were evaluated numerically. Numerical analysis and simulation results show that a higher performance gain can be achieved for all of the cases considered by employing multiple receive antennas at the CU, where performance improvements are due to the spatial diversity provided. In addition, results show that increased channel correlation has little impact on the sensing performance at high SNR, but helps to improve the sensing performance at low SNR. Sensing performance also increases as the Rician K-factor value increases. In the event of a mis-matched channel condition, simulation results illustrate that the sensing performance at a particular correlation coefficient or Rician K-factor depends heavily on the true channel and is relatively insensitive to the detector irrespective of the number of antennas employed at the CU. Therefore, a simple multi-antenna time-invariant detector could be employed to detect the PU transmission.

Based on the results obtained in Chapters 4 and 5, in Chapter 6, we derived theoretical expressions for the distribution of detection delay for a multi-antenna time-invariant CUSUM detector when the receive signal is transmitted over (independent and correlated) Rayleigh and (independent and correlated) Rician channels. In particular, we derived the expected value and variance of the received signal experiencing a correlated Rician channel, where the received signal is a product of two complex variables (the PU signal and correlated Rician channel) with additive noise. These are beneficial in deriving the approximate expression for the distribution of detection delay. The derivation of the approximate expressions for the detection delay distribution for all of the cases considered was based on the theory of Brownian motion with drift.

Results in Chapter 6 show that the Brownian motion approach provides a good approximation of the distribution of detection delay for multi-antenna quickest spectrum sensing at low SNR for all cases considered. In contrast, at high SNR, good approximations of the distribution of detection delay for all cases are achieved for long detection delays. Furthermore, channel correlation and LOS strength have little impact on the detection delay distribution at moderate and low SNR region. However, at high SNR, there are minor effects caused by varying ρ and K for short delays whereas for longer delays, the channel correlation and LOS strength have little impact on the distribution of detection delay. In addition, the employment of multiple receive antennas at the CU helps in reducing the likelihood of long detection delays. In practical cognitive radio systems, the limiting factor is not the rapid detection of strong PU signals but the more challenging detection of weak PU signals. Therefore, the most important case is when the received signal experiences low SNR, where there is a higher possibility of long detection delay. Hence, our method of approximating the distribution of detection delay is beneficial in analyzing the quickest spectrum sensing performance at such a low SNR.

7.2 FUTURE WORK

This research study may be extended in numerous ways and there are numerous additional open problems that are useful for future research work. Below are some suggestions of possible areas for future research:

1. **Further performance analysis**
   In Chapter 5, we derived a theoretical performance analysis for quickest spectrum sensing with multiple antenna CUs when the received signal experiences Gaussian and independent Rayleigh channels. These performance analyses could be further extended to the case when the received signal is transmitted over correlated Rayleigh or independent Rician channels. In particular, the upper bound and asymptotic worst-case detection delay could be derived for both cases. However, these derivations may be complicated to evaluate due

to the existence of multiple integrals in the Kullback-Leibler divergence. These multiple integrals are caused by the pdfs of the received signals in the presence of the PU and the log likelihood ratios for both correlated Rayleigh and Rician cases, which involve numerical integration. Nevertheless, analyzing the theoretical sensing performance for both cases is an interesting possibility as it would allow an investigation of the theoretical effect of channel correlation or the Rician K-factor on quickest spectrum sensing performance.

2. **Approximate density expressions**

   The pdfs of the received signal in the presence of the PU when the received signal experiences correlated Rayleigh and independent Rician channels, derived in Chapter 5, cannot be given in closed forms. Therefore, it would be helpful to explore methods to approximate the pdf expressions, which avoid numerical integration. This will be beneficial in deriving the theoretical performance analysis for the quickest spectrum sensing, discussed above in 1. However, finding approximate expressions for the pdfs is likely to be difficult and complicated by the fact that the density is a multidimensional function of σ_N^2 and the elements of the antenna correlation matrix, $\mathbf{R}$. Hence, deriving approximate pdf expressions for the correlated Rayleigh and independent Rician cases which avoid numerical integration would be an interesting future direction.

3. **Multi-antenna cooperative quickest spectrum sensing**

   The study of quickest spectrum sensing with multiple antennas in this thesis may be further extended to cooperative quickest spectrum sensing with multiple receive antennas at the CUs as a future research direction. This extension could possibly help the sensing performance at very low SNR (i.e. SNR=-20 dB), where the receive signal is very weak. Cooperative detection can be implemented in three different contexts based on the cooperation level: centralized, distributed and relay-assisted. Therefore, the extension of this research into cooperative sensing would need to take into account which type of cooperative sensing to employ.

4. **Application to distributed antenna system**

   Another possible area for future work is to apply the multi-antenna quickest spectrum sensing techniques, studied in this thesis, to distributed antenna systems (DASs). DASs have recently received considerable attention because they can enhance spectral efficiency and provide high data rate services [131–133]. Spectrum sensing in the case when the primary user is using a DAS has been studied in [134, 135], where the authors use energy detectors in performing cooperative spectrum sensing in DAS. So far, there has not been any research on quickest spectrum sensing in a DAS where the primary user is equipped with remote antenna units (RAUs). Therefore, developing quickest spectrum sensing, employing multiple antennas at the CU in the context of a mobile communication system can be an interesting future work. This proposed research work can then be possibly extended to cooperative sensing in DAS. Given that quickest spectrum sensing performs better than energy detection in terms of maximum average throughput [49], the study of non-cooperative/cooperative quickest spectrum sensing with multiple antennas in a DAS will significantly improve the performance of the current results in the literature.

5. **Reconfigurable antenna arrays**

   In Chapter 5, results show that sensing in the low SNR region is assisted by spatial correlation. However, data transmission traditionally prefers independent channels. Hence, there are competing demands on the antenna array and hence, developing a reconfigurable antenna array is a possible solution. However, designing such a reconfigurable antenna system poses a number of challenges. The antennas would need to be able to change the direction of the main lobe on a real time basis and at different frequencies.

6. **Multiuser MISO and MIMO cognitive radio network**

   In Chapter 5, we study the quickest spectrum performance when multiple receive antennas are employed at the cognitive user. It will be an interesting area of future research direction to study the quickest spectrum sensing performance in multiple sources or multiple antennas at the transmitter with one (MISO) or more (MIMO) antennas at the cognitive user. In this MISO scenario, the received signal at the CU can be written as $Y[i] = (H_1[i] \times S_1[i]) + (H_2[i] \times S_2[i]) + ... + (H_M[i] \times S_M[i]) + N[i]$, where M denotes the number of antennas at the primary user or number of sources at the transmitter. Hence, in order to detect any PU, the signal needs to be sum up at the CU, which gives $Y[i] = X[i] + N[i]$, where $X[i] = (H_1[i] \times S_1[i]) + (H_2[i] \times S_2[i]) + ... + (H_M[i] \times S_M[i])$. Therefore, in this MISO case, we have the same situation as in SISO case discussed in

Chapter 3 but the received signal at the CU, $Y[i]$ has more power compared to the SISO case. This large received signal power could be beneficial in detecting the PU quicker as the CUSUM statistic might take a small number of samples to cross the threshold. However, the signal model for the MISO case will be different from the SISO case because of the sum that appears in $Y[i]$. This work could also be extended to the MIMO case.

REFERENCES

[1] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Comput. Netw.*, vol. 50, no. 13, pp. 2127–2159, Sept. 2006.

[2] Y. Xiao and F. Hu, *Cognitive Radio Networks.* CRC Press, 2008.

[3] I. F. Akyildiz, W.-Y. Lee, and K. R. Chowdhury, "Crahns: Cognitive radio ad hoc networks," *Ad Hoc Netw.*, vol. 7, no. 5, pp. 810–836, July 2009.

[4] H. Jafarkhani, *Space-Time Coding: Theory and Practice.* Cambridge University Press, 2005.

[5] J. G. Proakis, *Digital communications*, 4th ed. McGraw-Hill, 2000.

[6] A. Khattab, D. Perkins, and M. Bayoumi, *Cognitive Radio Networks: From Theory to Practice.* Springer, 2012.

[7] B. Wang and K. J. R. Liu, "Advances in cognitive radio networks: A survey," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 5–23, Feb. 2011.

[8] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, pp. 116–130, 2009.

[9] M. Masonta, M. Mzyece, and N. Ntlatlapa, "Spectrum decision in cognitive radio networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1088–1107, 2013.

[10] E. Tragos, S. Zeadally, A. Fragkiadakis, and V. Siris, "Spectrum assignment in cognitive radio networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1108–1135, 2013.

[11] G. Staple and K. Werbach, "The end of spectrum scarcity [spectrum allocation and utilization]," *IEEE Spectrum*, vol. 41, no. 3, pp. 48–52, Mar. 2004.

[12] S.-H. Hwang and M.-J. Rim, "Adaptive operation scheme for quiet period in ieee 802.22 system," in *Proc. ICTC*, 2011, pp. 482–484.

[13] V. Valenta, R. MarsÌŇaÌĄlek, G. Baudoin, M. Villegas, M. Suarez, and F. Robert, "Survey on spectrum utilization in Europe: Measurements, analyses and observations," in *Proc. CROWNCOM*, June 2010, pp. 1–5.

[14] M. Lopez-Benitez and F. Casadevall, "On the spectrum occupancy perception of cognitive radio terminals in realistic scenarios," in *Proc. CIP*, June 2010, pp. 99–104.

[15] Y.-C. Liang, K.-C. Chen, G. Li, and P. Mahonen, "Cognitive radio networking and communications: An overview," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3386–3407, Sept. 2011.

[16] K. Sithamparanathan and A. Giorgetti, *Cognitive Radio Techniques: Spectrum Sensing, Interference Mitigation and Localization.* Artech House, 2012.

[17] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, 1999.

[18] H. Arslan, *Cognitive Radio, Software Defined Radio, and Adaptive Wireless Systems (Signals and Communication Technology).* Springer, 2007.

[19] Federal Communications Commission, "Notice of proposed rule making and order: Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," ET Docket No. 03-108, Feb. 2005.

[20] I. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 40–48, 2008.

[21] P. Morreale and K. Terplan, *CRC Handbook of Modern Telecommunications*, 2nd ed. CRC Press, 2010.

[22] K.-L. Du and M. N. S. Swamy, *Wireless Communication Systems: From RF Subsystems to 4G Enabling Technologies.* Cambridge University Press, 2010.

[23] K. J. R. Liu and B. Wang, *Cognitive Radio Networking and Security: A Game-Theoretic View.* Cambridge University Press, 2011.

[24] B. Wild and K. Ramchandran, "Detecting primary receivers for cognitive radio applications," in *Proc. DySPAN*, Nov 2005, pp. 124–130.

[25] E. Hossain and V. K. Bhargava, *Cognitive Wireless Communication Networks.* Springer, 2007.

[26] K.-C. Chen and R. Prasad, *Cognitive Radio Networks.* Wiley, 2009.

[27] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: Challenges and solutions," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, pp. 1–15, 2010.

[28] Y. Zeng and Y.-C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1784–1793, June 2009.

[29] F. R. Yu, *Cognitive Radio Mobile Ad Hoc Networks.* Springer, 2011.

[30] E. Biglieri, A. J. Goldsmith, L. J. Greenstein, N. B. Mandayam, and H. V. Poor, *Principles of Cognitive Radio.* Cambridge University Press, 2012.

[31] L. E. Doyle, *Essentials of Cognitive Radio.* Cambridge University Press, 2009.

[32] R. C. Qiu, Z. Hu, H. Li, and M. C. Wicks, *Cognitive Radio Communication and Networking: Principles and Practice.* Wiley, 2012.

[33] E. Hossain, D. Niyato, and Z. Han, *Dynamic Spectrum Access and Management in Cognitive Radio Networks*, 1st ed. Cambridge University Press, 2009.

[34] A. M. Wyglinski, M. Nekovee, and Y. T. Hou, *Cognitive Radio Communications and Networks: Principles and Practice.* Academic Press, 2010.

[35] B. A. Fette, *Cognitive Radio Technology.* Academic Press, 2009.

[36] L. Lai, Y. Fan, and H. Poor, "Quickest detection in cognitive radio: A sequential change detection framework," in *Proc. GLOBECOM*, Nov. 30 - Dec. 4, 2008, pp. 1–5.

[37] S. Zarrin and T. J. Lim, “Cooperative quickest spectrum sensing in cognitive radios with unknown parameters,” in *Proc. GLOBECOM*, Nov. 30 - Dec. 4, 2009, pp. 1–6.

[38] Q. Zou, S. Zheng, and A. Sayed, “Cooperative spectrum sensing via sequential detection for cognitive radio networks,” in *Proc. SPAWC*, 2009, pp. 121–125.

[39] H. Li, “Cyclostationary feature based quickest spectrum sensing in cognitive radio systems,” in *Proc. VTC-Fall*, Sept. 2010, pp. 1–5.

[40] F. Suratman, A. Tetz, and A. Zoubir, “Collaborative spectrum sensing using sequential detections: Soft decision vs. hard decision,” in *Proc. ICoICT*, 2013, pp. 1–6.

[41] H. Li, C. Li, and H. Dai, “Quickest spectrum sensing in cognitive radio,” in *Proc. CISS*, Mar. 2008, pp. 203–208.

[42] H. Li, H. Dai, and C. Li, “Collaborative quickest spectrum sensing via random broadcast in cognitive radio systems,” in *Proc. GLOBECOM*, Nov. 30 - Dec. 4, 2009, pp. 1–6.

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

[44] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall, 1993.

[45] Q. Zhao and J. Ye, “Quickest change detection in multiple on-off processes,” in *Proc. ICASSP*, Apr. 2009, pp. 3649–3652.

[46] ——, “Quickest detection in multiple on-off processes,” *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 5994–6006, 2010.

[47] H. Li, “Restless watchdog: Selective quickest spectrum sensing in multichannel cognitive radio systems,” *EURASIP J. Adv. Signal Process.*, vol. 2009, Mar. 2009.

[48] Q. Zhao and J. Ye, “When to quit for a new job: Quickest detection of spectrum opportunities in multiple channels,” in *Proc. MILCOM*, Nov. 2008, pp. 1–6.

[49] S. Zarrin and T. J. Lim, “Throughput-sensing tradeoff of cognitive radio networks based on quickest sensing,” in *Proc. ICC*, June 2011, pp. 1–5.

[50] H.-H. Chen and M. Guizani, *Next Generation Wireless Systems and Networks.* Wiley, 2006.

[51] B. S. R. A. Chockalingam, *Large* MIMO *Systems.* Cambridge University Press, 2013.

[52] G. Vitetta, D. P. Taylor, G. Colavolpe, F. Pancaldi, and P. A. Martin, *Wireless Communications: Algorithmic Techniques.* Wiley, 2013.

[53] A. Goldsmith, *Wireless Communications.* Cambridge University Press, 2005.

[54] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Prentice Hall, 2002.

[55] D. Tse, *Fundamentals of Wireless Communications.* Cambridge University Press, 2005.

[56] W. Stallings, *Wireless Communications & Networks*, 2nd ed. Prentice Hall, 2005.

[57] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 2nd ed. Wiley, 2005.

[58] E. Biglieri, *Coding for Wireless Channels.* Springer, 2005.

[59] A. Bhattacharya, *Digital Communication.* McGraw-Hill, 2005.

[60] S. Haykin, *Communication Systems*, 4th ed. Wiley, 2001.

[61] D. A. Guimaraes, *Digital Transmission: A Simulation-Aided Introduction with VisSim/Comm.* Springer, 2010.

[62] S. H. Choi, P. Smith, B. Allen, W. Q. Malik, and M. Shafi, "Severely fading MIMO channels: Models and mutual information," in *Proc. ICC*, June 2007, pp. 4628–4633.

[63] G. D. Durgin, T. S. Rappaport, and D. A. D. Wolf, "New analytical models and probability density functions for fading in wireless communications," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 1005–1015, June 2002.

[64] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. Wiley, 1995, vol. 2.

[65] A. Maaref and R. Annavajjala, "The gamma variate with random shape parameter and some applications," *IEEE Commun. Lett.*, vol. 14, no. 12, pp. 1146–1148, December 2010.

[66] H. Kobayashi, B. L. Mark, and W. Turin, *Probability, Random Processes, and Statistical Analysis: Applications to Communications, Signal Processing, Queueing Theory and Mathematical Finance*. Cambridge University Press, 2012.

[67] F. Khan, LTE *for* 4G *Mobile Broadband: Air Interface Technologies and Performance*. Cambridge University Press, 2009.

[68] U. Charash, "Reception through nakagami fading multipath channels with random delays," *IEEE Trans. Commun.*, vol. 27, no. 4, pp. 657–670, Apr 1979.

[69] B. Holter and G. Oien, "On the amount of fading in MIMO diversity systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2498–2507, Sept 2005.

[70] M. Jankiraman, *Space-time codes and* MIMO *systems*. Artech House, 2004.

[71] C. Oestges and B. Clerckx, MIMO *Wireless Communications*. Academic Press, 2007.

[72] A. F. Molisch, *Wireless Communications*, 2nd ed. Wiley, 2011.

[73] J. D. Gibson, *The Communications Handbook*, 2nd ed. CRC Press, 2002.

[74] G. Tsoulos, MIMO *System Technology for Wireless Communications*. CRC Press, 2006.

[75] H. Sasaoka, *Mobile Communications*. IOS Press, 2000.

[76] B. Vucetic and J. Yuan, *Space-time coding*. Wiley, 2003.

[77] S. Panic, M. Stefanovic, J. Anastasov, and P. Spalevic, *Fading and Interference Mitigation in Wireless Communications*. CRC Press, 2013.

[78] K. N. Le, *Orthogonal Frequency Division Multiplexing with Diversity for Future Wireless Systems*. Bentham Science Publishers, 2012.

[79] D. G. Brennan, "Linear diversity combining techniques," *Proc. IEEE*, vol. 91, pp. 331–356, Feb. 2003.

[80] J. Choi, *Optimal Combining and Detection*. Cambridge University Press, 2010.

[81] M. K. Simon and M.-S. Alouini, “A unified approach to the performance analysis of digital communication over generalized fading channels,” *Proc. IEEE*, vol. 86, no. 9, pp. 1860–1877, Sept. 1998.

[82] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, “Breaking spectrum gridlock with cognitive radios: An information theoretic perspective,” *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[83] J. D. Gibson, *Mobile Communications Handbook*, 3rd ed. CRC Press, 2012.

[84] T. Alpcan, H. Boche, M. L. Honig, and H. V. Poor, *Mechanisms and Games for Dynamic Spectrum Allocation.* Cambridge University Press, 2013.

[85] M. S. Obaidat, A. Anpalagan, and I. Woungang, *Handbook of Green Information and Communication Systems.* Academic Press, 2012.

[86] E. Yaacoub and Z. Dawy, *Resource Allocation in Uplink* OFDMA *Wireless Systems: Optimal Solutions and Practical Implementations.* Wiley, 2012.

[87] F.-L. Luo, *Digital Front-End in Wireless Communications and Broadcasting: Circuits and Signal Processing.* Cambridge University Press, 2011.

[88] X. Chu, D. Lopez-Perez, Y. Yang, and F. Gunnarsson, *Heterogeneous Cellular Networks: Theory, Simulation and Deployment.* Cambridge University Press, 2013.

[89] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, “A survey on spectrum management in cognitive radio networks,” *IEEE Commun. Mag.*, pp. 40–48, Apr. 2008.

[90] E. Yaacoub and Z. Dawy, “Enhancing the performance of OFDMA underlay cognitive radio networks via secondary pattern nulling and primary beam steering,” in *Proc. WCNC*, Mar. 2011, pp. 1476–1481.

[91] Y. Mei, “Information bounds and quickest change detection in decentralized decision systems,” *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2669–2681, July 2005.

[92] A. G. Tartakovsky and V. V. Veeravalli, “General asymptotic bayesian theory of quickest change detection,” *Theory Probab. Appl.*, vol. 49, no. 3, pp. 458–497, 2005.

[93] A. N. Shiryayev, "On optimum methods in quickest detection problems," *Theory Probab. Appl.*, vol. 8, pp. 22–46, 1963.

[94] G. Lorden, "Procedures for reacting to a change in distribution," *Ann. Math. Statist.*, vol. 42, no. 6, pp. 1897–1908, Dec. 1971.

[95] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100–115, 1954.

[96] A. Wald, "Sequential tests of statistical hypothesis," *Ann. Math. Statist.*, vol. 16, no. 2, pp. 117–186, June 1945.

[97] C.-D. Fuh, "Sprt and cusum in hidden markov models," *Ann. Statist.*, vol. 31, no. 3, pp. pp. 942–977, June 2003.

[98] J. Unnikrishnan, V. Veeravalli, and S. Meyn, "Minimax robust quickest change detection," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1604–1614, Mar. 2011.

[99] J. Duan and Y. Li, "Performance analysis of cooperative spectrum sensing in different fading channels," in *Proc. ICCET*, Apr. 2010, pp. V3–64–V3–68.

[100] S. Nallagonda, S. Suraparaju, S. D. Roy, and S. Kundu, "Performance of energy detection based spectrum sensing in fading channels," in *Proc. ICCCT*, Sept. 2011, pp. 575–580.

[101] E. Hanafi, P. A. Martin, P. J. Smith, and A. J. Coulson, "Performance of quickest spectrum sensing over various fading channels," in *Proc. AusCTW*, Jan. 29 - Feb. 1 2013, pp. 69–74.

[102] R. G. Gallager. Detection, decisions, and hypothesis testing. [Online]. Available: http://www.rle.mit.edu/rgallager/documents/6.262vcc8.pdf

[103] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 6th ed. Academic Press, 2000.

[104] J. Stewart, *Calculus*, 2nd ed. Brooks/Cole, 1991.

[105] A. J. Coulson, "Effect of the channel and channel model on dynamic blind spectrum sensing performance," in *Proc. AusCTW*, Jan. 31 - Feb. 3 2011, pp. 19–24.

[106] K.-i. Takizawa, S. Sasaki, J. Zhou, S. Muramatsu, and H. Kikuchi, "Online SNR estimation for parallel combinatorial SS systems in Nakagami fading channels," in *Proc. GLOBECOM*, vol. 2, Nov. 2002, pp. 1239–1243.

[107] W. C. Jakes, *Microwave Mobile Communications.* Wiley, 1974.

[108] E. Hanafi, P. A. Martin, P. J. Smith, and A. J. Coulson, "Extension of quickest spectrum sensing to multiple antennas and rayleigh channels," *IEEE Commun. Lett.*, vol. 17, pp. 625–628, Apr. 2013.

[109] D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes.* CRC Press, 1977.

[110] M. Liao, *Applied Stochastic Processes.* CRC Press, 2013.

[111] W. Feller, *An Introduction to Probability Theory and Its Application.* Wiley, 1968, vol. 1.

[112] D. Khoshnevisan, *Probability.* American Mathematical Society, 2007.

[113] G. Grimmet and D. Stirzaker, *Probability and Random Processes.* Oxford University Press, 2001.

[114] S. N. Ethier, *The Doctrine of Chances: Probabilistic Aspects of Gambling.* Springer, 2010.

[115] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed. Springer, 1991.

[116] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables*, 2nd ed. Kluwer Academic, 2002.

[117] R. M. Feldman and C. Valdez-Flores, *Applied Probability and Stochastic Processes*, 2nd ed. Springer, 2010.

[118] S.-H. Hwang, J.-H. Baek, and O. Dobre, "Spectrum sensing using multiple antenna-aided energy detectors for cognitive radio," in *Proc. CCECE*, May 2009, pp. 209–212.

[119] P. Bianchi, M. Debbah, M. Maida, and J. Najim, "Performance of statistical tests for single-source detection using random matrix theory," *IEEE Trans. Inf. Theory*, vol. 57, pp. 2400–2419, Apr. 2011.

[120] P. Wang, J. Fang, N. Han, and H. Li, "Multiantenna-assisted spectrum sensing for cognitive radio," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 1791–1800, May 2010.

[121] S. Nallagonda, S. Roy, and S. Kundu, "Performance of cooperative spectrum sensing in Rician and Weibull fading channels," in *Proc. INDICON*, 2011, pp. 1–5.

[122] E. Gismalla and E. Alsusa, "On the performance of energy detection using Bartlett's estimate for spectrum sensing in cognitive radio systems," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3394–3404, 2012.

[123] S. Kim, J. Lee, H. Wang, and D. Hong, "Sensing performance of energy detector with correlated multiple antennas," *IEEE Signal Process. Lett.*, vol. 16, no. 8, pp. 671–674, Aug. 2009.

[124] A. Ghasemi and E. Sousa, "Opportunistic spectrum access in fading channels through collaborative sensing," *J. Commun.*, vol. 2, no. 2, pp. 71–82, Mar. 2007.

[125] X. Yang, K. Lei, S. Peng, and X. Cao, "Sensing performance analysis of GLRT detector with multiple antennas: Does correlation help?" in *Proc. ICWMNN*, 2010, pp. 225–229.

[126] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Duxbury, 2002.

[127] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sept. 2001.

[128] G. Lorden, "On excess over the boundary," *Ann. Math. Statist.*, vol. 41, no. 2, pp. 520–527, 1970.

[129] F. Farrokhi, G. Foschini, A. Lozano, and R. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Commun. Lett.*, vol. 5, no. 3, pp. 85–87, Mar. 2001.

[130] H. Suraweera, G. Karagiannidis, and P. Smith, "Performance analysis of the dual-hop asymmetric fading channel," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2783–2788, June 2009.

[131] D. Wang, X. You, J. Wang, Y. Wang, and X. Hou, “Spectral efficiency of distributed MIMO cellular systems in a composite fading channel,” in *Proc. ICC*, May 2008, pp. 1259–1264.

[132] T. Ahmad, S. Al-Ahmadi, H. Yanikomeroglu, and G. Boudreau, “Downlink linear transmission schemes in a single-cell distributed antenna system with port selection,” in *Proc. VTC-Spring*, May 2011, pp. 1–5.

[133] T. Alade, H. Zhu, and H. Osman, “Performance analysis of distributed antenna system for high building wireless communication,” in *Proc. VTC-Spring*, May 2011, pp. 1–5.

[134] W. Yu, S. Yu, and X. Wang, “Optimal soft combination and cooperative sensing in distributed antenna systems,” in *Proc. VTC-Spring*, May 2011, pp. 1–4.

[135] S. Yu and X. Wang, “Joint spectrum sensing in distributed MIMO systems,” in *Proc. VTC-Spring*, May 2011, pp. 1–4.