2014

Advanced Statistical Signal Processing Methods in Sensing, Detection, and Estimation for Communication Applications

Hongting Zhang
Louisiana State University and Agricultural and Mechanical College

ADVANCED STATISTICAL SIGNAL PROCESSING METHODS IN SENSING, DETECTION, AND ESTIMATION FOR COMMUNICATION APPLICATIONS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The School of Electrical Engineering and Computer Sciences

by
Hongting Zhang
B.S., Harbin Institute of Technology, 2009
M.S., Louisiana State University, 2011
December 2014

To my family and friends

Acknowledgements

First, Please allow me to thank my major advisor, Dr. Hsiao-Chun Wu, for his continuous support in my Ph.D. program. Dr. Wu was always there to listen and to give advice. He showed me different approaches solving research problems and taught me how to accomplish any goal with persistency.

I also want to convey my sincere gratefulness to my other committee members, Dr. Xuebin Liang, Dr. Xin Li, Dr. Hongchao Zhang, and Dr. Wanjun Wang for sharing their valuable time and for giving me helpful information to complete the writing of this dissertation as well as the challenging research that lies behind it.

Besides, I would like to express my sincere gratitude to Dr. Hong Jiang in Alcatel Lucent Bell Labs, who has been mentoring my research and helping me all the time ever since my intern there in summer, 2012. His knowledge and inspiration really navigated my research ever since.

I also wish to thank my research team colleagues Dr. Kun Yan, Dr. Lu Lu, and Mr. Tian Xia who helped me in many ways.

Last but not least, I would like to thank my parents and husband. For their love, support and patience during the past four years it has taken me to graduate.

Table of Contents

List of Figures

Abstract

The applications of wireless communications and digital signal processing have dramatically changed the way we live, work, and learn over decades. The requirement of higher throughput and ubiquitous connectivity for wireless communication systems has become prevalent nowadays. Signal sensing, detection and estimation have been prevalent in signal processing and communications for many years. The relevant studies deal with the processing of information-bearing signals for the purpose of information extraction. Nevertheless, new robust and efficient signal sensing, detection and estimation techniques are still in demand since there emerge more and more practical applications which rely on them. In this dissertation work, several novel signal sensing, detection and estimation schemes are proposed for wireless communications applications, such as spectrum sensing, symbol-detection/channel-estimation, and encoder identification. The associated theories and practice in robustness, computational complexity, and overall system performance evaluation are also provided.

1 Introduction to Sensing, Detection and Estimation

In Chapter 1, the motivation of the research topic is represented, which is originated from the idea that a better statistical feature extraction will result in a more efficient and robust system design. Besides, a comprehensive literature review is also provided in this chapter. In Chapters 2 to 4, novel statistical signal processing techniques are proposed and presented for different communication system designs, namely the spectrum sensing scheme, the symbol-detection/channel-estimation approach, and the blind encoder identification technique.

1.1 Research Motivation and Applications

The applications of wireless communications and digital signal processing have dramatically changed the way we live, work, and learn over decades. The requirement of higher throughput and ubiquitous connectivity for wireless communication systems has become prevalent nowadays. When designing wireless communication networks, not only is it important to establish the desired functionalities for new applications, but also it is crucial to investigate how to achieve the optimal bandwidth, energy efficiency, etc., due to the scarcity of the respective resources. This investigation relies on interdisciplinary effort that encompasses areas of signal processing, telecommunications, control, and information theory. In order to improve the performance of modern communication networks, one may ask whether the current design of the

wireless network is optimal or not. If not, what is the fundamental limit on the performance improvement that can be achieved? And, how can this be achieved at minimum cost?

Statistical signal processing is an area of applied mathematics and algorithm design that treats signals as stochastic processes, dealing with their statistical properties. If these statistical properties can be properly utilized, the efficiency of the communication systems can be greatly improved. My previous research has been focusing on statistical signal processing for wireless communications, which provides promising answers to the above questions. Specifically, I have studied a number of fundamental statistical signal processing approaches from different aspects which have be demonstrated to improve the performances of the communication systems. The primary objective of my research is to investigate how new statistical signal properties can be extracted more effectively over different network models and scenarios and benefit wireless communication quality-of-service. In this dissertation, I will address a specific statistic signal feature, discussing its properties, and numerous applications, which include cognitive radios, pilot detections, channel estimations, encoder identification, and so on.

1.2 Literature Review

1.2.1 Spectrum Sensing

Cognitive radio (CR) is a promising solution to combat the scarcity of electromagnetic radio spectrum resource [1, 2, 3]. Traditionally, a large amount of spectrum bands have already been assigned to different users, (*primary users* or PUs), who have the exclusive right to use these bands. However, these licensed bands are actually not fully used either temporally or spatially. Thus, unlike the traditional spectrum allocation policy, cognitive radio allows any unauthorized user, (*secondary user*

or SU), to use the licensed bands whenever these bands are not occupied by PUs. Therefore, CR can effectively enhance the spectrum efficiency.

Spectrum sensing is the essential front-end mechanism for CR. The detection methods often used for single-reception spectrum sensing are *matched filtering* approach [4, 5], *feature detection* approach [6, 7], and *energy detection* approach [8, 9, 4, 10, 11, 12]. The matched filtering method can maximize the signal-to-noise ratio (SNR) inherently. However it is difficult to carry out the detection without signal information regarding the pilots and the frame structure. The feature detection method is primarily based on cyclostationarity, and it also relies on the given crucial statistical information about the PU signals. The energy detection method is the most popular one since it does not need any statistical information about the signal to be detected. Nevertheless, when the signal energy fluctuates substantially in time or noise power is large, it becomes quite difficult to distinguish between the absence and the presence of the PU signal(s) [4, 5]. In our previous work, we proposed a local spectrum-sensing method based on the Jarque-Bera (JB) statistics [13]. In this dissertation, we propose to extend this promising single-reception detection method to cooperative spectrum-sensing in both stationary and time-varying environments. As exhibited by our Monte Carlo simulation results, the performance of the JB statistics based spectrum-sensing technique is superior to that of the energy-detection based spectrum-sensing scheme.

In [9, 8], two optimal cooperative spectrum-sensing schemes were proposed subject to single-reception energy detection. In [9], the received signal energy estimates of all local detectors need to be sent to the fusion center (FC). Besides, the precise estimator is indispensable at each local detector to estimate the PU signal's strength and the noise variance. These information need to be sent to the FC as well. Thus, the FC can apply the *criterion of the deflection coefficient maximization* to determine

the optimal fusion weights. However, a large signal bandwidth is required for communications between the SUs and the FC, which is impractical. In [8], the square-law combined scalar of the signal energy experienced at each local detector is sent to the FC; then the PU signal power estimate can be established at the FC. Although the technique proposed in [8] does not require a large transmission bandwidth as [9], the underlying assumption that the actual noise power is given is not realistic. Therefore, in order to save transmission bandwidth and facilitate a novel totally blind cooperative spectrum-sensing scheme, in this dissertation, we assume that only local detection decisions are sent to the FC and no *a priori* knowledge of signals and noises at the local detectors is known to the FC. The optimal weights are obtained by simple counting without complicated PU signal strength or noise variance estimations.

Although the optimal data-fusion rule was first proposed in [14] for cooperative spectrum-sensing, the difficulty arises as the probabilities of miss detection and false alarm for each sensing node are required to be known prior to final decision (global detection). The existing estimators need to store all of the local decisions for a while to build the reliable ensemble averages as the aforementioned probabilities [1, 15, 16, 17]. They are obviously impractical, especially when the time-varying characteristics of the signal and the environment are conspicuous. In addition, the optimal data fusion rule cannot be implemented on-line if it relies on these ensemble-average probability estimators. In other words, they need large memory spaces to store the historical local decisions and they cannot adapt to fast time-variance emerging in the system. To tackle this problem, in this dissertation, we propose a novel on-line recursive estimator built upon a temporal *discount factor* so that one can adaptively estimate the essential parameters involved in the optimal data-fusion rule, based on our previous work [13, 17]. Thus, only four parameters are needed to be stored and updated at every sample time instant for each sensing node. Furthermore, by using this tem-

poral discount factor, the cooperative spectrum-sensing scheme can react and tack the time-varying environment more quickly. With this new mechanism, we establish a new on-line implementation scheme for the optimal fusion rule and facilitate a novel cooperative spectrum-sensing system using JB statistics, which can be applied to time-varying environment effectively. The theoretical analysis to demonstrate the advantage of our proposed system and to numerically determine the optimal discount factor is also provided in this dissertation. Simulation results also demonstrate that our new method is much more robust than other existing approaches [18, 17, 19].

1.2.2 Symbol Detection and Channel Estimation

Orthogonal frequency-division multiplexing (OFDM) is a prevalent modulation technology for carrying digital data on a large number of closely-spaced orthogonal subcarriers. The major advantage of OFDM over traditional single-carrier technologies is its capability of coping with severe channel conditions, such as attenuation of high-frequency components in the frequency spectrum of a long copper wire, narrowband interference and/or frequency-selective fading due to multipath medium, etc., in the absence of sophisticated equalizers. Therefore, recently, OFDM has become a very popular modulation technology for wideband wired and wireless communications, including digital television, digital audio broadcast, digital subscriber line (DSL), wireless networks, powerline networks, and fourth-generation (4G) mobile communications.

On the other hand, *pilot-symbol-assisted scheme* is commonly employed to help acquire the channel information, where training pilots and data symbols are placed on different subcarriers prior to OFDM modulation [20, 21, 22, 23]. However, these training pilots consume valuable bandwidth and thus reduce the data rate (spectral efficiency) as well. Instead, an alternative emerged as the *superimposed training* (ST) *scheme* in [24], where training pilots are added on top of data symbols prior to OFDM

modulation. The major advantage of the ST scheme is no loss in data rate. However, in addition to sacrificing useful transmission power for carrying superimposed pilots, there is another disadvantage due to the interference from data symbols to pilots, which would severely impair the later channel estimation at the receiver.

To combat the aforementioned problems, a novel *data-dependent superimposed-training* (DDST) technique for single-carrier communication systems was introduced in [25], where the (unknown) data-induced interference during channel estimation was eliminated by nulling some frequency tones of the information data at the transmitter. According to the simulation results in [25], the DDST technique can lead to a much better channel-estimation accuracy than the previous ST scheme.

Furthermore, the DDST scheme was extended to renovate OFDM systems in [26]. To avoid the permanent loss of certain data symbols due to the nulling operation on the corresponding subcarriers, the information data is first precoded and then a subset of the precoded data are nulled on the fixed, equal-spaced subcarriers where the training pilots can be inserted (placed) afterwards [26]. The precoder is used to spread each information symbol over all the subcarriers (with different weights) so as to increase the transmission diversity and mitigate the impairment caused by the above-stated nulling operation at the transmitter. This technique in [26] avoids any data-rate loss resulting from the insertion of the frequency-division multiplexed pilots at the cost of the transmitted signal distortion to some extent.

To further lessen the above-mentioned signal distortion, in this dissertation, we propose a new pilot insertion technique (PIT) incorporated with three pilot detection techniques (PDTs) to blindly and accurately detect the pilot positions for future OFDM systems [27, 28]. Considering the impact of the pilot positions on the time-domain signal waveform variations, at the transmitter, we propose to select the optimal pilot positions to minimize the distortion of the original transmitted sig-

nal, which is caused by the subcarrier-removal (nulling operation) in the frequency domain. Later, our proposed new OFDM receiver structure will be introduced in this dissertation, where three different blind pilot-detection techniques are designed without any *a priori* knowledge of the pilot positions dynamically selected by the transmitter. After the pilot positions are blindly detected, channel equalization will then be performed and information symbols can be recovered iteratively thereupon. Besides, rigorous theoretical analysis and Monte Carlo simulation results for various OFDM systems over multipath fading channels will also be presented.

1.2.3 Blind Encoder Identification

Adaptive modulation and coding (AMC) or link adaptation is broadly used in wireless communications to adapt the modulation, coding, or other signal/protocol parameters to the time-varying channel quality, such as the path-loss, the interference due to signals coming from other transmitters, the sensitivity of the receiver amplifier, the available transmitter power margin, and etc. AMC systems really improve the rate of transmission and/or the bit error rate, by exploiting the channel state information and selecting the best modulation and coding combination accordingly from the predefined modulation/coding candidate sets [29, 30, 31, 32]. Especially over fading channels, AMC systems exhibit great performance enhancements compared to the communication systems using fixed modulation/coding schemes [33].

However, in order to synchronously adapt the corresponding demodulation and decoding mechanisms at the receiver to the changes occurred at the transmitter, a monitor or similar mechanisms need to be undertaken to communicate between the transmitter and the receiver, which would obviously reduce the energy and spectrum efficiency. Lately, the blind identification approaches for the AMC systems have drawn tremendous research interest [30, 31, 34]. In [30] and [31], the blind channel-encoder identification methods (classification among a predefined candidate scheme

set) were proposed for binary low-density parity-check (LDPC) codes and binary convolutional codes, respectively. The mathematical formulae for the log-likelihood ratios (LLRs) of the syndrome *a posteriori* probabilities were also established over the binary Galois field or $\mathbb{GF}(2)$. However, the definition and the calculation of the log-likelihood ratio vectors (LLRVs) of the syndrome *a posteriori* probabilities over non-binary Galois fields are quite different from those over $\mathbb{GF}(2)$ in two ways. First, the LLRVs for the syndrome *a posteriori* probabilities over $\mathbb{GF}(q)$ are defined as $(q-1)$-dimensional vectors in contrast to real-valued scalars for the binary case [35]. Second, the calculation procedure for the LLRVs of the syndrome *a posteriori* probabilities is a recursion involving all the LLRVs of the received symbols' *a posteriori* probabilities and the parity-check relations (which will be introduced in the later context), while it only requires single-step (non-recursive) calculation for the binary case. Therefore, further study for the blind non-binary channel-encoder identification is needed. In [34], a blind channel-encoder identification scheme for non-binary convolutional codes was proposed in the absence of noise, which cannot be deemed realistically in practice.

In this dissertation, we propose a new *blind non-binary channel-encoder identification* method over $\mathbb{GF}(q)$ [36]. This new scheme will focus on the identification of Reed-Solomon (RS) codes over $\mathbb{GF}(q)$. However, it could also be easily extended to any other non-binary channel-encoder identification as long as there exist similar parity-check relations for the channel codes to be used. Our proposed blind encoder identification system involves a parameter estimator built upon the expectation-maximization (EM) algorithm followed by the blind identifier of the channel encoder subject to a predefined encoder candidate set. Besides, we also decrease the computational complexity for calculating the LLRVs of the syndrome *a posteriori* probabilities by simplifying the procedure in conjunction with a small-size look-up table. The

computational complexity comparison will be facilitated and the effectiveness of our proposed new scheme evaluated via Monte Carlo simulations will be demonstrated finally.

1.3 Notations

The statistical expectation is denoted by $\mathbb{E}\{\ \}$. A k-combination of an n-element set is denoted by $\mathbb{C}_k^n = \frac{n!}{k!(n-k)!}$, where $n!$ is an n factorial. For an arbitrary set $\mathcal{S}$, $|\mathcal{S}|$ is the size of $\mathcal{S}$. The event $\blacklozenge$ can denote either the occurrence of miss detection or the occurrence of false alarm in this dissertation. A vector is denoted by $\vec{a}$ and a matrix is denoted by $\widetilde{A}$. A *zero vector* $\vec{0}$ is a column vector whose entries are all zero. The transpose and the Hermitian adjoint of a matrix $\widetilde{A}$ are denoted by $\widetilde{A}^T$ and $\widetilde{A}^H$, respectively. $\widetilde{I}_N$ represents the identity matrix of size $N \times N$. An $N \times N$ diagonal matrix is represented by $\mathrm{diag}(a_0, a_1, \ldots, a_{N-1})$ such that the arguments inside the parentheses denote the diagonal elements in the corresponding order. The set of all positive integers, real numbers, and complex numbers are denoted by $\mathcal{Z}^+$, $\mathcal{R}$, and $\mathcal{C}$, respectively. The symbol $j \stackrel{\text{def}}{=} \sqrt{-1}$ is reserved throughout this dissertation. The operators $\mathfrak{Re}\{\ \}$ and $\mathfrak{Im}\{\ \}$ indicate the real and imaginary parts of the complex number inside the braces, respectively. $\lfloor b \rfloor$ is the largest integer which is smaller than or equal to b, where $b \in \mathcal{R}$. $\mathbb{N}(a, b)$ stands for the Gaussian process with mean a and variance b. $P_r\{\mathbf{A}\}$ and $p(\mathbf{A})$ represent the cumulative distribution function (CDF) and the probability density function (PDF) of an event $\mathbf{A}$, respectively. The Galois field of size q is denoted by $\mathbb{GF}(q)$, while $\oplus_q$, $\odot_q$, and $\otimes_q$ denote the addition, multiplication, and element-wise multiplication over $\mathbb{GF}(q)$, respectively.

2 Spectrum Sensing

2.1 Single-Reception Spectrum Sensing

In this section, the local (single reception) signal detection model for spectrum sensing will be introduced. The JB statistic feature is adopted in the proposed spectrum sensing system. The details are presented in the following subsections.

2.1.1 Signal Model for Spectrum Sensing

Denote the continuous-time received signal by $r_c(t)$ during the sensing period. The underlying signal from the primary users is denoted by $s_c(t)$ and $w_c(t)$ is the additive white Gaussian noise (AWGN). Hence, we obtain

$$r_c(t) \stackrel{\text{def}}{=} s_c(t) + \omega_c(t). \tag{2.1}$$

Assume that we are interested in the frequency band with the central frequency f_c and the bandwidth W. We sample the received signal at a sampling rate f_s, where $f_s \geq 2(f_c + W)$. Let $T_s = \frac{1}{f_s}$ be the *sampling period* and N be the *sample size*. For notational convenience, we denote

$$\begin{aligned} r(n) &\stackrel{\text{def}}{=} r_c(nT_s), \ \ n = 1, \ldots, N, \\ s(n) &\stackrel{\text{def}}{=} s_c(nT_s), \ \ n = 1, \ldots, N, \\ \omega(n) &\stackrel{\text{def}}{=} \omega_c(nT_s), \ \ n = 1, \ldots, N. \end{aligned} \tag{2.2}$$

According to [37], for the signal detection (local spectrum sensing) problem, there involve two hypotheses, namely H_0: signal is absent and H_1: signal is present. The

discrete-time received signals under these two hypotheses are given by

$$\begin{aligned} H_0 &: r(n) = \omega(n), \\ H_1 &: r(n) = s(n) + \omega(n), \end{aligned} \tag{2.3}$$

where $r(n)$ denotes the received signal samples, perhaps enduring the effects of path loss, multipath fading, and time dispersion, and $\omega(n)$ is the discrete-time AWGN with zero mean and variance σ^2. Here $s(n)$ can be considered as the superposition of the signals emitted from multiple primary users. When the received signal $r(n)$ consists of multiple sources (from either multiple independent sources or a single source signal traveling through multiple paths), it is usually modeled as the *correlated signal* [37]. It is assumed that signal and noise are uncorrelated with each other. The local spectrum sensing (or signal detection) problem is therefore to determine whether the signal $s(n)$ exists or not, based on the received signal samples $r(n)$ [37, 3].

2.1.2 Energy Based Sensing Algorithm

The energy detector senses spectrum holes by determining whether the primary signal is absent or present in a given frequency slot. The energy detector typically operates without prior knowledge of the primary signal parameters. Its key parameters, including detection threshold, number of samples, and estimated noise power, determine the detection performance. More specifically, the energy detector measures the energy associated with the received signal over a specified time duration and bandwidth. The measured value is then compared with an appropriately selected threshold to determine the presence or the absence of the primary signal.

Assume that the predetermined threshold is θ_0. Therefore, the energy based sensing algorithm can be given by

$$H_0 : \text{if } \sum_{n=1}^{N} |r(n)|^2 < \theta_0; \tag{2.4}$$

$$H_1 : \text{if } \sum_{n=1}^{N} |r(n)|^2 \geq \theta_0. \tag{2.5}$$

2.1.3 Jarqur-Bera (JB) Statistic Based Sensing Algorithm

The JB statistic based detection algorithm is used in this dissertation as an efficient single reception method. It involves three major aspects, namely (i) pre-processing, and (ii) JB-statistic based detection. They are summarized as follows.

Pre-Processing

The pre-processing step for transforming the received signal $r(n)$ into the frequency domain is the same as the HOS detection method [38]. The block diagram of the new spectrum sensing method is depicted in Fig. 2.1. When the signal $r(n)$ is received, first we multiply $r(n)$ by $e^{-j2\pi f_c n T_s}$ to down-convert it to the baseband. Then, this baseband signal is sent through a digital image rejection low-pass (LP) filter with bandwidth $BW_r = 8 \times 10^6 \times \frac{2\pi}{f_s}$ radians. The image rejection filter is placed in the receiver so that the image frequencies along with other unwanted signals are filtered out to enhance the signal quality. Next, the enhanced signal $r_2(n)$ is further multiplied by $e^{-j2\pi f_v n T_s}$. Then, the resulted signal $r_3(n)$ goes through the operations consisting of a down-sampler following a digital anti-aliasing filter whose bandwidth is given by $\text{BW}_a = \frac{N_{\text{FFT}}}{T_{\text{sensing}}} \times \frac{2\pi}{f_s}$,where N_{FFT} is the FFT window size, and $T_{\text{sensing}} = \frac{n}{f_s}$ is the sensing time. The down-sampling rate f_d is given by $f_d = \text{floor}\left(\frac{2\pi}{BW_a}\right)$, where the function "floor" is the operation to round $\frac{2\pi}{BW_a}$ to the nearest integer less than or equal to $\frac{2\pi}{BW_a}$. The down-sampled signal $r_5(n)$ is sent to a serial-to-parallel port

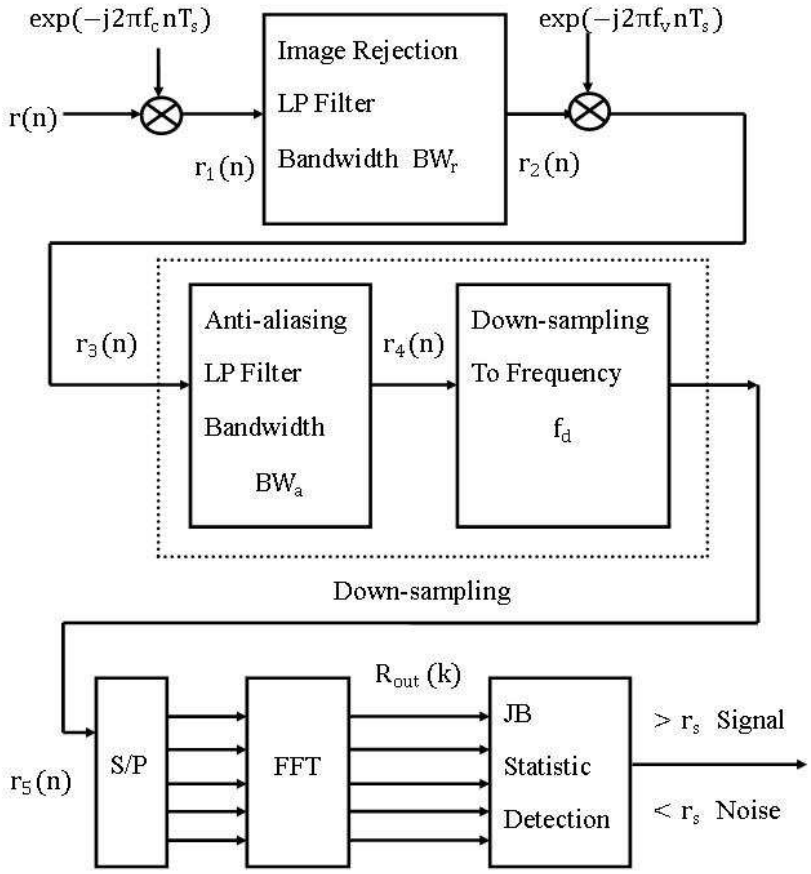

Figure 2.1. The (single-reception) spectrum sensing system diagram.

and then the $N_{\rm FFT}$-point FFT will be taken to result in a half-period FFT-sequence $R_{\rm out}(k)$, $k = 0, 1, \ldots, \frac{N_{\rm FFT}}{2} - 1$.

JB-Statistic Based Detection

In statistics, the Jarque-Bera test is a goodness-of-fit measure of departure from normality, based on the sample kurtosis and the sample skewness. The test statistic, JB, is defined as

$$\mathrm{JB} \stackrel{\text{def}}{=} \frac{n_s}{6}\left(\mathcal{S}^2 + \frac{(\mathcal{K}-3)^2}{4}\right), \tag{2.6}$$

where n_s is the number of observations (or degrees of freedom in general); $\mathcal{S}$ is the sample skewness and $\mathcal{K}$ is the sample kurtosis. They are defined as

$$\mathcal{S} \overset{\text{def}}{=} \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n_s}\sum_{l=1}^{n_s}(x_l - \bar{x})^3}{\left(\frac{1}{n_s}\sum_{l=1}^{n_s}(x_l - \bar{x})^2\right)^{3/2}}, \tag{2.7}$$

$$\mathcal{K} \overset{\text{def}}{=} \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n_s}\sum_{l=1}^{n_s}(x_l - \bar{x})^4}{\left(\frac{1}{n_s}\sum_{l=1}^{n_s}(x_l - \bar{x})^2\right)^{2}}, \tag{2.8}$$

where $\hat{\mu}_3$ and $\hat{\mu}_4$ are the estimates of the third and fourth central moments, respectively; x_l, $l = 1, \ldots, n_s$ are the observations; $\bar{x}$ is the sample mean and $\hat{\sigma}^2$ is the estimate of the second central moment or the variance.

Since $R_{\text{out}}(k)$, $k = 0, 1, \ldots, \frac{N_{\text{FFT}}}{2} - 1$ are complex-valued, the absolute values of $R_{\text{out}}(k)$, $k = 0, 1, \ldots, \frac{N_{\text{FFT}}}{2} - 1$ are used instead [39]. That is, invoke Eqs. (2.6), (2.7), and (2.8) to calculate the JB statistic of $|R_{\text{out}}(k)|$ and compare it with the threshold r_s to decide if there exists the signal $s(n)$. If JB $> r_s$, it infers that the signal exists; otherwise (JB $\leq r_s$), the signal is absent. The next subsection will present the theoretical analysis regarding how to select the appropriate threshold r_s.

2.1.4 Simulation Results: Energy Detector vs. JB Detector

In this simulation, we use the commonly-used microphone signal (as the PU signal) to benchmark the spectrum-sensing methods, whose details can be found in [13, 40]. The details of the single-reception spectrum sensing simulation set-up are in compliance with the IEEE 802.22 standard, which can also refer to [13, 40]. In the following, we will present the simulation results to compare our proposed JB-statistic-based spectrum-sensing method and the energy-based spectrum-sensing scheme in the low SNR scenario. We carry out the statistical averages over 2,000 Monte Carlo trials to quantify the detection performances. Figures 2.2, 2.3, 2.4, 2.5, and 2.6 delineate the corresponding ROC (receiver-operating characteristic) curves for the SNR values of -10, -15, -20, -25 dB and -27dB, respectively. According to Figures 2.2-2.6, it

is clear that when noise is large, our proposed JB-statistic-based spectrum-sensing technique (denoted by "JB based detector" in these figures) always outperforms the commonly-used energy-based spectrum-sensing technique (denoted by "Energy based detector" in the figures). Obviously, the performance margin is very large especially for very low SNR conditions. Since our proposed JB-statistic-based spectrum-sensing technique achieves the better local detection performance, we use this detector for all cooperative spectrum-sensing methods in comparison later on.

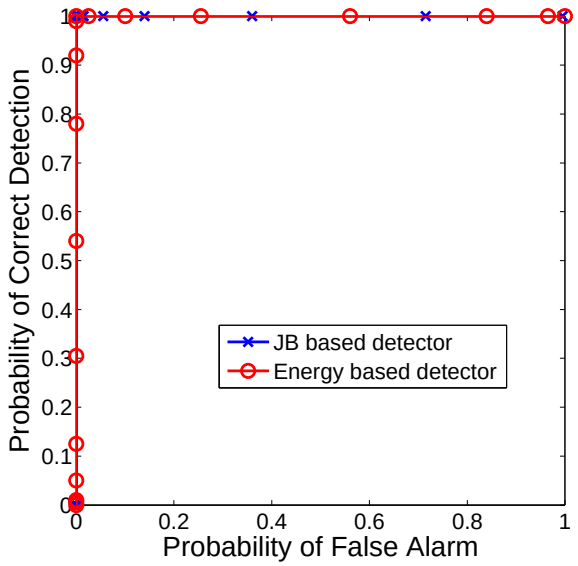

Figure 2.2. ROC curves for our proposed JB-statistic-based detector and the energy-based detector (SNR = -10 dB).

2.1.5 Conclusion

In this section, we propose to use the JB statistic as a feature in the single reception and proposed a novel robust spectrum sensing scheme. According to our Monte Carlo simulation results for the wireless microphone signals, the JB-statistics based detection method is more robust than the commonly used energy-detection based spectrum-sensing scheme in a very broad SNR range.

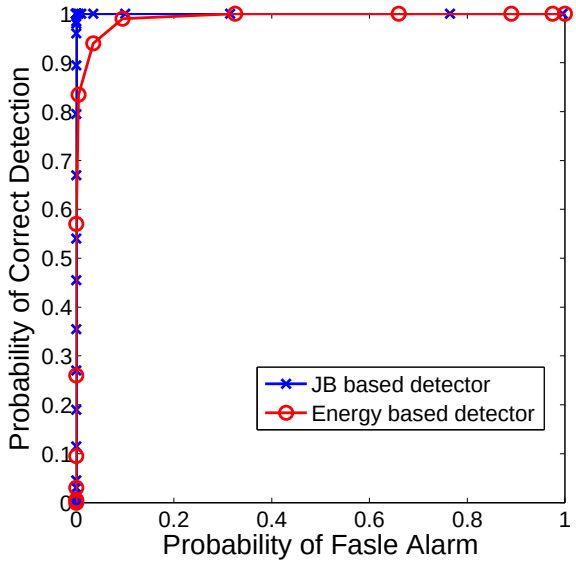

Figure 2.3. ROC curves for our proposed JB-statistic-based detector and the energy-based detector (SNR = -15 dB).

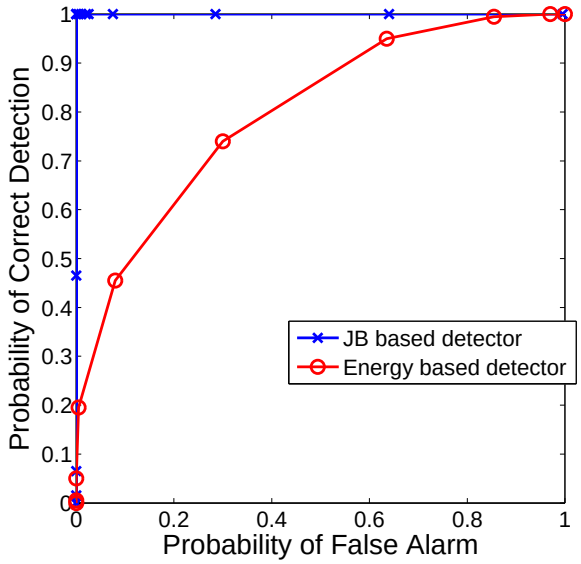

Figure 2.4. ROC curves for our proposed JB-statistic-based detector and the energy-based detector (SNR = -20 dB).

2.2 Cooperative Spectrum Sensing

When multiple receivers are available, the cooperative spectrum sensing methods are feasible for more reliable performance than local spectrum sensing schemes. In coop-

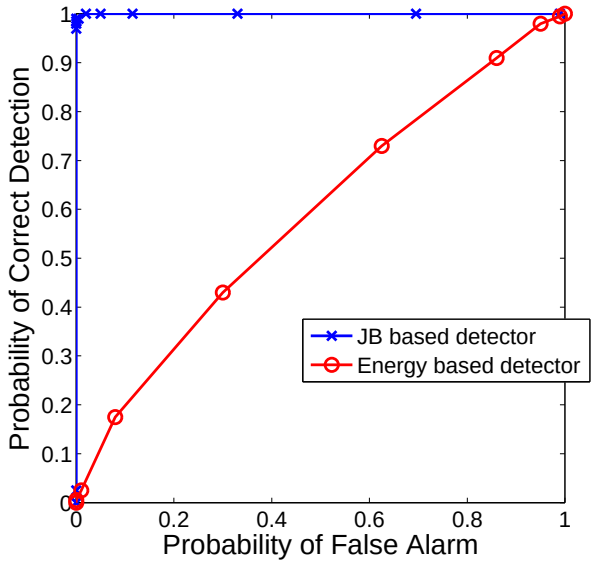

Figure 2.5. ROC curves for our proposed JB-statistic-based detector and the energy-based detector (SNR = -25 dB).

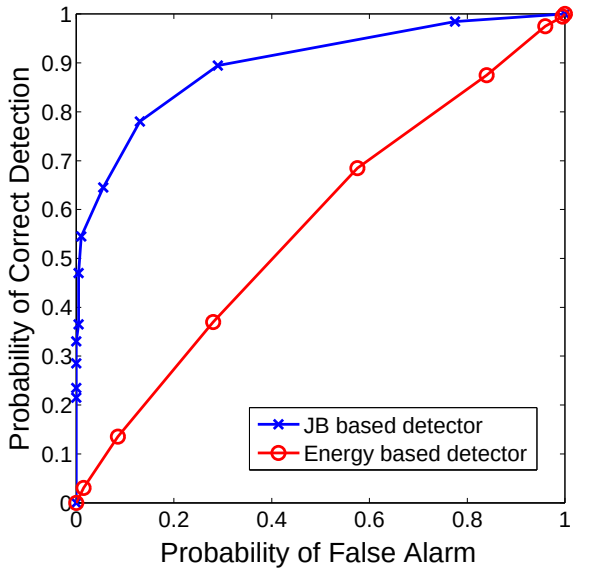

Figure 2.6. ROC curves for our proposed JB-statistic-based detector and the energy-based detector (SNR = -27 dB).

erative spectrum sensing, the data fusion mechanism is crucial. Usually, transmitting additional data, such as the likelihood ratio, credibility or raw detection data, to the center node can increase the reliability of the global decisions. However, this is not

feasible in practice since it requires transmission of sensor information in real time and a large communication bandwidth [14]. Thus, signal processing mechanisms are preferred to be performed at the local sensing nodes. Here, we consider that all the signal processing procedures are done at the local sensors and only the local decisions are transmitted to the center node to make a global decision. In such a condition, the optimal data fusion rule of the distributed detection system is given by [14]. Under this rule, the detection statistics can be formulated as the weighted sum of the local decisions, and the weights are the functions of the miss detection probabilities and the false alarm probabilities experienced at the local sensing nodes. This weighted-sum data-fusion rule can be summarized in the following subsections.

2.2.1 Optimal Data Fusion for Cooperative Spectrum Sensing

Consider a binary hypothesis test as follows:

$$\begin{aligned} H_0 &: \text{ PU signal is absent,} \\ H_1 &: \text{ PU signal is present.} \end{aligned} \tag{2.9}$$

The *a priori* probabilities of the two hypotheses are $P(H_0) = P_0$, and $P(H_1) = P_1$. Assume that there are M local sensing nodes, and the decisions are denoted by u_i, $i = 1, 2, \ldots, M$, where

$$u_i = \begin{cases} -1, & \text{if } H_0 \text{ is true,} \\ +1, & \text{if } H_1 \text{ is true.} \end{cases} \tag{2.10}$$

Also, we denote the false alarm probabilities and the miss detection probabilities by P_{F_i} and P_{M_i}, respectively, where $i = 1, 2, \cdots, M$.

At the center node, the optimal (global) decision rule is actually subject to the likelihood ratio test such that

$$\frac{P(u_1, \cdots, u_M|H_1)}{P(u_1, \cdots, u_M|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}, \tag{2.11}$$

where C_{00}, C_{01}, C_{10}, and C_{11} are the *decision costs*. The minimum probability of error criterion here is set as [41], namely, $C_{00} = C_{11} = 0$ and $C_{01} = C_{10} = 1$. Define $\overrightarrow{\mathbf{u}} \stackrel{\text{def}}{=} [u_1\ u_2\ \cdots\ u_M]$. Thus, Eq. (2.11) becomes

$$\frac{P(\overrightarrow{\mathbf{u}}|H_1)}{P(\overrightarrow{\mathbf{u}}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1}, \tag{2.12}$$

and the corresponding log-likelihood ratio test is

$$\log\left(\frac{P(\overrightarrow{\mathbf{u}}|H_1)}{P(\overrightarrow{\mathbf{u}}|H_0)}\right) + \log\left(\frac{P_1}{P_0}\right) \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{2.13}$$

Assume that the decisions of different local sensing detectors are independent. We get

$$\begin{aligned} P(\overrightarrow{\mathbf{u}}|H_1) &= \prod_{i=1}^{M} P(u_i|H_1) \\ &= \prod_{S_+}(1 - P_{M_i}) \cdot \prod_{S_-} P_{M_i}, \end{aligned} \tag{2.14}$$

where S_+ is the set of all i such that $u_i = +1$ and S_- is the set of all i such that $u_i = -1$. Similarly, we have

$$\begin{aligned} P(\overrightarrow{\mathbf{u}}|H_0) &= \prod_{i=1}^{M} P(u_i|H_0) \\ &= \left(\prod_{i\in S_+} P_{F_i}\right) \times \left(\prod_{i\in S_-}(1 - P_{F_i})\right). \end{aligned} \tag{2.15}$$

Substituting Eqs. (2.14) and (2.15) into Eq. (2.13), we have

$$\log\left(\frac{P_1}{P_0}\right) + \sum_{i\in S_+} \log\left(\frac{1 - P_{M_i}}{P_{F_i}}\right) + \sum_{i\in S_-} \log\left(\frac{P_{M_i}}{1 - P_{F_i}}\right) \underset{H_0}{\overset{H_1}{\gtrless}} 0, \tag{2.16}$$

which can also be expressed as

$$w_0 + \sum_{i=1}^{M} w_i\, u_i \underset{H_0}{\overset{H_1}{\gtrless}} 0, \tag{2.17}$$

where w_i $(i = 1, \ldots, M)$ is the weight of the i^{th} local sensing detector and w_0 is a function of the *a priori* probabilities. They can be given by

$$w_0 = \log\left(\frac{P_1}{P_0}\right), \tag{2.18}$$

and

$$w_i = \begin{cases} \log\left(\frac{1-P_{M_i}}{P_{F_i}}\right), & \text{if } u_i = +1, \\ \log\left(\frac{1-P_{F_i}}{P_{M_i}}\right), & \text{if } u_i = -1, \end{cases} \qquad i = 1, 2, \cdots, M. \tag{2.19}$$

To implement this optimal data fusion rule given by Eq. (2.15), one must know the weights in Eq. (2.17), which are determined by P_0, P_1, P_{M_i}, and P_{F_i}. However, these probabilities are not given in practice. Thus, we need to estimate these weights from the detection information (local decisions) we can get from the local sensing detectors.

2.2.2 Estimation of the Weights

The cooperative spectrum sensing is more reliable than the single-reception spectrum sensing. Therefore, we often use the global decision from the cooperative spectrum sensing as the ground truth, u_0, to estimate the probabilities of miss detection and false alarm. By continuously comparing the local decisions with the ground truth, the local probabilities of miss detection and false alarm can be estimated, so the weights in Eq. (2.17) can be adaptively updated.

For the i^{th} local sensing detector at the m^{th} moment, $\varepsilon_i(m)$ denotes the outcome, and $\varepsilon_i(m) \in \{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}$, where they are specified as four *states* below:

$$\begin{aligned} \varepsilon_1 &: \ u_0 = +1 \text{ and } u_i = +1, \\ \varepsilon_2 &: \ u_0 = -1 \text{ and } u_i = -1, \\ \varepsilon_3 &: \ u_0 = +1 \text{ and } u_i = -1, \\ \varepsilon_4 &: \ u_0 = -1 \text{ and } u_i = +1. \end{aligned} \tag{2.20}$$

Thus, we can get the *cumulative state* $\mathcal{C}_i(m)$ of the i^{th} local sensing detector at the m^{th} detection time slot. It is given by

$$\begin{aligned} \mathcal{C}_i(m) &\stackrel{\text{def}}{=} \sum_{k=1}^{m} \varepsilon_i(k) \\ &= \alpha_{1i}(m)\varepsilon_1 + \alpha_{2i}(m)\varepsilon_2 + \alpha_{3i}(m)\varepsilon_3 + \alpha_{4i}(m)\varepsilon_4, \end{aligned} \tag{2.21}$$

where α_{1i}, α_{2i}, α_{3i}, and α_{4i} are the cumulative times that ε_1, ε_2, ε_3, and ε_4 have occurred, respectively. Thus, we obtain

$$\begin{aligned} \hat{P}_{M_i}(m) &= \frac{\alpha_{3i}(m)}{\alpha_{1i}(m) + \alpha_{3i}(m)}, \\ \hat{P}_{F_i}(m) &= \frac{\alpha_{4i}(m)}{\alpha_{2i}(m) + \alpha_{4i}(m)}, \end{aligned} \tag{2.22}$$

and

$$\frac{\hat{P}_1(m)}{\hat{P}_0(m)} = \frac{\alpha_{1i}(m) + \alpha_{3i}(m)}{\alpha_{2i}(m) + \alpha_{4i}(m)}, \tag{2.23}$$

where $\hat{P}_{M_i}(m)$, $\hat{P}_{F_i}(m)$, $\hat{P}_1(m)$, and $\hat{P}_0(m)$ are the estimates for P_{M_i}, P_{F_i}, P_1, and P_0, respectively at the m^{th} detection time slot. The estimated weights in Eq. (2.17) at the m^{th} detection time slot can be determined as

$$\hat{w}_0(m) = \log\left(\frac{\hat{P}_1(m)}{\hat{P}_0(m)}\right) = \log\left(\frac{\alpha_{1i}(m) + \alpha_{3i}(m)}{\alpha_{2i}(m) + \alpha_{4i}(m)}\right), \tag{2.24}$$

and

$$\hat{w}_i(m) = \begin{cases} \log\left(\frac{1-\hat{P}_{M_i}(m)}{\hat{P}_{F_i}(m)}\right) = \log\left(\frac{\alpha_{1i}(m)}{\alpha_{4i}(m)}\right) - \hat{w}_0(m), \\ \qquad\qquad\qquad\qquad \text{if } u_i = +1, \\ \log\left(\frac{1-\hat{P}_{F_i}(m)}{\hat{P}_{M_i}(m)}\right) = \log\left(\frac{\alpha_{2i}(m)}{\alpha_{3i}(m)}\right) + \hat{w}_0(m), \\ \qquad\qquad\qquad\qquad \text{if } u_i = -1, \end{cases} \tag{2.25}$$

where $i = 1, 2, \cdots, M$.

2.2.3 Temporal Discount Factor

It is obvious that the estimated probabilities of miss detection and false alarm given by Eq. (2.22) will converge eventually when the environment is *stationary* with a

fixed SNR. However, this assumption is often unrealistic. When the environment of a certain local detector is time-varying, the cumulative states, which would have been misled by the history, could slow the convergence speeds of the estimated parameters. For example, if the noise of the i^{th} local sensing detector is time varying, the received signal in Eq. (2.3) should be modified as

$$\begin{aligned} H_0 &: r_i^{(m)}(n) = \upsilon_i^{(m)} \bar{\omega}_i(n), \\ H_1 &: r_i^{(m)}(n) = s_i^{(m)}(n) + \upsilon_i^{(m)} \bar{\omega}_i(n), \end{aligned} \tag{2.26}$$

where $\bar{\omega}_i(n) \propto \mathcal{N}(0,1), \forall i$ are *normalized* AWGN with zero mean and unity variance, and $\upsilon_i^{(m)}$ is a factor varying with respect to m, $m = 1, 2, \ldots$. Thus, the SNR of the i^{th} local sensing detector at the m^{th} sensing interval can be written as

$$\mathcal{SNR}_i^{(m)} \stackrel{\text{def}}{=} \frac{\sum_{n=1}^{N} \left| s_i^{(m)}(n) \right|^2}{\sum_{n=1}^{N} \left| \upsilon_i^{(m)} \bar{\omega}_i(n) \right|^2}, \qquad i = 1, 2, \ldots, M. \tag{2.27}$$

Therefore, a sudden SNR change at the m^{th} sensing interval at a certain local sensing node i could be formulated as a sudden change in the value of $\upsilon_i^{(m)}$. Assume that SNR is constant within a sensing interval, and sudden changes in SNR only occur between different sensing intervals.

When the environment is time-varying, the convergence speed (from the original probability of miss detection or false alarm to the new probability of miss detection or false alarm) of the algorithm in Section 2.2.2 would be quite slow. Especially when the cumulative states have been aggregated for a long time, any abrupt SNR change would make the *system trackability* fail.

A window of fixed length γ was used in [42] to retain the latest γ local decisions at each local detector and discard all the decisions in the past. Although this method can mitigate the time-varying problem to some extent, it treats all the γ recent decisions equally and the corresponding trackability would still be in concern. In order for

our proposed scheme to react promptly and accommodate the abrupt environmental changes, we adopt a *temporal discount factor*, ζ, from the *reinforcement learning* in [43] to pose a discount on the influence of the past cumulative states. Consequently, the influence of all local decisions will be discounted exponentially with time.

Hence, the cumulative state $\mathcal{C}_i^{(m)}$ can be rewritten as

$$\begin{aligned} \mathcal{C}_i^{(m)} &\stackrel{\text{def}}{=} \alpha'_{1i}(m)\varepsilon_1 + \alpha'_{2i}(m)\varepsilon_2 + \alpha'_{3i}(m)\varepsilon_3 + \alpha'_{4i}(m)\varepsilon_4 \\ &= \sum_{k=1}^{\mu} \left(\xi_i^{(\mathcal{S}_k^+)}\right) \zeta^{\mu-k} + \sum_{k=1}^{\nu} \left(\xi_i^{(\mathcal{S}_k^-)}\right) \zeta^{\nu-k} \end{aligned} \tag{2.28}$$

where the discount factor ζ satisfies $0 < \zeta \leq 1$. Note that $\mathcal{S}^+$ and $\mathcal{S}^-$ denote the sets of time slots corresponding to global decisions H_1 and H_0, respectively. They are

$$\mathcal{S}^+ \stackrel{\text{def}}{=} \left\{ m \,\middle|\, u_0^{(m)} = 1 \right\}, \tag{2.29}$$

$$\mathcal{S}^- \stackrel{\text{def}}{=} \left\{ m \,\middle|\, u_0^{(m)} = 0 \right\}, \tag{2.30}$$

and μ, ν are the sizes of $\mathcal{S}^+$, $\mathcal{S}^-$, respectively, such that $\mu = \left|\{m|u_0^{(m)} = 1\}\right|$ and $\nu = \left|\{m|u_0^{(m)} = 0\}\right|$, where $\mu + \nu = m$. We sort the elements of $\mathcal{S}^+$ and $\mathcal{S}^-$ both in ascending order. Thus, $\mathcal{S}_k^+$ and $\mathcal{S}_k^-$ stand for the k^{th} elements of the ordered sets $\mathcal{S}^+$ and $\mathcal{S}^-$, respectively. Thus, for ρ=1 or 3, $\alpha'_{\rho i}(m)$ is given by

$$\alpha'_{\rho i}(m) = \begin{cases} \zeta \times \alpha'_{\rho i}(m-1) + 1, & \text{if } u_0^{(m)} = 1 \text{ and } \xi_i^{(m)} = \varepsilon_\rho \\ \zeta \times \alpha'_{\rho i}(m-1), & \text{if } u_0^{(m)} = 1 \text{ and } \xi_i^{(m)} = \varepsilon_{4-\rho} \end{cases} \tag{2.31}$$

for ρ=2 or 4, $\alpha'_{\rho i}(m)$ is given by

$$\alpha'_{\rho i}(m) = \begin{cases} \zeta \times \alpha'_{\rho i}(m-1) + 1, & \text{if } u_0^{(m)} = 0 \text{ and } \xi_i^{(m)} = \varepsilon_\rho \\ \zeta \times \alpha'_{\rho i}(m-1), & \text{if } u_0^{(m)} = 0 \text{ and } \xi_i^{(m)} = \varepsilon_{6-\rho} \end{cases} \tag{2.32}$$

Consequently, the estimated probabilities and the corresponding estimated weights specified by Eqs. (2.24)-(2.25) can be modified easily by substituting $\alpha_{\rho i}(m)$ with $\alpha'_{\rho i}(m)$, where $\rho = 1, 2, 3, 4$.

Here ζ is used to control the relative influence of the past local decisions. In particular, a local decision received by the center node in the past is discounted exponentially. As we set $\zeta \to 1$, the past local decisions are emphasized more and more. When $\zeta = 1$, the adaptive algorithm here degenerates into the sample-average based estimation method in Section 2.2.2. Thus, by properly choosing the discount factor ζ, one may make the cooperative spectrum-sensing algorithm to adapt swiftly to different environmental changes.

2.2.4 Discount Factor Analysis

According to Eq. (2.22) in conjunction with the substitution of all $\alpha_{\rho i}(m)$'s with $\alpha'_{\rho i}(m)$'s, probability estimators for miss detection and false alarm at the center node are similar to each other. Therefore, in this section, we use $\blacklozenge$ to denote either one of these two events. In other words, $H_{\blacklozenge}$ denotes H_1 for miss detection analysis and $H_{\blacklozenge}$ denotes H_0 for false alarm analysis.

Lemma 1: When the discount factor ζ $(0 < \zeta < 1)$ is involved, the statistical expectation of the estimated probability of $\blacklozenge$ in the m^{th} time slot for the i^{th} local detector is the true probability of $\blacklozenge$ if the environment of the i^{th} local detector is stationary (i.e., $v_i^{(m)}$ is constant for all m).

Proof: Among the previous m successive time slots, there are N_1 time slots corresponding to $H_{\blacklozenge}$. For the i^{th} local detector, there are K_1 time slots of $\blacklozenge$ by the local spectrum sensing among these N_1 time slots. Thus, the true probability of $\blacklozenge$ at the m^{th} time slot for the i^{th} local detector is $\frac{K_1}{N_1}$. The estimated probability of $\blacklozenge$, $\hat{P}_{\blacklozenge_i}(m)$, at the m^{th} time slot is given by

$$\hat{P}_{\blacklozenge_i}(m) = \begin{cases} \frac{K_1}{N_1}, & \zeta = 1 \\ \frac{\psi(K_1,\zeta,N_1)}{1+\zeta+\zeta^2+\cdots+\zeta^{N_1-1}}, & 0 < \zeta < 1 \end{cases} \tag{2.33}$$

where $\psi(K_1, \zeta, N_1)$ is a polynomial consisting of K_1 terms. These K_1 terms can be randomly chosen from K_1 elements in $\Psi \stackrel{\text{def}}{=} \left\{1, \zeta, \zeta^2, \ldots, \zeta^{N_1-1}\right\}$. The probability of each element in Ψ to contribute to $\psi(K_1, \zeta, N_1)$ is

$$\begin{aligned} \mathcal{P} &= \frac{\mathbb{C}_{K_1-1}^{N_1-1}}{\mathbb{C}_{K_1}^{N_1}} \\ &= \frac{(N_1-1)!}{(K_1-1)!(N_1-K_1)!} \times \frac{K_1!(N_1-K_1)!}{N_1!} \\ &= \frac{K_1}{N_1}. \end{aligned} \tag{2.34}$$

Thus, the statistical expectation of the estimated probability of ♦ at the m^{th} time slot for the i^{th} local detector is given by

$$\begin{aligned} \mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\} &= \frac{\frac{K_1}{N_1} \times 1 + \frac{K_1}{N_1} \times \zeta + \cdots + \frac{K_1}{N_1} \times \zeta^{N_1-1}}{1 + \zeta + \zeta^2 + \cdots + \zeta^{N_1-1}} \\ &= \frac{K_1}{N_1}. \end{aligned} \tag{2.35}$$

□

Lemma 2: When the environment of the i^{th} local detector is stationary, the estimated probability of local ♦, $\hat{P}_{\blacklozenge_i}(m)$, can be upper- and lower-bounded. When $\zeta \to 1$, both bounds approach the true probability of local ♦ and they get tighter as ζ gets closer to 1.

Proof: From the proof of Lemma 1, we know that $\hat{P}_{\blacklozenge_i}(m) = \frac{\psi(K_1,\zeta,N_1)}{1+\zeta+\zeta^2+\cdots+\zeta^{N_1-1}}$. Obviously, if we choose the largest K_1 elements from Ψ and constitute the polynomial $\psi(K_1, \zeta, N_1)$, the upper bound for $\hat{P}_{\blacklozenge_i}(m)$ is obtained thereby. Similarly, if we choose the smallest K_1 elements from Ψ and constitute the polynomial $\psi(K_1, \zeta, N_1)$, the lower bound for $\hat{P}_{\blacklozenge_i}(m)$ can be acquired instead. Thus, we have

$$\frac{\zeta^{N_1-K_1}(1-\zeta^{K_1})}{1-\zeta^{N_1}} \leq \hat{P}_{\blacklozenge_i}(m) \leq \frac{1-\zeta^{K_1}}{1-\zeta^{N_1}}. \tag{2.36}$$

We know that the true probability of local ♦ at the m^{th} time slot is $P_{\blacklozenge_i}(m) = \frac{K_1}{N_1}$. Next, we will prove

$$\frac{\zeta^{N_1-K_1}(1-\zeta^{K_1})}{1-\zeta^{N_1}} \leq \frac{K_1}{N_1} \leq \frac{1-\zeta^{K_1}}{1-\zeta^{N_1}}. \tag{2.37}$$

Moreover, the difference between the upper and lower bounds decreases as $\zeta \to 1$.

First, we will prove $\frac{\zeta^{N_1-K_1}(1-\zeta^{K_1})}{1-\zeta^{N_1}}$ is a monotonically increasing function of ζ over $0 < \zeta < 1$. Taking the derivative of $\frac{\zeta^{N_1-K1}(1-\zeta^{K_1})}{1-\zeta^{N_1}}$ with respect to ζ, we get

$$\frac{\partial\left[\frac{\zeta^{N_1-K1}(1-\zeta^{K_1})}{1-\zeta^{N_1}}\right]}{\partial\zeta} = \frac{\zeta^{N_1-K_1-1}\left[N_1 - K_1 - N_1\zeta^{K_1} + K_1\zeta^{N_1}\right]}{(1-\zeta^{N_1})^2}, \tag{2.38}$$

where the derivative of $N_1 - K_1 - N_1\zeta^{K_1} + K_1\zeta^{N_1}$ with respect to ζ is

$$\frac{\partial\left[N_1 - K_1 - N_1\zeta^{K_1} + K_1\zeta^{N_1}\right]}{\partial\zeta} = N_1K_1\zeta^{N_1-1} - N_1K_1\zeta^{K_1-1} \leq 0, \tag{2.39}$$

over $0 < \zeta \leq 1$. Therefore, $N_1 - K_1 - N_1\zeta^{K_1} + K_1\zeta^{N_1}$ is a monotonically decreasing function of ζ over $0 < \zeta \leq 1$. Obviously,

$$\left[N_1 - K_1 - N_1\zeta^{K_1} + K_1\zeta^{N_1}\right]\Big|_{\zeta=1} = 0, \tag{2.40}$$

which indicates that $N_1 - K_1 - N_1\zeta^{K_1} + K_1\zeta^{N_1} \geq 0$ and $\frac{\zeta^{N_1-K1}(1-\zeta^{K_1})}{1-\zeta^{N_1}}$ is a monotonically increasing function of ζ over $0 < \zeta \leq 1$.

According to the L'Hospital's rule,

$$\lim_{\zeta\to 1}\frac{\zeta^{N_1-K_1}(1-\zeta^{K_1})}{1-\zeta^{N_1}} = \lim_{\zeta\to 1}\frac{(N_1-K_1)\zeta^{N_1-K_1-1} - N_1\zeta^{N_1-1}}{-N_1\zeta^{N_1-1}} = \frac{K_1}{N_1}. \tag{2.41}$$

Therefore, $\frac{\zeta^{N_1-K_1}(1-\zeta^{K_1})}{1-\zeta^{N_1}} - \frac{K_1}{N_1} \leq 0$.

A similar procedure can be performed to prove that $\frac{1-\zeta^{K_1}}{1-\zeta^{N_1}} \geq \frac{K_1}{N_1}$ and $\frac{1-\zeta^{K_1}}{1-\zeta^{N_1}}$ is a monotonically decreasing function of ζ over $0 < \zeta < 1$ and

$$\lim_{\zeta\to 1}\frac{1-\zeta^{K_1}}{1-\zeta^{N_1}} = \lim_{\zeta\to 1}\frac{-K_1\,\zeta^{K_1-1}}{-N_1\,\zeta^{N_1-1}} = \frac{K_1}{N_1}. \tag{2.42}$$

□

Lemma 3: When the environment of the i^{th} local detector is time-varying, the probability estimator for ♦ given by Eq. (2.22) with $\zeta = 1$ becomes biased on average.

Proof: Suppose that after N_1 time slots corresponding to $H_\blacklozenge$, the local environment of the i^{th} detector changes, and its probability of ♦ changes accordingly. After

the environment changes until the m^{th} time slot, we collect N_2 time slots corresponding to $H_{\blacklozenge}$ and K_2 time slots of $\blacklozenge$ at the i^{th} local detector. Hence, the probability of $\blacklozenge$ at the i^{th} local detector changes from $\frac{K_1}{N_1}$ to $\frac{K_2}{N_2}$ due to this environment variation. However, the estimated probability of $\blacklozenge$ at the m^{th} time slot is given by

$$\hat{P}_{\blacklozenge_i}(m) = \begin{cases} \frac{K_1+K_2}{N_1+N_2}, & \zeta = 1 \\ \frac{\psi_1(K_1,\zeta,N_1,N_2)+\psi_2(K_2,\zeta,N_2)}{1+\zeta+\zeta^2+\cdots+\zeta^{N_1+N_2-1}}, & 0 < \zeta < 1 \end{cases} \tag{2.43}$$

where $\psi_1(K_1, \zeta, N_1, N_2)$ is a polynomial consisting of K_1 terms. These K_1 terms are randomly chosen from the set $\Psi_1 \stackrel{\text{def}}{=} \left\{\zeta^{N_2}, \zeta^{N_2+1}, \ldots, \zeta^{N_1+N_2-1}\right\}$. Similarly, $\psi_2(K_2, \zeta, N_2)$ is a polynomial consisting of K_2 terms, which are randomly drawn from the set $\Psi_2 \stackrel{\text{def}}{=} \left\{1, \zeta, \ldots, \zeta^{N_2-1}\right\}$. Similar to the proof of Lemma 1, $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ is given by

$$\begin{aligned} &\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\} \\ &= \frac{\frac{K_1}{N_1} \times (\zeta^{N_1+N_2-1} + \zeta^{N_1+N_2-2} + \cdots + \zeta^{N_2}) + \frac{K_2}{N_2} \times (\zeta^{N_2-1} + \zeta^{N_2-2} + \cdots + 1)}{1 + \zeta + \zeta^2 + \cdots + \zeta^{N_1+N_2-1}} \end{aligned} \tag{2.44}$$

and

$$\begin{cases} \lim\limits_{\zeta \to 1} \mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\} = \frac{K_1+K_2}{N_1+N_2}, \\ \lim\limits_{\zeta \to 0} \mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\} = \frac{K_2}{N_2}. \end{cases} \tag{2.45}$$

When $\zeta = 1$, $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ is biased. When $\zeta = 0$, $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ is the same as the true probability of local $\blacklozenge$ at the m^{th} time slot.

Eqs. (2.44) and (2.45) can be extended to several successive environmental changes as well. Assume that a local detector endures $L-1$ SNR changes in series, and the corresponding probabilities of local $\blacklozenge$ are $\frac{K_1}{N_1}$, $\frac{K_2}{N_2}$, ..., $\frac{K_L}{N_L}$, respectively. In this scenario, the true probability of local $\blacklozenge$ is $\frac{K_L}{N_L}$ at this time. The expectation of the estimated probability of $\blacklozenge$ for the i^{th} local detector at the m^{th} time slot, $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$,

becomes

$$\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\} = \frac{\frac{K_1}{N_1}\times(\zeta^{N_1+\cdots+N_L-1}+\cdots+\zeta^{N_2+\cdots+N_L})+\cdots+\frac{K_L}{N_L}\times(\zeta^{N_L-1}+\zeta^{N_L-2}+\cdots+1)}{1+\zeta+\zeta^2+\cdots+\zeta^{N_1+\cdots+N_L-1}} \quad (2.46)$$

and

$$\begin{cases} \lim\limits_{\zeta\to 1}\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\} = \frac{K_1+\cdots+K_L}{N_1+\cdots+N_L}, \\ \lim\limits_{\zeta\to 0}\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\} = \frac{K_L}{N_L}. \end{cases} \quad (2.47)$$

When $\zeta = 1$, $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ is biased. When $\zeta = 0$, $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ is the same as the true probability of local $\blacklozenge$ at the m^{th} time slot. □

Lemma 4: Assume that the environment of the i^{th} local detector is time-varying. Since $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ is a monotonic function with respect to ζ over $0 < \zeta \leq 1$, the probability estimator for $\blacklozenge$ with $0 < \zeta < 1$ is more reliable (i.e., leading to a more accurate probability estimate) than that with $\zeta = 1$ given by Eq. (2.22) on statistical average.

Proof: The proof can be considered in two cases. First, consider one environmental change at the i^{th} local detector such that the probability of local $\blacklozenge$ varies from $\frac{K_1}{N_1}$ to $\frac{K_2}{N_2}$. We will prove that $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ given by Eq. (2.44) is a monotonic function with respect to ζ over $0 < \zeta \leq 1$. Suppose $x \stackrel{\text{def}}{=} \zeta^{N_2}(1+\zeta+\cdots+\zeta^{N_1-1}) = \frac{\zeta^{N_2}-\zeta^{N_1+N_2}}{1-\zeta}$ and $y \stackrel{\text{def}}{=} 1+\zeta+\cdots+\zeta^{N_2} = \frac{1-\zeta^{N_2}}{1-\zeta}$. The derivatives of x and y with respect to ζ are

$$\frac{\partial x}{\partial \zeta} = \frac{N_2\zeta^{N_2-1}-(N_1+N_2)\zeta^{N_1+N_2-1}-(N_2-1)\zeta^{N_2}+(N_1+N_2-1)\zeta^{N_1+N_2}}{(1-\zeta)^2}, \quad (2.48)$$

$$\frac{\partial y}{\partial \zeta} = \frac{-N_2\zeta^{N_2-1}+(N_2-1)\zeta^{N_2}+1}{(1-\zeta)^2}. \quad (2.49)$$

Thus, the expectation of the estimated probability of local $\blacklozenge$ at the m^{th} time slot becomes

$$\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\} = \frac{\frac{K_1}{N_1}x+\frac{K_2}{N_2}y}{x+y}. \quad (2.50)$$

Taking the derivative of $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ in Eq. (2.50) with respect to ζ, we get

$$\frac{\partial\left[\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}\right]}{\partial\zeta} = \frac{\left(\frac{K_2}{N_2}-\frac{K_1}{N_1}\right)\left(x\frac{\partial y}{\partial\zeta}-\frac{\partial x}{\partial\zeta}y\right)}{(x+y)^2}. \tag{2.51}$$

Substitute the formulae of x, y, $\frac{\partial x}{\partial\zeta}$, and $\frac{\partial y}{\partial\zeta}$ into $x\frac{\partial y}{\partial\zeta}-\frac{\partial x}{\partial\zeta}y$ to obtain

$$\begin{aligned} & x\frac{\partial y}{\partial\zeta}-\frac{\partial x}{\partial\zeta}y \\ &= \frac{\zeta^{N_2-1}(\zeta-1)\left(N_2+N_1\zeta^{N_1+N_2}-(N_1+N_2)\zeta^{N_1}\right)}{(1-\zeta)^3}. \end{aligned} \tag{2.52}$$

Taking the derivative of $N_2+N_1\zeta^{N_1+N_2}-(N_1+N_2)\zeta^{N_1}$ with respect to ζ, we obtain

$$\begin{aligned} & \frac{\partial\left[N_2+N_1\zeta^{N_1+N_2}-(N_1+N_2)\zeta^{N_1}\right]}{\partial\zeta} \\ &= N_1(N_1+N_2)\zeta^{N_1+N_2-1}-N_1(N_1+N_2)\zeta^{N_1-1} \\ &\leq 0. \end{aligned} \tag{2.53}$$

Thus, $N_2+N_1\zeta^{N_1+N_2}-(N_1+N_2)\zeta^{N_1}$ is a monotonically decreasing function with respect to ζ over $0<\zeta\leq 1$. Besides,

$$\left[N_2+N_1\zeta^{N_1+N_2}-(N_1+N_2)\zeta^{N_1}\right]\Big|_{\zeta=1}=0. \tag{2.54}$$

According to Eqs. (2.53) and (2.54), $N_2+N_1\zeta^{N_1+N_2}-(N_1+N_2)\zeta^{N_1}\geq 0$ for $0<\zeta\leq 1$, which means $x\frac{\partial y}{\partial\zeta}-\frac{\partial x}{\partial\zeta}y\leq 0$ for $0<\zeta\leq 1$.

Now we consider Eq. (2.51). When $\frac{K_1}{N_1}<\frac{K_2}{N_2}$, it is obvious that $\frac{K_1+K_2}{N_1+N_2}<\frac{K_2}{N_2}$. Thus, according to Eq. (2.45), $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ becomes a decreasing function with respect to ζ over $0<\zeta\leq 1$, and it decreases from $\frac{K_2}{N_2}$ (for $\zeta=0$) to $\frac{K_1+K_2}{N_1+N_2}$ (for $\zeta=1$) monotonically. Similar results can be drawn for $\frac{K_1}{N_1}>\frac{K_2}{N_2}$. Therefore, the performance of the probability estimator with the discount factor $0<\zeta<1$ is better than the sample-average estimator given by Eq. (2.22) with $\zeta=1$.

Second, assume that the i^{th} local detector endures $L-1$ environmental changes, say $L>2$, with the corresponding probabilities of local $\blacklozenge$ as $\frac{K_1}{N_1}$, $\frac{K_2}{N_2}$, ..., $\frac{K_L}{N_L}$, respectively. In this scenario, it would be very complicated to study the monotonicity of

$\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ given by Eq. (2.46). However, we may further assume that prior to the $(L-1)^{\text{th}}$ (most recent) environmental change, the estimated probability of local $\blacklozenge$ has already converged to $\frac{K'}{N'}$ (not necessarily the true probability). Thus, Eqs. (2.46) and (2.47) can be rewritten as

$$\begin{aligned}&\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}\\ &= \frac{\frac{K'}{N'}\times(\zeta^{N'+N_L-1}+\cdots+\zeta^{N_L})+\frac{K_L}{N_L}\times(\zeta^{N_L-1}+\zeta^{N_L-2}+\cdots+1)}{1+\zeta+\zeta^2+\cdots+\zeta^{N'+N_L-1}}\end{aligned} \tag{2.55}$$

and

$$\begin{cases}\lim\limits_{\zeta\to 1}\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}=\frac{K'+K_L}{N'+N_L},\\ \lim\limits_{\zeta\to 0}\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}=\frac{K_L}{N_L}.\end{cases} \tag{2.56}$$

Consequently, following exactly the identical proof for the case considering only one environmental change, the monotonicity of Eq. (2.55) over $0<\zeta\leq 1$ can be justified accordingly. When $\frac{K'}{N'}<\frac{K_L}{N_L}$, it is obvious that $\frac{K'+K_L}{N'+N_L}<\frac{K_L}{N_L}$. Thus, according to Eq. (2.56), $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ becomes a decreasing function with respect to ζ over $0<\zeta\leq 1$, and it decreases from $\frac{K_L}{N_L}$ (for $\zeta=0$) to $\frac{K'+K_L}{N'+N_L}$ (for $\zeta=1$) monotonically. Similar results can be drawn for $\frac{K'}{N'}>\frac{K_L}{N_L}$. Therefore, the performance of the probability estimator with the discount factor $0<\zeta<1$ is better than the sample-average estimator given by Eq. (2.22) with $\zeta=1$. □

From all the aforementioned lemmas, the summary is provided as follows. When the optimal data fusion rule stated in Section 2.2.2 is used, one needs to know the exact probabilities of miss detection and false alarm at the moment, or K_2, N_2 as mentioned above. However, in practice, these probabilities are not known since no one knows when and how the local SNR changes. Therefore, we propose to use the probability estimators in conjunction with a discount factor ζ. Lemmas 1-4 facilitate the theoretical analysis that how the choice of ζ will influence the

probability estimation accuracies. When the environment of the i^{th} local detector is stationary, as $\zeta \to 1$, the probability estimate of local $\blacklozenge$ will get close to the true probability. When the environment is time-varying, on statistical average, the probability estimate of local $\blacklozenge$ will approach the true probability as $\zeta \to 0$, while that of local $\blacklozenge$ will be biased as $\zeta \to 1$. In other words, the smaller the discount factor ζ, the better trackability the spectrum-sensing system. Therefore, the appropriate choice of ζ should be related to the tradeoff between the estimation accuracy and the system trackability.

Lemma 5: When one tries to minimize the mean square error with respect to the discount factor ζ subject to the tradeoff between estimation accuracy and system trackability, a proper choice of ζ is within the interval $(0.99,\ 1)$.

Proof: From the proof of Lemma 4, $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ is a monotonically decreasing function of ζ, and it changes from the true probability of local $\blacklozenge$, $\frac{K_2}{N_2}$ (for $\zeta = 0$), to a biased value $\frac{K_1+K_2}{N_1+K_2}$ (for $\zeta = 1$). In summary, there occurs an interesting phenomenon: when $\zeta \to 0$, $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ is more accurate, but the probability estimates spread over a broader range; when $\zeta \to 1$, $\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\}$ is less accurate, but the probability estimates spread over a narrower range. Thus, we may investigate the mean square error (MSE) performance of this probability estimator with respect to ζ to determine the appropriate discount factor.

Consider a temporal environmental change at the i^{th} local detector such that the probability of local $\blacklozenge$ varies from $\frac{K_1}{N_1}$ to $\frac{K_2}{N_2}$. The MSE of $\hat{P}_{\blacklozenge_i}(m)$ subject to a discount factor ζ is given by

$$\mathcal{MSE}(\zeta) \stackrel{\text{def}}{=} \mathbb{E}\left\{\left(\hat{P}_{\blacklozenge_i}(m) - \frac{K_2}{N_2}\right)^2\right\}. \tag{2.57}$$

It is easy to derive from Eq. (2.44) that

$$\mathbb{E}\left\{\hat{P}^2_{\blacklozenge_i}(m)\right\} = \frac{\psi_3}{\left(\frac{1-\zeta^{N_1+N_2}}{1-\zeta}\right)^2}, \tag{2.58}$$

where $\psi_3 \stackrel{\text{def}}{=} \frac{K_1(N_1-K_1)}{N_1(N_1-1)}\zeta^{2N_2}\left(\frac{1-\zeta^{2N_1}}{1-\zeta^2}\right) + \frac{K_1(K_1-1)}{N_1(N_1-1)}\zeta^{2N_2}\left(\frac{1-\zeta^{N_1}}{1-\zeta}\right)^2 + \frac{K_2(N_2-K_2)}{N_2(N_2-1)}\left(\frac{1-\zeta^{2N_2}}{1-\zeta^2}\right) + \frac{K_2(K_2-1)}{N_2(N_2-1)}\left(\frac{1-\zeta^{N_2}}{1-\zeta}\right)^2 + \frac{2K_1K_2}{N_1N_2}\zeta^{N_2}\left(\frac{(1-\zeta^{N_1})(1-\zeta^{N_2})}{(1-\zeta)^2}\right)$, for $0 < \zeta < 1$. The variance of $\hat{P}_{\blacklozenge_i}(m)$ is given by

$$\mathbb{E}\left\{\sigma^2_{\blacklozenge_i}\right\} \stackrel{\text{def}}{=} \mathbb{E}\left\{\hat{P}^2_{\blacklozenge_i}(m)\right\} - \mathbb{E}^2\left\{\hat{P}_{\blacklozenge_i}(m)\right\} = \psi_4 \Big/ \left(\frac{1-\zeta^{N_1+N_2}}{1-\zeta}\right)^2, \tag{2.59}$$

where

$$\begin{aligned}\psi_4 \stackrel{\text{def}}{=} & \frac{K_1(N_1-K_1)}{N_1^2(N_1-1)}\zeta^{2N_2}\left[N_1\frac{1-\zeta^{2N_1}}{1-\zeta^2} - \left(\frac{1-\zeta^{N_1}}{1-\zeta}\right)^2\right] \\ & + \frac{K_2(N_2-K_2)}{N_2^2(N_2-1)}\left[N_2\frac{1-\zeta^{2N_2}}{1-\zeta^2} - \left(\frac{1-\zeta^{N_2}}{1-\zeta}\right)^2\right].\end{aligned} \tag{2.60}$$

Therefore, the MSE defined by Eq. (2.57) becomes $\mathcal{MSE}(\zeta) = \mathbb{E}\left\{\sigma^2_{\blacklozenge_i}\right\} + \left[\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\} - \frac{K_2}{N_2}\right]^2$, which is

$$\mathcal{MSE}(\zeta) = \frac{\psi_4(1-\zeta)^2}{\left(1-\zeta^{N_1+N_2}\right)^2} + \left[\left(\frac{K_1}{N_1} - \frac{K_2}{N_2}\right)\frac{\zeta^{N_2}(1-\zeta^{N_1})}{(1-\zeta^{N_1+N_2})}\right]^2, \tag{2.61}$$

where ψ_4 is defined in Eq. (2.60).

In order to analyze the monotonicity of this complex expression of MSE, we assume $N_1 = N_2 = N$. Denote $\chi \stackrel{\text{def}}{=} \left(\mathbb{E}\left\{\hat{P}_{\blacklozenge_i}(m)\right\} - \frac{K_2}{N_2}\right)^2$, and $\frac{\partial\chi}{\partial\zeta}$ is given by

$$\frac{\partial\chi}{\partial\zeta} = \left(\frac{K_1-K_2}{N}\right)^2 \frac{2N\zeta^{2N-1}}{(1+\zeta^N)^3}. \tag{2.62}$$

Thus, Eq. (2.59) can be reformulated as

$$\begin{aligned}\mathbb{E}\left\{\sigma^2_{\blacklozenge_i}\right\} &= \frac{K_1(N-K_1)\zeta^{2N} + K_2(N-K_2)}{N^2(N-1)} \\ &\quad \times \frac{N(1+\zeta^2+\cdots+\zeta^{2N-2}) - (1+\zeta+\cdots+\zeta^N-1)^2}{(1+\zeta+\cdots+\zeta^{2N-1})^2}.\end{aligned} \tag{2.63}$$

Therefore, $\left.\frac{\partial\mathbb{E}\left\{\sigma^2_{\blacklozenge_i}\right\}}{\partial\zeta}\right|_{\zeta=0} = \frac{-2K_2(N-K_2)}{N(N-1)}$, and $\left.\frac{\partial\mathbb{E}\left\{\sigma^2_{\blacklozenge_i}\right\}}{\partial\zeta}\right|_{\zeta=1} = 0$. Besides, $\left.\frac{\partial\chi}{\partial\zeta}\right|_{\zeta=0} = 0$, and $\left.\frac{\partial\chi}{\partial\zeta}\right|_{\zeta=1} = \frac{(K_1-K_2)^2}{8N}$. Furthermore, N is a large integer, which means $\frac{\partial\chi}{\partial\zeta}$ will increase very abruptly when ζ gets very close to 1. Since both $\frac{\partial\chi}{\partial\zeta}$ and $\frac{\partial\mathbb{E}\left\{\sigma^2_{\blacklozenge_i}\right\}}{\partial\zeta}$ are continuous

functions over $0 < \zeta < 1$, there exists ζ, which is very close to 1, yielding the minimum MSE. In other words, $\mathcal{MSE}(\zeta)$ is a "bowl-shape" function over $0 < \zeta < 1$. A typical example is illustrated by Figure 2.7. Figure 2.7 (a) exhibits the MSE curves versus ζ for $N_1 = 1000$, $N_2 = 1000$, and different K_1, K_2 values. Figure 2.7 (b) is a zoom-in version of Figure 2.7 (a) around $\zeta \to 1$. From Figure 2.7, one may observe that $\mathcal{MSE}(\zeta)$ is a "bowl-shape" function over $0 < \zeta < 1$. When ζ is small, the mean square error drops down as ζ increases. When $\zeta \to 1$ and $\left|\frac{K_1}{N_1} - \frac{K_2}{N_2}\right| \neq 0$, $\mathcal{MSE}(\zeta)$ abruptly rises at a discount factor very close to $\zeta = 1$. This turning point appears closer to 1 when the true probability change $\left|\frac{K_1}{N_1} - \frac{K_2}{N_2}\right|$ becomes smaller. On the other hand, it can also be found that when $\left|\frac{K_1}{N_1} - \frac{K_2}{N_2}\right|$ is fixed and N_1, N_2 become larger, this turning point will appear closer to 1. Obviously, the discount factor $\zeta = 1$ is not a good choice in the minimum MSE sense. Of course, one can undertake an exhaustive search within a small interval around $\zeta = 1$ to find the optimal choice of ζ. However, the optimal discount factor depends on N_1, N_2, K_1, and K_2 but they are not available in practice. Empirically speaking, to approximately guarantee $\mathcal{MSE}(\zeta) \leq \frac{\mathcal{MSE}(0)}{10}$, ζ should be selected from the interval $(0.99, 1)$. □

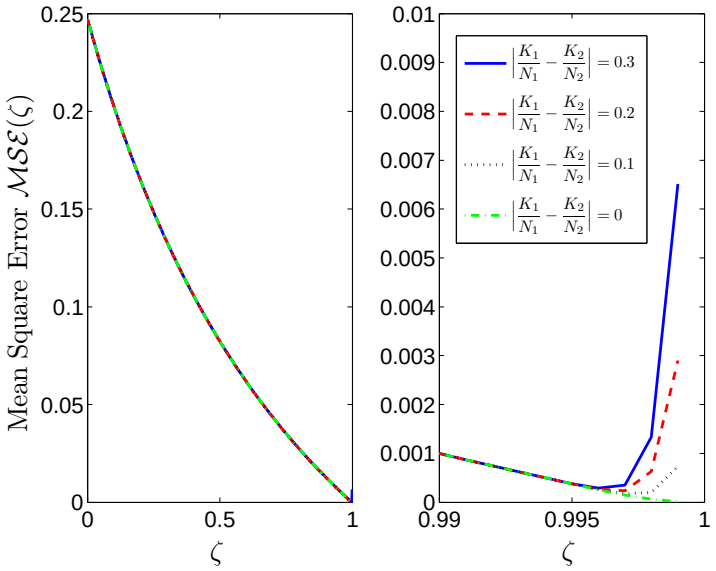

Figure 2.7. Mean square error with respect to ζ.

In order to justify the validity of the aforementioned MSE analysis, another typical example is illustrated by Figure 2.8. In Figure 2.8, we compare the simulated MSEs of the estimated probabilities of local miss detection resulting from Monte Carlo experiments with the theoretical MSEs by use of different temporal discount factors ζ. Suppose that the SNR at a certain local detector changes from -25 dB to -30 dB after 1000 sensing intervals ($N_1 = 1000$), and the probability of local miss detection is estimated after another 1000 sensing intervals ($N_2 = 1000$). Since the true values of K_1 and K_2 are unavailable in practice, we use the statistical mean values of K_1 and K_2 when the local SNR is -25dB and -30dB, respectively. We carry out one hundred Monte Carlo trials to calculate the the average simulated MSEs. It is obvious that the MSEs we obtain from the simulation results are very close to the theoretical MSEs according to Figure 2.8.

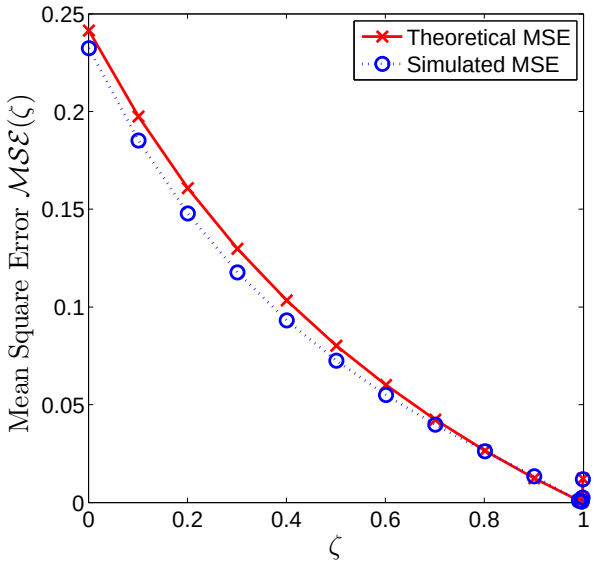

Figure 2.8. Comparison between the simulated and theoretical mean square errors.

2.2.5 Simulation Results

In the stationary environment, the *average risk* $\mathcal{R}$ is well known as a standard measure to compare the performances of signal detectors in the classical Bayesian hy-

pothesis theory. It is given by

$$\begin{aligned}\mathcal{R} \;\stackrel{\text{def}}{=}\; & P(H_0)\left[P(u_0=-1|H_0)\,C_{00}+P(u_0=+1|H_0)\,C_{10}\right] \\ & +P(H_1)\left[P(u_0=+1|H_1)\,C_{11}+P(u_0=-1|H_1)\,C_{01}\right]. \end{aligned} \tag{2.64}$$

Besides, we assume $C_{00}=C_{11}=0$ and $C_{10}=C_{01}=1$. Thus, Eq. (2.64) becomes

$$\begin{aligned}\mathcal{R} = & \; P(H_0)\,P(u_0=+1|H_0)+P(H_1)\,P(u_0=-1|H_1) \\ = & \; P_0\,P_F+P_1\,P_M, \end{aligned} \tag{2.65}$$

where P_F and P_M denote the *global false alarm probability* and the *global miss-detection probability*, respectively.

In the existing literature, the most commonly-used data fusion rules are the "OR" and "AND" rules [44]. We depict the average risk $\mathcal{R}$ (given by Eq. (2.65)) with respect to the global false alarm probability P_F in Figure 2.9 for three different mechanisms based on the "OR", "AND", and optimal data-fusion rules for the SNR value of -20 dB. According to Figure 2.9, our proposed cooperative spectrum-sensing method based on the optimal data-fusion rule leads to the superior performance compared to the two other schemes across all P_F conditions. Similar trends can also be observed for other SNR values.

In Section 2.2.3, we have introduced a new adaptive cooperative spectrum-sensing algorithm for the time-varying environments. In order to illustrate the influence of the proposed temporal discount factor, we compare the estimated local detection performances by use of different temporal discount factors (ζ= 0.9985, 0.999, 0.9995, and 1). The time-varying environment is established as follows. Suppose that the SNR value at a certain local detector suddenly changes from -25 dB to -30 dB at a certain time point (say, at iteration number 2,000), the estimated correct detection probabilities of this local detector by use of different temporal discount factors ζ are plotted in Figure 2.10. The new probability of detection for $SNR=-30$ is

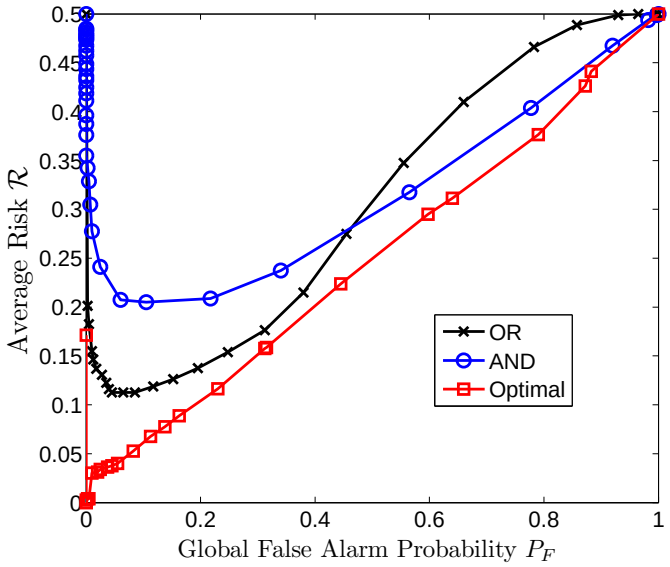

Figure 2.9. The average risks $\mathcal{R}$ versus the global false alarm probabilities P_F for the cooperative spectrum-sensing methods using the "OR", "AND", and optimal data-fusion rules.

approximately 59%. One can observe that the local detector with a temporal discount factor $0 < \zeta < 1$ will converge to the new probability of detection eventually, while the one with $\zeta = 1$ will converge to a biased value, as stated by Lemma 3. It is obvious that the local detector with a temporal discount factor $0 < \zeta < 1$ reacts more quickly to the environmental changes than the detector simply using the sample-average estimators (or $\zeta = 1$), and the estimates of the former detector converge to the new stable correct detection probability faster than the estimates of the latter scheme. Besides, the smaller this temporal discount factor, the better the corresponding trackability. In addition, the scheme using a fixed-length window is also simulated for comparison. The window lengths used in Figure 2.11 are $\gamma =$ 400, 600, 800, and 1000 samples. According to Figure 2.11, the fluctuations (misadjustments) of the estimated probabilities are very conspicuous even if large window sizes are used. The choice of an appropriate window length highly depends on how often the local SNR changes but in reality, no one can predict when the local SNR

changes. Hence the fixed-length window scheme is not robust at all in practice. Fur-

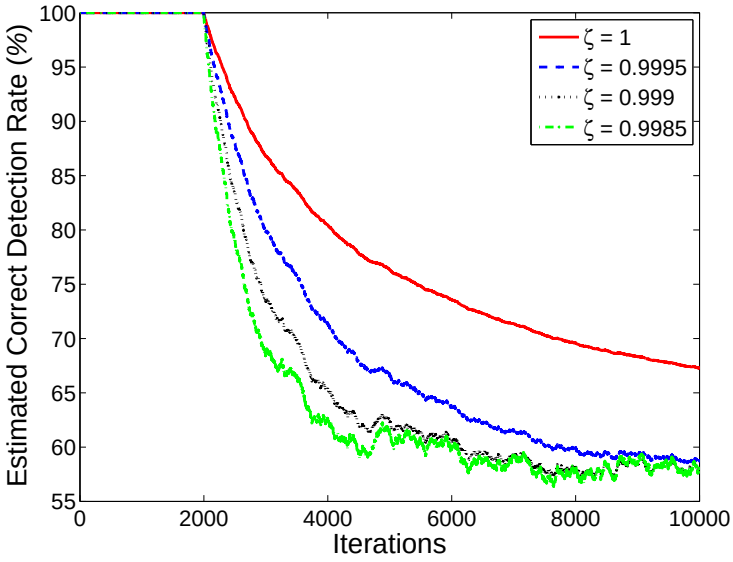

Figure 2.10. The convergence trends of the estimated (local) correct detection probabilities using different discount factors for an arbitrary sensing node.

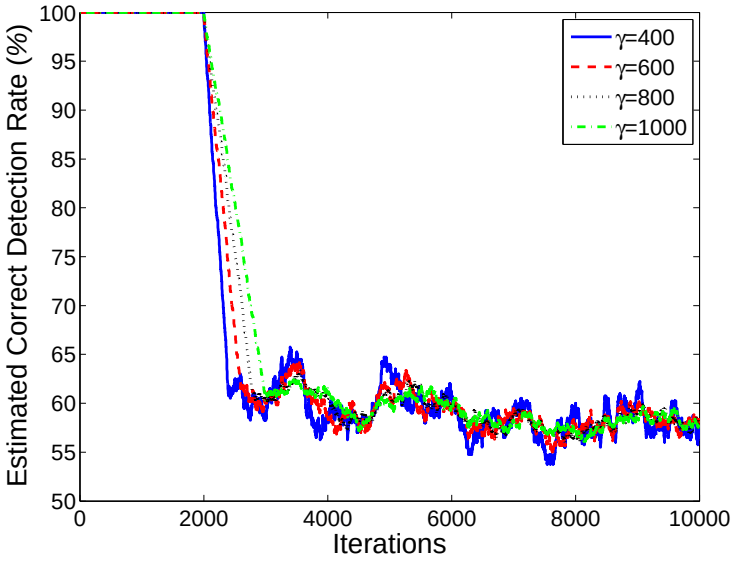

Figure 2.11. The convergence trends of the estimated (local) correct detection probabilities using different window sizes for an arbitrary sensing node.

thermore, we compare the cooperative spectrum-sensing performances in terms of *receiver operating characteristics* (ROC) curves subject to different discount factors,

namely $\zeta = 0.991$ and $\zeta = 1$. Three local detectors are used. Suppose that the source is at the origin point. The coordinates of the three local detectors are randomly generated by the computer at (10, 10), (6, 8) and (6, 9), respectively. We assume that the signal energies received at the local detectors, E_i $(i = 1, 2, 3)$, are inversely proportional to the squared distance, d_i^2. In other words, $\frac{E_i}{E_j} = \frac{d_j^2}{d_i^2}$ $(i, j = 1, 2, 3)$. Here we arbitrarily choose $N_1 = N_2 = 1000$, and the SNR changes (denoted by $\Delta\mathcal{SNR}_i^{(m)} \stackrel{\text{def}}{=} \mathcal{SNR}_i^{(m)} - \mathcal{SNR}_i^{(m-1)}$) across all local detectors are $\Delta\mathcal{SNR}_i^{(m)}$=−10 dB and $\Delta\mathcal{SNR}_i^{(m)}$=−12 dB, $\forall i$, where the original average SNR are −20 dB, −17.16 dB, and −17.67 dB, respectively. The ROC curves are delineated in Figure 2.12. It is obvious that the cooperative spectrum-sensing scheme using the discount factor $\zeta = 0.991$ outperforms that using the sample-average estimators or $\zeta = 1$ when the false alarm probability is not large. Finally, in Figure 2.13, we compare the coop-

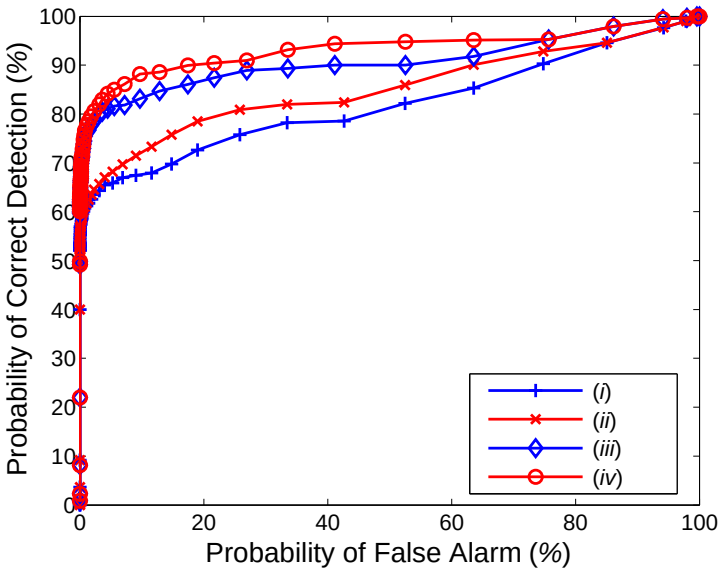

Figure 2.12. The ROC curves of the cooperative spectrum-sensing performances for (i) an SNR decrement by 12 dB ($\Delta\mathcal{SNR}_i^{(m)}$=−12 dB, $i = 1, 2, 3$) and $\zeta = 1$, (ii) an SNR decrement by 12 dB ($\Delta\mathcal{SNR}_i^{(m)}$=−12 dB, $i = 1, 2, 3$) and $\zeta = 0.991$, (iii) an SNR decrement by 10 dB ($\Delta\mathcal{SNR}_i^{(m)}$=−10 dB, $i = 1, 2, 3$) and $\zeta = 1$, and (iv) an SNR decrement by 10 dB ($\Delta\mathcal{SNR}_i^{(m)}$=−10 dB, $i = 1, 2, 3$) and $\zeta = 0.991$.

erative spectrum-sensing performances in terms of *receiver operating characteristics* (ROC) curves subject to discount factors, fixed-length windows, and the exact knowledge of K_1, N_1, K_2, N_2. Note that the best ROC performance can be achieved when the exact knowledge of K_1, N_1, K_2, N_2 is available. The simulation setup is the same as above with $\Delta\mathcal{SNR}_i^{(m)}$=−12 dB, $\forall i$. It is obvious that the ROC curves subject to different fixed-window sizes greatly overlap with each other, and they fall below the ROC curve subject to a discount factor $\zeta = 0.991$ (it means that the latter ROC performance is better). If we know the exact values of K_1, N_1, K_2, N_2, the actual optimal data-fusion rule can thus be undertaken, which yields a slightly better ROC performance than that subject to the discount factor $\zeta = 0.991$.

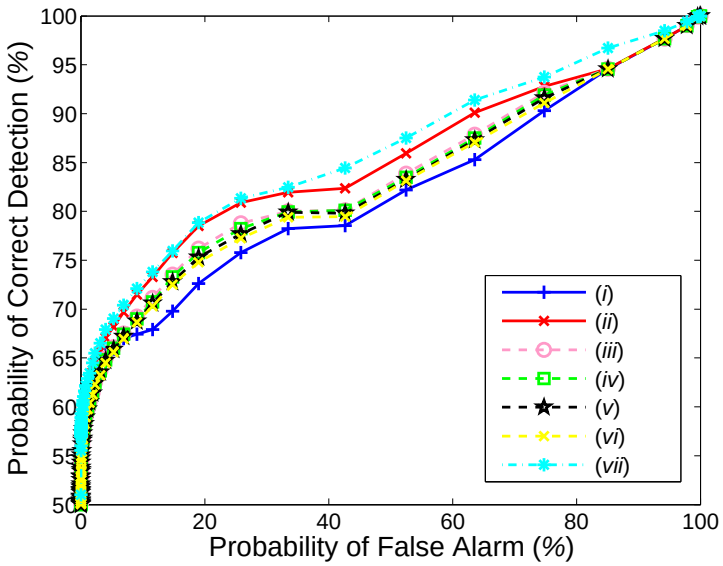

Figure 2.13. The ROC curves of the cooperative spectrum-sensing performances for an SNR decrement by 12 dB ($\Delta\mathcal{SNR}_i^{(m)}$=−12 dB, $i = 1, 2, 3$) subject to (i) a discount factor $\zeta = 1$, (ii) a discount factor $\zeta = 0.991$, (iii) a fixed-window length $\gamma = 400$, (iv) a fixed-window length $\gamma = 600$, (v) a fixed-window length $\gamma = 800$, (vi) a fixed-window length $\gamma = 1000$, and (vii) the complete knowledge of K_1, K_2, N_1, N_2.

2.2.6 Conclusion

In this chapter, a novel adaptive cooperative spectrum-sensing technique was proposed based on JB-statistics and the optimal data-fusion rule. By adopting a proper temporal discount factor, this new cooperative spectrum-sensing scheme can also adapt to time-varying environments effectively. The advantage of the new discount factor based probabilistic estimators is also theoretically investigated and the optimal discount factor value is facilitated. According to Monte Carlo simulation results for wireless microphone signals, our JB-statistic-based detection method is more robust than the commonly-used energy-based spectrum-sensing scheme over a broad variety of SNR conditions. Besides, our proposed new cooperative spectrum-sensing scheme can achieve a much lower average risk than other existing spectrum-sensing methods using "OR" and "AND" data fusion rules. In addition, this new cooperative spectrum-sensing scheme can greatly outperform the conventional cooperative spectrum-sensing method using sample-average estimators when any local detector suffers from an abrupt signal-to-noise ratio change. Therefore, this new cooperative spectrum-sensing mechanism would be a very promising solution to the future cognitive radio technology.

3 Symbol Detection and Channel Estimation for OFDM Systems

3.1 Problem Definition

The baseband OFDM transmitter (for generating digital OFDM modulated signal vectors) involving the proposed novel frequency-domain optimal pilot multiplexing scheme is shown by the block diagram in Figure 3.1. Assume that there are N subcarriers, N_p pilots, and $N = M \times N_p$, where $M \in \mathcal{Z}^+$, is the spacing between two adjacent pilots. According to Figure 3.1, the information bits are first mapped onto a normalized q-QAM (quadrature amplitude modulation) signal constellation Θ, where $\Theta \stackrel{\text{def}}{=} \{\alpha_1, \alpha_2, \ldots, \alpha_q\}$, and then multiplexed onto N orthogonal subcarriers to constitute an (unfiltered) OFDM symbol vector $\vec{\mathcal{S}}^{\text{u}} = [S_0^{\text{u}}, S_1^{\text{u}}, \ldots, S_{N-1}^{\text{u}}]^T$, where $\mathbb{E}\{|S_k^{\text{u}}|^2\} = \sigma_s^2 = 1$, for $k = 0, 1, \ldots, N-1$ (k is the subcarrier index). An $N \times N$ unitary Walsh Hadamard matrix $\widetilde{W}$ is then used to pre-code the signal vector $\vec{\mathcal{S}}^{\text{u}}$, resulting in a "filtered" signal vector $\vec{\mathcal{S}} = [S_0, S_1, \ldots, S_{N-1}]^T$. It is

$$\vec{\mathcal{S}} = \widetilde{W}\, \vec{\mathcal{S}}^{\text{u}}, \tag{3.1}$$

where

$$\widetilde{W}^H \widetilde{W} = \widetilde{I}_N. \tag{3.2}$$

According to [26], when the conventional frequency-domain pilot multiplexing technique (FDPMT) is used, pilots are inserted into the signal vector $\vec{\mathcal{S}}$ in the equal-spaced manner after the original subcarrier data at these positions are removed.

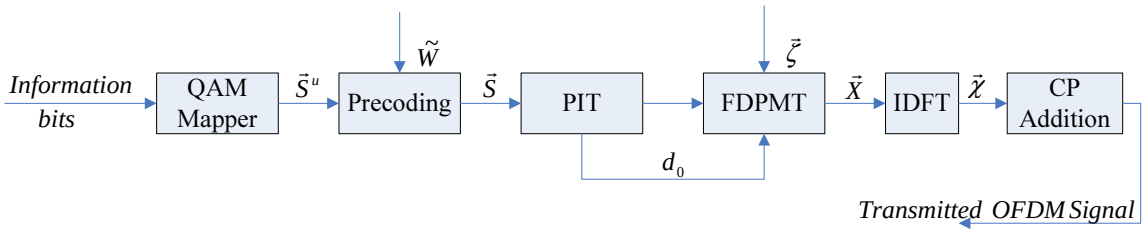

Figure 3.1. The new OFDM baseband transmitter using our proposed optimal dynamical pilot insertion technique.

Therefore, the actually transmitted OFDM signal in the frequency domain, $\vec{\mathbf{X}} \stackrel{\text{def}}{=} [X_0, X_1, \ldots, X_{N-1}]^T$, is given by

$$\vec{\mathbf{X}} = [C_0, S_1, \ldots, S_{M-1}, C_1, S_{M+1}, \ldots, S_{N-1}]^T, \tag{3.3}$$

or

$$X_k = \begin{cases} C_{\lfloor k/M \rfloor}, & \text{if } k \pmod M = 0 \\ S_k, & \text{otherwise} \end{cases} \tag{3.4}$$

Denote $\vec{\zeta} \stackrel{\text{def}}{=} [C_0, C_1, \ldots, C_{N_p-1}]^T$, which is composed by the *Chu sequence* of length N_p (refer to [45]). Its elements are used as the pilots such that

$$C_t = \begin{cases} \exp\left(j\pi\varsigma t^2/N_p\right), & \text{if } N_p \text{ is even} \\ \exp\left(j\pi\varsigma t(t+1)/N_p\right), & \text{if } N_p \text{ is odd} \end{cases} \tag{3.5}$$

where $t = 0, 1, \ldots, N_p - 1$, and ς is relatively prime to N_p. An N-point inverse discrete Fourier transform (IDFT) is then operated on $\vec{\mathbf{X}}$, resulting in the time-domain transmitted signal $\vec{\chi} \stackrel{\text{def}}{=} [x_0, x_1, \ldots, x_{N-1}]^T$ where

$$x_i = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k \, e^{j\frac{2\pi ik}{N}}, \tag{3.6}$$

for $i = 0, 1, \ldots, N-1$.

It is obvious that the OFDM subcarrier symbols S_0, S_M, …, $S_{(N_P-1)M}$ are eliminated from the signal vector $\vec{\mathcal{S}}$ for the further insertion of training pilots, which will definitely cause distortion to the time-domain transmitted signal.

3.2 New OFDM Transmitter

In order to minimize the aforementioned signal distortion caused by the FDPMT, we propose a novel robust pilot insertion technique here. First, we define

$$\begin{aligned}
\vec{\mathbf{X}}^m &\stackrel{\text{def}}{=} [X_0^m, X_1^m, \ldots, X_{N-1}^m]^T, && (3.7)\\
\vec{\chi}^m &\stackrel{\text{def}}{=} [x_0^{\mathrm{m}}, x_1^{\mathrm{m}}, \ldots, x_{N-1}^{\mathrm{m}}]^T, && (3.8)\\
\vec{\Phi}_m &\stackrel{\text{def}}{=} [S_m, S_{M+m}, S_{2M+m}, \ldots, S_{(N_P-1)M+m}]^T, && (3.9)\\
\vec{\Psi}_m &\stackrel{\text{def}}{=} [m, M+m, 2M+m, \ldots, (N_P-1)M+m]^T, && (3.10)
\end{aligned}$$

where

$$X_k^m \stackrel{\text{def}}{=} \begin{cases} S_k, & \text{if } k \pmod M \neq m \\ 0, & \text{if } k \pmod M = m \end{cases} \tag{3.11}$$

for $k = 0, 1, \ldots, N-1$, and

$$x_i^m \stackrel{\text{def}}{=} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k^m \, e^{j\frac{2\pi ik}{N}}, \tag{3.12}$$

for $i = 0, 1, \ldots, N-1$. Note that $m = 0, 1, \ldots, M-1$. Next, we define the distortion due to the subcarrier-removal at the positions indexed by the vector $\vec{\Psi}_m$, which is

$$\begin{aligned}
\Xi_m &\stackrel{\text{def}}{=} \sum_{i=0}^{N-1} |x_i - x_i^m|^2 \\
&= \sum_{i=0}^{N-1} \left| \frac{1}{\sqrt{N}} \sum_{k \in \vec{\Psi}_m} S_k e^{j\frac{2\pi ik}{N}} \right|^2 \\
&= \sum_{k \in \vec{\Psi}_m} |S_k|^2 \\
&\quad + \frac{1}{N} \sum_{i=0}^{N-1} \sum_{k \in \vec{\Psi}_m} \sum_{l \in \vec{\Psi}_m, l \neq k} S_k S_l^* e^{j\frac{2\pi i(k-l)}{N}},
\end{aligned} \tag{3.13}$$

where S_l^* is the complex conjugate of S_l. Note that

$$\begin{aligned}
&\frac{1}{N}\sum_{i=0}^{N-1}\sum_{k\in\vec{\Psi}_m}\sum_{l\in\vec{\Psi}_m, l\neq k} S_k S_l^* e^{j\frac{2\pi i(k-l)}{N}} \\
&= \frac{1}{N}\sum_{k\in\vec{\Psi}_m}\sum_{l\in\vec{\Psi}_m, l\neq k} S_k S_l^* \sum_{i=0}^{N-1} e^{j\frac{2\pi i(k-l)}{N}} \\
&= 0. \qquad (3.14)
\end{aligned}$$

According to Eqs. (3.13) and (3.14), we have $\Xi_m = \sum_{k\in\vec{\Psi}_m} |S_k|^2$. Therefore, the minimization of Ξ_m leads to

$$\begin{aligned}
d_0 &= \arg\min_m \ \Xi_m \\
&= \arg\min_m \sum_{k\in\vec{\Psi}_m} |S_k|^2 \\
&= \arg\min_m \ \vec{\Phi}_m^H \vec{\Phi}_m, \qquad (3.15)
\end{aligned}$$

where d_0 is called the *optimal pilot offset*. After d_0 is picked to minimize Ξ_m according to Eq. (3.15), one can construct the *pilot-position vector* $\vec{\Psi}_{d_0}$ and the *frequency-domain pilot-assisted OFDM symbol vector* $\vec{\mathbf{X}}$. They are given by

$$\vec{\Psi}_{d_0} \stackrel{\text{def}}{=} \Big[d_0, M+d_0, \ldots, (N_P-1)M+d_0\Big]^T, \qquad (3.16)$$

and

$$\vec{\mathbf{X}} = [X_0, X_1, \ldots, X_{N-1}]^T, \qquad (3.17)$$

where

$$X_k = \begin{cases} C_{\lfloor k/M \rfloor}, & \text{if } \ k \ (\text{mod } M) = d_0, \\ S_k, & \text{otherwise.} \end{cases} \qquad (3.18)$$

Note that $\vec{\mathbf{X}}$ given by Eq. (3.3) is a special case of Eq. (3.17) for $d_0 = 0$.

After an N-point IDFT is invoked for the signal vector $\vec{\mathbf{X}}$ given by Eq. (3.17), the resultant time-domain signal vector $\vec{\chi}$ will be further appended by a cyclic prefix (CP) of length N_{cp}. Thus the digital OFDM signal sequence in the baseband

can be constructed. The subsequent standard operations including lowpass filtering, digital-to-analog conversion, and frequency-up conversion in OFDM systems will be neglected here since they stay the same as usual. Assume that the medium is a multipath channel with the maximum delay spread up to L sampling (symbol) periods. We further assume $N_p \geq N_{cp} \geq L$ so that the interblock interference is absent [46].

It is easy to discover that if d_0 is always set to be 0 and this information is known to the receiver, the pilot insertion/detection scheme would be simplified to a special case, as suggested by [26], when no optimization is undertaken. When our proposed PIT highlighted by Eqs. (3.15)-(3.18) is considered, new blind pilot-detection methods and the corresponding new receiver structure need to be designed. Details are referred to Section 3.3.

3.3 New OFDM Receiver

The block diagram for the baseband OFDM receiver (dealing with digital signals) is illustrated in Figure 3.2 where our proposed blind pilot-detection mechanism is incorporated. Generally speaking, this receiver structure is composed by two major components: one is the channel estimator assisted by blind PDT (details will be discussed in Section 3.3.1); the other is the channel equalization incorporated with the iterative symbol reconstruction (details are referred to Section 3.3.2).

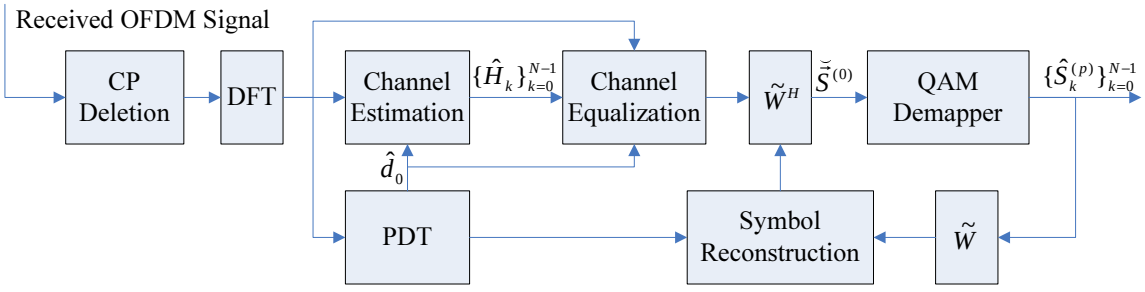

Figure 3.2. The new OFDM receiver using blind pilot detection techniques (corresponding to the transmitter given by Figure 3.1).

3.3.1 Channel Estimation aided by Different Blind Pilot Detection Techniques

At the receiver, the CP is first removed and then an N-point DFT is taken to reconstruct the received signal samples R_k in the frequency domain such that

$$R_k = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} r_i \, e^{-j\frac{2\pi ik}{N}}, \qquad k = 0, 1, \ldots, N-1, \tag{3.19}$$

where $\vec{\gamma} \stackrel{\text{def}}{=} [r_0, r_1, \ldots, r_{N-1}]^T$ is the received signal vector after the CP is removed, and the frequency-domain received signal vector is represented by $\vec{\Upsilon} \stackrel{\text{def}}{=} [R_0, R_1, \ldots, R_{N-1}]^T$. Since $N_{cp} \geq L$, it is obvious that

$$R_k = \begin{cases} H_k \, C_t + \Omega_k, & \text{if } \; k \; (\text{mod } M) = d_0 \\ H_k \, S_k + \Omega_k, & \text{otherwise} \end{cases} \tag{3.20}$$

where $t \stackrel{\text{def}}{=} \frac{k-d_0}{M}$, H_k is the k^{th} element of the channel frequency response (CFR) sequence vector, which is given by $\vec{\mathcal{H}} \stackrel{\text{def}}{=} [H_0, H_1, \ldots, H_{N-1}]^T \in \mathcal{C}^{N\times 1}$, and Ω_k is the k^{th} element of the complex additive white Gaussian noise (AWGN) sequence vector $\vec{\mathbf{\Omega}} \stackrel{\text{def}}{=} [\Omega_0, \Omega_1, \ldots, \Omega_{N-1}]^T \in \mathcal{C}^{N\times 1}$, each of whose entries has zero mean and variance σ_Ω^2. Consequently, the signal-to-noise ratio (SNR) is given by

$$\xi = 10 \; \log_{10} \left(\frac{\sigma_s^2}{\sigma_\Omega^2} \right). \tag{3.21}$$

We assume that both M and $\vec{\zeta}$ are fixed and known to the receiver. However, the pilot-position offset d_0 needs to be blindly located. In order to combat this problem, we first define a matrix $\widetilde{\mathbf{\Gamma}}$ as

$$\widetilde{\mathbf{\Gamma}} \stackrel{\text{def}}{=} \left[\vec{\mathbf{\Lambda}}_0, \vec{\mathbf{\Lambda}}_1, \cdots, \vec{\mathbf{\Lambda}}_m, \ldots, \vec{\mathbf{\Lambda}}_{M-1} \right] \in \mathcal{C}^{N_p \times M}, \tag{3.22}$$

where

$$\vec{\mathbf{\Lambda}}_m \stackrel{\text{def}}{=} \left[\frac{R_m}{C_0}, \cdots, \frac{R_{tM+m}}{C_t}, \cdots, \frac{R_{(N_p-1)M+m}}{C_{N_p-1}} \right]^T \in \mathcal{C}^{N_p \times 1}, \tag{3.23}$$

for $m = 0, 1, \ldots, M-1$ and $t = 0, 1, \ldots, N_p - 1$. Then, an N_p-point IDFT is invoked for each $\vec{\mathbf{\Lambda}}_m$, $m = 0, 1, \ldots, M-1$, such that

$$\lambda_{m,n} = \frac{1}{\sqrt{N_p}} \sum_{t=0}^{N_p-1} \Lambda_{m,t}\, e^{j\frac{2\pi nt}{N_p}}, \tag{3.24}$$

where $\Lambda_{m,t}$ is the t^{th} element of $\vec{\mathbf{\Lambda}}_m$, for $t = 0, 1, \ldots, N_p - 1$.

It is not hard to prove that when $m = d_0$ and $N_p > L$, $\lambda_{d_0,n} \approx \sqrt{N_p/N}\, h_n\, e^{-j\frac{2\pi n d_0}{N}}$ for high SNR, where h_n is the n^{th} coefficient of the channel impulse response (CIR).

On the other hand, when $m \neq d_0$, $\vec{\mathbf{\Lambda}}_m$ is just the sequence vector $\{H_k\, S_k\}_{k=0}^{N-1}$ divided by the Chu sequence vector $\vec{\zeta}$ given by Eq. (3.5) and then plus the AWGN, element-by-element. We assume that $\mathbb{E}[|H_k|^2] = 1$, $\forall k$, and the entries in $\vec{\mathcal{S}}^{\text{u}}$ are independently and identically distributed (i.i.d.). Therefore, as the inter-pilot spacing M increases, $\lambda_{m,n}$ approaches a zero-mean complex Gaussian process with variance $\sigma_s^2 + \sigma_\Omega^2$ according to the sampling theorem and the central limit theorem.

Note that we have to blindly estimate the actual pilot offset $m = d_0$ used by the transmitter. Consequently, we need to classify d_0 from $\{\lambda_{0,n}\}_{n=0}^{N_p-1}, \{\lambda_{1,n}\}_{n=0}^{N_p-1}, \ldots,$ $\{\lambda_{M-1,n}\}_{n=0}^{N_p-1}$ based on the fact that

$$\lambda_{m,n} \approx \begin{cases} \sqrt{\frac{N_p}{N}}\, h_n\, e^{-j\frac{2\pi n d_0}{N}} + \mathbb{N}(0, \sigma_\Omega^2), & \text{if } \ m = d_0 \\ \mathbb{N}(0,\ \sigma_\Omega^2 + \sigma_s^2), & \text{otherwise} \end{cases} \tag{3.25}$$

According to Eq. (3.25), we propose three different pilot-detection techniques to estimate the pilot offset d_0. The effectiveness evaluation and more insightful investigation will be presented in Section 3.3.3.

Variance Based Pilot Detection

According to Eq. (3.25), when $m \neq d_0$, $\{\lambda_{m,n}\}_{n=0}^{N_p-1}$ is approximately a complex Gaussian process with a larger variance $\sigma_\Omega^2 + \sigma_s^2$ than the case for $m = d_0$ (the variance is σ_Ω^2 instead). Therefore, we propose to use the amplitude variances of

$\{\lambda_{m,n}\}_{n=0}^{N_p-1}$ to estimate the pilot offset d_0. It yields

$$\hat{d}_0 = \underset{m\in\{0,1,\ldots,M-1\}}{\arg\max} \frac{1}{N_p}\sum_{n=0}^{N_p-1}\left(|\lambda_{m,n}| - \bar{\lambda}_m\right)^2, \tag{3.26}$$

where $\bar{\lambda}_m$ is the mean of the sequence $\{|\lambda_{m,n}|\}_{n=0}^{N_p-1}$ such that

$$\bar{\lambda}_m \overset{\text{def}}{=} \frac{1}{N_p}\sum_{n=0}^{N_p-1}|\lambda_{m,n}|. \tag{3.27}$$

Subspace Based Pilot Detection

Intuitively speaking, according to Eq. (3.25), the samples $\lambda_{m,n}$ involving the channel gain h_n form the signal subspace while other samples $\lambda_{m,n}$ span the noise subspace. In practice, however, the time-support for h_n is unknown at the receiver (a blind channel-filter length estimation method can be found in [47], though). We may simply define the space spanned by the last $N_p - N_{cp}$ samples of $\{\lambda_{d_0,n}\}_{n=0}^{N_p-1}$ as the noise subspace. Thus, the pilot offset d_0 can be blindly spotted as

$$\hat{d}_0 = \underset{m\in\{0,1,\ldots,M-1\}}{\arg\min} \sum_{n=N_{cp}}^{N_p-1}|\lambda_{m,n}|^2. \tag{3.28}$$

Jarqur-Bera Statistics Based Pilot Detection

By carefully observing the two processes given by Eq. (3.25), when $m = d_0$, since $N_p > L$, $\{|\lambda_{d_0,n}|\}_{n\in\{n'|h_{n'}\neq 0\}}$ constitutes a Gaussian-mixture process with L cluster means $\sqrt{N_p/N}\,h_n\,e^{-j\frac{2\pi n d_0}{N}}$, while $\{|\lambda_{m,n}|\}_{n\in\{n'|h_{n'}=0\}}$ constitutes a zero-mean complex Gaussian process with variance σ_Ω^2, for $n' = 0, 1, \ldots, N_p-1$. When $m \neq d_0$, as stated by Eq. (3.25), the entire series $\{\lambda_{m,n}\}_{n=0}^{N_p-1}$ is approximately a complex Gaussian process with variance $\sigma_\Omega^2+\sigma_s^2$. Therefore, we can classify the M sequences, $\{\lambda_{m,n}\}_{n=0}^{N_p-1}$ associated with different pilot offsets m, for $m = 0, 1, \ldots, M-1$, into these two classes (*channel fading coefficients bearing series* and *noise-only series*) by utilizing their distinguishable statistical characteristics.

In the literature [18, 48], the *Jarqur-Bera (JB) statistic test* is a goodness-of-fit measure of departure from normality, based on the *sample kurtosis* and the *sample skewness*. Since the larger the sample size, the more reliable this test will be, we propose to use both real and imaginary parts of the data samples independently to double the data amount and then carry out the JB-statistic test as follows:

$$\hat{d}_0 = \underset{m\in\{0,1,\ldots,M-1\}}{\arg\max} \ \mathfrak{JB}\left\{\{\lambda_{m,n}\}_{n=0}^{N_p-1}\right\}, \tag{3.29}$$

where $\mathfrak{JB}\{\ \}$ is the JB-statistic operator on a complex sample sequence inside the braces. More specifically, we denote $\vec{\eta}_m \overset{\text{def}}{=} \Big[\mathfrak{Re}\{\lambda_{m,0}\},\ldots,\mathfrak{Re}\{\lambda_{m,N_p-1}\},\mathfrak{Im}\{\lambda_{m,0}\},\ldots,\mathfrak{Im}\{\lambda_{m,N_p-1}\}\Big]^T \in \mathcal{R}^{2N_p\times 1}$, and the n^{th} element of $\vec{\eta}_m$ is $\lambda'_{m,n}$, for $n=0,1,\ldots,2N_p-1$. Thus, $\mathfrak{JB}\left\{\{\lambda_{m,n}\}_{n=0}^{N_p-1}\right\}$ in Eq. (3.29) can be expressed as

$$\begin{aligned}&\mathfrak{JB}\left\{\{\lambda_{m,n}\}_{n=0}^{N_p-1}\right\}\\ &\qquad\overset{\text{def}}{=} \frac{N_p}{3}\left(\mathfrak{S}^2\{\vec{\eta}_m\}+\frac{\left(\mathfrak{K}\{\vec{\eta}_m\}-3\right)^2}{4}\right),\end{aligned} \tag{3.30}$$

where $\mathfrak{S}^2\{\vec{\eta}_m\}$ and $\mathfrak{K}\{\vec{\eta}_m\}$ denote the sample skewness and the sample kurtosis of the data sequence vector $\vec{\eta}_m$. They are

$$\mathfrak{S}^2\{\vec{\eta}_m\} \ \overset{\text{def}}{=}\ \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{2N_p}\sum_{n=0}^{2N_p}(\lambda'_{m,n}-\hat{\bar{\eta}}_m)^3}{\left(\frac{1}{2N_p}\sum_{n=0}^{2N_p}(\lambda'_{m,n}-\hat{\bar{\eta}}_m)^2\right)^{3/2}}, \tag{3.31}$$

$$\mathfrak{K}^2\{\vec{\eta}_m\} \ \overset{\text{def}}{=}\ \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{2N_p}\sum_{n=0}^{2N_p}(\lambda'_{m,n}-\hat{\bar{\eta}}_m)^4}{\left(\frac{1}{2N_p}\sum_{n=0}^{2N_p}(\lambda'_{m,n}-\hat{\bar{\eta}}_m)^2\right)^{2}}, \tag{3.32}$$

where $\hat{\mu}_3$, $\hat{\mu}_4$, $\hat{\bar{\eta}}_m$, and $\hat{\sigma}^2$ are the sample estimates of the third central moment, the fourth central moment, the mean, and the variance of all entries in the vector $\vec{\eta}_m$, respectively. Note that since only L elements of $\{\lambda_{d_0,n}\}_{n=0}^{N_p-1}$ are not drawn from a Gaussian process (unlike the other N_p-L elements), our proposed classifier using JB-statistic will appear to be more accurate if the channel length L is relatively large.

If $\hat{d}_0 = d_0$, the CFR $\{\hat{H}_k\}_{k=0}^{N-1}$ is estimated by performing an N-point DFT on $\left\{\sqrt{\frac{N}{N_p}}\lambda_{\hat{d}_0,n}e^{j\frac{2\pi n\hat{d}_0}{N}}\right\}_{n=0}^{N_p-1}$ padded with $N-N_p$ zeros at the very end. Using our pro-

posed PDTs in this section, we can blindly detect the pilot positions at the receiver without any *a priori* information, so a novel dynamical PIT can be established to provide much more flexibility than the conventional FDPMT in [26]. Although we have considered DFT/IDFT-based interpolation technique for each data block, our proposed scheme can also be easily realized in the employment of other types of interpolation techniques such as Wiener interpolation, linear interpolation, or high-order interpolation methods (see [22, 23]).

3.3.2 Channel Equalization aided Iterative Symbol Reconstruction

After the pilot offset $\hat{d}_0$ and the CFR $\{\hat{H}_k\}_{k=0}^{N-1}$ are estimated, with the estimated pilot positions and the CFRs, the minimum-mean-square-error (MMSE) equalization of the symbols in the non-pilot frequency bins is performed. By nulling the subcarriers at the detected pilot positions, we can construct the received signal samples in the frequency domain such that

$$\breve{R}_k \stackrel{\text{def}}{=} \begin{cases} 0, & \text{if } \ k \ (\text{mod } M) = \hat{d}_0 \\ R_k, & \text{otherwise} \end{cases} \tag{3.33}$$

Denote the *MMSE equalizer matrix* by

$$\widetilde{G} \stackrel{\text{def}}{=} \text{diag}(g_0, g_1, \ldots, g_{N-1}), \tag{3.34}$$

where

$$g_k \stackrel{\text{def}}{=} \left(\hat{H}_k^* \, \sigma_s^2\right) \Big/ \left(|\hat{H}_k|^2 \sigma_s^2 + \sigma_\Omega^2\right) \tag{3.35}$$

and

$$\breve{\vec{\Upsilon}} \stackrel{\text{def}}{=} [\breve{R}_0, \ldots, \breve{R}_{N-1}]^T. \tag{3.36}$$

Therefore, the initial raw-data estimation directly from the received signal is given by

$$\breve{\vec{\mathcal{S}}}^{(0)} = \widetilde{W}^H \, \widetilde{G} \, \breve{\vec{\Upsilon}}, \tag{3.37}$$

and the corresponding initial symbol estimation is carried out as

$$\hat{S}_k^{(p)} = \arg\min_{\alpha\in\Theta} \left| \breve{S}_k^{(p)} - \alpha \right|, \qquad k = 0, 1, \ldots, N-1, \tag{3.38}$$

where $\hat{S}_k^{(p)}$ is the k^{th} information symbol estimate ($\hat{S}_k^{(p)}$ is also the k^{th} element of the vector $\hat{\vec{\mathcal{S}}}^{(p)}$), and its superscript (p) represents the iteration number p starting with $p = 0$; $\breve{S}_k^{(p)}$ denotes the k^{th} element of $\breve{\vec{\mathcal{S}}}^{(p)}$ (the iteration number p starts from 0 as well). Applying the *iterative amplitude reconstruction* (IAR) procedure for OFDM systems [49], we can iteratively reconstruct the information symbols as follows:

$$\breve{S}_k^{(p)} \overset{\text{def}}{=} \begin{cases} \breve{S}_k^{(p-1)}, & \text{if } \; k \pmod{M} = \hat{d}_0 \\ \breve{S}_k^{(0)}, & \text{otherwise} \end{cases} \tag{3.39}$$

where $\breve{\vec{\mathcal{S}}}^{(p-1)} = \widetilde{W}\hat{\vec{\mathcal{S}}}^{(p-1)}$. Then $\left\{\hat{S}_k^{(p)}\right\}_{k=0}^{N-1}$ is obtained according to Eq. (3.38). Repeat Eqs. (3.38) and (3.39) for more iterations ($p = 1, 2, 3, \ldots$). Heuristically speaking, only a few iterations ($p = 1$ or $p = 2$) are required for a reasonable convergence.

3.3.3 Simulation Comparisons for Different Schemes

In our Monte Carlo experiments, a block of 512 QPSK (quadrature phase-shift keying) information symbols are randomly generated in each trial. The pilot sequence length is chosen to be $N_p = 32$ (the *pilot overhead ratio* is thus 6.25%), the cyclic prefix length is $N_{cp} = 26$, and Doppler effect is absent. We will compare the OFDM systems involving our proposed three PIT/PDT schemes with the OFDM system using the conventional FDPMT suggested in [26], in terms of *correct pilot detection rate* and *symbol error rate*.

First, the CIR coefficients are randomly generated in accordance with an uncorrelated fading model such that they are characterized by the exponential power delay profile, i.e., $\sigma_l^2 = \exp(-4l/L)$, $l = 0, 1, \ldots, L-1$, where l is the channel-tap index. A thousand Monte Carlo trials are taken for statistical averages. In Figure 3.3, the

average correct (pilot) detection rates (for accurately spotting the true pilot offset d_0) versus different SNRs (ranging from -10 dB to 20 dB) are depicted for our proposed three different PDTs (denoted by "JB Detector", "Variance Detector", and "Subspace Detector" in the figures) subject to the channel length $L = 6$. It is clear that the JB-statistic based PDT leads to the best performance over the entire SNR range, while the subspace based PDT leads to the worst performance.

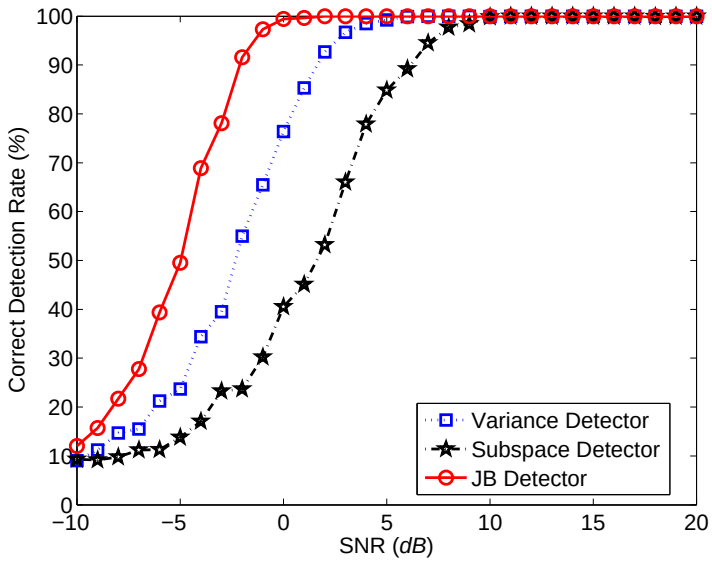

Figure 3.3. Average correct pilot-detection rates versus signal-to-noise ratios (SNRs) for the underlying OFDM transceiver depicted in Figures 3.1 and 3.2 subject to the channel length $L = 6$.

Then we keep all the simulation parameters except that the channel length is enlarged to $L = 26$ to redo Monte Carlo trials. The results are depicted in Figure 3.4. According to Figure 3.4, the JB-statistic based pilot detector still leads to the superior performance to the other two schemes. However, the variance based detector gives the worst performance for this scenario. The reason is that when the channel length gets large, the sample variances for the two conditions specified by Eq. (3.25) would become less and less distinguishable. On the other hand, the channel length would pose little impact on the subspace based detector.

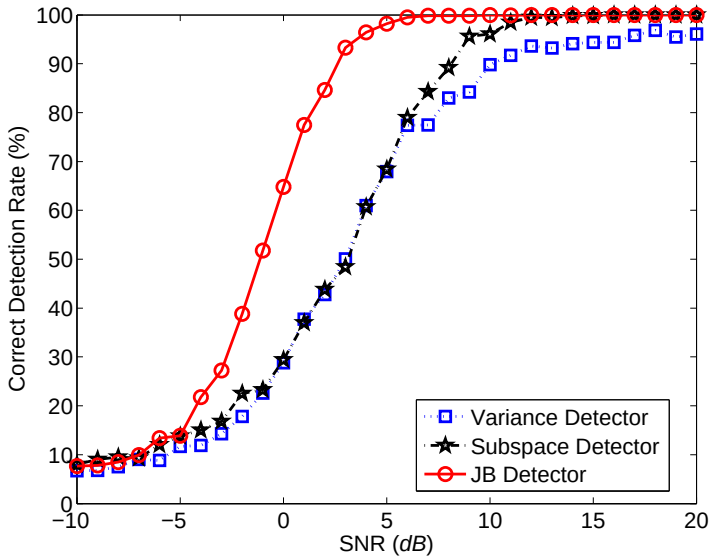

Figure 3.4. Average correct pilot-detection rates versus signal-to-noise ratios (SNRs) for the underlying OFDM transceiver depicted in Figures 3.1 and 3.2 subject to the channel length $L = 26$.

Note that only the last $N_p - N_{cp}$ elements in each sequence $\{\lambda_{m,n}\}_{n=0}^{N_p-1}$ are used to form the noise subspace, as stated by Section 3.3.1. If the CP length is equal to the number of pilots $N_{cp} = N_p$, the noise subspace has to diminish and no subspace based PDT is viable. Nevertheless, since most multipath channels in reality comply with an exponential power delay profile, the magnitudes of the last CIR coefficients are often quite small [50, 51]. Regardless of the mathematical restriction, a fixed small number of elements at the tail of $\{\lambda_{m,n}\}_{n=0}^{N_p-1}$ can still be used to form the noise subspace for detecting the pilot positions.

According to numerous simulations, it seems that the JB-statistic based pilot detector almost always leads to the highest correct detection rates. It is important to note that the JB-statistic involves the estimation of high-order central moments, which would not be so reliable when the sample size (N_p) is very small.

Moreover, the corresponding symbol error rates (SERs) to the above-mentioned OFDM systems are also compared in Figure 3.5. We also delineate the SER perfor-

mances for the OFDM receivers using the true dynamical pilot-position information (no PDT is necessary, denoted by "Ground Truth") and the conventional FDPMT with pilot insertion at the fixed positions (no pilot detection at the receiver, denoted by "Conventional FDPMT") [26], respectively. If there is no iterative symbol reconstruction, we write "$(p = 0)$", while only one iteration of symbol reconstruction is denoted by "$(p = 1)$". According to Figure 3.5, our proposed blind PDTs really can help the OFDM system to get very close to the performance resulting from the true pilot-position information. Besides, by minimizing the distortion caused by frequency-domain pilot-insertion, the SER performance of our proposed scheme is much better than that of the conventional FDPMT method. We have also found that the SER performances stay almost the same whether the total iteration number is selected as $p = 1$ or $p > 1$. It indicates that the iterative symbol reconstruction scheme is very effective even though very few iterations are executed.

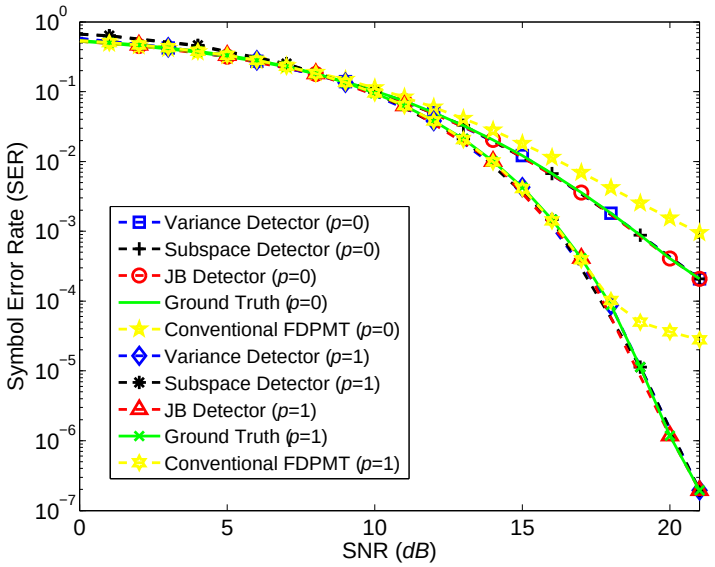

Figure 3.5. Symbol error rates versus signal-to-noise ratios (SNRs) for different OFDM systems where $N_p = 32$ is used in all OFDM transceivers.

In Figure 3.6, we also depict the SER performances for all of the aforementioned schemes as we demonstrated in Figure 3.5. We use exactly the same simulation set-up as before except that $N_p = 64$ is used here instead. In this scenario, twice information data are deleted in the frequency domain for pilot insertion. Intuitively, the more pilots, the more accurate pilot detection can be expected. However, although our proposed schemes using dynamical pilot insertion still lead to better SER performances than the conventional FDPMT method, the SER performances are worse than those resulting from $N_p = 32$ as shown in Figure 3.5. As a matter of fact, the more pilots are inserted, the more information symbols (deleted symbols due to pilot insertion) need to be reconstructed at the receiver; therefore the SER performances floor as the SNR gets large. This phenomenon clearly shows a tradeoff between the SER performance and the pilot-detection accuracy in our proposed schemes.

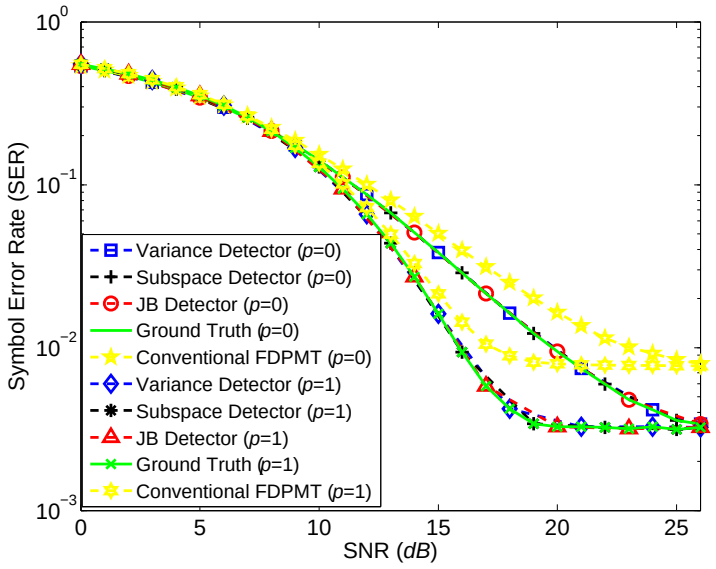

Figure 3.6. Symbol error rates versus signal-to-noise ratios (SNRs) for different OFDM systems where $N_p = 64$ is used in all OFDM transceivers.

3.3.4 Conclusion

In this chapter, a novel robust pilot insertion technique (PIT) and three corresponding blind pilot detection techniques (PDTs) were proposed for orthogonal frequency-division multiplexing systems. The proposed pilot insertion technique is optimal in the sense of minimum time-domain signal distortion caused by the subcarrier-removal in the frequency domain for pilot insertion. Furthermore, we design a new OFDM receiver structure including PDT assisted channel-estimation, channel equalization, and iterative symbol reconstruction. Numerous Monte Carlo simulation results demonstrate that our proposed new OFDM system using dynamical PIT and blind PDTs can achieve better symbol-error-rate performance than the conventional FDPMT (frequency-domain pilot-multiplexing technique) based OFDM system.

4 Blind Encoder Identification

4.1 AMC Transceiver with Blind Encoder Identification

The system diagram of an AMC transceiver involving our proposed new blind encoder identification mechanism is illustrated in Figure 4.1. Without loss of generality, we only consider the *baseband signaling*.

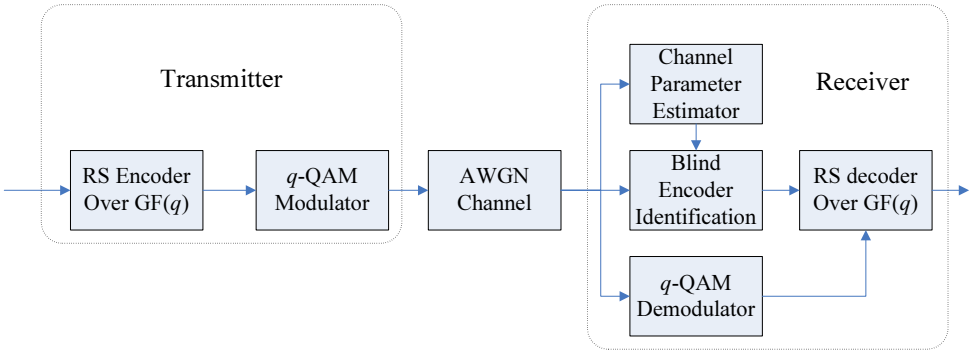

Figure 4.1. The system diagram of an AMC transceiver using our proposed new blind encoder identification mechanism.

4.1.1 Transmitter Model

At the transmitter, the information symbol sequences are encoded to generate RS codes over $\mathbb{GF}(q)$, whose encoder is selected from a predefined encoder candidate set to adapt to the current channel condition. The coded symbols are then modulated by q-QAM (quadrature amplitude modulation), and the modulated signal is transmitted through an additive white Gaussian noise (AWGN) channel. In this work, we only consider the RS codes over the binary extension fields, namely $\mathbb{GF}(2^{\mu})$ (μ can be an arbitrary positive integer, or $\mu \in \mathcal{Z}^{+}$), although our proposed encoder

identification scheme could be applied to other general fields as well. There are two different approaches to encode the RS codes. The first approach can be carried out in the time domain through the calculation of parity-check symbols, while the second approach can be carried out in the frequency domain through inverse Fourier transform. Throughout this section, we will focus on the inverse Fourier transform method (the second approach) to implement the RS encoder due to its low complexity. The details of the transmitter in Figure 4.1 are given as follows.

An i.i.d. (independently identically distributed) information sequence $\vec{b}_{k_\psi} = [b_1, \ldots, b_{k_\psi}]^T$ of length k_ψ over $\mathbb{GF}(q)$ is zero-padded at the end to construct an augmented sequence of length $q-1$, namely $\vec{b}_{q,\psi} = [b_1, \ldots, b_{k_\psi}, 0, \ldots, 0]^T$, where ψ is the index of a certain RS encoder drawn from the candidate set Ψ. It is obvious that the code rate of the encoder ψ is $R_\psi = k_\psi/(q-1)$. A $(q-1) \times (q-1)$ generator matrix $\widetilde{G}_q$ defined over $\mathbb{GF}(q)$ is then used to encode $\vec{b}_{q,\psi}$, resulting in a codeword $\vec{c}_\psi = [c_{1,\psi}, \ldots, c_{q-1,\psi}]^T \in \mathfrak{C}_\psi$, where $\mathfrak{C}_\psi$ denotes the set of codewords. Thus, $\mathfrak{C}_\psi$ is a subset of $[\mathbb{GF}(2^\mu)]^{(q-1)}$ where $q = 2^\mu$. This encoding process can also be expressed by

$$\vec{c}_\psi = \widetilde{G}_q \odot_q \vec{b}_{q,\psi}. \tag{4.1}$$

It is trivial to derive

$$\vec{c}_\psi \in \mathfrak{C}_\psi \Longleftrightarrow \widetilde{\Xi}_\psi \odot_q \vec{c}_\psi = \vec{0}, \tag{4.2}$$

where $\widetilde{\Xi}_{q,\psi}$ consists of the last $(q-1-k_\psi)$ rows of the matrix $\widetilde{F}_q$, where

$$\widetilde{F}_q \otimes_q \widetilde{G}_q = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}_{(q-1)\times(q-1)} . \tag{4.3}$$

In fact, when we encode the information sequence using an RS encoder over a binary extension field $\mathbb{GF}(2^{\mu})$ in the frequency domain, the matrices $\widetilde{F}_q$ and $\widetilde{G}_q$ are inverse Fourier transform matrix and Fourier transform matrix, respectively.

Now the codeword $\vec{c}_{\psi}$ will be mapped to a normalized q-QAM signalling constellation Ω, resulting in a modulated vector $\vec{t}_{\psi} = [t_{1,\psi},\ t_{2,\psi},\ \ldots,\ t_{q-1,\psi}]^T$, where $\mathbb{E}\{|t_{i,\psi}|^2\} = 1$ for $i = 1, 2, \ldots, q-1$. Suppose that the elements of $\mathbb{GF}(q)$ are α_0, α_1, α_2, ..., α_{q-1}, where $\alpha_0 = 0$ and s_0, s_1, s_2, ..., $s_{q-1} \in \mathcal{C}$ are the corresponding signaling constellations. Thus, $\Omega = \{s_0,\ s_1,\ \ldots,\ s_{q-1}\}$, and $t_{i,\psi} \in \Omega$, $\forall i$. Moreover, we assume that the received signal undergoes perfect time- and frequency-synchronization.

4.1.2 AWGN Channel

In this section, we assume that the modulated q-QAM signal will undergo the AWGN channel. The received base-band sequence within a code block as defined in the previous subsection is denoted by $\vec{r}_{\psi} = [r_{1,\psi}, r_{2,\psi}, \ldots, r_{q-1,\psi}]^T$, such that

$$\vec{r}_{\psi} = h\, e^{j\phi}\vec{t}_{\psi} + \vec{n}, \tag{4.4}$$

where h is the unknown channel gain (assume that the channel coherence time is larger than a code block period here), ϕ is the unknown phase offset, and $\vec{n}$ is the zero-mean complex AWGN vector of length $q-1$ with a variance σ^2 for its both real and imaginary parts. Consequently, the *signal-to-noise ratio* (SNR) per coded symbol of the modulated signal is given by

$$\varpi = \frac{h^2}{2\sigma^2}. \tag{4.5}$$

In order to evaluate the effect of different coding rates R_{ψ} over the same candidate set Ψ, the SNR per uncoded symbol of the modulated signal is given by

$$\zeta_{\psi} \stackrel{\text{def}}{=} \frac{\varpi}{R_{\psi}} = \frac{h^2}{2R_{\psi}\,\sigma^2}. \tag{4.6}$$

On the other hand, in order to evaluate the effect of different coding rates and different modulation orders, the SNR per uncoded symbol of the unmodulated signal is given by

$$\xi_{\psi,\mu} \stackrel{\text{def}}{=} \frac{\varpi/\mu}{R_\psi} = \frac{h^2}{2\mu R_\psi \sigma^2}. \tag{4.7}$$

4.1.3 Receiver Model

At a non-AMC receiver, the down-converted base-band signal will be fed into a q-QAM demodulator and then an appropriate RS decoder. However, in an AMC system, the encoder information is unknown to the receiver if no control channel is available. Therefore, according to Figure 4.1, we will pass $\vec{r}_\psi$ into our proposed blind encoder identification subsystem, which involves a channel parameter estimator and an encoder identifier. After the blind identification of the encoder information from the predefined candidate set has been accomplished, the corresponding RS decoder will be established accordingly. Meanwhile, $\vec{r}_\psi$ will also be passed into a q-QAM demodulator and then the decoder will be used to decode the demodulated symbols, resulting in the estimated information sequence $\hat{\vec{b}}_{k_\psi} = [\hat{b}_1, \ldots, \hat{b}_{k_\psi}]^T$.

First, the unknown channel parameters including the channel gain h, noise variance σ^2, and phase offset ϕ are blindly estimated using the EM algorithm according to [52] and [53], resulting in the estimated parameters, $\hat{h}$, $\hat{\sigma}^2$, and $\hat{\phi}$.

Blind Channel Encoder Identification

The LLRV of each received symbol's *a posteriori* probabilities is then calculated, and the LLRVs of the syndrome *a posteriori* probabilities are computed according to different RS parity-check relations. The RS encoder $\hat{\psi}$ subject to the maximum average LLR in the predefined candidate set Ψ will be selected as the one used at the transmitter and the corresponding decoder will be prepared accordingly. Our proposed new scheme will be manifested in details in Section 4.2.

$$\mathcal{L}_i(y) \stackrel{\text{def}}{=} \mathcal{L}_i(\iota \odot_q x_1 \oplus_q \kappa \odot_q x_2) = \ln \frac{P_r\{y=0\}}{P_r\{y=\alpha_i\}}$$
$$= \ln \frac{\sum_{\beta \in \mathbb{GF}(q)} \exp\left\{-\mathcal{L}(x_1=\beta) - \mathcal{L}(x_2 = \kappa^{-1} \odot_q \iota \odot_q \beta)\right\}}{\sum_{\beta \in \mathbb{GF}(q)} \exp\left\{-\mathcal{L}(x_1=\beta) - \mathcal{L}\Big(x_2 = \kappa^{-1} \otimes_q (\iota \odot_q \beta \oplus_q \alpha_i)\Big)\right\}}. \quad (4.11)$$

4.2 Our Proposed New Blind Channel Encoder Identification Method

Reed Solomon codes are a special family of the BCH (Bose, Ray-Chaudhuri and Hocquenghem) codes. The length of the RS code is one less than the size of the field where the symbols are defined [54]. Since the calculations have to be performed over the same Galois field, all the RS codes within the same predefined candidate set should be of the same length. The difference among the RS encoders in the same candidate set is the code rate, $R_\psi = k_\psi/(q-1)$. In order to formulate the LLRVs of the syndrome *a posteriori* probabilities from the parity-check relations and the LLRVs of the received base-band symbols $\vec{r}_\psi$, the basic log-likelihood algebra over $\mathbb{GF}(q)$ is introduced as follows.

4.2.1 Log-Likelihood over $\mathbb{GF}(q)$

Suppose that $\mathbb{GF}(q) = \{\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_{q-1}\}$, where $\alpha_0 = 0$. Unlike the binary case in which the LLR of a random variable (RV) is defined as a scaler, the LLRV of an RV over $\mathbb{GF}(q)$ is defined as a vector of length $q-1$, which is

$$\vec{\mathcal{L}}(x) \stackrel{\text{def}}{=} [\mathcal{L}(x=\alpha_1), \mathcal{L}(x=\alpha_2), \ldots, \mathcal{L}(x=\alpha_{q-1})], \quad (4.8)$$

where

$$\mathcal{L}\left(x=\alpha_i\right) \stackrel{\text{def}}{=} \ln \frac{P_r\{x=0\}}{P_r\{x=\alpha_i\}}, \quad i=1,\ldots,q-1. \quad (4.9)$$

Eq. (4.9) is the natural logarithm of the ratio between the probabilities of x taking values 0 and α_i. From now on, we will denote the i^{th} element of $\vec{\mathcal{L}}(x)$ by $\mathcal{L}_i(x),\ i = 1, 2, \ldots, q-1$ for notational convenience.

$$\vec{\mathcal{L}}(\gamma_{\psi,i}) \stackrel{\text{def}}{=} \boxplus\Big(\vec{\mathcal{L}}(c_{1,\psi}|r_{1,\psi}), \vec{\mathcal{L}}(c_{2,\psi}|r_{2,\psi}), \ldots, \vec{\mathcal{L}}(c_{q,\psi}|r_{q,\psi}), f_{i,1}, f_{i,2}, \ldots, f_{i,q-1}\Big) \quad (4.20)$$

Suppose that there are two RVs x_1 and x_2, and two elements ι and κ, all of which are defined over $\mathbb{GF}(q)$. Besides, the LLRVs of x_1 and x_2 are denoted by $\vec{\mathcal{L}}(x_1)$ and $\vec{\mathcal{L}}(x_1)$, respectively. We can thus define a new operator $\boxplus$ over $\mathbb{GF}(q)$ as

$$\boxplus\Big(\vec{\mathcal{L}}(x_1), \vec{\mathcal{L}}(x_2), \iota, \kappa\Big) \stackrel{\text{def}}{=} \vec{\mathcal{L}}(\iota \odot_q x_1 \oplus_q \kappa \odot_q x_2). \quad (4.10)$$

The i^{th} component of the right-hand-side (RHS) of Eq. (4.10) can be derived by Eq. (4.11) on top of next page, and $\mathcal{L}(x_1 = \beta)$ is defined as

$$\mathcal{L}(x_1 = \beta) \stackrel{\text{def}}{=} \mathcal{L}_i(x_1), \quad (4.12)$$

where $\alpha_i = \beta$, for $i = 1, \ldots, q-1$.

4.2.2 Blind Reed-Solomon Encoder Identification

Assume that the candidate set Ψ consists of M RS codes with code rates $R_\psi = k_\psi/(q-1)$, where $\psi = 1, 2, \ldots, M$. We sort these M codes according to their code rates in an ascending order, namely $0 < R_1 < R_2 < \cdots < R_M < 1$. Obviously,

$$\vec{\varsigma}_\psi \stackrel{\text{def}}{=} \widetilde{\Xi}_{q,1} \odot_q \vec{c}_\psi = [\delta_1, \ldots, \delta_{k_\psi - k_1}, 0, \ldots, 0]^T, \quad (4.13)$$

where $\widetilde{\Xi}_{q,1}$ consists of the last $q-1-k_1$ rows of the Galois field Fourier transform matrix $\widetilde{F}_q$, $\delta_i \neq 0$, for $i = 1, 2, \ldots, k_\psi - k_1$, the elements of $\vec{\varsigma}_\psi$, namely $\varsigma_{\psi,1}, \ldots, \varsigma_{\psi,q-1-k_1}$, are defined as code-checks, or *c-checks* in short, and the number of zeros at the end of $\vec{\varsigma}_\psi$ is ν. Therefore, the encoder identification problem becomes how to determine the number of zeros at the end of the *c-check vector* $\vec{\varsigma}_\psi$ such that

$$\hat{\psi} = i, \quad \text{subject to} \quad \nu = q - 1 - k_i. \quad (4.14)$$

We can also construct an index set $\mathfrak{P}_\psi$ for each RS code $\psi \in \Psi$, which consists of all the indices corresponding to the last $q - 1 - k_\psi$ rows of $\widetilde{F}_q$. It yields $\mathfrak{P}_\psi \stackrel{\text{def}}{=}$

$\{k_\psi + 1, k_\psi + 2, \ldots, q-1\}$. Suppose that the i^{th} row of $\widetilde{F}_q$ is denoted by $\vec{F}_{i,\psi} \stackrel{\text{def}}{=} [f_{i,1}, f_{i,2}, \ldots, f_{i,q-1}]$. Thus, we have

$$i \in \mathfrak{P}_\psi \Longleftrightarrow \vec{F}_{i,\psi} \odot_q \vec{c}_\psi = 0, \tag{4.15}$$

where $\vec{F}_{i,\psi} \odot_q \vec{c}_\psi$ can be expanded as

$$f_{i,1} \odot_q c_{1,\psi} \oplus_q f_{i,2} \odot_q c_{2,\psi} \oplus_q \cdots \oplus_q f_{i,q-1} \odot_q c_{q-1,\psi} = 0, \tag{4.16}$$

for $i \in \mathfrak{P}_\psi$.

Given a received base-band symbol $r_{j,\psi}$, $j = 1, \ldots, q-1$, the i^{th} element of the LLRV for $r_{j,\psi}$'s *a posteriori* probability can be formulated according to the Bayes's theorem as follows:

$$\begin{aligned}
\mathcal{L}_i(c_{j,\psi}|r_{j,\psi}) &= \mathcal{L}\left(c_{j,\psi} = \alpha_i | r_{j,\psi}\right) \\
&= \ln \frac{P_r\{c_{j,\psi} = 0 \,|\, r_{j,\psi}\}}{P_r\{c_{j,\psi} = \alpha_i \,|\, r_{j,\psi}\}} \\
&= \ln \frac{P_r\{r_{j,\psi} \,|\, c_{j,\psi} = 0\}}{P_r\{r_{j,\psi} \,|\, c_{j,\psi} = \alpha_i\}} + \ln \frac{P_r\{c_{j,\psi} = 0\}}{P_r\{c_{j,\psi} = \alpha_i\}} \\
&= \mathcal{L}(r_{j,\psi} \,|\, c_{j,\psi} = \alpha_i) + \mathcal{L}(c_{j,\psi} = \alpha_i) \\
&= \mathcal{L}_i(r_{j,\psi} \,|\, c_{j,\psi}) + \mathcal{L}_i(c_{j,\psi}),
\end{aligned} \tag{4.17}$$

where $\mathcal{L}_i(c_{j,\psi}) = 0$ because each symbol in any RS codeword is assumed to have equal probabilities of taking values α_i, for $i = 0, 1, \ldots, q-1$. Note that $\mathcal{L}_i(r_{j,\psi} \,|\, c_{j,\psi})$ can be calculated using the estimated channel parameters according to Section **??**. It is

$$\mathcal{L}_i(r_{j,\psi} \,|\, c_{j,\psi}) = \ln \frac{\exp\left\{-\frac{||r_{j,\psi} - \hat{h} e^{j\hat{\phi}} s_0||^2}{2\hat{\sigma}^2}\right\}}{\exp\left\{-\frac{||r_{j,\psi} - \hat{h} e^{j\hat{\phi}} s_i||^2}{2\hat{\sigma}^2}\right\}}, \tag{4.18}$$

where s_j is the modulated phaser corresponding to the coded symbol $\alpha_j \in \mathbb{GF}(q)$.

Define the "*check*" $\gamma_{\psi,i}$ as the symbol which equals the linear combination of the i^{th} row of $\widetilde{F}_q$ and the received base-band sequence $\vec{r}_\psi$ over $\mathbb{GF}(q)$. It is

$$\gamma_{\psi,i} \stackrel{\text{def}}{=} f_{i,1} \odot_q r_{1,\psi} \oplus_q f_{i,2} \odot_q r_{2,\psi} \oplus_q \cdots \oplus_q f_{i,q-1} \odot_q r_{q-1,\psi}, \tag{4.19}$$

for $i = k_1 + 1, \ldots, q-1$. If one wants to calculate the LLRV $\vec{\mathcal{L}}(\gamma_{\psi,i})$ of the syndrome *a posteriori* probability for a certain check, say $\gamma_{\psi,i}$ in Eq. (4.19), the formula can be obtained using Eq. (4.11) in a recursive manner, which is given by Eq. (4.20) on top of this page.

Heuristically speaking, according to the LLRV definition given by Eq. (4.8) and the parity-check relations given by Eq. (4.19), each element of $\vec{\mathcal{L}}(\gamma_{\psi,i})$ is expected to be a positive value for $i \in \mathfrak{P}_\psi$. Therefore, we define the average LLR, $\Upsilon_{\psi,i}$, by taking average of $\mathcal{L}_j(\gamma_{\psi,i})$ over $j = 1, 2, \ldots, q-1$, such that

$$\Upsilon_{\psi,i} \stackrel{\text{def}}{=} \frac{1}{q-1} \sum_{j=1}^{q-1} \mathcal{L}_j(\gamma_{\psi,i}). \tag{4.21}$$

Note that $\Upsilon_{\psi,i} > \Upsilon_{\psi,i'}$, for $i \in \mathfrak{P}_\psi$ and $i' \notin \mathfrak{P}_\psi$. Furthermore, the elements of $\vec{\mathcal{L}}(\gamma_{\psi,i'})$, for $i' \notin \mathfrak{P}_\psi$ would be sometimes positive and sometimes negative to result in a small magnitude or $|\Upsilon_{\psi,i'}| \to 0$.

The next step is to determine the set of i-indices for which $\Upsilon_{\psi,i}$ take large values. A straightforward approach would be employing a predefined threshold. However, this approach is not robust since different modulation/coding strategies and noise levels would significantly change the appropriate thresholds. Consequently, we propose to use a statistical test to tackle this problem. Suppose that the average LLR vector we obtain at the receiver is $\vec{\Upsilon}_\psi \stackrel{\text{def}}{=} [\Upsilon_{\psi,k_1+1},\ \Upsilon_{\psi,k_1+2},\ \ldots,\ \Upsilon_{\psi,q-1}]$. Assume that the distributions of $\Upsilon_{\psi,i}$ are two different Gaussian functions with different means and variances under the two hypotheses, namely H_1: $i \in \mathfrak{P}_\psi$ and H_0: $i \notin \mathfrak{P}_\psi$. Therefore, we have two conditional probability density functions:

$$p_{\Upsilon_{\psi,i}|H_0}(\Upsilon_{\psi,i}|H_0) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(\Upsilon_{\psi,i} - \varrho_1)^2}{2\sigma_1^2}\right), \tag{4.22}$$

$$p_{\Upsilon_{\psi,i}|H_1}(\Upsilon_{\psi,i}|H_1) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(\Upsilon_{\psi,i} - \varrho_2)^2}{2\sigma_2^2}\right), \tag{4.23}$$

where ϱ_1, ϱ_2, σ_1, and σ_2 can be easily estimated from $\Upsilon_{\psi,i}$ defined in Eq. (4.21) (details are omitted due to page limit). Hence we can obtain the likelihood ratio

$\Lambda(\Upsilon_{\psi,i})$ as given by

$$\Lambda(\Upsilon_{\psi,i}) \stackrel{\text{def}}{=} \frac{p_{\Upsilon_{\psi,i}|H_0}(\Upsilon_{\psi,i}|H_0)}{p_{\Upsilon_{\psi,i}|H_1}(\Upsilon_{\psi,i}|H_1)}. \tag{4.24}$$

It yields

$$\gamma_{\psi,i} = 0, \quad \text{if} \quad \Lambda(\Upsilon_{\psi,i}) < 1 \quad (H_1 \ \text{is} \ \text{true}), \tag{4.25}$$

$$\gamma_{\psi,i} \neq 0, \quad \text{if} \quad \Lambda(\Upsilon_{\psi,i}) \geq 1 \quad (H_0 \ \text{is} \ \text{true}). \tag{4.26}$$

4.2.3 Computational Complexity Reduction

In this subsection, we will introduce how the calculation of the LLRVs for the syndrome *a posteriori* probabilities, namely the calculation of Eqs. (4.11) and (4.20), can be simplified, since the computational complexity of our proposed new blind encoder identification scheme is mainly caused by this. First, the *Jacobi logarithm operator* between two real numbers τ_1 and τ_2, where $\tau_1, \tau_2 \in \mathcal{R}$, can be defined as

$$\tau_1 \star \tau_2 \stackrel{\text{def}}{=} \ln\left(e^{\tau_1} + e^{\tau_2}\right). \tag{4.27}$$

It is easy to discover that the Jacobi logarithm operator satisfies the associative law, that is

$$\begin{aligned} \tau_1 \star \tau_2 \star \tau_3 &= \ln\left(e^{\tau_1} + e^{\tau_2} + e^{\tau_3}\right) \\ &= (\tau_1 \star \tau_2) \star \tau_3. \end{aligned} \tag{4.28}$$

Thus, the summations in the numerator and the denominator of Eq. (4.11) both can be undertaken recursively. Generally speaking, the calculation of $\vec{\mathcal{L}}(\iota \odot_q x_1 \oplus_q \kappa \odot_q x_2)$ requires $2(q-1)^2$ real-valued additions, $2(q-1)^2$ Jacobi logarithm operations, $4q(q-1)$ multiplications over $\mathbb{GF}(q)$, and $q(q-1)$ additions over $\mathbb{GF}(q)$.

The Jacobi logarithm operator can also be formulated as

$$\tau_1 \star \tau_2 = \max(\tau_1, \tau_2) + \ln\left(1 + e^{-|\tau_1 - \tau_2|}\right), \tag{4.29}$$

where $\ln\left(1+e^{-|\tau_1-\tau_2|}\right) \in (0, \ln 2]$ is a relatively small number. When we use the Jacobi logarithm operation recursion to calculate the numerator and the denominator of Eq. (4.11), namely $\tau_1 \star \tau_2 \star \cdots \tau_q = (\cdots((\tau_1 \star \tau_2) \star \tau_3) \cdots \star \tau_q)$, some τ_i's, $i = 1, 2, \ldots, q$ will take large positive values. Therefore, $\tau_1 \star \tau_2 \star \cdots \tau_q$ can be further simplified by

$$\tau_1 \star \tau_2 \star \cdots \tau_q \approx \max(\tau_1, \tau_2, \ldots, \tau_q). \tag{4.30}$$

According to Eqs. (4.29) and (4.30), the computational complexity of $\vec{\mathcal{L}}(\iota \odot_q x_1 \oplus_q \kappa \odot_q x_2)$ becomes $2(q-1)^2$ real-valued additions, $4q(q-1)$ multiplications over $\mathbb{GF}(q)$, and $q(q-1)$ additions over $\mathbb{GF}(q)$.

Moreover, the multiplications/additions over the Galois field $\mathbb{GF}(q)$ can be easily implemented by two corresponding pre-calculated look-up tables over $\mathbb{GF}(q)$, respectively [55]. However, if we sort the elements of $\mathbb{GF}(q)$ such that $\alpha_0 = 0$, $\alpha_1 = \alpha^0$, $\alpha_2 = \alpha^1$, ..., $\alpha_{q-1} = \alpha^{q-2}$, where α is the primitive of $\mathbb{GF}(q)$, the computational complexity may be further reduced.

Since the first column of the Galois field Fourier transform matrix $\widetilde{F}_q$ consists of all 1's, we have $\iota_1 = 1$ (the subscript denotes the recursion number). Thus, in the first recursion, we need to compute $\mathcal{L}_i(f_{i,1} \odot_q r_{1,\psi} \oplus_q f_{i,2} \odot_q r_{2,\psi})$ where $\iota_1 = f_{i,1} = 1$, $\kappa_1 = f_{i,2}$, and ι_j, κ_j are actually the coefficients ι, κ defined in Eq. (4.11) (their additional subscript j indicates the j^{th} recursion). For the second recursion, we need to calculate $\mathcal{L}_i(f_{i,1} \odot_q r_{1,\psi} \oplus_q f_{i,2} \odot_q r_{2,\psi} \oplus_q f_{i,3} \odot_q r_{3,\psi})$ as given by

$$\begin{aligned}
&\mathcal{L}_i(f_{i,1} \odot_q r_{1,\psi} \oplus_q f_{i,2} \odot_q r_{2,\psi} \oplus_q f_{i,3} \odot_q r_{3,\psi}) \\
&= \mathcal{L}_i\Big(1 \odot_q (f_{i,1} \odot_q r_{1,\psi} \oplus_q f_{i,2} \odot_q r_{2,\psi}) \\
&\qquad \oplus_q f_{i,3} \odot_q r_{3,\psi}\Big),
\end{aligned} \tag{4.31}$$

where $\iota_2 = 1$. Similar results can be found for the subsequent recursions as well. Consequently, we have $\iota_i = 1$ for $i = 1, 2, \ldots, q-2$. Suppose that $\kappa^{-1} = \alpha^{\epsilon}$, where

$\epsilon \in \{0, 1, \ldots, q-2\}$, and $\vec{\mathcal{L}}(x_i) = \left[\frac{P_r\{x_i=0\}}{P_r\{x_i=\alpha^0\}}, \ldots, \frac{P_r\{x_i=0\}}{P_r\{x_i=\alpha^{q-2}\}}\right]$ for $i = 1, 2$. Thus, the numerator of Eq. (4.11) becomes

$$\begin{aligned}
&\ln \sum_{\beta \in \mathbb{GF}(q)} \exp\left\{-\mathcal{L}(x_1 = \beta) - \mathcal{L}(x_2 = \kappa^{-1} \odot_q \beta)\right\} \\
&\quad = -2 \star \left(-\frac{P_r\{x_1=0\}}{P_r\{x_1=\alpha^0\}} - \frac{P_r\{x_2=0\}}{P_r\{x_2=\alpha^{\epsilon}\}}\right) \star \cdots \\
&\qquad \star \left(-\frac{P_r\{x_1=0\}}{P_r\{x_1=\alpha^{q-2}\}} - \frac{P_r\{x_2=0\}}{P_r\{x_2=\alpha^{\epsilon-1}\}}\right).
\end{aligned} \tag{4.32}$$

Therefore, we can construct another LLRV for x_2 with $\kappa^{-1} = \alpha^{\epsilon}$ such that

$$\vec{\mathcal{L}}_{\epsilon}(x_2) = \left[\frac{P_r\{x_2 = 0\}}{P_r\{x_2 = \alpha^{\epsilon}\}}, \ldots, \frac{P_r\{x_2 = 0\}}{P_r\{x_2 = \alpha^{\epsilon-1}\}}\right], \tag{4.33}$$

which is a circularly right shift of $\vec{\mathcal{L}}(x_2)$ by a step size ϵ. According to Eq. (4.33), Eq. (4.32) can be reformulated as

$$\begin{aligned}
&\ln \sum_{\beta \in \mathbb{GF}(q)} \exp\left\{-\mathcal{L}(x_1 = \beta) - \mathcal{L}(x_2 = \kappa^{-1} \odot_q \beta)\right\} \\
&= -2 \star \vec{\mathcal{L}}(x_1) \bigstar \vec{\mathcal{L}}_{\epsilon}(x_2),
\end{aligned} \tag{4.34}$$

where $\bigstar$ is the element-wise Jacobi logarithm operator between two vectors. A similar approach can be taken to calculate the denominator of Eq. (4.11) as well. Hence, actually no look-up table for the multiplications over $\mathbb{GF}(q)$ is necessary.

Consequently, when one tries to directly calculate the LLRV of the syndrome *a posteriori* probabilities of a certain check, namely $\vec{\mathcal{L}}(\gamma_{\psi,i})$ in Eq. (4.20), it requires $2(q-1)^3$ real-valued additions, $2(q-1)^3$ Jacobi logarithm operations, $4q(q-1)^2$ multiplications over $\mathbb{GF}(q)$, and $q(q-1)^2$ additions over $\mathbb{GF}(q)$. Nevertheless, when the simplified method in Section 4.2.3 is used, the total computational complexity is reduced to $2(q-1)^3$ real-valued additions in conjunction with the storage of a $q \times q$ look-up table.

4.3 Simulation Results and Conclusion

The performance of our proposed novel blind channel encoder identification scheme is evaluated by Monte Carlo simulations, in terms of the probability that the receiver can correctly identify the encoder from a predefined candidate set given only one codeword block of received base-band sequence. Specifically, four encoders RS$(15, 5)$, RS$(15, 7)$, RS$(15, 9)$, RS$(15, 11)$ in $\mathbb{GF}(16)$, four encoders RS$(31, 15)$, RS$(31, 19)$, RS$(31, 23)$, RS$(31, 27)$ in $\mathbb{GF}(32)$, and four encoders RS$(63, 43)$, RS$(63, 47)$, RS$(63, 51)$, RS$(63, 55)$ in $\mathbb{GF}(64)$ are adopted here because they are popular. For each Galois field, all the four RS codes are used to form the candidate set Ψ. The phase offset ϕ is randomly generated within $(-\pi/4, \pi/4)$ restricted by the quadrature symmetry. The probabilities of correct identification are attained by taking average over five hundred Monte Carlo trials.

Figure 4.2 illustrates the correct identification rate versus the SNR per uncoded symbol of the modulated signals, ζ_ψ defined in Eq. (4.6), among the aforementioned four different RS codes over $\mathbb{GF}(16)$. The results indicate that the correct identification rate can approach 100% for each encoder as $\zeta_\psi \geq 20$ dB. Similar trends can also be found from Monte Carlo simulations using other RS codes over $\mathbb{GF}(32)$ and $\mathbb{GF}(64)$, which are delineated in Figures 4.3 and 4.4, respectively.

According to Figures 4.2-4.4, when SNR is fairly low, we can identify the RS encoder with the lowest code rate very well. In this situation, the likelihood ratio test facilitated by Eqs. (4.22)-(4.26) would almost always choose H_1 because $\varrho_1 \approx \varrho_2 \approx 0$ and $\sigma_1 > \sigma_2$ when ζ_ψ is relatively small. Thus, $\Lambda(\Upsilon_{\psi,i}) < 1$, which means all checks tend to be classified as 0. In other words, if the encoder with the lowest code rate is used at the transmitter, it will almost always be identified correctly when SNR is low. It actually makes sense in practice since when the channel SNR is low, the AMC system would very probably choose the encoder with a low code rate from the

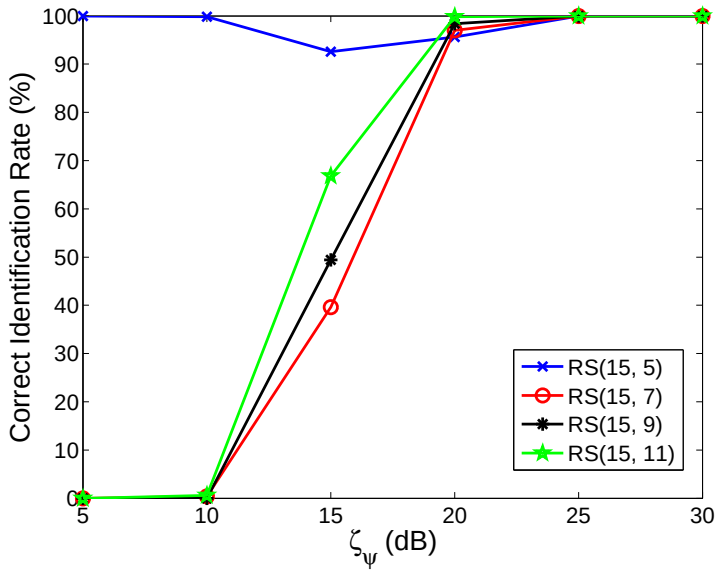

Figure 4.2. The correct identification rate with respect to ζ_ψ defined in Eq. (4.6) for four different RS codes over $\mathbb{GF}(16)$.

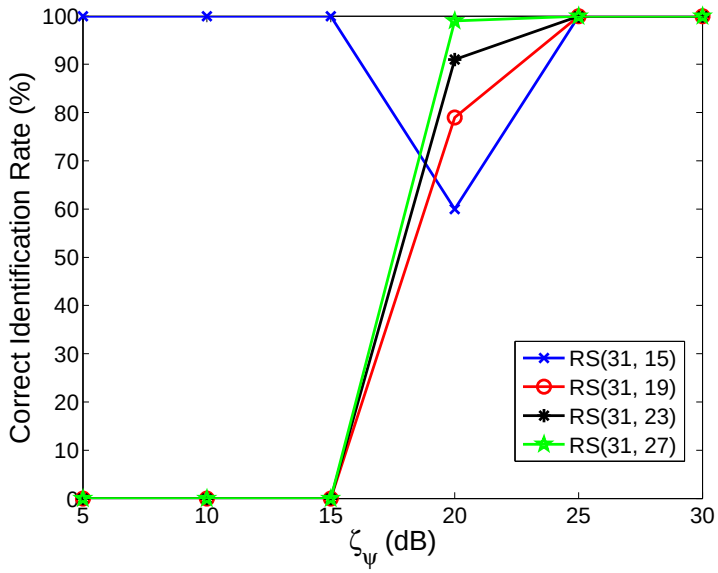

Figure 4.3. The correct identification rate with respect to ζ_ψ defined in Eq. (4.6) for four different RS codes over $\mathbb{GF}(32)$.

candidate set. Figure 4.5 depicts the correct identification rate versus the SNR per uncoded symbol of the unmodulated signals, $\xi_{\psi,\mu}$ defined in Eq. (4.7), for RS$(15, 9)$, RS$(31, 23)$, RS$(63, 51)$, RS$(127, 111)$, and RS$(255, 223)$. As a result, if one wants

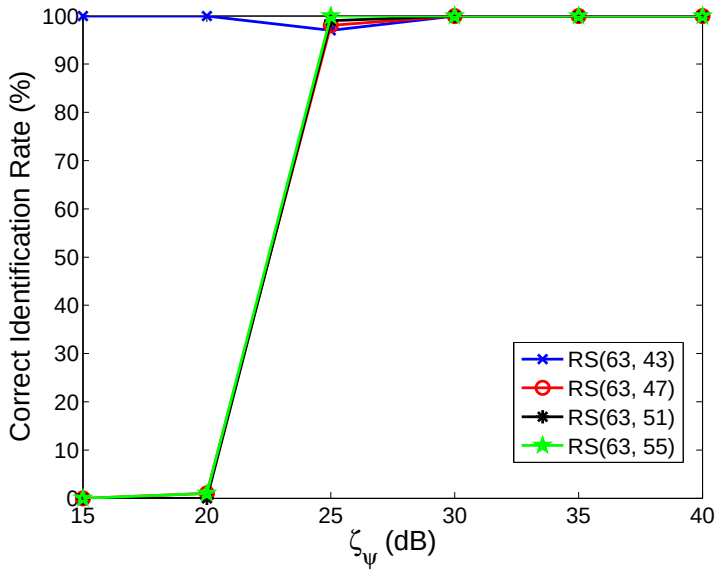

Figure 4.4. The correct identification rate with respect to ζ_ψ defined in Eq. (4.6) for four different RS codes over $\mathbb{GF}(64)$.

to achieve the same correct identification rate using our proposed new scheme, the required SNR is larger for the RS encoder over a larger Galois field.

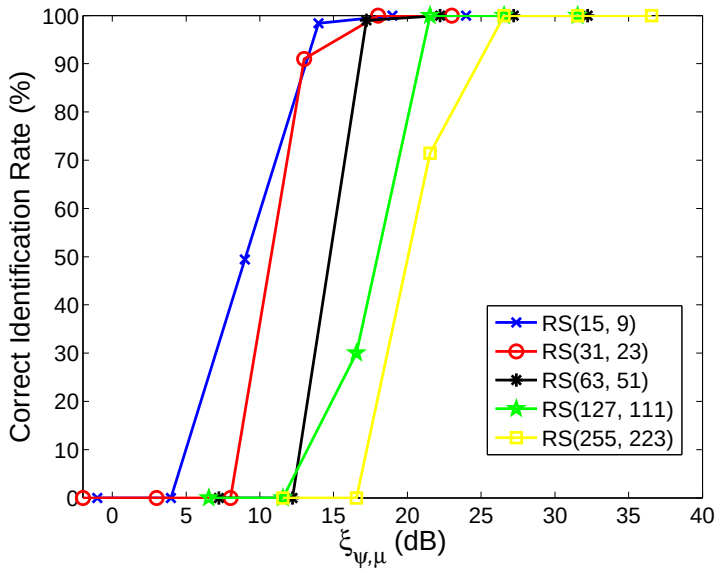

Figure 4.5. The correct identification rate with respect to $\xi_{\psi,\mu}$ defined in Eq. (4.7) for different RS codes over different Galois fields.

In this chapter, a novel blind identification scheme with low computational complexity was proposed for Reed-Solomon encoders over $\mathbb{GF}(q)$. The proposed new scheme is based on the average log-likelihood ratio of syndrome *a posteriori* probabilities, which is calculated from the LLRVs of the received base-band sequence and the corresponding parity-check relations, and the likelihood ratio test to determine the values of the checks. In addition, the needed calculation was also simplified to come up with a new fast computational procedure. The new method is very promising for the future cognitive radio technology or the next generation adaptive modulation and coding systems.

5 Conclusions

In this dissertation, we discussed the methodology of designing robust and efficient statistical signal sensing, detection and estimation algorithms which could be applied in wireless communication systems. The field of signal sensing, detection and estimation is concerned with the analysis of received signals to determine the presence or absence of signals of interest, to classify the useful statistical information, and to extract information either purposefully or inadvertently included in these signals.

Three novel and robust statistical signal processing algorithms are proposed for different communication applications, namely spectrum sensing, symbol-reconstruction/ channel-estimation and blind encoder identification. First, we proposed a novel adaptive cooperative spectrum-sensing scheme based on JB statistic single-reception spectrum-sensing technique. We also found that the commonly-used sample-average estimator for the cumulative weights in the data-fusion rule becomes unreliable in time-varying environments. To overcome this drawback, we adopt a temporal discount factor, which is crucial to the probability estimators. New theoretical analysis to justify the advantage of our proposed new estimators over the conventional sample-average estimators and to determine the optimal numerical value of the proposed discount factor is presented. The Monte Carlo simulation results are also provided to demonstrate the superiority of our proposed adaptive cooperative spectrum-sensing method in both stationary and time-varying environments.

Second, we proposed the frequency-domain pilot multiplexing techniques (FDPMTs) for the channel estimation and equalization in OFDM systems. A robust and effective pilot insertion and detection scheme is devised thereby. The information signal sequence resulting from the constellation mapper is spread over all subcarriers by a precoder and certain subcarriers can be nulled for the insertion of training pilots. These pilot positions are optimally selected to minimize the distortion of the transmitted time-domain signal (OFDM modulated signal) caused by the aforementioned subcarrier-removal at the corresponding pilot positions. The associated new receiver structure is also presented, where three different blind pilot-detection techniques are designed without any *a priori* knowledge of the pilot positions (based on sample variance, subspace decomposition, and Jarqur-Bera (JB) statistics, respectively), and the distorted data symbols can thus be iteratively reconstructed. Besides, rigorous theoretical analysis and Monte Carlo simulation results both demonstrate that our proposed new OFDM system using dynamical pilot positions is more robust than the conventional OFDM system using the fixed pilot positions over multipath fading channels.

Third, we proposed a novel blind channel-encoder identification scheme with low computational complexity for Reed-Solomon (RS) codes over Galois field $\mathbb{GF}(q)$, which could also be applied to other similar non-binary channel codes as well. Our proposed new scheme involves the estimation of the channel parameters using the expectation-maximization (EM) algorithm, the calculation of the log-likelihood ratio vectors (LLRVs) of the syndrome *a posteriori* probabilities over $\mathbb{GF}(q)$, and the identification of the non-binary RS encoder in use subject to the maximum average log-likelihood ratio (LLR) over the pre-selected candidate encoder set. Simulation results justify the effectiveness of this new mechanism.

In summary, we proposed different statistical signal processing methods for different wireless communication applications in this dissertation. It can be seen that when proper statistical signal processing schemes are adopted, the corresponding design of the wireless communication systems can be greatly simplified. However, it is obvious that more studies for different signal sensing, detection and estimation problems are needed in the future. Moreover, to evaluate each technique, more precise and more diversified criteria can be foreseen preferable by researchers. Moreover, for different signal sensing detection and estimation problems, more effective and reliable techniques are also in urgent need. Consequently, how to propose more robust, efficient and reliable signal sensing detection and estimation techniques would remain challenging in the scientific and engineering communities.

References

[1] Qihang Peng, Kun Zeng, Jun Wang, and Shaoqian Li. A distributed spectrum sensing scheme based on credibility and evidence theory in cognitive radio context. In *Proceedings of IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–5, September 2006.

[2] Adam Noel and Robert Schober. Convex sensing-reporting optimization for cooperative spectrum sensing. *IEEE Transactions on Wireless Communications*, 11(5):1900–1910, May 2012.

[3] Y.-C. Liang, Y. H. Zeng, E. C. Y. Peh, and A. T. Hoang. Sensing-throughput tradeoff for cognitive radio networks. *IEEE Transactions on Wireless Communication*, 7(4):1326–1337, April 2008.

[4] A. Sahai and D. Cabric. Spectrum sensing: fundamental limits and practical challenges. In *Proceedings of IEEE International Symposium on New Frontiers Dynamic Spectrum Access Networks*, November 2005.

[5] H. S. Chen, W. Gao, and D. G. Daut. Signature based spectrum sensing algorithms for IEEE 802.22 WRAN. In *Proceedings of IEEE International Conference on Communications*, pages 6487–6492, June 2007.

[6] S. Enserink and D. Cochran. A cyclostationary feature detector. In *Proceedings of Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 806–810, October-November 1994.

[7] Y. P. Lin and C. He. Subsection-average cyclostationary feature detection in cognitive radio. In *Proceedings of International Conference on Neural Networks and Signal Processing*, pages 604–608, July 2008.

[8] Kenta Umebayashi, Hironori Tsuchiya, and Yasuo Suzuki. Analysis of optimal weighted cooperative spectrum sensing with multiple antenna elements. *IEICE Transactions on Communications*, E95-B(10):3261–3269, October 2012.

[9] Bin Shen, Sana Ullah, and Kyungsup Kwak. Deflection coefficient maximization criterion based optimal cooperative spectrum sensing. *International Journal of Electronics and Communications*, 64(9):819–827, September 2010.

[10] Saman Atapattu, Chintha Tellambura, and Hai Jiang. Energy detection based cooperative spectrum sensing in cognitive radio networks. *IEEE Transactions on Wireless Communications*, 10(4):1232–1241, April 2011.

[11] S. M. Kay. *Fundamentals of statistical signal processing: detection theory*. Prentice-Hall, Upper Saddle River, New Jersey, 1998.

[12] S. J. Shellhammer, S. Shankar, R. Tandra, and J. Tomcik. Performance of power detector sensors of DTV signals in IEEE 802.22 WRANs. In *Proceedings of the Fist International Workshop on Technology and Policy for Accessing Spectrum*, August 2006.

[13] Lu Lu, Hsiao-Chun Wu, and S. S. Iyengar. A novel robust detection algorithm for spectrum sensing. *IEEE Journal on Seleted Areas in Communications*, 29(2):305–315, Feburary 2011.

[14] Z. Chair and P. K. Varshney. Optimal data fusion in multiple sensor detection systems. *IEEE Transactions on Aerospace and Electronic Systems*, 22(1):98–101, January 1986.

[15] Naghmeh Mansouri and Madjid Fathi. Simple counting rule for optimal data fusion. In *Proceedings of IEEE Conference on Control Applications*, volume 2, pages 1186–1191, June 2003.

[16] N. Ansari, J. G. Chen, and Y. Z. Zhang. Adaptive decision fusion for unequiprobable sources. In *IEE Proceedings of Radar, Sonar and Navigation*, volume 144, pages 105–111, June 1997.

[17] Hongting Zhang, Hsiao-Chun Wu, Lu Lu, and S. S. Iyengar. Adaptive cooperative spectrum sensing based on a novel robust detection algorithm. In *Proceedings of IEEE International Conference on Communications*, pages 3560–3564, June 2012.

[18] Hongting Zhang, Hsiao-Chun Wu, and Lu Lu. Analysis and algorithm for robust adaptive cooperative spectrum-sensing. *IEEE Transactions on Wireless Communications*, 13(2):618–629, February 2014.

[19] Hongting Zhang, Hsiao-Chun Wu, and Shih-Yu Chang. Analysis and algorithm for robust adaptive cooperative spectrum-sensing in time-varying environments. In *Proceedings of IEEE International Conference on Communications*, pages 2617–2621, June 2013.

[20] L. He, Y. Wu, and S. Ma. Superimposed training based channel estimation and data detection for OFDM amplify-and-forward cooperative systems under high mobility. *IEEE Transactions on Signal Processing*, 60(1):274–284, January 2012.

[21] A. G. Orozco-Lugo, M. M. Lara, and D. C. Mclernon. Channel estimation using implicit training. *IEEE Transactions on Signal Processing*, 52(1):240–254, January 2004.

[22] S. Coleri, M. Ergen, A. Puri, and A. Bahai. Channel estimation techiques based on pilot arrangement in OFDM systems. *IEEE Transactions on Broadcasting*, 48(3):223–229, September 2002.

[23] F. Sanzi, J. Sven, and J. Speidel. A comparative study of iterative channel estimators for mobile OFDM systems. *IEEE Transactions on Wireless Communications*, 12(5):849–859, September 2003.

[24] W. C. Huang, C. P. Li, and H. J. Li. On the power allocation and system capacity of OFDM systems using superimposed training schemes. *IEEE Transactions on Vehicular Technology*, 58(9):240–254, May 2009.

[25] M. Ghogho, D. McLernon, E. Alameda-Hernandez, and A. Swami. Channel estimation and symbol detection for block transmission using data-dependent superimposed training. *IEEE Signal Processing Letters*, 12(3):226–229, March 2005.

[26] Gaoqi Dou, Chunquan He, Congying Li, and Jun Gao. Channel estimation and symbol detection for OFDM systems using data-nulling superimposed pilots. *Electronics Letters*, 50(3):179–180, January 2014.

[27] Hongting Zhang and Hsiao-Chun Wu. Robust pilot detection techniques for channel estimation and symbol detection in OFDM systems. *accepted by IEEE Signal Processing Letters*, 2014.

[28] Hongting Zhang, Hsiao-Chun Wu, Hong Jiang, and Scott Huang. Robust pilot detection techniques for OFDM systems. In *accepted by IEEE Global Telecommunications Conference*, December 2014.

[29] Xiaozhou Huang, Hsiao-Chun Wu, and Yiyan Wu. Novel pilot-free adaptive modulation for wireless OFDM systems. *IEEE Transactions on Vehicular Technology*, 57(6):3863–3867, November 2008.

[30] Tian Xia and Hsiao-Chun Wu. Novel blind identification of LDPC codes using average LLR of syndrome *a posteriori* probability. In *Proceedings of IEEE International Conference on Intelligent Transport Systems Telecommunications*, pages 12–16, November 2012.

[31] R. Moosavi and E.G. Larsson. A fast scheme for blind identification of channel codes. In *Proc. IEEE Global Telecommunications Conference (GLOBECOM'2011)*, pages 1–5, Houston, TX, December 2011.

[32] Yonas G. Debessu, Hsiao-Chun Wu, and Hong Jiang. Novel blind encoder parameter estimation for turbo codes. *IEEE Communications Letters*, 16(12):1917–1920, December 2012.

[33] Andrea J. Goldsmith and Soon-Ghee Ghee Chua. Adaptive coded modulation for fading channels. *IEEE Transactions on Communications*, 46(5):595–602, May 1998.

[34] Y. Zrelli, M. Marazin, R. Gautier, and E. Rannou. Blind identification of convolutional encoder parameters over GF(2^m) in the noiseless case. In *Proc. IEEE International Conference on Computer Communications and Networks (ICCCN'2011)*, pages 1–5, August 2011.

[35] H. Wymeersch, H. Steendam, and M. Moeneclaey. Log-domain decoding of LDPC codes over GF(q). In *Proc. IEEE International Conference on Communications (ICC'2004)*, pages 772–776, Paris, France, June 2004.

[36] Hongting Zhang, Hsiao-Chun Wu, and Hong Jiang. Novel blind encoder identification of Reed-Solomon codes with low computational complexity. In *Proceedings of IEEE Global Telecommunications Conference*, pages 3294–3299, December 2013.

[37] Y. H. Zeng and Y.-C. Liang. Eigenvalue based spectrum sensing algorithms for cognitive radio. *IEEE Transactions on Communications*, 57(6):1784–1793, June 2009.

[38] A. N. Mody. Spectrum sensing of the DTV in the vicinity of the video carrier using higher order statistics, July 2007.

[39] Lu Lu, Hsiao-Chun Wu, and S. S. Iyengar. A novel robust detection algorithm for spectrum sensing. *IEEE Journal on Seleted Areas in Communications*, 29(2):305–315, Feb. 2011.

[40] C. Clanton, M. Kenkel, and Y. Tang. Wireless microphone signal simulation method, March 2007.

[41] Mourad Barkat. *Signal Detection and Estimation*. Atech House Inc., 2 edition, 2005.

[42] Chengqi Song and Qian Zhang. Sliding-window algorithm for asynchronous cooperative sensing in wireless cognitive networks. In *Proceedings of IEEE International Conference on Communications*, pages 3432–3436, May 2008.

[43] Tom M. Mitchell. *Machine Learning*. MIT Press and McGraw-Hill Companies, Inc., 1997.

[44] Kenta Umebayashi, Janne J. Lehtomäki, Takanao Yazawa, and Yasuo Suzuki. Efficient decision fusion for cooperative spectrum sensing based on OR-rule. *IEEE Transactions on Wireless Communications*, 11(7):2585–2595, July 2012.

[45] D. Chu. Polyphase codes with good periodic correlation properties. *IEEE Transactions on Information Theory*, 18(4):531–532, July 1972.

[46] Hsiao-Chun Wu. Analysis and characterization of intercarrier and interblock interferences for wireless mobile OFDM systems. *IEEE Transactions on Broadcasting*, 52(2):203–210, June 2006.

[47] X. Wang, H.-C. Wu, S. Y. Chang, Y. Wu, and J.-Y. Chouinard. Efficient non-pilot-aided channel length estimation for digital broadcasting receivers. *IEEE Transactions on Broadcasting*, 55(3):633–641, September 2009.

[48] Lu Lu, Hsiao-Chun Wu, and S.-S. Iyengar. A novel robust detection algorithm for spectrum sensing. *IEEE Journal on Selected Areas in Communications*, 29(2):305–315, February 2011.

[49] U. K. Kwon, D. Kim, K. Kim, and G. H. Im. Amplitude clipping and iterative reconstruction of STBC/SFBC-OFDM signals. *IEEE Signal Processing Letters*, 14(11):808–811, November 2007.

[50] M. R. Raghavendra and K. Giridhar. Improving channel estimation in OFDM systems for sparse multipath channels. *IEEE Signal Processing Letters*, 12(1):52–55, January 2005.

[51] J. Wang and J. Chen. Performance of wideband CDMA with complex spreading and imperfect channel estimation. *IEEE Journal on Selected Area in Communications*, 19(1):152–163, January 2001.

[52] W. Gappmair, M. Flohberger, and O. Koudelka. Moment-based estimation of the signal-to-noise ratio for oversampled narrowband signals. In *Mobile and Wireless Communications Summit, 2007. 16th IST*, pages 1–4, July 2007.

[53] W. Gappmair, R. Lopez-Valcarce, and C. Mosquera. Joint NDA estimation of carrier frequency/phase and SNR for linearly modulated signals. *IEEE Signal Processing Letters*, 17(5):517–520, May 2010.

[54] Peter Sweeney. *Error Control Coding: From Theory to Practice*. John Wiley & Sons, LTD, Baffins Lane, Chichester, West Sussex PO19 IUD, England, 2002.

[55] T. Clevorn and P. Vary. Low-complexity belief propagation by approximations with lookup-tables. In *Proc.* 5^{th} *International ITG Conference on Source and Channel Coding (SCC'2004)*, pages 211–215, Erlangen, Germany, January 2004.

List of Publications

1. **Hongting Zhang**, Hsiao-Chun Wu, and Lu Lu, "Analysis and Algorithm for Robust Adaptive Cooperative Spectrum Sensing", *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 618–629, Feb. 2014.

2. **Hongting Zhang**, and Hsiao-Chun Wu, "Robust Pilot Detection Techniques for Channel Estimation and Symbol Detection in OFDM Systems", *accepted by IEEE Signal Processing Letters*, 2014.

3. Lu Lu, **Hongting Zhang**, and Hsiao-Chun Wu, "Novel Energy-Based Localization Technique for Multiple Sources", *IEEE Systems Journal*, vol. 8, no. 1, pp. 142–150, Mar. 2014.

4. XiaoyuFeng, **Hongting Zhang**, and Hsiao-Chun Wu, "A New Approach for Optimal Multiple Watermarks Injection", *IEEE Signal Processing Letters*, vol. 18, no. 10, pp. 575–578, Oct. 2011.

5. Limei Guo, Hsiao-Chun Wu, **Hongting Zhang**, Tian Xia, Shahab Mehraeen, "Robust Optimization for Home-Load Scheduling under Price Uncertainty in Smart Grids", *accepted by Proc. International Conference on Computing, Networking and Communications (ICNC'2015)*, Feburary, 2015.

6. **Hongting Zhang**, Hsiao-Chun Wu, Hong Jiang, and Scott Huang, "Robust Pilot Detection Techniques for OFDM Systems", *accepted by Proc. IEEE Global Telecommunications Conference (Globecom'2014)*, Dec. 2014.

7. **Hongting Zhang**, Hsiao-Chun Wu, and Hong Jiang, "Novel Blind Encoder Identification of Reed-Solomon Codes with Low Computational Complexity", *Proc. IEEE Global Telecommunications Conference (Globecom'2013)*, Dec. 2013, pp. 3294–3299.

8. **Hongting Zhang**, Hsiao-Chun Wu, and Shih Yu Chang, "Analysis and Algorithm for Robust Adaptive Cooperative Spectrum-Sensing in Time-Varying Environments", *Proc. IEEE International Conference on Communications (ICC'2013)*, June. 2013, pp. 2617–2621.

9. **Hongting Zhang**, Hsiao-Chun Wu, and Shih Yu Chang, "New Fast Optimal Window Design Algorithm Based on the Eigen-Decomposition of the Symmetric Toeplitz Matrix", *Proc. IEEE International Conference on Communications (ICC'2013)*, June. 2013, pp. 4766–4770.

10. **Hongting Zhang** and Hsiao-Chun Wu, "Novel Fast MUSIC Algorithm for Spectral Estimation with High Subspace Dimension", *Proc. International Conference on Computing, Networking and Communications (ICNC'2013)*, Jan. 2013, pp. 474–478.

11. **Hongting Zhang**, Hsiao-Chun Wu, and Lu Lu, "Adaptive Cooperative Spectrum Sensing Based on a Novel Robust Detection Algorithm", *Proc. IEEE International Conference on Communications (ICC'2012)*, Jun. 2012, pp. 3560–3564.

12. **Hongting Zhang**, Hsiao-Chun Wu, Yiyan Wu, and Bo Rong, "Novel Fast Algorithm to Design Optimal Eigen Filters for Transmission Systems", *Proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB'2012)*, Jun. 2012, pp. 1–5.

13. Xiaoyu Feng, **Hongting Zhang**, Hsiao-Chun Wu, Yiyan Wu, "A New Approach for Optimal Multiple Watermarks Injection", *Proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB'2011)*, Jun. 2011, pp. 1–6.

Vita

Hongting Zhang received a Bachelor Engineering degree from Harbin Institute of Technology Electrical Engineering Department in 2009. She got a Master of Engineering degree from Louisiana State University in 2011. She was a research intern at Alcatel Lucent - Bell Labs in 2012 and wireless system intern at Broadcom Corporation in 2013. She is currently pursuing the degree of Doctor of Philosophy in the Department Electrical Engineering and Computer Science, Louisiana State University, Baton Rouge. Her research interests are in the areas of wireless communications and signal processing.

Zhang, Hongting

B.S. Electrical Engineering, Harbin Institute of Technology, 2009

M.S. Electrical Engineering, Louisiana State University, 2011

Doctor of Philosophy, Fall Commencement, 2014

Major: Electrical and Computer Engineering

ADVANCED STATISTICAL SIGNAL PROCESSING METHODS IN SENSING, DETECTION, AND ESTIMATION FOR COMMUNICATION APPLICATIONS

Dissertation directed by Professor Hsiao-Chun Wu