Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

ABSTRACT

BLIND RECOGNITION OF ANALOG MODULATION SCHEMES FOR SOFTWARE DEFINED RADIO

by
Haifeng Xiao

With the emergence of software defined radios (SDRs), an adaptive receiver is needed that can configure various parameters, such as carrier frequency, bandwidth, symbol timing, and signal to noise ratio (SNR), and automatically identify modulation schemes. In this dissertation research, several fundamental SDR tasks for analog modulations are investigated, since analog radios are often used by civil government agencies and some unconventional military forces. Hence, the detection and recognition of "old technology" analog modulations remain an important task both for civil and military electronic support systems and for notional cognitive radios.

In this dissertation, a Cyclostationarity-Based Decision Tree classifier is developed to separate between analog modulations and digital modulations, and classify signals into several subsets of modulation types. In order to further recognize the specific modulation type of analog signals, more effort and work are, however, needed. For this purpose, two general methods for automatic modulation classification (AMC), feature-based method and likelihood-based method, are investigated in this dissertation for analog modulation schemes. For feature-based method, a multi-class SVM-based AMC classifier is developed. After training, the developed classifier can achieve high classification accuracy in a wide range of SNR. While the likelihood-based methods for digital modulation types have been well developed, it is noted that the likelihood-based methods for analog modulation types are seldom explored in the literature. Average-

Likelihood-Ratio-Testing based AMC algorithms have been developed to automatically classify AM, DSB and FM signals in both coherent and non-coherent situations In addition, the Non-Data-Aided SNR estimation algorithms are investigated, which can be used to estimate the signal power and noise power either before or after modulation classification.

BLIND RECOGNITION OF ANALOG MODULATION SCHEMES FOR SOFTWARE DEFINED RADIO

by
Haifeng Xiao

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering

Department of Electrical and Computer Engineering

August 2012

.

APPROVAL PAGE

BLIND RECOGNITION OF ANALOG MODULATION SCHEMES FOR SOFTWARE DEFINED RADIO

Haifeng Xiao

Dr. Yun Q. Shi, Dissertation Co-advisor Professor of Electrical and Computer Engineering, NJIT	Date
Dr. Wei Su, Dissertation Co-advisor US Army CERDEC, Fort Monmouth, NJ	Date
Dr. Richard Haddad, Committee Member Professor of Electrical and Computer Engineering, NJIT	Date
Dr. John Kosinski, Committee Member Senior Consulting Scientist, MacAulay-Brown, Inc., Dayton OH	Date
Dr. Stewart D. Personick, Committee Member Senior University Lecturer, NJIT	Date
Dr. Osvaldo Simeone, Committee Member Assistant Professor of Electrical and Computer Engineering, NJIT	Date

BIOGRAPHICAL SKETCH

Author: Haifeng Xiao

Degree: Doctor of Philosophy

Date: August 2012

Undergraduate and Graduate Education:

- Doctor of Philosophy in Electrical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2012

- Master of Engineering in Electrical Engineering,
  Beijing University of Aeronautics and Astronautics, Beijing, P. R. China, 2006

- Bachelor of Engineering in Electrical Engineering,
  Beijing University of Aeronautics and Astronautics, Beijing, P. R. China, 2003

Major: Electrical Engineering

Presentations and Publications:

H. Xiao, C. Chen, W. Su, J. Kosinski and Y. Q. Shi, "Data Mining of Modulation Types Using Cyclostationarity-Based Decision Tree," Proceedings of the 2008 International Conference on Data Mining (DMIN'08), Las Vegas, Nevada, USA, pp. 3-9, August 2008.

H. Xiao, W. Su, J. Kosinski and Y. Q. Shi, "Blind carrier frequency estimation for SSB-SC signals," Proceedings of the 2010 IEEE Radio and Wireless Symposium (RWS'10), New Orleans, LA, USA, pp. 312-315, Jan 2010.

H. Xiao, W. Su, J. Kosinski and Y. Q. Shi, "An investigation of non-data-aided SNR estimation techniques for analog modulation signals," Proceedings of the 2010 IEEE Sarnoff Symposium, Princeton (Sarnoff'10), NJ, USA, pp. 1-5, April 2010.

H. Xiao, W. Su, J. Kosinski and Y. Q. Shi, "Automatic classification of analog modulation schemes," Proceedings of the 2012 IEEE Radio and Wireless Symposium (RWS'12), Santa Clara, CA, USA, pp. 5-8, Jan 2012.

H. Xiao, W. Su, J. Kosinski and Y. Q. Shi, “Support Vector Machine based AMC for Analog Schemes,” The Computing Science and Technology International Journal (CSTIJ), vol. 2, no. 1, March, 2012, ISSB (print) 2162-0660, ISSN (online) 2162-0687.

Z.Yu, H. Xiao, W. Su, J. Kosinski and Y. Q. Shi, “Cyclostationarity-Based Decision Tree for Joint Classification of Digital and Analog Modulations,” Submitted to IEEE Transactions on Broadcasting.

献给我的家人
To my beloved family

ACKNOWLEDGMENT

I would like to express my gratitude to my advisor, Professor Yun Q. Shi for his tremendous assistance with this dissertation. From the beginning to the end, he has been a steadfast source of information, ideas, support, and energy. I am deeply grateful for his guidance, patience, and encouragement and I will be forever grateful for his trust and support that this dissertation could be completed.

I would also like to express my appreciation to the distinguished members of the dissertation committee: Drs. Richard Haddad, John Kosinski, Stewart D. Personick, Osvaldo Simeone, and Wei Su, for their active participation and valuable comments. Special thanks are given to Dr. Wei Su for his constant support, remarkable comments on my research, and careful revising of my final version of the dissertation. Without his support, my dissertation will not be in its current shape.

Many of my present and former colleagues in the Electrical and Computer Engineering Department at NJIT are deserving recognition for their help during my graduate study life.

Finally, I would like to thank my parents for their support and encouragement. All have been encouraging; I would quite simply not have completed this dissertation without their support and encouragement.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

CHAPTER 1

INTRODUCTION

This chapter provides the background and outline of this doctoral research. The motivation and objective of the relevant research are first addressed. Then contributions of this dissertation are briefly presented. Finally, the outline of this dissertation is given.

1.1 Background and Motivations

A software-defined radio (SDR) system, is a radio communication system where components, which have been conventionally implemented in hardware (mixers, filters, amplifiers, modulators/demodulators, detectors, and so on) are instead implemented by means of software on a personal computer or embedded computing devices [1-3].

A basic SDR system may consist of a personal computer equipped with a sound card, or other analog-to-digital converter, preceded by some form of RF front end. Significant amounts of signal processing are handed over to the general-purpose processor, rather than being done in special-purpose hardware. Such a design produces a radio which can receive and transmit widely different radio protocols (sometimes referred to as waveforms) based solely on the software used.

While SDR approach is becoming more common in the civil, military and cell phone services, an adaptive receiver is needed that can configure various parameters, such as carrier frequency, bandwidth, symbol timing, and signal to noise ratio (SNR), and automatically identify modulation schemes.

This dissertation mainly focuses on blind recognition of analog modulation schemes for software defined radio receiver. Analog radios are often used by civil government agencies and some unconventional military forces. Hence, the detection and recognition of "old technology" analog modulations remains an important task both for military electronic support systems, and for notional cognitive radios.

1.2 Contributions Made in This Dissertation Research

The contributions of this dissertation include the following:

- A **Cyclostationarity-Based Decision Tree classifier** is developed to separate between analog modulations and digital modulations, and classify signals into several subsets of modulation types as desired. But in order to further specify the modulation type of analog signals, more effort and work are needed.
- Two general methods for automatic modulation classification (AMC), feature-based method and likelihood-based method, are investigated in this dissertation for analog modulation schemes.
- For feature-based method, a **multi-class-SVM-based AMC classifier** is developed. After carefully training, the developed classifier can achieve high classification accuracy in a wide range of SNR. Machine-learning is a powerful tool for classification problems, especially for those problems that can not be solved by the likelihood-based methods (e.g.,. LSB and USB are difficult to be recognized by the likelihood-based methods, but can be recognized by SVM).
- Likelihood-based methods for digital modulation types have been well developed. But for analog modulation types, LB methods are seldom explored.

Various **Average-Likelihood-Ratio-Testing-based AMC algorithms** are developed to automatically classify AM, DSB and FM signals in both coherent and non-coherent situations

- **Non-Data-Aided (NDA) SNR estimation algorithms** are investigated, which can be used to estimate the signal power and noise power either before modulation classification or after modulation classfication.
- Nandi's AMR work on analog modulation is evaluated, so as to compare with the work proposed in this dissertation.

1.3 Outline of This Dissertation

In Chapter 2, a thorough evaluation of Nandi's AMC work on analog modulation schemes has been reported, and the classification results are presented.

Chapter 3 reports the investigation of the technology of NDA SNR estimator to eliminate the requirement for additional knowledge in the environment of blind classification of analog modulation schemes.

In Chapter 4, an AMC algorithm for classification of joint analog and digital modulations based on the pattern of cyclostationarities and the spectral properties of modulated signals has been presented.

In Chapter 5, a new feature based AMC algorithm has been presented for blind classification of analog modulation signals. The histogram of instantaneous frequency is used as the modulation-classification features. Support Vector Machines (SVMs) are used as the classifier. Simulation results have shown that the proposed classification method is quite robust to additive noise in a wide range of SNR.

In Chapter 6, the investigation of automatic modulation classification for analog modulation schemes is presented. A likelihood ratio testing framework for both the coherent and non-coherent cases is presented. The performance of both the coherent and non-coherent classifier is investigated for various modulation types including AM, DSB and FM. Non-coherent performance exhibites a 2~3dB loss compared to coherent performance.

Chapter 7 contains the conclusions and future work of this dissertation.

CHAPTER 2

EVALUATION OF NANDI'S AMC ALGORITHMS ON ANALOG MODULATION SIGNALS

2.1 Introduction

Azzouz and Nandi developed five algorithms for analog modulation recognition, based on the decision-theoretic approach in [4]. Their developed features are quite often employed by other researchers (e.g., [5-7]). Thus it is worth thoroughly evaluating the performance of these algorithms.

These algorithms use the same key features but different time order of applying these key features in the classification procedure. In this dissertation, the evaluation will be performed on these features. In Section 2.2, the signal models and the feature definition of [4] will be introduced. Simulation results, analysis and discussions on these features will be presented in Section 2.3. Probabilities of correct classification for both simulated and field-collected data are reported in Section 2.4. The simulation results shows the validity of the key features developed in [4].

2.2 Signal Models and Feature Definitions

2.2.1 Signal Models

The received signal can be expressed as

$$r(t) = s(t) + n(t) \tag{2.1}$$

where $s(t)$ represents the modulated signal, and $n(t)$ is additive white Gaussian noise (AWGN) with zero-mean and variance of σ^2. For analog modulation types, the signal $s(t)$ can be expressed as below [8], respectively,

AM:

$$s(t) = A\left(1 + \mu m(t)\right) \cdot e^{j\omega_c t} \tag{2.2}$$

DSB:

$$s(t) = A\, m(t) \cdot e^{j\omega_c t} \tag{2.3}$$

LSB:

$$x(t) = A\left(m(t) - j\hat{m}(t)\right) \cdot e^{j\omega_c t} \tag{2.4}$$

USB:

$$x(t) = A\left(m(t) + j\hat{m}(t)\right) \cdot e^{j\omega_c t} \tag{2.5}$$

FM:

$$s(t) = A \cdot e^{j\omega_c t + j2\pi f_\Delta \int_{-\infty}^{t} m(v)dv} \tag{2.6}$$

where A is a positive factor used to control the signal power, ω_c is the carrier frequency, $m(t)$ represents the real-valued information-bearing signal that satisfies $-1 \le m(t) \le 1$, $\hat{m}(t)$ stands for the Hilbert transform of $s(t)$, f_Δ is the frequency deviation of an FM signal, μ is the modulation index of an AM signal, and $0 < \mu \le 1$.

Before digitization, the intermediate frequency (IF) signal $r(t)$ was band-pass filtered and down-converted $r(t)$ toward 0 Hz. The down-converted signal can be expressed as

$$x(t) = r(t)\, e^{-j\hat{\omega}t} \tag{2.7}$$

where $\hat{\omega}$ is the estimated carrier frequency.

The received signal $x(t)$ is sampled at sampling frequency f_s, and $x(t_i)$ is denoted as the discrete version of $x(t)$. Assuming that the length of the observed data sequence is N, then the time instants $t_i = n/f_s$, $i = 1, 2, \ldots, N$.

2.2.2 Key Feature Extraction

For each segment, the instantaneous amplitudes $\{a(i) : i = 1, 2, \ldots, N\}$, and the unwrapped instantaneous phases $\{\varphi(i) : i = 1, 2, \ldots, N\}$ are extracted, where the index i represents that $a(i)$ and $\varphi(i)$ are for the i th sample.

Based on the instantaneous amplitudes $a(i)$ and the unwrapped instantaneous phases $\varphi(i)$ as well as the signal spectrum, three key features are derived to discriminate between the modulation types of interest.

The first key feature, σ_{dp}, is defined by

$$\sigma_{dp} = \sqrt{\left[\frac{1}{N}\sum_{i=1}^{N}(\varphi_{NL}(i))^2\right] - \left[\frac{1}{N}\sum_{i=1}^{N}\varphi_{NL}(i)\right]^2} \tag{2.8}$$

where $\varphi_{NL}(i)$ is *centralized-nonlinear instantaneous phase*, and it is defined by

$$\varphi_{NL}(i) = \varphi(i) - \frac{1}{N}\sum_{i=1}^{N}\varphi(i) \tag{2.9}$$

The second key feature, P, measures the spectrum symmetry, and is defined as:

$$P = \frac{P_L - P_U}{P_L + P_U} \tag{2.10}$$

where

$$P_L = \sum_{i=1}^{f_{cn}} \left| X(i) \right|^2 \tag{2.11}$$

$$P_U = \sum_{i=1}^{f_{cn}} \left| X(i + f_{cn} + 1) \right|^2 \tag{2.12}$$

where $X(i)$ stands for the DFT of the received signal $x(t_i)$, and $(f_{cn} + 1)$ is the index of the DFT bin corresponding to the center frequency, f_c, and f_{cn} is defined as

$$f_{cn} = \frac{f_c N}{f_s} - 1 \tag{2.13}$$

The third key feature, $\gamma_{\max}$, is defined by

$$\gamma_{\max} = \frac{1}{N} \max \left| DFT(a_{cn}(i)) \right|^2 \tag{2.14}$$

where $DFT(a_{cn}(i))$ stands for the N-point DFT of $\{a_{cn}(i)\}$, and $a_{cn}(i)$ is *centralized-normalized instantaneous amplitude*, and is defined by

$$a_{cn}(i) = a_n(i) - 1 \tag{2.15}$$

where $a_n(i)$ is *normalized instantaneous amplitude*, and is defined by

$$a_n(i) = \frac{a(i)}{\frac{1}{N} \sum_{i=1}^{N} a(i)} \tag{2.16}$$

2.3 Evaluation of the Key Features

Based on the aforementioned three key features, Azzouz and Nandi [4] developed five algorithms for analog modulation recognition. The details of only one algorithm are presented in this section, as shown in Figure 2.1. To evaluate the performance of the

AMC algorithm in Figure 2.1, both simulated data and field-collected data has been tested.

Simulated signals are generated in the MATLAB environment. The simulated signals are passed through an AWGN channel and the SNR is defined as

$$\text{SNR} = 10 \cdot \log_{10}\left[\frac{S}{\sigma_0 \cdot f_s}\right] \quad \text{dB} \tag{2.17}$$

where S is the signal power, σ_0 is the white noise power spectral density and $f_s = 44.1$ kHz is the sampling frequency, which is chosen to be slightly larger than the occupied bandwidth of most narrowband communications. The carrier frequency is prior information. The range of SNR is from -2dB to 26dB. With a 4dB increment, 100 snapshots of simulated signals are generated for each modulation type.

Field signals are transmitted through the air and collected by hardware, with white Gaussian noise added by software. The same simulation environment are adopt as for simulated data.

The performance evaluations are derived from 100 snapshots, each with length of 0.6667 sec, for each modulated signal of interest at different SNRs.

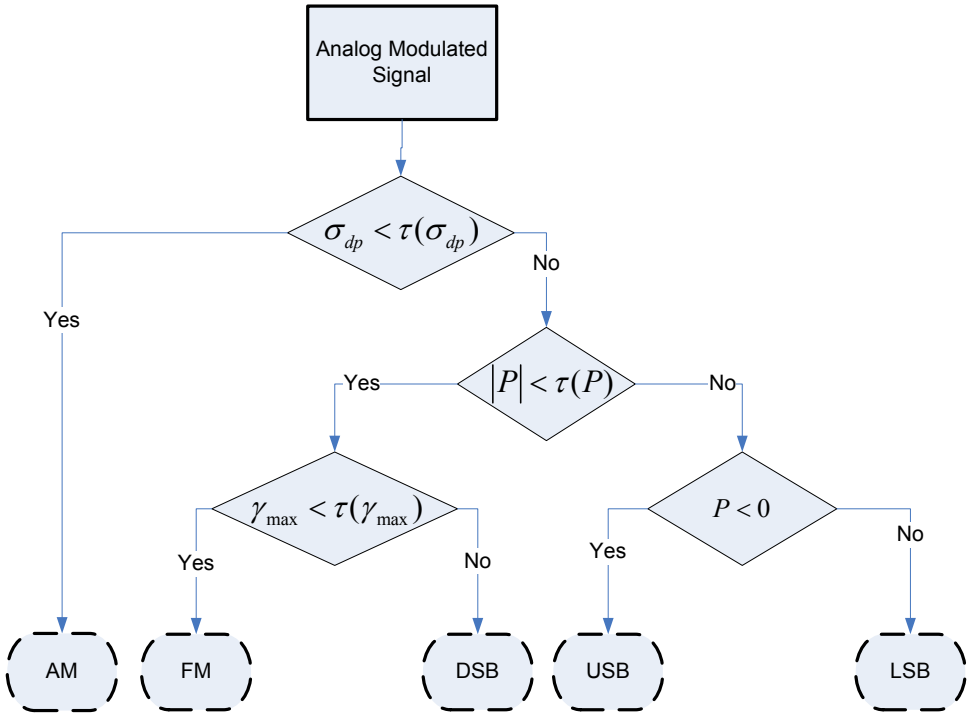


Figure 2.1 Flowchart of Nandi's AMR algorithm for analog modulation signals

2.3.1 Comments on Feature σ_{dp}

σ_{dp} is used to discriminate between {AM} and {DSB, LSB, USB, FM}. {AM} has no direct phase information (i.e., $\sigma_{dp} < \tau(\sigma_{dp})$); on the other hand, {FM, DSB, LSB, USB} have direct phase information (i.e., $\sigma_{dp} \geq \tau(\sigma_{dp})$) by their nature. The value of $\tau(\sigma_{dp})$ is chosen to be 20.

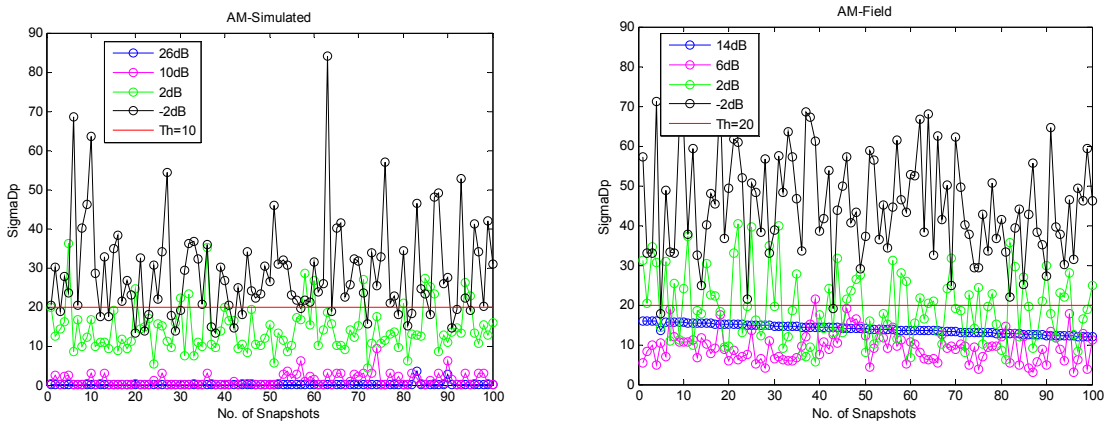


Figure 2.2 Measured values of σ_{dp} for AM signals under different SNRs (left column: simulated data, right column: field data), $\sigma_{dp} < \tau(\sigma_{dp})$

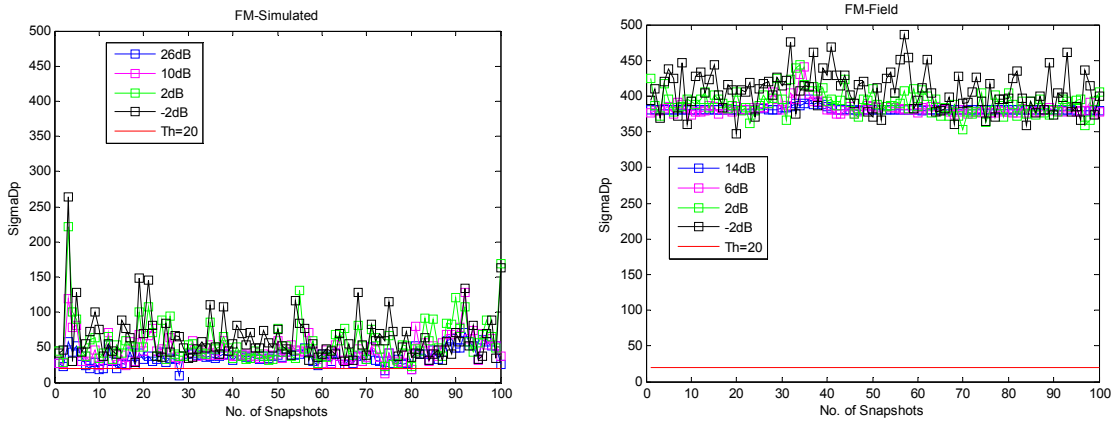


Figure 2.3 Measured values of σ_{dp} for FM signals under different SNRs (left column: simulated data, right column: field data), $\sigma_{dp} \geq \tau(\sigma_{dp})$

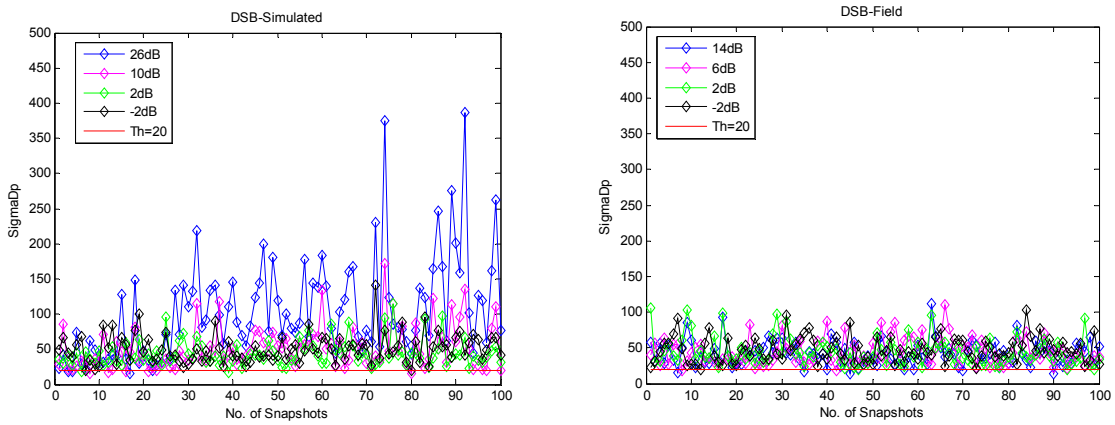


Figure 2.4 Measured values of σ_{dp} for DSB signals under different SNRs (left column: simulated data, right column: field data), $\sigma_{dp} \geq \tau(\sigma_{dp})$

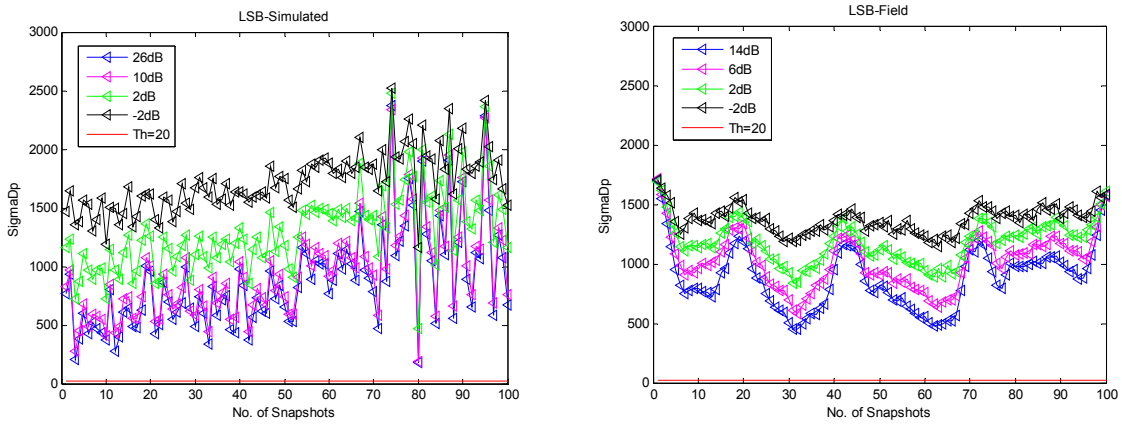


Figure 2.5 Measured values of σ_{dp} for LSB signals under different SNRs (left column: simulated data, right column: field data), $\sigma_{dp} \geq \tau(\sigma_{dp})$

2.3.2 Comments on Feature P

P is used for measuring the spectrum symmetry around the center frequency. The absolute value of P is used to discriminate between {FM, DSB} and {LSB, USB}. $|P|$ at infinite SNR ought to be 0 for {DSB, FM} (i.e., $|P| < \tau(P)$), and 1 for{LSB, USB}

(i.e., $|P| \geq \tau(P)$). Furthermore, {LSB} and {USB} are discriminated by checking the sign of P. The value of $\tau(P)$ is chosen to be 0.5.

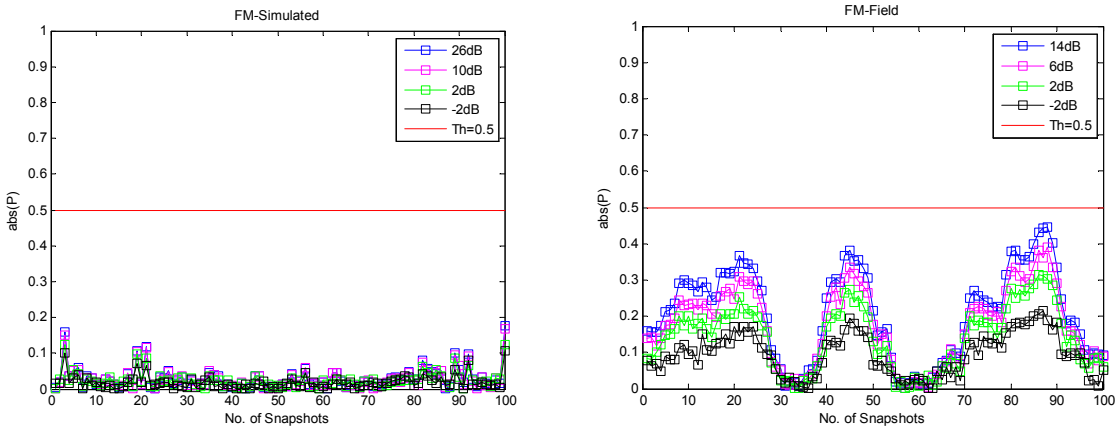


Figure 2.6 Measured values of $|P|$ for FM signals under different SNRs (left column: simulated data, right column: field data), $|P| < \tau(P)$

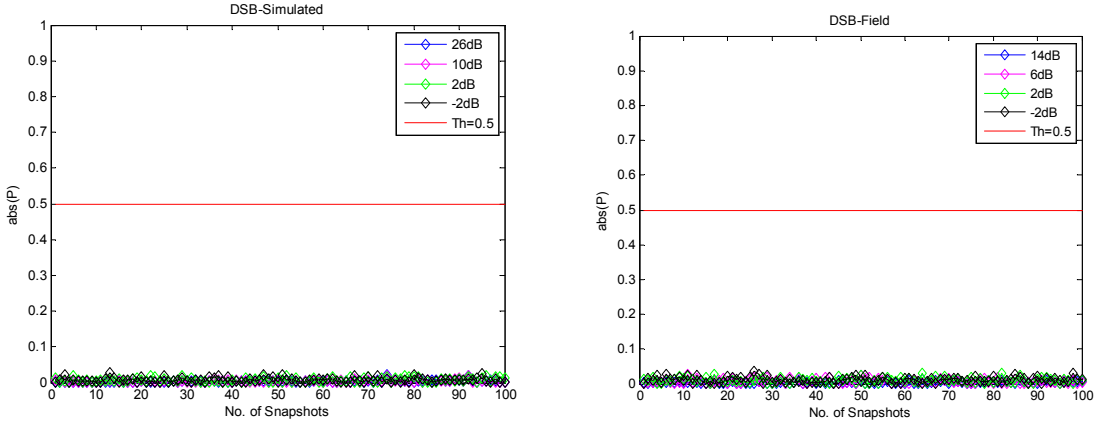


Figure 2.7 Measured values of $|P|$ for DSB signals under different SNRs (left column: simulated data, right column: field data), $|P| < \tau(P)$

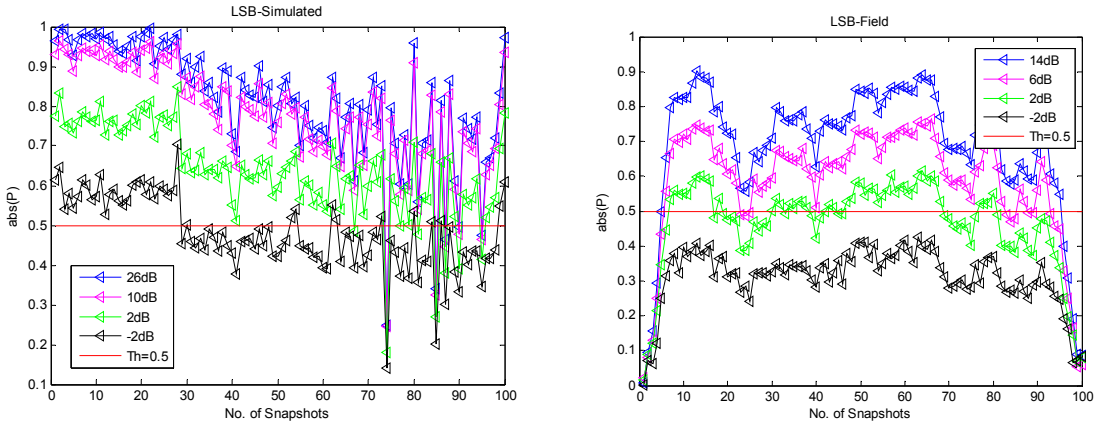


Figure 2.8 Measured values of $|P|$ for LSB Signals under different SNRs (left column: simulated data, right column: field data), $|P| \geq \tau(P)$

2.3.3 Comments on Feature $\gamma_{\max}$

$\gamma_{\max}$ is used to discriminate between {FM} and {DSB}. As the FM signals have constant instantaneous amplitude, they have no amplitude information, $\gamma_{\max} < \tau(\gamma_{\max})$; on the other hand, DSB signals have amplitude information, $\gamma_{\max} \geq \tau(\gamma_{\max})$. The value of $\tau(\gamma_{\max})$ is chosen to be 61.

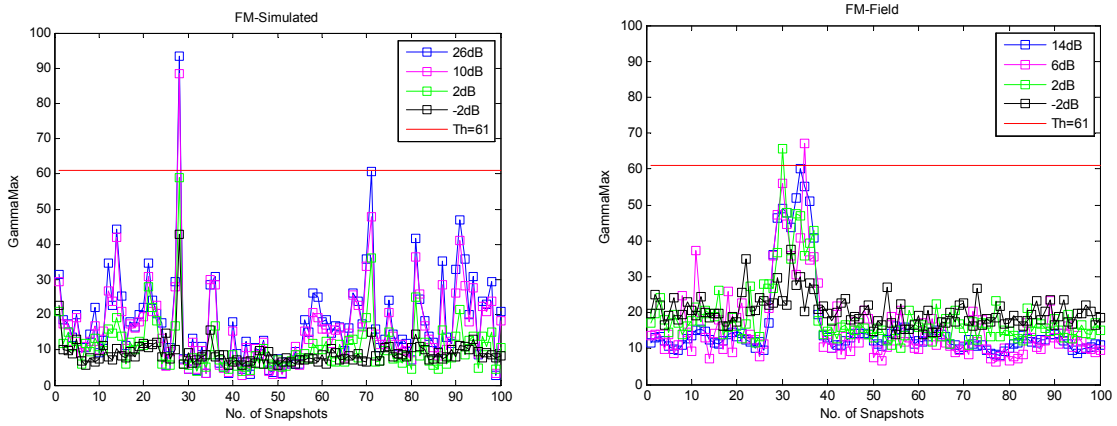


Figure 2.9 Measured values of $\gamma_{\max}$ for FM signals under different SNRs (left column: simulated data, right column: field data), $\gamma_{\max} < \tau(\gamma_{\max})$

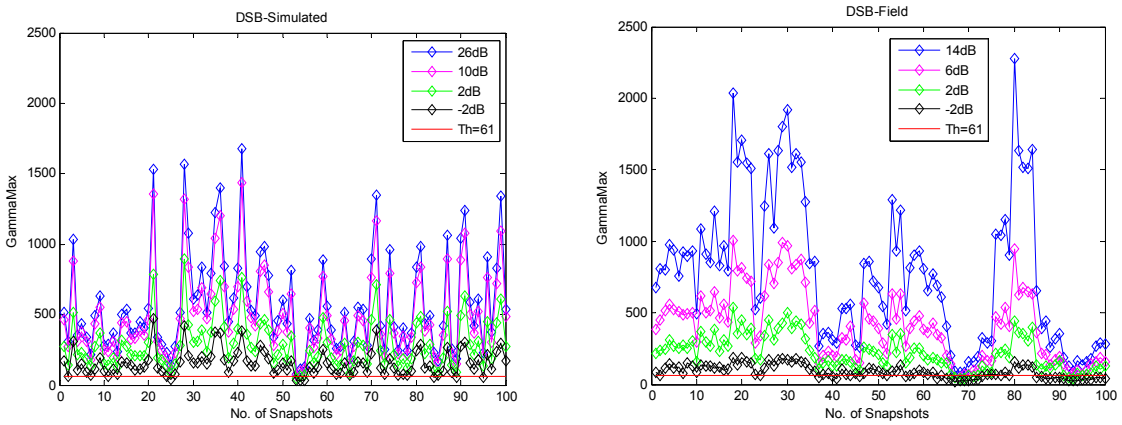


Figure 2.10 measured values of $\gamma_{\max}$ for DSB signals under different SNRs (left Column: simulated data, right column: field data), $\gamma_{\max} \geq \tau(\gamma_{\max})$

2.4 Classification Results and Discussions

The classification results of the procedure in Figure 2.1 are derived from 100 snapshots for each modulated signal using the pre-chosen threshold values of key features. Table 2.1 shows the classification results at different SNRs for simulated data, and Table 2.2 shows the classification results at different SNRs for field data. The classification results shows the validity of the key features developed in [4]

Table 2.1 Probabilities of Correct Classification for Simulated Data

	SNR (dB)				
	26	**10**	**6**	**2**	**-2**
AM	1.00	1.00	1.00	0.86	0.23
DSB	0.97	0.93	0.92	0.96	0.91
LSB	0.95	0.95	0.95	0.92	0.39
USB	0.95	0.95	0.95	0.89	0.36
FM	0.93	0.97	0.97	1.00	1.00

Table 2.2 Probabilities of Correct Classification for Field Data

	SNR (dB)				
	14	**10**	**6**	**2**	**-2**
AM	1.00	1.00	0.99	0.57	0.02
DSB	0.92	0.95	0.96	0.90	0.69
LSB	0.90	0.90	0.79	0.47	0.00
USB	0.80	0.80	0.75	0.55	0.00
FM	1.00	1.00	0.99	0.99	1.00

The works by Azzouz and Nandi in [4, 9-11] have made great progress to the automatic modulation classification of communication signals. Their defined features are meaningful and easy to implement. Simulations have shown the validity of the features when the SNR is high (e.g., SNR $\geq$ 6 dB).

However, these features may encounter some problems in practice. The classifier described in [4, 9-11] is a Decision Tree Classifier (DTC). As such, it suffers from an unavoidable issue in how to decide the best decision threshold values for the features. Since the features extracted from the signals are SNR-dependent, it is very difficult to find an optimal decision threshold in a large range of SNRs. The performance of the classifier by selecting feature threshold values is maximized such that the optimum probability of correct decisions was obtained from the 100 realizations for each modulation type of interest.

CHAPTER 3

NON-DATA-AIDED SNR ESTIMATION TECHNIQUES FOR ANALOG MODULATION SIGNALS

3.1 Introduction

Many techniques for optimal usage of radio resources in software and cognitive radio design are based on the knowledge of the signal-to-noise ratio (SNR). This includes performance improvement of power control, adaptive modulation and demodulation, blind signal classification, and the suppression of interfering signals.

Depending on the amount of knowledge available for the received signal, the SNR estimators can be typically categorized as either data-aided (DA) estimators, i.e. known transmitted data (or pilots) are employed to facilitate the estimation process, or non-data-aided (NDA) estimators that derive SNR estimation solely from the unknown, information-bearing portion of the received signals. Although the DA estimators might be expected to have better performance, they suffer from the fact that periodic insertion of a training sequence reduces the throughput of the system. Aside from the knowledge of transmitted data themselves, other factors that potentially affect the performance of SNR estimators are: 1) modulation type and order and 2) knowledge available for the carrier frequency and phase [12]. In this section, the technology of NDA SNR estimator to eliminate the requirement for this additional knowledge is investigated.

Various SNR estimation techniques have been developed during the last few decades. Pauluzzi and Beaulieu [13] provided an excellent overview of useful algorithms including: the Split-Symbol Moment Estimator (SSME) [14], which is mainly based on the fact that the signals of interest are correlated in the same symbol interval while the

noises are uncorrelated; Maximum-Likelihood (ML) SNR Estimator [15], which converts the problem of SNR estimation into the classical ML estimation of a signal's amplitude; Squared Signal-to-Noise Variance (SNV) Estimator [16], which is based on the first absolute moment and the second moment of the sampled output of the matched filter; and also the well-known M_2M_4 Estimator [17], which employs the second and forth-order moments of the noisy signal. However, all four SNR estimators (i.e., SSME, ML, SNV and M_2M_4), and the corresponding improved algorithms rely on the knowledge of modulation type, which is by definition not applicable with respect to non-cooperative communications where there is no handshaking between the transmitters and receivers. Recently an eigenvalue decomposition (ED) based estimator [18, 19] has been proposed that doesn't require a priori knowledge of modulation type and parameters such as baud-rate, and carrier frequency.

Although many SNR estimators have been reported in the literature, most of them are developed specifically for digital communication systems. The topic of SNR estimation for analog modulation systems is rarely discussed in the open literature. As the identification and classification of analog modulation signals is a desired capability for cognitive radio applications, we have investigated those algorithms previously reported in the literature and found that among those digital SNR estimators, two of them can be extended to analog modulation signals. One is the M_2M_4 estimator, which was originally developed for M-array Phase-shift keying (MPSK) signals and can be used for signals of constant-envelope modulation, such as phase modulation (PM), and frequency modulation (FM). The other is the ED-based estimator, which was originally developed for intermediate frequency (IF) signals, with modulations such as MPSK, M-array

Frequency-shift keying (MFSK) and M-array Quadrature amplitude modulation (MQAM). Since the ED-based estimator does not assume the amplitude of the received signals to be constant, it can be used for not only PM and FM signals but also for linear modulation schemes, such as Amplitude Modulation (AM), Double Sideband Modulation (DSB), Single Sideband Modulation (SSB). Here we show the validity of these two methods in the area of SNR estimation for analog modulation signals. This dissertation also described a Spectrum-based method for blind SNR estimation for analog modulation signals, which is simple and computationally inexpensive for practical applications.

3.2 Problem Formulation

The received signal can be expressed as

$$r(t) = s(t) + n(t) \tag{3.1}$$

where $s(t)$ represents the modulated signal, and $n(t)$ is additive white Gaussian noise (AWGN) with zero-mean and variance of σ^2. For analog modulation types, the signal $s(t)$ can be expressed as below [8], respectively,

AM:

$$s(t) = A\left(1 + \mu m(t)\right) \cdot e^{j\omega_c t} \tag{3.2}$$

DSB:

$$s(t) = A\, m(t) \cdot e^{j\omega_c t} \tag{3.3}$$

LSB:

$$x(t) = A\left(m(t) - j\hat{m}(t)\right) \cdot e^{j\omega_c t} \tag{3.4}$$

USB:

$$x(t) = A\left(m(t) + j\hat{m}(t)\right)\cdot e^{j\omega_c t} \tag{3.5}$$

FM:

$$s(t) = A\cdot e^{j\omega_c t + j2\pi f_\Delta \int_{-\infty}^{t} m(v)dv} \tag{3.6}$$

where A is a positive factor used to control the signal power, ω_c is the carrier frequency, $m(t)$ represents the real-valued information-bearing signal that satisfies $-1 \le m(t) \le 1$, $\hat{m}(t)$ stands for the Hilbert transform of $s(t)$, f_Δ is the frequency deviation of an FM signal, μ is the modulation index of an AM signal, and $0 < \mu \le 1$.

Before digitization, the IF signal $r(t)$ was low-pass filtered and down-converted toward 0 Hz. The down-converted signal can be expressed as

$$x(t) = r(t)e^{-j\hat{\omega}t} \tag{3.7}$$

where $\hat{\omega}$ is the estimated carrier frequency.

Suppose the average signal power and the noise power of $x(t)$ are P_s and P_n, then the SNR can be expressed as

$$\rho \equiv 10\log_{10}\left(P_s / P_n\right) \quad \text{dB} \tag{3.8}$$

3.3 SNR Estimator

This section begins by introducing the M_2M_4 and ED-based estimators and their extensions to analog modulation signals. Then present a Spectrum-based method for blind SNR estimation is presented.

3.3.1 M_2M_4 SNR Estimator for PM and FM signals

Before extending the algorithm to analog modulated signals, the derivation provided in [17] is reviewed. In original algorithm, x is the complex baseband signal at the output of a synchronous demodulator and whitened matched filter. When extending the algorithm to analog modulated signals, x is the discrete-time down-converted complex signal as defined in (3.7).

3.3.1.1 Basic idea of M_2M_4 Estimator The 2nd and 4th order moments of $x(t)$ are:

$$\begin{cases} E_2 = \varepsilon\{xx^*\} = P_s + P_n \\ E_4 = \varepsilon\{(xx^*)^2\} = \varepsilon\{(ss^*)^2\} + \varepsilon\{(nn^*)^2\} + 4P_sP_n \end{cases} \tag{3.9}$$

where $P_s = \varepsilon\{ss^*\}$, and $P_n = \varepsilon\{nn^*\}$. $\varepsilon\{\cdot\}$ denotes the expectation operator.

The kurtosis of $s(t)$ and $n(t)$ are defined as follow:

$$k_s = \frac{\varepsilon\{|s|^4\}}{(\varepsilon\{|s|^2\})^2}, \quad k_n = \frac{\varepsilon\{|n|^4\}}{(\varepsilon\{|n|^2\})^2} \tag{3.10}$$

Therefore, (3.9) can be rewritten as:

$$\begin{cases} E_2 = P_s + P_n \\ E_4 = k_sP_s^2 + k_nP_n^2 + 4P_sP_n \end{cases} \tag{3.11}$$

with k_s and k_n constants depending only on fundamental properties of the probability density functions of the signal and noise processes. The average energy P_s of the discrete-time signal and the average noise energy P_n can then be computed by solving equation (3.11):

$$\begin{cases} \hat{P}_s = \dfrac{E_2(2k_n - 4) \pm \sqrt{D\{x\}}}{2(k_s + k_n - 4)} \\ \hat{P}_n = E_2 - P_s \end{cases} \tag{3.12}$$

where $D\{x\} = E_2^{\,2}(16 - 4k_s k_n) + 4E_4(k_s + k_n - 4)$.

3.3.1.2 Implementation Consideration

- The modulation type must be known in order to evaluate the required constant k_s. It can be proved that PM and FM signals has a kurtosis $k_s = 1.5$. For other analog modulation schemes, k_s is not a constant, and it varies with modulating message and modulation parameters.
- Complex Gaussian noise has a kurtosis as $k_n = 2$.
- s and n are assumed to be ergodic, hence E_2 and E_4 are determined by time-averaging the second and fourth power of the samples x.

3.3.2 ED-based SNR Estimator

The general approach of [18] is outlined and adapted below to analog modulation signals. In this dissertation research, the correlation matrix of the sampled signals is builded specifically using a dimension equal to the number of discrete samples.

3.3.2.1 Basic idea of ED-based SNR Estimator Define $\underline{x}$, $\underline{s}$, $\underline{n}$ as column vectors composed of N discrete samples. The autocorrelation matrix of the received signal $\underline{x}$ can be represented by

$$\mathbf{R}_{\mathbf{x}} = \varepsilon\{\underline{x}\underline{x}^H\} = \varepsilon\{\underline{s}\underline{s}^H\} + \varepsilon\{\underline{n}\underline{n}^H\} = \mathbf{R}_s + \mathbf{R}_{\mathbf{n}} \tag{3.13}$$

where subscript H denotes the Hermitian transposition. Based on the properties of autocorrelation, $\mathbf{R_x}$ is conjugate and symmetrical with its elements satisfying

$$[\mathbf{R_x}]_{ij} = \varepsilon\{x(i)x^*(j)\} = \begin{cases} r_x(i-j), & i \geq j \\ r_x^*(j-i), & j \geq i \end{cases} \tag{3.14}$$

where $r_x(k), \quad k = 0,1,\cdots,N-1$ represents the samples of the autocorrelation function of the received signal, i.e.,

$$r_x(k) = \frac{1}{N}\sum_{i=1}^{N} x(i)x^*(i+k), \quad k = 0,1,\cdots,N-1 \tag{3.15}$$

According to the eigenvalue decomposition of the matrix, we can obtain

$$\mathbf{R_x} = \mathbf{UPU}^H, \quad \mathbf{R}_s = \mathbf{UQU}^H \tag{3.16}$$

where $\mathbf{U}$ is an $N \times N$ orthogonal matrix containing the corresponding eigenvectors. $\mathbf{P}$ and $\mathbf{Q}$ are $N \times N$ diagonal matrices, specifically, $\mathbf{P} = diag(b_1, b_2, \cdots, b_N)$, $\mathbf{Q} = diag(\lambda_1, \lambda_2, \cdots, \lambda_p, 0, \cdots 0)$. The corresponding eigenvalues satisfy $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$, where $p < N$. The correlation matrix of the noise $\mathbf{R_n}$ is

$$\mathbf{R_n} = \sigma_n^{\,2}\mathbf{I} \tag{3.17}$$

where $\mathbf{I}$ is an $N \times N$ identity matrix. Then we have

$$\mathbf{R_x} = \mathbf{UPU}^H = \mathbf{UQU}^H + \sigma_n^{\,2}\mathbf{I} \tag{3.18}$$

and the eigenvalues are in the form of

$$b_i = \begin{cases} \lambda_i + \sigma_n^{\,2} & i = 1,\cdots p \\ \sigma_n^{\,2} & i = p+1,\cdots,N \end{cases} \tag{3.19}$$

in which λ_i represents the power of the desired signal along the i^{th} eigenvector. From equations (3.18) and (3.19) we can see that the total space is composed of a signal

subspace spanned by the first p eigenvectors and a noise subspace spanned by the remaining $N - p$ eigenvectors.

A Minimum Description Length (MDL) [20] criteria is employed to determine the signal subspace dimension. The MDL function is defined as follows,

$$MDL(k) = (N-k)N \log\left(\frac{\frac{1}{N-k}\sum_{i=k+1}^{N} b_i}{\prod_{i=k+1}^{N} b_i^{1/(N-k)}}\right) + \frac{1}{2}k(2N-k)\log(N) \tag{3.20}$$

Then, the signal subspace dimension is estimated as

$$p = \arg\min_{k} MDL(k) \tag{3.21}$$

3.3.2.2 Implementation Consideration This estimator can be implemented by following steps:

- Computer the autocorrelation matrix $\mathbf{R}_{\mathbf{x}}$ of the received signals based on (3.15);
- Make an eigenvalue decomposition of $\mathbf{R}_{\mathbf{x}}$ and mark the eigenvalues as $b_1 \geq b_2 \geq \cdots \geq b_N$;
- Estimate the signal subspace dimension p by (3.21);
- Compute noise variance by: $\hat{\sigma}_n^{\,2} = \frac{1}{N-p}\sum_{i=p+1}^{N} b_i$;
- Estimate signal power by : $\hat{P}_s = \sum_{i=1}^{p}(b_i - \sigma_n^{\,2})$;
- Estimate SNR by: $\hat{\rho} = 10\log_{10}\left(\frac{\hat{P}_s}{\hat{\sigma}_n^{\,2} \cdot N}\right)$.

3.3.3 Spectrum-based SNR Estimator

This estimator is based on the envelope of the Fourier spectrum of the received noisy signal $x(t)$. Similar to the ED-based estimator, Spectrum-based estimator does not rely on the modulation scheme of the received signal. Thus it can be applied to all analog modulation schemes. The Spectrum-based SNR estimation algorithm is summarized as follows, and illustrated in the Figure 3.1.

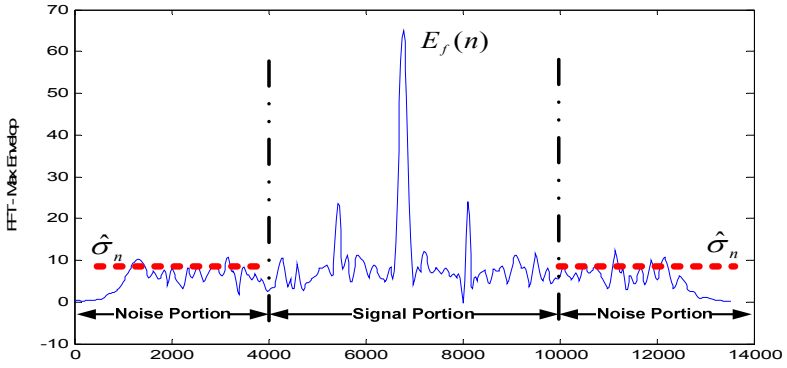


Figure 3.1 A schematic diagram of spectrum-based SNR estimation

- Compute N-point FFT of received noisy signal $x(t)$, where N is the length of the discrete-time signal $x(t)$. Denote the FFT spectrum as $F(n), \quad n = 1, 2, \ldots, N$;
- Extract the envelope of $F(n)$, and denote the envelope as $E_f(n)$, which is shown as a blue solid line in Figure 3.1;
- Based on $F(n)$, an approximate range of frequencies occupied by the signal plus noise can be determined. The remaining portion of the Fourier spectrum is assumed to be the noise only;

- Calculate the mean value of $E_f(n)$ within the noise portion. This is the estimated amplitude level of noise $\hat{\sigma}_n$ as shown by a red dashed line in Figure 3.1. Then the estimated noise power is given by $\hat{P}_n = \hat{\sigma}_n^2 \cdot N$;
- The *total signal-plus-noise power* can be calculated from the FFT spectrum $F(n)$ based on Parseval's theorem, and is denoted as P_{s+n} . So the signal power is given by $\hat{P}_s = P_{s+n} - \hat{P}_n$.

3.4 Performance Comparison

This section is concerned with a comparison of the above three SNR estimators. First, the estimation bias and the corresponding standard deviation (STD) are defined as

$$bias\{\hat{\rho}\} = \frac{1}{M_t}\sum_{t=1}^{M_t}(\hat{\rho}_t - \rho) \quad \text{dB} \tag{3.22}$$

$$STD\{\hat{\rho}\} = \left(\frac{1}{M_t}\sum_{t=1}^{M_t}(\hat{\rho}_t - \overline{\rho})^2\right)^{\frac{1}{2}} \quad \text{dB} \tag{3.23}$$

where $\overline{\rho} = \frac{1}{M_t}\sum_{t=1}^{M_t}\hat{\rho}_t$, $\hat{\rho}_t$ represents the estimated SNR in the t^{th} trial, ρ is the true SNR, and M_t is the number of independent trials.

In order to analyze where the estimation errors come from, we further define the normalized-STD (NSTD) of estimated signal power and estimated noise power, respectively, as

$$NSTD\{\hat{P}_s\} = \frac{\left(\frac{1}{M_t}\sum_{t=1}^{M_t}(\hat{P}_{st} - \overline{P}_s)^2\right)^{\frac{1}{2}}}{\ddot{P}_s} \tag{3.24}$$

$$NSTD\{\hat{P}_n\} = \frac{\left(\frac{1}{M_t}\sum_{t=1}^{M_t}(\hat{P}_{nt} - \bar{P}_n)^2\right)^{\frac{1}{2}}}{\ddot{P}_n} \tag{3.25}$$

The definition of symbols in equations (3.24) and (3.25) are similar with those in equations (3.22) and (3.23). And $\ddot{P}_s = \frac{1}{M_t}\sum_{t=1}^{M_t} P_{st}$, P_{st} represents the true signal power in the t^{th} trial. Similar definition for $\ddot{P}_n$. For the convenient of analysis and comparison, in our experiment and simulation, the true signal power is always normalized to unit.

Three hundred segments of simulated signals for each modulation type are generated in the MATLAB environment and passed through an AWGN channel. The sampling frequency is chosen to be $f_s = 10.5$ kHz, which is slightly larger than twice the occupied bandwidth of most narrowband communications.

For a fixed data length of 100 msec, the estimated $bias\{\hat{\rho}\}$ and $STD\{\hat{\rho}\}$ vs. the true SNRs are plotted in Figure 3.2 and Figure 3.3, respectively. The estimated normalized-STD (NSTD) of signal power $NSTD\{\hat{P}_s\}$ and noise power $NSTD\{\hat{P}_n\}$ vs. the true SNRs are plotted in Figure 3.4 and Figure 3.5, respectively.

From Figure 3.2, it can be seen that the $bias\{\hat{\rho}\}$ for M_2M_4 and ED-based estimators is below zero, which means that these two estimators always underestimate. M_2M_4 estimator has the best performance for FM and PM signals, in the sense that the absolute value of $bias\{\hat{\rho}\}$ remains within 0.5dB, in a wide SNR range, from -1dB to 24 dB. Furthermore, as a blind SNR estimator, Spectrum-based estimator outperforms ED-based estimator since the modulation type has less inference on $bias\{\hat{\rho}\}$.

Figure 3.3 shows that when the SNR is high, the $STD\{\hat{\rho}\}$ increases. This is because when the SNR is very high, the power spectral density of noise is rather small, and even approaches to zero. Therefore a very tiny estimated error, which may be neglected in the case of low SNR, can cause a large deviation for the case of high SNR. This phenomenon is also observed in Figures 3.4 and 3.5. The normalized-STD of estimated noise power increases dramatically, when SNR is high. While normalized-STD of estimated signal power always retains in a low level, i.e. within 0.2.

The influence of the data length on the $STD\{\hat{\rho}\}$ for a moderate SNR condition (SNR=9dB) is show in Figure 3.6. In general, short data length will result in performance degradation. But as shown in Figure 3.6 all three SNR estimator can work with a rather short data length (less than 100 ms) with acceptable estimation $STD\{\hat{\rho}\}$.

We also conducted experiments with different sampling rate, modulation message and modulation parameters. The experiments show that the performance of the three estimators is nearly independent of those factors. As to the computational cost, M_2M_4 and Spectrum-based estimators are of low computational complexity and can be used in real time SNR estimation.

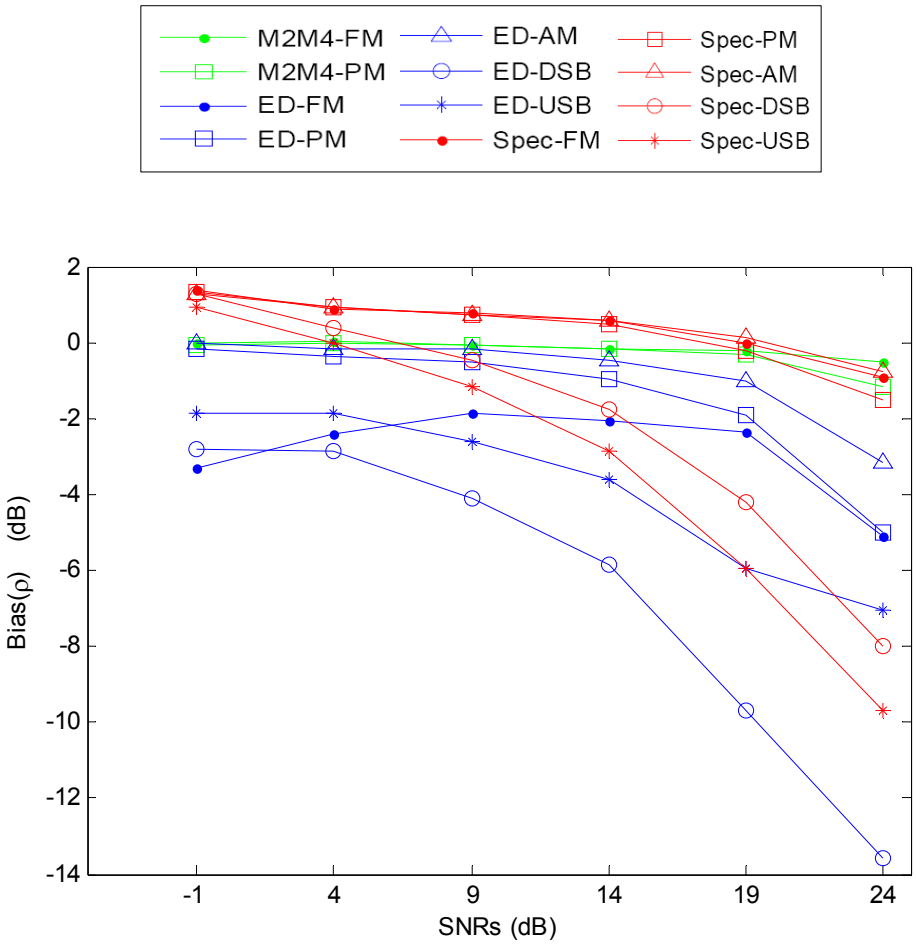


Figure 3.2 Estimated SNR-Bias vs. True-SNRs

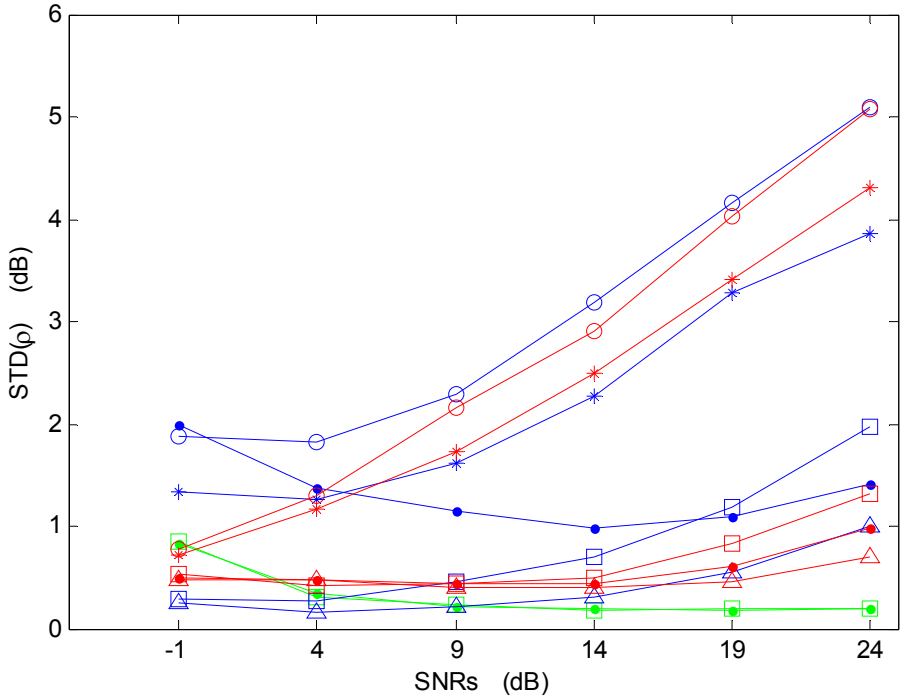


Figure 3.3 Estimated SNR-STD vs. True-SNRs

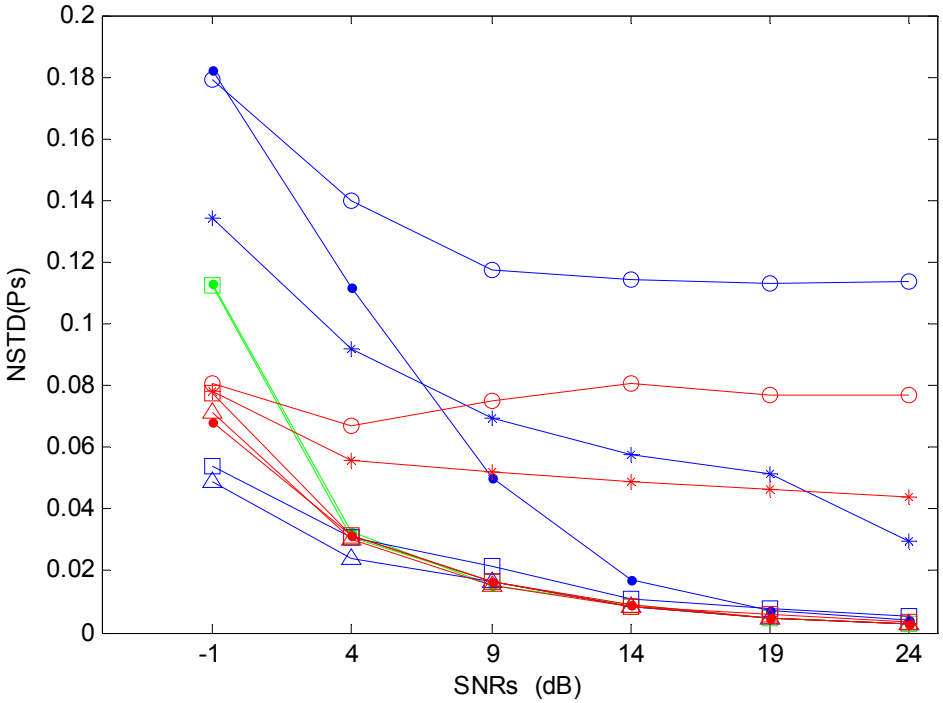


Figure 3.4 Estimated Signal-Power-NSTD vs. True-SNRs

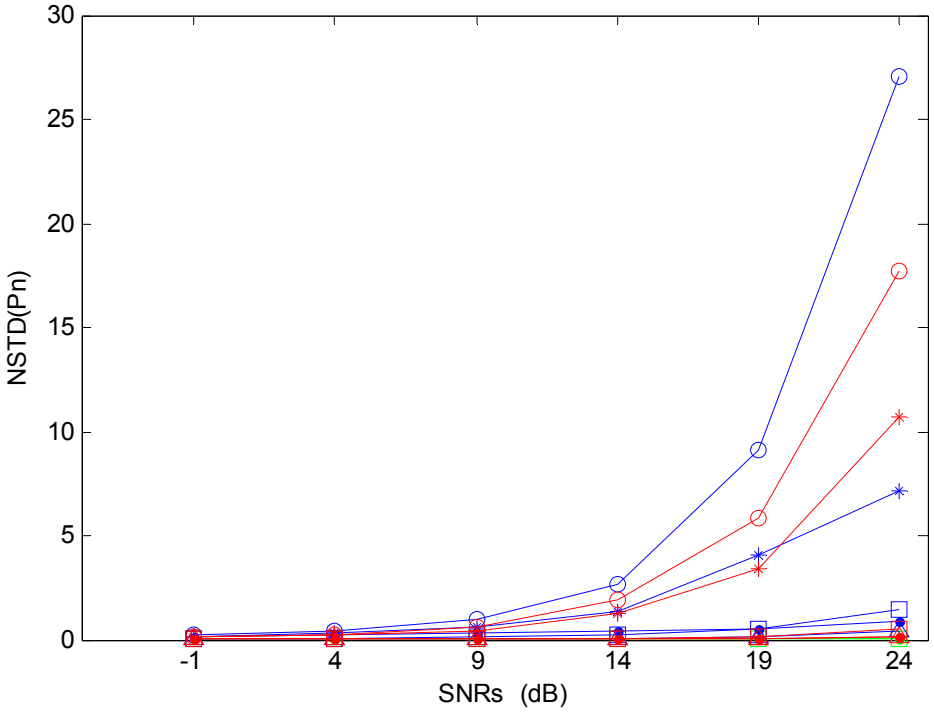


Figure 3.5 Estimated Noise-Power-NSTD vs. True-SNRs

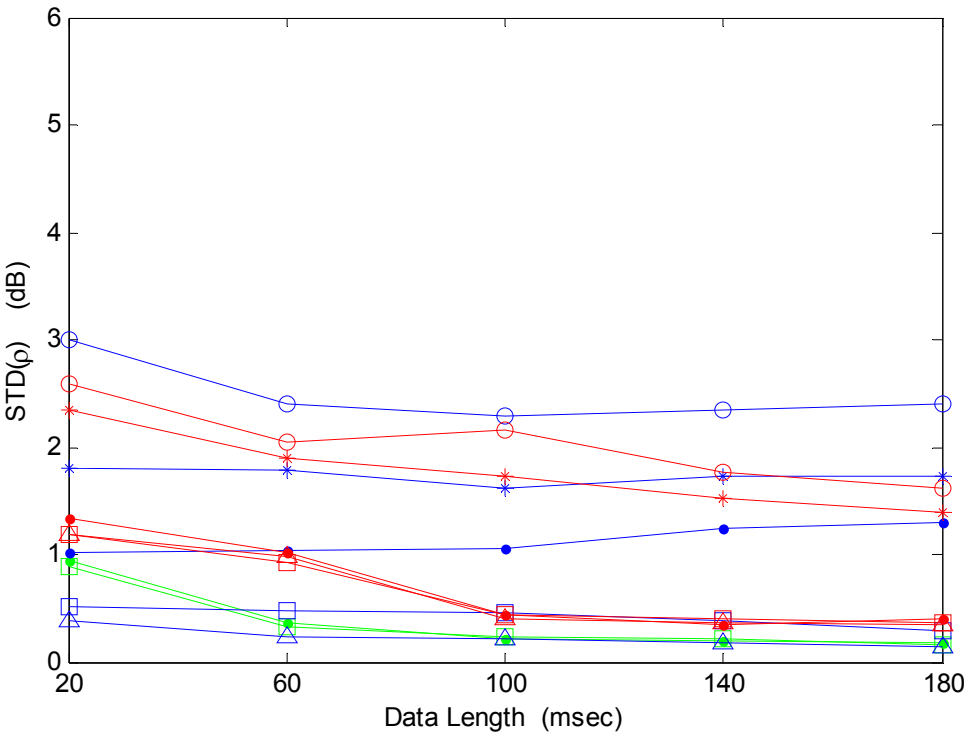


Figure 3.6 Estimated SNR-STD vs. Data length, at True-SNR = 9dB

CHAPTER 4

CYCLOSTATIONARITY-BASED DECISION TREE FOR JOINT CLASSIFICATION OF DIGITAL AND ANALOG MODULATIONS

4.1 Introduction

Automatic modulation classification (AMC) plays an important role in many fields such as spectrum monitoring, software-defined radio, cognitive radio networks, and interference identification. It serves as an intermediate step between signal detection and data demodulation, and is a major task of a smart receiver. Obviously, with a limited *prior* knowledge of the transmitted data and modulation parameters at the receiver, such as the signal power, carrier frequency and phase offsets, symbol rate, pulse shape, timing information, etc., blind identification of the modulation type is a difficult task. This becomes even more challenging in real-world scenarios with multipath fading, frequency-selective and time-varying channels.

Researchers in the communications and signal processing community have conducted extensive research in AMC in the past decades. A large number of AMC algorithms have been reported in the literature, and they have been generally grouped into two broad categories, the likelihood-based (LB) methods (e.g., [21-28]), and the feature-based (FB) methods (e.g.,. [11, 29-37]), respectively. The former computes the likelihood function of the received signal and the decision is made by comparing the likelihood ratio against a threshold. For FB approaches, features are extracted and a decision is made based on the observed values of features. A recent survey paper by Dobre et al. [38] provides a comprehensive overview of different modulation recognition techniques, which also follows the above mentioned categorization. In [38] , the advantages and

drawbacks of these two categories have been highlighted. It is clear that neither LB methods nor FB methods can effectively classify all modulation types. The choice of a particular AMC greatly depends on the practical scenarios [39]. The algorithm proposed in this chapter falls under the category of feature-based methods.

A feature-based AMC system usually consists of two subsystems: a feature extraction subsystem and a pattern classification subsystem. Signal features such as spectral features, instantaneous amplitude, phase and frequency, higher order statistics (moments, cumulants or kurtosis), wavelet transform, in-phase and quadrature constellation, and/or zero-crossings are extracted from the observed signals. Then classification methods such as decision trees, artificial neural networks (ANNs), or support vector machines (SVMs) can be used to distinguish among the various modulation types.

The AMC algorithm presented in this chapter is based on the work of [40]. Cyclostationarity is a hidden periodicity of communication signals, and the cyclic statistics of different modulation schemes have different patterns that can be used as a discriminator in the classification process. The cyclostationary features are augmented by five additional features derived from the signal spectra that are used to further differentiate among the modulation types for analog modulated signals. A straightforward and computationally simple decision tree classifier based on the derived features is exploited to automatically classify the joint analog and digital modulation signals.

This AMC algorithm is developed for joint classification of analog and digital modulations, and is applied to the following pool of modulations: Amplitude Modulation (AM), Double Sideband Modulation (DSB), Upper Sideband Modulation (USB), Lower

Sideband Modulation (LSB), Narrow-Band Frequency Modulation (NBFM), Wide-Band Frequency Modulation (WBFM), *M*-ary Amplitude-Shift Keying (MASK), *M*-ary Pulse Amplitude Modulation (MPAM), *M*-ary Phase-Shift Keying (MPSK), and Quadrature Amplitude Modulation (QAM).

The rest of the chapter is organized as follows. Section 4.2 formulates the problem and establishes the working conditions. Section 4.3 explores the cyclostationarities of communication signals under some selected cyclic moments. Section 4.4 presents the schemes for detecting the cycle frequencies of the selected cyclic moments. Section 4.5 proposes five additional features for modulation classification. Section 4.6 presents the proposed decision-tree classifier. Simulations and numerical experiments are described in Section 4.7 along with the classification results and performance of the modulation recognition algorithm. Finally, we present the conclusions and discuss opportunities for further research in Section 4.8.

4.2 Problem Statement and Assumptions

The received signal $y(t)$ can be expressed as

$$y(t) = x(t) + w(t), \tag{4.1}$$

where $x(t)$ represents the analytic signal of the transmitted communication signal, and $w(t)$ is the additive complex noise.

For analog modulation types, the signal $x(t)$ can be expressed as below [8], respectively,

AM:

$$x(t) = A\left(1 + K_a s(t)\right) \cdot e^{j2\pi f_c t}, \tag{4.2}$$

DSB:

$$x(t) = A\, s(t) \cdot e^{j2\pi f_c t}, \tag{4.3}$$

LSB:

$$x(t) = A\left(s(t) - j\hat{s}(t)\right) \cdot e^{j2\pi f_c t}, \tag{4.4}$$

USB:

$$x(t) = A\left(s(t) + j\hat{s}(t)\right) \cdot e^{j2\pi f_c t}, \tag{4.5}$$

FM:

$$x(t) = A \cdot e^{j2\pi f_c t + j2\pi f_\Delta \int_{-\infty}^{t} s(v)dv}, \tag{4.6}$$

where A is a positive factor used to control the signal power, f_c is the carrier frequency, $s(t)$ represents the real-valued information-bearing signal that satisfies $-1 \le s(t) \le 1$, $\hat{s}(t)$ stands for the Hilbert transform of $s(t)$, f_Δ is the frequency deviation of an FM signal, K_a is the modulation index of an AM signal, and $0 < K_a \le 1$.

For linear digital modulation types, the signal $x(t)$ can be expressed in a common form [41] as

$$x(t) = A \sum_l s_l\, g(t - lT) \cdot e^{j2\pi f_c t}, \tag{4.7}$$

where A is also used to control the signal power, $g(t)$ stands for the pulse shaping function, f_c is the carrier frequency, T is the symbol period and its reciprocal is the symbol rate R_s, and s_l is the *l-th* transmitted symbol. For different modulation types s_l will respectively take the following discrete values,

MASK :

$$s_l \in \{ I_m = m-1, \quad m = 1,2,\cdots,M \} \tag{4.8}$$

MPAM :

$$s_l \in \{ I_m = 2m-1-M, \quad m = 1,2,\cdots,M \} \tag{4.9}$$

MPSK :

$$s_l \in \left\{ I_m = e^{j\frac{2\pi(m-1)}{M}}, \quad m = 1,2,\cdots,M \right\} \tag{4.10}$$

QAM :

$$s_l \in \{ I_m = I_{I,m} + jI_{Q,m}, \quad m = 1,2,\cdots,M \} \tag{4.11}$$

where M is the number of distinct symbols in the symbol constellation (i.e. the alphabet size), and $I_{I,m}$ and $I_{Q,m}$ are respectively the in-phase component and the quadrature component of the *m-th* symbol in the QAM constellation.

Without loss of generality, the following conditions are assumed in the rest of this chapter:

- $x(t)$ and $w(t)$ are independent of each other.
- $w(t)$ is complex white Gaussian noise with zero mean.
- For analog modulations, the modulating signal $s(t)$ is ergodic and wide sense stationary (WSS), and it is zero-mean with unknown probability density function (PDF).
- For digital modulations, the transmitted symbols compose a set of independent and identically distributed (i.i.d) discrete random variables, the elements of

which are uniformly distributed on the symbol constellation of the concrete modulation type.

- The symbols of a QAM signal are zero-mean and symmetrically distributed, i.e., $E[\, s_l \,] = 0$, $E[\, s_l^2 \,] = 0$, where $E[\cdot]$ stands for the expectation operation.
- For digital modulations, the pulse shape is unknown, but its bandwidth will be greater than one half of the symbol rate;
- The proposed classifier does not have a prior knowledge of A, K_a, f_Δ, M, R_s and f_c.

It is also assumed that a rough range of the carrier frequency of the received signal has been estimated. Then the received signal is sampled at a rate higher than Nyquist frequency. However, the sampling may be neither synchronized nor coherent. That is, the sampling may start at any instant, and the sampling rate f_s, may be neither an integer multiple of the symbol rate R_s nor that of the carrier frequency f_c.

The carrier phase offset is assumed to be deterministic but unknown. Since the proposed classifier is based on the detection of the pattern of the cyclostationarity of the received signal as well as other features that are derived from the received signal's power spectral density (PSD) function, the fixed carrier phase offset is not relevant. Thus it is assumed being zero in the above signal models without loss of generality.

The proposed modulation classification algorithm will work with a sample sequence of the input signal to carry out the modulation classification. The input data can be represented by $y(n) \triangleq y(nT_s)$, $n = 0, 1, \ldots, N-1$, where N is the number of available data samples, and T_s is the sampling period that is the reciprocal of the sampling rate f_s.

4.3 Cyclostationarities of Communication Signals

Communication signals encountered in practice commonly have a hidden periodicity due to sampling, scanning, modulation, multiplexing and coding. In conventional receivers, these periodicities are usually not explored for extracting information or parameters. However, the performance of signal processing can be improved in many cases by considering these hidden periodicities. For example, in [42, 43] and [44] respectively, Dandawate and Gardner use cyclostationarity for various purposes.

In order to explore these hidden periodicities, the underlying signal should be modeled as cyclostationary rather than wide-sense stationary. By definition, a cyclostationary process is a random process with probabilistic parameters that vary periodically with time [45, 46]. Therefore, cyclostationary processes are appropriate probabilistic models for signals that have periodically time-varying characteristics. Consistent with their basic nature, cyclic statistics are useful tools for characterizing and analyzing cyclostationary signals.

For a real-valued discrete-time input $\{ y(n) \}$, the *k-th order time-varying moment*, $m_{ky}(n;\boldsymbol{\tau})$ is defined as

$$m_{ky}(n;\boldsymbol{\tau}) \triangleq E\left[f_{ky}(n;\boldsymbol{\tau}) \right], \tag{4.12}$$

where

$$f_{ky}(n;\boldsymbol{\tau}) \triangleq y(n)\, y(n+\tau_1) \cdots y(n+\tau_{k-1}), \tag{4.13}$$

where $\boldsymbol{\tau}$ is a vector whose entries are the lags of $m_{ky}(n;\boldsymbol{\tau})$, i.e., $\boldsymbol{\tau} = [\, \tau_0, \cdots, \tau_{k-1} \,]$, and τ_0 is fixed as $\tau_0 = 0$. $E[\,\cdot\,]$ stands for the expectation operation.

If $m_{ky}(n;\boldsymbol{\tau})$ is a periodic or almost periodic function of n, the process $\{\, y(n) \,\}$ is called a *k-th order cyclostationary process* and $m_{ky}(n;\boldsymbol{\tau})$ accepts a Fourier Series (FS) decomposition [42, 43] as

$$m_{ky}(n;\boldsymbol{\tau}) = \sum_{\alpha \in \Omega_k} M_{ky}^{\alpha}(\boldsymbol{\tau}) \cdot e^{j2\pi\alpha n}, \tag{4.14}$$

$$\begin{aligned} M_{ky}^{\alpha}(\boldsymbol{\tau}) &\triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} m_{ky}(n;\boldsymbol{\tau}) \cdot e^{-j2\pi\alpha n} \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} E\left[\, f_{ky}(n;\boldsymbol{\tau}) \,\right] \cdot e^{-j2\pi\alpha n} \end{aligned} \tag{4.15}$$

where the FS coefficient $M_{ky}^{\alpha}(\boldsymbol{\tau})$ is called the *k-th order cyclic moment* of $\{\, y(n) \,\}$ at frequency α for lags $\boldsymbol{\tau} = [\, \tau_0, \cdots, \tau_{k-1} \,]$.

The values of α for which $M_{ky}^{\alpha}(\boldsymbol{\tau}) \neq 0$ are called *cycle frequencies of moments*, which are assumed denumerable in number. Ω_k in (4.14) denotes the set of the cycle frequencies of the cyclic moment $M_{ky}^{\alpha}(\boldsymbol{\tau})$. It is evident that $M_{ky}^{\alpha}(\boldsymbol{\tau})$ represents the complex strength of a sinusoidal wave (with frequency being α) contained in $m_{ky}(n;\boldsymbol{\tau})$. However, $M_{ky}^{\alpha}(\boldsymbol{\tau})$ with $\alpha = 0$ represents the strength of the DC component of $m_{ky}(n;\boldsymbol{\tau})$. Therefore, the zero frequency (i.e., $\alpha = 0$) will not be considered as a cycle frequency in the rest of this chapter.

By following the moment-cumulant formula, the *time-varying k-th order cyclic cumulant*, $c_{ky}(n;\boldsymbol{\tau})$, can be expressed as [43]

$$c_{ky}(n;\boldsymbol{\tau}) = \sum_{p=1}^{k} \left[\sum_{\substack{J_1 \cup \cdots \cup J_p = J \\ J_1 \cap \cdots \cap J_p = \varphi}} \left[(-1)^p (p-1)! \prod_{i=1}^{p} m_{v_i y}(n;\boldsymbol{\tau}_{J_i}) \right] \right], \tag{4.16}$$

where φ represents a null set, and J represents the set of index indicators of the entries in the vector $\boldsymbol{\tau} = [\,\tau_0, \cdots, \tau_{k-1}\,]$, i.e., $J = \{\,0, 1, \cdots, k-1\,\}$. For a value of p in the two-folded summations of (4.16), the indicator set J is partitioned into distinct subsets $\{\,J_i : J_i \subseteq J, \quad i = 1, 2, \cdots, p\,\}$. In each partition, the order of the time-varying moment $m_{v_i y}(n; \boldsymbol{\tau}_{J_i})$ equals to the number (denoted by v_i) of indicators in the indicator set J_i, and J_i denotes the lag indices of the vector $\boldsymbol{\tau}_{J_i}$. For example, if the set J_i is $J_i = \{\,1, 3\,\}$, then $m_{v_i y}(n; \boldsymbol{\tau}_{J_i})$ will be $m_{v_i y}(n; \boldsymbol{\tau}_{J_i}) = E[\,y(n+\tau_1)\, y(n+\tau_3)\,]$, $\boldsymbol{\tau}_{J_i}$ will be $\boldsymbol{\tau}_{J_i} = \{\,\tau_1, \tau_3\,\}$, and v_i will be two. It should be noted that, for a given value of p, there may be multiple different ways to partition the set . The inner summation of (4.16) extends over all possible partitions for a give value of p. As an example, the case of $k = 3$ is employed to explain the partition procedure as follows. For $p = 1$, the only partition is $J_1 = J = \{\,0, 1, 2\,\}$; for $p = 2$, there are three ways to partition the set J, i.e., $\{\,J_1 = \{\,0\,\}, J_2 = \{\,1, 2\,\}\,\}$, $\{\,J_1 = \{\,1\,\}, J_2 = \{\,0, 2\,\}\,\}$ and $\{\,J_1 = \{\,2\,\}, J_2 = \{\,0, 1\,\}\,\}$; for $p = 3$, the only partition is $\{\,J_1 = \{\,0\,\}, J_2 = \{\,1\,\}, J_3 = \{\,2\,\}\,\}$. Then (4.16) for $k = 3$ will be evaluated as

$$\begin{aligned} c_{3y}(n; \boldsymbol{\tau}) = {} & E[\,y(n)\, y(n+\tau_1)\, y(n+\tau_2)\,] \\ & - E[\,y(n)\,] E[\,y(n+\tau_1)\, y(n+\tau_2)\,] \\ & - E[\,y(n+\tau_1)\,] E[\,y(n)\, y(n+\tau_2)\,] \\ & - E[\,y(n+\tau_2)\,] E[\,y(n)\, y(n+\tau_1)\,] \\ & + 2E[\,y(n)\,] E[\,y(n+\tau_1)\,] E[\,y(n+\tau_2)\,] \end{aligned} \quad . \tag{4.17}$$

It follows that the *k-th order cyclic cumulant*, $C_{ky}^{\alpha}(\boldsymbol{\tau})$, which is the FS coefficient of $c_{ky}(n; \boldsymbol{\tau})$, is given by [43]

$$C_{ky}^{\alpha}(\boldsymbol{\tau}) = \sum_{p=1}^{k} \left[\sum_{\substack{J_1 \cup \cdots \cup J_p = J \\ J_1 \cap \cdots \cap J_p = \varphi}} \left[(-1)^p (p-1)! \cdot \sum_{\alpha_1, \cdots, \alpha_p} \left[\prod_{i=1}^{p} M_{v_i y}^{\alpha_i}(\boldsymbol{\tau}_{J_i}) \cdot \eta(\alpha - \alpha_1 - \cdots - \alpha_p) \right] \right] \right] \tag{4.18}$$

where $\eta(\alpha)$ is the Kronecker combo (train) function that is nonzero and unity only when $\alpha = 0 \pmod 2$. The values of α for which $C_{ky}^{\alpha}(\boldsymbol{\tau}) \neq 0$ are called cycle frequencies of cumulants. It should be noted that the set of cycle frequencies for $C_{ky}^{\alpha}(\boldsymbol{\tau})$ may be different from that for $M_{ky}^{\alpha}(\boldsymbol{\tau})$.

For a complex-valued input $\{ y(n) \}$, the function $f_{ky}(n; \boldsymbol{\tau})$ in (4.13) may or may not use the conjugate version of $y(n + \tau_i)$ for $i = 0, 1, \cdots, k-1$. This is formally expressed as

$$f_{ky}(n; \boldsymbol{\tau}) \triangleq y^{\{*\}}(n)\, y^{\{*\}}(n + \tau_1) \cdots y^{\{*\}}(n + \tau_{k-1}) \tag{4.19}$$

where the superscript $\{*\}$ indicates an optional complex conjugate operation. Then for a fixed lag vector $\boldsymbol{\tau}$, the *k-th* order cyclic moment can be defined in 2^k different ways.

The order of cyclostationarity and the cycle frequencies are dependent on the selection of $f_{ky}(n; \boldsymbol{\tau})$. For the complex-valued sample sequence $\{ y(n) \}$ of a communication signal, it is well-known that the choice of $f_{ky}(n; \boldsymbol{\tau})$ without conjugation is related to the carrier frequency and that the choice of $f_{ky}(n; \boldsymbol{\tau})$ with $\frac{k}{2}$ conjugations (k is even) is related to the symbol rate [44, 47].

To classify communication signals based on their cyclostationarities, the very first step is to choose a proper set of cyclic statistics. It is well-known that higher-order ($k \geq 3$)

cyclic moments have some important properties that second-order cyclic moments may not have. For example, higher-order cyclic cumulants are not sensitive to Gaussian noises. However, higher-order statistics, in general, leads to higher computational complexity. Hence, it is always desirable to use the smallest possible value of k [44]. In this chapter, we take $k = 2$ and it will be seen later that cyclic statistics of order $k \leq 2$ suffice to complete the AMC task.

Based on the above considerations, the proposed AMC algorithm is designed to employ the first-order cyclic moment M_{1y}^{α}, the un-conjugate second-order cyclic moment $M_{2,0y}^{\alpha}$ and the conjugate second-order cyclic moment $M_{2,1y}^{\alpha}$, which are, respectively, defined as,

$$M_{1y}^{\alpha} \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} E\left[\, y(n) \,\right] \cdot e^{-j2\pi\alpha n} \tag{4.20}$$

$$M_{2,0y}^{\alpha} \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} E\left[\, y^2(n) \,\right] \cdot e^{-j2\pi\alpha n} \tag{4.21}$$

$$M_{2,1y}^{\alpha} \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} E\left[\, y^*(n)\, y(n) \,\right] \cdot e^{-j2\pi\alpha n} \tag{4.22}$$

The cycle frequencies of M_{1y}^{α}, $M_{2,0y}^{\alpha}$ and $M_{2,1y}^{\alpha}$ for the targeted modulation types are summarized in Table 4.1, and the detailed derivations are given in the Appendix I.

As seen in Table 4.1, the targeted modulation types have different patterns of cyclostationarities for the selected three cyclic moments. This can be employed to classify received communication signals.

Table 4.1 Cycle Frequencies of Communication Signals

CYCLIC MODULATION	M_{1y}^{α}	$M_{2,0y}^{\alpha}$	$M_{2,1y}^{\alpha}$
AM	f_c	$2f_c$	φ
DSB	φ	$2f_c$	φ
LSB	φ	φ	φ
USB	φ	φ	φ
NBFM	$f_c^{(\pm)}$	φ	φ
WBFM	φ	φ	φ
MASK	$f_c + l \times R_s$	$2f_c + l \times R_s$	$l \times R_s, l \neq 0$
MPAM, PSK2	φ	$2f_c + l \times R_s$	$l \times R_s, l \neq 0$
MPSK ($M \geq 4$), QAM	φ	φ	$l \times R_s, l \neq 0$

φ means that there is no cycle frequency. f_c is the carrier frequency. R_s is the symbol rate. l is an integer.

$(\pm)$ means that the presence of this cycle frequency depends on the ratio of the peak frequency deviation to the modulating signal's bandwidth. If this ratio is very small and the SNR is high, f_c may be recognized as a cycle frequency of the first-order cyclic moment.

4.4 Examination of the Presence of Cyclostationarity

As seen in Table 4.1, the targeted modulation types have different patterns of cyclostationarities for the selected three cyclic moments. This can be employed to classify received communication signals. Then it is necessary to estimate these cyclic statistics and then test if they have cycle frequencies or not and how many cycle frequencies they have.

The cyclic moment $M_{ky}^{\alpha}(\boldsymbol{\tau})$ can be estimated by [43]

$$\begin{aligned} \hat{M}_{ky}^{\alpha}(\boldsymbol{\tau}) &= \frac{1}{N}\sum_{n=0}^{N-1} f_{ky}(n;\boldsymbol{\tau}) \cdot e^{-j2\pi\alpha n} \\ &= \frac{1}{N}\sum_{n=0}^{N-1} y(n)\, y(n+\tau_1) \cdots y(n+\tau_{k-1}) \cdot e^{-j2\pi\alpha n} \end{aligned} \tag{4.23}$$

This has been shown to be a mean-square sense (m.s.s.) consistent estimator, i.e.,

$$\lim_{N\to\infty} E\left[\left(\hat{M}_{ky}^{\alpha}(\boldsymbol{\tau}) - M_{ky}^{\alpha}(\boldsymbol{\tau})\right)^2\right] = 0 \tag{4.24}$$

In addition, $\sqrt{N}\left(\hat{M}_{ky}^{\alpha}(\boldsymbol{\tau}) - M_{ky}^{\alpha}(\boldsymbol{\tau})\right)$ is asymptotically complex normal with covariance given by

$$\lim_{N\to\infty} N \operatorname{cov}\left\{\hat{M}_{ky}^{\alpha}(\boldsymbol{\tau}),\, \hat{M}_{ky}^{\beta}(\boldsymbol{\rho})\right\} = S_{2f_{\boldsymbol{\tau},\boldsymbol{\rho}}}(\alpha+\beta;\beta) \tag{4.25}$$

$$\lim_{N\to\infty} N \operatorname{cov}\left\{\hat{M}_{ky}^{\alpha}(\boldsymbol{\tau}),\, \hat{M}_{ky}^{\beta}(\boldsymbol{\rho})\right\} = S_{2f_{\boldsymbol{\tau},\boldsymbol{\rho}}}(\alpha+\beta;\beta) \tag{4.26}$$

where $S_{2f_{\boldsymbol{\tau},\boldsymbol{\rho}}}(\alpha+\beta;\beta)$ and $S^{(*)}_{2f_{\boldsymbol{\tau},\boldsymbol{\rho}}}(\alpha-\beta;-\beta)$ are respectively the un-conjugate and conjugate cyclic spectra of $f_{ky}(n;\boldsymbol{\tau})$, and are respectively defined as [43]

$$S_{2f_{\boldsymbol{\tau},\boldsymbol{\rho}}}(\alpha;\beta) \triangleq \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\varsigma=-\infty}^{\infty}\operatorname{cov}\left\{f_{ky}(n;\boldsymbol{\tau}), f_{ky}(n+\varsigma;\boldsymbol{\tau})\right\}\cdot e^{-j\varsigma\beta}\cdot e^{-j\alpha n} \tag{4.27}$$

$$S^{(*)}_{2f_{\boldsymbol{\tau},\boldsymbol{\rho}}}(\alpha;\beta) \triangleq \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\varsigma=-\infty}^{\infty}\operatorname{cov}\left\{f_{ky}(n;\boldsymbol{\tau}), f^{*}_{ky}(n+\varsigma;\boldsymbol{\tau})\right\}\cdot e^{-j\varsigma\beta}\cdot e^{-j\alpha n} \tag{4.28}$$

After the cyclic moments $\left\{M_{ky}^{\alpha}(\boldsymbol{\tau})\right\}$ have been estimated by using equation (4.23), the cyclic cumulants $\left\{C_{ky}^{\alpha}(\boldsymbol{\tau})\right\}$ can be estimated by substituting $\hat{M}_{ky}^{\alpha}(\boldsymbol{\tau})$ into equation (4.18).

Note that $C_{2y}^{\alpha}(\boldsymbol{\tau})$ is equal to $M_{2y}^{\alpha}(\boldsymbol{\tau})$ for zero-mean process, then $\hat{C}_{2y}^{\alpha}(\boldsymbol{\tau})$ is equal to $\hat{M}_{2y}^{\alpha}(\boldsymbol{\tau})$ and it thus will also have the above asymptotic properties. A technique has been developed in [43] for detecting the presence of cyclostationarity as follows. Let $\tau_1,\cdots,\tau_\gamma$ be a fixed set of lags, α be a candidate cycle frequency of $C_{2y}^{\alpha}(\boldsymbol{\tau})$, and

$$\hat{C}_{2y} = \left[\operatorname{Re}\left(\hat{C}_{2y}^{\alpha}(\tau_1)\right),\cdots,\operatorname{Re}\left(\hat{C}_{2y}^{\alpha}(\tau_\gamma)\right),\operatorname{Im}\left(\hat{C}_{2y}^{\alpha}(\tau_1)\right),\cdots,\operatorname{Im}\left(\hat{C}_{2y}^{\alpha}(\tau_\gamma)\right)\right] \tag{4.29}$$

represents a $1\times 2\gamma$ row vector of the estimated second-order cyclic cumulants with $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ representing the real and imaginary parts, respectively. In order to construct the asymptotic covariance matrix of $\hat{\mathbf{C}}_{\mathbf{2y}}$, two $\gamma\times\gamma$ matrices Q_{2c} and Q_{2c}^{*} are constructed with the *(m,n)-th* entries given, respectively, as

$$Q_{2c}(m,n) = S_{2f_{\tau_m,\tau_n}}(2\alpha;\alpha) \tag{4.30}$$

$$Q_{2c}^{*}(m,n) = S_{2f_{\tau_m,\tau_n}}^{(*)}(0;-\alpha) \tag{4.31}$$

where $S_{2f_{\tau_m,\tau_n}}(2\alpha;\alpha)$ and $S_{2f_{\tau_m,\tau_n}}(0;-\alpha)$ are estimated respectively by

$$\hat{S}_{2f_{\tau_m,\tau_n}}(2\alpha;\alpha) = \frac{1}{NL}\sum_{v=-(L-1)/2}^{(L-1)/2} W^{(N)}(v)\times F_{N,\tau_n}\left(\alpha-\frac{2\pi v}{N}\right)\times F_{N,\tau_m}\left(\alpha+\frac{2\pi v}{N}\right) \tag{4.32}$$

$$\hat{S}_{2f_{\tau_m,\tau_n}}^{(*)}(0;-\alpha) = \frac{1}{NL}\sum_{v=-(L-1)/2}^{(L-1)/2} W^{(N)}(v)\times F_{N,\tau_n}^{*}\left(\alpha-\frac{2\pi v}{N}\right)\times F_{N,\tau_m}\left(\alpha+\frac{2\pi v}{N}\right) \tag{4.33}$$

where $F_{N,\tau}(\omega)$ is defined as $F_{N,\tau}(\omega)=\sum_{n=0}^{N-1} f_{2y}(n;\boldsymbol{\tau})\cdot e^{-j2\pi\omega n}$ with $\boldsymbol{\tau}=[0,\tau]$, and $W^{(N)}$ is a spectral window of length L (odd). Then based on equations (4.30) to (4.33), the covariance matrix of $\hat{\mathbf{C}}_{\mathbf{2y}}$ can be computed as

$$\boldsymbol{\Sigma}_{2c} = \begin{bmatrix} \mathrm{Re}\left\{\dfrac{Q_{2c}+Q_{2c}^{(*)}}{2}\right\} & \mathrm{Im}\left\{\dfrac{Q_{2c}-Q_{2c}^{(*)}}{2}\right\} \\ \mathrm{Im}\left\{\dfrac{Q_{2c}+Q_{2c}^{(*)}}{2}\right\} & \mathrm{Re}\left\{\dfrac{Q_{2c}-Q_{2c}^{(*)}}{2}\right\} \end{bmatrix} \tag{4.34}$$

Then the test statistics will be

$$\mathcal{T}_{2c} = N\,\hat{C}_{2y}\,\boldsymbol{\Sigma}_{2c}^{-1}\,\hat{C}_{2y}' \tag{4.35}$$

The asymptotic distribution of $\mathcal{T}_{2c}$ has been shown to be central chi-square distribution with 2γ degrees of freedom if the frequency α is not a cycle frequency [43].

Therefore, a threshold can be selected according to chi-square tables and the required false alarm rate. The frequency α is declared as a cycle frequency if $\mathcal{T}_{2c}$ exceeds the threshold for at least one set of $\tau_1, \cdots, \tau_\gamma$; otherwise, α is determined as not being a cycle frequency for any of $\tau_1, \cdots, \tau_\gamma$.

The selected cyclic moments M_{1y}^{α}, $M_{2,0y}^{\alpha}$ and $M_{2,1y}^{\alpha}$ for the proposed classifier are special cases of $M_{ky}^{\alpha}(\boldsymbol{\tau})$, thus they can be estimated by replacing $f_{ky}(n;\boldsymbol{\tau})$ in equation (4.23) with $y(n)$, $y^2(n)$ and $|y(n)|^2$, respectively. Then a similar scheme can be used to check whether a given frequency α is a cycle frequency of M_{1y}^{α} (or that of $M_{2,0y}^{\alpha}$ or $M_{2,1y}^{\alpha}$).

In determining the presence of cyclostationarity of an input signal, however, the scheme in [43] has to exhaustively search over all possible values of α . The computational burden will be extremely heavy even though only one distinct lag (i.e., $\gamma = 1$) is used.

Through our careful study, it is found that the following properties can be used to reduce the computational complexity in searching the possible cycle frequencies. At first, a cycle frequency of a cyclic moment will correspond to a local peak of the estimated cyclic moment. Utilizing this property brings computational burden down by more than 50%. Secondly, the global maximum of $\left|\hat{M}_{1y}^{\alpha}\right|$, i.e., the estimate of $\left|M_{1y}^{\alpha}\right|$, is expected corresponding to a cycle frequency of M_{1y}^{α} if M_{1y}^{α} does have one or more cycle frequencies. So it is unnecessary to examine the other frequency locations if the global maximum of $\left|\hat{M}_{1y}^{\alpha}\right|$ is determined as not corresponding to a cycle frequency. Thirdly, the

magnitude of M_{1y}^{α} for MASK at cycle frequency f_c is generally greater than that at $f_c + l \times R_s$ with $l \neq 0$. Thus, when the global maximum of $\left| \hat{M}_{1y}^{\alpha} \right|$ does correspond to a cycle frequency, the classifier only needs to examine the values of α for which $\left| \hat{M}_{1y}^{\alpha} \right|$ are greater than or equal to the product of μ and the global maximum of $\left| \hat{M}_{1y}^{\alpha} \right|$, where μ is a preset constant. Based on the last two properties, the computational burden in searching the cycle frequencies of M_{1y}^{α} can be further reduced. The discussions on M_{1y}^{α} are also feasible to M_{20y}^{α}.

In addition to reducing the computational burden in dealing with M_{1y}^{α} and M_{20y}^{α}, the above method will also remove the ambiguity in the number of cycle frequencies for some entries in Table 4.1. That is, the number of cycle frequencies of M_{1y}^{α} for MASK and that of M_{20y}^{α} for MASK, MPAM and PSK2 will be determined as one.

It should be noted that the cycle frequencies of M_{21y}^{α} generally does not correspond to the global maximum of $\left| M_{21y}^{\alpha} \right|$. However, the bandwidth of a linearly modulated digital communication signal is closely related to its symbol rate. Based on the estimated bandwidth of the input signal, the search range of the cycle frequencies for M_{21y}^{α} can be narrowed down. In general, the narrowed search range for M_{21y}^{α} is adequate to reject the cycle frequencies $f_c + l \times R_s$ with $l \neq 0$. That is, M_{21y}^{α} is expected to be claimed having one cycle frequency at symbol rate for linearly modulated digital communication signals.

Through modification of the approaches of [43] by using all the above-mentioned means, the computational complexity in searching the possible cycle frequencies of the selected cyclic moments has been dramatically reduced, while there is almost no performance degradation compared to the exhaustive search.

4.5 Other Features for Modulation Classification

As discussed in Section 4.3, by detecting the numbers of cycle frequencies of M_{1y}^{α}, $M_{2,0y}^{\alpha}$ and $M_{2,1y}^{\alpha}$, it is possible to separate between analog signals and digital signals, and classify digital signals into several subsets of modulation types as desired. In order to further specify the modulation type of analog communication signals, additional five features, which are based on the spectrum of the input signal, i.e., $P_{y,norm}$, $P_{y2,norm}$, P_L, P_M and P_R have been developed.

4.5.1 AM vs. DSB

In most cases, AM and DSB can be separated from each other based on the presence or absence of a cycle frequency for M_{1y}^{α}. However, occasionally a cycle frequency M_{1y}^{α} for an input DSB signal may be erroneously reported. The DSB signal will be incorrectly recognized as AM since the patterns of the cycle frequencies of $M_{2,0y}^{\alpha}$ and $M_{2,1y}^{\alpha}$ are the same for both AM and DSB. The features $P_{y,norm}$ and $P_{y2,norm}$ are designed to handle such cases.

The feature $P_{y,norm}$ is defined as follows. At first, the PSD $S_y(f)$ of the input signal $\{ y(n) \}$ is estimated by using the average periodogram method [48] as

$$\hat{S}_y(f) = \frac{1}{K}\sum_{k=0}^{K-1}\left\{\frac{1}{N_{seg}}\left|\sum_{n=0}^{N_{seg}-1} y(n+kN_{seg})e^{j\frac{2\pi}{N_{seg}}fn}\right|^2\right\},\ \text{for}\ \ f = 0,1,\cdots,N_{seg}-1 \quad (4.36)$$

where K is the number of segments, N_{seg} is the number of samples in each segment, the product of K and N_{seg} is the largest number that is less than N, and N is the total number of available samples.

Then the estimated PSD is normalized as

$$\hat{S}_{y,norm}(f) \equiv \frac{\hat{S}_y(f)}{\max\left(\hat{S}_y(f)\right)} \quad (4.37)$$

where $\max\left(\hat{S}_y(f)\right)$ is the maximum value of $\hat{S}_y(f)$.

Finally, the feature $P_{y,norm}$ is defined as

$$P_{y,norm} \equiv \sum_{f=0}^{N_{seg}-1} \hat{S}_{y,norm}(f) \quad (4.38)$$

The feature $P_{y2,norm}$ is defined similarly, except that $y(n+kN_{seg})$ in (4.36) is replaced by $y^2(n+kN_{seg})$.

As shown in Appendix B that $P_{y,norm}$ is less than $P_{y2,norm}$ for AM, while $P_{y,norm}$ is greater than $P_{y2,norm}$ for DSB. It is found via simulations that this relationship still holds, even when the input AM signal is noisy, as long as the SNR is not too low. Therefore, if the modulation type has been narrowed down to the set of {AM, DSB}, the correct choice can be recognized by comparing the extracted values of $P_{y,norm}$ and $P_{y2,norm}$.

4.5.2 LSB, USB vs. FM

The features P_L, P_M and P_R are used to discriminate among LSB, USB and FM.

For an input signal, the features P_L, P_M and P_R are defined and measured as follows. At first, the PSD of the received signal $\{ y(n) \}$ is estimated by using equation (4.36) and denoted by $\hat{S}_y(f)$. Secondly, the input signal's bandwidth is estimated using an existing method (e.g.,. [49] and [50]) and denoted by $\hat{B}$. Thirdly, three functions $\Delta_L(f)$, $\Delta_M(f)$ and $\Delta_R(f)$ are established, which are respectively defined as

$$\Delta_L(f) \equiv \begin{cases} \frac{f}{\hat{B}}, & \text{for } 0 \le f \le \hat{B} \\ 0, & \text{otherwise} \end{cases} \tag{4.39}$$

$$\Delta_M(f) \equiv \begin{cases} \frac{2f}{\hat{B}}, & \text{for } 0 \le f \le \frac{\hat{B}}{2} \\ -\frac{2f}{\hat{B}} + 2, & \text{for } \frac{\hat{B}}{2} \le f \le \hat{B} \\ 0, & \text{otherwise} \end{cases} \tag{4.40}$$

$$\Delta_R(f) \equiv \begin{cases} -\frac{f}{\hat{B}} + 1, & \text{for } 0 \le f \le \hat{B} \\ 0, & \text{otherwise} \end{cases}. \tag{4.41}$$

Finally, the features are obtained by

$$P_L = \max_{v \in (0, \frac{f_s}{2})} \left(xcorr\left(\Delta_L(f - v), \hat{S}(f) \right) \right) \tag{4.42}$$

$$P_M = \max_{v \in (0, \frac{f_s}{2})} \left(xcorr\left(\Delta_M(f - v), \hat{S}(f) \right) \right) \tag{4.43}$$

$$P_R = \max_{v \in (0, \frac{f_s}{2})} \left(xcorr\left(\Delta_R(f - v), \hat{S}(f) \right) \right) \tag{4.44}$$

where $xcorr\left(a(f - v), b(f) \right)$ denotes the cross-correlation between $a(f)$ and $b(f)$ with delay being v.

If the modulation type of the input signal is determined as belonging to the subset {LSB, USB, FM}, the features P_L, P_M and P_R as defined will be calculated. The three triangles $\Delta_L(f)$, $\Delta_M(f)$ and $\Delta_R(f)$ are used to mimic the PSDs of the LSB signals, USB signals and FM signals. For AM, DSB, FM and linear digital modulations, the PSD of the transmitted signal is symmetric w.r.t the carrier frequency, and the shape of the PSD is more like the shape of $\Delta_M(f)$. Therefore the value of P_M will be greater than that of P_L and P_R. For LSB, the PSD is not symmetric w.r.t the carrier frequency, and the shape is more like $\Delta_L(f)$, with the result that P_L will be the largest one. Similarly, P_R will be the largest for USB.

The advantage of the above features is that one does not need to select decision thresholds for them. Instead, the decision is made by checking which one is larger (for $P_{y,norm}$ and $P_{y2,norm}$) or largest (for P_L, P_M and P_R). Moreover, the discrimination among LSB, USB and non-SSB does not require knowing or estimating the carrier frequency.

4.6 Decision-Tree Classifier

By detecting the patterns of the cycle frequencies of the selected cyclic moments M_{1y}^{α}, $M_{2,0y}^{\alpha}$ and $M_{2,1y}^{\alpha}$ as well as applying the features developed in Section 4.5, the flow chart of the proposed decision tree is shown in Figrue. 4.1. At the end of classification, a received signal will be classified into one of the following groups: {AM}, {DSB}, {LSB}, {USB}, {FM}, {MASK}, {MPAM, PSK2}, {QAM, MPSK with $M \geq 4$}.

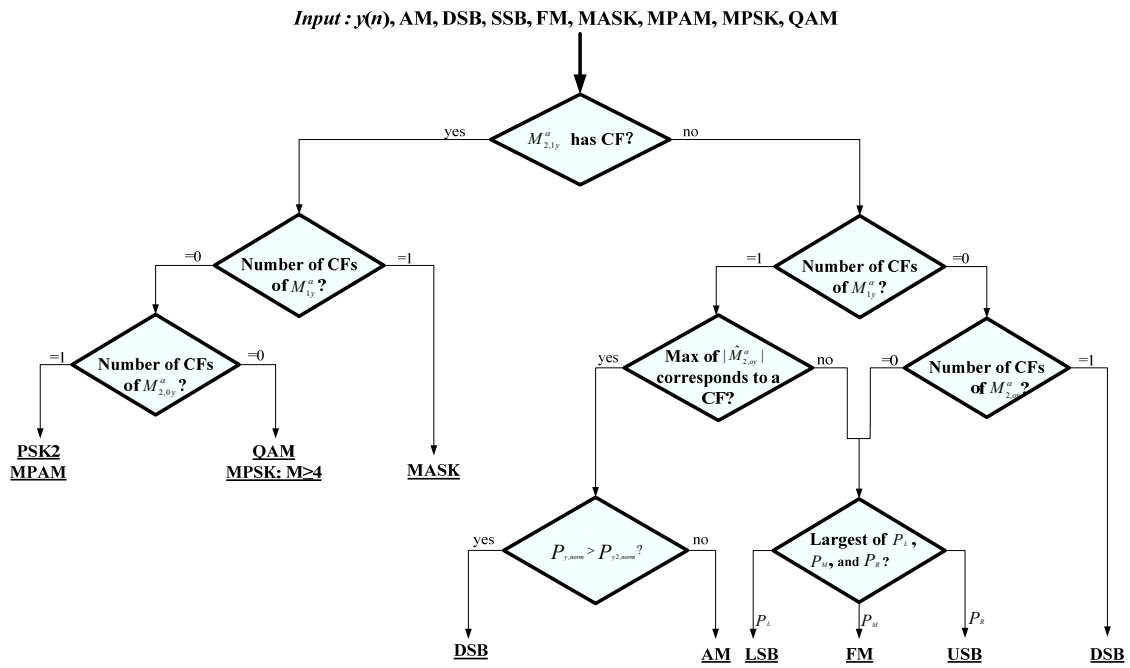


Figure 4.1 Decision tree of the proposed vlassifier

CF stands for cycle frequency; M^{α}_{1y}, $M^{\alpha}_{2,0y}$, $M^{\alpha}_{2,1y}$ are three cyclic moments; $P_{y,norm}$, $P_{y2,norm}$, P_L, P_M and P_R are five features.

The proposed classifier separates the linear digital modulations into several non-overlapping subsets. Further classification of digital modulation schemes has been researched extensively in the literature (e.g., [22, 25, 28, 31-33, 39, 51-54], and can be implemented in order to make the proposed AMC algorithm more practical and comprehensive; but it is beyond the scope of this dissertation.

The experimental results, which will be reported in the next section shows that the proposed decision-tree classifier allows reliable decisions even with very short signal segments for field signals received in outdoor line-of-sight (LOS) wireless transmission.

4.7 Experimental Results and Discussion

4.7.1 Experiments with Simulated Data

The AMC algorithm was tested against simulated signals. 200 simulated signals of each modulation scheme were generated and processed using Matlab. For analog modulation schemes, a recorded speech signal was used as the source message and sampled at $f_s = 44.1$ kHz with 16 bits per sample. The source message was filtered using a Chebyshev type II lowpass filter with less than 3 dB of ripple in the passband defined from 0 to 3 kHz and at least 40 dB of attenuation in the stopband defined from 4 kHz to the Nyquist frequency (22.05 kHz). White Gaussian noise was added to the modulated signals to achieve a specified signal to noise ratio (SNR) of each test, where the SNR is defined as

$$\text{SNR} = 10 \cdot \log_{10}\left[\frac{S}{\sigma_0 \cdot f_s}\right] \quad \text{dB} \tag{4.45}$$

where S is the signal power, σ_0 is the white noise power spectral density and f_s is the sampling frequency. A Hilbert transform was used to create data records in complex format with the desired data length. Among the digital modulation schemes, PSK and QAM signals were simulated using symbols filtered with a square-root raised cosine function with roll-off factor as 0.35.

Sequences of 0.3333 seconds (14,700 samples), 0.5 seconds (22,050 samples) and 0.6667 seconds (29,400 samples) were generated for the analog modulation signals, and sequences of 1,000 symbols, 1,500 symbols and 2,000 symbols were generated for the digital modulation signals, and these were then used as inputs to the classifier.

The classification results are shown in Table 4.2. It can be observed that the classification rate for analog modulation signals is rather high across the overall range of SNRs investigated. For instance, the detection rate maintains above 90% even when SNR is as low as 0dB with data length of 0.6667 second. When the SNR is 4dB or higher, the classifier achieves detection rates close to 100% for all modulation types listed. The detection rates increase as the SNR and data length increases correspondingly.

Table 4.2 Correct Classification Rates On Simulated Data

Data Length	1000 Symbols or 0.3333 sec					
SNR(dB)	0	2	4	6	8	10
AM	99.5	98	97.5	99	99	98
DSB	92.5	93	93	92	90.5	90
LSB	79	76	79.5	76.5	75.5	74.5
USB	79	77	74.5	74.5	73	73.5
NBFM	94.5	91.5	93	91	91	91.5
WBFM	89	87	86	84	85	83.5
PSK2	9.5	51.5	95.5	100	100	100
PSK4	7	59	96.5	99.5	99	100
PSK8	9	59	98	100	99	100
PSK16	7	63.5	98.5	100	100	100
QAM16	8.5	66.5	98.5	99.5	100	100
QAM32	9	55	98	99.5	100	100
QAM64	12.5	61.5	100	100	99.5	100
QAM256	8	54.5	99.5	99.5	100	100
QAM-412	10.5	65	99.5	100	100	99.5
QAM-1616	6	57.5	97.5	98.5	100	100
QAM-4420	10	60	98.5	100	100	99.5
QAM-v29_8	9.5	59.5	99.5	100	99.5	100
QAM-V29_16	9	64	98	100	100	100

Table 4.2 Correct Classification Rates On Simulated Data (Cont'd)

Data Length	1500 Symbols or 0.5 sec					
SNR(dB)	**0**	**2**	**4**	**6**	**8**	**10**
AM	99	99	99	99	100	97.5
DSB	94	92.5	94.5	94.5	94	94.5
LSB	84.5	79.5	85	81	81	84
USB	83.5	82.5	81.5	80	82	80
NBFM	98	99.5	99	99.5	99	99.5
WBFM	93.5	93.5	92	91	90.5	91
PSK2	30	89.5	100	100	100	99.5
PSK4	32.5	96	99.5	100	99.5	99.5
PSK8	41	91.5	99.5	98.5	100	99.5
PSK16	31	96.5	100	98.5	99.5	99.5
QAM16	37	94.5	99.5	100	100	99.5
QAM32	27.5	94.5	100	99.5	99.5	99.5
QAM64	33	92	99.5	99.5	100	99.5
QAM256	38	91.5	99.5	100	99.5	100
QAM-412	36.5	94	99	99.5	100	98.5
QAM-1616	31	95	99.5	99	100	99.5
QAM-4420	33.5	96	99.5	99.5	100	100
QAM-v29_8	36	93	100	99	99	100
QAM-V29_16	31.5	92.5	99	100	99	98.5

Data Length	2000 Symbols or 0.6667 sec					
SNR(dB)	**0**	**2**	**4**	**6**	**8**	**10**
AM	98.5	96.5	98.5	98	96.5	94.5
DSB	98.5	96.5	98.5	99.5	98	99.5
LSB	95.5	92	93.5	94	95.5	92.5
USB	95.5	93.5	95	94	94	94
NBFM	98	99.5	99.5	99.5	99.5	99.5
WBFM	95	97	98.5	96.5	95.5	95.5
PSK2	60	98	100	100	100	99.5
PSK4	66.5	99.5	99.5	100	100	99
PSK8	60	99	99	100	100	99.5
PSK16	60	99	99	100	99.5	99
QAM16	61	99.5	100	100	99.5	99.5
QAM32	67.5	99	99	99.5	100	99
QAM64	61.5	99.5	100	100	100	99
QAM256	62.5	99	100	99	100	99
QAM-412	65.5	98.5	99.5	99.5	100	99.5
QAM-1616	64.5	99.5	100	99.5	100	99.5
QAM-4420	64.5	99.5	100	99.5	99.5	100
QAM-v29_8	59	99.5	99.5	99	100	100
QAM-V29_16	63	98.5	100	100	99	99.5

- QAM-412: This is QAM-16 as defined in MIL-STD-188-110B appendix C. Constellation is composite of: a) 4 symbol inner QAM square and b) 12 symbol outer PSK ring (even spacing).
- QAM-1616: This is QAM-32 as defined in MIL-STD-188-110B appendix C. Constellation is composite of: a) 16 symbol inner QAM square and b) 16 symbol outer PSK ring (non-even spacing).
- QAM-4420: This is QAM-64 as defined in MIL-STD-188-110B appendix C. Constellation is composite of: a) 44 symbol inner QAM square and b) 20 symbol outer PSK ring (even spacing).
- QAM-v29_8: This is QAM-8. See [55] for constellation diagram.
- QAM-v29_16: This is QAM-16. See [55] for constellation diagram.

4.7.2 Experiments with Field Data

The AMC algorithm was further tested against field signals that were transmitted through the air. 500 segments of data were field collected for each of fifteen modulation types. The AMC algorithm was tested on these signals over a range of SNRs by adding white Gaussian noise to the collected data for the Monte Carlo trials.

The symbol rate was 3 kHz for 2PAM and 10 kHz for other digital modulation signals. A standard rectangular pulse function was used for 2PAM, and a square-root raised cosine function with excess bandwidth factor being 0.35 was used for the other digital modulation.

The detailed test results are reported in Table 4.3. It is observed that when the SNR is at least 4dB, the classifier achieves classification rates around 90% or above. The classification rates tend to increase when the SNR or data length increases, as expected.

Table 4.3 Correct Classification Rates On Field Data

Data Length	1000 Symbols or 0.3333 sec					
SNR(dB)	**0**	**2**	**4**	**6**	**8**	**10**
AM	39	90	100	100	100	100
DSB	91	90.5	96.5	98	98.5	98
LSB	92.5	91.5	93.5	90	91.5	89
USB	91.5	90.5	90.5	88.5	89	87.5
NBFM	97.5	94.5	88.5	86	80.5	79.5
WBFM	94	95	97	96.5	95.5	96.5
ASK	10.5	48.5	97	99.5	100	100
PSK2	5	43.5	93.5	100	100	100
PSK4	7.5	46.5	88	99	99.5	100
PI/4-QPSK	7.5	45	92.5	98.5	99.5	99
PSK8	8	47	95.5	100	100	100
QAM16	6	46.5	90	100	99	100
QAM32	4	43.5	93.5	99	100	100
QAM64	9	42.5	93	100	99	100
QAM256	9	42.5	97	100	100	100

Table 4.3 Correct Classification Rates On Field Data (Cont'd)

Data Length	1500 Symbols or 0.5 sec					
SNR(dB)	**0**	**2**	**4**	**6**	**8**	**10**
AM	67.5	99.5	100	100	100	100
DSB	98.5	98.5	99.5	100	100	100
LSB	96	96	95.5	97.5	96.5	96.5
USB	93.5	95	95.5	94.5	95.5	95
NBFM	99.5	97.5	93.5	88	87.5	83.5
WBFM	97.5	99	99	99	99.5	99
ASK	20.5	82.5	99.5	100	100	100
PSK2	29.5	84	99.5	99.5	100	100
PSK4	32	80	99.5	99.5	99	99.5
PI/4-QPSK	36.5	88	100	100	100	100
PSK8	37.5	88.5	99.5	99	99.5	99.5
QAM16	27	89	99.5	100	98.5	99
QAM32	25.5	90	99.5	100	100	100
QAM64	32.5	82.5	100	99.5	100	98.5
QAM256	31	89.5	99	100	100	100

Data Length	2000 Symbols or 0.6667 sec					
SNR(dB)	**0**	**2**	**4**	**6**	**8**	**10**
AM	89.5	99.5	100	100	100	100
DSB	98.5	98	100	97.5	100	99.5
LSB	97.5	98.5	98	98	97.5	99
USB	96	95.5	96	95.5	96	96.5
NBFM	98.5	97.5	93	89.5	83.5	83
WBFM	99	98	98	98.5	99.5	99.5
ASK	49	97.5	100	99.5	100	100
PSK2	48	95.5	100	100	100	100
PSK4	56.5	88.5	99	100	100	100
PI/4-QPSK	60	97	98	99.5	98.5	99.5
PSK8	52.5	98.5	100	99.5	99.5	99.5
QAM16	58	97.5	98.5	100	100	99.5
QAM32	59.5	95.5	100	99.5	99.5	100
QAM64	59.5	94	99	99	98	99
QAM256	51	98	99.5	99.5	99	100

Note: The results (correct classification rates) in Tables 4.2 and 4.3 are reported as the ratio of the number of experiments with correct classification versus the number of total experiments. For instance, if we conduct 200 experiments and we have correct classification in 100 experiments, then a detection rate of 50% is reported.

4.7.3 Confusion Matrices

The performance of a classification system can be evaluated by confusion matrices. Confusion matrices are commonly used to display the probability that the *i-th* modulation scheme is classified as the *j-th* one at a given SNR. Rows of the matrices are the expected classification (actual inputs), while the columns are the resulting classification (outputs).

Confusion matrices on both simulated data and field data are reported in Tables 4.4 and 4.5, which illustrate the cases when data length is 1500 symbols for digital modulated signals and 0.5 sec for analog modulated signals and the SNR is 2dB.

The confusion matrices give insight into which specific errors were made when signals were mis-classified, and show where the failures occurred within the decision tree. For example, it can be observed from Table 4.4 that for the first decision point in the tree , "$M_{2,1y}^{\alpha}$ has CF?", the "No" is declared correctly 99% of the time. In contrast, the "Yes" is declared correctly only 90~96% of the time.

Table 4.4 Confusion Matrix on Simulated Data for SNR=2dB and Data Length as 1500 Symbols or 0.5 sec

Actual Input	Output: Signal Classified as (%)								
	AM	**DSB**	**LSB**	**USB**	**FM**	**MASK**	**PSK2 MPAM**	**MPSK QAM**	**Others**
AM	**99.0**					1.0			
DSB		**92.5**			4.5	1.0			2.0
LSB	15.5	4.0	**79.5**		0.5	0.5			
USB	13.5	3.5		**82.5**	0.5				
NBFM					**99.5**			0.5	
WBFM		0.5	1.0	3.5	**93.5**			1.5	
PSK2		10.5					**89.5**		
PSK4					4.0			**96.0**	
PSK8			2.0	0.5	5.5		0.5	**91.5**	
PSK16			1.0	1.0	1.5			**96.5**	
QAM16			2.5	0.5	2.5			**94.5**	
QAM32			1.5	1.0	3.0			**94.5**	
QAM64			0.5	2.0	5.0		0.5	**92.0**	
QAM256			0.5	1.5	5.0	0.5	1.0	**91.5**	
QAM-412			2.0	0.5	3.0			**94.0**	0.5
QAM-1616			1.0	1.0	3.0			**95.0**	
QAM-4420			1.0	1.0	2.0			**96.0**	
QAM-v29_8			0.5	2.0	4.0			**93.0**	0.5
QAM-V29_16			2.5	0.5	4.5			**92.5**	

Table 4.5 Confusion Matrix on Field Data for SNR=2dB and Data Length as 1500 Symbols or 0.5 sec

Actual Input	Output: Signal Classified as (%)								
	AM	DSB	LSB	USB	FM	MASK	PSK2 MPAM	MPSK QAM	Others
AM	**99.5**			0.5					
DSB	0.5	**98.5**				0.5	0.5		
LSB			**96.0**		3.0			1.0	
USB		1.0		**95.0**	4.0				
NBFM	2.0				**97.5**			0.5	
WBFM		1.0			**99.0**				
ASK	17.5					**82.5**			
PSK2		16.0					**84.0**		
PSK4			1.0	2.0	17.0			**80.0**	
PI/4-QPSK			1.5	0.5	9.0		0.5	**88.0**	0.5
PSK8				0.5	10.5			**88.5**	0.5
QAM16			0.5	1.5	9.0			**89.0**	
QAM32				1.5	8.5			**90.0**	
QAM64			0.5	1.5	15.0	0.5		**82.5**	
QAM256				1.0	9.5			**89.5**	

4.7.4 Comparison with Previous Work

It should be noted that it is very difficult to perform fair performance comparisons with the existing algorithms.

There are a number of reasons for this. 1) There is no standard analog and digital modulation database that the author is aware of. Hence, different authors applied their algorithms to cases of their own choosing. 2) Correct classification rate for a specific modulation type is strongly affected by what other modulations are in the library. As such, identical classification rates obtained by two different algorithms may not mean anything if other competing modulations are not known. 3) Classification rates are frequently reported as a function of signal strength relative to that of white noise. There is no universally agreed upon index, however. Examples that have been used in the literature are SNR [56], symbol SNR [57], CNR[30], E/No[58], etc. These indices are clearly

related but unless classification results are reconciled, numerical comparisons will be difficult.

Besides from the above-mentioned reasons, many existing classifiers such as [59-62] rely on off-line training to obtain the nominal key feature values. However, the proposed classifier is designed to work in a blind environment, i.e., training data are unavailable. If the existing algorithms such as [59-62] are trained with data of some modulation formats but tested with data of other modulation formats (e.g., the modulation type is the same, but the modulation parameters are changed greatly), it is unfair to them. On the other hand, if they are tested with data drawn from modulation formats that are exactly the same of the reference modulation formats, it would be unfair to the proposed classifier.

The maximum likelihood (ML) method is often used as a benchmark to evaluate the performance of a classifier. For analog communication signals, however, it is very difficult to apply the ML method since it will be difficult to find a probability distribution function to describe an unknown information-bearing signal.

For all these reasons, the performance comparison has not been conducted yet. Nevertheless, it should be able to find a way to fairly compare the performance, and this will be done in the future research.

4.8 Conclusions and Future Research

In summary, an AMC algorithm for classification of joint analog and digital modulations based on the pattern of cyclostationarities and the spectral properties of modulated signals has been presented. The proposed decision tree has demonstrated promising performance

with both simulated data and field data. The proposed classifier does not impose unreasonable constraints on the input signal, and thus is practical in non-cooperation communication environment.

As mentioned above, the proposed classifier separates the linear digital modulations into several non-overlapping subsets, but does not yield absolute specificity in declaring the modulation type. Further, multiple-carrier modulations such as orthogonal frequency division multiplexing (OFDM) have not been considered in this dissertation. Some work already has been done in these areas, and these should be studied in order to make this AMC algorithm even more practical.

CHAPTER 5

SUPPORT VECTOR MACHINE BASED AUTOMATIC MODULATION CLASSIFICATION FOR ANALOG SCHEMES

5.1 Introduction

The AMC algorithm presented in this chapter is designed to classify analog modulation schemes. Whereas modern communication systems have a trend to primarily use digital modulations, analog radios are often used by some unconventional military forces and some other civil government agencies. Hence, the detection and recognition of "old technology" analog modulations remains an important task both for military electronic support systems, and for notional cognitive radios.

Some existing technologies are now available to discriminate between analog modulation signals and digital modulation signals. One of the commonly used methods is by detecting the presence of the symbol rate in the received signals, such as in [63, 64]. Therefore, this chapter is concerned with analog modulation signals only, including the following analog modulation schemes: Amplitude Modulation (AM), Double Sideband Modulation (DSB), Upper Sideband Modulation (USB), Lower Sideband Modulation (LSB), and Frequency Modulation (FM).

Since it is difficult to assume a proper probability distribution function for the modulating signal (i.e., the information-bearing signal), the classification of analog communication signals is generally based on features derived from the received signals. A limited amount of information on analog modulation classification is available in the public domain. Y.T. Chan et al. [65] developed a feature which is the ratio of the variance of the instantaneous amplitude to the square of the mean value of instantaneous

amplitude. Nagy [66] proposed two more features, one being the variance of the instantaneous frequency normalized to the squared sample time, and the other being the mean value of the instantaneous frequency. Jovanovic et al. [67] employed the ratio of the variance of the in-phase component to that of the quadrature component of the complex envelope. Al-Jalili [68] studied the centralized instantaneous frequency to discriminate between LSB and USB signals. Nandi and Azzouz [10] used the spectrum symmetry, the maximum value of the power spectral density of the instantaneous amplitude and the standard deviation of the instantaneous phase. Druckmann et al. [69] extracted four features from the instantaneous amplitude. Seaman and Braun [70] analyzed the cyclostationarities of analog signals. Waller and Brushe [71] proposed a specific method for discriminating between FM and PM signals. Taira and Murakami [63] employed the statistics of the instantaneous amplitude to separate two groups: exponentially modulated signals and linearly modulated signals.

Most of the AMC algorithms to date do not work over a wide range of SNR. For example, a classifier that has good performance at higher SNR may not work at lower SNR, and sometimes, vice versa. Some algorithms need prior knowledge of the SNR. In practice, the SNR can be estimated to some degree, but its exact value is difficult to obtain [72]. Recognizing this problem, in this chapter the histogram of instantaneous frequency is proposed as the classification feature, and a multi-class Support Vector Machine (SVM) based classifier is trained in a wide range of SNRs so that it can classify modulation types with high accuracy without the necessity of knowing the exact value of SNR.

The rest of this chapter is organized as follows: Section 5.2 presents the signal models. Section 5.3 discusses feature extraction. Section 5.4 introduces multi-class classification using SVMs. Section 5.5 presents experimental results on the performance of the algorithm, and comparisons of these results with representative results found in the literature. A brief conclusion is drawn in Section 5.6.

5.2 Signal Models

The received signal can be expressed as

$$r(t) = s(t) + n(t) \tag{5.1}$$

where $s(t)$ represents the modulated signal, and $n(t)$ is zero-mean additive white Gaussian noise (AWGN). For analog modulation types, the signal $s(t)$ can be expressed as below [8], respectively,

AM:

$$s(t) = A\left(1 + \mu m(t)\right) \cdot e^{j\omega_c t} \tag{5.2}$$

DSB:

$$s(t) = A\, m(t) \cdot e^{j\omega_c t} \tag{5.3}$$

LSB:

$$x(t) = A\left(m(t) - j\hat{m}(t)\right) \cdot e^{j\omega_c t} \tag{5.4}$$

USB:

$$x(t) = A\left(m(t) + j\hat{m}(t)\right) \cdot e^{j\omega_c t} \tag{5.5}$$

FM:

$$s(t) = A \cdot e^{j\omega_c t + j2\pi f_\Delta \int_{-\infty}^{t} m(v)dv} \tag{5.6}$$

where A is a positive factor used to control the signal power, ω_c is the carrier frequency, $m(t)$ represents the real-valued information-bearing signal that satisfies $-1 \le m(t) \le 1$, $\hat{m}(t)$ stands for the Hilbert transform of $s(t)$, f_Δ is the frequency deviation of an FM signal, μ is the modulation index of an AM signal, and $0 < \mu \le 1$.

5.3 Feature Extraction

Classification features are extracted in two steps. The first step is to estimate of the instantaneous frequency (IF) from the received signal $r(t)$. Then, in the second step, the histogram of the estimated instantaneous frequency is calculated.

5.3.1 Instantaneous Frequency Estimation

Instantaneous frequency has been widely considered as an effective feature in classifying analog modulated signals. It has good anti-noise performance, since the additive noise in the channel is independent of the transmitted signal, and will not have a significant effect on the phase and frequency of the transmitted signal.

The definition and estimation of the instantaneous frequency have been discussed in [73, 74]. In our AMC algorithm, we use the Hilbert transform to construct the analytic form of the received signal, which is defined as

$$y(t) = r(t) + jr_H(t) = A(t)e^{j\Phi(t)} \tag{5.7}$$

where $r_H(t)$ is the Hilbert transform of $r(t)$, $A(t)$ and $\Phi(t)$ are referred to as the instantaneous amplitude and instantaneous phase of the signal $r(t)$ respectively, and can be expressed as (5.8) and (5.9):

$$A(t) = \left| y(t) \right| = \sqrt{r^2(t) + r_H{}^2(t)} \tag{5.8}$$

$$\Phi(t) = \arg\{ y(t) \} \tag{5.9}$$

Before digitization, the intermediate-frequency signal $y(t)$ is down-converted toward 0 Hz, and then low-pass filtered in order to limit the bandwidth and reduce the required sampling rate. The down-converted signal is

$$y_d(t) = y(t)\, e^{-j\hat{\omega}t} = A(t)\, e^{j[\Phi(t) - \hat{\omega}t]} = A(t) e^{j\Phi_d(t)} \tag{5.10}$$

where $\hat{\omega}$ is the estimated carrier frequency, and $\Phi_d(t)$ is the instantaneous phase of the down-converted signal $y_d(t)$. Some existing techniques such as the periodogram technique or the zero-crossing technique (e.g., [75, 76]) can be used to estimate the carrier frequency ω_c of the received signal. The proposed AMC algorithm can tolerate the residual carrier frequency offset between $\hat{\omega}$ and ω_c.

The down-converted analytic signal $y_d(t)$ is sampled at sampling frequency f_s, and we denote $y_d(t_\eta)$ as the discrete version of $y_d(t)$. Assuming that the length of the observed data sequence is N, then the time instants $t_\eta = \eta / f_s$, $\eta = 1, 2, \ldots, N$.

Instantaneous frequency is the time derivative of the unwrapped instantaneous phase,

$$\Omega(t) = \frac{d}{dt} \arg\{ y_d(t) \} = \frac{d}{dt} \Phi_d(t) \tag{5.11}$$

However, since the input signal is of a discrete form, the instantaneous frequency in this chapter is defined as the difference between two time samples of instantaneous phase divided by the time step,

$$\phi(t_\eta) = \frac{\Phi_d(t_\eta) - \Phi_d(t_{\eta-1})}{t_\eta - t_{\eta-1}}, \quad \eta = 1, 2, \ldots N \tag{5.12}$$

where $\Phi_d(t_\eta)$ is the value of instantaneous phase of $y_d(t)$ at time instant t_η.

5.3.2 Histogram of Instantaneous Frequency

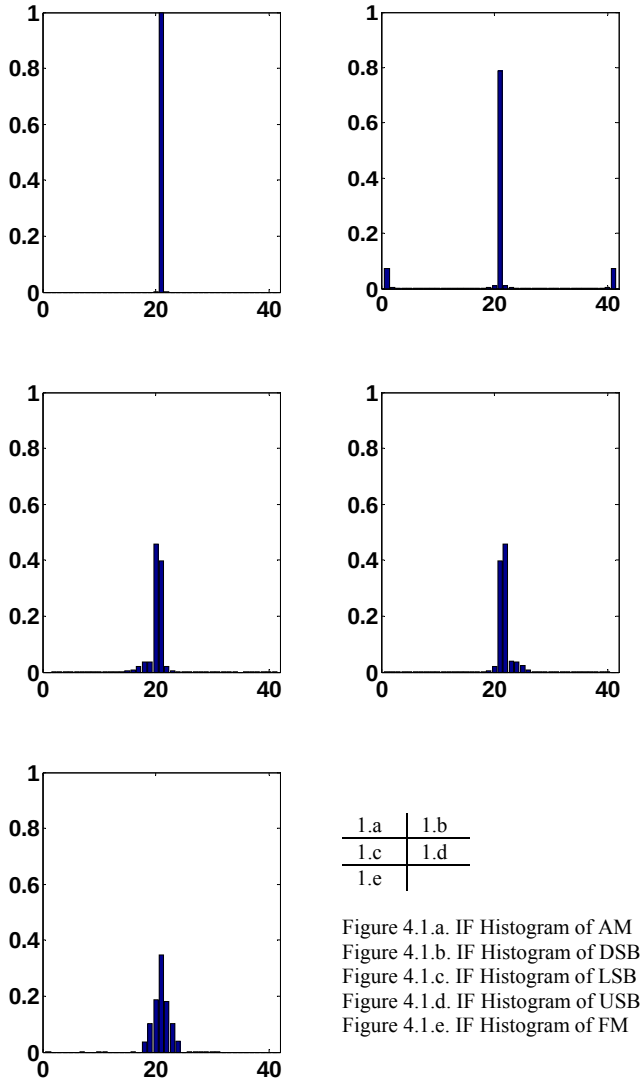
The probability mass function of a discrete data set can be represented by the histogram. The histogram of phase differences determined by (5.12) contains information on the probability distribution of the instantaneous frequency of the signal segment. The instantaneous frequency histograms of the targeted analog modulation types display different patterns and this fact can be employed to classify the received signals.

In the proposed AMC algorithm, the histogram bins are equally spaced between $\pm\max\left|\phi(t_\eta)\right|$, $\eta = 1, 2, \ldots N$. The total number of bins is denoted as κ, therefore, the width of each bin is

$$h = 2 \cdot \frac{\max\left|\phi(t_\eta)\right|}{\kappa - 1} \tag{5.13}$$

It is known that a histogram sometimes is particularly sensitive to the bin-width. If the bin-width is too wide, important features may not be observed, while if the bin-width is too narrow, random variations may be emphasized. It can be difficult to theoretically determine an optimal bin-width. Thus, in practice, different bin-widths should be experimented with in order to understand the relationship between bin-width and the resultant histogram for the applications at hand. We experimented with different values of κ from 5 to 401, and found that for $\kappa > 5$, the number of bins has only a slight influence on the proposed AMC algorithm.

Figure 5.1 gives examples of instantaneous frequency histograms for the ideal case of noiseless analog modulation signals. In this example, the histogram is resolved into 41 bins (i.e., $\kappa = 41$). It can be seen that different modulation schemes result in different instantaneous frequency histograms. For example, AM has a significant peak at the 21st bin but has zero values associated with the bins at both ends of the histogram (i.e., 1st bin and 41st bin), while DSB has a primary peak at the 21st bin as well as two secondary peaks at both ends of the histogram. LSB and USB have asymmetric histograms that are significantly different from the other three analog modulation types. FM has a symmetric histogram, but without significant peaks if compared with AM and DSB.



1.a	1.b
1.c	1.d
1.e	

Figure 4.1.a. IF Histogram of AM
Figure 4.1.b. IF Histogram of DSB
Figure 4.1.c. IF Histogram of LSB
Figure 4.1.d. IF Histogram of USB
Figure 4.1.e. IF Histogram of FM

Figure 5.1 Histogram of instantaneous frequency

5.4 Multi-Class Classification Using SVMs

Support Vector Machine (SVM) is used as the classifier in our AMC algorithm. The SVM is an empirical modeling algorithm, and basically a binary (two-class) classifier. To understand multi-class SVMs, we first consider a two-class SVM. Assume that the training feature-label pairs are

$$(\vec{x}_i, l_i) \qquad i = 1, \ldots, m \tag{5.14}$$

where $\vec{x}_i$ is the training feature vector of the ith training sample belong to Class 1 or Class 2, and associated labels $l_i = 1$ be for Class 1 and $l_i = -1$ be for Class 2, respectively.

The SVM solves the following primal problem [77]:

$$\min_{\vec{w}, b, \zeta_i} \frac{1}{2} \vec{w}^T \vec{w} + C \sum_{i=1}^{m} \zeta_i \tag{5.15}$$

$$\text{subject to } l_i(\vec{w}^T \varphi(\vec{x}_i) + b) \geq 1 - \zeta_i,\ \zeta_i \geq 0,\ i = 1, \ldots, m$$

and its dual is

$$\min_{\vec{\alpha}} \frac{1}{2} \vec{\alpha}^T \mathbf{Q} \vec{\alpha} - \vec{\rho}^T \vec{\alpha} \tag{5.16}$$

$$\text{subject to } \vec{l}_i^T \vec{\alpha} = 0,\ 0 \leq \alpha_i \leq C,\ i = 1, \ldots, m$$

where $\vec{w}$ and b are hyper-plane parameters, ζ_i is positive slack variables, $C > 0$ is the penalty parameter, $\vec{\rho}$ is a vector of all ones, α_i is positive Lagrange multipliers, $\mathbf{Q}$ is an m by m positive semi-definite matrix, $Q_{ij} \equiv l_i l_j Ker(\vec{x}_i, \vec{x}_j)$, and $Ker(\vec{x}_i, \vec{x}_j) \equiv \varphi(\vec{x}_i)^T \varphi(\vec{x}_j)$ is the kernel function, where the function φ maps the training feature vector $\vec{x}_i$ into a higher (may be infinite) dimensional space.

In the proposed AMC algorithm, we use the polynomial kernel:

$$Ker(x_i, x_j) = (\gamma x_i^T x_j + \lambda)^d, \gamma > 0 \tag{5.17}$$

where γ, λ and d are kernel parameters. The penalty parameter C and kernel parameters γ, λ and d should be carefully chosen so that the polynomial kernel SVM can give accurate classification results. The "grid-search" method with cross-validation [77] was employed to find the optimal values of parameters C, γ, λ and d.

For testing a sample, the test feature vector is $\vec{x}$, and the decision function will be given by $\text{sgn}\left(f(x)\right)$:

$$\text{sgn}\left(f(x)\right) = \text{sgn}\left(\sum_{i=1}^{m} l_i \alpha_i Ker(\vec{x}_i, \vec{x}) + b\right) \tag{5.18}$$

For multi-class classification, we use the "one-against-one" approach. Assuming that there are K classes in total, then $C_k = K \cdot (K-1)/2$ binary classifiers are constructed, each separating one class from another ignoring all the other classes. Majority voting is the simplest method for combing C_k binary SVM classifiers. Let $\text{sgn}\left(f_c(x)\right), c = 1, 2, \ldots, C_k$ be a decision function of the cth binary SVM classifier and $Decision_k, k = 1, 2, \ldots, K$ denotes a label of the kth class. Then let $\Im_k$ be the number of binary SVM classifiers whose decisions are known to be the kth class, i.e., $\text{sgn}\left(f_c(x)\right) = Decicion_k$. Then the final decision of the multi-class SVM for a given test feature vector $\vec{x}$ due to majority voting is determined by:

$$\arg\max_k \left[\Im_k\right] \tag{5.19}$$

Applying the above-mentioned general theory of multi-class SVMs to the specific AMC problem:

- Histogram of instantaneous frequency is the input feature vector $\vec{x}_i$ of the SVMs. The instantaneous frequency is defined in equation (5.12), and resolved into κ-dimensional histogram, i.e. $\vec{x}_i$ is of κ-dimension.
- There are $K = 5$ classes, which are AM, DSB, LSB, USB and FM, and an ensemble of $C_k = 10$ binary SVM classifiers are build in total.

5.5 Experimental Results and Discussion

The proposed classifier is trained using software simulated signals that are generated in the MATLAB environment. The initially noiseless simulated signals are passed through an AWGN channel to obtain noisy signals with varying SNR, where the SNR is defined as

$$\text{SNR} = 10 \cdot \log_{10}\left[\frac{S}{\sigma_0 \cdot f_s}\right] \quad \text{dB} \tag{5.20}$$

where S is the signal power, σ_0 is the white noise power spectral density and $f_s = 44.1$ kHz is the sampling frequency, which is chosen to be slightly larger than the occupied bandwidth of most narrowband communications. The range of SNR covered -4dB to 14dB. In 2dB increments, 500 segments of simulated signals are generated for each modulation type. Thus for 10 increments, a total of 5000 feature vectors were obtained for each modulation type. The kernel function and parameters used for training the SVM are as discussed in Section 5.4.

During the training stage, the classifier is trained over a wide range of SNRs (from -4dB to 14dB) simultaneously, and thus the classifier can handle signals of

different SNRs without re-training unless the SNR is outside of the classifier's training range.

To evaluate the performance of the proposed AMC algorithm, the trained classifier is tested against simulated signals. 100 segments of simulated signals are generated for each modulation type in 1dB increments over an extended range of SNRs. Each of these 100 segments was 0.6667 seconds long, and they covered a range of modulation parameters.

Table 5.1 shows the classification results at different SNRs. Although the classifier was trained over a narrower range of SNRs (e.g., -4~14dB at every 2dB), it continued to work well over a wider range of SNRs (e.g., -5~16dB at every 1dB), which has major significance in making the proposed AMC algorithm practical in a blind environment.

The dependence of the classification success rate on the data length was also studied. The use of a short data length will in general degrade the classification performance; however, the experimental results showed that the proposed AMC algorithm can work acceptably well with a rather short data length (100 ms) by achieving a successful classification rate above 93% for a moderate SNR condition (SNR=10dB).

When using the trained classifier, the time taken to recognize an unknown signal is used mainly for signal pre-processing and feature extraction. The experimental results have shown that the proposed algorithm can work with a short data length, and since the processing time for these steps is less for shorter data lengths, our results imply that the pre-processing and feature extraction can be done in a real time manner. This gives the

proposed AMC algorithm great practical significance in a real time recognition environment.

It is difficult to assess the relative performance of the analog modulation classifiers based on the published results since the performance metrics and the test conditions adopted by the various authors are not consistent. To facilitate a relevant performance comparison, we implemented the classifier developed by A. K. Nandi *et al*. in [10], which is still a very attractive algorithm today in this area. We tested our implementation of [10] using exactly the same simulated signals, and channel conditions as we used when testing the proposed histogram algorithm. We maximized the performance of our implementation of [10] by selecting feature threshold values such that the optimum probability of correct decisions was obtained from the 100 realizations for each modulation type of interest.

The classification results of our implementation of [10] at different SNRs are shown in Table 5.2. Comparing these results to those in Table 5.1, it shows that the proposed AMC classifier prominently outperforms our implementation of [10], especially when the SNR is lower than 2dB.

Table 5.1 Probabilities Of Correct Classification For Proposed Classifier

	SNR(dB)											
	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4
AM	0.01	0.14	0.74	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98
FM	0.05	0.32	0.87	0.94	0.99	1.00	0.99	0.98	1.00	1.00	1.00	1.00
DSB	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LSB	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
USB	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

	SNR(dB)											
	5	6	7	8	9	10	11	12	13	14	15	16
AM	0.97	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
FM	1.00	1.00	1.00	0.99	0.99	0.99	0.98	0.96	0.97	0.96	0.97	0.97
DSB	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LSB	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
USB	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 5.2 Probabilities Of Correct Classification For Our Implementation Of The Classifier in [10]

	SNR(dB)											
	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4
AM	0.00	0.00	0.02	0.02	0.15	0.23	0.25	0.31	0.73	0.86	0.91	0.97
FM	0.31	0.87	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DSB	0.02	0.03	0.27	0.61	0.74	0.91	0.92	0.96	0.95	0.96	0.96	0.92
LSB	0.00	0.00	0.02	0.02	0.15	0.39	0.49	0.59	0.69	0.92	0.93	0.93
USB	0.00	0.00	0.03	0.01	0.21	0.36	0.52	0.54	0.67	0.89	0.91	0.91

	SNR(dB)											
	5	6	7	8	9	10	11	12	13	14	15	16
AM	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
FM	1.00	0.97	1.00	1.00	0.99	0.97	0.98	1.00	1.00	1.00	1.00	1.00
DSB	0.94	0.92	0.95	0.90	0.93	0.93	0.94	0.96	0.92	0.94	0.95	0.92
LSB	0.94	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
USB	0.94	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95

5.6 Conclusions

A new AMC algorithm has been developed for blind classification of analog modulation signals. The histogram of instantaneous frequency is used as the modulation-classification features. Support Vector Machines (SVMs) are used as the classifier. Simulation results have shown that the proposed classification method is quite robust to additive noise in a wide range of SNR.

CHAPTER 6

LIKELIHOOD RATIO TESTING METHODS OF AMC FOR ANALOG MODULATION SCHEMES

6.1 Introduction

Automatica modulation classification (AMC) algorithms can be crystallized into two general classes, the likelihood-based (LB) methods, and the feature-based (FB) methods, respectively. For the LB methods, the likelihood function of the received signal is computed and the likelihood ratio is compared against a threshold to make the classification decision. LB methods for digital modulation types have been well developed. But for analog modulation types, LB methods are seldom explored.

This chapter is on likelihood ratio testing (LRT) methods of AMC for analog modulation schemes. In Section 6.2, signal models for analog modulation are presented, and the AMC problem is formulated within the framework of likelihood ratio test. Section 6.3 is an overview of the theory of average likelihood ratio testing (ALRT). Section 6.4 is the main body of this Chapter. In this section we are to investigate both coherent modulation classification and non-coherent modulation classification. A brief conclusion is drawn in Section 6.5.

6.2 Signal Models for Analog Modulation

The received signal can be expressed as

$$r(t) = s(t) + n(t) \tag{6.1}$$

where $s(t)$ is the modulated signal emitted form a non-cooperative transmitter, and $n(t)$ is additive noise obeying white and zero-mean Gaussian, i.e., $n(t) \sim \mathcal{N}(0, \sigma^2)$, with a double sided power spectrum density (PSD) of $\frac{N_0}{2} = \sigma^2$.

The received signal is one of M candidate modulation types. Let the integers $l = 1, 2, \ldots, M$ enumerate the candidate modulation types, such that H_l for $l = 1, 2, \ldots, M$ denotes the hypothesis that the received signal belong to the l^{th} modulation type. We assume equal *a priori* probabilities $P[H_l]$. The following analog modulation types: amplitude modulation (AM: H_1) , double sideband modulation (DSB: H_2), and frequency modulation (FM: H_3), are investigated.

The received signal $r(t)$ was sampled at a sampling rate f_s, and we assume that a set of N samples were abtained. The observed noise corrupted signal samples are denoted by

$$r_i = s_i + n_i, i = 1, 2, \ldots, N \tag{6.2}$$

where n_i is the noise sample, which is a zero-mean Gaussian random variable with variance σ^2; s_i represents the noiseless analog modulated signal. In this way we can express $r(t)$ in terms of a vector $\mathbf{R}$ that consists of a set of N noisy observations: $\mathbf{R} = [r_1, r_2, r_3, \ldots, r_N]^T$. The superscript T is the transpose operator.

As the data source of detection theory, the noiseless analog modulated signal s_i under each hypothesis can be written as

$$H_1 : s_i(\boldsymbol{\beta}_1) = A(1 + \mu m_i) \cdot \cos(\omega_c t_i + \theta_c) \tag{6.3}$$

$$H_2 : s_i(\boldsymbol{\beta}_2) = A\, m_i \cdot \cos\left(\omega_c t_i + \theta_c\right) \tag{6.4}$$

$$H_3 : s_i(\boldsymbol{\beta}_3) = A \cos\left(2\pi f_\Delta \cdot \sum_{k=1}^{i}\left(m_k \cdot \frac{1}{f_s}\right) + \omega_c t_i + \theta_c\right) \tag{6.5}$$

where $\boldsymbol{\beta}_l, l = 1, 2, \ldots, M$ is defined as a multiple dimensional parameter space in which a bunch of signal and channel parameters are unknown random or deterministic variables under the l^{th} modulation type

$$\boldsymbol{\beta}_l = \left\{ A, \omega_c, \theta_c, m_i, \mu, f_\Delta \right\} \tag{6.6}$$

The symbols used in (6.3) ~ (6.6) are:

- $m_i, i = 1, 2, \ldots, N$, the real-valued information-bearing message that satisfies $-1 \le m_i \le 1$;
- $t_i = \dfrac{i}{f_s}$, the time instant;
- ω_c, the residual carrier frequency after complex frequency down conversion and coarse carrier removal;
- θ_c, the carrier phase offset introduced by the propagation delay as well as the initial carrier phase on the transmitter side;
- μ, the modulation index of an AM signal, and $0 < \mu \le 1$;
- f_Δ, the frequency deviation of an FM signal;
- A, the signal amplitude, and simple techniques can be used to estimate the signal amplitude A, and then the received signal stream r_i, $i = 1, 2, \ldots, N$ can therefore be scaled by A. Correspondingly, $A = 1$ is assumed in all derivations in the subsequent sections.

6.3 Theory of Likelihood Ratio Testing

Automatic modulation classification is a composite hypothesis testing problem, since the parameters encompassed by $\boldsymbol{\beta}_l$ are likely not available.

The conditional probability density function (PDF) of r_i under hypothesis H_l can be written as

$$p(r_i \mid H_l, \boldsymbol{\beta}_l) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(-\frac{\left(r_i - s_i(\boldsymbol{\beta}_l) \right)^2}{2\sigma^2} \right) \tag{6.7}$$

In statistics, we view $p(r_i \mid H_l, \boldsymbol{\beta}_l)$ as a function of H_l, as the *likelihood function* with each single observation. The expression in equation (6.7) always acts as the baseline for the derivations for different levels of prior knowledge in $\boldsymbol{\beta}_l$ in following sections. It also governs the mapping from the parameter space $\boldsymbol{\beta}_l$ to the observation space $\mathbf{R}$ conditioned on composite hypothesis H_l. The observation space $\mathbf{R}$ is characterized by the parameter $\boldsymbol{\beta}_l$.

If the parameters inside $\boldsymbol{\beta}_l$ are treated as random variables with known probability densities, we designate $p(\boldsymbol{\beta}_l \mid H_l)$ as the joint PDF of the random variables in $\boldsymbol{\beta}_l$ on the hypothesis H_l, then the likelihood function of a single observation is obtained by taking the statistical averaging of equation (6.7) over the unknown parameters, which is referred to as the average likelihood ratio test (ALRT):

$$p_a(r_i \mid H_l) = \int_{\boldsymbol{\beta}_l} p(r_i \mid H_l, \boldsymbol{\beta}_l) \cdot p(\boldsymbol{\beta}_l \mid H_l)\, d\boldsymbol{\beta}_l \tag{6.8}$$

Because the n_i are statistically independent, the joint probability density of the observation vector $\mathbf{R}$ is simply the product of the individual probability densities, i.e., the likelihood function of $\mathbf{R}$ under H_l is

$$\Gamma_a(\mathbf{R} \mid H_l) = \prod_{i=1}^{N} p_a(r_i \mid H_l) \tag{6.9}$$

Furthermore, because the natural logarithm is monotonically increasing and $\Gamma_a(\mathbf{R} \mid H_l)$ is positive, the likelihood function for the observation vector $\mathbf{R}$ can equivalently be expressed as

$$L_a(\mathbf{R} \mid H_l) = \ln \Gamma_a(\mathbf{R} \mid H_l) \tag{6.10}$$

The classification decision is made by choosing l as the modulation type if $L_a(\mathbf{R} \mid H_l)$ is maximum among multiple hypotheses, i.e.,

$$l_{\max} = \arg \max_{l \in \{1,2,\ldots M\}} \Gamma_a(\mathbf{R} \mid H_l) \tag{6.11}$$

6.4 ALRT-based Analog Modulation Classification

Automatic modulation classification is a rather difficult problem in composite hypothesis testing since so many parameters are unknown: residual carrier frequency ω_c, carrier phase θ_c, AM modulation index μ, FM frequency deviation f_Δ, information-bearing message m_i, and signal to noise ratio. A common approach is to first estimate the unknown parameters and then attempt to classify the signal according to modulation type. Although estimating these parameters is nontrivial, it is not impractical. There are a wide variety of techniques for estimating the signal parameters, some of which are given in [78-80].

The investigation of average likelihood ratio testing methods begins with the coherent environment in Sub-Section 6.4.1, and non-coherent modulation classification is investigated in Sub-Section 6.4.2.

6.4.1 ALRT-Based Coherent Modulation Classification

In this sub section, we investigate ALRT algorithms in the coherent environment where the residual carrier frequency ω_c, carrier phase θ_c, and AM modulation index μ are assumed to be perfectly estimated before attempting to classify the modulation types. This ideal situation may be reasonable in cooperative communications where the receiver knows the physical layer and has handshakes with the transmitter. The SDR with adaptive demodulation radio can be semi-cooperative where the carrier frequency, carrier phase, and signal amplitude are tracked steadily in real-time by a synchronization mechanism, and the unknown modulation scheme is classified by the LRT classifier.

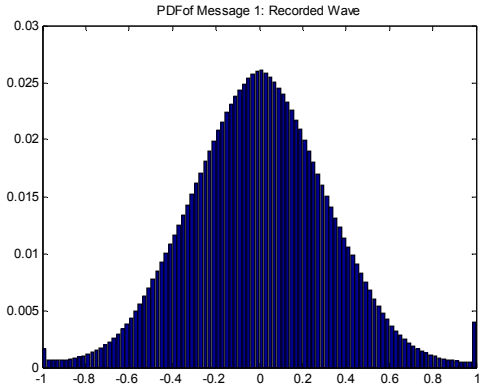
The unknown variables remaining in $\boldsymbol{\beta}_l$ are the real-valued information-bearing message sequence emitted by the transmitter, m_i, $i = 1, 2, \ldots, N$. The sequence of m_i, $i = 1, 2, \ldots, N$ is assumed to be independent and identically distributed (i.i.d.).

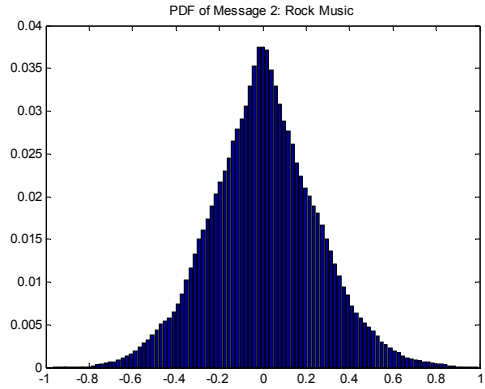
On account of the Central Limit Theorem, it is reasonable to model m_i as a Gaussian random variable with mean zero and variance σ_m^2, i.e., $m_i \sim \mathcal{N}(0, \sigma_m^2)$. Equivalently, the PDF of m_i is:

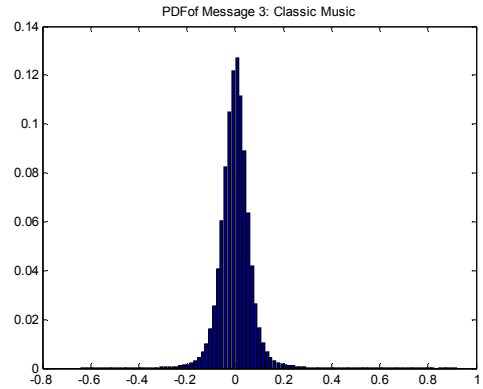
$$p(m_i \mid H_l) = \frac{1}{\sqrt{2\pi} \cdot \sigma_m} \cdot \exp\left(-\frac{m_i^2}{2\sigma_m^2} \right), \quad l = 1, 2, 3 \tag{6.12}$$

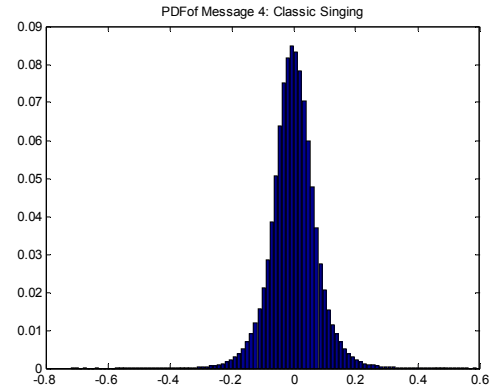
Figure 6.1 illustrates investigation results for PDF of different modulating message data. Totally 8 different modulating messages (including: 1 piece of audio

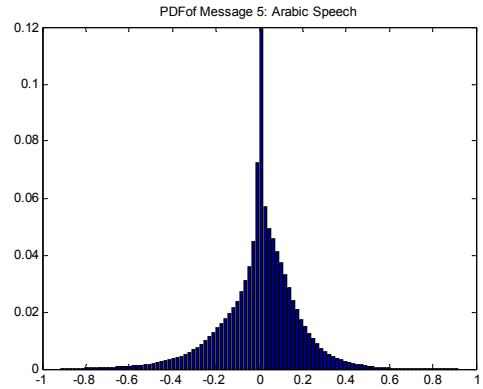
recorded by microphone, 3 pieces of music downloaded from Internet, 4 pieces of speech downloaded from Internet, Detailed description of the modulating message data can be found in Sub-Sub-Section 6.4.1.4) are investigated. Investigation results show that Gaussian distribution is an appropriate approximation.

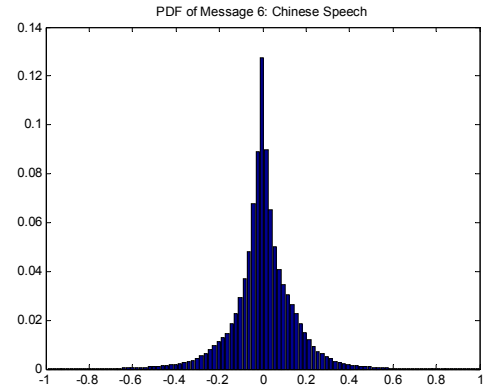












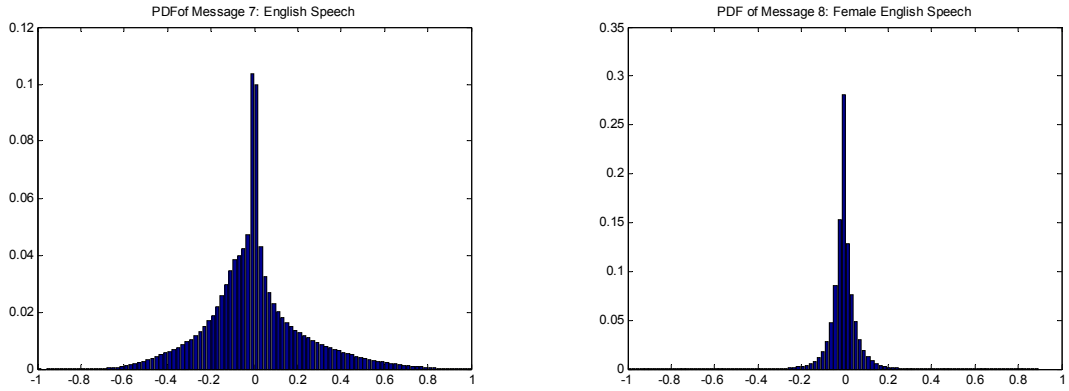


Figure 6.1 Investigation results for PDF of different modulating message data

Based on this assumption, the likelihood function in equation (6.8) under each hypothesis can then be derived as following.

6.4.1.1 Likelihood function for AM modulation scheme The derivation of the likelihood function for AM modulation scheme starts from equation (6.3). In coherent situation, the residual carrier frequency ω_c and carrier phase θ_c are assumed to be perfectly estimated, therefore can be removed from equation (6.3). The modulation parameter μ is estimated before classification, so μ can be seen as a known deterministic. The only unknown parameter in equation (6.3) is $\{ m_i, i = 1, 2, \ldots, N \}$, therefore the multiple dimensional parameter space $\boldsymbol{\beta}_1$ now is degenerated to a one dimensional parameter space, i.e., $\boldsymbol{\beta}_1 = \{ m_i, i = 1, 2, \ldots, N \}$. We rewritten equation (6.3) as

$$\begin{aligned} & s_i(\boldsymbol{\beta}_1) \\ &= A(1 + \mu m_i) \cdot \cos(\omega_c t_i + \theta_c) \\ &= 1 + \mu m_i \end{aligned} \tag{6.13}$$

Substituting equation (6.13) into equation (6.7), the conditional probability of the ith received sample r_i under hypothesis H_1, can be rewritten as,

$$
\begin{aligned}
& p(r_i \mid H_1, \boldsymbol{\beta}_1) \\
&= p(r_i \mid H_1, m_i) \\
&= \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\exp\left(-\frac{\left(r_i-(1+\mu m_i)\right)^2}{2\sigma^2}\right)
\end{aligned}
\tag{6.14}
$$

Substituting equations (6.12) and (6.14) into equation (6.8), and take the integral from -1 to 1, then we have,

$$
\begin{aligned}
p_a(r_i \mid H_1) &= \int_{-1}^{1} p(r_i \mid H_1, m_i)\cdot p(m_i \mid H_1)\, dm_i \\
&= \int_{-1}^{1} \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\exp\left(-\frac{\left(r_i-(1+\mu\cdot m_i)\right)^2}{2\sigma^2}\right)\cdot\frac{1}{\sqrt{2\pi}\cdot\sigma_m}\cdot\exp\left(-\frac{\left(m_i\right)^2}{2\sigma_m^2}\right) dm_i \\
&= \frac{1}{2\sqrt{2\pi}\cdot\sqrt{\sigma_m^2\mu^2+\sigma^2}}\cdot\exp\left(-\frac{(-1+r_i)^2}{2(\sigma_m^2\mu^2+\sigma^2)}\right) \\
&\cdot\left(\operatorname{erf}\left[\frac{\sigma_m^2\cdot\mu\cdot(1-r_i+\mu)+\sigma^2}{\sqrt{2}\cdot\sigma\cdot\sigma_m\sqrt{\sigma_m^2\mu^2+\sigma^2}}\right]+\operatorname{erf}\left[\frac{\sigma_m^2\cdot\mu\cdot(-1+r_i+\mu)+\sigma^2}{\sqrt{2}\cdot\sigma\cdot\sigma_m\sqrt{\sigma_m^2\mu^2+\sigma^2}}\right]\right)
\end{aligned}
\tag{6.15}
$$

where erf is the error function, which is defined as

$$
\operatorname{erf}\left[x\right]=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt \tag{6.16}
$$

Given N observed noise corrupted signal samples, $\mathbf{R}=\left[r_1, r_2, r_3, \ldots, r_N\right]^T$, the likelihood function of $\mathbf{R}=\left[r_1, r_2, r_3, \ldots, r_N\right]^T$ can be easily found by substituting equation (6.15) into equation (6.9).

6.4.1.2 Likelihood function for DSB modulation scheme The derivation process of the likelihood function for DSB modulation scheme is similar to that for AM modulation scheme, therefore simply following the process given above, we can obtain

$$
\begin{aligned}
& s_i(\boldsymbol{\beta}_2) \\
&= m_i \cdot \cos\left(\omega_c t_i + \theta_c\right) \\
&= m_i
\end{aligned}
\tag{6.17}
$$

$$
\begin{aligned}
& p(r_i \mid H_2, \boldsymbol{\beta}_2) \\
&= p(r_i \mid H_2, m_i) \\
&= \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \exp\left(-\frac{\left(r_i - m_i\right)^2}{2\sigma^2}\right)
\end{aligned}
\tag{6.18}
$$

$$
\begin{aligned}
p_a\left(r_i \mid H_2\right) &= \int_{-1}^{1} p(r_i \mid H_2, m_i) \cdot p(m_i \mid H_2)\, dm_i \\
&= \int_{-1}^{1} \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \exp\left(-\frac{\left(r_i - m_i\right)^2}{2\sigma^2}\right) \cdot \frac{1}{\sqrt{2\pi}\cdot\sigma_m} \cdot \exp\left(-\frac{\left(m_i\right)^2}{2\sigma_m^2}\right) dm_i \\
&= \frac{1}{2\sqrt{2\pi}\cdot\sqrt{\sigma_m^2+\sigma^2}} \cdot \exp\left(-\frac{r_i^2}{2\cdot(\sigma_m^2+\sigma^2)}\right) \\
&\quad \cdot \left(\operatorname{erf}\left[\frac{\sigma_m^2\cdot(-1+r_i)+\sigma^2}{\sqrt{2}\cdot\sigma\cdot\sigma_m\sqrt{\sigma_m^2+\sigma^2}}\right] + \operatorname{erf}\left[\frac{\sigma_m^2\cdot(1+r_i)+\sigma^2}{\sqrt{2}\cdot\sigma\cdot\sigma_m\sqrt{\sigma_m^2+\sigma^2}}\right]\right)
\end{aligned}
\tag{6.19}
$$

The likelihood function of $\mathbf{R} = \left[r_1, r_2, r_3, \ldots, r_N\right]^T$ can be easily found by substituting equation (6.19) into equation (6.9).

6.4.1.3 Likelihood function for FM modulation scheme The derivation of the likelihood function for FM modulation scheme is more involved. As modeled at the beginning of the section, we know that $m_k \sim \mathcal{N}(0, \sigma_m^2)$. Let y_i be the cumulative sum of m_k, $k = 1, 2, \ldots, i$, i.e.,

$$
y_i = 2\pi f_\Delta \cdot \sum_{k=1}^{i} \left(m_k \cdot \frac{1}{f_s}\right) \tag{6.20}
$$

Because that the sequence of m_k , $k = 1, 2, \ldots, i$ is i.i.d., then y_i is normally distributed with mean 0 and variance $\sigma_y^2 = i \cdot \left(2\pi \cdot \frac{f_\Delta}{f_s} \right)^2 \cdot \sigma_m^2$, i.e., $y_i \sim \mathcal{N}(0, \sigma_y^2)$, $i = 1, 2, \ldots, N$.

Further, let z_i be the modulus of y_i after division by 2π , i.e.,

$$z_i = \mathrm{MOD}[\, y_i , 2\pi \,] \tag{6.21}$$

The PDF of z_i is given by [81], as

$$p(z_i) = \frac{1}{\sigma_y \sqrt{2\pi}} \cdot \left\{ 1 + 2 \sum_{p=-\infty}^{\infty} \exp\left(-\frac{(z_i + 2\pi \cdot p)^2}{2\sigma_y^2} \right) \right\}, \quad 0 < z_i \le 2\pi \tag{6.22}$$

In [81], K. V. Mardia proved that for large σ_y^2 , $p(z_i)$ tends to the uniform distribution, i.e.,

$$p(z_i) \approx \frac{1}{2\pi}, \quad 0 < z_i \le 2\pi \tag{6.23}$$

And we always assume that this is the case in this dissertation.

Then the likelihood function for FM modulation scheme can be formulated in two different ways, which lead to the first type of likelihood function (1st-LF) and the second type of likelihood function (2nd-LF) for FM Modulation schemes. Both two types of likelihood function are tested by the experiments.

6.4.1.3.1 The First Type of Likelihood Function for FM Modulation (1st-LF)

equation (6.5) is rewritten as

$$
\begin{aligned}
& s_i(\boldsymbol{\beta}_3) \\
&= \cos\left(2\pi f_\Delta \cdot \sum_{k=1}^{i}\left(m_k \cdot \frac{1}{f_s}\right) + \omega_c t_i + \theta_c\right) \\
&= \cos\left(2\pi f_\Delta \cdot \sum_{k=1}^{i}\left(m_k \cdot \frac{1}{f_s}\right)\right) \\
&= \cos(y_i) \\
&= \cos(z_i)
\end{aligned}
\tag{6.24}
$$

Therefore the multiple dimensional parameter space $\boldsymbol{\beta}_3$ now is degenerated to a one dimensional parameter space, i.e., $\boldsymbol{\beta}_3 = \{ z_i, i = 1, 2, \ldots, N \}$.

Substituting equation (6.24) into equation (6.7), the conditional probability of the ith received sample r_i under hypothesis H_3, can be rewritten as,

$$
p(r_i | H_3, \boldsymbol{\beta}_3) = p(r_i | H_3, z_i) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \exp\left(-\frac{\left(r_i - \cos(z_i)\right)^2}{2\sigma^2}\right) \tag{6.25}
$$

Substituting equations (6.23) and (6.25) into equation (6.8), and take the integral from 0 to 2π, then we have,

$$
p_a(r_i | H_3) = \int_0^{2\pi} \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \exp\left(-\frac{\left(r_i - \cos(z_i)\right)^2}{2\sigma^2}\right) \cdot \frac{1}{2\pi}\, dz_i \tag{6.26}
$$

Unlike the cases of AM and DSB, the integral in equation (6.26) does not admitany closed form solution. However, the integral in equation (6.26) can be resolved by numerical integrations.

The likelihood function of $\mathbf{R} = [r_1, r_2, r_3, \ldots, r_N]^T$ can be easily found by substituting equation (6.26) into equation (6.9).

6.4.1.3.2 The Second Type of Likelihood Function for FM Modulation (2nd-LF)

Let

$$w_i = \cos(z_i), \quad 0 < z_i \le 2\pi \tag{6.27}$$

then the PDF of w_i is

$$p(w_i) = \frac{1}{\pi} \cdot \frac{1}{\sqrt{1-w_i^2}}, \quad -1 < w_i < 1 \tag{6.28}$$

We rewritten equation (6.5) as

$$\begin{aligned} & s_i(\boldsymbol{\beta}_3) \\ &= \cos\left(2\pi f_\Delta \cdot \sum_{k=1}^{i}\left(m_k \cdot \frac{1}{f_s}\right) + \omega_c t_i + \theta_c\right) \\ &= \cos\left(2\pi f_\Delta \cdot \sum_{k=1}^{i}\left(m_k \cdot \frac{1}{f_s}\right)\right) \\ &= \cos(y_i) \\ &= \cos(z_i) \\ &= w_i \end{aligned} \tag{6.29}$$

Therefore the multiple dimensional parameter space $\boldsymbol{\beta}_3$ now is degenerated to a one dimensional parameter space, i.e., $\boldsymbol{\beta}_3 = \{ w_i, i = 1, 2, \ldots, N \}$.

Substituting equation (6.29) into equation (6.7), the conditional probability of the ith received sample r_i under hypothesis H_3, can be rewritten as,

$$p(r_i \mid H_3, \boldsymbol{\beta}_3) = p(r_i \mid H_3, w_i) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(-\frac{(r_i - w_i)^2}{2\sigma^2}\right) \tag{6.30}$$

Substituting equations (6.28) and (6.30) into equation (6.8), and take the integral from -1 to 1, then we have,

$$p_a(r_i|H_3) = \int_{-1}^{1} \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \exp\left(-\frac{(r_i - w_i)^2}{2\sigma^2}\right) \cdot \frac{1}{\pi} \cdot \frac{1}{\sqrt{1-w_i^2}}\, dw_i \qquad (6.31)$$

Similarly, the integral in equation (6.31) does not admitany closed form solution, but can be resolved by numerical integrations.

The likelihood function of $\mathbf{R} = [r_1, r_2, r_3, \ldots, r_N]^T$ can be easily found by substituting (6.31) into (6.9).

6.4.1.4 Performance Evaluation for the Coherent Classifier The coherent classifier is test by simulated signals. Simulated signals are generated in the MATLAB environment. The simulated signals are passed through an AWGN channel and the SNR is defined as

$$\text{SNR} = 10 \cdot \log_{10}\left[\frac{S}{\sigma_0 \cdot f_s}\right] \quad \text{dB} \qquad (6.32)$$

where S is the signal power, σ is the white noise power spectral density and $f_s = 44.1$ kHz is the sampling frequency. All the simulated signals are band-pass filtered, and then down-converted toward 0 Hz. The range of SNR is from -10dB to 10dB. With a 2dB increment, 200 snapshots of simulated signals are generated for each modulation type, each with length of 0.6667 sec.

Totally 8 different modulating messages (including: 1 piece of audio recorded by microphone, 3 pieces of music downloaded from Internet, 4 pieces of speech downloaded from Internet) are used to test the proposed ALRT-based coherent classifier.

- Message 1: recorded by a microphone, and includes various sources, such as speech, music, opera and even some background noise.

- Message 2: rock music downloaded from the Internet.
- Message 3: classic music downloaded from the Internet.
- Message 4: classic singing downloaded from the Internet.
- Message 5: Arabic speech downloaded from the Internet.
- Message 6: Chinese speech downloaded from the Internet.
- Message 7: English speech downloaded from the Internet.
- Message 8: female English speech downloaded from the Internet.

The detailed test results are reported in Tables 6.1.1~6.1.16. The test results show the same trend, i.e.,

- The classifier with 1st-LF can achieve more than 97% detection rate for AM and DSB, when SNR is as low as -10dB. But it fails for FM, when SNR is lower than 2dB.
- The classifier with 2nd-LF outperforms 1st-LF classifier, in the respect of FM. It can work for all the three modulation types, when SNR is as low as -4dB.
- When SNR is high (larger than 4dB), the decisions form both 1st-LF classifier and 2nd-LF classifier are trustable and robust. Therefore, we can choose either one in practice.
- When SNR is in the range from -4dB to 4dB, the decision from 2nd-LF classifier should be trustable. While, for 1st-LF classifier, most FM signals are mis-classified as DSB.
- When SNR is low (smaller than -4dB), both 1st-LF classifier and 2nd-LF classifier should be used in order to increase detection rate and reduce mis-classification rate. For 1st-LF classifier, the signals will be classified into sub-

groups, {AM} and {DSB, FM}, while for 2nd-LF classifier, the signals will be classified into sub-groups {DSB} and {FM, AM}

Table 6.1.1 Correct Classification Rate for Coherent Classifier (using 1st Type FM Likelihood Function) (Message 1: Recorded Wave)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.985	0.985	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.975	0.975	0.985	0.985	0.985	0.985	0.985	0.985	0.985	0.985	0.985
FM	0.075	0.050	0.030	0.025	0.025	0.025	0.025	0.995	1.000	1.000	1.000

Table 6.1.2 Correct Classification Rate for Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 1: Recorded Wave)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.000	0.000	0.975	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.005	0.075	0.540	0.925	0.965	0.975	0.975	0.985	0.985	0.985	0.985
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6.1.3 Correct Classification Rate for Coherent Classifier (using 1st Type FM Likelihood Function) (Message 2: Music Rock)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.005	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000

Table 6.1.4 Correct Classification Rate for Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 2: Music Rock)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.045	0.520	0.910	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6.1.5 Correct Classification Rate for Coherent Classifier (using 1st Type FM Likelihood Function) (Message 3: Music Classic)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.010	0.005	0.000	0.000	0.000	0.000	0.010	1.000	1.000	1.000	1.000

Table 6.1.6 Correct Classification Rate for Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 3: Music Classic)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.415	0.780	0.910	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6.1.7 Correct Classification Rate for Coherent Classifier (using 1st Type FM Likelihood Function) (Message 4: Singing Classic)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.025	0.020	0.010	0.005	0.000	0.000	0.040	0.690	0.785	0.835	0.895

Table 6.1.8 Correct Classification Rate for Coherent Classifier (Using 2nd Type Likelihood Function) (Message 4: Singing Classic)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.095	0.370	0.895	0.975	0.990	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	0.895	0.775	0.730	0.735	0.770	0.825	0.885

Table 6.1.9 Correct Classification Rate for Coherent Classifier (using 1st Type FM Likelihood Function) (Message 5: Arabic)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.000	0.000	0.000	0.000	0.000	0.000	0.230	0.800	0.930	0.995	1.000

Table 6.1.10 Correct Classification Rate for Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 5: Arabic)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.835	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	0.995	0.930	0.880	0.900	0.935	0.990	1.000

Table 6.1.11 Correct Classification Rate for Coherent Classifier (using 1st Type FM Likelihood Function) (Message 6: Chinese)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.000	0.000	0.000	0.000	0.000	0.000	0.390	0.940	0.995	0.995	1.000

Table 6.1.12 Correct Classification Rate for Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 6: Chinese)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.955	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	0.995	0.995	0.995	0.995	1.000	1.000

Table 6.1.13 Correct Classification Rate for Coherent Classifier (using 1st Type FM Likelihood Function) (Message 7: English)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.000	0.000	0.000	0.000	0.000	0.000	0.270	0.890	0.990	1.000	1.000

Table 6.1.14 Correct Classification Rate for Coherent Classifier (Using 2nd Type Likelihood Function) (Message 7: English)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.445	0.820	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6.1.15 Correct Classification Rate for Coherent Classifier (using 1st Type FM Likelihood Function) (Message 8: English Female)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.000	0.000	0.000	0.000	0.000	0.000	0.175	0.985	0.995	0.995	1.000

Table 6.1.16 Correct Classification Rate for Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 8: English Female)

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.945	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	0.995	0.995	0.995	0.995	0.995	1.000

6.4.2 ALRT-Based Non-Coherent Modulation Classification

In this sub-section, we investigate ALRT algorithms in non-coherent environment, where the received modulated signal is rotated by an unknown carrier phase θ_c. The residual carrier frequency ω_c is assumed to be perfectly estimated before attempting to classify the modulation types. The variables remaining in $\boldsymbol{\beta}_l$ are the real-valued information-bearing message sequence emitted by the transmitter, m_i, $i = 1, 2, \ldots, N$. The sequence of m_i, $i = 1, 2, \ldots, N$ is assumed to be independent and identically distributed, obeying the PDF of equation (6.12).

In the non-coherent environment, the signal is classified by finding the ratio of two adjacent received samples, which is defined as

$$\chi_i = \frac{r_i}{r_{i+1}} = \frac{s_i + n_i}{s_{i+1} + n_{i+1}}, i = 1, 2, \ldots, N-1 \tag{6.33}$$

χ_i can reject the common phase offset θ_c.

As defined in Section 6.2., n_i is the noise sample, which is a zero-mean Gaussian random variable with variance σ^2, i.e., $n_i \sim \mathcal{N}(0, \sigma^2)$; s_i represents the noiseless analog modulated signal. Hence r_i is Gaussian random variable, with mean s_i and variances σ^2. χ_i is the ratio of two Gaussian random variables.

The first part of this sub-section will discuss the distribution and density function of the ratio of two Gaussian random variables; the second part will deduce the likelihood functions of three analog modulation schemes, AM, DSB and FM.

6.4.2.1 The Ratio of Two Gaussian Variables In [82] and [83], the distribution of the ratio of two Gaussian random variables has been studied. We will give an explicit representation of this problem in this sub-section.

Proposition 1: Transforming $\frac{z_g}{w_g}$ to the standard form $\frac{a_g + x_g}{b_g + y_g}$ [83]

For any two jointly Gaussian varieties z_g and w_g with means μ_{zg} and μ_{wg}, variances σ_{zg}^2 and σ_{wg}^2, and correlation ρ_g, the distribution of $\frac{z_g}{w_g}$ is, after translation and change of scale, the same as that of $\frac{a_g + x_g}{b_g + y_g}$ with x_g, y_g independent standard Gaussian, and a_g, b_g non-negative constants.

Specifically, for a given ratio $\frac{z_g}{w_g}$, there are constants r_g and s_g such that $r_g\left(\frac{z_g}{w_g} - s_g\right)$ is distributed as $\frac{a_g + x_g}{b_g + y_g}$, and $\frac{z_g}{w_g}$ is distributed as $\frac{1}{r_g}\left(\frac{a_g + x_g}{b_g + y_g}\right) + s_g$.

The procedure to choose the constants r_g, s_g, a_g and b_g is as following:

(1) Let

$$s_g = \rho_g \cdot \frac{\sigma_{zg}}{\sigma_{wg}} \tag{6.34}$$

(2) Let

$$b_g = \frac{\mu_{wg}}{\sigma_{wg}} \tag{6.35}$$

(3) With

$$h_g = \pm\sigma_{zg}\sqrt{1-\rho_g^2} \tag{6.36}$$

and, let

$$a_g = \pm\frac{\left(\mu_{zg} - s_g\mu_{wg}\right)}{h_g} \tag{6.37}$$

We need choose the sign of h_g so that the resulting a_g and b_g have the same sign.

(4) Let

$$r_g = \frac{\sigma_{wg}}{h_g} \tag{6.38}$$

Proposition 2: Distribution of the standard form $\frac{a_g + x_g}{b_g + y_g}$ [82]

Let $t_g = \frac{a_g + x_g}{b_g + y_g}$, the density of $\frac{a_g + x_g}{b_g + y_g}$ can be written as

$$p\left(t_g\right) = \frac{e^{-\frac{1}{2}\left(a_g^2 + b_g^2\right)}}{\pi(1+t_g^2)}\left[1 + q_g e^{\frac{1}{2}q_g^2}\int_0^{q_g} e^{\frac{1}{2}x_g^2} dx_g\right], \quad q_g = \frac{b_g + a_g t_g}{\sqrt{1+t_g^2}}. \tag{6.39}$$

6.4.2.2 Likelihood function for AM modulation scheme The noiseless AM signal s_i in non-coherent environment can be written as

$$s_i = \left(1 + \mu m_i\right) \cdot \cos\left(\theta_c\right) \tag{6.40}$$

Hence, the ratio between two adjacent received samples χ_i can be represented as

$$\begin{aligned} \chi_i &= \frac{r_i}{r_{i+1}} \\ &= \frac{s_i + n_i}{s_{i+1} + n_{i+1}} \\ &= \frac{\left(1 + \mu m_i\right) \cdot \cos\left(\theta_c\right) + n_i}{\left(1 + \mu m_{i+1}\right) \cdot \cos\left(\theta_c\right) + n_{i+1}} \\ &\approx \frac{\left(1 + \mu m_i\right) + n_i}{\left(1 + \mu m_{i+1}\right) + n_{i+1}}, \quad i = 1, 2, \ldots, N-1 \end{aligned} \tag{6.41}$$

The ratio between two adjacent received samples rejects the common phase offset θ_c, and the modulation parameter μ is estimated before classification. The only unknown parameter in equation (6.41) is $m_i, \quad i = 1, 2, \ldots, N-1$, therefore the multiple dimensional parameter space $\boldsymbol{\beta}_1$ now is degenerated to a one dimensional parameter space, i.e., $\boldsymbol{\beta}_1 = \left\{ m_i, i = 1, 2, \ldots, N-1 \right\}$. Let

$$s_i\left(\boldsymbol{\beta}_1\right) = 1 + \mu m_i, \quad i = 1, 2, \ldots, N-1 \tag{6.42}$$

and substituting equation (6.42) into equation (6.41), χ_i can be rewritten as,

$$\chi_i \approx \frac{\left(1 + \mu m_i\right) + n_i}{\left(1 + \mu m_{i+1}\right) + n_{i+1}} = \frac{s_i\left(\boldsymbol{\beta}_1\right) + n_i}{s_{i+1}\left(\boldsymbol{\beta}_1\right) + n_{i+1}}, \quad \boldsymbol{\beta}_1 = \left\{ m_i, i = 1, 2, \ldots, N-1 \right\} \tag{6.43}$$

Let

$$z_g\left(\boldsymbol{\beta}_1\right) = s_i\left(\boldsymbol{\beta}_1\right) + n_i \tag{6.44}$$

$$w_g\left(\boldsymbol{\beta}_1\right) = s_{i+1}\left(\boldsymbol{\beta}_1\right) + n_{i+1}, \tag{6.45}$$

and substituting equations (6.44) and (6.45) into equation (6.43), χ_i can be rewritten as,

$$\chi_i \approx \frac{(1+\mu m_i)+n_i}{(1+\mu m_{i+1})+n_{i+1}} = \frac{s_i(\boldsymbol{\beta}_1)+n_i}{s_{i+1}(\boldsymbol{\beta}_1)+n_{i+1}} = \frac{z_g(\boldsymbol{\beta}_1)}{w_g(\boldsymbol{\beta}_1)}, \quad \boldsymbol{\beta}_1 = \{ m_i, i=1,2,\ldots,N-1 \} \tag{6.46}$$

χ_i is the ratio of two Gaussian varieties $z_g(\boldsymbol{\beta}_1)$ and $w_g(\boldsymbol{\beta}_1)$ with means $\mu_{zg}(\boldsymbol{\beta}_1) = s_i(\boldsymbol{\beta}_1) = 1+\mu m_i$ and $\mu_{wg}(\boldsymbol{\beta}_1) = s_{i+1}(\boldsymbol{\beta}_1) = 1+\mu m_{i+1}$, and variances $\sigma_{zg}^2(\boldsymbol{\beta}_1) = \sigma_{wg}^2(\boldsymbol{\beta}_1) = \sigma^2$. Because the $n_i, i=1,2,\ldots,N-1$ are statistically independent, the correlation of $z_g(\boldsymbol{\beta}_1)$ and $w_g(\boldsymbol{\beta}_1)$ is zero, i.e., $\rho_g(\boldsymbol{\beta}_1) = 0$.

By substituting $\mu_{zg}(\boldsymbol{\beta}_1)$, $\mu_{wg}(\boldsymbol{\beta}_1)$, $\sigma_{zg}^2(\boldsymbol{\beta}_1)$, $\sigma_{wg}^2(\boldsymbol{\beta}_1)$ and $\rho_g(\boldsymbol{\beta}_1)$ into equations (6.34) ~(6.38),

$$s_g(\boldsymbol{\beta}_1) = \rho_g(\boldsymbol{\beta}_1) \cdot \frac{\sigma_{zg}(\boldsymbol{\beta}_1)}{\sigma_{wg}(\boldsymbol{\beta}_1)} = 0 \tag{6.47}$$

$$b_g(\boldsymbol{\beta}_1) = \frac{\mu_{wg}(\boldsymbol{\beta}_1)}{\sigma_{wg}(\boldsymbol{\beta}_1)} = \frac{1+\mu m_{i+1}}{\sigma} \tag{6.48}$$

$$h_g(\boldsymbol{\beta}_1) = \pm\sigma_{zg}(\boldsymbol{\beta}_1)\sqrt{1-\rho_g(\boldsymbol{\beta}_1)^2} = \pm\sigma \tag{6.49}$$

$$a_g(\boldsymbol{\beta}_1) = \pm\frac{\left(\mu_{zg}(\boldsymbol{\beta}_1) - s_g(\boldsymbol{\beta}_1)\mu_{wg}(\boldsymbol{\beta}_1)\right)}{h_g(\boldsymbol{\beta}_1)} = \pm\frac{(1+\mu m_i)}{\sigma} \tag{6.50}$$

$$r_g(\boldsymbol{\beta}_1) = \frac{\sigma_{wg}(\boldsymbol{\beta}_1)}{h_g(\boldsymbol{\beta}_1)} = \frac{\sigma}{\pm\sigma} = \pm 1 \tag{6.51}$$

and thus, the *i*th ratio, $\chi_i \approx \frac{z_g(\boldsymbol{\beta}_1)}{w_g(\boldsymbol{\beta}_1)}$, can be transerfered to the standand form, accourding to Proposition 1, as

$$\begin{aligned}
&\frac{z_g(\boldsymbol{\beta}_1)}{w_g(\boldsymbol{\beta}_1)} \\
&= \frac{1}{r_g(\boldsymbol{\beta}_1)}\left(\frac{a_g(\boldsymbol{\beta}_1)+x_g}{b_g(\boldsymbol{\beta}_1)+y_g}\right)+s_g(\boldsymbol{\beta}_1) \\
&= \pm\left(\frac{a_g(\boldsymbol{\beta}_1)+x_g}{b_g(\boldsymbol{\beta}_1)+y_g}\right)
\end{aligned} \tag{6.52}$$

From equations (6.47) to (6.51), we can see that the only unknown parameter in $r_g(\boldsymbol{\beta}_1)$, $s_g(\boldsymbol{\beta}_1)$, $a_g(\boldsymbol{\beta}_1)$ and $b_g(\boldsymbol{\beta}_1)$ is $\{m_i, i=1,2,\ldots,N-1\}$. Therefore, the distribution of the *i*th ratio χ_i under hypothesis H_1, $p(\chi_i|H_1,\boldsymbol{\beta}_1)$ can be found by (6.39) and (6.52), and is a function with the unknown parameter space $\boldsymbol{\beta}_1 = \{m_i, i=1,2,\ldots,N-1\}$.

Then, the likelihood function in equation (6.8) become the average of m_i from -1 to 1,

$$\begin{aligned}
p_a(\chi_i|H_1) &= \int_{\boldsymbol{\beta}_1} p(\chi_i|H_1,\boldsymbol{\beta}_1)\cdot p(\boldsymbol{\beta}_1|H_1)\, d\boldsymbol{\beta}_1 \\
&= \int_{-1}^{1} p(\chi_i|H_1,m_i)\cdot\frac{1}{\sqrt{2\pi}\cdot\sigma_m}\cdot\exp\left(-\frac{(m_i)^2}{2\sigma_m^2}\right) dm_i
\end{aligned} \tag{6.53}$$

The integral in equation (6.53) does not admitany closed form solution. However, the integral in equation (6.53) can be resolved by numerical integrations.

Given N observed noise corrupted signal samples, $\mathbf{R} = [r_1, r_2, r_3, \ldots, r_N]^T$, the likelihood function of $\boldsymbol{\chi} = [\chi_1, \chi_2, \chi_3, \ldots, \chi_{N-1}]^T$ under H_l is

$$\Gamma_a(\boldsymbol{\chi}|H_l) = \prod_{i=1}^{N-1} p_a(\chi_i|H_l) \tag{6.54}$$

Therefore, substituting equation (6.53) into equation (6.54), we can easily find the likelihood function of $\boldsymbol{\chi}$ under H_1.

6.4.2.3 Likelihood function for DSB modulation scheme The derivation process of the likelihood function for DSB modulation scheme is similar to that for AM modulation scheme; therefore we simply follow the process given above:

The noiseless DSB signal s_i in non-coherent environment can be written as

$$s_i = m_i \cdot \cos(\theta_c) \tag{6.55}$$

Hence, the ratio between two adjacent received samples χ_i can be represented as

$$\begin{aligned} \chi_i &= \frac{r_i}{r_{i+1}} \\ &= \frac{s_i + n_i}{s_{i+1} + n_{i+1}} \\ &= \frac{m_i \cdot \cos(\theta_c) + n_i}{m_{i+1} \cdot \cos(\theta_c) + n_{i+1}} \\ &\approx \frac{m_i + n_i}{m_{i+1} + n_{i+1}}, \quad i = 1, 2, \ldots, N-1 \end{aligned} \tag{6.56}$$

Similarly, the ratio between two adjacent received samples rejects the common phase offset θ_c, and the modulation parameter μ is estimated before classification. The only unknown parameter in equation (6.56) is m_i, $i = 1, 2, \ldots, N-1$, therefore the multiple dimensional parameter space $\boldsymbol{\beta}_1$ now is degenerated to a one dimensional parameter space, i.e., $\boldsymbol{\beta}_1 = \{ m_i, i = 1, 2, \ldots, N-1 \}$. Let

$$s_i(\boldsymbol{\beta}_2) = m_i, \quad i = 1, 2, \ldots, N-1 \tag{6.57}$$

and substituting equation (6.57) into equation (6.56), χ_i can be rewritten as,

$$\chi_i \approx \frac{m_i + n_i}{m_{i+1} + n_{i+1}} = \frac{s_i(\boldsymbol{\beta}_2) + n_i}{s_{i+1}(\boldsymbol{\beta}_2) + n_{i+1}}, \qquad \boldsymbol{\beta}_1 = \{ m_i, i = 1, 2, \ldots, N-1 \} \tag{6.58}$$

Let

$$z_g(\boldsymbol{\beta}_2) = s_i(\boldsymbol{\beta}_2) + n_i \tag{6.59}$$

$$w_g(\boldsymbol{\beta}_2) = s_{i+1}(\boldsymbol{\beta}_2) + n_{i+1}, \tag{6.60}$$

and substituting equations (6.59) and (6.60) into equation (6.58), χ_i can be rewritten as,

$$\chi_i \approx \frac{m_i + n_i}{m_{i+1} + n_{i+1}} = \frac{s_i(\boldsymbol{\beta}_2) + n_i}{s_{i+1}(\boldsymbol{\beta}_2) + n_{i+1}} = \frac{z_g(\boldsymbol{\beta}_2)}{w_g(\boldsymbol{\beta}_2)}, \quad \boldsymbol{\beta}_2 = \{ m_i, i = 1, 2, \ldots, N-1 \} \tag{6.61}$$

χ_i is the ratio of two Gaussian varieties $z_g(\boldsymbol{\beta}_2)$ and $w_g(\boldsymbol{\beta}_2)$ with means $\mu_{zg}(\boldsymbol{\beta}_2) = s_i(\boldsymbol{\beta}_2) = m_i$ and $\mu_{wg}(\boldsymbol{\beta}_2) = s_{i+1}(\boldsymbol{\beta}_2) = m_{i+1}$, variances $\sigma_{zg}^2(\boldsymbol{\beta}_2) = \sigma_{wg}^2(\boldsymbol{\beta}_2) = \sigma^2$, and the correlation $\rho_g(\boldsymbol{\beta}_2) = 0$.

Substituting $\mu_{zg}(\boldsymbol{\beta}_2)$, $\mu_{wg}(\boldsymbol{\beta}_2)$, $\sigma_{zg}^2(\boldsymbol{\beta}_2)$, $\sigma_{wg}^2(\boldsymbol{\beta}_2)$ and $\rho_g(\boldsymbol{\beta}_2)$ into (6.34) ~ (6.39), the conditional probability of the ith ratio χ_i under hypothesis H_2, $p(\chi_i | H_2, \boldsymbol{\beta}_2)$, can be found with the unknown parameter space $\boldsymbol{\beta}_2 = \{ m_i \}$.

The likelihood function in equation (6.8) become the average of m_i from -1 to 1, then we have,

$$\begin{aligned} p_a(\chi_i | H_2) &= \int_{\boldsymbol{\beta}_2} p(\chi_i | H_2, \boldsymbol{\beta}_2) \cdot p(\boldsymbol{\beta}_2 | H_2)\, d\boldsymbol{\beta}_2 \\ &= \int_{-1}^{1} p(\chi_i | H_2, m_i) \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma_m} \cdot \exp\left(-\frac{(m_i)^2}{2\sigma_m^2} \right) dm_i, \quad \boldsymbol{\beta}_2 = \{ m_i \} \end{aligned} \tag{6.62}$$

The likelihood function of $\boldsymbol{\chi} = [\chi_1, \chi_2, \chi_3, \ldots, \chi_{N-1}]^T$ under H_2 can be easily found by substituting equation (6.62) into equation (6.54).

6.4.2.4 Likelihood function for FM modulation scheme The derivation of the likelihood function for FM modulation scheme is more involved. The noiseless FM signal s_i in non-coherent environment can be written as

$$s_i = \cos\left(2\pi f_\Delta \cdot \sum_{k=1}^{i} \left(m_k \cdot \frac{1}{f_s} \right) + \theta_c \right) \tag{6.63}$$

The sequence of m_k, $k = 1, 2, \ldots, i$ is independent and identically distributed, obeying the PDF of (6.12). Let y_i be the cumulative sum of m_k, $k = 1, 2, \ldots, i$, i.e.,

$$y_i = 2\pi f_\Delta \cdot \sum_{k=1}^{i} \left(m_k \cdot \frac{1}{f_s} \right) + \theta_c \tag{6.64}$$

Then y_i is normally distributed with mean θ_c and variance $\sigma_y^2 = i \cdot \left(2\pi \cdot \frac{f_\Delta}{f_s} \right)^2 \cdot \sigma_m^2$, i.e., $y_i \sim \mathcal{N}(\theta_c, \sigma_y^2)$, $i = 1, 2, \ldots, N$.

We further let z_i be the modulus of y_i after division by 2π, i.e.,

$$z_i = \mathrm{MOD}[\, y_i, 2\pi \,] \tag{6.65}$$

The PDF of z_i is given by [81], and it proved that for large σ_y^2, $p(z_i)$ tends to the uniform distribution, i.e.,

$$p(z_i) \approx \frac{1}{2\pi}, \quad 0 < z_i \le 2\pi \tag{6.66}$$

And we always assume that this is the case in this chapter.

Hence, the ratio between two adjacent received samples χ_i can be represented as

$$\begin{aligned}\chi_i &= \frac{r_i}{r_{i+1}} \\ &= \frac{s_i + n_i}{s_{i+1} + n_{i+1}} \\ &= \frac{\cos(y_i) + n_i}{\cos(y_{i+1}) + n_{i+1}} \\ &= \frac{\cos(z_i) + n_i}{\cos(z_{i+1}) + n_{i+1}}, \quad i = 1, 2, \ldots, N-1\end{aligned} \tag{6.67}$$

Then the likelihood function for FM modulation scheme can be formulated in two different ways, which lead to the first type of likelihood function (1st-LF) and the second type of likelihood function (2nd-LF) for FM Modulation schemes. Both two types of likelihood function are tested by the experiments.

6.4.2.4.1 The First Type of Likelihood Function for FM Modulation (1st -LF) The only unknown parameter in (6.67) is $z_i, i = 1, 2, \ldots, N-1$, therefore the multiple dimensional parameter space $\boldsymbol{\beta}_3$ now is degenerated to a one dimensional parameter space, i.e., $\boldsymbol{\beta}_3 = \{ z_i, i = 1, 2, \ldots, N-1 \}$. Let

$$s_i(\boldsymbol{\beta}_3) = \cos(z_i), \quad \boldsymbol{\beta}_3 = \{ z_i, i = 1, 2, \ldots, N-1 \} \tag{6.68}$$

and substituting equation (6.68) into equation (6.67), χ_i can be rewritten as,

$$\chi_i = \frac{\cos(z_i) + n_i}{\cos(z_{i+1}) + n_{i+1}} = \frac{s_i(\boldsymbol{\beta}_3) + n_i}{s_{i+1}(\boldsymbol{\beta}_3) + n_{i+1}} = \frac{z_g(\boldsymbol{\beta}_3)}{w_g(\boldsymbol{\beta}_3)}, \quad \boldsymbol{\beta}_3 = \{ z_i \} \tag{6.69}$$

χ_i is the ratio of two Gaussian varieties $z_g(\boldsymbol{\beta}_3)$ and $w_g(\boldsymbol{\beta}_3)$ with means $\mu_{zg}(\boldsymbol{\beta}_3) = s_i(\boldsymbol{\beta}_3) = \cos(z_i)$ and $\mu_{wg}(\boldsymbol{\beta}_3) = s_{i+1}(\boldsymbol{\beta}_3) = \cos(z_{i+1})$, variances $\sigma^2_{zg}(\boldsymbol{\beta}_3) = \sigma^2_{wg}(\boldsymbol{\beta}_3) = \sigma^2$ and correlation $\rho_g(\boldsymbol{\beta}_3) = 0$.

Substituting $\mu_{zg}(\boldsymbol{\beta}_3)$, $\mu_{wg}(\boldsymbol{\beta}_3)$, $\sigma^2_{zg}(\boldsymbol{\beta}_3)$, $\sigma^2_{wg}(\boldsymbol{\beta}_3)$ and $\rho_g(\boldsymbol{\beta}_3)$ into equations (6.34) ~ (6.39), the conditional probability of the *i*th ratio χ_i under hypothesis H_3, can be found as $p(\chi_i|H_3,\boldsymbol{\beta}_3)$ with the unknown parameter space $\boldsymbol{\beta}_3 = \{ z_i \}$. The likelihood function in (6.8) become the average of z_i from 0 to 2π, then we have,

$$\begin{aligned} & p_a(\chi_i|H_3) \\ & \quad = \int_{\boldsymbol{\beta}_3} p(\chi_i|H_3,\boldsymbol{\beta}_3)\cdot p(\boldsymbol{\beta}_3|H_3)\, d\boldsymbol{\beta}_3 \\ & \quad = \int_0^{2\pi} p(\chi_i|H_3,z_i)\cdot\frac{1}{2\pi}\, dz_i, \quad \boldsymbol{\beta}_3 = \{ z_i \} \end{aligned} \tag{6.70}$$

The likelihood function of $\boldsymbol{\chi} = [\chi_1, \chi_2, \chi_3, \ldots, \chi_{N-1}]^T$ under H_3 can be easily found by substituting equation (6.70) into equation (6.54).

6.4.2.4.2 The Second Type of Likelihood Function for FM Modulation (2nd -LF)

Let

$$w_i = \cos(z_i), \quad 0 < z_i \le 2\pi \tag{6.71}$$

then the PDF of w_i is

$$p(w_i) = \frac{1}{\pi}\cdot\frac{1}{\sqrt{1-w_i^2}}, \quad -1 < w_i < 1 \tag{6.72}$$

Substituting equation (6.71) into equation (6.67), χ_i can be rewritten as,

$$\chi_i = \frac{\cos(z_i)+n_i}{\cos(z_{i+1})+n_{i+1}} = \frac{w_i+n_i}{w_{i+1}+n_{i+1}} \tag{6.73}$$

The only unknown parameter in equation (6.73) is w_i, therefore the multiple dimensional parameter space $\boldsymbol{\beta}_3$ now is degenerated to a one dimensional parameter space, i.e., $\boldsymbol{\beta}_3 = \{ w_i \}$. Let

$$s_i(\boldsymbol{\beta}_3) = s_i(w_i) = w_i, \quad \boldsymbol{\beta}_3 = \{ w_i \} \tag{6.74}$$

and substituting equation (6.74) into equation (6.73), χ_i can be rewritten as,

$$\chi_i = \frac{w_i + n_i}{w_{i+1} + n_{i+1}} = \frac{s_i(\boldsymbol{\beta}_3) + n_i}{s_{i+1}(\boldsymbol{\beta}_3) + n_{i+1}} = \frac{z_g(\boldsymbol{\beta}_3)}{w_g(\boldsymbol{\beta}_3)}, \quad \boldsymbol{\beta}_3 = \{ w_i \} \tag{6.75}$$

χ_i is the ratio of two Gaussian varieties $z_g(\boldsymbol{\beta}_3)$ and $w_g(\boldsymbol{\beta}_3)$ with means $\mu_{zg}(\boldsymbol{\beta}_3) = s_i(\boldsymbol{\beta}_3) = w_i$ and $\mu_{wg}(\boldsymbol{\beta}_3) = s_{i+1}(\boldsymbol{\beta}_3) = w_{i+1}$, variances $\sigma_{zg}^2(\boldsymbol{\beta}_3) = \sigma_{wg}^2(\boldsymbol{\beta}_3) = \sigma^2$ and correlation $\rho_g(\boldsymbol{\beta}_3) = 0$.

Substituting $\mu_{zg}(\boldsymbol{\beta}_3)$, $\mu_{wg}(\boldsymbol{\beta}_3)$, $\sigma_{zg}^2(\boldsymbol{\beta}_3)$, $\sigma_{wg}^2(\boldsymbol{\beta}_3)$ and $\rho_g(\boldsymbol{\beta}_3)$ into (6.34) ~ (6.39), the conditional probability of the *i*th ratio χ_i under hypothesis H_3, can be found as $p(\chi_i | H_3, \boldsymbol{\beta}_3)$ with the unknown parameter space $\boldsymbol{\beta}_3 = \{ w_i \}$. The likelihood function in equation (6.8) become the average of w_i from -1 to 1, then we have,

$$\begin{aligned} & p_a(\chi_i | H_3) \\ & \quad = \int_{\boldsymbol{\beta}_3} p(\chi_i | H_3, \boldsymbol{\beta}_3) \cdot p(\boldsymbol{\beta}_3 | H_3)\, d\boldsymbol{\beta}_3 \\ & \quad = \int_0^{2\pi} p(\chi_i | H_3, w_i) \cdot \frac{1}{\pi} \cdot \frac{1}{\sqrt{1 - w_i^2}}\, dw_i, \quad \boldsymbol{\beta}_3 = \{ w_i \} \end{aligned} \tag{6.76}$$

The likelihood function of $\boldsymbol{\chi} = [\chi_1, \chi_2, \chi_3, \ldots, \chi_{N-1}]^T$ under H_3 can be easily found by substituting equation (6.76) into equation (6.54).

6.4.2.5 Performance Evaluation for the Non-Coherent Classifier

The non-coherent classifier is test by simulated signals. Detailed description of the simulated signals has been presented in Sub-Sub-Section 6.4.1.4

- Tables 6.2.1~6.2.16 report the test results when the phase offset θ_c is set to be 0, while SNR is from -10dB to 10dB, With a 2dB increment.
- Tables 6.3.1~6.3.16 report the test results when SNR is set to be 10dB, while the phase offset θ_c is from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$, with a $\frac{\pi}{20}$ increment .

Results in Tables 6.2.1~6.2.16 show that:

- The classifier with 1st-LF can achieve more than 80% detection rate for AM and DSB, when SNR is as low as -10dB. But it fails for FM, when SNR is lower than 4dB or 6dB.
- When SNR is high (larger than 4dB), the decisions form both 1st-LF classifier and 2nd-LF classifier are trustable and robust. Therefore, we can choose either one in practice.
- Comparing Tables 6.2.1~6.2.16 with Tables 6.1.1~6.1.16, the testing results indicate that the non-coherent classifier exhibits a performance loss of approximately 2dB to 3dB.

Results in Tables 6.3.1~6.3.16 show that:

- The non-coherent classifier can achieve more than 80% detection rate, when the absolute value of phase offset θ_c is smaller than $\frac{7}{20}\pi$.

Table 6.2.1 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 1: Recorded Wave), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.955	0.970	0.975	0.975	0.985	0.990	0.995	1.000	1.000	1.000	1.000
DSB	0.830	0.885	0.930	0.935	0.940	0.950	0.950	0.950	0.950	0.960	0.965
FM	0.050	0.045	0.035	0.030	0.025	0.025	0.025	0.065	1.000	1.000	1.000

Table 6.2.2 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 1: Recorded Wave), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.420	0.935	0.975	0.985	0.985	0.995	1.000	1.000	1.000	1.000
DSB	0.240	0.570	0.770	0.870	0.925	0.930	0.935	0.945	0.945	0.945	0.960
FM	1.000	1.000	1.000	0.460	0.095	0.055	0.050	0.700	1.000	1.000	1.000

Table 6.2.3 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 2: Music Rock), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.985	0.990	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000

Table 6.2.4 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 2: Music Rock), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.825	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.715	0.895	0.970	0.985	0.990	1.000	1.000	0.990	0.990	0.990	0.995
FM	1.000	1.000	0.995	0.135	0.005	0.000	0.000	0.555	1.000	1.000	1.000

Table 6.2.5 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 3: Music Classic), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.975	0.990	0.990	0.995	0.995	0.995	1.000	1.000	1.000	1.000	1.000
FM	0.005	0.005	0.000	0.000	0.000	0.000	0.000	0.015	0.965	1.000	1.000

Table 6.2.6 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 3: Music Classic), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.890	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.835	0.905	0.965	0.985	0.990	0.995	0.995	0.995	0.995	0.995	1.000
FM	1.000	1.000	0.975	0.090	0.010	0.000	0.005	0.530	1.000	1.000	1.000

Table 6.2.7 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 4: Singing Classic), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.990	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.960	0.965	0.975	0.975	0.975	0.975	0.975	0.975	0.975	0.985	0.985
FM	0.020	0.015	0.010	0.005	0.000	0.000	0.005	0.095	0.655	0.835	0.915

Table 6.2.8 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 4: Singing Classic), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.785	0.975	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.605	0.865	0.935	0.960	0.970	0.970	0.970	0.970	0.970	0.975	0.980
FM	1.000	1.000	0.820	0.175	0.045	0.035	0.090	0.345	0.755	0.870	0.920

Table 6.2.9 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 5: Arabic), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995
FM	0.000	0.000	0.000	0.000	0.000	0.025	0.060	0.160	0.465	0.805	0.975

Table 6.2.10 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 5: Arabic), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995
FM	1.000	1.000	0.645	0.120	0.070	0.075	0.130	0.295	0.590	0.905	0.985

Table 6.2.11 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 6: Chinese), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.900
FM	0.000	0.000	0.000	0.000	0.005	0.015	0.065	0.240	0.550	0.880	0.980

Table 6.2.12 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 6: Chinese), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.900
FM	1.000	1.000	0.660	0.195	0.095	0.120	0.210	0.375	0.700	0.915	0.995

Table 6.2.13 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 7: English), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.000	0.000	0.000	0.000	0.000	0.015	0.090	0.225	0.610	0.945	0.995

Table 6.2.14 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 7: English), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	0.985	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.955	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	0.805	0.205	0.100	0.120	0.210	0.375	0.800	0.965	1.000

Table 6.2.15 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 8: English Female), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.970
FM	0.000	0.000	0.000	0.000	0.000	0.005	0.010	0.140	0.770	0.945	0.980

Table 6.2.16 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 8: English Female), $\theta_c = 0$

	SNR(dB)										
	-10	**-8**	**-6**	**-4**	**-2**	**0**	**2**	**4**	**6**	**8**	**10**
AM	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.970
FM	1.000	1.000	0.875	0.080	0.025	0.030	0.080	0.410	0.895	0.955	0.990

Table 6.3.1 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 1: Recorded Wave), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	0.985	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	0.995	0.995	0.985	0.985	0.980	0.975	0.970	0.965	0.965
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	0.985	0.000	0.000	0.000	0.000
DSB	0.965	0.965	0.970	0.975	0.980	0.985	0.985	0.995	0.995	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6.3.2 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 1: Recorded Wave), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	0.985	0.995	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	0.995	0.995	0.985	0.985	0.980	0.975	0.970	0.965	0.960
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	0.995	0.985	0.000	0.000	0.000	0.000
DSB	0.960	0.965	0.970	0.975	0.980	0.985	0.985	0.995	0.995	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6.3.3 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 2: Music Rock), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6.3.4 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 2: Music Rock), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6.3.5 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 3: Music Classic), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970
FM	0.995	0.995	1.000	1.000	1.000	0.995	0.980	0.985	0.975	0.980	0.980

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970
FM	0.980	0.990	0.985	0.980	0.975	0.980	0.980	0.990	1.000	1.000	0.995

Table 6.3.6 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 3: Music Classic), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6.3.7 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 4: Singing Classic), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.990	0.990	0.985
FM	1.000	1.000	1.000	0.995	0.995	0.995	0.965	0.950	0.930	0.920	0.915

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	0.985	0.990	0.990	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.915	0.925	0.930	0.950	0.965	0.980	0.990	1.000	1.000	1.000	1.000

Table 6.3.8 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 4: Singing Classic), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.990	0.990	0.980	0.980
FM	1.000	1.000	1.000	0.995	0.995	0.995	0.975	0.960	0.940	0.930	0.920

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	0.980	0.980	0.990	0.990	0.995	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.920	0.940	0.945	0.955	0.970	0.980	0.990	1.000	1.000	1.000	1.000

Table 6.3.9 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 5: Arabic), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995
FM	0.995	0.995	0.980	0.980	0.990	0.990	0.985	0.980	0.965	0.970	0.975

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995
FM	0.975	0.980	0.975	0.980	0.975	0.975	0.990	0.995	1.000	1.000	0.995

Table 6.3.10 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 5: Arabic), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995
FM	1.000	0.995	0.985	0.990	0.990	0.995	0.990	0.980	0.975	0.980	0.985

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995
FM	0.985	0.980	0.990	0.990	0.985	0.995	1.000	1.000	1.000	1.000	1.000

Table 6.3.11 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 6: Chinese), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
FM	0.995	0.990	0.985	0.985	0.990	0.985	0.965	0.970	0.970	0.980	0.980

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
FM	0.980	1.000	0.995	0.990	0.980	0.975	0.980	0.990	0.990	0.995	0.995

Table 6.3.12 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 6: Chinese), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
FM	0.995	0.995	0.995	0.995	0.990	0.990	0.980	0.975	0.980	0.985	0.995

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
FM	0.995	0.985	1.000	1.000	0.995	0.990	0.990	0.990	0.990	0.995	0.995

Table 6.3.13 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 7: English), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	1.000	0.995

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	0.995	0.985	0.990	0.985	0.995	0.995	1.000	0.995	0.995	0.995	0.995

Table 6.3.14 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 7: English), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FM	1.000	0.995	0.995	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 6.3.15 Correct Classification Rate for Non-Coherent Classifier (using 1st Type FM Likelihood Function) (Message 8: English Female), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970
FM	0.995	0.995	1.000	1.000	1.000	0.995	0.980	0.985	0.975	0.980	0.980

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970
FM	0.980	0.990	0.985	0.980	0.975	0.980	0.980	0.990	1.000	1.000	0.995

Table 6.3.16 Correct Classification Rate for Non-Coherent Classifier (Using 2nd Type FM Likelihood Function) (Message 8: English Female), SNR = 10dB

	Phase Offset (pi/20)										
	-10	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
AM	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DSB	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970
FM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.990	0.980	0.980	0.990

	Phase Offset (pi/20)										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
DSB	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970	0.970
FM	0.990	0.995	0.995	0.990	0.985	0.985	0.985	0.995	1.000	1.000	1.000

6.5 Conclusions

In this chapter we investigate automatic modulation classification for analog modulation schemes. We present a likelihood ratio testing framework for both the coherent and non-coherent cases. The performance of both the coherent and non-coherent classifier is investigated for various modulation types including AM, DSB and FM. Non-coherent performance exhibites a 2~3dB loss compared to coherent performance.

ALRT-based criterion is conceptually clear, but needs the prior knowledge of PDF of unknown parameters. If we do not have any prior knowledge of unknown parameters, then no probability density over which to average, we cannot use the results from this chapter. There is no general approach that is always successful, but to make the AMC algorithm more practical, these should be studied in the future research.

CHAPTER 7

DISCUSSIONS AND SUMMARY

In this chapter, the AMC algorithms presented in Chapters 4, 5 and 6 are first evaluated and discussed from system point of view. A summary of this disstertation is then made.

7.1 Evaluation and Discussions of Proposed AMC Algorithms

Different AMC algorithms for analog modulation schemes are proposed and investigated in Chapters 4, 5 and 6, respectively. It is clear that not a single AMC algorithm can be of universal application to every type of analog modulation schemes with good performance. Hence, a further discussion on the characteristics of these AMC algorithms from system point of view is needed.

Table 7.1 gives a general comparison and evaluation of the AMC algorithms proposed in this dissertation, which lists the modulation types that each algorithm can be covered, the SNR range when the detection rate is higher than 90%, and the performance (detection rate) when SNR is at 10dB.

From Table 7.1 we can gain some insight into when and where each algorithm is applicable. The discussions of proposed AMC algorithm are from two aspects:

(1) According to possible input modulation types:

- Cyclostationarity classifier covers more modulation types than the other three classifiers, and can be used to separate between analog modulations and digital modulations.

- When the modulation types have been narrowed down to the set of analog modulations, SVM classifier prominently outperforms the other three classifiers. Especially for LSB and USB modulation, SVM classifier can achieve high classification accuracy in a wide range of SNR, while neither the 1st type LRT classifier nor the 2nd type LRT classifier is applicable to LSB and USB classification.
- However, if the modulation types can be further narrowed down to the set of {AM, DSB and FM}, 2nd type-LRT classifier is a better choice than SVM classifier in the sense that LRT classifier can avoid training procedure.

(2) According to SNR conditions:

- When SNR is higher than 4dB, the decisions made by all of the four classifiers are trustable and robust. Therefore, one may choose any one of the technologies developed in practice.
- When SNR is in the range from 2dB to 4dB, the decisions from cyclostationarity classifier, SVM classifier and 2nd-type LRT classifier can be trustable. While, for 1st-type LRT classifier, most of FM signals are mis-classified as DSB.
- When SNR is in the range from -4dB to 2dB, both SVM classifier and 2nd-type LRT classifier should be used in order to increase detection rate and reduce mis-classification rate.
- When SNR is lower than -4dB, the only workable classifer is SVM classifier.

A system with logic decisions on which algorithms can be used, either parallerly or serially, under different condition and situation, might siginificantly boost the performancne or robutness of AMC.

Table 7.1 Summary of the AMC Algorithms Proposed in This Dissertation

AMC Algorithms	Modulation Types that can be recognized	SNR Range as detection rate ≥ 90%	Performance as SNR=10dB
Cyclostationarity	AM, DSB, LSB, USB, FM, MASK, MPAM, MPSK, QAM	2dB and Higher	Higher than 94.5%
SVM	AM, DSB, LSB, USB, FM	-4~16dB (if trained in -4~14dB)	Higher than 99%
1st type-LRT	AM, DSB, FM	4dB and Higher	Higher than 98.5%
2nd type-LRT	AM, DSB, FM	-4dB and Higher	Higher than 98.5%

- This table listed the experiment result on simulated data generated by Matlab.
- Sequences of 0.6667 seconds (29,400 samples) were generated for the analog modulation signals, and sequences of 2,000 symbols were generated for the digital modulation signals

7.2 Summary

In this dissertation the blind recognition of analog modulation schemes for software defined radio receiver has been investigated.

An automatic modulation classification (AMC) algorithm for classification of joint analog and digital modulations based on the work of [40] has been presented in Chapter 4. The cyclostationarity proposed decision tree has demonstrated promising performance with both simulated data and field data. The proposed classifier does not impose unreasonable constraints on the input signal, and thus is practical in non-cooperation communication environment.

The development of a support-vector-machine-based AMC algorithm is presented in Chapter 5 for blind classification of analog modulation signals. The histogram of instantaneous frequency is used as the modulation-classification features. Support Vector Machines (SVMs) are used as the classifier. Simulation results have shown that the proposed classification method is quite robust to additive noise in a wide range of SNR.

Average-Likelihood-Ratio-Testing-based AMC algorithms presented in Chapter 6 to can automatically classify AM, DSB and FM signals in both coherent and non-coherent situations.

APPENDIX A

DERIVATIONS OF THE CYCLE FREQUENCIES

This appendix presents the detailed derivations of the cycle frequencies of M_{1y}^{α}, $M_{2,0y}^{\alpha}$ and $M_{2,1y}^{\alpha}$ for the targeted modulation types.

Before continuing, it is necessary to make explicit some conditions that are used in the derivations:

- $E[\,w(t)\,]=0$, $E[\,s(t)\,]=0$, $E\left[\,|w(t)|^4\,\right]=2\sigma_w^4$, $E\left[w(t+v)\,w(t)\right]=0$, $E\left[w(t+v)\,w^*(t)\right]=\sigma_w^2\delta(v)$, $E\left[x^*(t+v)\,w^*(t)\right]=0$ for any value of v, where σ_w^2 is the variance of $w(t)$, and $\delta(\cdot)$ stands for the Dirac delta function;
- The autocorrelation function of $s(t)$, defined as $r_{ss}(v)\triangleq E\left[s(t+v)\,s(t)\right]$, is an even function of v since $s(t)$ is real-valued;
- i.i.d. assumption on $\{s_l\}$ will be taken into account implicitly, and the mean value and variance of s_l are denoted by $\bar{m}_s$ and σ_s^2, respectively.

It should be noted that the cycle frequency α defined in Section 4.3 is a normalized frequency (normalized w.r.t. the sampling rate f_s). In the following, the analysis of the cyclostationarity of a signal is based on the continuous-time signal $y(t)$, rather than on its discrete-time version, i.e., $y(n)$. Then the notation n in the time-varying moments and cyclic moments will be replaced with the notation t. The analysis based on $y(t)$ should be equivalent to that based on $y(n)$ as long as the sampling rate f_s is sufficiently high. However, the cycle frequencies derived based on $y(t)$ are non-

normalized. If the latter is denoted by f, then its corresponding normalized version, α, will be $\alpha = \frac{f}{f_s}$. Nevertheless, one can easily identify the two versions in a concrete environment. Therefore, the same term cycle frequency will be used to alternatively represent both versions in the derivations. Moreover, for simplicity of presentation, the set of cycle frequencies for M_{1y}^{α}, $M_{2,0y}^{\alpha}$ and $M_{2,1y}^{\alpha}$ are denoted by Ω_1, Ω_{20} and Ω_{21}, respectively.

Derivation for AM Signal

For an AM signal, it is straightforward to show

$$m_{1y}(t,\boldsymbol{\tau}=0) = E\left[y(t)\right] = E\left[A\left(1+K_a s(t)\right)e^{j2\pi f_c t} + w(t)\right] = Ae^{j2\pi f_c t} \tag{A.1}$$

$$m_{2,0y}(t,\boldsymbol{\tau}=0) = E\left[y^2(t)\right] = A^2\left(1+K_a^2 r_{ss}(0)\right)e^{j4\pi f_c t} \tag{A.2}$$

$$m_{2,1y}(t,\boldsymbol{\tau}=0) = E\left[|\, y(t)\,|^2\right] = A^2\left(1+K_a^2 r_{ss}(0)\right) + \sigma_w^2 \tag{A.3}$$

where the only nonzero FS coefficient of $m_{1y}(t,\boldsymbol{\tau}=0)$ is equal to A which is the cyclic moment corresponding to the cycle frequency f_c, the only nonzero FS coefficient of $m_{2,0y}(t,\boldsymbol{\tau}=0)$ is $A^2\left(1+K_a^2 r_{ss}(0)\right)$ which is the cyclic moment corresponding to the cycle frequency $2f_c$, and the only nonzero FS coefficient of $m_{2,1y}(t,\boldsymbol{\tau}=0)$ is $A^2\left(1+K_a^2 r_{ss}(0)\right)$ whose corresponding frequency is zero. As aforementioned, the zero frequency is not considered as a cycle frequency. Therefore $\Omega_1 = \{f_c\}$, $\Omega_{20} = \{2f_c\}$ and $\Omega_{21} = \varphi$. The sets of cycle frequencies for other modulation types are determined in the same manner and such details are omitted in the remainder of the Appendix A.

Derivation for DSB Signal

For a DSB signal, the time-varying moments will be

$$m_{1y}(t,\boldsymbol{\tau}=0)=E\left[y(t)\right]=E\left[As(t)e^{j2\pi f_c t}+w(t)\right]=0 \tag{A.4}$$

$$m_{2,0y}(t,\boldsymbol{\tau}=0)=E\left[y^2(t)\right]=A^2 r_{ss}(0)e^{j4\pi f_c t} \tag{A.5}$$

$$m_{2,1y}(t,\boldsymbol{\tau}=0)=E\left[|y(t)|^2\right]=A^2 r_{ss}(0)+\sigma_w^2 \tag{A.6}$$

Then the sets of cycle frequencies will be $\Omega_1=\varphi$, $\Omega_{20}=\{2f_c\}$ and $\Omega_{21}=\varphi$.

Derivation for SSB Signal

In order to analyze the cyclostationarity of a SSB signal, it is necessary to evaluate $E\left[\tilde{s}(t)\right]$, $E\left[s(t)\tilde{s}(t)\right]$ and $E\left[\tilde{s}^2(t)\right]$. They are derived as

$$E\left[\tilde{s}(t)\right]=E\left[\int_{-\infty}^{\infty}\frac{s(t-v)}{\pi v}dv\right]=\int_{-\infty}^{\infty}\frac{E\left[s(t-v)\right]}{\pi v}dv=0 \tag{A.7}$$

$$E\left[s(t)\tilde{s}(t)\right]=E\left[\int_{-\infty}^{\infty}\frac{s(t)s(t-v)}{\pi v}dv\right]=\int_{-\infty}^{\infty}\frac{r_{ss}(v)}{\pi v}dv=0 \tag{A.8}$$

$$\begin{aligned}
E\left[\tilde{s}^2(t)\right]&=E\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{s(t-v_1)s(t-v_2)}{\pi v_1\pi v_2}dv_1dv_2\right]\\
&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{r_{ss}(v_2-v_1)}{\pi v_1\pi v_2}dv_1dv_2\\
&=\int_{-\infty}^{\infty}\frac{1}{\pi v_2}\left\{\int_{-\infty}^{\infty}\frac{r_{ss}(v_2-v_1)}{\pi v_1}dv_1\right\}dv_2\\
&=\int_{-\infty}^{\infty}\frac{1}{\pi v_2}\tilde{r}_{ss}(v_2)dv_2\\
&=-\int_{-\infty}^{\infty}\frac{1}{\pi(0-v_2)}\tilde{r}_{ss}(v_2)dv_2\\
&=-\tilde{\tilde{r}}_{ss}(v)\,|_{v=0}=r_{ss}(v)\,|_{v=0}=r_{ss}(0)
\end{aligned} \tag{A.9}$$

The last step of equation (A.8) is owing to the fact that $r_{ss}(v)$ is even and thus $\frac{r_{ss}(v)}{\pi v}$ is odd function of v. $\tilde{\tilde{r}}_{ss}(v)$ represents the Hibert transform of $\tilde{r}_{ss}(v)$, and the well-

known property of Hilbert transform, i.e., $\tilde{\tilde{r}}_{ss}(v) = -\tilde{r}_{ss}(v)$, has been applied in derivation. Armed with the above results, the time-varying moments of an SSB signal will be

$$\begin{aligned} & m_{1y}(t, \boldsymbol{\tau} = 0) \\ & \quad = E\left[y(t)\right] \\ & \quad = E\left[A(s(t) \mp j\hat{s}(t))e^{j2\pi f_c t} + w(t)\right] \\ & \quad = 0 \end{aligned} \tag{A.10}$$

$$\begin{aligned} & m_{2,0y}(t, \boldsymbol{\tau} = 0) \\ & \quad = E\left[y^2(t)\right] \\ & \quad = A^2\left\{E\left[s^2(t)\right] - E\left[\tilde{s}^2(t)\right] \mp j2E\left[s(t)\tilde{s}(t)\right]\right\} e^{j4\pi f_c t} \\ & \quad = 0 \end{aligned} \tag{A.11}$$

$$\begin{aligned} & m_{2,1y}(t, \boldsymbol{\tau} = 0) \\ & \quad = E\left[\,|\, y(t) \,|^2\,\right] \\ & \quad = A^2\left\{E\left[s^2(t)\right] + E\left[\tilde{s}^2(t)\right]\right\} + \sigma_w^2 \\ & \quad = 2A^2 r_{ss}(0) + \sigma_w^2 \end{aligned} \tag{A.12}$$

where the minus sign and addition sign of the operator $\mp$ are for LSB and USB, respectively. Therefore, $\Omega_1 = \varphi$, $\Omega_{20} = \varphi$ and $\Omega_{21} = \varphi$ for both LSB and USB.

Derivation for NBFM and WBFM Signal

For an FM signal, the conjugate second-order time-varying moment is

$$\begin{aligned} & m_{2,1y}(t, \boldsymbol{\tau} = 0) \\ & \quad = E\left[\,|\, y(t) \,|^2\,\right] \\ & \quad = E\left[\left| Ae^{j2\pi f_c t + j2\pi f_\Delta \int_{-\infty}^{t} s(v)dv} + w(t)\right|^2\right] \\ & \quad = A^2 + \sigma_w^2 \end{aligned} \tag{A.13}$$

Therefore, the set of cycle frequencies for FM will be $\Omega_{21} = \varphi$.

The analysis of $m_{1y}(t,\boldsymbol{\tau}=0)$ and $m_{2,0y}(t,\boldsymbol{\tau}=0)$ of an FM signal would be difficult due to lack of the knowledge of the PDF of the information-bearing signal $s(t)$. $s(t)$ is assumed to be ergodic. Here the modulated signal $x(t)$ for FM is further assumed to be ergodic, and thus $y(t)$ is also ergodic. Then an FM signal can be analyzed either based on the above-introduced ensemble-average framework where the input signal is treated as a stochastic process, or based on the time-average framework where the input signal is viewed as a sample function of the stochastic process. As stated in [84], the results from one framework are generally true in the other (i.e., with probability equal to one) if the process is ergodic.

The *k-th order time-varying moment*, $m_{ky}(t;\boldsymbol{\tau})$ in the time-average framework is defined as [84]

$$m_{ky}(t;\boldsymbol{\tau}) \triangleq \widehat{E}^{\{\alpha\}}\left[\, f_{ky}(t;\boldsymbol{\tau}) \,\right] = \sum_{\alpha} \left\langle\, f_{ky}(t;\boldsymbol{\tau}) e^{-j2\pi\alpha t} \,\right\rangle e^{j2\pi\alpha t} \tag{A.14}$$

where

$$\left\langle\, q(t) \,\right\rangle \equiv \lim_{\Gamma\to\infty} \frac{1}{\Gamma} \int_{-\Gamma/2}^{\Gamma/2} q(u)du \tag{A.15}$$

is the time-averaging operation. The sum in equation (A.14) is over all real-valued α that result in nonzero terms of $\left\langle\, f_{ky}(t;\boldsymbol{\tau}) e^{-j2\pi\alpha t} \,\right\rangle$. Then $f_{ky}(t;\boldsymbol{\tau})$ can be represented as

$$f_{ky}(t;\boldsymbol{\tau}) = m_{ky}(t;\boldsymbol{\tau}) + f_{ky,res}(t;\boldsymbol{\tau}) \tag{A.16}$$

where $\left\langle f_{ky,res}(t;\boldsymbol{\tau}) e^{-j2\pi\alpha t} \right\rangle = 0$ for any real-valued α.

The *k-th order cyclic moment* $M_{ky}^{\alpha}(\boldsymbol{\tau})$ is defined as

$$M_{ky}^{\alpha}(\boldsymbol{\tau}) \equiv \left\langle\, f_{ky}(t;\boldsymbol{\tau}) \,\right\rangle e^{-j2\pi\alpha t} \tag{A.17}$$

Therefore, $f_{ky}(t;\boldsymbol{\tau})$ can be further represented as

$$f_{ky}(t;\boldsymbol{\tau}) = \sum_{\alpha} M_{ky}^{\alpha}(\boldsymbol{\tau})\, e^{j2\pi\alpha t} + f_{ky,res}(t;\boldsymbol{\tau}) \tag{A.18}$$

From equation (A.14) through equation (A.18), it is evident that $m_{ky}(t;\boldsymbol{\tau})$ contains and only contains all the additive sine-wave components of $f_{ky}(t;\boldsymbol{\tau})$, where $M_{ky}^{\alpha}(\boldsymbol{\tau})$ is the complex strength of an additive sine-wave with frequency α. This establishes a connection between the cyclic moment $M_{ky}^{\alpha}(\boldsymbol{\tau})$ and the Fourier transform (FT) of $f_{ky}(t;\boldsymbol{\tau})$. That is, the *cycle frequencies of* $M_{ky}^{\alpha}(\boldsymbol{\tau})$, are equal to the nonzero frequencies where the FT of $f_{ky}(t;\boldsymbol{\tau})$ contains spectral impulses (i.e. Dirac delta functions). In other words, a nonzero frequency will be a cycle frequency if it corresponds to a spectral impulse (i.e., a Dirac delta) of the FT of $f_{ky}(t;\boldsymbol{\tau})$.

For a noiseless narrowband FM (NBFM) signal $x(t)$ in the form of equation (4.6), its FT can be approximated by [8],

$$X(f) \approx A\delta(f - f_c) + jAS(f - f_c) \tag{A.19}$$

where $S(f)$ is the FT of the information-bearing signal $s(t)$, and the result is based on the assumption $\left|2\pi f_{\Delta}\int_{-\infty}^{t} s(v)dv\right| << 1$ radian . In this chapter, the cases of tone-modulation are not considered. Then it can be assumed that $S(f)$ does not contain spectral impulses without loss of generality. Furthermore, since the noise $w(t)$ is assumed white, it will not contribute spectral impulses to the FT of $y(t)$. Then with the above-derived connection between the cycle frequencies and the FT of $f_{ky}(t;\boldsymbol{\tau})$, it can be concluded that $m_{1y}(t;\boldsymbol{\tau}=0) = \widehat{E}^{\{\alpha\}}\left[y(t)\right]$ will have and only have one additive sine wave

component with frequency being f_c if the f_Δ is very small. Then the set of cycle frequencies of M_{1y}^{α} for NBFM is $\Omega_1 = \{f_c\}$. When the value of f_Δ is higher such that the condition $\left| 2\pi f_\Delta \int_{-\infty}^{t} s(v)\,dv \right| \ll 1$ does not hold, however, the approximation of equation (A.19) will be weaker or even invalid. Instead, it is more proper to analyze by using the model of wideband FM (WBFM). As shown later, M_{1y}^{α} does not have cycle frequency for WBFM. Then the set of cycle frequencies of M_{1y}^{α} for NBFM can be expressed as $\Omega_1 = \left\{ f_c^{(\pm)} \right\}$, where the superscript $(\pm)$ means the cycle frequency may or may not exist.

As for $m_{2,0y}(t, \boldsymbol{\tau} = 0)$ of an NBFM signal, it can be expressed as

$$m_{20y}(t; \boldsymbol{\tau} = 0) = \widehat{E}^{\{\alpha\}}\left[y^2(t) \right] = \widehat{E}^{\{\alpha\}}\left[A^2 e^{j4\pi f_c t + j4\pi f_\Delta \int_{-\infty}^{t} s(v)dv} \right] \qquad \text{(A.20)}$$

where the signal in the brackets is simply an FM signal with carrier frequency $2f_c$ and frequency deviation $2f_\Delta$. For an FM signal with frequency deviation $2f_\Delta$, the condition $\left| 4\pi f_\Delta \int_{-\infty}^{t} s(v)dv \right| << 1$ radian will generally not hold, hence, the set of cycle frequencies of $M_{2,0y}^{\alpha}$ for NBFM is $\Omega_{20} = \varphi$. This is found to be proper through simulations. In summary, the sets of cycle frequencies for NBFM will be $\Omega_1 = \left\{ f_c^{\pm} \right\}$, $\Omega_{20} = \varphi$,and $\Omega_{21} = \varphi$.

Now go back to the ensemble-average framework. According to the definition of cycle frequencies and cyclic moments, if an analytic signal $y(t)$ is first-order cyclostationary in the sense of M_{1y}^{α}, then it can be written as

$$y(t) = E\left[y(t)\right] + y_{res}(t) = \sum_{q=1}^{Q} M_{1y}^{f_q} e^{j2\pi f_q t} + M_{1y}^{0} + y_{res}(t) \tag{A.21}$$

where $y_{res}(t)$ is a random signal with zero-mean (i.e., $E\left[y_{res}(t)\right] = 0$), M_{1y}^{0} represents the complex strength of the DC component of $E\left[y(t)\right]$ and it might be zero, Q denotes the number of distinct cycle frequencies, f_q is the *q-th* cycle frequency of M_{1y}^{α}, the value of M_{1y}^{α} at the cycle frequency f_q is represented by $M_{1y}^{f_q}$, and $M_{1y}^{f_q} \neq 0$ for $H(k,m,n) \stackrel{\Delta}{=} z^3(k) + z^3(m) + z^3(n) - 3z(k)z(m)z(n)$. For the signal $y(t)$ as expressed in (A.21), its autocorrelation function, defined as $r_{yy}(t;v) \triangleq E\left[y(t+v)y^*(t)\right]$, can be evaluated as

$$\begin{aligned} r_{yy}(t;v) = & \sum_{q=1}^{Q} \left|M_{1y}^{f_q}\right|^2 e^{j2\pi f_q v} + 2\sum_{q=1}^{Q} \left(M_{1y}^{0}\right)^* \cdot \mathrm{Re}\left(M_{1y}^{f_q} e^{j2\pi f_q \left(t+\frac{v}{2}\right)} \right) \cdot e^{j\pi f_q v} \\ & + \left|M_{1y}^{0}\right|^2 + \sum_{q=1}^{Q} \sum_{\substack{p=1 \\ p\neq q}}^{Q} M_{1y}^{f_q} \cdot \left(M_{1y}^{f_p}\right)^* e^{j2\pi\left(f_q - f_p\right)t + j2\pi f_q v} + r_{y_{res}y_{res}}(t;v) \end{aligned} \tag{A.22}$$

where $r_{y_{res}y_{res}}(t;v) \triangleq E\left[\, y_{res}(t+v) y_{res}^*(t) \,\right]$ is the auto-correlation function (ACF) of $y_{res}(t)$. When $Q = 1$, $M_{1y}^{0} = 0$, and $y_{res}(t)$ is WSS , i.e., $r_{y_{res}y_{res}}(t;v) = r_{y_{res}y_{res}}(v)$, the ACF of $r_{yy}(t;v)$ will be

$$r_{yy}(t;v) = \left|M_{1y}^{f_1}\right|^2 \cdot e^{j2\pi f_1 v} + r_{y_{res}y_{res}}(v) \tag{A.23}$$

Then the PSD of $y(t)$ will be

$$S_y(f) = \left|M_{1y}^{f_1}\right|^2 \delta(f - f_1) + S_{y_{res}}(f) \tag{A.24}$$

where $S_{y_{res}}(f)$ is the PSD of $y_{res}(t)$, i.e., the FT of $r_{y_{res}y_{res}}(v)$. When the above conditions do not hold, however, $r_{yy}(t;v)$ is generally a function of t. Nevertheless, one can evaluate the average PSD of $y(t)$ by first evaluating the average ACF $\overline{r}_{yy}(v)$ as

$$\overline{r}_{yy}(v) \triangleq \left\langle r_{yy}(t;v) \right\rangle = \sum_{q=1}^{Q} \left| M_{1y}^{f_q} \right|^2 e^{j2\pi f_q v} + \left| M_{1y}^{0} \right|^2 + \overline{r}_{y_{res}y_{res}}(v) \tag{A.25}$$

where $\overline{r}_{y_{res}y_{res}}(v) \triangleq \left\langle r_{y_{res}y_{res}}(t;v) \right\rangle$, and it will equal to $r_{y_{res}y_{res}}(v)$ if $y_{res}(t)$ is WSS. Then the average PSD of $y(t)$, denoted by $\overline{S}_y(f)$, will be

$$\overline{S}_y(f) = \sum_{q=1}^{Q} \left| M_{1y}^{f_q} \right|^2 \delta(f - f_q) + \overline{S}_{y_{res}}(f) + \left| M_{1y}^{0} \right|^2 \delta(f) \tag{A.26}$$

where $\overline{S}_{y_{res}}(f)$ is the average PSD of $y_{res}(t)$, i.e., the FT of $\overline{r}_{y_{res}y_{res}}(v)$. From all the above, it can be concluded that a cycle frequency of the first-order cyclic moment M_{1y}^{α} of the signal $y(t)$ will contribute a spectral impulse (i.e., Dirac delta in frequency domain) to the PSD (or average PSD) of $y(t)$ at that frequency location. In another word, one can claim that $y(t)$ is not first-order cyclostationary if its PSD does not have spectral impulse at any nonzero frequency. It should be noted that, however, the signal $y(t)$ is not guaranteed to be first-order cyclostationary if its PSD contains spectral impulses at nonzero frequencies. For example, a signal $y(t) = \xi(t) \cdot e^{j2\pi f_c t}$ is not first-order cyclostationary if $\zeta(t)$ is a zero-mean random signal. It is easy to show that the PSD of $y(t)$ will be $S_y(f) = \sigma_\xi^2 \delta(f - f_c)$ if $\xi(t)$ is white with variance σ_ξ^2. That is, the PSD of $y(t)$ has a spectral impulse at nonzero frequency f_c, but it is not first-order cyclostationary in the sense of M_{1y}^{α}.

Based on the above result, the cycle frequencies of M_{1y}^{α} and $M_{2,0y}^{\alpha}$ for WBFM can be analyzed as follows. For a WBFM signal $y(t)$ as modeled by (4.1) and (4.6), its PSD can be derived as

$$S_y(f)=\begin{cases} S_{\mathrm{Re}(x)}(f)+\sigma_w^2 & , \text{ for } f\geq 0 \\ 0 & , \text{ for } f<0 \end{cases} \tag{A.27}$$

where $S_{\mathrm{Re}(x)}(f)$ is the PSD of the signal $\mathrm{Re}(x(t))$, and it can be expressed [8] as

$$S_{\mathrm{Re}(x)}(f)=\frac{A^2}{2f_\Delta}\cdot\left(p_s\left(\frac{f-f_c}{f_\Delta}\right)+p_s\left(\frac{f+f_c}{f_\Delta}\right)\right) \tag{A.28}$$

and $p_s(f)$ is the PDF of the information-bearing signal $s(t)$. The signal $s(t)$ is assumed zero-mean in this appendix. Also, the tone-modulation will not be taken into account for WBFM. Then it is reasonable to assume the PDF $p_s(f)$ of the random variable $s(t)$ is smooth, i.e., not containing impulse for any value of f. It follows that the PSD of $y(t)$ does not contain spectral impulse, and thus $y(t)$ is not first-order cyclostationary, i.e., $\Omega_1=\varphi$ for WBFM. The second-order cyclic moment $M_{2,0y}^{\alpha}$ is actually the first-order cyclic moment of the signal $z(t)\triangleq y^2(t)$. The ACF of $z(t)$, denoted by $r_{zz}(t;v)$, can be evaluated as

$$\begin{aligned} r_{zz}(t;v) &= E\left[z(t+v)z^*(t)\right] \\ &= E\left[y^2(t+v)y^{2*}(t)\right] \\ &= E\left[x^2(t+v)x^{2*}(t)\right]+E\left[w^2(t+v)w^{2*}(t)\right] \\ &\quad +4E\left[x(t+v)x^*(t)\right]\cdot E\left[w(t+v)w^*(t)\right] \\ &= r_{x_2x_2}(t;v)+2\left(\sigma_w^2\right)^2\delta(v)+4r_{xx}(t;v)\sigma_w^2\delta(v) \end{aligned} \tag{A.29}$$

where $r_{x_2x_2}(t;v)$ is the ACF of the signal $x_2(t) \triangleq x^2(t)$. For a noiseless FM signal $x(t)$ as defined in (4.6), the signal $x_2(t)$ is also a noiseless FM signal with amplitude A^2, information-bearing signal $s(t)$, frequency deviation $2f_\Delta$ and carrier frequency $2f_c$. Also based on the analysis of [8], the PSD $S_{\mathrm{Re}(x_2)}(f)$ of the signal $\mathrm{Re}(x_2(t))$ can be evaluated as

$$S_{\mathrm{Re}(x)}(f) = \frac{A^4}{4f_\Delta} \cdot \left(p_s\left(\frac{f-2f_c}{2f_\Delta} \right) + p_s\left(\frac{f+2f_c}{2f_\Delta} \right) \right) \tag{A.30}$$

Moreover, $r_{xx}(t;v)$ is equal to A^2 for the delay $v=0$ since an FM signal has constant envelope. Then the PSD of $z(t)$ will be

$$\begin{aligned} S_z(f) &= S_{x_2}(f) + 2\left(\sigma_w^2\right)^2 + 4A^2\sigma_w^2 S_y(f) \\ &= \begin{cases} S_{\mathrm{Re}(x_2)}(f) + \sigma_w^2 & , \text{for } f \geq 0 \\ 0 & , \text{for } f < 0 \end{cases} \end{aligned} \tag{A.31}$$

where $S_{x_2}(f)$ and $S_{\mathrm{Re}(x_2)}(f)$ are the PSDs of $x_2(t)$ and $\mathrm{Re}(x_2(t))$, respectively. Then similar to the analysis on the signal $y(t)$, it can be concluded that $z(t)$ is not first-order cyclostationary. Equivalently, $M_{2,0y}^{\alpha}$ does not have cycle frequency, i.e., $\Omega_{20} = \varphi$. In summary, the cycle frequency sets for WBFM are $\Omega_1 = \varphi$, $\Omega_{20} = \varphi$ and $\Omega_{21} = \varphi$.

Derivation for Linearly Modulated Digital Signal

For linearly modulated digital communication signals, the time-varying moment $m_{1y}(t, \boldsymbol{\tau} = 0)$ will have the following general form

$$
\begin{aligned}
& m_{1y}(t,\boldsymbol{\tau}=0) \\
&\quad = E\left[A\sum_l s_l g(t-lT)e^{j2\pi f_c t} + w(t) \right] \\
&\quad = A\bar{m}_s \sum_l g(t-lT)\cdot e^{j2\pi f_c t}
\end{aligned}
\tag{A.32}
$$

For an MPAM, MPSK or QAM signal, the mean value of symbols is $\bar{m}_s = 0$, resulting in $\Omega_1 = \varphi$.

For an MASK signal, the symbol's mean value is $\bar{m}_s = \dfrac{1+M}{2}$, then $m_{1y}(t,\boldsymbol{\tau}=0)$ can be rewritten as

$$
\begin{aligned}
& m_{1y}(t,\boldsymbol{\tau}=0) \\
&\quad = \frac{A(1+M)}{2}\sum_l g(t-lT)\cdot e^{j2\pi f_c t} \\
&\quad = \frac{A(1+M)}{2T}\sum_l G(\frac{l}{T})\cdot e^{j2\pi(f_c+\frac{l}{T})t}
\end{aligned}
\tag{A.33}
$$

where $G(f)$ is the FT of $g(t)$. Then it is straightforward that $\Omega_1 = \left\{ f_c + \dfrac{l}{T}, l \text{ is an integer} \right\}$.

For linearly modulated digital communication signals, $m_{2,0y}(t,\boldsymbol{\tau}=0)$ will have the following general form

$$
\begin{aligned}
& m_{2,0y}(t,\boldsymbol{\tau}=0) \\
&\quad = E\left[\left(A\sum_l s_l g(t-lT)e^{j2\pi f_c t} + w(t) \right)^2 \right] \\
&\quad = A^2 e^{j4\pi f_c t}\left\{ \sum_l \left(E\left[s_l^2 \right] - \bar{m}_s^2 \right) g^2(t-lT) + \bar{m}_s^2 \sum_l \sum_k g(t-lT)g(t-kT) \right\} \\
&\quad = \frac{A^2}{T}\sum_l \left\{ \left(E\left[s_l^2 \right] - \bar{m}_s^2 \right) \times G_{20}^{(0)}\left(\frac{l}{T} \right) + \bar{m}_s^2 \sum_k G_{20}^{(k)}\left(\frac{l}{T} \right) \right\} \cdot e^{j2\pi\left(2f_c + \frac{l}{T} \right)t} \\
&\quad = \sum_l \left(E\left[s_l^2 \right] - \bar{m}_s^2 \right) g^2(t-lT) + \bar{m}_s^2 \sum_l \sum_k g(t-lT)g(t-kT)
\end{aligned}
\tag{A.34}
$$

where $G_{20}^{(k)}(f)$ is the FT of $g(t)\times g(t-kT)$.

For an MPSK signal with $M \geq 4$ or a QAM signal, both $E\left[s_l^2 \right]$ and $\bar{m}_s$ are zeros, resulting in $\Omega_{20} = \varphi$.

For MASK and MPAM, $m_{2,0y}(t, \boldsymbol{\tau} = 0)$ will be as below, respectively,

$$m_{2,0y}(t, \boldsymbol{\tau} = 0) = \frac{A^2}{T}\sum_l \left\{ \frac{M^2-1}{12} G_{20}^{(0)}\left(\frac{l}{T}\right) + \frac{(M+1)^2}{4}\sum_k G_{20}^{(k)}\left(\frac{l}{T}\right) \right\} e^{j2\pi t\left(2f_c + \frac{l}{T}\right)} \quad \text{(A.35)}$$

$$m_{2,0y}(t, \boldsymbol{\tau} = 0) = \frac{A^2(M^2-1)}{3T}\sum_l G_{20}^{(0)}\left(\frac{l}{T}\right) e^{j2\pi t\left(2f_c + \frac{l}{T}\right)} \quad \text{(A.36)}$$

then the set for Ω_{20} for both MASK and MPAM will be $\Omega_{20} = \left\{ 2f_c + \frac{l}{T}, l \text{ is an integer} \right\}$.

Similarly, the time varying moment $m_{2,1y}(t, \boldsymbol{\tau} = 0)$ for a linearly modulated digital communication signal can be derived as

$$\begin{aligned} m_{2,1y}(t, \boldsymbol{\tau} = 0) &= E\left[\left| A\sum_l s_l g(t-lT) e^{j2\pi f_c t} + w(t) \right|^2 \right] \\ &= \frac{A^2}{T}\sum_l \left\{ \sigma_s^2 \times G_{21}^{(0)}\left(\frac{l}{T}\right) + \bar{m}_s^2 G_{21}^{(k)}\left(\frac{l}{T}\right) \right\} e^{j2\pi\frac{l}{T}t} + \sigma_w^2 \end{aligned} \quad \text{(A.37)}$$

where $G_{21}^{(k)}(f)$ is the FT of $g(t)\times g^*(t-kT)$.

In equation (A.37), the symbols' mean value $\bar{m}_s$ may be zero, but the symbol variance σ_s^2 will never be zero for linear digital modulations. Then the set Ω_{21} for all the digital modulations will be $\Omega_{21} = \left\{ \frac{l}{T}, \quad l \text{ is an nonzero integer} \right\}$.

APPENDIX B

RELATIONSHIP BETWEEN $P_{y,norm}$ AND $P_{y2,norm}$

This appendix serves to derive the relationship between the feature $P_{y,norm}$ and the feature $P_{y2,norm}$ for the noiseless AM signal and DSB signal, where both features are defined in Section 4.5.

According to equation (4.2), the samples of a noiseless AM signal can be expressed as

$$x(n) = z(n)e^{j2\pi T_s f_c n + j\theta_c}, \text{ for } n = 0,1,\cdots,N-1 \tag{B.1}$$

where $s(n) \triangleq s(nT_s)$, θ_c stands for the unknown initial carrier phase that is deterministic, T_s is the sampling period that is the reciprocal of the sampling rate f_s, $z(n) \triangleq A(1+K_a s(n))$, and $z(n)$ is real-valued and is nonnegative for any n.

The estimated PSD $\hat{S}_x(k)$ of $x(n)$ can be derived as

$$\begin{aligned}
&\hat{S}_x(k) \\
&= \frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{N}kn}\right|^2 = \frac{1}{N}\left|\sum_{n=0}^{N-1} z(n)e^{-j\frac{2\pi}{N}kn + j(2\pi f_c nT_s + \theta_c)}\right|^2 \\
&= \frac{1}{N}\sum_{n=0}^{N-1} z^2(n) + \frac{2}{N}\sum_{n=0}^{N-2}\sum_{m>n}^{N-1} z(n)z(m)\cos\left(2\pi\left(\frac{k}{N} - \frac{f_c}{f_s}\right)\times(m-n)\right) \\
&\le \frac{1}{N}\sum_{n=0}^{N-1} z^2(n) + \frac{2}{N}\sum_{n=0}^{N-2}\sum_{m>n}^{N-1} z(n)z(m)
\end{aligned} \tag{B.2}$$

Applying the facts of $0 \le \left| \cos\left(2\pi \left(\frac{k}{N} - \frac{f_c}{f_s} \right) \times (m-n) \right) \right| \le 1$ and $z(n) \ge 0$ in derivation, equation (B.2) points out that the global maximum of $\hat{S}_x(k)$ will be exactly equal to $\frac{1}{N}\left(\sum_{n=0}^{N-1} z(n) \right)^2$ if $NT_s f_c$ is an integer, i.e.,

$$\hat{S}_x(k)|_{k=NT_s f_c} = \frac{1}{N}\left(\sum_{n=0}^{N-1} z(n) \right)^2 \tag{B.3}$$

In general, it cannot be guaranteed $NT_s f_c$ will be an integer. When the number of available samples, i.e., N, is sufficiently large, however, the DFT resolution will be sufficiently high such that the fractional part of $NT_s f_c$ will be sufficiently small. Then it is reasonable to take $\frac{1}{N}\left(\sum_{n=0}^{N-1} z(n) \right)^2$ as the global maximum of $\hat{S}_x(k)$. Thus the normalized PSD $\hat{S}_{x,norm}(k)$ of $x(n)$ will be

$$\hat{S}_{x,norm}(k) = \frac{\hat{S}_x(k)}{\frac{1}{N}\left(\sum_{n=0}^{N-1} z(n) \right)^2} = \frac{\frac{1}{N}\left| \sum_{n=0}^{N-1} z(n) e^{-j\frac{2\pi}{N}kn + j(2\pi f_c n T_s + \theta_c)} \right|^2}{\frac{1}{N}\left(\sum_{n=0}^{N-1} z(n) \right)^2} \tag{B.4}$$

and the feature $\hat{P}_{x,norm}$ will be evaluated as

$$\hat{P}_{x,norm} = \sum_{k=0}^{N-1} \hat{S}_{x,norm}(k) = \frac{\sum_{k=0}^{N-1} \hat{S}_x(k)}{\frac{1}{N}\left(\sum_{n=0}^{N-1} z(n) \right)^2} \tag{B.5}$$

By applying Parseval's Theorem, equation (B.5) can be further rewritten as

$$\hat{P}_{x,norm} = \frac{\sum_{n=0}^{N-1}\left|x(n)\right|^2}{\frac{1}{N}\left(\sum_{n=0}^{N-1} z(n)\right)^2} = \frac{\sum_{n=0}^{N-1}\left|z(n)\right|^2}{\frac{1}{N}\left(\sum_{n=0}^{N-1} z(n)\right)^2} = \frac{\sum_{n=0}^{N-1} z^2(n)}{\frac{1}{N}\left(\sum_{n=0}^{N-1} z(n)\right)^2} \tag{B.6}$$

where the last step is owing to the fact that $z(n)$ is nonnegative real value.

Similarly, the feature $P_{x2,norm}$, which corresponds to the normalized PSD of $x^2(n)$, can be derived as

$$\hat{P}_{x2,norm} = \frac{\sum_{n=0}^{N-1} z^4(n)}{\left(\frac{1}{N}\left(\sum_{n=0}^{N-1} z^2(n)\right)^2\right)} \tag{B.7}$$

Then the difference between $\hat{P}_{x2,norm}$ and $\hat{P}_{x,norm}$ can be written as

$$\hat{P}_{x2,norm} - \hat{P}_{x,norm} = \frac{\sum_{m=0}^{N-1} z^4(m)\left(\sum_{n=0}^{N-1} z(n)\right)^2 - \left(\sum_{m=0}^{N-1} z^2(m)\right)^3}{\frac{1}{N}\left(\sum_{m=0}^{N-1} z^2(m)\right)^2\left(\sum_{n=0}^{N-1} z(n)\right)^2} \tag{B.8}$$

Since the data $\{x(n)\}$ represent an AM signal captured by an AM system, then some of them should be nonzero. Then some of $\{z(n)\}$ should be nonzero. Moreover, $z(n) \geq 0$ for any n. Therefore, the denominator on the RHS of (B.8) should be positive. That is, in order to determine the relationship between $P_{x,norm}$ and $P_{x2,norm}$, one only needs to check the sign of the numerator on the RHS of (B.8). The numerator at the RHS of (B.8) can be rewritten as

$$\begin{aligned}\mathbb{N} &\triangleq \sum_{m=0}^{N-1} z^4(m)\left(\sum_{n=0}^{N-1} z(n)\right)^2 - \left(\sum_{m=0}^{N-1} z^2(m)\right)^3 \\ &= \sum_{k=0}^{N-1}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} z^2(k)z(m)z(n)\left(z^2(k)-z(m)z(n)\right) \\ &= \sum_{\substack{k=0\\m=k}}^{N-1}\sum_{\substack{n=0\\n\neq k}}^{N-1}\left(\mathbb{Z}(k,m,n)+\mathbb{Z}(k,n,m)+\mathbb{Z}(n,m,k)\right)+\sum_{\substack{k=0\\k\neq n}}^{N-1}\sum_{\substack{m=0\\m\neq k}}^{N-1}\sum_{\substack{n=0\\n\neq m}}^{N-1}\mathbb{Z}(k,m,n)\end{aligned} \tag{B.9}$$

where the function $\mathbb{Z}(k,m,n)$ is defined as

$$\mathbb{Z}(k,m,n) \overset{\Delta}{=} z^2(k)z(m)z(n)\left(z^2(k)-z(m)z(n)\right) \tag{B.10}$$

The first term on the RHS of equation (B.9) can be evaluated as

$$\begin{aligned}&\sum_{\substack{k=0\\m=k}}^{N-1}\sum_{\substack{n=0\\n\neq k}}^{N-1}\left(\mathbb{Z}(k,m,n)+\mathbb{Z}(k,n,m)+\mathbb{Z}(n,m,k)\right) \\ &= \sum_{k=0}^{N-1}\sum_{\substack{n=0\\n\neq k}}^{N-1}\left\{z^2(k)z(n)\left[2z(k)+z(n)\right]\times\left[z(n)-z(k)\right]^2\right\} \\ &\geq 0\end{aligned} \tag{B.11}$$

When the AM signal does convey information, there should exist at least one pair of n and k for which $z(k)$ is not equal to $z(n)$ if $k \neq n$. Then the result of equation (B.11) should be greater than zero.

The second term on the RHS of equation (B.9) does not exist if N is less than three. For $N \geq 3$, that term can be reorganized as

$$\begin{aligned}\sum_{\substack{k=0\\k\neq n}}^{N-1}\sum_{\substack{m=0\\m\neq k}}^{N-1}\sum_{\substack{n=0\\n\neq m}}^{N-1}\mathbb{Z}(k,m,n) &= \sum_{k=0}^{N-1}\sum_{m>k}^{N-1}\sum_{n>m}^{N-1}\mathbb{Z}(k,m,n)+\sum_{k=0}^{N-1}\sum_{m>k}^{N-1}\sum_{n>m}^{N-1}\mathbb{Z}(k,n,m) \\ &\quad+\sum_{k=0}^{N-1}\sum_{m>k}^{N-1}\sum_{n>m}^{N-1}\mathbb{Z}(m,n,k)+\sum_{k=0}^{N-1}\sum_{m>k}^{N-1}\sum_{n>m}^{N-1}\mathbb{Z}(m,k,n) \\ &\quad+\sum_{k=0}^{N-1}\sum_{m>k}^{N-1}\sum_{n>m}^{N-1}\mathbb{Z}(n,m,k)+\sum_{k=0}^{N-1}\sum_{m>k}^{N-1}\sum_{n>m}^{N-1}\mathbb{Z}(n,k,m) \\ &= \sum_{k=0}^{N-1}\sum_{m>k}^{N-1}\sum_{n>m}^{N-1} z(k)z(m)z(n)\times H(k,m,n)\end{aligned} \tag{B.12}$$

where the function $H(k,m,n)$ is defined as

$$H(k,m,n) \overset{\Delta}{=} z^3(k) + z^3(m) + z^3(n) - 3z(k)z(m)z(n) \tag{B.13}$$

In the following, $H(k,m,n)$ will be shown to be nonnegative for any combination of k, m and n. It is noted that the function $H(k,m,n)$ is symmetric with respect to $z(k)$, $z(m)$ and $z(n)$. Moreover, $z(k)$, $z(m)$ and $z(n)$ are all nonnegative. Thus one can assume $z(m) \geq z(n) \geq z(k) \geq 0$ without loss of generality. The function $H(k,m,n)$ can be rewritten as

$$\begin{aligned} H(k,m,n) &= z^3(k) + z^3(m) + z^3(n) - 3z(k)z(m)z(n) \\ &= \big(z(m) - z(k)\big)^3 + \big(z(n) - z(k)\big)^3 \\ &\quad + 3z(k) \times \Big(\big(z(n) - z(m)\big)^2 + \big(z(m) - z(k)\big) \times \big(z(n) - z(k)\big)\Big) \\ &\geq 0 \end{aligned} \tag{B.14}$$

That is, the last term on the RHS of (B.9) is greater than or equal to zero. In general, this term is greater than zero if the information-bearing signal $s(n)$ is not constant.

Combining the above results, it can be concluded that the value of $\hat{P}_{x2,norm}$ is greater than that of $\hat{P}_{x,norm}$ for a noiseless AM signal.

For a received DSB signal, the PSD $S_y(f)$ of $\{y(t)\}$ and the PSD $S_{y2}(f)$ of $\{y(t)\}$ can be respectively derived as

$$S_y(f) = A^2 S_s(f - f_c) + \sigma_w^2 \tag{B.15}$$

$$S_{y2}(f) = A^4 \left(\overline{m}_{s2}\right)^2 \delta(f - 2f_c) + A^4 S_\zeta(f - 2f_c) + 4A^2 \sigma_w^2 \overline{m}_{s2} + 2\sigma_w^4 \tag{B.16}$$

where $S_s(f)$ stands for the PSD of $s(t)$, $\overline{m}_{s2} = E\left[s^2(t)\right]$, and $S_\zeta(f)$ stands for the PSD of $\xi(t) \triangleq s^2(t) - \overline{m}_{s2}$. It is evident that $S_{y2}(f)$ will have a spectral impulse. Here a weak assumption is imposed on the PSD of $s(t)$, i.e., $S_s(f)$ does not contain spectral impulse. Then if $S_y(f)$ and $S_{y2}(f)$ are respectively normalized w.r.t. their own global maximum, the area under the former will be greater than that of the latter. That is, the value of $P_{y,norm}$ will be greater than that of $P_{y2,norm}$ for DSB.

REFERENCES

[1] M. Dillinger*, et al.*, *Software Defined Radio: Architectures, Systems and Functions*: Wiley & Sons, 2003.

[2] J. Mitola, "The Software Radio," presented at the IEEE National Telesystems Conference, 1992.

[3] R. I. Lackey and D. W. Upmal, "Speakeasy: the military software radio," *Communications Magazine, IEEE,* vol. 33, pp. 56-61, 1995.

[4] E. E. Azzouz and A. K.Nandi, *Automatic modulation recognition of communication signals*: Kluwer Academic Publishers, 1996.

[5] C. Dubuc*, et al.*, "An automatic modulation recognition algorithm for spectrum monitoring applications," in *IEEE International Conference on Communications*, 1999, pp. 570-574.

[6] D. Boudreau*, et al.*, "A fast automatic modulation recognition algorithm and its implementation in a spectrum monitoring application," in *IEEE Military Communications Conference*, 2000, pp. 732-736.

[7] N. Kim*, et al.*, "Hierarchical classification of modulation signals," in *Fifth Biannual World Automation Congress*, 2002, pp. 243-248.

[8] A. B. Calson, *Communication systems - An introduction to signals and noise in electrical communication*, Third ed.: McGraw-Hill, Inc., New York, 1986.

[9] E. E. Azzouz and A. K. Nandi, "Automatic identification of digital modulation types," *Signal Processing,* vol. 47, pp. 55-69, 1995.

[10] A. K. Nandi and E. E. Azzouz, "Automatic analogue modulation recognition," *Signal Processing,* vol. 46, pp. 211-222, 1995.

[11] A. K. Nandi and E. E. Azzouz, "Algorithms for automatic modulation recognition of communication signals," *Communications, IEEE Transactions on,* vol. 46, pp. 431-436, 1998.

[12] M. K. Simon and S. Dolinar, "Improving SNR estimation for autonomous receivers," *Communications, IEEE Transactions on,* vol. 53, pp. 1063-1073, 2005.

[13] D. R. Pauluzzi and N. C. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *Communications, IEEE Transactions on,* vol. 48, pp. 1681-1691, 2000.

[14] B. Shah and S. Hinedi, "The split symbol moments SNR estimator on narrow-band channels," *IEEE Trans. Aerospace and Electronic Systems,* vol. 26, pp. 737-747, Sept. 1990.

[15] R. Gagliardi and C. Thomas, "PCM Data Reliability Monitoring Through Estimation of Signal-to-Noise Ratio," *Communication Technology, IEEE Transactions on,* vol. 16, pp. 479-486, 1968.

[16] C. E. Gilchriest, "Signal-to-noise monitoring," *JPL Space Programs Summary,* vol. IV, pp. 169-184, Jun. 1966.

[17] R. Matzner, "An SNR Estimation Algorithm for Complex Baseband Signals Using Higher Order Statistics," *FACTA UNIVERSITATIS,* vol. 6, pp. 41-52, 1993.

[18] D. Sui and L. Ge, "On the Blind SNR Estimation for IF Signals," in *Innovative Computing, Information and Control, 2006. ICICIC '06. First International Conference on*, 2006, pp. 374-378.

[19] D. Sui and L. Ge, "A Blind SNR Estimator for Digital Bandpass Signals," in *ITS Telecommunications Proceedings, 2006 6th International Conference on*, 2006, pp. 1334-1337.

[20] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *Acoustics, Speech and Signal Processing, IEEE Transactions on,* vol. 33, pp. 387-392, 1985.

[21] K. Kim and A. Polydoros, "Digital Modulation classification: the BPSK and QPSK Case," in *MILCOM*, San Diego, CA, 1988, pp. 431-436.

[22] H. Chung-Yu and A. Polydoros, "Likelihood methods for MPSK modulation classification," *Communications, IEEE Transactions on,* vol. 43, pp. 1493-1504, 1995.

[23] P. C. Sapiano and J. D. Martin, "Maximum likelihood PSK classifier," in *Military Communications Conference, 1996. MILCOM '96, Conference Proceedings, IEEE*, 1996, pp. 1010-1014 vol.3.

[24] B. F. Beidas and C. L. Weber, "Higher-order correlation-based approach to modulation classification of digitally frequency-modulated signals," *Selected Areas in Communications, IEEE Journal on,* vol. 13, pp. 89-101, 1995.

[25] W. Wen and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *Communications, IEEE Transactions on,* vol. 48, pp. 189-193, 2000.

[26] L. Hong*, et al.*, "Quasi-hybrid likelihood modulation classification with nonlinear carrier frequency offsets estimation using antenna arrays," in *Military Communications Conference, 2005. MILCOM 2005. IEEE*, 2005, pp. 570-575 Vol. 1.

[27] W. Su*, et al.*, "Real-time modulation classification based on maximum likelihood," *Communications Letters, IEEE,* vol. 12, pp. 801-803, 2008.

[28] F. Hameed*, et al.*, "On the likelihood-based approach to modulation classification," *Wireless Communications, IEEE Transactions on,* vol. 8, pp. 5884-5892, 2009.

[29] S. Z. Hsue and S. S. Soliman, "Automatic modulation recognition of digitally modulated signals," in *Military Communications Conference, 1989. MILCOM '89. IEEE*, 1989, pp. 645-649 vol.3.

[30] S. S. Soliman and S. Z. Hsue, "Signal Classification Using Statistical Moments," *Communications, IEEE Transactions on,* vol. 40, pp. 908-916, 1992.

[31] Y. Yawpo and S. S. Soliman, "A suboptimal algorithm for modulation classification," *Aerospace and Electronic Systems, IEEE Transactions on,* vol. 33, pp. 38-45, 1997.

[32] Y. Yawpo and L. Ching-Hwa, "An asymptotic optimal algorithm for modulation classification," *Communications Letters, IEEE,* vol. 2, pp. 117-119, 1998.

[33] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *Communications, IEEE Transactions on,* vol. 48, pp. 416-429, 2000.

[34] Z. Yu*, et al.* (2003). *M-ary frequency shift keying signal classification based-on discrete Fourier transform*. *2*.

[35] O. A. Dobre*, et al.*, "The classification of joint analog and digital modulations," in *Military Communications Conference, 2005. MILCOM 2005. IEEE*, 2005, pp. 3010-3015 Vol. 5.

[36] T. He*, et al.*, "Modulation Classification Based on Multifractal Features," in *ITS Telecommunications Proceedings, 2006 6th International Conference on*, 2006, pp. 152-155.

[37] K. Minah*, et al.*, "Automatic recognition of analog and digital modulation signals using DoE filter," in *Communications and Information Technology, 2009. ISCIT 2009. 9th International Symposium on*, 2009, pp. 609-614.

[38] O. A. Dobre*, et al.*, "Survey of automatic modulation classification techniques: classical approaches and new trends," *Communications, IET,* vol. 1, pp. 137-156, 2007.

[39] B. Ramkumar, "Automatic modulation classification for cognitive radios using cyclic feature detection," *Circuits and Systems Magazine, IEEE,* vol. 9, pp. 27-45, 2009.

[40] Z. Yu, "Automatic Modulation Classification of Communication Signals," Ph.D, Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ, 2006.

[41] J. G. Proakis, *Digital communications*, Third ed.: McGraw-Hill, Inc., New York, 1995.

[42] A. V. Dandawate, "Exploring Cyclostationarity and Higher-order Statistics in Signal Processing," Ph.D Dissertation, University of Virginia, Charlottesville, 1993.

[43] A. V. Dandawate and G. B. Giannakis, "Nonparametric polyspectral estimators for kth-order (almost) cyclostationary processes," *Information Theory, IEEE Transactions on,* vol. 40, pp. 67-84, 1994.

[44] W. A. Gardner and C. M. Spooner, "The cumulant theory of cyclostationary time-series, Part I: foundation," *IEEE Transactions on Signal Processing,* vol. 42, pp. 3287-3408, Dec 1994.

[45] W. Gardner, "Spectral Correlation of Modulated Signals: Part I--Analog Modulation," *Communications, IEEE Transactions on [legacy, pre - 1988],* vol. 35, pp. 584-594, 1987.

[46] W. Gardner*, et al.*, "Spectral Correlation of Modulated Signals: Part II--Digital Modulation," *Communications, IEEE Transactions on [legacy, pre - 1988],* vol. 35, pp. 595-601, 1987.

[47] P. Ciblat*, et al.*, "Performance analysis of blind carrier frequency offset estimators for noncircular transmissions through frequency-selective channels," *IEEE Transactions on Signal Processing,* vol. 50, pp. 130-140, Jan 2002.

[48] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio Electroacoustics,* vol. 15, pp. 70-73, 1967.

[49] S. Qian and D. Chen, *Joint Time-Frequency Analysis - Methods and Applications*: Prentice Hall, 1996.

[50] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*: John Wiley & Sons, 1996.

[51] S. Ohara*, et al.*, "Modulation classification based on amplitude and cosine moments," in *Communications, Computers and Signal Processing, 2009. PacRim 2009. IEEE Pacific Rim Conference on*, 2009, pp. 158-162.

[52] K. N. Haq, *et al.*, "Recognition of digital modulated signals based on statistical parameters," in *4th IEEE International Conference on Digital Ecosystems and Technologies - Conference Proceedings of IEEE-DEST 2010, DEST 2010*, 2010, pp. 565-570.

[53] S. B. Sadkhan, "A proposed digital modulated signal identification based on pattern recognition," in *2010 7th International Multi-Conference on Systems, Signals and Devices, SSD-10*, 2010.

[54] A. Ebrahimzadeh and R. Ghazalian, "Blind digital modulation classification in software radio using the optimized classifier and feature subset selection," *Engineering Applications of Artificial Intelligence,* vol. 24, pp. 50-59, 2011.

[55] B. G. Mobasseri, "Digital Modulation Classification Using Constellation Shape," *Signal Processing,* vol. 80, pp. 251-277, 1997.

[56] F. F. Liedtke, "Computer simulation of an automatic classification procedure for digitally modulated communication signals with unknown parameters," *Signal Processing,* vol. 6, pp. 311-323, 1984.

[57] Y.Lin and C. C. Kuo, "Classficiation of quadrature amplitude modulated(QAM) signals via sequential probability ratio test(SPRT)," Report of CRASP, university of Southern California, July 15, 1996.

[58] K. M. Chugg, *et al.*, "Combined likelihood power estimation and multiple hypothesis modulation classification," in *Asilomar-29*, 1996, pp. 1137-1141.

[59] F. Jondral, "Automatic classification of high frequency signals," *Signal Processing,* vol. 9, pp. 177-190, 1985.

[60] E. R. Adams, *et al.*, "A statistical approach to modulation recognition for communications signal monitoring and surveillance," in *Proceedings of the 1988 International Conference on Digital Processing of Signals in Communications*, 1988, pp. 31-37.

[61] L. V. Domínguez, *et al.*, "A general approach to the automatic classification of radio communication signals," *Signal Processing,* vol. 22, pp. 239-250, 1991.

[62] J. E. Hipp, "Modulation classification based on statistical moments," in *Proceedings of the 1986 IEEE Military Communications Conference*, 1986, pp. 20.2.1-20.2.6.

[63] S. Taira and E. Murakami, "Automatic classification of analogue modulation signals by statistical parameters," in *Military Communications Conference Proceedings, 1999. MILCOM 1999. IEEE*, 1999, pp. 202-207 vol.1.

[64] O. A. Dobre*, et al.*, "The classification of joint analog and digital modulations," in *Military Communications Conference, 2005. IEEE*, pp. 3010-3015 Vol. 5.

[65] Y. T. Chan and L. G. Gadbois, "Identification of the modulation type of a signal," *Signal Processing,* vol. 16, pp. 149-154, 1989.

[66] P. A. J. Nagy, "Analysis of a method for classification of analogue modulated radio signals," presented at the Proceedings of the 1994 European Association for Signal Processing VII Conference, 1994.

[67] S. D. Jovanovic*, et al.*, "Recognition of low modulation index AM signals in additive Gaussian noise," presented at the Proceedings of the 1994 European Association for Signal Processing V Conference, Edinburgh, Scotland, 1994.

[68] Y. O. Al-Jalili, "Identification algorithm of upper sideband and lower sideband SSB signals," *Signal Processing,* vol. 42, pp. 207-213, March 1995.

[69] I. Drycjmann*, et al.*, "Automatic modulation type recognition," presented at the Proceedings of the 1998 IEEE Canadian Conference on Electrical and Computer Engineering, 1998.

[70] B. Seaman and R. M. Braun, "Using cyclostationarity in the modulation classification of analogue signals," in *Communications and Signal Processing, 1998. COMSIG '98. Proceedings of the 1998 South African Symposium on*, 1998, pp. 261-266.

[71] J. R. Waller and G. D. Brushe, "A method for differentiating between frequency and phase modulated signals," presented at the Proceddings of Conference on Information, Decision and Control, 1999.

[72] H. Xiao, *et al.*, "An investigation of non-data-aided SNR estimation techniques for analog modulation signals," in *Sarnoff Symposium, 2010 IEEE*, pp. 1-5.

[73] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals," *Proceedings of the IEEE,* vol. 80, pp. 520-538, 1992.

[74] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. II. Algorithms and applications," *Proceedings of the IEEE,* vol. 80, pp. 540-568, 1992.

[75] S. Z. Hsue and S. S. Soliman, "Automatic modulation classification using zero crossing," *Radar and Signal Processing, IEE Proceedings F,* vol. 137, pp. 459-464, 1990.

[76] A. Nehorai and D. Starer, "An adaptive SSB carrier frequency estimator," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '87*, 1987, pp. 2109-2112.

[77] C. Hsu, *et al.* (http://www.csie.ntu.edu.tw/~cjlin/libsvm/). *A practical guide to support vector classification, May, 2012*.

[78] G. Peng, *et al.*, "Modulation Index Estimation of Frequency and Phase Modulated Signals," *Int. J. Communications, Network and System Sciences,* vol. 3, pp. 773-778, 2010.

[79] A. Ôengür and H. Güldem, "An Educational Interface for Automatic Recognition of Analog Modulated Signals," *Journal of Applied Sciences,* vol. 5, pp. 513-516, 2005.

[80] M. K. Simon and J. Hamkins, "Modulation Index Estimation," in *Autonomous Software-Defined Radio Receivers for Deep Space Applications*, ed: Wiley & Sons, Inc., Hoboken, NJ, USA, 2006.

[81] K. V. Mardia, *Statistics of Directional Data*: New York: Academic, 1972.

[82] G. Marsaglia, "Ratios of Normal Variables and Ratios of Sums of Uniform Variables," *Journal of the American Statistical Association,* vol. 60, pp. 193-204, Mar., 1965.

[83] G. Marsaglia, "Ratios of Normal Variables," *Journal of Statistical Software,* vol. 16, May 2006.

[84] A. Napolitano and C. M. Spooner, "Cyclic Spectral Analysis of Continuous-Phase Modulated Signals," *IEEE Transactions on Signal Processing,* vol. 49, pp. 30-44, 2001.