2007

Non-Cooperative Modulation Recognition Via Exploitation of Cyclic Statistics

Eric C. Like

Wright State University

Non-Cooperative Modulation Recognition Via Exploitation of Cyclic Statistics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Engineering

By

Eric C. Like
B.S.E.E., University of Texas at Austin, 2003

2007
Wright State University

Wright State University
SCHOOL OF GRADUATE STUDIES

December 14, 2007

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Eric C. Like ENTITLED Non-Cooperative Modulation Recognition Via Exploitation of Cyclic Statistics BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science in Engineering.

Zhiqiang Wu, Ph.D.
Thesis Director

Fred Garber, Ph.D.
Department Chair

Committee on
Final Examination

Zhiqiang Wu, Ph.D.

Fred Garber, Ph.D.

Kuan-Lun Chu, Ph.D.

Joseph F. Thomas, Jr., Ph.D.
Dean, School of Graduate Studies

ABSTRACT

Like, Eric. M.S.E., Department of Electrical Engineering, Wright State University, 2007. *Non-Cooperative Modulation Recognition Via Exploitation of Cyclic Statistics.*

This research proposes and evaluates a feature based modulation classification system designed to discriminate between AM, BFSK, OFDM, DS-CDMA, 4-ASK, 8-ASK, BPSK, QPSK, 8-PSK, 16- PSK, 16-QAM, and 64-QAM signals without a priori knowledge of critical signal parameters, including carrier frequency, symbol rate, or phase offset, among others. The classifier is based on the principles of cyclostationarity and leverages cyclic statistics to make its classification decision. The classification process is performed in a hierarchical process in order to exploit the lower variance of lower order cyclic statistics before making use of the higher order cyclic statistics. The initial classification is based on the estimated Spectral Coherence Function (SOF) of the received signal, followed by estimates of the signal's fourth- through eighth-order Cyclic Cumulants (CC). The performance of the classification system is evaluated under fading flat, two-path, and 20-path transmission channels, and multiantenna combining methods are exploited to increase the system performance.

Contents

List of Figures

Acknowledgements

There are many people I would like thank for their support and knowledge, without whom this research would not have been possible. Firstly, to my thesis advisor, Dr. Zhiqiang (John) Wu, thank you for your enthusiasm, patience, and guidance.

I would also like to express my sincere appreciation to my friends and family, thank you for the constant support and encouragement.

1 Introduction

1.1 Research Motivation

The demand for spectrum usage has seen a considerable increase in recent years. As a result, novel methods to maximize the use of the available spectrum have been proposed. One critical area is through the use of cognitive radio [1][2]. Traditionally, wireless devices wishing to access the spectrum were required to use the band licensed to them by the FCC. As the number of wireless users has increased, there has been a corresponding decrease in the amount of available spectrum. Cognitive radio seeks to relieve this burden by determining which areas of the spectrum are in use at a particular time. If a given band of the spectrum is not currently being used by the licensed user, that band could be used by another system. To achieve this mere power detection would suffice. However, a more efficient method to maximize the use of the available spectrum would be to not simply avoid frequency bands that are in use, but rather to limit the amount of in-band transmission down to an acceptably low level so as to avoid interfering with the primary user. Since different modulation schemes are able to tolerate different amounts of interference, the modulation scheme of the current user will have to be determined. In this case, merely detecting the presence of the signal will not be sufficient.

Modulation recognition has been a area of ongoing research for over two decades. As a result, there are numerous methods that have been developed to estimate the modulation

scheme of an unknown signal. Each method makes a set of assumptions in order to make a classification, and generally only operates reliably under the limited scenario for which it is designed. These assumptions range from knowledge of signal parameters, such as carrier frequency, symbol rate, or phase offset, to knowledge of the transmission channel model. Therefore, there is an ongoing need for a reliable modulation recognition system that can operate without the

1.2 Problem Statement

The goal of this research is to develop a non-cooperative modulation recognition system capable of reliably identifying the modulation scheme of a received signal. The system should not require knowledge of critical signal parameters, such as carrier frequency, symbol rate, or phase offset, among others.

1.3 Research Approach

Cyclostationary based approaches have been developed for modulation recognition, and have demonstrated considerable ability to distinguish between different modulation schemes without knowledge of critical signal parameters. In particular, the Spectral Coherence Function (SOF) has been shown to produce distinct features for lower level modulation schemes. The classifier proposed in [3] leveraged the SOF of received signals to classify BPSK, QPSK, FSK, MSK, and AM signals in AWGN channels at low SNR. However, the ability of the SOF to classify signals in multipath channels was not assessed, and the classification set is desired to be expanded to higher-level modulations. Additionally, the classifier assumed that any cyclic features would occur within the reduced subset of cycle frequencies for which the SOF is estimated. Thus, in order to ensure that all cyclic fea-

tures are identified, an prohibitive amount of data must be analyzed, leading to a highly inefficient classification system.

Higher order cyclic statistics have also been exploited for modulation recognition, where they have been shown to distinguish between higher order modulation schemes. In [4], cyclic cumulants (CC) were used to classify M-ASK, M-PSK, and M-QAM signals, of various orders M in the presence of flat fading channels. However, the classifier assumed precise knowledge of the symbol rate of the received signal, as well as that the carrier frequency of the signal had been perfectly removed and the signal centered at baseband.

The research performed in this thesis is focused on leveraging the benefits of both of the previous classifiers to produce a classification decision on a wide range of signals without the a priori knowledge of critical signal parameters. The system is designed to discriminate between designed AM, BFSK, OFDM, DS-CDMA, 4-ASK, 8-ASK, BPSK, QPSK, 8-PSK, 16- PSK, 16-QAM, and 64-QAM signals. Additionally, the performance of the classifier in the presence of multipath channels is assessed, and methods to increase the reliability of the classifier in these channels are investigated.

1.4 Thesis Organization

Chapter 2 introduces the basics of modulation recognition, and describes the general categories modulation classification schemes, where a feature based (FB) classification system is chosen for implementation employing cyclic statistics. The theory and background of cyclostationary statistics is given, where it is presented in the traditionary stochastic probabilistic (SP) framework as well as a fraction of time (FOT) probabilistic framework. The SOF and CC are presented as critical statistics to be used as features in the final classifier design are presented, and the impact that multipath channels have on their performance is presented.

In Chapter 3, methods to implement the estimation of the feature sets are investigated, and a suitable estimation process is obtained. The final modulation recognition system design is presented as a multistage classifier based on the SOF and CC, and its operation explained. Various methods to improve performance in multipath channels are presented, and a process is presented to increase the classification reliability.

Chapter 4 provides simulation results and analysis of the classification system under various channel conditions. The performance of the classifier is compared to the baseline classifier presented in [3] based on the SOF, as well as to the classifier given in [4] based on the CC.

Finally Chapter 5 presents a summary of the results, conclusions drawn from the research, and provides recommendations for possible future research.

2 Background

2.1 Introduction

This chapter presents an overview of the fundamental theory that will be used throughout the remainder of this thesis. Section 2.2 covers the various classes of modulation recognition algorithms that have been used in the past. In Section 2.3 the theory of cyclostationarity and how it applies to communication signals is presented, where it is given in the context of both the traditional stochastic process (SP) probabilistic framework as well as the fraction of time (FOT) probabilistic framework. In Section 2.4, various methods used to estimate the cyclic spectrum of signals are presented. The necessary generalization of cyclostationarity to higher order statistics is given in Section 2.5 before the chapter concludes with a discussion of the effects that multipath channels have on the estimation of cyclic statistics in in Section 2.6.

2.2 Modulation Recognition Algorithms

Modulation recognition has been a subject of considerable research for over two decades. Classification schemes can generally be divided into one of two broad categories - likelihood based (LB) approaches, which form tests based on the statistics of the received signal, and feature based (FB) approaches, which attempt to derive reliable estimators based on

critical features of the received signal.

2.2.1 Likelihood Based Tests

LB approaches attempt to explicitly model the probability distribution of the received signals. Depending on the degree of information known about the signals being discriminated, LB approaches can be further divided into three distinct techniques - average likelihood ratio tests (ALRT), generalized likelihood ratio tests (GLRT), and hybrid likelihood ratio tests (HLRT). In ALRT based methods, each of the unknown signal and channel parameters are treated as random variables with known probability density functions (pdf). These pdfs are used to compute the likelihood of each potential modulation scheme, and the modulation scheme corresponding to greatest likelihood is selected . Though this method results in the optimal classifier in the Bayesian sense, optimality can only be claimed when the pdfs are accurately estimated [5].

In many situations, the appropriate distribution of the signal and channel statistics are not known, resulting in a classifier that is ill-suited for the current situation. This problem can be resolved by treating the unknown parameters as deterministic but unknown variables. In this case, the best performance is obtained by the uniformly most powerful (UMP) test, when it exists [6]. If a UMP test does not exist, the maximum likelihood (ML) estimates of the unknown parameters can be computed. These ML estimates can be used in a likelihood ratio test, which gives the GLRT. This method does not require any assumptions about the signal or channel parameters, resulting in a classifier structure that will be applicable to a variety of environments [5].

ALRT based methods suffer from the necessity of having an accurate estimate of the signal and channel pdfs, and from the computationally intensive requirement of a multidimensional integration. GLRT based methods on the other hand are less computationally intensive, but maximization over the unknown data symbols can lead to the same value of

the likelihood function for nested signal constellations (e.g., BPSK and QPSK), which will yield incorrect signal classifications. The HLRT attempts to exploit the benefits of each scheme by modeling the data symbols as random variables, and the remaining signal and channel statistics as unknown deterministic variables. ML estimates of the various signal and channel parameters are formed as in the GLRT, while still averaging over the unknown data symbols as in ALRT. However, as the number of unknown parameters under consideration are increased, the computation required to find their ML estimates can become time consuming, rendering the algorithm unsuitable to operate in a real-time manner [7].

2.2.2 Feature Based Tests

Rather than attempting to explicitly model the underlying signal and channel pdfs, FB approaches attempt to extract critical statistics from received signals to make a classification based on the reduced data set. This can frequently be performed at a fraction of the complexity of LB systems. While FB methods are suboptimal in the Bayesian sense, they often provide near optimal performance close to that of the LB approaches. Furthermore, FB algorithms are commonly less sensitive to modeling errors which would significantly degrade the performance of LB algorithms [7].

In general, there is no well defined method specified to select appropriate features for classification purposes, resulting in a wide array of feature sets that have been proposed as suitable indicators of a signal's modulation scheme. Features that have been used include statistics derived from the instantaneous amplitude, phase and frequency of received signals, zero crossing intervals, wavelet transforms, amplitude and phase histograms, and constellation shapes, as well as many others [7, 8, 9].

However, even though FB methods do not explicitly make use of the pdf of the signal parameters, many do require a priori knowledge of critical signal parameters, such as the carrier frequency, carrier phase, symbol rate, or timing offset, among others. Additionally,

like many LB methods, FB methods frequently assume that the carrier frequency of the received signal has been removed and that the signal has been sampled at the symbol rate without timing errors prior to processing. Generally, knowledge of these statistics is not known a priori, and their necessity may render the classification algorithm ineffective.

Cyclostationary based (CS) based approaches have been investigated as a potential method for modulation recognition. CS based methods have been shown to be insensitive to unknown signal parameters and to preserve the phase information in the signal, a claim that cannot be made by methods based on purely stationary statistics. This allows the signal's parameters to be estimated from the cyclic statistics directly. In cases when the classification algorithm has been designed to be insensitive to these parameters, the unknown parameters can be neglected altogether. Furthermore, the effects due to multipath channels can be shown to introduce negligible effects to certain cyclic statistics, rendering them insensitive to non-ideal channel situations [10, 11].

Due to the inherent insensitivity of cyclic statistics to unknown signal parameters, as well as their robustness to channel effects, CS based methods are investigated as a modulation recognition scheme. The following section provides an overview of the fundamental theory behind cyclic statistics which will be used to derive a classification system.

2.3 Fundamentals of Cyclostationarity

The theory of cyclostationarity is based on the fact that certain random signals have statistics that vary periodically in time. These signals, known as cyclostationary signals (CS), are in contrast to stationary signals whose statistics remain constant in time. This process can be viewed in the stochastic process (SP) probabilistic framework, or in a simplified fraction of time (FOT) probabilistic framework. The SP framework is the more traditional method used in engineering analysis. However, for reasons which will be outlined below,

this framework is not always the most appropriate when dealing with cyclostationary signals. Additionally, when only one realization of a random signal is available, the need to abstract to the concept of an ensemble with an associated random process is not always necessary or even useful. In these instances, it can be more beneficial to employ the concept of the FOT framework, which is based on a single realization of the random signal [11]. The two frameworks are described in the following two sections, where it is noted that the two frameworks become equivalent under the assumption of cycloergodicity (CE).

2.3.1 Stochastic Process Probabilistic Framework

In the SP framework, the Nth-order probability distribution of a random signal $\mathbf{X}(t)$ is defined as [12]

$$\begin{aligned} &F_{\mathbf{X}(t+\tau_1)\cdots\mathbf{X}(t+\tau_{N-1})\mathbf{X}(t)}(\xi_1,\ldots,\xi_{N-1},\xi_N) \\ &\equiv P\{\mathbf{X}(t+\tau_1)\leq\xi_1,\ldots,\mathbf{X}(t+\tau_N)\leq\xi_{N-1},\mathbf{X}(t)\leq\xi_N\} \end{aligned} \tag{2.1}$$

which can be represented equivalently as [11]

$$\begin{aligned} &F_{\mathbf{X}(t+\tau_1)\cdots\mathbf{X}(t+\tau_{N-1})\mathbf{X}(t)}(\xi_1,\ldots,\xi_{N-1},\xi_N) \\ &\equiv E\left\{I\left[\xi_N-\mathbf{X}(t)\right]\prod_{k=1}^{N-1}I\left[\xi_k-\mathbf{X}(t+\tau_k)\right]\right\} \end{aligned} \tag{2.2}$$

where $I[\xi]$ is the event indicator or step function

$$I[\xi]\equiv\begin{cases}1, & \xi>0\\ \\ 0, & \xi\leq 0\end{cases} \tag{2.3}$$

A stochastic process $\mathbf{X}(t)$ is said to be Nth-order cyclostationary in the strict sense if its Nth and lower-order probability distribution is periodic in t with some period, T_0, such that

$$\begin{aligned} & F_{\mathbf{X}(t+\tau_1+T_0)\cdots\mathbf{X}(t+\tau_{N-1}+T_0)\mathbf{X}(t+T_0)}(\xi_1,\ldots,\xi_{N-1},\xi_N) \\ & = F_{\mathbf{X}(t+\tau_1)\cdots\mathbf{X}(t+\tau_{N-1})\mathbf{X}(t)}(\xi_1,\ldots,\xi_{N-1},\xi_N) \\ & \forall t\epsilon\Re \qquad \forall(\tau_1,\ldots,\tau_{N-1})\epsilon\Re^{N-1} \\ & \forall(\xi_1,\ldots,\xi_N)\epsilon\Re^N. \end{aligned} \tag{2.4}$$

The random process $\mathbf{X}(t)$ is said to be Nth-order wide sense cyclostationary if its nth-order/q-conjugate temporal moment functions (TMF)

$$\mathscr{R}_X(t,\boldsymbol{\tau})_{n,q} \equiv E\{\mathbf{X}^{(*)_n}(t)\prod_{i=1}^{n-1}\mathbf{X}^{(*)_i}(t+\tau_i)\} \tag{2.5}$$

are periodic with some period, T_0, such that

$$\begin{aligned} & \mathscr{R}_X(t+T_0,\boldsymbol{\tau})_{n,q} = \mathscr{R}_X(t,\boldsymbol{\tau})_{n,q} \\ & \forall t\epsilon\Re \qquad \forall(\tau_1,\ldots,\tau_{n-1})\epsilon\Re^{n-1} \\ & \forall n = 1,\ldots,N \qquad q = 0,\ldots,n \end{aligned} \tag{2.6}$$

where $\boldsymbol{\tau} \equiv [\tau_1,\ldots,\tau_{n-1}]$ and $(*)_i$ denotes ones of the q possible conjugations. The nth-order/q-conjugate TMF of $\mathbf{X}(t)$ will then permit a Fourier Series expansion

$$\mathscr{R}_X(t,\boldsymbol{\tau})_{n,q} = \sum_{\{\alpha\}} \mathscr{R}_X^{\alpha}(\boldsymbol{\tau})_{n,q} e^{j2\pi\alpha t} \tag{2.7}$$

where $\alpha = k/T_0$, and the Fourier coefficients are calculated from

$$\mathscr{R}_X^{\alpha}(\boldsymbol{\tau})_{n,q} \equiv \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} \mathscr{R}_X(t,\boldsymbol{\tau})_{n,q} e^{-j2\pi\alpha t} dt \tag{2.8}$$

where $\mathscr{R}_x^{\alpha}(\boldsymbol{\tau})_{n,q}$ is known as the Nth-order Cyclic Temporal Moment Function (CTMF). In the case where $\mathscr{R}_x(t,\boldsymbol{\tau})_{n,q}$ is composed of multiple, possibly incommensurate periodicities, the above equation can be reexpressed as

$$\mathscr{R}_X^{\alpha}(\boldsymbol{\tau})_{n,q} \equiv \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \mathscr{R}_X(t,\boldsymbol{\tau})_{n,q} e^{-j2\pi\alpha t} dt \tag{2.9}$$

and the sum in (2.7) is over the set of all α for which the CTMF is not identically zero as a function of $\boldsymbol{\tau}$. In either case, the set $\{\alpha\}$ is termed the set of cycle frequencies.

When the TMF is evaluated for $n = 2$ and $q = 1$, it is termed the autocorrelation function (AF) and, in order to preserve symmetry with respect to the lag variables, is redefined as

$$\begin{aligned} \mathscr{R}_X(t,\tau) &\equiv E\{\mathbf{X}(t+\tau/2)\mathbf{X}^*(t-\tau/2)\} \\ &= \sum_{\{\alpha\}} \mathscr{R}_X^{\alpha}(\tau) e^{j2\pi\alpha t} \end{aligned} \tag{2.10}$$

where $\mathscr{R}_x^{\alpha}(\tau)$ is termed the cyclic autocorrelation function (CAF). The Fourier transform of the CAF is denoted the Spectral Correlation Function (SCF), given by

$$\mathscr{S}_x^{\alpha}(f) = \int_{-\infty}^{\infty} \mathscr{R}_X^{\alpha}(\tau) e^{-j2\pi f\tau} d\tau \tag{2.11}$$

The SCF gets its name from the equivalent representation

$$\mathscr{S}_X^\alpha(f) = E\{\mathcal{X}(f+\alpha/2)\mathcal{X}^*(f-\alpha/2)\} \tag{2.12}$$

where $\mathcal{X}(f)$ is the Fourier transform of a sample path of $\mathbf{X}(t)$

$$\mathcal{X}(f) \equiv \int_{-\infty}^{\infty} x(t)e^{-2\pi ft}dt \tag{2.13}$$

From the above two equations, it can be seen that the SCF is a true measure of the correlation between the spectral components of $\mathbf{X}(t)$ in the stochastic sense.

To obtained a statistic that is independent of signal strength, a normalized version of the SCF is desired. The Spectral Coherence Function (SOF) is given as:

$$C_X^\alpha(f) = \frac{\mathscr{S}_X^\alpha(f)}{\left[\mathscr{S}_X^0\left(f+\frac{\alpha}{2}\right)^* \mathscr{S}_X^0\left(f-\frac{\alpha}{2}\right)\right]^{1/2}} \tag{2.14}$$

where the SOF is seen to be a proper coherence value with a magnitude in the range of $[0,1]$.

One significant drawback to applying the SP framework to periodic signals is observed by noting that purely stationary random processes may have periodic sample paths, resulting in behavior that cannot be predicted from traditional probabilistic analysis. One common example of this situation is a communication signal with a random phase component. This phenomenon can be removed by the assumption that the random process is cycloergodic (CE). However, this assumption is generally not valid for communication signals or other situations when the unknown phase of the carrier signal is treated as a random variable.

2.3.2 Fraction of Time Probabilistic Framework

In the FOT Probabilistic Framework, the underlying probability distribution is based on a single time series, without the generalization to the ensemble used in the SP framework. This is generally beneficial as it removes a level of conceptual complexity and is more applicable to cases where only one realization of the random signal is obtainable. In this case, the first-order FOT probability distribution of a persistent time signal is defined as

$$\begin{aligned} F^0_{x(t)}(\xi) &\equiv P\{x(t) \leq \xi\} \\ &\equiv E^0\left\{I\left[\xi - x(t)\right]\right\} \end{aligned} \tag{2.15}$$

where $E^0\left\{\cdot\right\}$ is the time-average operator defined by

$$E^0\left\{h(t)\right\} \equiv \lim_{Z\to\infty} \frac{1}{2Z} \int_{-Z}^{Z} h(t+t')dt' \tag{2.16}$$

The above concept of probabilities in this framework therefore correspond to the fraction of time that a signal takes on a certain value, leading to the name FOT Probabilistic Framework. The Nth-order FOT probability distribution is defined similarly as

$$\begin{aligned} &F^0_{x(t+\tau_1)\cdots x(t+\tau_{N-1})x(t)}(\xi_1, \ldots, \xi_{N-1}, \xi_N) \\ &\equiv P\{x(t+\tau_1) \leq \xi_1, \ldots, x(t+\tau_N) \leq \xi_{N-1}, x(t) \leq \xi_N\} \\ &\equiv E^0\left\{I\left[\xi_N - x(t)\right] \prod_{k=1}^{N-1} I\left[\xi_k - x(t+\tau_k)\right]\right\} \end{aligned} \tag{2.17}$$

.

It can be shown that the Nth-order FOT probability distribution defined in (2.17) is independent of t and is therefore said to be a stationary probabilistic model [13]. To ac-

commodate signals whose Nth-order products exhibit periodic components, the FOT Probabilistic Framework is extended into the Cyclic FOT (CFOT) Probabilistic Framework. Here, the Nth-order CFOT probability distribution of a persistent time signal is defined as

$$\begin{aligned} & F^{\{\alpha\}}_{x(t+\tau_1)\cdots x(t+\tau_{N-1})x(t)}(\xi_1,\ldots,\xi_{N-1},\xi_N) \\ & \equiv P\{x(t+\tau_1)\leq\xi_1,\ldots,x(t+\tau_N)\leq\xi_{N-1},x(t)\leq\xi_N\} \\ & \equiv E^{\{\alpha\}}\left\{I\left[\xi_N-x(t)\right]\prod_{k=1}^{N-1}I\left[\xi_k-x(t+\tau_k)\right]\right\} \end{aligned} \tag{2.18}$$

where $E^{\{\alpha\}}\{\cdot\}$ is referred to as the multiple sine wave extraction operator defined by

$$E^{\{\alpha\}}\{h(t)\}\equiv\sum_{\alpha\epsilon\{\alpha\}}E^{\alpha}\{h(t)\}, \tag{2.19}$$

$E^{\alpha}\{\cdot\}$ is the single sine wave extraction operator

$$\begin{aligned} E^{\alpha}\{h(t)\} & \equiv \lim_{Z\to\infty}\frac{1}{2Z}\int_{-Z}^{Z}h(t+t')e^{-j2\pi\alpha t'}dt' \\ & \equiv E^{0}\left\{h(t)e^{-j2\pi\alpha t}\right\}e^{j2\pi\alpha t}, \end{aligned} \tag{2.20}$$

and the summation in (2.19) is over the set of all α for which $E^{\alpha}\{\cdot\}$ is not identically zero. The multiple sine wave extraction operator defined in (2.19) is analogous to the traditional expectation operator in the SP Probabilistic Framework and is used to remove non-constant/non-periodic components of the signal, which are treated as random fluctuations. Additionally, the similarity between the final line of (2.18) and (2.2) is apparent.

Similar to the SP Probabilistic Framework, a time signal $x(t)$ in the CFOT Probabilistic Framework is said to be Nth-order cyclostationary in the strict sense if its Nth and

lower-order CFOT probability distribution is periodic in t with some period, T_0, such that

$$\begin{aligned} & F^{\{\alpha\}}_{x(t+\tau_1+T_0)\cdots x(t+\tau_{N-1}+T_0)x(t+T_0)}(\xi_1,\ldots,\xi_{N-1},\xi_N) \\ & = F^{\{\alpha\}}_{x(t+\tau_1)\cdots x(t+\tau_{N-1})x(t)}(\xi_1,\ldots,\xi_{N-1},\xi_N) \\ & \forall t\epsilon\Re \qquad \forall(\tau_1,\ldots,\tau_{N-1})\epsilon\Re^{N-1} \\ & \forall(\xi_1,\ldots,\xi_N)\epsilon\Re^N. \end{aligned} \tag{2.21}$$

and the time signal $x(t)$ is said to be Nth-order wide sense cyclostationary if its nth-order/q-conjugate temporal moment functions (TMF)

$$R_X(t,\boldsymbol{\tau})_{n,q} \equiv E^{\{\alpha\}}\{x^{(*)_n}(t)\prod_{i=1}^{n-1}x^{(*)_i}(t+\tau_i)\} \tag{2.22}$$

are periodic with some period, T_0, such that

$$\begin{aligned} & R_X(t+T_0,\boldsymbol{\tau})_{n,q} = R_X(t,\boldsymbol{\tau})_{n,q} \\ & \forall t\epsilon\Re \qquad \forall(\tau_1,\ldots,\tau_{n-1})\epsilon\Re^{n-1} \\ & \forall n=1,\ldots,N \qquad q=0,\ldots,n. \end{aligned} \tag{2.23}$$

The nth-order/q-conjugate TMF of $x(t)$ will then permit a Fourier Series expansion

$$R_X(t,\boldsymbol{\tau})_{n,q} = \sum_{\{\alpha\}} R^{\alpha}_X(\boldsymbol{\tau})_{n,q}e^{j2\pi\alpha t} \tag{2.24}$$

where $\alpha = k/T_0$ and the Fourier coefficients are calculated from

$$R_X^{\alpha}(\boldsymbol{\tau})_{n,q} \equiv \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} R_X(t,\boldsymbol{\tau})_{n,q} e^{-j2\pi\alpha t} dt \tag{2.25}$$

and $R_x^{\alpha}(\boldsymbol{\tau})_{n,q}$ is known as the Nth-order Cyclic Temporal Moment Function (CTMF). Those frequencies $\{\alpha\}$ for which the CTMF is not identically zero are termed the set of cycle frequencies. For signals where $R_x(t,\boldsymbol{\tau})_{n,q}$ is composed of multiple, possibly incommensurate periodicities, the equation above can be reexpressed as

$$R_X^{\alpha}(\boldsymbol{\tau})_{n,q} \equiv \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} R_X(t,\boldsymbol{\tau})_{n,q} e^{-j2\pi\alpha t} dt \tag{2.26}$$

.

For the special case of $n = 2$ and $q = 1$, the TMF is again termed the autocorrelation function (AF) and is redefined similarly as in the SP case as

$$R_X(t,\tau) \equiv E^{\{\alpha\}}\{x(t+\tau/2)x^*(t-\tau/2)\}$$

$$= \sum_{\{\alpha\}} R_X^{\alpha}(\tau) e^{j2\pi\alpha t} \tag{2.27}$$

where $R_x^{\alpha}(\tau)$ is termed the cyclic autocorrelation function (CAF). The Fourier transform of the CAF is denoted the Spectral Correlation Function (SCF), given by

$$S_X^{\alpha}(f) = \int_{-\infty}^{\infty} R_X^{\alpha}(\tau) e^{-j2\pi f\tau} d\tau. \tag{2.28}$$

The interpretation of (2.28) as a spectral correlation arises from the fact that it can be shown that it is obtainable from [14, 11]

$$S_X^\alpha(f) = \lim_{T\to\infty} \lim_{\Delta t\to\infty} \frac{1}{\Delta t} \int_{-\Delta t/2}^{\Delta t/2} X_T\left(t, f+\frac{\alpha}{2}\right) X_T^*\left(t, f-\frac{\alpha}{2}\right) dt \tag{2.29}$$

where X_T is the windowed Fourier of x(t):

$$X_T(t,f) = \frac{1}{\sqrt{T}} \int_{t-T/2}^{t+T/2} x(u) e^{j2\pi f u} du \tag{2.30}$$

Thus, the SCF can be seen to be a measure of the temporal correlation of different spectral components of $x(t)$. The SOF is then defined similarly as

$$C_X^\alpha(f) = \frac{S_X^\alpha(f)}{\left[S_X^0\left(f+\frac{\alpha}{2}\right)^* S_X^0\left(f-\frac{\alpha}{2}\right)\right]^{1/2}} \tag{2.31}$$

where the SOF is again a proper coherence value with a magnitude in the range of $[0, 1]$.

The definitions of the necessary cyclic statistics given in this section are analogous to those given in Section 2.3.1. The primary difference being that statistics in the CFOT Probabilistic Framework are based on a single sample path and therefore relieve the unnecessary abstraction to an ensemble with an associated random process. However, when the assumption of a CE is valid in the SP Probabilistic Framework, the two frameworks will be equivalent. In this case, the statistics presented in this section can be interpreted as methods to produce reliable estimates of the SP statistics when only a single sample path is available. For the two reasons stated above, and due to the fact that in practice only a single realization of the signal is generally available for analysis, the notation associated with the CFOT Probabilistic Framework will be followed for the remainder of this thesis with the underlying assumption that the notation and corresponding concepts from either framework could be applied.

2.3.3 Benefits of Cyclic Statistics

Cyclic statistics have several benefits for signal analysis over traditional stationary statistics. One benefit of cyclic statistics is that they contain information about critical signal parameters for communications signals, such as the carrier frequency and symbol rate. Communications signals produce cyclic features caused by underlying periodicities in the signal such as sampling, modulating, scanning, multiplexing, and coding. These features can be used to identify the modulation scheme of the signal.

Another benefit of cyclic statistics can be observed by evaluating the CTMF at $\{\alpha\} = \{0\}$. The resulting equation for the CTMF is given by

$$R_X^0(\tau) \equiv E^{\{0\}}\{x(t+\tau/2)x^*(t-\tau/2)\} \equiv R_X(\tau) \tag{2.32}$$

where $R_X(\tau)$ is the traditional autocorrelation of a stationary signal. Additionally, the SCF of a signal evaluated at $\{\alpha\} = \{0\}$ is given by

$$S_X^0(f) = \int_{-\infty}^{\infty} R_X^0(\tau)e^{-j2\pi f\tau}d\tau \equiv S_X(f) \tag{2.33}$$

where $S_X(f)$ is the traditional Power Spectral Density (PSD) of a stationary signals. The traditional AF and PSD then can be considered special cases of the CAF and SCF provided by cyclostationary analysis, where each is evaluated at $\alpha = 0$. Cyclostationary statistics therefore provide a generalization of traditional stationary statistics, while still incorporating the original content.

When cyclic statistics are computed for purely stationary signals, i.e. those signals that do not exhibit cyclostationarity, the set of cycle frequencies will be limited solely to $\{\alpha\} = \{0\}$. In this case, the CAF and SCF are given by

$$R_X^{\alpha}(\boldsymbol{\tau}) = \begin{cases} R_X(\boldsymbol{\tau}) & \alpha = 0 \\ 0 & \alpha \neq 0 \end{cases} \tag{2.34}$$

$$S_X^{\alpha}(f)^{\alpha}(\boldsymbol{\tau}) = \begin{cases} S_X(f) & \alpha = 0 \\ 0 & \alpha \neq 0 \end{cases} . \tag{2.35}$$

Thus, for stationary signals, the cyclic statistics are identically zero for all $\{\alpha\} \neq \{0\}$.

A primary benefit of cyclic statistics is their ability to separate signals with distinct cycle frequencies. When analyzing a signal composed of the sum of two or more signals with distinct cycle frequencies, the CAF and the SCF of each signal can be evaluated independently without interference due to the other signals at each of the distinct cycle frequencies [13]. In the case of the SCF, this means that even if the PSD of the signals overlap, the cyclic spectrum will still be non-overlapping and the distinct features of each signal will be preserved.

Therefore, in the common case of a signal in the presence of additive stationary noise, since the noise is stationary it will have no cyclic features other than at $\alpha = 0$. Therefore, with an extremely low SNR when the signal is buried in noise, the signal's PSD will be completely masked. However each of the signal's non-zero cyclic features can still be determined without any interference from the noise, and the signal's SCF will be unaffected for $\alpha \neq 0$.

The ability of the SCF to separate the spectrum of overlapping signals, coupled with the fact that it more readily provides information about the signal than the PSD, make it an ideal foundation for modulation recognition.

2.4 Estimation of Cyclic Statistics

The cyclic statistics presented in Section 2.3.2 are defined in terms of signals with an infinite time duration. In practice, only a finite time length of the signal is available for analysis. For this reason, estimates of the above statistics must be computed based on the available data.

There are generally two methods used for cyclic spectral analysis: temporal smoothing methods and frequency smoothing methods. Each method is based on the cyclic periodogram estimate of the SCF [14]

$$S_{X_T}^{\alpha}(t,f) = X_T\left(t, f+\frac{\alpha}{2}\right) X_T^*\left(t, f-\frac{\alpha}{2}\right) \tag{2.36}$$

where $X_T(t,f)$ is defined as

$$X_T(t,f) = \int_{-\infty}^{\infty} a_T(t-u)x(u)e^{-j2\pi fu}du \tag{2.37}$$

and $a_T(t)$ is a data tapering window of width T. The temporal resolution of the cyclic periodogram is equal to the duration over which $X_T(t,f)$ is computed, $\Delta t = T$. Similarly, the frequency resolution is given by the corresponding frequency resolution of $X_T(t,f)$, $\Delta f \approx 1/T$, resulting in a time frequency resolution product of $\Delta t \Delta f \approx 1$. In order to achieve statistical reliability, a substantial amount of temporal or frequency smoothing must be used, increasing the time resolution product, $\Delta t \Delta f \gg 1$.

Temporal smoothing methods are based on smoothing (2.36) in time, resulting in the time smoothed cyclic periodogram estimate of the SCF

$$S_{X_T}^{\alpha}(t,f)_{\Delta t} = \int_{-\infty}^{\infty} X_T\left(u, f+\frac{\alpha}{2}\right) X_T^*\left(u, f-\frac{\alpha}{2}\right) g_{\Delta t}(t-u)du \tag{2.38}$$

where $g_{\Delta t}(t)$ is a data tapering window of width $\Delta t \gg 1/\Delta f \approx T$, which becomes the new time resolution of 2.38.

Frequency smoothing methods on the other hand are based on smoothing (2.36) in frequency, resulting in the frequency smoothed cyclic periodogram estimate of the SCF

$$S^{\alpha}_{X_T}(t,f)_{\Delta f} = \int_{-\infty}^{\infty} X_T\left(t, v+\frac{\alpha}{2}\right) X_T^*\left(t, v-\frac{\alpha}{2}\right) g_{\Delta f}(f-v)dv \tag{2.39}$$

where $g_{\Delta f}(f)$ is a data tapering window of width $\Delta f \gg 1/\Delta t = 1/T$, with Δf the resulting frequency resolution of 2.39.

In each case, it can be shown that the cycle frequency resolution $\Delta\alpha \approx 1/\Delta t$, resulting in the equivalent requirement for statistical reliability, $\Delta f \gg \Delta\alpha$. Thus, the SCF estimate must be resolved much finer in cycle frequency than in spectral frequency in order to achieve a reliable estimate.

It can be shown that as the width of the data tapering windows increase, the estimates produced by the temporal and the spectral smoothing methods approach the ideal SCF

$$\begin{aligned} S^{\alpha}_X(f) &= \lim_{\Delta t\to\infty}\lim_{T\to\infty} S^{\alpha}_X(t,f)_{\Delta t} \\ &= \lim_{\Delta t\to\infty}\lim_{T\to\infty} S^{\alpha}_X(t,f)_{\Delta f} \end{aligned} \tag{2.40}$$

While both methods give similar approximations to the ideal SCF, each is better suited for different situations. Temporal smoothing methods are generally more computationally efficient for general cyclic spectral analysis. Frequency smoothing methods on the other hand are more efficient for producing high-accuracy estimates of the SCF for a relatively small set of known cycle frequencies [15, 16, 14]. This result will be explored further in Section 3.2.

2.5 Higher Order Cyclic Statistics

The methods outlined in the previous section can be used to efficiently estimate the SCF of received signals. However, these methods quickly become computationally infeasible when applied to estimating higher order cyclic polyspectrum. For this reason, the computation and analysis of higher order cyclic features is generally limited to the time domain. The resulting feature of choice for HOCS signals is based on the temporal cumulant (TC) of the received signal $x(t)$, given as

$$\begin{aligned} C_x(t,\boldsymbol{\tau})_{n,q} &= Cum\{x^{(*)_1}(t+\tau_1)\dots x^{(*)_{n-1}}(t+\tau_{n-1})x^{(*)_n}(t)\} \\ &= \sum_{P_Z=\{P_n\}} (-1)^{Z-1}(Z-1)! \prod_{z=1}^{Z} R_x(t,\boldsymbol{\tau}_z)_{n_z,q_z} \\ &= R_x(t,\boldsymbol{\tau})_{n,q} \; - \sum_{P_Z=\{P_n\}} \left[\prod_{z=1}^{Z} C_x(t,\boldsymbol{\tau}_z)_{n_z,q_z} \right] \end{aligned} \tag{2.41}$$

where $(*)_i$ denotes ones of q possible conjugations, P_Z is a set of Z distinct partitions $\{p_1,\dots,p_Z\}$ of $\{1,2,\dots,n\}$, $\{P_n\}$ is the set of all possible sets P_Z, and n_z and q_z correspond to the number of elements and the number of conjugated terms in the subset p_z, respectively. As shown in (2.41), when computing the TC the effect of lower order moments are removed, leaving the only remaining impact due to the current order. This process isolates the cyclic features present at an order n from those features that are merely composed of products of lower order features. As a result the signal's HOCS features can be investigated independently at each order.

Like the TMF, the TC is also a periodic function for cyclostationary signals with Fourier components given by

$$C_x^{\alpha}(\boldsymbol{\tau})_{n,q} = \lim_{T\to\infty} 1/T \int_{-T/2}^{T/2} C_x(t,\boldsymbol{\tau})_{n,q} e^{-j2\pi\alpha t} dt \tag{2.42}$$

$$C_x(t,\boldsymbol{\tau})_{n,q} = \sum_{\{\alpha\}} C_x^{\alpha}(\boldsymbol{\tau})_{n,q} e^{j2\pi\alpha t} \tag{2.43}$$

where $C_x^{\alpha}(\boldsymbol{\tau})$ is termed the cyclic cumulant (CC) of $x(t)$, and the sum in (2.43) is over the set of all α for which the CC is not identically zero as a function of $\boldsymbol{\tau}$, and $\{\alpha\}$ is again termed the set of cycle frequencies.

As in the case of the CTMF and the SCF, when the CC is evaluated at $\alpha = 0$, it reduces to the traditional stationary cumulant. Thus the stationary cumulant can be viewed as a special case of the CC evaluated at $\alpha = 0$. As a result, when a purely stationary signal is considered, its CC is given by

$$C_x^{\alpha}(\boldsymbol{\tau})_{n,q} = \begin{cases} C_x^0(\boldsymbol{\tau})_{n,q}, & \alpha = 0 \\ 0, & \alpha \neq 0 \end{cases} \tag{2.44}$$

One benefit of using cumulants for the analysis HOCS signals rather than moments is that cumulants can be shown to obey the following semi-linearity property:

$$C_y(t,\boldsymbol{\tau})_{n,q} = \sum_i C_{x_i}(t,\boldsymbol{\tau})_{n,q} \tag{2.45}$$

where

$$y(t) \equiv \sum_i x_i(t) \tag{2.46}$$

Thus the TC of a sum of signals is equal to the sum of the individual TCs of each signal. As a result, unlike higher order CTMFs, the CC preserves the separability of signals similarly to that observed in the SCF and CAF. When analyzing the CC of a signal composed of two or more signals that overlap in time and frequency, but which have higher order features at distinct cycle frequencies, the CC of each signal can be evaluated without interference due to the other signals.

Another benefit of using cumulants over higher order moments is that the higher order CC of gaussian random process is identically zero for all orders greater than two. The higher order CC of a signal received in the presence of noise is therefore identical to the CC of the signal without the noise at all cycle frequencies, regardless of the SNR.

2.6 Multipath Channel Affects

In practice wireless communication signals will be subject to multipath propagation, it is therefore useful to asses the impact these channels will have on the above statistics. Taking this into consideration, the CTMF of a signal $x(t)$ subject to multipath effects is given as [11]

$$R_Y^\alpha(\boldsymbol{\tau})_{n,q} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\prod_{j=1}^{n} h^{(*)_j}(\lambda_j) \right] R_X^\alpha(\boldsymbol{\tau} - \boldsymbol{\lambda})_{n,q} d\boldsymbol{\lambda}, \quad \boldsymbol{\lambda} \equiv [\lambda_1, \cdots, \lambda_n] \quad (2.47)$$

.

$$y(t) = x(t) \otimes h(t) \quad (2.48)$$

where $h(t)$ is the unknown channel response, and $H(f)$ is the Fourier Transform of $h(t)$. In the above equations, the CAF of the original signal is smoothed by the channel response in an n-fold convolution. This effect will in general severely distort the original CAF.

For the case of $n = 2$ and $q = 1$, the received CAF then becomes

$$R_Y^\alpha(\boldsymbol{\tau}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\lambda_1) h^*(\lambda_2) R_X^\alpha(\tau - \lambda_1 + \lambda_2) d\lambda_1 d\lambda_2 \quad (2.49)$$

resulting in a received SCF of

$$S_Y^\alpha(f) = H\left(f + \frac{\alpha}{2}\right) H^*\left(f - \frac{\alpha}{2}\right) S_x^\alpha(f) \tag{2.50}$$

The resulting SCF of the received signal may also be significantly distorted. However, when forming the SOF, by substituting (2.50) into (2.14) it is evident that the channel effects are removed, and the resulting SOF is equal to that of the original undistorted signal [11]. As a result, the SOF of the original signal is completely preserved in the presence of multipath channels, so long as no frequency of the signal of interest is completely nullified by the channel.

The CC of a signal undergoing multipath distortion is given as [11]

$$C_Y^\alpha(\boldsymbol{\tau})_{n,q} = \int_{-\infty}^{\infty} \left[\prod_{j=1}^{n} h^{(*)_j}(\lambda_j)\right] C_X^\alpha(\boldsymbol{\tau} - \boldsymbol{\lambda})_{n,q} d\boldsymbol{\lambda} \tag{2.51}$$

.

Without a normalized statistic for HOCS as the SOF, the CC is expected to experience severe distortion. It can be shown that when a signal is subject to multipath, its higher order cumulants have the effect of appearing more Gaussian like [17]. As a result, as the multipath becomes more severe, the magnitude of higher order CCs of the signal will decrease, and the estimation of the signal's higher order features will be negatively impacted.

2.7 Summary

This chapter presented Likelihood Based (LB) classifiers and Feature Based (FB) classifiers as the two general approaches to modulation recognition in Section 2.2, where each was introduced and key aspects was discussed. The underlying theory of cyclostationarity was introduced in Section 2.3 in both a stochastic process (SP) probabilistic framework as well as a fraction of time (FOT) probabilistic framework, where the benefits of using cyclic statistics were presented. The two basic methods for estimating second order cyclic statistics were introduced in Section 2.4, and the essential features for higher order cyclic statistics were given in Section 2.5. The chapter concluded with a discussion of the effects multipath channels have on the estimation of the above cyclic statistics.

3 Classifier Design

3.1 Introduction

This chapter presents the outline for a multistage modulation recognition system designed to discriminate between AM, BFSK, OFDM, DS-CDMA, 4-ASK, 8-ASK, BPSK, QPSK, 8-PSK, 16-PSK, 16-QAM, and 64-QAM modulation types. The system described is a feature based (FB) modulation recognition system utilizing the cyclic statistics defined in the previous chapter. It is implemented in a hierarchical approach to classify the signals using the smallest amount of required data possible, while simultaneously maximizing the reliability of the system. The proposed modulation classification system is a combination and extension of two current algorithms, modified to for efficiency and reliability.

In Section 3.2, various methods for estimating the spectral correlation function (SCF) of a signal are outlined, and a combination of estimation methods is chosen for implementation. Section 3.3 outlines the method for estimating higher order cyclic cumulants (CC), and a process designed to determine the carrier frequency and symbol rate for digital signals is presented. In Section 3.4 the final classifier structure is presented, and its sequence of operations defined. The chapter concludes by addressing methods to compensate for adverse affects due to multipath channels in Section 3.5.

3.2 Cyclic Spectral Estimation

As outlined in Section 2.4, the two main methods used for cyclic spectral analysis are based on temporal smoothing or frequency smoothing. While each method will produce similar estimates of the SCF, each is better suited for different applications. Temporal smoothing methods can be more computationally efficient for cyclic spectral analysis over the entire bi-frequency plane. Frequency smoothing methods on the other hand are more efficient for producing high-accuracy estimates of the SCF for a relatively small set of known cycle frequencies [18, 19, 20]. The methods required to produce each estimate are given in the following sections.

3.2.1 Temporal Smoothing Methods

As stated in Section 2.4, temporal smoothing methods for estimating the SCF are based on the time smoothed cyclic periodogram, given by (2.38). When computed digitally, the time smoothed cyclic periodogram can be expressed as

$$S_{X_T}^{\alpha_0}(n, f_0)_{\Delta t} = \sum_{r=-N/2}^{N/2-1} X_T(r, f_1) X_T^*(r, f_2) g_{\Delta t}(n-r) \tag{3.52}$$

$$X_T(n, f_k) = \sum_{r=-N'/2}^{N'/2-1} a_T(r) x(n-r) e^{-j2\pi f_k (n-r) T_s} \tag{3.53}$$

where T_s is the sampling rate, $f_k \in 1/T \cdot \{-N'/2, \ldots, N'/2-1\}$, $\alpha_0 \equiv f_1 - f_2$, $f_0 \equiv (f_1 + f_2)/2$, and N' and N are assumed to be even with $N > N'$. The time and frequency resolutions of the above estimate are given by $\Delta t = NT_s$ and $\Delta f = 1/T = 1/(N'T_s)$, respectively, resulting in a time frequency resolution product of $\Delta t \Delta f = N/N'$.

The above two equations can be computed efficiently on a digital device using a fast

Fourier transform (FFT) followed by a digital correlator. A schematic of the process used to estimate the SCF is shown in Figure 3.1.

To achieve statistical reliability, $\Delta t \Delta f \gg 1$ or equivalently $N \gg N'$, implying that $X_T(n, f)$ must be computed and correlated over many data segments each of length N' to produce the SCF image. However, since $\Delta\alpha \approx 1/\Delta t = 1/T$, the resulting cycle frequency resolution decreases below that provided by the FFTs used to compute (3.53). As a result, cyclic features of interest may occur at cycle frequencies other than those estimated in Fig 3.1, and will thus not be detected by the algorithm. To demonstrated this, the coverage diagram of the estimates produced by temporal smoothing methods is shown in Figure 3.2 with the resulting time and frequency resolution of each estimate given in Figure 3.3 [15, 18].

One solution to the above problem is to zero pad the signal processed by the FFT in (3.53) from a length of N' out to a length equal to the total number of samples N. While this will ensure all possible cycle frequencies are detected, it will require an excessive amount of additional computational burden for the classification system.

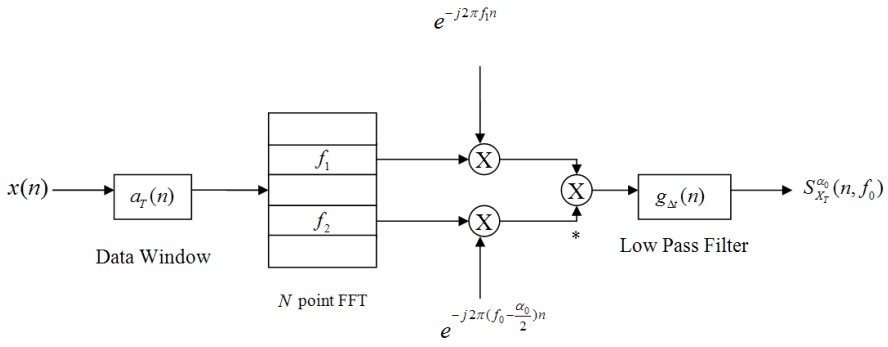

Figure 3.1: Time Smoothed Cyclic Periodogram Implementation

Another solution to the above problem is achieved by employing a variation of (3.52) known as the Strip Spectrum Correlation Algorithm (SSCA). The SSCA estimate of the

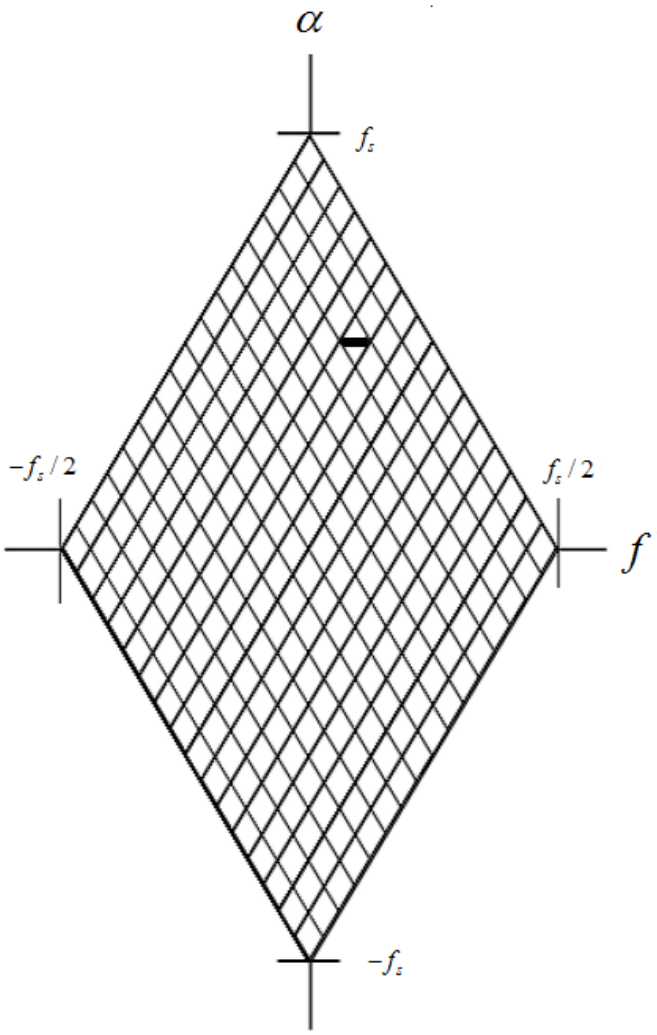

Figure 3.2: Coverage Diagram for the Time Smoothed Cyclic Periodogram in the Bi-frequency Plane. Estimates are only produced for features contained in a narrow band in the center of each diamond

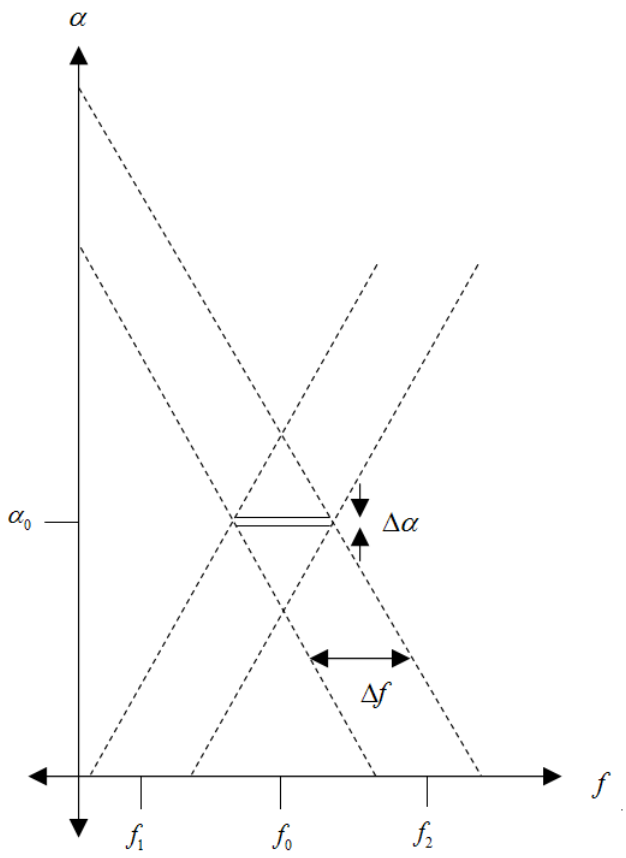

Figure 3.3: Time Smoothed Cyclic Periodogram SCF Resolution

SCF is given by [16]

$$S_{X_T}^{\alpha_0}(n, f_0)_{\Delta t} = \sum_{r=-N/2}^{N/2-1} X_T(r, f_k) x^*(r) g_{\Delta t}(n-r) e^{-j2\pi rq/N} \tag{3.54}$$

where $\alpha_0 = f_k + q\Delta\alpha$ and $f_0 = f_k/2 - q\Delta\alpha/2$. Equation (3.54) is the same form as (3.52) with the second FFT estimate $X_T^*(r, f)$ there replaced with $x^*(r)e^{-j2\pi rq/N}$.

The implementation of the SSCA is shown in Figure 3.4 below [18].

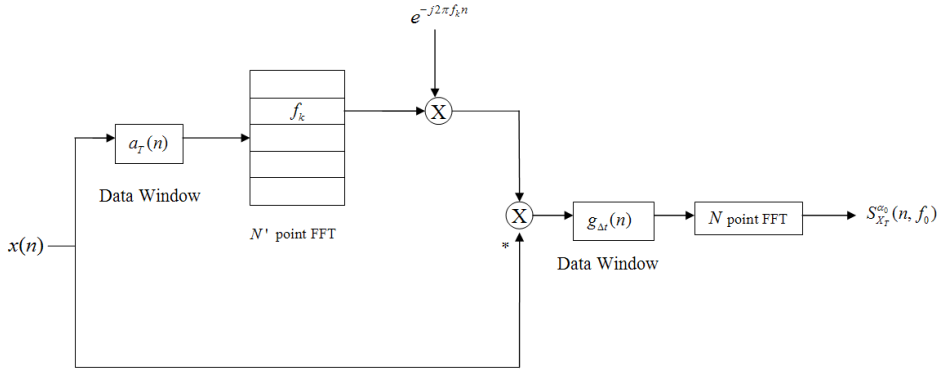

Figure 3.4: Strip Spectrum Correlation Algorithm Implementation

It can be shown that the SSCA produces strips of point estimates which lie along the line $\alpha = 2f_k - 2f_0$ in the bi-frequency plane with temporal resolution $\Delta t = N/f_s$, cycle frequency resolution $\Delta\alpha \approx 1/\Delta t = f_s/N$, and frequency resolution $\Delta f = 1/T = f_s/N'$. For statistical reliability, $\Delta t \Delta f \gg 1$ again implies $N \gg N'$. A representative coverage diagram produced by each strip of the SSCA is shown in Figure 3.5 and the resulting time and frequency resolution of an estimate produced by a strip is shown in Figure 3.6 [15, 18].

The SSCA implementation has been shown to produce highly efficient estimates of the SCF over the entire bi-frequency plane. However, the channelization of the signal performed by the initial FFT provides poor frequency resolution for the estimates. As a result, the SSCA is highly efficient at estimating the SCF over the entire region of support, but does not produce estimates with sufficiently fine precision in the frequency domain for

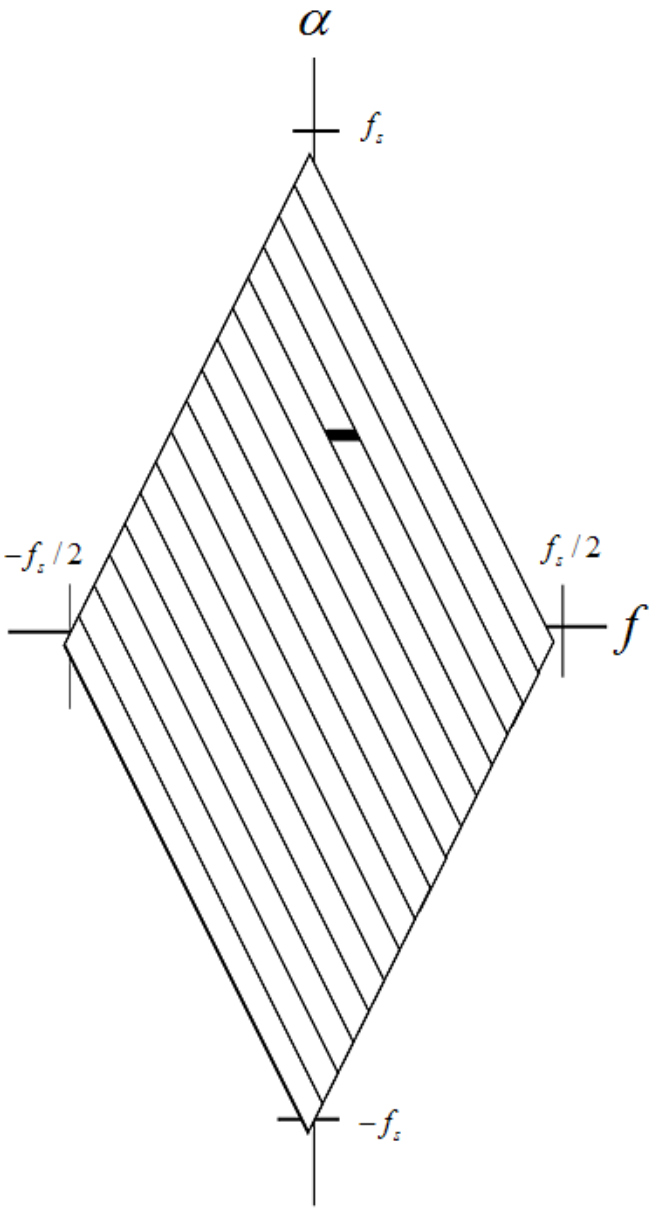

Figure 3.5: Coverage Diagram of the Strip Spectrum Correlation Algorithm in the Bi-frequency Plane with $N' = 16$

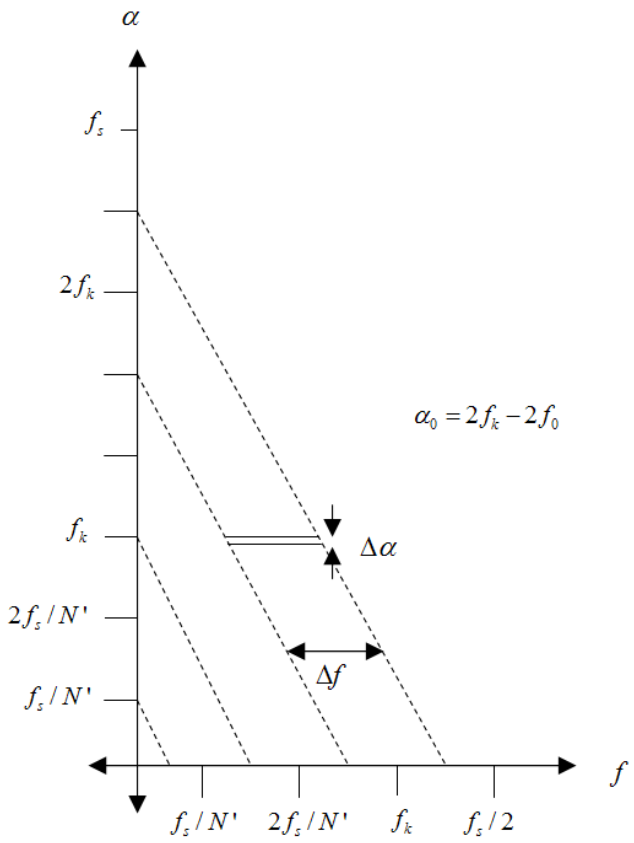

Figure 3.6: SCF Estimate Resolution of the Strip Spectrum Correlation Algorithm. Estimates are produced along diagonal strips in the bi-frequency plane

most applications. The SSCA can be used to estimate the cycle frequencies of interest, but another method is needed to obtain high accuracy estimates of the SCF. As will be discussed in the following section, frequency smoothing methods have been shown to produce estimates of the SCF with much finer frequency resolution than that provided by the SSCA, while still remaining computationally efficient when computed over a limited set of cycle frequencies [18].

3.2.2 Frequency Smoothing Methods

Frequency smoothing methods for estimating the SCF are based on the frequency smoothed cyclic periodogram, given by (2.39). When computed digitally, the frequency smoothed cyclic periodogram can be expressed as

$$S_{X_T}^{\alpha_0}(n, f_0)_{\Delta f} = \sum_{r=-N'/2}^{N'/2-1} X_T(n, f_1 + r/T) X_T^*(n, f_2 + r/T) a_{\Delta f}(r) \tag{3.55}$$

$$X_T(n, f_k) = \sum_{r=-N/2}^{N/2-1} g_T(r) x(n-r) e^{-j2\pi f_k (n-r) T_s} \tag{3.56}$$

where T_s is the sampling rate, $f_k \in 1/T \cdot \{-N/2, \ldots, N/2-1\}$, $\alpha_0 \equiv f_1 - f_2$, $f_0 \equiv (f_1 + f_2)/2$, and N' and N are assumed to be even with $N > N'$. The time and frequency resolutions of the above estimate are given by $\Delta t = T = NT_s$ and $\Delta f = N'/T = N'/(NT_s)$, respectively, resulting in a time frequency resolution product of $\Delta t \Delta f = N'$.

The above two equations can be computed efficiently by performing an FFT over the entire input signal, followed by a cross spectral correlation of the resulting estimated frequency components. A schematic of the process used to estimate the SCF is shown in Figure 3.7.

To achieve statistical reliability, $\Delta t \Delta f \gg 1$ implies $N' \gg 1$. The amount of spectral smoothing used becomes a design parameter that is subject to a trade-off between statistical reliability and spectral resolution. To achieve a statistically reliable estimate of the SCF, a large amount of spectral smoothing is desired. However, in order to resolve fine spectral features in the observed signal, the amount of spectral smoothing should be minimized. Therefore, the amount of spectral smoothing implemented should be chosen to achieve the largest amount of smoothing possible while still maintaining the ability to resolve the critical features of the signal being processed.

For $\Delta t \Delta f \gg 1$ the coverage of the frequency smoothed cyclic periodogram has characteristics identical to that of the time smoothed cyclic periodogram given by Figure 3.3. However, now the SCF can be estimated at any value of $\alpha_0 = \{0, \pm f_s/N, \pm 2f_s/N, \ldots, \pm f_s(1 - 1/N)\}$, which covers the entire bi-frequency plane, as can be seen when considering that the cycle frequency resolution of each estimate is $\Delta\alpha \approx 1/\Delta t = 1/(NT_s) = f_s/N$. To estimate the SCF over the entire region of support, the spectral correlation must be completed over all $2N - 1$ of these values of α_0. This leads to a computationally intensive estimation procedure. However, if the cycle frequencies of interest have been identified, a reliable estimate of the SCF can be obtained via frequency smoothing at a significantly reduced level of complexity.

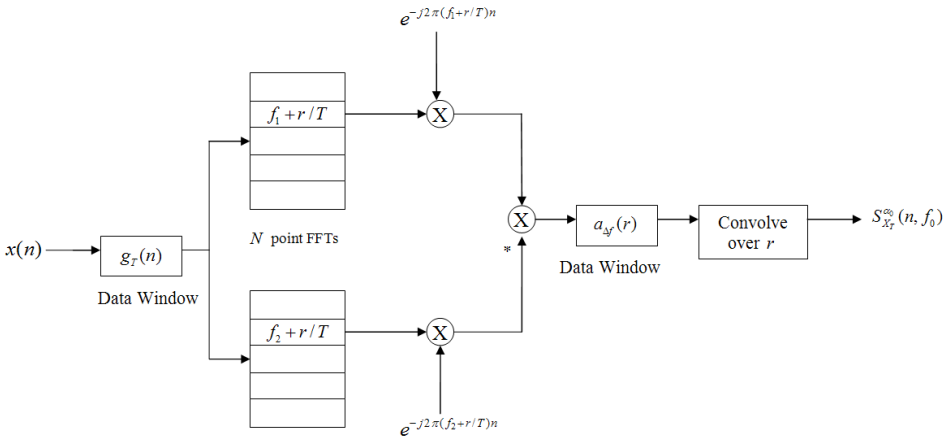

Figure 3.7: Freq Smoothed Cyclic Periodogram Implementation

3.2.3 Spectral Estimation Process

In [3], the SOF of received signals was shown to produce a highly reliable feature for modulation classification of received signals. However, the method described is based on the temporal smoothing method of Fig 3.1. As mentioned above, this method is highly inefficient, and is only able to locate features present at a limited set of cycle frequencies. If a signal has a feature located at a cycle frequency other than the limited set estimated, the feature will pass unnoticed.

In view of the results of the previous two sections, the spectral estimation procedure is implemented in a two stage process. In the first stage, the SSCA is implemented in order to estimate the cycle frequencies that contain features for the signal being analyzed. This is done by forming an estimate of the SOF from the SCF estimate, and choosing the N cycle frequencies with the largest features, where N is a user selectable value. After this set of cycle frequencies has been identified, the Frequency Smoothing SCF estimation method is used to produce a high accuracy estimate of the reduced set of cycle frequencies. This estimate of the SCF is then used to compute a high accuracy estimate of the SOF. For each stage, the estimate of the SOF is given by

$$C_X^{\alpha_0}(f_0) = \frac{S_{X_T}^{\alpha_0}(n, f_0)_\Delta}{\left[\left(S_{X_T}^0(n, f_1)_\Delta\right)^* \left(S_{X_T}^0(n, f_2)_\Delta\right)\right]^{1/2}} \tag{3.57}$$

where $S_{X_T}^{\alpha_0}(n, f_0)_\Delta$ is the SCF estimate produced in each stage, $f_0 = (f_1 + f_2)/2$, and $\alpha_0 = f_1 - f_2$ is the set of cycle frequencies for which the SCF is estimated. To account for the unknown phase of the SOF, the magnitude of $C_X^{\alpha_0}(f_0)$ is computed and used for classification. Example SOF estimates for several common modulation schemes are shown in Figures 3.8 through 3.12.

Once the SOF images are obtained, the image must be processed by the classifier and a classification decision made. However, the SOF feature produced is a three-dimensional

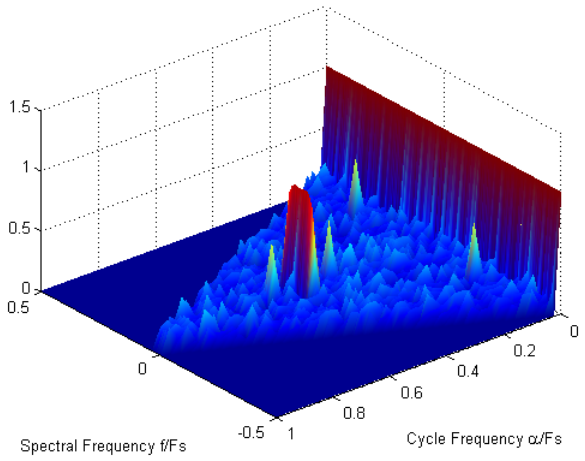

Figure 3.8: SOF of a BPSK signal in an AWGN Channel at 5 dB SNR, with $1/Ts = Fs/10$, $Fc = .25Fs$, #samples = 4092

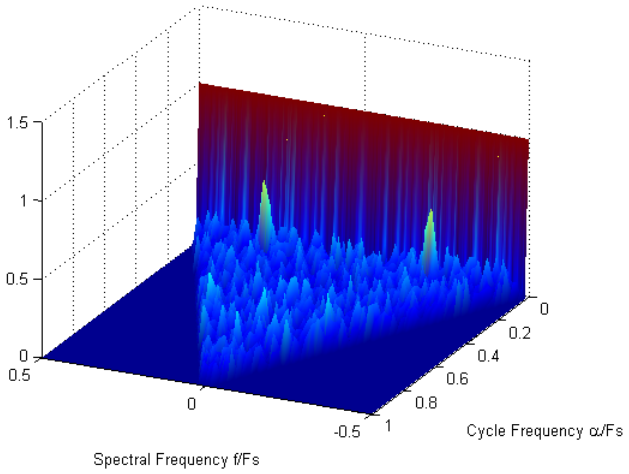

Figure 3.9: SOF of a QPSK signal in an AWGN Channel at 5 dB SNR, with $1/Ts = Fs/10$, $Fc = .25Fs$, #samples = 4092

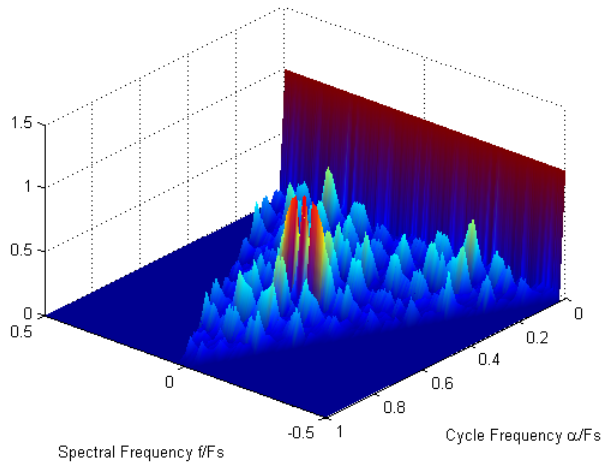

Figure 3.10: SOF of a BFSK signal in an AWGN Channel at 5 dB SNR, with $1/Ts = Fs/10$, $Fc = .25Fs$, #samples = 4092

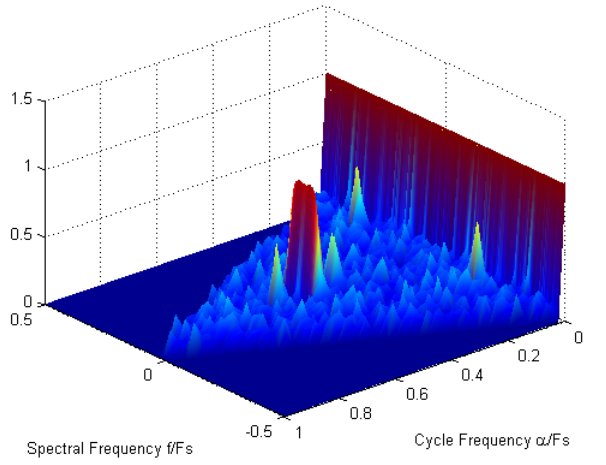

Figure 3.11: SOF of a 4-ASK signal in an AWGN Channel at 5 dB SNR, with $1/Ts = Fs/10$, $Fc = .25Fs$, #samples = 4092

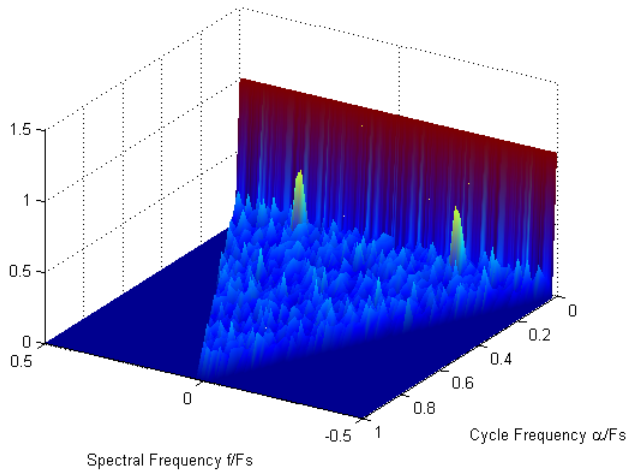

Figure 3.12: SOF of a 8-PSK signal in an AWGN Channel at 5 dB SNR, with $1/Ts = Fs/10$, $Fc = .25Fs$, #samples = 4092

image, and still presents an unreasonable amount of data for a classifier to operate on in real-time. Therefore, the data set must be further reduced to provide a more computationally manageable feature for processing. In [3] the authors proposed using merely the cycle frequency profile of the SOF as a feature. This reduced feature was able to classify modulation schemes with a moderately high degree of reliability. However, with a minimal increase in computational complexity, both the frequency profile as well as the cycle frequency profile can be used, creating a pseudo three dimensional image of the SOF which performs at a significantly higher degree of reliability for classification. The resulting features used for classification are then defined as the cycle frequency profile:

$$\overrightarrow{\alpha} = max_f[C_X^\alpha] \tag{3.58}$$

and the spectral frequency profile

$$\overrightarrow{f} = max_\alpha[C_X^\alpha] \tag{3.59}$$

These two feature vectors can then be processed to classify the modulation scheme of

an unknown signal.

The resulting SOF estimation processes is shown in Figure 3.13.

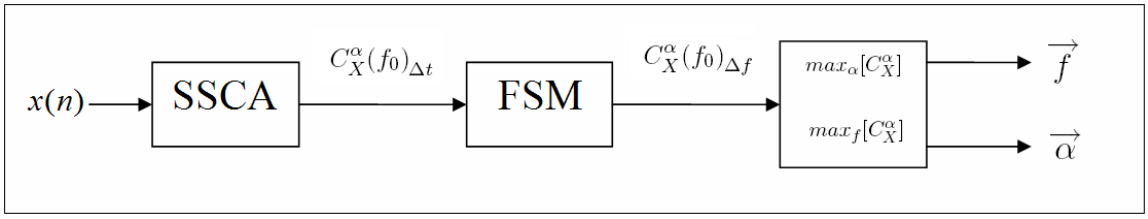

Figure 3.13: Multistage SOF Estimation Process

3.3 Higher Order Cyclic Features

The SOF produces highly distinct images for different classes modulation schemes. However, there are some modulation schemes that produce identical SOF images. Therefore, while the SOF is able to reliably classify many of the more fundamental classes of signals, it will not be able to distinguish between certain sets of digital schemes, or between different orders of a single modulation scheme. In particular, BPSK signals and ASK signals produce identical SOF images. Similarly, the SOF images of QAM signals are identical to those of higher order PSK signals [10]. As an example, compare the estimated SOF of the BPSK signal in Figure 3.8 with that of the 4-ASK signal shown in Figure 3.11, and of the QPSK signal in Figure 3.9 with that of the 8-PSK signal shown in Figure 3.12.

3.3.1 Cyclic Cumulant Feature Development

To discriminate between signals for which the SOF is unable higher order cyclic statistics (HOCS) may be employed. As outlined in Section 2.5, the cyclic cumulant (CC) is a higher order statistic capable of discerning cyclic features independently for various orders. Assuming that any analog or FSK signals have been properly identified by their SOF, the

remaining digital signals received by the classifier can be modeled as

$$x(t) = \Re\{\tilde{x}(t)\} \tag{3.60}$$

$$\tilde{x}(t) = e^{j2\pi f_c t} e^{j\phi} \sum_{-\infty}^{\infty} s_k p(t - kT_D - t_0) + \tilde{n}(t) \tag{3.61}$$

where $x(t)$ is the received signal, $\tilde{x}(t)$ is the analytic signal of $x(t)$, f_c is the carrier frequency, ϕ is the carrier phase offset, $p(t)$ is the pulse shape, T_D is the symbol period, t_0 is the signal time offset, $\tilde{n}(t)$ is additive Gaussian noise, and s_k is the kth digital symbol transmitted at time $t \in (kT_D - T_D/2, kT_D + T_D/2)$. Here, the symbols s_k are assumed to be zero-mean, identically distributed random variables.

By substituting (3.61) into (2.41) and (2.42), it can be shown that the resulting value of the CC is given by [4]:

$$\begin{aligned} C_x^{\alpha}(\boldsymbol{\tau})_{n,q} = \; & C_{s,n,q} T_D^{-1} e^{-j2\pi\beta t_0} e^{j(n-2q)\phi} e^{j2\pi f_c \sum_{u=1}^{n-1} (-)_u \tau_u} \\ & \times \int_{-\infty}^{\infty} p^{(*)_n}(t) \prod_{u=1}^{u=n-1} p^{(*)_u}(t + \tau_u) e^{-j2\pi\beta t} dt \end{aligned} \tag{3.62}$$

$$\alpha = \beta + (n - 2q) f_c, \quad \beta = k/T_D$$

where α is the cycle frequency being estimated, β is the cycle frequency of the equivalent baseband signal contributing to the estimate, the possible minus sign $(-)_u$ comes from one of the q conjugations, and $C_{s,n,q}$ is the nth-order/q-conjugate cumulant of the stationary discrete data sequence.

When computed digitally, (3.62) above can be expressed as

$$C_x^{\alpha_0}(\boldsymbol{\tau}_m)_{n,q} = \quad C_{s,n,q}\,\rho e^{-j2\pi\beta t_0} e^{j(n-2q)\phi} e^{j2\pi f_c \sum_{u=1}^{n-1}(-)_u \tau_{m,u}}$$

$$\times \sum_{r=-\infty}^{\infty} p^{(*)_n}(r) \prod_{u=1}^{u=n-1} p^{(*)_u}(r+\tau_{m,u}) e^{-j2\pi\beta r T_s} \tag{3.63}$$

$$\alpha_0 = \beta + (n-2q)f_c, \;\; \beta = k/T_D, \;\; \rho \equiv T_s/T_D$$

where $\boldsymbol{\tau}_m$ is now discrete valued.

Thus, the resulting value of the CC of the received signal is directly proportional to $C_{s,n,q}$. The value of $C_{s,n,q}$ is well-known for common modulation schemes, and is given in Table 3.1 below [4]. By leveraging knowledge of the distinct values of $C_{s,n,q}$ for different modulation schemes, the correct scheme of an unknown signal can be accurately determined from the CC of the signal.

Since the carrier frequency, phase, and signal time offset are all unknown a priori, their effect much be removed from the above statistic before it can be processed for classification. Since these terms only affect the phase of the above statistic, the magnitude of (3.62) can be taken to remove their effect. The resulting feature is given as

$$\Gamma_x(\alpha_0, \boldsymbol{\tau}_m)_{n,q} = \left| C_{s,n,q}\,\rho \times \sum_{r=-\infty}^{\infty} p^{(*)_n}(r) \prod_{u=1}^{u=n-1} p^{(*)_u}(r+\tau_{m,u}) e^{-j2\pi\beta r T_s} \right| \tag{3.64}$$

$$\alpha_0 = \beta + (n-2q)f_c, \;\; \beta = k/T_D, \;\; \rho \equiv T_s/T_D$$

3.3.2 Cyclic Cumulant Estimation Process

Assuming a raised cosine pulse shape, the maximum of the resulting function $\Gamma_x(\alpha_0, \boldsymbol{\tau}_m)_{n,q}$ has been shown to occur at $\boldsymbol{\tau}_m = \vec{\mathbf{0}}_{\boldsymbol{n}}$, where $\vec{\mathbf{0}}_{\boldsymbol{n}}$ is an n-dimensional zero vector. Further-

$C_{s,n,q}$	4-ASK	8-ASK	BPSK	Q-PSK
$C_{s,2,0}$	1	1	1	0
$C_{s,2,1}$	1	1	1	1
$C_{s,4,0}$	-1.36	-1.24	-2	1
$C_{s,4,1}$	-1.36	-1.24	-2	0
$C_{s,4,2}$	-1.36	-1.24	-2	-1
$C_{s,6,0}$	8.32	7.19	16	0
$C_{s,6,1}$	8.32	7.19	16	-4
$C_{s,6,2}$	8.32	7.19	16	0
$C_{s,6,3}$	8.32	7.19	16	4
$C_{s,8,0}$	-111.85	-92.02	-272	-34
$C_{s,8,1}$	-111.85	-92.02	-272	0
$C_{s,8,2}$	-111.85	-92.02	-272	34
$C_{s,8,3}$	-111.85	-92.02	-272	0
$C_{s,8,4}$	-111.85	-92.02	-272	-34
$C_{s,n,q}$	8-PSK	16-PSK	16-QAM	64-QAM
$C_{s,2,0}$	0	0	0	0
$C_{s,2,1}$	1	1	1	1
$C_{s,4,0}$	0	0	-0.68	-0.62
$C_{s,4,1}$	0	0	0	0
$C_{s,4,2}$	-1	-1	-0.68	-0.62
$C_{s,6,0}$	0	0	0	0
$C_{s,6,1}$	0	0	2.08	1.80
$C_{s,6,2}$	0	0	0	0
$C_{s,6,3}$	4	4	2.08	1.80
$C_{s,8,0}$	1	0	-13.98	-11.50
$C_{s,8,1}$	0	0	0	0
$C_{s,8,2}$	0	0	-13.98	-11.50
$C_{s,8,3}$	0	0	0	0
$C_{s,8,4}$	-33	-33	-13.98	-11.50

Table 3.1: Theoretical Stationary Cumulants[4]

more, at $\boldsymbol{\tau}_m = \vec{\mathbf{0}}_{\boldsymbol{n}}$, the function decreases with increasing k, indicating an optimal value of k to be 1 [19]. $\Gamma_x(\alpha, \boldsymbol{\tau})_{n,q}$ should then be evaluated at $\alpha_0 = 1/T_D + (n-2q)f_c$.

The CC of unknown signals have been used to reliably classify the modulation scheme of several signals in [4]. However, both the carrier frequency f_c and symbol rate $1/T_D$ were assumed to be precisely known a priori. Additionally the carrier frequency is assumed to have been perfectly removed prior to processing. In general, these values are not known a priori by a classifier, and a robust classification system should be able to perform without prior knowledge of these values. However, the desired value of α_0 used to evaluate the CC depends on both f_c and $1/T_D$. Therefore, the estimation of these two values will need to be performed during the classification process.

The value of α_0 can be derived by noting that cyclic features will only occur at intervals of $1/T_D$. For a raised cosine pulse, the magnitude of $\Gamma_x(\alpha_0, \boldsymbol{\tau}_m)_{n,q}$ obtains its largest value at $k = 0$, corresponding to a cycle frequency of $\alpha_0 = (n-2q)f_c$. The next largest peak occurs at $k = 1$, which is the desired cycle frequency. To estimate the desired value of α_0, all that is needed is to search for the cycle frequency corresponding to the largest cyclic feature, and evaluate the CC at an offset of $1/T_D$ from this location.

It has been shown that for a given number of signal samples available for classification, CCs produce less variances in their lower order estimates than for higher order estimates [19]. Therefore the lowest order CC capable of estimating $1/T_D$ should be used to achieve a more reliable estimate. The second-order/one-conjugate CC is therefore selected to estimate $1/T_D$, as all of the modulation schemes being considered will contain a feature at this cycle frequency. Using the value of $\alpha_0 = 1/T_D$ computed from the second-order CC, paired with the estimate of $\alpha_0 = (n-2q)f_c$ obtained for each CC, the computation of the value of $\alpha_0 = 1/T_D + (n-2q)f_c$ is straightforward.

The resulting values of the different order/conjugate pairs of the CCs can now be used to classify the signal further to discriminate between modulation schemes for which the

SOF was unable to differentiate. By referring to Table 3.1, the specific modulation type as well as its order can be determined from the expected values of $C_{s,n,q}$.

The classifier presented in [4] produced classification decision based on eighth-order CCs of the received signal. Given that the variance of the CC estimates increase with increasing order [19], a more reliable classification can be made by using the lower order CCs in the classification process. By implementing a hierarchical scheme, lower order CCs can perform an initial classification, followed by progressively higher order CCs to further refine the classification decision. In this way a more reliable estimate can be obtained than by using eighth-order CCs alone. Furthermore, in poor channel conditions, a hierarchical scheme is expected to be better able to distinguish between modulation families than a scheme based purely on a single higher order CC, due to the lower variance in the CCs.

3.3.3 Identification of OFDM Signals

As discussed above, the SOF cannot be leveraged to distinguish between signals with identical second order cyclostationary statistics (SOCS). The subcarriers in an OFDM system can be modeled as independent (though still phase coherent) modulated signals. As a result, the SCF of an OFDM signal will be the sum of the SCF of the individual subcarrier signals. However, because the spectrum of the individual subcarriers overlap, the value of the SOF will be reduced compared to if the signals did not overlap in frequency. In the presence of low SNR and a limited number of available samples for classification, the difference between the estimated SOF of OFDM signals and that of single carrier QAM and MPSK signals ($M > 2$) becomes negligible. Therefore, assuming independent data is transmitted on each subcarrier, OFDM signals cannot be reliably classified based on their estimated SOF.

One method to distinguish OFDM signals from the single carrier signals in question can be obtained by considering the fact that OFDM signals are composed of multiple time

varying signals. Assuming that the OFDM signal is composed of a large number of subcarriers modulated with independent data, the Central Limit Theorem states that the OFDM signal can be approximated as a Gaussian random signal [20]. Through the use of a simple normality test, the OFDM signals can therefore be reliably identified. Since Gaussian signals do not exhibit features for cumulants other than for their 2nd-order/1-conjugate CC, the CC features derived above to distinguish between the HOCS signals can also be used to classify an OFDM signal.

3.4 Classifier Design

The proposed classifier is designed to classify AM, BFSK, OFDM, DS-CDMA, 4-ASK, 8-ASK, BPSK, QPSK, 8-PSK, 16-PSK, 16-QAM, and 64-QAM modulation types. It is designed in a hierarchical approach to classify the signals using the smallest amount of required data possible, while simultaneously maximizing the reliability of the system. At each stage in the system, the signal's modulation scheme is either classified or grouped with similar schemes narrowed down into a smaller subset. The system is designed to require no knowledge of the received signal's carrier frequency, phase shift, or symbol rate, and only assumes the signal's presence has been identified, and that it is located within the bandwidth of interest.

The first stage of the classifier computes the SOF of the signal, and compresses the data into the feature vector composed of the concatenation of $\overrightarrow{\alpha}$ and $\overrightarrow{f}$. A neural network is designed to classify the SOF feature vector. Neural networks were chosen due to their relative ease of setup and use, as well as their ability to generalize to any carrier frequency or symbol rate. The system consists of five independent linear feed forward neural networks, each trained to classify a signal as either AM, BFSK, DS-CDMA, a linear modulation scheme with a real-valued constellation (BPSK, 4-ASK, 8-ASK), or a linear modulation scheme with a complex-valued constellation (OFDM, 8-PSK, 16-PSK, 16-QAM,

64-QAM). Each network has four neurons in its hidden layer and one neuron in its output layer, each layer with a hyperbolic tangent sigmoid transfer function. The inputs to each network are the concatenated profile vectors. A system diagram for this first stage is given in Figure 3.14.

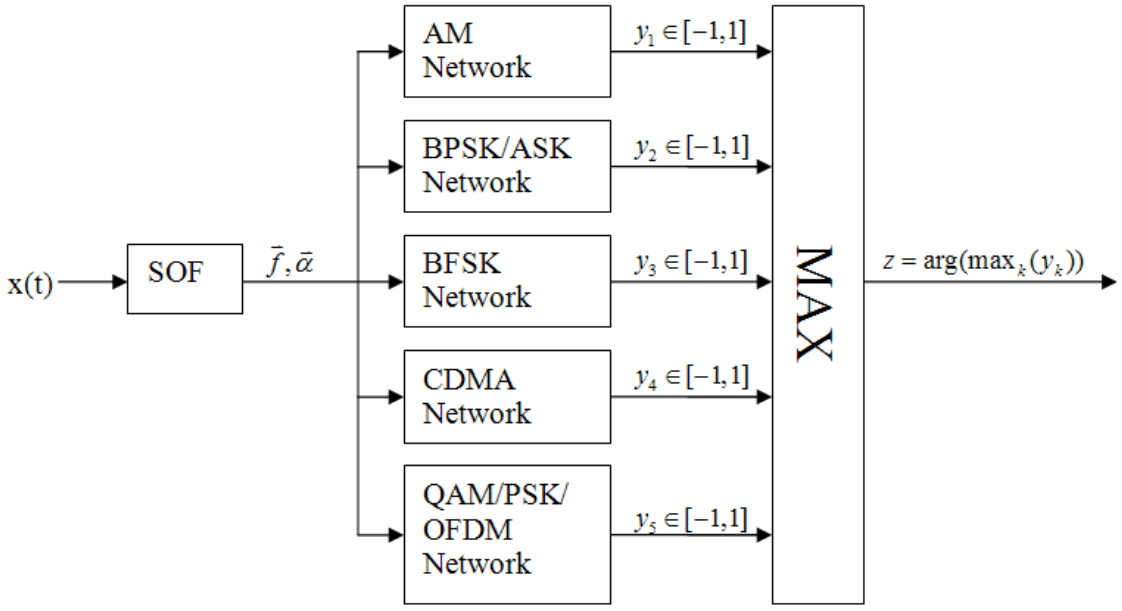

Figure 3.14: SOF-Based Classification System Diagram

The BPSK and ASK signals demonstrate identical SOF images, and are not distinguishable based on that metric alone. Similarly, the PSK and QAM signals have identical spectral components. As mentioned in the previous section, the OFDM signal is composed of potentially independently varying signals on each subchannel, which may or may not demonstrate SOCS. However, due to the overlapping nature of the subchannels in an OFDM system, the resulting SOF is decreased, resulting in a SOF image that resembles those of QAM and PSK signals. Additionally, the DS-CDMA scheme can be thought to look like a BPSK signal. However, due to the underlying periodicities incurred by both its symbol rate as well as its spreading code, it produces features not found in BPSK or QPSK signals. Thus it can be reliably classified by its SOF image without knowledge of

its spreading code.

After the initial classification stage described above, the HOCS based processing is also implemented in a hierarchical approach to maximize the ability to accurately determine the subclass of a signal before further narrowing the list of candidate modulations. This is a critical step since the variance of the CC estimates increase with increasing order [19]. Therefore, an attempt is made to classify a signal using the lowest order CC possible before proceeding to higher order CCs.

In each stage of the HOCS based classification, the feature vector used for classification is composed of the appropriate CCs estimated from the received signal:

$$\widehat{\Psi} = [\Gamma_x(1/T, \vec{\mathbf{0}}_{\boldsymbol{n}})_{n,q_1}, \ldots, \Gamma_x(1/T, \vec{\mathbf{0}}_{\boldsymbol{n}})_{n,q_n}] \tag{3.65}$$

where n refers to the appropriate order for the current stage. This vector is then compared to the expected vector obtained for each modulation type, defined similarly as:

$$\Psi^{(i)} = [\Gamma^{(i)}(1/T, \vec{\mathbf{0}}_{\boldsymbol{n}})_{n,q_1}, \ldots, \Gamma^{(i)}(1/T, \vec{\mathbf{0}}_{\boldsymbol{n}})_{n,q_n}] \tag{3.66}$$

where i corresponds to one of the M possible modulation schemes being considered by the current stage. The class corresponding to the feature vector with the minimum Euclidean distance from the estimated vector is selected. The processing is then handed off to the next stage until the final modulation scheme as been determined.

The network diagram of the system is shown in Figure 3.15. If the SOF network determined the signal to have a real-valued modulation scheme (BPSK, 4-ASK, 8-ASK), then it is handed off to the final classification stage using eighth order CCs. Otherwise, the fourth order CCs are used to classify the signal as being an OFDM signal or as having either a circular constellation (8-PSK, 16-PSK) or a square constellation (QPSK, 16-QAM,

62-QAM). For each signal class, the final stage of the classifier forms the feature vector $\widehat{\Psi}$ from the five eighth-order CCs of the received signal, except for OFDM signals which were already identified using fourth-order CCs.

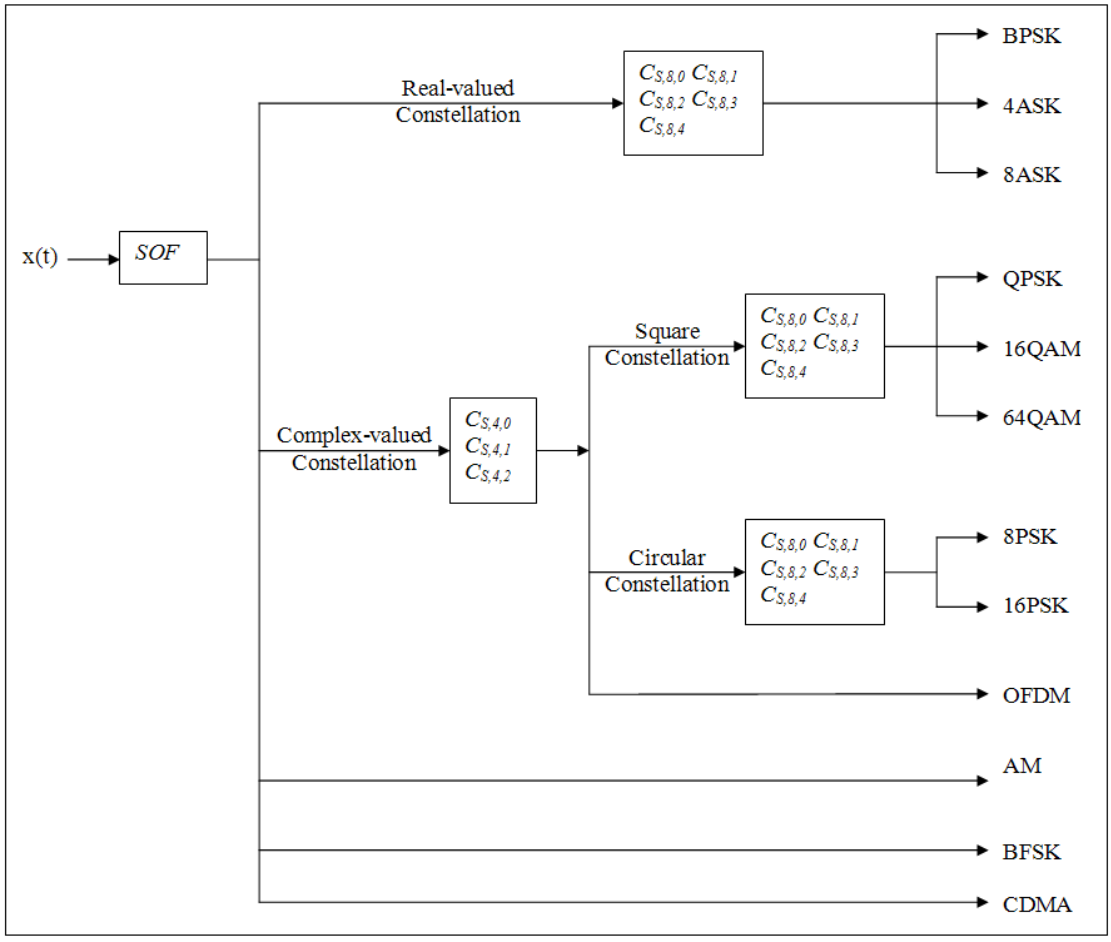

Figure 3.15: Modulation Classification System Diagram

3.5 Multipath Channel Compensation

In the presence of multipath fading channels, the received signal can be severely distorted. Several methods exist to exploit spacial diversity through the use of multiple receiver antennas. By assuming that the channel fades independently on each antenna, the signal received

on each can be combined in various ways to improve performance. The general equation for the received analytic signal undergoing multipath propagation is given by:

$$\tilde{y}(t) = \sum_{p=1}^{P} \kappa_p e^{j\theta_p} \tilde{x}(t - t_p) + \tilde{n}(t) \tag{3.67}$$

where $\kappa_p e^{\theta_p}$ is the channel response on path p, t_p is the delay of the pth path, and P is the total number of paths received by the classifier.

This can be separated into two general situations. In the first situation, the channel is varying sufficiently slowly so that it can be assumed to be static over the block of data be analyzed

If the signal is assumed to only be experiencing flat fading, the simplest combining method is to employ a selection combiner (SC). In [4], the effectiveness of an SC-based system was evaluated to combat the effects of flat fading for modulation recognition. By estimating the received power on each antenna, the signal on the antenna with the highest observed power can be selected for classification, while the others are discarded. When assuming that the noise on each antenna have identical powers, this choice will correspond to the signal with the largest SNR, which leads to an extremely simple implementation.

However, in the case of flat fading, a maximum ratio combiner (MRC) can also be implemented. In this case, the signal received from each antenna is weighted by its SNR before being summed with the signals from the other antennas. In practice, the value of the SNR can be estimated simply by using one of several methods [21] [22] [23]. However, for the signals to combine coherently, the unknown phase on each channel must be compensated for before adding them together. This can be performed by computing the correlation between signals from two channels, given by:

$$E\{\tilde{y}_1(t)\tilde{y}_2^*(t)\}$$

$$= E\{(\kappa_1 e^{j\theta_1}\tilde{x}(t) + \tilde{n}_1(t))(\kappa_2^* e^{-j\theta_2}\tilde{x}^*(t) + \tilde{n}_2^*(t))\} \tag{3.68}$$

$$= \sigma_{\tilde{x}}^2 \kappa_1 \kappa_2 e^{j(\theta_1 - \theta_2)}$$

where $\sigma_{\tilde{x}}^2$ is the power of the signal to be classified. From here, the relative phase difference is given as the phase of the resulting statistic:

$$\Delta\widehat{\theta} = \angle(\sigma_{\tilde{x}}^2 \kappa_1 \kappa_2 e^{j(\theta_1 - \theta 2)}) \tag{3.69}$$

The signal $\tilde{y}_2(t)$ can then be multiplied by $e^{j\Delta\widehat{\theta}}$ to align its phase with the phase of the first channel. This procedure can be repeated as necessary depending on the number of antennas employed.

An additional method proposed to compensate for channel corruption in the SOF computation is through a variant of the MRC. While the SOF was derived to be highly insensitive to channel distortion in Section 2.6, the SOF image obtained when in a deep fade can be significantly distorted by the additive noise components present, which will be amplified when forming the SOF from the SCF. The MRC variant described here then, attempts to compensate for this effect by combining weighted estimates of the SOF from each receiver. For this method, the SOF is computed independently for the signal received on each antenna. After the feature vectors $\overrightarrow{\alpha}$ and $\overrightarrow{f}$ are formed, they are each weighted by the SNR estimated on their respective antennas. Then each is summed, and the procedure follows as before. It is worth noting that this method can be utilized in any fading channel, without the necessity for the assumption of a flat fading channel.

The second general situation exists when the channel is not varying slow enough to be approximated static throughout the signal's evaluation. Since each of the classification

methods above attempts to estimate expected values of joint moments, they are quickly corrupted by a rapidly fading channel. The HOCS features are particularly sensitive since they require a greater amount of samples to converge, during which time the channel can vary drastically. The first stage SOF-based classifier is less sensitive to channel variations, thus providing greater incentive for its use as the first stage in the system.

3.6 Summary

This chapter described the design of a multistage modulation recognition system intended to discriminate between AM, BFSK, OFDM, DS-CDMA, 4-ASK, 8-ASK, BPSK, QPSK, 8-PSK, 16-PSK, 16-QAM, and 64-QAM modulation schemes. In Section 3.2, various methods to implement the estimators of the spectral correlation function (SCF) were presented, and a combination of two methods were chosen for implementation. The process used to estimate higher order cyclic cumulants (CC) was described in Section 3.3. The final classifier design is presented in Section 3.4, where its implementation in a hierarchical approach was described. The chapter concluded by addressing methods to compensate for adverse affects due to multipath channels in Section 3.5.

4 Analysis and Results

4.1 Introduction

This chapter describes the simulation and results obtained for the performance of the designed modulation classification system. The system is evaluated in a variety of channel conditions, without a priori knowledge of critical signal parameters. The performance of the classifier is compared to that of the baseline classifiers given in [3] and [4].

In Section 4.2 the simulation setup is described and the channel models through which the simluated signals are transmitted as given. Simulation results are given in Section 4.3.

4.2 Simulation Setup

Simulations were run with AM, BFSK, OFDM, DS-CDMA, 4-ASK, 8-ASK, BPSK, QPSK, 8-PSK, 16-PSK, 16-QAM, and 64-QAM modulated signals. Each of the digital signals were simulated with a IF carrier frequency uniformly distributed between .23 and .27 times the sampling rate, a symbol rate uniformly distributed between .16 and .24 times the sampling rate, and a raised cosine pulse shape with a 50% excess bandwidth, with the exception of the BFSK which was not modeled with a shaping filter. The analog signals were bandlimited using the same raised cosine filter. Additionally, the classifier's receive filter is assumed to be an ideal low-pass filter. Since the symbol rate is assumed to be unknown,

the digital signals were not sampled at an integer multiple of the symbol rates, but were sampled at a constant rate independent of the symbol rate and the IF carrier frequency.

The SOF neural-network based system was train and tested using 4096 samples, corresponding to an average of approximately 410 symbols for the digitally modulated signals. The HOCS based system was tested with 65536 samples for its classification decision, corresponding to an average of approximately 6500 symbols for the digitally modulated signals. The system was tested in a variety of channel conditions, with an SNR range of 0 dB to 15 dB. The channel models simulated include a flat fading channel, two-path fading channel, and a harsh 20-path fading channel. Each of the fading channels are simulated as:

1. Block fading: where the channel is assumed to be fading slowly enough so as to be approximately static over a single block of observed data
2. Fast fading: where the channel is assumed to be fading with each path maintaining a coherence value of 0.9 over 500 samples, approximately equal to 50 symbols for the digitally modulated signals.

Additionally, it is assumed that the SNR of the signal on each antenna can be accurately estimated, and that the channel phase offset between antennas is accurately determined for the block flat fading channel.

The system performance is measured by its probability of correct classification (Pcc), defined as the percentage of the total number of modulation classifications made that were accurate. The SOF based classifier from [3] using only the cycle frequency profile is simulated as a benchmark for comparison to the first stage of the proposed classifier. This demonstrates the advantage of using both the cycle frequency as well as the spectral frequency profile for the initial classification stage. The purely eighth-order CC feature vector from [4] is used as a benchmark for comparison to the proposed classifier from end to end. However, to achieve a fair comparison, the AM, DS-CDMA, and BFSK signals were ex-

cluded from consideration for this case since the purely eighth-order CC does not have the ability to classify signals of this type.

4.3 Results

The systems were first tested in the block flat fading channel. Here, the systems were simulated using a multi antenna approach. The initial SOF-based stage used the MRC-variant method outlined in Section 3.5, while the HOCS-based stage utilized traditional MRC. Figure 4.16 compares the performance of the first stage of the proposed classifier with that of its benchmark. Here the proposed classifier obtains a significant performance increase over the baseline. The initial stage of the proposed classifier achieves a classification rate of nearly 100% Pcc for all SNR levels of interest when using four antennas with the MRC-variant. Figure 4.17 compares the proposed classifier to its eighth-order CC counterpart. In this case, the proposed classifier achieves a gain of 3 dB SNR over the benchmark. It is also noteworthy that with the addition on only a single antenna, a considerable performance gain is achieved.

Next, the systems were tested in a two-path as well as a 20-path block fading channel. Here, the initial SOF based stage was again implemented with the MRC-variant, while the HOCS based systems used SC. The performance of the initial classification stage subject to the two-path channel is shown in Figure 4.18 and the results under the 20-path channel are shown in Figure 4.19. These figures demonstrate the robustness of the SOF against multipath channel effects, as it is subject to only a slight performance degradation as compared to the flat fading channel. The performance of the final classification decision is shown in Figure 4.20 for the two-path case and in Figure 4.21 for the multipath case. The performance of each system is significantly degraded from the performance under the flat fading channel. The performance is insufficient to classify a signal with a reasonable degree of reliability. However, a benefit to the multistage approach is that it utilizes lower-order CCs in

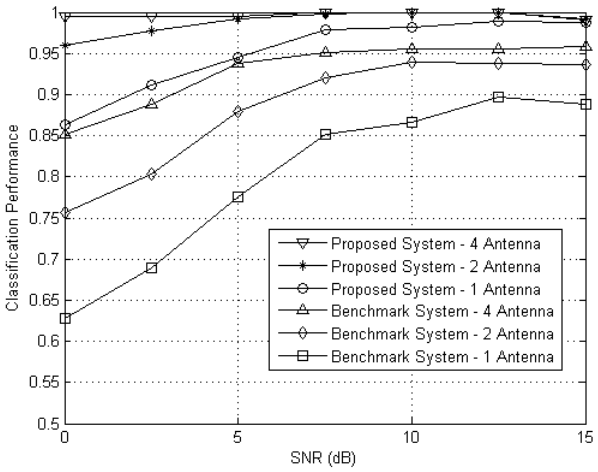

Figure 4.16: Classification Performance of proposed initial SOF-based classification stage and benchmark in a block flat fading channel using the MRC-variant combining scheme

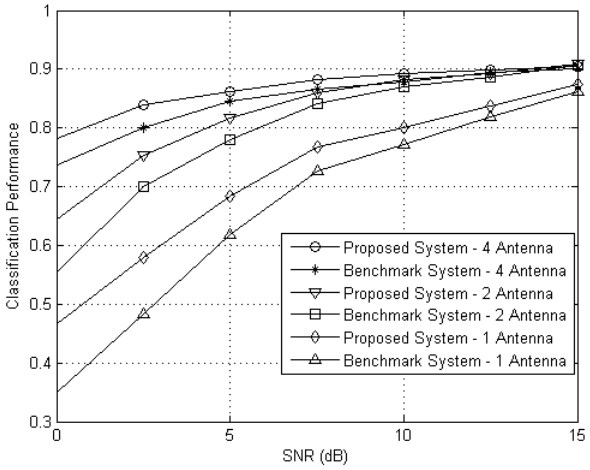

Figure 4.17: Classification Performance of proposed classifier and benchmark in a block flat fading channel using MRC

each decision stage, thus lowering the variance of the estimated statistic. While this doesn't achieve a significant benefit in the final classification stage, it does allow for a more reliable estimate of the family of the signal. This is demonstrated in Figures 4.22 and 4.23 where the ability of the two systems to classify the received signal as having a real-valued constellation (BPSK, 4ASK, or 8ASK), a square-constellation(QPSK, 16QAM, or 64 QAM), a circular constellation (8PSK or 16PSK), or as being an OFDM signal, where the other three signal types are not considered as the purely eighth-order CC feature vector is not capable of classifying them. Here, while it is noted that the number of antennas used does not affect the overall modulation-family classification performance, using the multistage approach does increase the observed classification performance by approximately two to three times.

Finally, the classifier performance was evaluated under the fast fading channels. The performance of the SOF based classifier under each of the fast fading channels is given in Figures 4.24 through 4.26. The initial stage of the classifier is again only moderately degraded, and the proposed classifier maintains a significant improvement over the baseline SOF based classifier. The final stage of the classifiers are unable to reliably determine the exact modulation scheme of the received signal under these harsh channel conditions, the multistage approach is still able to reliably determine the modulation family of the signal of interest. The system performance under the fast fading flat, two-path, and 20-path channels are shown in Figures 4.27 through 4.29.

The ability of the system to maintain a high degree of reliability in determining the modulation family is in part due to the insensitivity of the SOF to the multipath affect, as well as to the fewer number of required symbols that must be observed before a classification can be made. As this stage of the classifier requires significantly fewer observed symbols to make a classification, it is only moderately affected. While the purely eighth-order CC based classifier is drastically degraded, the proposed classifier is still able to produce a moderate gain in modulation class recognition, demonstrating the ability of the

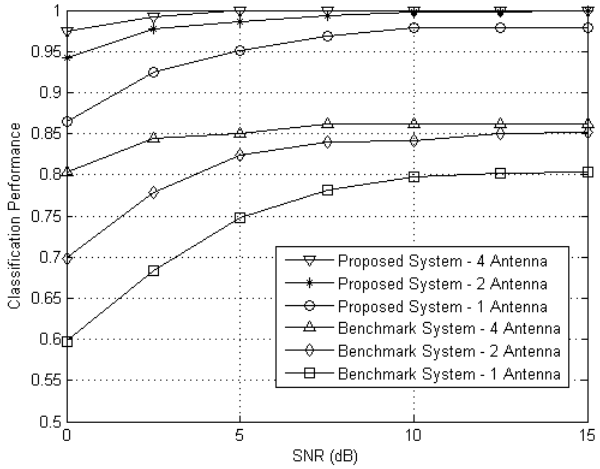

Figure 4.18: Classification Performance of proposed initial SOF-based classification stage and benchmark in a block two-path fading channel using the MRC-variant combining scheme

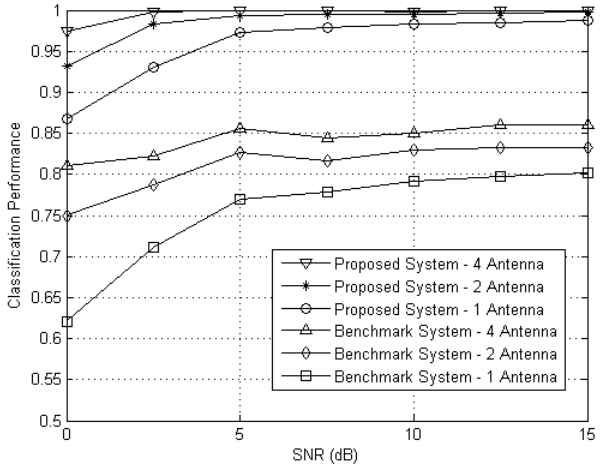

Figure 4.19: Classification Performance of proposed initial SOF-based classification stage and benchmark in a block 20-path fading channel using the MRC-variant combining scheme

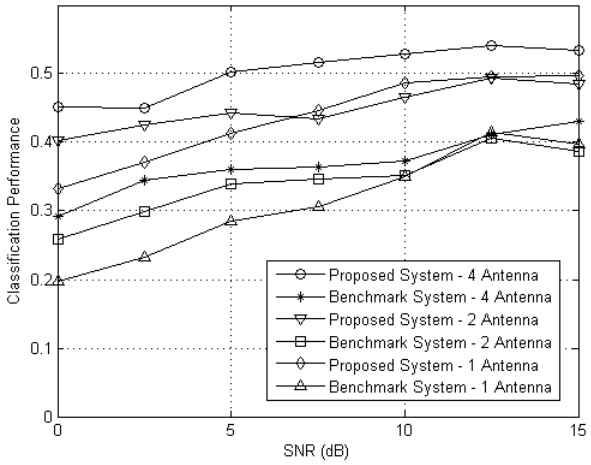

Figure 4.20: Classification Performance of proposed classifier and benchmark in a block two-path fading channel using SC

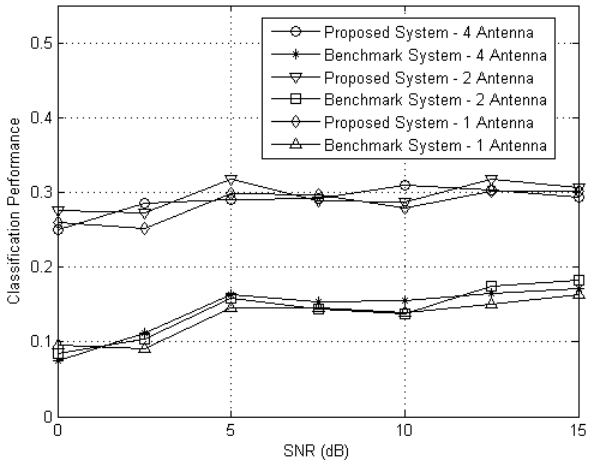

Figure 4.21: Classification Performance of proposed classifier and benchmark in a block 20-path fading channel using SC

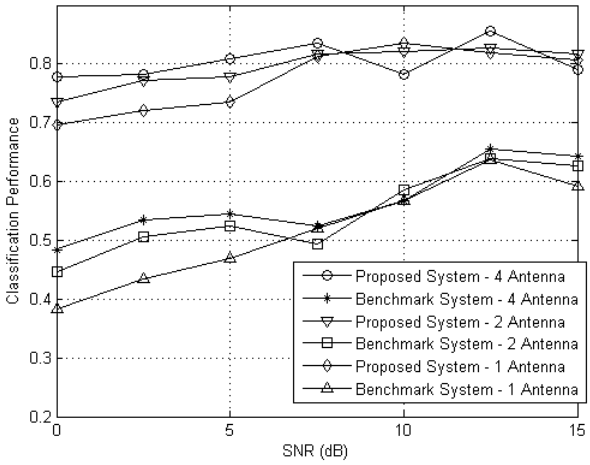

Figure 4.22: Ability of proposed classifier and benchmark to determine a signal's modulation family in a block two-path fading channel using SC

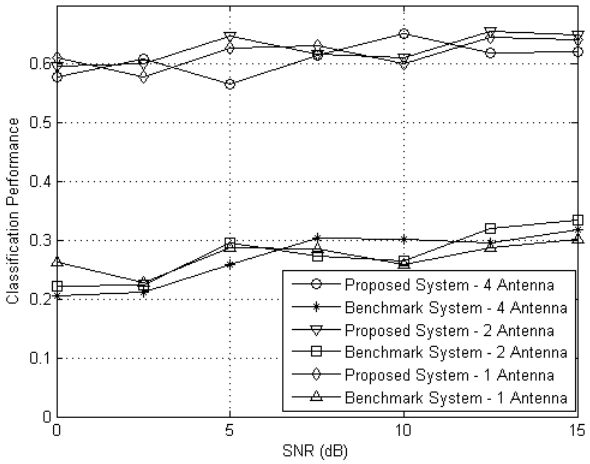

Figure 4.23: Ability of proposed classifier and benchmark to determine a signal's modulation family in a block 20-path fading channel using SC

lower order cumulants to reliably distinguish between lower order modulations even in the presence of multipath fading.

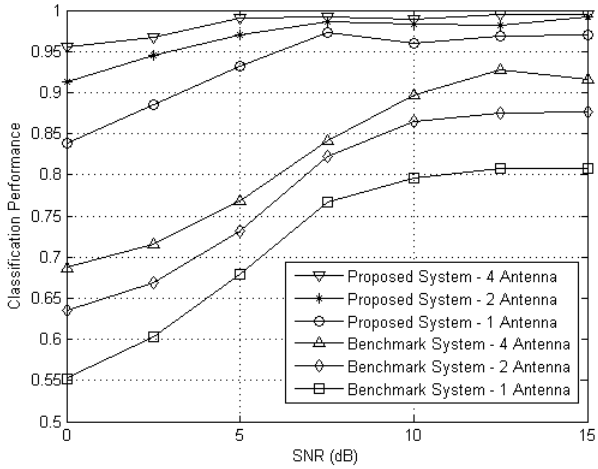

Figure 4.24: Classification Performance of proposed initial SOF-based classification stage and benchmark in a fast varying flat fading channel with using the MRC-variant combining scheme

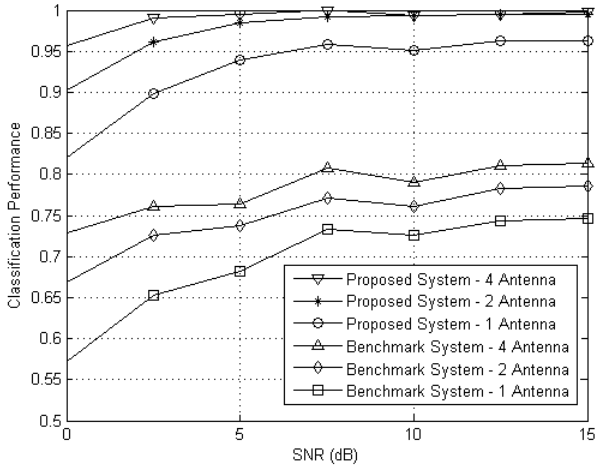

Figure 4.25: Classification Performance of proposed initial SOF-based classification stage and benchmark in a fast varying two-path fading channel using the MRC-variant combining scheme

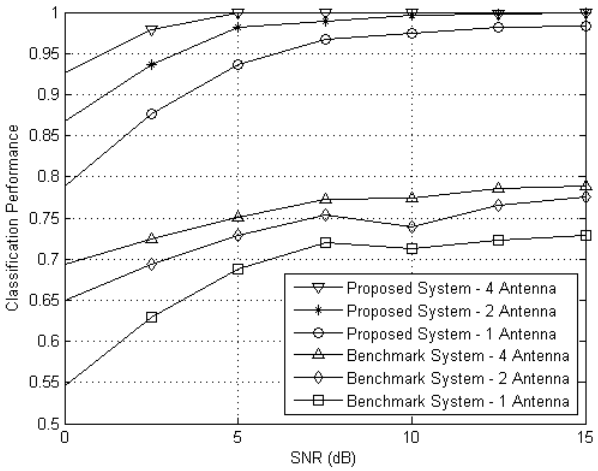

Figure 4.26: Classification Performance of proposed initial SOF-based classification stage and benchmark in a fast varying 20-path fading channel using the MRC-variant combining scheme

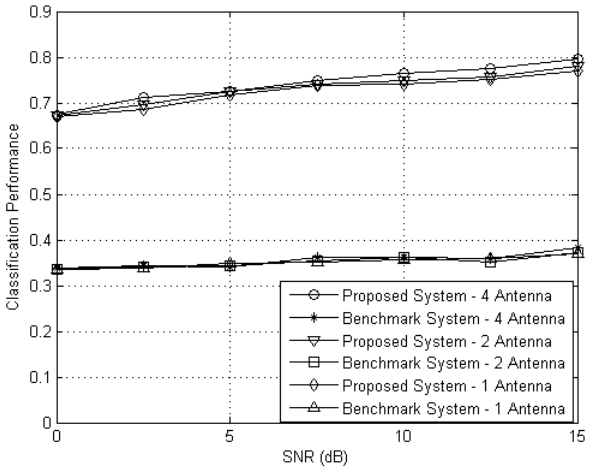

Figure 4.27: Ability of proposed classifier and benchmark to determine a signal's modulation family in a fast varying flat fading channel using SC

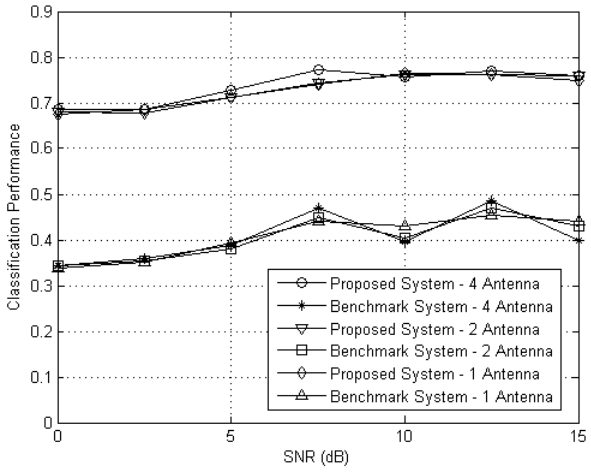

Figure 4.28: Ability of proposed classifier and benchmark to determine a signal's modulation family in a fast varying two-path fading channel using SC

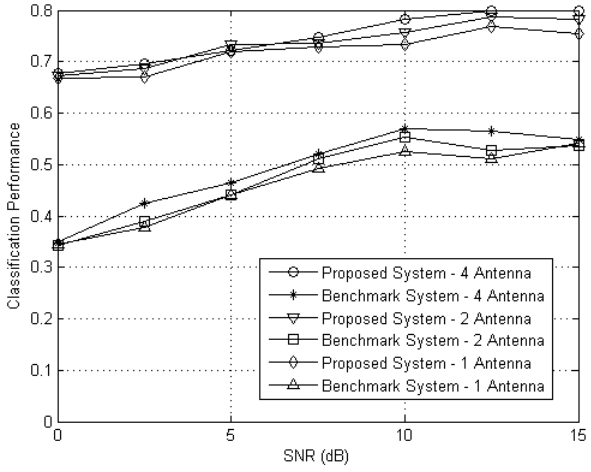

Figure 4.29: Ability of proposed classifier and benchmark to determine a signal's modulation family in a fast varying 20-path fading channel using SC

4.4 Summary

This chapter described the simulation setup under which the designed modulation classification system is evaluated. The system was evaluated in a variety of channel conditions, without a priori knowledge of critical signal parameters, and is compared to the baseline classifiers given in [3] and [4].

The proposed modulation classification system is capable of reliably distinguishing between classification systems in flat fading. The classification rate is further enhanced by the use of multiple receive antennas. As the number of antennas in increased, the performance of the baseline classifier approaches that of the multistage classifier designed here. In this situation, the multistage classifier can be used to trade hardware complexity for a modest gain in computational complexity. In the presence of multipath and fast fading channels, the system performance is significantly degrade as compared to the flat fading channel performance, even with the use of additional receive antennas. However, even though the final modulation scheme of the signal cannot be determined reliably, the classifier is still able to reliably determine the modulation family of the received signal, due to its hierarchical scheme. In contrast, the baseline classifier given in [4] produces significantly less reliable estimates of the modulation family under the same channel conditions.

5 Conclusion

5.1 Restatement of Research Goal

The goal of this research was to develop a non-cooperative modulation recognition system capable of reliably identifying the modulation scheme of a received signal. The system was desired to not require knowledge of critical signal parameters, such as carrier frequency, symbol rate, or phase offset, among others.

5.2 Conclusions

A modulation recognition system based cyclostationary statistics is presented, and its performance evaluated in a variety of channel conditions. The classifier is implemented in a multistage approach in order to leverage the reliability and lower variance of lower order cyclic statistics. Methods to increase the performance in the presence of fading channels are developed based on leveraging multiple receive antennas.

The resulting modulation classification system is capable of reliably determining the modulation scheme of received signals in moderate channel conditions. With the addition of multiple receive antennas, the performance of the classifier is shown to be increased even further. In multipath and fast fading channels, the final modulation discrimination capability of the classifier becomes unreliable, even when exploiting additional receive antennas.

However, even though the final modulation scheme of the signal cannot be determined reliably, the classifier is still able to reliably determined the modulation family of the received signal due to its hierarchical scheme, in contrast to the performance of the baseline classifier.

5.3 Recommendations for Future Research

Several areas remain open for further research and improvements to the proposed classifier design.

The primary area of necessary research is to increase the final classification performance in the presence of multipath and fast fading channels. These methods could include blind equalization of the received signal, or modification of the CC estimates to compensate for channel corruption. Additionally, methods aimed at reducing the required number of samples needed to obtain CC estimates would prove extremely useful, and would also improve classifier performance under moderately paced fading channels.

Extension of the classifier to additional signal types would be of added benefit so as to operate on a larger set of signals. Moreover, identification of the pulse shape of received signals would increase the precision of the classification, and would allow more reliable demodulation of the received signals once after their modulation scheme has been determined. Furthermore mitigating effects due to transmitter timing errors and phase jitter would increase the reliability of estimates made of lower fidelity systems.

Another area of necessary research includes extending the classification system to consider signals that are undergoing jamming, and the joint classification of multiple, possibly overlapping signals.

Finally, evaluation of the classifier with actual transmitted waveforms would be of benefit to prove a realized hardware solution is able to match the simulated performance.

Bibliography

[1] J. Mitola, *Cognitive Radio: An Integrated Agent Architecture For Software Defined Radio*, Ph.D. dissertation, KTH Royal Institute of Technology, Stockholm, Sweden, 2000.

[2] The proceedings of the *1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Baltimore, 2005.

[3] A. Fehske, J. Gaeddert and J. H. Reed, "A New Approach to Signal Classification Using Spectral Correlation and Neural Networks", *1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 144-150, Baltimore, 2005.

[4] O. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Selection Combining for Modulation Recognition in Fading Channels," in *IEEE MILCOM Proceedings*, 2005.

[5] P. Prokopios, A. Anastasopoulos, and A. Polydoros, "Likelihood Ratio Tests for Modulation Classification," in *IEEE MILCOM Proceedings*, pp.670674, Los Angeles, California, Oct. 2000.

[6] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I*, New York: Wiley, 1968.

[7] O Dobre, A Abdi, Y Bar-Ness, and W. Su, "A Survey of Automatic Modulation Classification Techniques: Classical Approaches and New Trends," *IEEE Proceedings on Communications*, 2006.

[8] E. Azzouz and A. Nandi, *Automatic Modulation Recognition of Communication Signals*, Kluwer Academic Publishers, 1996.

[9] W. Su and J. Kosinski, "Comparison and Modification of Automated Communication Modulation Recognition Methods", *IEEE MILCOM Proceedings*, October, 2002.

[10] W. A. Gardner, W. A. Brown and C.-K. Chen, Spectral Correlation of Modulated Signals: Part II - Digital Modulation, *IEEE Transactions on Communications*, Vol. 35, No. 6, pp. 595-601, June 1987.

[11] W. A. Gardner, *Cyclostationarity in Communications and Signal Processing*, New Jersey: IEEE Press, 1993.

[12] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a Century of Research," *IEEE Signal Processing*, Vol 86, No. 4, pp. 639-697, 2006.

[13] W. A. Gardner. *Statistical Spectral Analysis: A Nonprobabilistic Theory*. New Jersey: Prentice Hall, 1988.

[14] W. A. Gardner, Measurement of Spectral Correlation, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-34, No. 5. October 1986.

[15] Brown W.A., Loomis H.H., Digital Implementations of Spectral Correlation Analyzers, *IEEE Transaction on signal processing*, Vol. 41, No. 2, pp. 703-720 Feb. 1993.

[16] Roberts, R., Brown, W., Loomis, H., Computationally Efficient Algorithms for Cyclic Spectral Analysis, *IEEE Signal Processing Magazine*, Vol. 8, No. 2, pp. 38-49, April 1991.

[17] A. Swami and B. Sadler, Hierarchical Digital Modulation Classification Using Cumulants, *IEEE Transactions on Communication*, Vol. 48, pp.416-429, 2000.

[18] Statistical Signal Processing, Inc. *Spectral Correlation Analyzer: Users Guide*, Version 1.0. May 1997.

[19] O. Dobre, Y. Bar-Ness, and W. Su, Higher Order Cyclic Cumulants for High Order Modulation Classification, *IEEE MILCOM Proceedings*, 2003.

[20] W. A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems*. New York: Macmillan, 1986.

[21] D. Pauluzzi and N. Beaulieu, A Comparison of SNR Estimation Techniques for the AWGN Channel, *Proceedings for IEEE Pacific Rim Conf. Communications, Computers, and Signal Processing*, pp. 36-39, 1995.

[22] Z. Chen, B. Nowrouzian, and C. Zarowski, An Investigation of SNR Estimation Techniques Based on Uniform Cramer-Rao Lower Bound, *48th Midwest Symposium on Circuits and Systems*, 2005.

[23] A. Wiesel, J. Goldberg, and H. Messer-Yaron, SNR Estimation in Time-Varying Fading Channels, *IEEE Transactions on Communications*, Volume 54, 2006.