Signal Detection and Modulation Classification in Non-Gaussian Noise Environments

Venkata Gautham Chavali

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering

Claudio R. C. M. da Silva, Chair
R. Michael Buehrer
Steven W. Ellingson
Patrick R. Schaumont
Scott C. Leman

July 12, 2012
Blacksburg, Virginia

Keywords: Signal Detection, Modulation Classification, Non-Gaussian Noise

Signal Detection and Modulation Classification in Non-Gaussian Noise Environments

Venkata Gautham Chavali

(ABSTRACT)

Signal detection and modulation classification are becoming increasingly important in a variety of wireless communication systems such as those involving spectrum management and electronic warfare and surveillance, among others. The majority of the signal detection and modulation classification algorithms available in the literature assume that the additive noise has a Gaussian distribution. However, while this is a good model for thermal noise, various studies have shown that the noise experienced in most radio channels, due to a variety of man-made and natural electromagnetic sources, is non-Gaussian and exhibits impulsive characteristics. Unfortunately, conventional signal processing algorithms developed for Gaussian noise conditions are known to perform poorly in the presence of non-Gaussian noise. For this reason, the main goal of this dissertation is to develop statistical signal processing algorithms for the detection and modulation classification of signals in radio channels where the additive noise is non-Gaussian.

One of the major challenges involved in the design of these algorithms is that they are expected to operate with limited or no prior knowledge of the signal of interest, the fading experienced by the signal, and the distribution of the noise added in the channel. Therefore, this dissertation develops new techniques for estimating the parameters that characterize the additive non-Gaussian noise process, as well as the fading process, in the presence of unknown signals. These novel estimators are an integral contribution of this dissertation.

The signal detection and modulation classification problems considered here are treated as hypothesis testing problems. Using a composite hypothesis testing procedure, the unknown fading and noise process parameters are first estimated and then used in a likelihood ratio test to detect the presence or identify the modulation scheme of a signal of interest. The proposed algorithms, which are developed for different non-Gaussian noise models, are shown to outperform conventional algorithms which assume Gaussian noise conditions and also algorithms based on other impulsive noise mitigation techniques.

This dissertation has three major contributions. First, in environments where the noise can be modeled using a Gaussian mixture distribution, a new expectation-maximization algorithm based technique is developed for estimating the unknown fading and noise distribution parameters. Using these estimates, a hybrid likelihood ratio test is used for modulation classification. Second, a five-stage scheme for signal detection in symmetric α stable noise environments, based on a class of robust filters called the matched myriad filters, is presented. New algorithms for estimating the noise distribution parameters are also developed. Third, a modulation classifier is proposed for environments in which the noise can be modeled as a time-correlated non-Gaussian random process. The proposed classifier involves the use of a whitening filter followed by likelihood-based classification. A new H_∞ filter-based technique for estimating the whitening filter coefficients is presented.

Dedication

To my parents, my advisor Dr. Claudio da Silva, and to my dear friend Anup.

Acknowledgments

Firstly, I am very grateful to my advisor Dr. Claudio da Silva without whose valuable guidance and constant encouragement this work would not have been possible. Working with him has helped me understand the importance of being sincere, thorough, and methodical in one's approach to work. The last three years have been a wonderful learning experience and have greatly helped me grow and mature as a researcher. I will always be indebted to him for instilling in me the confidence to trust my own abilities.

I would like to thank the members of my PhD committee, Dr. Ellingson, Dr. Buehrer, Dr. Schaumont, and Dr. Leman, for their valuable suggestions and their time. I would also like to thank all my teachers at Virginia Tech and the Indian Institute of Technology Madras for teaching me the fundamental concepts which have greatly helped me in my research.

I would like to thank the Bradley Department of Electrical and Computer Engineering for providing me with the opportunity to pursue my PhD studies here at Virginia Tech. I am also very grateful for all the support and valuable inputs from the people at InterDigital Inc., the sponsors, in part, of this research.

Special thanks to my uncle Dr. L. Kameshwara Rao at the Indian Institute of Science, Bangalore, for the long discussions on graduate education.

I am very thankful to my friends and colleagues at Wireless@VT, especially Chris Headley, Daniel Jakubisin, and Sai Dhiraj for all their invaluable help and support. I would also like to thank Nancy Goad, Hilda Reynolds, and Cindy Hopkins for all their help over the last

three years.

I would further like to thank my wonderful friends in Blacksburg, particularly Dhrumeel Bakshi, Nitin Arora, Harpreet Singh Dhillon, Sushrutha Vigraham, Dinesh Datla, Mike Steiner, and Ambuj Agrawal for all the fun times here. I would also like to thank the wonderful people of the town of Blacksburg for all their warmth and their help in making this town a home away from home.

Finally, I would like to thank my parents and my sister for helping me become the person I am today and God for his blessings.

Contents

1 Introduction **1**

1.1 Outline of the Dissertation 4

1.2 Contributions 5

2 Noise Models **8**

2.1 Middleton's Model 9

2.1.1 Class A noise 10

2.1.2 Class B noise 12

2.1.3 Additional notes on Middleton's model 12

2.2 Symmetric α Stable Model 14

2.2.1 Model description 14

2.2.2 Comparison with Middleton's model 15

2.3 Time-Correlated Non-Gaussian Channels 17

2.3.1 Models for time-correlated non-Gaussian noise 18

2.4 Conclusions 20

3 Maximum-Likelihood Modulation Classification in Gaussian Mixture Noise Environments **21**

3.1 System Model 22

3.2 Likelihood-based Modulation Classifier 23

3.3 Estimation of the Unknown Channel and Noise Distribution Parameters . . . 25

3.4 Numerical Results and Discussion 29

3.5 Conclusion 33

3.6 Acknowledgments 34

4 Signal Detection in Symmetric α Stable Noise Environments **37**

4.1 System Model 38

4.2 The Myriad Filter 40

4.3 Proposed Detection Scheme 42

4.3.1 Estimation of noise distribution parameters 42

4.3.2 Matched myriad filter 47

4.3.3 Processing after the matched myriad filter 48

4.4 Numerical Results 50

4.5 Conclusion 57

4.6 Acknowledgments 58

5 Modulation Classification in Time-Correlated Non-Gaussian Noise Environments **63**

5.1 System Model 65

5.2 Likelihood-based modulation classification 67
5.3 Estimation of the AR parameters 70
5.3.1 The H_∞ filter 71
5.3.2 H_∞ filter-based estimation of AR parameters 73
5.3.3 Dual H_∞ filter-based estimation of the AR parameters 75
5.3.4 Choosing of the H_∞ filter parameters 78
5.4 Estimation of the channel coefficient and noise distribution parameter 79
5.5 Numerical Results 80
5.6 Conclusions 87

6 Conclusions **91**

A Collaborative Spectrum Sensing Based on a New SNR Estimation and Energy Combining Method **94**

A.1 System Model 95
A.2 Collaborative Spectrum Sensing Methods 97
A.3 Proposed Collaborative Spectrum Sensing Method 101
A.4 Numerical Results 104
A.5 Conclusions 108
A.6 Acknowledgments 109

B Asymptotic Normality of the Matched Myriad Filter Output **110**

B.1 Rectangular pulse shape 111

B.1.1 Asymptotic normality in the no transition-case 111

B.1.2 Asymptotic normality in the one transition-case 112

B.2 SRRC pulse shape 113

C Model Matching using Measured Impulsive Noise Statistics **115**

Bibliography **118**

List of Figures

3.1 Log-likelihood function of the observed data for 8-PSK (dotted lines) and 16-QAM (solid) and various noise parameters: $\lambda_2 = 0.05,\ \kappa = 75$ (square), $\lambda_2 = 0.05,\ \kappa = 75$ (star), $\lambda_2 = 0.1,\ \kappa = 100$ (diamond), $\lambda_2 = 0.2,\ \kappa = 100$ (circle), and $\lambda_2 = 0.25,\ \kappa = 100$ (plus). 30

3.2 Average bias in the estimates of the phase (in radians, dashed lines) and magnitude (unitless, solid lines) of the channel coefficient α. The average bias in the estimate $\hat{X}$ of the parameter X is defined as $E[|X - \hat{X}|]$. Number of observed symbols equal to 100 (diamond), 200 (square), and 500 (circle). . . 31

3.3 Normalized variance of the estimates of σ_1^2 (dashed lines) and σ_2^2 (solid). The normalized variance of the estimate $\hat{X}$ of the parameter X is defined as $Var[\hat{X}]/X$. Number of observed symbols equal to 100 (diamond), 200 (square), and 500 (circle). 32

3.4 Probability of correct classification of the proposed classifier for $\lambda_2 = 0.1$ and $\kappa = 100$. The performances of ideal classifiers derived for the non-Gaussian (star) and Gaussian noise cases (plus) are also shown. Number of observed symbols for parameter estimation equal to 100 (diamond), 200 (square), and 500 (circle). Number of symbols used for classification is 500. 33

3.5 Probability of correct classification of the proposed classifier for $\lambda_2 = 0.3$ and $\kappa = 250$. The performances of ideal classifiers derived for the non-Gaussian (star) and Gaussian noise cases (plus) are also shown. Number of observed symbols for parameter estimation equal to 100 (diamond), 200 (square), and 500 (circle). Number of symbols used for classification is 500. 34

3.6 Probability of correct classification of the proposed classifier for a Rician channel with parameters $K = 0.5$ and $E[|\alpha|^2] = 3$. $\lambda_2 = 0.1$ and $\kappa = 100$. The performances of ideal classifiers derived for the non-Gaussian (star) and Gaussian noise cases (plus) are also shown. Number of observed symbols for parameter estimation equal to 100 (diamond), 200 (square), and 500 (circle). Number of symbols used for classification is 500. 35

4.1 Proposed five-stage detection scheme. 42

4.2 Performance of the proposed estimators for SRRC pulse shape as a function of N_{ECF}. Bias of the estimates of α (circle) and γ (square) for the exact (dashed line) and alternative (solid line) methods. α, γ, and the signal power are equal to 1.4, 0.5, and 1, respectively. Roll off factor is equal to 0.75. The bias of the estimate $\hat{X}$ of X is defined as $E[|X - \hat{X}|]$. 52

4.3 Performance of the proposed estimators for SRRC pulse shape as a function of N_{ECF}. Standard deviation of the estimates of α (circle) and γ (square) for the exact (dashed line) and alternative (solid line) methods. α, γ, and the signal power are equal to 1.4, 0.5, and 1, respectively. Roll off factor is equal to 0.75. 53

4.4 Performance of the considered detectors for α and γ equal to 1.3 and 1, respectively. Adjusted SNR $\approx$ -7dB. Rectangular pulse shape. a) ideal detector (solid line); b) proposed detector with N_{ECF} equal to 50,000 (dash-dot line); c) proposed detector with N_{ECF} equal to 10,000 (dotted line); and d) ZMNL-based detector (dashed line). The number of symbols used for timing synchronization, parameter estimation in stage IV, and detection are 500, 1,000, and 250, respectively. 54

4.5 Performance of the four considered detectors for α and γ equal to 1.3 and 0.5, respectively. Adjusted SNR $\approx$ -2.5dB. SRRC pulse shape with roll off factor of 0.75. a) ideal detector (solid line); b) proposed detector - exact method (dash-dot line); c) proposed detector - alternative method (dotted line); and d) ZMNL-based detector (dashed line). The number of symbols used for timing synchronization, parameter estimation in stage IV, and detection are 500, 1,000, and 250, respectively. 55

4.6 Performance of the four considered detectors for α and γ equal to 1.2 and 0.75, respectively. Adjusted SNR $\approx$ -4.5dB. SRRC pulse shape with roll off factor of 0.75. a) ideal detector (solid line); b) proposed detector - exact method (dash-dot line); c) proposed detector - alternative method (dotted line); and d) ZMNL-based detector (dashed line). The number of symbols used for timing synchronization, parameter estimation in stage IV, and detection are 500, 1,000, and 250, respectively. 56

4.7 Performance of the four considered detectors for α, γ, and signal power equal to 1.1, 10, and 100, respectively. Adjusted SNR ≈-5dB. SRRC pulse shape with roll off factor of 0.75. a) ideal detector (solid line); b) proposed detector - exact method (dash-dot line); c) proposed detector - alternative method (dotted line); and d) ZMNL-based detector (dashed line). The number of symbols used for timing synchronization, parameter estimation in stage IV, and detection are 500, 1,000, and 250, respectively. 58

4.8 Performance of the four considered detectors for α, and γ equal to 0.8 and 1.5, respectively. Adjusted SNR ≈-10dB. SRRC pulse shape with roll off factor of 0.75. a) ideal detector (solid line); b) proposed detector - exact method (dash-dot line); c) proposed detector - alternative method (dotted line); and d) ZMNL-based detector (dashed line). The number of symbols used for timing synchronization, parameter estimation in stage IV, and detection are 500, 1,000, and 250, respectively. P = 18. 59

5.1 H_∞ filter-based dual estimation scheme for AR parameter estimation. 75

5.2 Average bias in the estimates of the AR(1) process parameter: $a[1] = -0.6$. The average bias in the estimate $\hat{X}$ of the parameter X is defined as $E[|X - \hat{X}|]$. Number of symbols for AR parameter estimation equal to 250 (dotted line), 500 (dashed line), and 1000 (solid line). Modulation scheme used is QPSK. 81

5.3 Average bias in the estimates of the AR(2) process parameters: $a[1] = -0.25$ (circle), $a[2] = -0.53$ (square). The average bias in the estimate $\hat{X}$ of the parameter X is defined as $E[|X - \hat{X}|]$. Number of symbols used in the estimation of the AR parameters is equal to 500 (dashed line) and 1,000 (solid line). Modulation scheme used is QPSK. 82

5.4 Average normalized autocorrelation at the whitening filter output: SNR = -10 dB (dashed line); SNR = 10 dB (dotted line). The average normalized autocorrelation is defined as $E[\frac{|R_X|}{\max(|R_X|)}]$, where R_X is the autocorrelation function. Number of symbols used in the estimation of the AR parameters is 1,000. AR(1) process - $a[2] = -0.6$ Also plotted is the average normalized correlation at the whitening filter input (solid line). 83

5.5 Average normalized autocorrelation at the whitening filter output: SNR = -10 dB (dashed line); SNR = 10 dB (dotted line). The average normalized autocorrelation is defined as $E[\frac{|R_X|}{\max(|R_X|)}]$, where R_X is the autocorrelation function. Number of symbols used in the estimation of the AR parameters is 1,000. AR(2) process : $a[1] = -0.25, a[2] = -0.53$. Also plotted is the average normalized correlation at the whitening filter input (solid line). . . . 84

5.6 Normalized variance of the estimates of σ_1^2 (dashed lines) and σ_1^2 (solid) when the whitening filter coefficients are estimated (circle) are perfectly known (square). The normalized variance of the estimate $\hat{X}$ of the parameter X is defined as $\frac{\text{var}(\hat{X})}{X}$. Modulation scheme considered is QPSK. AR(1) process : $a[1] = -0.6$. Number of symbols used in the estimation of the AR parameters and in the EM algorithm are 1,000 and 500, respectively. 85

5.7 Normalized variance of the estimates of σ_1^2 (dashed lines) and σ_1^2 (solid) when the whitening filter coefficients are estimated (circle) and are perfectly known (square). The normalized variance of the estimate $\hat{X}$ of the parameter X is defined as $\frac{\text{var}(\hat{X})}{X}$. Modulation scheme considered is QPSK. AR(2) process : $a[1] = -0.25, a[2] = -0.53$. Number of symbols used in the estimation of the AR parameters and in the EM algorithm are 1,000 and 500, respectively. . . 86

5.8 Average bias in the estimates of the phase (in radians, dashed lines) and magnitude (unit less, solid lines) of the channel coefficient α when the whitening filter coefficients are estimated (circle) and are perfectly known (square). The average bias of the estimate $\hat{X}$ of the parameter X is defined as $E[|\hat{X} - X|]$. Modulation scheme considered is QPSK. AR(1) process : $a[1] = -0.6$. Number of symbols used in the estimation of the AR parameters and in the EM algorithm are 1,000 and 500, respectively. 87

5.9 Probability of correct classification of the proposed classifier. AR(1) process with $a[1] = -0.6$ and two-term Gaussian mixture model with $\lambda_1 = 0.9$ and $\frac{\sigma_2^2}{\sigma_1^2} = 100$. The number of symbols used for AR parameter estimation and the EM algorithm are 1,000 and 500 (circle) and 500 and 1,000 (square), respectively. The performances of "ideal" classifier (solid line), the classifier developed in Chapter 3 for white non-Gaussian noise with 1,500 symbols for parameter estimation (dashed line), and the classifier developed for Gaussian noise conditions in [1] with 1,500 symbols for parameter estimation (dotted line) are also shown. Number of symbols used for classification is 500. 88

5.10 Probability of correct classification of the proposed classifier. AR(2) process with $a[1] = -0.25$, $a[2] = -0.53$ and two-term Gaussian mixture model with $\lambda_1 = 0.9$ and $\frac{\sigma_2^2}{\sigma_1^2} = 100$. The number of symbols used for AR parameter estimation and the EM algorithm are 1,000 and 500 (circle) and 500 and 1,000 (square), respectively. The performances of "ideal" classifier (solid line), the classifier developed in Chapter 3 for white non-Gaussian noise with 1,500 symbols for parameter estimation (dashed line), and the classifier developed for Gaussian noise conditions in [1] with 1,500 symbols for parameter estimation (dotted line) are also shown. Number of symbols used for classification is 500. 89

5.11 Probability of correct classification of the proposed classifier (circle). AR(2) process with $a[1] = -0.25$, $a[2] = -0.53$, three-term Gaussian mixture model with proportions $\{0.098, 0.018, 0.002\}$ and variances $\{\sigma_1^2 = 10, \sigma_2^2 = 50, \sigma_3^2 = 1075\}$. The number of symbols used for AR parameter estimation and the EM algorithm are 250 and 500 (circle) and 1,000 and 5,000 (square), respectively. The performances of "ideal" classifier (solid line), the classifier developed in Chapter 3 for white non-Gaussian noise with 750 symbols for parameter estimation (dashed line), and the classifier developed for Gaussian noise conditions in [1] with 750 symbols for parameter estimation (dotted line) are also shown. Number of symbols used for classification is 500. 90

A.1 Performance of the EGC (dotted lines), WLC (dashed lines), and OC methods (solid lines) for $M = 2$ (SNR values equal to -1 and -5 dB, circle) and 4 (-1, -5, -8 and -10 dB, diamond) radios. The observation window of each radio has a total of $K = 50$ samples. 99

A.2 Illustration of the operation of the proposed collaborative spectrum sensing method. In this figure, N_{SI} and N_{off} are equal to 4 and 8, respectively. . . . 100

A.3 Instantaneous and estimated SNR values for a particular realization of the proposed simulation environment. The primary user is off from 10s to 20s. The sensing interval and N_{SI} are equal to 100 ms and 4, respectively. 103

A.4 Performance of the propose (dash-dot lines), EGC (dotted lines) and OC (solid lines) methods for $M = 2$ (diamond), 3 (star) and 4 (circle) radios. The sensing interval and N_{SI} are equal to 100 ms and 9, respectively. 105

A.5 Performance of the proposed method for N_{SI} equal to 4 (circle), 9 (diamond), and 19 (star). The sensing interval is equal to 100 ms. For comparison, the performance of the EGC (dotted line) and OC (solid line) methods are also shown. The collaborative network consists of $M = 4$ radios. 106

A.6 Performance of weighted energy combining during the 1^{st} (circle), 5^{th} (diamond), 9^{th} (star) sensing intervals after an interval in which SNR estimation is performed. For comparison, the performances of the EGC (dotted line) and OC (solid line) methods are also shown. The number of radios, sensing interval, and N_{SI} are equal to 4, 100 ms, and 9, respectively. 107

A.7 Performance of proposed method when weighted energy combining is performed by using the OC (dashed line) and WLC (dash-dot line) methods. For comparison, the performances of the EGC (dotted line) and OC (solid line) methods are also shown. The number of radios, sensing interval, and N_{SI} are equal to 3, 100 ms, and 9, respectively. 108

C.1 Statistics of the measured impulsive noise and the noise generated using the proposed model. Measurements made at 2.44GHz. The following parameters were used to generate the noise from the proposed model: 3-term Gaussian mixture model with $\lambda_1 = 0.98, \lambda_2 = 0.018, \lambda_3 = 0.002, \sigma_1^2 = 10, \sigma_2^2 = 50, \sigma_3^2 = 1075$; and AR(2) process with $a[1] = -0.25, a[2] = -0.53$. 117

List of Tables

3.1 Confusion matrix of the proposed classifier for $\lambda_2 = 0.1$ and $\kappa = 100$. Average received SNR = 5dB 35

3.2 Confusion matrix of the proposed classifier for $\lambda_2 = 0.3$ and $\kappa = 250$. Average received SNR = 5dB 36

4.1 Adjusted SNR (in dB) for different values of α, γ, and signal power. 50

4.2 Performance of the proposed estimator for rectangular pulse shape. (Bias, Standard deviation) of $\hat{\alpha}$ and $\hat{\gamma}$ are shown for different combinations of α, γ, and signal power. The bias of the estimate $\hat{X}$ of X is defined as $E[|X - \hat{X}|]$. 60

4.3 Performance of the proposed estimator for SRRC pulse shape (roll off factor = 0.75) - Exact method. (Bias, Standard deviation) of $\hat{\alpha}$ and $\hat{\gamma}$ are shown for different combinations of α, γ, and signal power. The bias of the estimate $\hat{X}$ of X is defined as $E[|X - \hat{X}|]$. 61

4.4 Performance of the proposed estimator for SRRC pulse shape (roll off factor = 0.75) - Alternative method. (Bias, Standard deviation) of $\hat{\alpha}$ and $\hat{\gamma}$ are shown for different combinations of α and γ. The bias of the estimate $\hat{X}$ of X is defined as $E[|X - \hat{X}|]$. 62

5.1 Proposed H_∞ filter-based dual estimation procedure for AR parameter estimation. 77

5.2 Confusion matrix of the proposed classifier. Received SNR = 6dB. AR(2) process with $a[1] = -0.25$, $a[2] = -0.53$ and two-term Gaussian mixture model with $\lambda_1 = 0.9$ and $\frac{\sigma_2^2}{\sigma_1^2} = 100$. The number of symbols used for AR parameter estimation and the EM algorithm are 500 and 1,000 (square), respectively. Number of symbols used for classification is 500. 89

5.3 Confusion matrix of the proposed classifier. Received SNR = 6dB. AR(2) process with $a[1] = -0.6$ and two-term Gaussian mixture model with $\lambda_1 = 0.9$ and $\frac{\sigma_2^2}{\sigma_1^2} = 100$. The number of symbols used for AR parameter estimation and the EM algorithm are 500 and 1,000 (square), respectively. Number of symbols used for classification is 500. 90

C.1 Statistics of the measured impulsive noise and the noise generated using the proposed model. Measurements made at 2.44GHz. The following parameters were used to generate the noise from the proposed model: 3-term Gaussian mixture model with $\lambda_1 = 0.98, \lambda_2 = 0.018, \lambda_3 = 0.002, \sigma_1^2 = 10, \sigma_2^2 = 50, \sigma_3^2 = 1075$; and AR(2) process with $a[1] = -0.25, a[2] = -0.53$. 116

Chapter 1

Introduction

Signal detection and classification are two of the most fundamental problems in signal processing, with numerous applications in areas such as communications, speech recognition, biomedicine, and image processing, among many others [2]. Signal detection refers to the problem of detecting the presence of a signal of interest from noisy observations while classification involves the identification of certain signal features. In this dissertation, we consider modulation classification, which is the problem of identifying the modulation format of the signal of interest[1].

Algorithms for signal detection and classification are essential components of a variety of wireless communication systems. For example, there has been a great deal of interest of late in developing intelligent radio systems which can more efficiently utilize the limited radio frequency spectrum. In these systems, through *dynamic spectrum access* (DSA), secondary users are allowed to access primary user spectrum on a dynamic and non-interfering basis [3]. For secondary users to be able to identify potential spectral opportunities, fast and accurate algorithms for signal detection and classification are critical. In addition, in multi-mode software radios and in adaptive modulation and coding systems, radio receivers must have knowledge of the modulation scheme used by the transmitter prior to signal demodulation.

[1]From here on, the term "classification" is used to refer to modulation classification.

The signaling overhead associated with the transmission of this modulation information can be reduced through the use of blind modulation classification techniques, thereby resulting in an increase in the data throughput [4]. Also, in military communications, electronic warfare and surveillance systems frequently scan the radio spectrum for detection and possible interception of hostile signals. Reliable algorithms for detection and classification are of vital importance in these scenarios [5].

A large number of algorithms have been developed in the literature for the detection and classification of communication signals. For a comprehensive survey of such algorithms, the reader is referred to [6] and [7] for detection and [8] for classification algorithms. The significant majority of these algorithms have been developed under the assumption that the additive noise experienced in the radio channel has a Gaussian distribution. While this is a very good model for thermal noise, various studies have shown that radio channels are affected by electromagnetic activity from a variety of sources, and that the combined noise due to these sources is non-Gaussian and has impulsive characteristics [9–20]. Examples of processes that are well modeled by non-Gaussian/impulsive models include atmospheric noise [21], co-channel interference in wireless networks [11], spurious emissions from electronic equipment like microwave ovens [16], clocks and buses of laptop and desktop computers [17], LCD monitors [22], elevators [19], and automobile ignitions [20], among numerous others. This combined noise, often referred to as "electro magnetic interference" and "radio frequency interference" is simply denoted as "noise" in this dissertation.

Signal processing algorithms developed for Gaussian noise environments typically perform significantly worse when non-Gaussian noise is present [17], [23–27]. This is due, in part, to "the lack of robustness of linear and quadratic type signal processing procedures to many types of non-Gaussian statistical behavior" [23]. For example, due to the impulsive nature of the additive non-Gaussian noise experienced in radio channels, a host of popular (quadratic type) signal processing techniques such as energy detectors [28], Kalman filters [29], and other least squares-based estimation procedures [30] are known to show a severe degradation in performance.

The aim of this dissertation, therefore, is to develop new statistical signal processing algorithms for reliable detection and classification of signals in practical radio channels with additive non-Gaussian noise.

One of the major challenges involved in the design of these algorithms is that they are expected to operate with limited or no prior knowledge of the signal parameters, the fading experienced by the signal, and the distribution of the noise added in the channel. This is because, while radios typically estimate the channel and noise distribution parameters *after* signal demodulation, signal detection and classification have to be performed in the preprocessing stage, that is, *before* signal demodulation [8]. Therefore, in this dissertation, blind techniques are designed for estimating the parameters of the non-Gaussian noise distribution and the state of the fading process[2]. The detectors and classifiers previously developed in the literature for non-Gaussian noise environments commonly assume that these parameters are either known a priori or estimated using training methods [24], [25], [27], [32], [33]. The design of these blind estimators is therefore is a major contribution of this dissertation.

With the aim of developing optimal or near-optimal detection and classification algorithms, likelihood-based techniques are adopted, since they minimize the probability of error [8]. In these techniques, the problem at hand is first formulated as a hypothesis testing problem. Following a *composite hypothesis testing* procedure, the unknown state of the fading process and the noise distribution parameters are estimated, and a likelihood ratio test is then used to choose the appropriate hypothesis [2].

In this dissertation, the signals of interest are assumed to be digital amplitude-phase modulated; that is, the considered modulation schemes are amplitude-shift keying (ASK), phase-

[2]The term "blind" refers to the fact that no training symbols are used in the estimation process. However, other signal parameters like carrier frequency, symbol period, pulse shape, and bandwidth are assumed to be known. This is a valid assumption for potential usage scenarios. For example, IEEE 802.22 radios that make spectrum access decisions by relying on spectrum sensing must be able to detect analog and digital TV signals [31]. The signal parameters used by the TV systems can be found in their standards. The assumption that the signal parameters are known is widely used in the literature for the design of detection and classification algorithms. See [6], [8], [25], and [32], among many others. For applications in which this assumption may not hold, these parameters should be incorporated into the design of the detector/classifier as nuisance parameters.

shift keying (PSK), and quadrature-amplitude modulation (QAM) schemes, which are used in a variety of wireless communication systems. Additionally, different statistical-physical models for the additive non-Gaussian noise are considered. These include Middleton's Class A model (in which the noise is modeled as having a Gaussian mixture distribution) [12] and the symmetric α stable model [9]. The proposed detection and classification algorithms, and the underlying estimators, are analyzed in detail. Also, the performance gains obtained over algorithms developed for Gaussian noise conditions and over other impulsive noise mitigation techniques are demonstrated.

In addition to developing various signal processing algorithms for non-Gaussian noise environments, a technique for improving the reliability of signal detection in spectrum sensing applications through *radio collaboration* is also presented. More specifically, in the Appendix of this dissertation, a collaborative spectrum sensing approach which exploits spatial diversity by *intelligently* combining the sensing data provided by multiple radios is proposed. In the proposed approach, the sensing data provided by each radio is weighted in proportion to its reliability. The novelty of this work lies in exploiting the correlation of the sensing data provided by each radio over successive sensing intervals to keep complexity of the weight estimation process at acceptable levels. It is shown that the proposed method offers an excellent tradeoff between reliability and complexity compared with conventional collaborative sensing methods.

1.1 Outline of the Dissertation

Chapter 2 presents different models available in the literature for the additive noise that affects radio channels. Two popular statistical-physical models, namely the Middleton's model and the symmetric α stable model, are examined in detail. The problem of modeling *time-correlated* non-Gaussian noise environments is also discussed.

In Chapter 3, a likelihood-based technique for the classification of digital amplitude-phase

modulated signals in flat fading channels with non-Gaussian noise is presented. The additive noise is modeled by a Gaussian mixture distribution. Also developed, is a novel technique for blind estimation of the fading state and the noise distribution parameters, which is based on the use of a variant of the expectation-maximization (EM) algorithm. It is shown that the proposed classifier outperforms a likelihood-based classifier designed for Gaussian noise conditions.

A five-stage scheme for signal detection in symmetric α stable noise is presented in Chapter 4. The proposed detector is based on the use of a matched myriad filter, which is known to be an effective technique for dealing with non-Gaussian/impulsive noise. New empirical characteristic function (ECF)-based algorithms for estimating the noise distribution parameters, in the presence of an additional unknown signal, are also developed. Results are presented which show that the proposed detection scheme outperforms a popular zero-memory non-linearity (ZMNL)-based scheme.

The problem of classification in time-correlated non-Gaussian noise environments is addressed in Chapter 5. The proposed classifier involves the use of a whitening filter followed by likelihood-based classification. A new technique for estimating the whitening filter coefficients, based on the use of the robust H_∞ filter, is also developed. The gain in performance obtained over classifiers which assume that the noise process is white is demonstrated.

A summary of the research and the conclusions are presented in Chapter 6. Finally, as previously mentioned, a collaborative spectrum sensing algorithm which exploits both the spatial diversity of multiple radios and the correlation in time of the sensing data provided by each radio in order to improve the reliability of spectrum sensing is presented in Appendix A.

1.2 Contributions

The major contributions of this dissertation are:

1. Chapter 3: Design of a likelihood-based classifier for identifying the modulation scheme of an unknown digital amplitude-phase modulated signal in a flat-fading channel with non-Gaussian noise. The non-Gaussian noise is modeled using a Gaussian mixture distribution. This research includes:

   (a) the development of a blind EM-based algorithm for estimating the unknown state of the fading process and the noise distribution parameters,

   (b) proposal of a hybrid likelihood ratio test-based modulation classifier, and

   (c) performance analysis of the proposed estimator and modulation classifier for different fading models and noise distribution parameters.

   Publication: V. G. Chavali and C. R. C. M. da Silva, "Maximum-likelihood classification of digital amplitude-phase modulated signals in flat fading non-Gaussian channels," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2051-2056, Aug. 2011.

2. Chapter 4: A novel matched myriad filter-based scheme for the detection of digital amplitude-phase modulated signals in symmetric α stable distributed noise environments. This research includes:

   (a) design of new ECF-based techniques for estimating the noise distribution parameters in the presence of an unknown signal,

   (b) proposal of a novel five-stage scheme for the detection of signals in non-Gaussian noise environments,

   (c) a detailed analysis of the asymptotic properties of the output of a matched myriad filter, and

   (d) comparison of the proposed detection scheme with a ZMNL-based scheme.

   Publication: V. G. Chavali and C. R. C. M. da Silva, "Detection of digital amplitude-phase modulated signals in symmetric alpha-stable noise," *IEEE Trans. Commun.*, accepted for publication.

3. Chapter 5: Development of a likelihood-based technique for classification of digital amplitude-phase modulated signals in time-correlated non-Gaussian noise environments. This research includes:

   (a) design of a new classifier which involves the use of a whitening filter followed by likelihood-based classification,

   (b) development of a novel H_∞ filter based-technique for estimating the whitening filter coefficients, and

   (c) performance analysis of the proposed classifier and comparison with a classifier developed for white noise conditions.

4. Appendix A: A practical collaborative spectrum sensing algorithm based on weighted energy combining is developed for Gaussian noise environments. This research includes:

   (a) proposal of a new collaborative spectrum sensing algorithm that exploits the correlation of the sensing data provided by a given radio in successive sensing intervals to keep the computational complexity of weighted energy combining at acceptable levels and

   (b) performance analysis which shows that the proposed method offers an excellent tradeoff between reliability and complexity, when compared with conventional collaborative sensing methods.

   Publication: V. G. Chavali and C. R. C. M. da Silva, "Collaborative spectrum sensing based on a new SNR estimation and energy combining method," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 4024-4029, Oct. 2011.

Chapter 2

Noise Models

In realistic environments, radios coexist with a variety of sources of electromagnetic radiation which interfere with their operation. These sources, which introduce additive noise into the system, can include atmospheric noise [9], co-channel interference in wireless networks [11], and incidental radiation from various man-made devices, such as automobile ignitions [12], elevators [15], microwave ovens [16], clocks and busses of computers [17], and LCD monitors [22], among many others. Various measurement campaigns conducted in a variety of environments concur that this additive noise is well modeled by non-Gaussian processes and in most cases exhibits impulsive characteristics [15–22].

For the development of signal processing algorithms which offer optimal or near-optimal performance in radio environments, accurate models of the additive non-Gaussian noise are necessary. In this chapter, two of the most popular models for the first-order statistics of the noise, namely the Middleton's model (Section 2.1) and the symmetric α stable model (Section 2.2), are presented. The popularity of these models stems from the fact that these are statistical-physical models, that is, they are derived by combining "appropriate physical and statistical descriptions of general (radio) noise environments" [12]. Consequently, these models can accurately represent the statistics of the noise experienced in a wide variety of environments. Additionally, in this chapter, a discussion on the available models for time-

correlation (second order statistics) of the additive noise is also presented[1] (Section 2.3).

2.1 Middleton's Model

Middleton proposed a statistical-physical approach to model the non-Gaussian noise generated by various natural and man-made sources of electromagnetic activity in realistic environments [12]. The first-order statistics, which include the probability density function (pdf) and the characteristic function, of the envelope and the instantaneous amplitude of the noise were first derived in [12] and later updated in [13]. This model has been used in telecommunications and other domains including underwater acoustics, radar, and optics (see [13] and references therein).

The main reason for the wide popularity of Middleton's model is its *canonical* nature, i.e., the analytical form of the model is invariant of the nature of the noise sources, the noise waveforms, and propagation environments [34]. Another attractive feature of this model is that its parameters represent different features of the radio environment such as the source distribution, propagation properties, and beam patterns.

In deriving this model, the following assumptions are made about the physical environment of interest[3]:

1. An infinite number of potential noise sources are assumed to be present in the environment.

2. The location and the noise emission times of the sources are assumed to be independent and Poisson distributed in space and time, respectively.

3. The noise waveforms emitted from the different sources have the same form; however, their amplitude, durations, and frequencies may be randomly distributed.

[1]First-order statistical models, like the Middleton's model and the symmetric α stable model, do not convey any information regarding the time correlation of the noise process.

[3]See [12] for a more detailed description of all the assumptions made in Middleton's model.

4. The noise process is assumed to be stationary, i.e., the underlying physical mechanisms, such as the average number of emitters and the path loss exponent, are time invariant over the period of observation.

5. The output of the receiver's front end is considered to be narrowband.

According to this model, the additive noise is classified into three broad categories, Class A, Class B and Class C. Classification into these categories is based on the response of the receiver's front end (RF/IF filtering stages) to the noise.

2.1.1 Class A noise

Class A noise was originally defined to constitute noise which "is typically narrower spectrally than the receiver in question" [12]. It was later redefined as the noise which "produces *ignorable transients* in the typical receiver" [35] meaning that the steady state response to the filtering process dominates the transients generated at the front end[4]. Effectively, this means that the noise is undistorted by the receiver front end, which is often the case when the noise has a bandwidth smaller than the receiver bandwidth. The necessary and sufficient condition for Class A noise is given by [35]

$$T_I \Delta f_R \gg 1, \tag{2.1}$$

where T_I is the duration of a typical emission from a noise source and Δf_R is the effective bandwidth of the receiver. For example, if $\Delta f_R = 6$ MHz (digital TV), then $T_I \gg 1/\Delta f_R \approx 0.2$ μs.

Let $X(t)$ represent the instantaneous amplitude of the Class A noise. According to this model, the probability density function of the normalized instantaneous amplitude x is given

[4]As stated in [34], "Strictly speaking, we must always have a 'build up' and transient decay period generated by the front-end stages of the receiver, when the incoming signal first appears, and next terminates. For Class A interference these transient periods are negligible vis-a-vis the incoming (noise) emission's duration."

by

$$p(x) = \exp(-A_A) \sum_{m=0}^{\infty} \frac{A_A^m}{m! \sqrt{4\pi\hat{\sigma}_{mA}^2}} \exp\Big(-\frac{x^2}{4\hat{\sigma}_{mA}^2}\Big), \tag{2.2}$$

where $x = \frac{X(t)}{\sqrt{\Omega_{2A}(1+\Gamma'_A)}}$ and $2\hat{\sigma}_{mA}^2 = \frac{\frac{m}{A_A}+\Gamma'_A}{1+\Gamma'_A}$.

Class A noise is completely defined by the following three parameters:

- A_A is the *overlap index* which is defined as the number of noise emissions per second times the mean duration of the typical emission (T_I). The smaller the value A_A, the more impulsive the noise is. The noise becomes more Gaussian as A_A increases,
- $\Gamma'_A \equiv \frac{\sigma_G^2}{\Omega_{2A}}$ is the *Gaussian factor*, which is defined as the ratio of the average power of the Gaussian component (σ_G^2) to the average power of the non-Gaussian component of the noise, and
- Ω_{2A} is the average power of the non-Gaussian component.

Noise emitted from various man-made devices such as dielectric heaters, soldering machines and plastic welders [12], modems [36], and microwave ovens [16] have been shown to be accurately modeled using the Middleton's Class A model . Additionally, it was shown through statistical analysis that co-channel interference in cellular and Wi-Fi networks [11] and microcell environments [37] can be well modeled using the Middleton's Class A model.

Middleton's Class A model has been extensively used to model impulsive noise in the design of various signal processing algorithms for channel estimation [33], channel equalization [38], signal detection [39], modulation classification [40], and signal demodulation [41], among others. Various techniques have been proposed for accurate estimation of the Class A model parameters [42], [43].

2.1.2 Class B noise

Class B noise is characterized by significant transients or "ringing" of the receiver [12] which generally occurs when the noise bandwidth is broader than the receiver bandwidth. For Class B, the transient response in the filtering stages of the front end dominates the steady state response leading to time broadening of the noise waveforms. In order to take into account the distortions caused by this broadening, Class B noise must be handled differently from Class A noise. Formally, the necessary and sufficient condition for Class B noise is given by [35]

$$T_I \Delta f_R \ll 1. \tag{2.3}$$

A detailed discussion on the derivation of the first-order statistics of the Class B noise model is presented in [12]. The pdf of the Class B noise does not have a single closed form expression and is represented using a pair of functions (one for small values of the noise amplitude and one for larger values). The Class B noise model has six parameters. The physical meaning of these parameters are explained in sufficient detail in [12], [44], while various (empirical and analytical) methods for their estimation are presented in [42].

The first-order statistics of the noise from natural sources like lightning discharges in the atmosphere and man-made sources such as automobile ignitions and fluorescent lights, among several others can be accurately represented using the Class B noise model [12].

2.1.3 Additional notes on Middleton's model

1. Class C noise, which is a mixture of both Class A and Class B noise, can be reduced to a Class B form for practical analysis [13], [42].

2. Very commonly, Middleton's model is associated with impulsive noise (spikes). However, it must be noted that the model is canonical in terms of waveforms, which can be continuous wave trains and structured pulses, for example, in addition to impulses.

In the context of Middleton's model, the word "impulsive" means that the probability of obtaining large instantaneous values of noise is higher when compared to Gaussian noise. Class B noise generally constitutes short duration pulses (wide bandwidth) and hence is more impulsive in nature when compared to Class A noise.

3. While the derived statistics of the Class A and Class B noise provide an insightful description about the noise processes involved, other vital properties required for the design of optimal receivers, such as the time correlation of the noise process, are not considered in Middleton's work. For this reason, a largely popular assumption made in the design of communication receivers is that the noise process is white [39]. Furthermore, as stated in [45], "by assuming independent samples, we obtain and upper bound on the performance of the truly optimum detector [receiver], since if the samples were dependent, the detector [receiver] would make use of the information contained in this dependence to reduce the interference [noise]. That is, the performance of the optimal detector [receiver] for dependent samples can be no worse than the performance of the optimum detector [receiver] for independent samples (for the same continuous detection time from which the discrete samples were taken)." A detailed discussion on this assumption of "effective independence" is provided in [45]. The independence of the noise samples has been exploited to design communication receiver structures optimized for Class A noise (see for example [33], [39], [46], among others).

4. The Class A distribution has the form of a Gaussian mixture model with infinite number of terms. Gaussian mixture models with number of terms $N \leq 4$ ($N = 2$ being the most popular) can be a good approximation to Middleton's Class A noise [25], [39], [47]. For this reason, in this dissertation, when dealing with Class A environments, an N-term Gaussian mixture distribution is used to model the non-Gaussian noise.

5. The complex nature of the Class B noise model makes the design of signal processing algorithms using this model infeasible. For this reason, as an alternative, simpler models such as the symmetric α stable distributions are commonly used in the literature.

(See Section 2.2.2 for details.)

2.2 Symmetric α Stable Model

A strong justification for the use of Gaussian distribution in modeling additive noise is the Central Limit Theorem which states that the sum of independent and identically distributed (i.i.d) random processes *with finite variances* converges to a Gaussian random process. The more powerful Generalized Central Limit Theorem states that the limiting distribution of a sum of i.i.d random processes is a stable distribution [48]. Stable distributions have tails heavier than a Gaussian distribution and also have infinite variance. These two characteristics of the stable distribution make it a popular model to represent the non-Gaussian noise which exists in realistic environments and exhibits impulsive behavior. Stable distributions have also been used to model various phenomena in areas such as physics, astrology, hydrology and biology (see [49], [50] and references therein). In this section, the focus is on symmetric α stable (SαS) distributions, which are a subclass of stable distributions with skewness parameter (β) equal to zero.

2.2.1 Model description

Under appropriate assumptions about the spatial distribution of the noise sources, their emission characteristics and the propagation conditions in a physical environment, it has been shown in [9] that the SαS distribution can be used to model the first-order statistics of the noise affecting radio receivers. As in Middleton's work, this model assumes that the noise sources are independent and Poisson distributed in space and time, and the source density and path loss model have an inverse power relationship with the distance to the receiver. However, various other assumptions are made with the aim of deriving a simpler and tractable model. A detailed analysis of the derivation and the features of this statistical-physical model are presented in [9].

The characteristic function of an SαS distribution is given by

$$\Phi(\omega) = \exp(j\delta\omega - \gamma|\omega|^{\alpha}), \tag{2.4}$$

where

- α is the *characteristic exponent* with $0 < \alpha \leq 2$. It is a measure of the heaviness of the tails of the distribution with smaller α signifying heavier tails. The SαS distribution reduces to the Gaussian and Cauchy distributions for $\alpha = 2$ and $\alpha = 1$, respectively.
- δ is known as the *location parameter*. When $1 < \alpha \leq 2$, δ is the mean; and when $0 < \alpha \leq 1$, δ is the median. When an SαS distribution is used to model non-Gaussian noise, δ is typically taken to be zero.
- γ is the *scale parameter*, which is also known as the *dispersion*. It determines the spread of the distribution around the location parameter. The dispersion plays a role analogous to that of the variance in the Gaussian distribution (since variance is not finite for SαS distributions).
- An additional parameter β known as the *skewness parameter* ($-1 \leq \beta \leq 1$) also exists for stable distributions. It is a measure of the asymmetry of the distribution. For SαS distributions, β is equal to zero.

2.2.2 Comparison with Middleton's model

1. In order to obtain a canonical model, in Middleton's derivations, considerable attention was given to details like the structure of the noise waveforms and source distributions [12]. As a result, Middleton's model, especially the Class B model, is difficult to derive and is highly complex. In contrast, the SαS models are derived from simplifying assumptions on the radio environment. More specifically, the beam patterns of the

receiver antenna and the noise sources are assumed to be non-directional and the noise sources are assumed to be isotropically distributed in space. These assumptions make the derivation easier and the model less complex.

2. As discussed earlier, Middleton's Class B model for the noise amplitude has two characteristic functions (or pdfs), one for amplitude values smaller than a threshold and one for larger values. As this threshold approaches infinity, the Class B model can be represented by a single characteristic function which has the same form as the characteristic function of the SαS distribution [13]. Therefore, the Middleton's Class B model can be approximated with an SαS distribution when the amplitude threshold tends to infinity [9]. In [13], it has been stated that Middleton's Class B model can be considered to be the sum of a SαS and a Gaussian process.

3. In a majority of the signal processing algorithms designed for Class B environments, due to the complicated form of the noise model, the impact of the Gaussian component is commonly ignored and the noise is modeled as having an SαS distribution. This is because, as stated in [51], "the effects of impulsive [SαS] component, will generally, but not always, be more detrimental than those of the finite-variance [Gaussian] component".

Various experimental studies have confirmed that the SαS distribution can accurately model impulsive noise. For example, it has recently been shown that the SαS distribution models the noise emitted from the clocks and busses of laptop and desktop computers with great accuracy [17]. The first-order statistics of the co-channel interference in multiple access systems such as femtocell networks, ad-hoc networks, among others, can be well modeled using the SαS distribution [11], [50], [52]. Atmosphere noise can also be modeled using this distribution [9].

Examples of algorithms that have been designed for SαS distributed noise processes include various source localization [53], direction of arrival estimation [54], signal detection [55], and signal demodulation techniques [56].

2.3 Time-Correlated Non-Gaussian Channels

In the previous sections, the non-Gaussian nature of the additive noise experienced by receivers in realistic environments was discussed and two well-known models for its first-order statistics were presented. With this knowledge, and the assumption that the noise samples at the output of a receiver are independent, various signal processing algorithms have been developed for optimal or near-optimal performance in non-Gaussian noise (see [23], [39], among many others). While this assumption of independence is convenient for the design of tractable algorithms, it contradicts observations made in various studies that the non-Gaussian noise exhibits time correlation (see [14–16], and [45] among others). The time correlation exists because of several reasons. For example, as observed in [15], the power radiated by non-Gaussian noise sources might not be constant in the frequency band of interest. Correlation can also be caused by the spreading of the noise waveforms due to filtering of the (wideband) noise at the receiver [57] or due to bursts of impulsive interference [16].

In these situations, receivers designed under the assumption that the noise process is white will only offer sub-optimal performance. This is because, as stated in [58], the optimal receiver structures depend on complete knowledge of the noise statistics. This includes the second-order statistics (which represent the time correlation of the noise process) in addition to the first-order statistics.

The design of optimal receivers in time-correlated non-Gaussian environments is challenging for two main reasons. First, there is a lack of statistical-physical models which accurately represent the time correlation of the non-Gaussian noise under a wide variety of environments. Second, even if these models were available, obtaining analytically tractable algorithms that take into account both the correlation and the non-Gaussian behavior of the noise process would not be an easy task.

Various models for time-correlated non-Gaussian noise have been proposed in literature which allow for the design of tractable receiver algorithms. It must be noted, however, that

most of these models are not physically motivated. The most popular models available in the literature are now discussed.

2.3.1 Models for time-correlated non-Gaussian noise

1. *Auto Regressive Moving Average (ARMA) models*:

   A correlated non-Gaussian random sequence $\{n_k\}$ can be generated by passing an i.i.d random sequence $\{e_k\}$ through an ARMA(P, Q) process,

   $$n_k = c + e_k + \sum_{j=1}^{Q} b_j e_{k-j} + \sum_{i=1}^{P} a_i n_{k-i}. \tag{2.5}$$

   One reason for the popularity of the ARMA model is its ability to closely match a variety of correlation statistics using only a few parameters [59].

   In modeling non-Gaussian noise, the i.i.d sequence $\{e_k\}$ is commonly assumed to have either a Gaussian mixture distribution [57], [60–62] or an SαS distribution [56], [63], [64], so that the first-order statistics of this model are compatible with the statistical-physical models available in the literature. For the correlation, MA [60], [63], [64], AR [57], [61], and ARMA models [62], [56] have been used.

   Various statistical signal processing algorithms for signal detection [57], [61], [63], signal demodulation [60], and direction of arrival estimation [65] in colored non-Gaussian noise environments have been proposed using this model.

2. *Multi-Dimensional Gaussian Mixture Model*:

   The pdf of an M-term N-dimensional Gaussian mixture model can be represented as

   $$f(\mathbf{n}) = \sum_{i=1}^{M} \frac{w_i}{\sqrt{(2\pi)^N}|\boldsymbol{\Sigma_i}|^{1/2}} \exp - \Big\{ \frac{1}{2} (\mathbf{n} - \boldsymbol{\mu_i})' \boldsymbol{\Sigma}_i^{-1} (\mathbf{n} - \boldsymbol{\mu_i}) \Big\}, \tag{2.6}$$

   where $\boldsymbol{\mu_i}$, $\boldsymbol{\Sigma_i}$ and w_i are the mean vector, covariance matrix, and the weights of each of

the M terms, respectively. This model matches the statistical-physical model developed for spatially correlated impulsive noise environments in [66]. An appealing feature of multi-dimensional Gaussian mixture models is that they approximate the statistics of a various other correlated non-Gaussian random processes with high fidelity provided that enough number of terms are used [67].

Techniques for signal detection [47] and diversity combining [68], among others, have been developed which use multi-dimensional Gaussian mixture models to model time-correlated non-Gaussian noise.

3. *Compound Gaussian model*:

   In this model, the noise process is represented as a product of a Gaussian (possibly correlated) random process $g(t)$ and a non-negative random process $s(t)$.

$$n(t) = s(t)g(t) \tag{2.7}$$

   The compound Gaussian model is a very generic noise model whose marginal pdf can be set to either the Middleton's Class A model or the SαS model, by appropriately choosing the random process $s(t)$ [69].

   When dealing with compound Gaussian process, it is commonly assumed that process $s(t)$ is invariant over the noise observation period, thus resulting in a Spherically Invariant Random Processes. These processes are commonly used to model colored non-Gaussian noise due to atmospheric effects [21], radar clutter echoes [70], and multi-access interference in cellular networks [71], among others.

4. Other time correlation models used in the literature include m-dependent noise [72], ϕ-mixing noise [58], [73], transformation noise [74], [75], and long range dependence [14], [76].

2.4 Conclusions

In this chapter, an analysis of various models used to characterize the non-Gaussian noise affecting radio receivers in realistic environments was presented. Two popular statistical-physical models, namely the Middleton's model and the SαS model, were discussed in detail. In addition, models for time-correlated non-Gaussian noise were also studied. The models discussed in this chapter will be used in the rest of the dissertation for the design of signal detection and modulation classification algorithms in non-Gaussian noise environments.

Chapter 3

Maximum-Likelihood Modulation Classification in Gaussian Mixture Noise Environments

Techniques for modulation classification can be divided into two broad categories, *likelihood*-based and *feature*-based [8]. Most of the classifiers developed for non-Gaussian noise conditions are feature-based; that is, they exploit modulation dependent features of the signal, such as cumulants [77], [78], for classification. While feature-based classifiers are generally easier to implement, they are sub-optimal. Likelihood-based classifiers, on the other hand, are optimal in the Bayesian sense, as they minimize the probability of classification error [2].

Different likelihood-based classifiers have been developed for the case in which the additive noise is modeled to have a Gaussian distribution. In these classifiers, techniques like the average likelihood ratio test (ALRT) [79], the generalized likelihood ratio test (GLRT) [80], and the quasi hybrid likelihood ratio test (qHLRT) [81] are used to deal with the unknown channel state and the noise distribution parameters. The reader is referred to [8] for a survey of these techniques. However, the only likelihood-based classifier available in the literature

for the non-Gaussian noise case was proposed in [40], where it was assumed that the channel state and noise distribution are known to the classifier.

In this chapter, a likelihood-based technique is proposed for the classification of digital amplitude-phase modulated signals in flat-fading channels with non-Gaussian noise. The additive noise is modeled by a Gaussian mixture distribution (with a known number of terms). Also developed is a blind algorithm for estimating the *unknown* channel state and noise distribution parameters. With these estimates, the signal is classified using a hybrid likelihood ratio test. Results are presented which show that the proposed classifier's performance approaches that of the ideal classifier with perfect knowledge of the fading state and noise distribution.

The remaining sections of this chapter are organized as follows. Section 3.1 presents the system model. The proposed hybrid likelihood ratio test-based classifier is developed in Section 3.2, while the blind technique for parameter estimation is derived in Section 3.3. The numerical results are presented in Section 3.4 and the conclusions are drawn in Section 3.5.

3.1 System Model

The data conveyed in the observed signal is assumed to be mapped onto a digital amplitude-phase constellation $\mathcal{S}_m$ (unknown to the classifier). The symbols $\{s_m[k] \in \mathcal{S}_m\}_{k=1}^{K}$ represent the received data symbols and are assumed to be independent and uniformly distributed among the N_m constellation points that define $\mathcal{S}_m$. Assuming perfect recovery of symbol timing and sampling at the symbol rate, the low-pass equivalent of the signal at the matched filter's output is

$$r[k] = \alpha s_m[k] + w[k], \qquad k = 1, 2, \ldots, K, \tag{3.1}$$

where α is a *complex* factor used to represent both the flat fading experienced by the signal and the unknown power and carrier phase of the transmitted signal. Despite the fact that α is used to represent different channel and signal parameters, this variable is referred to

as the "channel coefficient". The channel is considered to be slowly varying, and hence α is taken to be constant over observation interval. The pulse shape is assumed to be known satisfies the Nyquist intersymbol interference criterion.

In (3.1), $\{w[k]\}_{k=1}^{K}$ is a set of complex random variables that represents the additive noise. As previously discussed, the noise in most radio channels is known from experimental studies to be non-Gaussian, due in part to the impulsive nature of man-made and natural electromagnetic noise. The pdf of the noise sample $w[k]$ is therefore given by the N-term Gaussian mixture

$$p(w) = \sum_{n=1}^{N} \frac{\lambda_n}{\pi\sigma_n^2} \exp\left(-\frac{|w|^2}{\sigma_n^2}\right), \tag{3.2}$$

where λ_n is the probability that $w[k]$ is chosen from the n^{th} term in the pdf, with $0 \leq \lambda_n \leq 1$ and $\sum_{n=1}^{N} \lambda_n = 1$. The number of terms N is assumed to be known. This model is chosen because the Gaussian mixture density closely approximates Middleton's canonical Class A noise model [12], [25] and other symmetric pdfs [82]. In addition, (3.2) includes the Gaussian pdf as a special case. The random variables $\{w[k]\}_{k=1}^{K}$ are assumed to be i.i.d. – a common assumption for analytical purposes (see, for example, [23], [82–84]).

In the analysis that follows, the channel coefficient α and the noise distribution parameters $\{\lambda_n\}_{n=1}^{N}$ and $\{\sigma_n\}_{n=1}^{N}$ are modeled as *deterministic, unknown variables.*

3.2 Likelihood-based Modulation Classifier

Let H_m be the hypothesis that the modulation scheme (that is, the amplitude-phase constellation) of the received signal is $\mathcal{S}_m$, $m = 1, 2, \ldots, M$. Likelihood-based modulation classification is a composite hypothesis testing problem in which the hypothesis that maximizes the (log-)likelihood of the matched filter's output $\{r[k]\}_{k=1}^{K}$ is chosen,

$$\hat{H} = \arg\max_{H_m} \log p\left(r[1], \ldots, r[K] | H_m\right). \tag{3.3}$$

Due to the assumption that the symbols $\{s_m[k]\}_{k=1}^{K}$ and the noise samples $\{w[k]\}_{k=1}^{K}$ are two sets of i.i.d. random variables (with different statistics), the signal samples $\{r[k]\}_{k=1}^{K}$ are also i.i.d. for a given hypothesis H_m. In this case, the Total Probability Theorem can be used to show that $p\left(r[1],\ldots,r[K]|\, H_m\right) = \prod_{k=1}^{K}\sum_{i=1}^{N_m} p\left(r[k]|\, s_{mi}[k]\right) P\left(s_{mi}[k]|\, H_m\right)$ [85], where $s_{mi}[k]$ is the i^{th} constellation point of the m^{th} modulation scheme. Substituting this result and (3.2) into the log-likelihood function, and assuming that all constellation points of a modulation scheme have the same *a priori* probability, $P\left(s_{mi}[k]|\, H_m\right) = 1/N_m$, yields

$$\log\left(p\left(r[1],\ldots,r[K]|\, H_m\right)\right) = \sum_{k=1}^{K}\log\left\{\frac{1}{N_m}\sum_{i=1}^{N_m}\sum_{n=1}^{N}\frac{\lambda_n}{\pi\sigma_n^2}\exp\left(-\frac{|r[k]-\alpha s_{mi}[k]|^2}{\sigma_n^2}\right)\right\}. \quad (3.4)$$

It is seen in (3.3) and (3.4) that the classifier requires knowledge of α, $\{\lambda_n\}_{n=1}^{N}$, and $\{\sigma_n\}_{n=1}^{N}$. However, classifiers typically have no knowledge of the fading experienced by the signal and the distribution of the noise added in the channel. In this classifier, the hybrid likelihood ratio test (HLRT) is adopted, which does not require knowledge of the unknowns' pdf, as in the ALRT, and it does not fail in uniquely classifying nested constellations (such as 16-QAM and 32-QAM), as in the GLRT [8]. In the HLRT, the unknown symbols are averaged out, as in the ALRT, and α, $\{\lambda_n\}_{n=1}^{N}$, and $\{\sigma_n\}_{n=1}^{N}$ are estimated by using maximum-likelihood estimation, as in the GLRT.

As the unknowns are estimated with no knowledge of the received signal's modulation scheme, using a standard composite hypothesis testing procedure [2, Chapter 6], one set of estimates is obtained for each hypothesis (that is, for each possible modulation scheme). Let the estimates of α, $\{\lambda_n\}_{n=1}^{N}$, and $\{\sigma_n\}_{n=1}^{N}$ obtained assuming that H_m is true be denoted by α^m, $\{\lambda_n^m\}_{n=1}^{N}$, and $\{\sigma_n^m\}_{n=1}^{N}$, respectively. Using this definition, the final form of the proposed classifier is

$$\hat{H} = \arg\max_{H_m}\sum_{k=1}^{K}\log\left\{\frac{1}{N_m}\sum_{i=1}^{N_m}\sum_{n=1}^{N}\left\{\frac{\lambda_n^m}{\pi\left(\sigma_n^m\right)^2}\ \exp\left(-\frac{|r[k]-\alpha^m s_{mi}[k]|^2}{\left(\sigma_n^m\right)^2}\right)\right\}\right\}. \quad (3.5)$$

3.3 Estimation of the Unknown Channel and Noise Distribution Parameters

The EM algorithm is an iterative method for the numerical evaluation of maximum-likelihood estimates [86]. This algorithm is very useful in problems where the joint use of observed data $\boldsymbol{o}$ (in this case, $\boldsymbol{o} = \{r[k]\}_{k=1}^{K}$) and *hidden* data $\boldsymbol{h}$ (to be defined) simplifies the maximization of the likelihood function. Areas of application include speech recognition and the estimation of mixture distribution parameters, among others. The observed data along with the hidden data constitute the complete data $\boldsymbol{c}$, so that $\boldsymbol{c} = \{\boldsymbol{o}, \boldsymbol{h}\}$.

The EM algorithm consists of two steps. First, in the expectation step (E-step), the expectation of the log-likelihood of the complete data given the observed data is evaluated. Second, in the maximization step (M-step), new estimates of the unknowns are obtained by maximizing the expectation computed in the E-step. The EM algorithm has the property that the log-likelihood function of the observed data increases monotonically with each iteration. However, it does not guarantee convergence to the global maximum when the log-likelihood has multiple local maxima, with the point of convergence depending on the initial estimates. Detailed discussions of the convergence of the EM algorithm can be found in [87] and [88].

A new set of variables $\{y[k]\}_{k=1}^{K}$ is introduced, where $y[k]$ is associated with the k^{th} signal sample $r[k]$. The value of $y[k]$ indicates the term of the mixture which models the noise component of $r[k]$; thus, for an N-term mixture distribution, $y[k] \in \{1, 2, ..., N\}$. For example, using (3.2), $y[k] = 1$ means that the noise component of $r[k]$ is a zero-mean Gaussian with variance σ_1^2. Using this definition, the hidden data is taken to be $\boldsymbol{h} = \{y[k], s_m[k]\}_{k=1}^{K}$, where $s_m[k]$ is defined in (3.1). Therefore, the complete data is given by $\boldsymbol{c} = \{r[k], y[k], s_m[k]\}_{k=1}^{K}$. (The hidden data is also often defined as $\boldsymbol{h} = \{y[k]\}_{k=1}^{K}$ or $\boldsymbol{h} = \{s_m[k]\}_{k=1}^{K}$ – examples include [30], [89].)

Let $\boldsymbol{\theta} = \{\alpha, \lambda_1, \lambda_2, \ldots, \lambda_N, \sigma_1, \sigma_2, \ldots, \sigma_N\}$ be the vector of variables to be estimated, and

$\boldsymbol{\theta^p} = \{\alpha^p, \lambda_1^p, \lambda_2^p, \ldots, \lambda_N^p, \sigma_1^p, \sigma_2^p, \ldots, \sigma_N^p\}$ be the current estimate of $\boldsymbol{\theta}$. The proposed algorithm for the estimation of the channel state and noise distribution parameters is obtained as follows:

1) E-step. The conditional expectation of the log-likelihood of $\boldsymbol{c}$ is defined as [86]

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta^p}|H_m) = E[\log(p(\boldsymbol{c}|\boldsymbol{\theta}, H_m))|\boldsymbol{o}, \boldsymbol{\theta^p}, H_m] = \sum_{\boldsymbol{h}} \log(p(\boldsymbol{c}|\boldsymbol{\theta}, H_m)) P(\boldsymbol{h}|\boldsymbol{o}, \boldsymbol{\theta^p}, H_m). \quad (3.6)$$

Recalling that the variables $\{r[k]\}_{k=1}^K$, $\{y[k]\}_{k=1}^K$, and $\{s_m[k]\}_{k=1}^K$ are sets of i.i.d. random variables (with different statistics), $\log(p(\boldsymbol{c}|\boldsymbol{\theta}, H_m))$ in (3.6) can be written as

$$\begin{aligned}
\log(p(\boldsymbol{c}|\boldsymbol{\theta}, H_m)) &= \log\left(\prod_{k=1}^{K} p\left(r[k], y[k], s_m[k]|\boldsymbol{\theta}, H_m\right)\right) \\
&= \sum_{k=1}^{K} \log\left(p(r[k]|y[k], s_m[k], \boldsymbol{\theta}) \times P(y[k], s_m[k]|\boldsymbol{\theta}, H_m)\right) \\
&= \sum_{k=1}^{K} \log\left(\frac{\lambda_{y[k]}}{N_m}\, p(r[k]|y[k], s_m[k], \boldsymbol{\theta})\right) \qquad (3.7)
\end{aligned}$$

where, using the model presented in Section 3.1, $P(y[k], s_m[k]|\boldsymbol{\theta}, H_m) = P(y[k]|\boldsymbol{\theta}) P(s_m[k]|H_m)$ and $P(s_m[k]|H_m) = 1/N_m$ for $s_m[k] \in \mathcal{S}_m$. The probability that the noise component of $r[k]$ is modeled by the $y[k]^{th}$ term of the mixture distribution, $P(y[k]|\boldsymbol{\theta})$, is denoted by $\lambda_{y[k]}$.

The probability $P(\boldsymbol{h}|\boldsymbol{o}, \boldsymbol{\theta^p}, H_m)$ in (3.6) is obtained by using Bayes' Theorem, and is given by

$$P(\boldsymbol{h}|\boldsymbol{o}, \boldsymbol{\theta^p}, H_m) = \prod_{i=1}^{K} P(y[i], s_m[i]|r[i], \boldsymbol{\theta^p}, H_m) = \prod_{i=1}^{K} \frac{\lambda_{y[i]}^p}{N_m} \frac{p(r[i]|y[i], s_m[i], \boldsymbol{\theta^p})}{p(r[i]|\boldsymbol{\theta^p}, H_m)}. \quad (3.8)$$

Substituting (3.7) and (3.8) in (3.6) and rearranging the summations, we have

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta^p}|H_m) = \sum_{k=1}^{K} \sum_{\boldsymbol{h}} \left(\log \left(\frac{\lambda_{y[k]}}{N_m}\, p(r[k]|y[k], s_m[k], \boldsymbol{\theta}) \right) \prod_{i=1}^{K} \frac{\lambda^p_{y[i]}}{N_m} \frac{p(r[i]|y[i], s_m[i], \boldsymbol{\theta^p})}{p(r[i]|\boldsymbol{\theta^p}, H_m)} \right). \tag{3.9}$$

The l^{th} term in the first summation of (3.9) can be simplified as

$$\begin{aligned} T_l &= \sum_{y[l]} \sum_{s_m[l]} \left(\log \left(\frac{\lambda_{y[l]}}{N_m}\, p(r[l]|y[l], s_m[l], \boldsymbol{\theta}) \right) \frac{\lambda^p_{y[l]}}{N_m} \frac{p(r[l]|y[l], s_m[l], \boldsymbol{\theta^p})}{p(r[l]|\boldsymbol{\theta^p}, H_m)} \right. \\ &\quad \left. \times \sum_{\{y[k]\}_{k=1}^{K} - y[l]} \sum_{\{s_m[k]\}_{k=1}^{K} - s_m[l]} \prod_{i=1, i\neq l}^{K} \frac{\lambda^p_{y[i]}}{N_m} \frac{p(r[i]|y[i], s_m[i], \boldsymbol{\theta^p})}{p(r[i]|\boldsymbol{\theta^p}, H_m)} \right). \end{aligned} \tag{3.10}$$

Noticing that $\sum_{y[i]} \sum_{s_m[i]} P(y[i], s_m[i]|r[i], \boldsymbol{\theta^p}, H_m) = 1$, (3.9) reduces to

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta^p}|H_m) = \sum_{k=1}^{K} \sum_{y[k]} \sum_{s_m[k]} \log \left(\frac{\lambda_{y[k]}}{N_m}\, p(r[k]|y[k], s_m[k], \boldsymbol{\theta}) \right) \frac{\lambda^p_{y[k]}}{N_m} \frac{p(r[k]|y[k], s_m[k], \boldsymbol{\theta^p})}{p(r[k]|\boldsymbol{\theta^p}, H_m)}. \tag{3.11}$$

From (3.1) and (3.2), the pdfs in (3.11) are given by

$$\begin{aligned} p(r[k]|y[k], s_m[k], \boldsymbol{\theta}) &= \frac{1}{\pi \sigma^2_{y[k]}} \exp\left(-\frac{|r[k] - \alpha s_m[k]|^2}{\sigma^2_{y[k]}} \right), \\ p(r[k]|y[k], s_m[k], \boldsymbol{\theta^p}) &= \frac{1}{\pi \left(\sigma^p_{y[k]} \right)^2} \exp\left(-\frac{|r[k] - \alpha^p s_m[k]|^2}{\left(\sigma^p_{y[k]} \right)^2} \right), \end{aligned}$$

and, by using the Total Probability Theorem,

$$p(r[k]|\boldsymbol{\theta^p}, H_m) = \frac{1}{N_m} \sum_{s_m[k]} \sum_{n=1}^{N} \frac{\lambda^p_n}{\pi \left(\sigma^p_n \right)^2} \exp\left(-\frac{|r[k] - \alpha^p s_m[k]|^2}{\left(\sigma^p_n \right)^2} \right).$$

2) M-step. New estimates of the unknown variables are obtained by determining

$$\boldsymbol{\theta^p} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta^p}|H_m). \tag{3.12}$$

Unfortunately, due to the non-trivial coupling between the expressions for the channel coefficient and variances of the Gaussian mixture distribution, the simultaneous update of all the variables' estimates in the M-step is difficult in this formulation. In order to alleviate this problem, a variant of the EM algorithm known as the Expectation/Conditional Maximization (ECM) algorithm is used [90]. In this algorithm, the complicated M-step is replaced by computationally simpler conditional maximization stages, while the E-step remains unchanged. As shown in [88, Chapter 5], the ECM algorithm preserves the monotone convergence property of the EM algorithm. Also, while the ECM algorithm "typically converges more slowly than the EM algorithm in terms of number of iterations" [88], this algorithm "can always be constructed to converge at the same or approximately the same rate as the EM algorithm" [91].

Using the ECM procedure, the variables to be estimated are divided into two sets: $\boldsymbol{\theta}_1 = \{\alpha\}$ and $\boldsymbol{\theta}_2 = \{\lambda_1, \ldots, \lambda_N, \sigma_1, \ldots, \sigma_N\}$. The function $Q(\boldsymbol{\theta}, \boldsymbol{\theta^p}|H_m)$ is first maximized with respect to $\boldsymbol{\theta}_1$, with $\boldsymbol{\theta}_2$ fixed at its current estimate value, $\boldsymbol{\theta}_2^p = \{\lambda_1^p, \ldots, \lambda_N^p, \sigma_1^p, \ldots, \sigma_N^p\}$. By taking the derivative of (3.11) with respect to α, the estimate $\boldsymbol{\theta}_1^p = \{\alpha^p\}$ is updated by using

$$\alpha^p = \frac{\sum_k \sum_{y[k]} \sum_{s_m[k]} \frac{r[k]s_m^*[k]}{\left(\sigma_{y[k]}^p\right)^2} \frac{\lambda_{y[k]}^p}{N_m} \frac{p(r[k]|y[k], s_m[k], \boldsymbol{\theta^p})}{p(r[k]|\boldsymbol{\theta^p}, H_m)}}{\sum_k \sum_{y[k]} \sum_{s_m[k]} \frac{|s_m[k]|^2}{\left(\sigma_{y[k]}^p\right)^2} \frac{\lambda_{y[k]}^p}{N_m} \frac{p(r[k]|y[k], s_m[k], \boldsymbol{\theta^p})}{p(r[k]|\boldsymbol{\theta^p}, H_m)}}.$$

The set $\boldsymbol{\theta}_2^p$ is then updated with the result of maximizing $Q(\boldsymbol{\theta}, \boldsymbol{\theta^p}|H_m)$ with respect to $\boldsymbol{\theta}_2$, while keeping $\boldsymbol{\theta}_1$ fixed at $\boldsymbol{\theta}_1^p$. By now taking the derivative of (3.11) with respect to λ_n and σ_n, the estimates λ_n^p and σ_n^p are updated by using, respectively,

$$\lambda_n^p = \frac{1}{K} \sum_k \sum_{s_m[k]} \frac{\lambda_n^p}{N_m} \frac{p(r[k]|y[k] = n, s_m[k], \boldsymbol{\theta^p})}{p(r[k]|\boldsymbol{\theta^p}, H_m)}$$

and

$$\sigma_n^p = \sqrt{\frac{\sum_k \sum_{s_m[k]} |r[k] - \alpha^p s_m[k]|^2 \frac{\lambda_n^p}{N_m} \frac{p(r[k]|y[k]=n, s_m[k], \boldsymbol{\theta^p})}{p(r[k]|\boldsymbol{\theta^p}, H_m)}}{\sum_k \sum_{s_m[k]} \frac{\lambda_n^p}{N_m} \frac{p(r[k]|y[k]=n, s_m[k], \boldsymbol{\theta^p})}{p(r[k]|\boldsymbol{\theta^p}, H_m)}}}.$$

As discussed in Section 3.2, one set of estimates of α, $\{\lambda_n\}_{n=1}^N$, and $\{\sigma_n\}_{n=1}^N$ is obtained for each hypothesis. Therefore, the EM-based estimator is performed for each possible modulation scheme. The obtained estimates, denoted by α^m, $\{\lambda_n^m\}_{n=1}^N$, and $\{\sigma_n^m\}_{n=1}^N$ when H_m is assumed true, are then used by the classifier, as seen in (3.5).

3.4 Numerical Results and Discussion

In the results that follow, the modulation schemes considered are BPSK, QPSK, 8-PSK, and 16-QAM. (Note that the signal constellations of the BPSK, QPSK, and 8-PSK schemes are nested [2], [8].) The modulation schemes are assumed to be equally likely. Unless otherwise noted, the amplitude of the channel coefficient α is assumed to be Rayleigh distributed, with $E[|\alpha|^2] = 2$, and its phase uniformly distributed in $(0, 2\pi]$. The number of terms in the Gaussian mixture distribution is taken to be $N = 2$. This case is widely used in the literature (see, among others, [23], [25], [82], and [84]) as a model for impulsive noise. In this case, the first and second terms of the mixture represent the thermal noise (with variance σ_1^2 and proportion λ_1) and the impulsive noise (with variance $\sigma_2^2 \gg \sigma_1^2$ and proportion $\lambda_2 = 1 - \lambda_1$) components, respectively. Unless specified, the values of λ_2 and of the ratio of the variances of the impulsive noise and thermal noise components, $\kappa = \sigma_2^2/\sigma_1^2$, are set to 0.1 and 100, respectively. The received signal-to-noise ratio (SNR) is defined as the ratio of the average received signal power and the variance of the thermal noise component[5]; that is, the received SNR is equal to $E[|\alpha|^2]E[|s_m[k]|^2]/\sigma_1^2$.

[5]The SNR is also often defined for the $N = 2$ case as the ratio of signal power to the noise power, where the noise power is equal to $(\lambda_1\sigma_1^2 + \lambda_2\sigma_2^2)$; that is, the SNR is equal to $\left(E[|\alpha|^2]E[|s_m[k]|^2]\right)/(\lambda_1\sigma_1^2 + \lambda_2\sigma_2^2)$. For the arbitrary N case ($N > 2$) there is not a well established definition for SNR. However, the two definitions used for the $N = 2$ case can be generalized to an arbitrary N. For example, the SNR could be defined as the ratio of signal power to the noise power, with the noise power being equal to $(\lambda_1\sigma_1^2 + \lambda_2\sigma_2^2 + \ldots + \lambda_N\sigma_N^2)$.

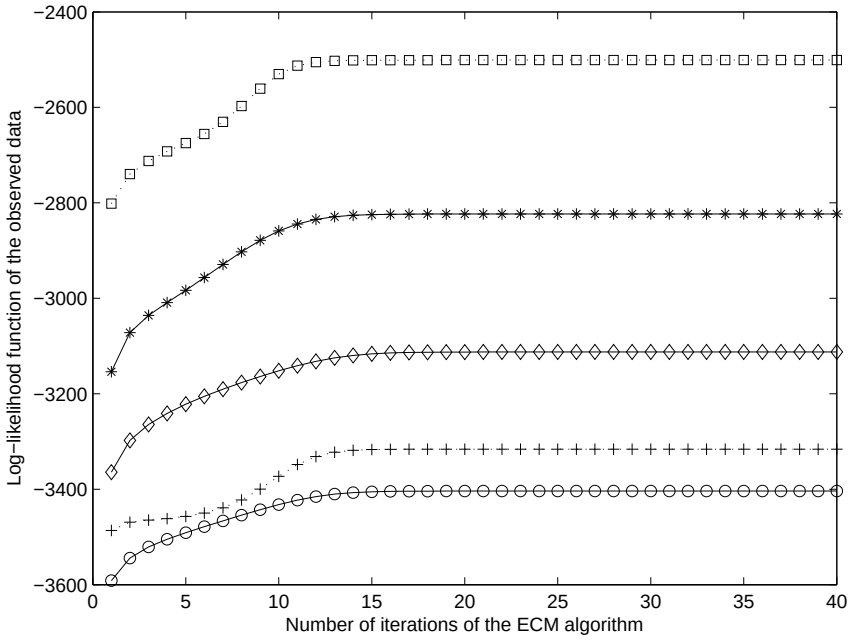

Figure 3.1: Log-likelihood function of the observed data for 8-PSK (dotted lines) and 16-QAM (solid) and various noise parameters: $\lambda_2 = 0.05,\ \kappa = 75$ (square), $\lambda_2 = 0.05,\ \kappa = 75$ (star), $\lambda_2 = 0.1,\ \kappa = 100$ (diamond), $\lambda_2 = 0.2,\ \kappa = 100$ (circle), and $\lambda_2 = 0.25,\ \kappa = 100$ (plus).

As discussed in Section 3.3, the EM and ECM algorithms have the property that the log-likelihood function of the observed data increases monotonically to some value with each iteration. This result can be seen in Fig. 3.1 for the proposed formulation given different modulation schemes and noise distribution parameters. In this implementation, parameter initialization was performed according to the guidelines given in [82]. Also, the estimator was assumed to have converged when the change in the estimates' values over successive iterations was less than or equal to 0.1%.

The performance of the proposed algorithm for the blind estimation of the channel state and noise distribution parameters is shown in Figs. 3.2 and 3.3. In Fig. 3.2, the average bias in the estimates of the amplitude and phase of the channel coefficient is shown for various numbers of observed symbols. The normalized variance of the estimates of σ_1^2 and σ_2^2 is shown in Fig. 3.3. As expected, the performance of the estimator improves with increasing

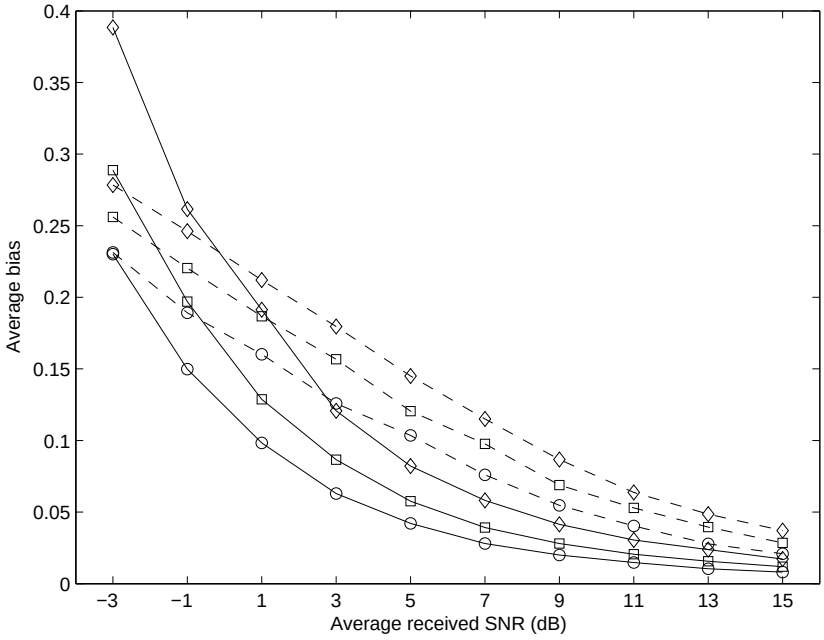

Figure 3.2: Average bias in the estimates of the phase (in radians, dashed lines) and magnitude (unitless, solid lines) of the channel coefficient α. The average bias in the estimate $\hat{X}$ of the parameter X is defined as $E[|X - \hat{X}|]$. Number of observed symbols equal to 100 (diamond), 200 (square), and 500 (circle).

the SNR and/or the number of observed symbols. It is seen that for given values of SNR and the number of observed symbols, the normalized variance of the estimate of σ_1^2 is always smaller than that of σ_2^2. This is because impulsive noise is present in only 10% of the observed symbols ($\lambda_2 = 0.1$).

The proposed classifier's performance is shown in Figs. 3.4-3.6 for various noise parameter values and fading models (Rayleigh and Rician). For reference, the performance of the ideal ("genie"-aided) classifier with perfect knowledge of the channel state and noise distribution is also plotted. This classifier is given by (3.5) with the true values (instead of noisy estimates) of α, $\{\lambda_n\}_{n=1}^N$, and $\{\sigma_n\}_{n=1}^N$ used. Similarly, the performance of the likelihood-based classifier formulated for *Gaussian* noise channels when used in a *Gaussian mixture* noise environment is also presented. This case corresponds to a situation in which there is a mismatch between the noise model used in the classifier design and the actual statistics of the noise added in

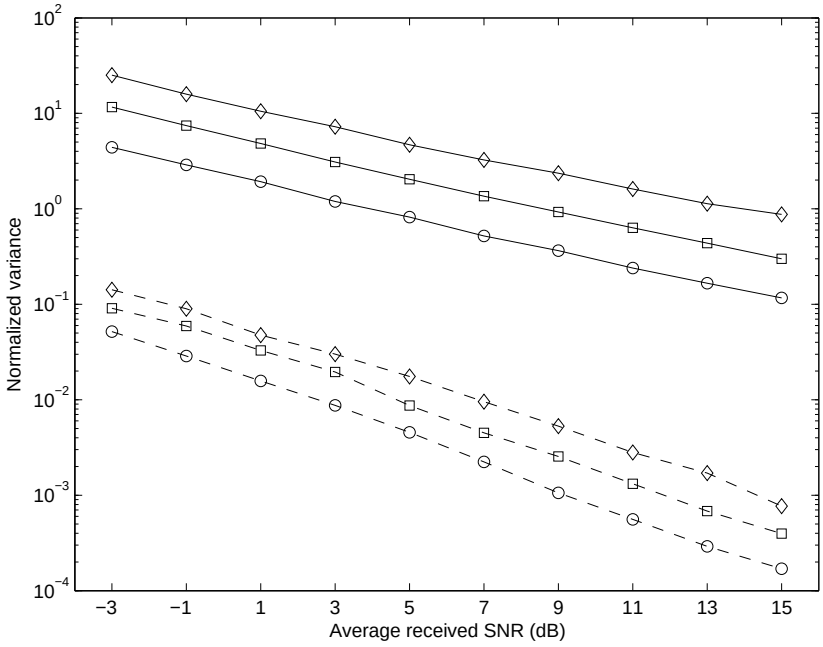

Figure 3.3: Normalized variance of the estimates of σ_1^2 (dashed lines) and σ_2^2 (solid). The normalized variance of the estimate $\hat{X}$ of the parameter X is defined as $Var[\hat{X}]/X$. Number of observed symbols equal to 100 (diamond), 200 (square), and 500 (circle).

the channel. This classifier is given by (3.5) with $N = 1$ and variance $\lambda_1\sigma_1^2 + \lambda_2\sigma_2^2$. The parameters α, $\{\lambda_n\}_{n=1}^N$, and $\{\sigma_n\}_{n=1}^N$ are also assumed to be known.

It is seen in Figs. 3.4-3.6 that the performance of the proposed classifier approaches that of the ideal classifier designed for Gaussian mixture noise in all cases. It is also seen that the gap in performance between the proposed classifier and the classifier designed for Gaussian channels is always significant, and that, by comparing Figs. 3.4 and 3.5, this gap increases as the noise becomes "more impulsive." It can also be seen that the performance improves for increasing number of symbols used in the estimation process, as the estimation accuracy increases.

Additionally, Tables 3.1 and 3.2 present the confusion matrix of the proposed classifier for noise distribution parameters. It can be observed from these tables that, for a given value of the received SNR, the BPSK scheme is most easily identifiable, while the 16-QAM

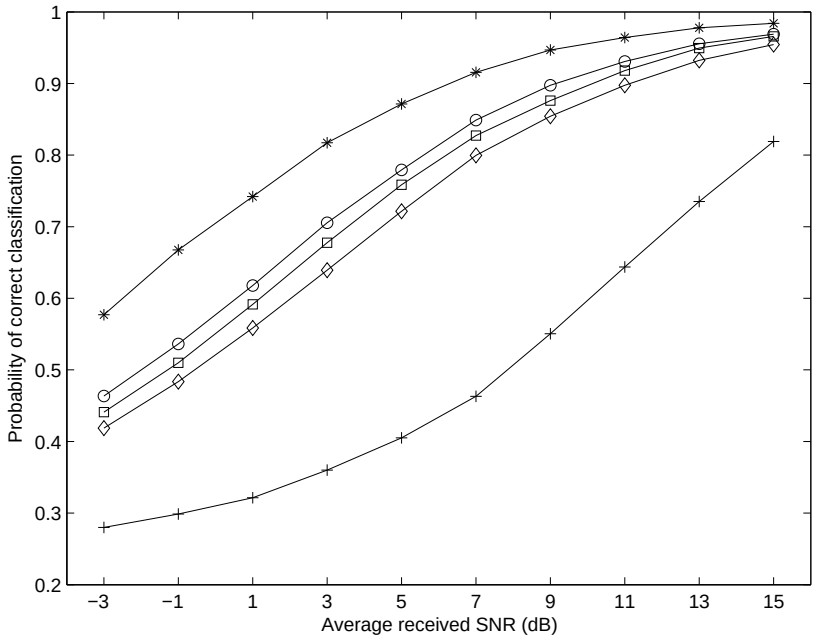

Figure 3.4: Probability of correct classification of the proposed classifier for $\lambda_2 = 0.1$ and $\kappa = 100$. The performances of ideal classifiers derived for the non-Gaussian (star) and Gaussian noise cases (plus) are also shown. Number of observed symbols for parameter estimation equal to 100 (diamond), 200 (square), and 500 (circle). Number of symbols used for classification is 500.

scheme is the most difficult to identify.

3.5 Conclusion

A new algorithm for the classification of digital amplitude-phase modulated signals in flat fading channels with non-Gaussian noise was proposed. In this derivation, the additive noise was modeled by a Gaussian mixture distribution, which is a widely used model for different noise environments, including heavy-tailed and Gaussian noise. The proposed likelihood-based classifier utilizes a variant of the EM algorithm to estimate the channel and noise parameters without the aid of training symbols. Numerical results show that the proposed classifier achieves near-optimum performance when sufficient observation symbols are avail-

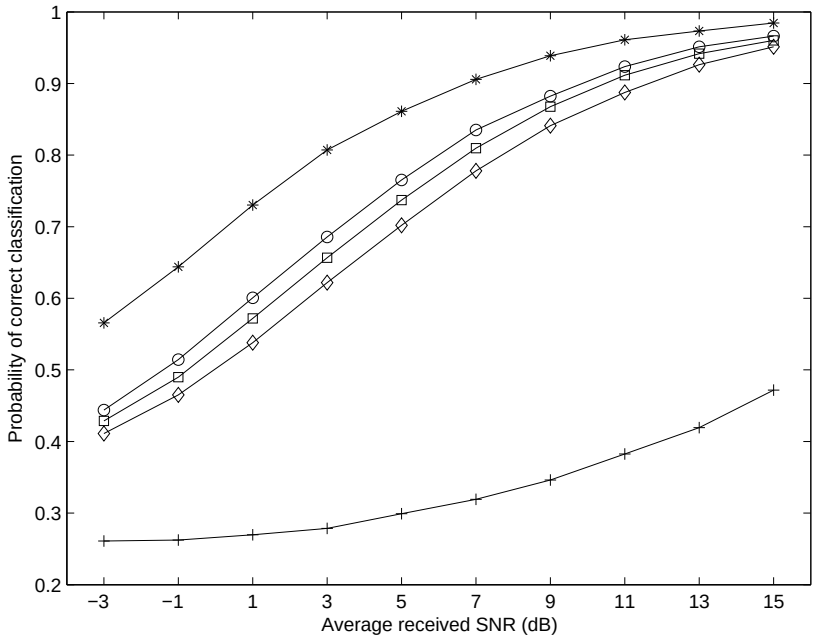

Figure 3.5: Probability of correct classification of the proposed classifier for $\lambda_2 = 0.3$ and $\kappa = 250$. The performances of ideal classifiers derived for the non-Gaussian (star) and Gaussian noise cases (plus) are also shown. Number of observed symbols for parameter estimation equal to 100 (diamond), 200 (square), and 500 (circle). Number of symbols used for classification is 500.

able for parameter estimation. It was also observed that the classifier designed for Gaussian channels performs significantly worse than the proposed classifier when Gaussian mixture noise is present.

3.6 Acknowledgments

This chapter, in part, is a reprint of the material as it appears in "Maximum-likelihood classification of digital amplitude-phase modulated signals in flat fading non-Gaussian channels," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2051-2056, Aug. 2011. The dissertation author was the primary author and Prof. Claudio da Silva directed and supervised the research which forms the basis of this chapter.

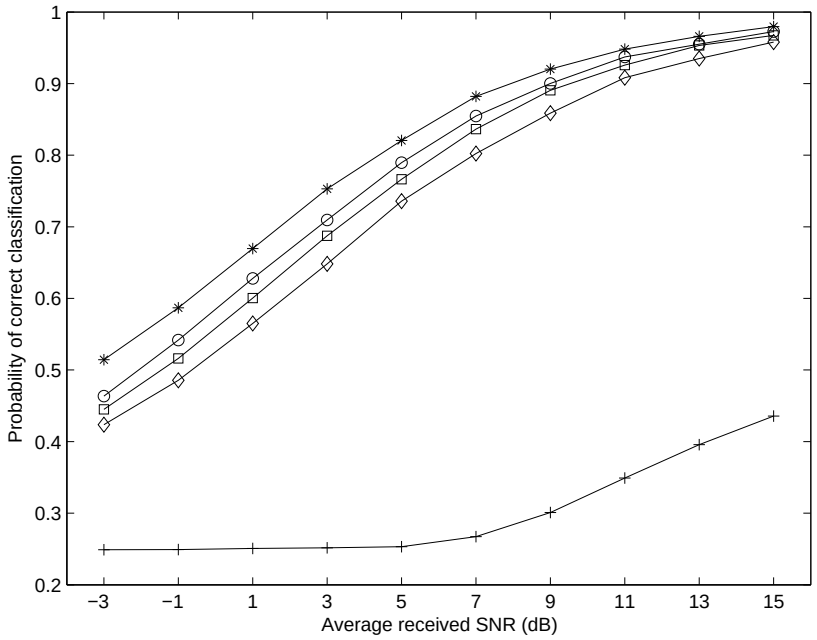

Figure 3.6: Probability of correct classification of the proposed classifier for a Rician channel with parameters $K = 0.5$ and $E[|\alpha|^2] = 3$. $\lambda_2 = 0.1$ and $\kappa = 100$. The performances of ideal classifiers derived for the non-Gaussian (star) and Gaussian noise cases (plus) are also shown. Number of observed symbols for parameter estimation equal to 100 (diamond), 200 (square), and 500 (circle). Number of symbols used for classification is 500.

Table 3.1: Confusion matrix of the proposed classifier for $\lambda_2 = 0.1$ and $\kappa = 100$. Average received SNR = 5dB

	BPSK	QPSK	8PSK	16-QAM
BPSK	0.974	0.005	0.013	0.008
QPSK	0.055	0.741	0.105	0.099
8PSK	0.060	0.051	0.835	0.054
16-QAM	0.080	0.192	0.147	0.583

Table 3.2: Confusion matrix of the proposed classifier for $\lambda_2 = 0.3$ and $\kappa = 250$. Average received SNR = 5dB

	BPSK	QPSK	8PSK	16-QAM
BPSK	0.971	0.006	0.017	0.006
QPSK	0.058	0.732	0.108	0.102
8PSK	0.063	0.061	0.820	0.056
16-QAM	0.081	0.205	0.173	0.541

Chapter 4

Signal Detection in Symmetric α Stable Noise Environments

Signal detection is a fundamental problem in signal processing with a variety of applications in wireless communications. In military applications, for example, electronic warfare and surveillance systems must reliably detect hostile signals [5]. Signal detection is also a possible enabling technique for the realization of DSA systems [31]. The design of optimal signal detection schemes requires knowledge of transmitted signal parameters, fading experienced by the signal, and the distribution of the noise added in the channel. However, in various applications, such as those previously mentioned, detectors operate with limited or no knowledge of these factors. Additionally, in most scenarios of interest, detection schemes must be reliable even at very low SNR levels. For example, the IEEE 802.22 standard requires reliable detection of signals which are 15 dB below the noise floor [31]. These issues make signal detection a challenging task.

A large number of detection schemes have been developed over the years for the case in which the additive noise is a Gaussian process [6], [7]. However, as previously discussed, the additive noise experienced in most radio channels is a non-Gaussian process. Algorithms designed for optimal performance in Gaussian noise typically perform significantly worse

when non-Gaussian noise is present [13], [17], [23], [30]. For this reason, various algorithms have been developed to effectively deal with non-Gaussian noise [24], [92], [93], [94]. One such technique, which has been gaining popularity in the design of wireless receivers, is the *myriad filter* [17], [94], [95]. Myriad filter is a non-linear operator known to be of great value for signal processing in impulsive noise environments, especially when the noise is modeled by the SαS distribution (defined in Section 2.2) [94], [95]. In this chapter, we propose a *matched* myriad filter-based solution to the problem of signal detection in SαS noise.

The main contributions of this chapter are: 1) a new five-stage scheme for the detection of digital amplitude-phase modulated signals in SαS noise; 2) new ECF-based algorithms for the estimation of noise distribution parameters in the presence of an unknown signal; 3) a detailed analysis of the asymptotic properties of the matched myriad filter output; and 4) performance analysis of the proposed detection scheme and its comparison with that of a ZMNL-based scheme.

The different sections of this Chapter are organized as follows. The system model is presented in Section 4.1. The myriad filter is introduced in Section 4.2 and the proposed detection scheme is detailed in Section 4.3. Numerical results and conclusions are presented in Section 4.4 and 4.5, respectively. Additionally, a detailed analysis of the asymptotic properties of the matched myriad filter output in presented in Appendix B.

4.1 System Model

Signal detection can be modeled as a binary hypothesis testing problem with hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ denoting the absence and presence of the signal of interest, respectively. In the proposed model, the low-pass equivalent of the received signal is given by

$$\mathcal{H}_0 \;:\; r(t) = n(t) \tag{4.1}$$

$$\mathcal{H}_1 \;:\; r(t) = s(t) + n(t), \text{ where } s(t) = \alpha_{ch} \sum_{m=-\infty}^{\infty} s_m p(t - mT - \epsilon T). \tag{4.2}$$

In (4.2), the complex random variables $\{s_m\}_{m=-\infty}^{\infty}$ are the transmitted data symbols, which are assumed to be independent and uniformly distributed among the $N_{\mathcal{S}}$ constellation points of a known digital amplitude-phase constellation $\mathcal{S}$ with $E[|s_m|^2] = 1$. The real-valued pulse shape $p(t)$ and the symbol duration T are assumed to be known. The timing offset of the received signal, due to the lack of a time reference at the receiver, is denoted by ϵT. Without loss of generality, we assume that $0 \leq \epsilon < 1$. The channel is assumed to be slowly varying and, in particular, the (deterministic and unknown) parameter $\alpha_{ch}(= \alpha_R + j\alpha_I)$ in (4.2), which represents both the flat fading experienced by the signal and the unknown power and carrier phase of the transmitted signal, is assumed to remain constant during the observation interval.

As previously discussed, the noise experienced in most radio channels is non-Gaussian, due to the impulsive nature of man-made and natural electromagnetic noise. In this chapter, the additive noise is modeled using the SαS distribution, which is known to accurately model the amplitude of noise processes that produce "significant transients at the receiver" and/or that are spectrally wider than the receiver's front-end [13]. It has been shown that co-channel interference [11] and the radio frequency interference generated by clocks and busses of computers [17] are well modeled using the SαS distribution.

The SαS distribution is defined by its characteristic function, since a closed form expression for its probability density function (pdf) is not always available. The characteristic function $\Phi_\eta(\omega)$ of an SαS random variable η is given by [9]

$$\Phi_\eta(\omega) = E[\exp(j\omega\eta)] = \exp(-j\omega\delta - \gamma|\omega|^\alpha), \tag{4.3}$$

where α, γ, and δ are the characteristic exponent, dispersion, and location, respectively. For a more detailed description of these parameters, the reader is referred to Section 2.2.

It is assumed that the noise process $n(t)$ is white, a common assumption made for analytical purposes (see [23], [94], [96], [97], among others). Additionally, as the complex baseband model is used, we assume that the samples of the noise process $n(t)$, denoted by $n = n_R + jn_I$,

have a bivariate isotropic SαS distribution whose characteristic function is given by [9, Ch. 9]

$$\Phi_n(\omega_1, \omega_2) = E[\exp(j(\omega_1 n_R + \omega_2 n_I))] = \exp\left(-\gamma(\omega_1^2 + \omega_2^2)^{\frac{\alpha}{2}}\right). \tag{4.4}$$

4.2 The Myriad Filter

The design of optimal receivers is challenging, if not impractical, when the noise is SαS distributed. This is mainly due to the fact that this distribution does not have a closed form pdf (except when $\alpha = 1$ or 2). Consequently, various sub-optimal algorithms have been proposed in the literature. These include Cauchy receivers [24], locally optimal receivers [32], and Gaussian-tailed ZMNLs [92]. In addition, *myriad filters* are becoming increasingly popular because of certain optimality properties they possess in SαS environments [98]. These filters are based on the concept of sample myriad. The sample myriad is an M-estimator for the location δ of a sequence of independent and identically distributed (i.i.d.) samples $\{x_i\}_{i=1}^N$, and is given by [95]

$$\beta_K = \arg\min_{\beta} \sum_{i=1}^{N} \log[K^2 + (x_i - \beta)^2]. \tag{4.5}$$

As seen in (4.5), the sample myriad has a freely tunable parameter K, known as the *linearity parameter*, whose value plays a significant role in the estimator's behavior. The sample myriad approaches the sample mean and the sample mode as $K \to \infty$ and $K \to 0$, respectively [98].

For SαS distributions, the sample myriad provides the maximum-likelihood estimate of the location δ when α approaches 0 and when it is equal to 1 and 2 by setting K to be 0, γ, and ∞, respectively [98]. Using this fact, the value of K can be adjusted based on the parameters α and γ to obtain near-optimal estimates for δ. Because no closed-form expression is available for the optimal value of K (that is, the value of K which minimizes the variance of the estimates), various approximate expressions have been proposed. The

expression proposed by Gonzalez and Arce in [98], for example, given by

$$K = \sqrt{\left(\frac{\alpha}{2-\alpha}\right)}\gamma^{1/\alpha}, \tag{4.6}$$

has been shown to "consistently provide efficient results in a variety of conditions" [95]. Alternatively, the expression

$$K = \left(-0.66 + 0.44e^{1.28\alpha} + 7.62 \times 10^{-34}e^{39.24\alpha}\right)\gamma^{1/\alpha}, \tag{4.7}$$

proposed in [99], has been shown to be more accurate than (4.6), especially for $\alpha < 1$ and/or $\gamma > 1$ [100]. It is important to note that in order to use either of these expressions, the parameters α and γ of the underlying SαS distribution must either be known or estimated[6].

In the development of robust signal processing techniques for impulsive noise environments, by relying on a generalized formulation of the myriad, the concept of myriad filters has been extended to matched myriad filters, weighted myriad filters, and scaled weighted myriad filters [96], [102], [103]. Matched myriad filters, for example, developed for use in digital communication receivers, take into account the pulse shape of the transmitted signal [96]. Matched myriad filters are more robust than conventional matched filters, which are known to *show significant performance degradation in the presence of impulsive noise* [96], [104]. The matched myriad filter performs the non-linear operation given by [96]

$$\beta_K = \arg\min_{\beta} \sum_{i=1}^{N} \log\left[K^2 + |p_i|^2\left|\frac{r_i}{p_i} - \beta\right|^2\right]. \tag{4.8}$$

In (4.8), $\{r_i\}_{i=1}^N$ and $\{p_i\}_{i=1}^N$ are samples of the received signal $r(t)$ and of the pulse shape $p(t)$, respectively, where $r_i = r(t)|_{t=iT_s}$, $p_i = p(t)|_{t=iT_s}$, and T_s is the sampling period.

Further details on the matched myriad filter are given in Section 4.3.2. In addition, we

[6]When the noise distribution is not known, or when there is a mismatch between the values of the SαS distribution parameters used in the algorithm design and their actual values, the parameter K can be obtained by using auxiliary measures like median absolute deviation and percentile coefficient of kurtosis. The reader is referred to [101], and references therein, for details.

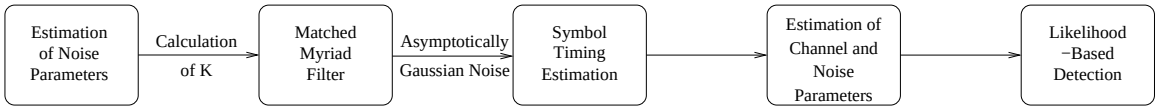

Figure 4.1: Proposed five-stage detection scheme.

derive important asymptotic properties of this filter in Appendix B, where we show that the matched myriad filter's output is, or can be approximated to be, asymptotically Gaussian.

4.3 Proposed Detection Scheme

In this section, a five-stage matched myriad filter-based scheme for the detection of digital amplitude-phase modulated signals in SαS noise is presented. As depicted in Fig. 4.3, in the first stage, the unknown parameters α and γ of the SαS distributed noise are estimated using an ECF-based algorithm. The received signal is passed through a matched myriad filter in the second stage by using the parameter K calculated from the estimated values of α and γ. The filtered signal is then used in the third stage to calculate the timing offset ϵ. After this stage, as shown in this section, the considered problem reduces to detecting a distorted (scaled and rotated) digital signal in noise that can be approximated as i.i.d. Gaussian. Following standard composite hypothesis testing procedures, unknown channel and noise distribution parameters are estimated (fourth stage) and using these estimates, a likelihood ratio test is used (fifth stage) to make a decision on the presence of the signal of interest. In the sub-sections that follow, we present and analyze the individual stages of the proposed detection scheme.

4.3.1 Estimation of noise distribution parameters

Standard parameter estimation algorithms typically cannot be used when the noise is SαS distributed. Maximum-likelihood estimators, for example, are infeasible due to the lack of a closed form expression for the noise pdf. Also, as the moments E$[|n|^m]$ of an SαS random

variable n are non-existent when $m \geq \alpha$ [9], standard method-of-moments estimators may not be applied directly to this case. For this reason, alternative algorithms, such as those based on ECFs [93] and on the concepts of covariation [9], fractional lower order statistics [97], and normalized cumulants, moments, and cross-correlations [105], are typically used.

The problem of estimating the parameters of an SαS distribution has been extensively studied for the case in which the observed signal has an SαS distribution, such as the signal under the $\mathcal{H}_0$ hypothesis given by (4.1). In this case, algorithms based on sample fractiles [100], [106], extreme value theory [107], and ECFs [108], for example, have been proposed. However, to the best of the authors' knowledge, estimation of the SαS distribution parameters when the observed signal is the sum of an SαS distributed signal with a non-SαS distributed signal, such as the signal under the $\mathcal{H}_1$ hypothesis given by (4.2), has not been addressed. In this sub-section, *new* ECF-based algorithms are developed for this case. ECF-based estimators are chosen because the characteristic function of an SαS random variable has a simple closed-form expression and due to the fact that these estimators can exploit the independence of $s(t)$ and $n(t)$.

For a sequence of i.i.d. samples $\{x_i\}_{i=1}^{N_{ECF}}$, the ECF $\hat{\Phi}_x(\omega) = \frac{1}{N_{ECF}} \sum_{i=1}^{N_{ECF}} \exp(j\omega x_i)$, is an unbiased and low-complexity estimator of the characteristic function $\Phi_x(\omega) = E[\exp(j\omega x)]$ [109]. The idea behind ECF-based estimators is the minimization of a distance measure between $\hat{\Phi}_x(\omega)$ and $\Phi_x(\omega)$ [109]. In our approach, we adopt the discrete ECF method, where $\hat{\Phi}_x(\omega)$ is computed at discrete values of ω and then fitted to $\Phi_x(\omega)$ using a least squares technique [108]. We now present ECF-based estimators for α and γ under the $\mathcal{H}_0$ and $\mathcal{H}_1$ hypotheses.

Estimation of noise distribution parameters - $\mathcal{H}_0$ case

In this case, the samples $\{r_i\}_{i=1}^{N_{ECF}}$, where $r_i = r(t)|_{t=iT}$ and $r(t)$ is given by (4.1), are i.i.d. with characteristic function given by (4.4). The ECF of the complex samples $\{r_i =$

$r_{iR} + jr_{iI}\}_{i=1}^{N_{ECF}}$ is computed as [109]

$$\hat{\Phi}_r(\omega_1, \omega_2) = \frac{1}{N_{ECF}} \sum_{i=1}^{N_{ECF}} \exp(j(\omega_1 r_{iR} + \omega_2 r_{iI})). \tag{4.9}$$

The estimates of α and γ are obtained by first calculating the ECF of the received signal at discrete values $\{\omega_{1q}, \omega_{2q}\}_{q=1}^{Q}$ and then solving the system of equations given by

$$\log\left(-\log\left(\hat{\Phi}_r\left(\omega_{1q}, \omega_{2q}\right)\right)\right) = \log(\gamma) + \frac{\alpha}{2}\log\left(\omega_{1q}^2 + \omega_{2q}^2\right), \quad q = 1, 2, 3, .., Q, \tag{4.10}$$

using the linear least squares technique presented in [110, Ch. 8].

Estimation of noise distribution parameters - $\mathcal{H}_1$ case

Exploiting the independence of $s(t)$ and $n(t)$ in (4.2), the characteristic function of $r(t)$ is given by

$$\Phi_r(\omega_1, \omega_2) = \Phi_s(\omega_1, \omega_2)\Phi_n(\omega_1, \omega_2), \tag{4.11}$$

where $\Phi_s(\omega_1, \omega_2)$ is the characteristic function of $s(t)$ and $\Phi_n(\omega_1, \omega_2)$ is given by (4.4). It must be noted that $\Phi_s(\omega_1, \omega_2)$ is *unknown* to the receiver and is a function of the modulation scheme, pulse shape, symbol duration, and the unknown parameters α_{ch} and ϵ. In the following, ECF-based estimators are presented for rectangular and square root raised cosine (SRRC) pulse shapes.

Rectangular pulse shape: In this case, sampling the received signal $r(t)$, given by (4.2), at symbol rate results in an i.i.d. sequence $\{r_i\}_{i=1}^{N_{ECF}}$ where $r_i = s_i^{rect} + n_i$, $s_i^{rect} = \alpha_{ch} s(t)|_{t=iT} = \alpha_{ch} s_i$, and $n_i = n(t)|_{t=iT}$. For rectangular pulses, the samples s_i^{rect} are not a function of ϵ and there is no inter-symbol interference (ISI). As the symbols s_i are assumed to be equally likely, the characteristic function of s_i^{rect} is given by

$$\Phi_s^{rect}(\omega_1, \omega_2) = \frac{1}{N_{\mathcal{S}}} \sum_{s_i = s_R + js_I \in \mathcal{S}} \exp\left[j(\omega_1(\alpha_R s_R - \alpha_I s_I) + \omega_2(\alpha_R s_I + \alpha_I s_R))\right]. \tag{4.12}$$

As seen in (4.12), $\Phi_s^{rect}(\omega_1, \omega_2)$ is a function of the modulation scheme used. For example, for QPSK signaling, it can be shown that

$$\Phi_s^{rect}(\omega_1, \omega_2) = \cos\left(\frac{\omega_1\alpha_R + \omega_2\alpha_I}{\sqrt{2}}\right)\cos\left(\frac{\omega_1\alpha_I - \omega_2\alpha_R}{\sqrt{2}}\right). \tag{4.13}$$

The characteristic function $\Phi_r^{rect}(\omega_1, \omega_2)$ is obtained by substituting (4.4) and (4.13) in (4.11). The ECF $\hat{\Phi}_r(\omega_1, \omega_2)$, given by (4.9), is calculated at different frequency values, and the unknown parameters α, γ, and α_{ch} are obtained by solving the system of equations

$$\hat{\Phi}_r(\omega_{1q}, \omega_{2q}) = \cos\left(\frac{\omega_{1q}\alpha_R + \omega_{2q}\alpha_I}{\sqrt{2}}\right)\cos\left(\frac{\omega_{1q}\alpha_I - \omega_{2q}\alpha_R}{\sqrt{2}}\right)\exp\left(-\gamma\left(\omega_{1q}^2 + \omega_{2q}^2\right)^{\frac{\alpha}{2}}\right), \tag{4.14}$$

for $q = 1, 2, 3..., Q$, using a non-linear least squares technique [110, Ch. 8]. While (4.14) is valid only for QPSK, it is straightforward to extend the proposed approach to other modulation schemes.

SRRC pulse shape - Exact method: When the SRRC pulse shape is used, the estimation problem becomes more difficult because the samples are now a function of ϵ and of the ISI resulting from the lack of synchronization. This is evident from the expression for the information component of the received signal at $t = 0$, $s^{srrc} = s(t)|_{t=0}$, given by

$$s^{srrc} = \alpha_{ch}\sum_{m=-\infty}^{\infty} s_m p(-mT-\epsilon T) = \sum_{m=-\infty}^{\infty}[(\alpha_R s_{mR}-\alpha_I s_{mI})+j(\alpha_R s_{mI}+\alpha_I s_{mR})]p(-mT-\epsilon T),$$

where s_{mR} and s_{mI} are the real and imaginary parts of the symbol s_m. For ease of analysis, we assume that $p(t) = 0$ for $t \geq LT$ and $t \leq -LT$. This is a reasonable assumption since the tails of an SRRC pulse approach zero. To calculate the characteristic function of $s(t)$, we assume that ϵ is a random variable uniformly distributed between 0 and 1. This makes $s(t)$ in (4.2) a stationary random process, and hence its characteristic function can be calculated

by averaging over the random variables $\{s_m\}_{m=-L}^{L-1}$ and ϵ, resulting in

$$\Phi_s^{srrc}(\omega_1,\omega_2) = E\left[\exp\left(j\sum_{m=-L}^{L-1}\big(\omega_1(\alpha_R s_{mR}-\alpha_I s_{mI})+\omega_2(\alpha_R s_{mI}+\alpha_I s_{mR})\big)f(m+\epsilon)\right)\right]$$
$$=\int_0^1\prod_{m=-L}^{L-1}\left[\frac{1}{M}\sum_{s_m}\exp(j((\alpha_R\omega_1+\alpha_I\omega_2)s_{mR}+(\alpha_R\omega_2-\alpha_I\omega_1)s_{mI})f(m+\epsilon))\right]d\epsilon, \quad (4.15)$$

where $f(x)=\frac{\frac{4\beta}{\pi}[cos((1+\beta)\pi x)+\frac{sin((1-\beta)\pi x)}{4\beta x}]}{1-16\beta^2x^2}$ is the SRRC pulse. The characteristic function of the received signal $\Phi_r^{srrc}(\omega_1,\omega_2)$ is obtained by substituting (4.4) and (4.15) in (4.11).

The ECF $\hat{\Phi}_r(\omega_1,\omega_2)$ is computed from samples $\{r_i\}_{i=1}^{N_{ECF}}$ obtained through a non-uniform sampling procedure. In this procedure, the i^{th} sample r_i is obtained by sampling $r(t)$ at $t=2(i-1)LT+\epsilon_i T$, where ϵ_i is a random variable which varies independently from sample to sample and is uniformly distributed between 0 and 1. The time offset $2LT$ between successive samples ensures that they are approximately independent, while the (random) offset $\epsilon_i T$ accounts for the assumption in the derivation of (4.15) that ϵ is a uniform random variable. In the proposed algorithm, the estimates are taken to be the values that minimize the residual sum of squares $\sum_{q=1}^{Q}|\hat{\Phi}_r(\omega_{1q},\omega_{2q})-\Phi_r^{srrc}(\omega_{1q},\omega_{2q})|^2$. Due to the lack of a closed form expression for $\Phi_r^{srrc}(\omega_1,\omega_2)$, α and γ are obtained numerically by performing a grid search.

SRRC pulse shape - Alternative method: Under certain conditions, such as when there is significant ISI and/or a higher order modulation scheme is used, the samples of $s(t)$ can be approximated to have a complex Gaussian distribution. In this case, their characteristic function is given by $\Phi_s^{gauss}(\omega_1,\omega_2)=\exp\left(-\gamma_g\left(\omega_1^2+\omega_2^2\right)\right)$, where γ_g is the signal power. By substituting this equation and (4.4) in (4.11), the characteristic function of the received signal is obtained as

$$\Phi_r^{gauss}(\omega_1,\omega_2)=\exp\left(-\gamma_g\left(\omega_1^2+\omega_2^2\right)\right)\exp\left(-\gamma\left(\omega_1^2+\omega_2^2\right)^{\frac{\alpha}{2}}\right). \quad (4.16)$$

The ECF $\hat{\Phi}_r(\omega_1,\omega_2)$ is computed using the non-uniform sampling-based procedure described

for the exact method. With a closed form expression for $\Phi_r^{gauss}(\omega_1, \omega_2)$, non-linear least squares techniques, such as the one presented in [111], can be used to solve for the noise distribution parameters. While the alternative method is easier to implement when compared to the exact method, its performance is expected to be worse since it relies on an approximation.

The ECF-based estimators for α and γ for the case in which the observed signal is given by the sum of an SαS distributed signal with a non-SαS distributed signal are an important contribution of the chapter. With these estimates, an appropriate expression, such as (4.6) or (4.7), is used in order to calculate the parameter K. It is worth noting that in signal detection problems, using a standard composite hypothesis testing procedure [2], unknown parameters of the observed signal's pdf (such as α and γ in our model) are estimated *before* the receiver has made a decision on the presence of a signal. Therefore, the proposed detector obtains one set of estimates for α and γ (and, consequently, one value of K) for each hypothesis.

4.3.2 Matched myriad filter

The proposed detection scheme uses a *non-linear* matched myriad filter, which is known to be more robust than *linear* matched filters to impulsive noise [104]. Let $\{r_i^\tau\}_{i=1}^N$ denote the samples of $r(t)$ corresponding to a running window beginning at time instant τ, i.e. $r_i^\tau = r(t)|_{t=\tau+(i-1)T_s}$, where T_s is the sampling period given by $T_s = T/P$ and P is an arbitrary oversampling factor. The output of the matched myriad filter at time instant τ is calculated as

$$\hat{\theta}_N^\tau = \arg\min_\theta \sum_{i=1}^{N} \log\left[K^2 + |p_i|^2 \left|\frac{r_i^\tau}{p_i} - \theta\right|^2\right], \tag{4.17}$$

where the coefficients $\{p_i\}_{i=1}^N$ are given by $p_i = p(t)|_{t=iT_s}$. The duration of the sliding window is equal to the length of the pulse used for pulse shaping. For rectangular pulses, N is equal to P; while for SRRC pulses, N is taken to be $2LP$.

A detailed discussion on the asymptotic properties of $\hat{\theta}_N^\tau$ is presented in Appendix B.

For rectangular pulse shapes, a theoretical analysis is developed which shows that $\hat{\theta}_N^{\tau}$ is asymptotically Gaussian. For SRRC pulse shapes, a similar analysis is unfortunately mathematically intractable. However, as shown in Appendix B, based on well known results available in the literature and on our own numerical analysis, the matched myriad filter output can be approximated to be asymptotically Gaussian. Exploiting this asymptotic behavior, standard signal processing techniques designed for optimal/near-optimal performance in Gaussian noise are therefore used in the subsequent stages of the proposed detection scheme.

4.3.3 Processing after the matched myriad filter

Symbol timing synchronization

In the third stage of the proposed scheme, using the output signal of the matched myriad filter $\hat{\theta}_N^{\tau}$, the timing offset ϵ is estimated with the square timing recovery algorithm [112]. However, it must be noted that any synchronization algorithm designed for Gaussian noise can be used. The estimate $\hat{\epsilon}$ of the timing offset ϵ is given by

$$\hat{\epsilon} = \sum_{k=1}^{N_{synch}} |r_k^{my}|^2 e^{j2\pi(k-1)/P}, \tag{4.18}$$

where $r_k^{my} = \hat{\theta}_N^{\tau}|_{\tau=kT_s}$ and N_{synch} is the number of samples used to obtain the estimate.

After synchronization, the samples $\{r_k^{my}\}$ are downsampled to obtain the sequence $\{r_k^{syn}\}$ at the symbol rate. As the matched myriad filter output is, or can be approximated to be, asymptotically Gaussian, as shown in Appendix B, r_k^{syn} can be written as

$$\mathcal{H}_0 \;:\; r_k^{syn} = \eta_k^0 \tag{4.19}$$

$$\mathcal{H}_1 \;:\; r_k^{syn} = \alpha_{ch} s_k + \eta_k^1, \tag{4.20}$$

where $\{\eta_k^0\}$ and $\{\eta_k^1\}$ are sequences of zero mean i.i.d. Gaussian random variables with

unknown variances ν_0 and ν_1, respectively. Thus, the original problem has been reduced to the simpler problem of detecting a digital amplitude-phase modulated signal in Gaussian noise. A composite hypothesis testing procedure is used which involves estimation of the unknown parameters ν_0, ν_1, and α_{ch} followed by a likelihood ratio test using the estimated parameters [2].

Estimation of channel and noise parameters after filtering

In this stage, the unknown parameters ν_0, ν_1, and α_{ch} are estimated. The estimate of the parameter ν_0 is obtained as the variance of the samples $\{r_k^{syn}\}$ under the assumption that the hypothesis $\mathcal{H}_0$ is true. Likewise, assuming $\mathcal{H}_1$ to be true, the estimates of ν_1 and $|\alpha_{ch}|$ (the magnitude of α_{ch}) are obtained using a method-of-moments technique [113], while $\angle\alpha_{ch}$ (the phase of α_{ch}) is estimated using the M-power phase synchronizer [112]. It is important to note again that any other algorithm designed to estimate channel and noise parameters in Gaussian noise could be used.

Likelihood-based detection

Using the estimates $\hat{\nu}_0$, $\hat{\nu}_1$, and $\hat{\alpha}_{ch}$ of the unknown parameters ν_0, ν_1, and α_{ch}, respectively, a likelihood ratio test is used to make a decision on the presence of the signal. The decision rule is given by

$$D = \frac{\frac{1}{(\pi\hat{\nu}_1)^{N_{det}/2}} \prod_{k=1}^{N_{det}} \sum_{m=1}^{N_{\mathcal{S}}} \frac{1}{N_{\mathcal{S}}} \exp\left(-\frac{1}{\hat{\nu}_1}|r_k^{syn} - \hat{\alpha}_{ch} S_m|^2\right)}{\frac{1}{(\pi\hat{\nu}_0)^{N_{det}/2}} \exp\left(-\frac{1}{\hat{\nu}_0}\sum_{k=1}^{K}|r_k^{syn}|^2\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)}, \tag{4.21}$$

where $\{S_m\}_{m=1}^{N_{\mathcal{S}}}$ are the constellation points, N_{det} is the number of symbols used in the detection stage, and $P(\mathcal{H}_i)$ is the *a priori* probability of the hypothesis $\mathcal{H}_i$.

Table 4.1: Adjusted SNR (in dB) for different values of α, γ, and signal power.

Signal power	$\downarrow \gamma$ - $\alpha \rightarrow$	0.6	1.0	1.4	1.8
10	0.5	12.50	8.85	7.40	7.10
	1.0	2.40	2.90	3.15	3.70
	1.5	-3.50	-0.65	0.70	1.80
	2.0	-7.65	-3.00	-1.10	0.40
4	0.5	8.50	4.85	3.40	3.10
	1.0	-1.60	-1.10	-0.85	-0.30
	1.5	-7.50	-4.65	-3.30	-2.20
	2.0	-11.65	-7.00	-5.10	-3.60
1	0.5	2.50	-1.15	-2.60	-2.90
	1.0	-7.60	-7.10	-6.85	-6.30
	1.5	-13.50	-10.65	-9.30	-8.20
	2.0	-17.65	-13.00	-11.10	-9.60

4.4 Numerical Results

In the results that follow, the modulation scheme of $s(t)$ is assumed to be QPSK. The amplitude of α_{ch} is assumed to be Rayleigh distributed, and its phase uniformly distributed in $[0, 2\pi)$. The complex noise samples having a bivariate isotropic SαS distribution are generated using the technique developed in [114]. To calculate the output of the matched myriad filter, the fixed-point search algorithm [115] was used. As the conventional definition of SNR cannot be used when the noise is SαS distributed (because its variance is infinite), we use the definition proposed in [92], where the SNR is defined as the ratio of the signal power (which, for the model presented in Section 4.1, is equal to $E[|\alpha_{ch}|^2]$) to the noise power (measured after passing the noise through a data-adaptive Gaussian-tailed ZMNL). The values of this *adjusted* SNR for different noise distribution parameters and signal power levels used in this section are shown in Table 4.1.

We first present the performance of the proposed ECF-based algorithms used to estimate α and γ under hypothesis $\mathcal{H}_1$. The ECF is calculated at $Q = 15$ different frequency values,

and the number of samples used to compute the ECF, N_{ECF}, is $10,000$, unless otherwise noted. While there is no established procedure for selecting the frequencies $\{\omega_{1q}, \omega_{2q}\}_{q=1}^{Q}$ to be used, it was observed that choosing lower values (typically < 1) results in better estimates, since the variance of the ECF is smaller at these frequencies. Similar observations were made in [93]. In our numerical results, the frequency values are randomly distributed between 0 and 1.

The performance of the proposed estimator for rectangular pulses is presented in Table 4.2 for different noise distribution parameters and signal power levels. It can be observed that for fixed values of α and γ, as the signal power increases, the bias and standard deviation of $\hat{\alpha}$ and $\hat{\gamma}$ decrease for most cases shown, as expected. In a few cases, when the increase in the adjusted SNR (or signal power) values is not significant, the accuracy of the estimates remains approximately the same. Similarly, when the signal power is fixed, as γ reduces and/or α reduces, the estimation accuracy of $\hat{\gamma}$ improves. In the estimation of α, both the bias and the standard deviation of the estimate first increase as the true value of α increases, but, in some cases, the estimator actually improves after a certain value of α. This behavior was previously observed in [116] ($\mathcal{H}_0$ case). Similar trend is observed in the estimation of α for different values of γ; in this case, the estimator performance first improves but then degrades after a certain value of γ.

Tables 4.3 and 4.4 show the performance of the exact and alternative estimation methods, respectively, proposed for SRRC pulses. (Results for the alternative method are only presented for the lowest signal power because the simplifying assumption made in its derivation results in noticeable performance loss when higher adjusted SNR values are considered.) It is seen in Table 4.3 that the exact method provides good estimates for all adjusted SNR values, as expected. Also, as seen in Table 4.4, the performance of the alternative method quickly degrades as α increases. This is because the simplifying assumption has more of a deteriorating impact on the estimator performance in this case. Finally, by comparing Tables 4.3 and 4.4, it is observed that the exact method outperforms the alternative method for the majority of cases considered.

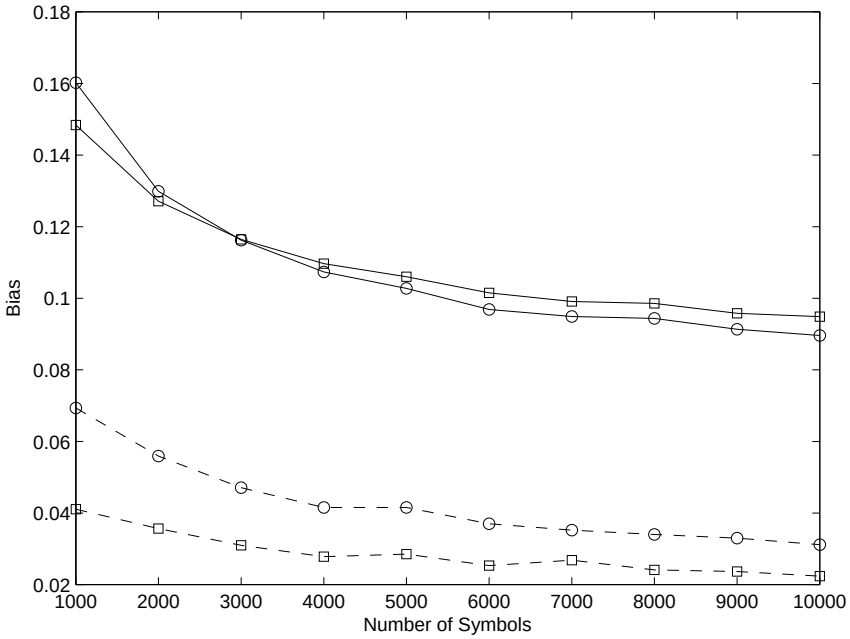

Figure 4.2: Performance of the proposed estimators for SRRC pulse shape as a function of N_{ECF}. Bias of the estimates of α (circle) and γ (square) for the exact (dashed line) and alternative (solid line) methods. α, γ, and the signal power are equal to 1.4, 0.5, and 1, respectively. Roll off factor is equal to 0.75. The bias of the estimate $\hat{X}$ of X is defined as $E[|X - \hat{X}|]$.

Figs. 4.2 and 4.3 show the bias and standard deviation, respectively, of $\hat{\alpha}$ and $\hat{\gamma}$ for different values of N_{ECF} when an SRRC pulse is used. As expected, as the value of N_{ECF} increases, the performance of both estimators (exact and alternative) improve. Also, it can be seen that the bias of the estimates obtained using the exact method approaches zero as N_{ECF} increases.

It can be observed in Tables 4.2-4.4 and in Fig. 4.2 that the proposed estimators are biased for a finite observation interval. This is due to the fact that the proposed estimators make use of non-linear least squares procedures, which might result in biased estimates when the number of samples used is finite. This result has been observed in [108] and [111], among others. As the number of samples used in the estimation process goes to infinity, it is observed that the proposed estimators for rectangular pulse shape and the exact method for the SRRC

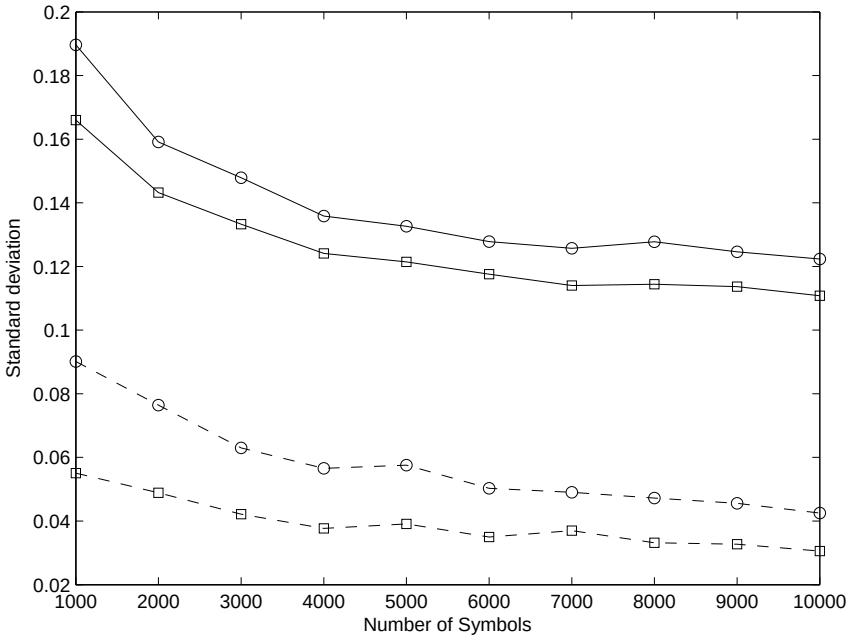

Figure 4.3: Performance of the proposed estimators for SRRC pulse shape as a function of N_{ECF}. Standard deviation of the estimates of α (circle) and γ (square) for the exact (dashed line) and alternative (solid line) methods. α, γ, and the signal power are equal to 1.4, 0.5, and 1, respectively. Roll off factor is equal to 0.75.

pulse shape, similarly to other non-linear least squares estimation procedures (see [108], [111], and [117] among others), provide unbiased estimates. Due to the approximation made in its derivation, the alternative method only asymptotically provides unbiased estimates in some cases.

The performance of the proposed five-stage detection scheme is presented in Figs. 4.4-4.8, for different pulse shapes and noise distribution parameter values. Unless otherwise noted, the oversampling factor P and the signal power are taken to be 10 and 1, respectively, and the value of K is obtained from the estimates $\hat{\alpha}$ and $\hat{\gamma}$ by using (4.6). In each of these figures, the proposed detector is compared with two other detectors:

1. An ideal detector with perfect knowledge of the noise distribution parameters α and γ and of the symbol timing offset ϵ. The performance of this detector is obtained by

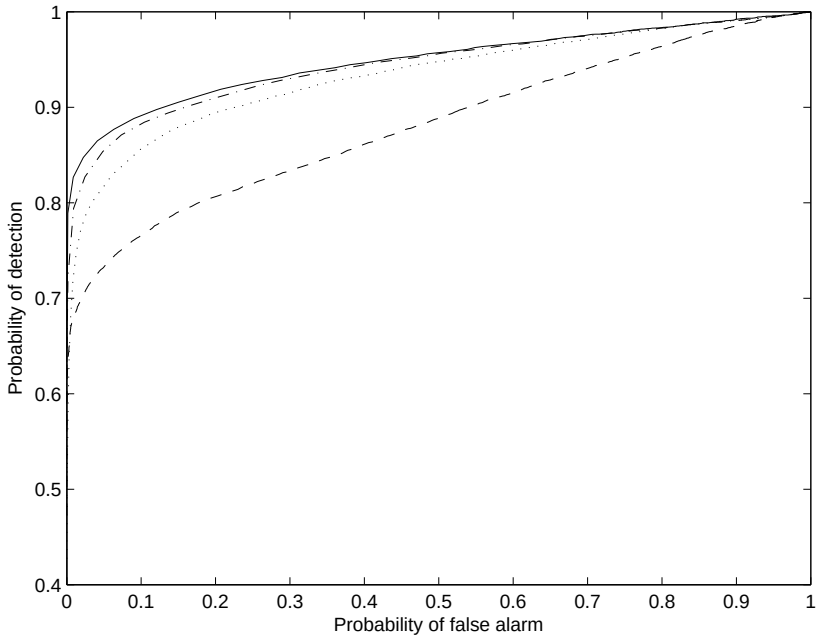

Figure 4.4: Performance of the considered detectors for α and γ equal to 1.3 and 1, respectively. Adjusted SNR $\approx$ -7dB. Rectangular pulse shape. a) ideal detector (solid line); b) proposed detector with N_{ECF} equal to 50,000 (dash-dot line); c) proposed detector with N_{ECF} equal to 10,000 (dotted line); and d) ZMNL-based detector (dashed line). The number of symbols used for timing synchronization, parameter estimation in stage IV, and detection are 500, 1,000, and 250, respectively.

replacing the estimated values of α, γ, and ϵ with their true values.

2. A ZMNL-based detection scheme. In this case, the first two stages of the proposed scheme are replaced by a ZMNL followed by a matched filter. The ZMNL used is the popular data-adaptive Gaussian-tailed ZMNL which has been utilized in a variety of signal processing applications, like direction of arrival estimation [92] and signal detection [57]. The last three stages of the ZMNL-based scheme remain the same as those of the proposed scheme.

Fig. 4.4 shows the performance of the proposed detection scheme when a rectangular pulse is considered for different values of N_{ECF}. It can be seen that the performance of the proposed scheme is very close to that of the ideal detector for both values of N_{ECF}. This is

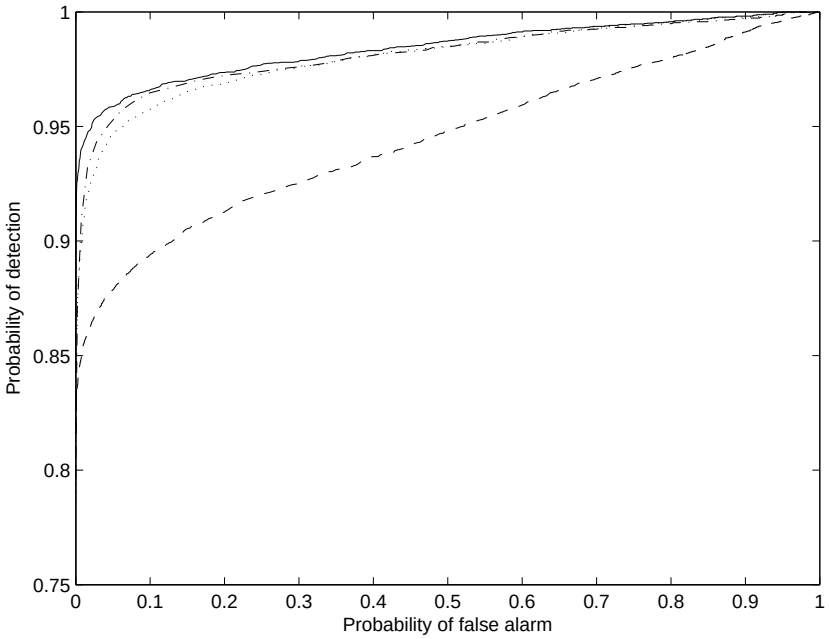

Figure 4.5: Performance of the four considered detectors for α and γ equal to 1.3 and 0.5, respectively. Adjusted SNR $\approx$ -2.5dB. SRRC pulse shape with roll off factor of 0.75. a) ideal detector (solid line); b) proposed detector - exact method (dash-dot line); c) proposed detector - alternative method (dotted line); and d) ZMNL-based detector (dashed line). The number of symbols used for timing synchronization, parameter estimation in stage IV, and detection are 500, 1,000, and 250, respectively.

due to the small bias and standard deviation of the estimates obtained with the proposed estimator. It can also be observed that the myriad filter-based detectors outperform the ZMNL-based scheme.

The performance of the proposed detector when an SRRC pulse shape is used is presented in Figs. 4.5 and 4.6 for different adjusted SNR values. The SRRC pulse is truncated to have a duration of $4T$. The performance of the proposed scheme is shown when both the exact and alternative methods are used in the first stage to estimate the noise distribution parameters. It can be seen that the myriad filter-based schemes show a noticeable performance improvement over the ZMNL-based scheme, especially for low probability of false alarm values. It is also seen that the detector gives approximately the same performance

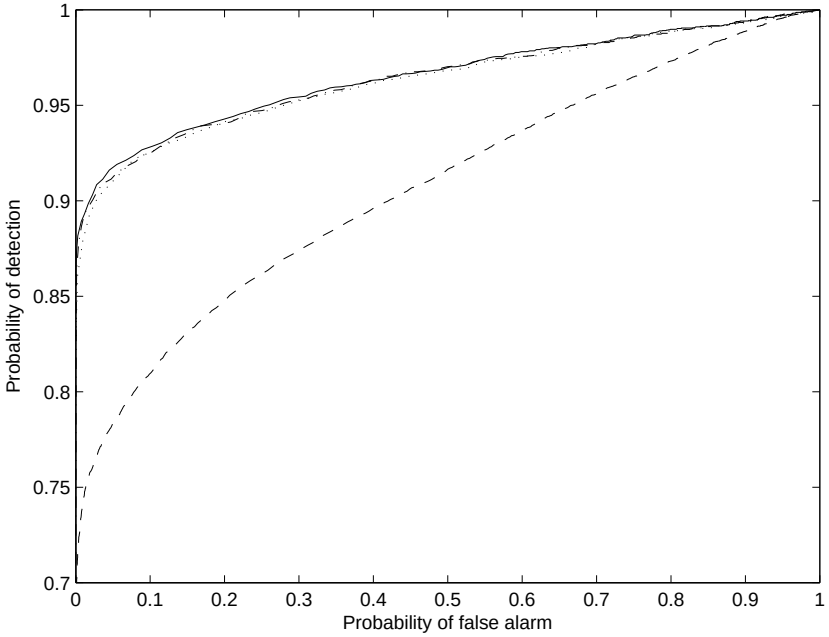

Figure 4.6: Performance of the four considered detectors for α and γ equal to 1.2 and 0.75, respectively. Adjusted SNR $\approx$ -4.5dB. SRRC pulse shape with roll off factor of 0.75. a) ideal detector (solid line); b) proposed detector - exact method (dash-dot line); c) proposed detector - alternative method (dotted line); and d) ZMNL-based detector (dashed line). The number of symbols used for timing synchronization, parameter estimation in stage IV, and detection are 500, 1,000, and 250, respectively.

for both estimators (exact and alternative methods) in the considered (low adjusted SNR value) scenarios.

Figs. 4.7 and 4.8 show the performance of the proposed detection scheme for environments in which the noise power is very high ($\gamma = 10$) and highly impulsive ($\alpha = 0.8$), respectively. To obtain the results shown in Fig. 4.7, a scaling of the frequencies $\{\omega_{1q}, \omega_{2q}\}_{q=1}^{Q}$ was necessary to account for the high value of γ. Specifically, in this case, the frequency values are randomly distributed between 0 and 0.2. (The necessity for scaling the frequencies for high values of γ is explained in [93].) Also, in the results shown in Figs. 4.7 and 4.8, the value of K was calculated by using (4.7). It is seen from Figs. 4.7 and 4.8 that the myriad filter-based schemes significantly outperform the ZMNL-based detector, thereby demonstrating the robustness of

the proposed detection scheme to different channel and noise conditions.

The improvement in performance of the proposed technique can be attributed to the near optimal performance of the matched myriad filter in the presence of SαS noise. More specifically, consider the output of the third stage of the detection scheme (after timing synchronization and sampling at symbol rate). The use of a matched myriad filter in the proposed scheme leads to signal estimates with variance very close to the optimal variance (variance of the maximum-likelihood estimator), as observed in [99]. However, this is not the case for the ZMNL-based technique, which uses a linear matched filter (high variance estimates of the signal are obtained). It is this property of the matched myriad filter which allows the proposed detection scheme to easily distinguish between the two hypotheses under consideration, thus leading to better detection performance. It must be noted that effective matched myriad filter operation requires the knowledge of K, which in turn depends on the noise distribution parameters. These parameters can be estimated using the technique developed in Section 4.3.1.

4.5 Conclusion

In this chapter, we proposed a five-stage scheme for the detection of digital amplitude-phase modulated signals in SαS distributed noise. New ECF-based algorithms for the estimation of noise distribution parameters in the presence of an additional unknown signal were also presented. The proposed detection scheme is based on the use of a matched myriad filter, which is a very effective technique for signal processing when the noise is SαS distributed. Numerical results show that the proposed detection scheme outperforms a popular ZMNL-based detection scheme. An analysis of matched myriad filters was also presented in Appendix B which shows that the output of these filters is, or can be approximated to be, asymptotically Gaussian.

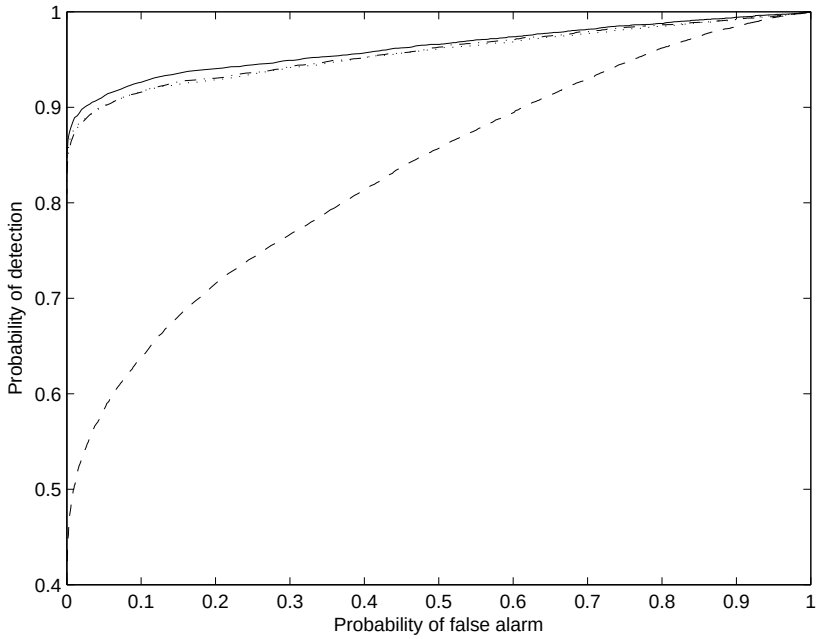

Figure 4.7: Performance of the four considered detectors for α, γ, and signal power equal to 1.1, 10, and 100, respectively. Adjusted SNR $\approx$-5dB. SRRC pulse shape with roll off factor of 0.75. a) ideal detector (solid line); b) proposed detector - exact method (dash-dot line); c) proposed detector - alternative method (dotted line); and d) ZMNL-based detector (dashed line). The number of symbols used for timing synchronization, parameter estimation in stage IV, and detection are 500, 1,000, and 250, respectively.

4.6 Acknowledgments

This chapter, in part, is a reprint of the material as it appears in "Detection of digital amplitude-phase modulated signals in symmetric alpha-stable noise," accepted for publication in *IEEE Trans. Commun.* The dissertation author was the primary author and Prof. Claudio da Silva directed and supervised the research which forms the basis of this chapter.

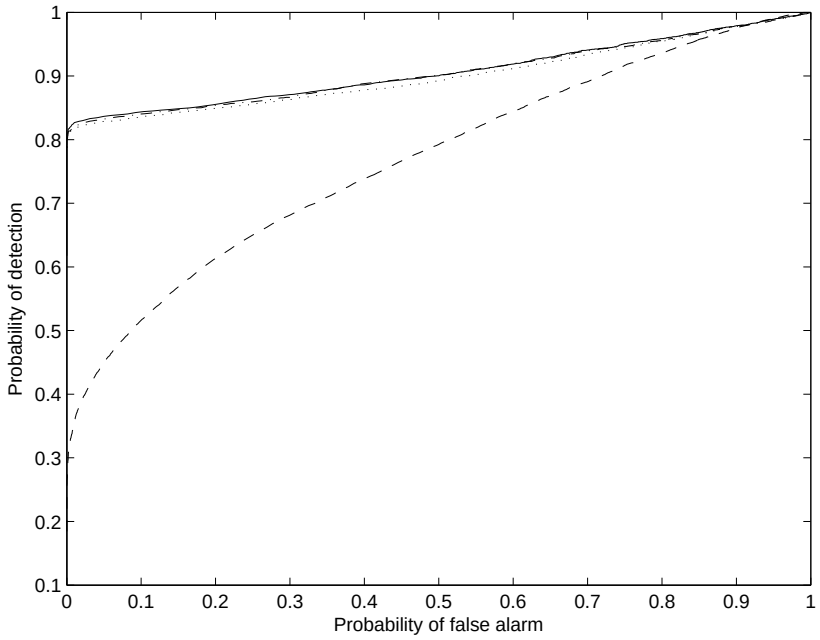

Figure 4.8: Performance of the four considered detectors for α, and γ equal to 0.8 and 1.5, respectively. Adjusted SNR $\approx$-10dB. SRRC pulse shape with roll off factor of 0.75. a) ideal detector (solid line); b) proposed detector - exact method (dash-dot line); c) proposed detector - alternative method (dotted line); and d) ZMNL-based detector (dashed line). The number of symbols used for timing synchronization, parameter estimation in stage IV, and detection are 500, 1,000, and 250, respectively. P = 18.

Table 4.2: Performance of the proposed estimator for rectangular pulse shape. (Bias, Standard deviation) of $\hat{\alpha}$ and $\hat{\gamma}$ are shown for different combinations of α, γ, and signal power. The bias of the estimate $\hat{X}$ of X is defined as $E[|X - \hat{X}|]$.

Signal power	$\downarrow \gamma$ - $\alpha \rightarrow$		0.6	1.0	1.4	1.8
10	0.5	$\hat{\alpha}$	(0.030, 0.047)	(0.036, 0.054)	(0.043, 0.064)	(0.047, 0.067)
		$\hat{\gamma}$	(0.016, 0.030)	(0.018, 0.032)	(0.022, 0.039)	(0.028, 0.050)
	1	$\hat{\alpha}$	(0.025, 0.038)	(0.030, 0.045)	(0.035, 0.051)	(0.036, 0.049)
		$\hat{\gamma}$	(0.027, 0.041)	(0.032, 0.050)	(0.039, 0.065)	(0.048, 0.079)
	1.5	$\hat{\alpha}$	(0.027, 0.043)	(0.031, 0.048)	(0.036, 0.051)	(0.036, 0.049)
		$\hat{\gamma}$	(0.044, 0.072)	(0.051, 0.081)	(0.064, 0.097)	(0.077, 0.118)
	2.0	$\hat{\alpha}$	(0.033, 0.058)	(0.035, 0.057)	(0.040, 0.059)	(0.039, 0.052)
		$\hat{\gamma}$	(0.072, 0.124)	(0.081, 0.125)	(0.100, 0.147)	(0.119, 0.169)
4	0.5	$\hat{\alpha}$	(0.032, 0.050)	(0.040, 0.059)	(0.050, 0.073)	(0.051, 0.074)
		$\hat{\gamma}$	(0.019, 0.033)	(0.024, 0.038)	(0.033, 0.052)	(0.046, 0.067)
	1	$\hat{\alpha}$	(0.029, 0.043)	(0.035, 0.052)	(0.042, 0.060)	(0.041, 0.053)
		$\hat{\gamma}$	(0.034, 0.051)	(0.043, 0.063)	(0.060, 0.085)	(0.079, 0.107)
	1.5	$\hat{\alpha}$	(0.031, 0.046)	(0.037, 0.057)	(0.044, 0.060)	(0.041, 0.050)
		$\hat{\gamma}$	(0.055, 0.077)	(0.070, 0.096)	(0.096, 0.125)	(0.124, 0.153)
	2.0	$\hat{\alpha}$	(0.038, 0.062)	(0.044, 0.068)	(0.051, 0.068)	(0.044, 0.052)
		$\hat{\gamma}$	(0.089, 0.134)	(0.111, 0.153)	(0.148, 0.181)	(0.185, 0.212)
1	0.5	$\hat{\alpha}$	(0.034, 0.050)	(0.048, 0.068)	(0.068, 0.088)	(0.071, 0.082)
		$\hat{\gamma}$	(0.021, 0.028)	(0.031, 0.041)	(0.053, 0.063)	(0.086, 0.087)
	1	$\hat{\alpha}$	(0.032, 0.045)	(0.046, 0.066)	(0.064, 0.074)	(0.059, 0.057)
		$\hat{\gamma}$	(0.039, 0.050)	(0.059, 0.073)	(0.100, 0.104)	(0.151, 0.131)
	1.5	$\hat{\alpha}$	(0.037, 0.051)	(0.054, 0.073)	(0.070, 0.075)	(0.059, 0.050)
		$\hat{\gamma}$	(0.066, 0.080)	(0.101, 0.114)	(0.162, 0.148)	(0.225, 0.168)
	2.0	$\hat{\alpha}$	(0.047, 0.060)	(0.066, 0.084)	(0.079, 0.080)	(0.063, 0.051)
		$\hat{\gamma}$	(0.109, 0.122)	(0.162, 0.163)	(0.241, 0.193)	(0.317, 0.213)

Table 4.3: Performance of the proposed estimator for SRRC pulse shape (roll off factor = 0.75) - Exact method. (Bias, Standard deviation) of $\hat{\alpha}$ and $\hat{\gamma}$ are shown for different combinations of α, γ, and signal power. The bias of the estimate $\hat{X}$ of X is defined as $E[|X - \hat{X}|]$.

Signal power	$\downarrow \gamma$ - $\alpha \rightarrow$		0.6	1.0	1.4	1.8
10	0.5	$\hat{\alpha}$	(0.033, 0.044)	(0.040, 0.051)	(0.044, 0.062)	(0.047, 0.068)
		$\hat{\gamma}$	(0.018, 0.021)	(0.018, 0.021)	(0.020, 0.023)	(0.021, 0.026)
	1	$\hat{\alpha}$	(0.022, 0.033)	(0.029, 0.040)	(0.034, 0.049)	(0.038, 0.050)
		$\hat{\gamma}$	(0.035, 0.031)	(0.039, 0.032)	(0.041, 0.038)	(0.045, 0.038)
	1.5	$\hat{\alpha}$	(0.023, 0.034)	(0.024, 0.036)	(0.030, 0.042)	(0.037, 0.052)
		$\hat{\gamma}$	(0.054, 0.046)	(0.053, 0.050)	(0.066, 0.056)	(0.063, 0.056)
	2.0	$\hat{\alpha}$	(0.026, 0.035)	(0.027, 0.040)	(0.036, 0.049)	(0.036, 0.049)
		$\hat{\gamma}$	(1.998, 0.068)	(2.003, 0.067)	(1.991, 0.084)	(2.001, 0.077)
4	0.5	$\hat{\alpha}$	(0.031, 0.045)	(0.040, 0.061)	(0.044, 0.062)	(0.032, 0.045)
		$\hat{\gamma}$	(0.018, 0.023)	(0.025, 0.046)	(0.026, 0.036)	(0.026, 0.040)
	1	$\hat{\alpha}$	(0.024, 0.039)	(0.029, 0.041)	(0.033, 0.049)	(0.029, 0.045)
		$\hat{\gamma}$	(0.030, 0.049)	(0.032, 0.046)	(0.042, 0.066)	(0.038, 0.053)
	1.5	$\hat{\alpha}$	(0.024, 0.036)	(0.021, 0.030)	(0.027, 0.036)	(0.028, 0.037)
		$\hat{\gamma}$	(0.043, 0.059)	(0.035, 0.049)	(0.048, 0.068)	(0.052, 0.069)
	2.0	$\hat{\alpha}$	(0.024, 0.037)	(0.022, 0.033)	(0.029, 0.040)	(0.026, 0.035)
		$\hat{\gamma}$	(0.053, 0.073)	(0.050, 0.069)	(0.063, 0.085)	(0.057, 0.081)
1	0.5	$\hat{\alpha}$	(0.026, 0.035)	(0.033, 0.043)	(0.031, 0.041)	(0.023, 0.035)
		$\hat{\gamma}$	(0.017, 0.026)	(0.022, 0.031)	(0.026, 0.039)	(0.027, 0.040)
	1	$\hat{\alpha}$	(0.021, 0.028)	(0.023, 0.032)	(0.024, 0.031)	(0.022, 0.031)
		$\hat{\gamma}$	(0.026, 0.035)	(0.030, 0.044)	(0.036, 0.050)	(0.034, 0.047)
	1.5	$\hat{\alpha}$	(0.015, 0.023)	(0.017, 0.022)	(0.020, 0.028)	(0.020, 0.026)
		$\hat{\gamma}$	(0.035, 0.044)	(0.033, 0.043)	(0.037, 0.052)	(0.043, 0.058)
	2.0	$\hat{\alpha}$	(0.018, 0.025)	(0.019, 0.024)	(0.021, 0.027)	(0.022, 0.030)
		$\hat{\gamma}$	(0.045, 0.059)	(0.043, 0.057)	(0.047, 0.061)	(0.051, 0.069)

Table 4.4: Performance of the proposed estimator for SRRC pulse shape (roll off factor = 0.75) - Alternative method. (Bias, Standard deviation) of $\hat{\alpha}$ and $\hat{\gamma}$ are shown for different combinations of α and γ. The bias of the estimate $\hat{X}$ of X is defined as $E[|X - \hat{X}|]$.

Signal power	$\downarrow \gamma$ - $\alpha \rightarrow$		0.6	1.0	1.4	1.8
1	0.5	$\hat{\alpha}$	(0.023, 0.029)	(0.051, 0.066)	(0.091, 0.119)	(0.229, 0.256)
		$\hat{\gamma}$	(0.029, 0.035)	(0.053, 0.070)	(0.107, 0.128)	(0.253, 0.217)
	1	$\hat{\alpha}$	(0.017, 0.022)	(0.036, 0.046)	(0.059, 0.076)	(0.144, 0.160)
		$\hat{\gamma}$	(0.042, 0.053)	(0.076, 0.100)	(0.151, 0.180)	(0.409, 0.347)
	1.5	$\hat{\alpha}$	(0.015, 0.020)	(0.029, 0.034)	(0.047, 0.059)	(0.115, 0.122)
		$\hat{\gamma}$	(0.071, 0.069)	(0.092, 0.116)	(0.186, 0.223)	(0.540, 0.449)
	2.0	$\hat{\alpha}$	(0.014, 0.019)	(0.028, 0.033)	(0.041, 0.051)	(0.105, 0.110)
		$\hat{\gamma}$	(0.071, 0.088)	(0.115, 0.147)	(0.220, 0.268)	(0.689, 0.564)

Chapter 5

Modulation Classification in Time-Correlated Non-Gaussian Noise Environments

Various studies have shown that the additive noise experienced in most radio channels is non-Gaussian and also exhibits significant time correlation [14–16], [45]. Time correlation can arise due to the narrowband filtering of uncorrelated noise at the receivers [57] and bursts of impulsive interference [16], among other reasons. Additionally, as observed in [15], the power radiated by non-Gaussian noise sources is not constant over wide bandwidths. Despite this fact, due to the complicated nature of the analysis required, no statistical-physical models are available for the second-order statistics (time correlation) of the non-Gaussian noise processes [118]. Popular statistical-physical models, including the Middleton's model [12] and the symmetric α stable model [9], characterize only the first-order statistical properties of the additive noise.

In part due to the lack of statistical-physical models for time-correlated non-Gaussian noise environments, the majority of signal processing algorithms designed for practical radio channels are developed by taking into account only the first-order statistics of the noise

process, while assuming that the noise process is white. See [17], [23], and [25], among many others. However, design of optimal detection/classification algorithms requires complete knowledge of the noise statistics, including its second-order statistics [58]. Algorithms designed under the assumption that the noise process is white will not result in optimal performance in practical scenarios where the additive noise is time-correlated [57], [14].

In this chapter, a new algorithm is presented for the classification of digital amplitude-phase modulated signals in flat-fading channels with the additive noise modeled as a time-correlated non-Gaussian random process. The Gaussian mixture distribution is used to model the first-order statistics of the additive noise, while the time correlation of the noise is modeled using an AR process. The AR model is chosen because "it can describe a large set of correlation patterns with only a few number of parameters" [59]. Additionally, in this chapter, a new technique for blind estimation of the AR parameters of the noise process using the robust H_∞ filter is also developed.

In the proposed classifier, the estimated AR parameters are used as coefficients of a whitening filter. The whitening filter is used because, as discussed in what follows, the likelihood-based classification of the whitened signal (with approximately independent noise) offers a significant reduction in complexity over the likelihood-based classification of the received signal with correlated noise. At the output of the whitening filter, estimates of the unknown fading state and noise distribution parameters are obtained using a variant of the EM algorithm and a hybrid likelihood ratio test is used to identify the modulation scheme of the received signal. Performance analysis shows that the proposed classifier outperforms the classifier developed under the assumption that the noise process is white.

The remaining sections of this chapter are organized as follows. The system model is presented in Section 5.1, while the proposed likelihood-based classifier is developed in Section 5.2. A new H_∞ filter-based technique for blind estimation of the AR parameters is developed in Section 5.3. In Section 5.4, a variant of the EM algorithm, proposed in Chapter 3, is used for estimating the state of the fading process and the Gaussian mixture

distribution parameters. Numerical results and conclusions are presented in Sections 5.5 and 5.6, respectively.

5.1 System Model

We assume that the data conveyed in the received signal is mapped onto a digital amplitude-phase constellation $\mathcal{S}_m$ (unknown to the classifier). Let $\{s_m[k]\}_{k=1}^{K}$ denote the received data symbols, assumed to be independent and uniformly distributed among the N_m constellation points that define $\mathcal{S}_m$. Assuming perfect recovery of symbol timing and sampling at the symbol rate, the low-pass equivalent of the received signal at the matched filter's output is given by

$$r[k] = \alpha s_m[k] + n[k],\ k = 1, 2, 3, ..., K. \tag{5.1}$$

In (5.1), α is a (deterministic and unknown) complex factor used to represent both the flat fading experienced by the signal and the unknown power and carrier phase of the transmitted signal. Despite the fact that α is used to represent different channel and signal parameters, we refer to this variable as the "channel coefficient". It should be noted from (5.1) that the channel coefficient is assumed to be constant over the duration of the observation window. (That is, the channel is assumed to be slowly varying.) The pulse shape used is assumed to satisfy the Nyquist intersymbol interference criterion.

Also in (5.1), $\{n[k]\}_{k=1}^{K}$ is a set of complex random variables used to represent the additive noise. As previously discussed, the additive noise experienced by most radio channels is non-Gaussian and correlated in time. In this chapter, the non-Gaussian behaviour of the noise is modeled using the Gaussian mixture distribution. This model is very popular as it closely approximates Middleton's Class A model (a statistical-physical model for man-made and natural electromagnetic noise in radio environments) [12], [25] and other symmetric distributions [82]. It has also been shown that noise from sources like microwave ovens [16], modems [36], and co-channel interference in cellular networks [11], among others, can be

accurately modeled using the Gaussian mixture distribution. This model has been used, for example, in the development of signal processing algorithms for multiuser-detection [23], channel estimation [33], and classification [40]. Additionally, an AR process is used to model the correlation among the noise samples. The AR model is one of the most popular models for correlated time-series data available in the literature since it can describe a variety of correlation structures using only a few parameters [59]. It has been widely used to model the noise correlation of typical radio propagation environments. See, for example, [57], [59], and [62]. As shown in Appendix C, the parameters of the proposed noise model can be tuned to very closely approximate the statistics of the impulsive noise obtained from the measurement campaign reported in [15], further justifying the use of this model.

The correlated noise samples $\{n[k]\}_{k=1}^{K}$ are generated by passing a sequence of i.i.d random variables $\{e[k]\}_{k=1}^{K}$ (called the driving noise) through an AR(P) filter with coefficients $\{a[i]\}_{i=1}^{P}$. That is,

$$n[k] = -\sum_{i=1}^{P} a[i]n[k-i] + e[k]. \tag{5.2}$$

The driving noise is assumed to have a N-term Gaussian mixture distribution with pdf given by

$$f(e) = \sum_{i=1}^{N} \frac{\lambda_i}{\pi\sigma_i^2} \exp\left(-\frac{|e|^2}{\sigma_i^2}\right), \tag{5.3}$$

where λ_i and σ_i are the proportion and the standard deviation of the i^{th} term, respectively. It must be noted that the pdf of the noise samples $\{n[k]\}_{k=1}^{K}$ also has the form of a Gaussian mixture distribution. This is because the AR process in a linear process and any linear transformation of Gaussian mixture random variables results in another Gaussian mixture random variable [88]. (The pdf of the process $\{n[k]\}_{k=1}^{K}$ is infact an infinite term Gaussian mixture model with the proportion and variance of each term depending on the AR parameters $\{a[i]\}_{i=1}^{P}$ and the driving noise distribution parameters $\{\lambda_i, \sigma_i\}_{i=1}^{N}$.)

In the analysis that follows, the channel coefficient α, the AR parameters $\{a[i]\}_{i=1}^{P}$, and the Gaussian mixture distribution parameters $\{\lambda_i, \sigma_i\}_{i=1}^{N}$ are modeled as *deterministic,*

unknown variables. That is, the classifier operates with no prior knowledge of the fading and the noise statistics, a reasonable assumption given that these statistics vary in both time and space. Furthermore, the number of terms of the Gaussian mixture model (N) and the order of the AR process (P) are assumed to be known[7].

5.2 Likelihood-based modulation classification

Let H_m be the hypothesis that the m^{th} modulation scheme (that is, the constellation $\mathcal{S}_m$) was transmitted. Maximum-likelihood modulation classification is a composite hypothesis testing problem in which the hypothesis that maximizes the (log-)likelihood of the received symbols $\boldsymbol{r}= \{r[k]\}_{k=1}^{K}$ is chosen,

$$H_{ML} = \arg\max_{H_m} \log\left(p\left(\boldsymbol{r}|H_m\right)\right). \tag{5.4}$$

Using the Total Probability Theorem, the likelihood function $p\left(\boldsymbol{r}|H_m\right)$ can be written as

$$p\left(\boldsymbol{r}|H_m\right) = \sum_{\boldsymbol{s}_m} p(\boldsymbol{r}|\boldsymbol{s}_m)\mathrm{P}(\boldsymbol{s}_m), \tag{5.5}$$

where $\boldsymbol{s}_m = \{s_m[k]\}_{k=1}^{K}$. Using the system model defined in Section 5.1, (5.5) can be simplified as

$$p\left(\boldsymbol{r}|H_m\right) = \frac{1}{N_m^K} \sum_{s_m[1]} \sum_{s_m[2]} .. \sum_{s_m[K]} p(\boldsymbol{r}|\boldsymbol{s}_m) = \frac{1}{N_m^K} \sum_{s_m[1]} \sum_{s_m[2]} .. \sum_{s_m[K]} p(\boldsymbol{n}), \tag{5.6}$$

[7]The value of $N \leq 4$ is commonly used in the literature, with $N = 2$ being the most popular choice, to model non-Gaussian noise pdf [23], [25], [30]. Additionally, the value of P can be chosen to be a large value. In such a case, if the chosen value of P is greater than its true value, the technique developed for estimating the AR parameters would estimate the "unneeded" coefficients as zero. However, if the assumption of known values of N and P is not reasonable for the scenarios under consideration, these parameters must be incorporated into the analysis as nuisance parameters. In such a case, alternative algorithms such as those based on Markov Chain Monte Carlo techniques [119], [120] or Dirichlet processes [121] must be used.

where $\boldsymbol{n}= \{n[k] = r[k]-\alpha s_m[k]\}_{k=1}^{K}$ is the AR process defined in (5.2). Using Bayes Theorem, the density function $p(\boldsymbol{n})$ is given by

$$p(\boldsymbol{n}) = p(n[K], \ldots, n[1]) = \left(\prod_{k=P+1}^{K} p(n[k]|n[k-1], \ldots, n[1]) \right) p(n[P], \ldots, n[1]). \qquad (5.7)$$

For an AR(P) process, $p(n[k]|n[k-1], \ldots, n[1]) = p(n[k]|n[k-1], \ldots, n[k-P])$. Additionally, $p(n[P], \ldots, n[1])$ can be written as

$$p(n[P], \ldots, n[1]) = \prod_{k=1}^{P} p\left(n[k]|n[k-1], \ldots, n[k-P]\right), \qquad (5.8)$$

where, without loss of generality, it is assumed that the values of $n[k]$ are equal to zero for $k \leq 0$. From (5.7) and (5.8), $p(\boldsymbol{n}) = \prod_{k=1}^{K} p\left(n[k]|n[k-1], \ldots, n[k-P]\right)$, and finally, using (5.2) and (5.3), $p(\boldsymbol{n})$ can be written as

$$p(\boldsymbol{n}) = \prod_{k=1}^{K} p\left(n[k]|n[k-1], \ldots, n[k-P]\right) = \prod_{k=1}^{K} f\left(n[k] + \sum_{i=1}^{K} a[i] n[K-i] \right). \qquad (5.9)$$

where $f(e)$ is defined in (5.3). Substituting (5.6) and (5.9) in (5.4), the maximum-likelihood modulation classifier is given by

$$H_{ML} = \arg\max_{H_m} \log\left(\frac{1}{N_m^K} \sum_{s_m[1]} \cdots \sum_{s_m[K]} \prod_{k=1}^{K} f\Big((r[k] - \alpha s_m[k]) + \sum_{i=1}^{P} a[i](r[k-i] - \alpha s_m[k-i]) \Big) \right). \qquad (5.10)$$

As can be seen from (5.10), the computation of the log-likelihood function under each hypothesis involves the summation over a very large number $\left(= N_m^K\right)$ of terms, thus making the maximum-likelihood classification of the received symbols $\boldsymbol{r}$ infeasible.

To overcome this problem, we develop an alternative technique for modulation classification. In this technique, the AR parameters $\{a[i]\}_{i=1}^{P}$ are used to implement a whitening

filter. The whitening filter makes the noise component of the signal used for classification approximately independent, thereby resulting in a classifier of significantly lesser complexity when compared to the classifier in (5.10), as shown below. For the noise model under consideration (5.2), the whitening filter used is a $(P+1)$-tap FIR filter with coefficients $\{1, a[1], \ldots, a[P]\}$. The output of the whitening filter $\tilde{\boldsymbol{r}} = \{\tilde{r}[k]\}_{k=1}^{K}$ is given by

$$\tilde{r}[k] = r[k] + \sum_{i=1}^{P} a[i]r[k-i] = \alpha(s_m[k] + \sum_{i=1}^{P} a[i]s_m[k-i]) + e[k], \tag{5.11}$$

for $k = 1, \ldots, K$. Likelihood-based classification of the whitened symbols $\tilde{\boldsymbol{r}}$ results in the following classifier

$$H_{wML} = \arg\max_{H_m} \log\left(p\left(\tilde{\boldsymbol{r}}|H_m\right)\right), \tag{5.12}$$

with the likelihood function $p\left(\tilde{\boldsymbol{r}}|H_m\right)$, *assuming independent symbols*, given by

$$p\left(\tilde{\boldsymbol{r}}|H_m\right) = \prod_{k=1}^{K} \frac{1}{N_m^{P+1}} \sum_{s_m[k]} \cdots \sum_{s_m[k-P]} f\left(\tilde{r}[k] - \alpha\left(s_m[k] + \sum_{i=1}^{P} a[i]s_m[k-i]\right)\right). \tag{5.13}$$

Substituting (5.13) in (5.12), the likelihood-based classifier at the output of the whitening filter is given by

$$H_{wML} = \arg\max_{H_m} \sum_{k=1}^{K} \log\left(\frac{1}{N_m^{P+1}} \sum_{s_m[k]} \cdots \sum_{s_m[k-P]} f\left(\tilde{r}[k] - \alpha\left(s_m[k] + \sum_{i=1}^{P} a[i]s_m[k-i]\right)\right)\right). \tag{5.14}$$

As seen from (5.14), the proposed classifier has significantly lesser complexity when compared to the classifier in (5.10) (the summation of N_m^K terms is replaced by K summations over N_m^{P+1} terms).

In the proposed approach to classification, the whitening filter in (5.11) requires the knowledge of the AR parameters $\{a[i]\}_{i=1}^{P}$. In addition, the classifier in (5.14), requires that the channel coefficient α and the noise distribution parameters $\{\lambda_i, \sigma_i\}_{i=1}^{N}$ are estimated. These parameters are, however, not known to the receiver a priori. For this reason, in this

chapter, we develop a new H_∞ filter-based technique for estimating the AR parameters. To estimate the channel coefficient α and the noise distribution parameters $\{\lambda_i, \sigma_i\}_{i=1}^N$, the EM algorithm-based technique presented in Chapter 3 is used.

The proposed classifier can be summarized as follows. First, blind estimates $\{\hat{a}[i]\}_{i=1}^P$ of the AR parameters are obtained from the received symbols $\boldsymbol{r}$ using the H_∞ filter based-technique described in Section 5.3. The received symbols are then passed through a whitening filter with coefficients $\{1, \hat{a}[1], \hat{a}[2], \ldots, \hat{a}[P]\}$ to obtain whitened symbols $\tilde{\boldsymbol{r}}$. The channel coefficients α and the noise distribution parameters $\{\lambda_i, \sigma_i\}_{i=1}^N$ are estimated at the output of the whitening filter using a variant of the EM algorithm (Section 5.4). Based on a standard composite hypothesis testing procedure [2, Chapter 6], one set of estimates for α and $\{\lambda_i, \sigma_i\}_{i=1}^N$ is obtained for each hypothesis (that is, for each possible modulation scheme). Let the estimates of α and $\{\lambda_i, \sigma_i\}_{i=1}^N$ obtained assuming that H_m is true be denoted by α^m and $\{\lambda_i^m, \sigma_i^m\}_{i=1}^N$, respectively. The final form of the proposed classifier is then given by

$$H_{wML} = \arg\max_{H_m} \sum_{k=1}^{K} \log\Bigg(\frac{1}{N_m^{P+1}} \sum_{s_m[k]} \cdots \sum_{s_m[k-P]} \sum_{i=1}^{N} \frac{\lambda_i^m}{\pi \left(\sigma_i^m\right)^2} \times \exp\Bigg(- \frac{\left| \tilde{r}[k] - \alpha \left(s_m[k] + \sum_{i=1}^{P} \hat{a}[i] s_m[k-i] \right) \right|^2}{(\sigma_i^m)^2} \Bigg) \Bigg). \quad (5.15)$$

5.3 Estimation of the AR parameters

In order to estimate the AR parameters $\{a[i]\}_{i=1}^P$, we first observe that the received symbols in (5.1) are obtained as the sum of the AR process $\boldsymbol{n}$ $\left(= \{n[k]\}_{k=1}^K\right)$ with an additional unknown process $\alpha \boldsymbol{s}_m$, where $\boldsymbol{s}_m = \{s_m[k]\}_{k=1}^K$. The received symbols $\boldsymbol{r}$ $\left(= \{r[k]\}_{k=1}^K\right)$ therefore constitute a noisy or a *corrupted* AR process. Various techniques are available in the literature for estimating the AR parameters from such corrupted AR processes. These include algorithms based on the modified Yule-Walker equations [122], noise compensated least squares algorithms [123], and Kalman filters [124], among others. However, these

techniques are known to suffer from a significant degradation in performance when the underlying pdfs are non-Gaussian and possess impulsive characteristics [59], [29], [125], as in the case of interest. Additionally, techniques developed to estimate the AR parameters in non-Gaussian environments, such as those based on the EM algorithm [126], cannot be easily extended to the case where an additional signal is present [127]. This is especially true when there is no a priori information about the distributions of the AR process and the unknown additional signal [128].

In this section, we develop a technique for blind estimation of the AR parameters based on the use of the H_∞ filter. The H_∞ filter is a robust estimator for the state of a linear dynamic system. It was originally designed as an alternative to the traditional Kalman filter, which is ineffective in the presence of outliers (like in the case of impulsive noise) and noise uncertainty [129]. While the Kalman filter minimizes the mean square error, the H_∞ filter is a *minimax* estimator which minimizes the peak error power in the frequency domain [130]. It must be noted that like the Kalman filter, the H_∞ filter is also a sequential estimator. One of the main reasons for the popularity of the H_∞ filter is that, unlike the Kalman filter and its extensions, no prior knowledge of the statistics of the processes involved, in this case $\alpha \boldsymbol{s}_m$ and $\boldsymbol{n}$, is required (although this information can be used if available) [29], [125]. The only assumption made is that the processes have finite energies over the observation interval. The H_∞ filters have been used to develop a variety of robust signal processing algorithms for channel estimation [131], equalization [132], and signal demodulation [133], among others, for non-Gaussian noise environments.

5.3.1 The H_∞ filter

Consider a linear dynamic system with state-space model given by

$$\text{State Equation} \quad : \quad X[k+1] = F_k X[k] + W[k] \tag{5.16}$$

$$\text{Observation Equation} \quad : \quad Y[k] = H_k X[k] + V[k], \tag{5.17}$$

for $k = 1, 2, \ldots, K$, where $X[k]$ is the state vector and $Y[k]$ is the observation vector. The H_∞ filter provides a robust technique to estimate the vector $Z[k] = L_k X[k]$, obtained by a linear operation on the state $X[k]$, where L_k is a known matrix. (For example, to estimate the state vector $X[k]$, L_k is taken to be an identity matrix.) The state noise and the observation noise are denoted by $W[k]$ and $V[k]$, respectively. The matrices L_k, F_k, and H_k are defined by the problem under consideration.

The aim of H_∞ filtering is to obtain the estimate $\hat{Z}[k]$ of $Z[k]$ which minimizes the H_∞ norm given by [129]

$$J_\infty = \max_{W[k],V[k],X[0]} \frac{\sum_{k=0}^{K} |Z[k] - \hat{Z[k]}|^2_{S_k}}{|X[0] - \hat{X[0]}|^2_{P_0^{-1}} + \sum_{k=0}^{K}(|W[k]|^2_{Q_k^{-1}} + |V[k]|^2_{R_k^{-1}})}, \tag{5.18}$$

where S_k, P_0, Q_k, and R_k are symmetric positive definite matrices chosen by the user [129]. (The choice of these matrices is further discussed in Section 5.3.4.) The notation $|A|^2_B$ denotes $A^H B A$, where A^H is the Hermitian of A. The parameter P_0 is a measure of how close the initial estimate $\hat{X}[0]$ is to the true value $X[0]$, while Q_k and R_k play roles analogous to the covariances of the state and of the observation noise in a Kalman filter, respectively.

Because the direct minimization of J_∞ is not tractable, the following suboptimal criterion is commonly considered [129]- [134]

$$J_\infty = \max_{W[k],V[k],X[0]} \frac{\sum_{k=0}^{K} |Z[k] - \hat{Z[k]}|^2_{S_k}}{|X[0] - \hat{X}[0]|^2_{P_0^{-1}} + \sum_{k=0}^{K}(|W[k]|^2_{Q_k^{-1}} + |V[k]|^2_{R_k^{-1}})} < \gamma^2, \tag{5.19}$$

where γ^2 is a user specified bound. The solution for (5.19) is given by $\hat{Z}[k] = L_k \hat{X}[k]$ [134],

where

$$\hat{X}[k+1] = F_k\hat{X}[k] + K_{k+1}U[k+1], \tag{5.20}$$

$$U[k+1] = Y[k+1] - H_{k+1}F_k\hat{X}[k], \tag{5.21}$$

$$K_{k+1} = P_{k+1}H_{k+1}^H[R_{k+1} + H_{k+1}P_{k+1}H_{k+1}^H]^{-1}, \text{ and} \tag{5.22}$$

$$P_{k+1} = F_kP_kF_k^H - F_k[I + \left(H_k^HR_k^{-1}H_k - \gamma^{-2}\bar{S}_k\right)^{-1}P_k^{-1}]^{-1}P_kF_k^H + Q_k, \tag{5.23}$$

with $\bar{S}_k = L_k^HS_kL_k$ [129]. In addition, the solution $\hat{Z}[k]$ of (5.19) exists if and only if the following condition is satisfied at each time instant k,

$$P_k^{-1} - \gamma^{-2}\bar{S}_k + H_k^HR_k^{-1}H_k > 0. \tag{5.24}$$

5.3.2 H_∞ filter-based estimation of AR parameters

In this subsection, we develop a H_∞ filter-based technique for estimating the AR parameters $\{a[i]\}_{i=1}^P$ from the corrupted AR process $\boldsymbol{r}$. We first observe that, using (5.1) and (5.2), the corrupted AR process $\boldsymbol{r}$ can be written as

$$r[k] = -\sum_{i=1}^{P} a[i]r[k-i] + \alpha\left(s_m[k] + \sum_{i=1}^{P} a[i]s_m[k-i]\right) + e[k]. \tag{5.25}$$

Using (5.25), together with the fact that the AR parameters are assumed to be time-invariant, the state-space equations for the H_∞ filter are given by

$$\text{State Equation} \;:\; X[k+1] = X[k] \tag{5.26}$$

$$\text{Observation Equation} \;:\; r[k] = H_kX[k] + V[k], \tag{5.27}$$

where $X[k] = \left[a[1], a[2], \ldots, a[P]\right]^T$, $H_k = \left[-r[k-1], -r[k-2], \ldots, -r[k-P]\right]$, and $V[k] = e[k] + \alpha\big(s_m[k] + a[1]s_m[k-1] + \ldots a[P]s_m[k-P]\big)$. The estimates of the AR parameters

are obtained by using $Z[k] = X[k]$ in (5.20)-(5.23).

In this procedure, the AR parameters are estimated by observing the corrupted AR process $\boldsymbol{r}$, as seen in the observation equation (5.27). However, various studies have shown that estimating the AR parameters in this manner (that is, directly from a corrupted AR process) results in highly biased (and hence unsatisfactory) estimates [135], [136]. In these studies, it has been suggested that a more effective approach would be to estimate the AR process $\boldsymbol{n}$ from the corrupted AR process $\boldsymbol{r}$ and then compute the AR parameters from the estimated AR process $\hat{\boldsymbol{n}}$. Such a procedure, which involves the estimation of both the state (in this case $\boldsymbol{n}$) and the underlying parameters (in this case $\{a[i]\}_{i=1}^{P}$) of the system, is known as a *dual estimation procedure* [137].

The dual estimation procedure involves the use of two sequential estimators operating in parallel, one for estimating the state and one estimating the parameters of the system under consideration. This approach can be understood as a generalized version of the EM algorithm, where the E-step is replaced by the state estimator and the M-step is replaced by the parameter estimator [138]. While the dual estimation procedure does not have the convergence properties of the EM algorithm, it has the advantage of being computationally less complex and requiring lesser memory due to its sequential nature [127]. This procedure has been used in a applications such as channel estimation [135] and speech enhancement [136], among various others.

In the following subsection, a dual estimation procedure for estimating the AR parameters $\{a[i]\}_{i=1}^{P}$ from the corrupted AR process $\boldsymbol{r}$ with the help of two H_∞ filters operating in parallel is presented[8].

[8]At this point, we are only interested in estimating the AR parameters to be used as the coefficients of the whitening filter. For this reason, in this procedure, only the AR parameters are considered to be the unknown parameters. A technique which does not require the knowledge of the other unknown parameters (channel coefficient and Gaussian mixture model parameters) is used to obtain the AR parameter estimates. The estimation of all the unknown parameters simultaneously using the dual estimation procedure is difficult due to the non-Gaussian distribution of the processes involved.

5.3.3 Dual H_∞ filter-based estimation of the AR parameters

This is a sequential two-step approach (that is, two steps are performed at every time instant) for estimating the AR parameters. In the first step, estimates of the AR process $\boldsymbol{n}$ (state) are obtained from the corrupted AR process $\boldsymbol{r}$, while in the second step the AR parameters are calculated from the estimated AR process $\hat{\boldsymbol{n}} = \{\hat{n}[k]\}_{k=1}^{K}$. It must be noted, however, that the first step requires the knowledge of the AR parameters to estimate $\hat{\boldsymbol{n}}$, and consequently, in this approach, the two steps described above have to be implemented in parallel, as shown in Fig. 5.1.

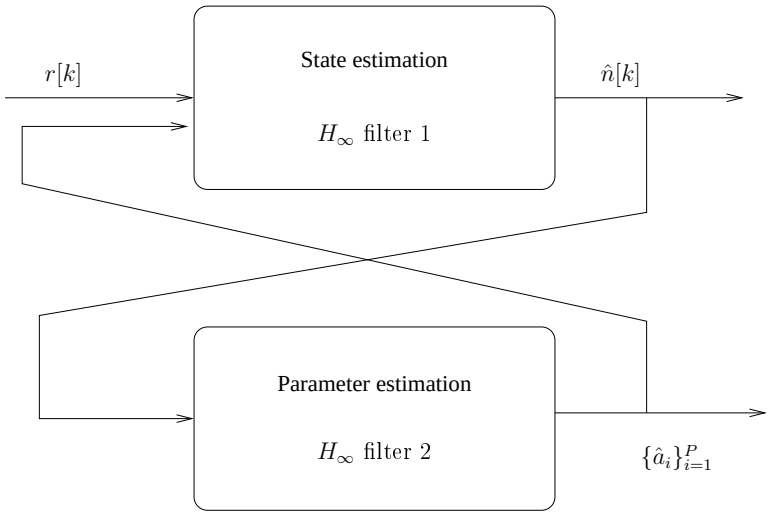

Figure 5.1: H_∞ filter-based dual estimation scheme for AR parameter estimation.

Due to the non-Gaussian distribution of the noise and the lack of knowledge of the statistics of the processes involved, H_∞ filters are used for implementing both the state estimation and the parameter estimation steps. The state-space equations for the H_∞ filters used in these steps, at the k^{th} time instant, are:

State estimation

In the first step, the most recent estimates of the AR parameters $\{\hat{a}[i]\}_{i=1}^{P}$ are used to obtain the estimate $\hat{n}[k]$. Taking the state vector $X_1[k]$ to be $X_1[k] = [n[k], n[k-1], \ldots, n[k-P]]^T$,

the state-space equations for this filter are

$$X_1[k+1] = F_{1k}X_1[k] + W_1[k] \tag{5.28}$$

$$r[k] = H_{1k}X_1[k] + V_1[k], \tag{5.29}$$

where $V_1[k] = \alpha s_m[k]$, $H_{1k} = [1, 0, 0, \ldots, 0]$, $W_1[k] = \big[e[k+1], 0, 0, \ldots, 0\big]^T$, and

$$F_{1k} = \begin{bmatrix} -\hat{a}[1] & -\hat{a}[2] & \cdots & -\hat{a}[P] & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \tag{5.30}$$

The parameter of interest $n[k]$ is given by $n[k] = Z_1[k] = L_{1k}X_1[k]$ with $L_{1k} = [1, 0, 0, \ldots, 0]$.

Parameter estimation

In this step, the filtered estimate $\hat{n}[k]$ $(= \hat{Z}_1[k])$ is used to update the AR parameter estimates $\{\hat{a}[i]\}_{i=1}^{P}$. Since the AR parameters are assumed to be time-invariant, the state equation of this H_∞ filter can be written as

$$X_2[k+1] = X_2[k], \text{ where } X_2[k] = \big[\hat{a}[1], \hat{a}[2], \ldots, \hat{a}[P]\big]^T. \tag{5.31}$$

For the observation equation in this step, using (5.20) in the state estimation step, we have

$$\hat{n}[k] = L_{1k}\hat{X}_1[k] = L_{1k}F_{1(k-1)}\hat{X}_1[k-1] + L_{1k}K_{1k}U_1[k]$$

Using $L_{1k} = [1, 0, 0, \ldots, 0]$ and (5.30), the observation equation is given by

$$\hat{n}[k] = H_{2k}X_2[k] + V_2[k], \text{ where } H_{2k} = \hat{X}_1[k-1]^T \text{ and } V_2[k] = L_{1k}K_{1k}U_1[k]. \tag{5.32}$$

The sub-optimal H_∞ filtering criterion (5.19) for the two filters with state space equations given by (5.28)-(5.31) and (5.32) are

$$J_{1\infty} = \max_{e[k],s_m[k],X[0]} \frac{\sum_{k=0}^{K} |n[k] - \hat{n}[k]|^2_{S_{1k}}}{|X_1[0] - \hat{X}_1[0]|^2_{P_{10}^{-1}} + \sum_{k=0}^{K}(|e[k]|^2_{Q_{1k}^{-1}} + |\alpha s_m[k]|^2_{R_{1k}^{-1}})} < \gamma_1^2 \text{ and} \tag{5.33}$$

$$J_{2\infty} = \max_{W_2[k],V_2[k],X[0]} \frac{\sum_{k=0}^{K} |X_2[k] - \hat{X}_2[k]|^2_{S_{2k}}}{|X_2[0] - \hat{X}_2[0]|^2_{P_{20}^{-1}} + \sum_{k=0}^{K}(|V_2[k]|^2_{R_{2k}^{-1}})} < \gamma_2^2. \tag{5.34}$$

Since the statistics of $n[k]$, $e[k]$, $s_m[k]$, and $X_2[k]$ are assumed to be time invariant, the time-dependencies of S_{1k}, Q_{1k}, R_{1k}, and S_{2k} can be dropped [129]. The dual estimation procedure for estimating the AR parameters, obtained by applying (5.20)-(5.23) to the two H_∞ filters, is presented in Table 5.1.

Table 5.1: Proposed H_∞ filter-based dual estimation procedure for AR parameter estimation.

For $k = 0$ to K

State Estimation Step

$U_1[k+1] = r[k+1] - H_{1(k+1)} F_{1k} \hat{X}_1[k]$

$P_{1(k+1)} = F_{1k} P_{1k} F_{1k}^H - F_{1k}[I + \left(H_{1k}^H R_{1k}^{-1} H_{1k} - \gamma^{-2} \bar{S}_{1k}\right)^{-1} P_{1k}^{-1}]^{-1} P_{1k} F_{1k}^H + Q_{1k}$,

$K_{1(k+1)} = P_{1(k+1)} H_{1(k+1)}^H [R_{1(k+1)} + H_{1(k+1)} P_{1(k+1)} H_{1(k+1)}^H]^{-1}$,

$\hat{n}[k+1] = L_{1(k+1)} \hat{X}_1[k+1]$

Adjusting the weight $Q_{1(k+1)}$

$M_{k+1} = P_{1(k+1)} - F_{1(k+1)} P_{1k} F_{1(k+1)}^H + K_{1(k+1)} |U_1[k+1]|^2 K_{1(k+1)}^H$

$Q_{1(k+1)} = \frac{k}{k+1} Q_{1k} + \frac{1}{k+1} M_{k+1}$

Parameter Estimation Step

$H_{2(k+1)} = \hat{X}_1[k]^T$, $F_{2k} = I$

$U_2[k+1] = \hat{n}[k+1] - H_{2(k+1)} \hat{X}_2[k]$

$P_{2(k+1)} = P_{2k} - [I + \left(H_{2k}^H R_{2k}^{-1} H_{2k} - \gamma^{-2} \bar{S}_{2k}\right)^{-1} P_{2k}^{-1}]^{-1} P_{2k}$,

$K_{2(k+1)} = P_{2(k+1)} H_{2(k+1)} [R_{2(k+1)} + H_{2(k+1)} P_{2(k+1)} H_{2(k+1)}^H]^{-1}$,

$\hat{X}_2[k+1] = \hat{X}_2[k] + K_{2(k+1)} U_2[k+1]$

end

5.3.4 Choosing of the H_∞ filter parameters

The user-dependent parameters which are to be chosen include $\{S_1,\ P_{10},\ Q_1, R_1, \gamma_1\}$ and $\{S_2, P_{20}, Q_2, R_2, \gamma_2\}$. Since, at each step, all the estimates are to be obtained with the same accuracy, without loss of generality, S_1 and S_2 are taken to be identity matrices [129]. With no priori information being available on the initial states, the corresponding weighting parameters P_{10} and P_{20} are typically chosen to be large positive definite matrices [125]. The parameters γ_1 and γ_2 are chosen such that the condition (5.24) is met for both the filters. Accordingly, at the $(k+1)^{th}$ time instant, γ_1 and γ_2 are chosen so as to satisfy [136]

$$\gamma_1^2 > \max\left\{\text{eig}\left(\bar{S}_1\left(P_{1k}^{-1} + H_{1k}^H R_1^{-1} H_{1k}\right)^{-1}\right)\right\} \tag{5.35}$$

$$\gamma_2^2 > \max\left\{\text{eig}\left(\bar{S}_2\left(P_{2k}^{-1} + H_{2k}^H R_{2k}^{-1} H_{2k}\right)^{-1}\right)\right\} \tag{5.36}$$

where $\max\{(\text{eig}(M)\}$ denotes the maximum eigenvalue of the matrix M. The parameters Q_i and R_i play roles that are analogous to the covariance of the state noise and the observation noise in a Kalman filter. Hence, they are dependent on the noise statistics. However, one of the major advantages of using H_∞ filters is that their performance are insensitive to the lack of knowledge of these covariances [129]. For example, it was observed in [29] that even for values of Q_k and R_k with deviations of ± 10 dB from the true variance, the H_∞ filter has a negligible loss in performance. In this chapter, while the values of R_i are chosen to be constants (identity matrices), the parameter Q_i is updated using the technique proposed in [135] and [136].

In this section, a new H_∞ filter-based dual estimation procedure was proposed for blind estimation of the AR parameters of a correlated non-Gaussian noise process in the presence of an unknown signal.

5.4 Estimation of the channel coefficient and noise distribution parameter

The estimates $\{\hat{a}[i]\}_{i=1}^{P}$ of the AR parameters obtained in Section 5.3 are used as the coefficients of a whitening filter to ideally remove the correlation among of the noise component of the received symbols $\{r[k]\}_{k=1}^{K}$. The whitened symbols $\{\tilde{r}[k]\}_{k=1}^{K}$ at the output of this filter are given by

$$\tilde{r}[k] = r[k] + \sum_{i=1}^{P} \hat{a}[i]r[k-1] = \alpha\tilde{s}_m[k] + \tilde{e}[k]. \tag{5.37}$$

In (5.37), the noise component $\{\tilde{e}[k]\}_{k=1}^{K}$ can be approximated as i.i.d random variables with their distribution given by (5.3). If the AR parameters were perfectly estimated, we would have $\{\tilde{e}[k] = e[k]\}_{k=1}^{K}$. The information component in (5.37) consists of symbols $\{\tilde{s}_m[k]\}_{k=1}^{K}$ from a new constellation $\tilde{\mathcal{S}}_m$. The constellation points of $\tilde{\mathcal{S}}_m$ are obtained as $\tilde{s}_m = s_m^0 + \sum_{i=1}^{P} \hat{a}[i]s_m^i$ for all possible combinations of $\{s_m^i \in \mathcal{S}_m\}_{i=0}^{P}$. The number of constellation points in $\tilde{\mathcal{S}}_m$ is therefore equal to N_m^{P+1}.

In order to estimate the channel coefficient and Gaussian mixture model parameters $\left\{\alpha, \{\lambda_i, \sigma_i\}_{i=1}^{N}\right\}$, denoted by $\boldsymbol{\theta}$, assuming that the whitened symbols $\{\tilde{r}[k]\}_{k=1}^{K}$ are independent, the ECM algorithm proposed in Chapter 3 is used. In this case, for a given hypothesis $\mathcal{H}_m$, the iterative procedure for obtaining *maximum-likelihood* estimates of $\boldsymbol{\theta}$ is as follows. The estimate of α at the $(t+1)^{th}$ iteration is obtained by using

$$\alpha^{t+1} = \frac{\sum_k \sum_{y[k]} \sum_{\tilde{s}_m[k]} \frac{\tilde{r}[k]\tilde{s}_m^*[k]}{\left(\sigma_{y[k]}^t\right)^2} \frac{\lambda_{y[k]}^t}{\tilde{N}_m} \frac{p(\tilde{r}[k]|y[k],\tilde{s}_m[k],\boldsymbol{\theta^t})}{p(\tilde{r}[k]|\boldsymbol{\theta^t},H_m)}}{\sum_k \sum_{y[k]} \sum_{\tilde{s}_m[k]} \frac{|\tilde{s}_m[k]|^2}{\left(\sigma_{y[k]}^t\right)^2} \frac{\lambda_{y[k]}^p}{\tilde{N}_m} \frac{p(\tilde{r}[k]|y[k],\tilde{s}_m[k],\boldsymbol{\theta^t})}{p(\tilde{r}[k]|\boldsymbol{\theta^t},H_m)}} \tag{5.38}$$

where $\boldsymbol{\theta^t} = \left\{\alpha^t, \{\lambda_n^t, \sigma_n^t\}_{n=1}^{N}\right\}$ is the estimate of $\boldsymbol{\theta}$ after the t^{th} iteration. After estimating α^{t+1} using (5.38), $\boldsymbol{\theta^t}$ is updated as $\boldsymbol{\theta^t} = \left\{\alpha^{t+1}, \{\lambda_n^t, \sigma_n^t\}_{n=1}^{N}\right\}$. The estimates λ_i and σ_i for

$i = 1, 2, \ldots, N$, are then updated by using, respectively,

$$\lambda_i^{t+1} = \frac{1}{K} \sum_k \sum_{\tilde{s}_m[k]} \frac{\lambda_i^t}{\tilde{N}_m} \frac{p(\tilde{r}[k]|y[k] = i, \tilde{s}_m[k], \boldsymbol{\theta^t})}{p(\tilde{r}[k]|\boldsymbol{\theta^t}, H_m)} \tag{5.39}$$

and

$$\sigma_i^{t+1} = \sqrt{\frac{\sum_k \sum_{\tilde{s}_m[k]} |\tilde{r}[k] - \alpha^t \tilde{s}_m[k]|^2 \frac{\lambda_i^t}{\tilde{N}_m} \frac{p(\tilde{r}[k]|y[k]=i, \tilde{s}_m[k], \boldsymbol{\theta^t})}{p(\tilde{r}[k]|\boldsymbol{\theta^t}, H_m)}}{2 \sum_k \sum_{\tilde{s}_m[k]} \frac{\lambda_i^t}{\tilde{N}_m} \frac{p(\tilde{r}[k]|y[k]=i, \tilde{s}_m[k], \boldsymbol{\theta^t})}{p(\tilde{r}[k]|\boldsymbol{\theta^t}, H_m)}}}. \tag{5.40}$$

These iterations are repeated until convergence. While the ECM algorithm does not guarantee convergence to the maximum-likelihood estimates (global maximum of the log-likelihood function), it converges to a local maximum, with the point of convergence depending on the initial estimates. Detailed discussions of the convergence of the ECM algorithm can be found in [88] and [91].

As discussed in Section 5.2, one set of estimates of $\left\{\alpha^m, \{\lambda_i^m, \sigma_i^m\}_{i=1}^N\right\}$ is obtained for each hypothesis. These estimates, along with the estimates of the whitening filter coefficients $\{\hat{a}[i]\}_{i=1}^P$, are then used by the classifier in (5.15) for identifying the modulation scheme from the whitened symbols $\{\tilde{r}[k]\}_{k=1}^K$.

5.5 Numerical Results

In the results that follow, the modulation schemes considered are BPSK, QPSK, and 8-PSK (with each of them assumed to be equally likely). The amplitude of the channel coefficient α is assumed to be Rayleigh distributed with $E[|\alpha|^2] = 1$, and its phase uniformly distributed in $(0,\ 2\pi]$. In the N-term Gaussian mixture model defined in (5.3), based on common nomenclature available in the literature, the first term (with proportion λ_1 and variance σ_1^2) is used to represent the thermal noise component, while the other terms model the impulsive noise component [23], [25], [82]. Unless specified, the number of terms in the Gaussian mixture distribution is taken to be $N = 2$ and λ_1 and $\frac{\sigma_2^2}{\sigma_1^2}$ are chosen to be 0.9 and 100,

respectively. This case is widely used in the literature (see, among others, [23], [25], [82], [84]) as a model for impulsive noise. We define the received SNR as the ratio of the average received signal power and the variance of the thermal noise component; that is, the received SNR is equal to $E[|\alpha|^2]/\sigma_1^2$.

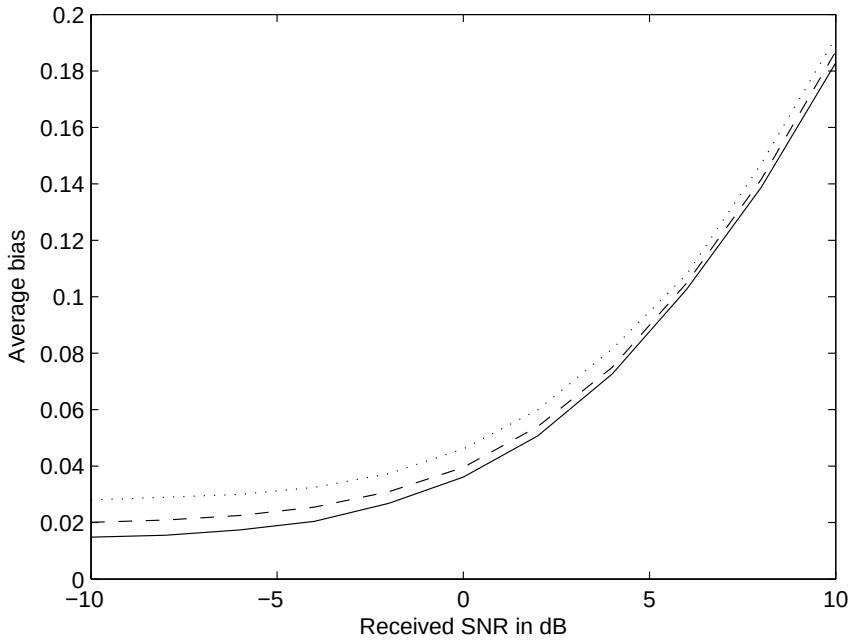

Figure 5.2: Average bias in the estimates of the AR(1) process parameter: $a[1] = -0.6$. The average bias in the estimate $\hat{X}$ of the parameter X is defined as $E[|X - \hat{X}|]$. Number of symbols for AR parameter estimation equal to 250 (dotted line), 500 (dashed line), and 1000 (solid line). Modulation scheme used is QPSK.

The performance of the proposed H_∞ filter-based dual estimation procedure for estimating the AR parameters is presented in Figs. 5.2 and 5.3. The bias of the estimated AR parameters is obtained as a function of the received SNR for two separate cases: Fig. 5.2, an AR(1) filter $\{1, -0.6\}$; and Fig. 5.3, an AR(2) filter $\{1, -0.25, -0.53\}$. It can be seen in Figs. 5.2 and 5.3 that as the received SNR decreases, the bias in the estimation of the AR parameters reduces. This is because, as the received SNR decreases, the relative strength of the AR process (noise) in the received signal increases and consequently the accuracy of the AR parameter estimation increases. In addition, the performance of the estimator improves

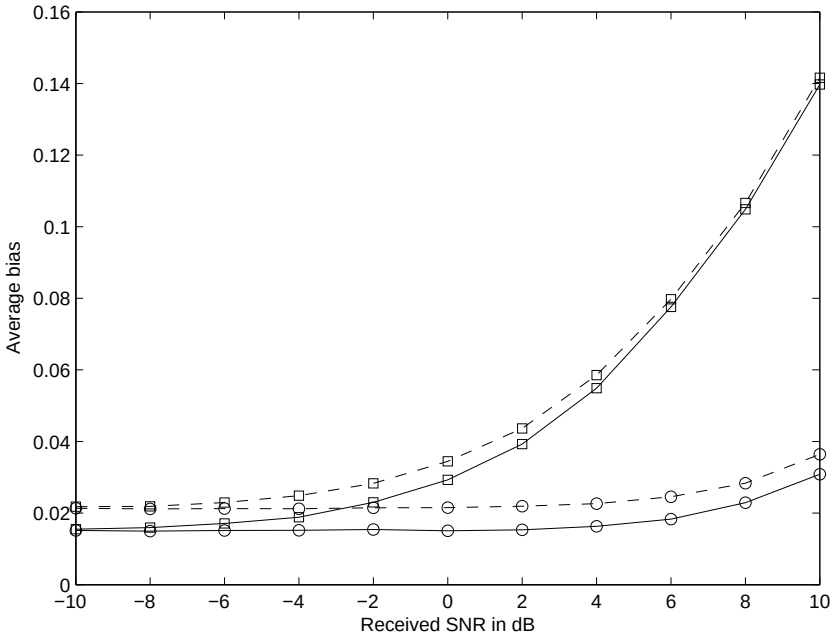

Figure 5.3: Average bias in the estimates of the AR(2) process parameters: $a[1] = -0.25$ (circle), $a[2] = -0.53$ (square). The average bias in the estimate $\hat{X}$ of the parameter X is defined as $E[|X - \hat{X}|]$. Number of symbols used in the estimation of the AR parameters is equal to 500 (dashed line) and 1,000 (solid line). Modulation scheme used is QPSK.

as the number of symbols used in the estimation process increases, as expected.

It must be noted that the AR parameter estimation is performed without prior knowledge of the signal and noise statistics. For this reason, the parameters of the two H_∞ filters used in the dual estimation procedure, which would ideally depend on these statistics, are instead chosen as discussed in Section 5.3.4. The bias in the AR parameter estimation observed in Figs. 5.2 and 5.3 is a direct consequence of this mismatch in the H_∞ filter parameters. Despite the bias in the AR parameter estimation, the performance of the proposed classifier is close to that of the "ideal" classifier (as seen in the results presented later in this section).

The estimated AR parameters are used as the coefficients of a whitening filter to reduce the correlation among the noise component of the received signal. To demonstrate the efficacy of the whitening filter, Figs. 5.4 and 5.5 show the average normalized correlation

of the noise component of the signal before and after the whitening filter. As can be seen in Figs. 5.4 and 5.5, the reduction in the correlation of the noise component of the signal is noticeable. It is also observed that the whitening filter has a better performance at low values of the received SNR. This is because the estimates of the AR parameters are more accurate at these SNRs, as previously discussed.

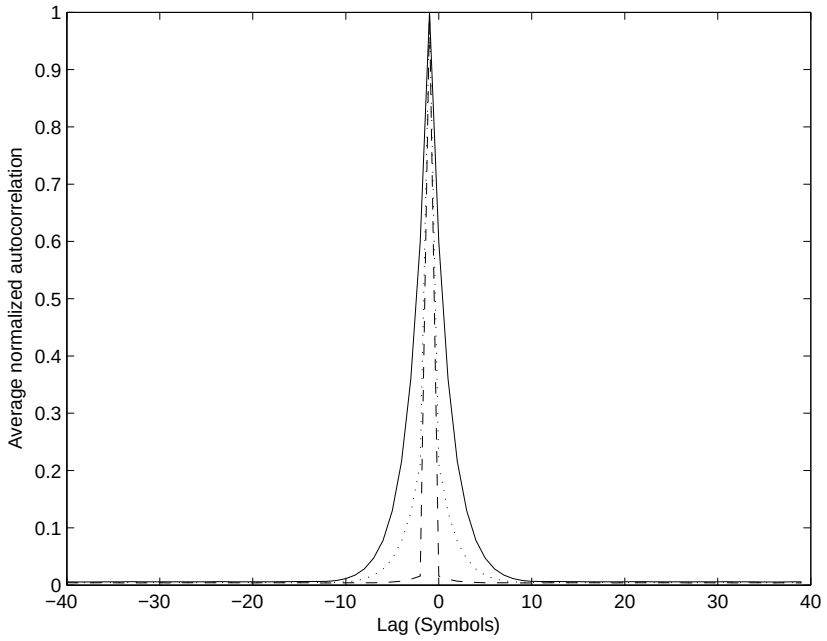

Figure 5.4: Average normalized autocorrelation at the whitening filter output: SNR = -10 dB (dashed line); SNR = 10 dB (dotted line). The average normalized autocorrelation is defined as $E[\frac{|R_X|}{\max(|R_X|)}]$, where R_X is the autocorrelation function. Number of symbols used in the estimation of the AR parameters is 1,000. AR(1) process - $a[2] = -0.6$ Also plotted is the average normalized correlation at the whitening filter input (solid line).

The performance of the ECM algorithm used to estimate the channel coefficient and noise distribution parameters after the whitening filter is shown in Figs. 5.6-5.8. In Figs. 5.6 and 5.7, the normalized variances of the estimates of σ_1^2 and σ_2^2 are shown when the whitening filter coefficients are obtained using the technique described in Section 5.3. Also shown are the normalized variances when these coefficients are assumed to be perfectly known. It can be seen from Figs. 5.6 and 5.7 that as the received SNR increases, the normalized variance

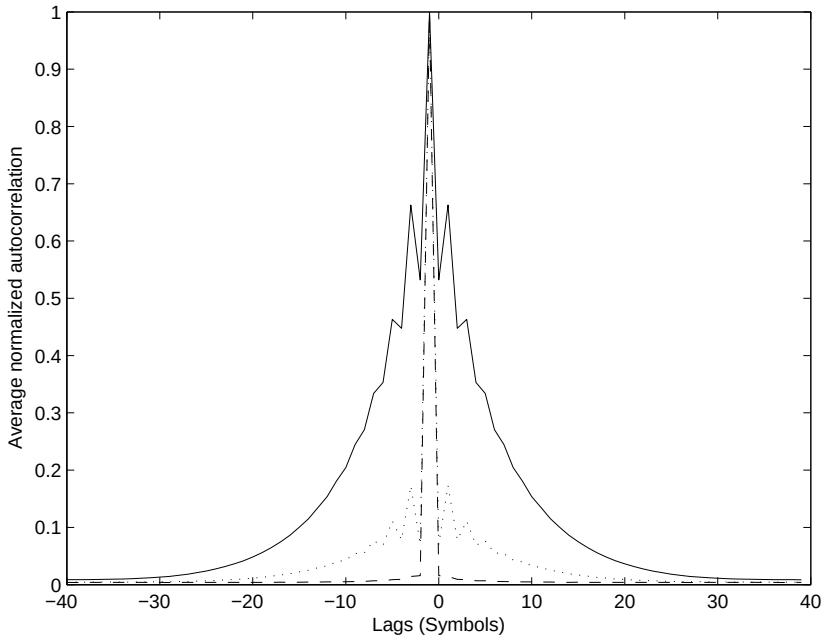

Figure 5.5: Average normalized autocorrelation at the whitening filter output: SNR = -10 dB (dashed line); SNR = 10 dB (dotted line). The average normalized autocorrelation is defined as $E[\frac{|R_X|}{\max(|R_X|)}]$, where R_X is the autocorrelation function. Number of symbols used in the estimation of the AR parameters is 1,000. AR(2) process : $a[1] = -0.25, a[2] = -0.53$. Also plotted is the average normalized correlation at the whitening filter input (solid line).

of both σ_1^2 and σ_2^2 decreases in almost all cases, as expected. For low values of the received SNR, the estimation performance for the case in which the whitening filter coefficients are estimated is approximately the same as the estimation performance when these coefficients are assumed to be known. As the received SNR increases, the gap in performance increases, since the accuracy of AR parameter estimation reduces at higher values of the received SNRs. In Fig. 5.7, additionally, the normalized variance of σ_1^2 increases at higher values of SNRs due to the increasing bias of the AR parameter estimation. Similar trends are also observed for the estimates of the channel coefficient α. The average bias in the estimation of the magnitude and the phase of α reduces as the received SNR increases, as seen in Fig. 5.8.

Figs. 5.9 and 5.10 show the performance of the proposed classifier for different AR model parameters. (The corresponding confusion matrices are presented for Tables 5.2 and 5.3).

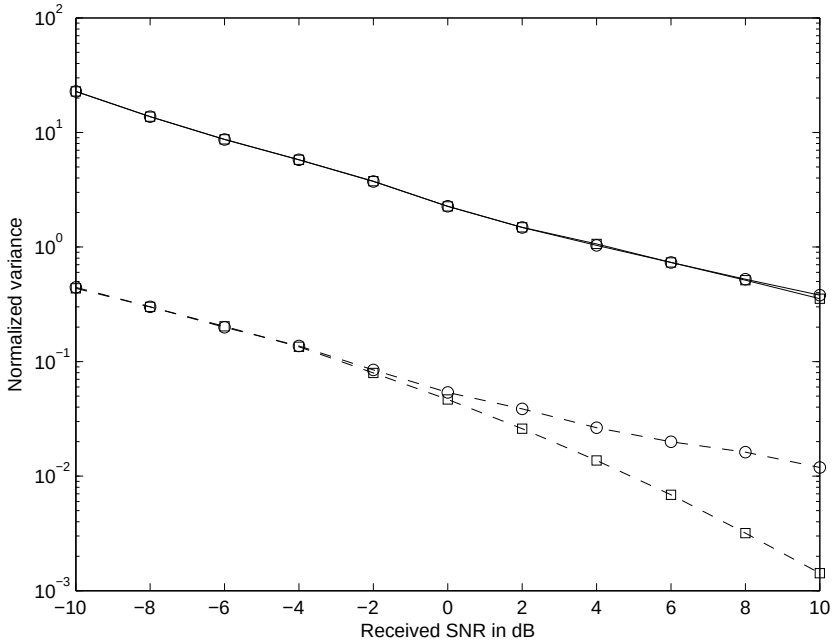

Figure 5.6: Normalized variance of the estimates of σ_1^2 (dashed lines) and σ_1^2 (solid) when the whitening filter coefficients are estimated (circle) are perfectly known (square). The normalized variance of the estimate $\hat{X}$ of the parameter X is defined as $\frac{\mathrm{var}(\hat{X})}{X}$. Modulation scheme considered is QPSK. AR(1) process : $a[1] = -0.6$. Number of symbols used in the estimation of the AR parameters and in the EM algorithm are 1,000 and 500, respectively.

Additionally, Fig. 5.11 shows the performance of the proposed classifier, for the noise model parameters which correspond to the statistics of the impulsive noise measured in [15]. (Please refer to Appendix C for details.) For reference, the proposed classifier is compared to three other classifiers

1. The "ideal" classifier with perfect knowledge of the AR parameters, the Gaussian mixture model parameters, and the channel coefficient. This classifier is given by (5.15) with the true values (instead of noisy estimates) of $\{a[i]\}_{i=1}^{P}$, α, and $\{\lambda_i, \sigma_i\}_{i=1}^{N}$ used.

2. The likelihood-based classifier formulated for non-Gaussian noise channels in Chapter 3 where the noise is assumed to be white and hence no whitening filter is used.

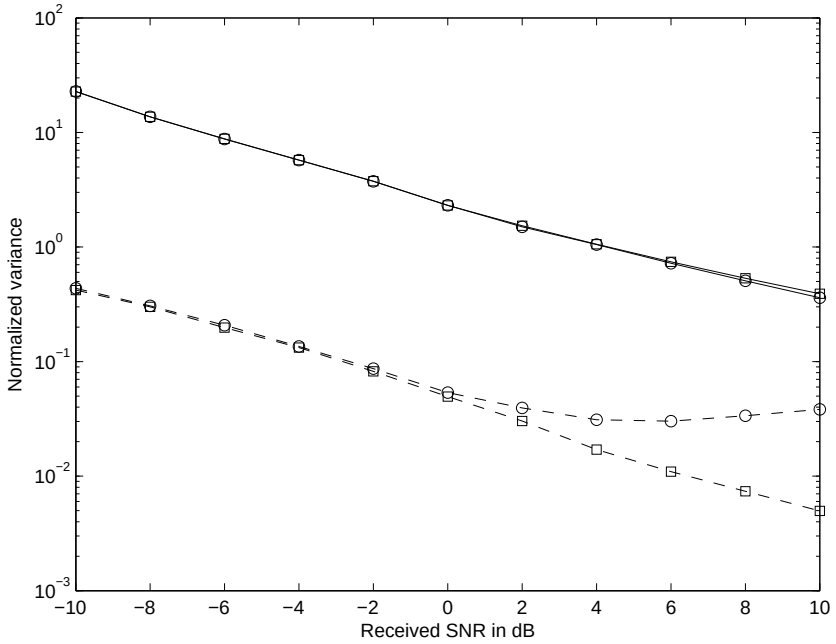

Figure 5.7: Normalized variance of the estimates of σ_1^2 (dashed lines) and σ_1^2 (solid) when the whitening filter coefficients are estimated (circle) and are perfectly known (square). The normalized variance of the estimate $\hat{X}$ of the parameter X is defined as $\frac{\mathrm{var}(\hat{X})}{X}$. Modulation scheme considered is QPSK. AR(2) process : $a[1] = -0.25, a[2] = -0.53$. Number of symbols used in the estimation of the AR parameters and in the EM algorithm are 1,000 and 500, respectively.

3. The likelihood-based classifier developed for Gaussian noise conditions in [1]. In this case the channel coefficient and the variance of the Gaussian noise required by the classifier are estimated using a method-of-moments-based technique.

As can be seen in these figures that the performance of the proposed classifier is close to the performance of the "ideal" classifier. Additionally, it can be seen in Fig. 5.11 that the gap in performance between the proposed and the "ideal" classifier reduces as the number of symbols used in the estimation process increases. Figs. 5.9-5.11 also show that the proposed classifier performs better than the classifier designed assuming that the noise process is white and significantly outperforms the classifier developed for Gaussian noise conditions.

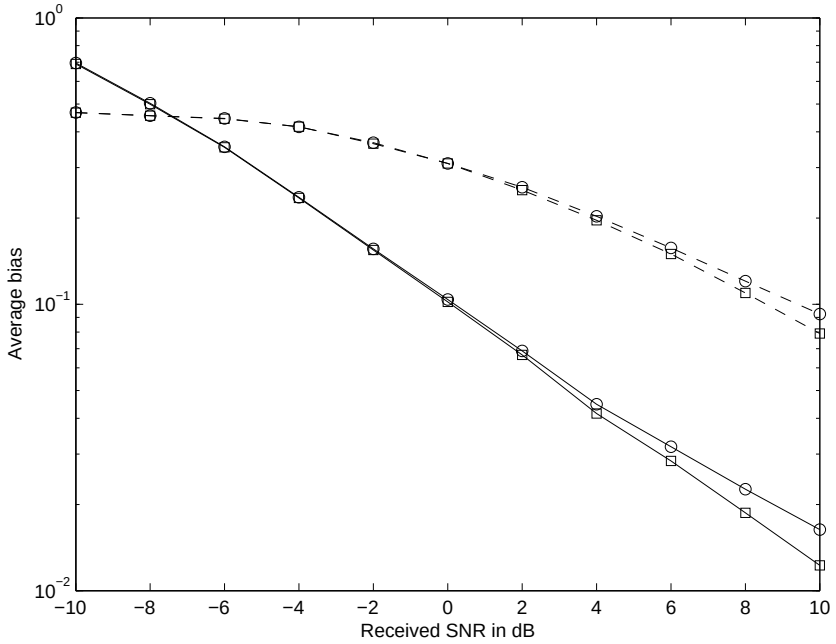

Figure 5.8: Average bias in the estimates of the phase (in radians, dashed lines) and magnitude (unit less, solid lines) of the channel coefficient α when the whitening filter coefficients are estimated (circle) and are perfectly known (square). The average bias of the estimate $\hat{X}$ of the parameter X is defined as $E[|\hat{X} - X|]$. Modulation scheme considered is QPSK. AR(1) process : $a[1] = -0.6$. Number of symbols used in the estimation of the AR parameters and in the EM algorithm are 1,000 and 500, respectively.

5.6 Conclusions

In this chapter, a likelihood-based technique was presented for the classification of digital amplitude-phase modulated signals in flat fading channels in time-correlated non-Gaussian noise. The first-order statistics of the additive noise were modeled using a Gaussian mixture model and the noise correlation was modeled using an AR process. Because the maximum-likelihood modulation classifier is computationally infeasible for the noise model considered, the proposed classifier involves the use of a whitening filter followed by likelihood-based classification of the whitened signal leading to a significant reduction in the complexity. A new algorithm for blind estimation of the unknown AR parameters was developed using a H_∞ filter-based dual estimation procedure, while the channel coefficient and the Gaussian

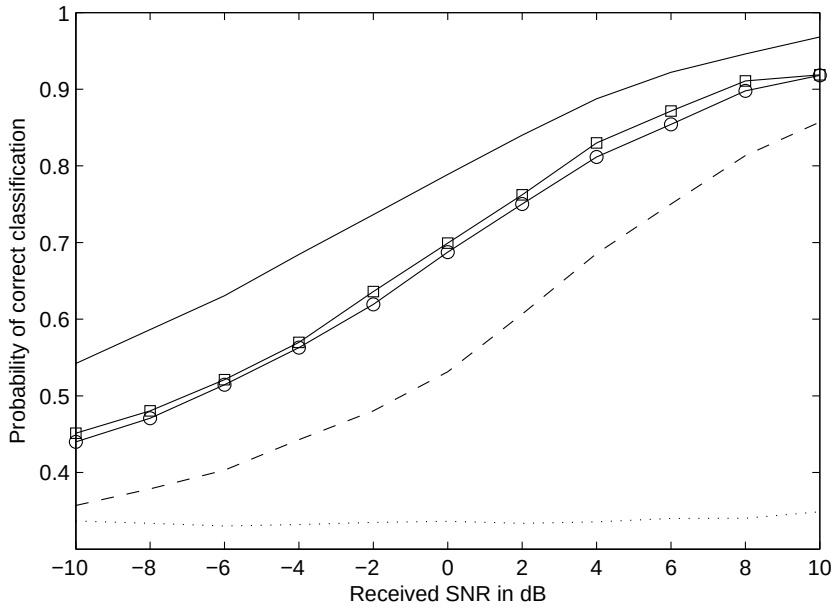

Figure 5.9: Probability of correct classification of the proposed classifier. AR(1) process with $a[1] = -0.6$ and two-term Gaussian mixture model with $\lambda_1 = 0.9$ and $\frac{\sigma_2^2}{\sigma_1^2} = 100$. The number of symbols used for AR parameter estimation and the EM algorithm are 1,000 and 500 (circle) and 500 and 1,000 (square), respectively. The performances of "ideal" classifier (solid line), the classifier developed in Chapter 3 for white non-Gaussian noise with 1,500 symbols for parameter estimation (dashed line), and the classifier developed for Gaussian noise conditions in [1] with 1,500 symbols for parameter estimation (dotted line) are also shown. Number of symbols used for classification is 500.

mixture model parameters were estimated using a variant of the EM algorithm proposed in Chapter 3. Numerical results were presented which show that the proposed classifier easily outperforms the classifier developed under the assumption that the noise process is white.

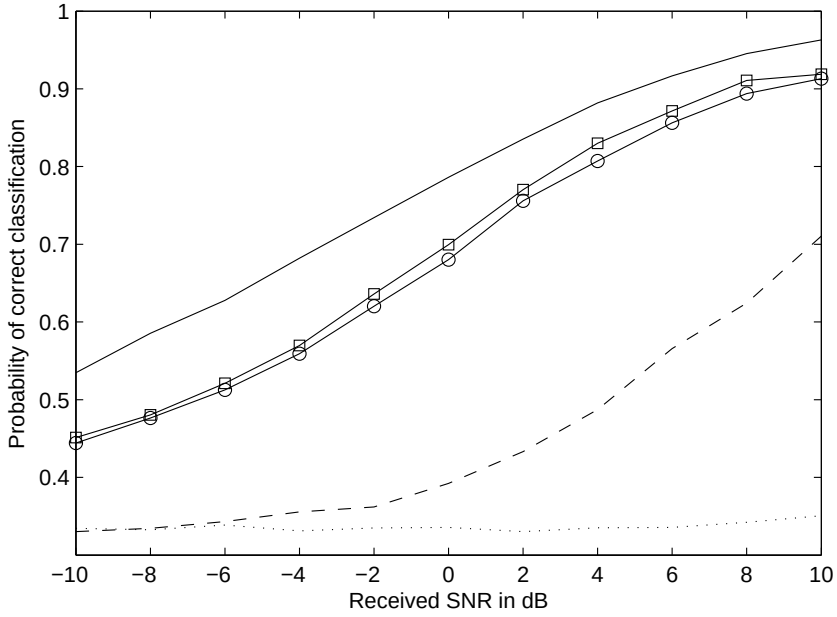

Figure 5.10: Probability of correct classification of the proposed classifier. AR(2) process with $a[1] = -0.25$, $a[2] = -0.53$ and two-term Gaussian mixture model with $\lambda_1 = 0.9$ and $\frac{\sigma_2^2}{\sigma_1^2} = 100$. The number of symbols used for AR parameter estimation and the EM algorithm are 1,000 and 500 (circle) and 500 and 1,000 (square), respectively. The performances of "ideal" classifier (solid line), the classifier developed in Chapter 3 for white non-Gaussian noise with 1,500 symbols for parameter estimation (dashed line), and the classifier developed for Gaussian noise conditions in [1] with 1,500 symbols for parameter estimation (dotted line) are also shown. Number of symbols used for classification is 500.

Table 5.2: Confusion matrix of the proposed classifier. Received SNR = 6dB. AR(2) process with $a[1] = -0.25$, $a[2] = -0.53$ and two-term Gaussian mixture model with $\lambda_1 = 0.9$ and $\frac{\sigma_2^2}{\sigma_1^2} = 100$. The number of symbols used for AR parameter estimation and the EM algorithm are 500 and 1,000 (square), respectively. Number of symbols used for classification is 500.

	BPSK	QPSK	8PSK
BPSK	0.990	0.005	0.005
QPSK	0.061	0.827	0.112
8PSK	0.074	0.124	0.802

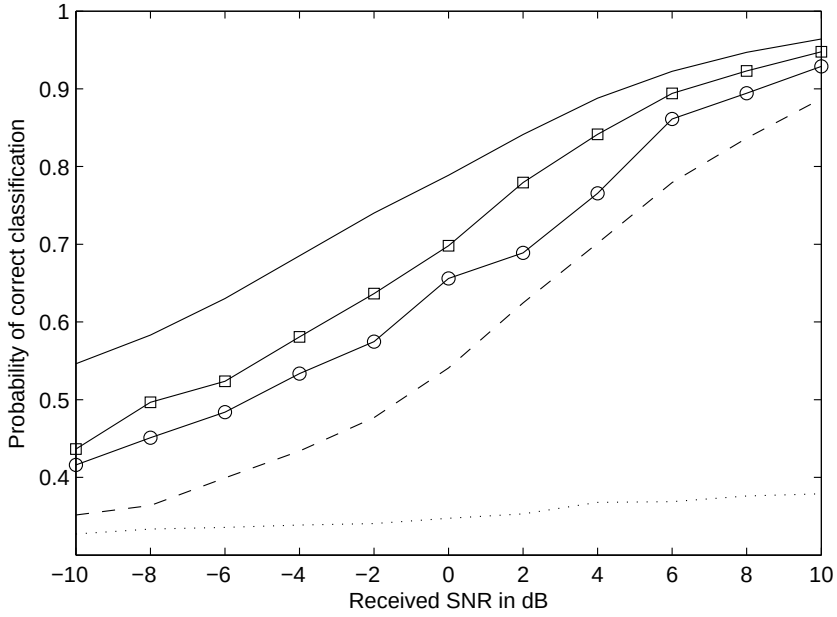

Figure 5.11: Probability of correct classification of the proposed classifier (circle). AR(2) process with $a[1] = -0.25$, $a[2] = -0.53$, three-term Gaussian mixture model with proportions $\{0.098, 0.018, 0.002\}$ and variances $\{\sigma_1^2 = 10, \sigma_2^2 = 50, \sigma_3^2 = 1075\}$. The number of symbols used for AR parameter estimation and the EM algorithm are 250 and 500 (circle) and 1,000 and 5,000 (square), respectively. The performances of "ideal" classifier (solid line), the classifier developed in Chapter 3 for white non-Gaussian noise with 750 symbols for parameter estimation (dashed line), and the classifier developed for Gaussian noise conditions in [1] with 750 symbols for parameter estimation (dotted line) are also shown. Number of symbols used for classification is 500.

Table 5.3: Confusion matrix of the proposed classifier. Received SNR = 6dB. AR(2) process with $a[1] = -0.6$ and two-term Gaussian mixture model with $\lambda_1 = 0.9$ and $\frac{\sigma_2^2}{\sigma_1^2} = 100$. The number of symbols used for AR parameter estimation and the EM algorithm are 500 and 1,000 (square), respectively. Number of symbols used for classification is 500.

	BPSK	QPSK	8PSK
BPSK	0.988	0.004	0.008
QPSK	0.088	0.784	0.128
8PSK	0.056	0.128	0.816

Chapter 6

Conclusions

In this dissertation, different statistical signal processing algorithms were developed for signal detection and modulation classification in radio channels where the additive noise is non-Gaussian. The design of these algorithms is essential since conventional signal processing approaches, developed under the assumption that the noise has a Gaussian distribution, are known to perform poorly in the presence of non-Gaussian noise.

One of the most novel aspect of the proposed algorithms is that they were developed with limited or no prior knowledge of the signal of interest, the state of the fading process, and the parameters which characterize the noise distribution. This was achieved through the design of new techniques for blind estimation of the parameters of the fading process and the non-Gaussian noise distribution. This is as opposed to the majority of detection and classification algorithms available in the literature for non-Gaussian noise environments which either assume to have prior knowledge of these unknown parameters or estimate them using training methods.

In this dissertation, two popular statistical-physical models for non-Gaussian noise were considered, namely the Middleton's Class A model (in which the additive noise is modeled as a Gaussian mixture distribution) and the symmetric α stable model. The Class A model is

used to represent the amplitude of the noise processes which produce "ignorable transients" at the receiver and/or that are spectrally narrower than the receiver's front end. The SαS distribution, on the other hand, is known to accurately model the noise processes that produce "significant transients at the receiver" and/or that are spectrally wider than the receiver's front end. In addition, time-correlated non-Gaussian noise environments were also considered by modeling the noise correlation using an AR process.

More specifically, in Chapter 3, a maximum-likelihood modulation classifier was developed for the classification of digital amplitude-phase modulated signals in flat fading channels with non-Gaussian noise. The additive noise was modeled by a Gaussian mixture distribution. Maximum-likelihood estimates of the unknown fading state and the noise distribution parameters were obtained using a novel EM-algorithm based technique. The classification was performed using a hybrid likelihood ratio test and the performance of the proposed classifier was shown to be significantly better than the conventional likelihood-based classifier developed for Gaussian noise conditions. It was also shown that the performance of the proposed classifier approaches the performance the ideal classifier (assuming prior knowledge of the unknown parameters) as the number of symbols used in the estimation process increases.

In Chapter 4, a new scheme for the detection of digital amplitude-phase modulated signals in radio environments where the additive noise can be modeled using the SαS distribution was presented. The proposed detection scheme involves the use of a matched myriad filter which in known to have certain optimality properties when dealing with SαS distributions. Also derived were new ECF-based methods to estimate the noise distribution parameters in the presence of an unknown signal. The proposed detection scheme was shown to outperform a popular detection scheme which uses a data adaptive ZMNL for mitigating the impulsive noise.

A new technique for the classification of digital amplitude-phase modulated signal in time-correlated non-Gaussian environments was presented in Chapter 5. The pdf of the non-Gaussian noise was modeled using a Gaussian mixture distribution and the time correlation

was modeled using an AR model. The proposed classifier involves the use of a whitening filter which is followed by likelihood-based classification. A new technique for estimating the AR parameters of the non-Gaussian noise (which are also the coefficients of the whitening filter) using a H_∞ filter based dual estimation procedure was also derived. A hybrid likelihood ratio test was proposed at the output of the whitening filter, which uses the estimates of the fading state and the Gaussian mixture model parameters obtained using the ECM algorithm developed in Chapter 3. Numerical results presented showed that the proposed classifier offers a significant gain in performance over a classifier developed under the assumption that the non-Gaussian noise process is white.

In addition to the design of algorithms for signal detection and classification in non-Gaussian noise environments (Chapters 3-5), in Appendix A, a new method to improve the reliability of signal detection in spectrum sensing applications using radio collaboration was also developed. In the proposed method, a spectrum sensing decision is made by *weighting* the sensing data provided by multiple radios based on their reliability. The novelty of the proposed method lies in exploiting the correlation of the sensing data provided by each radio over multiple sensing intervals to keep the complexity of the weight estimation process at acceptable levels. It was demonstrated that the proposed method offers an excellent tradeoff between reliability and complexity when compared to traditional collaborative spectrum sensing approaches.

The algorithms proposed in this dissertation can be used for signal detection and classification in a variety of applications. For example, these algorithms can be used in DSA systems for accurate identification of potential spectral opportunities. In military communications, these algorithms can be used for the detection and interception of hostile signals which are extremely important in electronic warfare and surveillance systems. Additionally, the throughput of adaptive modulation and coding systems and systems using multi-mode software radios can be improved through the use of blind modulation classification algorithms like the ones presented in this dissertation.

Appendix A

Collaborative Spectrum Sensing Based on a New SNR Estimation and Energy Combining Method

Different spectrum regulatory authorities and commercial companies have been promoting the concept of DSA for its potential to allow for a more efficient use of the radio spectrum [3], [31]. In DSA systems, improvements in spectral efficiency is achieved by allowing secondary users to access vacant primary user frequency bands on a dynamic and non-interfering basis. One of the possible enabling technologies for these systems is *spectrum sensing*, which involves the use of fast and accurate algorithms for signal detection to identify the vacant primary user spectrum. The reliability of spectrum sensing can be increased by using collaborative methods which exploit spatial diversity by combining, in an advantageous manner, the sensing data provided by multiple radios [139–150].

As shown in [139–141], among others, in order for collaborative spectrum sensing to be effective, the data provided by each radio should be *weighted* in proportion to their reliability during data combining. This is because in realistic scenarios, the received SNR value at each radio can vary greatly, for example, due to path-loss and shadowing. The

processing requirement of the weight estimation process (which may require SNR estimation, for example), along with the fact that the weights must be constantly updated (due to the time-varying nature of wireless channels), however, can make weighted data combining infeasible in practical systems. For this reason, it is commonly assumed in the literature that propagation conditions are approximately the same for all radios and, consequently, the sensing data provided by all radios are weighted equally. See, for example, [142–145].

In the following, a practical collaborative sensing method is presented which exploits the correlation of the sensing data provided by a given radio in successive sensing intervals to keep the requirements of weighted data combining at acceptable levels. In the proposed method, during the period when a primary user is detected, the costly weight estimation process is performed only every $N_{SI}+1$ sensing intervals (and not in every interval, as in conventional collaborative methods). During the same period, in the intervals when SNR estimation is not performed, the radios only obtain received energy measurements, which are combined with weights determined by using the *most recent estimate* of each radio's received SNR value. It is demonstrated that the proposed method offers an excellent trade-off between reliability and complexity.

A.1 System Model

Spectrum sensing can be modeled as a binary hypothesis testing problem with hypothesis $\mathcal{H}_0$ and $\mathcal{H}_1$ denoting the absence and presence of a primary user, respectively. In the proposed model, the low-pass equivalent of the k^{th} sample of the received signal at the m^{th} radio is written as

$$\begin{aligned} \mathcal{H}_0 &: r_m[k] = w_m[k] \\ \mathcal{H}_1 &: r_m[k] = \alpha_m s[k] + w_m[k], \end{aligned} \tag{A.1}$$

for $k = 1, 2, \ldots, K$ and $m = 1, 2, \ldots, M$; that is, the observation window of each radio has K samples and sensing is performed with M radios. Sampling at the symbol rate is assumed.

In (A.1), the complex variables $\{s[k]\}_{k=1}^{K}$ represent the (deterministic and unknown) primary user's signal[9], and the complex variables $\{w_m[k]\}_{k=1}^{K}$ represent the additive noise, assumed to be independent and identically distributed (i.i.d.) circularly symmetric Gaussian random variables with zero mean and variance σ_n^2. Also in (A.1), α_m is a (deterministic and unknown) complex variable used to represent both the flat fading experienced by the signal arriving at the m^{th} radio, and the unknown power and carrier phase of the transmitted signal. For this reason, without loss of generality, $E[|s[k]|^2]$ is assumed to be unitary. Note from (A.1) that the channel is assumed to remain constant over the duration of the K observed samples.

Using the Central Limit Theorem [152], it can be shown that the received energy measurement of each radio, $E_m = \sum_{k=1}^{K} |r_m[k]|^2$, can be approximated[10] to have the following conditional probability density function (pdf) for a suitably large K [147],

$$\mathcal{H}_0 \;:\; E_m \sim \mathcal{N}(K\sigma_n^2, K\sigma_n^4) \tag{A.2}$$

$$\mathcal{H}_1 \;:\; E_m \sim \mathcal{N}\left(K(1+\gamma_m)\sigma_n^2, K(1+2\gamma_m)\sigma_n^4\right), \tag{A.3}$$

where the notation $E_m \sim \mathcal{N}(\mu, \sigma^2)$ is used to represent that E_m is a (real) Gaussian random variable with mean value μ and variance σ^2. In (A.3), $\gamma_m = |\alpha_m|^2/\sigma_n^2$, is the received SNR value of the primary user signal at the m^{th} radio. (As shown in [147], the mean and variance

[9]The assumption in (A.1) that only one primary user exists under $\mathcal{H}_1$ is valid for potential usage scenarios. For example, 802.22 radios that rely on sensing must be able to reliably detect TV signals. Spatial multiplexing usually guarantees that only the transmission of a single TV station can be detected (or that it is significantly stronger than signals from other co-channel stations) in a particular channel at a given location. In applications in which this assumption does not hold and an unknown number of primary users are present under $\mathcal{H}_1$, an appropriate detection technique (for example, based on Random Set Theory [151]) would have to be used by the collaborating radios.

[10]To verify this approximation, the estimated pdf of energy measurements obtained numerically through simulation was compared to the Gaussian pdf given by (A.3) by using the goodness of fit statistic R-square (also known as the coefficient of determination) [153]. The value of R-square (which is a measure of the correlation between two pdfs) was found to be greater than 0.99 for values of K as low as 50 for a wide range of SNR values of practical interest, thus validating the approximation.

values of the pdf given by (A.3) are a function of the primary user's modulation scheme. The values used in (A.3) correspond to a PSK signal with modulation order greater than 2.)

A.2 Collaborative Spectrum Sensing Methods

In the literature, collaborative sensing methods based on energy combining are popular because of their ease of implementation and low complexity. See, among others, [139–146], [148], [150]. In these methods, each radio measures the received energy of the primary user signal, and then relays this information to a fusion center (which may be one of the collaborating radios or a separate entity). At the fusion center, a final decision metric, D, is formed by combining the received energy measurements in an advantageous manner. In this section, three well-known energy combining-based collaborative methods are analyzed. Based on this analysis, a new collaborative spectrum sensing method is proposed in Section A.3.

Equal gain combining (EGC): In this method, the fusion center sums the received energy measurements and compares the result to a threshold. The decision rule is given by

$$D_{EGC} = \sum_{m=1}^{M} E_m \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{EGC}. \tag{A.4}$$

In the EGC method, all radios are weighted equally irrespective of the reliability of their sensing data. Therefore, it is expected that this method will have poor performance unless the received SNR values of the collaborating radios are approximately the same. Despite this fact, this method is widely adopted because it does not require weight estimation [142–145].

Weighted linear combining (WLC): The decision metric in this method is a weighted linear combination of the received energy measurements. The decision rule can be written as

$$D_{WLC} = \sum_{m=1}^{M} w_m E_m \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{WLC}. \tag{A.5}$$

As determining the *optimal* set of weights $\{w_m\}_{m=1}^M$ that minimize the probability of detection error of the WLC method is not straightforward, a heuristic technique was proposed in [146] which "provides near-optimal solutions for general systems". For the considered system model, assuming that the collaborating radios' observations are conditionally independent (given the hypothesis H_i), the weights, as proposed by this heuristic technique, are given by

$$w_m = \frac{\gamma_m}{1+2\gamma_m}. \tag{A.6}$$

Optimal combining (OC): This method provides the optimal detection performance for the case in which the radios relay energy measurements to the fusion center. Assuming that the collaborating radios' observations are conditionally independent (given the hypothesis H_i), the likelihood ratio test can be written as $\left(\prod_{m=1}^M p(E_m|\mathcal{H}_1)\right) / \left(\prod_{m=1}^M p(E_m|\mathcal{H}_0)\right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta$ [2]. Using (A.2) and (A.3), and after some algebraic manipulation, the decision rule is obtained as[11]

$$D_{OC} = \sum_{m=1}^{M} w_{m,1} E_m^2 + w_{m,2} E_m \gtrless \eta_{OC}, \tag{A.7}$$

where

$$w_{m,1} = \frac{2\gamma_m}{1+2\gamma_m} \quad \text{and} \quad w_{m,2} = \frac{-2\gamma_m K \sigma_n^2}{1+2\gamma_m}. \tag{A.8}$$

As seen in (A.5) and (A.7), in the WLC and OC methods, *the obtained energy measurements are weighted in proportion to their reliability.* For this reason, the WLC and OC methods have the potential to significantly outperform the EGC method in practical sce-

[11] Under the Neyman-Pearson criterion, the decision rule's threshold is set to yield the desired probability of false alarm P_{fa}. For the model presented in Section A.1, the decision metric of linear combining methods, including the EGC and WLC, is Gaussian under each hypothesis ($\mathcal{N}(\mu_i, \sigma_i^2)$ under $\mathcal{H}_i$). In this case, the threshold η_{NP} and the probability of detection P_d are obtained as $\eta_{NP} = \mu_0 + \sigma_0 Q^{-1}(P_{fa})$ and $P_d = Q\left((\eta_{NP} - \mu_1)/\sigma_1\right)$, respectively, where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\frac{u^2}{2}) \mathrm{d}u$. For the OC method, the decision metric, (A.7), is given by a quadratic form in Gaussian variables. While the moment generating function of quadratic forms in Gaussian variables has a closed-form expression [154], closed-form expressions for the pdf are only available for some particular cases of the quadratic form [2], [154]. To the best of the author's knowledge, there is no closed-form expression for the pdf of the decision metric in (A.7) when the variables have pdfs given by (A.2) and (A.3). In this case, a numerical algorithm (such as the ones proposed in [155] and [156], among others) should be used to obtain P_d and P_{fa} and to determine η_{NP}.

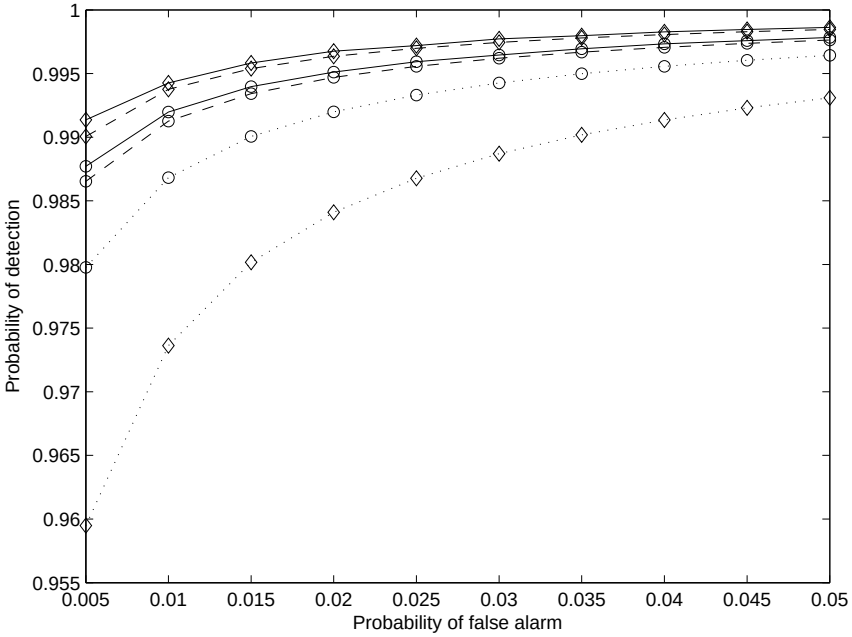

Figure A.1: Performance of the EGC (dotted lines), WLC (dashed lines), and OC methods (solid lines) for $M = 2$ (SNR values equal to -1 and -5 dB, circle) and 4 (-1, -5, -8 and -10 dB, diamond) radios. The observation window of each radio has a total of $K = 50$ samples.

narios (at the cost of having to estimate the combining weights, which, as seen in (A.6) and (A.8), requires SNR estimation).

The performance of the three considered methods is presented in Fig. A.1. As expected, it is seen in this figure that the weighted combining methods outperform the EGC method for the two cases considered. It is also seen in Fig. A.1 that, while it is expected that the performance of collaborative sensing would improve with an increase in the number of radios, the performance of the EGC method is actually worse when four radios are used compared to that of two radios in the scenario considered. This is because the reliability of the two added radios (SNR values of -8 and -10 dB) is low compared to that of the existing radios (-1 and -5 dB), and the EGC method weights all radios equally. As the WLC and OC methods weight the measurements in proportion to their reliability, this behavior is not observed in Fig. A.1 for the weighted methods.

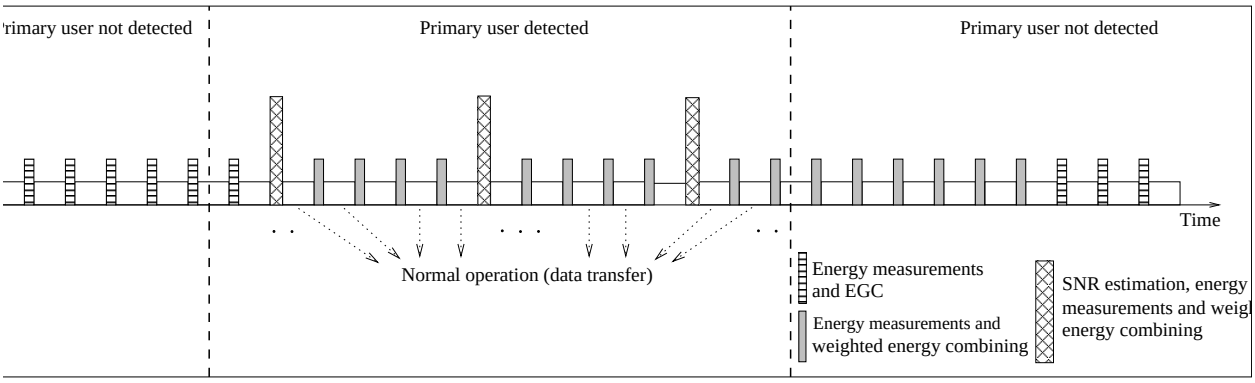

Figure A.2: Illustration of the operation of the proposed collaborative spectrum sensing method. In this figure, N_{SI} and N_{off} are equal to 4 and 8, respectively.

The improved performance of weighted combining results from the fact these methods use the *received SNR value at each radio as a reliability measure* of the sensing data it provides. Unfortunately, SNR estimation is a time and processing intensive procedure (in comparison with measuring the received signal energy, for example), which requires the estimation of signal and noise parameters [157]. For example, a maximum-likelihood-based SNR estimator [157] requires the radio to first synchronize with the received signal in order to estimate the transmitted symbols and then estimate the received signal power. The noise power is then estimated by subtracting the estimated signal power from the total received power. Therefore, the time and processing requirements of the weight estimation process can make weighted data combining infeasible for practical systems. In the next section, a practical method is proposed, which keeps the requirements of weighted data combining at acceptable levels, providing an excellent trade-off between reliability and complexity compared with conventional collaborative methods.

(It is worth noting that while the use of SNR estimates as a reliability measure in collaborative spectrum sensing has been previously considered [139], [140], [146], [148], the problem of SNR estimation in the context of collaborative sensing is still underdeveloped. The proposed method proposed offers a possible solution to this problem.)

A.3 Proposed Collaborative Spectrum Sensing Method

In wireless systems, signals travel from transmitter to receiver over multiple reflective paths, causing fluctuations in the received signal's amplitude. Due to changes in the channel (resulting from the mobility of the radios, for example), the random signal fluctuations caused by *multipath propagation* are time-variant. The time correlation of such fluctuations is partially characterized by the so-called coherence time, which is an estimate of the expected time during which the multipath fading can be considered as constant. In addition, the random signal power attenuation due to *path loss* and *shadowing* can also be correlated in time. For example, when the random change in the position of a receiver in a given time interval is relatively small compared to its distance to the transmitter, the power attenuation due to path loss is approximately constant during this time interval. Also, correlated shadow fading has been reported to occur in various scenarios of interest, including cellular systems, indoor environments, and indoor-outdoor links [158].

The proposed collaborative spectrum sensing method exploits the correlation of the sensing data provided by a given radio in successive sensing intervals to keep the requirements of weighted data combining at acceptable levels. In the proposed method, during the period when a primary user is detected (that is, when it is possible to estimate the received SNR value and thus the reliability of the data provided by each radio), the costly weight estimation process is performed only every $N_{SI} + 1$ sensing intervals (and not in every interval, as in conventional methods). During the same period, in the intervals when SNR estimation is not performed, the collaborating radios only obtain energy measurements, which are combined with weights determined by using the *most recent estimate* of each radio's received SNR value.

The proposed collaborative spectrum sensing method, which has its operation illustrated in Fig. A.2, consists of the following steps:

1. When collaborative sensing is initiated, each radio measures the received energy of the

primary user signal and relays its measurement to the fusion center. SNR estimates are not obtained at this point in time because the sensing network has not made a decision yet whether a primary user is present (SNR estimates would be valid) or not (estimates would be invalid). (Note that in the absence of a primary user signal, the SNR estimate is *ideally* equal to zero.) The fusion center makes a decision by using the EGC method.

2. If a primary user signal is not detected, the collaborating radios obtain received energy measurements in all sensing intervals until a primary user is detected (and the method goes to step 3). During this time, EGC is used by the fusion center. (Recall that if a primary user signal is not present, the observations used by the radios are thermal noise samples. In this case, it is impossible for the radios to estimate their "reliability" if a primary user was present, and weighted data combining is impractical.)

3. If or when a primary user signal is detected, each radio relays to the fusion center its estimated received SNR value and measured received energy in the sensing interval following the detection. Using the received SNR values, the combining weights are obtained by using (A.8). Given the weights and measured received energy values, the fusion center uses the OC method, with decision rule given by (A.7), to track the presence of the primary user. As shown in Fig. A.2, in the N_{SI} sensing intervals after an interval in which SNR estimates are obtained, the radios only obtain received energy measurements. In these sensing intervals, the fusion center uses (A.7) with combining weights determined by using the *most recent estimate* of each radio's received SNR value. In the proposed method, SNR estimation is not performed in all sensing intervals in order to keep the computational complexity of weighted data combining (which, as discussed in Section A.2, is dominated by the costly SNR estimation process) at acceptable levels. SNR estimation (along with the update of the combining weights) is performed every $N_{SI} + 1$ sensing intervals.

4. If the primary user is no longer detected during step 3, the most recent combining

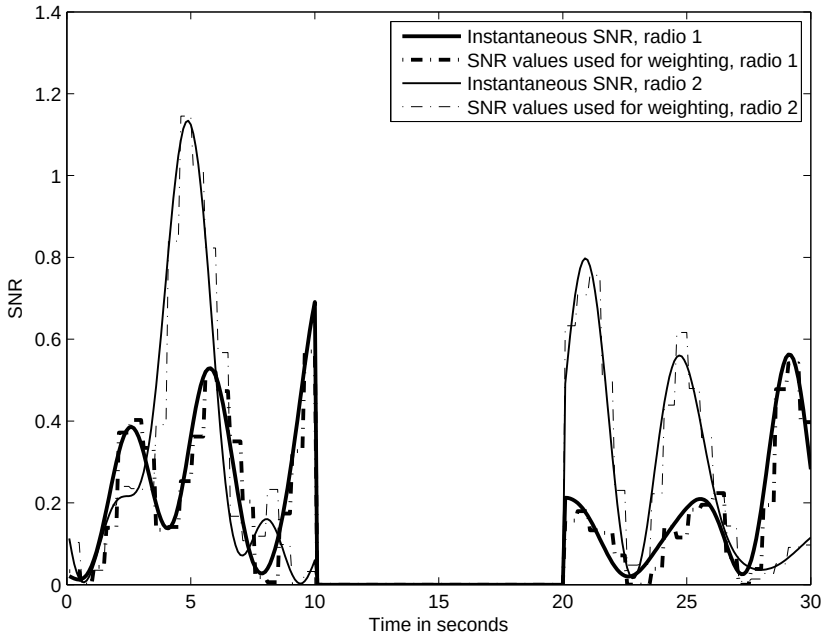

Figure A.3: Instantaneous and estimated SNR values for a particular realization of the proposed simulation environment. The primary user is off from 10s to 20s. The sensing interval and N_{SI} are equal to 100 ms and 4, respectively.

weights are used for a pre-determined number of sensing intervals N_{off}, after which they are considered obsolete. The algorithm then reverts to step 2.

As the data provided by a given radio in successive sensing intervals is likely to be correlated in scenarios of practical interest, the used weights will not be significantly different from the "ideal" ones if system parameters (sensing interval, N_{SI}, and N_{off}) are properly chosen. For example, using the simulation model described in Section A.4, the estimated SNR value (updated every $N_{SI}+1$ sensing intervals) and the true SNR value for a particular realization is shown in Fig. A.3. It can be seen that, due to the channel correlation, the estimated SNR value is approximately the same as the true SNR value until a new estimate is obtained. As a result, as shown in Section A.4, the proposed method's performance closely approaches that of the "ideal" OC method.

It is worth noting that OC was used in the proposed method because it provides optimal

detection performance when radios relay conditionally independent energy measurements to the fusion center. However, any other weighted combining scheme, such as the WLC, could have been used. Finally, it can also be noted that the proposed method fits well with the (single radio) sensing procedure defined by the IEEE 802.22 standard, which has two stages: fast and fine sensing [31]. In the fast sensing stage, which has a short sensing window ($\leq$ 1ms), a low complexity method, such as energy detection, is used. The fine sensing stage, which has a longer sensing window ($\approx$ 25ms), uses more complex methods, such as cyclic feature detection. Therefore, the different sensing intervals of the proposed method could be incorporated into the fast (energy measurements only) and fine (SNR estimation) stages of the IEEE 802.22 framework.

A.4 Numerical Results

In order to analyze the performance of the proposed method, a collaborative sensing scenario with a stationary primary user and M mobile radios is considered. The sensing data obtained by the radios is transmitted to the fusion center through a low-capacity channel that, due to the use of forward error correction and protocols such as ARQ, is assumed to be error-free. The movement of the radios is modeled by a random walk process [159] with a speed uniformly chosen between 0 and 1 m/s. The received SNR value is assumed to vary, due to path loss, between 0 and -15 dB depending on the distance from the primary user. Also, each radio is assumed to experience time-varying Rayleigh fading with coherence time equal to 1s [160]. The primary user is assumed to transmit a QPSK signal at 500 kbps. SNR estimates and received energy measurements are obtained by using a total of 1000 and 50 samples, respectively.

For SNR estimation, a variant of the EM is used. Specifically, the ECM algorithm derived in Section 3.3, with the number of terms in the mixture model (N) and the number of possible modulation schemes (M) set equal to one is used. This algorithm is used because

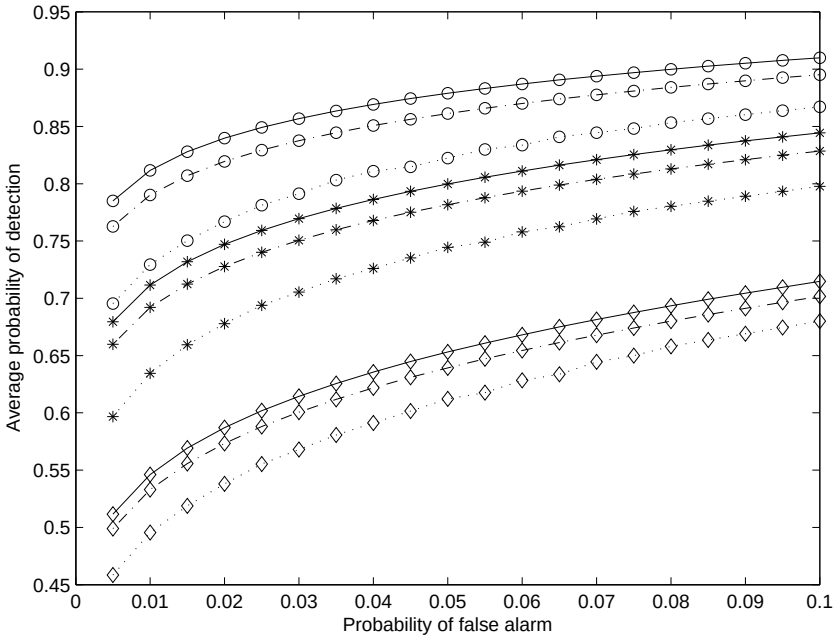

Figure A.4: Performance of the propose (dash-dot lines), EGC (dotted lines) and OC (solid lines) methods for $M = 2$ (diamond), 3 (star) and 4 (circle) radios. The sensing interval and N_{SI} are equal to 100 ms and 9, respectively.

it is a practical implementation of the (optimal) maximum-likelihood-based SNR estimator. However, the proposed method does not require the use of a particular estimator, and any other algorithm, such as the ones presented in [157], could have been used.

Note that while the channel was assumed to remain constant over the observation window in Section A.1, the simulated channels are actually time-varying. However, the simulated channel of a given radio is time-correlated, and it is approximately constant during the sensing interval (the normalized maximum absolute deviation of the channel gain magnitude was found to be on average less than 0.1%). For this reason, the assumption made in the system model is valid.

The performance of the proposed, EGC, and OC methods is shown in Fig. A.4 for different numbers of radios. It is seen in this figure that the performance of the proposed method closely approximates that of the OC method for all cases, despite the fact that the proposed

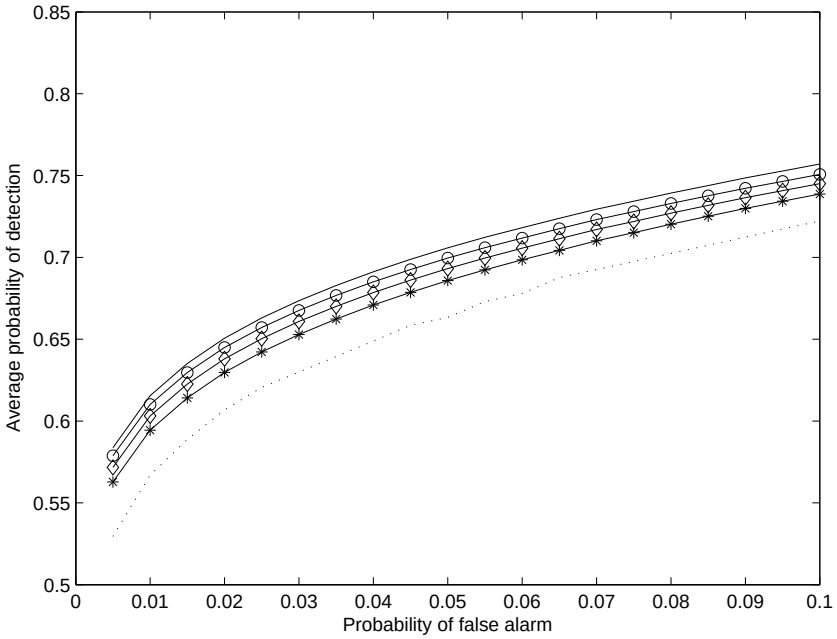

Figure A.5: Performance of the proposed method for N_{SI} equal to 4 (circle), 9 (diamond), and 19 (star). The sensing interval is equal to 100 ms. For comparison, the performance of the EGC (dotted line) and OC (solid line) methods are also shown. The collaborative network consists of $M = 4$ radios.

method performs SNR estimation, in this particular implementation, in only 10% of the sensing intervals. In the remaining 90% of the intervals, the proposed method requires only (low-complexity) received energy measurements. Furthermore, there is a noticeable difference in performance between the proposed and EGC methods, and this difference increases as more radios are used.

The frequency in which SNR estimation is performed affects both the performance and the computational complexity of the proposed method. Fig. A.5 shows the performance of the proposed method for different N_{SI} values. As expected, it is seen that the difference in performance between the proposed and OC methods decreases with an increase in the frequency at which SNR estimation is performed. However, this performance gain comes at the cost of more frequent SNR estimations.

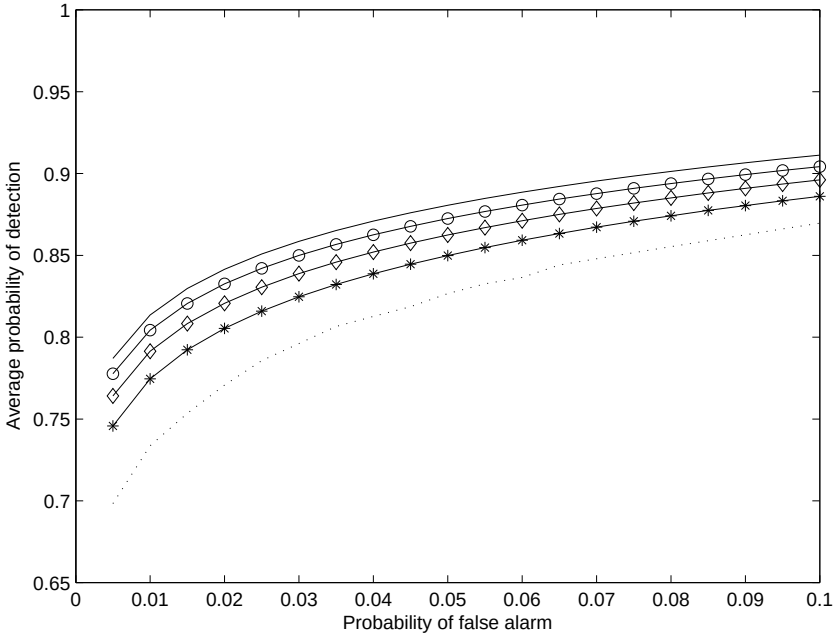

Figure A.6: Performance of weighted energy combining during the 1st (circle), 5th (diamond), 9th (star) sensing intervals after an interval in which SNR estimation is performed. For comparison, the performances of the EGC (dotted line) and OC (solid line) methods are also shown. The number of radios, sensing interval, and N_{SI} are equal to 4, 100 ms, and 9, respectively.

Fig. A.6 shows the performance of the proposed method in the first, fifth, and ninth intervals after an interval in which SNR estimation was performed (that is, the weights used during data combining were obtained by using SNR values estimated one, five, and nine intervals ago, respectively). As the time separation between the sensing interval in which the SNR estimates were obtained and the one in which data combining is performed increases, the difference between the "ideal" weights and the "dated" weights tends to increase and hence the sensing performance degrades, as seen in this figure. However, on average, the performance of the proposed method closely approximates that of the OC method, as shown in Figs. A.4 and A.5.

As discussed in Section A.3, OC was used in the proposed method because it provides optimal detection performance when radios relay conditionally independent energy measure-

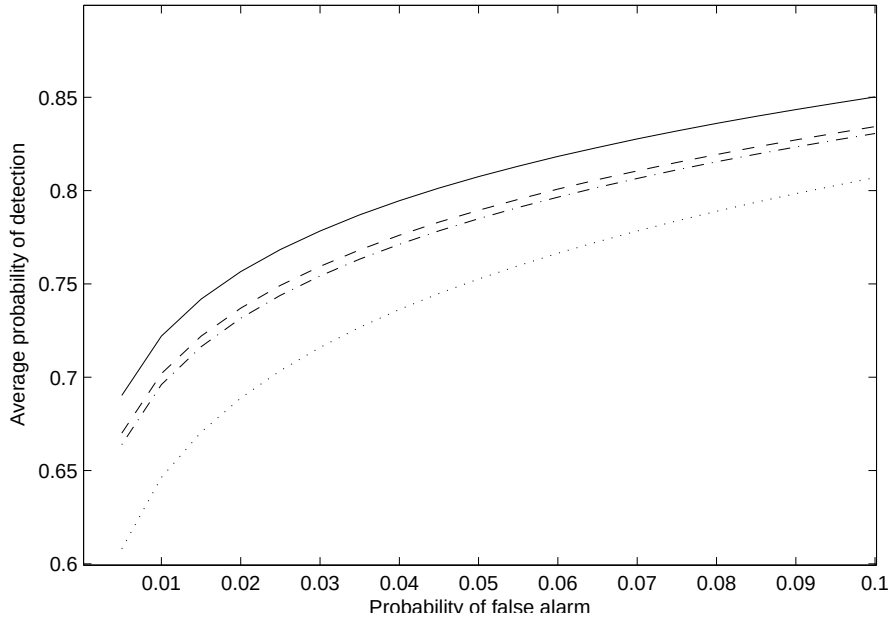

Figure A.7: Performance of proposed method when weighted energy combining is performed by using the OC (dashed line) and WLC (dash-dot line) methods. For comparison, the performances of the EGC (dotted line) and OC (solid line) methods are also shown. The number of radios, sensing interval, and N_{SI} are equal to 3, 100 ms, and 9, respectively.

ments to the fusion center. However, any other weighted combining scheme, such as the WLC, could have been used. The performance of proposed method when weighted energy combining is performed by using the OC and WLC combining methods is shown in Fig. A.7. It is seen that the proposed method using either OC or WLC performs very well, with the OC-based method marginally outperforming the WLC-based method, as expected.

A.5 Conclusions

A practical collaborative sensing method was proposed which exploits the correlation of the data provided by a given radio in successive sensing intervals. In the proposed method, during the period when a primary user is detected, the costly weight estimation process is performed only every $N_{SI}+1$ sensing intervals. Weighted energy combining is performed by using the most recent estimate of each radio's SNR as a reliability measure of the sensing

data it provides. Numerical results show that the performance of the proposed method closely approximates that of the OC method if system parameters are properly chosen.

A.6 Acknowledgments

This chapter, in part, is a reprint of the material as it appears in "Collaborative spectrum sensing based on a new SNR estimation and energy combining method," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 4024-4029, Oct. 2011. The dissertation author was the primary author and Prof. Claudio da Silva directed and supervised the research which forms the basis of this chapter.

Appendix B

Asymptotic Normality of the Matched Myriad Filter Output

Finite sample properties of M-estimators, such as the matched myriad filter presented in Section 4.2, are typically analytically intractable, and hence asymptotic methods play an important role in the analysis of their behavior [161]. In the following, the asymptotic behavior of the matched myriad filter's output is discussed under the hypothesis $\mathcal{H}_1$ for the rectangular and the SRRC pulse shapes. The corresponding result for the null hypothesis $\mathcal{H}_0$ follows directly from the provided analysis. As described in Section 4.3.2, we consider the oversampled sequence $\{r_i^\tau\}_{i=1}^N$, corresponding to an observation window beginning at a time instant τ, given by

$$r_i^\tau = S_i^\tau + n_i, \text{ where } S_i^\tau = \alpha_{ch} \sum_{m=-\infty}^{\infty} s_m p\left(\tau + (i-1)T_s - mT - \epsilon T\right). \tag{B.1}$$

In this Appendix, we show that the output of the matched myriad filter, given by (4.17), is, or can be approximated to be, asymptotically Gaussian.

B.1 Rectangular pulse shape

For the rectangular pulse shape case, the matched pulse coefficients $\{p_i\}_{i=1}^{N}$, given by $p_i = p(t)|_{t=iT_s}$, are all equal to 1. Therefore, the matched myriad filter output is given by

$$\hat{\theta}_N^{\tau} = \arg\min_{\theta} \sum_{i=1}^{N} \log\left[K^2 + |r_i^{\tau} - \theta|^2\right]. \tag{B.2}$$

It should be noted that since the duration of the sliding window is equal to the symbol period for rectangular pulses, there can be at most one symbol transition in a window. Therefore, we consider two different cases, one in which no symbol transition occurs in the observation window, and the other where there is exactly one symbol transition in the observation window.

B.1.1 Asymptotic normality in the no transition-case

Since the pulse is rectangular and there are no symbol transitions, S_i^{τ} in (B.1) is the same for all i (say S^{τ}). The problem of calculating $\hat{\theta}_N^{\tau}$ in (B.2) reduces to calculating the M-estimate of the location of a set of i.i.d. random variables r_i^{τ} using the cost function $\log(K^2 + x^2)$. This is a well researched problem in robust statistics, and it can be shown that, under mild regularity conditions, the M-estimates obtained are consistent and asymptotically Gaussian with distribution [162]

$$\hat{\theta}_N^{\tau} \sim \mathcal{N}\left(S^{\tau}, \frac{\int \psi^2(x) f_X(x) dx}{N(\int \psi'(x) f_X(x) dx)^2}\right), \tag{B.3}$$

where $\psi(x) = d\log(K^2 + x^2)/dx = 2x/(K^2 + x^2)$ and $f_X(x)$ is the noise pdf. It can be observed that the asymptotic variance in (B.3) is independent of the information component $s(t)$, and is in fact the same as that of the signal absence ($\mathcal{H}_0$) case – see (8) in [99], for example.

B.1.2 Asymptotic normality in the one transition-case

While the asymptotic normality of the matched myriad filter output in the no transition-case is a well known result, to the best of the authors' knowledge, no such result is available for the one transition-case. In this case, given that a symbol transition occurs at a time κT $(0 < \kappa < 1)$ from the beginning of the sliding window (that is, after κN samples), the samples $\{r_i^\tau\}_{i=1}^N$ are given by $r_i^\tau = S_0^\tau + n_i$, for $i = 1, 2, ..\kappa N$, and $r_i^\tau = S_1^\tau + n_i$, for $i = \kappa N + 1, ..., N$, where S_0^τ and S_1^τ denote the information component before and after the symbol transition, respectively. The output $\hat{\theta}_N^\tau$, obtained by equating the first order derivative of the expression on the right hand side of (B.2) to zero [162], is given by

$$\frac{1}{N}\sum_{i=1}^{N}\psi\left(r_i^\tau - \hat{\theta}_N^\tau\right) = \frac{\kappa}{\kappa N}\sum_{i=1}^{\kappa N}\psi\left(r_i^\tau - \hat{\theta}_N^\tau\right) + \frac{1-\kappa}{(1-\kappa)N}\sum_{i=\kappa N+1}^{N}\psi\left(r_i^\tau - \hat{\theta}_N^\tau\right) = 0, \quad \text{(B.4)}$$

where $\psi(x) = 2x/(K^2 + x^2)$. Note that $\left\{r_i^\tau - \hat{\theta}_N^\tau\right\}_{i=1}^{\kappa N}$ and $\left\{r_i^\tau - \hat{\theta}_N^\tau\right\}_{i=\kappa N+1}^{N}$ are two different sets of i.i.d. random variables. Applying the weak law of large numbers [161] to the summations in (B.4), as $N \to \infty$, we have $\kappa E[\psi(n + S_0^\tau - \hat{\theta}_N^\tau)] + (1-\kappa)E[\psi(n + S_1^\tau - \hat{\theta}_N^\tau)] = 0$, whose solution we denote by θ_τ. Considering the Taylor series expansion of (B.4) about θ_τ, we have

$$\frac{1}{N}\sum_{i=1}^{N}\psi\left(r_i^\tau - \theta_\tau\right) + \frac{(\hat{\theta}_N^\tau - \theta_\tau)}{N}\sum_{i=1}^{N}\psi'\left(r_i^\tau - \theta_\tau\right) + \frac{(\hat{\theta}_N^\tau - \theta_\tau)^2}{2N}\sum_{i=1}^{N}\psi''\left(r_i^\tau - \theta_\tau\right) = 0. \quad \text{(B.5)}$$

Rearranging the terms of (B.5) yields

$$\sqrt{N}(\hat{\theta}_N^\tau - \theta_\tau) = \frac{-\frac{1}{\sqrt{N}}\left[\sum_{i=1}^{\kappa N}\psi\left(r_i^\tau - \theta_\tau\right) + \sum_{i=\kappa N+1}^{N}\psi\left(r_i^\tau - \theta_\tau\right)\right]}{\frac{1}{N}\sum_{i=1}^{N}\psi'\left(r_i^\tau - \theta_\tau\right) + \frac{(\hat{\theta}_N^\tau - \theta_\tau)}{2N}\sum_{i=1}^{N}\psi''\left(r_i^\tau - \theta_\tau\right)}. \quad \text{(B.6)}$$

Consider the denominator on the right-hand side of (B.6). The second term approaches 0 for large N since $\hat{\theta}_N \to \theta_\tau$, while the first term can be written as

$$\frac{1}{N}\sum_{i=1}^{N}\psi'\left(r_i^\tau - \theta_\tau\right) = \frac{\kappa}{\kappa N}\sum_{i=1}^{\kappa N}\psi'\left(n_i + S_0^\tau - \theta_\tau\right) + \frac{1-\kappa}{(1-\kappa)N}\sum_{i=\kappa N+1}^{N}\psi'\left(n_i + S_1^\tau - \theta_\tau\right). \quad \text{(B.7)}$$

Similar to (B.4), for large N, (B.7) can be written as $\kappa E[\psi'(n+S_0^\tau-\theta_\tau)]+(1-\kappa)E[\psi'(n+S_1^\tau-\theta_\tau)]$ using the weak law of large numbers. (Note that in making these comments, we assume certain mild regularity conditions to hold, such as the boundedness of $\frac{1}{2N}\sum_{i=1}^{N}\psi''\left(r_i^\tau - \theta_\tau\right)$ and the asymptotic consistency of $\hat{\theta}_N^\tau$.) For the numerator, as previously observed, the elements constituting each of the two summations are two sets of i.i.d. random variables. Using the Central Limit Theorem, the sum of these sequences, and thereby the numerator itself, is asymptotically Gaussian. Also, the mean of the numerator is equal to $\sqrt{N}[\kappa E[\psi\left(n + S_0^\tau - \theta_\tau\right)] + (1-\kappa)E[\psi\left(n + S_1^\tau - \theta_\tau\right)] = 0$. Based on these observations, we have that, for large N, the denominator and the numerator of (B.6) are a constant and a zero mean Gaussian random variable, respectively. On further simplification, the asymptotic distribution of $\hat{\theta}_N^\tau$ is given by

$$\hat{\theta}_N^\tau \sim \mathcal{N}(\theta_\tau, V), \text{ where } V = \frac{\kappa\text{var}[\psi(n + S_1 - \theta_\tau)] + (1-\kappa)\text{var}[\psi(n + S_0 - \theta_\tau)]}{N\left(\kappa E[\psi'(n + S_1 - \theta_\tau)] + (1-\kappa)E[\psi'(n + S_0 - \theta_\tau)]\right)^2}. \quad \text{(B.8)}$$

In this analysis, we have shown that the output of a matched myriad filter $\hat{\theta}_N^\tau$ for the rectangular pulse shape is asymptotically Gaussian. While the proposed analysis is presented assuming that the numbers are real, an extension to the case of complex numbers follows directly.

B.2 SRRC pulse shape

For this pulse shape, the samples $\{r_i^\tau\}_{i=1}^N$ are no longer i.i.d. due to ISI and to the fact that the observed sequence is not stationary. For this reason, the analysis developed for rectangular

pulses does not apply to this case. However, the asymptotic properties of M-estimators of the form (4.17) have been extensively studied in the literature when the observed signal is given by

$$r_i^\tau = \theta p_i + n_i,\ i = 1, 2, 3, ..., N, \tag{B.9}$$

where $\{n_i\}_{i=1}^N$ are i.i.d. variables. The *asymptotic normality of* $\hat{\theta}_N^\tau$ *when the observed signal is given by (B.9) is well established* [161], [162]. Recall that the sample myriad is an M-estimator.

Our problem of interest can be seen as an M-estimation problem with a misspecified model, where the signal samples are not i.i.d. variables. While it could be expected that an M-estimator would behave erratically due to such model mismatch, it was shown in [161], for example, that this is not the case for a maximum-likelihood estimator, which is a particular case of the M-estimator. In fact, the maximum-likelihood estimates in this case have been shown to be asymptotically Gaussian [161]. In addition, asymptotic properties of M-estimators under non-i.i.d. conditions is a very popular research subject, and the asymptotic normality of these estimators has been proved for different dependencies and non-identical noise distribution conditions. See, among others, [163] and [164] and references therein. Based on these observations, together with our own numerical analysis, the matched myriad filter output for the SRRC pulse shape can be approximated to be asymptotically Gaussian.

Appendix C

Model Matching using Measured Impulsive Noise Statistics

As discussed in Chapter 5, various measurement campaigns have shown that the noise experienced in practical radio channels is both non-Gaussian and correlated in time. One such popular measurement campaign was conducted by Blackard *et al.* to study the impulsive noise in indoor wireless environments [15]. In [15], in order to characterize both the instantaneous and the time dependent statistics of the observed impulsive noise, the distributions of the following three parameters of the impulsive noise were obtained from the noise measurements:

1. pulse amplitude, which is the magnitude of the observed impulsive noise,
2. pulse duration, which is the duration of the noise pulse with amplitude above a specified threshold level, and
3. pulse spacing, defined as the time elapsed between two consecutive noise bursts which exceed a specified threshold.

In these measurements, the mean of the measured noise amplitude was chosen as the threshold. The complimentary cumulative distribution functions (CCDF)[12] of the three measured noise parameters are given in [15].

In order to make the statistics of the noise realizations obtained with the statistical model given by (5.2) and (5.3) to closely approximate those reported in [15], we empirically matched the CCDFs of three impulsive noise parameters described above which are obtained from the proposed noise model to the measured values of these statistics. Table C.1 and Fig. C.1 show the CCDFs of these impulsive noise parameters obtained from the measurements and from the proposed noise model for a specific combination of model parameters $(\{a[i]\}_{i=1}^{P}, \{\lambda_i, \sigma_i\}_{i=1}^{N})$. It can be seen that the statistics of the simulated noise are in close agreement with the statistics reported in [15], for values of CCDF ≥ 0.1.

Table C.1: Statistics of the measured impulsive noise and the noise generated using the proposed model. Measurements made at 2.44GHz. The following parameters were used to generate the noise from the proposed model: 3-term Gaussian mixture model with $\lambda_1 = 0.98, \lambda_2 = 0.018, \lambda_3 = 0.002, \sigma_1^2 = 10, \sigma_2^2 = 50, \sigma_3^2 = 1075$; and AR(2) process with $a[1] = -0.25, a[2] = -0.53$.

CCDF	Pulse amplitude (dB above thermal noise floor)	Pulse duration (1 unit = 40ns)	Pulse spacing (1 unit = 40ns)
100	(1, 1)	(2, 2)	(2, 2)
50	(18, 17.5)	(3, 3)	(5, 4)
10	(24, 22.7)	(7, 7)	(10, 13)
1	(29, 26.4)	(15, 15)	(21, 25)
0.1	(31, 32.3)	(36, 22)	(47, 37)
0.01	(37, 37.1)	(74, 38)	(105, 49)
0.001	(45, 40)	(168, 44)	(320, 62)

[12]The complimentary cumulative distribution function is defined as (1 - CDF)x100, where CDF is the cumulative distribution function

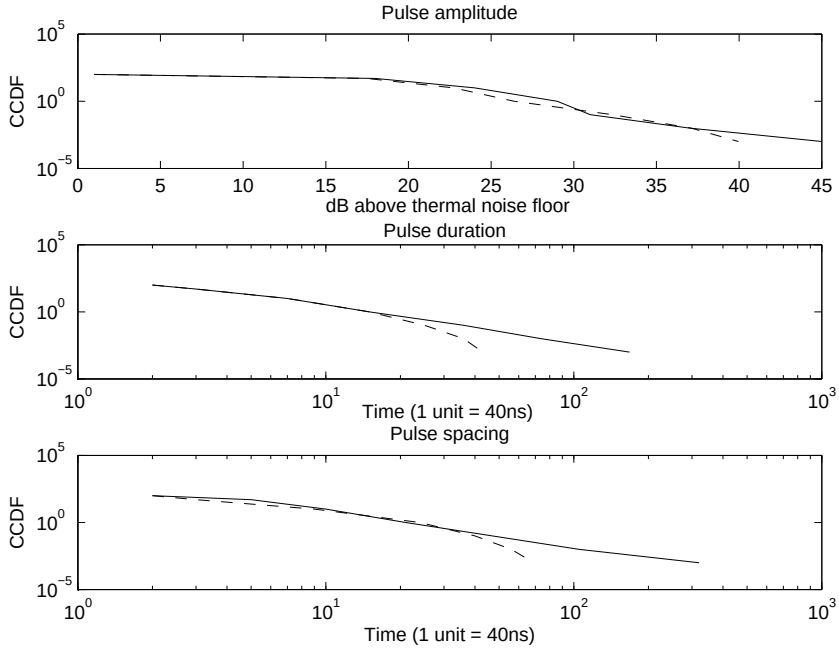

Figure C.1: Statistics of the measured impulsive noise and the noise generated using the proposed model. Measurements made at 2.44GHz. The following parameters were used to generate the noise from the proposed model: 3-term Gaussian mixture model with $\lambda_1 = 0.98, \lambda_2 = 0.018, \lambda_3 = 0.002, \sigma_1^2 = 10, \sigma_2^2 = 50, \sigma_3^2 = 1075$; and AR(2) process with $a[1] = -0.25, a[2] = -0.53$.

Bibliography

[1] O. A. Dobre and F. Hameed, "Likelihood-based algorithms for linear digital modulation classification in fading channels," in *Proc. IEEE Canadian Conf. Elect. Comput. Eng.*, Ottawa, Canada, 2006, pp. 1347–1350.

[2] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory.* Prentice-Hall, 1993.

[3] "Unlicensed operation in TV broadcast bands and additional spectrum for unlicensed devices below 900 MHz and in the 3 GHz band," Federal Communications Commission, Washington, DC., Docket 10-174, Sept. 2010.

[4] J. L. Xu, W. Su, and M. Zhou, "Software-defined radio equipped with rapid modulation recognition," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1659–1667, May 2010.

[5] R. A. Poisel, *Introduction to Communication Electronic Warfare Systems.* Artech House, 2008.

[6] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, Mar. 2009.

[7] L. M. Garth and H. V. Poor, "Detection of non-Gaussian signal: A paradigm for modern statistical signal processing," *Proc. IEEE*, vol. 82, no. 7, pp. 1061–1095, July 1994.

[8] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "A survey of automatic modulation classification techniques: Classical approaches and new developments," *IET Commun.*, vol. 1, no. 2, pp. 137–156, Apr. 2007.

[9] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*. Wiley, 1995.

[10] E. N. Skomal, *Man-made Radio Noise*. Van Nostrand Reinhold, 1978.

[11] K. Gulati, B. L. Evans, J. G. Andrews, and K. R. Tinsley, "Statistics of co-channel interference in a field of Poisson and Poisson-Poisson clustered interferers," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6207–6222, Dec. 2010.

[12] D. Middleton, "Statistical-physical model for electromagnetic interference," *IEEE Trans. Electromagn. Compat.*, vol. EMC-19, no. 3, pp. 106–127, Aug. 1977.

[13] ——, "Non-Gaussian noise models in signal processing for telecommunications: New methods and results for Class A and Class B noise models," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.

[14] X. Yang and A. P. Petropulu, "Co-channel interference modeling and analysis in a Poisson field of interferers in wireless communications," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 64–76, Jan. 2003.

[15] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 11, no. 7, pp. 991–1001, Sept. 1993.

[16] H. Kanemoto, S. Miyamoto, and N. Morinaga, "A study on modeling of microwave oven interference and optimum reception," in *Proc. IEEE Int. Symp. Electromagnetic Compatibility*, Denver, CO, 1998, pp. 57–62.

[17] M. Nassar, K. Gulati, A. K. Sujeeth, N. Aghasadeghi, B. L. Evans, and K. R. Tinsley, "Mitigating near-field interference in laptop embedded wireless transceivers," *J. Signal Process. Syst.*, vol. 63, no. 1, pp. 1–12, Apr. 2011.

[18] T. Erpek, M. A. McHenry, and A. Stirling, "DSA operational parameters with wireless microphones," in *Proc. IEEE Symp. New Frontiers Dynamic Spectrum Access Netw.*, Singapore, 2010, pp. 1–11.

[19] T. K. Blankenship and T. S. Rappaport, "Characteristics of impulsive noise in the 450 MHz band in hospitals and clinics," *IEEE Trans. Antennas Propag.*, vol. 46, no. 2, pp. 194–203, Feb. 1998.

[20] M. G. Sanchez, L. de Haro, M. C. Ramon, A. Mansilla, C. M. Ortega, and D. Oliver, "Impulsive noise measurements and characterization in a UHF digital TV channel," *IEEE Trans. Electromagn. Compat.*, vol. 41, no. 2, pp. 124–136, May 1999.

[21] H. M. Hall, "A new model for impulsive phenomena: Application to atmospheric-noise communication channels," Stanford University, CA, Tech. Rep. SEL-66-052, Sept. 1966.

[22] J. Shi, A. Bettner, G. Chinn, K. Slattery, and X. Dong, "A study of platform EMI from LCD panels: Impact on wireless, root causes and mitigation methods," in *Proc. IEEE Int. Symp. Electromagn. Compat.*, vol. 3, Portland, OR, 2006, pp. 626–631.

[23] X. Wang and H. V. Poor, "Robust multiuser detection in non-Gaussian channels," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 289–305, Feb. 1999.

[24] G. A. Tsihrintzis and C. L. Nikias, "Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an α-stable process," *IEEE Trans. Commun.*, vol. 43, no. 2, pp. 904–914, Mar. 1995.

[25] K. S. Vastola, "Threshold detection in narrow-band non-Gaussian noise," *IEEE Trans. Commun.*, vol. 32, no. 2, pp. 134–139, Feb. 1984.

[26] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, no. 3, pp. 433–481, Mar. 1985.

[27] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*. Springer Verlag, 1988.

[28] F. Moghimi, A. Nasri, and R. Schober, "Adaptive l_p-norm spectrum sensing for cognitive radio networks," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1934–1945, July 2011.

[29] J. Cai, X. Shen, and J. W. Mark, "Robust channel estimation for OFDM wireless communication systems - an H_∞ approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 2060–2071, Nov. 2004.

[30] R. S. Blum, R. J. Kozick, and B. M. Sadler, "An adaptive spatial diversity receiver for non-Gaussian interference and noise," *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2100–2111, Aug. 1999.

[31] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. J. Shellhammer, and W. Caldwell, "IEEE 802.22: The first cognitive radio wireless regional area network standard," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 130–138, Jan. 2009.

[32] E. E. Kuruoglu, W. J. Fitzgerald, and P. J. W. Rayner, "Near optimal detection of signals in impulsive noise modeled with a symmetric α-stable distribution," *IEEE Commun. Lett.*, vol. 2, no. 10, pp. 282–284, Oct. 1998.

[33] A. E. El-Mahdy, "Adaptive signal detection over fast frequency-selective fading channels under Class-A impulsive noise," *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1110–1113, July 2005.

[34] D. Middleton and A. D. Spaulding, "A tutorial review of elements of weak signal detection in non-Gaussian EMI environments," Institute for Telecommunication Sceinces, Boulder, CO, Tech. Rep. 86-196, 1986.

[35] D. Middleton, "Canonical and quasi-canonical probability models of Class A interference," *IEEE Trans. Electromagn. Compat.*, vol. EMC-25, no. 2, pp. 76–103, May 1983.

[36] M. S. Britton and M. L. Scholz, "A practical utilization of Middleton EMI models: Automated modelling, parameter estimation and optimisation," Electron. Surveillance Res. Lab., Edinburgh, Australia, Tech. Rep. 86-196, 1995.

[37] R. Prasad, A. Kegel, and A. de Vos, "Performance of microcellular mobile radio in a co-channel interference, natural, and man-made noise environment," *IEEE Trans. Veh. Technol.*, vol. 42, no. 1, pp. 33–40, Feb. 1993.

[38] D. Zheng-Cong, T. Bin, and L. Ke, "Adaptive MIMO time-varying channel equalization using particle filtering," in *Proc. Int. Conf. Wireless Commun., Networking Mobile Computing*, Wuhun, China, 2005, pp. 123–126.

[39] A. D. Spaulding and D. Middleton, "Optimum reception in an impulsive interference environment - Part I: Coherent detection," *IEEE Trans. Commun.*, vol. 25, no. 9, pp. 910–923, Sept. 1977.

[40] A. E. El-Mahdy, "Classification of MFSK signals over time-varying flat correlated fading channels under Class-A impulsive noise environment," *IEE Proc.-Commun.*, vol. 151, no. 6, pp. 619–626, Dec. 2004.

[41] M. Hamza, H. T. Huynh, and P. Fortier, "Optimal detection of QAM in a man-made noise environment," in *Proc. IEEE Veh. Technol. Conf.*, Houston, TX, 1999, pp. 1301–1306.

[42] D. Middleton, "Procedures for determining the parameters of the first-order canonical models of Class A and Class B electromagnetic interference," *IEEE Trans. Electromagn. Compat.*, vol. EMC-21, no. 3, pp. 190–208, Aug. 1979.

[43] S. M. Zabin and H. V. Poor, "Efficient estimation of Class A noise parameters via the EM algorithm," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 60–72, Jan. 1991.

[44] D. Middleton, "Canonical non-Gaussian noise nodels: Their implications for measurement and for prediction of receiver performance," *IEEE Trans. Electromagn. Compat.*, vol. EMC-21, no. 3, pp. 209–220, Aug. 1979.

[45] A. D. Spaulding and D. Middleton, "Optimum reception in an impulsive interference environment," Office of Telecommunications, Boulder, CO, Tech. Rep. 75-67, 1975.

[46] R. Haring and A. J. H. Vinck, "Performance bounds for optimum and suboptimum reception under Class-A impulsive noise," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1130–1136, July 2002.

[47] P. A. Delaney, "Signal detection in multivariate Class-A interference," *IEEE Trans. Commun.*, vol. 43, no. 2, pp. 365–373, Feb. 1995.

[48] L. Breiman, *Probability*. Addison-Wesley, 1968.

[49] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: Stable processes and their applications," *Proc. IEEE*, vol. 81, no. 7, pp. 986–1010, July 1993.

[50] J. Ilow and D. Hatzinakos, "Impulsive noise modeling with stable distributions in fading environments," in *Proc. IEEE Signal Process. Workshop Stats. Signal Array Process.*, Corfu, Greece, 1996, pp. 140–143.

[51] A. Swami and B. M. Sadler, "On some detection and estimation problems in heavy-tailed noise," *Signal Processing*, vol. 82, no. 12, pp. 1829–1846, Dec. 2002.

[52] J. Ilow and D. Hatzinakos, "Analytic alpha-stable noise modeling in a Poisson field of interferers or scatterers," *IEEE Trans. Signal Process.*, vol. 46, no. 6, pp. 1601–1611, June 1998.

[53] P. Tsakalides and C. L. Nikias, "Wideband array signal processing with alpha-stable distributions," in *Proc. IEEE Military Commun. Conf.*, San Diego, CA, 1995, pp. 135–139.

[54] A. Swami and B. Sadler, "TDE, DOA and related parameter estimation problems in impulsive noise," in *Proc. IEEE Signal Process. Workshop Higher-Order Statistics*, Banff, Canada, 1997, pp. 273–277.

[55] G. A. Tsihrintzis and C. L. Nikias, "Data-adaptive algorithms for signal detection in sub-Gaussian impulsive interference," *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1873–1878, July 1998.

[56] ——, "Incoherent receivers in alpha-stable impulsive noise," *IEEE Trans. Signal Process.*, vol. 43, no. 9, pp. 2225–2229, Sept. 1995.

[57] B. M. Sadler, "Detection in correlated impulsive noise using fourth-order cumulants," *IEEE Trans. Signal Process.*, vol. 44, no. 11, pp. 2793–2800, Nov. 1996.

[58] H. V. Poor and J. B. Thomas, *Signal detection in dependent non-Gaussian noise.* JAI Press, 1993, vol. 2.

[59] D. Sengupta and S. M. Kay, "Efficient estimation of parameters for non-Gaussian autoregressive processes," *IEEE Trans. Acoustics Speech Signal Process.*, vol. 37, no. 6, pp. 785–794, June 1989.

[60] Y. C. Eldar and A. Yeredor, "Finite-memory denoising in impulsive noise using Gaussian mixture models," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 48, no. 11, pp. 1069–1077, Nov. 2001.

[61] S. M. Kay and D. Sengupta, "Detection in incompletely characterized colored non-Gaussian noise via parametric modeling," *IEEE Trans. Signal Process.*, vol. 41, no. 10, pp. 3066–3070, Oct. 1993.

[62] B. M. Sadler and G. B. Giannakis, "Estimation and detection in non-Gaussian noise using higher order statistics," *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2729–2741, Oct. 1994.

[63] I. J. Kim, S. R. Park, I. Song, J. Lee, H. Kwon, and S. Yoon, "Detection schemes for weak signals in first-order moving average of impulsive noise," *IEEE Trans. Veh. Technol.*, vol. 56, no. 1, pp. 126–133, Jan. 2007.

[64] C. L. Brown and A. M. Zoubir, "Locally optimum and rank-based known signal detection in correlated alpha-stable interference," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, Istanbul, Turkey, 2000, pp. 53–56.

[65] B. M. Sadler, G. B. Giannakis, and S. Shamsunder, "Noise subspace techniques in non-Gaussian noise using cumulants," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 3, pp. 1009–1018, July 1995.

[66] K. F. McDonald and R. S. Blum, "A statistical and physical mechanisms-based interference and noise model for array observations," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 2044–2056, July 2000.

[67] R. S. Blum, Y. Zhang, B. M. Sadler, and R. J. Kozick, "On the approximation of correlated non-Gaussian noise pdfs using Gaussian mixture models," in *Proc. Conf. Appl. Heavy Tailed Distributions Econom., Eng., Statist.*, Washington, DC., 1999, pp. 1–11.

[68] Y. Zhang and R. S. Blum, "An adaptive receiver with an antenna array for channels with correlated non-Gaussian interference and noise using the SAGE algorithm," *IEEE Trans. Signal Process*, vol. 48, no. 7, pp. 2172–2175, July 2000.

[69] S. Buzzi, E. Conte, A. de Maio, and M. Lops, "Optimum diversity detection over fading dispersive channels with non-Gaussian noise," *IEEE Trans. Signal Process*, vol. 49, no. 4, pp. 767 –776, Apr. 2001.

[70] K. D. Ward, C. J. Baker, and S. Watts, "Maritime surveillance radar. Part 1: Radar scattering from the ocean surface," *Proc. Inst. Elect. Eng. Pt. F, Commun., Radar Signal Process*, vol. 137, pp. 51–62, Apr. 1990.

[71] E. S. Sousa, "Interference modeling in a direct-sequence spread-spectrum packet radio network," *IEEE Trans. Commun.*, vol. 38, no. 7, pp. 1475–1482, Sept. 1990.

[72] H. V. Poor and J. B. Thomas, "Memoryless discrete-time detection of a constant signal in m-dependent noise," *IEEE Trans. Inf. Theory*, vol. 25, no. 1, pp. 54–61, Jan. 1979.

[73] D. R. Halverson and G. Wise, "Discrete-time detection in ϕ-mixing noise," *IEEE Trans. Inf. Theory*, vol. 26, no. 2, pp. 189–198, Mar. 1980.

[74] O. C. Au and J. B. Thomas, "On transformation noise: Properties and modeling," *J. Franklin Institute*, vol. 330, no. 4, pp. 707–720, July 1993.

[75] A. Martinez, P. Swaszek, and J. Thomas, "Locally optimal detection in multivariate non-Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 30, no. 6, pp. 815–822, Nov. 1984.

[76] X. Yang, H. V. Poor, and A. P. Petropulu, "Memoryless discrete time signal detection in long-range dependent noise," *IEEE Trans. Signal Process.*, vol. 52, no. 6, pp. 1607–1619, June 2004.

[77] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 416–429, Mar. 2000.

[78] O. A. Dobre, Y. Bar-Ness, and W. Su, "Robust QAM modulation classification algorithm using cyclic cumulants," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Atlanta, GA, 2004, pp. 745–748.

[79] J. A. Sills, "Maximum-likelihood modulation classification for PSK/QAM," in *Proc. IEEE Military Communications Conference*, vol. 1, 1999, pp. 217–220.

[80] P. Panagiotou, A. Anastasopoulos, and A. Polydoros, "Likelihood ratio tests for modulation classification," in *Proc. IEEE Military Communications Conference*, vol. 2, 2000, pp. 670–674.

[81] W. C. Headley and C. R. C. M. da Silva, "Asynchronous classification of digital amplitude-phase modulated signals in flat-fading channels," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 7–12, Jan. 2011.

[82] R. J. Kozick and B. M. Sadler, "Maximum-likelihood array processing in non-Gaussian noise with Gaussian mixtures," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3520–3535, Dec. 2000.

[83] D. W. J. Stein, "Detection of random signals in Gaussian mixture noise," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1788–1801, Nov. 1995.

[84] A. Nasri and R. Schober, "Performance of BICM-SC and BICM-OFDM systems with diversity reception in non-Gaussian noise and interference," *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3316–3327, Nov. 2009.

[85] W. Wei and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 189–193, Nov. 2000.

[86] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc. (Series B)*, vol. 39, no. 1, pp. 1–38, Jan. 1977.

[87] C. F. J. Wu, "On the convergence properties of the EM algorithm," *Ann. Stat.*, vol. 11, no. 1, pp. 95–103, MAr. 1983.

[88] G. J. McLachlan and T. Krishnan, *The EM algorithm and extensions.* Wiley-Interscience, 2008.

[89] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Trans. Commun.*, vol. 42, no. 7, pp. 2406–2413, July 1994.

[90] X. L. Meng and D. B. Rubin, "Maximum likelihood estimation via the ECM algorithm: A general framework," *Biometrika*, vol. 80, no. 2, pp. 267–278, June 1993.

[91] J. Sexton and A. R. Swensen, "ECM algorithms that converge at the rate of EM," *Biometrika*, vol. 87, no. 3, pp. 651–662, Sept. 2000.

[92] A. Swami, "Non-Gaussian mixture models for detection and estimation in heavy-tailed noise," in *Proc. Intl. Conf. Acoustics, Speech Signal Process.*, vol. 6, Istanbul, Turkey, 2000, pp. 3802–3805.

[93] J. Ilow and D. Hatzinakos, "Applications of the empirical characteristic function to estimation and detection problems," *Signal Process.*, vol. 65, no. 2, pp. 199–219, Mar. 1998.

[94] J. G. Gonzalez, "Robust techniques for wireless communications in non-Gaussian environments," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. of Delaware, Newark, DE, 1997.

[95] G. R. Arce, *Nonlinear Signal Processing: A Statistical Approach.* Wiley, 2005.

[96] J. G. Gonzalez, D. Griffith, and G. R. Arce, "Matched myriad filters for robust communications," in *Proc. Conf. Inform. Sci. Syst.*, Princeton, NJ, 1996, pp. 1–6.

[97] X. Ma and C. L. Nikias, "Joint estimation of time delay and frequency delay in impulsive noise using fractional lower order statistics," *IEEE Trans. Signal Process.*, vol. 44, no. 11, pp. 2669–2687, Nov. 1996.

[98] J. G. Gonzalez and G. R. Arce, "Optimality of the myriad filter in practical impulsive-noise environments," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 438–441, Feb. 2001.

[99] H.-S. Lim, T.-C. Chuah, and H.-T. Chuah, "On the optimal alpha-K curve of the sample myriad," *IEEE Signal Process. Lett.*, vol. 14, no. 8, pp. 545–548, Aug. 2007.

[100] A. A. Roenko, V. V. Lukin, and I. Djurovic, "Two approaches to adaptation of sample myriad to characteristics of SαS distribution data," *Signal Process.*, vol. 90, no. 7, pp. 2213–2123, July 2010.

[101] ——, "An overview of adaptive robust DFT," *EURASIP J. Adv. Signal Process.*, vol. 2010, pp. 1–17, June 2010.

[102] S. Kalluri and G. R. Arce, "A general class of nonlinear normalized adaptive filtering algorithms," *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2262–2272, Aug. 1999.

[103] ——, "Robust frequency-selective filtering using weighted myriad filters admitting real-valued weights," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2721–2733, Nov. 2001.

[104] H.-S. Lim, T.-C. Chuah, and H.-T. Chuah, "Robust chip-matched myriad filter-based multiuser receiver for impulsive channels," *Signal Process.*, vol. 88, no. 5, pp. 1216–1232, May 2008.

[105] A. Swami and B. Sadler, "Parameter estimation for linear alpha-stable processes," *IEEE Signal Process. Lett.*, vol. 5, no. 2, pp. 48–50, Feb. 1998.

[106] J. H. McCulloch, "Simple consistent estimators of stable distribution parameter," *Commun. Stat. - Simulation Computation*, vol. 15, no. 4, pp. 1109–1136, 1986.

[107] G. A. Tsihrintzis and C. L. Nikias, "Fast estimation of the parameters of alpha-stable impulsive interference," *IEEE Trans. Signal Process.*, vol. 44, no. 6, pp. 1492–1503, June 1996.

[108] I. A. Koutrouvelis, "Regression-type estimation of parameters of stable laws," *J. Amer. Stat. Assoc.*, vol. 75, no. 372, pp. 918–928, Dec. 1980.

[109] J. Yu, "Empirical characteristic function estimation and its applications," *Econometric Reviews*, vol. 23, no. 2, pp. 93–123, 2004.

[110] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice-Hall, 1993.

[111] R. Brcich and A. Zoubir, "Estimation and detection in a mixture of symmetric alpha stable and Gaussian interference," in *Proc. IEEE Sig. Process. Workshop Higher Order Stat.*, Caesarea, Israel, 1999, pp. 219–223.

[112] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers.* Plenum Press, 1997.

[113] F. Hameed, O. A. Dobre, and D. C. Popescu, "On the likelihood-based approach to modulation classification," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5884–5892, Dec. 2009.

[114] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance.* Chapman and Hill, 1994.

[115] S. Kalluri and G. R. Arce, "Fast algorithms for weighted myriad computation by fixed-point search," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 159–171, Jan. 2000.

[116] E. E. Kuruoglu, "Density parameter estimation of skewed alpha-stable distributions," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2192–2201, Oct. 2001.

[117] G. A. F. Seber and C. J. Wild, *Nonlinear Regression.* Wiley, 1989.

[118] K. J. Sangston and K. R. Gerlach, "Coherent detection of radar targets in a non-gaussian background," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 30, no. 2, pp. 330–340, Apr. 1994.

[119] S. Richardson and P. J. Green, "On Bayesian analysis of mixtures with an unknown number of components," *J. R. Stat. Soc. (Ser. B)*, vol. 59, no. 4, pp. 731–792, Apr. 1997.

[120] M. Stephens, "Bayesian analysis of mixtures with an unknown number of components an alternative to reversible jump methods," *Ann. Stat.*, vol. 28, no. 1, pp. 40–54, Jan. 2011.

[121] C. E. Rasmussen, "The infinite Gaussian mixture model," in *Advances in Neural Information Processing Systems*. MIT Press, 2000.

[122] B. Friedlander, "Instrumental variable methods for ARMA spectral estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 2, pp. 404–415, Apr. 1983.

[123] W.-R. Wu and P.-C. Chen, "Adaptive AR modeling in white Gaussian noise," *IEEE Trans. Signal Process.*, vol. 45, no. 5, pp. 1184–1192, May 1997.

[124] S. Gannot, D. Burshtein, and E. Weinstein, "Iterative and sequential Kalman filter-based speech enhancement algorithms," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 2, pp. 373–385, July 1998.

[125] X. Shen and L. Deng, "A dynamic system approach to speech enhancement using the H_∞ filtering algorithm," *IEEE Trans. Speech Audio Process.*, vol. 7, no. 4, pp. 391–399, July 1997.

[126] S. M. Verbout, J. M. Ooi, J. T. Ludwig, and A. V. Oppenheim, "Parameter estimation for autoregressive Gaussian-mixture processes: The EMAX algorithm," *IEEE Trans. Signal Process.*, vol. 46, no. 10, pp. 2744–2756, Oct. 1998.

[127] E. Weinstein, A. V. Oppenheim, and M. Feder, "Signal enhancement using single and multi-sensor measurements," Res. Lab. Electron., Massachusetts Inst. Technol., Cambridge, MA, Tech. Rep. 560, Nov. 1990.

[128] D. Sengupta and S. M. Kay, "Parameter estimation and GLRT detection in colored non-Gaussian autoregressive processes," *IEEE Trans. Acoustics, Speech Signal Process.*, vol. 38, no. 10, pp. 1661–1676, Oct. 1990.

[129] D. Simon, *Optimal State Estimation*. Wiley, 2006.

[130] M. J. Grimble and A. E. Sayed, "Solution of the H_∞ optimal linear filtering problem for discrete-time systems," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 38, no. 7, pp. 1092–1104, July 1990.

[131] P. Xu, J. K. Wang, F. Qi, and X. Song, "Space-alternating generalised expectation-maximisation-based H-infinity channel estimator for multiple-input multiple-output-orthogonal frequency division multiplexing systems," *IET Commun.*, vol. 5, no. 14, pp. 2068–2074, Sept. 2011.

[132] B. Hassibi, A. T. Erdogan, and T. Kailath, "MIMO linear equalization with an H_∞ criterion," *IEEE Trans Signal Process.*, vol. 54, no. 2, pp. 499–511, Feb. 2006.

[133] H. Kulatunga and V. Kadirkamanathan, "Multiple H-infinity filter based deterministic sequence estimation in non-Gaussian channels," *IEEE Signal Process. Lett.*, vol. 13, no. 4, pp. 185–188, Apr. 2006.

[134] B. Hassibi, A. H. Sayed, and T. Kailath, "Linear estimation in krein spaces - part ii: Applications," *IEEE Trans. Automat. Control*, vol. 41, no. 1, pp. 34–49, Jan. 1996.

[135] A. Jamoos, E. Grivel, N. Christov, and M. Najim, "Estimation of autoregressive fading channels based on two cross-coupled H_∞ filters," *Signal Image Video Process.*, vol. 3, no. 3, pp. 209–216, Sept. 2009.

[136] D. Labarre, E. Grivel, M. Najim, and N. Christov, "Dual H_∞ algorithms for signal processing - application to speech enhancement," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5195–5208, Nov. 2007.

[137] E. Wan and A. Nelson, "Dual extended Kalman filter methods," in *Kalman Filtering and Neural Networks.* Wiley, 2001.

[138] Z. Chen, "Bayesian filtering: From Kalman filters to particle filters, and beyond," Adaptive Syst. Lab., McMaster Univ., Hamilton, Canada, Tech. Rep., 2003.

[139] F. E. Visser, G. J. M. Janssen, and P. Pawelczak, "Multinode spectrum sensing based on energy detection for dynamic spectrum access," in *Proc. IEEE Veh. Technol. Conf.*, Singapore, 2008, pp. 1394–1398.

[140] M. I. B. Shahid and J. Kamruzzaman, "Weighted soft decision for cooperative sensing in cognitive radio networks," in *Proc. IEEE Int. Conf. Netw.*, New Delhi, India, 2008, pp. 1–6.

[141] S. S. Jeong, W. S. Jeon, and D. G. Jeong, "Collaborative spectrum sensing for multiuser cognitive radio systems," *IEEE Trans. Veh. Technol.*, vol. 58, no. 5, pp. 2564–2569, June 2009.

[142] S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative sensing among cognitive radios," in *Proc. IEEE Int. Conf. Commun.*, Istanbul, Turkey, 2006, pp. 1658–1663.

[143] A. Ghasemi and E. S. Sousa, "Oppurtunistic spectrum access in fading channels through collaborative sensing," *J. Commun.*, vol. 2, no. 2, pp. 71–82, Mar. 2007.

[144] A. Taherpour, Y. Norouzi, M. Nasiri-Kenari, A. Jamshidi, and Z. Zeinalpour-Yazdi, "Asymptotically optimum detection of primary user in cognitive radio networks," *IET Commun.*, vol. 1, no. 6, pp. 1138–1145, Dec. 2007.

[145] Y. Chen, "Analytical performance of collaborative spectrum sensing using censored energy detection," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3856–3865, Dec. 2010.

[146] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 28–40, Feb. 2008.

[147] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[148] J. Ma, G. Zhao, and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4502–4507, Nov. 2008.

[149] E. C. Y. Peh, Y.-C. Liang, Y. L. Guan, and Y. Zeng, "Optimization of cooperative sensing in cognitive radio networks: A sensing-throughput tradeoff view," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 5294–5299, Nov. 2009.

[150] T. Do and B. L. Mark, "Joint spatial-temporal spectrum sensing for cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 7, pp. 3480–3490, Sept. 2010.

[151] E. Biglieri and M. Lops, "Multiuser detection in a dynamic environment – Part I: User identification and data detection," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3158–3170, Sept. 2007.

[152] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 2002.

[153] A. S. Fotheringham and D. C. Knudsen, *Goodness-of-fit Statistics*. Geo Books, 1987.

[154] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables*. Marcel Dekker, 1992.

[155] C. W. Helstrom, "Computing the distribution of sums of random sine waves and of rayleigh-distributed random variables by saddle-point integration," *IEEE Trans. Commun.*, vol. 45, no. 11, pp. 1487–1494, Nov. 1997.

[156] E. Biglieri, G. Caire, G. Taricco, and J. Ventura-Traveset, "Computing error probabilities over fading channels: A unified approach," *Europ. Trans. Telecommun.*, vol. 9, no. 1, pp. 15–25, Jan. 1998.

[157] D. R. Pauluzzi and N. C. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1681–1691, Oct. 2000.

[158] S. S. Szyszkowicz, H. Yanikomeroglu, and J. S. Thompson, "On the feasibility of wireless shadowing correlation models," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4222–4236, Nov. 2010.

[159] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Commun. Mobile Comput.*, vol. 2, no. 5, pp. 483–502, Aug. 2002.

[160] Y. Li and X. Huang, "The simulation of independent Rayleigh faders," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1503–1514, Sept. 2002.

[161] A. W. van der Vaart, *Asymptotic Statistics*. Cambridge Univ. Press, 1998.

[162] H. V. Poor, *An Introduction to Signal Detection and Estimation*. Springer, 1994.

[163] W. B. Wu, "M-estimation of linear models with dependent errors," *Ann. Stat.*, vol. 35, no. 2, pp. 495–521, Apr. 2007.

[164] R. Beran, "Robust estimation in models for independent non-identically distributed data," *Ann. Stat.*, vol. 10, no. 2, pp. 415–428, Jun. 1982.