

1-1-2002

## Modulation classification of digital communication signals

Visalakshi S. Ramakonar  
*Edith Cowan University*

Follow this and additional works at: <https://ro.ecu.edu.au/theses>



Part of the [Electrical and Computer Engineering Commons](#)

---

### Recommended Citation

Ramakonar, V. S. (2002). *Modulation classification of digital communication signals*.  
<https://ro.ecu.edu.au/theses/752>

This Thesis is posted at Research Online.  
<https://ro.ecu.edu.au/theses/752>

# Edith Cowan University

## Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.
- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author's moral rights contained in Part IX of the Copyright Act 1968 (Cth).
- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

EDITH COWAN UNIVERSITY  
LIBRARY

# **Modulation Classification of Digital Communication Signals**

Visalakshi Saratha Devi Ramakonar

A Thesis Submitted in Partial Fulfilment of the Requirements for the  
Degree of Doctor of Philosophy  
at the  
School of Engineering & Mathematics  
Edith Cowan University



Principal Supervisor: Associate Professor Abdesselam Bouzerdoum  
Associate Supervisor: Dr. Daryoush Habibi

October 2002

To My Husband

## USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

# ABSTRACT

Modulation classification of digital communications signals plays an important role in both military and civilian sectors. It has the potential of replacing several receivers with one universal receiver. An automatic modulation classifier can be defined as a system that automatically identifies the modulation type of the received signal given that the signal exists and its parameters lie in a known range. This thesis addresses the need for a universal modulation classifier capable of classifying a comprehensive list of digital modulation schemes. Two classification approaches are presented: a decision-theoretic (DT) approach and a neural network (NN) approach.

First classifiers are introduced that can classify ASK, PSK, and FSK signals. A decision tree is designed for the DT approach and a NN structure is formulated and trained to classify these signals. Both classifiers use the same key features derived from the intercepted signal. These features are based on the instantaneous amplitude, instantaneous phase, and instantaneous frequency of the intercepted signal, and the cumulants of its complex envelope. Threshold values for the DT approach are found from the minimum total error probabilities of the extracted key features at SNR of 20 to -5dB. The NN parameters are found by training the networks on the same data. The DT and NN classifiers are expanded to include CPM signals. Signals within the CPM class are also added to the classifiers and a separate decision tree and new NN structure are found for these signals. New key features to classify these signals are also introduced.

The classifiers are then expanded further to include multiple access signals, followed by QAM, PSK8 and FSK8 signals. New features are found to classify these signals. The final decision tree is able to accommodate a total of fifteen different modulation types. The NN structure is designed in a hierarchical fashion to optimise the classification performance of these fifteen digital modulation schemes.

Both DT and NN classifiers are able to classify signals with more than 90% accuracy in the presence of additive white Gaussian within SNR ranging from 20 to 5dB. However,

the performance of the NN classifier appears to be more robust as it degrades gradually at the SNRs of 0 and  $-5\text{dB}$ . At  $-5\text{dB}$ , the NN has an overall accuracy of 73.58%, whereas the DT classifier achieves only 47.3% accuracy. The overall accuracy of the NN classifier, over the combined SNR range of 20 to  $-5\text{dB}$ , is 90.7% compared to 84.56% for the DT classifier.

Finally, the performances of these classifiers are tested in the presence of Rayleigh fading. The DT and NN classifier structures are modified to accommodate fading and again, new key features are introduced to accomplish this. With the modifications, the overall accuracy of the NN classifier, over the combined SNR range of 20 to  $-5\text{dB}$  and 120Hz Doppler shift, is 87.34% compared to 80.52% for the DT classifier.

# Declaration

---

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution of higher education; and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signature .

Date 13/02/03



# ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Abdesselam Bouzerdoum. I am deeply grateful for his guidance, encouragement and support. Without his knowledge and patience throughout, it is unlikely that this thesis would have come to fruition. I am very fortunate to have had his supervision.

I would also like to thank Dr Daryoush Habibi for his support and encouragement throughout the duration of my research.

My gratitude also goes to the School of Engineering and Mathematics, Edith Cowan University, for all the financial and equipment support provided during the course of this study.

I would like to thank the DSTO (Defence Science and Technology Organization) for giving me the opportunity to work onsite with their research team in Adelaide and for providing me with financial support. In particular, I would like to thank John Kitchen from the DSTO who provided me with all the technical facilities and support during my stay in Adelaide.

I would like to express my gratitude to my colleagues at Edith Cowan University, namely, Geoffrey Alagoda, Joe Austin-Crowe, Andrew Ehrhardt, Edward Gluszak, David Lucas, and Alexander Rassau for their encouragement, support and advice.

My gratitude goes to my family and friends who have supported and encouraged me throughout the trying times of this research. In particular, I would like to thank my husband, Kuhan who unwaveringly stood by me and provided me with the motivation and encouragement to complete this research.

Finally I would like to thank my dear Godfather who has sustained me in all ways throughout the years of this research. My deepest gratitude goes to him.

# CONTENTS

LIST OF FIGURES .....	xi
LIST OF TABLES .....	xvii
GLOSSARY OF TERMS .....	xxiii
 CHAPTER 1 .....	 1
Introduction .....	1
1.1 Objectives of the Thesis .....	1
1.2 Major Contributions of the Thesis .....	2
1.3 Organization of the Thesis .....	3
1.4 Publications Arising From PhD Research .....	5
 CHAPTER 2 .....	 7
Classification of Digital Modulation Schemes: A Review .....	7
2.1. Introduction .....	7
2.2. Maximum Likelihood Approach.....	7
2.2.1. General Maximum Likelihood Methods.....	8
2.2.2. MPSK Classifier Based on the Exact Phase Distribution .....	10
2.2.3. Classifiers Based on the Likelihood Functions.....	10
2.2.4. CPM Classification using ML Function .....	13
2.2.5. Classification Using Constellation Shape as A Robust Signature .....	13
2.3. Pattern Recognition Approach .....	14
2.3.1. Envelope-Based Methods .....	14
2.3.2. Higher-Order Statistical Methods .....	18
2.3.3. Other Methods.....	22
2.3.4. Classification using Neural Networks.....	27
2.4. Conclusions.....	28
 CHAPTER 3 .....	 31
Decision Theory .....	31
3.1. Introduction .....	31
3.2. Classification Decision Theory .....	31
3.2.1. The Bayes Decision Rule for Minimum Error.....	32
3.2.2. Bayes Error .....	33
3.2.3. Threshold Determination .....	35
3.2.4. Classifier Accuracy .....	39
3.2.5. Confidence Intervals for Discrete – Valued Hypotheses .....	40
3.2.6. Statistical Significance Versus Statistical Power .....	41
3.3. Conclusions .....	43
 CHAPTER 4 .....	 45

Classification Using Feedforward Artificial Neural Networks (ANNs).....	45
4.1. Introduction.....	45
4.2. Artificial Neural Networks.....	45
4.2.1. The Artificial Neuron Model .....	46
4.2.2. Activation Function Types.....	47
4.2.3. ANN Architectures .....	48
4.2.4. Learning Process.....	50
4.3. Classification Using Neural Networks.....	51
4.3.1. Learning and Generalisation .....	53
4.3.2. Bias and Variance Composition of the Prediction Error.....	53
4.3.3. Example of Classification Using Neural Networks .....	54
4.4. Conclusions.....	58
CHAPTER 5 .....	59
Modulation Classification of ASK, FSK, and PSK Signals .....	59
5.1 Introduction.....	59
5.2 Analytic Signal Representation.....	59
5.2.1 Hilbert Transform .....	60
5.2.2 Complex Envelope.....	61
5.2.3 Representations of Digital Modulation Schemes.....	62
5.3 Key Feature Extraction .....	68
5.3.1 Cumulant Key Features.....	68
5.3.2 Other Key Features .....	69
5.3.3 Explanation for Key Feature Selection .....	70
5.4 Decision – Theoretic Modulation Classification Method.....	71
5.4.1 Threshold Determination .....	72
5.4.2 Dependency of Key Feature Selection on Minimum Probability of Error ...	75
5.4.3 Receiver Operating Characteristic (ROC) Curves.....	80
5.5 Modulation Classification Using Artificial Neural Networks .....	82
5.5.1 Neural Network Structure .....	83
5.5.2 Training the Network.....	84
5.6 Performance Analysis .....	85
5.6.1 DT Classifier Results .....	85
5.6.2 NN Classifier Results and Comparison With DT Classifier.....	85
5.6.3 Comparison with Azzouz and Nandi's Classifier.....	88
5.7 Conclusions.....	90
CHAPTER 6 .....	93
Classification of Continuous Phase Modulated Signals .....	93
6.1 Introduction.....	93
6.2 Continuous Phase Modulated (CPM) Signals.....	94
6.3 CPM Signal Classification using DT Approach .....	96
6.3.1 Discrimination of CPM Signals From Other Signals: DT Approach .....	96
6.3.2 Threshold Determination .....	97
6.3.3 Classification of Signals Within the CPM Signal Class (DT Approach) .....	99
6.3.4 CPM Receivers .....	101

6.3.5	Key Feature Derivation .....	102
6.3.6	DT CPM Classification Method .....	105
6.3.7	Threshold Determination .....	105
6.4	Neural Network Classifier .....	110
6.4.1	Neural Network Structure .....	111
6.4.2	Training the Network .....	112
6.4.3	NN Classification Within the CPM Signal Class.....	113
6.5	Results .....	114
6.5.1	DT Classifier Performance Results.....	115
6.5.2	NN Classifier Performance With Comparison to DT Classifier Results ...	116
6.6	Conclusions .....	123
CHAPTER 7 .....		125
Classification of Multiple Access Signals .....		125
7.1	Introduction .....	125
7.2	Multiple Access Communication Systems .....	125
7.2.1	Direct Sequence Spread Spectrum (DS-SS) .....	127
7.2.2	Frequency Hopped Spread Spectrum (FH SS).....	132
7.2.3	Time Division Multiple Access (TDMA).....	133
7.3	Classification Procedure (DT Approach).....	134
7.3.1	Key Feature Derivation for Signal Classification .....	136
7.4	Threshold Determination .....	138
7.4.1	Dependency of Key Feature Selection on Minimum Probability of Error	143
7.5	Neural Network Classifier .....	143
7.5.1	Neural Network Structure .....	144
7.5.2	Training the Network .....	146
7.6	Performance Analysis .....	146
7.6.1	DT Classifier Results .....	146
7.6.2	Neural Network Classifier Results.....	147
7.7	Conclusions .....	152
CHAPTER 8 .....		153
Classification of PSK8, FSK8 and QAM Signals.....		153
8.1	Introduction .....	153
8.2	Signal Representation .....	153
8.3	DT Classification Procedure .....	157
8.3.1	Derivation of Key Features .....	157
8.3.2	Threshold Determination .....	160
8.3.3	Dependency of key feature selection on minimum probability of error ....	168
8.4	NN Classifier.....	169
8.4.1	Neural Network Structure .....	170
8.4.2	Training the Network .....	171
8.5	Performance Analysis of DT and NN Classifiers in the Presence of White Gaussian Noise.....	173
8.5.1	Performance Results for DT Classifier .....	173
8.5.2	Performance Results of NN Classifier .....	173

8.6	Conclusions .....	178
<b>CHAPTER 9 .....</b>		
	Classification of Digitally Modulated Signals in the Presence of Rayleigh Fading .....	181
9.1	Introduction .....	181
9.2	Classification in the Presence of Rayleigh Fading Channels .....	181
9.2.1	Introduction to Fading Channels .....	181
9.2.2	Characterisation of Fading Multipath Channels .....	182
9.2.3	Rayleigh Fading .....	183
9.2.4	DT Performance in Rayleigh Fading Conditions .....	185
9.3	Decision Tree Modifications for Rayleigh Fading .....	186
9.3.1	PSK, BPSK DS-SS, and QPSK DS-SS Signal Classification .....	186
9.3.2	TDMA Classification .....	187
9.3.3	FSK Signal Classification .....	187
9.3.4	Threshold Determination .....	188
9.3.5	Dependency of Key Feature Selection on Minimum Probability of Error .....	196
9.4	NN Classifier Modifications for Rayleigh Fading .....	197
9.4.1	Modified Neural Network Structure to Accommodate Rayleigh Fading .....	197
9.4.2	Neural Network Classifier for AWGN and Rayleigh Fading Channel .....	197
9.5	Performance Analysis of DT and NN Classifiers in the Presence of Rayleigh Fading .....	197
9.6	Conclusions .....	202
<b>CHAPTER 10 .....</b>		
	Conclusion .....	207
10.1	Introduction .....	207
10.2	Suggestions for Further Work .....	208
<b>Appendix A .....</b>		
A.1	Confusion Matrices for DT Classifier .....	211
A.2	Confusion Matrices for NN Classifier .....	213
<b>Appendix B .....</b>		
B.1	Confusion Matrices for DT Classifier .....	215
B.1.1	Classification of CPM Signals .....	215
B.1.2	Classification of Signals within the CPM Signal Class .....	217
B.2	Confusion Matrices for NN Classifier .....	219
B.2.1	NN Classification of CPM Signals .....	219
B.2.2	NN Classification of Signals Within the CPM Class .....	221
<b>Appendix C .....</b>		
C.1	Confusion Matrices for DT Classifier .....	227
C.2	Confusion Matrices for NN Classifier .....	233
<b>Appendix D .....</b>		
D.1	Confusion Matrices for DT Classifier .....	239

D.2 Confusion Matrices for NN Classifier ..... 245

Appendix E ..... 251

    E.1 Confusion Matrices for DT Classifier (Rayleigh Fading) ..... 251

    E.4 Confusion Matrices for NN Classifier (Rayleigh Fading)..... 257

REFERENCES..... 263



# LIST OF FIGURES

## CHAPTER 2

Figure 2.1. General maximum likelihood classifier..... 8

Figure 2.2. General pattern recognition system. .... 14

## CHAPTER 3

Figure 3.1. Functional blocks of signal classification..... 32

Figure 3.2. Bayes decision rule for minimum error [Fukunaga, 1990]..... 34

Figure 3.3. Example of Bayes decision rule for minimum error. .... 35

Figure 3.4. Example of posterior probabilities for two classes of digitally modulated signals (FSK4 and FSK8)..... 38

## CHAPTER 4

Figure 4.1. Neuron model. .... 47

Figure 4.2. Types of activation functions..... 48

Figure 4.3. A Layered feed-forward neural network. .... 50

Figure 4.4. Taxonomy of the learning process..... 51

Figure 4.5. Probability density function for Class 1. .... 55

Figure 4.6. Probability density function for Class 2. .... 56

Figure 4.7. Scatter plot of classes  $\omega_1$  and  $\omega_2$  Showing decision boundary..... 56

## CHAPTER 5

Figure 5.1. Useful features of ASK2 modulation, carrier frequency  $F_c = 150\text{kHz}$ . .... 65

Figure 5.2. Useful features of ASK4 modulation, carrier frequency  $F_c = 150\text{kHz}$ . .... 65

Figure 5.3. Useful features of PSK2 modulation. .... 66

Figure 5.4. Useful features of PSK4 modulation. .... 66

Figure 5.5. Useful features of FSK2 modulation. .... 67

Figure 5.6. Useful features of FSK4 modulation. .... 67



Figure 5.7. Instantaneous phase values for ASK2 and ASK4 signals. ....	71
Figure 5.8. Decision tree for classification of digital modulated signals.....	72
Figure 5.9. Total error probability for the key feature $\gamma_{maxf}$ .....	73
Figure 5.10. Total error probability for the key feature $ \hat{C}_{21} $ . ....	73
Figure 5.11. Total error probability for the key feature $ \hat{C}_{40} $ . ....	74
Figure 5.12. Total error probability for the key feature $\sigma_{fn}$ .....	74
Figure 5.13. Total error probability for the key feature $\mu_{dp}$ .....	75
Figure 5.14. ROC curves for the key feature $\sigma_{fn}$ to separate FSK2 and FSK4. ....	81
Figure 5.15. ROC curves for the key feature $\mu_{dp}$ to separate ASK2 and ASK4. ....	82
Figure 5.16. Neural network structure for modulation classifier.....	83
Figure 5.17. Overall accuracy of the NN and DT classifiers at different SNRs. ....	86
Figure 5.18. Classification accuracy of DT classifier and NN classifier at 20dB SNR....	86
Figure 5.19. Classification accuracy of DT classifier and NN classifier at 15dB SNR....	87
Figure 5.20. Classification accuracy of DT classifier and NN classifiers at 10dB SNR..	87
Figure 5.21. Classification accuracy of DT classifier and NN classifier at 5dB SNR.....	87
Figure 5.22. Classification accuracy of DT classifier and NN classifier at 0dB SNR.....	88
Figure 5.23. Classification accuracy of DT classifier and NN classifier at -5dB SNR. ...	88
Figure 5.24. Comparison of results of proposed classifiers with A&N for SNR 20dB... 91	
Figure 5.25. Comparison of results of proposed classifiers with A&N for SNR 15dB....	91
Figure 5.26. Comparison of results of proposed classifiers with A&N for SNR 10dB....	92

## CHAPTER 6

Figure 6.1. Useful features of CPM modulation. ....	96
Figure 6.2. Decision tree for classification of digital signals including CPM .....	98
Figure 6.3. Total error probability for the key feature $\gamma_{maxf}$ .....	98
Figure 6.4. PSD of binary CPM with different pulse shapes ( $h = 0.5$ ) [Proakis, 1995]. 104	
Figure 6.5. Smoothed PSD of LREC signals ( $L=1$ and $L=2$ ) at SNR of 20dB. ....	106
Figure 6.6. Close up of PSD in Figure 1 around the peak. ....	106
Figure 6.7. Decision tree for CPM signals.....	107
Figure 6.8. Total error probability for the key feature $\sigma_{fn}$ for Decision A.....	108
Figure 6.9. Total error probability for the key feature $\sigma_a$ for Decisions B and C .....	108

Figure 6.10. Total error probability for features in Decisions D and E .....	109
Figure 6.11. Neural network structure for modulation classification. ....	113
Figure 6.12. Neural network structure for classification of CPM signals.....	114
Figure 6.13. Second NN structure to classify signals within the CPM class.....	116
Figure 6.14. Graphical comparison of overall performance of classifiers.....	117
Figure 6.15. Classification accuracy of DT classifier and NN classifier at 20dB SNR. .	118
Figure 6.16. Classification accuracy of DT classifier and NN classifier at 15dB SNR..	119
Figure 6.17. Classification accuracy of DT classifier and NN classifier at 10dB SNR..	119
Figure 6.18. Classification accuracy of DT classifier and NN classifier at 5dB SNR....	119
Figure 6.19. Classification accuracy of DT classifier and NN classifier at 0dB SNR....	120
Figure 6.20. Classification accuracy of DT classifier and NN classifier at -5dB SNR. .	120
Figure 6.21. Classification accuracy of DT and NN classifiers for CPM 20dB SNR. ...	121
Figure 6.22. Classification accuracy of DT and NN classifiers for CPM at 15dB SNR.	121
Figure 6.23. Classification accuracy of DT and NN classifiers for CPM at 10dB SNR.	122
Figure 6.24. Classification accuracy of DT and NN classifiers for CPM at 5dB SNR. .	122
Figure 6.25. Classification accuracy of DT and NN classifiers for CPM at 0dB SNR. .	122
Figure 6.26. Classification accuracy of DT and NN classifiers for CPM at -5dB SNR.	123
Figure 6.27. Graphical comparison between overall performance of CPM classifiers..	124

## CHAPTER 7

Figure 7.1. Useful features of BPSK DS-SS modulation.....	131
Figure 7.2. Useful features of QPSK DS-SS modulation. ....	132
Figure 7.3. Useful features of FH SS modulation. ....	134
Figure 7.4. Time division multiple access (TDMA). ....	135
Figure 7.5. Useful features of TDMA modulation.....	135
Figure 7.6. Smoothed power spectral density for PSK2 and BPSK DS-SS signals. ....	137
Figure 7.7. Flowchart for identification of digital modulation schemes.....	139
Figure 7.8. Total error probability for the key feature $\gamma_{\min}$ . ....	140
Figure 7.9. Total error probability for the key feature $\sigma_{ap}$ .....	141
Figure 7.10. Total error probability for the key feature $\gamma_{\min}$ . ....	141
Figure 7.11. Total error probability for separation of TDMA and FH SS.....	142

Figure 7.12. Neural network structure for modulation classifier.....	147
Figure 7.13. Classification accuracy of DT classifier and NN classifier at 20dB SNR.....	148
Figure 7.14. Classification accuracy of DT classifier and NN classifier at 15dB SNR.....	148
Figure 7.15. Classification accuracy of DT classifier and NN classifier at 10dB SNR.....	149
Figure 7.16. Classification accuracy of DT classifier and NN classifier at 5dB SNR.....	149
Figure 7.17. Classification accuracy of DT classifier and NN classifier at 0dB SNR.....	150
Figure 7.18. Classification accuracy of DT classifier and NN classifier at -5dB SNR. ....	150
Figure 7.19. The overall classification accuracy of the NN and DT classifiers .....	151

## CHAPTER 8

Figure 8.1. Useful features of PSK8 modulation.....	155
Figure 8.2. Useful features of QAM8 modulation. ....	155
Figure 8.3. Useful features of QAM16 modulation.....	156
Figure 8.4. Useful features of FSK8 modulation.....	156
Figure 8.5. Decision tree for identification of digital modulation scheme. ....	158
Figure 8.6. Total error probability for the key feature $\sigma_{dp}$ .....	161
Figure 8.7. Total error probability for separation of PSK4 and PSK8.....	162
Figure 8.8. ROC curves for separation of PSK4 and PSK8 .....	162
Figure 8.9. Total error probability for the key feature $\mu_{dp}$ for ASK2 and QAM8.....	163
Figure 8.10. Total error probability for the key feature $\mu_{dp}$ for ASK4 and QAM16. ....	163
Figure 8.11. Total error probability for the key feature $L_{diff}$ with bandlimitation 100kHz. 164	
Figure 8.12. ROC curves for the key feature $L_{diff}$ to separate FSK8 and FSK4 .....	165
Figure 8.13. Total error probability for the key feature $L_{diff}$ with bandlimitation 200kHz. 165	
Figure 8.14. ROC curves for the key feature $L_{diff}$ with bandlimitation 200kHz.....	166
Figure 8.15. Total error probability for the key feature $\sigma_{fn}$ with bandlimitation 200kHz...167	
Figure 8.16. ROC curves for the key feature $\sigma_{fn}$ with bandlimitation 200kHz.....	167
Figure 8.17. Neural network structure for digital modulation classification. ....	172
Figure 8.18. Graphical comparison of overall performance between classifiers.....	175
Figure 8.19. Classification accuracy of DT classifier and NN classifier at 20dB SNR.....	175
Figure 8.20. Classification accuracy of DT classifier and NN classifier at 15dB SNR.....	176
Figure 8.21. Classification accuracy of DT classifier and NN classifier at 10dB SNR.....	176

Figure 8.22. Classification accuracy of DT classifier and NN classifier at 5dB SNR.....	177
Figure 8.23. Classification accuracy of DT classifier and NN classifier at 0dB SNR.....	177
Figure 8.24. Classification accuracy of DT classifier and NN classifier at -5dB SNR. ...	178

## CHAPTER 9

Figure 9.1. Modulation classifier performance with Rayleigh fading for SNR 20dB.....	185
Figure 9.2. Total Error probability for the key feature $\gamma_{minf}$ with 120Hz Doppler shift. ....	188
Figure 9.3. Total error probability for the key feature $\sigma_{ap}$ with 120Hz Doppler shift. ....	189
Figure 9.4. Total error probability for the key feature $\sigma_{ap}$ in a Gaussian channel. ....	189
Figure 9.5. ROC curves for the key feature $\sigma_{ap}$ with 120Hz Doppler shift .....	190
Figure 9.6. Total error probability for the key feature $P_{min}$ with 120Hz Doppler shift. ....	191
Figure 9.7. ROC curves for the key feature $P_{min}$ with 120Hz Doppler shift. ....	191
Figure 9.8. Total error probability for the key feature $\gamma_{maxf}$ with 120Hz Doppler shift. ....	192
Figure 9.9. Total error probability for the key feature $\sigma_{fn}$ , with and without fading .....	192
Figure 9.10. Total error probability for the key feature $L_{diff}$ with 120Hz Doppler shift. ..	193
Figure 9.11. ROC curves for the key feature $L_{diff}$ with 120Hz Doppler shift .....	193
Figure 9.12. Total error probability for the key feature $\gamma_{maxf}$ with 120Hz Doppler shift. ..	194
Figure 9.13. ROC curves for the key feature $\gamma_{maxf}$ with 120Hz Doppler shift .....	195
Figure 9.14. Modified decision tree to accommodate signals with Rayleigh fading .....	200
Figure 9.15. Modified neural network structure for signals with Rayleigh fading.....	201
Figure 9.16. Graphical comparison of overall performance between classifiers.....	203
Figure 9.17. Classification accuracy of DT classifier and NN classifier at 20dB SNR.....	203
Figure 9.18. Classification accuracy of DT classifier and NN classifier at 15dB SNR.....	204
Figure 9.19. Classification accuracy of DT classifier and NN classifier at 10dB SNR.....	204
Figure 9.20. Classification accuracy of DT classifier and NN classifier at 5dB SNR.....	205
Figure 9.21. Classification accuracy of DT classifier and NN classifier at 0dB SNR.....	205
Figure 9.22. Classification accuracy of DT classifier and NN classifier at -5dB SNR .....	206
Figure 9.23. Comparison of performance between classifiers with and without fading....	206



# LIST OF TABLES

## CHAPTER 3

Table 3.1. The relationship of different error probabilities for difference in two means...	42
Table 3.2. The relationship of different error probabilities for classification .....	43

## CHAPTER 5

Table 5.1. Summary of key feature thresholds and error probabilities.....	76
Table 5.2. Total minimum error probability for different scenarios of Decision 1. ....	77
Table 5.3 Total minimum error probability for Scenarios 1 – 4 of Decision 2 .....	78
Table 5.4. Total minimum error probability for Scenarios 5 – 7 of Decision 2.....	79
Table 5.5. Total minimum error probability for Decision 3, 4 and 5.....	80
Table 5.6. DT and NN classifier accuracy and 95% confidence intervals.....	86

## CHAPTER 6

Table 6.1. Total minimum error probability for Decision 1. ....	99
Table 6.2. Summary of key feature values and corresponding threshold values.....	110
Table 6.3. Total minimum error probability for Decisions A, B and C.....	110
Table 6.4. Total minimum error probability for Decision D and Decision E .....	111
Table 6.5 DT and NN classifier accuracy and 95% confidence intervals.....	118
Table 6.6 Comparison of DT and NN classifiers for CPM signals.....	124

## CHAPTER 7

Table 7.1. CDMA 7-bit Gold code set [Ramakonar, 1996]. ....	130
Table 7.2. Summary of key feature thresholds and error probabilities. ....	142
Table 7.3. Total minimum error probability for Decisions 1 – 4.....	143
Table 7.4. DT and NN classifier accuracy and 95% confidence intervals.....	151

CHAPTER 8

Table 8.1. Summary of key feature thresholds and error probabilities..... 168

Table 8.2. Total minimum error probability for Decisions 1 – 5..... 169

Table 8.3. DT and NN classifier accuracy and 95% confidence intervals..... 174

CHAPTER 9

Table 9.1. Summary of key feature thresholds and error probabilities..... 195

Table 9.2. Total minimum error probability for Decisions 1 – 3..... 198

Table 9.3. Total minimum error probability for Decisions 4 – 7..... 199

Table 9.4. DT and NN classifier accuracy and 95% confidence intervals with fading. . 202

APPENDIX A

Table A.1. DT classifier confusion matrix for signals at SNR = 20dB (test set)..... 211

Table A.2. DT classifier confusion matrix for signals at SNR = 15dB (test set)..... 211

Table A.3. DT classifier confusion matrix for signals at SNR = 10dB (test set)..... 211

Table A.4. DT classifier confusion matrix for signals at SNR = 5dB (test set)..... 212

Table A.5. DT classifier confusion matrix for signals at SNR = 0dB (test set)..... 212

Table A.6. DT classifier confusion matrix for signals at SNR = -5dB (test set). .... 212

Table A.7. NN classifier confusion matrix for signals at SNR = 20dB (test set). .... 213

Table A.8. NN classifier confusion matrix for signals at SNR = 15dB (test set). .... 213

Table A.9. NN classifier confusion matrix for signals at SNR = 10dB (test set). .... 213

Table A.10. NN classifier confusion matrix for signals at SNR = 5dB (test set). .... 214

Table A.11. NN classifier confusion matrix for signals at SNR = 0dB (test set).. .... 214

Table A.12. NN classifier confusion matrix for signals at SNR = -5dB (test set)..... 214

APPENDIX B

Table B.1. DT classifier confusion matrix for signals at SNR = 20dB (test set)..... 215

Table B.2. DT classifier confusion matrix for signals at SNR = 15dB (test set)..... 215

Table B.3. DT classifier confusion matrix for signals at SNR = 10dB (test set)..... 216

Table B.4. DT classifier confusion matrix for signals at SNR = 5dB (test set)..... 216

Table B.5. DT classifier confusion matrix for signals at SNR = 0dB (test set)..... 216

Table B.6. DT classifier confusion matrix for signals at SNR = -5dB (test set).....	217
Table B.7. DT classifier confusion matrix for signals at SNR = 20dB.....	217
Table B.8. DT classifier confusion matrix for signals at SNR = 15dB.....	217
Table B.9. DT classifier confusion matrix for signals at SNR = 10dB.....	218
Table B.10. DT classifier confusion matrix for signals at SNR = 5dB.....	218
Table B.11. DT classifier confusion matrix for signals at SNR = 0dB.....	218
Table B.12. DT classifier confusion matrix for signals at SNR = -5dB. ....	219
Table B.13. NN classifier confusion matrix for signals at SNR = 20dB (test set). ....	219
Table B.14. NN classifier confusion matrix for signals at SNR = 15dB (test set). ....	219
Table B.15. NN classifier confusion matrix for signals at SNR = 10dB (test set). ....	220
Table B.16. NN classifier confusion matrix for signals at SNR = 5dB (test set). ....	220
Table B.17. NN classifier confusion matrix for signals at SNR = 0dB (test set). ....	220
Table B.18. NN classifier confusion matrix for signals at SNR = -5dB (test set). ....	221
Table B.19. Neural network 1 confusion matrix for signals at SNR = 20dB. ....	221
Table B.20. Neural network 1 confusion matrix for signals at SNR = 15dB. ....	221
Table B.21. Neural network 1 confusion matrix for signals at SNR = 10dB. ....	222
Table B.22. Neural network 1 confusion matrix for signals at SNR = 5dB. ....	222
Table B.23. Neural network 1 confusion matrix for signals at SNR = 0dB. ....	222
Table B.24. Neural network 1 confusion matrix for signals at SNR = -5dB. ....	223
Table B.25. Neural network 2 confusion matrix for signals at SNR = 20dB. ....	223
Table B.26. Neural network 2 confusion matrix for signals at SNR = 15dB. ....	223
Table B.27. Neural network 2 confusion matrix for signals at SNR = 10dB. ....	224
Table B.28. Neural network 2 confusion matrix for signals at SNR = 5dB. ....	224
Table B.29. Neural network 2 confusion matrix for signals at SNR = 0dB. ....	224
Table B.30. Neural network 2 confusion matrix for signals at SNR = -5dB. ....	225

## APPENDIX C

Table C.1. DT classifier confusion matrix for signals at SNR = 20dB (test set).....	227
Table C.2. DT classifier confusion matrix for signals at SNR = 15dB (test set).....	228
Table C.3. DT classifier confusion matrix for signals at SNR = 10dB (test set).....	229
Table C.4. DT classifier confusion matrix for signals at SNR = 5dB (test set).....	230



Table C.5. DT classifier confusion matrix for signals at SNR = 0dB (test set). .....	231
Table C.6. DT classifier confusion matrix for signals at SNR = -5dB (test set). .....	232
Table C.7. NN classifier confusion matrix for signals at SNR = 20dB (test set). .....	233
Table C.8. NN classifier confusion matrix for signals at SNR = 15dB (test set). .....	234
Table C.9. NN classifier confusion matrix for signals at SNR = 10dB (test set). .....	235
Table C.10. NN classifier confusion matrix for signals at SNR = 5dB (test set). .....	236
Table C.11. NN classifier confusion matrix for signals at SNR = 0dB (test set). .....	237
Table C.12. NN classifier confusion matrix for signals at SNR = -5dB (test set). .....	238

## APPENDIX D

Table D.1. DT classifier confusion matrix for signals at SNR = 20dB (test set). .....	239
Table D.2. DT classifier confusion matrix for signals at SNR = 15dB (test set). .....	240
Table D.3. DT classifier confusion matrix for signals at SNR = 10dB (test set). .....	241
Table D.4. DT classifier confusion matrix for signals at SNR = 5dB (test set). .....	242
Table D.5. DT classifier confusion matrix for signals at SNR = 0dB (test set). .....	243
Table D.6. DT classifier confusion matrix for signals at SNR = -5dB (test set). .....	244
Table D.7. NN classifier confusion matrix for signals at SNR = 20dB (test set). .....	245
Table D.8. NN classifier confusion matrix for signals at SNR = 15dB (test set). .....	246
Table D.9. NN classifier confusion matrix for signals at SNR = 10dB (test set). .....	247
Table D.10. NN classifier confusion matrix for signals at SNR = 5dB (test set). .....	248
Table D.11. NN classifier confusion matrix for signals at SNR = 0dB (test set). .....	249
Table D.12. NN classifier confusion matrix for signals at SNR = -5dB (test set). .....	250

## APPENDIX E

Table E.1 DT classifier confusion matrix for signals at SNR = 20dB with fading .....	251
Table E.2. DT classifier confusion matrix for signals at SNR = 15dB with fading .....	252
Table E.3. DT classifier confusion matrix for signals at SNR = 10dB with fading. ....	253
Table E.4. DT classifier confusion matrix for signals at SNR = 5dB with fading. ....	254
Table E.5. DT classifier confusion matrix for signals at SNR = 0dB with fading. ....	255
Table E.6. DT classifier confusion matrix for signals at SNR = -5dB with fading .....	256
Table E.7. NN classifier confusion matrix for signals at SNR = 20dB with fading .....	257

Table E.8. NN classifier confusion matrix for signals at SNR = 15dB with fading .....	258
Table E.9. NN classifier confusion matrix for signals at SNR = 10dB with fading. ....	259
Table E.10. NN classifier confusion matrix for signals at SNR = 5dB with fading. ....	260
Table E.11. NN classifier confusion matrix for signals at SNR = 0dB with fading. ....	261
Table E.12. NN classifier confusion matrix for signals at SNR = -5dB with fading. ....	262



# GLOSSARY OF TERMS

ALF	Average likelihood function
ALF	Average likelihood function
ALLF	Average log-likelihood function
ALLR	Average log-likelihood ratio
ALRT	Average likelihood ratio test
AM	Amplitude modulation
ANN	Artificial neural network
ASK	Amplitude shift keying
AWGN	Additive white Gaussian noise
BPSK	Binary phase shift keying
CDMA	Code division multiple access
CNR	Carrier-to-noise ratio
CPFSK	Continuous phase frequency shift keying
CPM	Continuous phase modulation
CW	Carrier wave
DFT	Discrete Fourier transform
DS	Direct sequence
DSB	Double sideband
DT	Decision – theoretic
FH SS	Frequency hopped spread spectrum
FM	Frequency modulation
FSK	Frequency Shift keying
GLRT	Generalised likelihood ratio test
GMSK	Gaussian minimum shift keying
HCS	Half cycle sinusoid
HWT	Haar wavelet transform
IF	Infrared
ISI	Intersymbol interference

k-NN	k-Nearest neighbour
LF	Likelihood function
LLF	Log-likelihood function
LMS	Least mean square
LSB	Lower sideband
MAP	Maximum a posteriori
ML	Maximum likelihood
MLE	Maximum likelihood estimation
MSK	Minimum shift keying
NN	Neural network
OBD	Optimal brain damage
OFDM	Orthogonal frequency division multiplexing
OOK	On-off keying
OQPSK	Offset quadrature phase shift keying
PBC	Phase based classifier
PLL	Phase-locked loop
PSD	Power spectral density
PSK	Phase shift keying
PSP	Per-surviving processor
QAM	Quadrature amplitude modulation
qLLR	quasi-log-likelihood ratio
QPSK	Quadrature phase shift keying
RC	Raised cosine
REC	Rectangular pulse
SLC	Square law classifier
SNR	Signal-to-noise ratio
SOSE	Sum of squared envelopes
SS	Spread Spectrum
SSB	Single sideband
TDMA	Time division multiple access
USB	Upper sideband

# CHAPTER 1

## Introduction

Modulation identification plays an important part in both covert and overt operations. The main aim in communication intelligence (COMINT) applications is the perfect monitoring of the intercepted signals. The modulation type of the intercepted signal is one of the parameters that affects perfect monitoring. In the past, radar and communication systems have relied on operator interpreted measured parameters to classify and identify signals. In modern warfare, there are dense electromagnetic environments and automatic processing techniques are required for rapid response. Therefore, automatic modulation classification is necessary. Modulation classification exploits several classical communication disciplines that include detection and estimation. It has recently attracted interest from both the military and commercial sectors due to its capability of replacing several receivers with one universal receiver. This has practical application for example in a network environment where it is required for an incoming signal to be routed to an appropriate processor. An automatic modulation classifier can be defined as a system that automatically identifies the modulation type of the received signal given that the signal exists and its parameters lie in a known range. This chapter is organised as follows: first the objectives of the thesis are described next, followed by the major contributions of the thesis. A description of the thesis organization is presented in Section 1.3, and finally the publications arising from this research are listed in Section 1.4.

### 1.1 Objectives of the Thesis

There has been some research conducted into the area of automatic modulation recognition by Azzouz and Nandi [Azzouz and Nandi, 1996]. They proposed modulation classifiers capable of recognising certain analogue signals and a limited number of digital

modulation schemes. However, they tested their algorithms on signals with SNR values greater than or equal to 10dB. With the advent of new technology using digital transmission, the proposed modulation classifiers described in this thesis are designed for digital communication signals only. Therefore, the objectives of this thesis are:

- To design a modulation classifier that is able to classify a comprehensive list of digital modulation schemes
- To use two types of classifier implementations - the decision-theoretic (DT) approach and the neural network (NN) approach.
- To be able to classify digitally modulated signals in the presence of additive white Gaussian noise (AWGN) at SNR values down to -5dB.
- To design classifiers that can handle a different environment than the AWGN channel, such as a Rayleigh fading channel.

## **1.2 Major Contributions of the Thesis**

The major contributions of this thesis are:

- A new decision tree design for the classification of ASK, PSK and FSK signals using different key features from those proposed by Azzouz and Nandi.
- The decision-theoretic and neural network modulation classifier capabilities are extended to include the classification of CPM, BPSK DS-SS, QPSK DS-SS, FH SS, TDMA, FSK8, PSK8, QAM8, and QAM16 signals.
- New key features are found to classify these signals.
- Signals within the CPM class can be classified as full response, partial response or GMSK and an associated decision tree is developed for that. Also two neural network designs have been proposed for the classification of CPM signals.
- The classifiers' capabilities are extended to signals affected by Rayleigh fading. The developed decision tree and neural network are modified to accommodate faded signals with a Doppler spread of 120Hz.

### **1.3 Organization of the Thesis**

The structure of the thesis is as follows:

1. Chapter 2 gives an overview of literature related to digital modulation classification. The different techniques related to modulation classification will be described and the modulation types that can be classified by each technique will be discussed. The main classification techniques discussed are the maximum likelihood approach, the pattern recognition approach, and the neural network implementations.
2. Chapter 3 describes the theory behind the classification process for the decision-theoretic approach. The Bayes decision rule for minimum error is described as well as a method to derive the Bayes error. Threshold determination is described using an example. Classifier accuracy, confidence intervals, statistical power and statistical significance are also discussed.
3. Chapter 4 presents the theory behind the classification process for the neural network implementation. The general concepts of neural networks including the different classes of neural networks and their structures, training algorithms and learning paradigms are discussed. An example of classification using neural networks is also presented.
4. In chapter 5, the classification of ASK, PSK, and FSK signals is presented using the decision-theoretic and neural network approaches. New key features are introduced and an alternate decision tree design is devised. This new tree is compared to the design of [Azzouz and Nandi, 1996] and the performances of both these DT classifiers are compared. A neural network using the same key features used in the devised decision tree is also designed and tested. The performances of both the NN and DT classifiers are compared and some conclusions are made.
5. Chapter 6 expands the DT and NN modulation classifiers proposed in chapter 5 to accommodate continuous phase modulated (CPM) signals. These classifiers are able to distinguish between CPM signals and other modulation types (ASK, PSK,



and FSK). The classifiers can also identify signals within the CPM class - the signals are recognised as partial response, full response or Gaussian minimum shift keying (GMSK) signals. The performances of the DT and NN classifiers are also compared with some concluding remarks.

6. Chapter 7 presents an extension to the capabilities of the modulation classifiers described in chapter 6 to include multiple access signals. These signals are direct sequence spread spectrum (DS SS) or code division multiple access (CDMA), frequency hopped spread spectrum (FH SS), and time division multiple access (TDMA). They are very commonly used in the military for their low probability of interception and also in civilian areas, such as in mobile networks to reduce call dropouts and interference. These different types of signals are included in the modulation classification algorithms, which employ the decision theoretic and neural network approaches. The results are compared and presented for each classifier with some conclusions.
7. Chapter 8 completes the development of the modulation classifier structure. In this chapter, PSK8, FSK8 QAM8, and QAM16 signals are added to the modulation classifiers. These modulation classification algorithms employ the decision theoretic and neural network approaches. This results in two types of modulation classifiers that are capable of distinguishing fifteen types of digitally modulated signals. The performances of the DT and NN classifiers are tested and compared in the presence of additive white Gaussian noise (AWGN). Estimates of the classification accuracy are derived for SNR (signal-to-noise ratio) ranging from 20dB to -5dB.
8. Chapter 9 tests the performances of the developed classifiers in the presence of Rayleigh fading. Both classifier structures are modified slightly to accommodate fading and the performances of the modified classifiers are compared to the results in an AWGN channel.

9. Chapter 10 presents some concluding remarks about the thesis. Some suggestions for further research are also presented in this chapter.
10. The thesis also includes a number of appendices where tables of results are presented for the classifiers developed in Chapters 5 - 9.

#### **1.4 Publications Arising From PhD Research**

1. Ramakonar, V., Habibi, D. and Bouzerdoun, A., "Automatic Recognition of Digitally Modulated Communications Signals". *Proceedings of ISSPA '99*, pp. 753-756, August 1999.
2. Arulampalam, G., Ramakonar, V., Bouzerdoun, A., and Habibi, D., "Classification of Digital Modulation Schemes Using Neural Networks", *Proceedings of ISSPA '99*, pp. 649-652, August 1999.
3. Ramakonar, V., Habibi, D. and Bouzerdoun, A., "Classification of bandlimited FSK4 and FSK8 signals". *Proceedings of ISSPA 2001*, August 2001.
4. Ramakonar, V., Habibi, D. and Bouzerdoun, A., "New Algorithm for Classification in Rayleigh Fading Channels of Spread Spectrum Communications Signals". *Proceedings of ISC 2001*, pp 530-535, November 2001.
5. Ramakonar, V., Habibi, D. and Bouzerdoun, A. "New Methods for Classification of CPM and Spread Spectrum Communication Signals". *Communications World (Electrical and Computer Engineering Series)*, pp303- 308, 2001.



# CHAPTER 2

## Classification of Digital Modulation Schemes: A Review

### 2.1. Introduction

This chapter presents a review of the literature relevant to modulation classification. There have been a number of articles published in this area, which describe classifiers that can recognise a limited number of modulation types [Azzouz, 1998; Wei, 2000; Jondral, 1994; Swami, 2000]. There is no comprehensive reference for a classifier encompassing many modulation schemes. This serves as the motivation for this thesis.

There are many types of modulation classification methods and a description of each scheme and relevant publications will be presented. The chapter is organised as follows: Classification using the maximum likelihood approach is outlined first followed by a description of the pattern recognition approach to classification. Finally modulation recognition using neural networks is discussed.

### 2.2. Maximum Likelihood Approach

With the maximum likelihood (ML) approach, the classification is viewed as a multiple hypothesis testing problem. This is where a hypothesis,  $\omega_i$ , is arbitrarily assigned to the  $i$ th modulation type of  $m$  possible types. The conditional pdf  $p(X|\omega_i)$ , ( $i = 1, \dots, m$ ) determines the ML classification and  $X$  is an observation, eg a sampled frequency component. If the observation sequence  $X[k]$ ,  $k = 1, \dots, n$  is independent and identically distributed (i.i.d), the likelihood function (LF),  $L(X|\omega_i)$  can be expressed as [Stark, 1994]:

$$p(X | \omega_i) = \prod_{k=1}^n p(X[k] | \omega_i) = L(X | \omega_i) \quad (2.1)$$

The ML classifier outputs the  $j$ th modulation type based on the observation whenever  $L(X|\omega_j) > L(X|\omega_i)$ ,  $j \neq i$ ;  $j, i = 1, \dots, m$ . The log-likelihood function (LLF) can be used if the likelihood function is exponential due to the monotonic nature of the exponent function. It is common for the expressions of the pdf to be approximations and assume prior information of the symbol rate and SNR. Therefore quasi-optimal rules are defined. A general ML classifier is shown in Figure 2.1.

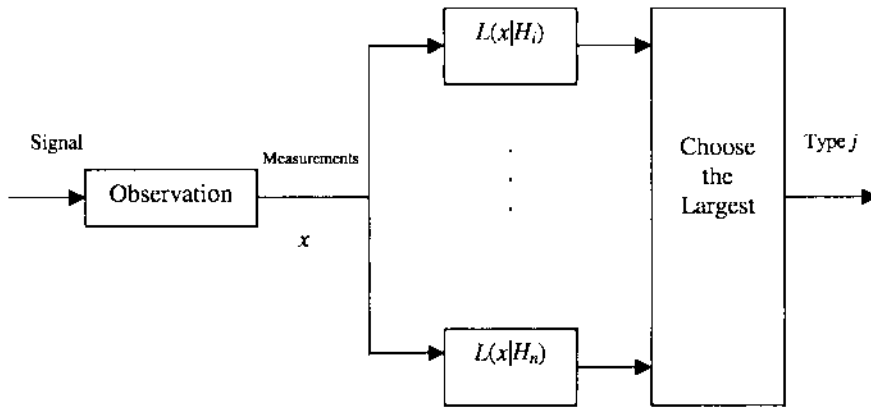


Figure 2.1. General maximum likelihood classifier.

We will outline general maximum likelihood classification techniques first, followed by a description of a ML classifier capable of recognising MPSK schemes based on the exact phase distribution. Classifiers that are based on the likelihood functions are then presented and we will describe how the ML function is used to classify continuous phase modulation (CPM). Finally a ML classifier using constellation shape is discussed.

### 2.2.1. General Maximum Likelihood Methods

A classifier capable of recognising digital amplitude modulated signals was proposed by Wei and Mendel [Wei, 1995]. The method was based on the ML approach and is applicable to any constellation based modulation type in an additive white Gaussian noise (AWGN) channel. The theoretical performance of the ML classifier under ideal conditions

was reported and this in turn serves as an upper bound of performance for any classifier. It was assumed that all signal parameters are known.

The classifier in [Wei, 1995] was extended to include PSK and QAM signals in [Wei, 2000]. It was shown that the I-Q domain data were sufficient statistics for modulation classification. A generic formula for the error probability of a ML classifier was obtained and an asymptotic performance study was carried out. The theoretical performance was derived under an ideal situation where all signal parameters as well as the noise power are known. The data symbols are independent and the pulse shape is rectangular. The classifier can accommodate any finite set of distinct constellations with zero error rate when the number of data symbols approaches infinity. Simulations were performed with SNR ranging from 0 to 15 dB.

A maximum likelihood classifier for QAM and PSK signals was proposed by Sills [Sills, 1999]. The classifier algorithms were designed for coherent and noncoherent conditions. The algorithm's performance was evaluated for PSK2, PSK4, PSK8, QAM16, QAM32 and QAM64 signals and compared with a psuedo maximum-likelihood noncoherent classification technique in terms of error rate, false alarm rate, and computational complexity. It was stated that the coherent ML classifier makes less than one error in ten across all six modulation types provided that the SNR is greater than or equal to 10dB. For the noncoherent ML classifier, there is less than one error in ten across the tested modulation types for SNR greater than or equal to 13dB. It was found that using a large number of symbols in the likelihood ratio reduces the probability of error and probability of false alarm.

A general ML classifier based on an approximation of the likelihood function was developed by Boiteau and Le Martret [Boiteau, 1998]. Equations were derived for the case of linear modulation and applied to MPSK signals. It was shown that the tests are a generalisation of the previous methods using the ML approach discussed in section 2.2.3. It was found that the likelihood function of an observation can be approximated by measuring the correlation between the higher-order statistic (or true temporal) and the

empirical. Thus, this type of classifier provides a theoretical foundation for systems that exploit cyclostationary properties to classify signals as well as many other empirical classification systems.

### **2.2.2. MPSK Classifier Based on the Exact Phase Distribution**

The classification of MPSK signals using an asymptotic optimal algorithm has been achieved by Yang and Liu [Yang, 1998]. The same results but with slightly different test statistics were also published earlier by Yang and Soliman in [Yang, 1991] and [Yang, 1997]. The exact phase distribution of a received MPSK signal was expressed in terms of the Fourier series expansion to develop the classification algorithm. The classifier was capable of recognising CW, BPSK, QPSK, and 8PSK signals. A multiple hypothesis classification rule was developed using the maximum a posteriori (MAP) probability rule, which was consequently reduced to a ML classifier. This is because the hypotheses were assumed equally likely. The SNR was assumed to be known and the classifier was shown to outperform the classifiers proposed in [Yang, 1991] and [Yang, 1997].

### **2.2.3. Classifiers Based on the Likelihood Functions**

There are six publications based on classification using likelihood functions that will be discussed. The first article describes a quasi log-likelihood classifier. The second publication performs a comparison on the performance of a  $M^{\text{th}}$ -law classifier and a  $q_M$ -rule classifier. Thirdly, an average log-likelihood classifier is discussed, followed by a description of a multiple hypothesis classifier. The fifth article, describes classification in unknown ISI environments using a LF. Finally we outline classification (based on likelihood functions) of QAM signals using the DFT of phase histograms.

#### **2.2.3.1. Quasi Log-Likelihood Ratio Classifier**

Polydoros and Kim [Polydoros, 1990] derive and analyse optimal and suboptimal decision rules for the detection of constant envelope quadrature digital modulations in the presence of noise. No timing or frequency uncertainty was assumed and signal parameters such as carrier frequency, initial phase, symbol rate and SNR were assumed to be known. The

effect of various stochastic models for the carrier phase was examined. The modulation classifier was for BPSK/QPSK signals based on an approximation of the likelihood function. A comparison between three classifiers for MPSK signals was introduced. The three classifiers were a phase based classifier (PBC) that is based on the phase histogram, a square law classifier (SLC) that is based on the fact that squaring an MPSK signal results in another MPSK signal with  $M/2$  states, and finally the quasi-log-likelihood ratio (qLLR) which is derived by approximating the likelihood ratio functions of phase modulated digital signals in white gaussian noise. The authors have proved analytically that the last method performs better than conventional phase-based and square-law classifiers particularly for lower signal-to-noise ratios (SNR).

#### **2.2.3.2. $M^{\text{th}}$ - law Classifier Versus $q_M$ -Rule**

A maximum likelihood classifier based on the likelihood function of MPSK and MQAM signals in AWGN noise was proposed by Hwang and Polydoros [Hwang, 1991]. Simplified versions of the likelihood function for each modulation type are represented by the  $q_M$  statistic. The  $q_M$  classifier is similar to a synchronous pulse-shaped matched filter. Its performance was compared with other  $M^{\text{th}}$ -law methods and the correct classification probability was found by having a long observation time ( $N \gg 1$  symbols) and estimated for a low SNR ( $\text{SNR} \ll 0\text{dB}$ ). For the  $M^{\text{th}}$ -law classifier to have comparable performance, the SNR had to be more than 2dB greater than the SNR for the  $q_M$  classifier. All signal parameters such as symbol rate, initial phase, carrier frequency and SNR were assumed to be known and the  $q_M$  classifier was only valid for SNR less than 0dB.

#### **2.2.3.3. Average Log-Likelihood Ratio Classifier**

The low SNR methods in [Hwang, 1991] were modified to accommodate higher SNR by Long, Chugg and Polydoros [Long, 1994]. The  $Q_M$  – rule, based on the average log-likelihood ratio (ALLR) was developed. An approximate expression for the pdf of the  $Q_M$  statistic was also developed for medium and high SNR environments. It was found that the approximation of the ALLR had better performance than the  $q_M$  rule in [Hwang, 1991]. The performance was evaluated for four different cases including CPFSK interference. All



signal parameters were assumed to be known and the classifier was developed for binary hypothesis testing.

#### ***2.2.3.4. Multiple Hypothesis Classifier***

The maximum likelihood classifier had been extended to estimate power and threshold setting automatically by Long, Chugg and Polydoros [Chugg, 1995]. The classifier also included more than two hypothesised modulation types. These modulation types were BPSK/QPSK/OQPSK. The modulation classification was based on the average likelihood function (ALF) and the threshold setting was based on the quasi log-likelihood ratio test. An estimate for the signal power based on the maximum likelihood function was derived. It was found that a reliable power estimate is hard to obtain when only in-band measurements are available.

#### ***2.2.3.5. Classification in Unknown ISI Environments***

A classification method for signals affected by intersymbol interference (ISI) was proposed by Lay and Polydoros [Lay, 1995]. It was assumed that the channel impulse response is known. An average likelihood ratio test (ALRT) and a generalised likelihood ratio test (GLRT) were derived. The channel identification was carried out simultaneously using per-surviving processing (PSP). Simulations were carried out for 16-ary digital modulations in known and unknown channels. It was found that the ALRT outperforms the GLRT but requires explicit knowledge of the signal power and noise variance of the channel. On the other hand, the GLRT only requires the ML estimate of the transmitted data. The Viterbi algorithm reduces the computational load of the decision statistics considerably, however the simultaneous classification and channel estimation is a time consuming task that may affect the classification tests detrimentally. This classifier was developed for binary hypothesis testing and it is assumed that all signal parameters are available except the impulse response.

#### **2.2.3.6. QAM Classification using DFT of Phase Histogram Combined With Modulus Information**

A method to classify various QAM signal constellations by analysing the DFT of the phase histogram and applying the magnitude distribution has been proposed by Schreyogg and Reichert [Schreyogg, 1997]. The likelihood functions were derived as well as a rule to combine them for classification. The LF was phase based and the pdf of the DFT bins of the phase histogram was used to derive the function. A LF based on the modulus was also derived and computed from the pdf of the constellations magnitude. The performance of the classifier was evaluated for a few different QAM constellations as well as BPSK, QPSK, and 8PSK signals.

#### **2.2.4. CPM Classification using ML Function**

Classification of CPM signals according to their modulation indices has been reported in [Chung, 1994] and [Huang, 1992]. Two classification rules based on the log likelihood function (LLF) for CPM signals in low SNR were proposed in [Huang, 1992]. The signals are passed through an AWGN channel and the classifier can differentiate two single-index CPM signals with different modulation indices  $h_1$  and  $h_2$ . The rule  $e(h_1, h_2)$  is equivalent to an energy comparator and the second rule  $c(h_1, h_2)$  has an original form. It was found that the second rule performs better than the first rule with short observations.

#### **2.2.5. Classification Using Constellation Shape as A Robust Signature**

A classification technique which uses the signal constellation shape as a stable modulation signature was proposed by Mobasseri, [Mobasseri, 1999]. The algorithm was designed for an AWGN channel and accounts for the presence of carrier recovery errors. The recovered constellations were modelled by binomial nonhomogenous spatial random fields. Experimental results were shown for various modulation standards including V.29, V.29\_fallback, PSK8 and QAM16. It was stated that the PSK8 signal can be correctly classified 90% of the time at SNR of 0dB. It was also stated that for the V.29, the classifier achieves performance levels exceeding 90% in the presence of large peak phase lock error and SNR of 3dB.

### 2.3. Pattern Recognition Approach

Generally, a pattern recognition system consists of sensing, feature extraction and decision procedures [Tou, 1974]. Each measurement, observation or pattern vector  $x = (X[1], X[2], \dots, X[n])^T$  describes a certain characteristic of the object or pattern. The size of the pattern vectors can be reduced because they often contain redundant information. This reduction in size of the pattern vectors is referred to as the feature extraction or preprocessing stage. The decision procedure can be a neural network, decision function, or distance function. The block diagram of a pattern recognition system is shown in Figure 2.2.

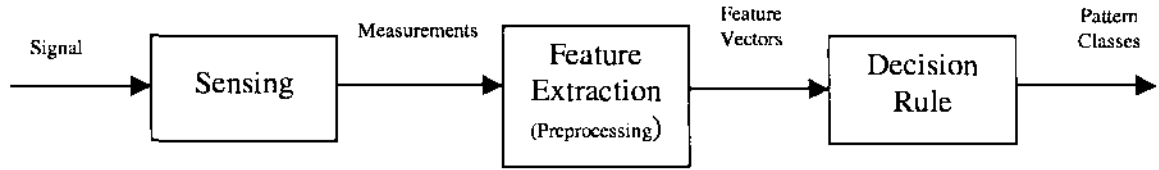


Figure 2.2. General pattern recognition system.

This section will be divided into three parts. The first part describes a pattern recognition approach using envelope-based methods, the second part describes a classification technique based on higher-order statistics, and the final part describes other methods of classification using the pattern recognition approach.

#### 2.3.1. Envelope-Based Methods

Classification using envelope-based methods can be accomplished using the ratio of different envelope statistics or deviations of instantaneous properties. Both techniques will now be discussed.

##### 2.3.1.1. Ratio of Different Envelope Statistics

A classifier based on the ratio ( $R$ ) of the variance of the envelope to the square of the mean of the envelope has been proposed in [Chan, 1989]. The classification method was based on four modulation types (AM, DSB, FM and SSB) and the equations for  $R$  were derived as a function of the carrier-to-noise ratio (CNR). The signal was classified according to where

the value of  $R$  lies. It was found that the length of the signal segment and the computation time were short making this method desirable for real time applications.

The classifier in [Chan, 1989] was extended to include more features based on the analytic envelope and on an approximation of the envelope, extracted by different means that did not require the Hilbert transform. The authors in [Druckmann, 1998] employ ratios of different statistics of these two envelopes to extract key features. The classification rule uses two features and the success rate of classification was reported to be 99% for a CNR of 10dB. This method, however, was not suitable for complex envelope representation.

An adaptive technique for classifying some types of digital modulations (ASK2, PSK2, PSK4 and FSK2) was introduced by DeSimio and Glenn [DeSimio, 1988]. Key features were derived from the signal envelope, signal spectra, the square of the signal, and the fourth power of the signal. These key features are the mean and the variance of the envelope, the magnitude and location of the two largest peaks in the signal spectrum, the magnitude of the spectral component at twice the carrier frequency of the signal squared, and the magnitude of the spectral component at four times the carrier frequency. The classification procedure was as follows:

- 1) feature vectors extraction,
- 2) weight vectors generation for each signal type, and
- 3) modulation classification.

The LMS algorithm is an adaptive technique that was used to generate decision functions. Also, the decision rule used was similar to that applied to pattern recognition algorithms. The classifier was trained using the values of the extracted key features at 20dB SNR. The classifier has the ability to discriminate between PSK2 and PSK4 signals at an SNR of 5dB.

#### ***2.3.1.2. Deviations of Instantaneous Properties***

One of the first authors to publish on modulation classification was Liedtke [Liedtke, 1984]. His work covers modulation recognition of digital signals. The modulation schemes covered were ASK2, PSK2, PSK4, PSK8, FSK2 and CW and use the universal

demodulator technique. To distinguish between different signals, key features such as the amplitude histogram, the phase difference histogram, the frequency histogram, the frequency variance and the amplitude variance were used. The classification procedure involves approximate signal bandwidth estimation, signal demodulation and parameter extraction, statistical computation and finally automatic modulation classification. It was claimed that when the signal's parameters are exactly known, the signal could be recognised at SNR values greater than or equal to 18dB.

Nandi and Azzouz have devised two algorithms for modulation classification [Azzouz, 1995]. Their algorithm encompasses both analog and digital signals. The first algorithm uses the decision-theoretic (DT) approach in which a set of decision criteria is developed for identifying different types of modulation. The second algorithm utilises artificial neural networks (ANN) as a new approach to modulation recognition [Azzouz, 1998]. Through simulations it was found that for the decision-theoretic algorithm, the overall success rate was over 94% at SNR of 15dB. The ANN algorithm had an overall success rate of over 96% at SNR of 15dB. All key features were considered simultaneously in the ANN approach whereas with the DT approach, each feature was considered one at a time against a certain threshold value. The success rate depends on the order of the features in the branches of the decision tree. This would imply that the ANN approach gives better results and this was found to be true. This research is based on these authors' work as it serves as a foundation to design a classifier that is capable of recognising a large range of modulation types.

A modulation recogniser using a pattern recognition approach was proposed by Jondral [Jondral, 1984]. The key features were extracted from the instantaneous amplitude, frequency and phase. These key features are the instantaneous amplitude, phase difference and frequency histograms. The received signal was divided into two adjacent sets called the learning set and the test set and the signal segment length was 4096 samples. Real signals had been used and the classification success rate was greater than or equal to 90% except for SSB (83%) and FSK4 (88%).

Aisbett [Aisbett, 1997] had developed a classifier that utilises signal parameters  $A^2$ ,  $AA'$  and  $A^2\theta'$  where  $A$  is the signal envelope derivative and  $\theta'$  is the instantaneous frequency. The key features used were the peak and tail values of the parameters  $A^2$ ,  $AA'$  and  $A^2\theta'$  as well as the variance of the squared instantaneous amplitude minus its squared mean. The types of signals that could be classified were ASK2, PSK2, DSB, AM, FM, and CW. The performance was claimed to be good for signals with higher SNR.

A modulation classifier that can recognise analog and digital signals has been proposed by Dominiguez et al [Dominiguez, 1991]. The recogniser can differentiate ASK2, ASK4, PSK2, PSK4, FSK2, FSK4, AM, DSB, FM, SSB, CW and noise. The number of samples per segment needed for the performance evaluation is 3000. There were three subsystems in the recogniser:

1. pre analysis,
2. features extraction,
3. classifier subsystem.

The key features were extracted from the histograms of the instantaneous amplitude, frequency and phase. It was claimed that for SNR values greater than or equal to 40dB, all modulation types were classified correctly. At SNR of 10dB no digitally modulated signals were classified correctly.

A modulation recogniser for multichannel systems was introduced by Nagy [Nagy, 1994]. With this modulation detector, the analysed signal was divided into individual components and each signal component was classified using a single – tone classifier. The signals that can be recognised are CW, ASK2, PSK2, PSK4 and FSK2. The classification process was as follows:

1. Each signal component in the estimated amplitude spectrum is detected and filtered, for eg the FSK2 signal is considered as two correlated ASK2 signals.
2. The differential phase is calculated to discriminate between the different types of single – tone signals.
3. Finally, all ASK2 signals are correlated to detect the FSK2 signals.

The single – tone classifier carries out the following tasks:

1. The amplitude histogram is used to discriminate the ASK2 signal from the CW, PSK2 and PSK4 signals.
2. The phase histogram is used to distinguish between CW, PSK2 and PSK4 signals.

It was stated that CW, PSK2 and PSK4 signals have been classified with greater than 98% success rate at 10dB SNR and the ASK2 signal with success rate of 87%.

A modulation recogniser for AM, FM, CW, ASK2 and FSK2 signals has been proposed by Martin [Martin, 1990]. The key features were extracted from the IF signal spectrum, its derivative and the instantaneous amplitude. These key features were the signal bandwidth, amplitude histogram and the relationship between spectral components. The signals have been classified with a success rate greater than 90% except for FM with 80% success rate.

Taira and Murukami [Taira, 1999] describe a modulation classification technique for analogue modulated signals including phase continuous FSK signals. For the discrimination between frequency modulation signals and amplitude modulation signals and for classification among the amplitude modulated signals, the statistical parameters of the signal envelope were used. For classification among the frequency modulated signals, the compactness of instantaneous frequency distribution was used. It was reported that good classification possibility has been ascertained by simulation when SNR is greater than or equal to 10dB. Discrimination between analogue and digital modulation schemes was accomplished via block processing.

### **2.3.2. Higher-Order Statistical Methods**

Six different methods for classification using higher-order statistics will be discussed in the following sections. The first method employs higher-order statistics to classify MPSK signals. The second method exploits the differences in higher-order moments, while the third method utilises cyclic temporal cumulants for classification. The fourth technique discussed uses time-domain higher-order correlations to classify FSK signals. Fourth-order cumulants are used to recognise certain digital modulation schemes in the fifth reviewed publication, and finally the time-average of the complex envelope is employed in the sixth classification method using higher-order statistics.

#### **2.3.2.1. Even Moment Based MPSK Classifier**

Soliman and Hsue [Soliman, 1992] investigate signal classification using statistical moments. The type of signals they considered were  $M$ -ary PSK signals. They show that for  $M$ -ary PSK signals, the  $n$ th moment ( $n$  even), of the phase of the signal is a monotonic increasing function of  $M$ . From this, an analytic expression for the probability of a misclassification was derived. A decision rule and a general hypothesis test were also developed. The classification procedure was as follows:

- 1) the instantaneous phase is extracted,
- 2) even order moments are calculated,
- 3) threshold comparison, and
- 4) modulation recognition.

All the signal parameters were assumed to be known. The performance of the algorithm was demonstrated by two examples. It was found that the eighth moment is adequate to identify BPSK signals with reasonable performance at low CNR. The suggested algorithm was compared to the qLLR method in [Polydoros, 1990] and also the square-law and phase-based methods. The qLLR method outperformed the proposed algorithm at low CNR but the latter was comparable to the square-law classifier and was better than the phase-based classifier. However, the qLLR classifier is only valid at CNR less than 0dB and can only be used to distinguish between BPSK and QPSK signals whereas the moments algorithm is more general.

Yang and Soliman [Yang, 1995] modified the modulation detector in [Soliman, 1992] by approximating the probability distribution function of the instantaneous phase. Instead of using the Tikhonov probability density function to approximate the exact phase distribution, the Fourier series expansion was used. The modification improved the results by 2dB for 99% success rate of modulation recognition. Also the computation for the  $n^{\text{th}}$  order moments was simpler than that proposed in [Soliman, 1992].

#### **2.3.2.2. The $M^{\text{th}}$ -law Based Classifier**

A classification method, which exploits the differences in the higher-order moment-spaces of the discrete-time modulating process, was proposed in [Reichert, 1992]. The carrier



frequency and symbol rate were assumed to be unknown so these differences in the higher-order moments contribute spectral lines associated with these unknown parameters. The spectral lines were detected by periodogram analysis and their existence, position and amplitude contribute to robust key features. It was possible to classify ASK2, PSK2, PSK4, MSK and FSK2. A complete statistical analysis of the classification performance was reported in terms of the probability of detection and false alarm rate. The theoretical performance figures were verified with simulations. A disadvantage of this method is that it is unsuitable for the complex envelope representation and the periodogram analysis is quite complex.

#### ***2.3.2.3. Cyclic Multi-correlation Based MPSK Classifier***

A multiple hypothesis QAM modulation classifier, utilising decision theory, was proposed by Marchand et al [Marchand, 1997 Marchand, 1998]. The same features have also been mentioned in [Le Martret, 1997] but a slightly different structure has been used. The proposed feature comprises a combination of fourth-order and squared second-order cyclic temporal cumulants. This combination was used to counter the uncertainty in the signal power. Simulations were carried out for 4QAM, 16QAM, and 64QAM signals. The performance was evaluated for SNR of 5dB and 10dB and it was found that the success rate was poor for sample sizes less than 1024 symbols. It was also stated that the authors are the only people to classify QAM signals exploiting cyclostationary properties.

#### ***2.3.2.4. Time-Domain Higher-Order Correlation MFSK Classifier***

A modulation classifier for MFSK signals was proposed by Beidas and Weber [Beidas, 1995]. The classifier was used to distinguish between MFSK signals and is based on the time – domain higher-order correlations. Two types of classifiers were presented: channelised and non-channelised. The channelised classifier was made up of a bank of matched filters and a set of successive correlators. Each matched filter was tuned to one of the designated frequency locations. In the non-channelised classifier, each signal was divided into three adjacent subbands – lower, middle and high. There were also three parallel processes assigned to each subband. Three algorithms were considered for the non-

channelised classifier which were: 1) a first-order correlation based classifier where three energy processors and three correlators were used, 2) a second-order correlation based classifier (type 1) where six correlators were used and 3) a second-order correlation based classifier (type 2) where three energy processors and six correlators were used. The log-likelihood function compared to a suitable threshold was used to decide about the number of levels of an MFSK signal. It was stated that the non-channelised classifiers can detect exact frequency locations perfectly.

#### **2.3.2.5. Classification Using Fourth-Order Cumulants**

Swami and Sadler devise a classification method based on elementary fourth-order cumulants for digital modulation schemes [Swami, 2000]. These statistics are said to characterise the shape of the distribution of the noisy baseband I and Q samples. It was shown that cumulant-based classification is particularly effective when used in a hierarchical scheme. This enables separation into subclasses at low SNR with small sample size which makes it appropriate for a preliminary classifier. The computational complexity is of order  $N$ , where  $N$  is the number of complex baseband samples. This method has been shown to be robust in the presence of carrier phase and frequency offsets and can be implemented recursively. Theoretical arguments were verified with simulation results and compared with existing approaches. The modulation schemes that can be classified are M-ASK, M-PSK and QAM signals. Results show that the classifier performs with 100% accuracy for 500 samples and SNR of 10dB even in the presence of carrier phase and frequency offsets.

Akmouche has proposed a classifier that discriminates single carrier modulations from multi-carrier modulations of OFDM type [Akmouche, 1999]. It was stated that multi-carrier methods are asymptotically Gaussian and therefore the proposed detector uses the statistical test of [Giannakis, 1990] based on fourth-order cumulants. The test was adapted by Akmouche to the specific case of digital modulations which reduces the algorithm complexity. Simulations were provided and show that for the worse case (filtered QAM-256 versus 32-OFDM), the detector achieves a probability of detection  $P_D$  of 0.99 for a probability of false alarm  $P_{FA}$  equivalent to 0.01.

#### ***2.3.2.6. Classification Based on Time Average of Complex Envelope.***

Rosti has addressed the feature extraction process of modulation classification [Rosti, 1998]. Useful characteristics and representations of communications signals were presented as well as the relevant knowledge of statistical signal processing. First and second order statistics of digital modulated signals were studied and a novel feature was proposed. This novel feature was based on the time average of the complex envelope representation of the digital signal. Previous methods and this novel feature were compared by investigating their discrimination performance through Matlab simulations. Modulation types that can be classified are: AM, DSB, SSB, FM, CW, PSK2, PSK4, FSK2 and FSK4.

#### **2.3.3. Other Methods**

Other methods for modulation classification using the pattern recognition approach will be outlined in the following sections. The first method employs the zero-crossing technique for classification. The second approach uses a modulation model and the third method classifies signals using distance functions. CPM signals are classified in the fourth publication using the sum of squared envelopes. The fifth technique discussed utilises time-frequency methods for signal recognition and the sixth and seventh approaches employ the discrete Fourier transform and the Wavelet transform respectively, for classification. Power moment matrices are employed in the eighth publication, and finally spread spectrum signals are classified using a modulation domain measurement technique.

##### ***2.3.3.1. Classification Using Zero-Crossings***

Hsue and Soliman use a zero-crossing technique for classification and report the findings in [Soliman, 1989] and [Soliman, 1990]. The zero-crossing sampler has the advantage of providing accurate phase transition information over a wide frequency range. The modulation recognition was achieved by utilising features such as phase difference and zero-crossing histograms. Signal parameters such as zero-crossing variance, carrier-to-noise ratio (CNR) and carrier frequency were estimated. The modulation detection was achieved by the following steps:

- 1) extraction of the zero-crossing sequence, the zero-crossing interval sequence and the zero-crossing interval difference sequence,

- 2) inter-symbol transition detection and carrier frequency estimation, and finally
- 3) modulation detection.

The zero-crossing sequence, the zero-crossing difference sequence and the zero-crossing interval difference sequence were all used to derive the phase and frequency information. The modulation type was decided from the variance of the zero-crossing interval sequence as well as the phase and frequency histograms. The types of signals considered were CW, MPSK and FSK signals. The recogniser first distinguishes between single-tone (CW and MPSK) and multi-tone signals (FSK) by comparing the variance of the zero-crossing difference sequence in the non-weak intervals of the signal with a suitable threshold. Then the number of levels ( $M$ ) in a single-tone signal was found by measuring the similarity of the normalised phase difference histogram. The number of levels in a multi-tone signal was found based on the number of hills in the zero-crossing interval difference histogram. From simulation results the authors found that a reasonable average probability of correct classification was possible for CNR greater than or equal to 15 dB.

Callaghan et al [Callaghan, 1985] have utilised the envelope and zero-crossing characteristics of the intercepted signal in their modulation classifier. A phase-locked loop (PLL) was used for carrier recovery in the weak intervals of the signal segment. In signals with modulation types such as MPSK, AM and DSB, the carrier frequency may be absent or severely suppressed and this is equivalent to having a signal with a low SNR. Therefore a high SNR is not required for accurate frequency estimation if a PLL is used for carrier recovery. If the receiver was not perfectly tuned to the carrier frequency then the performance of the recogniser deteriorated. The types of signals that could be recognised are AM, FM, FSK2, and CW. For correct recognition, the SNR must be greater than or equal to 20dB. The noise on the weak intervals of the signal segment caused incorrect estimate of the instantaneous frequency, and thus DSB and MPSK signals could not be discriminated.

Petrovic et al [Petrovic, 1989] has designed a modulation recogniser based on the zero-crossing rate and parameter variations of the AM detector output. In addition the parameter

variations in the FM detector output were also considered. The classification procedure was as follows:

1. AM and FM demodulation,
2. key features extraction,
3. modulation classification.

The signals that could be recognised are ASK2, FSK2, AM, FM, CW and SSB. For the FM detector output, both a narrow band and a wide band FM detection were performed. The key features were extracted from the AM detector output and it was stated that the results from preliminary tests with real signals show the successfulness of the classifier.

#### ***2.3.3.2. Classification Based on the Modulation Model***

Another modulation recogniser for digital modulation types was introduced by Assaleh et al [Assaleh, 1992]. The types of signals that could be recognised are CW, PSK2, PSK4, FSK2 and FSK4. The classification method uses a signal representation known as the modulation model. The modulation model was formed via autoregressive spectrum modelling. The key features were derived from the averaged spectrum of the instantaneous frequency. These key features are the mean and standard deviation of the averaged instantaneous frequency, the height of the spikes in the differential instantaneous frequency and the mean and standard deviation of the instantaneous bandwidth. It was claimed that the success rate for the different modulation types is greater than 99% at a SNR of 15dB.

#### ***2.3.3.3. Classification Based on Distance Functions***

A classification technique that uses the counts of the signals falling into different parts of the signal plane was proposed by Huo and Donoho [Huo, 1998]. The advantage of using the number of counts as a key feature is that the computation time is much faster than methods based on higher-order statistics and likelihood methods. To find the optimal place to partition the signal plane, the multinomial distributed Hellinger distance was maximised for two candidate modulation types. The performance of the classifier was evaluated for 4QAM and 6PSK and it was found that the proposed algorithm is dependent on the orientation of the symbols in the signal space. This makes this method suitable for binary classification only.

#### **2.3.3.4. Classification of CPM**

Two 2CPFSK signal classifiers based on the sum of squared envelopes (SOSE) were proposed in [Chung, 1994]. The classifiers were envelope based and developed for both single and multi-index CPM signals. In the first method, a variety of modulation sets were classified using an appropriately adjusted threshold. The second method was based on the approximate maximum likelihood estimation (MLE) of the index pattern derived from the SOSE and can be used for an infinite number of index sets. The proposed algorithm was compared to the LLF method in [Huang, 1992] and it was found that the LLF method performs better. However, the SOSE method was more robust at lower SNR.

#### **2.3.3.5. Time-Frequency Methods**

A new technique for feature extraction of modulation recognition based on the pattern recognition approach was proposed by Ketterer, Jondral and Costa [Ketterer, 1999]. The new algorithm exploits the Margenau-Hill distribution, autoregressive modelling and amplitude variations to detect phase shifts, frequency shifts, and amplitude shifts respectively. This method requires no a priori information about the signal and can classify PSK2, PSK4, PSK8, PSK16, FSK2, FSK4, QAM8, and OOK signals. The authors recommend this method in a general non-cooperative environment and state that their method is also computationally inexpensive. Simulations were carried out on synthetic and “real world” short-wave signals. Results indicated that this approach is robust against noise up to an SNR of around 10dB, where an overall success rate greater than 94% is obtained.

#### **2.3.3.6. Classification Using Discrete Fourier Transform**

A signal classification method using the discrete Fourier transform (DFT) was proposed in [Lallo, 1999]. To classify a signal the following steps were taken:

1. The carrier frequencies are obtained from the DFT of the signal.
2. The symbol rate is found once the carrier is known.
3. The amplitude and phase values of discontinuous functions using Euler’s formulae are calculated from the DFT.

4. The calculated phase and amplitude distributions are used for modulation analysis for each carrier frequency.

Tests were carried out over the telephone network and GSM radio for PSK2, QAM16, QAM40, QAM60 and FH signals. It was stated that satisfactory results are obtained.

#### **2.3.3.7. *Classification using the Wavelet Transform***

Ho, Prokopiw and Chan [Ho, 1995] proposed a modulation classifier that uses the wavelet transform for the identification problem. The application of the wavelet transform resulted in distinctive patterns for different types, which enabled simple processing for identification. Three classes of modulation types were investigated: FSK, PSK, M-ary PSK and M-ary FSK. The relevant statistics for the identification schemes were derived and simulations show that in most cases there is less than 8% error at around 15dB carrier-to-noise ratio (CNR) with 100 symbols.

Hong and Ho extended the classifier in [Ho, 1995] to include QAM signals [Hong, 1999]. The identifier consisted of two branches and a decision block. It computed the  $|HWT|$  of an input signal with and without amplitude normalisation. It then used median filters to remove the peaks in the  $|HWT|$ 's, calculated the variances of the median filter outputs, and made the decision of the input modulation type by comparing the variances from the two branches with thresholds. The relevant statistics for optimum threshold selection were derived. Simulations show that the percentage of correct identification was higher than 97% with 50 observations when the CNR was not lower than 5dB.

#### **2.3.3.8. *Modulation Classification Using Power Moment Matrices***

A new approach for modulation classification was proposed by Hero and Hadinejad-Mahram [Hero, 1998]. The method was based on a pattern recognition technique previously applied to word spotting problems in binary images. In this approach, a large number of spatial moments are arranged in a symmetric positive definite matrix for which eigendecomposition and noise subspace processing methods can be applied. The resultant denoised moment matrix has entries which are used in place of the raw moments for improved pattern classification. The authors generalised the moment matrix technique to

grey scale images and applied the technique to discrimination between M-ary PSK and QAM constellations in signal space. Invariance to unknown phase angle and signal amplitude was achieved by representing the in-phase and quadrature components of the signal in the complex plane and computing joint moments of normalised magnitude and phase components.

#### **2.3.3.9. *Classification of Spread Spectrum Signals***

A modulation classifier based on the modulation domain measurement technique was proposed by Schneider and Chu [Schneider, 1991]. The implementation of this technique allows modulation analysis even with spread spectrum signals such as frequency hopping or direct sequence. Modulation analysis includes phase, frequency, time and amplitude of BPSK, QPSK, 8PSK, 16QAM, communication type signals with hop and pulse, and Barker or chirp radar type signals. The signals were generated with a psuedo random sequence and eye patterns were formed by the use of frequency trigger. By using software demodulation, a coherent local oscillator is not required but the phase result will be relative. Curve fitting algorithms were demonstrated with the mentioned modulation schemes. With 500 MHz bandwidth, the amplitude noise floor was reported to be  $-73$  dBm and the frequency/phase/timing sensitivity was  $-60$  dBm. It was stated that this technique is applicable at any carrier frequency where down conversion to the required spread spectrum bandwidth can be implemented.

#### **2.3.4. *Classification using Neural Networks***

Two methods of classification using neural networks (NN) will be described. The first method uses a hierarchical structure to achieve classification and the second classifier employs a backpropagation NN to recognise a variety of signals.

##### **2.3.4.1. *Hierarchical Neural Network***

Louis and Sehier [Louis, 1994] introduce a methodology for building neural networks for modulation classification based on a hierarchical approach and a priori knowledge to speed



up the learning phase. Superiority over a single, large, fully connected network was demonstrated. This approach reduces the complexity of the system in order to improve generalisation. Reduced sensitivity to initial conditions allows automation of the learning phase and simulation results showed the superiority of the hierarchical approach. The modulation types that can be classified are PSK2, PSK4, PSK8, FSK2, FSK4, FSK8, QAM16, QAM64, OQPSK, and MSK. The hierarchical NN classifier was compared with the conventional backpropagation learning, the k-Nearest Neighbour (k-NN) classifier and the binary decision tree. Classification success rates were as high as 90% with a SNR ranging from 0dB to 50dB.

#### **2.3.4.2. Classification of Spectral Features**

Ghani and Lamontagne have used a backpropagation neural network for modulation classification [Ghani, 1993]. The modulation types that can be recognised are: AM, FM, QPSK, USB, LSB, FSK1, FSK2, BPSK, and CW. A variety of spectral pre-processors were investigated for feature extraction. For the given training and test sets, the Welch periodogram was found to give the best results. Simulation results showed that the neural network algorithm can match or even outdo the performance of conventional k-Nearest Neighbour (k-NN) classifiers. The overall classification success rate was greater than 97%. Furthermore, the optimisation of selected neural networks was demonstrated using the optimal brain damage (OBD) pruning technique.

## **2.4. Conclusions**

This chapter has covered the various modulation classification techniques found in recent literature. Most of these classification techniques are restricted to a few modulation types. The motivation for this thesis is to develop a classification algorithm that encompasses a range of digital modulation types. The classification method chosen is based on deviations of instantaneous properties (similar to Nandi and Azzouz's work) using the decision-theoretic and neural network approaches. The reason for this choice is the ability for the decision tree to be expanded to accommodate larger numbers of digital modulation schemes. The neural network is also another feasible method for classification and can be easily implemented once the key features are identified. Therefore this thesis presents

modulation classification algorithms based on the decision-theoretic approach and neural networks respectively for a comprehensive list of digital modulation schemes.



# CHAPTER 3

## Decision Theory

### 3.1. Introduction

In this chapter the theory behind the classification process will be described. The purpose of classification is to determine to which category or class a given sample or signal belongs. An observation vector consists of a set of numbers that can be obtained through a measurement process. The observation vector is the input to a decision rule where a sample is assigned to one of the given classes. We assume that the observation vector is a random vector whose conditional density function depends on its class. In the case of modulation classification, the observation vector consists of samples of particular key features that have been extracted from the intercepted signal. If the conditional density function for each class is known, then the classification problem becomes a problem in statistical hypothesis testing.

The organization of this chapter is as follows. First, a description of classification decision theory is presented. This theory includes a description of Bayes error and the Bayes decision rule for minimum error. Threshold determination is then discussed followed by a discussion on classifier accuracy, confidence intervals, and statistical significance. Finally, some conclusions are presented.

### 3.2. Classification Decision Theory

Classification of signals involves three main processes, which are shown in Figure 3.1. These processes are [Azzouz and Nandi, 1996]:

- Pre-processing - which involves extracting key features from the intercepted signal as well as signal isolation and segmentation.
- Training and learning phase - A “training set” of data is used to adjust the classifier structure for optimum performance.

- Test phase – A “test set” of data is used to decide about the modulation type of a particular signal.

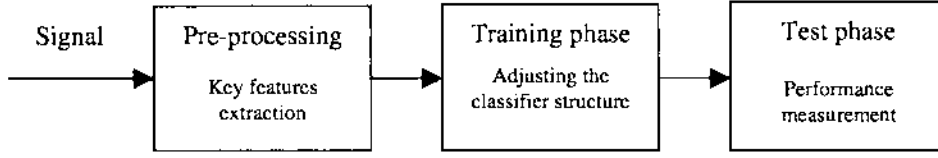


Figure 3.1. Functional blocks of signal classification.

Assuming that the pre-processing phase is completed (ie the key features are extracted), the next step is to adjust the classifier structure with training data. One of the functions of the training phase is to determine the best classification hypothesis, given the observed training data,  $X$  (or key features). In other words, we want the most probable hypothesis (which modulation type the signal is most likely to be classified as), given the data and the *a priori* probabilities. The Bayes decision rule for minimum error is used to determine the most probable hypothesis and is outlined in the next section. A two-class problem is discussed regarding the decision rule, which arises because each sample (or signal) belongs to one of two classes  $\omega_1$  or  $\omega_2$ . The conditional density functions and the *a priori* probabilities are assumed to be known.

### 3.2.1. The Bayes Decision Rule for Minimum Error

If  $X$  is an observation vector, the purpose is to determine whether the intercepted signal belongs to  $\omega_1$  or  $\omega_2$ . A decision rule based on posterior probabilities may be written as follows [Fukunaga, 1990]:

$$q_1(X) \begin{matrix} > \\ < \end{matrix} q_2(X) \quad \begin{matrix} \omega_1 \\ \omega_2 \end{matrix} \quad (3.1)$$

where  $q_i(X)$  is the *a posteriori* probability of  $\omega_i$  given  $X$ . Equation (3.1) indicates that if the probability of  $\omega_1$  given  $X$  is larger than the probability of  $\omega_2$ ,  $X$  is classified as  $\omega_1$ , and vice versa. The *a posteriori* probability  $q_i(X)$  can be calculated from the *a priori* probability  $P_i$  and the conditional density function  $P(X/\omega_i)$ , using Bayes theorem, as

$$q_i(X) = P(\omega_i / X) = \frac{P(X / \omega_i) P_i}{p(X)} \quad (3.2)$$

where  $p(X)$  is the mixture density function. Since  $p(X)$  is positive and common to both sides of the inequality, the decision rule of (3.1) can be expressed as

$$P(X / \omega_1) P_1 \underset{\omega_2}{\overset{\omega_1}{>}} P(X / \omega_2) P_2 \quad (3.3)$$

or

$$\ell(X) = \frac{P(X / \omega_1) \underset{\omega_2}{\overset{\omega_1}{>}} P_2}{P(X / \omega_2) \underset{\omega_2}{\overset{\omega_1}{<}} P_1} = \tau \quad (3.4)$$

The term  $\ell(X)$  is called the likelihood ratio and  $P(X/\omega_1)/P(X/\omega_2)$  is the threshold value of the likelihood ratio for the decision. For the classification of signals in this thesis, it is assumed that the *a priori* probabilities are equal for all intercepted signals. Therefore Equation (3.4) can be written as follows

$$P(X / \omega_1) \underset{\omega_2}{\overset{\omega_1}{>}} P(X / \omega_2) \quad (3.5)$$

Equation (3.3), (3.4), or (3.5) is called the Bayes test for minimum error.

### 3.2.2. Bayes Error

In general, any decision rule does not lead to perfect classification. To evaluate the performance of a decision rule, the probability of error (the probability that a sample is assigned to the wrong class) must be calculated. The conditional error given  $X$ , is denoted  $r(X)$ . It is found by the decision rule of (3.1) as either  $q_1(X)$  or  $q_2(X)$ , whichever is smaller. That is

$$r(X) = \min[q_1(X), q_2(X)] \quad (3.6)$$

The total error, which is called the Bayes error, is calculated by  $E\{r(X)\}$ .

$$\begin{aligned} \varepsilon &= E\{r(X)\} = \int r(X) p(X) dX \\ &= \int \min[P_1 P(X / \omega_1), P_2 P(X / \omega_2)] dX \\ &= P_1 \int_{L_2} P(X / \omega_1) dX + P_2 \int_{L_1} P(X / \omega_2) dX \\ &= P_1 \varepsilon_1 + P_2 \varepsilon_2 \end{aligned} \quad (3.7)$$

where

$$\varepsilon_1 = \int_{L_2} p_1(X) dX \quad \text{and} \quad \varepsilon_2 = \int_{L_1} p_2(X) dX \quad (3.8)$$

The integral regions  $L_1$  and  $L_2$  are the regions where  $X$  is classified as  $\omega_1$  and  $\omega_2$ . In  $L_1$ ,  $P_1P(X/\omega_1) > P_2P(X/\omega_2)$  and therefore  $r(X) = P_2P(X/\omega_2)/p(X)$ . Similarly for  $L_2$ ,  $r(X) = P_1P(X/\omega_1)/p(X)$  because  $P_2P(X/\omega_2) > P_1P(X/\omega_1)$ . In (3.8), two types of errors are defined: one results from misclassifying samples from  $\omega_1$  and the other results from misclassifying samples from  $\omega_2$ . The total error is the weighted sum of these two errors.

An example of this decision rule for a simple one-dimensional case is shown in Figure 3.2. In the diagram,  $p_1(X)$  represents  $P(X/\omega_1)$  and  $p_2(X)$  represents  $P(X/\omega_2)$ . The decision boundary is set at  $x = t$  where  $P_1p_1(X) = P_2p_2(X)$ , and  $x < t$  and  $x > t$  are assigned to  $L_1$  and  $L_2$  respectively. The resulting errors are  $P_1\varepsilon_1 = B + C$ ,  $P_2\varepsilon_2 = A$ , and  $\varepsilon = A + B + C$ , where  $A$ ,  $B$ , and  $C$  indicate the areas. For example,

$$B = \int_t^{t'} P_1 p_1(x) dx \quad (3.9)$$

This decision rule gives the smallest probability of error and this can be shown by referring to Figure 3.2. If the boundary is moved from  $t$  to  $t'$ , the new  $\omega_1$  and  $\omega_2$  regions are  $L'_1$  and  $L'_2$  respectively. The resulting errors are  $P_1\varepsilon'_1 = C$ ,  $P_2\varepsilon'_2 = A + B + D$ , and  $\varepsilon' = A + B + C + D$  which is larger than  $\varepsilon$  by  $D$ . The choice of the threshold,  $t$ , for the decision rule is very important to ensure the minimum probability of error. Therefore threshold determination is discussed in the next section.

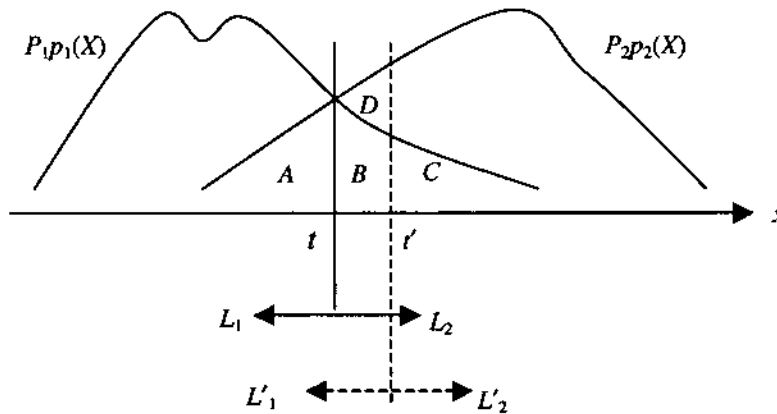


Figure 3.2. Bayes decision rule for minimum error [Fukunaga, 1990].

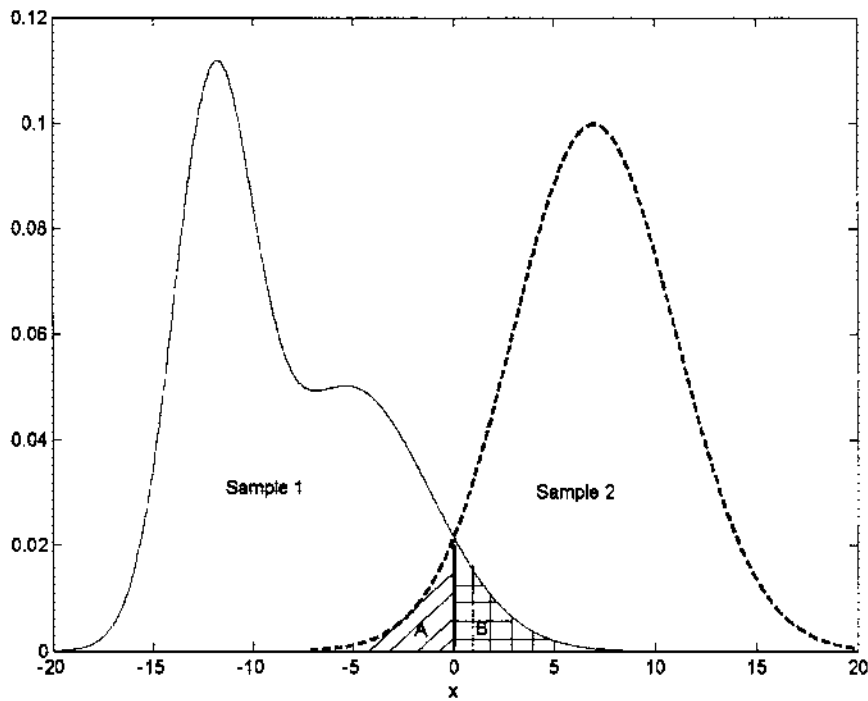


Figure 3.3. Example of Bayes decision rule for minimum error.

An example to illustrate the Bayes decision rule is shown in Figure 3.3. The density functions are normal with one function ( $p_1(X)$ ) having a mean ( $m$ ) of  $-7$  and standard deviation ( $s$ ) of  $6$  indicated by the solid line. The other function ( $p_2(X)$ ) has mean  $m = 5.85$  and standard deviation  $s = 5$  which is shown by the dashed line in Figure 3.3. Both functions intersect at  $x = 0$  making this value a likely threshold. The error probabilities can be found by integrating the functions for regions A and B respectively. The probability of error ( $\epsilon_1$ ) for  $p_1(X)$  is indicated by region B and is calculated to be  $0.11$  ( $10.68\%$ ). The probability of error ( $\epsilon_2$ ) for  $p_2(X)$  is calculated to be  $0.12$  ( $12.23\%$ ) and is represented by region A.

### 3.2.3. Threshold Determination

The threshold value,  $t$ , can be determined in three ways: One method is to find the threshold that gives the minimum probability of error as shown in section 3.2.2. The second method is to use the Bayes decision rule for minimum cost. This method is used when the misclassification of different samples have different consequences, i.e. the cost of misclassifying samples is different. The third method is to estimate the threshold from the density estimates of the posterior probabilities of the sample data.



This method is used when the true posterior probabilities are not known. These methods will now be described.

### 3.2.3.1. The Bayes Decision Rule for Minimum Cost

Minimising the error probability is often not the best criterion to design a decision rule because the misclassifications of  $\omega_1$  and  $\omega_2$  samples may have different consequences. An example of this is the misclassification of a cancer patient to a normal patient. This decision may be more detrimental than if the normal patient was misclassified as a cancer patient. Therefore it is appropriate to assign a cost to each situation as follows: Let

$$c_{ij} = \text{cost of deciding } X \in \omega_i \text{ when } X \in \omega_j \quad (3.10)$$

Then the conditional cost of deciding  $X \in \omega_i$  given  $X$ ,  $r_i(X)$ , is

$$r_i(X) = c_{i1}q_1(X) + c_{i2}q_2(X) \quad (3.11)$$

The decision rule and the resulting conditional cost given  $X$ ,  $r(X)$ , are

$$r_1(X) \begin{matrix} < \\ > \end{matrix} \begin{matrix} \omega_1 \\ \omega_2 \end{matrix} r_2(X) \quad (3.12)$$

and

$$r(X) = \min[r_1(X), r_2(X)] \quad (3.13)$$

The total cost of this decision is

$$\begin{aligned} r &= E\{r(X)\} = \int \min[r_1(X), r_2(X)] dX \\ &= \int \min[c_{11}q_1(X) + c_{12}q_2(X), c_{21}q_1(X) + c_{22}q_2(X)] p(X) dX \\ &= \int \min[c_{11}P_1p_1(X) + c_{12}P_2p_2(X), c_{21}P_1p_1(X) + c_{22}P_2p_2(X)] dX \quad (3.14) \\ &= \int_{L_1} [c_{11}P_1p_1(X) + c_{12}P_2p_2(X)] dX + \int_{L_2} [c_{21}P_1p_1(X) + c_{22}P_2p_2(X)] \end{aligned}$$

where  $L_1$  and  $L_2$  are determined by the decision rule in (3.12). The boundary that minimises  $r$  in (3.14) can be found by rewriting (3.14) as a function of  $L_1$  only. This

can be done by replacing  $\int_{L_2} p_i(X) dX$  with  $1 - \int_{L_1} p_i(X) dX$  since  $L_1$  and  $L_2$  do not overlap and cover the entire domain. Thus,

$$r = (c_{21}P_1 + c_{22}P_2) + \int_{L_1} [(c_{11} - c_{21})P_1p_1(X) + (c_{12} - c_{22})P_2p_2(X)] dX \quad (3.15)$$

We must choose  $L_1$  such that  $r$  is minimised. Suppose, for a given value of  $X$ , that the integrand of (3.15) is negative. The value  $r$  can be decreased by assigning  $X$  to  $L_1$ . If the integrand is positive,  $r$  can be decreased by assigning  $X$  to  $L_2$ . Thus the minimum cost decision rule is to assign to  $L$  those  $X$ 's and only those  $X$ 's, for which the integrand of (3.15) is negative. The decision rule is called the Bayes test for minimum cost and can be described by the following equation:

$$(c_{12} - c_{22})P_2p_2(X) \underset{\omega_2}{\overset{\omega_1}{>}} (c_{21} - c_{11})P_1p_1(X) \quad (3.16)$$

or

$$\frac{p_1(X)}{p_2(X)} \underset{\omega_2}{\overset{\omega_1}{<}} \frac{(c_{12} - c_{22})P_2}{(c_{21} - c_{11})P_1} = \tau \quad (3.17)$$

By comparing (3.17) with (3.4), it can be seen that Bayes test for minimum cost is a likelihood ratio test with a different threshold from (3.4), and that the selection of the cost functions is equivalent to changing the *a priori* probabilities:  $P_1$  and  $P_2$ . Equation (3.17) is equal to (3.4) for the special case when the cost functions are equal as shown by

$$c_{21} - c_{11} = c_{12} - c_{22} \quad (3.18)$$

This is called the symmetrical cost function where the cost becomes the probability of error. In the case of modulation classification, this condition holds true because the cost of misclassification is the same for all signals.

### 3.2.3.2. Posterior Probability Estimation

Most of the decision theorems assume that the density functions are known. However, it is common in practice to be unsure of the density functions and therefore it is necessary to estimate the functions using an unstructured approach. This approach is

called nonparametric estimation where the density function is estimated locally by a small number of neighbouring samples. This results in a less reliable estimate with a larger bias and variance than the parametric methods.

There are two common nonparametric estimation methods: one is called the Parzen density estimate and the other is the k-nearest neighbour density estimate; the two techniques are very similar. Once the estimated density functions have been derived, the technique outlined in section 3.2.2 can be used to determine the threshold. Accurate density estimation is very hard to achieve. However, the goal is to design a classifier and evaluate its performance – not to accurately estimate the density itself. For further information on these methods, the reader may refer to [Fukunaga, 1990].

Another method is to estimate the posterior probability directly from the sample data using models [Ripley, 1996]. Figure 3.4 shows an example of the direct modelling of the posterior probabilities. The two classes are digitally modulated signals where one class belongs to FSK4 signals and the other class belongs to FSK8 signals. The probabilities have been estimated from simulated data and the value  $X$  is a particular feature that has been extracted from each signal for classification. We assume that the *a priori* probability of each signal is equal.

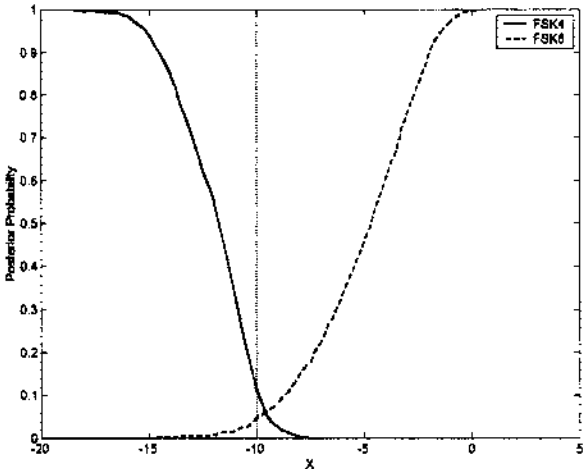


Figure 3.4. Example of posterior probabilities for two classes of digitally modulated signals (FSK4 and FSK8).

The threshold can be determined by choosing a value in the centre of where the two posterior probabilities cross into the error regions. Note that the centre is taken from where the error regions reach a saturation point. In this case, a suitable estimated threshold is  $-10$  and is shown by the dashed line in Figure 3.4.

This method can be illustrated again with our previous example shown in Figure 3.3. It can be seen from the figure that the best threshold value to choose is  $0$  because this is where the two posterior density functions cross into the error regions.

Once the threshold has been determined, the classifier can be evaluated in terms of accuracy and confidence intervals.

#### 3.2.4. Classifier Accuracy

The *kappa* ( $\kappa$ ) coefficient is usually used to evaluate the classifier accuracy compared to chance classification. It is a measure of the difference between the actual agreement between reference data and the classifier and the chance agreement between the reference data and a random classifier [Bouzerdoun, 2001]. The  $\kappa$  coefficient is calculated as

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (3.19)$$

where  $P_o$  is the observed accuracy and  $P_e$  is the chance agreement. When the sample sizes are equal, the chance classification derivation is simply dividing  $1$  by the number of groups. When the groups are unequal, the proportional chance criterion is used and is defined as

$$C_{PRO} = p^2 + (1 - p)^2 \quad (3.20)$$

where  $p$  is the proportion of individuals in the first group and  $1-p$  is the proportion of individuals in the second group. This criterion is obviously biased towards the group with the largest proportion of samples. Hair et al [Hair, 1996] suggest that the classification accuracy be at least one-fourth greater than that achieved by the chance accuracy. For a chance agreement it is expected that  $\kappa = 0$ , whereas for a true

agreement,  $\kappa = 1$ . The next step is to examine the probable error in this accuracy estimate.

#### **3.2.4.1. Sample Error and True Error**

To find the probable error, it is necessary to distinguish carefully between two notions of error. One is the error rate of the hypothesis over the sample data that is available and the other is the error rate of the hypothesis over the entire unknown distribution  $D$  of examples. These are called the sample error and true error respectively [Mitchell, 1997].

The sample error of a hypothesis with respect to some sample  $S$  of occurrences drawn from  $X$ , is the fraction of  $S$  that it misclassifies. The true error of a hypothesis is the probability that it will misclassify a single randomly drawn instance from the distribution  $D$ . In the case of the modulation classification technique used in this thesis, the sample error can only be calculated for the data that we have on hand. The ultimate aim is to find the true error because this is the error that we can expect to apply to future samples. Therefore we need to know how good an estimate of the true error is provided by the sample error. The answer to this is provided in the next section.

#### **3.2.5. Confidence Intervals for Discrete – Valued Hypotheses**

Suppose, we wish to estimate the true error for some discrete valued hypothesis  $H$ , based on its observed sample error over a sample  $S$  where [Mitchell, 1997]:

- the sample  $S$  contains  $n$  examples drawn independent of one another, and independent of  $H$ , according to the probability distribution  $D$ .
- $n \geq 30$
- hypothesis  $H$  has  $r$  errors over these  $n$  examples (ie., the error probability:  $error_S(H) = r/n$ ).

Under these conditions, it is possible to make the following statements due to statistical theory:

1. Given no other information, the most probable value of the true error  $error_D(H)$  is the statistical error  $error_S(H)$ .
2. With approximately 95% probability, the true error  $error_D(H)$  lies in the interval

$$\left[ error_s - 1.96\sqrt{\frac{error_s(100 - error_s)}{n}}, error_s + 1.96\sqrt{\frac{error_s(100 - error_s)}{n}} \right] \quad (3.21)$$

This interval is called the 95% confidence interval estimate for the true error. This expression is an approximation of the confidence interval and works well when

$$n \cdot error_s(H)(1 - error_s(H)) \geq 5 \quad (3.22)$$

The 95% confidence interval for the example shown in Figure 3.3 can be calculated as follows: For sample 1, the probability of error  $\varepsilon_1$  is 10.68%, by substituting this value into equation (3.21) the corresponding confidence interval is [8.34,13.01]. Similarly for sample 2, the error  $\varepsilon_2$  is 12.23% and the 95% confidence interval is [9.55,14.91].

Other factors to consider when we only have a sample distribution and not the whole distribution are statistical significance and statistical power, which will be discussed in the following section.

### 3.2.6. Statistical Significance Versus Statistical Power

Since it is rare to obtain the entire population of occurrences, we are forced to draw statistical inferences from a randomly drawn sample from that population. Interpreting statistical inferences requires that the acceptable levels of error be specified. The most common approach is to specify the level of Type I error, also known as the alpha ( $\alpha$ ) level. The Type I error is the probability of rejecting the null hypothesis when actually true. In other words, it is the chance of the test showing statistical significance when it is actually not present (“false positive”). By specifying an alpha level, the allowable limits for error are set because we are specifying the probability of concluding that significance exists when it actually does not.

An associated error known as the Type II error or beta ( $\beta$ ) is also determined when setting the Type I error. The beta is the probability of failing to reject the null hypothesis when it is actually false. Another probability that arises is called the power of the statistical inference test and is defined as  $1 - \beta$ . Power is the probability of correctly rejecting the null hypothesis when it should be rejected. In other words, it is the probability that statistical significance will be indicated if present. The relationship

of the different error probabilities for the hypothetical setting of testing for the difference in two means is shown below [Hair, 1996]:

Table 3.1. The relationship of the different error probabilities in the hypothetical setting of testing for the difference in two means.

Statistical Decision	Reality	
	H <sub>0</sub> : No difference in two means	H <sub>A</sub> : Difference in two means
	H <sub>0</sub> : No difference	H <sub>A</sub> : Difference
	1 - $\alpha$	$\beta$ Type II error
	$\alpha$ Type I error	1 - $\beta$ Power

H<sub>0</sub> is the null hypothesis and H<sub>A</sub> is the alternative hypothesis. It can be seen that specifying alpha establishes the level of statistical significance. In other words, it is the level of power that controls the probability of success in finding the differences if they exist. There is a trade-off in trying to reduce the different error types. Reducing the Type I error also reduces the power of the statistical test. Thus there must be a balance between the level of alpha and the resulting power. Power is also determined by three main factors:

1. Effect size – The probability of achieving statistical significance is also based on the actual magnitude of the effect of interest (eg, a difference of means between two groups or the correlation between variables) in the population, called effect size. It is expected that a larger effect is more likely to be found and thus affect the power of the test.
2. Alpha – It has already been stated that as alpha increases, the power also decreases. Thus as the chance of finding an incorrect significant effect reduces, the probability of correctly finding an effect also decreases.
3. Sample Size – At any given alpha level, increased sample size always produces greater power of the test. But there is a danger that increasing the sample size will produce too much power. This means that by increasing the sample size, smaller and smaller effects will be found to be statistically significant, until at very large sample sizes, almost any effect is significant. Therefore it is best to be aware that the sample size can affect the statistical test by either making it

too sensitive with very large sample sizes or not sensitive enough (at small sample sizes).

Referring to our example in Figure 3.1, the different error probabilities for classifying sample 1 and sample 2 are shown in Table 3.2.

Table 3.2. The relationship of the different error probabilities in the hypothetical setting of testing for classification of sample 1 and sample 2.

		Reality	
		H <sub>0</sub> : Sample 1	H <sub>A</sub> : Sample 2
Statistical Decision	H <sub>0</sub> : Sample 1	87.77%	10.68%
	H <sub>A</sub> : Sample 2	12.23% Type I error	89.32% Type II error Power

The level of significance for the classification accuracy can be tested using a *t* test. The formula for a two-group analysis of equal size is [Hair, 1996]

$$t = \frac{p - 0.5}{\sqrt{\frac{0.5 * (1.0 - 0.5)}{N}}} \tag{3.23}$$

where *p* is the proportion correctly classified and *N* is the sample size. The formula can be adjusted for use with more groups and unequal sample sizes.

The level of significance, *t* for sample 1 in our example in Figure 3.3 is calculated as 22.24 for *p* = 0.89 and *N* = 800. Similarly, for sample 2, the calculated *t* statistic is 21.37 with *p* = 0.88 and *N* = 800. The optimum *t* value is 28.28 when all values in the sample are correctly classified. Therefore it can be concluded that the level of significance for both samples is acceptable.

### 3.3. Conclusions

This chapter has outlined the underlying nature, concepts and approach to classification. The methodological concepts were clarified by presenting the basic guidelines for its application and interpretation. An example was presented and this outlined the major points needed to be familiar in applying Baye’s classification. The next chapter will



outline the theory behind neural networks, which are used as another method for classification. The proceeding chapters will demonstrate the theory outlined in this chapter and the next chapter, applied to digital modulation classification. The signals to be classified will be identified as well as key features extracted from the intercepted signal. A thorough analysis and interpretation of the various classification functions derived will also be presented.

# CHAPTER 4

## Classification Using Feedforward Artificial Neural Networks (ANNs)

### 4.1. Introduction

This chapter presents a brief introduction of classification theory based on artificial neural networks. *Artificial Neural Networks* (ANNs), or *Neural Networks* (NNs) for short, are another tool that will be used for classification of digital modulation schemes. ANNs use the pattern recognition approach to modulation classification. This approach is different to the *decision-theoretic* (DT) approach, where instead of a suitable threshold being chosen for each decision, the threshold at each neuron (node) is chosen automatically and adaptively. Also in the DT approach, each key feature is considered one at a time, whereas in the ANN algorithm, all key features are considered simultaneously. Therefore, it is implied that the ANN approach may perform better than the DT approach because the probability of a correct decision is not based on the time order of the key features.

The organization of this chapter is as follows. The general concepts of artificial neural networks, including the different classes of neural networks and their structures, learning paradigms and training algorithms are presented first. The next section presents a discussion on classification using neural networks, with an example, followed by some concluding remarks in the final section. This chapter serves as a building block to the digital modulation classifiers described in proceeding chapters that are based on neural networks.

### 4.2. Artificial Neural Networks

A neural network is a computational structure inspired by the study of biological neural processing [Rao and Rao, 1995]. The processing power in biological neural structures has brought about the study of these structures to help organise human made computing

structures. ANNs are a means to organise synthetic neurons to solve the same kind of difficult problems in the same way that the human brain may.

ANNs resemble the brain in two respects:

- Knowledge is acquired by the network through a learning process (learning algorithm).
- Interneuron connection strengths known as synaptic weights are used to store the “knowledge” [Haykin, 1999].

The learning algorithm modifies the synaptic weights in a prescribed fashion, based on the learning information presented, so as to achieve a particular objective.

Neural networks have better performance over conventional technologies in areas which include data segmentation, data compression, robust pattern detection, adaptive control, optimisation and scheduling, database mining, and complex mapping. Neural networks are advantageous because they offer specific processing advantages, such as nonlinear processing, adaptive learning, self-organisation, ability to handle contextual information and fault tolerance via redundant information coding. They also offer real time operation, they are universal information processors, have a neurobiological analogy, and can be implemented in VLSI.

Some applications of neural networks besides signal classification are:

- Financial prediction
- Control of nuclear power plants
- Coronary heart disease risk assessment
- Face recognition, etc.

#### **4.2.1. The Artificial Neuron Model**

The most common artificial neuron model is shown in Figure 4.1. It has 3 basic elements:

- Synapses or connecting links - Each synapse has associated with it a weight or strength,  $w$ . The signal  $x_j$  at input of synapse  $j$  connected to neuron  $k$  is multiplied by the synaptic weight  $w_{kj}$ .

- **Adder** -The adder is a linear combiner for summing the weighted input signals,  $w_{kj}x_j$ , and its output  $v_k$  is given by

$$v_k = \sum_{j=1}^p w_{kj}x_j \quad (4.1)$$

- **Activation function** - The activation function,  $\phi(v)$ , is the relationship between adder output and the final neuron output. It is often a non-linear function, thereby limiting the amplitude of the neuron output. The nonlinearity also helps in feature extraction. Normally, a constant threshold or bias value ( $\theta$ ) is also added, resulting in the following equation:

$$y_k = \phi(v_k - \theta_k) \quad (4.2)$$

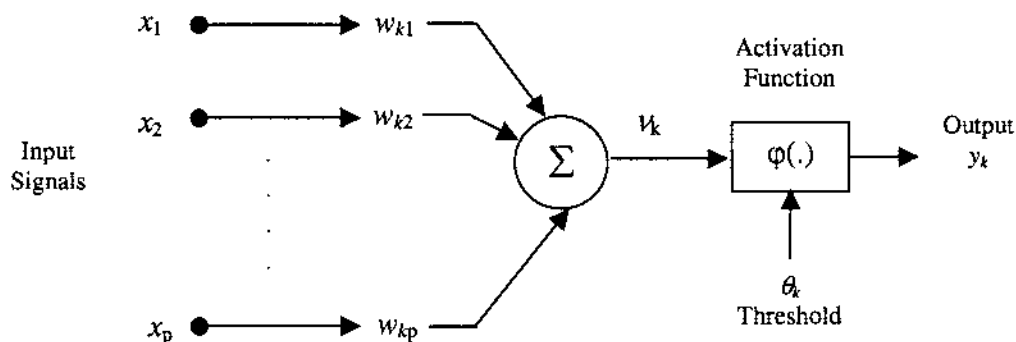


Figure 4.1. Neuron model.

#### 4.2.2. Activation Function Types

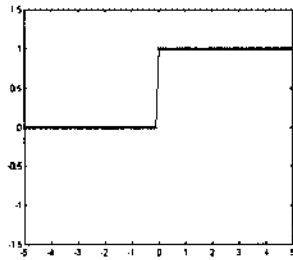
There are many commonly used activation functions. Some examples are the threshold function, linear function, piecewise linear function, and *sigmoid* functions such as the *logistic* function (logsig) and *hyperbolic tangent* (tansig) function. These functions are shown in Figure 4.2. The sigmoid and linear functions are the most popular because they are continuously differentiable, a very important criteria for most of the training algorithms.

The logistic function is described as

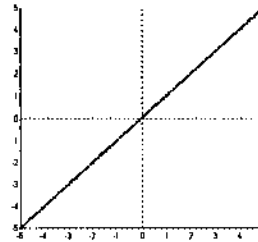
$$\varphi(v) = \frac{1}{1 + e^{-av}} \quad (4.3)$$

and the hyperbolic tangent is represented by

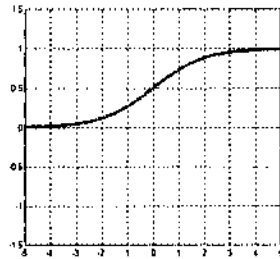
$$\varphi(v) = \tanh\left(\frac{v}{2}\right) = \frac{1 - e^{-v}}{1 + e^{-v}} \quad (4.4)$$



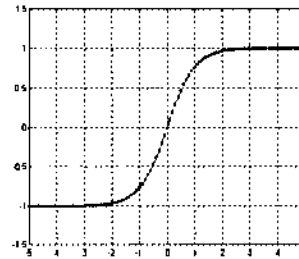
a) Threshold



b) Linear function



c) Logistic function



d) Hyperbolic tangent

Figure 4.2. Types of activation functions.

#### 4.2.3. ANN Architectures

Usually a number of neurons are connected together to form a neural network. A distinct structure of neurons in a network is called a neural network architecture. The neural network architecture is closely linked to the learning algorithm used to train the network. There are four general classes of network architectures [Haykin 1999]:

#### 1) Single-layer Feedforward Networks

These networks have only feedforward connections and have only a single layer of computing nodes (doesn't include input layer).

#### 2) Multilayer Feedforward Networks

The multilayer networks have one or more hidden layers of computing nodes. These layers can be fully or partially connected.

#### 3) Recurrent Networks

Recurrent networks have at least one feedback loop. They may be with or without hidden neurons, and normally have delay elements in the feedback loops.

#### 4) Lattice Structures

Lattice structures consist of a one, two or higher dimensional array of neurons. A set of source nodes feeds the lattice.

In this research only layered feed-forward network, called *multilayer perceptrons* (MLP), are considered for classification of digital modulation signals. An MLP consists of subgroups or layers of processing elements; each layer makes independent calculations and passes the resultant output to another layer which in turn makes calculations and passes the result to another layer and so on. The final output of the network is determined by a subgroup of one or more processing elements, called the output neurons. Each processing element makes its computations based on a weighted sum of its inputs.

The first layer is called the input layer, the last layer is called the output layer and the layers in between are called the hidden layers. The processing elements are referred to as artificial neurons because they are seen to be similar to neurons in the human brain. Figure 4.3 is a typical layered feed-forward network comprising three layers: input, output and one hidden layer. The neurons are represented with circular nodes.

The input consists of a vector  $\mathbf{x}$  whose input elements enter the network through the weight matrix  $\mathbf{W}$ . The weighted values at the synapses of a neuron are fed to a summing junction whose sum is  $\langle \mathbf{W}, \mathbf{x} \rangle$ , i.e., the dot product of the weight vector and the input vector. The

hidden layer neurons have a bias  $\theta$ , which is summed with the weighted inputs to form the net input  $v$ . The output of each unit,  $y$ , is found by feeding the net input  $v$  as an argument to the activation function  $\phi$ . The output is given by:

$$y = \phi(\mathbf{w} \times \mathbf{x} + \theta) \quad (4.5)$$

The transfer function as well as the weighted sum of inputs from the neural network determines the internal activation or raw output of a neuron.

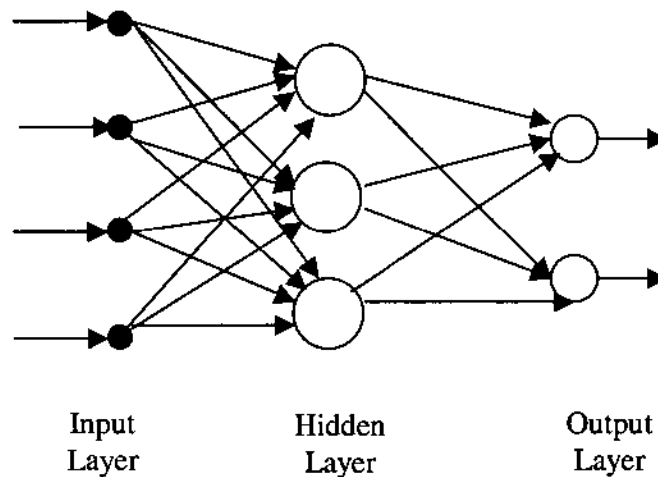


Figure 4.3. A Layered feed-forward neural network.

#### 4.2.4. Learning Process

The weights used on the connections between different layers have much significance. If the network is run with one set of weights, the network is said to have had no learning. If we start with one set of weights, run the network, modify some or all the weights and then run the network again with the new set of weights, the process is called training the network and the network is said to have learned. The learning process for neural networks can be outlined as follows:

1. The environment stimulates the neural network.
2. The neural network undergoes changes as a result of stimulation.
3. Then, the neural network responds in a new way to the environment because of changes to its internal structure.

The changes made to the NN are in terms of changes to the synaptic weights in the form:

$$w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n) \quad (4.6)$$

The calculation of  $\Delta w_{kj}$  is obtained from the learning algorithm to be used, which is a set of rules for the solution of the problem. In addition to learning algorithms, the learning process can be subdivided into learning paradigms: supervised learning, unsupervised learning and reinforcement learning. The learning paradigm refers to the manner in which the NN (learning machine) relates to its environment. For example, in supervised learning, the network interacts with a teacher by receiving a feedback signal indicating the desired outputs; whereas, in unsupervised learning the network only receives inputs from the environment with no indication as to what should be the desired outputs. Figure 4.4 shows a taxonomy of the learning process (adapted from [Haykin, 1999]). For a detailed description of learning paradigms and learning algorithms, the reader is referred to one of many neural network textbooks [Haykin, 1999]

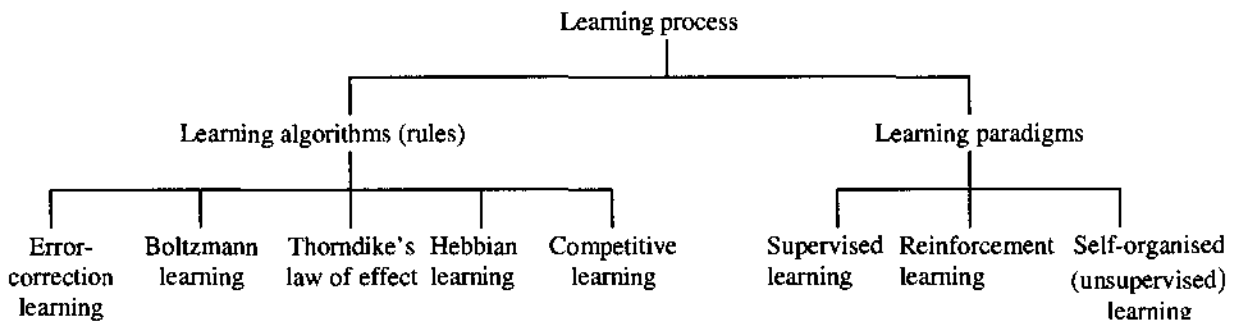


Figure 4.4. Taxonomy of the learning process.

### 4.3. Classification Using Neural Networks

Neural networks have emerged as an important tool for classification. There are many advantages of using NNs for classification:

- NNs can adjust themselves to the data without any explicit specification of functional or distributional form for the underlying model [Zhang, 2000].
- NNs can approximate any function with arbitrary accuracy. Since any classification procedure looks for a functional relationship between the group



membership and the attributes, or key features, of the object, it is important to accurately identify this underlying function.

- NNs are non-linear models, making them flexible in modelling real world complex relationships.
- Finally, NNs are able to estimate the posterior probabilities  $P(\omega_i/X)$ . Chapter 3 discussed how the posterior probabilities provide the basis for establishing the classification rule and performing statistical analysis. (In this chapter we will represent  $X$  as  $\mathbf{x}$  – the input vector).

For classification and regression, the operation of a NN can be interpreted as a mapping  $F : R^d \rightarrow R^M$ , where a  $d$ -dimensional input  $\mathbf{x}$ , is submitted to the network and an  $M$ -vectored network output  $y$  is obtained to make the classification decision. The network is typically built so that the mean squared error (MSE) is minimised. From least squares estimation theory [Papoulis, 1965], the mapping function  $F : \mathbf{x} \rightarrow y$  which minimises the expected squared error

$$E[y - F(\mathbf{x})]^2 \quad (4.7)$$

is the conditional expectation of  $y$  given  $\mathbf{x}$ .

$$F(\mathbf{x}) = E[y | \mathbf{x}] \quad (4.8)$$

With regards to classification, the desired output  $y$  is a vector of binary values and is the  $j$ th basis vector  $e_j = (0, \dots, 0, 1, 0, \dots, 0)^T$  if  $\mathbf{x} \in$  group  $j$ . Hence the  $j$ th element of  $F(\mathbf{x})$  is given by

$$\begin{aligned} F_j(\mathbf{x}) &= E[y_j | \mathbf{x}] \\ &= 1 \cdot P(y_j = 1 | \mathbf{x}) + 0 \cdot P(y_j = 0 | \mathbf{x}) \\ &= P(y_j = 1 | \mathbf{x}) \\ &= P(\omega_j | \mathbf{x}) \end{aligned} \quad (4.9)$$

In other words, the least squares estimate for the mapping function in a classification problem is exactly the posterior probability.

Neural networks are universal approximators [Cybenko, 1989] and can approximate any function arbitrarily closely (in theory). However, the mapping function represented by a network is not perfect due to the local minima problem, finite training data when training

the neural network, and suboptimal network structure. Therefore the posterior probabilities provided by neural networks are estimates of the true posteriors.

The link between neural networks and statistical pattern classifiers is the estimation of the posterior probabilities. However, it is not possible to make a direct comparison since NNs are generally non-linear while statistical methods are basically linear. If we appropriately code the desired output of the membership values, we may let neural networks directly model some discriminant functions. For example, in a two-group classification problem, if the desired output is coded as 1 if the sample is from class 1 or  $-1$  if the sample is from class 2, then from (4.9), the neural network estimates the following discriminant function:

$$g(x) = P(\omega_1 | \mathbf{x}) - P(\omega_2 | \mathbf{x}) \quad (4.10)$$

The classification rule is to assign  $\mathbf{x}$  to class  $\omega_1$  if  $g(\mathbf{x}) > 0$  or  $\omega_2$  if  $g(\mathbf{x}) < 0$ .

#### **4.3.1. Learning and Generalisation**

As we have described earlier, learning is the ability to approximate the underlying behaviour adaptively from the training data and generalisation is the ability to predict well beyond the training data [Zhang, 2000]. Overfitting occurs when the neural network fits the training sample very well but has poor generalisation capability for predicting future samples. Powerful data fitting or function approximation capability of the neural network further contributes to overfitting. Underfitting occurs when the network does not fit the training sample enough and therefore future samples cannot be predicted accurately. Overfitting and underfitting can be analysed through the bias-plus-variance decomposition of the prediction error.

#### **4.3.2. Bias and Variance Composition of the Prediction Error**

A thorough analysis of the relationship between learning and generalisation in neural networks based on the concepts of model bias and model variance can be found in [Geman, 1992]. A data-driven model may be too dependent on the specific data and have a large variance, on the other hand, a model which is less dependent on the data may represent the true functional relationship and have a large bias. Bias and variance are often incompatible and if one is reduced, it will cause the other to increase. Therefore a trade-off is necessary

in building a useful NN classifier. For example, if we consider a two-group classification problem in which the binary output variable  $y \in \{0,1\}$  is related to a set of input variables (the feature vector)  $\mathbf{x}$  by

$$y = F(\mathbf{x}) + \varepsilon \quad (4.11)$$

where  $F(\mathbf{x})$  is the target or underlying function and  $\varepsilon$  is assumed to be a zero-mean random variable. From (4.8) and (4.9), the target function is the conditional expectation of  $y$  given  $\mathbf{x}$ , that is

$$F(\mathbf{x}) = E(y | \mathbf{x}) = P(\omega_1 | \mathbf{x}) \quad (4.12)$$

If we have a training set  $T$  of size  $N$ , we need to find an estimate,  $f(\mathbf{x};T)$ , of  $F(\mathbf{x})$  so that the overall estimation error can be minimised. The most commonly used performance measure is the mean square error (MSE) defined as

$$\begin{aligned} \text{MSE} &= E[(y - f(\mathbf{x};T))^2] \\ &= E[(y - F(\mathbf{x}))^2] + (f(\mathbf{x};T) - F(\mathbf{x}))^2 \end{aligned} \quad (4.13)$$

Notice that the MSE depends on the particular data set  $T$  which means that any change in the data set and/or sample size may result in a change in the estimation function and hence the estimation error. Since the training data is random, the overall prediction error of the model can be written as

$$E_T \{E[(y - f(\mathbf{x};T))^2]\} = E[(y - F(\mathbf{x}))^2] + E_T [(f(\mathbf{x};T) - F(\mathbf{x}))^2] \quad (4.14)$$

Where  $E_T$  denotes the expectation over all possible random samples of size  $N$ .

Further information on bias and variance as well as methods for reducing the prediction error can be found in [Zhang, 2000]. An example of how NNs are used for classification is shown in the next section.

### 4.3.3. Example of Classification Using Neural Networks

Suppose we have two classes of overlapping two-dimensional normally distributed samples, labelled class 1 and class 2. Let  $\omega_1$  and  $\omega_2$  denote the set of events for which a random vector  $\mathbf{x}$  belongs to class 1 and class 2, respectively. The conditional probability for class 1 can be expressed as

$$P(\mathbf{x} | \omega_1) = \frac{1}{2\pi\sigma_1^2} \exp\left(-\frac{1}{2\sigma_1^2} \|\mathbf{x} - \boldsymbol{\mu}_1\|^2\right) \quad (4.15)$$

where  $\boldsymbol{\mu}_1$  is the mean vector ( $\boldsymbol{\mu}_1 = [2, 0]^T$ ) and  $\sigma_1^2$  is the variance ( $\sigma_1^2 = 4$ ).

The conditional probability for class 2 is

$$P(\mathbf{x} | \omega_2) = \frac{1}{2\pi\sigma_2^2} \exp\left(-\frac{1}{2\sigma_2^2} \|\mathbf{x} - \boldsymbol{\mu}_2\|^2\right) \quad (4.16)$$

where  $\boldsymbol{\mu}_2$  is the mean vector ( $\boldsymbol{\mu}_2 = [0, 0]^T$ ) and  $\sigma_2^2$  is the variance ( $\sigma_2^2 = 1$ ).

Furthermore, both classes are assumed to have equal prior probabilities,  $P_1 = P_2 = 0.5$ .

The probability density functions for class 1 and class 2 are shown in Figure 4.5 and Figure 4.6 respectively. The scatter plot of classes 1 and 2 is shown in Figure 4.7. Class  $\omega_1$  is represented by the 'o' symbol and class  $\omega_2$  is represented by the '+' symbol. The decision boundary is shown and its derivation will now be discussed.

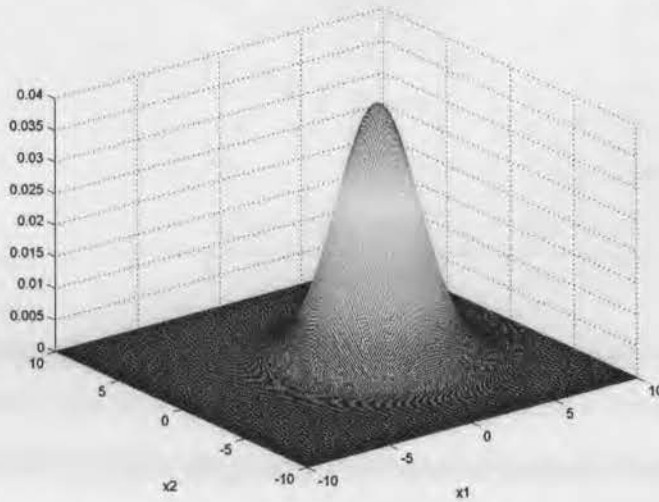


Figure 4.5. Probability density function for Class 1.

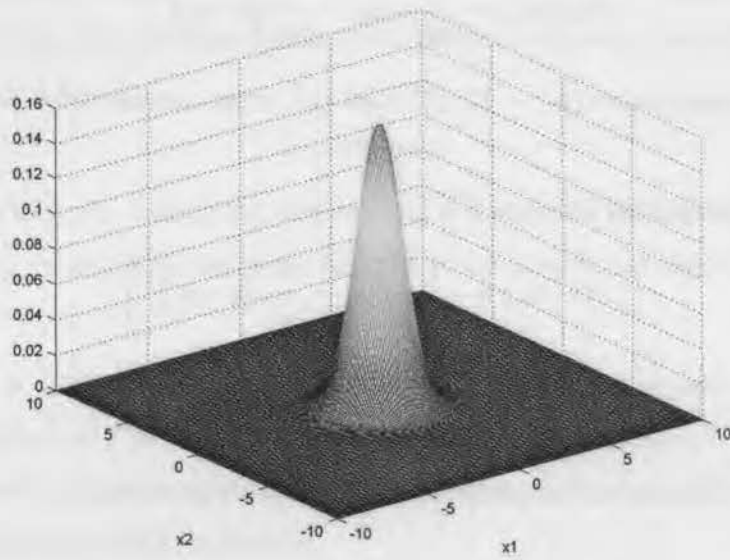


Figure 4.6. Probability density function for Class 2.

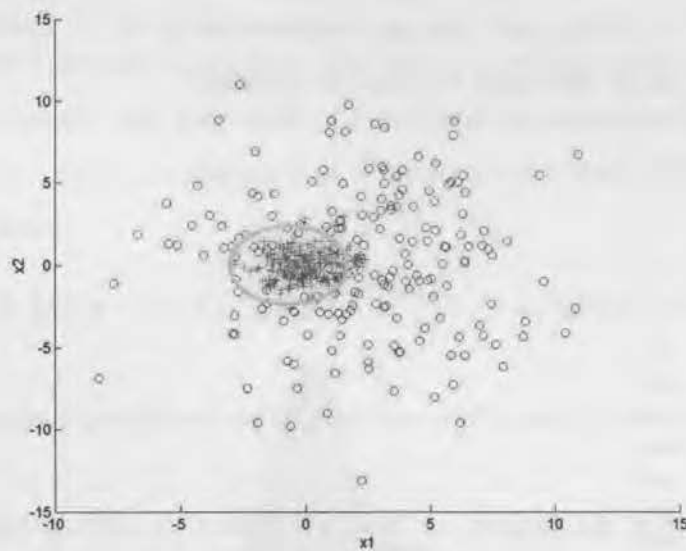


Figure 4.7. Scatter plot of classes  $\omega_1$  and  $\omega_2$  Showing decision boundary.

The optimum decision boundary is found by applying the likelihood ratio test as described in Chapter 3:

$$\ell(\mathbf{x}) \underset{\omega_2}{\overset{\omega_1}{>}} = \tau \quad (4.17)$$

Recall equation (3.4), where the likelihood ratio is defined as

$$\ell(\mathbf{x}) = \frac{P(\mathbf{x}/\omega_1)}{P(\mathbf{x}/\omega_2)} \quad (4.18)$$

The threshold is defined as

$$\tau = \frac{P_2}{P_1} = 1 \quad (4.19)$$

Therefore for our example, the optimum decision boundary is defined as

$$\ell(\mathbf{x}) = \frac{\sigma_2^2}{\sigma_1^2} \exp\left(-\frac{1}{2\sigma_1^2}\|\mathbf{x} - \boldsymbol{\mu}_1\|^2 + \frac{1}{2\sigma_2^2}\|\mathbf{x} - \boldsymbol{\mu}_2\|^2\right) = 1 \quad (4.20)$$

or equivalently

$$\frac{1}{\sigma_2^2}\|\mathbf{x} - \boldsymbol{\mu}_2\|^2 - \frac{1}{\sigma_1^2}\|\mathbf{x} - \boldsymbol{\mu}_1\|^2 = 4\ln\left(\frac{\sigma_1}{\sigma_2}\right) \quad (4.21)$$

Using straightforward manipulations, the decision boundary defined by (4.21) can be redefined as

$$\|\mathbf{x} - \mathbf{x}_c\|^2 = r^2 \quad (4.22)$$

$$\mathbf{x}_c = \frac{\sigma_1^2 \boldsymbol{\mu}_2 - \sigma_2^2 \boldsymbol{\mu}_1}{\sigma_1^2 - \sigma_2^2} \quad (4.23)$$

and

$$r^2 = \frac{\sigma_2^2 \sigma_1^2}{\sigma_1^2 - \sigma_2^2} \left[ \frac{\|\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1\|^2}{\sigma_1^2 - \sigma_2^2} + 4\ln\left(\frac{\sigma_1}{\sigma_2}\right) \right] \quad (4.24)$$

Equation (4.24) represents a circle with centre  $\mathbf{x}_c$  and radius  $r$ . Let  $\Omega_2$  define the region lying inside the circle. The classification rule may then be stated as follows:

*Classify the observation vector  $\mathbf{x}$  as belonging to class  $\omega_2$  if  $\mathbf{x} \in \Omega_2$  and to class  $\omega_1$  otherwise.*

For our example, we have a circular decision boundary whose centre is located at  $\mathbf{x}_c = [-2/3, 0]^T$  and has radius  $r \approx 2.34$ . This decision boundary is shown in Figure 4.7.

The probability of error for class 1 is  $\varepsilon_1 \approx 0.2642$ . Similarly, the probability of error for class 2 is  $\varepsilon_2 \approx 0.1056$ . The total error assuming both classes have equal priors is:  
 $\varepsilon = 0.5\varepsilon_1 + 0.5\varepsilon_2 \approx 0.1849$ . Therefore the probability of correct classification,  $P_c \approx 0.8151$ .

We will now compare these values derived theoretically with the performance of a neural network trained with simulated data of the same distributions outlined in equations (4.15) and (4.16). We generate 500 samples from each class and simulate a feed-forward MLP NN with two input neurons (because our data is two dimensional) and two output neurons corresponding to the two classes. The network has one hidden layer with 2 neurons. The probability of error for class 1 is found to be 0.2660 and the probability of class 2 is found to be 0.0082. The total probability of error is then calculated to be 0.1740. These figures are comparable to the theoretical error rates calculated previously showing that the simulated neural network performs according to the theory.

#### **4.4. Conclusions**

This chapter has covered the general theory and concepts of NNs. The neuron model and NN structures have been described and the learning process and associated algorithms have also been presented. Statistical decision theory applied to neural networks has been presented with an example of classification using neural networks. The following chapters will describe different digital modulation classifiers. These classifiers are based either on the decision-theoretic approach or neural networks. The contents of this chapter will be applied in the latter case.

# CHAPTER 5

## Modulation Classification of ASK, FSK, and PSK Signals

### 5.1 Introduction

In this chapter, modulation classification techniques utilising the decision-theoretic (DT) and neural network (NN) approaches are used to classify ASK2, ASK4, PSK2, PSK4, FSK2, and FSK4 signals. These signals have already been treated in [Azzouz and Nandi, 1996]. However, an alternative algorithm with a different decision tree is proposed here and new key features are introduced. This is the first of three chapters addressing classification of digitally modulated signals. Each chapter builds on the classifier of the preceding chapter by treating a different set of digital modulated signals, culminating in a full classifier for all digitally modulated signals. The classification is achieved through a DT approach or a NN algorithm. Before the classifier systems are discussed, the next section introduces the analytic signal representation of digitally modulated ASK, FSK and PSK signals. In Section 5.3, the key features of these signals are introduced. The two classification methods, namely the decision-theoretic and the neural network approach for recognising the different modulation types, are described in Sections 5.4 and 5.5. Section 5.6 presents and compares the performances of the different classifiers. Finally, Section 5.7 presents discussion and concluding remarks.

### 5.2 Analytic Signal Representation

The digital processing of broadband signals requires a high sampling rate. This means that the processing speed and the memory size must be increased. All the processing of the received data vector must be completed before the arrival of the next data segment. The bandwidth of the signal in practice will be minimal to keep the sampling rate low. If the signal  $x(t)$  is real, then from hermitian symmetry  $X(f) = X^*(-f)$ , where  $X(f)$  is the Fourier



transform. This means that the whole information content of the signal can be found in one half of the signal spectrum. Thus, any real signal can be represented by its right half spectrum, called the *analytic representation*. The digital processing of the analytic signal requires half the sampling rate that is needed for the broad band real signal but the same amount of memory is needed because the derived samples are complex.

### 5.2.1 Hilbert Transform

The spectral redundancy can be removed using the Hilbert transform, which gives the analytic representation of the signal. By applying the signal,  $x(t)$ , to a quadrature filter,  $F_Q$ , with impulse response  $r_Q(t)$  and complex gain  $G_Q(f)$  we get the Hilbert transform,  $y(t)$ . This can be written as

$$y(t) = F_Q\{x(t)\} = x(t) * r_Q(t) = \int_{-\infty}^{\infty} x(t - \theta) r_Q(\theta) d\theta \quad (5.1)$$

where

$$r_Q(t) = \frac{1}{\pi} \quad (5.2)$$

By substituting (5.2) into (5.1), the Hilbert transform can be expressed as:

$$y(t) = P.V. \int_{-\infty}^{\infty} \frac{x(t - \theta)}{\pi\theta} d\theta \quad (5.3)$$

where *P.V.* is the principle value of the integral. The complex gain of the quadrature filter is

$$G_Q(f) = \frac{Y(f)}{X(f)} = -j \operatorname{sgn}(f) \quad (5.4)$$

where *sgn* is the *signum* function. The analytic signal,  $z(t)$ , is the representation of the right half spectrum of a real signal  $x(t)$ . We can obtain  $z(t)$ , by applying  $x(t)$  to an analytizing filter,  $F_A$ . This filter is made up of an identity filter  $F_I$  and a quadrature filter  $F_Q$ . This can be shown as:

$$F_A = F_I + jF_Q \quad (5.5)$$

The impulse response of  $F_I$  is  $r_I(t) = \delta(t)$  and the complex gain is  $G_I(f) = 1$ . For  $F_Q$ , the impulse response and complex gain are given by (5.2) and (5.4), respectively. Hence, the analytic signal,  $z(t)$ , can be expressed as

$$z(t) = F_A\{x(t)\} = [F_I + jF_Q]\{x(t)\} = x(t) + jy(t) \quad (5.6)$$

It can be seen that the analytic signal,  $z(t)$ , is a complex function where its real part is the real signal  $x(t)$  and the imaginary part  $y(t)$  is the Hilbert transform of the signal  $x(t)$ . The spectrum of the analytic signal,  $Z(f)$  is given by

$$Z(f) = X(f) + jY(f) = [1 + \text{sgn}(f)]X(f) = 2U(f)X(f) \quad (5.7)$$

where  $U(f)$  is the unit step function in the frequency domain and is defined by

$$U(f) = \begin{cases} 1 & \text{if } f > 0 \\ \frac{1}{2} & \text{if } f = 0 \\ 0 & \text{otherwise} \end{cases} \quad (5.8)$$

### 5.2.2 Complex Envelope

The complex envelope,  $\alpha(t)$ , of a real signal,  $x(t)$ , can be derived from the analytic representation as follows

$$\alpha(t) = z(t)e^{-j2\pi f_c t} \quad (5.9)$$

where  $f_c$  is some arbitrary frequency. In the case of a narrowband signal,  $f_c$  is taken as the carrier frequency. From equations (5.6) and (5.9), the complex envelope can be expressed as

$$\alpha(t) = m(t) + jn(t) \quad (5.10)$$

where

$$m(t) = x(t) \cos(2\pi f_c t) + y(t) \sin(2\pi f_c t) \quad (5.11)$$

and

$$n(t) = y(t) \cos(2\pi f_c t) - x(t) \sin(2\pi f_c t) \quad (5.12)$$

$x(t)$  can be reconstructed from  $m(t)$  and  $n(t)$  using the following analytic form

$$x(t) = m(t) \cos(2\pi f_c t) - n(t) \sin(2\pi f_c t) \quad (5.13)$$

The instantaneous amplitude and instantaneous phase of a signal can be found from either the complex envelope representation in (5.10) or the analytic signal in (5.6). The instantaneous amplitude,  $a(t)$ , is defined as

$$a(t) = |z(t)| = \sqrt{x^2(t) + y^2(t)} = |\alpha(t)| = \sqrt{m^2(t) + n^2(t)} \quad (5.14)$$

The instantaneous phase,  $\phi(t)$ , can be calculated from the analytic expression as

$$\phi(t) = \begin{cases} \tan^{-1}[y(t)/x(t)] & \text{if } x(t) > 0, y(t) > 0 \\ \pi - \tan^{-1}[y(t)/x(t)] & \text{if } x(t) < 0, y(t) > 0 \\ \pi/2 & \text{if } x(t) = 0, y(t) = 0 \\ \pi + \tan^{-1}[y(t)/x(t)] & \text{if } x(t) < 0, y(t) < 0 \\ 3\pi/2 & \text{if } x(t) = 0, y(t) < 0 \\ 2\pi - \tan^{-1}[y(t)/x(t)] & \text{if } x(t) > 0, y(t) < 0 \end{cases} \quad (5.15)$$

$\phi(t)$  can be calculated from the complex envelope, except the linear phase component, due to the carrier frequency (i.e.,  $\phi(t) = \arg\{z(t)\} = \arg\{\alpha(t)\} + 2\pi f_c t$ ), is not present in the complex envelope representation because of the down-conversion.

The instantaneous frequency  $f(t)$  follows as

$$f(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt} \quad (5.16)$$

### 5.2.3 Representations of Digital Modulation Schemes

This section explains the digital modulation schemes considered for classification in this chapter, namely amplitude shift keying (ASK), phase shift keying (PSK) and frequency shift keying (FSK). The signals will be outlined with a graphical representation of the relevant features.

#### 5.2.3.1 Amplitude Shift Keying (ASK)

The ASK signal is represented as [Couch, 2001]:

$$s(t) = A_c m(t) \cos \omega_c t \quad (5.17)$$

where  $m(t)$  is a unipolar baseband data signal and  $A_c$  is a constant representing the power level. The complex envelope is given by:

$$\alpha(t) = A_c m(t) \quad (5.18)$$

For the binary case:

$$m(t) = \begin{cases} 1 & 0 \leq t \leq T_b, \text{ or} \\ 0 & 0 \leq t \leq T_b \end{cases} \quad (5.19)$$

where  $T_b$  is the bit duration ( $=1/R_b$ ),  $R_b$  is the bit rate. The instantaneous amplitude can be expressed as:

$$a(t) = |m(t)| = \begin{cases} 0 & \text{if } m(t) = 0 \\ 1 & \text{if } m(t) = 1 \end{cases} \quad (5.20)$$

and the instantaneous phase is

$$\phi(t) = 0 \quad (5.21)$$

It can be seen from Figure 5.1 and Figure 5.2 that the instantaneous amplitude looks like the bit stream, while the instantaneous phase and frequency are zero.

### 5.2.3.2 Phase Shift Keying (PSK)

The PSK signal is represented as [Couch, 2001]:

$$s(t) = A_c \cos[\omega_c t + D_p m(t)] \quad (5.22)$$

where  $m(t)$  is a bipolar baseband signal having peak values of  $\pm 1$  and a rectangular pulse shape (for convenience) and  $D_p$  is a the modulation index of the PSK signal.

The complex envelope is given by:

$$\alpha(t) = A_c e^{j\theta(t)} = x(t) + jy(t) \quad (5.23)$$

where the values of  $x$  and  $y$  are:

$$\begin{aligned} x_i &= A_c \cos \theta_i \\ y_i &= A_c \sin \theta_i \end{aligned} \quad (5.24)$$

for the permitted phase angles  $\theta_i$ ,  $i = 1, 2, \dots, M$  of the PSK signal. For PSK2,  $M = 2$ , for PSK4,  $M = 4$  and for PSK8,  $M = 8$ . For the binary case ( $M = 2$ ), we let  $D_p = \pi/2$  to give the maximum power in the signal [Couch, 2001] and the complex envelope becomes:

$$\alpha(t) = jm(t) \quad (5.25)$$

The instantaneous amplitude is

$$a(t) = |m(t)| = 1 \quad (5.26)$$

and the instantaneous phase is

$$\phi(t) = \begin{cases} -\pi/2 & \text{if } m(t) = -1 \\ \pi/2 & \text{if } m(t) = 1 \end{cases} \quad (5.27)$$

Therefore, the instantaneous frequency is zero. These attributes of PSK modulation are shown in Figure 5.3 and Figure 5.4.

### 5.2.3.3 Frequency Shift Keying (FSK)

The FSK signal is represented by

$$s(t) = A_c \cos \left[ \omega_c t + D_f \int_{-\infty}^t m(\lambda) d\lambda \right] \quad (5.28)$$

$$s(t) = \text{Re} \{ g(t) e^{j\omega_c t} \} \quad (5.29)$$

where

$$\alpha(t) = A_c e^{j\theta(t)} \quad (5.30)$$

and

$$\theta(t) = D_f \int_{-\infty}^t m(\lambda) d\lambda \quad \text{for FSK} \quad (5.31)$$

where  $m(t)$  is a baseband digital signal. Although  $m(t)$  is discontinuous at the switching time, the phase function  $\theta(t)$  is continuous because  $\theta(t)$  is proportional to the integral of  $m(t)$ . The instantaneous amplitude and phase are given by

$$a(t) = A_c \quad (5.32)$$

$$\phi(t) = D_f \int_{-\infty}^t m(\lambda) d\lambda \quad (5.33)$$

The instantaneous frequency is given by

$$f(t) = \frac{1}{2\pi} \frac{d\phi}{dt} = \frac{1}{2\pi} D_f m(t) \quad (5.34)$$

The attributes of FSK modulation are shown in Figure 5.5 and Figure 5.6.

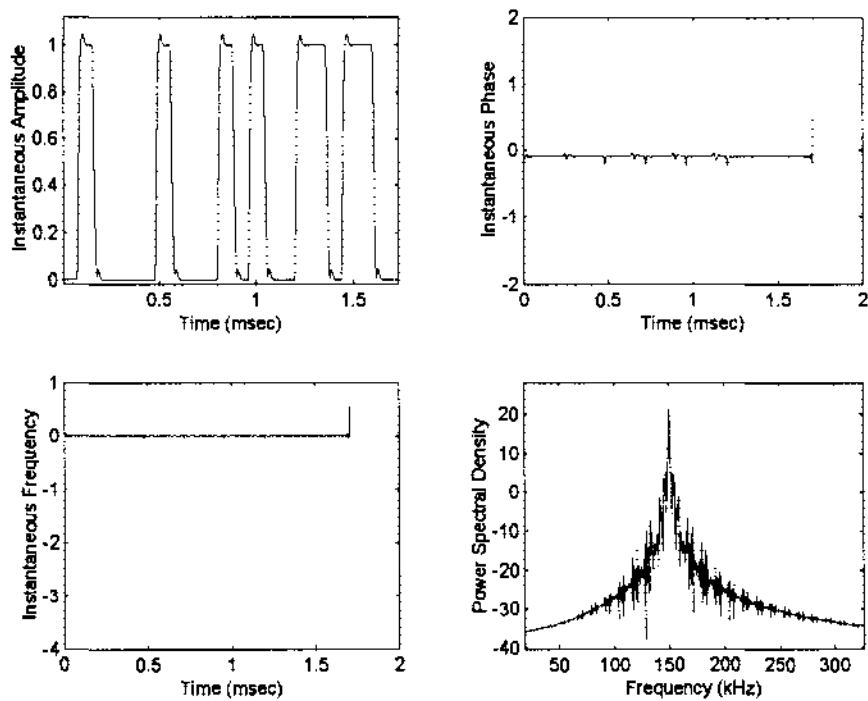


Figure 5.1. Useful features of ASK2 modulation, carrier frequency  $F_c = 150\text{kHz}$ .

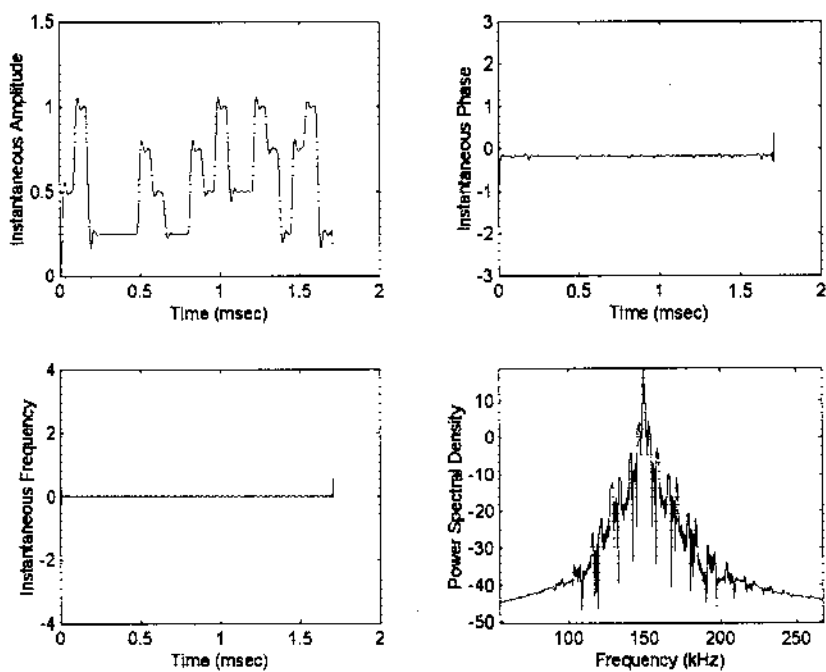


Figure 5.2. Useful features of ASK4 modulation, carrier frequency  $F_c = 150\text{kHz}$ .

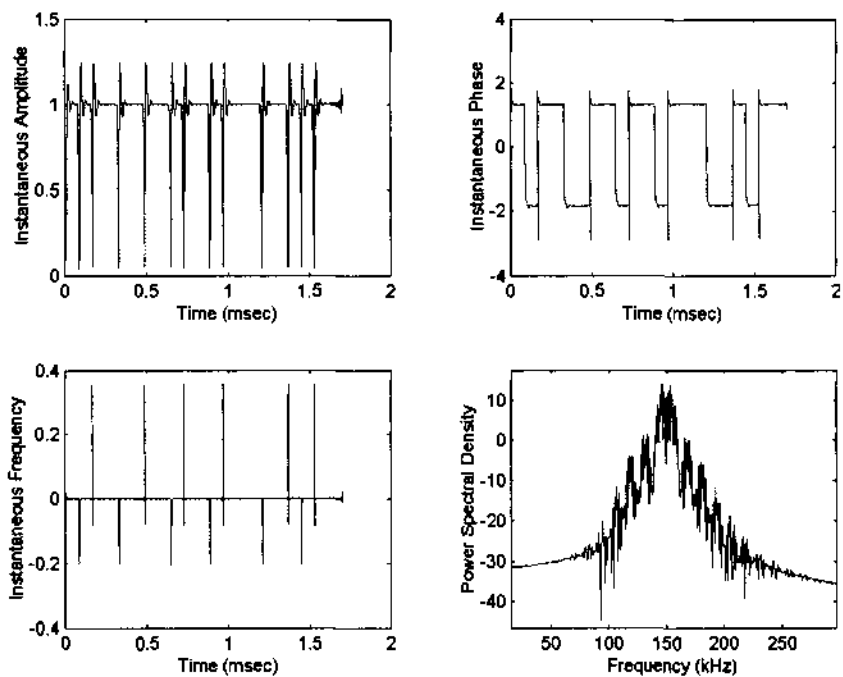


Figure 5.3. Useful features of PSK2 modulation.

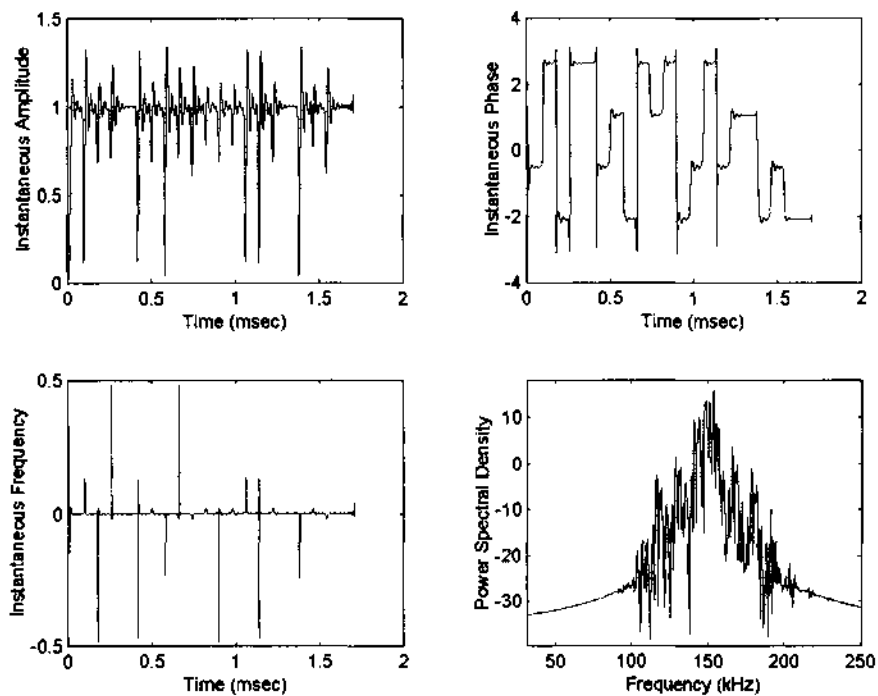


Figure 5.4. Useful features of PSK4 modulation.

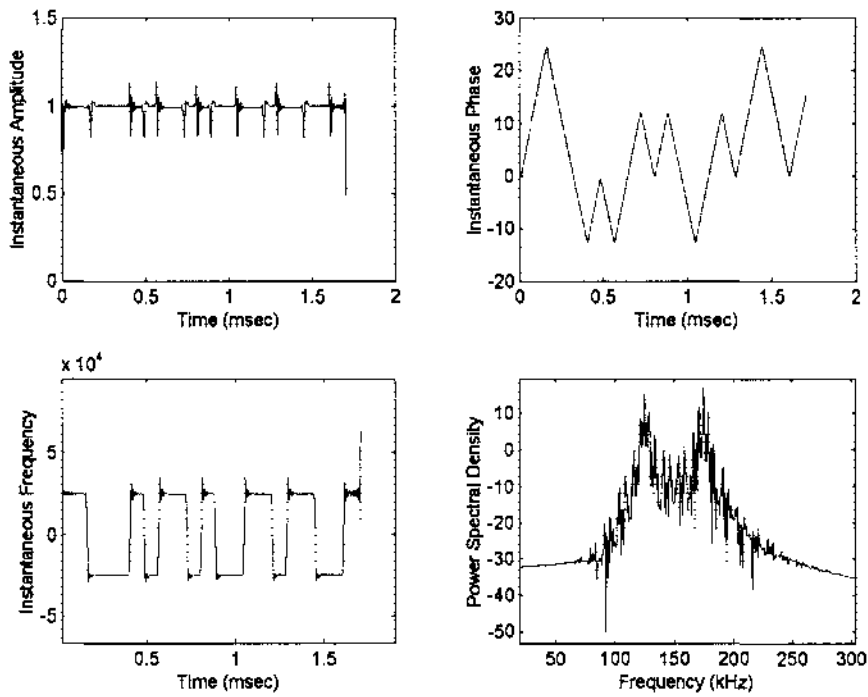


Figure 5.5. Useful features of FSK2 modulation.

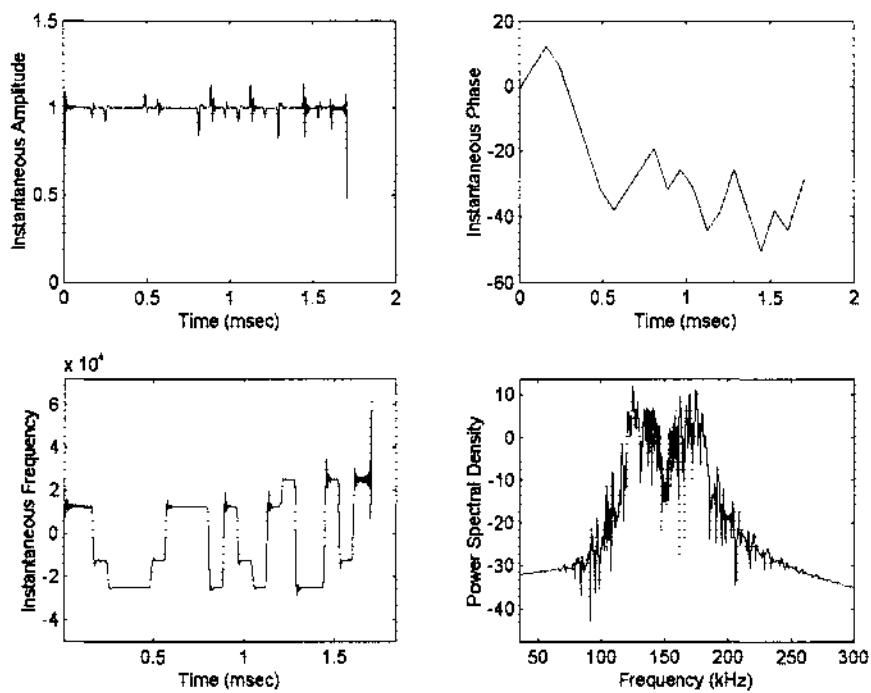


Figure 5.6. Useful features of FSK4 modulation.



### 5.3 Key Feature Extraction

The procedure for digital signal classification is based on the method outlined in [Azzouz and Nandi, 1996]. The intercepted signal with length  $K$  seconds, is sampled at a rate  $f_s$  and divided into  $M$  successive frames. Each frame is  $N_s$  samples long ( $N_s = 2048$ ), which is equivalent to 1.76ms. This results in  $M (=Kf_s/N)$  frames. A set of key features is extracted from each frame to decide the type of modulation. These key features are derived from the complex envelope of the signal  $y(t)$ , the instantaneous amplitude  $A(t)$ , the instantaneous phase  $\phi(t)$  and the instantaneous frequency  $f(t)$  of the intercepted signal. The key features from a particular segment (frame) are used to classify the segment as a certain modulation type.

Five key features are used in this modulation classification approach: two key features were discussed in [Swami and Sadler, 2000] and the other three are introduced here. The key features discussed by Swami and Sadler are based on higher-order cumulants of the signal and are described in section 5.3.1. The other three key features, introduced in this chapter, are discussed in section 5.3.2.

#### 5.3.1 Cumulant Key Features

The complex envelope of the intercepted signal is represented by  $y(n)$ . For a complex valued stationary random process, the second-order cumulants can be written in two ways depending on the placement of the conjugation operator

$$C_{20} = E[y^2(n)] \quad \text{and} \quad C_{21} = E[|y(n)|^2] \quad (5.35)$$

where  $E$  denotes the expectation operation. Similarly, the fourth-order cumulants can be defined in one of three ways:

$$\begin{aligned} C_{40} &= \text{cum}(y(n), y(n), y(n), y(n)) \\ C_{41} &= \text{cum}(y(n), y(n), y(n), y^*(n)) \\ C_{42} &= \text{cum}(y(n), y(n), y^*(n), y(n)) \end{aligned} \quad (5.36)$$

### 5.3.1.1 Sample Estimates

The cumulants in (5.35) and (5.36) can be approximated with the sample estimates of the corresponding moments [Swami and Sadler, 2000]. Assuming that  $y(n)$  is zero mean, the sample estimates of second-order cumulants are given by

$$\begin{aligned}\hat{C}_{21} &= \frac{1}{N} \sum_{n=1}^N |y(n)|^2 \\ \hat{C}_{20} &= \frac{1}{N} \sum_{n=1}^N y^2(n)\end{aligned}\tag{5.37}$$

The superscript ‘ $\wedge$ ’ denotes a sample average. The estimates of the fourth-order cumulants are

$$\begin{aligned}\hat{C}_{40} &= \frac{1}{N} \sum_{n=1}^N y^4(n) - 3\hat{C}_{20}^2 \\ \hat{C}_{41} &= \frac{1}{N} \sum_{n=1}^N y^3(n)y^*(n) - 3\hat{C}_{20}\hat{C}_{21} \\ \hat{C}_{42} &= \frac{1}{N} \sum_{n=1}^N |y(n)|^4 - |\hat{C}_{20}|^2 - 2\hat{C}_{21}^2\end{aligned}\tag{5.38}$$

The two cumulants features used in the modulation classifier are  $|\hat{C}_{21}|$  and  $|\hat{C}_{40}|$ .

### 5.3.2 Other Key Features

The other three new key features introduced in this chapter are the mean of the instantaneous phase,  $\mu_{dp}$ , the maximum value (measured in dB) of the power spectral density (PSD) of the normalised instantaneous frequency,  $\gamma_{maxf}$ , and the standard deviation of the normalised instantaneous frequency,  $\sigma_{fn}$ .

The key feature  $\mu_{dp}$  is defined as

$$\mu_{dp} = \frac{1}{C} \sum_{A_n(i) > a_t} \phi_{NL}(i)\tag{5.39}$$

where  $\phi_{NL}(i)$  is the value of the non-linear component of the instantaneous phase at time instants  $t = i/f_s$ ,  $C$  is the number of samples in the intercepted frame  $\{\{\phi(i)\}\}$  for which  $A_n(i) > a_t$ , and  $a_t$  is a threshold for  $A(t)$  below which the estimation of the instantaneous phase is very sensitive to noise.

The key feature  $\gamma_{maxf}$ , measured in dB's, is given by:

$$\gamma_{maxf} = 10 \log_{10} (\max |DFT(f_n)|^2) \quad (5.40)$$

where  $f_n$  is the normalised instantaneous frequency of the signal defined by  $f_n = f(t)/R_s$ .  $R_s$  is the symbol rate and  $f(t)$  is the instantaneous frequency.

The standard deviation of the normalised instantaneous frequency is evaluated over the non-weak segments of the received signal:

$$\sigma_{f_n} = \sqrt{\frac{1}{C} \left[ \sum_{A_n(i) > a_t} f_n^2(i) \right] - \left[ \frac{1}{C} \sum_{A_n(i) > a_t} f_n(i) \right]^2} \quad (5.41)$$

where  $f_n$  is the normalised instantaneous frequency,  $C$  is the number of samples in  $\{f_n(i)\}$  for which  $A_n(i) > a_t$ ,  $A_n(i) = A(i)/m_a$  where  $m_a$  is the average value of the instantaneous amplitude over one frame and  $a_t$  is a threshold for  $A_n(i)$  below which the estimation of the instantaneous phase is very sensitive to noise. It is found in [Azzouz and Nandi, 1996], that a suitable threshold is  $a_t = 1$ .

### 5.3.3 Explanation for Key Feature Selection

The key feature  $\gamma_{maxf}$  is used to discriminate between FSK2 and FSK4 as one group and ASK2, ASK4, PSK2 and PSK4 as the second group. Since ASK2, ASK4, PSK2 and PSK4 signals possess little or no frequency information, their power spectral density values measured in dB will be very small. On the other hand FSK2 and FSK4 signals possess some frequency information; therefore, their PSD values of the normalised instantaneous frequency will be larger.

The feature  $\mu_{dp}$  is used to distinguish between ASK2 and ASK4 signals. Both types of signals possess very little phase information. However, ASK2 signals possess slightly larger instantaneous phase values than ASK4 signals. Therefore, the mean of the instantaneous phase is a good feature to separate ASK2 and ASK4 signals. This can be inferred by inspecting Figure 5.7, which shows a close up of the instantaneous phase for ASK2 and ASK4 signals.

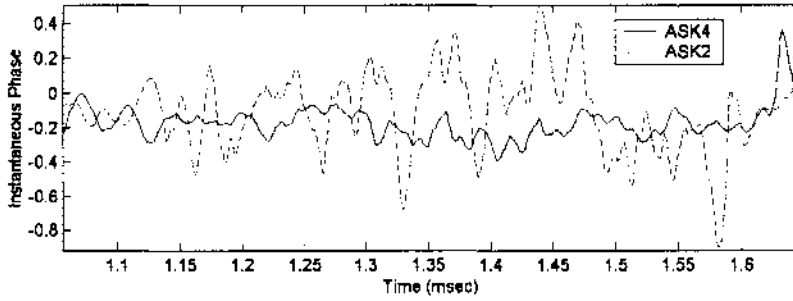


Figure 5.7. Instantaneous phase values for ASK2 and ASK4 signals.

The key feature  $|\hat{C}_{21}|$  is used to separate signals with phase information (PSK2, PSK4) from those with no phase information (ASK2, ASK4). It is found that the values of  $|\hat{C}_{21}|$  for PSK signals are greater than the values for ASK signals.

The key feature  $|\hat{C}_{40}|$  is used to separate PSK2 signals from PSK4 signals. By referring to Table I in [Swami and Sadler, 2000], the theoretical values of the fourth order cumulants are  $-2$  and  $1$  for PSK2 and PSK4 signals respectively. Therefore the absolute values should be around  $2$  for PSK2 signals and  $1$  for PSK4 signals.

The key feature  $\sigma_{fn}$  is used to separate FSK2 and FSK4 signals. In FSK2 signals, the symbols are represented by one of two frequency values situated at  $(f_c + 2R_s)$  and  $(f_c - 2R_s)$ . With FSK4 signals, there are an additional two frequency values that a symbol may be represented as situated at  $(f_c + R_s)$  and  $(f_c - R_s)$ . Since these values are smaller, the key feature values for FSK4 are also generally smaller than for FSK2.

## 5.4 Decision – Theoretic Modulation Classification Method

In the decision theoretic approach, a decision tree is constructed that has as its leaf nodes one of the desired modulation types; a flowchart depicting the final classification procedure is shown in Figure 5.8. The incoming signal segment is categorized as one of two possible sets of signals by comparing a key feature of the signal with a certain threshold. The threshold for each feature is chosen so that the number of correct decisions made is optimal. The determination of the thresholds is outlined next in Section 5.4.1.

### 5.4.1 Threshold Determination

The key feature thresholds are chosen so that the probability of a correct decision is obtained from 400 realisations of each modulation type at signal to noise ratio (SNR) ranging from 20 to 5dB. A set of modulation types is separated into two disjoint subsets, A and B, by the decision rule defined in equation (3.1) of Chapter 3. Note that here we use the alternate notation  $A = \omega_1$  and  $B = \omega_2$ , and assume that the priors  $P(A)$  and  $P(B)$  are equal,  $P(A) = P(B) = 0.5$ . The optimum threshold is chosen such that the Bayes error is minimised as described in Chapter 3.

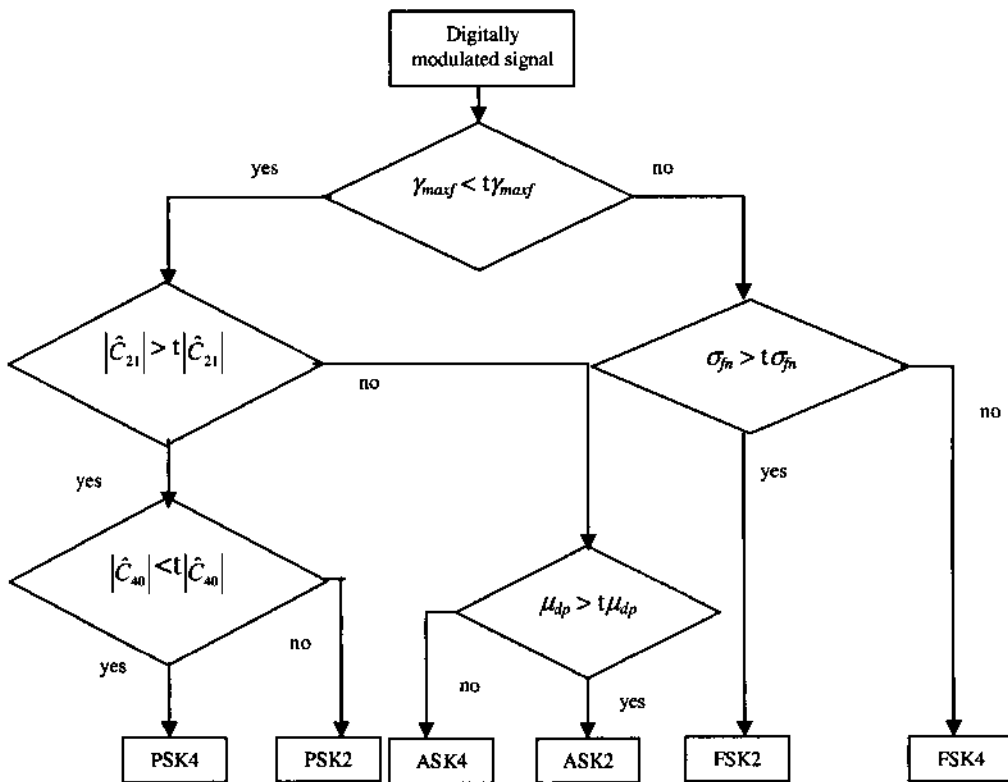


Figure 5.8. Decision tree for classification of digital modulated signals. The first decision separates signals with frequency information (right side of tree – FSK) from signals with little or no frequency information (left side of tree – ASK and PSK). The signals with no frequency information are then separated into signals with phase information (PSK) and signals with little or no phase information (ASK).

The estimated total error probability for the key feature  $\gamma_{maxf}$  is shown in Figure 5.9 to separate subset A (FSK2 and FSK4) and subset B (ASK2, ASK4, PSK2 and PSK4). It can be seen that a good choice for the threshold  $t\gamma_{maxf}$  is  $-40\text{dB}$  where the total minimum error probability is 0 for the SNR range of  $20\text{dB}$  to  $-5\text{dB}$ .

The estimated total error probability for the key feature  $|\hat{C}_{21}|$  is shown in Figure 5.10 for subset A (PSK2 and PSK4) and subset B (ASK2 and ASK4). The relevant threshold  $t|\hat{C}_{21}|$  is chosen to be 0.93 where the total minimum error probability is 0.0021 for the SNR range of  $20\text{dB}$  to  $5\text{dB}$ .

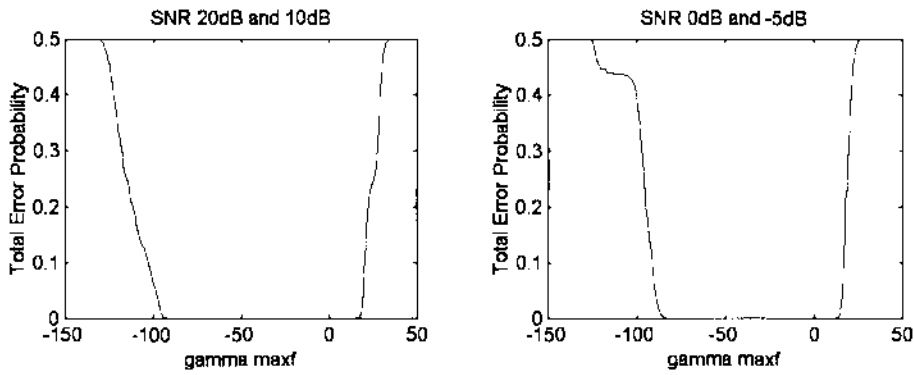


Figure 5.9. Total error probability for the key feature  $\gamma_{maxf}$  for SNR range of  $20\text{dB}$  to  $-5\text{dB}$ , for FSK2 and FSK4 (subset A) and ASK2, ASK4, PSK2 and PSK4 (subset B).

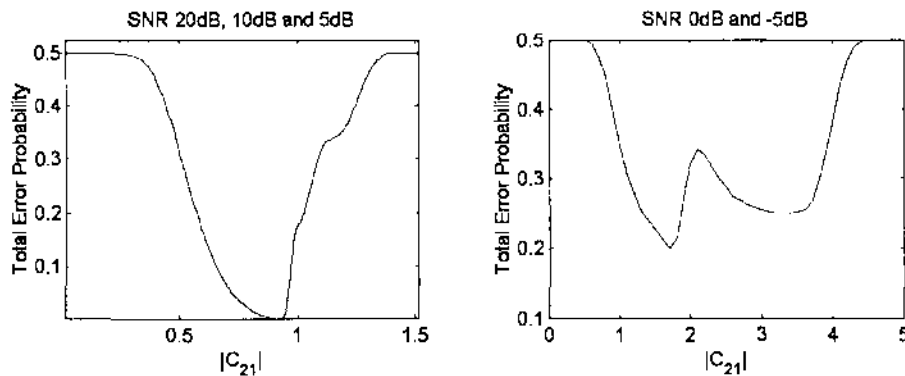


Figure 5.10. Total error probability for the key feature  $|\hat{C}_{21}|$ , for SNR range of  $20\text{dB}$  to  $-5\text{dB}$ , for PSK2 and PSK4 (subset A) and ASK2 and ASK4 (subset B).

The estimated total error probability for the key feature  $|\hat{C}_{40}|$  is shown in Figure 5.11 for subset A (PSK4) and subset B (PSK2). The total minimum probability of error is 0.0008 at the threshold  $t|\hat{C}_{40}| = 1.35$  for SNR range of 20dB to 5dB. The same threshold value gives the minimum error probability for SNR range of 0 to -5dB.

For FSK2 (subset A) and FSK4 (subset B), the total error probability for the key feature  $\sigma_{fn}$  is shown in Figure 5.12. The relevant threshold  $t\sigma_{fn}$  is chosen to be 1.65 where the total minimum error probability is 0.03 for the SNR range of 20dB to 5dB.

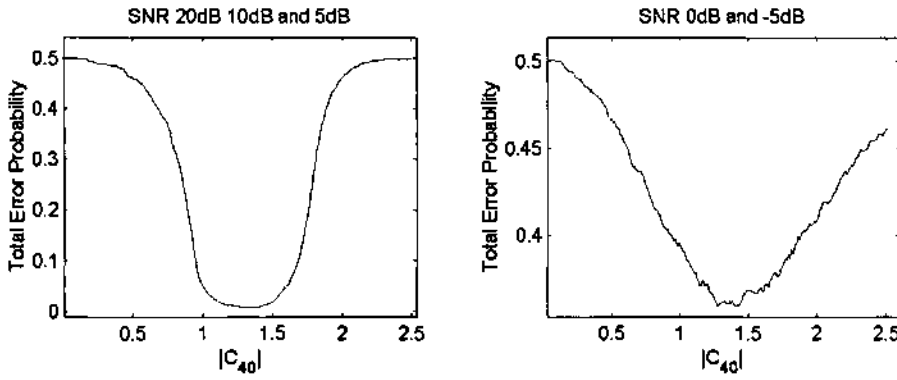


Figure 5.11. Total error probability for the key feature  $|\hat{C}_{40}|$ , for SNR range of 20dB to -5dB, for PSK4 (subset A) and PSK2 (subset B).

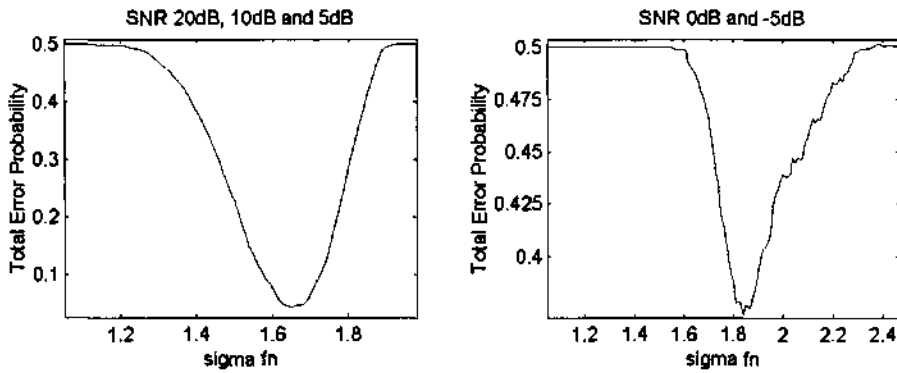


Figure 5.12. Total error probability for the key feature  $\sigma_{fn}$ , for SNR range of 20dB to -5dB, for FSK2 (subset A) and FSK4 (subset B).

The total error probability for the classification of ASK2 (subset A) and ASK4 (subset B) using the key feature  $\mu_{dp}$  is shown in Figure 5.13. The optimum threshold  $t_{\mu_{dp}}$  is chosen to be -0.125 for the SNR range of 20dB to -5dB.

A summary of the key feature threshold values and minimum error probabilities for the SNR range of 20dB to -5dB is shown in Table 5.1. These threshold values are used to discriminate between groups of signals as shown in Figure 5.8. A compromise must be made between the threshold values at higher and lower SNR. The threshold must be chosen so that the overall classification error is minimised. Here, we choose the threshold values that minimise the error probability between 20 and 5 dB; thus, the key feature thresholds  $t_{\gamma_{maxf}}$ ,  $t_{\sigma_{fn}}$ ,  $t_{\mu_{dp}}$ ,  $t_{|\hat{C}_{21}|}$  and  $t_{|\hat{C}_{40}|}$  are -40, 1.65, -0.125, 0.93 and 1.35, respectively.

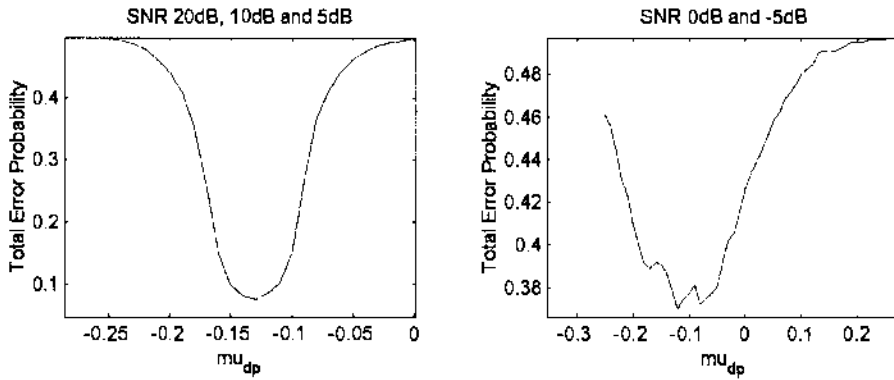


Figure 5.13. Total error probability for the key feature  $\mu_{dp}$ , for SNR range of 20dB to -5dB, for ASK2 (subset A) ASK4 (subset B).

#### 5.4.2 Dependency of Key Feature Selection on Minimum Probability of Error

The selection of a particular key feature for a specific decision is dependent on the minimum error probability. For example, the reason why the key feature  $\gamma_{maxf}$  is chosen for the first decision in Figure 5.8 is because it minimises the total error probability for that decision. In this section we examine the decision made at every stage of the decision tree and explain why a particular key feature is chosen at that stage, starting from the top of the tree.



Table 5.1. Summary of key feature thresholds and error probabilities.

Key Feature Threshold	SNR 20dB to 5dB		SNR 0dB to -5dB	
	Optimum Threshold	Minimum Error Probability	Optimum Threshold	Minimum Error Probability
$t\gamma_{maxf}$	-40	0	-40	0
$t\sigma_{fn}$	1.65	0.0300	1.84	0.3725
$t\mu_{dp}$	-0.125	0.0800	-0.12	0.3700
$t \hat{C}_{21} $	0.93	0.0021	1.7	0.1984
$t \hat{C}_{40} $	1.35	0.0008	1.35	0.3600

#### 5.4.2.1 Decision 1

At every stage of the decision tree, there are many possible scenarios that must be considered, depending on how the signals are grouped together. Let's call the grouping of FSK2 and FSK4 (subset A) and ASK2, ASK4, PSK2 and PSK4 (subset B) as *scenario 1-1* (scenario 1 of decision 1). Table 5.2 illustrates the total error probabilities of the different key features, along with the appropriate threshold (shown in brackets). The key feature that minimises the probability of error is chosen and the minimum error is shown in bold typeface. The errors are calculated from data based on the SNR range of 20dB to -5dB. Note that for scenario 1-1, the key feature  $\sigma_{fn}$  could also have been chosen since this feature also has a total minimum error probability of 0.

Alternatively, we could use the key feature  $|\hat{C}_{21}|$  for the first decision in the tree. Then the two groups of signals to be separated would be FSK and PSK signals in one group (subset A) and ASK signals in the other (subset B) which we will call *scenario 2-1*. The total error probabilities for this scenario are also shown in Table 5.2. It can be seen that the key feature  $|\hat{C}_{21}|$  minimises the error probability between these two groups of signals. However, this feature is not chosen for the first decision because the overall minimum error is not as small as that of scenario 1-1.

If we separated the signals into subset A consisting of PSK2 and PSK4 signals and subset B comprising ASK2, ASK4, FSK2, and FSK4 signals, which we refer to as *scenario 3-1*, the feature  $|\hat{C}_{40}|$  minimises the total error probability. However, as can be seen in Table 5.2, the total minimum error for this decision is still not as small as the error in scenario 1-1. Therefore, this feature is not used for the first decision.

Table 5.2. Total minimum error probability for different scenarios of Decision 1 for combined SNR range of 20dB to -5dB (threshold values are shown in brackets).

Key Feature	Total Minimum Error Probability ( <i>Scenario 1-1</i> )	Total Minimum Error Probability ( <i>Scenario 2-1</i> )	Total Minimum Error Probability ( <i>Scenario 3-1</i> )
$\gamma_{maxf}$	<b>0</b> (-40)	0.3333 (36.2)	0.1788 (-104.45)
$\mu_{dp}$	0.1473 (0.2)	0.3333 (-125)	0.3333 (100.7)
$ \hat{C}_{21} $	0.3185 (1.49)	<b>0.0794</b> (0.92)	0.2869(1.0850)
$ \hat{C}_{40} $	0.1089 (0.1)	0.3333 (1.1)	<b>0.1178</b> (0.7)
$\sigma_{fn}$	0 (1.0)	0.3333 (0.5)	0.3333 (0.5)

#### 5.4.2.2 Decision 2

The second decision in Figure 5.8 separates ASK signals from PSK signals using the key feature  $|\hat{C}_{21}|$ ; we call this *scenario 1-2* (Scenario 1 of Decision 2). Table 5.3 shows that the key feature  $|\hat{C}_{21}|$  minimises the total error probability for this decision.

Another possibility is to separate ASK2 (subset A) from ASK4, PSK2, and PSK4 (subset B); we call this *scenario 2-2*. It can be seen from Table 5.3 that although the key feature  $\mu_{dp}$  minimises the total error probability, it is still much higher than that of scenario 1-2.

The third scenario (*scenario 3-2*) has subset A consisting of ASK4, and subset B comprising ASK2, PSK2, and PSK4. It can be seen from Table 5.3 that the key feature

$|\hat{C}_{21}|$  minimises the probability of error. However, this scenario is not chosen because its minimum error is still larger than that of scenario 1-2.

By assigning PSK4 to one class and ASK2, ASK4, and PSK2 to the other class, we define *scenario 4-2*, where the minimum error probabilities for each feature are shown in the fifth column of Table 5.3. It can be observed that the feature  $\mu_{dp}$  minimises the total error probability but this minimum error is still not as small as that of scenario 1-2.

Table 5.3 Total minimum error probability for Scenarios 1 – 4 of Decision 2 for combined SNR range of 20dB to –5dB (threshold values are shown in brackets).

Key Feature	Total Minimum Error Probability ( <i>Scenario 1-2</i> )	Total Minimum Error Probability ( <i>Scenario 2-2</i> )	Total Minimum Error Probability ( <i>Scenario 3-2</i> )	Total Minimum Error Probability ( <i>Scenario 4-2</i> )
$\gamma_{maxf}$	0.3227 (-104.1)	0.2475 (-91.35)	0.2500 (-90)	0.2800 (-113)
$\mu_{dp}$	0.1968 (-0.145)	<b>0.2295 (-0.11)</b>	0.3025 (0)	<b>0.2310 (-0.2950)</b>
$ \hat{C}_{21} $	<b>0.1275 (0.93)</b>	0.2500 (0.2)	<b>0.2250 (0.52)</b>	0.3000 (2)
$ \hat{C}_{40} $	0.1300 (0.65)	0.2338 (0.1)	0.2445 (0.4)	0.2416 (2.55)
$\sigma_{fn}$	0.5000 (0)	0.2500 (-0.01)	0.2500 (-0.1)	0.2500 (-0.1)

*Scenario 5-2* is defined by separating PSK2 from PSK4, ASK2, and ASK4. The minimum error probabilities for this class are shown in Table 5.4. It can be seen that the minimum error occurs for the feature  $|\hat{C}_{40}|$  and though the error is low, it is still higher than in scenario 1-2.

Another possibility is to define subset A as ASK2 and PSK2 and subset B as ASK4 and PSK4. The corresponding calculated errors are shown in the third column of Table 5.4 labelled as *scenario 6-2*. The feature that minimises the error probability is  $\mu_{dp}$ . However, this scenario has too large an error probability to be considered for this decision.

The final scenario (*scenario 7-2*) defines subset A as ASK2 and PSK4 and subset B as ASK4 and PSK2. The feature that minimises the error probability for the SNR range of 20dB to -5dB is  $|\hat{C}_{40}|$ . This scenario is not feasible due to the high error risk as can be seen in Table 5.4.

Table 5.4. Total minimum error probability for Scenarios 5 – 7 of Decision 2 for combined SNR range of 20dB to -5dB (threshold values are shown in brackets).

Key Feature	Total Minimum Error Probability ( <i>Scenario 5-2</i> )	Total Minimum Error Probability ( <i>Scenario 6-2</i> )	Total Minimum Error Probability ( <i>Scenario 7-2</i> )
$\gamma_{maxf}$	0.2100 (-122)	0.3840 (-99.5)	0.3850 (-99.75)
$\mu_{dp}$	0.2500 (-0.3)	<b>0.2900 (-0.13)</b>	0.3800 (-0.13)
$ \hat{C}_{21} $	0.2500 (3)	0.4800 (0.65)	0.4710 (0.6)
$ \hat{C}_{40} $	<b>0.1300 (1.1)</b>	0.3100 (1.3)	<b>0.3150 (1.35)</b>
$\sigma_{fn}$	0.2500 (0.01)	0.5000 (0.01)	0.5000 (0.01)

#### 5.4.2.3 Decision 3, Decision 4 and Decision 5

The following decisions carry on from scenarios 1-1 and 1-2 of decisions 1 and 2, as this path gives the smallest error probability. The next decision we will examine in the classification tree separates FSK2 signals from FSK4 signals. The key feature chosen for this decision (which we will call decision 3) is  $\sigma_{fn}$  and the reason for this is that it minimises the total error probability compared to the other key features, as presented in Table 5.5.

Decision 4 separates ASK2 and ASK4 signals. The total minimum error probabilities using each feature for this decision are shown in Table 5.5. It can be seen from this table that the feature that gives the smallest error is  $\mu_{dp}$ .

The final decision separates PSK2 from PSK4 signals (Decision 5). The feature found to minimise the total error probability is  $|\hat{C}_{40}|$ . The other error probabilities for each feature with respect to Decision 5 are shown in Table 5.5.

Table 5.5. Total minimum error probability for Decision 3, Decision 4, and Decision 5 at combined SNR range of 20dB to -5dB (threshold values are shown in brackets).

Key Feature	Total Minimum Error Probability (Decision 3)	Total Minimum Error Probability (Decision 4)	Total Minimum Error Probability (Decision 5)
$\gamma_{maxf}$	0.3350 (22)	0.4004 (-110.45)	0.4328 (-122)
$\mu_{dp}$	0.4875 (30)	<b>0.1673 (-0.125)</b>	0.3250 (0.4)
$ \hat{C}_{21} $	0.4700 (1.0)	0.3871 (0.52)	0.4800 (1.0)
$ \hat{C}_{40} $	0.4825 (0)	0.4618 (0.55)	<b>0.1204 (1.35)</b>
$\sigma_{fn}$	<b>0.1925 (1.65)</b>	0.5000 (1.5)	0.5000 (1.5)

### 5.4.3 Receiver Operating Characteristic (ROC) Curves

A receiver operating characteristic (ROC) curve describes the tradeoff between maximising the probability of a correct decision ( $P_D$  – Detection Probability) and minimising the probability of an incorrect decision ( $P_{FA}$  – False Alarm). By considering two sets of modulation types  $A$  and  $B$ , we can call these two classification possibilities the null hypothesis  $H$  and the alternative hypothesis  $K$ . They are commonly written in the form:

$$H : x = \text{subset } A$$

$$K : x = \text{subset } B$$

where  $x$  is a particular key feature value.

The probability of false alarm is a function of the key feature threshold value  $t_x$  given by  $P_{FA} = P(R_K|H)$ . The probability of a correct decision  $P_D = P(R_K|K)$ . The plot of the pair  $P_{FA} = P_{FA}(t_x)$  and  $P_D = P_D(t_x)$  over the range of thresholds  $-\infty < t_x < \infty$  produces a ROC curve. Good features have ROC curves with desirable properties such as negative curvature, monotone increase in  $P_D$  as  $P_{FA}$  increases, and high slope of  $P_D$  at the point  $(P_{FA},$

$P_D) = (0,0)$ . The aim is to find ways to test between  $K$  and  $H$  that push the ROC curve towards the upper left corner, where  $P_D$  is high for low  $P_{FA}$ .

The ROC curves for the key feature  $\sigma_{fn}$  that separates FSK2 (subset A) from FSK4 (subset B) are shown in Figure 5.14 for SNR range of 20dB to -5dB. The curves show the detection probability of FSK2 (subset A) and the false alarm probability of FSK4 (subset B). By examining the ROC curves for  $\text{SNR} \geq 5\text{dB}$ , we can see that for the chosen threshold value,  $t\sigma_{fn}$  (indicated by 'x') has a detection probability ( $P_D$ ) of 0.9795 and false alarm probability ( $P_{FA}$ ) of 0.1203 at 5dB SNR.

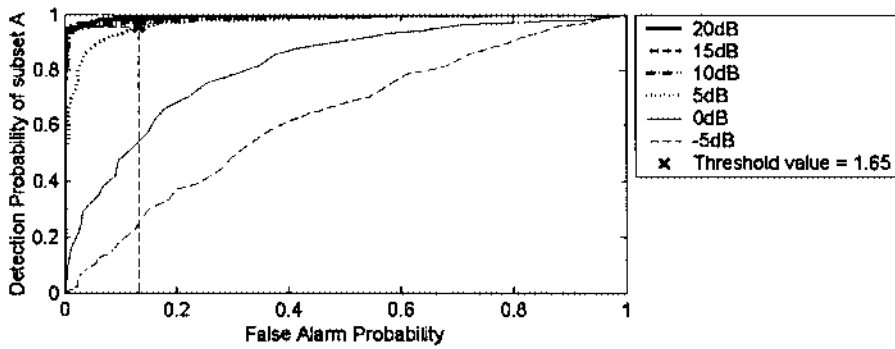


Figure 5.14. ROC curves for the key feature  $\sigma_{fn}$  to separate FSK2 (subset A) and FSK4 (subset B) signals for SNR range of 20dB to -5dB.

The ROC curves for the key feature  $\mu_{dp}$  that separates ASK2 (subset A) from ASK4 (subset B) are shown in Figure 5.14 for SNR range of 20dB to -5dB. For higher SNR,  $P_D$  (detection probability) is high for low  $P_{FA}$  (false alarm probability). By examining the ROC curves for  $\text{SNR} \geq 10\text{dB}$ , we can see that for the chosen threshold value,  $t\mu_{dp}$  (represented by the 'x') has a minimum  $P_D$  (detection probability) of 0.9509 and  $P_{FA}$  (false alarm probability) of 0.0013.

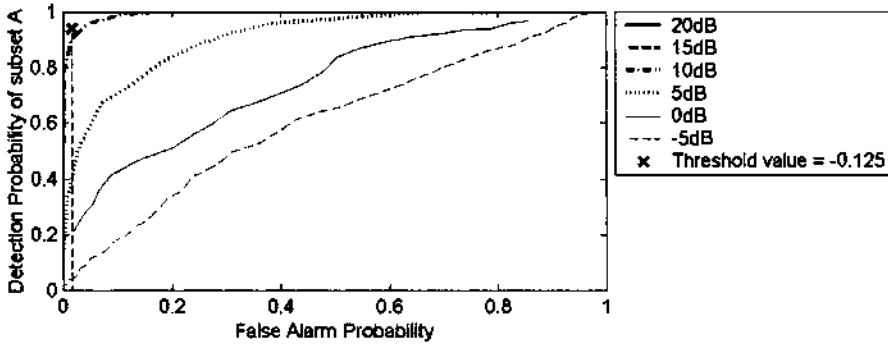


Figure 5.15. ROC curves for the key feature  $\mu_{dp}$  to separate ASK2 and ASK4 signals for SNR range of 20dB to -5dB.

Figure 5.14 and Figure 5.15 show that the chosen key features give very low false alarm rates at very high detection rates. The ROC curves for the remaining decisions in the tree are not presented because the error probability is 0 for SNRs of 20dB to 10dB.

## 5.5 Modulation Classification Using Artificial Neural Networks

The classification of ASK2, ASK4, PSK2, PSK4, FSK2, and FSK4 signals has been shown using the decision-theoretic approach. Classification can also be achieved using artificial neural networks. A neural network classifier will be proposed and compared to the decision theoretic classifier.

Simulations are carried out in Matlab using the neural network toolbox functions. The same key features used in the decision theoretic algorithm are used as the input datasets for the ANN algorithm. These key features are  $\sigma_{fn}$ ,  $\gamma_{maxf}$ ,  $\mu_{dp}$ ,  $|\hat{C}_{21}|$  and  $|\hat{C}_{40}|$ . The key features are normalised to the range  $-1$  to  $1$ , then passed to the neural network. This normalisation is performed to make the training of the network more efficient [Demuth and Beale, 1998] because the inputs have large differences in magnitude and it is also proven in [Azzouz and Nandi, 1996] that normalisation significantly improves the performance of the ANN classifier.

### 5.5.1 Neural Network Structure

Figure 5.16 presents the selected neural network structure for modulation classification of ASK, FSK and PSK signals. It consists of three subnetworks. The first has five inputs, corresponding to the five normalised key features, and four output neurons corresponding to ASK, PSK2, PSK4, and FSK. The other two subnetworks are used to differentiate between ASK2 and ASK4, and FSK2 and FSK4, respectively. The structure that is chosen for the first subnetwork consists of one hidden layer with ten neurons. Twenty versions of this structure are trained and tested to find the network parameters that give the best performance.

For the classification of ASK2 and ASK4, the chosen network structure has one input, corresponding to the key feature  $\mu_{dp}$ , one hidden layer with ten neurons and two output neurons, corresponding to ASK2 and ASK4 signals. Twenty versions of this network structure were trained and tested to find the best network parameters. The network to classify FSK2 and FSK4 has two inputs, corresponding to the features  $\sigma_{fn}$  and  $\gamma_{maxf}$ , one hidden layer with ten neurons, and one output neuron. Again, twenty versions of this network structure were trained and tested to find the best performance.

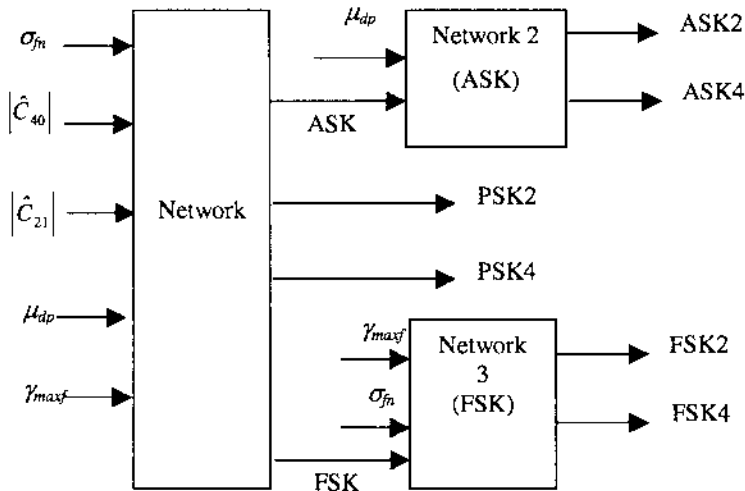


Figure 5.16. Neural network structure for modulation classifier.



The hidden layers in all network structures use the nonlinear *tan-sigmoid* (hyperbolic tangent) activation function because this enables better feature extraction and normally leads to a smaller network [Arulampalam, 1999]. The tan-sigmoid function also generally allows the network to learn faster [Haykin, 1999]. This approach is in contrast with other approaches using neural networks for modulation classification where a *log-sigmoid* function is used in the first hidden layer and a linear function is used in the second hidden layer [Azzouz, 1996] and [Azzouz, 1998]. The output layer uses the log-sigmoid activation function since the ideal outputs should be 1 (true) and 0 (false) for all other outputs. The full network structure is shown in Figure 5.16.

### 5.5.2 Training the Network

The large network is trained using the conjugate gradient method due to its fast training speed and the two smaller networks are trained using the Levenberg-Marquardt (LM) algorithm. This algorithm is currently one of the fastest training algorithms and approaches second-order training speeds [Demuth, 1998]; however, it requires a large amount of memory, which can slow it down significantly with large networks and/or a large amount of training data.

The networks are trained using 200 samples from each modulation type. The networks are also tested and validated using a separate set of 200 samples of each modulation type at different SNR values. While training, a mean square error performance goal is given and a cross validation set is used to stop the training early if overfitting occurs to maintain a good generalisation performance [Haykin, 1999], [Demuth, 1998]. The target values for true and false are offset from 1 and 0 (limit values for log-sigmoid function) to 0.8 and 0.1, respectively to improve the speed of convergence [Haykin, 1999]. The fast convergence properties of the LM algorithm in addition to offsetting the limit values, allows the network to be trained for a maximum of only 1000 epochs. For the larger network using the conjugate training algorithm, a maximum number of 500 epochs is sufficient for training. This is in orders of magnitude less than 250,000 epochs used in [Azzouz, 1996] and [Azzouz, 1998]. It is found that training the network with a mix of samples with SNR

ranging from 20dB to -5dB gives the best overall performance over a good spread of SNR values.

## 5.6 Performance Analysis

The performance results are derived from 200 realisations of each modulation type. The carrier frequency, sampling rate and the symbol rate are given values of 150kHz, 1200kHz and 12.5kHz respectively. The digital symbol sequence is randomly generated and the MPSK, MASK, and MFSK signals are generated using the expressions from Table 3.1 in [Azzouz and Nandi, 1996].

### 5.6.1 DT Classifier Results

The results for the test set of the DT approach are summarised in Table 5.6 for the SNR range of 20dB to -5dB; Table A.1- Table A.6 in Appendix A present the accuracy of the DT classifier on each modulation type. These results indicate that most types of the digital modulation schemes considered can be correctly classified with accuracy greater than 89% for  $\text{SNR} \geq 5\text{dB}$ . Figure 5.18 to Figure 5.23 show the same result graphically for SNR range of 20dB to -5dB.

### 5.6.2 NN Classifier Results and Comparison With DT Classifier

The performance results of the NN classifier are summarised in Table 5.6 and Figure 5.17. Figure 5.18 to Figure 5.23 show the performance of the NN and DT for different modulation types at various SNRs. It can be seen from Table 5.6 and Figure 5.17 that the NN and DT perform comparatively for most modulation types for SNR greater than 10dB. However, for SNRs of 5dB, 0dB and -5dB the NN classifier outperforms the DT classifier. The 0.95 confidence intervals on the accuracy of the DT and NN classifiers, shown in Table 5.6, indicate that these differences are significant. This may be due to the fact that more than one key feature is used in NN classification, whereas, the DT classifier uses only one feature per decision. The confusion matrices showing the results of the NN classifier are shown in Appendix A in Table A.7 – Table A.12 for SNR ranging from 20dB to -5dB, respectively.

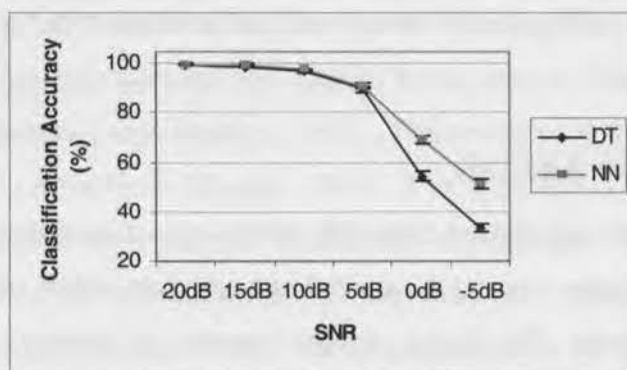


Figure 5.17. Overall accuracy of the NN and DT classifiers at different SNRs.

Table 5.6. DT and NN classifier accuracy and 95% confidence intervals.

SNR	DT Classifier		NN Classifier	
	Accuracy	95% Confidence Interval	Accuracy	95% Confidence Interval
20dB	99.50%	[99.22, 99.78]	99.83%	[99.67, 100.0]
15dB	98.83%	[98.40, 99.26]	99.67%	[99.44, 99.90]
10dB	96.71%	[95.99, 97.42]	97.58%	[96.97, 98.20]
5dB	89.12%	[88.62, 91.04]	90.85%	[89.70, 92.00]
0dB	54.38%	[52.38, 56.37]	69.24%	[67.39, 71.08]
-5dB	33.42%	[31.53, 35.30]	51.10%	[49.10, 53.10]
Overall	78.58%	[77.91, 79.25]	84.71%	[83.88, 85.54]

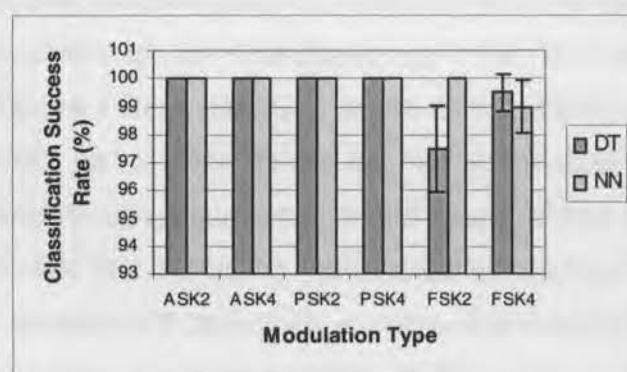


Figure 5.18. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 20dB SNR.

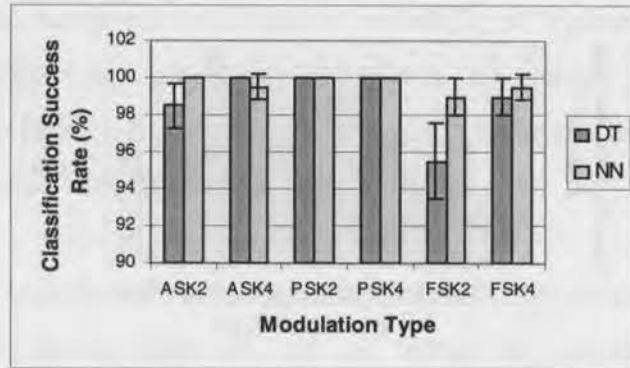


Figure 5.19. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 15dB SNR.

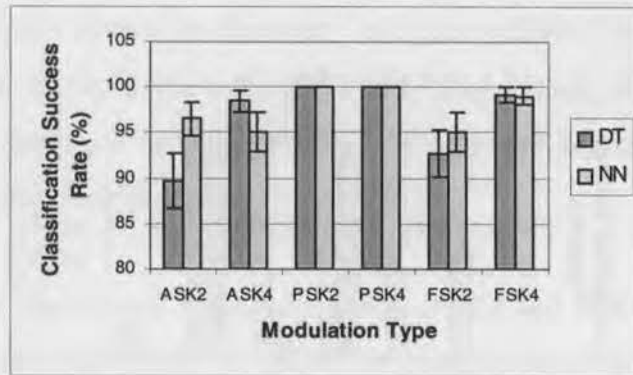


Figure 5.20. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 10dB SNR.

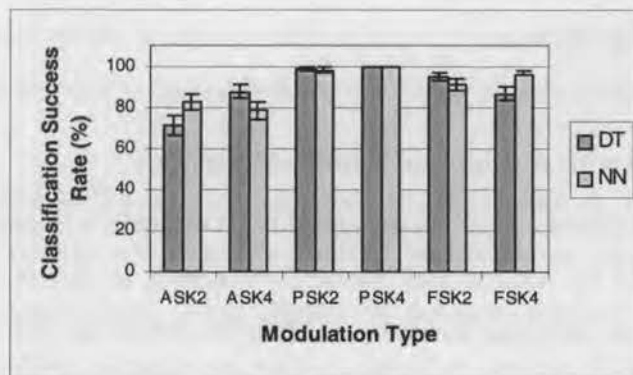


Figure 5.21. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 5dB SNR.

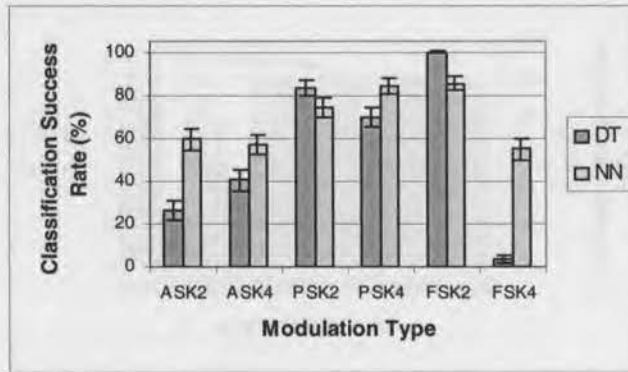


Figure 5.22. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 0dB SNR.

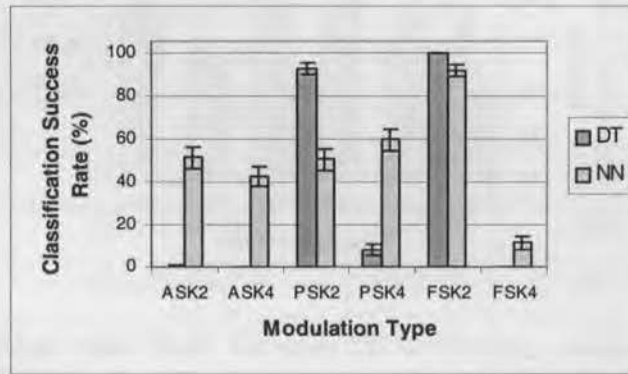


Figure 5.23. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at -5dB SNR.

### 5.6.3 Comparison with Azzouz and Nandi's Classifier

The DT and NN modulation classifiers presented in this chapter will now be compared with the classifier proposed by Azzouz and Nandi. By referring to Figure 3.4 in [Azzouz and Nandi, 1996], it can be seen that the order of classification in the tree structure is slightly different from the proposed tree structure in Figure 5.8. Both structures start by separating the signals with frequency information from those that do not possess frequency information. A&N use the key feature  $\gamma_{max}$  (maximum value of the PSD of the normalised instantaneous amplitude of the signal), whereas the key feature used in the proposed

classifier is  $\gamma_{maxf}$ . The reason for this is that the probability of separation at this point in the tree using  $\gamma_{maxf}$  is 100% whereas it is only 99.6% if the key feature  $\gamma_{max}$  is used (Table 3.3 [Azzouz and Nandi, 1996]). By referring to Figure 5.9, it can be seen that the two sets of signals can be separated with 100% accuracy.

To separate signals with absolute phase information (PSK4) from those that do not possess absolute phase information, A&N use the key feature  $\sigma_{ap}$  (standard deviation of the absolute value of the centred non-linear component of the instantaneous phase evaluated over the non-weak intervals of the signal segment). The key feature proposed in this chapter for the same task is  $|\hat{C}_{40}|$ . It is found in [Swami and Sadler, 2000] that features based on cumulants are immune to frequency and phase offsets. This theory is tested by adding a fixed phase offset of  $\pi/8$  to the PSK2 and PSK4 signals. It is found that the key feature  $\sigma_{ap}$  suffers variations in value, whereas the proposed key feature  $|\hat{C}_{40}|$  suffers no variation when a phase offset is present.

To separate signals that possess phase information (PSK2 and PSK4) from those signals that do not (ASK2 and ASK4), the key feature proposed in this chapter is  $|\hat{C}_{21}|$ . A&N use the key feature  $\sigma_{dp}$  (defined as the standard deviation of the direct value of the centred non-linear component of the instantaneous phase evaluated over the non-weak intervals of the signal segment) for this purpose. However, this feature is not immune to phase variations and may cause inaccuracies in results when phase or frequency offset is present. The feature  $|\hat{C}_{21}|$  is based on cumulants and is therefore robust against phase variations.

To separate ASK2 and ASK4 signals, A&N use the key feature  $\sigma_{aa}$  (standard deviation of the absolute value of the normalised – centred instantaneous amplitude of the signal segment). The key feature used in this chapter for the same purpose is  $\mu_{dp}$ . The proposed DT modulation classifier outperforms A&N's classifier for the SNR of 20dB and 15dB. For the SNR of 10dB, the performance drops slightly. Though this feature is based on phase, there is not much variation in value if there is a phase offset because no information is contained in the phase of the signal. It is also found in later chapters that this feature is robust in environments such as Rayleigh fading channels.

Finally, to separate FSK2 and FSK4 signals, the key feature used in this chapter is  $\sigma_{fn}$ . A&N used the key feature  $\sigma_{af}$  (standard deviation of the absolute value of the normalised – centred instantaneous frequency evaluated over the non-weak intervals of the signal segment) in their modulation classifier. There is not much difference in performance for both key features; however,  $\sigma_{fn}$  provides another alternative to separate the FSK signals.

The comparison in results for SNR of 20dB, 15dB and 10dB are shown in Figure 5.24 to Figure 5.26, respectively. To be fair, the threshold values for the compared proposed modulation classifier are derived from SNR of 20dB and 10dB only, as A&N have done. It can be seen that the results for the DT classifier are on par or slightly better than the classifier proposed by A&N, except for the FSK signals where the performance of the DT classifier is slightly inferior. The NN classifier results are also presented for comparison. A&N's NN classifier performs similarly to their DT classifier, hence the results are not shown. Our NN has been trained with SNRs ranging from 20dB to –5dB. Therefore, had it been trained with data from 20dB and 10dB only, as in [Azzouz and Nandi, 1996], the performance of our NN classifier would have been even better.

## 5.7 Conclusions

This chapter has introduced a modulation classifier that is capable of classifying six different digital modulation schemes. The decision-theoretic approach is used for classification. Key features are extracted from the incoming signal and these features are used to determine the modulation type by comparing the key feature values with a specific threshold. Azzouz and Nandi [Azzouz and Nandi, 1996] have used a similar approach to classify these particular signals; however, different key features and a different tree structure are used here. The key features introduced in this chapter are more robust against variations such as phase offsets. The performance of the decision-theoretic classifier introduced in this chapter is very good with an overall classification success rate of greater than 89% for  $\text{SNR} \geq 5\text{dB}$ . A neural network classifier based on the same key features as the DT approach is also proposed. The results of the NN classifier and the DT classifier are compared. It is found that the NN classifier performs slightly better than the DT approach

for higher SNR and much better for lower SNR. This is due to the fact that the main network is trained with all the key features. This is in contrast with the DT approach where only one key feature is used per decision and the threshold boundaries are only linear. These classifiers serve as a base for this thesis where the ultimate aim is to develop a digital modulation classifier capable of recognising a wide range of digital modulation schemes. The next chapter expands the classifiers discussed in this chapter to accommodate continuous phase modulated signals.

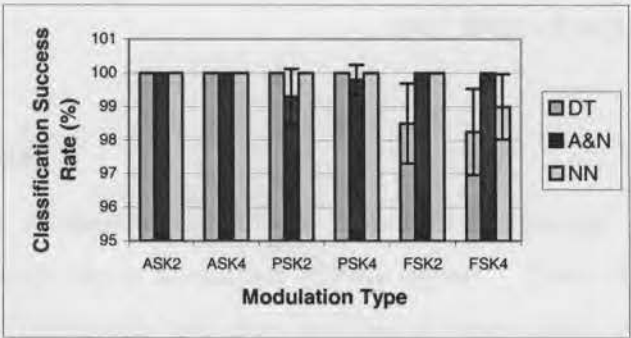


Figure 5.24. Comparison of results of proposed DT and NN classifiers with Azzouz and Nandi’s (A&N) classifier for SNR 20dB.

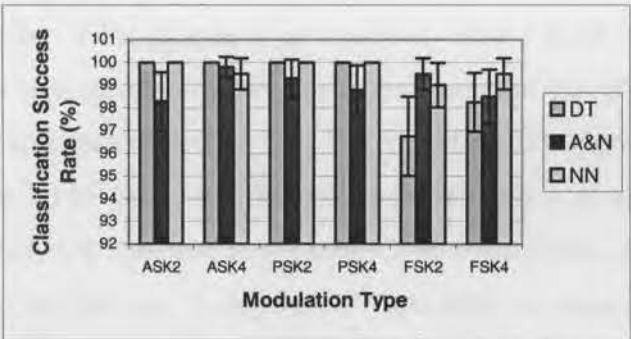


Figure 5.25. Comparison of results of proposed DT and NN classifiers with Azzouz and Nandi’s (A&N) classifier for SNR 15dB.



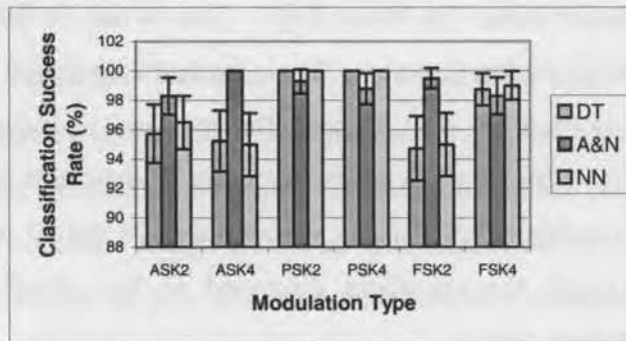


Figure 5.26. Comparison of results of proposed DT and NN classifiers with Azzouz and Nandi's (A&N) classifier for SNR 10dB.

# CHAPTER 6

## Classification of Continuous Phase Modulated Signals

### 6.1 Introduction

In this chapter, the modulation classifiers proposed in Chapter 5 are expanded to accommodate continuous phase modulated (CPM) signals. These classifiers are able to distinguish between CPM signals and other modulation types (ASK, PSK, and FSK). The classifiers can also identify signals within the CPM class - the signals are classified as partial response, full response or Gaussian minimum shift keying (GMSK) signals. The decision-theoretic (DT) approach and neural network (NN) algorithms are compared and results are presented for signal-to-noise-ratios (SNRs) of 20dB, 15dB, 10dB, 5dB, 0dB, and -5dB. The organization of the chapter is as follows. First a brief introduction to CPM signals is given in the next section, followed by a description of classifying CPM signals in general using the DT approach in Section 6.3. We extend the DT approach to classification of signals within the CPM class, for this key features and a novel decision tree are proposed. A neural network classifier to separate CPM signals from ASK, PSK, and FSK signals is proposed in Section 6.4. A separate NN classifier to classify signals within the CPM class is also proposed. The performance results are presented for the DT and NN classifiers and a comparison is made between them in section 6.5 followed by the conclusion of the chapter.

## 6.2 Continuous Phase Modulated (CPM) Signals

Recent publications concerning many techniques for automatic modulation classification have covered different digital signals. However signals with memory, such as CPM, have not been considered using DT and NN methods. In this chapter, CPM signals will be added to the existing modulation classifiers described in Chapter 5, but first a brief introduction to CPM signals is presented.

Continuous phase modulated signals are a class of signals that have memory incorporated in the modulation scheme. These signals have constant amplitude and carry the transmitted information in the phase [Proakis, 1995]. CPM signals are a subset of a class of signals known as continuous-phase FSK (CPFSK). The CPM signal can be described by

$$s(t) = \sqrt{\frac{2\xi}{T}} \cos[2\pi f_c t + \phi(t; I) + \phi_o] \quad (6.1)$$

where  $\xi$  is the signal energy,  $f_c$  is the carrier frequency,  $\phi_o$  is the initial phase of the carrier, and  $\phi(t; I)$  is the time varying phase of the carrier, defined as

$$\begin{aligned} \phi(t; I) &= 4\pi T f_d \int_{-\infty}^t d(\tau) d\tau \\ &= 4\pi T f_d \int_{-\infty}^t \left[ \sum_n I_n g(\tau - nT) \right] d\tau \end{aligned} \quad (6.2)$$

where  $f_d$  is the frequency deviation. Note that the integral of  $d(\tau)$  is continuous even though  $d(\tau)$  is discontinuous. Evaluating the integral in (6.2) gives the phase of the carrier in the interval  $nT \leq t \leq (n+1)T$ :

$$\phi(t; I) = 2\pi \sum_{k=-\infty}^n I_k h_k q(t - kT), \quad nT \leq t \leq (n+1)T \quad (6.3)$$

where  $\{h_k\}$  is a sequence of modulation indices,  $\{I_k\}$  is the sequence of  $M$ -ary information symbols chosen from the alphabet  $\pm 1, \pm 3, \dots, \pm(M-1)$  and  $q(t)$  is a normalised waveform shape which may be represented as the integral of some pulse  $g(t)$

$$q(t) = \int_0^t g(\tau) d\tau \quad (6.4)$$

When  $h_k = h$  for all  $k$ , the modulation index is fixed for all symbols. When the modulation index varies then the signal is referred to as a *multi-h* CPM signal. If  $g(t) = 0$  for  $t > T$ , the CPM signal is called full response CPM. If  $g(t) \neq 0$  for  $t > T$  then the signal is called partial response CPM.

For  $h = 2k/p$  where  $k$  and  $p$  have no common factors, the phase  $\phi(t; I)$  during the interval  $nT \leq t \leq (n+1)T$  can be written as

$$\phi(t; I) = 2\pi h \sum_{k=n-L+1}^n I_k q(t - kT) + \theta_n \quad (6.5)$$

where

$$\theta_n = \left[ h\pi \sum_{k=-\infty}^{n-L} I_k \right] \text{mod } 2\pi \quad (6.6)$$

$h$  is the modulation index and  $\theta_n$  is the memory of all symbols up to time  $(n-1)T$ .

When  $h = 0.5$ , the complex envelope is given by [Couch, 2001]

$$\alpha(t) = A_c e^{\pm j\pi t / (2T_b)} = x(t) + jy(t) \quad (6.7)$$

where  $T_b$  is the bit rate, the  $\pm$  signs denote the possible polarity of the data during the  $(0, T_b)$  interval, and

$$x(t) = A_c \cos(\pm \pi t / 2T_b), \quad 0 < t < T_b \quad (6.8)$$

$$y(t) = A_c \sin(\pm \pi t / 2T_b), \quad 0 < t < T_b \quad (6.9)$$

The instantaneous amplitude and phase are:

$$a(t) = A_c \quad (6.10)$$

$$\phi(t) = \tan^{-1}[y(t)/x(t)] \quad (6.11)$$

These features are shown in Figure 6.1.

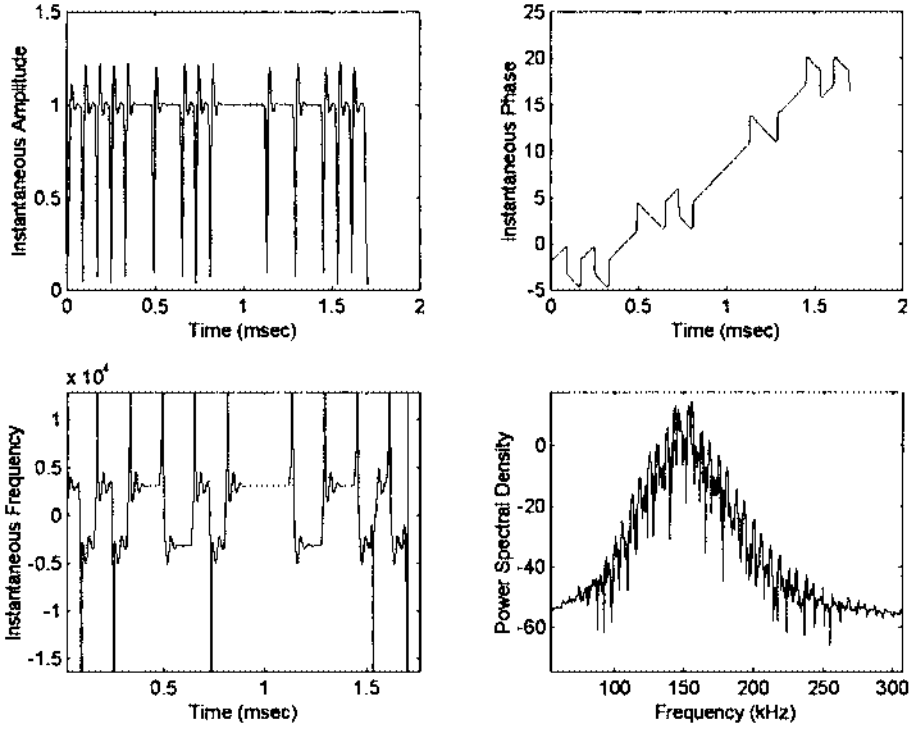


Figure 6.1. Useful features of CPM modulation.

## 6.3 CPM Signal Classification using DT Approach

### 6.3.1 Discrimination of CPM Signals From Other Signals: DT Approach

In this section, we extend the capability of the digital modulation classifier presented in Chapter 5 to cope with signals that have memory incorporated in their modulation scheme. With the decision-theoretic approach, the same classification procedure is used as in Chapter 5. However, to derive the appropriate key features a number of steps must be

taken. First the CPM signal is classified by the existing tree in Chapter 5. The signal is classified as FSK4; therefore, we know the decision has to be made between CPM and FSK4 signals. We can see from the plots of the instantaneous frequency in Figure 6.1 and Figure 5.6 (in Chapter 5) that CPM signals have smaller frequency values than FSK4 signals. In CPM the frequency separation is  $1/2T$ , which is the minimum frequency separation that is necessary to ensure orthogonality of the signals over the interval  $T$  [Proakis, 1995]. For FSK4 the frequency separation will be larger, thus the frequency values and hence the PSD values will be greater.

Therefore, the existing key feature  $\gamma_{maxf}$  is used to distinguish between these two types of signals because FSK4 signals contain more frequency information than CPM signals. The decision tree for this modulation classifier is shown in Figure 6.2. This is the same decision tree as in Chapter 5, except there is an additional decision added after the separation of FSK2 from FSK4.

### 6.3.2 Threshold Determination

The key feature thresholds are chosen so that the probability of a correct decision obtained from 400 realisations of each modulation type at the signal to noise ratio (SNR) range of 20dB to -5dB is maximised.

The optimum threshold  $t\gamma_{maxf2}$  is chosen such that the Bayes error is minimised as described in Chapter 3. The total error probability  $\varepsilon$ , is estimated directly from the sample data. The total error probability for the key feature  $\gamma_{maxf}$  at the SNR range of 20dB to -5dB is shown in Figure 6.3. for subset A (FSK4) and subset B (CPM). It can be seen that a good choice for the threshold  $t\gamma_{maxf2}$  is 15.7 where the total minimum error is 0 for SNR range of 20dB to 5dB. For the SNR range of 0dB to -5dB, the total minimum error at the same threshold is 0.112.

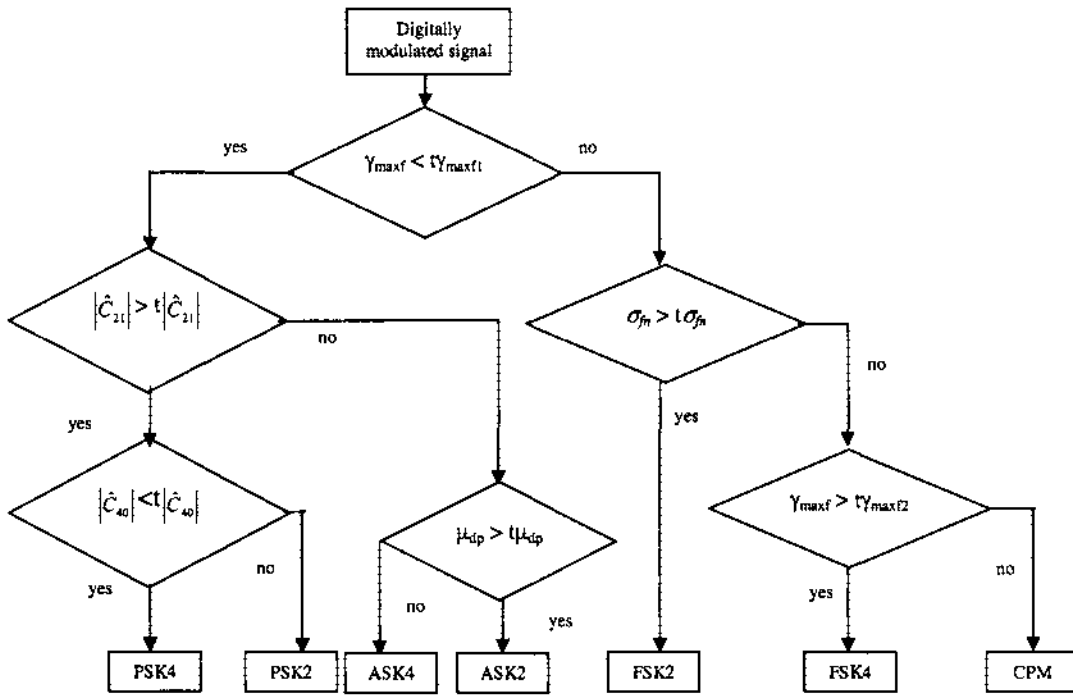


Figure 6.2. Decision tree for classification of digital modulation schemes including CPM signals. The first decision separates signals with frequency information (right side of tree – FSK and CPM) from signals with little or no frequency information (left side of tree – ASK and PSK). The signals with frequency information are divided into FSK and CPM. The signals with no frequency information are then separated into signals with phase information (PSK) and signals with little or no phase information (ASK).

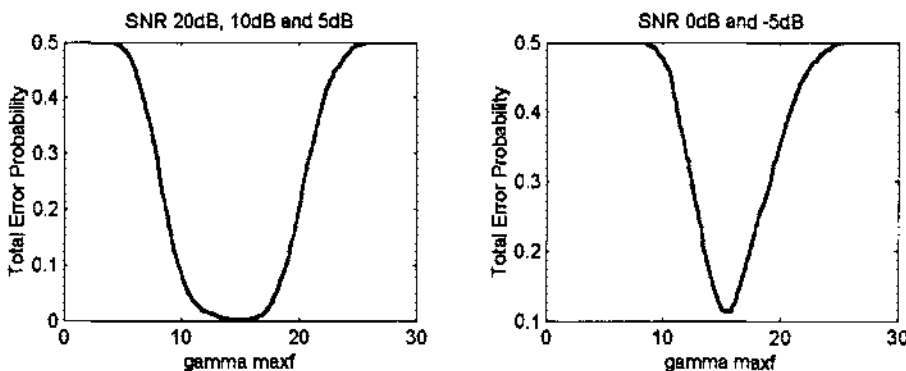


Figure 6.3. Total error probability for the key feature  $\gamma_{maxf}$ , for SNR range of 20dB to 5dB, for FSK4 (subset A) and CPM (subset B).

The key feature  $\gamma_{maxf}$  is chosen to separate FSK4 and CPM, rather than the other existing features, because this feature minimises the total error probability for that decision. This is illustrated in Table 6.1. The estimated minimum error probability for SNR = 20, 10, and 5dB is 0, and it is 0.112 for SNR = 0, and -5dB. Therefore, the total minimum error probability over the SNR range of 20dB to -5dB is estimated to be 0.0448 at the threshold value of 15.7.

### 6.3.3 Classification of Signals Within the CPM Signal Class (DT Approach)

In this subsection, we describe the classification of CPM signals within the CPM class. The proposed modulation classifier categorises the incoming CPM signal as full response, partial response or GMSK (Gaussian minimum shift keying). As described in section 6.2,  $L$  describes the pulse width for each pulse. If  $L = 1$ , the signal is defined as full response (the length of the pulse is equal to the period of the signal ( $T$ )) and if  $L > 1$ , the signal is defined as partial response (the length of the pulse is greater than the period of the signal). There are four pulse shapes that we examine here:

- LREC (rectangular pulse shape)
- LRC (raised cosine)
- HCS (half cycle sinusoid)
- GMSK (Gaussian minimum shift keying)

Table 6.1. Total minimum error probability for Decision 1: classification of FSK4 (subset A) and CPM (subset B) at combined SNR range of 20dB to -5dB (threshold values are shown in brackets).

Key Feature	Total Minimum Error Probability
$\gamma_{maxf}$	<b>0.0448 (15.7)</b>
$\mu_{dp}$	0.4390 (-1.6)
$ \hat{C}_{21} $	0.3650 (1)
$ \hat{C}_{40} $	0.4050 (0.04)
$\sigma_{fn}$	0.0850 (1.2)



These pulse shapes are described as follows.

For LREC:

$$g(t) = \begin{cases} \frac{1}{2LT}; & 0 \leq t \leq LT \\ 0; & \text{otherwise} \end{cases} \quad (6.12)$$

If  $L = 1$ , the signal is known as MSK or Minimum shift keying.

For LRC:

$$g(t) = \begin{cases} \frac{1}{2LT} \left[ 1 - \cos\left(\frac{2\pi t}{LT}\right) \right]; & 0 \leq t \leq LT \\ 0; & \text{otherwise} \end{cases} \quad (6.13)$$

For HCS:

$$g(t) = \begin{cases} \frac{\pi}{4LT} \sin\left(\frac{\pi t}{LT}\right); & 0 \leq t \leq LT \\ 0; & \text{otherwise} \end{cases} \quad (6.14)$$

For GMSK:

$$g(t) = \frac{1}{2T} \left[ Q\left(2\pi B_o \frac{t - \frac{T}{2}}{\sqrt{\ln 2}}\right) - Q\left(2\pi B_o \frac{t + \frac{T}{2}}{\sqrt{\ln 2}}\right) \right]; \quad 0 \leq B_o T \leq \infty \quad (6.15)$$

where  $Q(t)$  is the complementary error function  $\text{erfc}$ .

The notation for each pulse is denoted by the value of  $L$  followed by the pulse description. For example to show an LRC pulse shape with a pulse width of  $4T$  we denote this as 4RC. Since, the signals within the CPM class are very similar, it is only possible to differentiate between  $L = 1$ ,  $L = 2$ , and GMSK for  $\text{SNR} \geq 10\text{dB}$ . The full response and partial response signals are made up of a mixture of the three pulse shapes (LREC, LRC, and HCS). For example, the partial response signal consists of a mixture of 2REC, 2RC, and HCS where  $L = 2$ .

In the next section, the CPM receiver structure is described for both partial response and full response signals. The reason that the receive structure is discussed, is because it is explained that a CPM receiver designed for a particular pulse shape can accommodate any other pulse shape without sacrificing performance. This is the reason why the classifier is designed to recognise only three categories (partial response, full response and GMSK).

#### **6.3.4 CPM Receivers**

The purpose behind discriminating between CPM signals is so that the appropriate demodulator can be chosen to extract the desired transmitted information. We describe two types of receivers for CPM signals. The first receiver is for full response signals and the second receiver applies to partial response signals. Although it seems that there is one receiver for each type of CPM scheme, this is not the case. Tailoring the receiver to a specific CPM signal may simplify the receive structure. However, it is shown in [Swensson, 1994] that any CPM receiver can apply to all pulse shapes for both full and partial response without sacrificing the performance.

##### **6.3.4.1 ML Receiver for CPFSK**

The first receiver presented in [Anderson, 1986] is an optimum ML coherent receiver for CPFSK (CPM with a 1REC pulse shape). This receiver makes a decision about one symbol only, based on observation of a sequence of consecutive symbols. Although the receiver is for CPFSK detection it can be applied to CPM schemes with any pulse shape and any modulation index  $h$  provided that the CPM signal is full response ( $L = 1$ ). Also the receiver structure can be simplified for the special case of full response CPM with  $M = 2$  and  $h = 0.5$  (eg MSK).

##### **6.3.4.2 Optimum Viterbi Receivers**

The ML receiver for CPFSK (described in subsection 6.3.4.1) can apply to partial response ( $L > 1$ ) CPM signals but the receiver structure becomes unreasonably complex. Therefore, a general receiver for partial response CPM is used. The ML sequence estimation is done by means of a Viterbi processor. The metric (correlation between the received signal and an estimated signal over the  $n^{\text{th}}$  symbol interval) is calculated in a bank of linear filters

which are sampled every symbol interval. In other words, the receiver correlates the received signal over one symbol interval with all possible transmitted alternatives over that symbol interval. The complexity grows exponentially with signal memory. The limiting factors are the number of states  $S = pM^{L-1}$  and the number of filters  $F = 2M^L$  for calculating the metrics. For many cases with long smoothing pulses, the optimum receiver can be approximated by a receiver based on a shorter pulse shape  $g_r(t)$  of length  $L_r < L$  so that the complexity is reduced. It is shown in [Svensson, 1984] that the loss in error probability is very small when for example a binary 4RC signal is received in a 2REC receiver.

The key feature derivation is described in the next section, which includes an overview of the power spectra of CPM signals (since one of the key features is based on the PSD of the signal).

### 6.3.5 Key Feature Derivation

The key features used to distinguish between CPM signals are:

- $L_{diff}$ , which is the value of the smoothed PSD of the received signal at the carrier frequency of 150kHz. It is defined as:

$$L_{diff} = 10 \log_{10} (|DFT(s)|^2) \Big|_{f=150kHz} \quad (6.16)$$

where  $s$  is the received signal segment.

- $\sigma_a$ , which is the standard deviation of normalised-centred instantaneous amplitude and is defined by:

$$\sigma_a = \sqrt{\frac{1}{N_s} \left[ \sum_{i=1}^{N_s} A_{cn}^2(i) \right] - \left[ \frac{1}{N_s} \sum_{i=1}^{N_s} A_{cn}(i) \right]^2} \quad (6.17)$$

where  $A_{cn}(i)$  is the value of the normalised-centred instantaneous amplitude at time instants  $t = i/f_s$  ( $i = 1, 3, \dots, N_s$ ) and  $f_s$  is the sampling frequency.

$$A_{cn}(i) = A_n(i) - 1 \quad \text{where} \quad A_n(i) = \frac{A(i)}{m_a} \quad (6.18)$$

$m_a$  is the mean instantaneous amplitude evaluated over one segment

$$m_a = \frac{1}{N_s} \sum_{i=1}^{N_s} A(i) \quad (6.19)$$

Normalisation is necessary to compensate for the channel gain [Azzouz and Nandi, 1996].

- $\sigma_m$  is the standard deviation of the normalised instantaneous frequency, evaluated over the non-weak segments of the intercepted signal and was defined in Chapter 5, equation (5.41).

These key features are proposed for the following reasons:

- The bandwidth occupancy of CPM depends on the modulation index  $h$ , the pulse shape  $g(t)$  and the number of signals  $M$ . In general, small values of  $h$  result in the CPM signal having relatively small bandwidth occupancy, whereas large values of  $h$  result in large bandwidth occupancy. The use of smoother pulse shapes such as LRC results in smaller bandwidth occupancy. An example taken from [Proakis, 1995] is shown in Figure 6.4 where the power density spectrum is shown for binary CPM with different partial response raised cosine (LRC) pulses and  $h = 0.5$ . The power spectrum for an MSK signal is also shown for comparison. It can be seen that as  $L$  increases, the pulse  $g(t)$  becomes smoother and hence the corresponding spectral occupancy of the signal decreases. Therefore  $L_{diff}$  can be used to separate partial response CPM from full response CPM. Since the receivers mentioned in section 6.3.4 apply to all pulse shapes, it is only necessary to distinguish between full and partial response CPM signals. By inspecting the spectral performance of full and partial response schemes in [Anderson et al, 1986] it can be concluded that increasing the pulse duration  $L$  leads to a more compact PSD with side lobes that fall off more smoothly. Therefore a key feature may be the value of the PSD of the signal at some particular frequency. Figure 6.5 shows the smoothed PSD of two LREC signals with  $L = 1$  and  $L = 2$ . Figure 6.6 shows a close up of the PSD around the peak. It can be seen that for  $L = 2$ , the PSD side lobes fall off more quickly and the value of the PSD around the peak is less than that for  $L = 1$ . It can be seen from Figure 6.6 that for the frequency value of 150kHz (carrier frequency), the PSD values for both signals can be separated. The partial response schemes should have

lower PSD values therefore this key feature, can be used to separate partial response CPM from full response CPM.

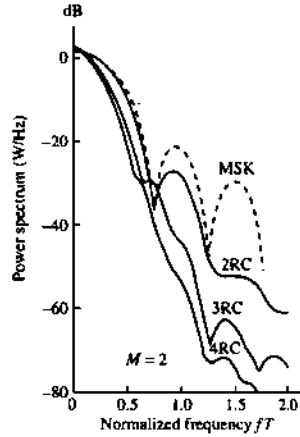


Figure 6.4. PSD of binary CPM with different pulse shapes ( $h = 0.5$ ) [Proakis, 1995].

- $\sigma_{f_n}$  is used to separate the signals at SNR of 10dB from signals at SNR of 20dB and 15dB. Since signals within the CPM class are so similar, it is very difficult to separate them at low SNR values. However, at SNR values greater than or equal to 10dB, it is possible to discriminate between full response, partial response and GMSK. Therefore, when deriving threshold values, we do not consider lower SNR values.
- $\sigma_a$  is used to distinguish between partial response signals and full response signals at SNR of 10dB. It is found that the feature  $L_{diff}$  is not sufficient to discriminate between these two classes for SNR values less than 15dB. The feature  $\sigma_a$  can be used to distinguish between partial and full response. This is due to the fact that although CPM signals have constant amplitude, there are slight variations in the instantaneous amplitude of partial response and full response signals. In general, partial response signals having lower instantaneous amplitude values than full response signals. The same key feature is used to separate GMSK signals from partial and full response signals. However separate threshold values are used for the SNR of 10dB and for the SNR range of 20dB to 15dB. It is found that in general,

GMSK signals have lower instantaneous amplitude values than full and partial response signals.

The DT classifier structure is discussed in the next section.

### 6.3.6 DT CPM Classification Method

The incoming CPM signal is categorised as full response, partial response or GMSK. This is because the receiver structure will be simplified if there is a separate design for each category. As mentioned earlier, the full and partial response signals consist of a mixture of the three pulse shapes (LREC, LRC, and HCS). The decision tree depicting the classification procedure is shown in Figure 6.7. A description of the threshold determination is presented in the next section.

### 6.3.7 Threshold Determination

The key feature thresholds are chosen such that the Bayes error is minimised as described in Chapter 3. The threshold values are obtained using 200 realisations of each modulation type at the SNR range of 20dB to 10dB (for the first decision – Decision A). For the left hand side of the decision tree, all threshold values are obtained using data of SNR 10dB. Similarly for the right hand side of the tree, the thresholds are found from data of SNR range 20dB to 15dB.

For Decision A, the desired threshold value for the feature  $\sigma_{fn}$  to separate SNR of 10dB from 20dB and 15dB can be found in Figure 6.8. It can be seen that an appropriate value for  $t\sigma_{fn}$  is 0.446 which has a corresponding minimum error of 0.0104.

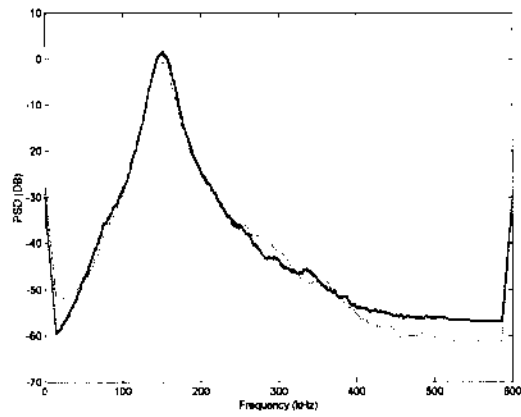


Figure 6.5. Smoothed PSD of LREC signals ( $L=1$  and  $L=2$ ) at SNR of 20dB.

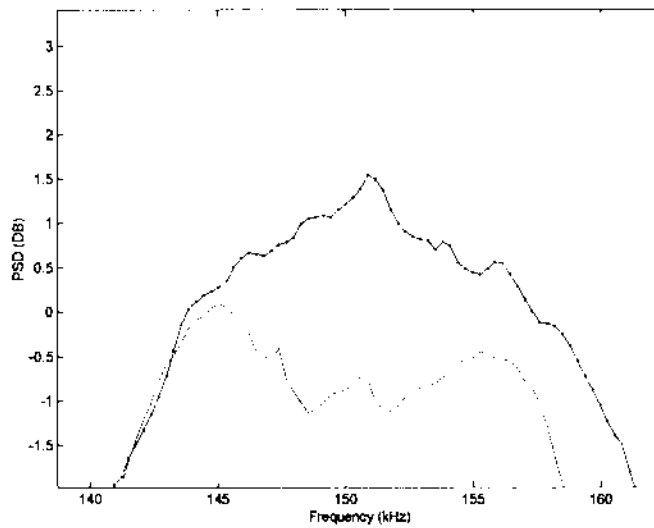


Figure 6.6. Close up of PSD in Figure 1 around the peak.

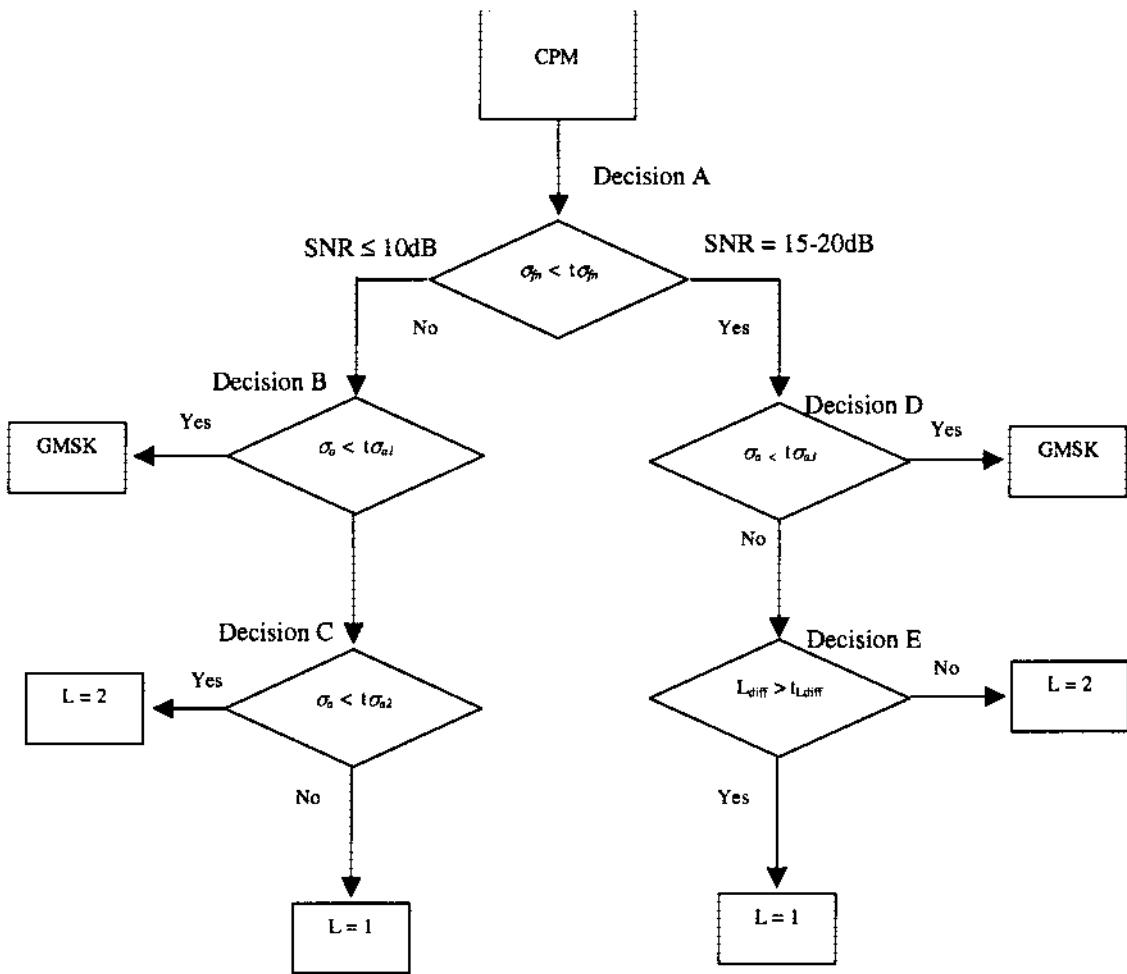


Figure 6.7. Decision tree for CPM signals. The first decision separates the signals at SNR of 10dB from signals with 15-20dB SNR (Decision A). If  $\sigma_m > t\sigma_m$ , then the signal SNR is less than or equal to 10dB. If this condition is satisfied, the next decision separates GMSK from partial and full response CPM (Decision B). Finally, partial and full response signals are classified in Decision C. If the signal is of SNR greater than 10dB, the next decision separates GMSK from partial and full response CPM (Decision D). The final decision classifies the signal as either full response or partial response CPM (Decision E).



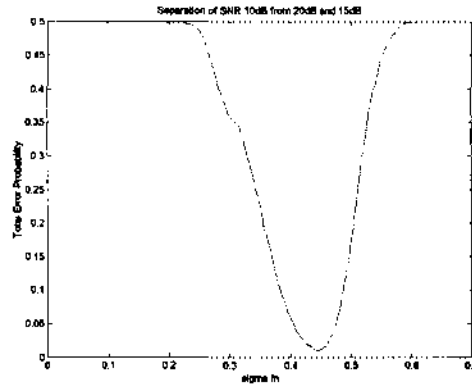


Figure 6.8. Total error probability for the key feature  $\sigma_h$  for Decision A (SNR of 20dB, 15dB and 10dB).

The estimated total error probability for the key feature  $\sigma_a$  that separates GMSK from full and partial response CPM at SNR of 10dB (Decision B) is 0.0563. This corresponds to the threshold value  $t\sigma_{a1}$  of 0.256. Similarly, for Decision C which separates full and partial response CPM at SNR of 10dB, the appropriate threshold value is  $t\sigma_{a2} = 0.275$ . This corresponds to a total minimum error probability of 0.2625. These values can be confirmed by referring to Figure 6.9.

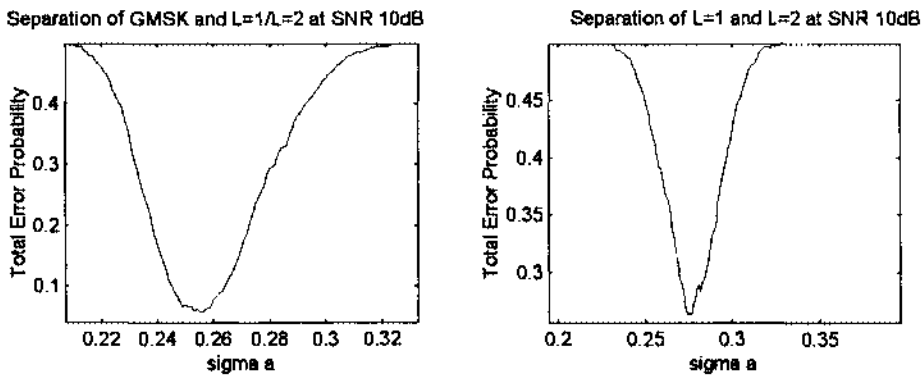


Figure 6.9. Total error probability for the key feature  $\sigma_a$  for Decision B and Decision C (SNR 10dB).

The threshold value  $t\sigma_{a,3}$  is also found in Figure 6.10 to separate GMSK from  $L = 1$  and  $L = 2$  for SNR values greater than or equal to 15dB (Decision D). An appropriate threshold is 0.175, which corresponds to a total minimum error probability of 0.015.

Decision E separates  $L = 1$  and  $L = 2$  at SNR greater than 10dB. The optimum threshold value for  $tL_{diff}$ , is found to be 0.4 from Figure 6.10, where the total minimum error probability is 0.2158.

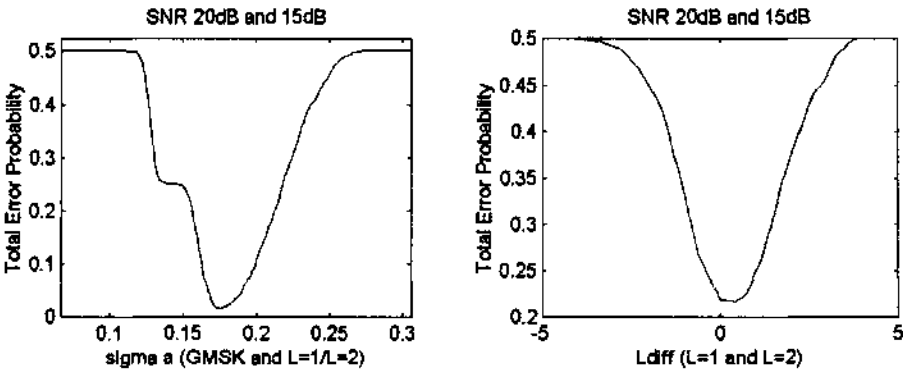


Figure 6.10. Total error probability for the key features  $tL_{diff}$  and  $t\sigma_{a,3}$  for Decision D and Decision E (SNR 20dB and 15dB).

A summary of the key feature values and their corresponding threshold values and minimum error probabilities are given in Table 6.2. The reason why the chosen key features are used for each decision (as described in this section) rather than the other existing key features is that these key features minimise the total error probability. This can be observed in Table 6.3 and Table 6.4 where each decision has the corresponding minimum error probability for every existing key feature. The chosen key feature is shown in bold and its associated threshold value is shown in brackets.

The next section outlines a NN classifier capable of recognising CPM signals. A NN classifier that classifies signals within the CPM class is also presented.

Table 6.2. Summary of key feature values and corresponding threshold values

Key Feature Threshold	Threshold Value	Total Minimum Error Probability
$t\sigma_{fn}$	0.4460	0.0104 (SNR 20dB - 10dB)
$t\sigma_{aa1}$	0.2560	0.0563 (SNR 10dB)
$t\sigma_{aa2}$	0.2750	0.2625 (SNR 10dB)
$t\sigma_{aa3}$	0.1750	0.015 (SNR 20dB and 15dB)
$tL_{diff}$	0.4000	0.2158 (SNR 20dB and 15dB)

Table 6.3. Total minimum error probability for Decision A (SNR 20dB, 15dB and 10dB), Decision B (SNR of 10dB), and Decision C (SNR of 10dB). Threshold values are shown in brackets.

Key Feature	Total Minimum Error Probability (Decision A)	Total Minimum Error Probability (Decision B)	Total Minimum Error Probability (Decision C)
$\sigma_{fn}$	<b>0.0104 (0.4460)</b>	0.5 (0.2)	0.4675 (0.52)
$\sigma_{aa}$	0.1268 (0.253)	<b>0.0563 (0.2560)</b>	<b>0.2625 (0.2750)</b>
$L_{diff}$	0.4146 (0)	0.3713 (2.1)	0.2817 (1.10)
$\gamma_{maxf}$	0.4629 (4.7)	0.5 (2.0)	0.3608 (7.10)
$\mu_{dp}$	0.4879 (12.0)	0.3812 (1.0)	0.3092 (8.0)
$ \hat{C}_{21} $	0.1414 (1.0)	0.2963 (1.06)	0.4892 (1.06)
$ \hat{C}_{40} $	0.4504 (0.07)	0.0903 (0.5)	0.4942 (0.626)

## 6.4 Neural Network Classifier

The classification of ASK2, ASK4, PSK2, PSK4, FSK2, FSK4, and CPM signals has been shown using the decision theoretic approach. A neural network classifier capable of classifying these same seven signals will be proposed in this section. In the succeeding

section, the performance of the NN classifier will be compared to the performance of the decision theoretic approach.

Table 6.4. Total minimum error probability for Decision D (SNR 20dB and 15dB) and Decision E (SNR 20dB and 15dB). Threshold values are shown in brackets.

Key Feature	Total Minimum Error Probability (Decision D)	Total Minimum Error Probability (Decision E)
$\sigma_{fn}$	0.1513 (0.36)	0.2983 (0.2)
$\sigma_{aa}$	<b>0.015 (0.175)</b>	0.33 (0.3)
$L_{diff}$	0.3331 (1.4)	<b>0.2158 (0.4)</b>
$\gamma_{maxf}$	0.4519 (8.7)	0.3838 (5.0)
$\mu_{dp}$	0.5 (-100)	0.2962 (9)
$ \hat{C}_{21} $	0.125 (0.97)	0.5 (-5.0)
$ \hat{C}_{40} $	0.0188 (0.5)	0.4933 (0.1)

Simulations are carried out in Matlab using the neural network toolbox functions. The same key features used in the decision theoretic algorithm are used as inputs to the NN algorithm. These key features are  $\gamma_{maxf}$ ,  $\mu_{dp}$ ,  $|\hat{C}_{21}|$ ,  $\sigma_{fn}$ , and  $|\hat{C}_{40}|$ . The key features are normalised to the range -1 to 1, then passed to the neural network.

The next subsection presents the NN structure of the classifier capable of recognising CPM signals. The training of this network is discussed in subsection 6.4.2. The NN structures for classification of signals within the CPM class are presented in subsection 6.4.3 with a discussion on the training of these NN structures.

#### 6.4.1 Neural Network Structure

The neural network structure is selected to have five inputs, corresponding to the five normalised key features, and four output neurons corresponding to ASK, PSK2, PSK4, and

FSK/CPM. Two other networks are used to differentiate between ASK2 and ASK4 signals on one hand, and FSK2, FSK4 and CPM signals on the other. The structure that is chosen for the large network consists of one hidden layer with twelve neurons. Twenty versions of this structure are tested to find the optimum network that gives the best performance.

For the classification of ASK2 and ASK4, the chosen network structure has one input corresponding to the key feature  $\mu_{dp}$ , and two output neurons corresponding to ASK2 and ASK4 signals. There is one hidden layer with ten neurons, and twenty versions of this network structure are tested to find the optimum performance.

The network to classify FSK2, FSK4, and CPM has two inputs corresponding to the features  $\sigma_{fn}$  and  $\gamma_{maxf}$  and three output neurons corresponding to the three types of signals. There is one hidden layer with twelve neurons, and twenty versions of this network structure are also tested to find the optimum performance.

The hidden layers in all network structures use the nonlinear *tan-sigmoid* (hyperbolic tangent) activation function and the output layer uses the *log-sigmoid* activation function. These functions are chosen for the same reasons as explained in the previous chapter. The full network structure is shown in Figure 6.11.

#### 6.4.2 Training the Network

The large network is trained using the conjugate gradient method due to its fast training speed and the two smaller networks are trained using the Levenberg-Marquardt (LM) algorithm. All networks are trained using 200 samples from each modulation type. The network is also tested and validated using a separate set of 200 samples of each modulation type. The target values for true and false are offset from 1 and 0 (limit values for log-sigmoid function) to 0.9 and 0.1 respectively as outlined in the previous chapter. The training data is a mix of samples of SNR range 20dB to -5dB.

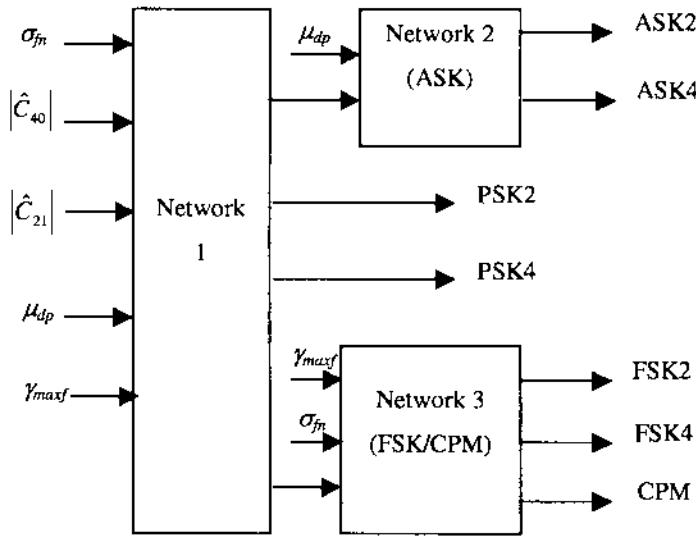


Figure 6.11. Neural network structure for modulation classification of ASK, PSK, FSK and CPM signals.

### 6.4.3 NN Classification Within the CPM Signal Class

Two neural network classifiers are proposed for the classification of partial response, full response, and GMSK signals. The same key features used in the decision theoretic algorithm are used as the inputs to the ANN algorithm. These key features are  $L_{diff}$ ,  $\sigma_a$ , and  $\sigma_m$ , which are normalised to the range  $-1$  to  $1$ , then passed to the neural network.

The first neural network structure is selected to have three inputs, corresponding to the three normalised key features, and three output neurons corresponding to the three CPM signal types. There are two hidden layers with seven neurons in the first layer and five neurons in the second layer. Twenty versions of this structure are tested to find the optimum network that gives the best performance.

The hidden layers in all networks use the nonlinear *tan-sigmoid* (hyperbolic tangent) activation function and the output layer uses a linear activation function. The full network structure is shown in Figure 6.12.

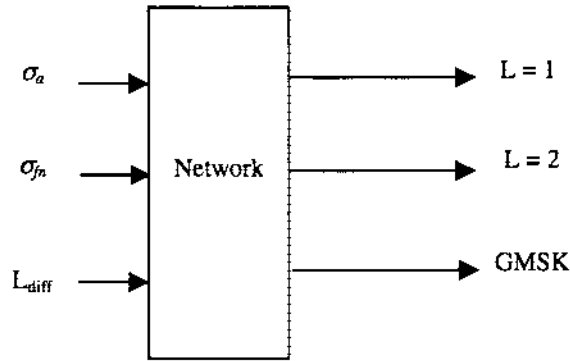


Figure 6.12. Neural network structure for classification of signals within the CPM class.

The network is trained using the Levenberg-Marquardt (LM) algorithm with 200 samples from each modulation type. The network is also tested and validated using a separate set of 200 samples of each modulation type. The target values for true and false are offset from 1 and 0 (limit values for *log-sigmoid* function) to 0.9 and 0.1, respectively. The training data is a mix of samples of SNR 20dB 15dB, and 10dB. This is necessary because the data is highly dependent on SNR as was shown for the DT classifier.

The second neural network structure is made up of three separate networks. Each network is trained with data of SNR 20dB, 15dB, and 10dB respectively. All networks have three input neurons corresponding to the three key features  $L_{diff}$ ,  $\sigma_a$ , and  $\sigma_{fm}$ , and three output neurons corresponding to the three categories  $L = 1$ ,  $L = 2$  and GMSK. The first sub-network is trained with data of 20dB SNR, and has one hidden layer with seven neurons. The second sub-network is trained with data of 15dB SNR, and has one hidden layer with ten neurons. Finally, the third sub-network has one hidden layer with ten neurons and is trained with data of 10dB SNR. The three networks are arranged in parallel and the modulation type with the maximum output is chosen as shown in Figure 6.13.

## 6.5 Results

The results for the DT classifiers are presented first. The NN results are then presented and compared to those of the DT classifiers for signals within the CPM class as well as CPM signals as one modulation type.

## 6.5.1 DT Classifier Performance Results

### 6.5.1.1 Results for DT Classification of ASK, PSK, FSK, and CPM Signals

The performance results of the DT classifier are derived from 200 realisations of each modulation type. The carrier frequency, sampling rate and the symbol rate are given the values of 150kHz, 1200kHz and 12.5kHz, respectively. The digital symbol sequence is randomly generated. The simulation results for the test set based on 200 realisations are given in Appendix B, Table B.1 – Table B.6, for the SNR range of 20dB to -5dB, respectively. The graphical representations of these results are shown in Figure 6.15 - Figure 6.20 for the SNR range of 20dB to -5dB. The results from the NN classifier are also shown for comparison. These results indicate that all types of the digital modulation schemes considered can be correctly classified with greater than 98% success rate for an SNR greater than 5dB. For lower SNR, the performance drops, as can be expected.

### 6.5.1.2 Results for DT Classification Within CPM Class

Simulations are carried out to classify full response CPM signals (consisting of a combination of LREC, LRC and HCS), partial response signals (also comprising LREC, LRC, HCS), and GMSK signals. For all signals  $h = 0.5$  and  $M = 2$ . For the partial response signals,  $L = 2$ . The carrier frequency, sampling rate and the symbol rate are given values of 150kHz, 1200kHz and 12.5kHz, respectively. The digital symbol sequence is randomly generated.

The graphical results for each CPM classification type are shown in Figure 6.21 - Figure 6.26 for the SNR range of 20dB to -5dB. The results for the NN classifier are also shown for comparison. The 95% confidence interval is also shown in all figures by the error bars. The confusion matrices for the DT classifier are shown in Appendix B in Table B.7 – Table B.12. It can be seen that the performance drops dramatically for partial response CPM of SNR less than 10dB because the classifier is trained with data of SNR 20dB, 15dB, and 10dB. However, the performance degradation is not an issue because for many cases of partial response CPM, the optimum receiver can be approximated by a receiver based on full response CPM as explained in subsection 6.3.4.2. It is implied in [Svensson, 1984] that



the loss in error probability is very small when for example a binary 2RC signal is received in a REC receiver.

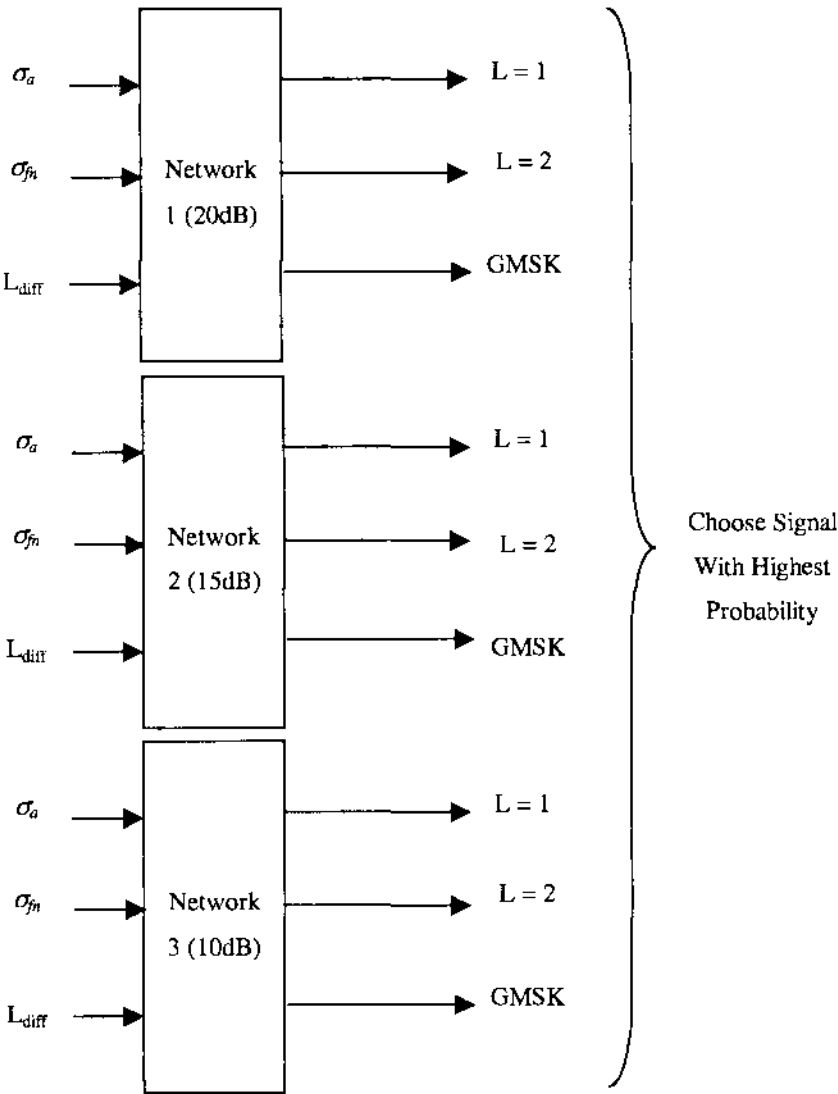


Figure 6.13. Second NN structure to classify signals within the CPM class.

### 6.5.2 NN Classifier Performance With Comparison to DT Classifier Results

#### 6.5.2.1 Results for NN Classification of ASK, PSK, FSK, and CPM Signals

The performance results of the NN classifier at SNR range of 20dB to -5dB are given in Figure 6.15 - Figure 6.20, respectively. The results for the DT classifier are also shown for

comparison with 95% confidence interval. It can be seen that the NN performs well, with 100% success rate for most modulation types with SNR of 20dB, 15dB and 10dB. The 0.95 confidence intervals on the classification accuracy of the DT and NN classifiers for the SNR range of 20dB to -5dB are shown in Table 6.5. It can be seen that the NN classifier performance is slightly better than that of the DT classifier for the SNR range of 20dB to 5dB. This is mainly because the key features have been chosen well enough so that there is minimal overlap between classes. This can be confirmed by referring to the graph of SNR versus classifier accuracy in Figure 6.14. For lower SNR, the NN performs much better than the DT approach because the NN can develop a decision boundary that is not restricted to being linear as in the DT approach. The confusion matrices showing the results of the NN classifier are shown in Appendix B, Table B.13 – Table B.18, for the range of SNR from 20dB to -5dB, respectively.

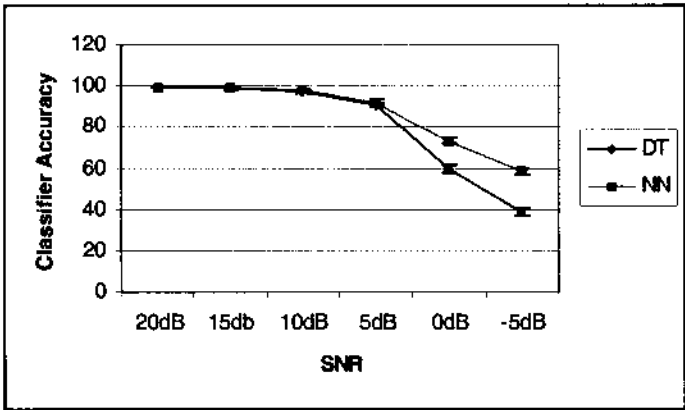


Figure 6.14. Graphical comparison of overall performance between the NN-based and DT-based classifiers with 95% CI for ASK, PSK, FSK, and CPM signals.

### 6.5.2.2 NN Classifier Results for Within CPM Class

The results of classification of CPM signals are shown in Figure 6.21 - Figure 6.26 for SNR range of 20dB to -5dB. The 95% confidence interval is shown on all figures by the error bars. Since the signals are classified as full response, partial response or GMSK, the input data is a combination of the signals from each classification type. For instance, the test data for  $L = 1$  is combined from the full response signals LREC, LRC, and HCS, and likewise for  $L = 2$ . It can be seen that the performance is good for both networks for SNR greater

comparison with 95% confidence interval. It can be seen that the NN performs well, with 100% success rate for most modulation types with SNR of 20dB, 15dB and 10dB. The 0.95 confidence intervals on the classification accuracy of the DT and NN classifiers for the SNR range of 20dB to -5dB are shown in Table 6.5. It can be seen that the NN classifier performance is slightly better than that of the DT classifier for the SNR range of 20dB to 5dB. This is mainly because the key features have been chosen well enough so that there is minimal overlap between classes. This can be confirmed by referring to the graph of SNR versus classifier accuracy in Figure 6.14. For lower SNR, the NN performs much better than the DT approach because the NN can develop a decision boundary that is not restricted to being linear as in the DT approach. The confusion matrices showing the results of the NN classifier are shown in Appendix B, Table B.13 – Table B.18, for the range of SNR from 20dB to -5dB, respectively.

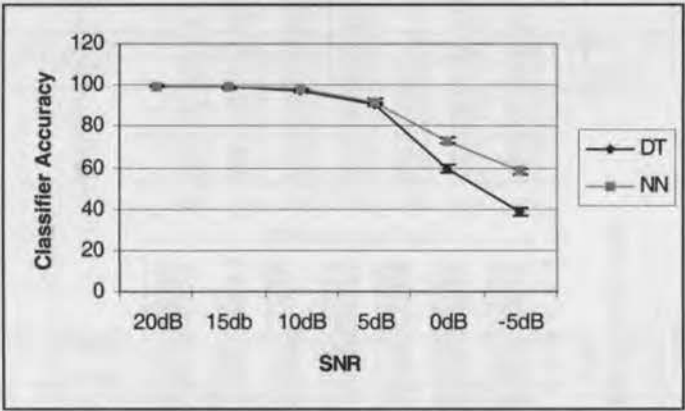


Figure 6.14. Graphical comparison of overall performance between the NN-based and DT-based classifiers with 95% CI for ASK, PSK, FSK, and CPM signals.

### 6.5.2.2 NN Classifier Results for Within CPM Class

The results of classification of CPM signals are shown in Figure 6.21 - Figure 6.26 for SNR range of 20dB to -5dB. The 95% confidence interval is shown on all figures by the error bars. Since the signals are classified as full response, partial response or GMSK, the input data is a combination of the signals from each classification type. For instance, the test data for  $L = 1$  is combined from the full response signals LREC, LRC, and HCS, and likewise for  $L = 2$ . It can be seen that the performance is good for both networks for SNR greater

than or equal to 10dB as the NNs are trained with data of SNR 20dB, 15dB, and 10dB. With lower SNR values, the performance drops, as can be expected.

Table 6.5 DT and NN classifier accuracy and 95% confidence intervals for ASK, PSK, FSK, and CPM signals.

SNR	DT Classifier		NN Classifier	
	Accuracy	95% Confidence Interval	Accuracy	95% Confidence Interval
20dB	99.57%	[99.34, 99.81]	99.5%	[99.24, 99.76]
15dB	99%	[98.63, 99.37]	99.28%	[98.97, 99.60]
10dB	97.17%	[96.56, 97.79]	97.86%	[97.32, 98.39]
5dB	90.46%	[89.38, 91.55]	92.08%	[91.08, 93.08]
0dB	59.61%	[57.79, 61.42]	73.05%	[71.41, 74.70]
-5dB	39.07%	[37.26, 40.88]	58.16%	[56.33, 59.98]
Overall	80.82%	[77.91, 79.25]	86.66%	[86.14, 87.17]

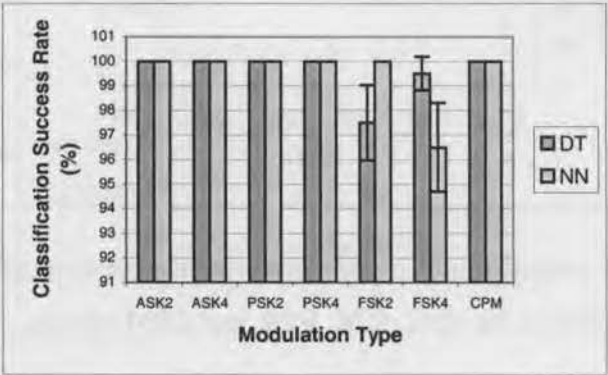


Figure 6.15. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for ASK, PSK, FSK, and CPM signals at 20dB SNR.

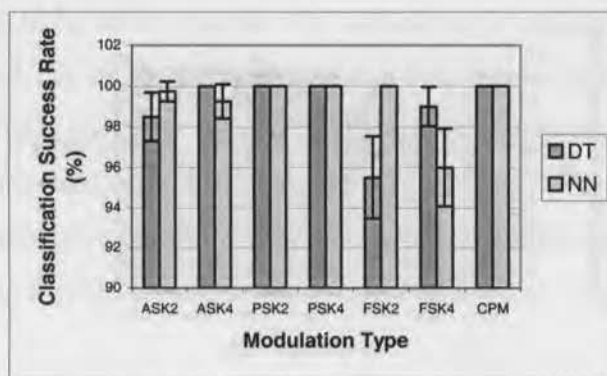


Figure 6.16. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for ASK, PSK, FSK, and CPM signals at 15dB SNR.

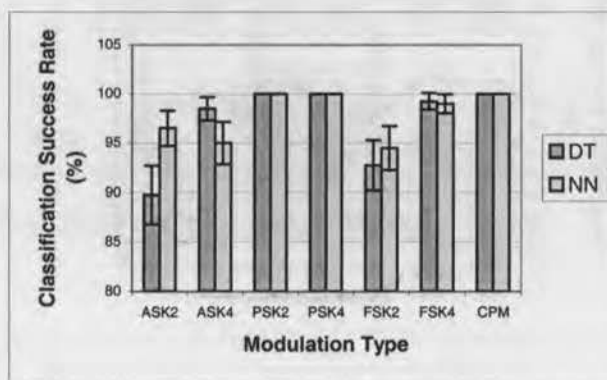


Figure 6.17. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for ASK, PSK, FSK, and CPM signals at 10dB SNR.

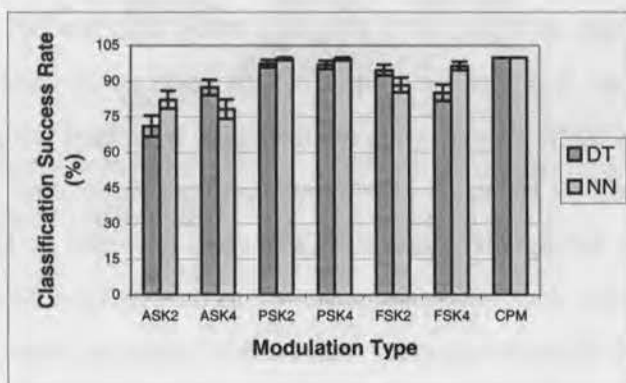


Figure 6.18. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for ASK, PSK, FSK, and CPM signals at 5dB SNR.

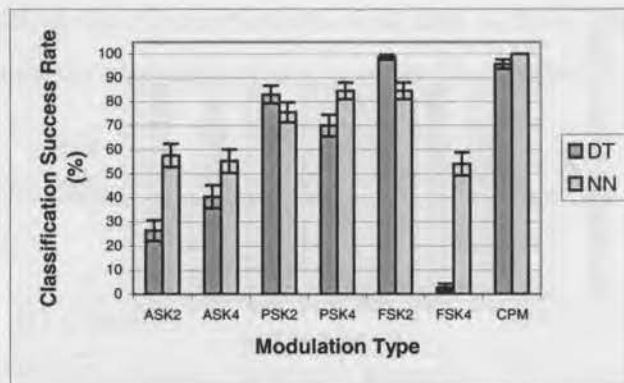


Figure 6.19. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for ASK, PSK, FSK, and CPM signals at 0dB SNR.

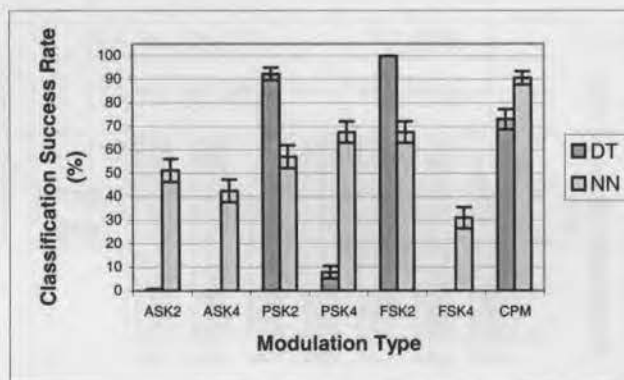


Figure 6.20. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for ASK, PSK, FSK, and CPM signals at -5dB SNR.

The second NN structure in Figure 6.13 performs better than the NN structure in Figure 6.12. This is due to the fact that the former NN is made up of three separate networks trained with individual SNR values, whereas the latter is trained with data of combined SNR ranging from 20dB to 10dB. A comparison between the overall performance of the DT and NN classifiers for signals within the CPM class is shown in Table 6.6. It can be seen that the DT classifier does not perform as well as the NN classifier. This may be due to the fact that the NN decision boundary may be non linear and based on more than one key feature, whereas the decision boundary for the DT classifier is linear and based on one key feature. However, for lower SNR values, the second NN structure outperforms both the first NN structure and the DT classifier. This is due to the fact that the signal with the

highest success rate from three separate NN structures is chosen. In other words, we choose the output of the network that gives the best results out of the three separate networks. The confusion matrices of the NN classifiers are shown in Appendix B, Table B.19 – Table B.24, for the first NN structure and, Table B.25 – Table B.30, for the second NN structure. A graphical comparison between the overall performance of the DT and NN classifiers is shown in Figure 6.27 for the SNR range of 20dB to -5dB.

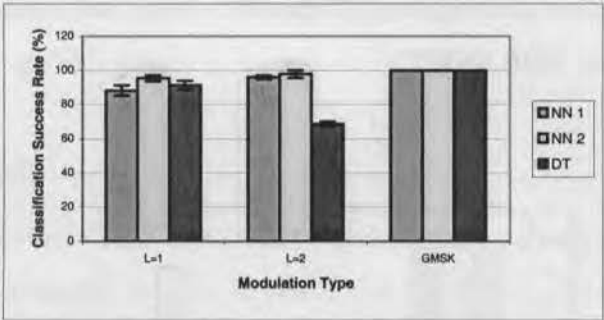


Figure 6.21. Classification accuracy of DT classifier (dark bars) and NN classifiers (light bars) for CPM signals at 20dB SNR.

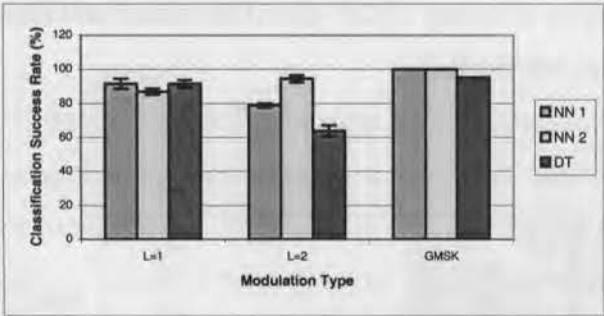


Figure 6.22. Classification accuracy of DT classifier (dark bars) and NN classifiers (light bars) for CPM signals at 15dB SNR.

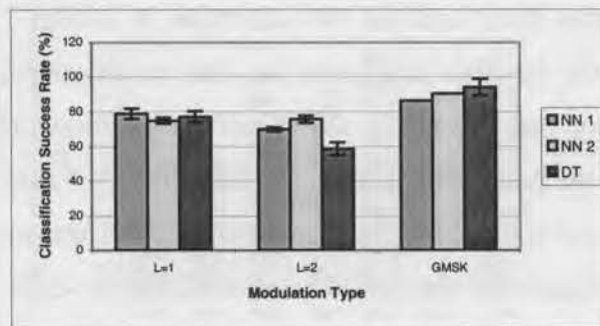


Figure 6.23. Classification accuracy of DT classifier (dark bars) and NN classifiers (light bars) for CPM signals at 10dB SNR.

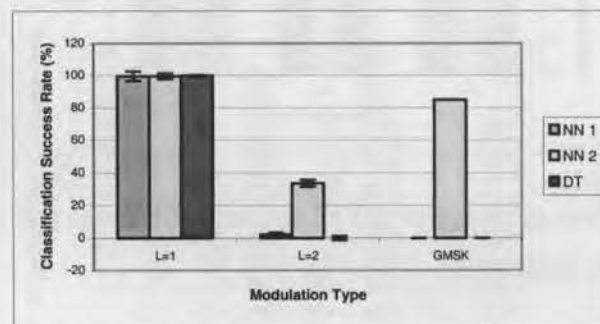


Figure 6.24. Classification accuracy of DT classifier (dark bars) and NN classifiers (light bars) for CPM signals at 5dB SNR.

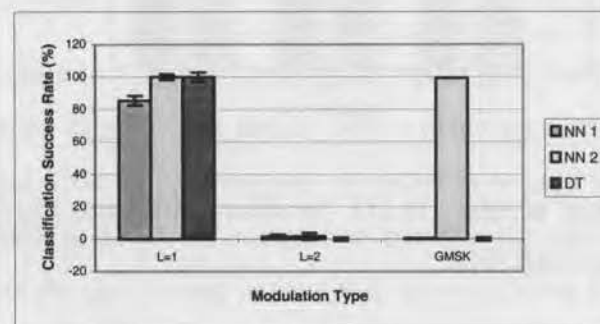


Figure 6.25. Classification accuracy of DT classifier (dark bars) and NN classifiers (light bars) for CPM signals at 0dB SNR.



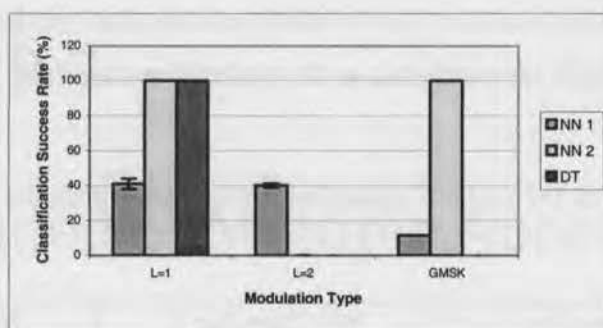


Figure 6.26. Classification accuracy of DT classifier (dark bars) and NN classifiers (light bars) for CPM signals at -5dB SNR.

## 6.6 Conclusions

In this chapter we considered a DT classifier capable of classifying ASK, FSK, PSK, and CPM signals. A NN classifier is also proposed that is capable of recognising these same signals. The performance of both classifiers are compared. Both the DT and NN modulation recognisers perform well, even for low SNR values. The NN outperforms the DT classifier due to the fact that the NN may use a non-linear decision boundary with more than one key feature for classification; whereas, the DT classifier uses one key feature per decision with a linear decision boundary.

For signals within the CPM class, by differentiating between partial response, full response and GMSK signals, the receiver structure chosen to detect the classified signal will be less complex than a receiver designed for all CPM signals. A decision tree is designed using data of SNR range 20dB to 10dB. This is because the signals within the CPM class are very similar and differentiation at lower SNR becomes very difficult. In addition to a DT approach, two NN classifiers for classification of signals within the CPM class are also proposed. The first NN is trained with data of SNR ranging from 20dB to 10dB. The second NN structure is made up of three parallel sub-networks. Each network is trained with SNR of 20dB, 15dB, and 10dB, respectively. It is found that the NN approach outperforms the DT classifier for most SNR values. This may be due to the fact that for the DT classifier, the threshold decision boundary is linear, whereas the NN classifier may

have a non-linear decision boundary that is based on more than one key feature. The next chapter introduces multiple access signals to the modulation classifier.

Table 6.6 Comparison of DT and NN classifiers for signals within the CPM class for SNR range 20dB to -5dB

SNR	DT Classifier		NN Classifier 1		NN Classifier 2	
	Accuracy	95% Confidence Interval	Accuracy	95% Confidence Interval	Accuracy	95% Confidence Interval
20dB	83.45%	[81.50,85.40]	94.82%	[93.66,95.98]	97.83%	[97.07,98.60]
15dB	83.58%	[81.64,85.52]	90.17%	[88.61,91.73]	93.83%	[92.57,95.09]
10dB	76.75%	[74.54,78.96]	78.5%	[76.35,80.65]	80.44%	[78.37,82.52]
5dB	33.33%	[30.86,35.80]	33.95%	[31.47,36.43]	72.72%	[70.39,75.06]
0dB	33.33%	[30.86,35.80]	29.17%	[26.79,31.55]	67.06%	[64.59,69.52]
-5dB	33.33%	[30.86,35.80]	30.95%	[28.53,33.37]	66.67%	[64.20,69.14]
Overall	57.30%	[54.71,59.89]	59.59%	[58.54,60.64]	79.76%	[78.90,80.62]

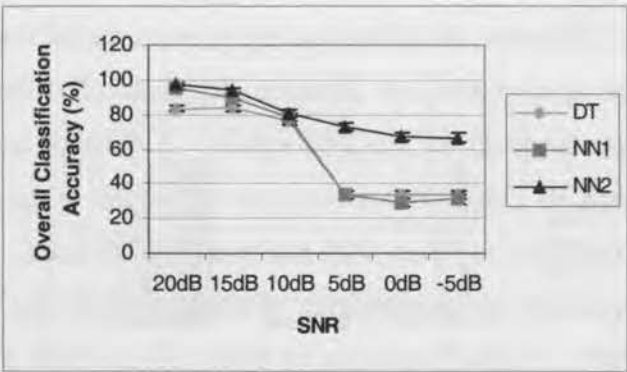


Figure 6.27. Graphical comparison of overall performance between the NN-based and DT-based classifiers (for within the CPM Class) with 95% CI.

# CHAPTER 7

## Classification of Multiple Access Signals

### 7.1 Introduction

This chapter presents an extension to the capabilities of the modulation classifiers described in Chapter 6 to include multiple access signals. These signal's modulation types are: direct sequence spread spectrum (DS SS) or code division multiple access (CDMA), frequency hopped spread spectrum (FH SS), and time division multiple access (TDMA). They are very commonly used in the military for their low probability of interception and also in the civilian areas in mobile networks to reduce call dropouts and interference. We include these different types of signals in the modulation classification algorithms, which employ the decision theoretic and neural network approaches. Results are compared and presented for SNR of 20dB, 15dB, 10dB, 5dB, 0dB, and -5dB.

The chapter is organised as follows: a brief introduction to multiple access signals is presented in section 7.2 followed by a description of the DT classification procedure in section 7.3. Threshold determination is discussed in section 7.4 and in section 7.5, a NN classifier is introduced to classify the same multiple access signals as the DT classifier. A discussion of the performance of both the DT and NN classifiers is presented in section 7.6 with results followed by concluding remarks in section 7.7.

### 7.2 Multiple Access Communication Systems

Multiple access communication systems have a large number of users sharing a common communication channel to transmit information to a receiver. The common channel may be the up-link in a satellite communication system, or some frequency band in the radio spectrum that is used by multiple users to communicate with a radio receiver.

One method for creating multiple subchannels for multiple access is to divide the time duration  $T_f$ , called the frame duration, into  $N$  nonoverlapping subintervals, each of duration  $T_f/N$ . Then to transmit information, each user is assigned to a particular time slot within each frame. This multiple access method is called *Time Division Multiple Access* (TDMA) and is commonly used in data and digital transmission. TDMA works well when the data transmitted is constant. Problems arise when the data becomes bursty. This is when there are periods of no data being transmitted and where these periods are greater than the periods of information transmission. This can be the case in a mobile cellular communications system carrying digitised voice, since speech signals contain long periods of silence. In these cases TDMA tends to be inefficient because there are wasted time slots when no data is being transmitted. This inefficiency limits the number of simultaneous users. TDMA will be discussed in more detail in section 7.2.3.

An alternative to TDMA is to allow more than one user to share a channel by using *direct-sequence spread spectrum signals* (DS-SS). It is given its name because the transmission bandwidth is much greater than the minimum bandwidth required to transmit the digital information. Each user is assigned a unique code or signature sequence that allows the user to spread the information signal across the frequency channel. The signals from various users are separated at the receiver by cross correlation of the received signal with each of the possible spreading codes. These codes are designed to have relatively small cross correlations so that there is no interference between users. This multiple access method is known as *code division multiple access* (CDMA). For a signal, to be defined as spread spectrum, the system must have the following characteristics [Peterson, 1995].

1. The transmitted signal energy must occupy a bandwidth which is larger than the information bit rate (usually much larger) and which is approximately independent of the information bit rate.
2. Demodulation must be accomplished, in part, by correlation of the received signal with a replica of the signal used in the transmitter to spread the information signal.

There are two types of SS, direct-sequence (DS) and frequency hopped (FH). These types will be described in more detail in subsections 7.2.1 and 7.2.2 respectively.

### 7.2.1 Direct Sequence Spread Spectrum (DS-SS)

The spectrum of a data-modulated signal can be spread by modulating the signal a second time by a very wideband spreading signal. The second modulation method is usually digital phase modulation. The spreading signal is chosen so that demodulation of the signal by an unintended receiver is made as hard as possible. Therefore the spreading signal is chosen specifically for the intended receiver to demodulate. Also if there is jamming, the intended receiver will still be able to discriminate between the data signal and jamming due to this property. A direct-sequence (DS) spread spectrum signal is one in which the bandwidth spreading is achieved by direct modulation of a data-modulated carrier by a wide-band spreading signal or code.

#### 7.2.1.1 Binary Phase Shift Keying Direct Sequence Spread Spectrum (BPSK DS-SS)

The simplest form of DS spread spectrum uses BPSK as the spreading modulation. The BPSK DS-SS signal can be mathematically represented as a multiplication of the carrier by a function  $c(t)$  which takes on values of  $\pm 1$ . Consider a constant envelope data-modulated signal  $s(t)$  defined by :

$$s(t) = A_c \cos[\omega_c t + \theta(t)] \quad (7.1)$$

where  $\theta(t)$  is the data phase modulation and  $\omega_c$  is the radian frequency. The bandwidth of this signal is usually between one-half and twice the data rate before DS spreading. The signal is multiplied by a function  $c(t)$  representing the spreading waveform, and the resulting transmitted waveform is:

$$s(t) = A_c c(t) \cos[\omega_c t + \theta(t)] \quad (7.2)$$

$$\theta(t) = D_p m(t) \quad (7.3)$$

where  $m(t)$  is a bipolar baseband signal having peak values of  $\pm 1$  and a rectangular pulse shape (for convenience) and  $D_p$  is the modulation index of the BPSK signal. The signal has a transmission delay  $T_d$ , is transmitted along a distortionless path, and is received with additive Gaussian noise and/or some other type of interference.

### 7.2.1.2 Spreading Codes

The waveform  $c(t)$  used to spread and despread the data-modulated carrier is usually generated using a shift register. This waveform  $c(t)$  is a pseudo-random code known as a PN sequence. This PN sequence is periodic with noise-like properties which makes the spread-spectrum signal hard to intercept. Each user in the CDMA system has a unique PN sequence assigned to them. Because users will be transmitting messages simultaneously, the PN code sequences must be mutually orthogonal so that interference from other users is avoided [Sklar, 1988]. For the spread-spectrum system to operate effectively, the PN codes  $c(t-T_d)$  must be determined initially and then tracked by the receiver. To achieve this,  $c(t)$  is chosen to have a two-valued autocorrelation function. The ideal spreading code would be an infinite sequence of equally likely random bits, however this is not possible in practice.

The most widely known PN sequences are the maximal-length shift-register sequences ( $m$ -sequences) which have a length of

$$n = 2^m - 1 \text{ bits} \quad (7.4)$$

They are generated by an  $m$ -stage shift register with linear feedback. The sequence is periodic with period  $n$  and each period of the sequence contains  $2^{m-1}-1$  zeros and  $2^{m-1}$  ones. It is desirable in a CDMA system to have a low cross-correlation between a pair of sequences. The number of  $m$  sequences generated by the shift register with low cross correlation values is too small for CDMA purposes. Therefore it has been found that Gold and Kasami sequences have better cross-correlation properties.

### 7.2.1.3 Gold and Kasami Sequences

It was found by Gold and Kasami that certain pairs of  $m$  sequences of length  $n$  have a three-valued cross correlation function with the values  $\{-1, -t(m), t(m)-2\}$  where

$$t(m) = \begin{cases} 2^{(m+1)/2} + 1 & (\text{odd } m), \\ 2^{(m+2)/2} + 1 & (\text{even } m), \end{cases} \quad (7.5)$$

For example, if  $m = 5$ , then  $t(5) = 2^3 + 1 = 9$ . The three possible values of the periodic cross-correlation function are then  $\{-1, -9, 7\}$  and the maximum magnitude of the cross-correlation for the pair of  $m$ -sequences is nine. Two  $m$ -sequences of length  $n$ , with periodic

cross-correlation taking on values of  $\{-1, -t(m), t(m)-2\}$ , are called preferred sequences. From a pair of preferred sequences where  $a = [a_1 a_2 \dots a_n]$  and  $b = [b_1 b_2 \dots b_n]$ , a sequence of length  $n$  can be constructed by taking the modulo-2 sum of  $a$  with the  $n$  cyclically shifted versions of  $b$  or vice versa. The resulting new periodic sequences have period  $n = 2^m - 1$ . By including the original sequences,  $a$  and  $b$ , we have a total of  $n + 2$  sequences called Gold sequences.

Kasami sequences have cross-correlation and autocorrelation values from the set  $\{-1, -(2^{m/2} + 1), 2^{m/2} - 1\}$ . The sequences are constructed by beginning with an  $m$ -sequence  $a$ , and forming a binary sequence  $b$  by taking every  $2^{m/2} + 1$  bit of  $a$ . This sequence,  $b$ , has a period  $n = 2^{m/2} - 1$ . Then by taking  $n = 2^m - 1$  bits of the sequences  $a$  and  $b$ , a new set of sequences is formed by modulo-2 adding the bits from  $a$  and  $b$  and all  $2^{m/2} - 2$  cyclic shifts of the bits from  $b$ . By including  $a$  in the set, a set of  $2^{m/2}$  Kasami sequences of length  $n = 2^m - 1$  is obtained.

In this thesis, Gold codes from a set of orthogonal Gold codes are used in simulations. These sequences are 7-bits in length and can accommodate up to 9 users in a CDMA scheme. The set of Gold codes is shown in Table 7.1.

The complex envelope is found by referring to the PSK2 signal and is represented by [Couch, 2001] as

$$\alpha(t) = A_c m(t) c(t) \quad (7.6)$$

The pulse width of  $c(t)$  is denoted by  $T_c$  and is called a chip interval.

The instantaneous amplitude and phase is:

$$a(t) = |m(t)c(t)| = 1 \quad (7.7)$$

$$\phi(t) = \begin{cases} -\pi/2 & \text{if } m(t)c(t) = -1 \\ \pi/2 & \text{if } m(t)c(t) = 1 \end{cases} \quad (7.8)$$

These features of BPSK DS-SS modulation are shown in Figure 7.1.

Table 7.1. CDMA 7-bit Gold code set [Ramakonar, 1996].

User, k	7-bit Gold Code Sequence
0	1 -1 1 1 1 -1 -1
1	1 1 -1 1 1 -1 1
2	-1 1 1 -1 1 -1 1
3	-1 1 -1 1 1 -1 -1
4	1 1 -1 -1 1 1 -1
5	1 -1 -1 -1 -1 -1 1
6	-1 1 1 -1 -1 1 -1
7	1 1 1 1 -1 1 1
8	-1 -1 -1 1 1 1 1

#### 7.2.1.4 Quadrature Phase Shift Keying Direct Sequence Spread Spectrum (QPSK DS-SS)

Quadrature-phase shift keying is advantageous because it allows simultaneous transmission on two carriers which are in-phase quadrature and this conserves spectrum. This means that for the same total transmitted power, the same bit error probability is achieved using one-half the transmission bandwidth. Bandwidth efficiency is not very important in low probability of detection and antijam applications. QPSK is used in spread spectrum applications due to the fact that it is less sensitive to some types of jamming and more difficult to detect using feature detectors in low probability of detection applications.

The QPSK DS-SS signal can be represented by

$$s(t) = A_c c_1(t) \cos[\omega_c t + \theta(t)] - A_c c_2(t) \sin[\omega_c t + \theta(t)] \quad (7.9)$$

$$\theta(t) = D_p m(t) \quad (7.10)$$

where  $c_1(t)$  and  $c_2(t)$  are the in-phase and quadrature spreading waveforms which are assumed only to take on values of  $\pm 1$ . The two terms of the QPSK spread-spectrum signal in equation (7.9) are identical, except for amplitude and a possible phase shift to the BPSK spread-spectrum signal in equation (7.2). Therefore, since the two signals are orthogonal, the power spectrum of the QPSK signal equals the algebraic sum of the two power spectra.

The complex envelope is given by



$$\alpha(t) = A_c m(t) c_1(t) - A_c m(t) c_2(t) \quad (7.11)$$

The instantaneous amplitude and phase are, respectively:

$$a(t) = 1 \quad (7.12)$$

$$\phi(t) = \begin{cases} 0 & \text{if } m(t)c_1(t) = -1 \\ \pi/2 & \text{if } m(t)c_1(t) = 1 \\ \pi & \text{if } m(t)c_2(t) = 1 \\ 3\pi/2 & \text{if } m(t)c_2(t) = -1 \end{cases} \quad (7.13)$$

The useful attributes of a QPSK DS-SS signal are shown in Figure 7.2.

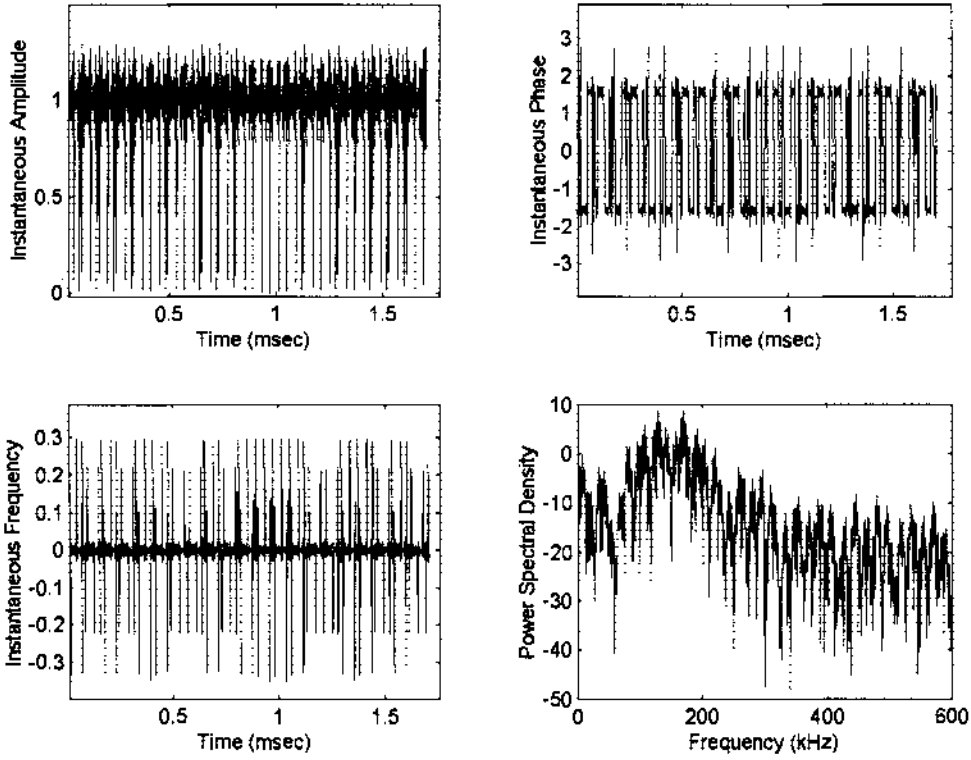


Figure 7.1. Useful features of BPSK DS-SS modulation.

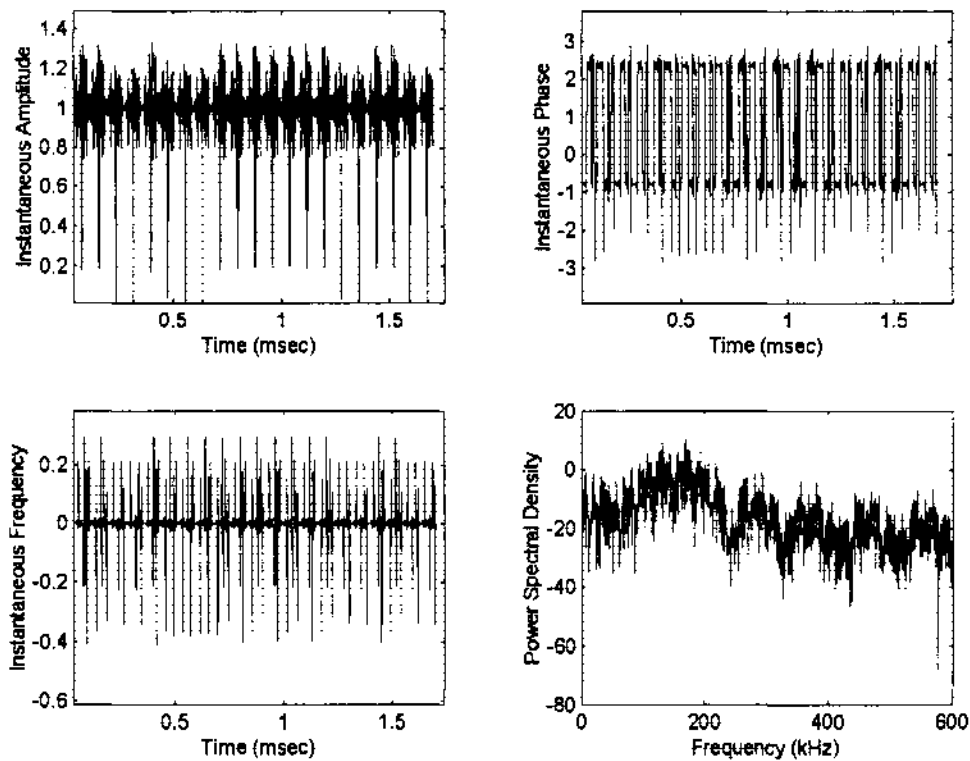


Figure 7.2. Useful features of QPSK DS-SS modulation.

### 7.2.2 Frequency Hopped Spread Spectrum (FH SS)

Another method used to widen the spectrum of the data-modulated carrier is to change the frequency of the carrier periodically. Each carrier frequency is chosen from a set of  $2^k$  frequencies which are spaced approximately one width of the data modulation bandwidth apart. The spreading code is used to control the sequence of carrier frequencies and thus does not directly modulate the data-modulated carrier. This modulation scheme is named frequency hopped (FH) spread spectrum because it appears as if the transmitted signal is hopping from one carrier frequency to another. The frequency hopping is removed in the receiver by down-converting (mixing) with a local oscillator signal which is hopping synchronously with the received signal.

### 7.2.2.1 Coherent Slow-Frequency Hop Spread Spectrum

In most cases of this type of modulation, the frequency hopping is done noncoherently. However, it is theoretically possible to have a fully coherent FH system. The frequency synthesiser output is a sequence of tones of duration  $T_c$ , so  $h_T(t)$  can be written as

$$h_T(t) = \sum_{n=-\infty}^{\infty} 2p(t - nT_c) \cos(\omega_n t + \phi_n) \quad (7.14)$$

where  $p(t)$  is a unit amplitude pulse of duration  $T_c$ , starting at time zero, and  $\omega_n$  and  $\phi_n$  are the radian frequency and phase during the  $n$ th frequency-hop interval. The radian frequency  $\omega_n$  is taken from a set of  $2^k$  frequencies. In a DS spread spectrum system, the spreading sequence was used one bit at a time. In contrast the FH system uses  $k$  bits of the spreading code at a time. The transmitted signal is the data-modulated carrier up-converted to a new frequency ( $\omega_b + \omega_n$ ) for each FH chip and is represented as

$$s_t(t) = \left[ s_d(t) \sum_{n=-\infty}^{+\infty} 2p(t - nT_c) \cos(\omega_n t + \phi_n) \right]_{\substack{\text{sum freq.} \\ \text{components}}} \quad (7.15)$$

The complex envelope is denoted as [Couch, 2001]:

$$\alpha(t) = \alpha_m(t) \alpha_c(t) \quad (7.16)$$

where  $\alpha_m(t)$  is the complex envelope of the information signal and  $\alpha_c(t)$  is of FM type where there are  $M = 2^k$  hop frequencies determined by the  $k$ -bit words obtained from the spreading code waveform  $c(t)$ .

The useful features of FH SS modulation are shown in Figure 7.3.

### 7.2.3 Time Division Multiple Access (TDMA)

In TDMA,  $M$  signals or users share the same frequency channel for a short duration of time called a time slot as shown in Figure 7.4. Sometimes unused time regions are inserted between adjacent slot assignments to allow for time uncertainty between signals. These time regions are called guard times and act as buffer zones to reduce interference. In a typical TDMA satellite application, time is segmented into intervals called frames. Each frame is further partitioned into time slots which are assigned to each user. The frame

structure repeats so that a fixed TDMA assignment constitutes one or more slots that periodically appear during each frame time. Some useful features of TDMA are shown in Figure 7.5.

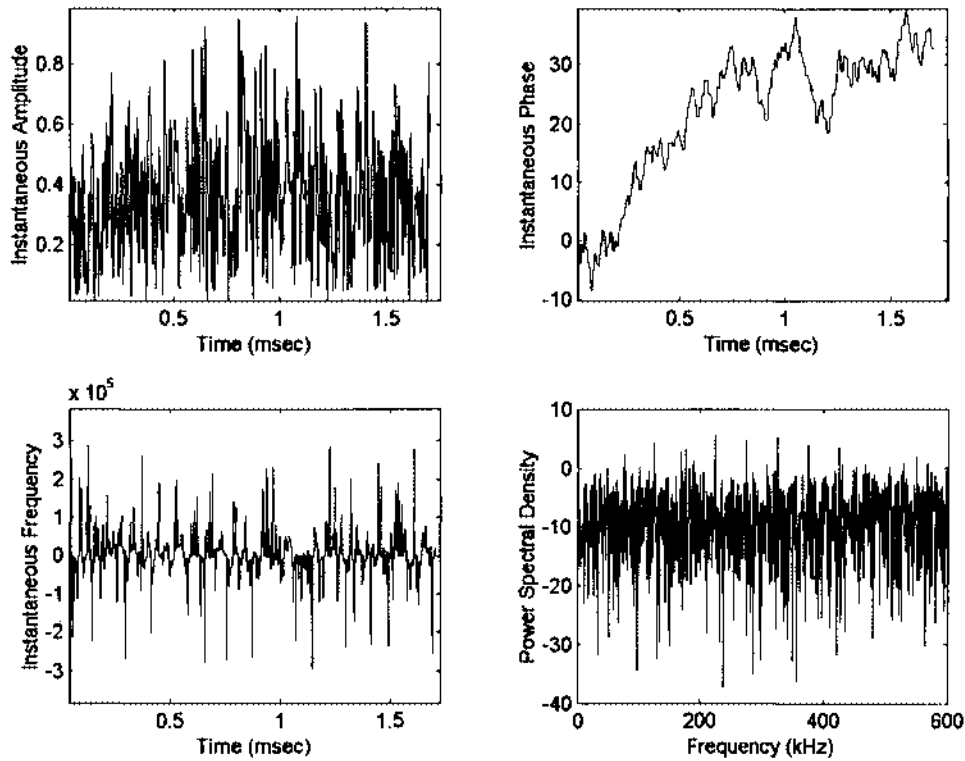


Figure 7.3. Useful features of FH SS modulation.

### 7.3 Classification Procedure (DT Approach)

The procedure for digital signal classification is based on the method outlined in Chapter 5 and Chapter 6. Key features are derived from the power spectral density and the instantaneous frequency of the intercepted signal. The following signals are added to the modulation classifier:

- BPSK DS-SS
- QPSK DS-SS
- FH SS
- TDMA

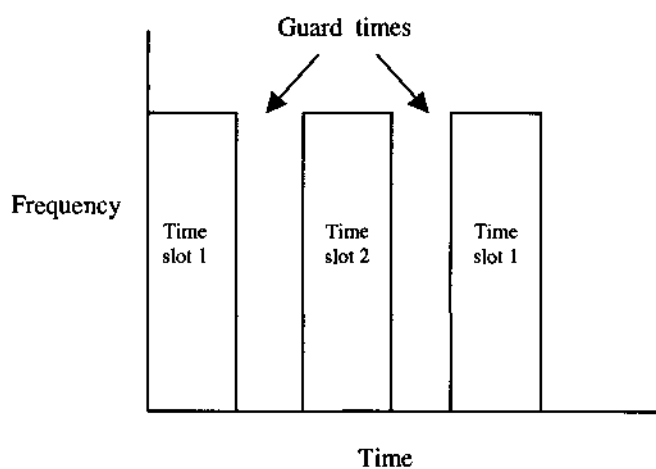


Figure 7.4. Time division multiple access (TDMA).

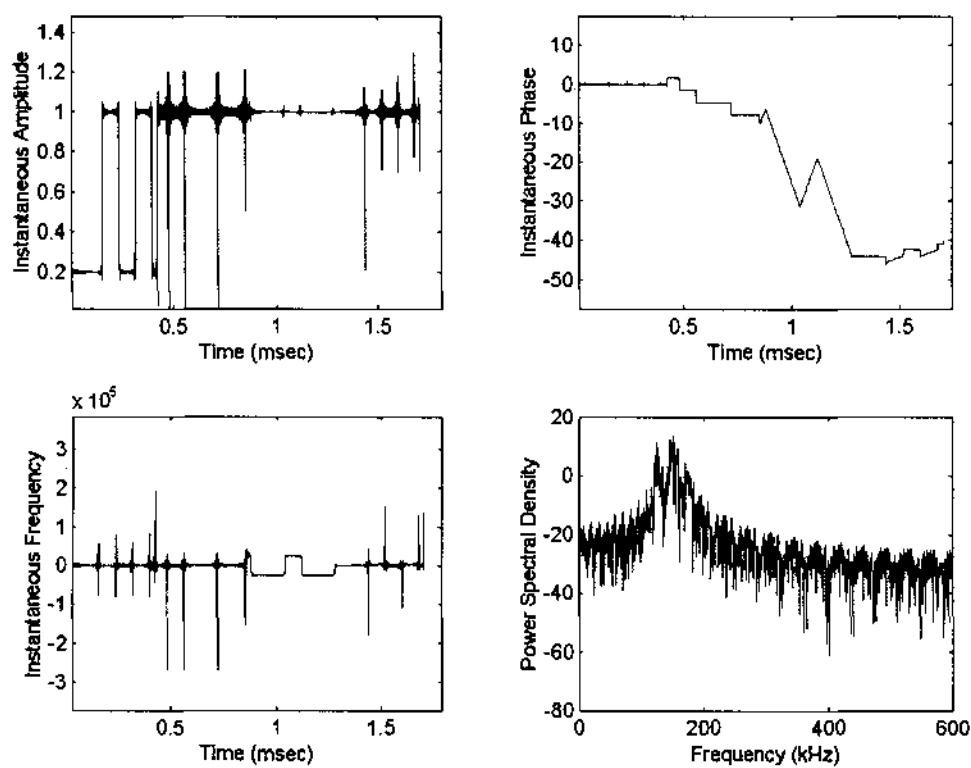


Figure 7.5. Useful features of TDMA modulation.

### 7.3.1 Key Feature Derivation for Signal Classification

To derive the appropriate key features for signal classification, the same method in Chapter 6 is used. The new signals are passed through the existing classifier and each signal is classified as a modulation type already defined in the tree. To find the actual modulation type of a particular signal, a decision node is added to the tree to distinguish between the modulation type that the signal is classified as and the actual modulation type of the new signal. We discuss BPSK DS-SS classification in subsection 7.3.1.1 and QPSK DS-SS classification in subsection 7.3.1.2. The classification of FH SS and TDMA signals are discussed in subsections 7.3.1.3 and 7.3.1.4 respectively. The decision tree depicting the classification procedure is shown in

Figure 7.7

#### 7.3.1.1 BPSK DS-SS Signal Classification

To derive the appropriate key feature for the classification of the BPSK DS-SS signal, the existing tree in Chapter 6 is utilised. The BPSK DS-SS signal is classified as a PSK2 signal, therefore we know the decision has to be made between the BPSK DS-SS signal and a PSK2 signal. By observing the smoothed power spectral densities of both signals in Figure 5.3 and Figure 7.1, it can be seen that the power of the BPSK DS-SS signal is spread due to the addition of the spreading sequence. In contrast, the PSK2 signal has most of the power centered around the carrier frequency. Figure 7.6 shows the smoothed PSD for one signal segment for PSK2 and BPSK DS-SS signals. It can be seen that for the PSK2 signal, the power drops off dramatically at frequencies further from the carrier frequency. However for the BPSK DS-SS signal, this degradation is not so steep because of the addition of more frequencies by the spreading sequence.

Therefore to separate these two types of signals, a new key feature  $\gamma_{min}$  is introduced which is the minimum value of the smoothed power spectral density (PSD) and is defined as:

$$\gamma_{min} = 10 \log_{10} (\min |DFT(s(t))|^2) \quad (7.17)$$

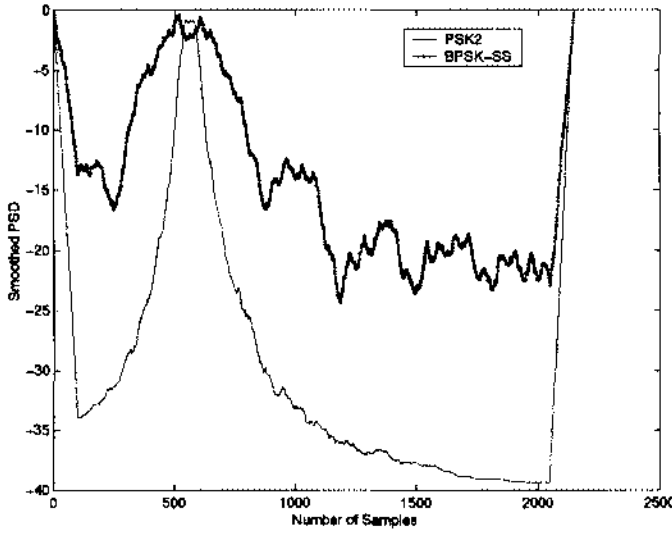


Figure 7.6. Smoothed power spectral density for PSK2 and BPSK DS-SS signals.

#### 7.3.1.2 QPSK DS-SS Signal Classification

The resulting tree after the addition of the BPSK DS-SS signal is used to derive the appropriate key feature for the classification of the QPSK DS-SS signal. The signal is classified as a BPSK DS-SS signal, therefore we know the decision has to be made between the QPSK DS-SS signal and a BPSK DS-SS signal. By observing the instantaneous phase of both signals in Figure 7.1 and Figure 7.2, it is found that the QPSK DS-SS signal has a slightly larger range of phase values. Therefore, a feature based on instantaneous phase is a logical choice. The key feature chosen to differentiate between BPSK DS-SS and QPSK DS-SS is  $\sigma_{ap}$ . This feature is defined in [Azzouz and Nandi, 1996] as the standard deviation of the absolute value of the non-linear component of the instantaneous phase, evaluated over the non-weak segments of the received signal. It is found that the QPSK DS-SS signal has higher standard deviation values due to the larger range of the instantaneous phase.

#### 7.3.1.3 FH SS Signal Classification

To derive the appropriate key feature for classification, we exploit the fact that the FH SS signal has frequency information. By inspecting the decision tree, we intuitively know that

the FH SS signal will lie in the right hand segment of the tree with the FSK and CPM signals. By using information found previously by classifying BPSK DS-SS and QPSK DS-SS signals we can predict that a decision will be made between the FH SS signal and the other signals with frequency information. By observing the smoothed power spectral densities of these signals in Figures 5.5, 5.6, 6.1, and 7.3, it can be seen that FH SS signals also have a greater power spread than FSK and CPM signals due to the addition of the spreading sequence. Therefore the key feature  $\gamma_{min}$  is also used to classify the FH SS signal.

#### 7.3.1.4 TDMA Signal Classification

To derive the appropriate key feature for classification, the TDMA signal is classified by the resulting tree after the addition of the FH SS signal. The TDMA signal is classified as a FH SS signal, therefore we know the decision has to be made between the TDMA signal and a FH SS signal.

It is found that the values of  $|\hat{C}_{21}|$  are higher for TDMA signals than for FH SS signals.

This is probably due to the fact that the TDMA signal is made up of a mixture of signals and there is no spreading sequence used in the modulation process. Therefore the key feature  $|\hat{C}_{21}|$  is used to separate TDMA and FH SS signals.

### 7.4 Threshold Determination

The same method in Chapter 3 is used to determine the thresholds  $t_{\gamma_{min1}}$ ,  $t_{\gamma_{min2}}$ ,  $t_{|\hat{C}_{21}|_2}$  and  $t_{\sigma_{ap}}$ . The key feature thresholds are chosen so that the probability of a correct decision is obtained from 400 realisations of each modulation type at the signal to noise ratio (SNR) range of 20dB to -5dB. A set of modulation types is separated into two non-overlapping subsets (A and B). The optimum threshold is chosen such that the Bayes error is minimised as described in Chapter 3. The total error probability  $\epsilon$ , is estimated directly from the sample data.



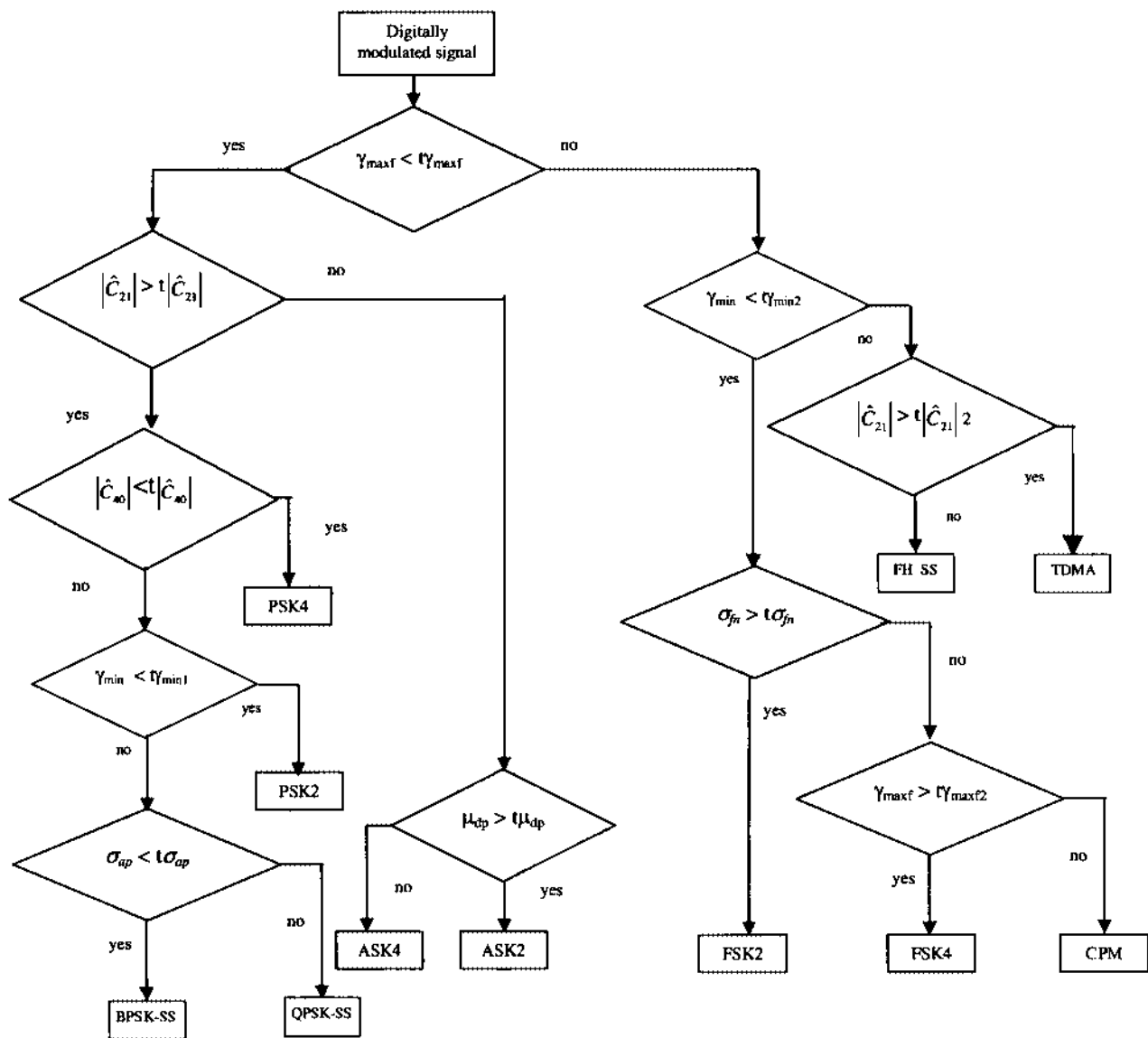


Figure 7.7. Flowchart for identification of digital modulation schemes. The first decision in the tree splits the modulation types into two groups: signals with frequency information (right hand side of tree) and signals without frequency information (left hand side of tree). The signals with frequency information are further split into multiple access signals (FH SS and TDMA) and FSK/CPM signals. The signals without frequency information are divided into signals with amplitude information (ASK) and signals with phase information (PSK, QPSK DS-SS and BPSK DS-SS). The signals with phase information are further divided into multiple access signals (BPSK DS-SS and QPSK DS-SS) and PSK signals.

The total error probability for the key feature  $\gamma_{min}$  at SNR range of 20dB to -5dB is shown in Figure 7.8 for subset A (PSK2) and subset B (BPSK DS-SS and QPSK DS-SS). It can be seen that a good choice for the threshold  $t\gamma_{min1}$  is -30.5 where the total minimum error is 0.0027 for the SNR range of 20dB to -5dB.

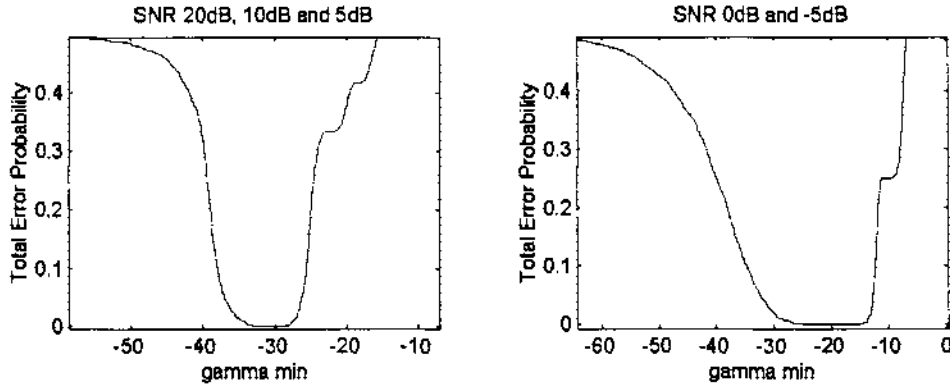


Figure 7.8. Total error probability for the key feature  $\gamma_{min}$ , at SNR range of 20dB to -5dB, for PSK2 (subset A) and BPSK DS-SS and QPSK DS-SS (subset B).

The total error probability for the key feature  $\sigma_{ap}$  for the SNR range of 20dB to -5dB is shown in Figure 7.9 for subset A (BPSK DS-SS) and subset B (QPSK DS-SS). An appropriate choice for the threshold  $t\sigma_{ap}$  is 0.76 where the total minimum error is 0 for the SNR range of 20dB to 5dB and 0.00062 for SNR range of 0dB to -5dB at the same threshold value.

To separate subset A (FSK2, FSK4, CPM) and subset B (FH SS and TDMA) using the key feature  $\gamma_{min}$ , the threshold is found by referring to the error probabilities in Figure 7.10 for the SNR range of 20dB to -5dB. It can be seen that a good choice for the threshold  $t\gamma_{min2}$  is -32.3, where the total minimum error for the SNR range of 20dB to -5dB is 0.0158.

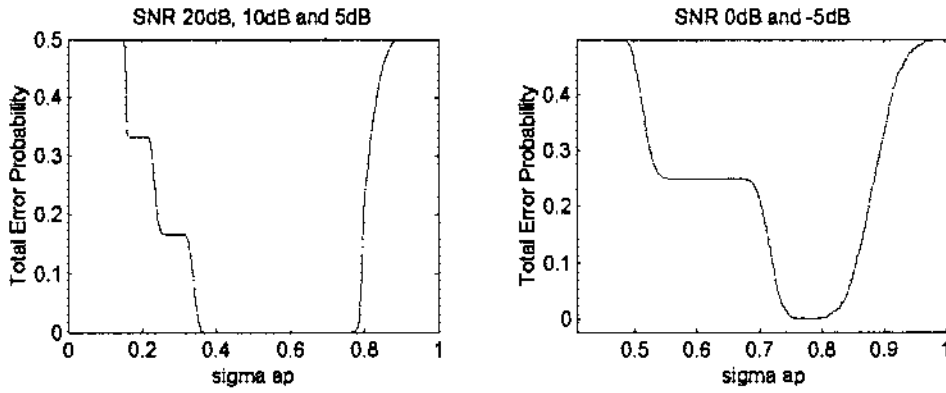


Figure 7.9. Total error probability for the key feature  $\sigma_{ap}$ , at SNR range of 20dB to -5dB, for BPSK DS-SS (subset A) and QPSK DS-SS (subset B).

The total error probability for the key feature  $|\hat{C}_{21}|$  is shown in Figure 7.11 for subset A (TDMA) and subset B (FH SS). It can be seen that a good choice for the threshold  $t|\hat{C}_{21}|_2$  is 0.5 where the total minimum error is 0 for the SNR range of 20dB to -5dB.

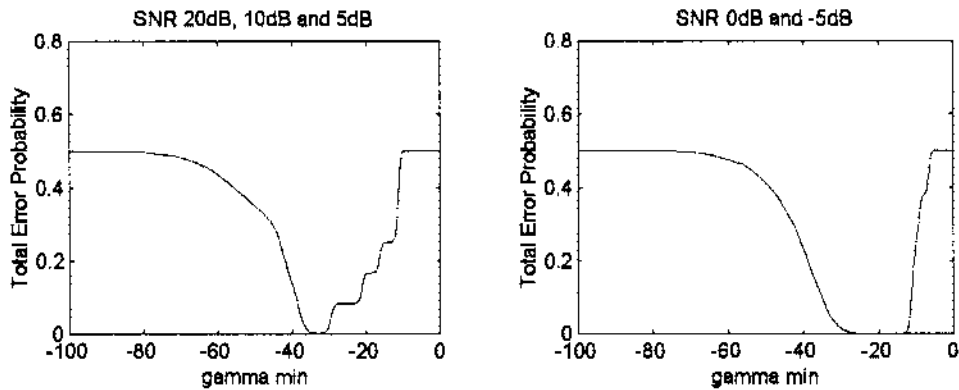


Figure 7.10. Total error probability for the key feature  $\gamma_{min}$ , at SNR range of 20dB to -5dB, for FSK2, FSK4, CPM (subset A) and FH SS and TDMA (subset B).

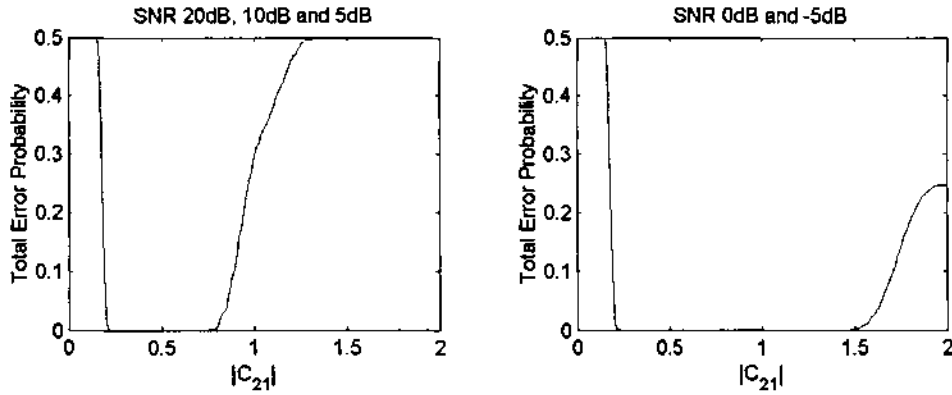


Figure 7.11. Total error probability for the key feature  $\left| \hat{C}_{21} \right|$ , at SNR range of 20dB to -5dB, for TDMA (subset A) and FH SS (subset B).

A summary of the key feature thresholds and their corresponding error probabilities for the SNR range of 20dB to -5dB is shown in Table 7.2. A compromise must be made between the threshold values at higher and lower SNR. The threshold must be chosen so that the overall classification error is minimised. Therefore, the optimum values for the key feature thresholds  $t\gamma_{min1}$ ,  $t\gamma_{min2}$ ,  $t\left| \hat{C}_{21} \right|_2$  and  $t\sigma_{ap}$  are -30.5, -33.3, 0.5 and 0.76, respectively.

Table 7.2. Summary of key feature thresholds and error probabilities.

Key Feature Threshold	SNR 20dB to 5dB		SNR 0dB to -5dB	
	Optimum Threshold	Minimum Error Probability	Optimum Threshold	Minimum Error Probability
$t\gamma_{min1}$	-30.5	0	-24.5	0
$t\gamma_{min2}$	-33.3	0	-25	0
$t\left  \hat{C}_{21} \right _2$	0.5	0	0.5	0
$t\sigma_{ap}$	0.76	0	0.76	0.00062

### 7.4.1 Dependency of Key Feature Selection on Minimum Probability of Error

The reason why the key features in the previous section are chosen over the other existing key features is because they minimise the total error probability for each decision. We will call the decision separating PSK2 (subset A) and BPSK DS-SS and QPSK DS-SS (subset B) decision 1. Decision 2 separates subset A (FSK2, FSK4, CPM) and subset B (FH SS) and decision 3 distinguishes TDMA (subset A) from FH SS (subset B). Finally, we define decision 4 as the classification of BPSK DS-SS (subset A) and QPSK DS-SS (subset B).

We can see from Table 7.3 that the key features that have been chosen minimise the total error probability (shown in bold) for each decision for the SNR range of 20dB to -5dB. The structure of the NN classifier is discussed in the next section.

Table 7.3. Total minimum error probability for Decisions 1 – 4 for combined SNR range of 20dB to -5dB (threshold values are shown in brackets).

Key Feature	Total Minimum Error Probability (Decision 1)	Total Minimum Error Probability (Decision 2)	Total Minimum Error Probability (Decision 3)	Total Minimum Error Probability (Decision 4)
$\gamma_{maxf}$	0.3333 (-85.7)	0.2500 (3.2)	0.0286 (19.3)	0.2140 (-24.28)
$\mu_{dp}$	0.2000 (0.8)	0.2500 (100)	0.1393 (37)	0.0313 (0.1)
$ \hat{C}_{21} $	0.3450 (1.76)	0.2020 (0.93)	<b>0 (0.5)</b>	0.4640 (3.1)
$ \hat{C}_{40} $	0.3335 (0.085)	0.2500 (0)	0.0085 (0.0266)	0.3420 (1.69)
$\sigma_{fn}$	0.3333 (0.1)	0.4000 (0.8)	0.1000 (1.8)	0.5000 (0)
$\sigma_{ap}$	0.0196 (0.95)	0.2500 (0)	0.2000 (14.8)	<b>0.0002 (0.76)</b>
$\gamma_{min}$	<b>0.0027 (-30.5)</b>	<b>0.0158 (-32.3)</b>	0.2000 (-13)	0.0950 (-24.3)

### 7.5 Neural Network Classifier

A neural network classifier is proposed that is based on the DT classifier described in section 7.3. This NN classifier is capable of recognising the same twelve signals (ASK2,

ASK4, PSK2, PSK4, FSK2, FSK4, CPM, BPSK DS-SS, QPSK DS-SS, FH SS, and TDMA) that are discriminated by the DT classifier. The same key features used in the decision-theoretic algorithm are used as inputs to the NN algorithm. These key features are  $\sigma_{ap}$ ,  $\gamma_{\max f}$ ,  $\mu_{dp}$ ,  $|\hat{C}_{21}|$ ,  $\gamma_{\min}$ ,  $\sigma_{fn}$ , and  $|\hat{C}_{40}|$ .

The key features are normalised to the range  $-1$  to  $1$ , then passed to the neural network. The NN structure will be described in subsection 7.5.1 and the training of the network is discussed in subsection 7.5.2.

### 7.5.1 Neural Network Structure

The neural network is a hierarchical structure based on the decision tree in Figure 7.7. It is found that this hierarchical structure results in better performance because it is made up of smaller networks. This is in contrast to one large network that is higher in complexity and takes longer to train. The accuracy of the classification will be poorer because the NN will have to classify all twelve signals at the same time. Smaller networks, however, have less output neurons and therefore generally perform better because the probability of discrimination is higher with a smaller number of signals.

The first network separates the signals with frequency information (ASK2, ASK4, PSK2, PSK4, BPSK DS-SS, and QPSK DS-SS) from those signals that do not possess any frequency information (FSK2, FSK4, CPM, FH SS, and TDMA). There are two inputs corresponding to the two key features  $\gamma_{\max f}$  and  $\gamma_{\min}$  and two output neurons assigned to the two sets of signals. Three network structures are tested with the simplest structure having one hidden layer consisting of two neurons. The performance of this network is good but the second network gives better results. The latter has two hidden layers with four neurons in each layer. However, a third tested structure is chosen as the optimum network for its simplicity as well as superior performance. This structure has one hidden layer comprising four neurons and performs as well as the more complex structure with two hidden layers.

The second network classifies ASK, PSK2, PSK4, BPSK DS-SS, and QPSK DS-SS signals. This network has five input neurons corresponding to the key features  $\sigma_{ap}$ ,  $\mu_{dp}$ ,

$|\hat{C}_{21}|$ ,  $\gamma_{\min}$ , and  $|\hat{C}_{40}|$ . There are also five output neurons representing the five signal types.

Two neural network structures are tested with the first structure having two hidden layers with four neurons in each layer. The performance of this network is mediocre. The second network structure that is tested has good performance and consists of one hidden layer with fifteen neurons. This network is chosen for its better results and twenty versions of this structure are tested to find the one that gives the optimum performance.

The third network in the hierarchy has two inputs corresponding to the key features  $\gamma_{\max f}$  and  $\gamma_{\min}$  and three output neurons corresponding to the remaining five signals: FSK/CPM, FH SS, and TDMA. The network structure that is chosen has one hidden layer with seven neurons. Twenty versions of this structure are tested to find the optimum performance.

For the classification of ASK2 and ASK4, the chosen network structure has one input corresponding to the key feature  $\mu_{dp}$ , and two output neurons corresponding to ASK2 and ASK4 signals. There is one hidden layer with ten neurons, and twenty versions of this network structure are tested to find the optimum performance.

The network to classify FSK2, FSK4, and CPM has two inputs corresponding to the features  $\sigma_{fh}$  and  $\gamma_{\max f}$  and three output neurons corresponding to the three types of signals. There is one hidden layer with twelve neurons, and twenty versions of this network structure are tested to find the optimum performance.

The hidden layers in all network structures use the nonlinear *tan-sigmoid* (hyperbolic tangent) activation function and the output layer uses the *log-sigmoid* activation function as explained in the previous chapters.

The full network structure is shown in Figure 7.12. In general it is found that the smaller structures are the optimum choice for the following reasons [Arulampalam, 1999]:

- The small structures are the least complex and therefore are the fastest to train since they contain the least number of synapses.

- Smaller structures also minimize the danger of overfitting and loss of generalization ability since they have the least “memory”.
- The larger networks have lower success rate due to their poorer generalization ability.

These reasons affirm that the hierarchical structure is the best choice for the neural network implementation of the modulation classifier.

### **7.5.2 Training the Network**

The same procedure used in the previous chapters is implemented to train the networks. The Levenberg-Marquardt (LM) algorithm using 200 samples from each modulation type is applied and the network is also tested and validated using a separate set of 200 samples from each modulation type. Training is carried out with data of SNR range 20dB to -5dB.

## **7.6 Performance Analysis**

The performance results are derived from 200 realisations of each modulation type. The carrier frequency, sampling rate and the symbol rate are given values of 150kHz, 1200kHz and 12.5kHz, respectively. The digital symbol sequence is randomly generated and the first Gold code sequence in Table 7.1 is used as the spreading sequence. The TDMA signal consists of an ASK2 signal, a PSK2 signal, and FSK2 signal and an MSK signal. Each signal has duration of 512 samples per frame and each frame is 2048 samples long. The DT classifier results are discussed in subsection 7.6.1. The NN performance is discussed in subsection 7.6.2 and a comparison with the DT classifier is included.

### **7.6.1 DT Classifier Results**

The simulation results for the test set for the modulation recogniser based on 200 realisations are shown in Figure 7.13 - Figure 7.18, for SNR 20dB to -5dB, respectively. The results of the NN classifier presented in the next section are also shown for comparison as well as the 95% confidence interval. The confusion matrices for the DT classifier are presented in Appendix C, Table C.1 – Table C.6. These results indicate that all types of the digital modulation schemes considered can be correctly classified with more than 98% success rate for SNR greater than or equal to 10dB. Seven of the eleven signals can be correctly classified with nearly 100% accuracy even at SNR of 5dB, however the



performance drops for SNR values of 0dB and -5dB as can be expected. Despite the drop in performance for lower SNR, the accuracy is still greater than 50%.

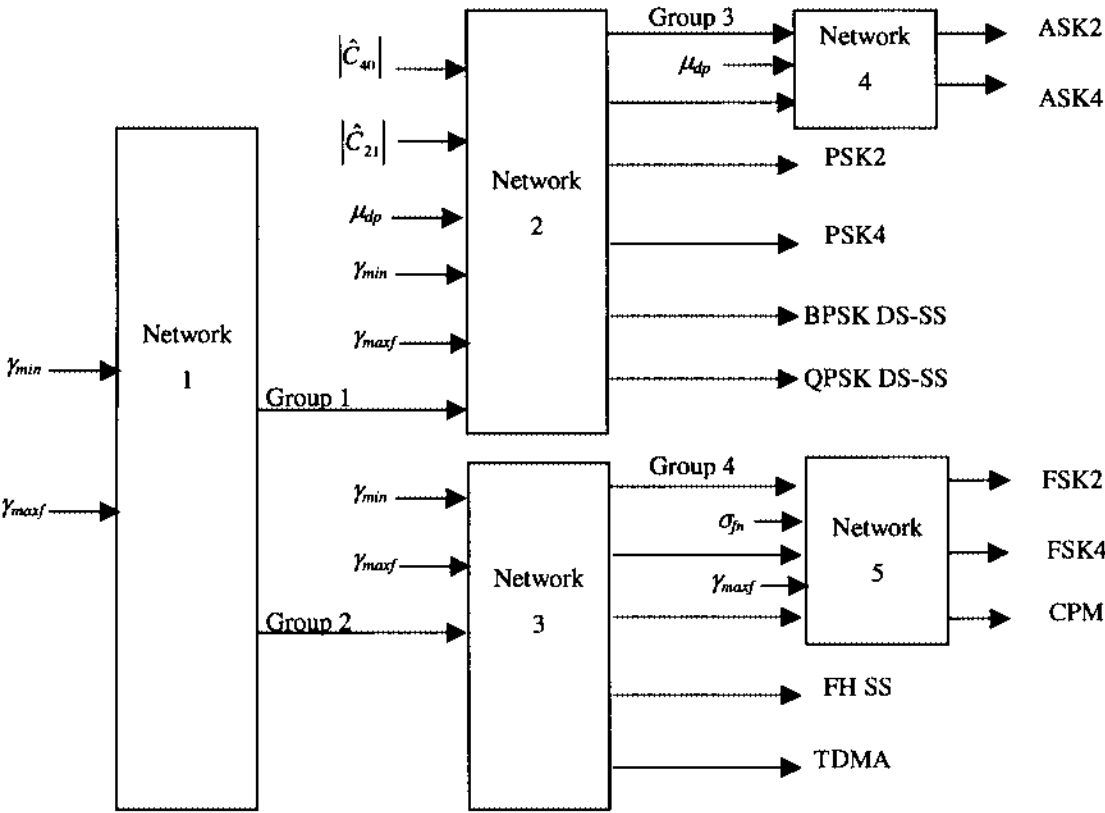


Figure 7.12. Neural network structure for modulation classifier.

### 7.6.2 Neural Network Classifier Results

The results outlining the NN and DT classifier performances are shown in Figure 7.13 - Figure 7.18 for SNR of 20dB to -5dB, respectively. It can be observed that the performances of both classifiers are very good for SNR greater than or equal to 10dB. For SNR of 5dB, the performance drops a little, but is still very good. For lower SNR, the performance drops more and it can be seen that the NN outperforms the DT classifier for SNR of 0dB and -5dB. This is probably because the NN can derive a non-linear decision boundary with many key features whereas the DT classifier is restricted by a linear decision

boundary with one key feature per decision. The confusion matrices are shown in Appendix C, Table C.7 – Table C.12, for the SNR range of 20dB to -5dB inclusive.

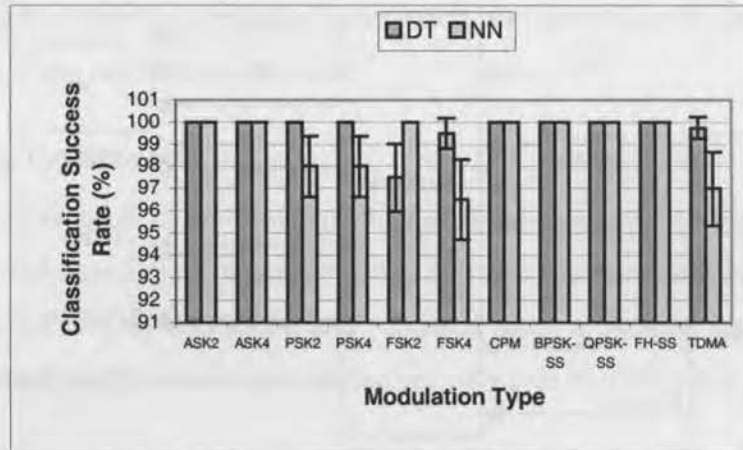


Figure 7.13. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 20dB SNR.

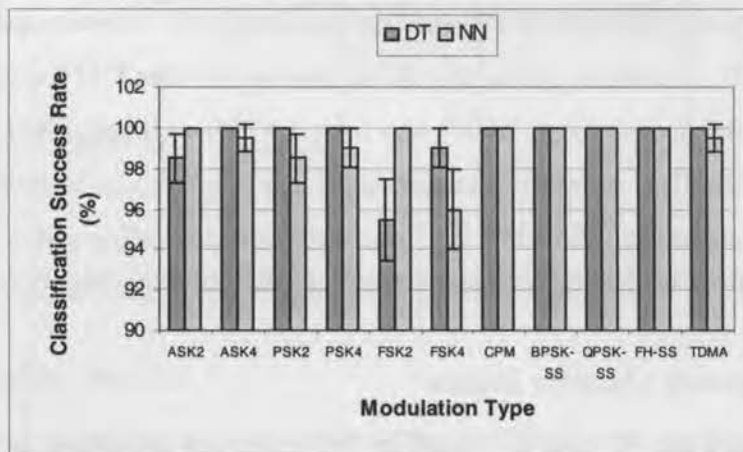


Figure 7.14. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 15dB SNR.

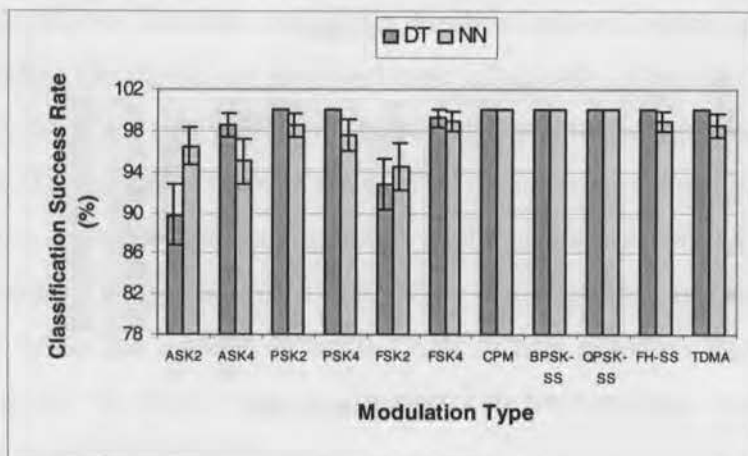


Figure 7.15. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 10dB SNR.

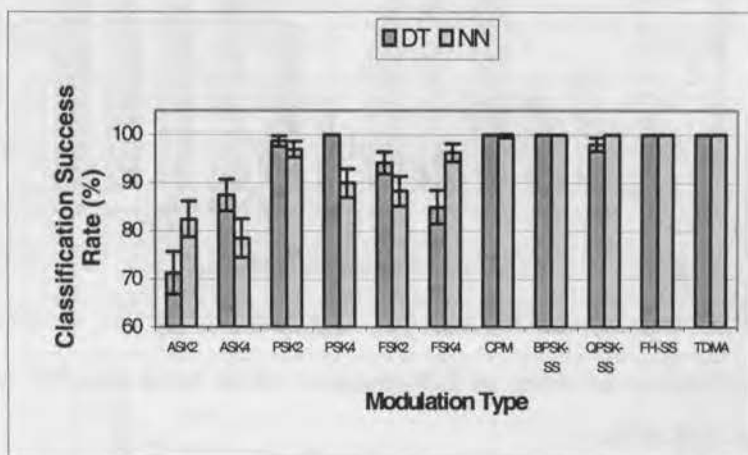


Figure 7.16. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 5dB SNR.

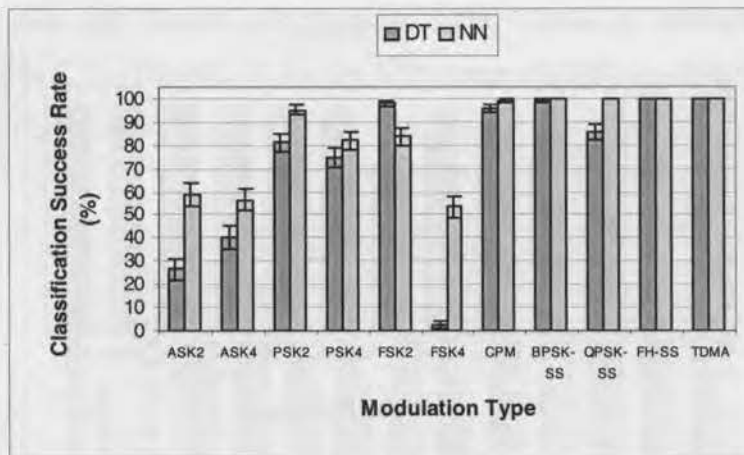


Figure 7.17. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 0dB SNR.

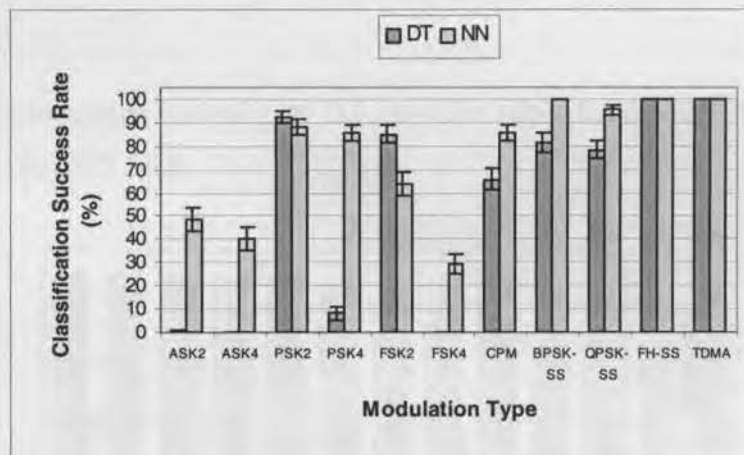


Figure 7.18. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at -5dB SNR.

A comparison of the overall success rates for the DT and NN classifiers is shown in Table 7.4, including the 95% confidence intervals. A graphical comparison of the overall success rates for the DT and NN classifiers is shown in Figure 7.19 which also has the 95% confidence intervals included. It can be inferred that the NN performance is generally similar to the performance of the DT algorithm for SNR greater than or equal to 5dB. For lower SNR, the NN classifier outperforms the DT classifier considerably. This may be due to the fact that the DT approach has hard decisions, meaning the thresholds are linear. On

the other hand, the NN classifier may have threshold regions which are not necessarily linear and therefore the signals are separated more effectively. Once the key features have been identified, the NN is able to learn the classifications directly from the training data. In contrast to the DT approach, there is no need to determine a classification algorithm or threshold values. The hierarchical approach to the neural network structure allows the formation of smaller networks, which have faster training times because the number of output classes within the network is small. This in turn produces higher success rates, which indicates that the neural network approach can accommodate even more signals if necessary without sacrificing performance.

Table 7.4. DT and NN classifier accuracy and 95% confidence intervals.

SNR	DT Classifier		NN Classifier	
	Accuracy	95% Confidence Interval	Accuracy	95% Confidence Interval
20dB	99.70%	[99.54, 99.86]	99.05%	[98.76, 99.33]
15dB	99.36%	[99.13, 99.60]	99.32%	[99.08, 99.56]
10dB	98.20%	[97.81, 98.60]	98.05%	[97.64, 98.45]
5dB	94.07%	[93.37, 94.77]	93.87%	[93.16, 94.57]
0dB	73.14%	[71.83, 74.45]	84.43%	[83.35, 85.49]
-5dB	55.59%	[54.12, 57.06]	76.08%	[74.82, 77.34]
Overall	86.68%	[86.27, 87.09]	91.80%	[91.46, 92.13]

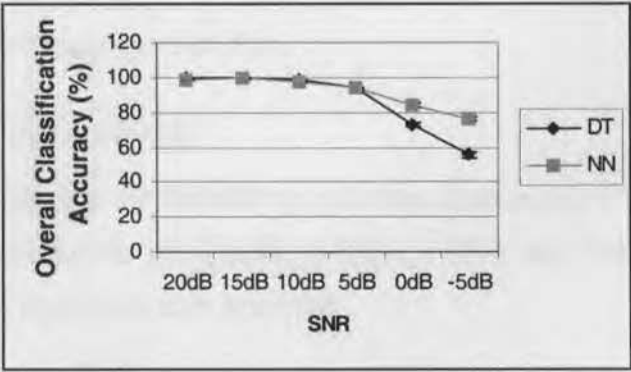


Figure 7.19. The overall classification accuracy of the NN and DT classifiers versus the SNR.

## 7.7 Conclusions

In this chapter, multiple access signals have been introduced and included as part of the modulation classifiers' recognisable signals. The multiple access signals used were BPSK DS-SS, QPSK DS-SS, FH SS, and TDMA. A new key feature,  $\gamma_{\min}$  was introduced and this particular key feature was used to identify the BPSK DS-SS, QPSK DS-SS, FH SS, and TDMA signals. The QPSK DS-SS signal was differentiated from the BPSK DS-SS signal using the key feature  $\sigma_{ap}$ , which was first introduced in [Azzouz and Nandi, 1996]. Suitable threshold values were calculated for the DT classifier and the results presented showed that the spread spectrum signals could be classified with approximately 100% success rate even at SNR as low as 5dB. The NN classifier was based on a hierarchical structure, which was found to give better results because the networks were smaller and gave better accuracy. The results of the DT and NN classifiers were compared and it was found that both classifiers performed comparatively equally, except for SNR below 5dB where the NN outperformed the DT classifier. This was possibly due to the NN's better generalisation capabilities and non-linear decision boundaries.

# CHAPTER 8

## Classification of PSK8, FSK8 and QAM Signals

### 8.1 Introduction

In this chapter, PSK8, FSK8, QAM8, and QAM16 signals are added to the modulation classifiers. These modulation classification algorithms employ the decision-theoretic and neural network approaches. This results in two types of modulation classifiers that are capable of distinguishing a very wide range of digitally modulated signals. The results for the DT and NN classifiers are presented and compared for SNR ranging from 20dB to -5dB. The performance is also tested for signals undergoing Rayleigh fading.

The structure of this chapter is as follows. In section 8.2 we describe the signals that are added to the modulation classifiers as well as their useful features. In section 8.3 we discuss the DT classifier implementation including the tree structure and threshold determination. Section 8.4 outlines the NN classifier implementation with the addition of the new signals. The results for both classifiers are presented in section 8.5 and a comparison between the performance of the DT and NN classifiers is made. Finally, we present concluding remarks in section 8.6.

### 8.2 Signal Representation

The signals that are added to the modulation classifiers discussed in Chapter 8 are described in this section. These signals are QAM8, QAM16, PSK8, and FSK8. The key features associated with these signals are also described

A well-known technique to reduce the bandwidth of a signal is to employ  $M$ -ary phase shift keying (MPSK) modulation. Instead of transmitting one bit of information per channel symbol period,  $k = \log_2 M$  bits are sent during each symbol period. The use of  $M$ -ary

symbols allows the data rate to be increased  $k$  times within the same bandwidth. Therefore for a fixed data rate, the use of  $M$ -ary PSK reduces the required bandwidth by a factor  $k$  [Sklar, 1988]. The representation of PSK signals has been shown in Chapter 5 and some useful features of PSK8 signals are shown in Figure 8.1. A QPSK signal consists of two independent amplitude modulated signals that are 90 degrees out of phase. The signal has amplitude levels of  $\pm 1$ . QAM is a logical extension of QPSK in that the signal also consists of two independently amplitude modulated signals. The only difference is that the signal can have  $k$ -bit symbols instead of amplitude levels of just +1 and -1. Therefore QAM signals can be viewed as combined amplitude and phase modulation.

The corresponding signal can be expressed as

$$s(t)x(t)\cos\omega_c t - y(t)\sin\omega_c t, \quad (8.1)$$

where  $x(t)$  and  $y(t)$  are the information bearing signal amplitudes of the quadrature carriers.

The complex envelope is given by [Couch, 2001]

$$\alpha(t) = x(t) + jy(t) = R(t)e^{j\theta(t)} \quad (8.2)$$

The instantaneous amplitude and phase are

$$a(t) = |R(t)| \quad (8.3)$$

$$\phi(t) = \tan^{-1}[y(t)/x(t)] \quad (8.4)$$

These features are shown in Figure 8.2 and Figure 8.3 for QAM8 and QAM16 modulation.

The representations FSK signals have been described in Chapter 5 and some useful features of FSK8 signals are shown in Figure 8.4.



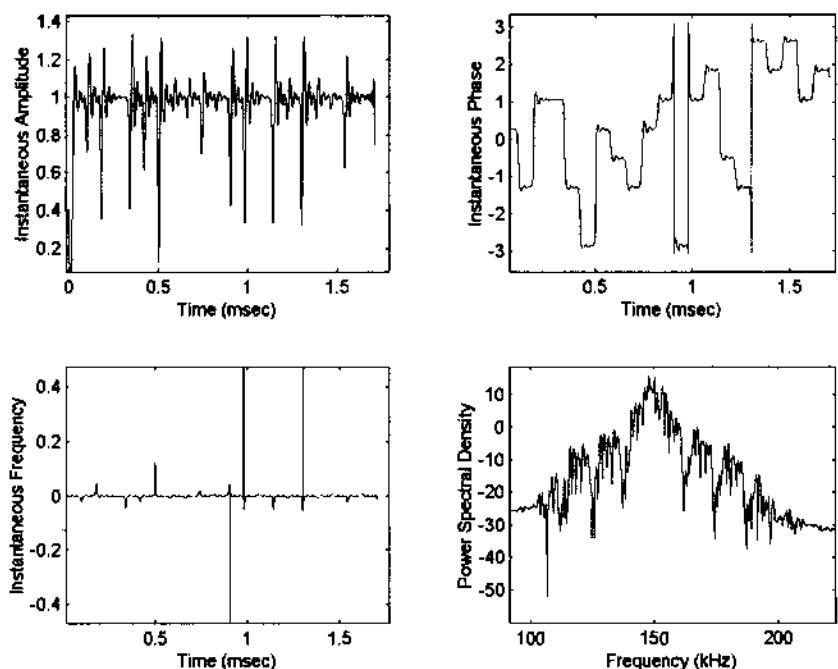


Figure 8.1. Useful features of PSK8 modulation.

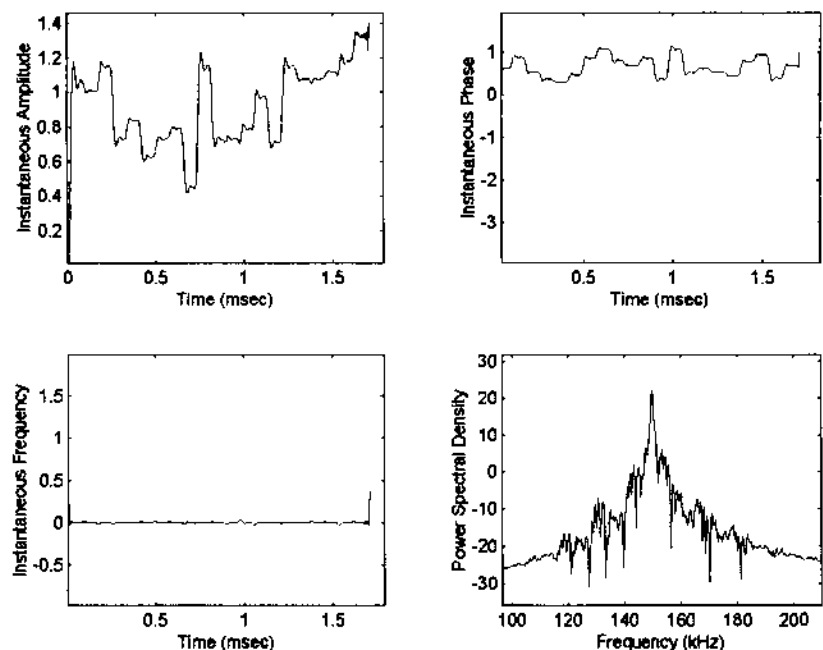


Figure 8.2. Useful features of QAM8 modulation.

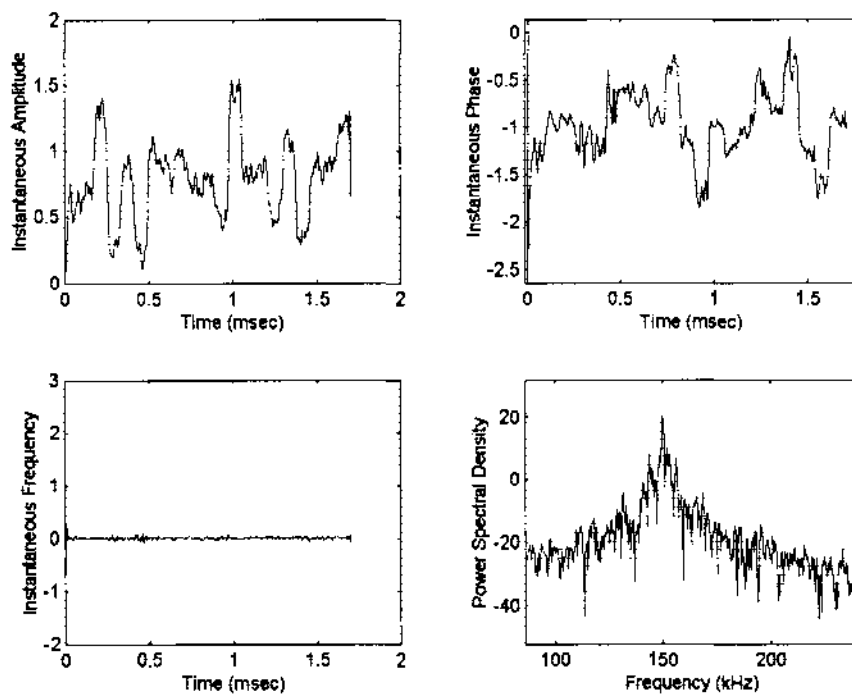


Figure 8.3. Useful features of QAM16 modulation.

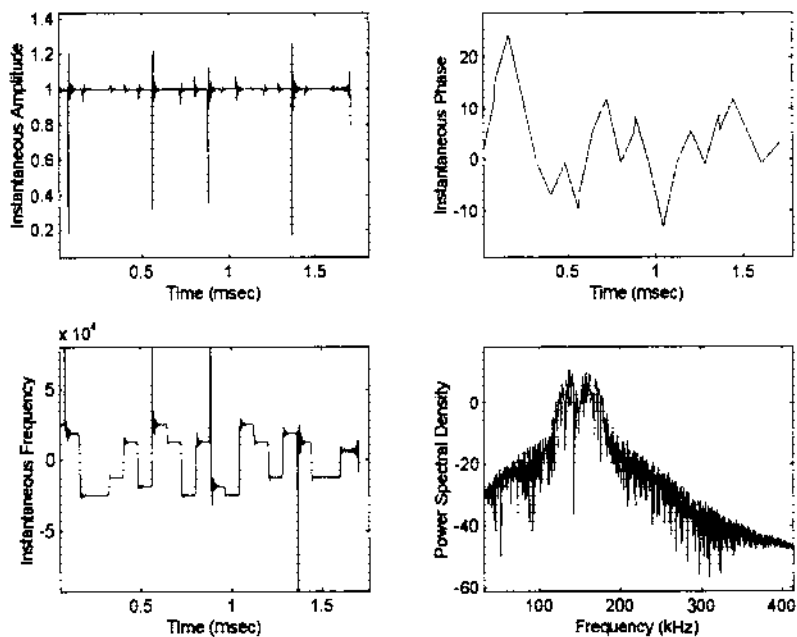


Figure 8.4. Useful features of FSK8 modulation.

### 8.3 DT Classification Procedure

This section outlines the procedure for digital signal classification that is based on the method outlined in Chapters 5-7. Firstly, the key features are derived from the instantaneous amplitude, the instantaneous phase, the instantaneous frequency, the smoothed power spectral density, and the fourth order cumulants of the intercepted signal. A description of the threshold values is presented followed by a discussion of how the key feature selection is dependent on the minimum error probability. A flowchart depicting the classification procedure for all digital modulation schemes is shown in Figure 8.5.

The first decision in the tree separates the signals with frequency information (right side of the tree) from signals with little or no frequency information (left side of the tree). The signals with frequency information are further divided into multiple access signals (FH SS and TDMA) and FSK/CPM signals. The signals with no frequency information are divided into signals with phase information (PSK and BPSK/QPSK DS SS) and signals with little or no phase information (ASK and QAM). The signals with phase information are split into multiple access signals (BPSK/QPSK DS SS) and PSK signals (PSK2, PSK4, PSK8). Finally, QAM signals are separated from ASK signals.

#### 8.3.1 Derivation of Key Features

To derive the appropriate key features, the new signals (PSK8, FSK8, QAM8, and QAM16) are passed through the existing classifier and each signal is classified as a modulation type already defined in the tree. To find the actual modulation type of a particular signal, a decision node is added to the tree to distinguish between the modulation type that the signal is classified as and the actual modulation type of the new signal.

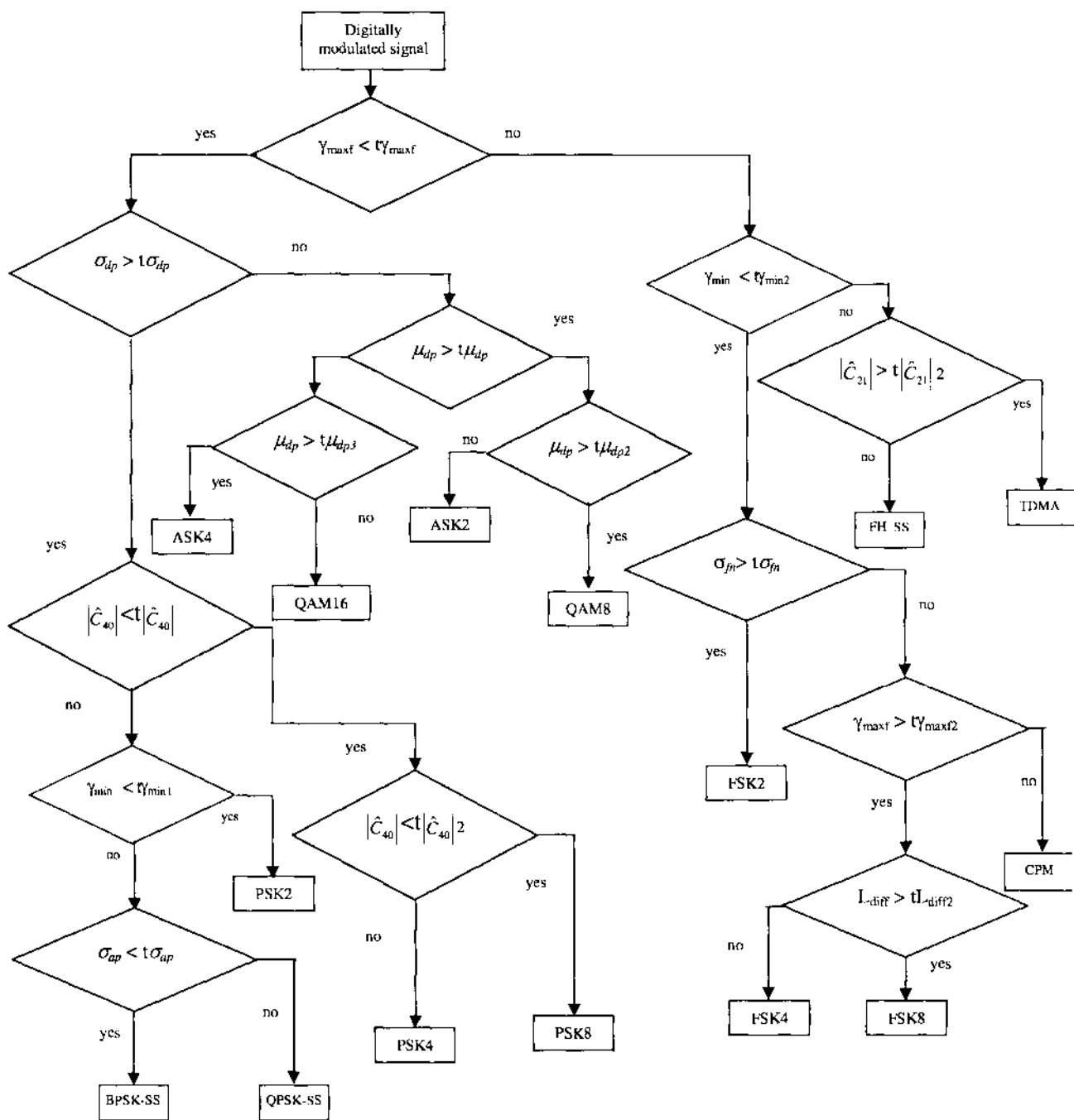


Figure 8.5. Decision tree for identification of digital modulation schemes. Refer to section 8.3 for an explanation of the tree.

### 8.3.1.1 QAM8 Signal Classification

To derive the appropriate key feature for classification, the QAM signal (namely a QAM8 signal) is classified by the existing tree in Chapter 7. The QAM8 signal is classified 76.5% as an ASK2 signal, 20% as a PSK4 signal, and 3.5% as a PSK2 signal. These results are not sufficient to add the QAM8 signal to the existing tree as it is. In other words, the key feature  $|\hat{C}_{21}|$  is not adequate for classification of QAM signals; therefore, the tree structure has to be modified slightly. The key feature  $|\hat{C}_{21}|$  is replaced by the key feature  $\sigma_{dp}$  which is the standard deviation of the direct value of the non-linear component of the instantaneous phase, evaluated over the non-weak segments of the received signals and is defined in [Azzouz and Nandi, 1996]. This key feature is used to separate signals with phase information (PSK signals) from those with no phase information (ASK signals). A suitable threshold for  $\sigma_{dp}$  is determined using the previous methods and is outlined in subsection 8.3.2. After the addition of this new key feature, the QAM8 signal is passed through the classifier again. This time the signal is classified 100% as an ASK2 signal so we know the decision has to be made between QAM8 and ASK2 signals. By observing the instantaneous phase plots for both signals, it can be seen that QAM8 signals possess some phase information since they are a combination of amplitude and phase modulation. Therefore the existing key feature  $\mu_{dp}$  is used to separate QAM8 and ASK2 signals.

### 8.3.1.2 QAM16 Signal Classification

The QAM16 signal is classified by the tree after the QAM8 signal has been added. It is found that the QAM16 signal is classified as an ASK4 signal. Since the decision is now to be made between QAM16 and ASK4 signals, we observe from Figure 5.2 and Figure 8.3 that the instantaneous phase values for ASK4 signals lie around zero and the instantaneous phase values for the QAM16 signal lie around  $-1$ . Therefore the key feature  $\mu_{dp}$  can be used to differentiate between these two modulation types.

### 8.3.1.3 PSK8 Signal Classification

To derive the appropriate key feature for classification, the PSK8 signal is classified by the existing tree after the addition of the QAM signal. The PSK8 signal is classified as a PSK4 signal, therefore we know the decision has to be made between the PSK8 signal and a PSK4 signal. To discriminate between these two signals, the existing key feature  $|\hat{C}_{40}|$  is used. By referring to Table I in [Swami and Sadler, 2000], it can be seen that the theoretical values of the fourth order cumulants for PSK4 signals are 1.0 and the values of the fourth order cumulants for PSK( $> 4$ ) are around 0.0. Therefore we can use the key feature  $|\hat{C}_{40}|$  to separate PSK4 and PSK8 signals. Another advantage of this key feature is that it is not affected by phase offsets.

### 8.3.1.4 FSK8 Signal Classification

The FSK8 signal is classified by the existing tree after the addition of the PSK8 and QAM signals, to obtain an appropriate key feature for classification. The signal is classified as FSK4; therefore we know the decision has to be made between FSK8 and FSK4 signals. The bandlimiting of the signals causes the FSK8 and FSK4 signals to have very similar characteristics and the separation of these two signals becomes almost impossible using the methods used in previous chapters. The key feature  $L_{\text{diff}}$  can be used to separate FSK4 and FSK8 to some degree but the performance is not satisfactory. Therefore, we increase the bandwidth of the FSK signals from 100kHz to 200kHz. This greatly improves the results of the classifier. However by increasing the bandwidth of the signals, the threshold value for the key feature  $\sigma_{f_n}$  to differentiate FSK2 (subset A) and FSK4 and CPM (subset B) must be modified and this modification is outlined in the next section.

## 8.3.2 Threshold Determination

As explained in previous chapters, the key feature thresholds are chosen so that the probability of a correct decision is obtained from 400 realisations of each modulation type at the SNR range of 20dB to -5dB. A set of modulation types is separated into two non-overlapping subsets (A and B). The optimum threshold is chosen so that the Bayes error is minimised as discussed in Chapter 3.

The total error probability for groups A and B are plotted and the threshold is chosen where the minimum error occurs. The total error probability for the key feature  $\sigma_{dp}$  is shown in Figure 8.6 for subset A (ASK2, ASK4, QAM8, and QAM16) and subset B (PSK2, PSK4, PSK8, BPSK DS-SS, and QPSK DS-SS). It can be seen that an appropriate value for the threshold  $t\sigma_{dp}$  is 1.1 where the total minimum error is 0.000166 for the SNR range of 20dB to 5dB. For the SNR of 0dB and -5dB, the total minimum error is 0.057 at the same threshold value.

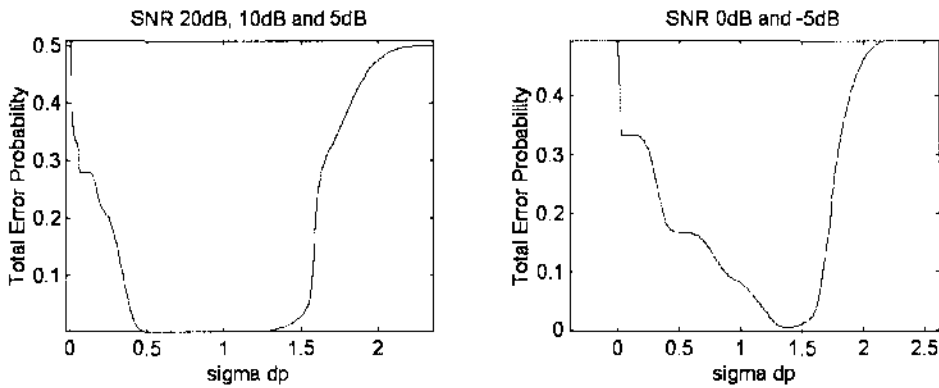


Figure 8.6. Total error probability for the key feature  $\sigma_{dp}$ , at SNR range of 20dB to -5dB, for ASK2, ASK4, QAM8, and QAM16 (subset A) and PSK2, PSK4, PSK8, BPSK DS-SS, and QPSK DS-SS (subset B).

The total error probability for the key feature  $|\hat{C}_{40}|$  to separate subset A (PSK4) and subset B (PSK8) is shown in Figure 8.7. By observation, an appropriate value for the threshold  $t|\hat{C}_{40}|_2$  is chosen to be 0.59 where the total minimum error probability is 0.0775 for the SNR range of 20dB to 5dB. For lower SNR of 0dB and -5dB, the total minimum error is 0.4652 at the same threshold value.

The ROC curves for the key feature  $|\hat{C}_{40}|$  that separates PSK4 (subset A) from PSK8 (subset B) are shown in Figure 8.8 for SNR range of 20dB to -5dB. The curves show the detection probability of subset A (PSK4) and false alarm probability of subset B (PSK8).

By examining the ROC curves for  $\text{SNR} \geq 10\text{dB}$ , we can see that the chosen threshold value,  $t|\hat{C}_{40}|_2$  (indicated by 'x') has a detection probability ( $P_D$ ) of 0.9075 and false alarm probability ( $P_{FA}$ ) of 0.0325 at 10dB SNR.

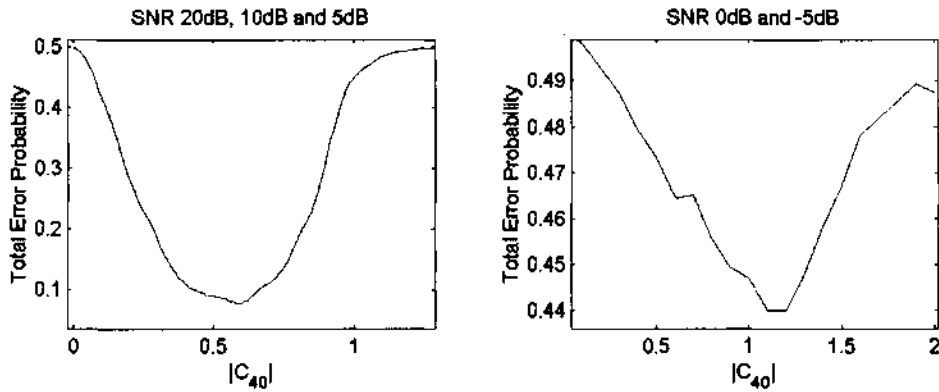


Figure 8.7. Total error probability for the key feature  $|\hat{C}_{40}|$ , at SNR range of 20dB to -5dB, for PSK4 (subset A) and PSK8 (subset B).

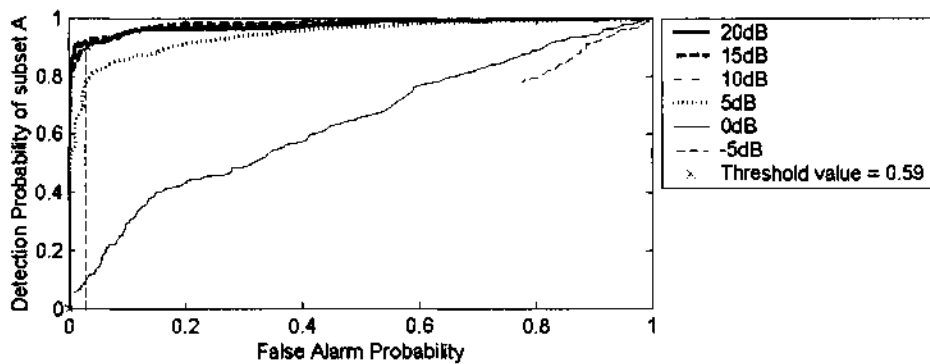


Figure 8.8. ROC curves for the key feature  $|\hat{C}_{40}|$  to separate PSK4 (subset A) and PSK8 (subset B) signals for SNR range of 20dB to -5dB.

For the separation of ASK2 from QAM8, the key feature  $\mu_{dp}$  is used. The total error probability for the key feature  $\mu_{dp}$  is shown in Figure 8.9 for subset A (ASK4) and subset B



(QAM8). From this figure we can infer that an appropriate choice for the threshold  $t\mu_{dp2}$  is 0.19 where the total minimum error probability is 0.0035 for the SNR range of 20dB to -5dB.

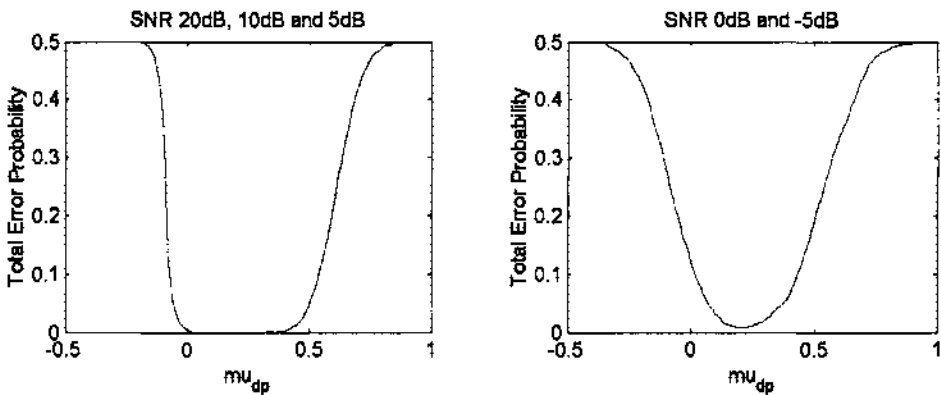


Figure 8.9. Total error probability for the key feature  $\mu_{dp}$ , at SNR range of 20dB to -5dB, for ASK2 (subset A) and QAM8 (subset B).

Similarly, to find the threshold value  $t\mu_{dp3}$  to separate QAM16 and ASK4, we examine the total error probability plot in Figure 8.10. It can be seen that at the threshold value of -0.46, the minimum error is 0 for the SNR range of 20dB to 5dB. For the SNR values of 0dB and -5dB, the minimum error is 0.0056 at the same threshold value.

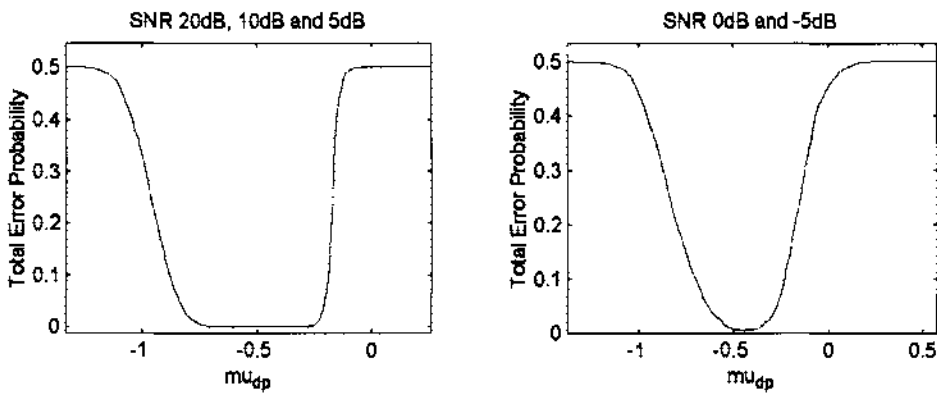


Figure 8.10. Total error probability for the key feature  $\mu_{dp}$ , at SNR range of 20dB to -5dB, for ASK4 (subset A) and QAM16 (subset B).

The threshold value for the key feature  $L_{diff}$  is found from the total error probability plotted in Figure 8.11 for subset A (FSK8) and subset B (FSK4). When the bandwidth of the FSK signals is 100kHz, an appropriate choice for the threshold value  $tL_{diff2}$  is 0.3 where the minimum error is 0.2096 for the SNR range of 20dB to 5dB.

The ROC curves for the key feature  $L_{diff}$  that separates FSK8 (subset A) from FSK4 (subset B) are shown in Figure 8.12 for SNRs 20dB, 15dB and 10dB. The curves show the detection probability of subset A (FSK8) and false alarm probability of subset B (FSK4). The bandwidth of the FSK signals is 100kHz. It can be observed that the ROC curves are not of a desirable form because the detection probability ( $P_D$ ) is not very high for low false alarm probability ( $P_{FA}$ ). Therefore, it is necessary to increase the bandwidth of the FSK signals to 200kHz to improve performance.

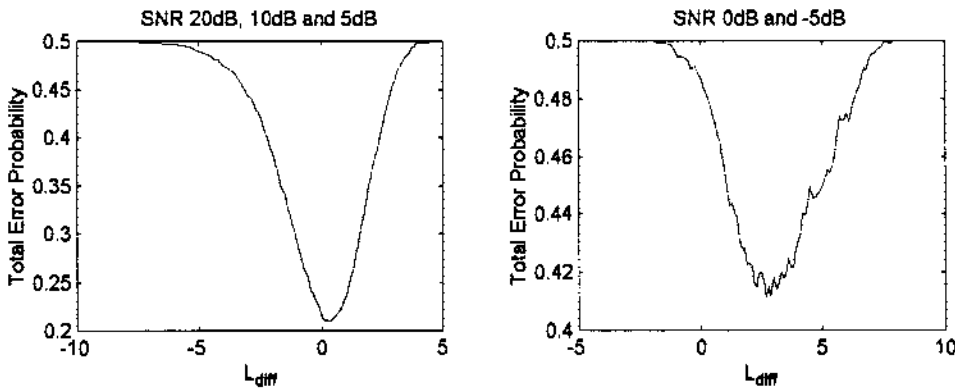


Figure 8.11. Total error probability for the key feature  $L_{diff}$ , at SNR range of 20dB to -5dB, for FSK8 (subset A) and FSK4 (subset B) bandlimited to 100kHz.

The total error probability for the key feature  $L_{diff}$  is shown in Figure 8.13 for subset A (FSK4) and subset B (FSK8). The total minimum error probability is 0.1004 and occurs at the threshold value  $tL_{diff2} = -7.1$  for the SNR range of 20dB to 5dB when the bandwidth of the FSK signals is increased to 200kHz.

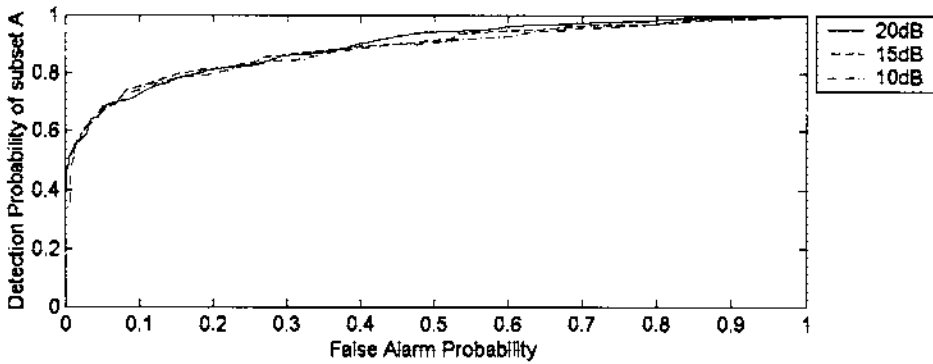


Figure 8.12. ROC curves for the key feature  $L_{diff}$  to separate FSK8 (subset A) and FSK4 (subset B) signals (bandlimited to 100kHz) for SNRs 20dB, 15dB and 10dB.

The ROC curves for the key feature  $L_{diff}$  that separates FSK8 (subset A) from FSK4 (subset B) are shown in Figure 8.14 for the SNR range of 20dB to -5dB. The curves show the detection probability of subset A (FSK8) and false alarm probability of subset B (FSK4). By observing the ROC curves for  $SNR \geq 5dB$ , we can see that the chosen threshold value  $t_{L_{diff2}}$  (indicated by 'x') has a minimum detection probability ( $P_D$ ) of 0.9688 and false alarm probability ( $P_{FA}$ ) of 0.1825 at 5dB SNR.

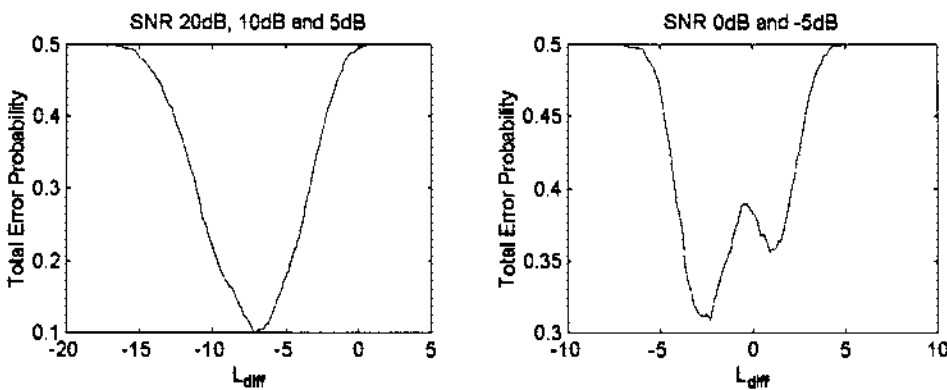


Figure 8.13. Total error probability for the key feature  $L_{diff}$ , at SNR range of 20dB to -5dB, for FSK4 (subset A) and FSK8 (subset B) bandlimited to 200kHz.

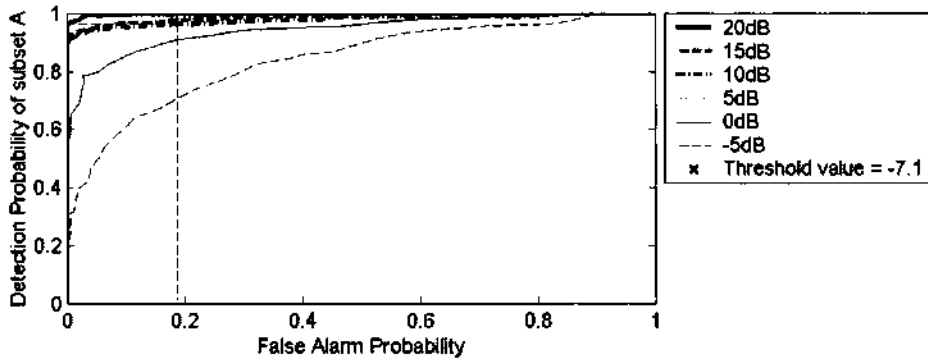


Figure 8.14. ROC curves for the key feature  $L_{diff}$  to separate FSK8 (subset A) and FSK4 (subset B) signals (bandlimited to 200kHz) for SNR range of 20dB to -5dB.

The new threshold value ( $t\sigma_{fn}$ ) that separates subset A (FSK2) and subset B (FSK4, FSK8, and CPM) is found in Figure 8.15 from the plotted total error probability. The appropriate choice for  $t\sigma_{fn}$  is 3.2 when the bandwidth of the FSK signals is increased to 200kHz. This gives a minimum error probability of 0.0225 for the SNR range of 20dB to 5dB. For SNR of 0dB and -5dB, the minimum error probability is 0.236 at the same threshold value.

The corresponding ROC curves for the key feature  $\sigma_{fn}$  that separates FSK2 (subset A) from FSK4, FSK8, and CPM (subset B) are shown in Figure 8.16 for the SNR range of 20dB to -5dB. The curves show the detection probability of subset A (FSK2) and false alarm probability of subset B (FSK4, FSK8, and CPM). By examining the ROC curves for SNR  $\geq 5$ dB, we can see that the chosen threshold value  $t\sigma_{fn}$  (indicated by 'x') has a detection probability of 0.94 and false alarm probability of 0.0058 when the bandwidth of the FSK signals is increased to 200kHz at 5dB SNR.

Note that for all the ROC curves in this chapter, both classes are equally important and we are not trying to bias one class against the other. The optimum threshold is only dependent on the total minimum error probability for the SNR range of 20dB to -5dB.

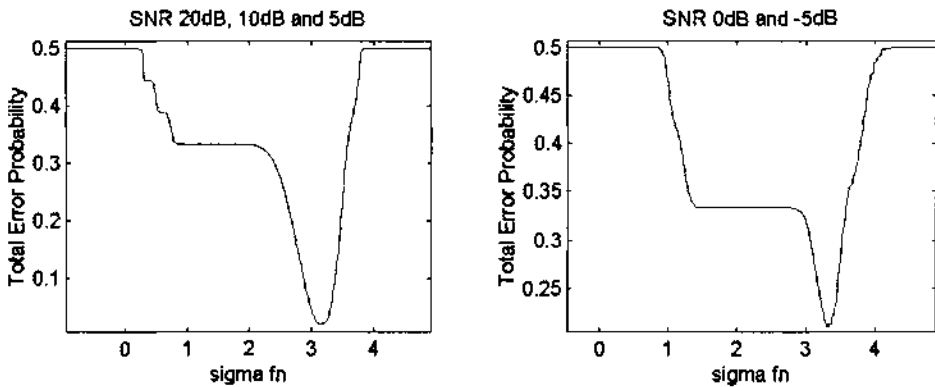


Figure 8.15. Total error probability for the key feature  $\sigma_{fn}$ , at SNR range of 20dB to -5dB, for FSK4, FSK8, and CPM (subset B) and FSK2 (subset A) bandlimited to 200kHz.

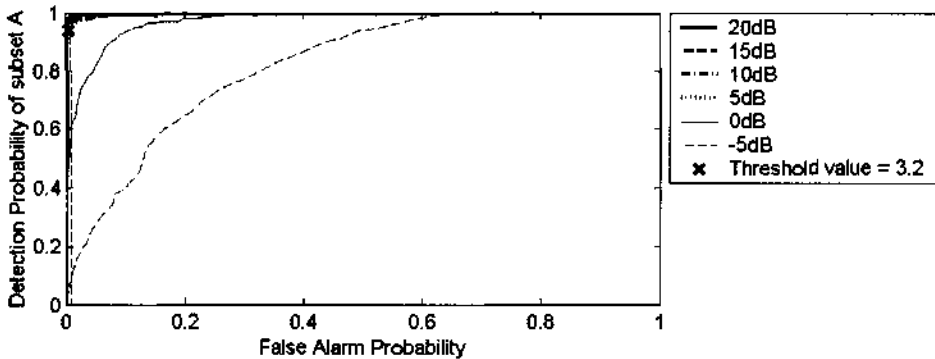


Figure 8.16. ROC curves for the key feature  $\sigma_{fn}$  to separate FSK2 (subset A) and FSK4, FSK8, and CPM (subset B) signals (bandlimited to 200kHz) for SNR range of 20dB to -5dB.

A summary of the key feature values and their relevant thresholds for the SNR range of 20dB to -5dB is shown in Table 8.1. The final threshold values are  $t_{\sigma_{dp}} = 1.1$ ,  $t_{|\hat{C}_{40}|_2} = 0.59$ ,  $t_{\mu_{dp2}} = 0.19$ ,  $t_{\mu_{dp3}} = -0.46$ ,  $t_{L_{diff2}} = -7.1$ , and  $t_{\sigma_{fn}} = 3.2$ .

Table 8.1. Summary of key feature thresholds and error probabilities.

Key Feature Threshold	SNR 20dB to 5dB		SNR 0dB to -5dB	
	Optimum Threshold	Minimum Error Probability	Optimum Threshold	Minimum Error Probability
$t_{\sigma_{dp}}$	1.1	0.000166	1.34	0.0077
$t_{ \hat{C}_{40} _2}$	0.59	0.0775	1.16	0.4350
$t_{\mu_{dp2}}$	0.19	0	0.19	0.0088
$t_{\mu_{dp3}}$	-0.46	0	-0.46	0.0056
$t_{L_{diff2}}$	-7.1	0.1004	-2.3	0.3094
$t_{\sigma_{fn}}$	3.2	0.0225	3.3	0.2090

### 8.3.3 Dependency of key feature selection on minimum probability of error

The reason why the key features in the previous section are chosen over the other existing key features is because they minimise the total error probability for each decision. We will call the decision separating ASK2, ASK4, and QAM (subset A) and PSK2, PSK4, PSK8, BPSK DS-SS, and QPSK DS-SS (subset B) decision 1. Decision 2 separates PSK4 (subset A) and PSK8 (subset B) and decision 3 distinguishes ASK2 (subset A) from QAM8 (subset B). Decision 4 is defined as the classification of ASK4 (subset A) and QAM16 (subset B) and finally, Decision 5 separates FSK4 (subset A) and FSK8 (subset B).

We can see from Table 8.2 that the key features that have been chosen minimise the total error probability for each decision for the SNR range of 20dB to -5dB.

### 8.4 NN Classifier

This section introduces a NN classifier that can recognise the same fifteen signals as the DT classifier described in section 8.3. The input datasets for the NN are the same key features used in the DT algorithm. These key features are:  $\gamma_{maxf}$ ,  $\gamma_{min}$ ,  $\sigma_{ap}$ ,  $\sigma_{dp}$ ,  $\sigma_{fn}$ ,  $\mu_{dp}$ ,  $L_{diff}$ ,  $|\hat{C}_{21}|$ , and  $|\hat{C}_{40}|$ . All key features are normalised to the range  $-1$  to  $1$ , then passed to the neural network. This section will describe the NN structure followed by a description of how the NN and its subnets are trained.

Table 8.2. Total minimum error probability for Decisions 1 – 5 for combined SNR range of 20dB to  $-5$ dB (threshold values are shown in brackets).

Key Feature	Total Minimum Error Probability (Decision 1)	Total Minimum Error Probability (Decision 2)	Total Minimum Error Probability (Decision 3)	Total Minimum Error Probability (Decision 4)	Total Minimum Error Probability (Decision 5)
$\gamma_{maxf}$	0.3640 (-104)	0.5000 (-92)	0.435(-108.3)	0.428(-109.3)	0.4132 (26)
$\mu_{dp}$	0.2504 (0)	0.4088 (0.044)	<b>0.0035 (0.19)</b>	<b>0.0022 (-0.46)</b>	0.4875 (-28)
$ \hat{C}_{21} $	0.0962 (0.94)	0.4968 (4.2)	0.2748 (0.7)	0.2717 (0.65)	0.4830 (1.0)
$ \hat{C}_{40} $	0.2208 (1.448)	<b>0.2325 (0.59)</b>	0.2145 (0.58)	0.2495 (0.4)	0.483 (0.02)
$\sigma_{fn}$	0.2600 (0)	0.5000 (0)	0.4989 (0)	0.4680 (0)	0.3277 (2.76)
$\sigma_{ap}$	0.1560 (0.52)	0.4305 (0.91)	0.2655 (0.26)	0.2285 (0.28)	0.4925 (410)
$\gamma_{minf}$	0.3003 (-32)	0.4284 (-43)	0.4150 (-50)	0.4231 (-45)	0.4509 (-45.7)
$L_{diff}$	0.278 (-0.37)	0.4910 (1.0)	0.3855 (-5.0)	0.3855 (-5)	<b>0.2602 (-7.1)</b>
$\sigma_{dp}$	<b>0.0229 (1.1)</b>	0.4769 (1.85)	0.303 (0.226)	0.303 (0.226)	0.4931 (410)

### 8.4.1 Neural Network Structure

The developed network is based on a seven-network structure. Each network is a feedforward network, commonly referred to as a multi-layer perceptron (MLP). The first network has two inputs corresponding to the two key features  $\gamma_{min}$  and  $\gamma_{max}$  and two output neurons corresponding to two groups of signals, which are:

1. FSK2, FSK4, FSK8, FH SS, TDMA, and CPM
2. QPSK-SS, BPSK-SS, PSK2, PSK4, PSK8, QAM8, QAM16, ASK2 and ASK4.

It is found that by dividing the signals into these two groups initially, results in optimum performance. Three structures are tested and it is found that the simplest structure giving the best performance has one hidden layer consisting of four neurons. Twenty versions of this structure are tested to find the one that gives the best performance. After the initial first network structure is designed, the other network structures can be derived from the decision tree to separate the signals. In all the networks described, all hidden layer use the nonlinear *tan-sigmoid* (hyperbolic tangent) function and the output layers are linear activation functions. Also, twenty versions of each network structure are examined to find the one that gives the best performance.

The second network separates the signals into two groups and therefore this network has two output neurons. The first group consists of signals with little or no phase information (ASK2, ASK4, QAM8, and QAM16 signals) and the second group consists of signals with phase information (PSK2, PSK4, PSK8, BPSK-SS, and QPSK-SS signals). The network has three input neurons corresponding to the key features  $\sigma_{dp}$ ,  $\sigma_{up}$ , and  $|\hat{C}_{40}|$ . It is found that the simplest structure that gives the best performance has one hidden layer with ten neurons.

The third network has one input corresponding to the key feature  $\mu_{dp}$  and three output neurons corresponding to ASK2 and ASK4 signals (as one group) QAM8, and QAM16 signals. Three networks are tested and the structure that gives the best performance has one hidden layer with four neurons.



The fourth tested network classifies PSK2, PSK4, PSK8, BPSK DS-SS, and QPSK DS-SS signals. This network has four input neurons corresponding to the key features  $\gamma_{\min}$ ,  $\sigma_{ap}$ ,  $\sigma_{dp}$ , and  $|\hat{C}_{40}|$ , and five output neurons corresponding to the five modulation types. It is found that the optimum structure in terms of simplicity and performance has two hidden layers with seven neurons in the first layer and four neurons in the second layer.

The fifth network has one input neuron corresponding to the key feature  $\mu_{dp}$  and two output neurons representing ASK2 and ASK4 signals respectively. The optimum structure is found to have one hidden layer with ten neurons. The sixth network classifies TDMA, FH SS, CPM, and FSK signals. Therefore, there are four output neurons and four input neurons corresponding to the key features:  $\gamma_{\max f}$ ,  $\gamma_{\min}$ ,  $\sigma_{fn}$ , and  $|\hat{C}_{21}|$ . The simplest structure giving the best performance has one hidden layer with four neurons.

The final network classifies FSK2, FSK4, and FSK8 signals and consists of three output neurons and three input neurons corresponding to the key features  $\gamma_{\max f}$ ,  $\sigma_{fn}$ , and  $L_{\text{diff}}$ . The network structure giving the best performance has one hidden layer with ten neurons.

The full network structure is shown in Figure 8.17 and it has been shown in Chapter 7 that smaller network structures give better performance. This is why the hierarchical layout is a better choice than one large network that must discriminate between all fifteen signals simultaneously.

#### 8.4.2 Training the Network

All networks are trained using the Levenberg-Marquardt (LM) algorithm using 200 samples from each modulation type except Network 1, which is trained using the conjugate gradient function. Each network is also tested and validated using a separate set of 200 samples of each modulation type as described in previous chapters. The training data is a mix of samples with SNR ranging from 20dB to -5dB.

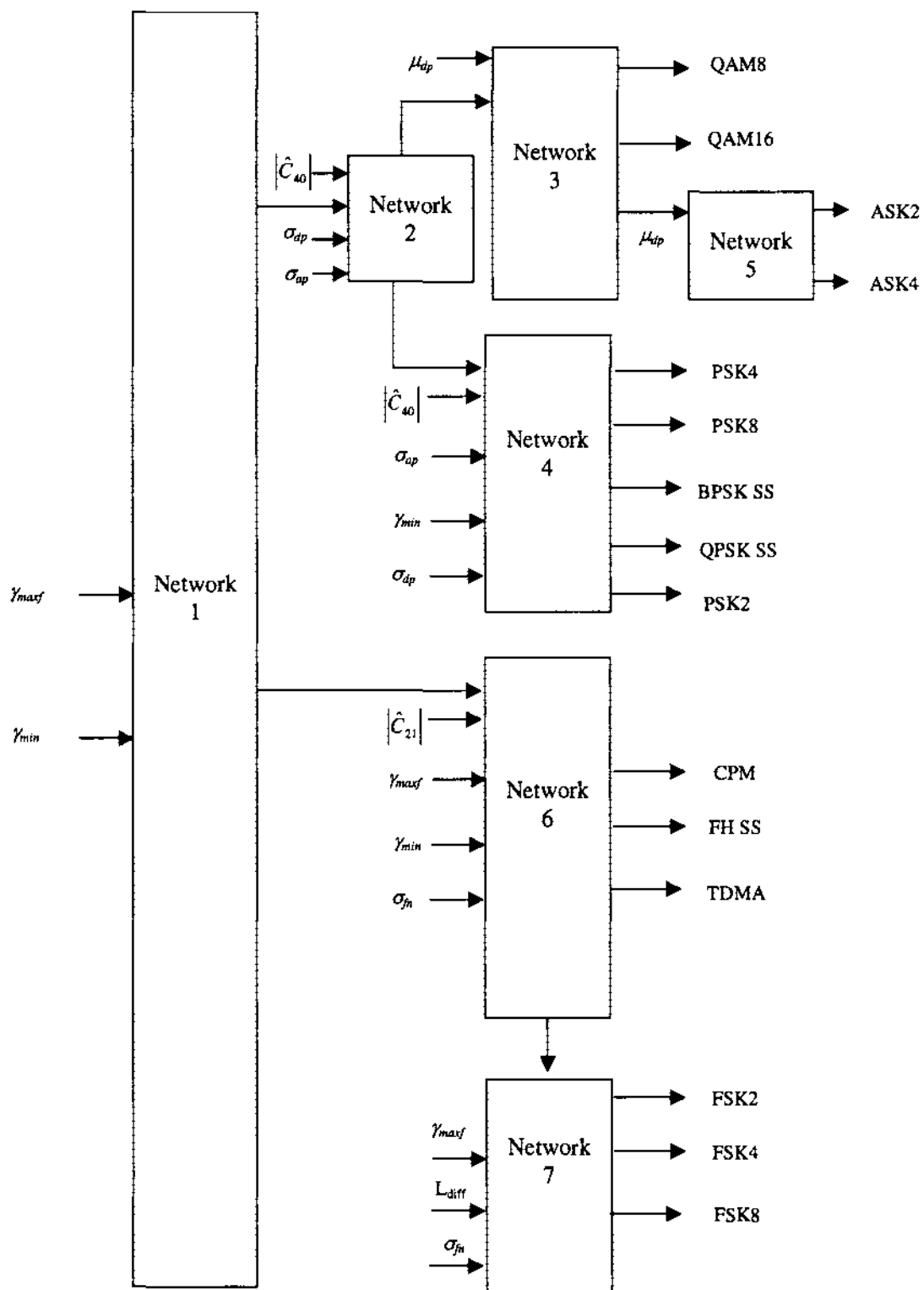


Figure 8.17. Neural network structure for digital modulation classification.

## **8.5 Performance Analysis of DT and NN Classifiers in the Presence of White Gaussian Noise**

In this section, we first present the results for the DT classifier, followed by the results of the NN classifier in the presence of Gaussian noise. A comparison is made between the accuracy of these two types of classifiers and the 95% confidence interval is also included.

### **8.5.1 Performance Results for DT Classifier**

The results for the DT classifier are derived from 200 realisations of each modulation type. The carrier frequency, sampling rate and the symbol rate are given values of 150kHz, 1200kHz and 12.5kHz, respectively. The digital symbol sequence is randomly generated.

The simulation results for the test set of the digital modulation recogniser for all signals based on 200 realisations are given in Appendix D, Table D.1 – Table D.6, for the SNR range of 20dB to –5dB, respectively. It can be seen that the performance of the classifier for SNR less than 5dB is much poorer for most signals. However for SNR values greater than or equal to 5dB, the results indicate that all types of the digital modulation schemes considered can be correctly classified with greater than 93% overall success.

The graphical representation of the performance of the modulation classifier for all modulation types is shown in Figure 8.19 - Figure 8.24 for an SNR range of 20dB to –5dB. The results are compared with the results from the NN classifier, discussed in the section.

### **8.5.2 Performance Results of NN Classifier**

The performance results of the NN classifier for the SNR range of 20dB to –5dB are given in Figure 8.19 - Figure 8.24 inclusive. The results for the DT classifier are also shown for comparison with the 95% confidence interval. It can be observed that the NN has good performance with a success rate of over 93% for the SNR range of 20dB to 5dB. The classifier still performs very well for SNR of 0dB at nearly 84% overall accuracy. This is because the network is trained with data of SNR range of 20dB to -5dB. A tabular comparison between the results from the DT approach and the NN approach is shown in Table 8.3. A graphical comparison between the overall success rates of both classifiers

over the SNR range of 20dB to -5dB is also shown in Figure 8.18. In general, the overall classifier accuracies for the DT and NN algorithms are similar for SNR values greater than or equal to 5dB. However, the NN classifier outperforms the DT classifier considerably at SNR 0dB and -5dB. This is probably due to the fact that the DT classifier has a linear decision boundary based on one key feature whereas the NN has the option of having non-linear decision boundaries based on more than one key feature. The confusion matrices are shown in Appendix D, Table D.7 – Table D.12, for the SNR range of 20dB to -5dB.

The NN approach is dependent on the DT approach in terms of key feature selection and hierarchical network selection. By referring to the decision tree in Figure 8.5, it can be seen that the neural network structure in Figure 8.17 is based on the decision tree. The key features relevant to a particular section of the decision tree serve as inputs to the corresponding network. For example, by referring to Network 3, the signals of interest are ASK, QAM8, and QAM16. If we observe the decision tree we can determine that the relevant key feature is  $\mu_{dp}$  and this is the input to Network 3.

Table 8.3. DT and NN classifier accuracy and 95% confidence intervals.

SNR	DT Classifier		NN Classifier	
	Accuracy	95% Confidence Interval	Accuracy	95% Confidence Interval
20dB	97.02%	[96.57, 97.45]	97.84%	[97.47, 98.21]
15dB	96.87%	[96.43, 97.31]	97.94%	[97.58, 98.30]
10dB	96.72%	[96.27, 97.17]	97.27%	[96.86, 97.68]
5dB	93.80%	[93.19, 94.41]	93.67%	[93.05, 94.29]
0dB	74.98%	[73.89, 76.08]	83.90%	[82.97, 84.83]
-5dB	47.37%	[46.10, 48.63]	73.58%	[72.46, 74.69]
Overall	84.46%	[84.08, 84.83]	90.70%	[90.40, 91.00]

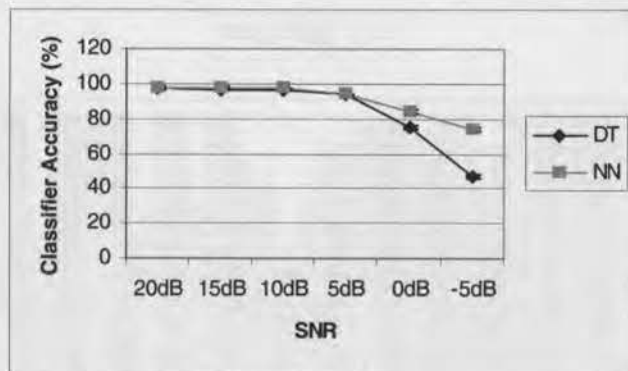


Figure 8.18. Graphical comparison of overall performance between the NN-based and DT-based classifiers.

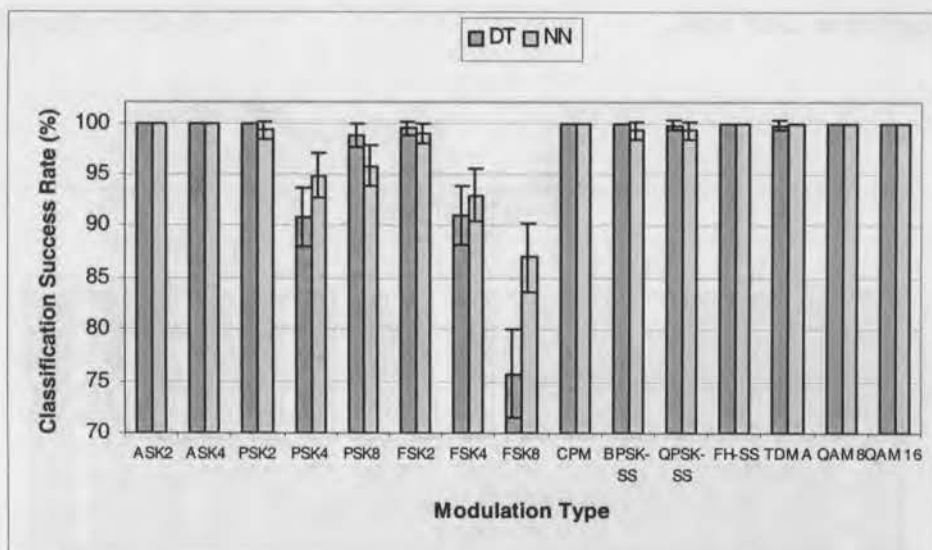


Figure 8.19. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 20dB SNR.

The next chapter outlines the performance of the DT and NN classifiers with signals affected by Rayleigh fading. The classifiers are modified to accommodate fading and the performance of these modified classifiers are compared to the classification performance of signals in an AWGN channel.



# CHAPTER 9

## Classification of Digitally Modulated Signals in the Presence of Rayleigh Fading

### 9.1 Introduction

In this chapter, the performance of the DT and NN classifiers described in Chapter 8 will be evaluated under the conditions of a Rayleigh fading channel. The classifiers will be modified so that they can perform optimally whether fading is present or not. In section 9.2 we discuss the classification of digitally modulated signals in the presence of Rayleigh fading beginning with an introduction to Rayleigh fading channels. The modifications in the decision tree to accommodate fading are discussed in section 9.3. Similarly, the modifications to the NN classifier in the presence of fading are outlined in section 9.4. The performance of both classifiers in the presence of fading are discussed in section 9.5 and a comparison between the performance in an AWGN channel and a fading channel is made.

### 9.2 Classification in the Presence of Rayleigh Fading Channels

The DT and NN classifiers described in Chapter 8 will be tested under the conditions of a Rayleigh fading channel. An introduction to fading channels will first be presented, followed by the results of the classification performance in a fading environment.

#### 9.2.1 Introduction to Fading Channels

In the 1920s, experiments were carried out with mobile communications at VHF frequencies. From the results of these experiments (carried out at about 50MHz) it was found that there was a very hostile propagation environment, particularly in urban centers.



Moving the vehicle over a few metres resulted in dramatic changes to the received signal's strength. The signal varied from excellent quality to no signal. The mobile or indoor radio channel is characterised by multipath reception. The received signal is a summation of the direct line of sight radio wave as well as a large number of reflected radio waves. These reflected waves interfere with the direct wave, which causes significant degradation in the strength of the signal.

In most communication systems, the channel is modelled as a linear time-invariant system. This model consists of a delay term proportional to the propagation delay between the channel modulator and channel demodulator. The transfer function consists of a frequency independent magnitude less than one that is proportional to the propagation loss. The channel is usually considered to be corrupted by AWGN which is adequate for deep space communication channels. However for many radio channels such as high-frequency (HF) long-distance communications via the ionosphere, microwave communications and mobile communications, the AWGN channel is an oversimplified model. In these three channels, the received signal has been shown experimentally to undergo fading. In addition, there are other types of fading channels such as very high frequency (VHF) communication channels between an aircraft and a synchronous satellite relay [Bond and Meyer, 1966] and line of sight (LOS) microwave links [Jakes, 1978], which undergo fading due to the formation of tropospheric inversion layers. This allows multiple transmission paths between the transmitter and receiver.

### **9.2.2 Characterisation of Fading Multipath Channels**

If an impulse is sent over a time-varying multipath channel, the received signal might appear as a train of pulses. Thus one characteristic of a multipath channel is the time spread introduced in the transmitted signal. A second characteristic is due to the time variations in the structure of the medium and as a result, the nature of the multipath varies with time. Thus if an impulse is sent over a channel, over and over again, we would observe changes in the received pulse train, such as changes in the size of individual pulses, changes in the relative delays among the pulses and changes in the number of pulses in the

pulse train [Bond and Meyer, 1966]. We can examine the effects of the channel on a transmitted signal represented by:

$$s(t) = \text{Re}[s_i(t)e^{j2\pi f_c t}] \quad (9.1)$$

Assuming that there are multiple propagation paths, there is a propagation delay and an attenuation factor. Thus the received bandpass signal may be expressed as:

$$x(t) = \sum_n \alpha_n(t) s(t - \tau_n(t)) \quad (9.2)$$

where  $\alpha_n(t)$  is the attenuation factor for the signal received on the  $n$ th path and  $\tau_n(t)$  is the propagation delay for the  $n$ th path. The equivalent lowpass received signal is:

$$r_i(t) = \sum_n \alpha_n(t) e^{-j2\pi f_c \tau_n(t)} s_i(t - \tau_n(t)) \quad (9.3)$$

The equivalent lowpass channel is described by the time-variant impulse response

$$c(\tau; t) = \sum_n \alpha_n(t) e^{-j2\pi f_c \tau_n(t)} \delta(\tau - \tau_n(t)) \quad (9.4)$$

When there are a large number of paths, the central limit theorem can be applied. Thus, the received signal  $r_i(t)$  can be modelled as a complex valued Gaussian random process which implies that the impulse response  $c(\tau; t)$  is also a complex-valued Gaussian random process in the  $t$  variable.

Rayleigh fading occurs when the impulse response  $c(\tau; t)$  is modelled as a zero-mean complex-valued Gaussian process. The envelope  $|c(\tau; t)|$  at any instant  $t$  is Rayleigh distributed. Ricean fading occurs when there are fixed scatterers or signal reflectors in the medium as well as randomly moving scatterers. The mean of the impulse response will not be zero and the envelope  $|c(\tau; t)|$  will have a Ricean distribution.

### 9.2.3 Rayleigh Fading

Rayleigh fading occurs on time varying multipath channels such as when the medium is time varying as in under-sea acoustic transmission. It can also occur with radio transmission through the upper atmosphere, mobile radio where the receiver and transmitter are in motion and indoor radio transmission where moving people cast shadows. In the case of mobile radio, the distances along the multiple propagation paths are changing and the

receiver observes the Doppler shifted versions of the transmitted signal. We can model the reception as [Lee and Messerschmidt, 1988]:

$$E(t) = \text{Re} \left\{ \sum_{n=1}^N \left( A_n e^{j\theta(t)} \right) e^{j\omega_n t} \right\} \quad (9.5)$$

where the amplitudes  $A_n$  vary slowly with time and hence they are considered to be fixed. The phases are varying rapidly because if there are vehicles on the move involved (as is often the case), the vehicle motions are large with respect to the transmitting wavelength. The phase can be modelled as:

$$\theta_n(t) = \omega_n t + \phi_n \quad (9.6)$$

where  $\phi_n$  are fixed random phases uniformly distributed from 0 to  $2\pi$  and the frequency offset  $\omega_n$  is the Doppler frequency shift due to the motion of the vehicle. The Doppler shift for a wave incident in the direction of motion is

$$\omega_n = \frac{2\pi}{\lambda} v \quad (9.7)$$

where  $\lambda$  is the wavelength and  $v$  is the velocity of the vehicle.

If equation (9.5) is written in terms of the real and imaginary parts of the complex exponentials, the resulting expression, in terms of quadrature components, is

$$E(t) = C(t) \cos \omega_c t - S(t) \sin \omega_c t \quad (9.8)$$

$$C(t) = \sum_{n=1}^N A_n \cos(\omega_n t + \phi_n), \quad S(t) = \sum_{n=1}^N A_n \sin(\omega_n t + \phi_n) \quad (9.9)$$

Since the terms in the summation are independent random variables, the baseband random processes  $C(t)$  and  $S(t)$  are approximately Gaussian according to the central limit theorem. The approximation becomes more accurate as the number of interferers  $N$  becomes large. Thus  $E(t)$  is Gaussian and the envelope is:

$$R(t) = \sqrt{C^2(t) + S^2(t)} \quad (9.10)$$

which has a Rayleigh distribution,

$$f_R(r) = \begin{cases} \frac{r}{\sigma} e^{-\frac{r^2}{2\sigma}}, & r \geq 0 \\ 0, & r < 0 \end{cases} \quad (9.11)$$

#### 9.2.4 DT Performance in Rayleigh Fading Conditions

Simulations are carried out in Matlab in a similar fashion to the simulations for the Gaussian channel. A Rayleigh fading channel is introduced instead of the Gaussian channel. The effect of the fading channel on the modulation classifiers is investigated for the DT and NN approaches.

The modulation classifier performance is evaluated for a Doppler spread of 120Hz. The key features and their corresponding threshold values for all signals do not change. The results for the 120Hz Doppler spread for SNR of 20dB are presented in Figure 9.1. The Doppler spread is chosen to be 120Hz as this is a reasonable value for mobile communications; in [Oon and Steele, 1997] Doppler frequencies of 150Hz were used.

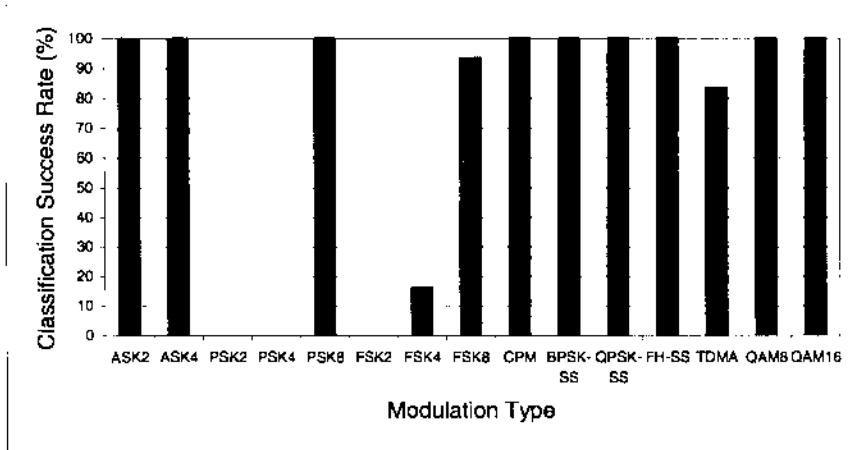


Figure 9.1. Modulation classifier performance in the presence of Rayleigh fading for SNR 20dB.

It can be seen from Figure 9.1 that the rate of classification is very bad for PSK2, PSK4, FSK2 and FSK4 signals with 120Hz Doppler spread. The performance also drops for the TDMA signal. Therefore, the decision tree is modified to accommodate signals undergoing fading.

### 9.3 Decision Tree Modifications for Rayleigh Fading

The decision tree in Figure 8.5. has key features based on cumulants to classify the PSK and TDMA signals. These cumulants are calculated using the complex envelopes of the signals. When fading is present, the envelope of the signal also diminishes in amplitude. Therefore, the features based on cumulants suffer performance degradation. To combat this limitation, new key features are introduced to classify the PSK and TDMA signals. This will be outlined in subsections 9.3.1 and 9.3.2, respectively.

The Doppler frequency also causes frequency shifts in the signal and this affects the classification of FSK signals. Therefore, the tree is modified with an additional key feature to improve the performance of classification of FSK signals in the presence of Rayleigh fading. This will be presented in subsection 9.3.3. The relevant threshold derivation is shown in subsection 9.3.4 and a discussion of the dependency of key feature selection on the minimum error probability is presented in subsection 9.3.5.

#### 9.3.1 PSK, BPSK DS-SS, and QPSK DS-SS Signal Classification

In the presence of Rayleigh fading, it is not possible to distinguish sufficiently between PSK2, BPSK DS-SS, and QPSK DS-SS signals as one group and PSK4 and PSK8 signals as another group, with the current tree structure. Therefore, the decision tree is modified to separate the spread spectrum signals from the PSK signals and then separate PSK2 from PSK4 and PSK8. To classify the spread spectrum signals, the key feature  $\gamma_{minf}$  is used. The feature  $\gamma_{minf}$  is used because it can distinguish between signals with more frequency information (such as spread spectrum signals) from signals with little or no frequency information (PSK signals).

To distinguish between PSK2 signals as one group and PSK4 and PSK8 signals as another group in the presence of Rayleigh fading, the key feature  $\sigma_{ap}$  is used. Rayleigh fading also introduces phase shifts to the signal. However, it is found that the effect on the overall instantaneous phase is not substantial because the faded signal gradually shifts alternately out of phase and then back in phase. Therefore, although this feature is based on phase, the phase offsets introduced by the fading should not affect the classification performance

drastically. The feature  $\sigma_{ap}$  is used because it can distinguish between signals with absolute phase information (PSK4 and PSK8) from signals without absolute phase information (PSK2).

To discriminate between PSK4 and PSK8 signals, the phase histogram is used. The histogram is formed using 50 bins and the value of the histogram at the 28<sup>th</sup> bin is found which corresponds to a phase of  $\pi/4$ . PSK4 signals use four phases ( $\pm\pi/2$  and  $\pm\pi$ ) to transmit information and PSK8 signals use eight phases ( $\pm\pi/4$ ,  $\pm\pi/2$ ,  $\pm3\pi/4$  and  $\pm\pi$ ). The 28<sup>th</sup> bin should contain no values for a signal with only four phases. Hence this feature  $P_{min}$  is used to separate PSK4 and PSK8 signals.

### 9.3.2 TDMA Classification

To separate TDMA signals from FH SS signals, the key feature  $\gamma_{maxf}$  is used. This feature is found to be suitable because in Chapter 7 it gave an overall error probability of 0.0286 for the SNR range of 20dB to -5dB for the threshold value of 19.3. This feature is also less susceptible to fading because it is not amplitude dependent.

### 9.3.3 FSK Signal Classification

To distinguish between FSK2 as one group and FSK4, FSK8, and CPM as another group, the existing key feature  $L_{diff}$  is used. This is because when fading is present, the key feature  $\sigma_{fn}$  is not sufficient to separate these signals due to the frequency shifts. However, since the same key feature is used to separate FSK4 and FSK8 signals when there is no fading present, another key feature must be used simultaneously. This is to ensure that the feature  $L_{diff}$  is only used when fading is present to separate FSK2 as one group and FSK4 and FSK8 as the other group. It is found that when the FSK signals undergo Rayleigh fading at 120Hz Doppler frequency, the feature  $\sigma_{fn}$  is able to determine whether fading is present.

The key feature  $\gamma_{maxf}$  is used to separate FSK4 and FSK8 signals when Rayleigh fading is present. This is because FSK8 signals have the same maximum instantaneous frequency values as FSK4 but they also have four other frequency values which are smaller due to the

eight frequency levels implemented in FSK8. Therefore, the values of  $\gamma_{maxf}$  should in general be greater for FSK4 signals than for FSK8 signals.

### 9.3.4 Threshold Determination

The total error probability for the key feature  $\gamma_{minf}$  for subset A (BPSK DS-SS and QPSK DS-SS) and subset B (PSK2, PSK4, PSK8) is shown in Figure 9.2. It can be seen that the appropriate value for the threshold  $t\gamma_{minf3}$  is  $-29.9$  where the total minimum error probability is 0.0009 for the SNR range of 20dB to  $-5$ dB.

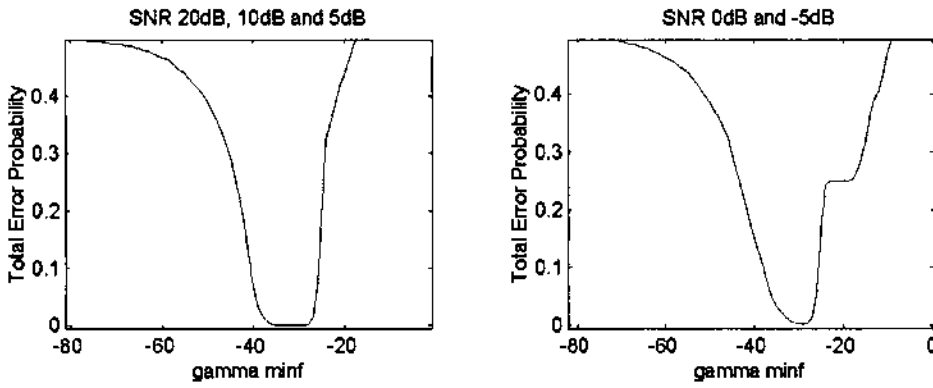


Figure 9.2. Total Error probability for the key feature  $\gamma_{minf}$ , at SNR range of 20dB to  $-5$ dB, for BPSK DS-SS and QPSK DS-SS (subset A) and PSK2, PSK4, and PSK8 (subset B) with fading and 120Hz Doppler shift.

The total error probability for the key feature  $\sigma_{ap}$  is shown in Figure 9.3 for subset A (PSK2) and subset B (PSK4 and PSK8). It can be observed from the figure that a good choice for the threshold  $t\sigma_{ap2}$  is 1.08 where the total minimum error is 0.0106 for the SNR range of 20dB to 5dB and 0.2784 for the SNR of 0dB and  $-5$ dB. It is found that the feature  $\sigma_{ap}$  is not affected by Rayleigh fading. Therefore this feature is sufficient to discriminate PSK2, PSK4 and PSK8 even in a Gaussian channel. This is demonstrated in Figure 9.4 where the optimum threshold is also 1.08. The minimum error probability for SNR range of 20dB to 5dB is 0.0063. The minimum error probability for SNR of 0dB and  $-5$ dB at the same threshold is 0.28.

The ROC curves for this key feature are shown in Figure 9.5 for the SNR range of 20dB to -5dB. The curves show the detection probability of subset A (PSK2) and false alarm probability of subset B (PSK4 and PSK8). By examining the ROC curves for  $\text{SNR} \geq 5\text{dB}$ , we can see that the chosen threshold value  $t\sigma_{ap2}$  (indicated by 'x') has a detection probability ( $P_D$ ) of 0.9875 and false alarm probability ( $P_{FA}$ ) of 0.0113 at 5dB SNR.

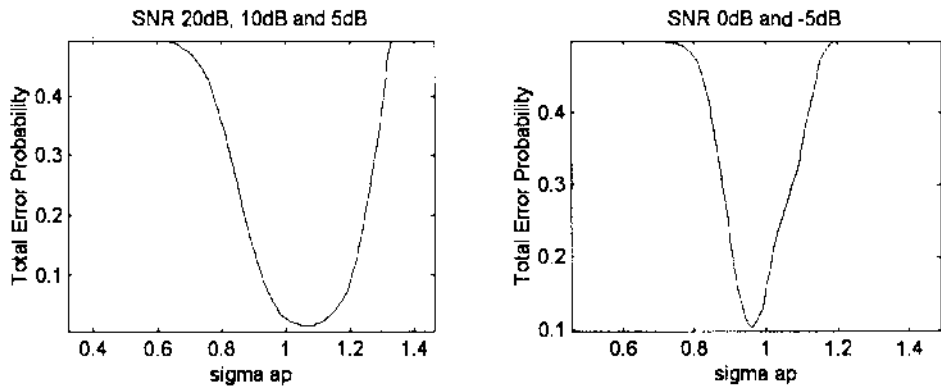


Figure 9.3. Total error probability for the key feature  $\sigma_{ap}$  , at SNR range of 20dB to -5dB, for PSK2 (subset A) and PSK4 and PSK8 (subset B) with fading and 120Hz Doppler shift.

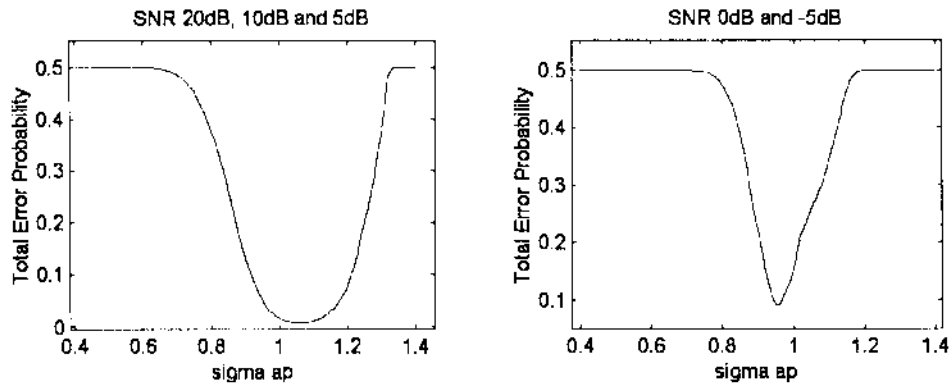


Figure 9.4. Total error probability for the key feature  $\sigma_{ap}$  , at SNR range of 20dB to -5dB, for PSK2 (subset A) and PSK4 and PSK8 (subset B) in a Gaussian channel.



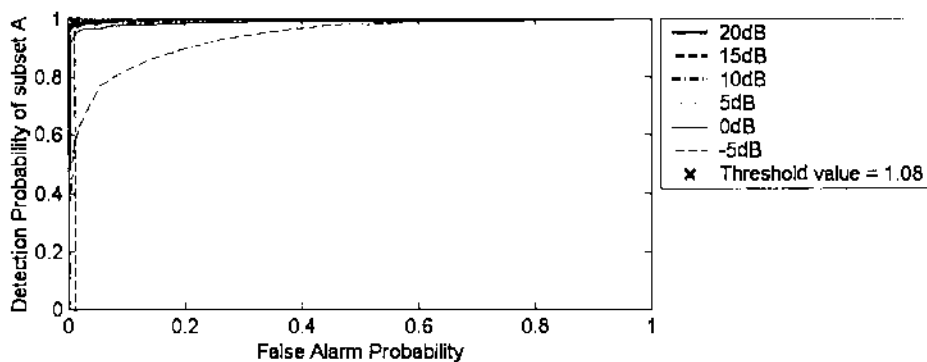


Figure 9.5. ROC curves for the key feature  $\sigma_{ap}$  to separate PSK2 (subset A) and PSK4 and PSK8 (subset B) signals with fading and 120Hz Doppler shift for SNR range of 20dB to -5dB.

The threshold value  $tP_{\min}$  is found from the total minimum error probability for the key feature  $P_{\min}$ . The total error probability is plotted in Figure 9.6 for PSK8 (subset A) and PSK4 (subset B) signals. It can be observed from the figure that a good choice for the threshold  $tP_{\min}$  is 26 where the minimum error is 0.1329 for the SNR range of 20dB to 5dB and 0.4744 for SNR of 0dB and -5dB. The ROC curves for this key feature are shown in Figure 9.7 for the SNR range of 20dB to -5dB. The curves show the detection probability of subset A (PSK8) and false alarm probability of subset B (PSK4). By examining the ROC curves for  $SNR \geq 10dB$ , we can see that the chosen threshold value  $tP_{\min}$  (indicated by 'x') has a detection probability ( $P_D$ ) of 0.7950 and false alarm probability ( $P_{FA}$ ) of 0.02 at 10dB SNR.

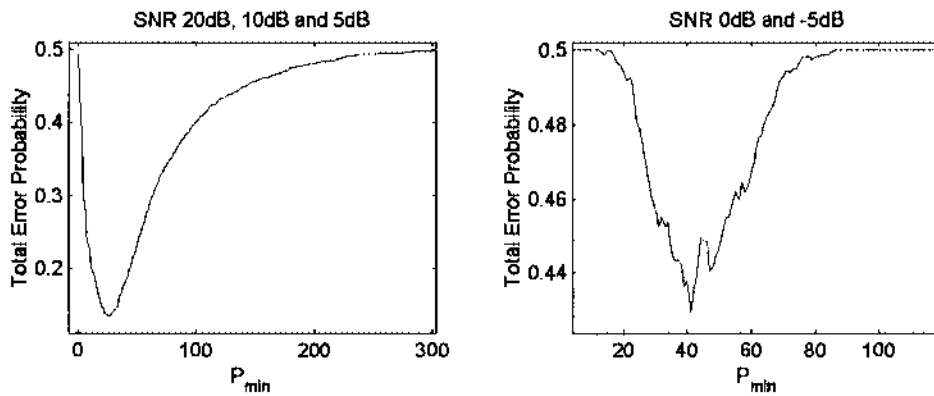


Figure 9.6. Total error probability for the key feature  $P_{min}$ , at SNR range of 20dB to -5dB, for PSK8 (subset A) and PSK4 (subset B) with fading and 120Hz Doppler shift.

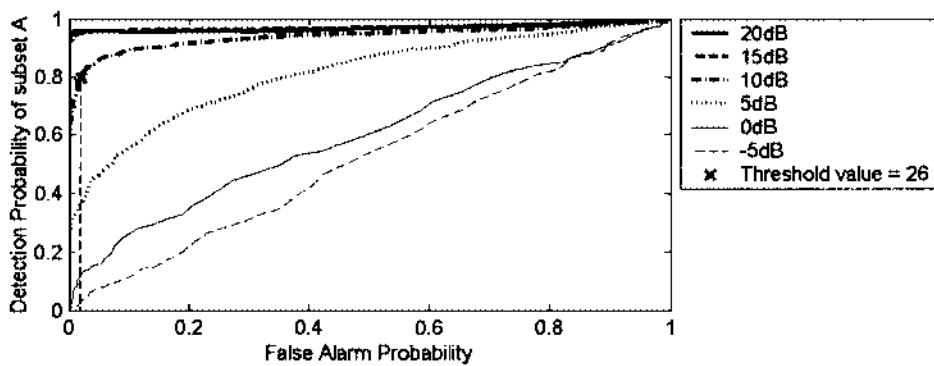


Figure 9.7. ROC curves for the key feature  $P_{min}$  to separate PSK8 (subset A) and PSK4 (subset B) signals with fading and 120Hz Doppler shift.

To separate FH SS signals from TDMA signals when fading is present, the key feature  $\gamma_{maxf}$  is used. The total error probability for the SNR range of 20dB to -5dB is shown in Figure 9.8. It is found that the optimum threshold value  $\gamma_{maxf3}$  is 44, which gives a minimum error probability of 0 for the SNR range of 20dB to -5dB.

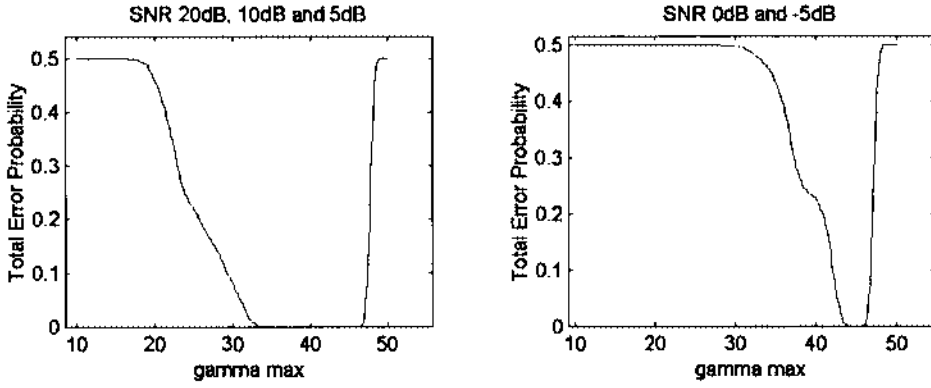


Figure 9.8. Total error probability for the key feature  $\gamma_{maxf}$ , at SNR range of 20dB to -5dB, for TDMA (subset A) and FH SS (subset B) with fading and 120Hz Doppler shift.

To determine whether fading is present for FSK signals, the key feature  $\sigma_{fn}$  is used. The total error probability for subset A (FSK2, FSK4, and FSK8 when fading is not present) and subset B (FSK2, FSK4, and FSK8 when fading is present) is shown in Figure 9.9. The optimum threshold occurs at  $t\sigma_{fn2} = 2.4$  where the total minimum error probability is 0.0629 for the SNR range of 20dB to 5dB.

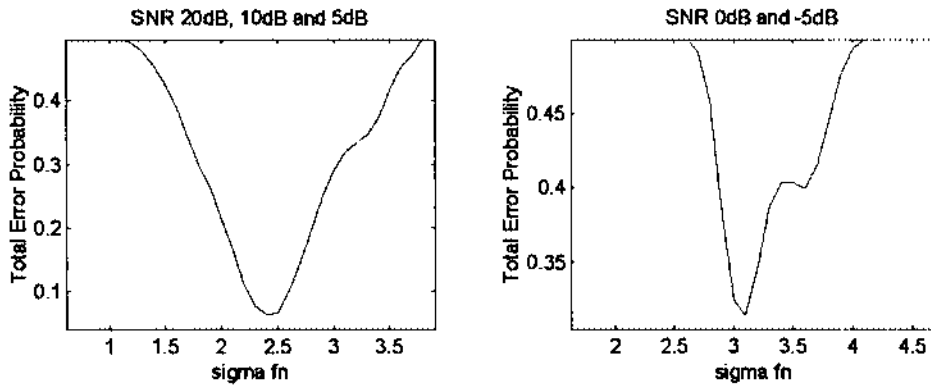


Figure 9.9. Total error probability for the key feature  $\sigma_{fn}$ , at SNR range of 20dB to -5dB, for FSK2, FSK4, FSK8, and CPM with no fading (subset A) and FSK2, FSK4, FSK8, and CPM (subset B) with fading and 120Hz Doppler shift.

The total error probability for the key feature  $L_{diff}$  is shown in Figure 9.10 for subset A (FSK2) and subset B (FSK4 and FSK8). It is found that the optimum threshold  $t_{L_{diff}3}$  is  $-7$  where the minimum error probability is 0.0592 for the SNR range of 20dB to 5dB and 0.4440 for SNR of 0dB and  $-5$ dB.

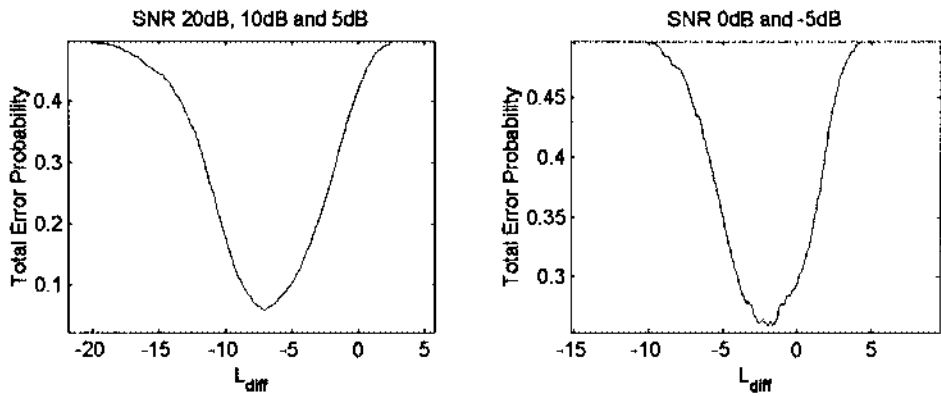


Figure 9.10. Total error probability for the key feature  $L_{diff}$ , at SNR range of 20dB to  $-5$ dB, for FSK2 (subset A) and FSK4, FSK8, and CPM (subset B) with fading and 120Hz Doppler shift.

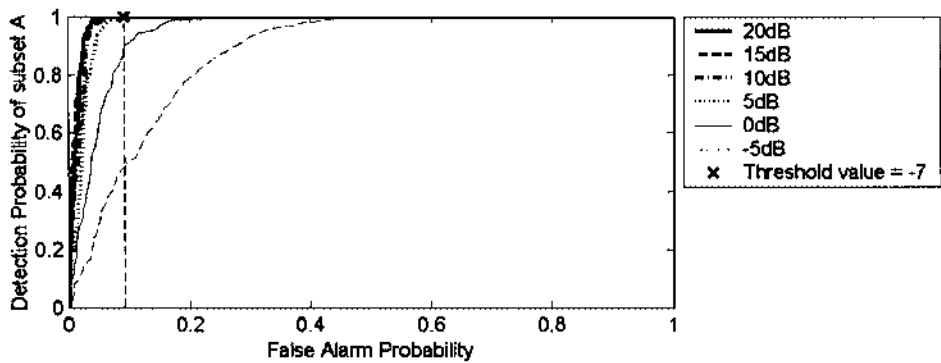


Figure 9.11. ROC curves for the key feature  $L_{diff}$  to separate FSK2 (subset A) and FSK4, FSK8, and CPM (subset B) signals with fading and 120Hz Doppler shift for SNR range of 20dB to  $-5$ dB.

The ROC curves for the key feature  $L_{diff}$  are shown in Figure 9.11 for the SNR range of 20dB to  $-5$ dB. The curves show the detection probability of FSK2 (subset A) and the false

alarm probability of FSK4, FSK8, and CPM (subset B). By examining the ROC curves for  $\text{SNR} \geq 10\text{dB}$ , we can see that the chosen threshold value  $t_{L_{\text{diff3}}}$  (indicated by 'x') has a detection probability ( $P_D$ ) of 1.0 and false alarm probability ( $P_{FA}$ ) of 0.0917 at 10dB SNR.

To classify FSK4 (subset A) and FSK8 (subset B) signals, the threshold value,  $t_{\gamma_{\text{maxf4}}}$ , is determined from the total error probability shown in Figure 9.12. The minimum error probability is 0.3942 for the SNR range of 20dB to 5dB, corresponding to a threshold value of  $t_{\gamma_{\text{maxf4}}} = 20$ . When fading is present, it becomes much harder to separate FSK signals as can be seen from the total error probability in Figure 9.12. The ROC curves for the key feature  $\gamma_{\text{maxf}}$  are shown in Figure 9.13 for the SNR range of 20dB to -5dB. The ROC curves show the difficulty in separating FSK4 and FSK8 signals since the probability of detection ( $P_D$ ) of FSK4 (subset A) is not so high for low probability of false alarm ( $P_{FA}$ ) of FSK8 (subset B). By examining the curves for  $\text{SNR} \geq 10\text{dB}$ , we can see that the chosen threshold value  $t_{\gamma_{\text{maxf4}}}$  (indicated by 'x') has a detection probability ( $P_D$ ) of 0.81 and false alarm probability ( $P_{FA}$ ) of 0.5125 at 10dB SNR.

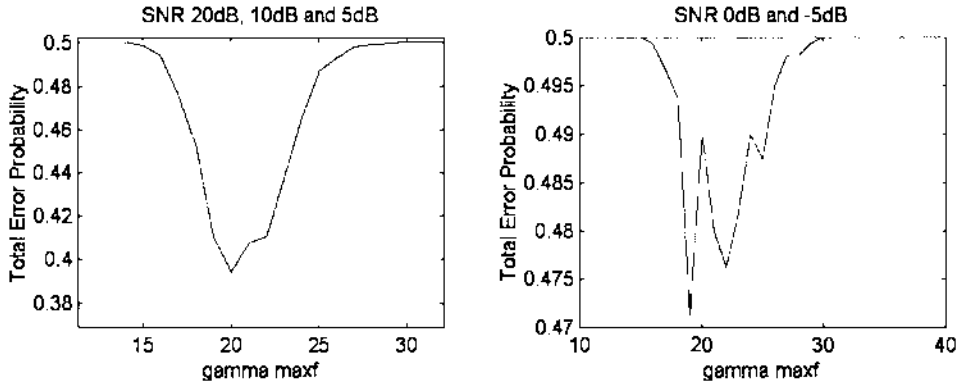


Figure 9.12. Total error probability for the key feature  $\gamma_{\text{maxf}}$ , at SNR range of 20dB to -5dB, for FSK4 (subset A) and FSK8 (subset B) with fading and 120Hz Doppler shift.

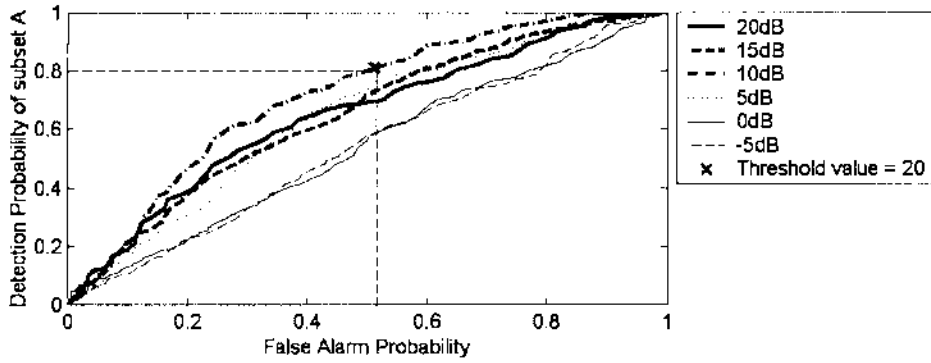


Figure 9.13. ROC curves for the key feature  $\gamma_{maxf}$  to separate FSK4 (subset A) and FSK8 (subset B) signals with fading and 120Hz Doppler shift for SNR range of 20dB to -5dB.

A summary of the key features for the modified decision tree to accommodate Rayleigh fading is shown in Table 9.1. The total minimum error probabilities and corresponding optimum thresholds for the SNR ranges of 20dB to 5dB and 0dB to -5dB are also shown.

Table 9.1. Summary of key feature thresholds and error probabilities.

Key Feature Threshold	SNR 20dB to 5dB		SNR 0dB to -5dB	
	Optimum Threshold	Minimum Error Probability	Optimum Threshold	Minimum Error Probability
$t\gamma_{minf3}$	-33	0	-29	0.0014
$tP_{min}$	26	0.1329	41	0.4294
$t\sigma_{fn2}$	2.4	0.0629	3.1	0.3140
$t\gamma_{maxf4}$	20	0.3942	19	0.4713
$tL_{diff3}$	-7	0.0592	-2	0.2584
$t\gamma_{maxf3}$	44	0	44	0
$t\sigma_{ap2}$	1.08	0.0106	0.98	0.0925

Therefore the optimum threshold values are:  $t\sigma_{ap2} = 1.08$ ,  $tP_{min} = 26$ ,  $t\sigma_{fn2} = 2.4$ ,  $t\gamma_{maxf3} = 44$ ,  $tL_{diff3} = -7$ ,  $t\gamma_{minf3} = -29.9$ , and  $t\gamma_{maxf4} = 20$ .

### 9.3.5 Dependency of Key Feature Selection on Minimum Probability of Error

The reason why the key features in the previous section are chosen over the other existing key features is that they minimise the total error probability for each decision. We will call the decision separating BPSK DS-SS, and QPSK DS-SS (subset A) from PSK2, PSK4 and PSK8, (subset B) decision 1. Decision 2 separates subset A (PSK8) and subset B (PSK4) and decision 3 distinguishes between fading being present for FSK2, FSK4, FSK8, and CPM (subset A) and fading not being present for the same signals (subset B). Decision 4 is defined as the classification of FSK2 (subset A) and FSK4, FSK8, and CPM (subset B) and Decision 5 separates FSK4 (subset A) and FSK8 (subset B). Decision 6 is defined as the classification of TDMA (subset A) and FH-SS (subset B) and decision 7 is the classification of PSK2 (subset A) and PSK4 and PSK8 (subset B). We can see from Table 9.2 and Table 9.3 that the key features that have been chosen minimise the total error probability for each decision for the SNR range of 20dB to -5dB.

It can be seen in decision 5, that the feature  $L_{diff}$  has the smallest error probability. However, this feature is not chosen because it is very sensitive to fading for this particular decision and will vary for different Doppler frequencies. Therefore, the feature  $\gamma_{maxf}$  is chosen for this decision instead because it is unaffected by fading. The modified decision tree to accommodate signals undergoing Rayleigh fading with Doppler spread of 120Hz is shown in

Figure 9.14. The first modification to the tree occurs where FH SS and TDMA are separated by the key feature  $\gamma_{maxf}$  if fading is present. The second modification occurs where the feature  $\sigma_{fn}$  is used to determine if fading is present for FSK signals. If fading is present the feature  $L_{diff}$  separates FSK2 from FSK4, FSK8, and CPM. The third change to the decision tree is where the feature  $\sigma_{ap}$  separates PSK2 from PSK4 and PSK8 when fading is present. Similarly, the feature  $P_{min}$  is used instead of  $|\hat{C}_{40}|$  to separate PSK4 from PSK8 in a fading channel. The final change in the tree occurs where FSK4 and FSK8 are separated by the feature  $\gamma_{maxf}$  in the presence of fading. The NN classifier modifications to accommodate fading are presented in the next section.

## 9.4 NN Classifier Modifications for Rayleigh Fading

The modulation classifier performance using the NN approach is investigated for a Doppler spread of 120Hz. The neural network structure shown in Figure 8.17 is slightly modified to accommodate fading signals. The modified structure is described in subsection 9.4.1. Subsection 9.4.2 describes how the NN classification works when the channel is unknown. This network structure is trained for modulation types with SNR range 20dB to -5dB.

### 9.4.1 Modified Neural Network Structure to Accommodate Rayleigh Fading

The neural network structure shown in Figure 9.15 has some modifications within the sub-nets. All of the modified networks have been trained with signals undergoing Rayleigh fading with Doppler spread 120Hz and SNR range of 20dB to -5dB. Network four is modified to have the input key feature  $P_{\min}$  instead of the feature  $\sigma_{dp}$ . Network six is retrained with data of SNR range of 20dB to -5dB that has been affected by Rayleigh fading. The structure giving the optimum performance is modified to have one hidden layer with seven neurons. The remaining sub-nets remain the same.

### 9.4.2 Neural Network Classifier for AWGN and Rayleigh Fading Channel

When a signal is intercepted and the channel is unknown, the signal can be simultaneously passed through the networks in 8.17 and Figure 9.15. The classification can be achieved by choosing the signal with the highest success rate from the two networks. The results of the DT and NN classifiers in the presence of fading will be discussed in the next section.

## 9.5 Performance Analysis of DT and NN Classifiers in the Presence of Rayleigh Fading

A graphical comparison between the overall classification accuracy of the DT and NN classifiers in the presence of fading is shown in Figure 9.16. It can be inferred that the NN implementation generally performs better than the DT algorithm. The performance results for the classification of each modulation type using the DT and NN approaches are shown in Figure 9.17 - Figure 9.22, for the SNR range of 20dB to -5dB, respectively. It can be seen that most of the signals can be classified with success rates greater than 80% for SNR greater than or equal to 5dB with the exception of FSK8 signals.



Table 9.2. Total minimum error probability for Decisions 1 – 3 for combined SNR range of 20dB to –5dB (threshold values are shown in brackets).

Key Feature	Total Minimum Error Probability (Decision 1)	Total Minimum Error Probability (Decision 2)	Total Minimum Error Probability (Decision 3)
$\gamma_{maxf}$	0.2953 (-102.6)	0.5000 (-92)	0.2724 (24.2)
$\mu_{dp}$	0.3938 (0.34)	0.4053 (0.08)	0.4502 (44)
$ \hat{C}_{21} $	0.1786 (0.96)	0.4762 (0.6)	0.2385 (0.9)
$ \hat{C}_{40} $	0.0378 (1.07)	0.2667 (0.1)	0.4267 (0.02)
$\sigma_{fn}$	0.3748 (0)	0.5000 (0)	<b>0.2378 (2.4)</b>
$\sigma_{ap}$	0.1823 (0.82)	0.4342 (0.91)	0.3892 (435)
$\gamma_{minf}$	<b>0.0009 (-29)</b>	0.4133 (-45.4)	0.4206 (-39)
$L_{diff}$	0.3221 (-2.02)	0.4787 (-2.9)	0.4173 (-5.0)
$\sigma_{dp}$	0.4466 (1.57)	0.4700 (1.91)	0.3866 (435)
$P_{min}$	0.2015 (13.72)	<b>0.2515 (26)</b>	0.4821 (12)

It can also be observed that the DT classifier performs better than the NN classifier for some signals such as FSK signals when a fading channel is present. With the DT classifier, the performance of the FSK4 signal gets worse while the FSK8 classification performance gets better with decreasing SNR. This is because the values of the key feature  $\gamma_{maxf}$  are very similar for FSK4 and FSK8 signals and therefore it is harder to separate these two signals in the presence of fading, so if the performance of one signal is good, the performance of the other suffers.

The NN classifier performs better than the DT classifier for signals such as PSK2, BPSK-SS, and QPSK-SS signals. For all other signals, the results for the NN classifier and the DT classifier are comparable when fading is present. By referring to the results for the Gaussian channel in Chapter 8, it can be seen that both the NN and DT classifier performances suffer in the presence of Rayleigh fading.

Table 9.3. Total minimum error probability for Decisions 4 – 7 for combined SNR range of 20dB to –5dB (threshold values are shown in brackets).

Key Feature	Total Minimum Error Probability (Decision 4)	Total Minimum Error Probability (Decision 5)	Total Minimum Error Probability (Decision 6)	Total Minimum Error Probability (Decision 7)
$\gamma_{maxf}$	0.2451 (23)	<b>0.4325 (20)</b>	<b>0 (44)</b>	0.2562 (-98.7)
$\mu_{dp}$	0.2500 (100)	0.4635 (-1.0)	0.0002 (455)	0.2255 (0.57)
$ \hat{C}_{21} $	0.2500 (0.35)	0.4373 (0.8)	0.3028 (0.98)	0.3813 (1.4)
$ \hat{C}_{40} $	0.3332 (0.4)	0.5000 (0)	0.0539 (0.16)	0.2587 (0.38)
$\sigma_{fn}$	0.2500 (4)	0.4443 (1.7)	0 (9.5)	0.2532 (0)
$\sigma_{ap}$	0.2500 (200)	0.4902 (460)	0 (250)	<b>0.0893 (1.08)</b>
$\gamma_{minf}$	0.2500 (-32)	0.4650 (-44)	0.1440 (-14)	0.3333 (-30)
$L_{diff}$	<b>0.1325 (-7)</b>	0.3762 (-1.0)	0.0011 (-4.434)	0.3333 (10)
$\sigma_{dp}$	0.2500 (100)	0.4900 (465)	0 (250)	0.3150 (2)
$P_{min}$	0.2500 (175)	0.4900 (70)	0.2300 (20)	0.3333 (357)

A tabular comparison of the overall performance of the DT and NN classifiers is shown in Table 9.4. It can be seen that the NN classifier performs on par with the DT classifier for SNR greater than 5dB. However, for lower SNR, the NN classifier outperforms the DT classifier considerably. The confusion matrices showing the results for the DT classifier are in Appendix E, Table E.1 – Table E.6, and the results for the NN classifier are in Table E.7 – Table E.12 for the SNR range of 20dB to –5dB respectively.

A graphical comparison between the NN and DT classifiers in both AWGN and fading environments is shown in Figure 9.23. In general, it can be observed that the NN classifiers perform slightly better than the DT classifiers for both AWGN and fading channels. Also, the performance of both classifiers degrades in the presence of fading.

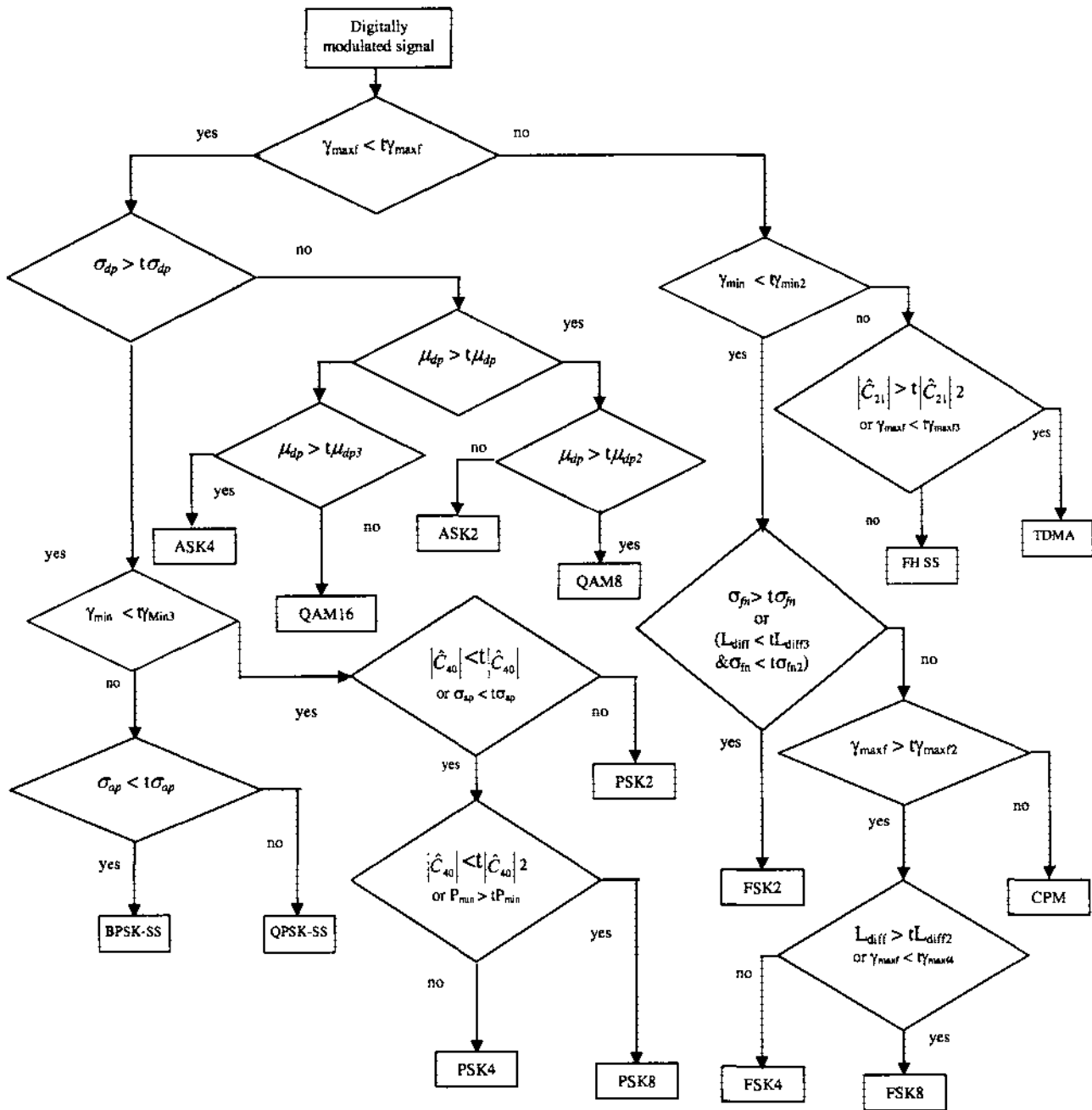


Figure 9.14. Modified decision tree to accommodate signals in the presence of Rayleigh fading. Refer to section 9.3.5 for an explanation of the modifications to the tree.

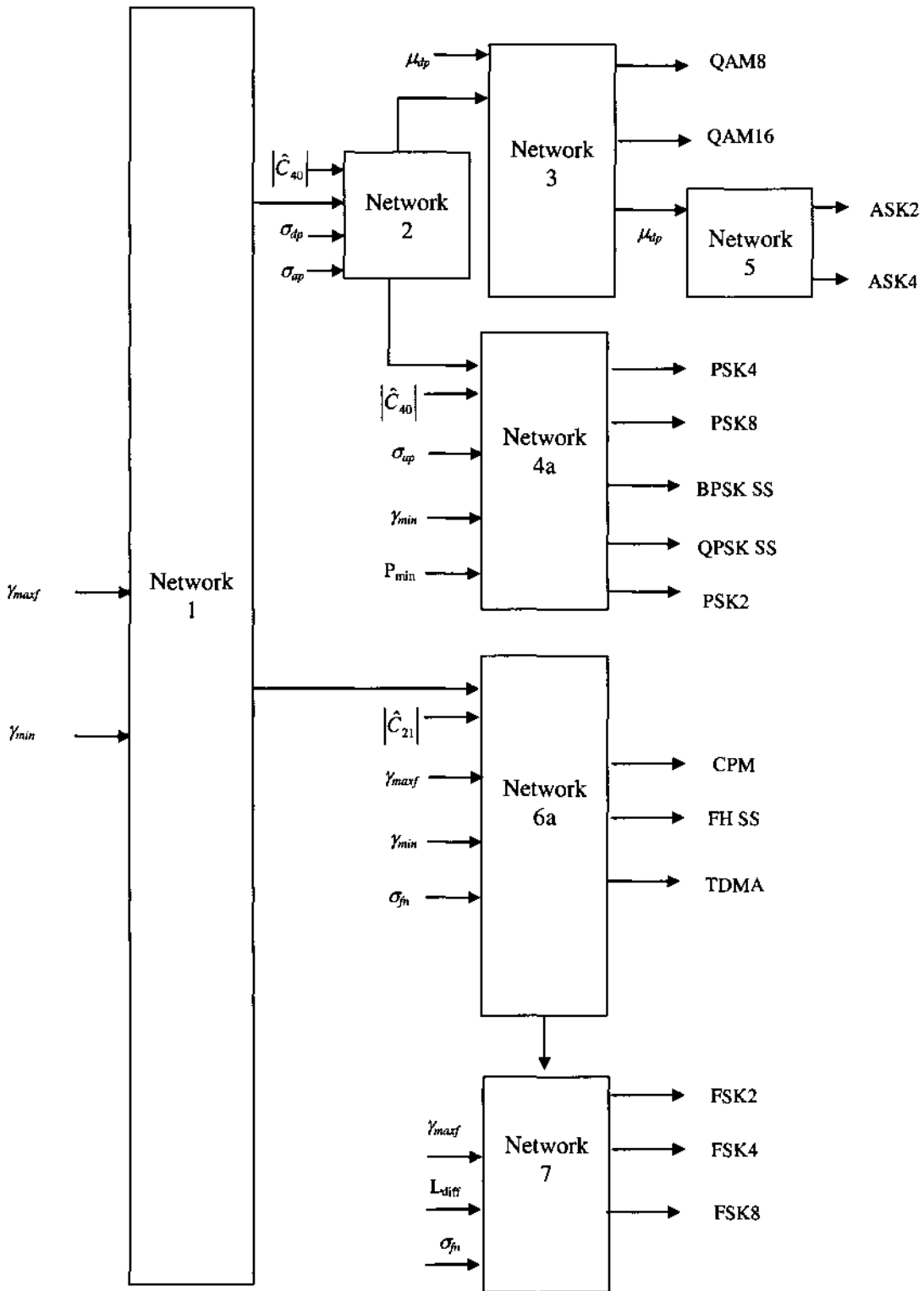


Figure 9.15. Modified neural network structure for digital modulation classifier in the presence of Rayleigh fading.

Table 9.4. DT and NN classifier accuracy and 95% confidence intervals in the presence of Rayleigh fading.

SNR	DT Classifier		NN Classifier	
	Accuracy	95% Confidence Interval	Accuracy	95% Confidence Interval
20dB	91.97%	[91.28, 92.65]	93.8%	[93.19, 94.41]
15dB	92.13%	[91.45, 92.81]	94.98%	[94.42, 95.53]
10dB	91.5%	[90.79, 92.21]	93.81%	[93.20, 94.42]
5dB	84.1%	[83.17, 85.03]	89.93%	[89.17, 90.69]
0dB	75.42%	[74.33, 76.51]	81.35%	[80.36, 82.34]
-5dB	50.43%	[49.17, 51.70]	70.19%	[69.03, 71.34]
Overall	80.93%	[80.52, 81.33]	87.34%	[87.00, 87.69]

## 9.6 Conclusions

The classifiers presented in Chapter 8 were tested in the presence of Rayleigh fading. It was found that the classifiers had to be modified slightly to accommodate fading channels. The performance of the DT modulation classifier suffered significantly for some signals (namely PSK and FSK) with a Doppler spread of 120Hz. This resulted in some modifications to the existing decision tree with the addition of new key features. The performance of the DT classifier with signals undergoing fading was good for SNR down to 10dB. However the classification of FSK8 signals was poor. For SNR of 5dB, the performance dropped significantly for PSK4, FSK2, and FSK8 signals.

The neural network classifier was tested with a Doppler spread of 120Hz. Certain sub-net structures were retrained with modified structures and inputs to improve classification performance. The performance was good for SNR greater than 10dB, however for lower SNR, the results were only marginally satisfactory. The results were compared to the DT classifier and it was found that for certain signals, the NN outperforms the DT classifier whereas for other signals, the DT classifier gives better results. However, the NN has

better overall performance than the DT approach over the SNR range of 20dB to -5dB. The classifiers' performances in a fading environment were also compared to the performances in an AWGN channel. It was found that the presence of fading causes the classification performance to suffer significantly more than in an environment where fading is not present.

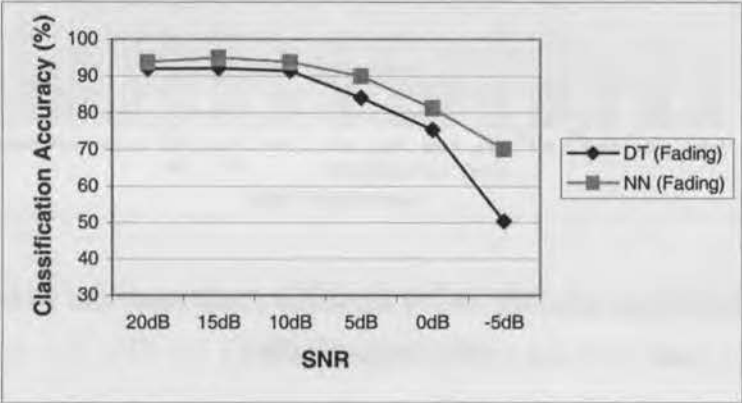


Figure 9.16. Graphical comparison of overall performance between the NN-based and DT-based classifiers for Rayleigh fading.

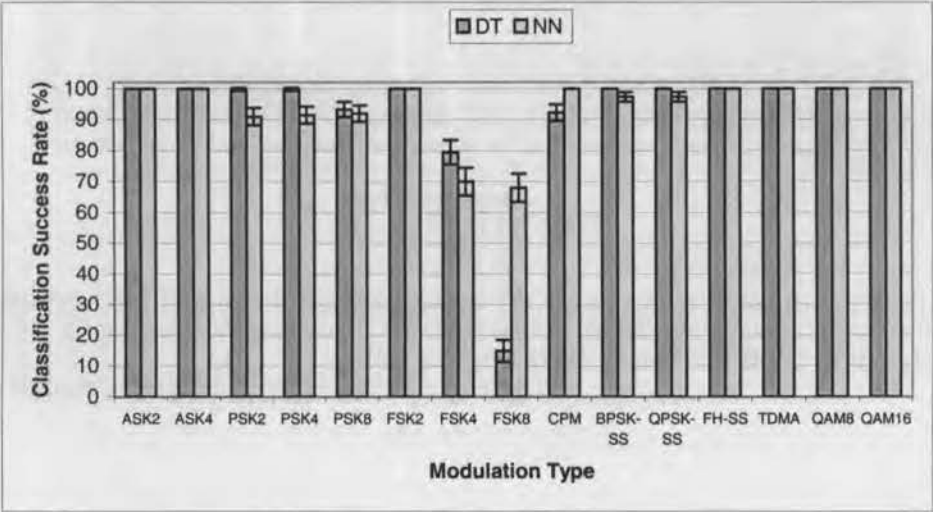


Figure 9.17. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 20dB SNR and 120Hz Doppler shift.

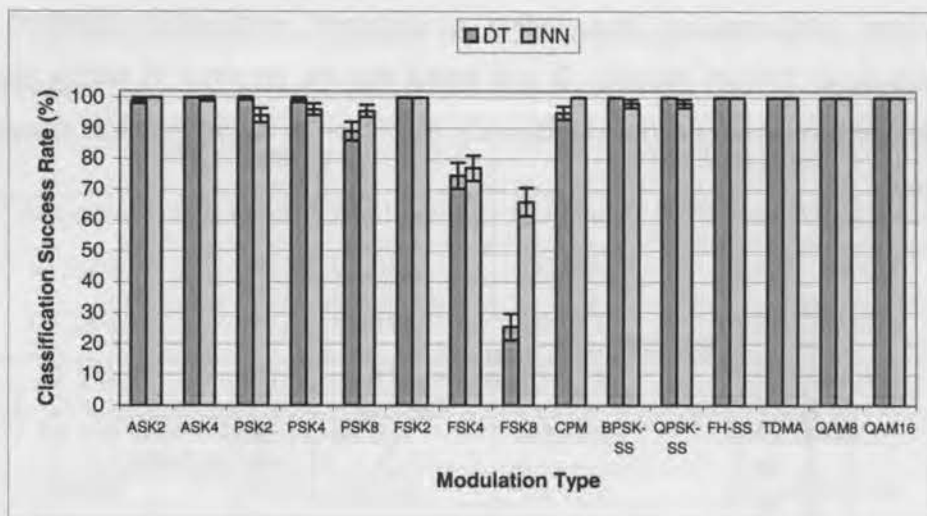


Figure 9.18. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 15dB SNR and 120Hz Doppler shift.

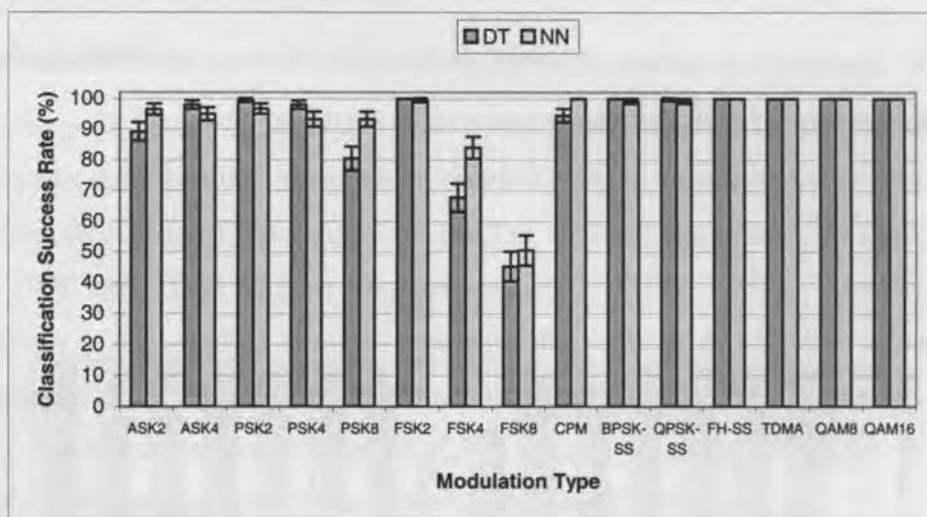


Figure 9.19. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 10dB SNR and 120Hz Doppler shift.

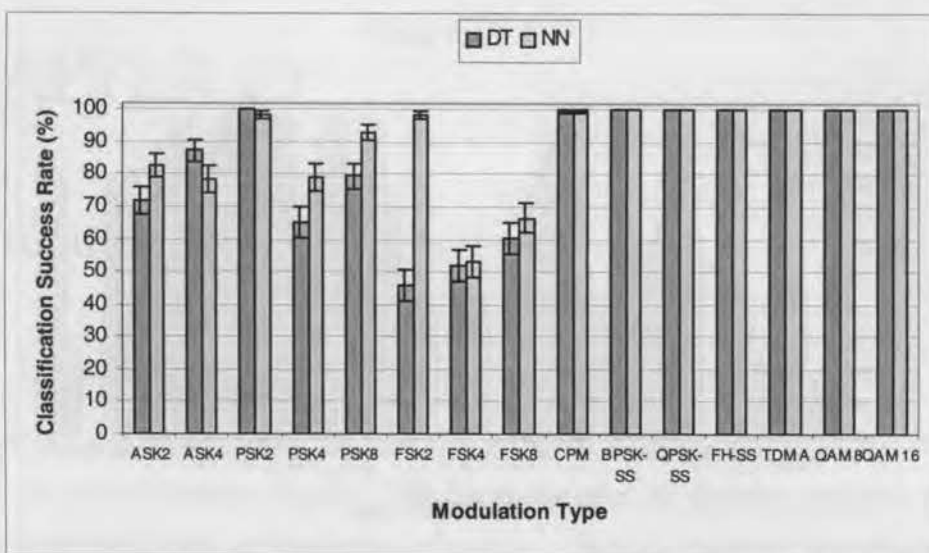


Figure 9.20. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 5dB SNR and 120Hz Doppler shift.

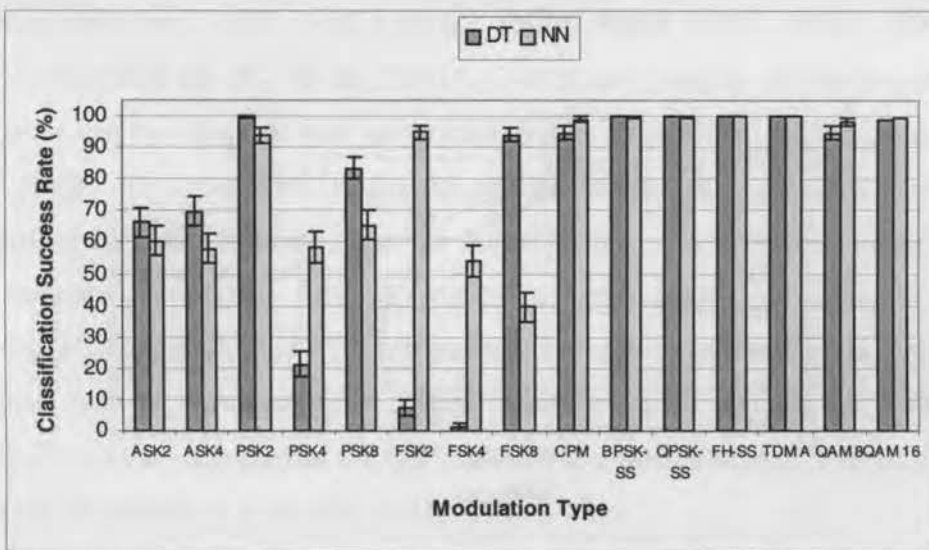


Figure 9.21. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at 0dB SNR and 120Hz Doppler shift.



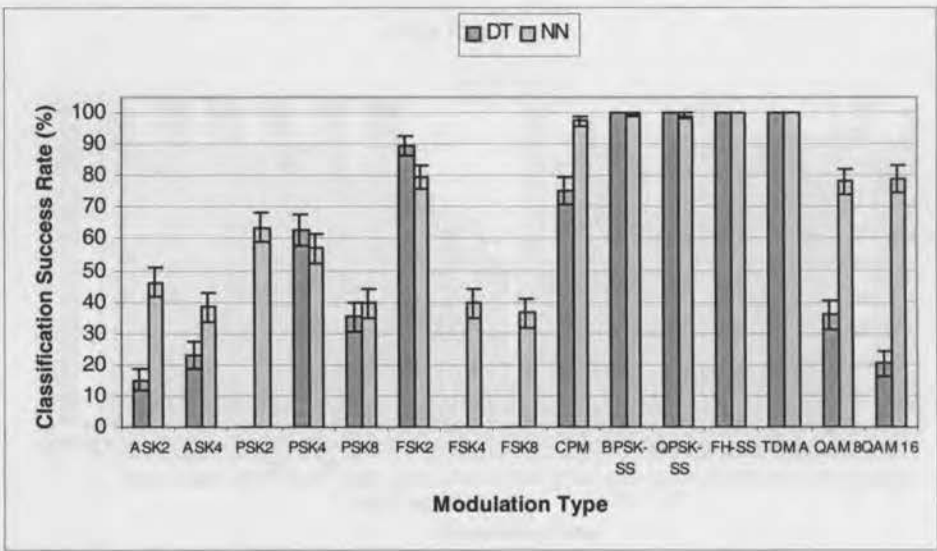


Figure 9.22. Classification accuracy of DT classifier (dark bars) and NN classifier (light bars) for signals at -5dB SNR and 120Hz Doppler shift.

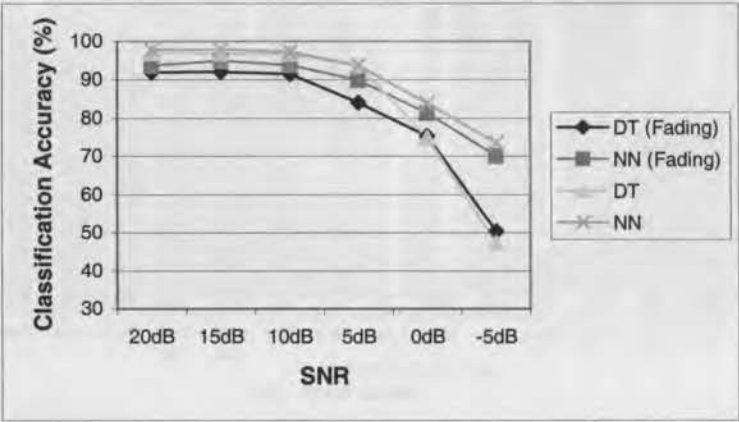


Figure 9.23. Comparison of overall performance between the NN-based and DT-based classifiers for Rayleigh fading and AWGN channels.

# CHAPTER 10

## Conclusion

### 10.1 Introduction

A framework has been presented in this thesis for the classification of digital modulation schemes of communications signals. The focus has been on decision theoretic and neural network implementations of modulation classifiers. New key features have been proposed to classify the signals and for the first time, fifteen different digital modulation types can be classified by one type of modulation classifier. These modulation schemes have been added to the modulation classifiers gradually in Chapters 5, 6, 7 and 8. The fifteen modulation types are: ASK2, ASK4, PSK2, PSK4, PSK8, FSK2, FSK4, FSK8, CPM, BPSK DS-SS, QPSK DS-SS, FH-SS, TDMA, QAM8, and QAM16. It has been shown that these signals can be classified with accuracies greater than 95% for SNR greater than or equal to 10dB. For lower SNR values the performance drops as can be expected. It is found that as more modulation types are added to the classifiers, the classification of signals becomes increasingly difficult, particularly with signals belonging to the same family (eg FSK2, FSK4, and FSK8 signals). For NN classifiers, it is found that a hierarchical network structure gives better results as more signals are added to the classifier. The overall accuracy of the NN classifier, over the combined SNR range of 20 to -5dB, is 90.7% compared to 84.56% for the DT classifier.

The performance of the DT and NN classifiers were also tested in the presence of Rayleigh fading with 120Hz Doppler shift. It was found that fading mainly affects key features which are dependent on the complex envelope of the signal or power spectral density. Some modifications to the classifiers had to be made so that they were capable of classifying signals in both an AWGN environment and Rayleigh fading environment. The

performance is generally slightly worse for fading channels compared to AWGN channels. With the modifications, the overall accuracy of the NN classifier, over the combined SNR range of 20 to -5dB and 120Hz Doppler shift, is 87.34% compared to 80.52% for the DT classifier.

A point to consider is that the classification accuracy is based on a single segment of the intercepted signal. In a real life situation, the decision will be based on a number of segments and therefore we expect the classification accuracy to improve; a signal classification accuracy of over 50% (from all the segments) will probably guarantee the correct recognition of the modulation type.

There are many factors which have not been addressed with regards to modulation classification in this thesis. There is also room for improvement with the classification techniques that have been discussed in this thesis. These issues are outlined in the next section.

## **10.2 Suggestions for Further Work**

This thesis is a first attempt at classifying a large range of digital modulation schemes and therefore leaves much room for improvement. Some suggestions for further research and improvement are listed below in no particular order:

1. Investigation into making threshold values dynamically changing with varying SNR would greatly improve the performance of the DT classifier.
2. The former point follows on from the suggestion of finding methods to determine the SNR of the unknown signal so that more accurate threshold values can be used to classify the signal.
3. Investigation into making key features more robust against varying SNR would improve the performance of classification of signals with very low SNR.
4. For the classification of FSK signals, the bandlimitation greatly hinders the classification performance. Therefore, further research into finding features that are not greatly affected by bandwidth would help improve performance.

5. It is found that features that are robust against phase offsets (such as features based on cumulants) are sensitive to fading channels. The converse also applies with features such as  $\sigma_{ap}$  that are robust in the presence of fading, are affected by phase offsets and variations. Therefore, further investigation into feature extraction and robustness under different conditions should be carried out to improve performance.
6. The effects of signal delay have not been examined and more work can be carried out regarding phase and frequency offsets and variations.
7. More research into the effects of Rayleigh fading channels with different Doppler frequencies can be made. Different types of fading, e.g. Ricean fading, can also be examined as well as other channel environments that would affect classification.
8. Investigation into NN structures to give better performance should also be made. Factors such as: training algorithms, number of training epochs, minimum error, number of layers and number of neurons in each layer, input features, outputs, and hierarchical structures should all be examined further. New NN technologies can also be researched further to improve performance.
9. Research into finding more features that can be extracted from communication signals should be done. Features such as wavelets, for example, can be investigated further to improve classification performance.
10. Analog communications signals can also be added to the classifiers discussed in this thesis. Azzouz and Nandi have designed a classifier incorporating a limited number of analog and digital signals in [Azzound and Nandi, 1996] and this is a good starting point.

Finally, it is hoped that the methods developed in this thesis can be extended to other applications, such as the design of a universal receiver.



# Appendix A

This appendix presents the results of the DT and NN classifiers described in Chapter 5.

## A.1 Confusion Matrices for DT Classifier

Table A.1. DT classifier confusion matrix for signals at SNR = 20dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	100%	-	-	-	-	-
ASK4	-	100%	-	-	-	-
PSK2	-	-	100%	-	-	-
PSK4	-	-	-	100%	-	-
FSK2	-	-	-	-	97.5%	2.5%
FSK4	-	-	-	-	0.5%	99.5%

Table A.2. DT classifier confusion matrix for signals at SNR = 15dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	98.5%	1.5%	-	-	-	-
ASK4	-	100%	-	-	-	-
PSK2	-	-	100%	-	-	-
PSK4	-	-	-	100%	-	-
FSK2	-	-	-	-	95.5%	4.5%
FSK4	-	-	-	-	1%	99%

Table A.3. DT classifier confusion matrix for signals at SNR = 10dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	89.75%	10.25%	-	-	-	-
ASK4	1.5%	98.5%	-	-	-	-
PSK2	-	-	100%	-	-	-
PSK4	-	-	-	100%	-	-
FSK2	-	-	-	-	92.75%	7.25%
FSK4	-	-	-	-	0.75%	99.25%

Table A.4. DT classifier confusion matrix for signals at SNR = 5dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	71.25%	27%	-	1.75%	-	-
ASK4	12%	87.75%	-	0.5%	-	-
PSK2	-	-	98.75%	1.25%	-	-
PSK4	-	-	-	100%	-	-
FSK2	-	-	-	-	94.75%	5.25%
FSK4	-	-	-	-	13.25%	86.75%

Table A.5. DT classifier confusion matrix for signals at SNR = 0dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	26.5%	11%	0.75%	61.75%	-	-
ASK4	17%	40.5%	-	42.5%	-	-
PSK2	-	-	81.25%	18.75%	-	-
PSK4	-	-	25.5%	74.5%	-	-
FSK2	-	-	-	-	99.75%	0.25%
FSK4	-	-	-	-	96.25%	3.75%

Table A.6. DT classifier confusion matrix for signals at SNR = -5dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	0.25%	0.25%	38.5%	61%	-	-
ASK4	-	0%	31%	69%	-	-
PSK2	-	-	92.25%	7.75%	-	-
PSK4	-	-	92%	8%	-	-
FSK2	-	-	-	-	100%	-
FSK4	-	-	-	-	100%	-

## A.2 Confusion Matrices for NN Classifier

Table A.7. NN classifier confusion matrix for signals at SNR = 20dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	100%	-	-	-	-	-
ASK4	-	100%	-	-	-	-
PSK2	-	-	100%	-	-	-
PSK4	-	-	-	100%	-	-
FSK2	-	-	-	-	100%	-
FSK4	-	-	-	-	1%	99%

Table A.8. NN classifier confusion matrix for signals at SNR = 15dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	100%	-	-	-	-	-
ASK4	0.5%	99.5%	-	-	-	-
PSK2	-	-	100%	-	-	-
PSK4	-	-	-	100%	-	-
FSK2	-	-	-	-	99%	1%
FSK4	-	-	-	-	0.5%	99.5%

Table A.9. NN classifier confusion matrix for signals at SNR = 10dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	96.5%	3.5%	-	-	-	-
ASK4	5%	95%	-	-	-	-
PSK2	-	-	100%	-	-	-
PSK4	-	-	-	100%	-	-
FSK2	-	-	-	-	95%	5%
FSK4	-	-	-	-	1%	99%



Table A.10. NN classifier confusion matrix for signals at SNR = 5dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	82.29%	17.46%	0.25%	-	-	-
ASK4	21.45%	78.3%	0.25%	-	-	-
PSK2	-	-	100%	-	-	-
PSK4	2%	-	-	98%	-	-
FSK2	-	-	-	-	91%	9%
FSK4	-	-	-	-	4.5%	95.5%

Table A.11. NN classifier confusion matrix for signals at SNR = 0dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	59.44%	38.81%	1%	0.75%	-	-
ASK4	41.26%	56.98%	1%	0.75%	-	-
PSK2	-	-	71%	19.5%	-	-
PSK4	1.5%	-	7.5%	91%	-	-
FSK2	-	-	-	-	85.5%	14.5%
FSK4	-	-	-	-	45%	55%

Table A.12. NN classifier confusion matrix for signals at SNR = -5dB (test set).

Simulated Modulation Type	Deduced Modulation Type					
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
ASK2	51.15%	45.36%	2.75%	0.75%	-	-
ASK4	54.04%	42.46%	2.75%	0.75%	-	-
PSK2	-	-	53%	38.5%	-	8.5%
PSK4	-	-	27%	66.5%	0.5%	6%
FSK2	-	-	-	-	91.5%	8.5%
FSK4	-	-	-	-	88.5%	11.5%

# Appendix B

This Appendix presents the results of the DT and NN classifiers described in Chapter 6.

## B.1 Confusion Matrices for DT Classifier

### B.1.1 Classification of CPM Signals

Table B.1. DT classifier confusion matrix for signals at SNR = 20dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	100%	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-
PSK4	-	-	-	100%	-	-	-
FSK2	-	-	-	-	97.5%	2.5%	-
FSK4	-	-	-	-	0.5%	99.5%	-
CPM	-	-	-	-	-	-	100%

Table B.2. DT classifier confusion matrix for signals at SNR = 15dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	98.5%	1.5%	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-
PSK4	-	-	-	100%	-	-	-
FSK2	-	-	-	-	95.5%	4.5%	-
FSK4	-	-	-	-	1%	99%	-
CPM	-	-	-	-	-	-	100%

Table B.3. DT classifier confusion matrix for signals at SNR = 10dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	89.75%	10.25%	-	-	-	-	-
ASK4	1.5%	98.5%	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-
PSK4	-	-	-	100%	-	-	-
FSK2	-	-	-	-	92.75%	7.25%	-
FSK4	-	-	-	-	0.75%	99.25%	-
CPM	-	-	-	-	-	-	100%

Table B.4. DT classifier confusion matrix for signals at SNR = 5dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	71.25%	27%	-	1.75%	-	-	-
ASK4	12%	87.75%	-	0.5%	-	-	-
PSK2	-	-	97.5%	2.5%	-	-	-
PSK4	-	-	3%	97%	-	-	-
FSK2	-	-	-	-	94.75%	5.25%	-
FSK4	-	-	-	-	13.25%	85.25%	1.25%
CPM	-	-	-	-	-	-	100%

Table B.5. DT classifier confusion matrix for signals at SNR = 0dB (test set)

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	26.5%	11%	0.75%	61.75%	-	-	-
ASK4	17%	40.5%	-	42.5%	-	-	-
PSK2	-	-	83%	17%	-	-	-
PSK4	-	-	30%	70%	-	-	-
FSK2	-	-	-	-	99.75%	0.25%	-
FSK4	-	-	-	-	96.25%	2.75%	1%
CPM	-	-	-	-	-	4.25%	95.75%

Table B.6. DT classifier confusion matrix for signals at SNR = -5dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	0.25%	0.25%	38.5%	61%	-	-	40.5%
ASK4	-	0%	31%	69%	-	-	31%
PSK2	-	-	92.25%	7.75%	-	-	-
PSK4	-	-	92%	8%	-	-	-
FSK2	-	-	-	-	100%	-	-
FSK4	-	-	-	-	100%	-	-
CPM	-	-	-	-	0.25%	26.75%	73%

### B.1.2 Classification of Signals within the CPM Signal Class

Table B.7. DT classifier confusion matrix for signals at SNR = 20dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	91.5%	8.5%	-
Partial Response CPM (L=2)	31.33%	68.67%	-
GMSK	-	-	100%

Table B.8. DT classifier confusion matrix for signals at SNR = 15dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	90.83%	8.5%	-
Partial Response CPM (L=2)	36%	64%	-
GMSK	-	-	100%

Table B.9. DT classifier confusion matrix for signals at SNR = 10dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	78.17%	21.83%	-
Partial Response CPM (L=2)	30.17%	69.83%	-
GMSK	-	-	100%

Table B.10. DT classifier confusion matrix for signals at SNR = 5dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	100%	-	-
Partial Response CPM (L=2)	100%	-	-
GMSK	4%	-	96%

Table B.11. DT classifier confusion matrix for signals at SNR = 0dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	100%	-	-
Partial Response CPM (L=2)	100%	-	-
GMSK	21.5%	-	78.5%

Table B.12. DT classifier confusion matrix for signals at SNR = -5dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	100%	-	-
Partial Response CPM (L=2)	100%	-	-
GMSK	49.75%	-	50.25%

## B.2 Confusion Matrices for NN Classifier

### B.2.1 NN Classification of CPM Signals

Table B.13. NN classifier confusion matrix for signals at SNR = 20dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	100%	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-
PSK4	-	-	-	100%	-	-	-
FSK2	-	-	-	-	100%	-	-
FSK4	-	-	-	-	2.5	96.5%	1%
CPM	-	-	-	-	-	-	100%

Table B.14. NN classifier confusion matrix for signals at SNR = 15dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	99.75%	-	0.25%	-	-	-	-
ASK4	0.5%	99.25%	0.25%	-	-	-	-
PSK2	-	-	100%	-	-	-	-
PSK4	-	-	-	100%	-	-	-
FSK2	-	-	-	-	100%	-	-
FSK4	-	-	-	-	2%	96%	2%
CPM	-	-	-	-	-	-	100%

Table B.15. NN classifier confusion matrix for signals at SNR = 10dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	96.5%	3.5%	-	-	-	-	-
ASK4	5%	95%	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-
PSK4	-	-	-	100%	-	-	-
FSK2	-	-	-	-	94.5%	5.5%	-
FSK4	-	-	-	-	1%	99%	-
CPM	-	-	-	-	-	-	100%

Table B.16. NN classifier confusion matrix for signals at SNR = 5dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	82.29%	17.46%	0.25%	-	-	-	-
ASK4	21.45%	78.3%	0.25%	-	-	-	-
PSK2	-	0.5%	99.5%	-	-	-	-
PSK4	-	0.5%	-	99.5%	-	-	-
FSK2	-	-	-	-	88.5%	11.5%	-
FSK4	-	-	-	-	3.5%	96.5%	-
CPM	-	-	-	-	-	-	100%

Table B.17. NN classifier confusion matrix for signals at SNR = 0dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	57.63%	37.62%	1.25%	3.5%	-	-	-
ASK4	40%	55.24%	1.25%	3.5%	-	-	-
PSK2	-	9%	75.5%	15.5%	-	-	-
PSK4	-	6%	9.5%	84.5%	-	-	-
FSK2	-	-	-	-	84.5%	15.5%	-
FSK4	-	-	-	-	54%	46%	-
CPM	-	-	-	-	-	-	100%

Table B.18. NN classifier confusion matrix for signals at SNR = -5dB (test set).

Simulated Modulation Type	Deduced Modulation Type						
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM
ASK2	51.15%	45.36%	2.75%	0.75%	-	-	-
ASK4	54.04%	42.46%	2.75%	0.75%	-	-	-
PSK2	-	-	57%	43%	-	-	-
PSK4	-	-	32.5%	67.5%	-	-	-
FSK2	-	-	-	-	67.5%	32.5%	-
FSK4	-	-	-	-	69%	31%	-
CPM	-	-	-	-	-	9.5%	90.5%

### B.2.2 NN Classification of Signals Within the CPM Class

Table B.19. Neural network 1 confusion matrix for signals at SNR = 20dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	88.83%	11.17%	-
Partial Response CPM (L=2)	3.83%	96.17%	-
GMSK	-	-	100%

Table B.20. Neural network 1 confusion matrix for signals at SNR = 15dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	91.5%	8.5%	-
Partial Response CPM (L=2)	20.83%	79%	0.17
GMSK	-	-	100%



Table B.21. Neural network 1 confusion matrix for signals at SNR = 10dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	79%	20.5%	0.5%
Partial Response CPM (L=2)	25.17%	70%	4.83%
GMSK	0.5%	13%	86.5%

Table B.22. Neural network 1 confusion matrix for signals at SNR = 5dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	99.67%	0.33%	-
Partial Response CPM (L=2)	97.85%	2.17%	-
GMSK	97.5%	2.5%	0%

Table B.23. Neural network 1 confusion matrix for signals at SNR = 0dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	85.5%	13.8%	0.67%
Partial Response CPM (L=2)	87.5%	12.5%	-
GMSK	86.5%	13%	0.5%

Table B.24. Neural network 1 confusion matrix for signals at SNR = -5dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	41.17%	45.33%	13.5%
Partial Response CPM (L=2)	46%	40.17%	13.83%
GMSK	53.5%	35%	11.5%

Table B.25. Neural network 2 confusion matrix for signals at SNR = 20dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	95.5%	4.5%	-
Partial Response CPM (L=2)	2%	98%	-
GMSK	-	-	100%

Table B.26. Neural network 2 confusion matrix for signals at SNR = 15dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	87%	13%	-
Partial Response CPM (L=2)	5.5%	94.5%	-
GMSK	-	-	100%

Table B.27. Neural network 2 confusion matrix for signals at SNR = 10dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	75%	24.67%	0.33%
Partial Response CPM (L=2)	20.5%	75.83%	3.67%
GMSK	1%	8.5%	90.5%

Table B.28. Neural network 2 confusion matrix for signals at SNR = 5dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	99.67%	0.33%	-
Partial Response CPM (L=2)	0.5%	33.5%	66%
GMSK	-	15%	85%

Table B.29. Neural network 2 confusion matrix for signals at SNR = 0dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	100%	-	-
Partial Response CPM (L=2)	-	1.07%	98.33%
GMSK	-	0.5%	99.5%

Table B.30. Neural network 2 confusion matrix for signals at SNR = -5dB.

Simulated Modulation Type	Deduced Modulation Type		
	Full Response CPM (L=1)	Partial Response CPM (L=2)	GMSK
Full Response CPM (L=1)	100%	-	-
Partial Response CPM (L=2)	100%	-	-
GMSK	-	-	100%



# Appendix C

This Appendix presents the results of the DT and NN classifiers described in Chapter 7.

## C.1 Confusion Matrices for DT Classifier

Table C.1. DT classifier confusion matrix for signals at SNR = 20dB (test set).

Simulated Modulation Type	Deduced Modulation Type										
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA
ASK2	100%	-	-	-	-	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-	-	-	-	-
PSK4	-	-	-	100%	-	-	-	-	-	-	-
FSK2	-	-	-	-	97.5%	2.5%	-	-	-	-	-
FSK4	-	-	-	-	0.5%	99.5%	-	-	-	-	-
CPM	-	-	-	-	-	-	100%	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	100%	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	100%	-	-
FH-SS	-	-	-	-	-	-	-	-	-	100%	-
TDMA	-	-	-	-	-	0.25%	-	-	-	-	99.75%

Table C.2. DT classifier confusion matrix for signals at SNR = 15dB (test set).

Simulated Modulation Type	Deduced Modulation Type										
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA
ASK2	98.5%	1.5%	-	-	-	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-	-	-	-	-
PSK4	-	-	-	100%	-	-	-	-	-	-	-
FSK2	-	-	-	-	95.5%	4.5%	-	-	-	-	-
FSK4	-	-	-	-	1%	99%	-	-	-	-	-
CPM	-	-	-	-	-	-	100%	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	100%	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	100%	-	-
FH-SS	-	-	-	-	-	-	-	-	-	100%	-
TDMA	-	-	-	-	-	-	-	-	-	-	100%

Table C.3. DT classifier confusion matrix for signals at SNR = 10dB (test set).

Simulated Modulation Type	Deduced Modulation Type										
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA
ASK2	89.75%	10.25%	-	-	-	-	-	-	-	-	-
ASK4	1.5%	98.5%	-	-	-	-	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-	-	-	-	-
PSK4	-	-	-	100%	-	-	-	-	-	-	-
FSK2	-	-	-	-	92.75%	7.25%	-	-	-	-	-
FSK4	-	-	-	-	0.75%	99.25%	-	-	-	-	-
CPM	-	-	-	-	-	-	100%	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	100%	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	100%	-	-
FH-SS	-	-	-	-	-	-	-	-	-	100%	-
TDMA	-	-	-	-	-	-	-	-	-	-	100%



Table C.4. DT classifier confusion matrix for signals at SNR = 5dB (test set).

[illegible]

Table C.5. DT classifier confusion matrix for signals at SNR = 0dB (test set).

Simulated Modulation Type	Deduced Modulation Type										
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA
ASK2	26.5%	11%	0.75%	61.75%	-	-	-	-	-	-	-
ASK4	17%	40.5%	-	42.5%	-	-	-	-	-	-	-
PSK2	-	-	81.25%	18%	-	-	-	-	0.75%	-	-
PSK4	-	-	25%	74.5%	-	-	-	0.25%	0.25%	-	-
FSK2	-	-	-	-	98%	0.25%	-	-	-	-	1.75%
FSK4	-	-	-	-	92.5%	2.75%	1%	-	-	-	3.75%
CPM	-	-	-	-	-	4.25%	95.75%	-	-	-	-
BPSK-SS	-	-	-	0.75%	-	-	-	99.25%	-	-	-
QPSK-SS	-	-	-	14%	-	-	-	-	86%	-	-
FH-SS	-	-	-	-	-	-	-	-	-	100%	-
TDMA	-	-	-	-	-	-	-	-	-	-	100%

Table C.6. DT classifier confusion matrix for signals at SNR = -5dB (test set).

[illegible]

## C.2 Confusion Matrices for NN Classifier

Table C.7. NN classifier confusion matrix for signals at SNR = 20dB (test set).

Simulated Modulation Type	Deduced Modulation Type										
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA
ASK2	99.5%	0.5%	-	-	-	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-	-	-	-	-
PSK4	-	-	-	100%	-	-	-	-	-	-	-
FSK2	-	-	-	-	100%	-	-	-	-	-	-
FSK4	-	-	-	-	2.5%	97.5%	-	-	-	-	-
CPM	-	-	-	-	-	-	100%	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	100%	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	100%	-	-
FH-SS	-	-	-	-	-	-	-	-	-	100%	-
TDMA	-	-	-	-	-	-	-	-	-	-	100%

Table C.8. NN classifier confusion matrix for signals at SNR = 15dB (test set).

Simulated Modulation Type	Deduced Modulation Type										
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA
ASK2	99%	-	1%	-	-	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-	-	-	-	-
PSK4	-	-	-	100%	-	-	-	-	-	-	-
FSK2	-	-	-	-	100%	-	-	-	-	-	-
FSK4	-	-	-	-	0.5%	99.5%	-	-	-	-	-
CPM	-	-	-	-	-	-	100%	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	100%	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	100%	-	-
FH-SS	-	-	-	-	-	-	-	-	-	100%	-
TDMA	-	-	-	-	-	-	-	-	-	-	100%

Table C.9. NN classifier confusion matrix for signals at SNR = 10dB (test set).

Simulated Modulation Type	Deduced Modulation Type										
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA
ASK2	98.5%	-	1.5%	-	-	-	-	-	-	-	-
ASK4	0.5%	99.5%	-	-	-	-	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-	-	-	-	-
PSK4	-	-	-	100%	-	-	-	-	-	-	-
FSK2	-	-	-	-	99.5%	0.5%	-	-	-	-	-
FSK4	-	-	-	-	1%	98.5%	0.5%	-	-	-	-
CPM	-	-	-	-	-	-	100%	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	100%	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	100%	-	-
FH-SS	-	-	-	-	-	-	-	-	-	99.5%	0.5%
TDMA	-	-	-	-	-	-	-	-	-	-	100%

Table C.10. NN classifier confusion matrix for signals at SNR = 5dB (test set).

Simulated Modulation Type	Deduced Modulation Type										
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA
ASK2	94%	-	1%	4.5%	-	-	-	0.5%	-	-	-
ASK4	0.5%	99.5%	-	-	-	-	-	-	-	-	-
PSK2	-	-	87%	8%	-	-	-	0.5%	4.5%	-	-
PSK4	-	-	-	100%	-	-	-	-	-	-	-
FSK2	-	-	-	-	99%	0.5%	-	-	-	-	0.5%
FSK4	-	-	-	-	2%	92.5%	4%	-	-	-	1.5%
CPM	-	-	-	-	-	-	100%	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	100%	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	100%	-	-
FH-SS	-	-	-	-	-	-	-	-	-	100%	-
TDMA	-	-	-	-	-	-	-	-	-	-	100%

Table C.11. NN classifier confusion matrix for signals at SNR = 0dB (test set).

Simulated Modulation Type	Deduced Modulation Type										
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA
ASK2	38%	-	-	55%	-	-	-	7%	-	-	-
ASK4	0.5%	99.5%	-	-	-	-	-	-	-	-	-
PSK2	-	-	2%	94.5%	-	-	-	-	3.5%	-	-
PSK4	-	-	-	99.5%	-	-	-	-	0.5%	-	-
FSK2	-	-	-	-	98.5%	-	-	-	-	-	1.5%
FSK4	-	-	-	-	1%	75%	18.5%	-	-	-	5.5%
CPM	-	-	-	-	-	2%	97.5%	-	-	0.5%	-
BPSK-SS	-	-	-	-	-	-	-	100%	-	-	-
QPSK-SS	-	-	-	-	-	-	-	42.5%	57.5%	-	-
FH-SS	-	-	-	-	-	-	-	-	-	100%	-
TDMA	-	-	-	-	-	-	-	-	-	-	100%



Table C.12. NN classifier confusion matrix for signals at SNR = -5dB (test set).

Simulated Modulation Type	Deduced Modulation Type										
	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA
ASK2	0%	-	1%	86.5%	-	-	-	3%	95%	-	-
ASK4	2.5%	97.5%	-	-	-	-	-	-	-	-	-
PSK2	-	-	55%	1.5%	-	-	-	-	43.5%	-	-
PSK4	-	0.5%	1.5%	55%	-	-	-	-	43%	-	-
FSK2	-	-	-	-	89%	-	-	-	-	-	11%
FSK4	-	-	-	-	2%	68%	9.5%	-	-	1.5%	19%
CPM	-	-	-	-	-	20%	65%	-	-	11.5%	3.5%
BPSK-SS	-	-	-	-	-	-	-	-	100%	-	-
QPSK-SS	-	-	-	-	-	-	-	-	100%	-	-
FH-SS	-	-	-	-	-	-	-	-	-	100%	-
TDMA	-	-	-	-	-	-	-	-	-	-	100%

# Appendix D

This Appendix presents the results of the DT and NN classifiers described in Chapter 8.

## D.1 Confusion Matrices for DT Classifier

Table D.1. DT classifier confusion matrix for signals at SNR = 20dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	100%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-	-	-	-	-	-	-	-	-
PSK4	-	-	-	90.75%	9.25%	-	-	-	-	-	-	-	-	-	-
PSK8	-	-	-	1.25%	98.75%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	99.5%	0.5%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	9%	91%	-	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	-	25.25%	75.75%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	-	-	99.75%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table D.2. DT classifier confusion matrix for signals at SNR = 15dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	98.5%	15%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	0.25%	-	99.75%	-	-	-	-	-	-	-	-	-	-	-	-
PSK4	0.25%	-	-	90%	9.75%	-	-	-	-	-	-	-	-	-	-
PSK8	-	-	-	2.25%	97.75%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	97.75%	2.25%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	7.25%	92.75%	-	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	-	23.5%	76.5%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	-	-	100%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table D.3. DT classifier confusion matrix for signals at SNR = 10dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	89.75%	10.25%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	1.5%	98.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-	-	-	-	-	-	-	-	-
PSK4	0.25%	-	-	89.75%	10%	-	-	-	-	-	-	-	-	-	-
PSK8	-	-	-	3.25%	96.75%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	98.5%	1.5%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	5.25%	94.75%	-	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	-	17%	82.75%	-	-	-	-	0.25%	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	-	-	100%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table D.4. DT classifier confusion matrix for signals at SNR = 5dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	73%	27%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	12%	88%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	-	0.25%	98.5%	-	1.25%	-	-	-	-	-	-	-	-	-	-
PSK4	-	-	-	84.5%	15.5%	-	-	-	-	-	-	-	-	-	-
PSK8	-	-	-	7.25%	92.75%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	95.5%	2.5%	0.75%	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	4%	79.75%	16.25%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	0.25%	2.5%	96%	-	-	-	-	1.25%	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-
QPSK-SS	-	-	-	1%	-	-	-	-	-	-	99%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table D.5. DT classifier confusion matrix for signals at SNR = 0dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	64.5%	35.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	30.75%	69.25%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	-	-	81.25%	1.25%	16.75%	-	-	-	-	-	0.75%	-	-	-	-
PSK4	-	-	25%	54%	20.5%	-	-	-	-	-	0.5%	-	-	-	-
PSK8	-	-	9.25%	55.25%	35.5%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	90%	-	-	-	-	-	-	10%	-	-
FSK4	-	-	-	-	-	71.5%	0%	22.5%	-	-	-	-	6%	-	-
FSK8	-	-	-	-	-	38.75%	-	52.5%	-	-	-	-	8.75%	-	-
CPM	-	-	-	-	-	-	-	6%	93.75%	-	-	-	0.25%	-	-
BPSK-SS	-	-	-	0.75%	-	-	-	-	-	99.25%	-	-	-	-	-
QPSK-SS	-	-	-	15.25%	-	-	-	-	-	-	84.75%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table D.6. DT classifier confusion matrix for signals at SNR = -5dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	16.5%	13.25%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	22.25%	32%	10.5%	15.25%	19.5%	-	-	-	-	0.25%	-	-	-	-	0.25%
PSK2	-	-	84%	1%	7%	-	-	-	-	-	8%	-	-	-	-
PSK4	-	-	85%	6.5%	1.75%	-	-	-	-	-	6.75%	-	-	-	-
PSK8	-	-	83%	8.5%	1.25%	-	-	-	-	7.25%	-	-	-	-	-
FSK2	-	-	-	-	-	72%	-	-	-	-	-	-	28%	-	-
FSK4	-	-	-	-	-	78.5%	0%	-	-	-	-	-	28%	-	-
FSK8	-	-	-	-	-	74.75%	-	0%	-	-	-	-	25.25%	-	-
CPM	-	-	-	-	-	-	-	21%	68.75%	-	-	-	10.25%	-	-
BPSK-SS	-	-	-	16.75%	1.5%	-	-	-	-	81.75%	-	-	-	-	-
QPSK-SS	-	-	-	17.75%	6.25%	-	-	-	-	-	76%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	38.75%	9.25%	2.75%	-	-	-	-	2.25%	-	-	-	47%	-
QAM16	-	-	55.25%	15%	3.75%	-	-	-	-	-	1.25%	-	-	-	24.75%

## D.2 Confusion Matrices for NN Classifier

Table D.7. NN classifier confusion matrix for signals at SNR = 20dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	100%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	0.7%	-	99.3%	-	-	-	-	-	-	-	-	-	-	-	-
PSK4	0.7%	-	-	94.83%	4.97%	-	-	-	-	-	-	-	-	-	-
PSK8	0.7%	-	-	3.48%	95.82%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	99%	1%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	6.5%	93%	0.5%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	-	13%	87%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	0.7%	-	-	-	-	-	-	-	-	99.3%	-	-	-	-	-
QPSK-SS	0.7%	-	-	-	-	-	-	-	-	-	99.3%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%



Table D.8. NN classifier confusion matrix for signals at SNR = 15dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	100%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	0.5%	99.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	0.7%	-	98.31%	0.5%	0.5%	-	-	-	-	-	-	-	-	-	-
PSK4	0.7%	-	-	93.34%	5.96%	-	-	-	-	-	-	-	-	-	-
PSK8	0.7%	-	-	3.97%	95.33%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	98.5%	1.5%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	3.5%	95%	1.5%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	-	9.5%	90.5%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	0.7%	-	-	-	-	-	-	-	-	99.3%	-	-	-	-	-
QPSK-SS	0.7%	-	-	-	-	-	-	-	-	-	99.3%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table D.9 NN classifier confusion matrix for signals at SNR = 10dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	96.5%	3.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	5%	95%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	0.4%	-	99.1%	-	0.5%	-	-	-	-	-	-	-	-	-	-
PSK4	0.4%	-	-	93.62%	5.98%	-	-	-	-	-	-	-	-	-	-
PSK8	0.4%	-	-	9.96%	89.64%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	96%	4%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	0.5%	94.5%	5%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	-	4.5%	95.5%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	0.4%	-	-	-	-	-	-	-	-	99.6%	-	-	-	-	-
QPSK-SS	0.4%	-	-	-	-	-	-	-	-	-	99.6%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table D.10. NN classifier confusion matrix for signals at SNR = 5dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	82.5%	17.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	21.5%	78.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	0.2%	-	98.3%	-	1.5%	-	-	-	-	-	-	-	-	-	-
PSK4	0.2%	-	-	86.83%	13%	-	-	-	-	-	-	-	-	-	-
PSK8	0.2%	-	-	15%	84.83%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	89%	11%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	-	88%	12%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	-	2.5%	97.5%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	0.2%	-	-	-	-	-	-	-	-	99.8%	-	-	-	-	-
QPSK-SS	0.2%	-	-	-	-	-	-	-	-	-	99.8%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table D.11. NN classifier confusion matrix for signals at SNR = 0dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	60.35%	39.4%	-	-	-	-	-	-	-	-	-	-	-	-	0.25%
ASK4	41.89%	57.86%	-	-	-	-	-	-	-	-	-	-	-	-	0.25%
PSK2	0.3%	-	95.71%	1.5%	1%	-	-	-	-	-	1.5%	-	-	-	-
PSK4	0.3%	-	2.5%	68.29%	28.5%	-	-	-	-	-	0.5%	-	-	-	-
PSK8	0.3%	-	2.5%	59.32%	37.39%	-	-	-	-	-	0.5%	-	-	-	-
FSK2	-	-	-	-	-	79.5%	20.5%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	9.5%	73%	17.5%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	0.5%	12.5%	87%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	0.3%	-	-	-	-	-	-	-	-	99.7%	-	-	-	-	-
QPSK-SS	0.3%	-	-	-	-	-	-	-	-	-	99.7%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table D.12. NN classifier confusion matrix for signals at SNR = -5dB (test set).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	50.8%	45.1%	3.6%	-	-	-	-	-	-	-	-	-	-	-	0.5%
ASK4	53.7%	42.2%	3.6%	-	-	-	-	-	-	-	-	-	-	-	0.5%
PSK2	0.1%	-	72.4%	13.5%	10%	-	-	-	-	-	4%	-	-	-	-
PSK4	0.1%	-	4.5%	65.9%	26.97%	-	-	-	-	-	2.5%	-	-	-	-
PSK8	0.1%	-	5%	59.94%	32%	-	-	-	-	-	3%	-	-	-	-
FSK2	-	-	-	-	-	79.5%	20.5%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	9.5%	73%	17.5%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	0.5%	12.5%	87%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	0.1%	-	-	-	-	-	-	-	-	99.9%	-	-	-	-	-
QPSK-SS	0.1%	-	2.5%	5.5%	-	-	-	-	-	-	91.9%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

# Appendix E

This Appendix presents the results of the DT and NN classifiers described in Chapter 9.

## E.1 Confusion Matrices for DT Classifier (Rayleigh Fading)

Table E.1 DT classifier confusion matrix for signals at SNR = 20dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	100%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	0.25%	-	99.75%	-	-	-	-	-	-	-	-	-	-	-	-
PSK4	-	-	-	99.75%	0.25%	-	-	-	-	-	-	-	-	-	-
PSK8	-	-	-	6.75%	93.25%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	100%	-	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	16.25%	79.5%	4.25%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	6.5%	78.5%	15%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	7.75%	-	-	92.25%	-	-	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	-	-	100%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table E.2. DT classifier confusion matrix for signals at SNR = 15dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	99%	1%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	0.25%	-	99.75%	-	-	-	-	-	-	-	-	-	-	-	-
PSK4	0.5%	0.25%	-	99.25%	-	-	-	-	-	-	-	-	-	-	-
PSK8	0.25%	-	-	9.75%	89%	-	-	-	-	-	-	-	-	1%	-
FSK2	-	-	-	-	-	100%	-	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	16.25%	74.5%	9%	0.25%	-	-	-	-	-	-
FSK8	-	-	-	-	-	5%	69.5%	25.5%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	5%	-	-	95%	-	-	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	-	-	100%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table E.3. DT classifier confusion matrix for signals at SNR = 10dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	89.25%	16.75%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	2%	98%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	0.5%	-	99.5%	-	-	-	-	-	-	-	-	-	-	-	-
PSK4	-	-	-	98%	2%	-	-	-	-	-	-	-	-	-	-
PSK8	-	-	-	19%	80.5%	-	-	-	-	-	-	-	-	-	0.5%
FSK2	-	-	-	-	-	100%	-	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	13.5%	67.75%	18.75%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	7.5%	46.75%	45.25%	0.5%	-	-	-	-	-	-
CPM	-	-	-	-	-	5.5%	-	-	94.5%	-	-	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-
QPSK-SS	-	-	-	0.25%	-	-	-	-	-	-	99.75%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%



Table E.4. DT classifier confusion matrix for signals at SNR = 5dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	71.75%	28.25%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	12.75%	87.25%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	-	-	100%	-	-	-	-	-	-	-	-	-	-	-	-
PSK4	-	-	-	65.25%	34.75%	-	-	-	-	-	-	-	-	-	-
PSK8	-	-	-	20.5%	79.5%	-	-	-	-	-	-	-	-	-	0.5%
FSK2	-	-	-	-	-	46%	50.25%	3.25%	-	-	-	-	0.5%	-	-
FSK4	-	-	-	-	-	8.25%	52%	38.5%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	5.25%	32.25%	60.25%	2.25%	-	-	-	-	-	-
CPM	-	-	-	-	-	0.5%	-	-	99.5%	-	-	-	-	-	-
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	-	-	100%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table E.5. DT classifier confusion matrix for signals at SNR = 0dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type															
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16	
ASK2	66.25%	33.75%	-	-	-	-	-	-	-	-	-	-	-	-	-	
ASK4	30.25%	69.75%	-	-	-	-	-	-	-	-	-	-	-	-	-	
PSK2	0.25%	-	99.75%	-	-	-	-	-	-	-	-	-	-	-	-	
PSK4	-	-	-	21.25%	78.75%	-	-	-	-	-	-	-	-	-	-	
PSK8	-	-	-	16.25%	83.5%	-	-	-	-	-	0.25%	-	-	-	-	
FSK2	-	-	-	-	-	7.5%	20.5%	70.5%	0.25%	-	-	-	1.25%	-	-	
FSK4	-	-	-	-	-	1.25%	1.25%	96%	-	-	-	-	1.25%	-	-	
FSK8	-	-	-	-	-	1.25%	2%	94%	-	-	-	-	2.75%	-	-	
CPM	-	-	-	-	-	-	-	5.5%	94.5%	-	-	-	-	-	-	
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	
QPSK-SS	-	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-	
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	
QAM8	0.25%	-	-	-	-	-	-	-	-	-	-	-	-	99.5%	-	
QAM16	0.5%	-	-	0.25%	0.25%	-	-	-	-	-	-	-	-	-	99%	

Table E.6. DT classifier confusion matrix for signals at SNR = -5dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	15%	6.75%	-	71.25%	-	-	-	-	-	-	-	-	-	0.75%	0.25%
ASK4	28.25%	23%	-	48.25%	-	-	-	-	-	-	0.5%	-	-	-	-
PSK2	-	-	0%	6%	32.25%	-	-	-	-	-	1.75%	-	-	-	-
PSK4	-	-	-	62.75%	36%	-	-	-	-	-	1.25%	-	-	-	-
PSK8	-	-	-	64.5%	35.25%	-	-	-	-	-	0.25%	-	-	-	-
FSK2	-	-	-	-	-	89.5%	-	0.25%	-	-	-	-	10.25%	-	-
FSK4	-	-	-	-	-	88%	-	-	-	-	-	-	12%	-	-
FSK8	-	-	-	-	-	86.75%	-	0%	-	-	-	-	13.25%	-	-
CPM	-	-	-	-	-	-	0.25%	22.25%	75%	-	-	-	2.5%	-	-
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	-	-	100%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	1.25%	-	-	24.25%	38.5%	-	-	-	-	0.25%	-	-	-	35.75%	-
QAM16	-	2.25%	-	36.5%	41%	-	-	-	-	-	-	-	-	-	20.25%

## E.4 Confusion Matrices for NN Classifier (Rayleigh Fading)

Table E.7. NN classifier confusion matrix for signals at SNR = 20dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	100%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	-	100%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	2.7%	-	90.98%	2.92%	3.41%	-	-	-	-	-	-	-	-	-	-
PSK4	2.7%	-	-	91.46%	5.84%	-	-	-	-	-	-	-	-	-	-
PSK8	2.7%	-	2.92%	1.46%	91.95%	-	-	-	-	0.97%	-	-	-	-	-
FSK2	-	-	-	-	-	100%	-	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	5%	70%	25%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	2%	30%	68%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	2.7%	-	-	-	-	-	-	-	-	97.3%	-	-	-	-	-
QPSK-SS	2.7%	-	-	-	-	-	-	-	-	-	97.3%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table E.8. NN classifier confusion matrix for signals at SNR = 15dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	100%	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	0.5%	99.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	1.9%	-	94.18%	2.45%	1.47%	-	-	-	-	-	-	-	-	-	-
PSK4	1.9%	-	-	96.14%	1.96%	-	-	-	-	-	-	-	-	-	-
PSK8	1.9%	-	-	2.45%	95.65%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	100%	-	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	5.5%	77%	17.5%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	1%	33%	66%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	1.9%	-	-	-	-	-	-	-	-	98.1%	-	-	-	-	-
QPSK-SS	1.9%	-	-	-	-	-	-	-	-	-	98.1%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table E.9. NN classifier confusion matrix for signals at SNR = 10dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	96.5%	3.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	5%	95%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	0.8%	-	96.72%	1.49%	0.99%	-	-	-	-	-	-	-	-	-	-
PSK4	0.8%	-	-	93.25%	5.95%	-	-	-	-	-	-	-	-	-	-
PSK8	0.8%	-	0.5%	5.46%	93.25%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	99.5%	0.5%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	6%	84%	10%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	3.5%	46%	50.5%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	100%	-	-	-	-	-	-
BPSK-SS	0.8%	-	-	-	-	-	-	-	-	99.2%	-	-	-	-	-
QPSK-SS	0.8%	-	-	-	-	-	-	-	-	-	99.2%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table E.10. NN classifier confusion matrix for signals at SNR = 5dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	82.5%	17.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	21.5%	78.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	-	-	98.5%	0.5%	1%	-	-	-	-	-	-	-	-	-	-
PSK4	-	-	0.5%	79%	20.5%	-	-	-	-	-	-	-	-	-	-
PSK8	-	-	1%	6%	93%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	98.5%	1.5%	-	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	7%	53%	40%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	6.5%	27%	66.5%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	99.5%	-	-	-	0.5%	-	-
BPSK-SS	-	-	-	-	-	-	-	-	-	100%	-	-	-	-	-
QPSK-SS	-	-	-	-	-	-	-	-	-	-	100%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	-	-	-	-	-	-	-	-	-	-	-	-	-	100%	-
QAM16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100%

Table E.11. NN classifier confusion matrix for signals at SNR = 0dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	60.5%	39.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
ASK4	42%	58%	-	-	-	-	-	-	-	-	-	-	-	-	-
PSK2	0.3%	-	93.72%	2%	3.99%	-	-	-	-	-	-	-	-	-	-
PSK4	0.3%	-	3.49%	58.32%	37.89%	-	-	-	-	-	-	-	-	-	-
PSK8	0.3%	-	3.49%	30.91%	65.3%	-	-	-	-	-	-	-	-	-	-
FSK2	-	-	-	-	-	95%	4%	1%	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	21.5%	54%	24.5%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	15%	45.5%	39.5%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	99%	-	-	-	1%	-	-
BPSK-SS	0.3%	-	-	-	-	-	-	-	-	99.7%	-	-	-	-	-
QPSK-SS	0.3%	-	-	-	-	-	-	-	-	-	99.7%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	2%	-	-	-	-	-	-	-	-	-	-	-	-	98%	-
QAM16	-	0.5%	-	-	-	-	-	-	-	-	-	-	-	-	99.5%



Table E.12. NN classifier confusion matrix for signals at SNR = -5dB (Doppler spread = 120 Hz).

Simulated Modulation Type	Deduced Modulation Type														
	ASK2	ASK4	PSK2	PSK4	PSK8	FSK2	FSK4	FSK8	CPM	BPSK-SS	QPSK-SS	FH-SS	TDMA	QAM8	QAM16
ASK2	46.2%	53.3%	-	-	-	-	-	-	-	-	-	-	-	0.5%	-
ASK4	61.1%	38.4%	-	-	-	-	-	-	-	-	-	-	-	0.5%	-
PSK2	0.1%	-	63.4%	14.49%	21.98%	-	-	-	-	-	-	-	-	-	-
PSK4	0.1%	-	7.99%	56.9%	34.47%	-	-	-	-	0.5%	-	-	-	-	-
PSK8	0.1%	-	11.99%	47.45%	39.5%	-	-	-	-	1%	-	-	-	-	-
FSK2	-	-	-	-	-	79.5%	19.5%	1%	-	-	-	-	-	-	-
FSK4	-	-	-	-	-	35.5%	39.5%	25%	-	-	-	-	-	-	-
FSK8	-	-	-	-	-	32%	31.5%	36.5%	-	-	-	-	-	-	-
CPM	-	-	-	-	-	-	-	-	97.5%	-	-	-	2.5%	-	-
BPSK-SS	0.1%	-	-	-	0.5%	-	-	-	-	99.4%	-	-	-	-	-
QPSK-SS	0.1%	-	1%	-	-	-	-	-	-	-	98.9%	-	-	-	-
FH-SS	-	-	-	-	-	-	-	-	-	-	-	100%	-	-	-
TDMA	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
QAM8	11%	-	-	-	-	-	-	-	-	-	-	-	-	89%	-
QAM16	-	11%	-	-	-	-	-	-	-	-	-	-	-	-	89%

# REFERENCES

1. Aisbett, J., "Automatic modulation recognition using time domain parameters", *Signal Processing*, vol. 13, pp 323-328, Oct 1987.
2. Akmouche, W., "Detection of Multicarrier Modulations Using 4<sup>th</sup>-Order Cumulants", *Proceedings of IEEE MILCOM '99*, pp. 432-436, 1999.
3. Anderson, J.B., Aulin, T., and Sundberg, C.E., *Digital Phase Modulation*. New York: Plenum Press. 1986.
4. Arulampalam, G., Ramakonar, V., Bouzerdoun, A., and Habibi, D., "Classification of Digital Modulation Schemes Using Neural Networks", *Proceedings of ISSPA '99*, pp. 649-652, August 1999.
5. Assaleh, K., Farrell, K. and Mammone, R.J., "A new method of modulation classification for digitally modulated signals." *MELCOM 92, Communication, Fusing, Command, Control, and Intelligence*, pp30.5.1-30.5.5, October 1992.
6. Azzouz, E.E and Nandi, A.K., "Algorithms for automatic modulation recognition of communication signals", *IEEE Trans. Commun*, vol. 46, pp431-436, April 1998.
7. Azzouz, E.E. and Nandi, A.K., "Automatic identification of digital modulations", *Signal Processing*, vol. 47, pp55-69, Nov 1995.
8. Azzouz, E.E. and Nandi, A.K., *Automatic Modulation Recognition of Communication Systems*. Netherlands: Kluwer Academic Publishers, 1996.
9. Beidas, B.F. and Weber, C.L., "Higher – order correlation based approach to modulation classification for digitally modulated signals", *IEEE Journal on Selected Areas in Communications*, vol. 13, No. 1, January 1995.
10. Boiteau, D. and Le Martret, C., "A General Maximum Likelihood Framework for Modulation Classification", *Proceedings of ICASSP '98*, vol. 4, pp. 2165-2168, May 1998.
11. Bond, F.E and Meyer, H.F., "Fading and Multipath Considerations in Aircraft/Satellite Communications Systems," *Communication Satellite Systems Technology*, R.B. Marsten, ed. (New York: Academic Press: 1966).
12. Bouzerdoun, A., "A Hierarchical Classifier for Multispectral Satellite Imagery", *IEICE Transactions on Electronics*, vol. E84-C, No. 12, pp 1952 – 1958, Dec 2001.
13. Callaghan, T.G., Pery, J.L. and Tjho, J.K., "Sampling and algorithms aid modulation recognition", *Microwaves RF*, vol. 24, pp 117-119, Sep 1985.
14. Chan, Y.T., Gadbois, L.G., "Identification of the Modulation Type of a Signal", *Signal Processing*, vol. 16, no. 2, pp. 149-154, February 1989.
15. Chugg, K.M, Long, C-S., and Polydoros, A., "Combined Likelihood Power Estimation and Multiple Hypothesis Modulation Classification", *Proceedings of ASILOMAR-29*, vol. 2, pp 1137-1141, November 1995.
16. Chung, C-D. and Polydoros, A., "Envelope-Based Classification Schemes for Continuous-Phase Binary Frequency Shift-Keyed Modulations", *Proceedings of IEEE MILCOM '94*, vol. 3, pp. 796-800, October 1994.
17. Cybenko, G., "Approximation by Superpositions of a Sigmoidal Function", *Math. Contr. Signals Syst.*, vol. 2, pp.303-314, 1989.

18. Demuth, H., and Beale, M., *Neural Network Toolbox for Use with Matlab: User's Guide*. The Mathworks inc, (3<sup>rd</sup> ed), 1998.
19. DeSimio, M.P. and Glenn, E.P., "Adaptive generation of decision functions for classification of digitally modulated signals". *NAECON*, pp1010-1014, 1988.
20. Dominguez, L.V., Borrallo, L.V. and Garcia, J.P., "A general approach to the automatic classification of radio communication systems", *Signal Processing*, vol. 22, pp 239–250, Mar 1991.
21. Drukkman, I.m Plotkin, E.I. and Swamy, M.N.S., "Automatic Modulation Type Recognition", *IEEE Canadian Conference on Electrical and Computer Engineering*, vol.. 1, pp 65-68, May 1998.
22. Freund, J.E. and Walpole, R.E., *Mathematical Statistics (4<sup>th</sup> Edition)*, Englewood Cliffs, N.J: Prentice Hall, 1987.
23. Fukunaga, K. *Introduction to Statistical Pattern Recognition*. Sandiego, CA: Academic Press Inc, 1990.
24. Geman, S., Bienenstock, E., and Doursat, T., "Neural Networks and the Bias/Variance Dilemma," *Neural Comput.*, vol5, pp. 1-58, 1992.
25. Ghani, N., Lamontagne, R., "Neural Networks Applied to the Classification of Spectral Features for Automatic Modulation Recognition", *Proceedings of IEEE MILCOM '93*, pp. 111-115, 1993.
26. Hagan, M., Demuth, H. and Beale, M., *Neural Network Design*. Boston MA: PWS Publishing Company, 1996.
27. Hair, F.J et al, *Multivariate Data Analysis (4<sup>th</sup> Edition)*. Englewood Cliffs New Jersey: Prentice Hall, 1995.
28. Haykin, S., *Neural Networks: A Comprehensive Foundation*. New Jersey: Prentice Hall, 2<sup>nd</sup> edition, 1999.
29. Hero, A.O., Hadinejad-Mahram, H., "Digital Modulation Classification Using Power Moment Matrices", *Proceedings of IEEE MILCOM '98*, pp. 3285-3288, 1998.
30. Ho, K.C., Prokopiw, W., and Chan, Y.T., "Modulation Identification by the Wavelet Transform", *Proceedings of IEEE MILCOM '95*, pp. 886-890, 1995.
31. Hong, L., Ho, K.C., "Identification of Digital Modulation Types Using the Wavelet Transform", *Proceedings of IEEE MILCOM '99*, 1999.
32. Huang, C.Y. and Polydoros, A., "Advanced Methods for Digital Quadrature and Offset Modulation Classification", *Proceedings of IEEE MILCOM '91*, vol. 2, pp. 841-845, November 1991.
33. Huang, C.Y., and Polydoros, A., "Two Small-SNR Classification Rules for CPM", *Proceedings of IEEE MILCOM '91*, vol. 2, pp. 841-845, November 1991.
34. Huo, X. and Donoho, D., "A Simple and Robust Modulation Classification Method via Counting", *Proceedings of ICASSP '98*, vol. 6, pp 3289-3292, May 1998.
35. Jakes, W.C., "An Approximate Method to Estimate an Upper Bound on the Effects of Multipath Delay Distortion on Digital Transmission," *IEEE Int. Conf. Commun.*, pp 47.11-47.1.5, June 1978.
36. Jondral, F., "Automatic classification of high frequency signals", *Signal Processing*, vol. 6, pp 311-323, Aug 1984.
37. Ketterer, H., Jondral, F., Costa, A.H., " Classification of Modulation Modes Using Time-Frequency Methods", *Proceedings of IEEE ICASSP '99*, March 1999.

38. Lallo, P., "Signal Classification by Discrete Fourier Transform", *Proceedings of IEEE MILCOM '99*, 1999.
39. Lay, N.E. and Polydoros, A., "Modulation Classification of Signals in Unknown ISI Environments", *Proceedings of MILCOM '95*, vol. 1, pp. 170-174, November 1995.
40. Le Martret, C. and Boiteau, D.M., "Modulation Classification by Mean of Different Orders Statistical Moments", *Proceedings of IEEE MILCOM '97*, vol. 3, pp. 1387-1391, November 1997.
41. Lee, W.C. and Messerschmidt, D.G., *Digital Communication*. Boston: Kluwer Academic Publishers, 1988.
42. Liedtke, F.F., "Computer Simulation of an automatic classification procedure for digitally modulated communication signals with unknown parameters". *Signal Processing*, vol. 6, pp. 311-323, August 1984.
43. Long, C.S., Chugg, K.M., Polydoros, A., "Further Results in Likelihood Classification of QAM Signals", *Proceedings of IEEE MILCOM '94*, vol. 1, pp. 57-61, October 1994.
44. Louis, C., and Sehier, P., "Automatic Modulation Recognition With a Hierarchical Neural Network.", *Proceedings of IEEE MILCOM '94*, pp. 713-717, October 1994.
45. Marchand, P., Le Martret, C., and Lacoume, J-L., "Classification of Linear Modulations by a Combination of Different Orders Cyclic Cumulants", *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics*, pp. 47-51, July 1997.
46. Marchand, P., Le Martret, C., and Lacoume, J-L., "Modulation Classification Based on a Maximum-Likelihood Receiver in the Cyclic-HOS Domain", *Proceedings of EUSIPCO '98*, vol. 4, pp. 2409-2412, September 1998.
47. Martin, A., "A signal analysis and classification strategy for implementation in an EW communication receiver", *Fifth International conference on Radio Receiver and Associated Systems*, pp. 222-226, Jul 1990.
48. Mitchell, T., *Machine Learning*, New York: McGraw-Hill, 1997.
49. Mobasser, B.G., "Constellation Shape as a Robust Signature for Digital Modulation Recognition", *Proceedings of IEEE MILCOM '99*, 1999.
50. Nagy, P.A.J., "A modulation classifier for multi-channel systems and multi-transmitter situations", *MILCOM '94 Conference*, 1994.
51. Oon, T. B. and Steele, R., 'Maximum Likelihood Channel Estimation of Flat Rayleigh Multiuser CDMA Channels', *IEE Colloquium on CDMA Techniques and Applications for 'Third Generation Mobile Systems'*, pp. 5/1 to 5/7, May 1997.
52. Papoulis, A., *Probability Random Variables, and Stochastic Processes*. New York: McGraw Hill Book Company, 1965.
53. Peterson, R.L., Ziemer, R.E., and Borth, D.E., *Introduction to Spread Spectrum Communications*. New Jersey: Prentice-Hall, 1995.
54. Petrovic, P.M, Krsmanovic, Z. B. and Remenski, N.K., "An automatic VHF signal classifier", *Mediterranean Electro technical Conference, MELECON*, pp. 385-387, 1989.
55. Polydoros, A. and Kim K., "On the detection and classification of quadrature digital modulations in broad-band noise". *IEEE Transactions on Communications*, vol. 40, pp. 1199-1211, August 1990.
56. Proakis, J.G., *Digital Communications*. New York: McGraw-Hill, 1995.

57. Ramakonar, V., SIMULINK Implementation of a CDMA Transmitter. *Honours thesis dissertation*, Edith Cowan University, Perth, Western Australia, 1996.
58. Rao, V. and Rao, H., *C++ Neural Networks and Fuzzy Logic*. New York: MIS Press, (2<sup>nd</sup> ed), 1995.
59. Rao, V. and Rao, H., *C++ Neural Networks and Fuzzy Logic*. New York: MIS Press, (2<sup>nd</sup> ed), 1995.
60. Reichart, J., "Automatic Classification of Communication Signals Using Higher Order Statistics", *Proceedings of ISCAPSP '92*, vol. 5, pp. 221-224, March 1992.
61. Rhode and Schwarz Broadcasting Division, "Fading Channel Simulation in DVB", [www.rohde-schwarz.com/www/appnotes\\_files.nsf/file/7BM05\\_1E.pdf/\\$file/7BM05\\_1E.pdf](http://www.rohde-schwarz.com/www/appnotes_files.nsf/file/7BM05_1E.pdf/$file/7BM05_1E.pdf), pp 3-13, December 2001.
62. Ripley, B. D., "Flexible non-linear approaches to classification", *From Statistics to Neural Networks. Theory and Pattern Recognition Applications*, 1994.
63. Ripley, B. D., *Pattern Recognition and Neural Networks*. Cambridge U.K.: Cambridge University Press, 1996.
64. Rosti, A.V., "Statistical Methods in Modulation Classification", *Masters Thesis Dissertation*, Tampere University of Technology, Tampere, Finland, December, 1998.
65. Schneider, R.F. and Chu, D.C., "Modulation Recognition of Spread Spectrum Signals Using Modulation Domain Measurements", *Proceedings of IEEE MILCOM '91*, pp. 35.1.1. - 35.1.8., 1991.
66. Schreyogg, C., and Reichert, J., "Modulation Classification of QAM Schemes using the DFT of Phase Histogram Combined with Modulus Information", *Proceedings of IEEE MILCOM '97*, vol. 3, pp. 1372-1376, November 1997.
67. Sills, J.A., "Maximum-Likelihood Modulation Classification for PSK/QAM", *Proceedings of IEEE MILCOM '99*, 1999.
68. Sklar, B., *Digital Communications: Fundamental and Applications*. New Jersey: Prentice Hall, 1988.
69. Soliman, S.S. and Hsue, S., "Automatic modulation classification using zero crossing" *IEE Proceedings of Radar and Signal Processing*, vol. 137, pp459-464, December 1990.
70. Soliman, S.S. and Hsue, S., "Automatic modulation Recognition of Digitally Modulated Signals", *Proceedings of IEEE MILCOM '89*, vol. 3, pp 645-649, October 1989.
71. Soliman, S.S. and Hsue, S., "Signal classification using statistical moments" *IEEE Trans. Commun.*, vol. 40, pp908-916, May 1992.
72. Stark, H. and Woods, J.W., *Probability, Random Processes, and Estimation Theory for Engineers*, 2<sup>nd</sup> ed., Prentice - Hall, 1994.
73. Sundberg, C.E., "Continuous Phase Modulation", *IEEE Communications Magazine*, vol. 24, pp 25-38, Apr 1986.
74. Sundberg, C.E., "Continuous Phase Modulation", *IEEE Communications Magazine*, vol. 24, pp 25-38, Apr 1986.
75. Svensson, A., "Receivers for CPM," Doctoral Thesis Telecommunication Theory, University of Lund, Sweden, May 1984.
76. Swami, A. and Sadler, B.M., "Hierarchical Digital Modulation Classification Using Cumulants", *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 416-429, March 2000.

77. Taira, S. and Murukami, E., "Automatic Classification of Analogue Modulation Signals by Statistical Parameters", *proceedings of IEEE MILCOM '99*, pp. 202-207, 1999.
78. Tou, J.T., and Gonzales, R.C., *Pattern Recognition Principles*, Addison Wesley, 1974.
79. Wang, W., *Communications Toolbox User's Guide for Use With Matlab*. MA: The Mathworks Inc, 1996.
80. Wei, W. and Mendel, J.M., "Maximum-Likelihood Classification for Digital Amplitude-Phase Modulations", *IEEE Transactions on Communications*, vol. 48, no. 2, pp. 189-193, February 2000.
81. Wei, W. and Mendel, J.M., "A New Maximum-Likelihood Method for Modulation Classification", *Proceedings of ASILOMAR-29*, vol. 2, pp. 1132-1136, November 1995.
82. Yang, Y. and Liu, C-H., "An Asymptotic Optimal Algorithm for Modulation Classification", *IEEE Communications Letters*, vol.2, no. 5, pp. 117-119, May 1998.
83. Yang, Y. and Soliman, S.S., "A Suboptimal Algorithm for Modulation Classification", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no 1, pp. 38-45, January 1997.
84. Yang, Y. and Soliman, S.S., "An improved moment – based algorithm for signal classification", *Signal Processing*, vol. 43, pp 231-244, May 1995.
85. Zhang, P.G., "Neural Networks for Classification: A Survey", *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, vol. 30, no 4, pp451-462, November 2000.